# Smart Antenna Systems for Mobile Communications

## FINAL REPORT

Ivica Stevanović, Anja Skrivervik and Juan R. Mosig
January 2003

Laboratoire d'Electromagnétisme et d'Acoustique

Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne Suisse
http://lemawww.epfl.ch/

ECOLE POLYTECHNIQUE
FEDERALE DE LAUSANNE

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Global demand for voice, data and video related services continues to grow faster than the required infrastructure can be deployed. Despite huge amount of money that has been spent in attempts to meet the need of the world market, the vast majority of people on Earth still do not have access to quality communication facilities. The greatest challenge faced by governments and service providers is the "last-mile" connection, which is the final link between the individual home or business users and worldwide network. Copper wires, traditional means of providing this "last-mile" connection is both costly and inadequate to meet the needs of the bandwidth intensive applications. Coaxial cable and power line communications all have technical limitations. And fiber optics, while technically superior and widely used in backbone applications, is extremely expensive to install to every home or business user. This is why more and more the wireless connection is being seen as an alternative to quickly and cost effectively meeting the need for flexible broadband links [1].

The universal and spread use of mobile phone service is a testament to the public's acceptance of wireless technology. Many of previously non-covered parts of the world now boast of quality voice service thanks in part to the PCS (Personal Communications Service) or cellular type wireless systems. Over the last few years the demand for service provision via the wireless communication bearer has risen beyond all expectations. At the end of the last century more than 20 million users in the United States only utilized this technology [2]. At present the number of cellular users is growing annually by approximately 50 percent in North America, 60 percent in western Europe, 70 percent in Australia and Asia and more than 200 percent in South America.

The proliferation of wireless networks and an increase in the bandwidth required has led to shortages in the scarcest resource of all, the finite number of radio frequencies that these devices use. This has increased the cost to obtain the few remaining licenses to use these frequencies and the related infrastructure costs required to provide these services.

In a majority of currently deployed wireless communication systems, the objective is to sell a product at a fair price (the product being information transmission) [3]. From a technical point of view, information transmission requires resources in the form of power and bandwidth. Generally, increased transmission rates require increased power and bandwidth independently of medium. While, on the one hand, transmission over wired segments of the links can generally be performed independently for each link (if we ignore the cross-talk in land lines) and, on the other hand, fibers are excellent at confining most of the useful information (energy) to a small region in space, wireless transmission

is much less efficient. Reliable transmission over relatively short distances in space requires a large amount of transmitted energy, spread over large regions of space, only a very small portion of which is actually received by the intended user. Most of the wasted energy is considered as interference to other potential users of the system.

Somewhat simplistically, the maximum range of such systems is determined by the amount of power that can be transmitted (and therefore received) and the capacity is determined by the amount of spectrum (bandwidth) available. For a given amount of power (constrained by regulation or practical considerations) and a fixed amount of bandwidth (the amount one can afford to buy) there is a finite (small) amount of capacity (bits/sec/Hz/unit-area, really per unit-volume) that operators can sell to their customers, and a limited range over which customers can be served from any given location. Thus, the two basic problems that arise in such systems are:

1. How to acquire more capacity so that a larger number of customers can be served at lower costs maintaining the quality at the same time, in areas where demand is large (spectral efficiency).
2. How to obtain greater coverage areas so as to reduce infrastructure and maintenance costs in areas where demand is relatively small (coverage).

In areas where demand for service exceeds the supply operators have to offer, the real game being played is the quest for capacity. Unfortunately, to date a universal definition of capacity has not evolved. Free to make their own definitions, operators and consumers have done so. To the consumer, it is quite clear that capacity is measured in the quality of each link he gets and the number of times he can successfully get such a link when he wants one. Consumers want the highest possible quality links at the lowest possible cost. Operators, on the other hand, have their own definitions of capacity in which great importance is placed on the number of links that can simultaneously be established. Since the quality and number of simultaneous links are inversely related in a resource-constrained environment, operators lean towards providing the lowest possible quality links to the largest possible number of users. The war wages on: consumers are wanting better links at lower costs, and operators are continually trying to maximize profitability providing an increasing number of lower quality links at the highest acceptable cost to the consumer. Until the quest for real capacity is successful, the battle between operators and their consumers over capacity, the precious commodity that operators sell to consumers, will continue.

There are many situations where coverage, not capacity, is a more important issue. Consider the rollout of any new service. Prior to initiating the service, capacity is certainly not a problem - operators have no customers. Until a significant percentage of the service area is covered, service cannot begin. Clearly, coverage is an important issue during the initial phases of system deployment. Consider also that in many instances only an extremely small percentage of the area to be served is heavily populated. The ability to cover the service area with a minimum amount of infrastructure investment is clearly an important factor in keeping costs down.

As it is often painfully obvious to operators, the two requirements, increased capacity and increased range, conflict in most instances. While up to recently used technology can provide for increased range in some cases and up to a limit increased capacity in other cases, it rarely can provide both simultaneously.

The International Mobile Telecommunications-2000 (IMT2000) and the European Universal Mobile

Telecommunications System (UMTS) are two systems among the others that have been proposed to take wireless communications into this century [2]. The core objective of both systems is to take the "personal communications user" into new information society where mass-market low-cost telecommunications services will be provided. In order to be universally accepted, these new networks have to offer mobile access to voice, data and multimedia facilities in an extensive range of operational environments, as well as economically supporting service provision in environments conventionally served by other wired systems. None of the proposals that include improved air interface and modulation schemes, deployment of smaller radio cells with combinations of different cell types in hierarchical architectures, and advanced signal processing, fully exploit the multiplicity of spatial channels that arises because each mobile user occupies a unique spatial location. Space is truly one of the final frontiers when it comes to new generation wireless communication systems. Spatially selective transmission and reception of RF energy promises substantial increases in wireless system capacity, coverage and quality. That this is certainly the case is attested to by the significant number of companies that have been recently brought the products based on such concepts to the wireless market place. Filtering in the space domain can separate spectrally and temporally overlapping signals from multiple mobile units. Thus, the spatial dimension can be exploited as a hybrid multiple access technique complementing frequency-division multiple access (FDMA), time-division MA (TDMA) and code-division MA (CDMA). This approach is usually referred to as space-division multiple access (SDMA) and enables multiple users within the same radio cell to be accommodated on the same frequency and time slot, as illustrated in Fig. 1.1.

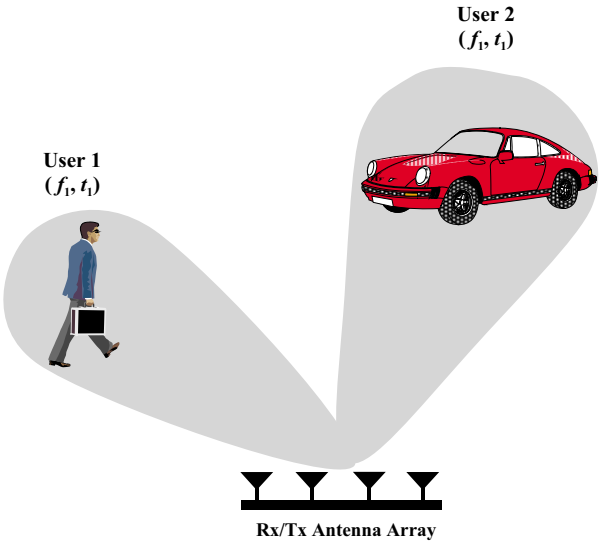



**Figure 1.1:** SDMA concept.

Realization of this filtering technique is accomplished using smart antennas, which are effectively antenna systems capable of modifying its time, frequency and spatial response. By exploiting the spatial domain via smart antenna systems, the operational benefits to the network operator can be summarized as follows:

- Capacity enhancement. SDMA with smart antennas allows for multiple users in a cell to use the same frequency without interfering with each other since the Base Station smart antenna

beams are sliced to keep different users in separate beams at the same frequency.

- Coverage extension. The increase in range is due to a bigger antenna gain with smart antennas. This would also mean that fewer Base Stations might be used to cover a particular geographical area and longer battery life in mobile stations.
- Ability to support high data rates.
- Increased immunity to "near-far" problems.
- Ability to support hierarchical cell structures.

## 1.1 Evolution from Omnidirectional to Smart Antennas

An antenna in a telecommunications system is the port through which radio frequency (RF) energy is coupled from the transmitter to the outside world for transmission purposes, and in reverse, to the receiver from the outside world for reception purposes [4]. To date, antennas have been the most neglected of all the components in personal communications systems. Yet, the manner in which radio frequency energy is distributed into and collected from space has a profound influence upon the efficient use of spectrum, the cost of establishing new personal communications networks and the service quality provided by those networks. The goal of the next several sections is to answer to the question "Why to use anything more than a single omnidirectional (no preferable direction) antenna at a base station?" by describing, in order of increasing benefits, the principal schemes for antennas deployed at base stations.

### 1.1.1 Omnidirectional Antennas

Since the early days of wireless communications, there has been the simple dipole antenna, which radiates and receives equally well in all directions (direction here being referred to azimuth). To find its users, this single-element design broadcasts omnidirectionally in a pattern resembling ripples radiation outward in a pool of water (Fig. 1.2).

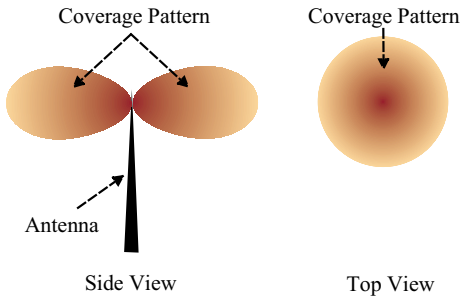



**Figure 1.2:** Omnidirectional Antennas and coverage patterns.

While adequate for simple RF environments where no specific knowledge of the users' whereabouts is either available or needed, this unfocused approach scatters signals, reaching desired users with only a small percentage of the overall energy sent out into the environment [5]. Given this limitation, omnidirectional strategies attempt to overcome environmental challenges by simply boosting the power level of the signals broadcast. In a setting of numerous users (and interferers), this makes a bad

situation worse in that the signals that miss the intended user become interference for those in the same or adjoining cells. In uplink applications (user to base station), omnidirectional antennas offer no preferential gain for the signals of served users. In other words, users have to shout over competing signal energy. Also, this single-element approach cannot selectively reject signals interfering with those of served users and has no spatial multipath mitigation or equalization capabilities. Therefore, omnidirectional strategies directly and adversely impact spectral efficiency, limiting frequency reuse. These limitations of broadcast antenna technology regarding the quality, capacity, and geographic coverage of wireless systems prompted an evolution in the fundamental design and role of the antenna in a wireless system.

### 1.1.2 Directional Antennas and Sectorized Systems

A single antenna can also be constructed to have certain fixed preferential transmission and reception directions. Sectorized antenna system take a traditional cellular area and subdivide it into sectors that are covered using directional antennas looking out from the same base station location (Fig. 1.3). Operationally, each sector is treated as a different cell in the system, the range of which can be greater than in the omni directional case, since power can be focused to a smaller area. This is commonly referred to as antenna element gain. Additionally, sectorized antenna systems increase the possible reuse of a frequency channel in such cellular systems by reducing potential interference across the original cell. As many as six sectors have been used in practical service, while more recently up to 16 sectors have been deployed [1]. However, since each sector uses a different frequency to reduce co-channel interference, handoffs (handovers) between sectors are required. Narrower sectors give better performance of the system, but this would result in to many handoffs.

While sectorized antenna systems multiply the use of channels, they do not overcome the major disadvantages of standard omnidirectional antennas such as filtering of unwanted interference signals from adjacent cells.

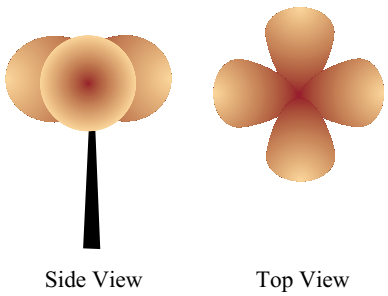



**Figure 1.3:** Sectorized antenna system and coverage pattern.

### 1.1.3 Diversity Systems

Wireless communication systems are limited in performance and capacity by three major impairments as shown in (Fig. 1.4) [6]. The first of these is multipath fading, which is caused by multiple paths that the transmitted signal can take to the receive antenna. The signals from these paths add with different phases, resulting in a received signal amplitude and phase that vary with antenna location, direction and polarization as well as with time (with movement in the environment). The second impairment

is delay spread, which is the difference in propagation delays among the multiple paths. When the delay spread exceeds about 10 percent of the symbol duration, significant intersymbol interference can occur, which limits the maximum data rate. The third impairment is co-channel interference. Cellular systems divide the available frequency channels into channel sets, using one channel set per cell, with frequency reuse (e.g. most TDMA systems use a frequency reuse factor of 7). This results in co-channel interference, which increases as the number of channel sets decreases (i.e. as the capacity of each cell increases). In TDMA systems, the co-channel interference is predominantly from one or two other users, while in CDMA systems there are typically many strong interferers both within the cell and from adjacent cells. For a given level of co-channel interference (channel sets), capacity can be increased by shrinking the cell size, but at the cost of additional base stations. We define the diversity gain (which is possible only with multipath fading) as the reduction in the required average output signal-to-noise ratio for a given BER with fading. All these concepts will be analyzed in more detail in the following chapters.

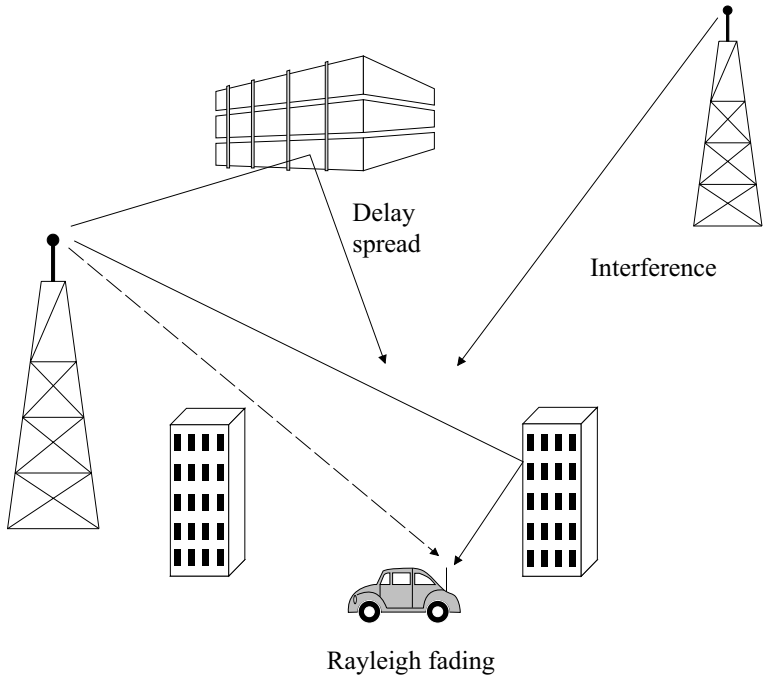



**Figure 1.4:** Wireless system impairments.

There are three different ways to provide low correlation (diversity gain): spatial, polarization and angle diversity.

For spatial diversity, the antennas are separated far enough for low fading correlation. The required separation depends on the angular spread, which is the angle over which the signal arrives at the receive antennas. With handsets, which are generally surrounded by other objects, the angular spread is typically $360^0$, and quarter-wavelength spacing of the antennas is sufficient. This also holds for base station antennas in indoor systems. For outdoor systems with high base station antennas, located above the clutter, the angular spread may be only a few degrees (although it can be much higher in urban areas), and a horizontal separation of 10-20 wavelengths is required, making the size of the antenna array an issue.

For polarization diversity, two orthogonal polarizations are used (they are often $\pm 45^0$). These orthogonal polarizations have low correlation, and the antennas can have a small profile. However, polarization diversity can only double the diversity, and for high base station antennas, the horizontal polarization can be $6-10$ dB weaker than the vertical polarization, which reduces the diversity gain.

For angle diversity, adjacent narrow beams are used. The antenna profile is small, and the adjacent beams usually have low fading correlation. However, with small angular spread, when the received signal is mainly arriving on one beam, the adjacent beams can have received signal levels more than 10 dB weaker than the strongest beam, resulting in small diversity gain.

Fig. 1.5 shows three antenna diversity options with four antenna elements for a $120^0$ sectorized system. Fig. 1.5(a) shows spatial diversity with approximately seven wavelengths ($7\lambda$) spacing between elements (3.3 m at 1900 MHz). A typical antenna element has an 18 dBi gain with a $65^0$ horizontal and $8^0$ vertical beamwidths. Figure Fig. 1.5(b) shows two dual polarization antennas, where the antennas can be either closely spaced ($\lambda/2$) to provide both angle and polarization diversity in a small profile, or widely spaced ($7\lambda$) to provide both spatial and polarization diversity. The antenna elements shown are $45^0$ slant polarization antennas, which are also commonly used, rather than vertically and horizontally polarized antennas. Finally, Fig. 1.5(c) shows a closely spaced ($\lambda/2$) vertically polarized array, which provides angle diversity in a small profile.

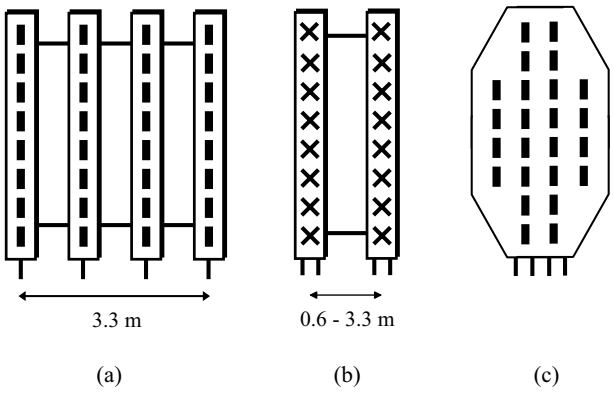



**Figure 1.5:** Antenna diversity options with four antenna elements: (a) spatial diversity; (b) polarization diversity with angular and spatial diversity; (c) angular diversity.

Diversity offers an improvement in the effective strength of the received signal by using one of the following two methods

- Switched diversity. Assuming that at least one antenna will be in a favorable location at a given moment, this system continually switches between antennas (connects each of the receiving channels to the best serving antenna) so as always to use the element with the highest signal power.
- Diversity combining. This approach corrects the phase error in two multipath signals and effectively combines the power of both signals to produce gain. Other diversity systems, such as maximal ratio combining systems, combine outputs of all the antennas to maximize the ratio of combined received signal energy to noise.

The diversity antennas merely switch operation from one working element to the other. Although

this approach mitigates severe multipath fading, its use of one element at a time offers no uplink gain improvement over any other single-element approach. The diversity systems can be useful in environments where fading is the dominant mechanism for signal degradation. In environments with significant interference, however, the simple strategies of locking onto the strongest signal or extracting maximum signal power from the antennas are clearly inappropriate and can result in crystal-clear reception of an interferer at the expense of the desired signal.

The need to transmit to numerous users more efficiently without compounding the interference problem led to the next step of the evolution antenna systems that intelligently integrate the simultaneous operation of diversity antenna elements.

## 1.2 Smart Antenna Systems

### 1.2.1 Catalogue of definitions

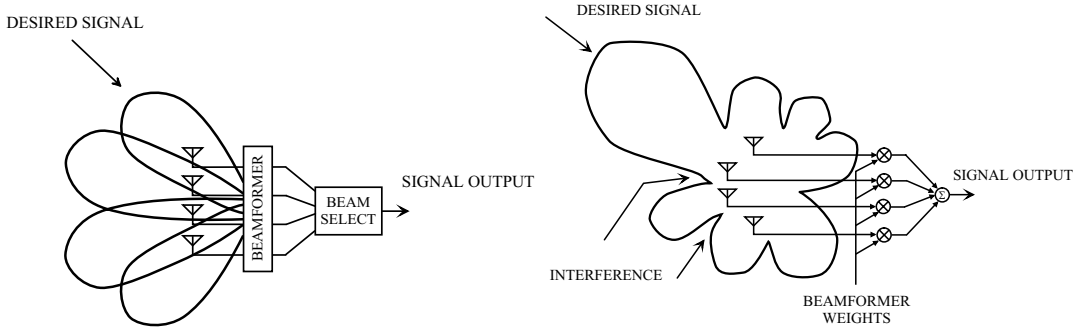
In this section the three definitions most frequently found in literature are listed. The only difference between them is in the way in which different types of Smart Antenna Systems are categorized.

**First Definition [7]**

A smart antenna is a phased or adaptive array that adjusts to the environment. That is, for the adaptive array, the beam pattern changes as the desired user and the interference move, and for the phased array, the beam is steered or different beams are selected as the desired user moves.

**Phased array or multibeam antenna** consists of either a number of fixed beams with one beam turned on towards the desired signal or a single beam (formed by phase adjustment only) that is steered towards the desired signal.

**Adaptive antenna array** is an array of multiple antenna elements with the received signals weighted and combined to maximize the desired signal to interference and noise (SINR) ratio. This means that the main beam is put in the direction of the desired signal while nulls are in the direction of the interference.




**(a)** Phased array. **(b)** Adaptive array.

**Figure 1.6:** Smart antenna systems definition.

**Second Definition [5, 8, 9]**

A smart antenna system combines multiple antenna elements with a signal processing capability to optimize its radiation and/or reception pattern automatically in response to the signal environment.

Smart antenna systems are customarily categorized as either switched beam or adaptive array systems.

**Switched beam antenna system** form multiple fixed beams with heightened sensitivity in particular directions. These antenna systems detect signal strength, choose from one of several predetermined, fixed beams, and switch from one beam to another as demand changes throughout the sector. Instead of shaping the directional antenna pattern with the metallic properties and physical design of a single element (like a sectorized antenna), switched beam systems combine the outputs of multiple antennas in such a way as to form finely sectorized (directional) beams with more spatial selectivity than it can be achieved with conventional, single element approaches.

**Adaptive antenna array systems** represent the most advanced smart antenna approach to date. Using a variety of new signal-processing algorithms, the adaptive system takes advantage of its ability to effectively locate and track various types of signals to dynamically minimize interference and maximize intended signal reception.

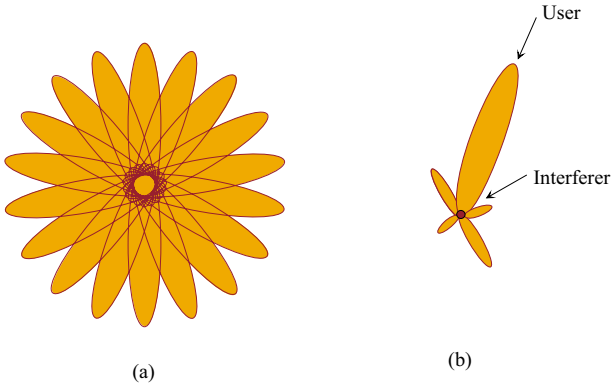



**Figure 1.7:** Switched beam system coverage patterns (a) and Adaptive array coverage (b).

**Third Definition [10, 11]**

Smart Antennas are arrays of antenna elements that change their antenna pattern dynamically to adjust to the noise, interference in the channel and mitigate multipath fading effects on the signal of interest.

The difference between a smart (adaptive) antenna and "dumb" (fixed) antenna is the property of having an adaptive and fixed lobe-pattern, respectively. The secret to the smart antennas' ability to transmit and receive signals in an adaptive, spatially sensitive manner is the digital signal processing capability present. An antenna element is not smart by itself; it is a combination of antenna elements to form an array and the signal processing software used that make smart antennas effective. This shows that smart antennas are more than just the "antenna", but rather a complete transceiver concept.

Smart Antenna systems are classified on the basis of their transmit strategy, into the following three types ("levels of intelligence"):

- Switched Beam Antennas
- Dynamically-Phased Arrays
- Adaptive Antenna Arrays

**Switched Beam Antennas**
Switched beam or switched lobe antennas are directional antennas deployed at base stations of a cell. They have only a basic switching function between separate directive antennas or predefined beams of an array. The setting that gives the best performance, usually in terms of received power, is chosen. The outputs of the various elements are sampled periodically to ascertain which has the best reception beam. Because of the higher directivity compared to a conventional antenna, some gain is achieved. Such an antenna is easier to implement in existing cell structures than the more sophisticated adaptive arrays, but it gives a limited improvement.

**Dynamically-Phased Arrays**
The beams are predetermined and fixed in the case of a switched beam system. A user may be in the range of one beam at a particular time but as he moves away from the center of the beam and crosses over the periphery of the beam, the received signal becomes weaker and an intra cell handover occurs. But in dynamically phased arrays, a direction of arrival (DoA) algorithm tracks the user's signal as he roams within the range of the beam that's tracking him. So even when the intra-cell handoff occurs, the user's signal is received with an optimal gain. It can be viewed as a generalization of the switched lobe concept where the received power is maximized.

**Adaptive Antenna Arrays**
Adaptive antenna arrays can be considered the smartest of the lot. An Adaptive Antenna Array is a set of antenna elements that can adapt their antenna pattern to changes in their environment. Each antenna of the array is associated with a weight that is adaptively updated so that its gain in a particular look-direction is maximized, while that in a direction corresponding to interfering signals is minimized. In other words, they change their antenna radiation or reception pattern dynamically to adjust to variations in channel noise and interference, in order to improve the SNR (signal to noise ratio) of a desired signal. This procedure is also known as 'adaptive beamforming' or 'digital beamforming'.

Conventional mobile systems usually employ some sort of antenna diversity (e.g. space, polarization or angle diversity). Adaptive antennas can be regarded as an extended diversity scheme, having more than two diversity branches. In this context, phased arrays will have a greater gain potential than switched lobe antennas because all elements can be used for diversity combining.

### 1.2.2 Relative Benefits/Tradeoffs of Switched Beam and Adaptive Array Systems

In the previous section three different definitions of Smart Antenna Systems, most commonly found in literature, are listed. However, the second definition, in which Smart Antenna Systems are divided into Switched Beam and Adaptive Array antenna systems, will be taken as a reference throughout this report. In this definition, the adaptive array antennas are subdivided into two classes: the first is the phased array antennas where only the phase of the currents is changed by the weights, and the second class are adaptive array antennas in strict sense, where both the amplitude and the phase of the currents are changed to produce a desired beam.

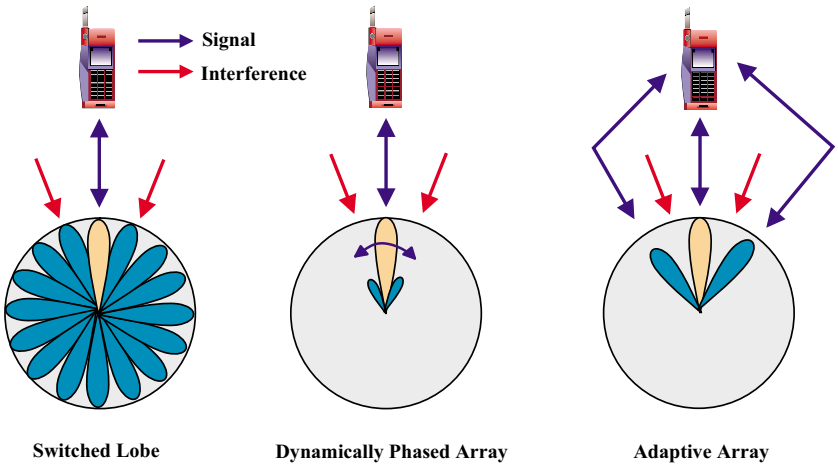



**Figure 1.8:** Different smart antenna concepts.

In terms of radiation patterns, switched beam is an extension of the cellular sectorization method in which a typical sectorized cell site has three 120-degree macro-sectors. The switched beam approach further subdivides macro-sectors into several micro-sectors thus improving range and capacity. Each micro-sector contains a predetermined fixed beam pattern with the greatest sensitivity located in the center of the beam and less sensitivity elsewhere. The design of such systems involves high-gain, narrow azimuth beam width antenna elements.

The switched beam system selects one of several predetermined fixed-beam patterns (based on weighted combinations of antenna outputs) with the greatest output power in the remote user's channel. RF or baseband DSP hardware and software drive these choices. The system switches its beam in different directions throughout space by changing the phase differences of the signals used to feed the antenna elements or received from them. When the mobile user enters a particular macro-sector, the switched beam system selects the micro-sector containing the strongest signal. Throughout the call, the system monitors signal strength and switches to other fixed micro-sectors as required.

All switched beam systems provide similar benefits even though the various systems utilize different hardware and software designs [9]. When compared to conventional sectored cells, switched beam systems can increase the range of a base station by anywhere from 20 to 200% depending on the circumstances. The additional coverage can save an operator substantial amounts in infrastructure costs and allow them to lower prices for consumers while remaining profitable.

There are, however, limitations to switched beam systems. Because beams are predetermined, the signal strength varies as the user moves through the sector. As a mobile unit moves towards the far azimuth edges of a beam, the signal strength can degrade rapidly before the user is switched to another micro-sector. Another limitation occurs because a switched beam system does not distinguish between a desired signal and interfering ones. If the interfering signal is at approximately the center of the selected beam and the user is away from the center of the selected beam, the interfering signal can be enhanced far more than the desired signal. In these cases, the quality for the user is degraded.

The adaptive antenna systems take a different approach. By adjusting to an RF environment as it changes (or the spatial origin of signals), adaptive antenna technology can dynamically alter the signal patterns to optimize the performance of the wireless system.

The adaptive approach utilizes sophisticated signal processing algorithms to continuously distinguish between desired signals, multipath and interfering signals as well as calculate their directions of arrival. This approach continuously updates its beam pattern based on changes in both the desired and interfering signal locations. The ability to smoothly track users with main lobes and interferers with nulls insures that the link budget is constantly maximized (there are neither micro-sectors nor pre-defined patterns).

This effect is similar to a person's hearing. When one person listens to another, the brain of the listener collects the sound in both ears, combines it to hear better, and determines the direction from which the speaker is talking. If the speaker is moving , the listener, even if his or her eyes are closed, can continue to update the angular position based solely on what he or she hears. The listener also has the ability to tune out unwanted noise, interference and focus on the conversation at hand.

Fig. 1.9 illustrates the beam patterns that each system might choose in the face of a signal of interest and two co-channel interferers in the positions shown. The switched beam system is shown in red on the left while the adaptive system is shown in blue on the right. The green lines delineate the signal of interest while the yellow lines display the direction of the co-channel interfering signals. Both systems have directed the lobe with the most gain in the general direction of the signal of interest, although the adaptive system has chosen more accurate placement, providing greater signal enhancement. Similarly, the interfering signals arrive at places of lower gain outside the main lobe, but again the adaptive system has placed these signals at the lowest possible gain points and better insures that the main signal received maximum enhancement while the interfering signals receive maximum suppression.

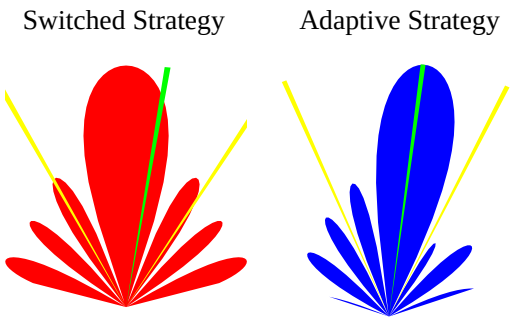



**Figure 1.9:** Beamforming lobes and nulls that Switched Beam (red) and Adaptive Array (blue) systems might choose for identical user signals (green line) and co-channel interferers (yellow lines).

Fig. 1.10 illustrates the relative coverage area for conventional sectorized, switched beam and adaptive antenna systems. Both types of smart antenna systems provide significant gains over conventional sectorized system. The low level of interference on the left represents a new wireless system with lower penetration levels. The significant level of interference on the right represents either a wireless system with more users or one using more aggressive frequency re-use patterns. In this scenario, the interference rejection capability of the adaptive system provides significantly more coverage than either the conventional or switched beam systems.

Another significant advantage of the adaptive antenna systems is the ability to "create" spectrum.

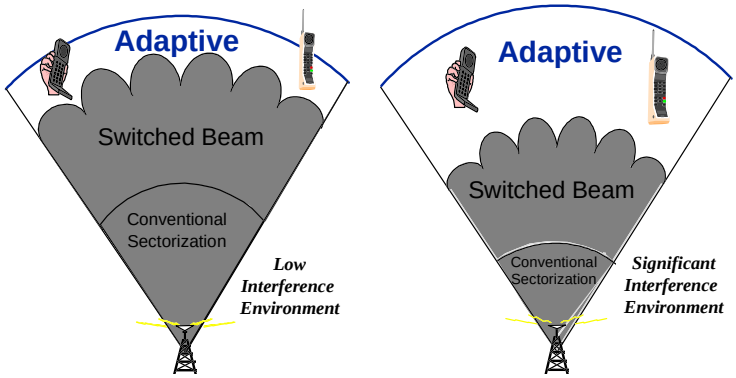



**Figure 1.10:** Coverage patterns for switched beam and adaptive array antennas.

Because of the accurate tracking and robust interference rejection capabilities, multiple users can share the same conventional channel within the same cell. System capacity increases through lower inter-cell frequency re-use patterns as well as intra-cell frequency re-use. The Fig. 1.11 shows how adaptive antenna approach can be used to support two users on the same conventional channel at the same time in the same cell. The blue beam pattern is used to communicate with the user on the left. The yellow pattern is used to talk with the user on the right. The red lines delineate the actual direction of each signal. Notice as the signals travel down the red line toward the base station, the yellow signal arrives at a blue null or minimum gain point and vice versa. As the users move, beam patterns are constantly updated to insure these positions. The right plot shows how the beam patterns have dynamically changed to insure maximum signal quality as one user moves towards the other.

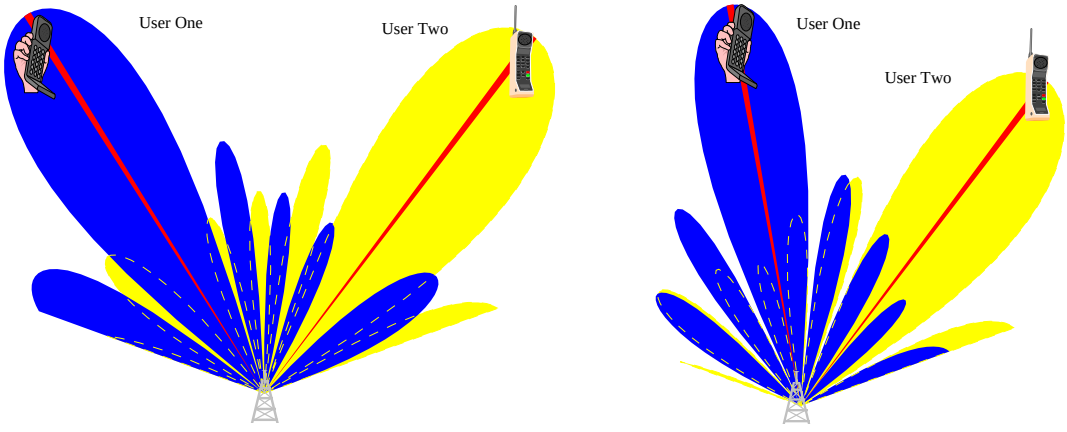



**Figure 1.11:** Fully adaptive spatial processing supporting two users on the same conventional channel simultaneously in the same cell.

The ability to continuously change the beam pattern with respect to both lobes and nulls separates the adaptive approach from the switched type. As interfering signals move throughout the sector, the

switched beam pattern is not altered because it only responds to movements in the signal of interest. In fact, when an interfering signal begins to approach the signal of interest and enters the gain of the main lobe, the interfering signal will be processed identically to the desired signal and signal to interference ratio will degrade accordingly. In contrast, the adaptive system is able to continue to distinguish between the signal and the interferer and allow them to get substantially closer than in the switched beam system while maintaining enhanced signal to interference ratio levels. The most sophisticated adaptive smart antenna systems will hand-over any two co-channel users, whether they are inter-cell or intra-cell, before they get too close and begin to interfere with each other. The benefits and tradeoffs of switched beam and adaptive array systems can be summarized as follows

- **Integration** - Switched beam systems are traditionally designed to retrofit widely deployed cellular system. They have been commonly implemented as an add-on or appliqué technology that intelligently addresses the needs of mature networks. In comparison, adaptive array systems have been deployed with a more fully integrated approach that offers less hardware redundancy than switched beam systems but require new build-out.

- **Range/Coverage** - Switched beam systems can increase base station range from 20 to 200% over conventional sectored cells, depending on environmental circumstance and hardware/software used. The added coverage can save an operator a substantial infrastructure costs and means lower prices for consumers. Also, the dynamic switching from beam to beam conserves capacity because the system does not send all signals in all directions. In comparison, adaptive array systems can cover a broader, more uniform area with the same power levels as a switched beam system.

- **Interference Suppression** - Switched beam antennas suppress interference arriving from directions away from the active beam's center. Because beam patterns are fixed, however, actual interference rejection is often the gain of the selected communication beam pattern in the interferer's direction. Also, they are normally used only for reception because of the system's ambiguous perception of the location of the received signal (the consequences of transmitting in the wrong beam being obvious). Also, because their beams are predetermined, sensitivity can occasionally vary as the user moves through the sector. Switched beam solutions work best in minimal to moderate co-channel interference and have difficulty in distinguishing between a desired signal and an interferer. If the interfering signal is at approximately the center of the selected beam, the interfering signal can be enhanced far more than the desired signal. Adaptive antenna approach offers more comprehensive interference rejection. Also, because it transmits an infinite, rather than finite number of combinations, its narrower focus creates less interference to neighboring users than a switched-beam approach.

- **Cost/Complexity** - In adaptive antenna technology more intensive signal processing via DSP's is needed and at the same time the installation costs are higher when compared to switched beam antennas.

### 1.2.3 Smart Antenna Evolution

All the levels of intelligence described in the previous sections are technologically realizable today. Until recently, cost barriers have prevented their use in commercial systems. The advent of low cost

digital signal and general-purpose processors and innovative algorithms have made smart antenna systems practical at a time where spectrally efficient solutions are an imperative. In the domain of personal and mobile communications, an evolutionary path in the utilization of smart antennas towards gradually more advanced solution can be established. The "levels of intelligence" in the previous section describe the level of technological development, while the steps described here can be regarded as part of a system evolution. The evolution can be divided into three phases:

- Smart antennas are used on uplink only (uplink meaning that the user is transmitting and the base station is receiving). By using a smart antenna to increase the gain at the base station, both the sensitivity and range are increased. This concept is called high sensitivity receiver (HSR) and is in principle not different from the diversity techniques implemented in mobile communication systems.

- In the second phase, directed antenna beams are used on the downlink direction (base station transmitting and user receiving) in addition to HSR. In this way, the antenna gain is increased both on uplink and downlink, which implies a spatial filtering in both directions. The method is called spatial filtering for interference reduction (SFIR). It is possible to introduce this in second-generation systems. In GSM, which is a TDMA/FDMA system this interference reduction results in an increase of the capacity or the quality in the system. This is achieved by either allowing a tighter re-use factor and thereby a higher capacity, or to keep the same re-use factor but with a higher SNR level and signal quality. In CDMA based systems, due to non-orthogonality between the codes at the receiver, the different users will interfere with each other. This is called Multiple Access Interference (MAI) and its effect is a reduction of the capacity in the CDMA network. An interference reduction provided by smart antennas translates directly into a capacity or quality increase in CDMA networks.

- The last stage in the development is the full space division multiple access (SDMA). This implies that more than one user can be allocated to the same physical communications channel simultaneously in the same cell separated by angle. It is a separate multiple access method, but is usually combined with other multiple access methods (FDMA, TDMA, CDMA). In a hard-limited system like GSM, SDMA allows more than 8 full-rate users to be served in the same cell on the same frequency at the same time by exploiting the spatial domain. CDMA does not have a similar hard-limit on the number of users. Instead, it is the multiple access interference (MAI) due to the non-orthogonality of the channel codes that limits the number of users. This flexibility inherent in CDMA systems allows the interference reduction to be translated into either more users in the system, higher bit rates for the existing users, improved quality for the existing users at the same bit-rates, extended cell range for the same number of users at the same bit rates, or any arbitrary combination of these.

# Chapter 2

# System Elements of a Smart Antenna

In this chapter the basic principle behind smart antennas is explained. In the first two sections the block diagrams of smart antenna receiving and transmitting systems are presented. In the last section the fundamental concepts of antenna arrays are presented.

## 2.1 Smart Antenna Receiver

Fig. 2.1 shows schematically the elements of the reception part of a smart antenna. The antenna array contains $M$ elements. The $M$ signals are being combined into one signal, which is the input to the rest of the receiver (channel decoding, etc.).

As the figure shows, the smart antenna reception part consists of four units. In addition to the antenna itself it contains a radio unit, a beam forming unit and a signal processing unit [12].

The array will often have a relatively low number of elements in order to avoid unnecessarily high complexity in the signal processing. Fig. 2.2 shows four examples of different array geometries. The first two structures are used for beamforming in the horizontal plane (azimuth) only. This will normally be sufficient for outdoor environments, at least in large cells. The first example (a) shows an one–dimensional linear array with uniform element spacing of $\triangle x$. This structure can perform beamforming in azimuth angle within an angular sector. This is the most common structure due to its low complexity. The second example (b) shows a birds eye view of a circular array with angular element spacing of $\triangle\phi = 2\pi/M$. This structure can perform beamforming in all azimuth angles. The last two structures are used for performing two–dimensional beamforming, in both azimuth and elevation angles. This may be desirable for indoor or dense urban environments. The front view of a two–dimensional linear array with horizontal element spacing of $\triangle x$ and vertical element spacing of $\triangle y$. Beamforming in the entire space, within all angles, requires some sort of cubic or spherical structure. The fourth example (d) shows a cubic structure with element separations of $\triangle x$, $\triangle y$ and $\triangle z$.

The radio unit consists of down–conversion chains and (complex) analog-to-digital converters (A/D). There must be $M$ down-conversion chains, one for each of the array elements.

The signal processing unit will, based on the received signal, calculate the complex weights $w_1, \ldots, w_M$ with which the received signal from each of the array elements is multiplied. These weights will decide the antenna pattern in the uplink direction (which will be shown in more detail later). The weights

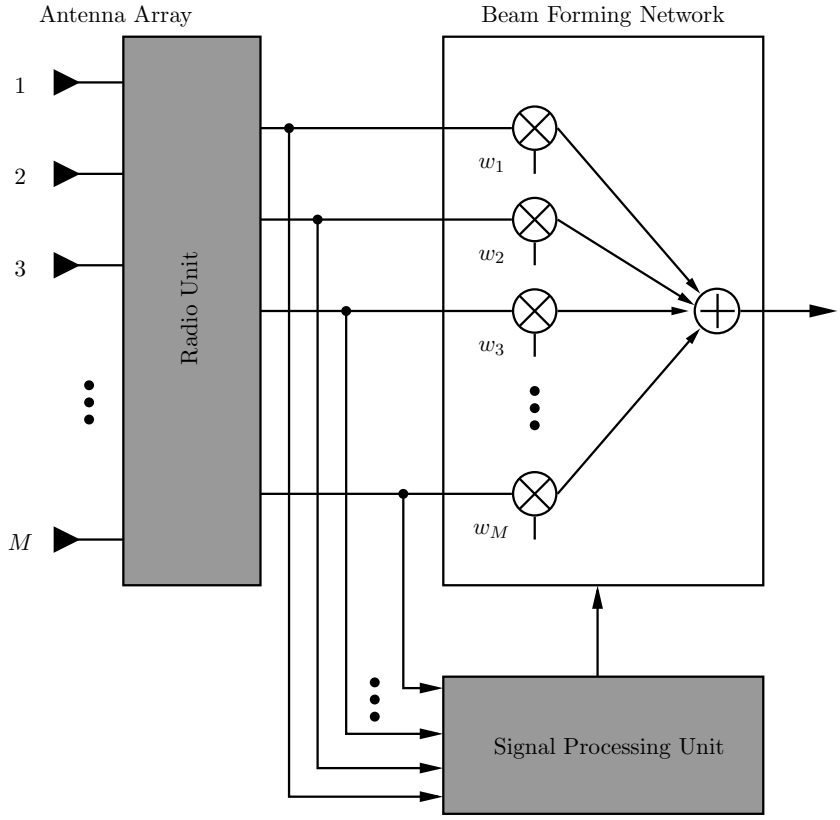



**Figure 2.1:** Reception part of a smart antenna.

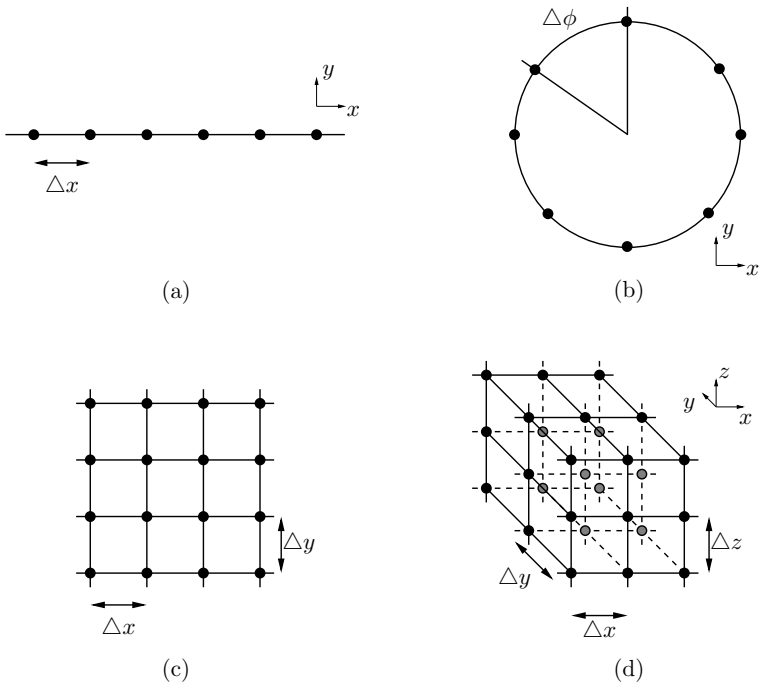



**Figure 2.2:** Different array geometries for smart antennas. (a) uniform linear array, (b) circular array, (c) 2 dimensional grid array and (d) 3 dimensional grid array.

can be optimized from two main types of criteria: maximization of received signal from the desired user (e.g. switched beam or phased array) or maximization of the SIR by suppressing the signal from interference sources (adaptive array). In theory, with $M$ antenna elements one can "null out" $M-1$ interference sources, but due to multipath propagation this number will normally be lower.

The method for calculating the weights will differ depending on the type of optimization criterion. When *switched beam* (SB) is used, the receiver will test all the pre-defined weight vectors (corresponding to the beam set) and choose the one giving the strongest received signal level. If the *phased array* approach (PA) is used, which consists of directing a maximum gain beam towards the strongest signal component, the *direction-of-arrival* (DoA) is first estimated and then the weights are calculated. A number of well documented methods exist for estimating the DoA and will be presented later.

If maximization of SIR is to be done (AA), the optimum weight vector (of dimension $M$) $\mathbf{W}_{\text{opt}}$ can be computed using a number of algorithms such as optimum combining and others that will be shown in the following.

When the beam forming is done digitally (after A/D), the beam forming and signal processing units can normally be integrated in the same unit (Digital Signal Processor, DSP). The separation in Fig. 2.1 is done to clarify the functionality. It is also possible to perform the beam forming in hardware at radio frequency (RF) or intermediate frequency (IF).

## 2.2 Smart Antenna Transmitter

The transmission part of the smart antenna is schematically very similar to the reception part. An illustration is shown in Fig. 2.3. The signal is split into $M$ branches, which are weighted by the complex weights $w_1, \ldots, w_M$ in the beam forming unit. The weights, which decide the radiation pattern in the downlink direction, are calculated as before by the signal processing unit. The radio unit consists of D/A converters and the up converter chains. In practice, some components, such as the antennas themselves and the DSP will of course be the same as on reception.

The principal difference between uplink and downlink is that no knowledge of the spatial channel response is available on downlink. In a time division duplex (TDD) system the mobile station and base station use the same carrier frequency only separated in time. In this case the weights calculated on uplink will be optimal on downlink if the channel does not change during the period from uplink to downlink transmission. However, this can not be assumed to be the case in general, at least not in systems where the users are expected to move at high speed. If frequency division duplex (FDD) is used, the uplink and downlink are separated in frequency. In this case the optimal weights will generally not be the same because of the channel response dependency on frequency.

Thus optimum beamforming (i.e., AA) on downlink is difficult and the technique most frequently suggested is the geometrical approach of estimating the direction-of-arrival (DoA). The assumption is directional reciprocity, i.e., the direction from which the signal arrived on the uplink is the direction in which the signal should be transmitted to reach the user on downlink. The strategy used by the base station is to estimate the DoA of the direction (or directions) from which the main part of the user signal is received. This direction is used on downlink by choosing the weights $w_1, \ldots, w_M$ so that the radiation pattern is a lobe or lobes directed towards the desired user. This is similar to Phased Array Systems. In addition, it is possible to position zeros in the direction towards other users so that the interference suffered by these users is minimized. Due to fading on the different signal paths, it has

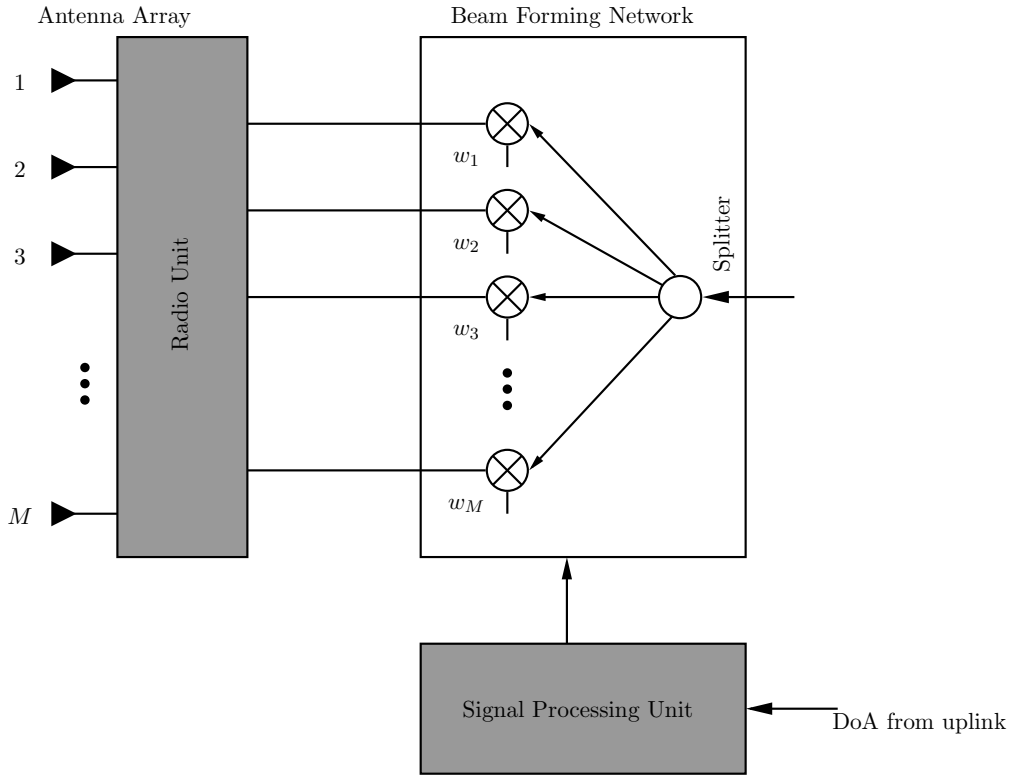



**Figure 2.3:** Transmission part of a smart antenna.

been suggested to choose the downlink direction based on averaging the uplink channel over a period of time. This will however be sub-optimum compared to the uplink situation where knowledge about the instantaneous radio channel is available.

It should be stressed that in the discussion above it is assumed that the interferers observed by the base stations are mobile stations and that the interferers observed by the mobile stations are base stations. This means that when the base station on transmission positions zeros in the direction towards other mobile stations than the desired one, it will reduce the interference suffered by these mobiles. If, however, the interferers observed by mobiles are other mobiles, as maybe the case, there will be a much more fundamental limitation in the possibility for interference reduction at the mobile.

## 2.3 Fundamentals of Antenna Arrays

An antenna array has spatially separated sensors whose output are fed into a weighting network or a beamforming network as shown in Fig. 2.1 and Fig. 2.3. The antenna array can be implemented as a transmitting or a receiving array. There are many assumptions made in analyzing an antenna array, they are as follows [13]:

- All signals incident on the receiving antenna array are composed of finite number of plane waves. These plane waves result from the direct as well as the multipath components.
- The transmitter and the objects that cause multipaths are in the far-field of the antenna array.
- The sensors are placed closely so that the amplitudes of the signals received at any two elements

of the antenna array do not differ significantly.

- Each sensor is assumed to have the same radiation pattern and the same orientation.
- The mutual coupling between the antenna elements is assumed to be negligible.

An antenna array with its coordinates is illustrated in Fig. 2.4.

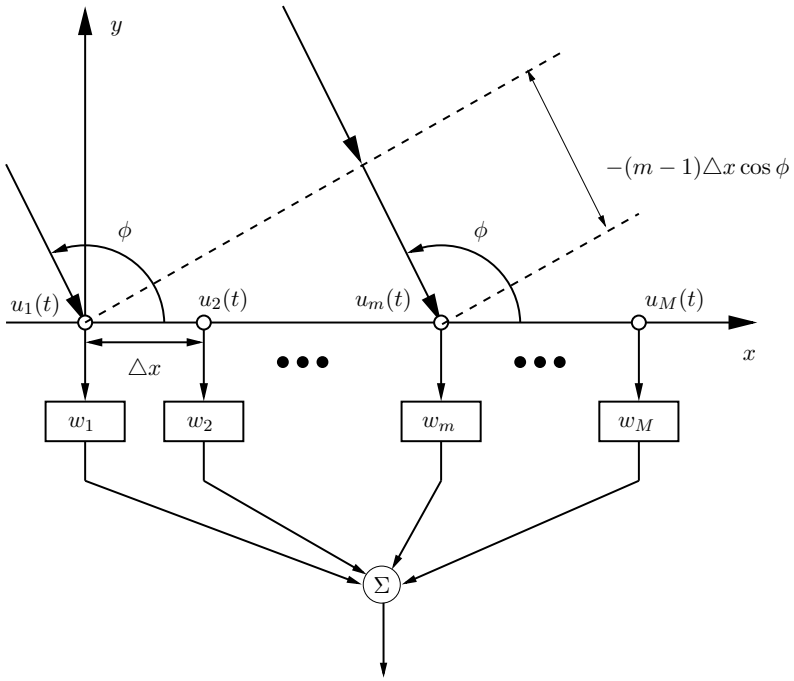



**Figure 2.4:** Illustration of plane wave incident from an angle $\phi$ on an uniform linear array (ULA) with inter-element spacing of $\triangle x$.

### 2.3.1 Theoretical model for an antenna array

An antenna array can be arranged in any arbitrary fashion, but the most preferred geometries are linear and circular geometries. Linear geometry is simpler to implement than the circular geometry, but the disadvantage is the symmetry (ambiguity) of the radiation pattern about the axis along the endfire, which is not the case in circular array. Linear array with uniformly spaced sensors is the most commonly used structure.

The array as shown in Fig. 2.5 has a reference element at the origin and the coordinates of the $m$th antenna element are marked as ($x_m$, $y_m$, $z_m$). The signal as it travels across the array undergoes a phase shift. The phase shift between the signal received at the reference element and the signal received at the element $m$ is given by

$$\triangle\gamma_m = \gamma_m(t) - \gamma_1(t) = -\beta x_m \cos\phi \sin\theta - \beta y_m \sin\phi \sin\theta - \beta z_m \cos\theta, \tag{2.1}$$

where $\beta = 2\pi/\lambda$ is the propagation constant in free space. This relation holds for a narrowband signal, in this case a signal whose modulated bandwidth is much less than the carrier frequency. The narrowband assumption allows us to assume that the only difference between the signal present at

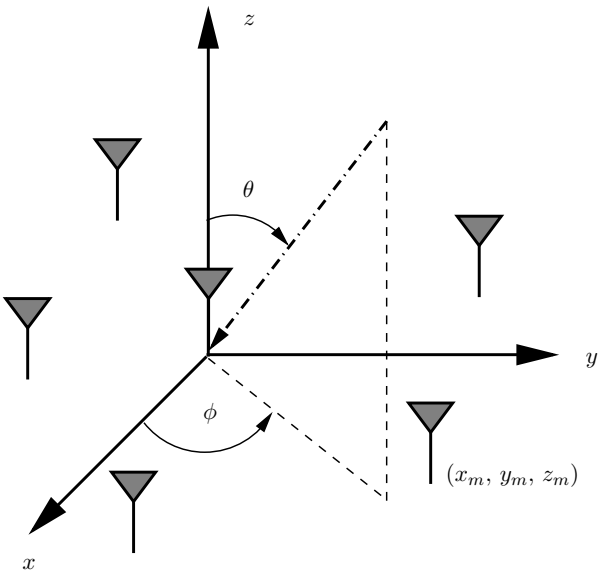



**Figure 2.5:** Illustration of the coordinates of an antenna array.

different elements of the array is the phase shift induced by the extra distance traveled and is not significantly affected by the modulation during this time. The reference plane is assumed to lie on $z = 0$. Since the distance between the transmitting and receiving antenna is larger than the distance between the heights of the receiving and transmitting antenna, a wave reaching the antenna array can be assumed to come along the horizon or with $\theta = 90^0$. Therefore, we will describe the direction-of-arrival (DoA) of each plane wave using only azimuth coordinate $\phi$. From (2.1) it can be seen that any variation in the array element height $z_m$ does not affect the phase difference between the reference element and element $m$. Therefore, we may consider only $x$ and $y$ offsets from the reference element.

Consider a transmitted narrowband signal in complex envelope representation

$$u_m(t) = A_m(t)e^{\mathrm{j}\gamma_m(t)}, \tag{2.2}$$

where $A_m(t)$ is the magnitude and $\gamma_m(t)$ is the phase of the signal. The vector containing these signals is called the *data* or the *illumination factor*

$$\mathbf{u}(t) = [u_1(t) \quad u_2(t) \quad \dots \quad u_M]. \tag{2.3}$$

A complex quantity $a_m(\phi)$ is defined as the ratio between the signal received at the antenna element $m$ and the signal received at the reference element when a plane wave is incident on the array and it is given by

$$a_m(\phi) = e^{-\mathrm{j}\beta(x_m \cos\phi + y_m \sin\phi)}. \tag{2.4}$$

If a single plane wave is incident on the antenna array, then

$$u_m(t) = u_1(t)a_m(\phi). \tag{2.5}$$

The response of an antenna array to a traveling single plane wave coming at an angle $\phi$ is defined as

the steering vector

$$\mathbf{a}(\phi) = \begin{bmatrix} 1 \\ a_2(\phi) \\ \dots \\ a_M(\phi) \end{bmatrix} = \begin{bmatrix} 1 \\ e^{-\mathrm{j}\beta(x_2 \cos\phi + y_2 \sin\phi)} \\ \dots \\ e^{-\mathrm{j}\beta(x_M \cos\phi + y_M \sin\phi)} \end{bmatrix}. \tag{2.6}$$

The collection of the steering vectors for all angles for a given frequency is known as the array manifold. The array manifold must be carefully measured to calibrate the array for direction finding experiments.

For narrowband adaptive beamforming, each array element output is multiplied by a complex weight $w_i^*$ modifying the phase and amplitude relation between the branches, and summed to give

$$\begin{aligned} v(t) &= u_1(t) \sum_{m=1}^{M} w_m^* e^{-\mathrm{j}\beta(x_m \cos\phi + y_m \sin\phi)} = \\ &= [w_1^* \quad w_2^* \quad \dots \quad w_M^*] \begin{bmatrix} 1 \\ e^{-\mathrm{j}\beta(x_2 \cos\phi + y_2 \sin\phi)} \\ \dots \\ e^{-\mathrm{j}\beta(x_M \cos\phi + y_M \sin\phi)} \end{bmatrix} u_1(t) = \mathbf{w}^{\mathrm{H}} \mathbf{u}(t). \end{aligned} \tag{2.7}$$

The response of the array (uniform linear array of isotropic elements) with the weighting network is called the array factor and it's defined as

$$\mathrm{AF}(\phi) = \frac{v(\phi)}{\max[v(\phi)]} = \mathbf{w}^{\mathrm{H}} \mathbf{a}(\phi). \tag{2.8}$$

The weighting network in an antenna array can be fixed or varying. In an adaptive array, the weights are adapted by minimizing certain criterion to maximize the signal-to-interference plus noise ratio (SINR) at the output of the array. Hence, the weighting network is very similar to a finite-impulse response (FIR) filter, where the time samples are replaced by spatial samples. The weighting network is therefore called spatial filter.

### 2.3.2 Array geometry and element spacing

The inter-element spacing between the antenna elements is an important factor in the design of an antenna array. If the elements are more than $\lambda/2$ apart, then the grating lobes appear which degrades the array performances.

Mutual coupling as an effect that limits the inter-element spacing of an array. If the elements are spaced closely (typically less than $\lambda/2$), the coupling effects will be larger and generally tend to decrease with increase in the spacing. Therefore, the elements have to be far enough to avoid mutual coupling and the spacing has to be smaller than $\lambda/2$ to avoid grating lobes. For all practical purposes, a spacing of $\lambda/2$ is preferred.

# Chapter 3

# Channel Model

In order to evaluate the performance of a smart antenna system, it is necessary to have detailed knowledge of the channel and the channel parameters. This is because the propagation channel is the principal contributor to many of the problems and limitations that beset mobile radio systems.

The propagation of radio signals on both the forward (base station to mobile) and reverse (mobile to base station) links is affected by the physical channel in several ways. In this chapter we review such effects and present detailed models to describe channel behavior.

A signal propagating through the wireless channel usually arrives at the destination along a number of different paths, referred to as multipaths. These paths arise from scattering, reflection, refraction or diffraction of the radiated energy of objects that lie in the environment. The received signal is much weaker than the transmitted signal due to phenomena such as mean propagation loss, slow fading and fast fading. The mean propagation loss comes from square-law spreading, absorption by water and foliage and the effect of ground reflections. Mean propagation loss is range dependent and changes very slowly even for fast mobiles. Slow fading results from a blocking effect by buildings and natural features and is also known as long-term fading, or shadowing. Fast fading results from multipath scattering in the vicinity of the mobile. It is also known as short-term fading or Rayleigh fading, for reasons explained below. Multipath propagation results in the spreading of the signal in different dimensions. These are the delay (or time) spread, Doppler (or frequency) spread and angle spread (see Fig. 3.1). These spreads have significant effects on the signal. The mean path loss, slow fading, fast fading, Doppler, delay and angle spread are the main channel effects [14] and are described in the following sections.

## 3.1 Mean Path Loss

The mean path loss describes the attenuation of a radio signal in a free space propagation situation, due to isotropic power spreading, and is given by the famous inverse square low (or Friis free space link equation) [15]

$$P_r = P_t \left( \frac{\lambda}{4\pi d} \right)^2 G_t G_r, \tag{3.1}$$

where $P_r$ and $P_t$ are the received and transmitted powers, $\lambda$ is the radio wavelength, $d$ is the range and $G_t$ and $G_r$ are the gains of the transmit and receive antennas respectively. In cellular environments,

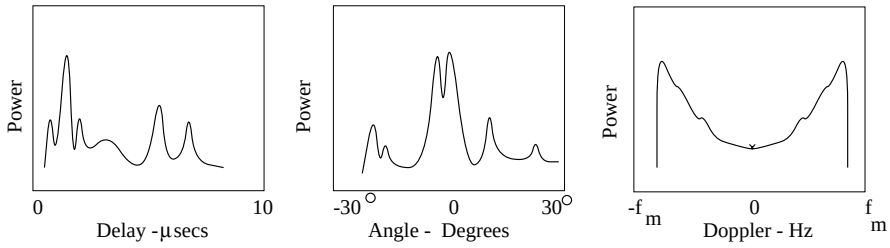



**Figure 3.1:** The radio channel induces spreading in several dimensions [14].

the main path is often accompanied by a surface reflected path which may destructively interfere with the primary path. Specific models have been developed that consider this effect and the path loss model can be given as

$$P_r = P_t \left( \frac{h_t h_r}{d^2} \right)^2 G_t G_r, \tag{3.2}$$

where $h_t$ and $h_r$ are the effective heights of the transmit and receive antennas respectively. Note that this particular path loss model follows an inverse fourth power law. In fact, depending on the environment, the path loss exponent may vary from 2.5 to 5.

## 3.2 Fading

In addition to path loss, the received signal exhibits fluctuations in signal level called fading. As these variations represent the change of the strength of the electrical field as a function of the distance from the transmitter, a mobile user will experience variation in time. The signal level of the continuous-time received signal — whose variations we can call signal fading — is typically composed of two multiplicative components, $\alpha_s$ and $\alpha_r$, as follows

$$\alpha(t) = \alpha_s(t)\alpha_r(t). \tag{3.3}$$

$\alpha_s(t)$ is called slow fading and represents the long-term time variations of the received signal, whereas $\alpha_r(t)$ represents the short-term (or multipath) fading. The slow fading $\alpha_s(t)$ is the envelope of the signal level $\alpha(t)$. We will explain how the fading affects the signal model later. In the following different types of fading will be explained.

### 3.2.1 Slow fading

Slow fading is caused by long-term shadowing effects of buildings or natural features in the terrain. It can also be described as the local mean of a fast fading signal (see below). The statistical distribution of the local mean has been studied experimentally and was shown to be influenced by the antenna height, the operating frequency and the type of environment. It is therefore difficult to predict. However, it has been observed that when all the above mentioned parameters are fixed, then the received signal power averaged over Rayleigh fading approaches a normal distribution when plotted in a logarithmic scale (i.e, in dB's). Such a distribution is called log-normal and it is described by the

following probability-density function

$$p(x) = \begin{cases} \frac{1}{\sqrt{\pi}\sigma x} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}, & x > 0 \\ 0, & x < 0 \end{cases} . \tag{3.4}$$

In the above equation, $x$ is a random variable representing the slow signal level fluctuation and $\mu$ and $\sigma$ are the mean and standard deviation of $x$ expressed in decibels, respectively. A typical value for the standard deviation of shadowing distribution is 8 dB.

### 3.2.2 Fast fading

The short-time fading $\alpha_r(t)$ corresponds to the rapid fluctuations of the received signal in space. It is caused by the scattering of the signal off objects near the moving mobile. If we assume that a large number of scattered wavefronts with random amplitudes and angles of arrival arrive at the receiver with phases uniformly distributed in [0, $2\pi$), then the in-phase and quadrature phase components of the vertical component of the electric field $E_z$ can be shown to be Gaussian processes. The envelope of the received signal has a Rayleigh density function given by

$$p(y) = \begin{cases} \frac{y}{\sigma^2} e^{-\frac{y^2}{2\sigma^2}}, & y > 0 \\ 0, & y < 0 \end{cases} . \tag{3.5}$$

If there is a direct path present, then it will no longer be a Rayleigh distribution but becomes a Rician distributed instead. The corresponding probability density function is given by

$$p(y) = \begin{cases} \frac{y}{\sigma^2} e^{-\frac{y^2+s^2}{2\sigma^2}} J_0\left(\frac{ys}{\sigma^2}\right), & y > 0 \\ 0, & y < 0 \end{cases} , \tag{3.6}$$

where $s^2$ is the mean power of the direct (line-of-sight) path and $J_p$ is the modified $p$-th order Bessel function of the first kind. Note that in the absence of a direct path ($s^2 = 0$) the Rician pdf reduces to the Rayleigh pdf.

## 3.3 Doppler Spread: Time-Selective Fading

When the mobile is in motion, the radio signal at the receiver experiences a shift in the frequency domain (also called Doppler shift), the amplitude of which depends on the path direction of arrival. In the presence of surrounding scatterers with multiple directions, a pure tone is spread over a finite spectral bandwidth. In this case, the Doppler power spectrum is defined as the Fourier transform of the time autocorrelation of the channel impulse response, and the Doppler spread is the support of the Doppler power spectrum. Assuming scatterers distributed uniformly n angle, the Doppler power spectrum is given by the so-called classical spectrum

$$S(f) = \frac{3\sigma^2}{2\pi f_m} \left[ 1 - \left( \frac{f - f_c}{f_m} \right)^2 \right]^{-1/2}, \quad f_c - f_m < f < f_c + f_m, \tag{3.7}$$

where $f_m = v/\lambda$ is the maximum Doppler shift, $v$ is the mobile velocity, $f_c$ is the carrier frequency and $\sigma^2$ is the signal variance.

When there is a dominant source of energy coming from a particular direction (as in line-of-sight situations) the expression for the spectrum needs to be corrected accordingly to the Doppler shift of the dominant path $f_D$

$$S(f) + B\delta(f - f_D), \tag{3.8}$$

where $B$ denotes the ratio of direct to scattered path energy.

The Doppler spread causes the channel characteristics to change rapidly in time, giving rise to the so-called time selectivity. The coherence time during which the fading channel can be considered as constant, is inversely proportional to the Doppler spread. A typical value of the Doppler spread in a macro-cell environment is about 175 Hz at 100 km/h in the 1900 MHz frequency band. A large Doppler spread makes good channel tracking as an essential feature of the receiver design.

## 3.4 Delay Spread: Frequency-Selective Fading

Multipath propagation is often characterized by several versions of the transmitted signal arriving at the receiver with different attenuation factors and delays. The spreading in the time domain is called delay spread and is responsible for the selectivity of the channel in the frequency domain (different spectral components of the signal carry different powers). The coherence bandwidth, which is the maximum range of frequencies over which the channel response can be viewed as constant (or the maximum frequency separation for which the frequency domain channel responses at two frequency shifts remain strongly correlated), is inversely proportional to the delay spread. Significant delay spread may cause strong inter-symbol interference which makes necessary the use of a channel equalizer.

## 3.5 Angle Spread: Space-Selective Fading

Angle spread at the receivers refers to the spread of angles of arrival of the multipaths at the antenna array. Likewise, angle spread at the transmitter refers to the spread of departure angles of the multipaths. The angle of arrival (or departure) of a path can be, in some cases, statistically related to the path delay. Angle spread causes space-selective fading, which means that signal amplitude depends on the spatial location of the antenna. Space-selective fading is characterized by the coherence distance. The larger the angle spread, the shorter the coherence distance. Once again, coherence distance represents the maximum spatial separation for which the channel responses at two antennas remain strongly correlated.

## 3.6 Multipath Propagation

Multipath scattering underlies the three spreading effects described above (mobile motion is also required to produce Doppler spread). It is important to understand the types of scatterers and their contribution to channel behaviour.

### 3.6.1 Macro-cells

A macro-cell is characterized by a large cell radius (up to a few tens of kilometers) and a base station located above the roof top. In macro-cell environments, the signal energy received at the base station comes from three main scattering sources: scatterers local to the mobile, remote dominant scatterers

and scatterers local to the base (see Fig. 3.2 for illustration). The following description refers to the reverse link but applies to the forward link as well.

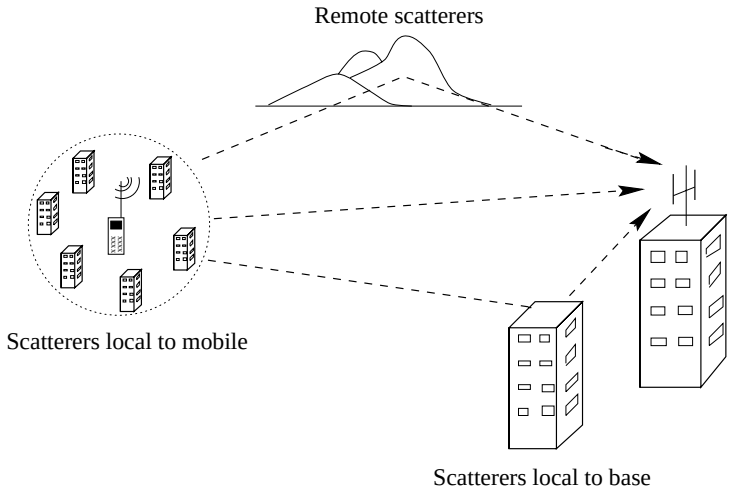



**Figure 3.2:** Each type of scatterer introduces specific channel spreading characteristics.

The scatterers local to the mobile are those located a few tens of meters from the hand-held terminal. When the terminal is in motion, these scatterers give rise a Doppler spread which causes time selective fading. Because of the small scattering radius, the paths that emerge from the vicinity of the mobile and reach the base station show a small delay spread and a small angle spread.

Of the paths emerging from the local-to-mobile scatterers, some reach remote dominant scatterers, like hills or high rise buildings, before eventually traveling to the base station. These paths will typically reach the base with medium to large angle and delay spreads (depending of course on the number and locations of these remote scatterers.

Once these multiple wavefronts reach the vicinity of the base station, they are usually further scattered by local structures such as buildings or other structures that are close to the base. These scatterers local to the base can cause large angle spread, therefore they can cause severe space-selective fading.

### 3.6.2 Micro-cells

Micro-cells are characterized by highly dense built-in areas, and by the user's terminal and base being relatively close (a few hundred meters). The base antenna has a low elevation and is typically below the roof top, causing significant scattering in the vicinity of the base. Micro-cell situations make the propagation difficult to analyze and the macro-cell model described earlier do not necessarily hold anymore. Very high angle spreads along with small delay spreads are likely to occur in this situations The Doppler spread can be as high as in macro-cells, although the mobility of the user is expected to be limited, due to the presence of mobile scatterers.

### 3.6.3 Typical channel parameters

The number of paths present in a cellular environment will depend both on the transmission bandwidth as well as on the carrier frequency. Measurements in macrocells indicate that in a GSM system (200 kHz bandwidth, 900 MHz carrier frequency) up to 6 to 12 paths may be present. The GSM system is characterized by a very short symbol period (3.7 $\mu$s), a short time slot (0.577 ms) and a high channel bandwidth. Since the delay spread in hilly terrain and urban areas can be much larger than the symbol period, severe ISI will be present and, hence, the channel is highly frequency selective. On the other hand, as the time slot is short, we do not expect significant channel variations during the time slot.

In the IS-136 (American TDMA standard for mobile communications) system on the other hand, the situation is reversed: the symbol period is 41.6 $\mu$s, the time slot is 6.66 ms and the bandwidth is much smaller (30 kHz). We therefore have negligible ISI as the symbol period is large compared to the delay spread and frequency selectivity of the channel is low. For high Doppler spread, the coherence time is smaller than the time-slot duration, indicating significant channel variation within the slot.

Typical channel delay, angle, and one-sidded Doppler (at 1800 MHz carrier frequency) spreads are given in Table 3.1.

| Environment | Delay Spread | Angle Spread | Doppler Spread |
|---|---|---|---|
| Flat Rural (Macro) | 0.5 $\mu$s | 1 deg | 190 Hz |
| Urban (Macro) | 5 $\mu$s | 20 deg | 120 Hz |
| Hilly (Macro) | 20 $\mu$s | 30 deg | 190 Hz |
| Microcell (Mall) | 0.3 $\mu$s | 120 deg | 10 Hz |
| Picocell (Indoors) | 0.1 $\mu$s | 360 deg | 5 Hz |

**Table 3.1:** Typical delay, angle and Doppler spreads in cellular radio systems [15].

## 3.7 Parametric Channel Model

A complete and accurate understanding of propagation effects in the radio channel requires a detailed description of the physical environment. The *specular model*, to be presented below, only provides a simplified description of the physical reality. However, it is useful as it describes the main channel effects and it provides the means for a simple and efficient mathematical treatment. In this model, the multiple elementary paths are grouped according to a (typically low) number of $L$ main path clusters, each of which contains paths that have roughly the same mean angle and delay. Since the paths in these clusters originate from different scatterers, the clusters typically have near independent fading. Based on this model, the continuous time channel response from a single transmit antenna to the $i$-th antenna of the receiver can be written as

$$f_i(t) = \sum_{l=1}^{L} a_i(\theta_l)\alpha_l^R(t)\delta(t-\tau_l), \tag{3.9}$$

where $\alpha_l^R(t)$, $\theta_l$ and $\tau_l$ are respectively, the fading (including mean path loss, slow and fast fading), angle, and delay of the $l$-th receive path cluster. Note that this model also includes the response of

the $i$-th antenna to a path from direction $\theta_l$, denoted by $a_i(\theta_l)$. Note that in the situation where the path cluster assumption is not acceptable, other channel models, called *diffuse* channel models are more appropriate.

# Chapter 4

# Signal Model for TDMA

Having seen the channel characteristics and having defined the channel model in the previous chapter, we can now define a signal model for space-time processing applications. We assume that antenna arrays are used at the base station only. We therefore assume that the mobile has a single antenna. This leads to four interesting channel configurations [14]

1. Reverse Link
   (a) SU-SIMO: Single user (SU) with single antenna input (SI) at mobile and multiple antenna output (MO) at the base station
   (b) MU-SIMO: Multi user (MU) with single antenna input (SI) at each mobile and multiple antenna composite output (MO) at the base station
2. Forward link
   (a) SU-MISO: Single user with multiple antenna input (MI) at the base station and single antenna output (SO) at the mobile unit
   (b) MU-MISO: Multiuser (MU) with multiple antenna composite input (MI) at the base and single antenna output (SO) at each mobile

Before we begin developing the signal models, some general remarks on the channel are in order. An important issue to be addressed is the principle of reciprocity. This principle implies that the channel is identical on the forward and reverse links as long as the channel is measured at the same frequency and at the same time instant. In time-division duplexing (TDD) systems, the principle of reciprocity applies as long as the "ping-pong" time (the time separating a receive from a transmit frame) is very small compared to the channel coherence time. In frequency-division duplexing (FDD) systems (most wireless systems are FDD), the separation between the forward and reverse-link frequencies is about 5 % of the main frequency. This means that the principle of reciprocity cannot be used. However, given the small frequency separation, the forward and reverse channel will share many common features. In a specular multipath channel as described in the previous chapter, the paths used by both links can be assumed to be identical. Therefore the number of paths, the path delays, and path angles (arrival/departure) are the same for both links. However, the path amplitudes and phases will not be the same on both links and they will in fact be uncorrelated. Also within any one link, the different

paths fade independently. Since fading appears as a multiplicative gain, the channel appears to be uncorrelated between the forward and reverse links.

In this chapter we will study the signal models for the four cases mentioned above. We focus on nonspread modulation techniques used in TDMA systems. CDMA signal models will be discussed later.

## 4.1 Reverse Link SU-SIMO

This model refers to the case when a single user transmits the information signal and it is received at a base station with multiple antennas. The baseband signal, $x_i(t)$ received by the base station at the $i$-th element of an $M$ element antenna array is given by

$$x_i(t) = \sum_{l=1}^{L} a_i(\theta_l)\alpha_l^R(t)u(t-\tau_l) + n_i(t), \tag{4.1}$$

where $L$ is the number of multipaths, $a_i(\theta_l)$ is the response of the $i$-th element to the $l$-th path from direction $\theta_l$, $\alpha_l^R(t)$ is the complex path fading, $\tau_l$ is the path delay, $n_i(t)$ is the additive noise present in the channel and $u(\cdot)$ is the transmitted signal that depends on the modulation waveform and the information data stream. In the IS-54 TDMA standard, $u(\cdot)$ is a $\pi/4$ shifted differential quadrature phase-shift keying (DQPSK), gray-coded signal that is modulated using a pulse with square-root raised cosine spectrum with excess of bandwidth of 0.35. In GSM, a Gaussian minimum-shift keying (GMSK) modulation is used.

For linear modulation (i.e., binary phase-shift keying, BPSK), the transmitted signal has the following form in baseband

$$u(t) = \sum_k g(t-kT)s(k), \tag{4.2}$$

where $g(\cdot)$ is the pulse shaping waveform, $T$ is a single period and $s(k)$ represents the information bits. Note that in the case of a non linear modulation scheme such as the GMSK used in GSM system, linear approximations are assumed to hold.

In the above model we have assumed that the inverse signal bandwidth is large compared to the travel time across the array. For example, in GSM the inverse signal bandwidth is 5 $\mu$s, whereas the travel time across the array is at most a few ns. This is widely referred to as the "narrowband" assumption in array processing. The signal bandwidth is a sum of the modulation bandwidth and the Doppler spread, with the latter being comparatively negligible. Therefore, the complex envelopes of the signals received by different antennas from a given path are identical except for phase and amplitude differences that depend on the path angle-of-arrival, array geometry and the element pattern. This angle-of-arrival dependent phase and amplitude response at the $i$-th element is $a_i(\theta_l)$.

We collect all the element responses to a path arriving from angle $\theta_l$ into an $M$-dimensional vector, called the array response vector defined as

$$\mathbf{a}(\theta_l) = [a_1(\theta_l)\ a_2(\theta_l)\ \ldots\ a_M(\theta_l)]^{\mathrm{T}}, \tag{4.3}$$

where superscript T denotes transpose of a vector or matrix. In array-processing literature the array vector $\mathbf{a}(\theta)$ is also known as the steering vector (Chapter 2).

We can rewrite the array output at the base station as

$$\mathbf{x}(t) = \sum_{l=1}^{L} \mathbf{a}(\theta_l)\alpha_l^R(t)u(t-\tau_l) + \mathbf{n}(t), \tag{4.4}$$

where

$$\mathbf{x}(t) = [x_1(t)\ x_2(t)\ \ldots\ x_M(t)]^{\mathrm{T}},$$
$$\mathbf{n}(t) = [n_1(t)\ n_2(t)\ \ldots\ n_M(t)]^{\mathrm{T}}$$

and $\mathbf{x}(t)$, $\mathbf{a}(\theta_l)$ and $\mathbf{n}(t)$ are $M$-dimensional complex vectors. The fading, $|\alpha_l^R(\cdot)|$, is Rayleigh or Rician distributed depending on the propagation model.

## 4.2 Reverse Link MU-SIMO

In this model, multiple mobiles transmit their information signals and they arrive at a base station that uses an antenna array to separate individual signals. The MU-SIMO can be easily obtained from the SU-SIMO model. Assuming there are $Q$ users, the composite received signal at the antenna array is a sum of the signals from $Q$ mobiles

$$\mathbf{x}(t) = \sum_{q=1}^{Q}\sum_{l=1}^{L} \mathbf{a}(\theta_{lq})\alpha_{lq}^R(t)u_q(t-\tau_{lq}) + \mathbf{n}(t), \tag{4.5}$$

where we have indexed each user signal, and corresponding path delay, angle and fading parameters by the user index $q$.

## 4.3 Forward Link SU-MISO

In this model, the base station uses an antenna array to transmit an information signal to a single mobile, In the forward link the paths that couple the signal to mobile will be the same as the paths available in the reverse link. However, the space-time processing of user signals is carried out at the transmitter before the signal is launched into the channel. Due to the effect of base-station antennas and transmit processing, the radiated signal will be directional. Therefore, the transmit beam pattern will selectively weight the energy coupled into each of the multipaths and in the extreme, some paths may not be excited at all. If we use space-only processing with a beamforming weight vector $\mathbf{w}$, the transmitted signal in each path will remain $u(t)$. Then the signal received by the mobile will be the sum of different path signals.

With the simplifying assumption of space-only processing, the received baseband signal at the mobile station, $x(t)$, is given by

$$x(t) = \sum_{l=1}^{L} \mathbf{w}^{\mathrm{H}}\mathbf{a}(\theta_l)\alpha_l^F(t)u(t-\tau_l) + n(t), \tag{4.6}$$

where $\mathbf{w}$ is the transmit weight vector that influences how the transmitted signal couples into the channel and superscript H denotes complex conjugate transpose. The path delay $\tau_l$ and angle parameters $\theta_l$ are the same as those of the reverse link. $\alpha_l^F(t)$ is complex fading on the forward link which in (fast ping-pong time) TDD systems will be identical to the reverse-link complex fading, $\alpha_l^R(t)$. In an FDD system $\alpha_l^F(t)$ and $\alpha_l^R(t)$ will have the same statistics but will generally be uncorrelated with each other.

If space-time processing is used in transmit, the delivered signal will have the form

$$x(t) = \sum_i \mathbf{w}_i^{\mathrm{H}} \mathbf{x}(t - t_i), \tag{4.7}$$

where $\{t_i\}$ is the set of time delays used for temporal processing. Hence, the signal model for this case becomes

$$x(t) = \sum_i \mathbf{w}_i^{\mathrm{H}} \sum_{l=1}^{L} \mathbf{a}(\theta_l) \alpha_l^F (t - t_i) u(t - \tau_l - t_i) + n(t). \tag{4.8}$$

## 4.4 Forward Link MU-MISO

In the multi-user case, the base station wishes to communicate with $Q$ users, simultaneously and in the same frequency band. This can be done by superposing, on each of the transmit antennas, the signals given by $Q$ beamformers $\mathbf{w}_1, \ldots, \mathbf{w}_Q$. At the $m$-th user, the received signal waveform contains the signal sent to that user, plus an interference from signals intended to all other users. This gives

$$x_m(t) = \sum_{q=1}^{Q} \mathbf{w}_q \sum_{l=1}^{L_q} \mathbf{a}(\theta_{lq}) \alpha_{lq}^F(t) u_q(t - t_{lq}) + n_m(t). \tag{4.9}$$

Note that each information signal $u_q(t)$ couples into the $L_q$ paths of the of the $q$-th user through the corresponding weight vector $\mathbf{w}_q$, for all $q$.

As in SU-MISO, the forward-link path parameters are related to the reverse-link path parameters. If space-time processing is used, the corresponding signal model can be easily derived from (4.8).

## 4.5 Discrete-Time Signal Model

The channel model described above uses physical path parameters such as path gain, delay, and angle of arrival. When the received signal is sampled at the receiver at symbol (or higher) rate, such a model is inconvenient to use. For linear modulation schemes, it is more convenient to use a "symbol response" channel model.

Such a discrete-time signal model can be easily obtained as follows. We address the SU-SIMO case, but the approach can be easily extended to the other cases. Let the continuous-time output from the receive antenna array, $x(t)$, be sampled at the symbol rate at instants $t = t_0 + kT$. Then the output at time instant $k$ may be written as

$$\mathbf{x}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{n}(k), \tag{4.10}$$

where $\mathbf{H}$ is the symbol response channel (an $M \times P$ matrix) that captures the effects of the array response, symbol waveform and path fading. $M$ is the number of antennas, $P$ is the channel length in symbol periods and $\mathbf{n}(k)$ is the sampled vector of additive noise. Note that $\mathbf{n}(k)$ may be colored in space and time, as will be shown later. $\mathbf{H}$ is assumed to be time invariant, that is $\alpha^R$ is constant. $\mathbf{s}(k)$ is a vector of $P$ consecutive elements of the data sequence and is defined as

$$\mathbf{s}(k) = \begin{bmatrix} s(k) \\ \vdots \\ s(k - P + 1) \end{bmatrix}. \tag{4.11}$$

It can be shown that the element at the position $(i, j)$ in matrix $\mathbf{H}$ is given by

$$[\mathbf{H}]_{ij} = \sum_{l=1}^{L} a_i(\theta_l)\alpha_l^R g((M_d + \triangle - j)T - \tau_l), \quad i = 1, \ldots, M,\ j = 1, \ldots, N, \tag{4.12}$$

where $M_d$ is the maximum integer path delay and $2\triangle T$ is the duration of the pulse-shaping waveform $g(t)$ (see (4.2)). The last equation can be written in a more compact form as

$$\mathbf{H} = \sum_{l=1}^{L} \mathbf{a}(\theta_l)\alpha_l^R \mathbf{g}^{\mathrm{T}}(\tau_l), \tag{4.13}$$

where

$$\mathbf{g}(\tau_l) = [g((M_d + \triangle - 1)T - \tau_l) \ldots g((M_d + \triangle - N)T - \tau_l)]^{\mathrm{T}}. \tag{4.14}$$

The signal model in (4.10) is simple but rich and allows the application of many techniques developed in other contexts. Note that (4.12) indicates that, given the multipath parameters, we can easily find $\mathbf{H}$. On the other hand, determining the multipath parameters given $\mathbf{H}$ is more complex task.

Note, once again, that GSM uses GMSK modulation where the transmitted signal is not a linear map of the underlying data sequence. In such a case, the signal model in (4.10) cannot be strictly used. In practice, a reasonable linear approximation of GMSK is possible and the model described above remains applicable.

To account for the presence of CCI, (4.10) can be generalized to

$$\mathbf{x}(k) = \sum_{q=1}^{Q} \mathbf{H}_q \mathbf{s}_q(k) + \mathbf{n}(k), \tag{4.15}$$

where $Q$ denotes the number of users and $q$ the user index.

Now, if we collect $N$ consecutive shapshots of $\mathbf{x}(\cdot)$ corresponding to time instants $k, \ldots, k + N - 1$, and neglect for a moment the intereference, we get

$$\mathbf{X}(k) = \mathbf{H}\mathbf{S}(k) + \mathbf{N}(k), \tag{4.16}$$

where $\mathbf{X}(k)$, $\mathbf{S}(k)$ and $\mathbf{N}(k)$ are defined as

$$\begin{aligned} \mathbf{X}(k) &= [\mathbf{x}(k)\ \ldots\ \mathbf{x}(k + N - 1)] \qquad (M \times N), \\ \mathbf{S}(k) &= [\mathbf{s}(k)\ \ldots\ \mathbf{s}(k + N - 1)] \qquad (P \times N), \\ \mathbf{N}(k) &= [\mathbf{n}(k)\ \ldots\ \mathbf{n}(k + N - 1)] \qquad (M \times N). \end{aligned}$$

Note that $\mathbf{S}(k)$ is a Toeplitz data matrix (has the same elements along diagonals).

(4.16) is the general linear data model that will be used when Space-Time processing algorithms are to be introduced. It expresses the received signals as a convolution of the transmitted data and the channel.

In space-time filtering, a space-time filter or equalizer $\mathbf{W}$ is used and it can be represented as

$$\mathbf{W}(k) = \begin{bmatrix} w_{11}(k) & \ldots & w_{1N}(k) \\ \vdots & \ldots & \vdots \\ w_{M1}(k) & \ldots & w_{MN}(k) \end{bmatrix} = [\mathbf{w}_1(k)\ \ldots\ \mathbf{w}_N(k)]. \tag{4.17}$$

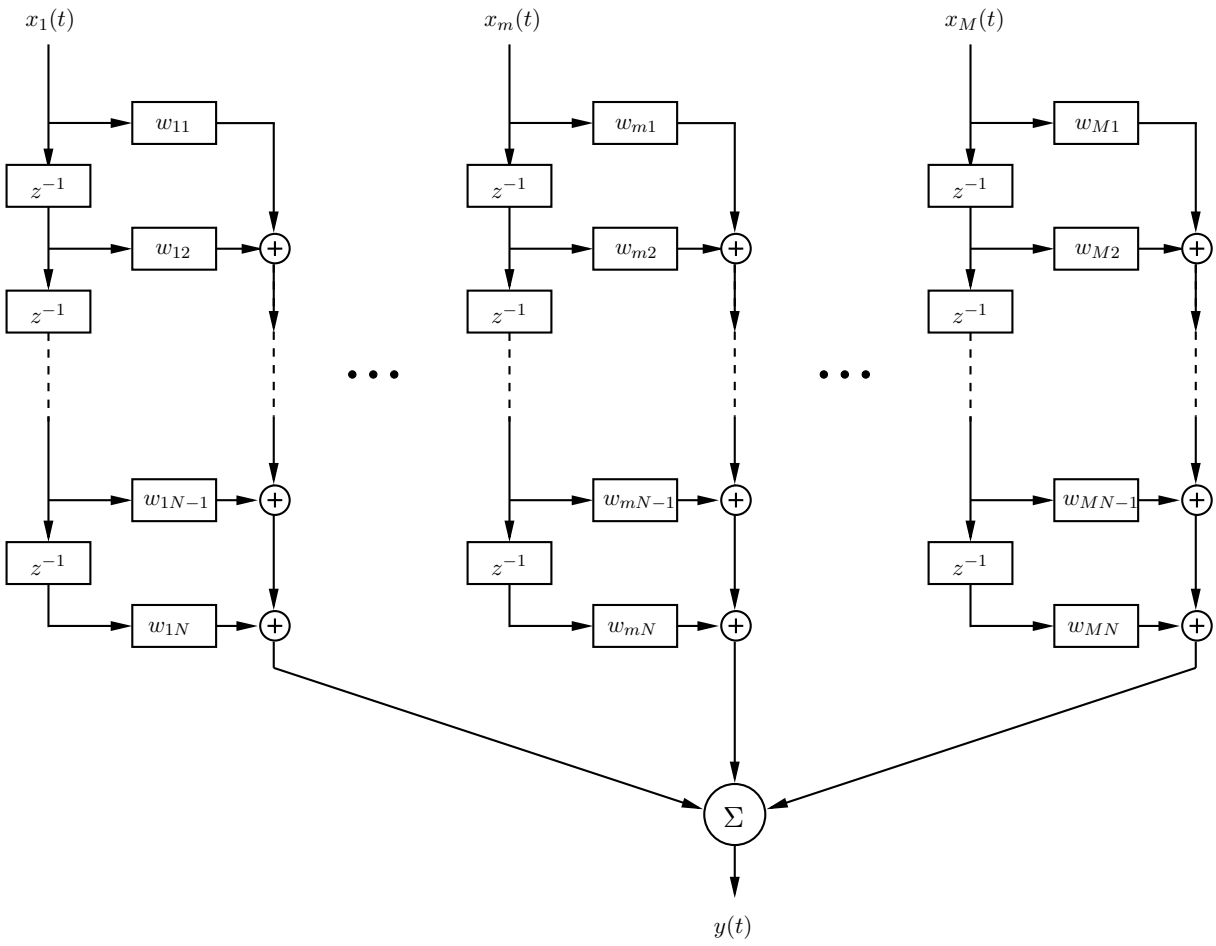



**Figure 4.1:** Structure of Space Time Beamformer.

Every row of the matrix $\mathbf{W}(k)$ corresponds to an element in an array for exits of a different tapped delay lines being at the same level (Fig. 4.1).

In order to obtain a convenient formulation for the ST filter output, we introduce the quantities $W(k)$ and $X(k)$ as follows

$$X(k) = \text{vec}(\mathbf{X}(k)) \qquad (MN \times 1), \tag{4.18a}$$

$$W(k) = \text{vec}(\mathbf{W}(k)) \qquad (MN \times 1), \tag{4.18b}$$

where the operator vec$(\cdot)$ is defined as

$$\text{vec}([\mathbf{v}_1 \ \ldots \ \mathbf{v}_M]) = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_M \end{bmatrix}. \tag{4.19}$$

The scalar equalizer output $y(k)$ can be written as

$$y(k) = W^H(k)X(k) = \text{trace}(\mathbf{W}^{\text{H}}(k)\mathbf{X}(k)). \tag{4.20}$$

Having described the signal and interference model, we can now study how the output of the antenna array can be processed to maximize signal demodulation performance.

# Chapter 5

# Overview of TDMA Adaptive Processing Methods

## 5.1 Digital Beam Forming with TDMA

Digital technology makes TDMA a practical access technology. In a TDMA system each user is apportioned a channel periodically for a brief period of time. The users' transmissions are therefore intermittent in nature, a condition that can only be accommodated by a digital transmitter that can store its source bits and then send them out at a transmission speed higher than that at which they are generated.

The DBF configurations for a base station system using TDMA in both the forward and reverse link cases are shown in Fig. 5.1 and Fig. 5.2, respectively. The reverse (forward) link system consists of [16]

1. An $M$-element antenna array
2. $M$ receiver (transmitter) modules
3. $M$ ADCs (DACs)
4. An RDBF (TDBF) network
5. $L$ digital demodulators (modulators)
6. $L$ digital demultiplexers (multiplexers)

In the reverse link case, beamforming is applied to the received signals from the $M$ antenna elements to produce $L$ reverse link beam outputs. Each beam output from the beamforming network will be demodulated and time-demultiplexed into $N$ message signals. That is, the number of simultaneous users can be up to $N \times L$. In the forward link case, the $N \times L$ message signals to be transmitted are arranged into $L$ groups. The $N$ messages in each group are time-multiplexed into a binary stream. A particular digital modulation scheme is then applied to the time-division multiplex (TDM) digital signal. Digital beamforming is applied to the $L$ modulated signals to form $L$ forward link beams. The digital outputs of the beamforming network are converted to analog signals for transmission.

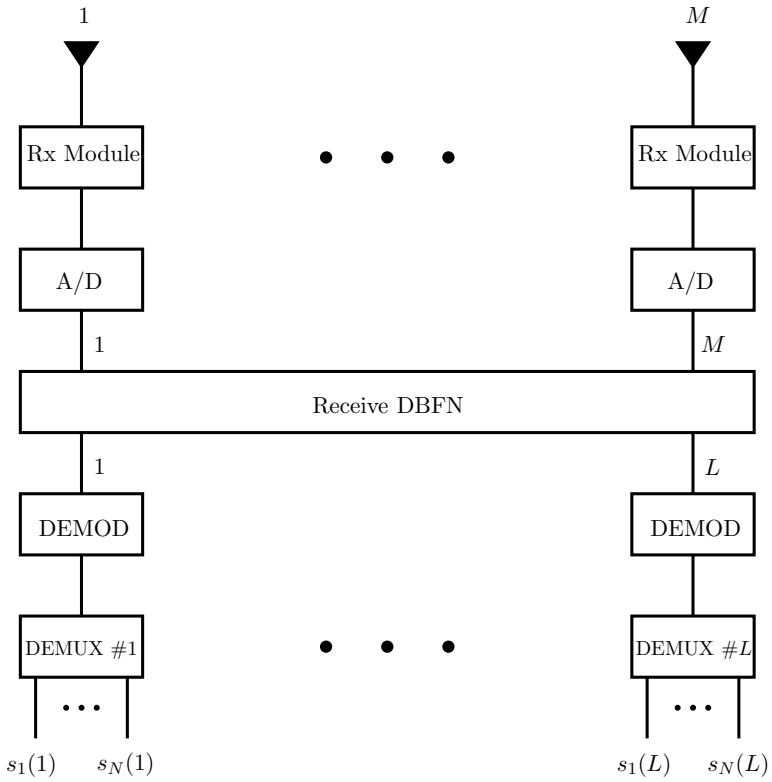



**Figure 5.1:** Reverse link DBF configuration for a TDMA system.

The digital beamforming network must process the signals sequentially (that is, frame by frame). In the forward link case, for example, within a TDMA frame $t_1$, the beamforming network produces $L$ beams for the $L$ message signals $\{s_1(1), s_1(2), \ldots, s_1(L)\}$, within another time frame $t_2$ it produces another $L$ beams for the message signal $\{s_2(1), s_2(2), \ldots, s_2(L)\}$ and so on. There will be problems if the processing time $t_p$ required by the beamforming network is longer than the TDMA time frame $t_i$. Therefore, it requires a very fast digital beamforming network. An alternative to using very high speed beamforming network is to use a bank of digital beamforming networks, which operate at relatively low speed but in a near-parallel fashion. As shown in Fig. 5.3 at the beginning of TDMA frame $t_i$, the upper switch is at position 1 to load signals $\{s_i(1), s_i(2), \ldots, s_i(L)\}$ into the prebuffer #1. At the beginning of TDMA frame $t_{i+1}$, the switch is at position 2 to load signals $\{s_{i+1}(1), s_{i+1}(2), \ldots, s_{i+1}(L)\}$ into the prebuffer #2 and so on. On the other hand, the lower switch is placed at position 1 after $t_p$ with respect to the beginning of TDMA frame $t_i$ to send the $M$ processed signals to the DAC. The lower switch is placed at position 2 after $t_p$ with respect to the beginning of TDMA frame $t_{i+1}$, and so on. The implementation of this scheme requires additional hardware such as buffers and additional software such as synchronization control.

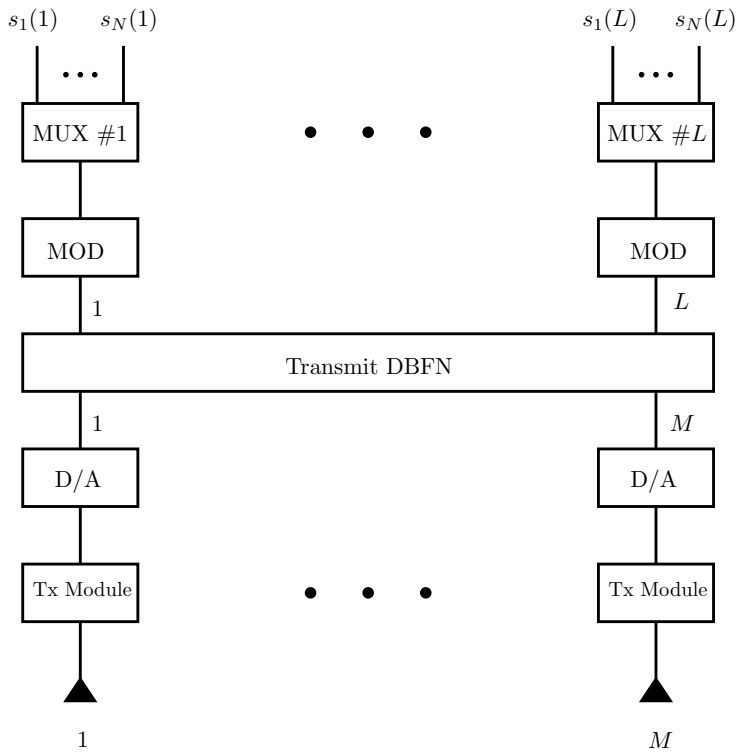



**Figure 5.2:** Forward link DBF configuration for a TDMA system.

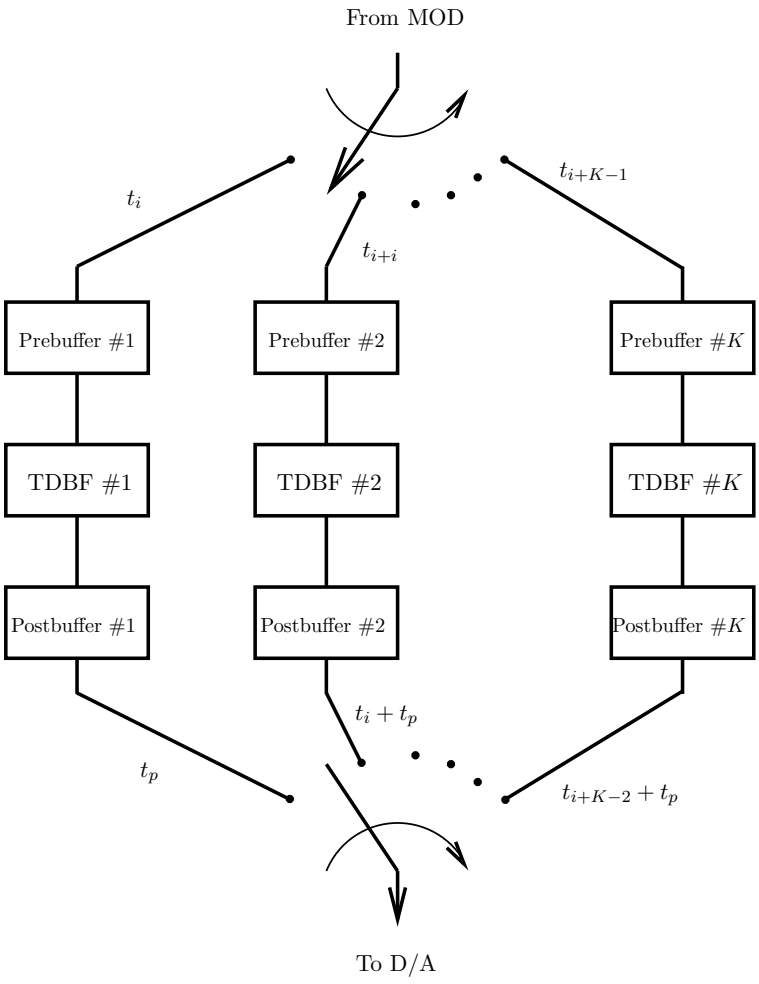



**Figure 5.3:** Multiple digital beamforming networks for TDMA applications. Signal flow structure.

## 5.2 Reverse Link Processing

The reverse link (or uplink) is the communication link from the mobile user to the base-station. It is assumed that a smart antenna is only employed at the base station and not in the mobile radio. Therefore it is the base station task to employ spatially selective reception in order to separate the desired signals from interferences. This task is called reverse link (or uplink) processing [17].

There are many varieties of smart antenna algorithms and structures proposed for the cellular applications in the literature. In this chapter a classification of smart antenna algorithms and receivers for narrowband signal processing is given in order to provide a comprehensive general picture for orientation in the large number of proposed methods and technical solutions. Smart antenna receivers can be classified as smart antennas with processing in the space domain only and space-time (ST) smart antenna receivers with processing in space and time domains simultaneously (see Fig. 5.4).

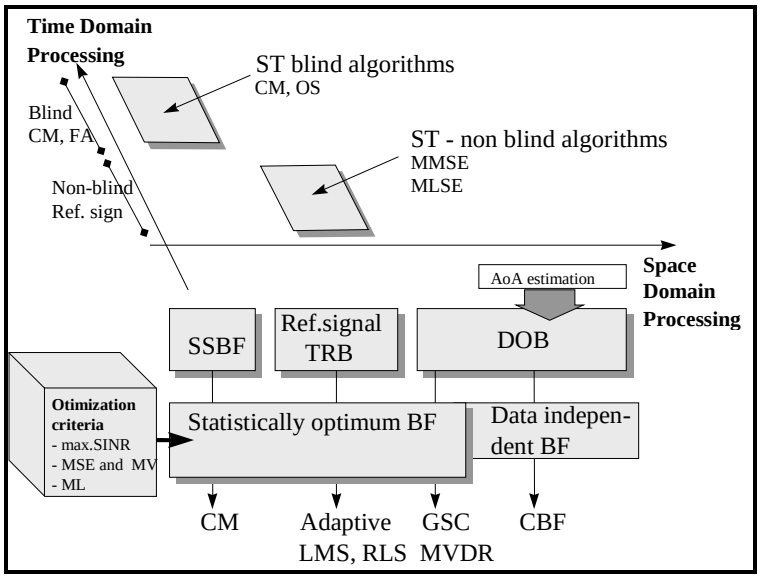



**Figure 5.4:** SA receiver classification [18].

Many of spatial domain only and ST algorithms include optimization procedures in their structure. The most popular optimization criteria are: maximum signal to interference plus noise ration (SINR), squared function based criteria such as minimum square error (MSE), minimum variance (MV) and maximum likelihood (ML).

According to the side information used for beamforming, the appropriate signal processing methods can be classified into three main categories

1. **Direction of Arrival Based Methods**. Methods based on spatial structure, that is, angle of arrival (AoA) of the incoming signal
2. **Training Signal Based Methods** that use a known *training signal* or code to achieve the signal processing objective
3. **Signal Structure Based Methods**, which exploit the temporal and/or spectral properties of the received signals.

The methods available in each of the three main categories mentioned above can be again grouped into

1. **Conventional methods** that use only the received data to achieve the desired signal processing objective (the structure of the channel/data is not used explicitly)

2. **Parametric methods** that use both, the received data and knowledge of the channel/data model to achieve the desired signal processing objective.

Because parametric methods exploit the knowledge of the underlying model, their performance depends strongly on the validity of the model itself. However, if the model is valid, then the parametric methods easily outperform the conventional methods. Most modern signal processing methods are parametric.

### 5.2.1 Direction of Arrival Based Methods

As mentioned before, DoA methods exploit the information in the *steering* vector $\mathbf{a}(\theta)$, which is used to estimate the direction of arrivals of the signals impinging on the sensor array. The estimated direction of arrivals are then used to determine the weights in the beamforming network. In the following, an overview will be given about the three main direction of arrival based methods, namely, conventional beamforming methods (based on optimization techniques as mean square error, signal to noise ratio), maximum likelihood estimation and the so-called subspace-based methods (which include MUSIC and ESPRIT).

#### Conventional methods

The first attempt to automatically localize signal sources using antenna arrays was through beamforming methods. The idea is to steer the array in one direction at a time and measure the output power. The steering locations which result in maximum power yield the DoA estimates.

Let the signal model be given by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{5.1}$$

where $\mathbf{A}$ is the $M \times L$ matrix whose columns are the array response vectors for each wavefront (assuming no delay spread):

$$\mathbf{A} = [\mathbf{a}(\theta_1) \ \ldots \ \mathbf{a}(\theta_L)],$$

$\mathbf{s}(t)$ contains the $L$ fading signals

$$\mathbf{s}(t) = [\alpha_1(t)u_1(t-\tau_1) \ \ldots \ \alpha_L(t)u_L(t-\tau_L)]^{\mathrm{T}}$$

and

$$u_l(t) = \sum_k s_l(k)g(t-kT).$$

The output of the smart antenna is given by

$$y(t) = \mathbf{w}^{\mathrm{H}}\mathbf{x}(t). \tag{5.2}$$

Given $N$ time samples of the signal $y(1), \ldots, y(N)$ the output power can be measured by

$$P(\mathbf{w}) = \frac{1}{N}\sum_{k=1}^{N}|y(k)|^2 = \frac{1}{N}\sum_{k=1}^{N}\mathbf{w}^{\mathrm{H}}\mathbf{x}(k)\mathbf{x}^{\mathrm{H}}(k)\mathbf{w} = \mathbf{w}^{\mathrm{H}}\hat{\mathbf{R}}_{xx}\mathbf{w}, \tag{5.3}$$

where $\hat{\mathbf{R}}_{xx}$ is a natural estimate of the covariance matrix

$$\hat{\mathbf{R}}_{xx} = \frac{1}{N}\sum_{k=1}^{N}\mathbf{x}(k)\mathbf{x}^{\mathrm{H}}(k).$$

Now, different beamforming approaches correspond to different choices of the weighting vector. An excellent review of beamforming methods is given in [19, 20].

A simple and widely used approach is the Least Mean Square Error (MSE) performance measure, which is formulated as

$$\min_{w} E\left\{\left(d(k) - \mathbf{w}^{\mathrm{H}}x(k)\right)^2\right\}, \tag{5.4}$$

where $d(k)$ is the desired response of the smart antenna output and $E\{\cdot\}$ the mathematical expectation. The solution to the above stated minimization problem is the well known Wiener-Hopf solution and is given by

$$\mathbf{w} = \mathbf{R}_{xx}^{-1}\mathbf{r}_{xd}, \tag{5.5}$$

where the crosscorelation vector $\mathbf{r}_{xd}$ is defined as

$$\mathbf{r}_{xd} = E\{\mathbf{x}(k)d^*(k)\} = \mathbf{a}(\theta)E\{s(k)d^*(k)\}. \tag{5.6}$$

Of course, ideally the desired signal is $s(k)$. Setting $d(k) = s(k)$, the above equation becomes

$$\mathbf{r}_{xd} = \mathbf{a}(\theta)E\{s(k)s^*(k)\} = \alpha\mathbf{a}(\theta). \tag{5.7}$$

The constant $\alpha$ is the power of the transmitted signal $s(k)$, but basically it just scales the output of the smart antenna. Setting $\alpha = 1$, the solution for the beamforming network weights is simply

$$\mathbf{w} = \mathbf{R}_{xx}^{-1}\mathbf{a}(\theta). \tag{5.8}$$

Inserting equation (5.8) into (5.3) and using the autocorrelation matrix estimate $\hat{\mathbf{R}}_{xx}$ the classical spatial spectrum is obtained as

$$P(\theta) = \mathbf{a}^{\mathrm{H}}(\theta)\hat{\mathbf{R}}_{xx}\mathbf{a}(\theta). \tag{5.9}$$

MSE algorithm does not require squaring, averaging or differentiation and hence can be implemented in most practical systems. However, the drawback is that the weights for this algorithm take a long time to converge. Other choices for the weight vector $\mathbf{w}$ are possible and are based on other performance measures such as

- Signal to noise ratio (SNR)
- Maximum noise variance (MNV)

However, for conventional methods, the solutions are all basically the same. The conventional methods obtain $P(\theta)$ as the spatial analogue of the classical periodogram in temporal time series analysis. The classical periodogram suffers from the fact that its standard deviation is approximately as large as the quantity to be estimated. Therefore in general it can be said that the resolution of these methods is poor because it is simply an extension of classical Fourier-based spectral analysis to sensor array data.

**Maximum Likelihood (ML) Method**

Let the noise is modeled as a stationary Gaussian white random process (mean value being zero and standard deviation being $\sigma$) and signal waveforms are deterministic (arbitrary) and unknown. As a consequence of these assumptions, the observation vector $\mathbf{x}(t)$ is going to be white Gaussian random process with mean $\mathbf{A}(\theta)\mathbf{s}(t)$ and covariance matrix $\sigma^2\mathbf{I}$. The likelihood function is the probability density function (pdf) of all observations given the unknown parameters. The pdf of one measurement vector $\mathbf{x}(t)$ of the dimension $M$ is the complex $M$-variate Gaussian

$$\frac{1}{(\pi\sigma^2)^M} e^{-\|\mathbf{x}(t)-\mathbf{A}\mathbf{s}(t)\|^2/\sigma^2}, \tag{5.10}$$

where $\|\cdot\|$ denotes Euclidean norm. Since the measurements are independent, given $N$ samples, the likelihood function is obtained as

$$L(\theta, \mathbf{s}(t), \sigma^2) = \prod_{k=1}^{N} (\pi\sigma^2)^{-M} e^{-\|\mathbf{x}(k)-\mathbf{A}\mathbf{s}(k)\|^2/\sigma^2}. \tag{5.11}$$

As indicated above, the unknown parameters in the likelihood function are the signal parameters $\theta$, the signal waveforms $\mathbf{s}(t)$ and the noise variance $\sigma^2$. The ML estimates of these unknowns are calculated as the maximizing arguments of $L(\theta, \mathbf{s}(t), \sigma^2)$, the rationale being that these values make the probability of the observations as large as possible. For convenience, the ML estimates are alternatively defined as the minimizing arguments of the negative log-likelihood function $-\log(L(\theta, \mathbf{s}(t), \sigma^2))$. Normalizing by $N$ and ignoring the parameter-independent $M \log \pi$ term, we get

$$l(\theta, \mathbf{s}(t), \sigma^2) = M \log \sigma^2 + \frac{1}{\sigma^2 N} \sum_{k=1}^{N} \|\mathbf{x}(k) - \mathbf{A}\mathbf{s}(k)\|^2, \tag{5.12}$$

Obviously, the estimate for the signal waveform is

$$\hat{\mathbf{s}}(t) = \mathbf{A}^+ \mathbf{x}(k), \tag{5.13}$$

where $\mathbf{A}^+$ is the pseudo-inverse of $\mathbf{A}$. To calculate $\hat{\sigma}^2$ it is necessary to take the derivative of the log-likelihood function and set the result equal to zero, that is

$$\hat{\sigma}^2 = \frac{1}{MN} \sum_{k=1}^{N} \|\mathbf{x}(k) - \mathbf{A}\mathbf{s}(k)\|^2. \tag{5.14}$$

If $\hat{\mathbf{s}}(k)$ is inserted in the above equation, then $\hat{\sigma}^2$ becomes

$$\hat{\sigma}^2 = \frac{1}{MN} \sum_{k=1}^{N} \|\mathbf{x}(k) - \mathbf{A}\mathbf{A}^+ \mathbf{x}(k)\|^2 = \frac{1}{M} \text{trace} \left\{ \mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_{xx} \right\}, \tag{5.15}$$

where $\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}\mathbf{A}^+$ is idempotent and hermitian. Substituting (5.15) and (5.13) into (5.11) shows that maximum likelihood parameters are obtained by solving the following minimization problem

$$\hat{\theta} = \arg \left\{ \min_{\theta} \text{trace} \left\{ \mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_{xx} \right\} \right\}. \tag{5.16}$$

The interpretation is that the measurements $\mathbf{x}(t)$ are projected onto a model subspace orthogonal to all anticipated signal components, and a power measurement $\frac{1}{N} \sum_{k=1}^{N} \|\mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{x}(k)\|^2 = \text{trace} \left\{ \mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{R}}_{xx} \right\}$

is evaluated. The energy should clearly be smallest when the projector indeed removes all the true signal components, that is, when $\hat{\theta} = \theta_0$. Since only a finite number of noisy samples is available, the energy is not perfectly measured and $\hat{\theta}$ will deviate from $\theta_0$. However, if the scenario is stationary, the error will converge to zero as the number of samples is increased to infinity. This remains valid for correlated or even coherent signals, although the accuracy in finite samples is somewhat dependent upon signal correlations. To calculate ML estimates, the non-linear optimization (5.16) has to be solved numerically which results in high computational complexity.

Maximum likelihood estimation is a parametric method and hence its resolution is not limited as is the case for the conventional methods. The ML estimator presented here can be classified as a deterministic ML estimator, because the impinging multipath rays of both, the desired signal and the interferers are modeled deterministically. It is also possible to model the interfering sources as colored noise. In [21] such a stochastic ML estimator has been shown to have a better accuracy then the corresponding deterministic ML estimates.

**Subspace Based Methods**

All the subspace based methods are based on the eigenvector decomposition of the covariance matrix

$$\mathbf{R}_{xx} = E\left\{\mathbf{x}(t)\mathbf{x}^{\mathrm{H}}(t)\right\}. \tag{5.17}$$

Taking into account (5.1), we get for the covariance matrix

$$\mathbf{R}_{xx} = \mathbf{A}E\left\{\mathbf{s}(t)\mathbf{s}^{\mathrm{H}}(t)\right\}\mathbf{A}^{\mathrm{H}} + E\left\{\mathbf{n}(t)\mathbf{n}^{\mathrm{H}}(t)\right\}. \tag{5.18}$$

Denote the covariance matrix of $\mathbf{s}(t)$ with $\mathbf{R}_{ss}$. Assuming the noise is Gaussian, the covariance matrix of $\mathbf{n}(t)$ is $\sigma^2\mathbf{I}$. Therefore $\mathbf{R}_{xx}$ can now be written as

$$\mathbf{R}_{xx} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I}. \tag{5.19}$$

Because $\mathbf{R}_{xx}$ is a positive definite, hermitian matrix, we can use singular value decomposition (SVD) to get

$$\mathbf{R}_{xx} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}, \tag{5.20}$$

with $\mathbf{U}$ unitary and $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \lambda_2 \ldots, \lambda_M\}$ a diagonal matrix of real eigenvalues ordered such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M \geq 0$. If a vector $\mathbf{x}$ is orthogonal to $\mathbf{A}^{\mathrm{H}}$, then it is an eigenvector of $\mathbf{R}_{xx}$ with eigenvalue $\sigma^2$, because

$$\mathbf{R}_{xx}\mathbf{x} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^{\mathrm{H}}\mathbf{x} + \sigma^2\mathbf{I} = \sigma^2\mathbf{x}. \tag{5.21}$$

There are $M - L$ linearly independent such vectors ($M$ being the number of elements in the array and $L$ being the number of impinging array fronts). Since the remaining eigenvalues are all larger than $\sigma^2$, we can partition the eigenvalue/vector pairs into noise eigenvectors (corresponding to eigenvalues $\lambda_{L+1} = \ldots = \lambda_M = \sigma^2$ ) and signal eigenvectors (corresponding to eigenvalues $\lambda_1 \geq \ldots \geq \lambda_L \geq \sigma^2$). Hence, we can write

$$\mathbf{R}_{xx} = \mathbf{U}_s\mathbf{\Lambda}_s\mathbf{U}_s^{\mathrm{H}} + \mathbf{U}_n\mathbf{\Lambda}_n\mathbf{U}_n^{\mathrm{H}}, \tag{5.22}$$

where $\mathbf{\Lambda}_n = \sigma^2\mathbf{I}$. Since all noise eigenvectors are orthogonal to $\mathbf{A}$, the columns of $\mathbf{U}_s$ must span the range space of $\mathbf{A}$ whereas those of $\mathbf{U}_n$ span its orthogonal complement (the null space of $\mathbf{A}^{\mathrm{H}}$). The projection operators onto these signal and noise subspaces are defined as

$$\mathbf{P}_{\mathbf{A}} = \mathbf{A}\mathbf{A}^{+} = \mathbf{U}_s(\mathbf{U}_s^{\mathrm{H}}\mathbf{U}_s)^{-1}\mathbf{U}_s^{\mathrm{H}} = \mathbf{U}_s\mathbf{U}_s^{\mathrm{H}}, \tag{5.23a}$$

$$\mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{I} - \mathbf{A}\mathbf{A}^{+} = \mathbf{U}_n(\mathbf{U}_n^{\mathrm{H}}\mathbf{U}_n)^{-1}\mathbf{U}_n^{\mathrm{H}} = \mathbf{U}_n\mathbf{U}_n^{\mathrm{H}} \tag{5.23b}$$

A. The MUSIC algorithm

The simplest of the algorithms that are based on the above stated subspace decomposition is the MUSIC (MUltiple SIgnal Classification) algorithm. As noted previously, the structure of the exact covariance matrix with the spatial white noise assumption implies that its spectral decomposition can be expressed as

$$\mathbf{R}_{xx} = \mathbf{A}\mathbf{U}_{ss}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I} = \mathbf{U}_s\mathbf{\Lambda}_s\mathbf{U}_s^{\mathrm{H}} + \sigma^2\mathbf{U}_n\mathbf{U}_n^{\mathrm{H}}, \tag{5.24}$$

where, assuming $\mathbf{A}\mathbf{U}_{ss}\mathbf{A}^{\mathrm{H}}$ to be of full rank, the diagonal matrix $\mathbf{\Lambda}_s$ contains the $L$ largest eigenvalues. Since the eigenvectors in $\mathbf{U}_n$ (the noise eigenvectors) are orthogonal to $\mathbf{A}$, we have

$$\mathbf{U}_n^{\mathrm{H}}\mathbf{a}(\theta) = 0, \qquad \theta \in \{\theta_1, \ldots, \theta_L\}. \tag{5.25}$$

To allow for unique DoA estimates, the array is usually assumed to be unambiguous, that is, any collection of $M$ steering vectors corresponding to distinct DoAs $\eta_k$ forms a linearly independent set $\{a(\eta_1), \ldots, a(\eta_M)\}$, where $L < M$. If $\mathbf{a}(\cdot)$ satisfies these conditions and $\mathbf{R}_{ss}$ has full rank, then $\mathbf{A}\mathbf{U}_{ss}\mathbf{A}^{\mathrm{H}}$ is also of full rank. It then follows that $\theta_1, \ldots, \theta_L$ are the only possible solutions to the relation in (5.25), which could therefore be used to exactly locate the DoAs.

In practice, an estimate $\hat{\mathbf{R}}_{xx}$ of the covariance matrix is obtained and its eigenvectors are separated into the signal and noise eigenvectors. The orthogonal projector onto the noise subspace is estimated as

$$\hat{\mathbf{P}}_{\mathbf{A}}^{\perp} = \hat{\mathbf{U}}_n\hat{\mathbf{U}}_n^{\mathrm{H}}. \tag{5.26}$$

The MUSIC spatial spectrum is then defined as

$$P_L(\theta) = \frac{\mathbf{a}^{\mathrm{H}}(\theta)\mathbf{a}(\theta)}{\mathbf{a}^{\mathrm{H}}(\theta)\hat{\mathbf{P}}_{\mathbf{A}}^{\perp}\mathbf{a}(\theta)}. \tag{5.27}$$

Although $P_L(\theta)$ is not a true spectrum in any sense (it is merely the distance between two subspaces), it exhibits peaks in the vicinity of the true DoAs as suggested by (5.25). The performance improvement of the MUSIC estimator was so significant that it became an alternative to most existing methods. In particular, it follows from the above reasoning that estimates of an arbitrary accuracy can be obtained if the data collection time is sufficiently long or the SNR is adequately high, and the signal model is sufficiently accurate. Thus, in contrast to the beamforming techniques, the MUSIC algorithm provides statistically consistent estimates. Its important limitation is still the failure to resolve closely spaced signals in small samples at low SNR scenarios. This loss of resolution is even more pronounced for highly correlated signals. In the limiting case of coherent signals, the property (5.25) is violated and the method fails to yield consistent estimates.

The extensions of the MUSIC algorithms are comprised of Root-MUSIC (for ULA), Constrained MUSIC (incorporates the knowledge of the known source to improve the estimates of the unknown source direction) and Beam-Space MUSIC (where the array data are passed through a beam-forming processor before applying MUSIC).

B. The ESPRIT algorithm

ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) [20] is a computationally efficient and robust method of DoA estimation. It uses two identical arrays in the sense that

array elements need to form matched pairs with an identical displacement vector, that is, the second element of each pair ought to be displaced by the same distance and in the same direction relative to the first element.

This, however, does not mean that one has to have two separate arrays. The array geometry should be such that the elements could be selected to have this property. For example, a ULA of four identical elements with an interelement spacing $d$ may be thought of as two arrays of three matched pairs, one with the first three elements and one with the last three elements such that the first and second elements form one pair, the second and third elements form another pair, and so on. The two arrays are displaced by the distance $d$. The way that ESPRIT exploits this subarray structure for DoA estimation is now briefly described.

Let the signals induced on the $l$-th pair due to a narrowband source in direction $\theta$ be denoted by $x_l(t)$ and $y_l(t)$. The phase difference between these two signals depends upon the time taken by the plane wave arriving from the source under consideration to travel from one element to the other. As the two elements are separated by the displacement $\triangle_0$, it follows that

$$y_l(t) = x_l(t)e^{\mathrm{j}2\pi\triangle_0 \cos\theta}, \tag{5.28}$$

where $\triangle_0$ is measured in wavelengths. Note that $\triangle_0$ is the magnitude of the displacement vector. This vector sets the reference direction, and all angles are measured with reference to this vector. Let the array signals received by the two arrays be denoted by $\mathbf{x}(t)$ and $\mathbf{y}(t)$. These are given by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}_x(t) \tag{5.29}$$

and

$$\mathbf{y}(t) = \mathbf{A}\mathbf{\Phi}\mathbf{s}(t) + \mathbf{n}_y(t), \tag{5.30}$$

where $\mathbf{A}$ is a $M$ by $L$ matrix, with its columns denoting the $L$ steering vectors corresponding to $L$ directional sources associated with the first subarray and $\mathbf{\Phi}$ is a $L$ by $L$ diagonal matrix, with its $l$-th diagonal element given by

$$\mathbf{\Phi}_{ll} = e^{\mathrm{j}2\pi\triangle_0 \cos\theta_l}. \tag{5.31}$$

where $\mathbf{s}(t)$ denotes $L$ source signals induced on a reference element and $\mathbf{n}_x$ and $\mathbf{n}_y$, respectively, denote the noise induced on the elements of the two subarrays. Comparing the equations for $\mathbf{x}(t)$ and $\mathbf{y}(t)$, it follows that the steering vectors corresponding to $L$ directional sources associated with the second subarray are given by $\mathbf{A}\mathbf{\Phi}$.

Let $\mathbf{U}_x$ and $\mathbf{U}_y$ denote two $M$ by $L$ matrices with their columns denoting the $L$ eigenvectors corresponding to the largest eigenvalues of the two array correlation matrices $\mathbf{R}_{xx}$ and $\mathbf{R}_{yy}$, respectively. As these two sets of eigenvectors span the same $L$-dimensional signal space, it follows that these two matrices $\mathbf{U}_x$ and $\mathbf{U}_y$ are related by a unique nonsingular transformation matrix $\psi$, that is

$$\mathbf{U}_x\psi = \mathbf{U}_y. \tag{5.32}$$

Similarly, these matrices are related to steering vector matrices $\mathbf{A}$ and $\mathbf{A}\mathbf{\Phi}$ by another unique nonsingular transformation matrix $\mathbf{T}$, as the same signal subspace is spanned by these steering vectors. Thus

$$\mathbf{U}_x = \mathbf{A}\mathbf{T} \tag{5.33}$$

and

$$\mathbf{U}_y = \mathbf{A}\mathbf{\Phi}\mathbf{T}. \tag{5.34}$$

Substituting for $\mathbf{U}_x$ and $\mathbf{U}_y$ and taking into account the fact that $\mathbf{A}$ is of full rank, one obtains

$$\mathbf{T}\psi\mathbf{T}^{-1} = \mathbf{\Phi}, \tag{5.35}$$

which states that the eigenvalues of $\psi$ are equal to the diagonal elements of $\mathbf{\Phi}$ and that the columns of $\mathbf{T}$ are eigenvectors of $\psi$. This is the main relationship in the development of ESPRIT. It requires an estimate of $\psi$ from the measurements of $\mathbf{x}(t)$ and $\mathbf{y}(t)$. An eigendecomposition of $\psi$ gives its eigenvalues, and by equating them to $\mathbf{\Phi}$ leads to the DoA estimates

$$\theta_l = \arccos\left\{\frac{\arg(\lambda_l)}{2\pi\triangle_0}\right\}, \qquad l = 1, \ldots, L. \tag{5.36}$$

How one obtains an estimate of $\psi$ from the array signal measurements efficiently has led to many versions of ESPRIT. One version, referred to as TLS ESPRIT, is summarized below.

1. Make measurements from two identical subarrays, which are displaced by $\triangle_0$. Estimate the two array correlation matrices from the measurements and find their eigenvalues and eigenvectors.

2. Find the number of directional sources $L$ using available methods.

3. Form the two matrices with their columns being the $L$ eigenvectors associated with the largest eigenvalues of each correlation matrix. Let these be denoted by $\mathbf{U}_x$ and $\mathbf{U}_y$. For an ULA, this could be done by first forming an $M$ by $L$ matrix $\mathbf{U}$ selecting its columns as the $L$ eigenvectors associated with the largest eigenvalues of the estimated array correlation matrix of the full array of $M$ elements. Then select the first $K < M$ rows of $\mathbf{U}$ to form $\mathbf{U}_x$ and the last of its $K$ rows to form $\mathbf{U}_y$.

4. Form a $2L$ by $2L$ matrix

$$\begin{bmatrix} \mathbf{U}_x^{\mathrm{H}} \\ \mathbf{U}_y^{\mathrm{H}} \end{bmatrix} [\mathbf{U}_x \quad \mathbf{U}_y]$$

   and find its eigenvectors. Let these eigenvectors be the columns of a matrix $\mathbf{V}$.

5. Partition $\mathbf{V}$ into $4L$ by $L$ matrices as

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}.$$

6. Calculate the eigenvalues $\lambda_l$, $l = 1, \ldots, L$ of the matrix

$$-\mathbf{V}_{11}\mathbf{V}_{22}^{-1}.$$

7. Estimate the angle of arrival $\theta_l$ using (5.36).

**Discussion on DoA based methods**

In DoA based methods it is assumed that the angular spread of the received signals is relatively small. DoA based techniques are analytically more tractable but these methods need calibration. Namely, the antenna elements are not to be identical and they are mutually coupled. Since DoA based methods exploit the knowledge of the steering vector $\mathbf{a}(\theta)$, mutual coupling and difference of antenna elements have to be included into the steering vector so that DoA methods can work properly. Because usually this data is not known beforehand, it has to be estimated very accurately by employing the array calibration.

Also AoA estimation requires that the number of signals wavefronts $L$ including CCI signals should be less than the number of antenna elements $M$. Loosely speaking, DoA methods performance depends on the ratio between signal angular spread and number of impinging wavefronts. Also the level of CCI suppression very much depends on the angular distribution of the interfering and desired signals. One of the main advantages of DoA based methods is that AoA estimated at the reverse link (uplink) can be directly translated to the forward link (downlink) in systems with TDD but DoA methods performance can be seriously degraded in the presence of coherent multipath producing signal cancellation at the array output. A critical assumption of the most AoA estimation algorithms is that the number of incident signals should be strictly less than the number of antenna elements. This requirement can be relaxed if the properties of incident signals are exploited.

### 5.2.2 Training Signal Methods

In many mobile communication systems such as GSM and IS-54, explicit training signals are inserted into the data bursts. These training signals can be used to estimate the beamformer or the channel for each transmitted signal. There are several different approaches that may be taken when training signals are available. Conventional methods, for example, use the training signal and the received signal vector $\mathbf{x}(k)$ to determine a beamformer that separates the impinging signals. Maximum likelihood estimation can be used to jointly estimate DoAs and the channels, an interesting special case being the type of single snapshot algorithm described below. Maximum likelihood estimation can also be used to estimate the channels alone, ignoring any knowledge of the steering vector $\mathbf{a}(\theta)$.

**Conventional Methods**

If a desired response, $d(t)$, is given, it can be used to calculate the weight vector $\mathbf{w}$. Similarly to Conventional DoA based methods, the MSE (Wiener-Hopf) solution for the weight vector $\mathbf{w}$ is stated here

$$\mathbf{w} = \mathbf{R}_{xx}^{-1}\mathbf{r}_{xd}. \tag{5.37}$$

The weight vector $\mathbf{w}$ can then be used to separate the transmitted signal from interferences. Again, other choices for the weight vector $\mathbf{w}$ are possible and are based on performance measure such as the SNR or the MNV. However, the conventional beamformer does not take into account the impulse response of the channel and therefore is not appropriate as a stand-alone for most mobile communication problems. Especially in the case of fading channels, the channel must be estimated and its effects reversed.

### Maximum Likelihood Method

It is usually assumed that the training signals from the interfering mobiles are temporally white. The interferers can then be modeled not deterministically, but stochastically, as noise which is spatially colored. No assumption is made about the number of interferers or their channels. Thus the discrete vector channel model used here can be represented as

$$\mathbf{x}(k) = \mathbf{H}\mathbf{b}(k) + \mathbf{n}(k), \tag{5.38}$$

where during the training period, the data $\mathbf{b}$ is known. In a similar way as in the case of DoA based Maximum Likelihood algorithm we can say that the observation vector is going to be Gaussian process with mean $\mathbf{H}\mathbf{b}(t)$ and with covariance matrix $\mathbf{Q}$. The estimation problem is then to jointly estimate $\mathbf{H}$ and $\mathbf{Q}$.

Using the negative log-likelihood function, the estimates are given by

$$\left[\hat{\mathbf{H}},\, \hat{\mathbf{Q}}\right]: \qquad \arg\left\{\min_{\mathbf{H},\mathbf{Q}}\left(\log(\det(\mathbf{Q})) + \frac{1}{N}\sum_{k=1}^{N}\left(\mathbf{x}(k) - \mathbf{H}\mathbf{b}(k)\right)^{\mathrm{H}}\mathbf{Q}^{-1}\left(\mathbf{x}(k) - \mathbf{H}\mathbf{b}(k)\right)\right)\right\}, \tag{5.39}$$

where $M$ is the dimension of the vector $\mathbf{x}(t)$ and $N$ is the number of time samples measured.

Using the trace property $\mathbf{x}^{\mathrm{H}}\mathbf{y} = \mathrm{trace}(\mathbf{y}\mathbf{x}^{\mathrm{H}})$, the above minimization problem can be rewritten as

$$\begin{aligned}\left[\hat{\mathbf{H}},\, \hat{\mathbf{Q}}\right]: \quad \arg\Bigg\{\min_{\mathbf{H},\mathbf{Q}}\Bigg(&\log(\det(\mathbf{Q})) + \mathrm{trace}\left(\mathbf{Q}^{-1}\hat{\mathbf{R}}_{xx}\right) - \\ &-\mathrm{trace}\left(\mathbf{Q}^{-1}\mathbf{H}\hat{\mathbf{R}}_{bx}\right) - \mathrm{trace}\left(\mathbf{H}^{\mathrm{H}}\mathbf{Q}^{-1}\hat{\mathbf{R}}_{xb}\right) + \mathrm{trace}\left(\mathbf{H}^{\mathrm{H}}\mathbf{Q}^{-1}\mathbf{H}\hat{\mathbf{R}}_{bb}\right)\Bigg)\Bigg\}\end{aligned} \tag{5.40}$$

Differentiating this with respect to $\mathbf{H}$ and $\mathbf{Q}$ and setting both equations to zero, the following two estimators are derived

$$\hat{\mathbf{H}} = \hat{\mathbf{R}}_{xb}\hat{\mathbf{R}}_{bb}^{-1}, \tag{5.41a}$$

$$\hat{\mathbf{Q}} = \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{xb}\hat{\mathbf{R}}_{bb}^{-1}\hat{\mathbf{R}}_{xb}^{\mathrm{H}}. \tag{5.41b}$$

The estimates $\hat{\mathbf{H}}$ and $\hat{\mathbf{Q}}$ can then be used in a sequence estimator.

### Space-Time processing

Time-only processing corresponds to equalizers that use a weighted sum of signal samples and space-only processing corresponds to simple beamforming that uses a weighted sum of antenna outputs. Space-time processing combines time-only and space-only processing (see Fig. 4.1).

A popular optimality criterion in space-time processing is maximum likelihood (ML) usually referred to as maximum likelihood sequence estimation (MLSE). Space-time MLSE (ST-MLSE) seeks to estimate the data sequence that is most likely to have been sent given the received vector signal. Another frequently used criterion is *minimum mean-square error* (MMSE). In ST-MMSE we obtain an estimate of the transmitted signal as a space-time weighted sum of the received signal and seek to minimize the means-square error between the estimate and the true signal at every time instant. We present ST-MLSE and ST-MMSE in a form that is a space-time extension of already described ML and MS algorithms.

A. ST-MLSE

With the channel model described by (4.16), we assume that the noise $\mathbf{N}$ is spatially and temporally white and Gaussian and that there is no interference. The MLSE problem can be shown to reduce to finding $\mathbf{S}$ so as to satisfy the following criterion

$$\arg\left\{\min_{\mathbf{S}} \|\mathbf{X}-\mathbf{HS}\|_F^2\right\}, \tag{5.42}$$

where the channel $\mathbf{H}$ is assumed to be known and $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. This is a generalization of the standard MLSE problem where the channel is now defined in space and in time. We can therefore use a space-time generalization of the well known Viterbi Algorithm (VA) to carry out the search in (5.42) efficiently. In the presence of CCI, which is likely to be both spatially and temporally correlated (due to delay spread), the MLSE criterion can be reformulated with a new metric to address the new problem. However, the temporal correlation complicates the implementation of the Viterbi equalizer, making MLSE less attractive in the presence of CCI with delay spread.

B. ST-MMSE

In the presence of CCI with delay spread, an ST-MMSE receiver is more attractive. This receiver combines the input in space and time to generate an output that minimizes the (squared) error between itself and the desired signal.

The scalar equalizer output $y(k)$ can be written as (4.20)

$$y(k) = W^H(k)X(k) = \text{trace}(\mathbf{W}^{\text{H}}(k)\mathbf{X}(k)), \tag{5.43}$$

where trace$(\cdot)$ denotes the trace of a matrix.

The ST-MMSE filter chooses the space-time filter weights to achieve the MMSE, that is

$$\arg\left\{\min_{W} E\|W^{\text{H}}X(k)-s(k-\zeta)\|^2\right\}, \tag{5.44}$$

where $\zeta$ is a delay chosen to center the space-time filter (the choice of this parameter strongly affects performance). The solution to this minimum mean square problem follows from the well-known projection theorem

$$E\left(X(k)(X^{\text{H}}(k)W - s^*(k-\zeta))\right) = 0. \tag{5.45}$$

This leads to

$$W = \left\{E\left(X(k)X^{\text{H}}(k)\right)\right\}^{-1} E\left(X(k)s^*(k-\zeta)\right), \tag{5.46}$$

where superscript $*$ denotes complex conjugate. If the interference and noise are independent of the signal, the transmitted bit sequence is white, $M < N$ and

$$E\left(X(k)s^*(k-\zeta)\right) = [0\ \ldots\ 0\,\text{vec}^{\text{T}}(\mathbf{H})\,0\ \ldots\ 0]^{\text{T}} = \overline{H}, \tag{5.47}$$

where the number of zeros preceding and succeeding $\text{vec}^{\text{T}}(\mathbf{H})$ depends on the choice of $\zeta$. Defining the space-time $MN \times MN$ covariance matrix $\mathbf{R}_{xx} = E(XX^{\text{H}})$, (5.46) takes the familiar form of

$$W = \mathbf{R}_{xx}^{-1}\overline{H}. \tag{5.48}$$

Note that when $M = N$, $\overline{H} = \text{vec}(\mathbf{H})$.

A number of techniques are available in order to solve (5.48) adaptively. Popular algorithms include the celebrated LMS or the recursive-least-square algorithm.

In multiuser MMSE we usually estimate each user signal separately using the single user MMSE processor given in (5.48). In this case the MMSE treats other user signals as interference with unknown structure. Multiuser techniques need either training signals or blind channel estimates for all the users. The multiple training signals should be designed to have low cross-correlation properties so as to minimize cross coupling in the channel estimates.

**Discussion on Training Signal Methods**

Beamforming based on training signal or time reference beamforming (TRB) is a computationally effective method at the expense of spectrum efficiency. Spatial information such as AoA or array manifold is not necessary. Depending on the particular system and/or scenario used, the reference signal may consists of a priori known signal multiplexed in frequency or in time with a useful signal or a reconstructed signal obtained from the detected symbols. The second approach is more attractive from the tracking requirements but the beamforming and detection are more interdependent. The use of training signals requires prior carrier and signal recovery, which is made difficult by the presence of CCI. In some cellular systems this technique is not applicable due to problems with obtaining a reference. Iterative adaptive algorithms such MSE, RLS (Recursive Least Squares), and DMI (Direct Matrix Inversion) algorithm are used for tracking. TRB overcomes interference by nulling its spatial signatures and shows greater robustness in the mobile environment where channel characteristics are continuously varying. Colored training sequences in GSM and an user dedicated pilot in UMTS can be used as a reference signal.

TRB technique optimally combines multipath components to increase SNR and to reduce the effect of fading. In a non-coherent multipath both DoA based and training signal methods have the same ability to overcome ISI since two different symbols for the same user will be uncorrelated and will look like noise. Instead of calibration in DoA based methods, the TRB techniques require accurate synchronization and to achieve the best performance when the delay spread is low. Unlike DoA there are no means of obtaining a transmitting weight vector for downlink beamforming with TRB from the information provided at the receiver in FDD. However, as mentioned before, the training sequences consume spectrum resource. In GSM, for example, 20% of the bits are dedicated to training.

### 5.2.3 Temporal Structure Methods

A signal $s(t)$ transmitted by a mobile radio has a rich temporal structure that can be used to improve the estimator performance. There are different types of temporal structure inherent to the transmitted signal, for example

- Finite Alphabet (FA)
- Constant Modulus (CM)

Temporal structure methods rely on this type of information to separate and equalize desired signals and interferers. Unlike the DoA based methods, the information contained in the array manifold is not used.

### Finite Alphabet

This approach is based on the finite alphabet (FA) property of the transmitted signals. The FA approach tries to fit the received data to the unknown channel and multi-user data. The discrete-time channel model from section 4.5 is used, which is given by

$$\mathbf{x}(k) = \sum_{q=1}^{Q} \mathbf{H}_q \mathbf{s}_q(k) + \mathbf{n}(k), \tag{5.49}$$

where $Q$ is the number of users and $\mathbf{x}(k)$ is the received signal for one of them. If there are $N$ snapshots of the received signal, the channel model can be put as

$$\mathbf{X} = \hat{\mathbf{H}}\mathbf{S} + \mathbf{N}, \tag{5.50}$$

where now $\mathbf{X}$ and $\mathbf{N}$ are $M \times N$ matrices, $\hat{\mathbf{H}}$ is a $M \times PQ$ matrix and $\mathbf{S}$ is $PQ \times N$ matrix. Here $P$ is the maximum memory of the channel. Both $\hat{\mathbf{H}}$ and $\mathbf{S}$ are unknown and the additive noise is assumed to be white and Gaussian. Obviously, the joint ML estimates for the channel and the data matrix are then given by the following minimization problem where the feasible set of the data matrix $\mathbf{S}$ is constrained to the known *finite alphabet*

$$\min_{\hat{\mathbf{H}}, \mathbf{S} \in \mathcal{FA}} \|\mathbf{X} - \hat{\mathbf{H}}\mathbf{S}\|_F^2, \tag{5.51}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. This is a joint ML problem where both the channel and data are unknown. The FA property allows us to solve the equation and estimate both $\hat{\mathbf{H}}$ and $\mathbf{S}$. Since the ML criterion is separable with respect to the unknowns, one approach to minimize the cost function is by alternating projections. Starting with an initial estimate of $\hat{\mathbf{H}}$, we minimize (5.51) with respect to $\mathbf{S}$ keeping $\hat{\mathbf{H}}$ constant. This is a data-detection problem. With an estimate of $\mathbf{S}$ an improved estimate of $\hat{\mathbf{H}}$ can be obtained by minimizing (5.51) with respect to $\hat{\mathbf{H}}$ keeping $\mathbf{S}$ fixed. This is a standard LS problem. This iterative process is continued until a fixed point is reached. The estimated symbols are then projected onto the finite alphabet, which removes ambiguities in the solution. The FA approach is fairly involved mathematically and the detailed solution is therefore not presented here. For details refer to [22, 23].

In [23, 24] an FA approach for identifying frequency selective FIR channels carrying multiple signals is presented. The approach uses oversampling of the received signal by a factor $\eta$ (number of samples per symbol).

The FA approach is computationally fairly complex, because a multidimensional least squares approach has to be employed to find a solution.

### Constant Modulus

In many wireless applications, the transmitted waveform has a constant envelope. A typical example of a constant-envelope waveform is the GMSK modulation used in the GSM cellular system, which has the following general form

$$u(t) = e^{\mathrm{j}(\omega t + \phi(t))}, \tag{5.52}$$

where $\phi(t)$ is a Gaussian-filtered phase output of an MSK signal.

The constant modulus (CM) algorithm has its origins in temporal (SISO) equalization techniques. The idea is to penalize deviations of the equalizer output $y(t)$ from a constant modulus. Therefore the CM algorithm minimizes a cost function of the form

$$J(\mathbf{w}) = E\{(|y(k)|^p - R_p)^2\}, \tag{5.53}$$

with respect to the weight vector $\mathbf{w}$ of the equalizer. In the above equation, $p$ is a positive integer, and $R_p$ is a positive real constant defined by

$$\frac{E\{|s(k)|^{2p}\}}{E\{|s(k)|^p\}}. \tag{5.54}$$

The resulting LMS algorithm is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \mathbf{x}(k) y^*(k)(|y(k)|^2 - R_p). \tag{5.55}$$

When using smart antenna systems, case with multiple users $Q$ becomes interesting. We consider a linear spatio-temporal equalization structure (see section 4.5)

$$\mathbf{W} = [W_1 \dots W_Q] \qquad (MN \times Q) \tag{5.56}$$

where $W_q$, $q = 1, \dots, Q$, denotes the spatio-temporal equalizer corresponding to the $q$-th signal. Then the $Q \times 1$ equalized output at time instant $k$ can be written as

$$\mathbf{y}(k) = \mathbf{W}^{\mathrm{H}}(k) X(k), \tag{5.57}$$

where $X(k)$ is defined as in (4.16). The Multiuser CM algorithm (MU-CM) determines $\mathbf{W}$ by minimizing a CM type cost function. The cost function proposed is the sum of a CM term and a cross-correlation term. The CM term penalizes the deviations of the equalized signals' magnitudes from a CM, whereas the cross-correlation term penalizes the correlation between them. The cost function is given by

$$\min_{\mathbf{W}} J(\mathbf{W}) = E\left\{\sum_{q=1}^{Q}(|y_q|^2 - R_p)^2\right\} + 2\sum_{l,n=1;\, l\neq n}^{Q}\sum_{\zeta=\zeta_1}^{\zeta_2}|r_{ln}(\zeta)|^2, \tag{5.58}$$

where $r_{ln}(\zeta)$ is the cross-correlation function between users $l$ and $n$ defined as

$$r_{ln} = E\left(y_l(k) y_n^*(k-\zeta)\right) \tag{5.59}$$

and $\zeta_1$, $\zeta_2$ are integers that depend on the channel delay spread. The corresponding stochastic-gradient algorithm has the form [14]

$$\mathbf{W}(k+1) = \mathbf{W}(k) - \mu[\hat{\triangle}_1(k) \ \dots \ \hat{\triangle}_Q(k)], \tag{5.60}$$

where

$$\triangle_q(k) = E\left\{(|y_q(k)|^2 - R_p) y_q^*(k) X(k)\right\} + \sum_{l=1;\, l\neq q}^{Q}\sum_{\zeta=\zeta_1}^{\zeta_2} r_{qn}(\zeta) E\{y_l^*(k-\zeta) X(k)\} \tag{5.61}$$

and $\hat{\triangle}_j$ is an estimate of $\triangle_j$ based on instantaneous values or sample averaging.

**Discussion on Temporal Structure Methods**

Temporal structure methods rely on the temporal structure contained in the transmitted signals. It was seen, that there are different types of temporal structure. Man-made signals transmit symbols from a finite alphabet and usually the symbols have a constant modulus. Generally, temporal structure methods do not make explicit use of available knowledge about spatial structure. Not using the knowledge about spatial structure has both positive and negative effects:

- No calibration of the array is needed (+)
- Temporal structure methods are not restricted by a maximum number of DoAs that can be estimated (+)
- The approach works better in scenarios with large angle spread (+)
- In scenarios with small angle spread, spatial structure methods could out-perform temporal structure methods, because available information is not used in order to estimate the unknown parameters (–)

Temporal Structure Methods vs. Training Signal Methods

In general, temporal structure methods are quite similar to the ML approach ignoring spatial structure. The advantages of temporal structure methods over training signal methods are

- No training sequence is needed, which saves available spectrum
- The channel is tracked during the whole duration of a burst
- No synchronization is required

## 5.3 Forward Link Processing

It was seen before, that a smart antenna can be used to separate several co-channel signals arriving from different angles. It is assumed, that a smart antenna is only employed at the base station and not in the mobile radio. Therefore, it is the base station's task to employ spatially selective transmission while the mobile radio's reception system is not spatially selective. The task of spatially selective transmission is called forward link or downlink processing.

The type of forward link processing used strongly depends on whether the communication system uses frequency division duplex (FDD) or time division duplex (TDD). This is, because in most FDD systems the reverse link and forward link fading may be considered independent, whereas in TDD systems the reverse link and forward link channels can be considered reciprocal. Hence, in TDD systems reverse link channel information may be used to achieve spatially selective transmission. In FDD systems, the reverse link channel information cannot be used directly and other types of forward link processing have to be considered.

In this section we first introduce the forward link discrete signal model, then describe forward link channel estimation methods and end with a brief discussion on forward link algorithms.

### 5.3.1 Forward Link Discrete Signal Model

Following the development of the discrete signal model, we can immediately write down the received signal model at a single mobile as

$$x(k) = \mathbf{w}^{\mathrm{H}}\mathbf{H}^{F}\mathbf{s}(k) + z(k) + n(k), \tag{5.62}$$

where $\mathbf{H}^F$ is the forward channel (the $i$-th row of $\mathbf{H}^F$ is the sampled response at the mobile due to a transmitted symbol from the $i$-th antenna at the base station), $z(k)$ is the interference arriving at the mobile from out-of-cell co-channel base station and $n(k)$ is the thermal noise at the mobile.

Likewise, in the multiuser case, the base station transmits $Q$ user signals, one to each mobile. Due to the channel characteristics, perfect isolation of these signals will not be possible. The signal received by the $q$-th mobile will contain its own intended signal along with signals intended for the other users within the cell plus CCI from other cells. It is therefore given by

$$x_q(k) = \mathbf{w}_q^{\mathrm{H}}\mathbf{H}_q^{F}\mathbf{s}_q(k) + \sum_{p=1,\, p\neq q}^{Q} \mathbf{w}_p^{\mathrm{H}}\mathbf{H}_p^{F}\mathbf{s}_p(k) + z_q(k) + n_q(k), \tag{5.63}$$

where $\sum_{p=1,\, p\neq q}^{Q} \mathbf{w}_p^{\mathrm{H}}\mathbf{H}_p^{F}\mathbf{s}_p(k)$ is the intracell CCI, $z_q(k)$ is intercell CCI, and $n_q(k)$ is the thermal noise.

### 5.3.2 Estimating the Forward Channel

There are three general approaches for estimating the forward channel $\mathbf{H}^F$

#### Channel Estimation in TDD

In a TDD system, if the duplexing time is small compared to the coherence time of the channel, both channels are the same and the base station can use its estimate of the reverse channel as the forward channel $\mathbf{H}^F = \mathbf{H}^R$.

#### Channel Estimation in FDD

In FDD systems, the forward and reverse channels can be potentially very different. This difference arises from differences in instantaneous complex path gains. Keeping this in mind, we can find the following methods to partially estimate the forward from the reverse channel

- Zero delay and angle spread: In these conditions the channel reduces to a single vector $\mathbf{a}(\theta)$. In this case the forward and reverse channels are identical up to a scalar factor.
- Zero-angle spread, nonzero delay spread: The reverse channel can be factored into spatial and temporal components $\mathbf{H}^R = \mathbf{a}h^R$, where $h^R$ is the temporal channel, which is the same at every antenna. The forward link channel will share a common $\mathbf{a}$ (for small antenna arrays) but not the temporal channel, that is $\mathbf{H}^F = \mathbf{a}h^F$ where $h^F$ is unknown.
- Zero delay spread, nonzero angle spread: The reverse channel now has a pure spatial component, $\mathbf{H}^R = \sum_{l=1}^{L} \alpha_l^R \mathbf{a}(\theta_l)$. The spatial channel is clearly different on both links while the (zero-delay spread) temporal channel is of course the same on both links.

- Finite delay and angle spread: The forward channels are different both spatially and temporally. However, the path delays and angles are same on both links. Using (4.13) and assuming no manifold ambiguities, we get that the two channels share common column or row subspaces: this structure can be used effectively for forward-link processing

**Channel estimation using feedback**

A direct approach to estimating the forward channel is to feedback the signal from the mobile unit and then use either a blind or training signal method for estimating the channel. To illustrate this approach, assume that we transmit a training signal through one antenna at a time from the base station to the mobile. Feedback signal from the mobile can then be used to solve an LS problem and estimate the channel from each base station antenna to the mobile.

### 5.3.3 Forward Link ST Processing

Once the forward link channel is known it is possible to develop space-time processing algorithms.

Unlike the reverse link, blind algorithms are applicable only when the mobiles can feed back the received signal to the base station. A more attractive approach might be to estimate the forward channel from the reverse channel as discussed above and then use training signal methods. We illustrate with one example each, the single and multiuser cases.

**Single-user MMSE**

The role of forward link processing in single user environments is to choose $\mathbf{w}$ so as to maximize the delivered signal level and minimize ISI at the mobile unit while minimizing CCI at the other susceptible co-channel mobiles in the other cells. The CCI at the reference mobile is not controlled by the base station under reference but is generated by other base stations transmitting to their own mobiles. Reducing CCI at one mobile requires the cooperation of the other base stations. Therefore we choose $\mathbf{w}$ such that

$$\min_{\mathbf{w}} \left\{ E\|\mathbf{w}^{\mathrm{H}}\mathbf{H}^{F}\mathbf{s}(k) - s(k-\zeta)\|^2 + C\sum_{q=1}^{Q-1} \mathbf{w}^{\mathrm{H}}\mathbf{H}_q^{F}\mathbf{H}_q^{F\mathrm{H}}\mathbf{w} \right\}, \tag{5.64}$$

where $C$ is a parameter that balances the ISI reduction at the reference mobile and CCI reduction at other co-channel mobiles, and $\zeta$ is again an integer delay that has to be chosen so as to center the filter output and optimize performance. $Q-1$ is the number of susceptible outer cell mobiles. $\mathbf{H}_q^F$ is the channel from the base station to the $q$-th outer cell mobile. To solve the (5.64) we need to know the forward link channel $\mathbf{H}^F$ to the reference mobile and $\mathbf{H}_q^F$ to co-channel mobiles. Finally, by using a vector $\mathbf{w}$ we have assumed spatial processing only. In general, ST processing will be used and we need to redefine $\mathbf{w}$ appropriately.

**Multi-user MMSE**

The role of forward link processing in multiuser environments is to chose $\mathbf{w}_q$ so as to maximize signal level while minimizing ISI and CCI at each mobile. If any CCI originates from other cell base stations, then the forward-link processing at the reference base station must reduce CCI generation at other cell mobiles and likewise benefit from CCI-reducing efforts at the desired mobile from other base stations.

If we assume that co-channel users are allowed to operate in the same cell and we ignore outer-cell CCI, we can formulate the problem as follows. We chose $\mathbf{w}_q$, $q = 1, \ldots, Q$ such that the average (or worst case) cost at each mobile is minimized. the cost at the $q$-th mobile becomes

$$\min_{\mathbf{w}_q} \left\{ E\|\mathbf{w}_q^{\mathrm{H}}\mathbf{H}_q^F\mathbf{s}_q(k) - s_q(k-\zeta)\|^2 + C \sum_{p=1;\, p\neq q}^{Q} \mathbf{w}_p^{\mathrm{H}}\mathbf{H}_q^F\mathbf{H}_q^{F\mathrm{H}}\mathbf{w}_p \right\}, \tag{5.65}$$

This leads to coupled problems that need nonlinear programming solutions. Also, by using a vector, $\mathbf{w}_q$, we have assumed spatial processing only. Again, in general, space-time processing can be used and we need to redefine $\mathbf{w}_q$ appropriately.

# Chapter 6

# Overview of CDMA Adaptive Processing Methods

The beamforming structures discussed in previous chapters were for narrow-band signals (with exception of space-time algorithms). The phase shift required by each weight $w_m$ depends on the differential propagation distance between each antenna element and the wavelength of the signal. Since the phase shift is a function of both the frequency and the desired beam direction while the weights only adjust the desired beam direction, it is evident that the weights are optimal in the chosen direction only at a centre frequency. However, in wideband systems, lower frequency signal components have less phase shift for a given propagation distance, whereas higher frequency signal components have a greater phase shift as they travel the same lengths. It is therefore necessary to introduce different phase responses for the different frequency components of the wideband signal, in order to provide flat frequency response within the signal bandwidth. The tapped-delay line allows each element to have a phase response that varies with frequency, resulting in a flat frequency response within the signal bandwidth. Such wideband arrays are therefore also called space-time receivers (see Fig. 4.1). An important special case of the wideband array is the spatial processing CDMA Rake receiver, which offers some of the benefits of the wideband array at a complexity level which is closer to the narrowband array.

In code division multiple access (CDMA) systems the users operate in the same frequency and time channel whereas in TDMA the users are separated in time. In direct sequence CDMA (DS-CDMA) each user has a unique spreading code. This code operates at a chip rate $P$ times greater than the information data rate. $P$ typically lies in the range of 100 to 1000 [14]. The DS-CDMA link therefore needs a large bandwidth channel that can be shared by multiple users. The user codes can be designed to be orthogonal or quasiorthogonal. The spreading code can be viewed as a complex symbol waveform with a large time-bandwidth product (approximately $P$), whereas in TDMA the time-bandwidth product of the symbol waveform is small (approximately 1).

If orthogonal codes are used and there is no multipath, code orthogonality can be maintained, thus users do not interfere with each other and signal detection is noise limited. Walsh codes are popular orthogonal codes, however, the number of Walsh codes is only equal to the code length, $P$. On the other hand, if we use quasi-orthogonal codes (whose number can be much larger) or orthogonal codes with multipath, the users interfere with each other and the detection becomes interference limited.

This multiple-access interference (MAI) that arises from other users is reduced by the processing gain $P$ during the detection process.

In essence, the presence of multiple co-channel users and the use of complex spreading code instead of symbol waveforms distinguish CDMA from TDMA signal processing. With the introduction of multiuser, TDMA signal processing (enabled by multiple antennas) there is increased convergence of TDMA and CDMA processing. In this chapter we first discuss digital beamforming systems for CDMA, then introduce the DS-CDMA signal model before briefly describing CDMA signal processing.

## 6.1 Digital Beam Forming with CDMA

The implementation of digital beam forming in a CDMA system is different from that in a TDMA system. In a TDMA system beamforming is carried out on a frame-by-frame basis for the entire frequency band. In a CDMA system, beamforming is carried out continuously for the entire CDMA frequency band. Moreover, the choice of DBF configuration depends on the type of a CDMA system, namely, a synchronous or asynchronous CDMA.

In a synchronous CDMA system, the information bit duration of each user signal in the system is time-aligned at the base station. The DBF configurations for a base station system in both the reverse link and the forward link cases are shown in Fig. 6.1 and Fig. 6.2, respectively. The reverse link (forward link) system consists of

1. An $M$-element antenna array
2. $M$ receiver (transmitter) modules
3. $M$ ADCs (DACs)
4. $N$ digital demodulator (modulator) banks, each of which consists of $M$ correlators and samplers
5. $N$ RDBF (TDBF) networks.

In the reverse link (uplink) case, a beamforming network is required for user signals that use an identical code. This requires that beamforming is carried out following the demodulation process (or code filtering). If each beamforming network produces $L$ beam outputs, the number of simultaneous users can be up to $N \times L$. In the forward link (downlink) case, the $N \times L$ message signals to be transmitted are arranged into $N$ groups. Beamforming is applied to the $L$ messages in each group. The beamforming outputs are spread with a particular code. The code-division multiplexed (CDM) signals are combined and converted to analog signals for transmission. In these configurations, the beamforming networks operate at the bit rate. However, it should be noted that the demodulation and modulation processes must be both linear and coherent to preserve the phase information that is required for beamforming.

In order for DBF to be implemented in an asynchronous CDMA system, the above configuration must be modified. As shown in Fig. 6.3, the receive beamforming networks must be placed before the demodulators and the transmit beamforming networks placed after the modulators. That is, the alternative reverse link (forward link) system consists of

1. An $M$-element antenna array

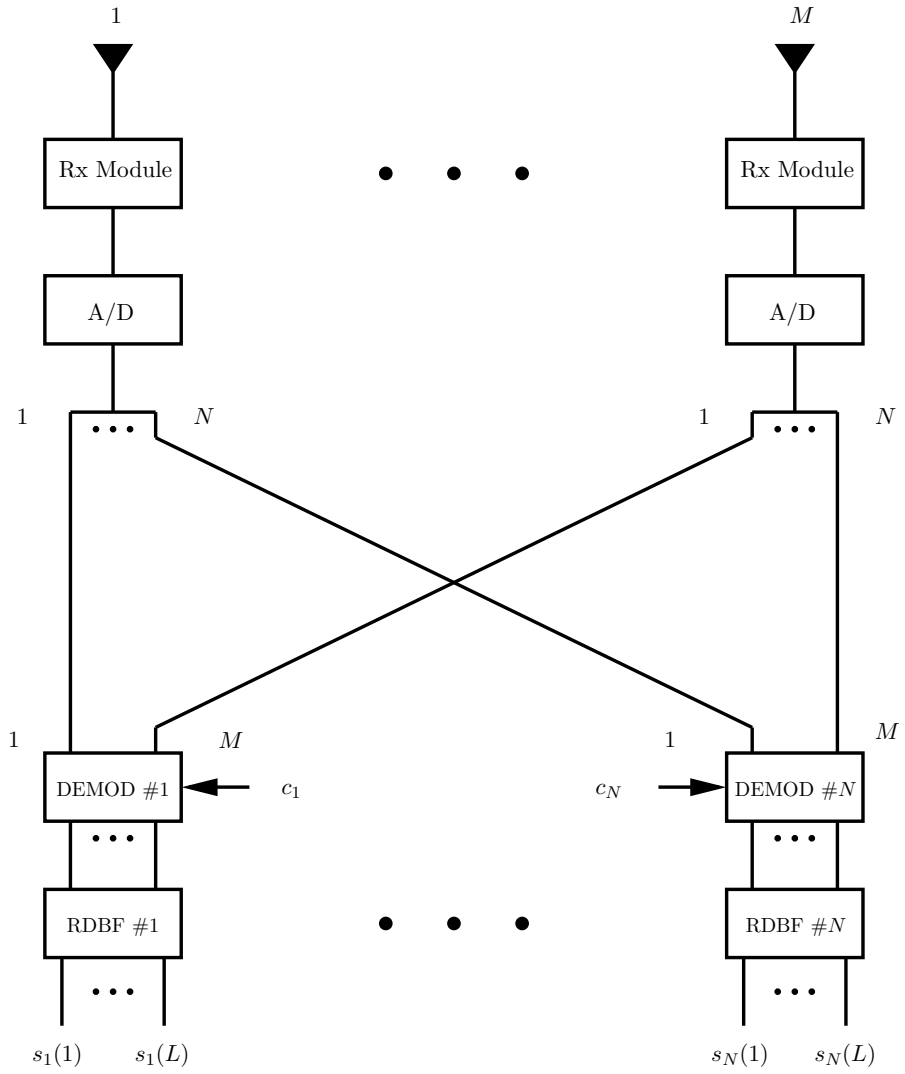



**Figure 6.1:** Reverse link DBF configuration for a CDMA system.

2. $M$ receiver (transmitter) modules

3. $M$ ADCs (DACs)

4. $M$ RDBF (TDBF)

5. $N \times L$ digital demodulators (modulators), each of which consists of a correlator and a sampler.

In Fig. 6.3 the time index $t_l$ for $c_m$ denotes the fact that the information bit duration of the user signals do not have to be synchronized. In these configurations, the demodulation and modulation processes can be linear or nonlinear, and coherent or noncoherent. Furthermore, $c_m(t_l)$ can be used as the reference to carry out adaptive beamforming. However, it is required that the beamforming networks operate at least at the chip rate.

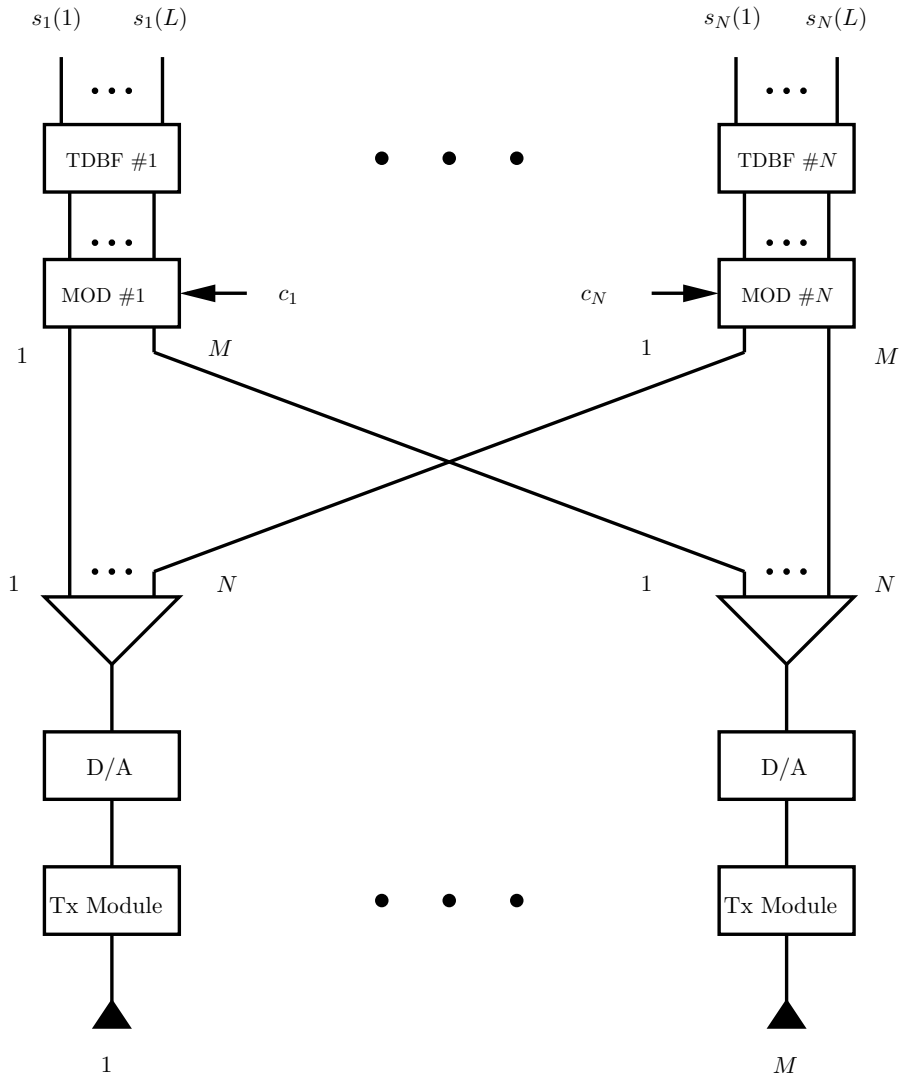
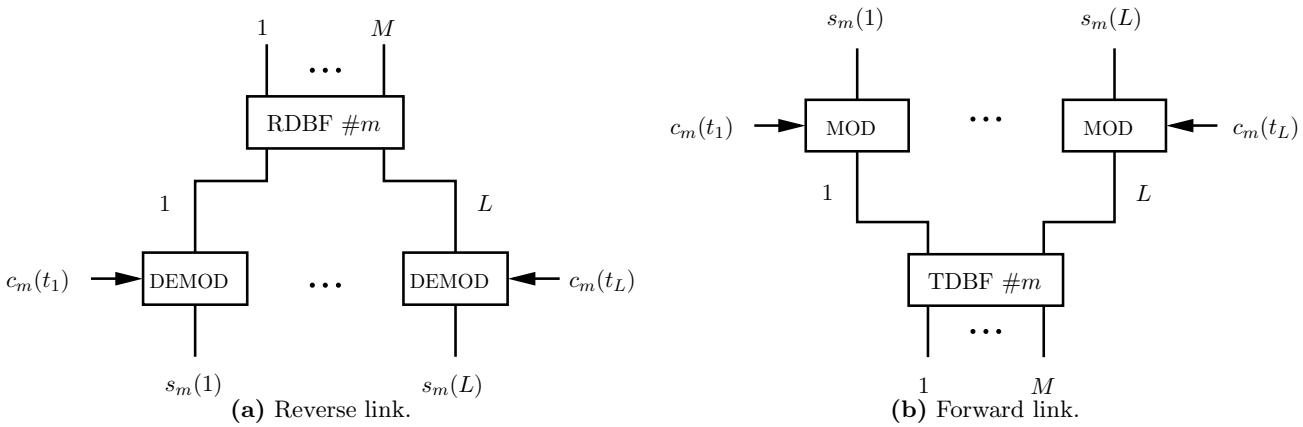



**Figure 6.2:** Forward link DBF configuration for a CDMA system.




**(a)** Reverse link. **(b)** Forward link.

**Figure 6.3:** Alternative DBF configurations for a CDMA system.

## 6.2 Signal Model

A simplified model for a received continuous-time signal in a DS-CDMA system has the following form

$$x(t) = \sum_{q=1}^{Q} x_q(t) + n(t), \tag{6.1}$$

where $x_q(t)$ is the contribution from the $q$-th user, $Q$ is the number of users sharing the same channel and $n(t)$ is the additive noise. Assuming a repeated spreading code, $x_q(t)$ can be expressed as

$$x_q(t) = \sum_{k=-\infty}^{+\infty} s_q(t) p_q(t - kT) + n_q(t), \tag{6.2}$$

where $\{s_q\}$ is the information bit stream (typically $s_q(k) = \pm 1$), $T$ is the symbol period and $p_q(t)$ is the overall channel given by the convolution of the impulse response of the channel and the spreading waveform, i.e.

$$p_q(t) = h_q(t) \otimes c_q(t), \tag{6.3}$$

where $\otimes$ denotes convolution, $h_q(\cdot)$ is the channel impulse response, and $c_q(\cdot)$ is the $q$-th user's spreading waveform given by

$$c_q(t) = \sum_{n=0}^{P-1} \gamma_q(n) g(t - nT_c), \tag{6.4}$$

where $P$ is the spreading factor or processing gain, $T_c = T/P$ is the chip period, $\{\gamma_q(n) \,|\, n = 0, \ldots, P-1\}$ is the $q$-th user's spreading code and $g(t)$ is the chip waveform. For a specular channel with $L$ paths, $h_q(t)$ can be written as

$$h_q(t) = \sum_{l=1}^{L} \alpha_{lq} \delta(t - \tau_{lq}), \tag{6.5}$$

where $\alpha_{lq}$ and $\tau_{lq}$ are the path coefficients and path delays, respectively.

## 6.3 Reverse Link Processing

A number of receiver structures have been proposed in CDMA and will be broadly summarized in this section. First we give an overview of Time-Only processing methods before finishing with Space-Time processing [14].

### 6.3.1 Time-Only Processing

A set of sufficient statistics for the detection of the user data is given by the outputs of a bank of matched filters, each matched to an individual user's channel, $p_q(t)$

$$r_q(k) = \int_{kT}^{(k+1)T} x(t) p_q^*(t - kT) \mathrm{d}t. \tag{6.6}$$

We can classify the detection problem into a single-user and a multiuser case. In a single-user detection we assume that each user's signal is detected independently of the others. In the multiuser case, all users are detected jointly. Each of these cases can be further classified into a number of subclasses with tradeoffs in complexity and receiver performance. We begin with the single-user receiver.

**Single-user approach**

Single-user receivers work on the assumption that the MAI is unknown random white noise while in fact MAI arises from other user signals and is not really unknown. The penalty for this assumption in DS-CDMA networks is the need for accurate power control wherein the received power from all users must be kept equal. If this is not ensured, the system capacity degrades rapidly. The advantage of single-user receiver is its simplicity and robustness. Its disadvantage is the need for accurate power control also known as the near-far problem. In the following, two different versions of the single-user receiver are described.

A. Simple Correlator

In this receiver, the channel is assumed to have a single path and therefore the channel $p_q(t)$ reduces to $c_q(t)$ and the correlator output is given by (Fig. 6.4)

$$r_q(k) = \int_{kT}^{(k+1)T} x(t)c_q(t-kT)\mathrm{d}t. \tag{6.7}$$

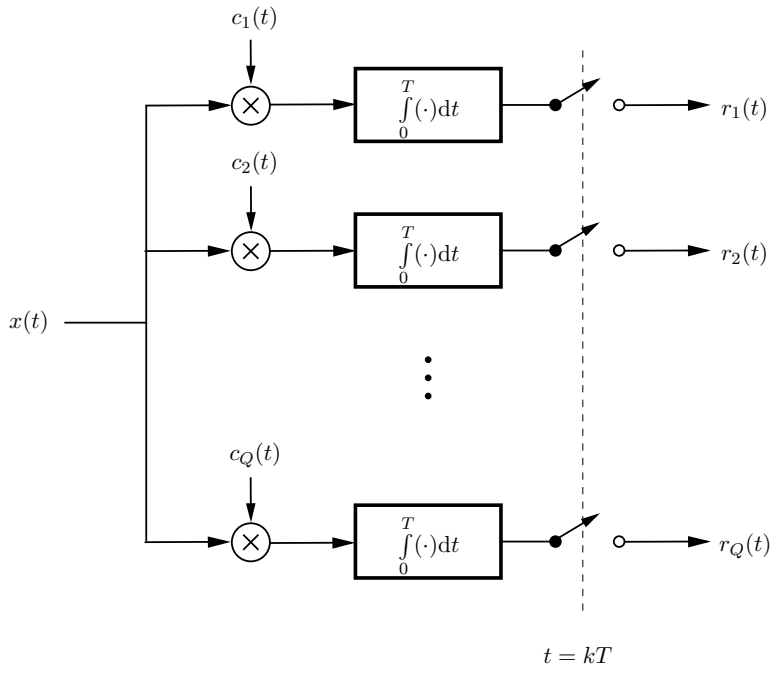



**Figure 6.4:** Simple Correlator.

The bit decisions are given by

$$\hat{s}_q(k) = \mathrm{dec}\left(r_q(k)\right), \tag{6.8}$$

where dec$(\cdot)$ is a decision (threshold) operation. The simple correlator will indeed be optimal if the following conditions are met

1. the codes are orthogonal, that is
$$\int_0^T c_i(t)c_j(t)\mathrm{d}t = \delta_{ij}, \tag{6.9}$$
where $\delta_{ij}$ denotes the Kronecker delta;
2. perfect symbol synchronization is achieved;
3. no multipath is present;
4. the additive noise is white and Gaussian.

However, if any of these conditions are not met, the receiver in (6.8) will be suboptimal. For example, if quasi-orthogonal spreading codes are used, this receiver will not be optimal due to the presence of MAI. Simple correlators are popular when the multipath is negligible.

B. The RAKE Receiver

When the channel exhibits significant multipath with delay spread larger than one chip period, an improved receiver can be designed to match this channel. Note that in DS-CDMA, multipath has both a positive and a negative effect. On one hand, the independently fading paths can be a valuable source of diversity. On the other hand, the multipath introduces interpath interference similar to ISI in TDMA and in the case where orthogonal codes are used, it also introduces MAI.

A popular single-user receiver in the presence of multipath is the RAKE combiner, first proposed by Price and Green in 1958 [25]. The RAKE receiver uses multiple correlators, one for each path, and the outputs of the correlators (called fingers) are then combined into a single output to maximize the SNR. If the combining (complex) weights are matched to the channel response $\alpha_{lq}$ at these fingers, we get a coherent RAKE receiver that is identical to a matched receiver (Fig. 6.5)

$$r_q(k) = \int_{kT+\tau_1(q)}^{(k+1)T+\tau_2(q)} x(t)p_q^*(t-kT)\mathrm{d}t = \sum_{l=1}^{L} \alpha_{lq}^* \int_{kT+\tau_{lq}}^{(k+1)T+\tau_{lq}} x(t)c_q(t-kT-\tau_{lq})\mathrm{d}t, \tag{6.10}$$

where (see (6.5))
$$\tau_1(q) = \min_l\{\tau_{lq}\}, \qquad \tau_2(q) = \max_l\{\tau_{lq}\}.$$

Another variation is the incoherent RAKE, in which the fingers are combined after detection (i.e., after removing channel phase information). The RAKE combining is maximal ratio when the weights are proportional to the path gain, and equal gain if the weights are all set to be equal. It is also possible to use combining weights to maximize SINR rather than SNR. In general, with few fingers and a large number of users, the MAI is likely to resemble uncorrelated noise and the two criteria will result in receivers with similar performance.

The use of a RAKE receiver assumes that the channel impulse response (path delay, amplitude and phase) is known. In practice, this is estimated by a tracker that estimates the path delay and computes its complex amplitude. The channel can be estimated using spreading codes and pilot or training signals. We can also estimate the channel with knowledge of spreading codes alone. A number of variations of the RAKE have been proposed and this is an area of active research.

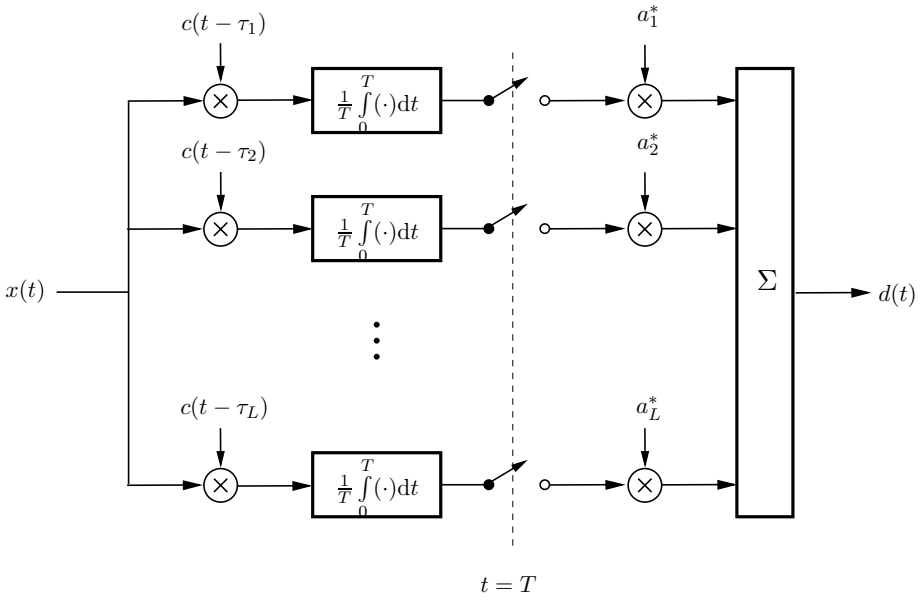



**Figure 6.5:** Coherent 1D RAKE receiver.

### Multiuser approach

A. Optimal (Maximum Likelihood) Detector

The optimal multiuser detector in the absence of delay spread satisfies the following criterion

$$\hat{s}(k) = \arg \min_{s\in\{-1,1\}^Q} \int_{kT}^{(k+1)T} \left[ x(t) - \sum_{q=1}^{Q} s_q(k) p_q(t-kT) \right]^2 \mathrm{d}t. \tag{6.11}$$

The multiuser receiver that implements this equation can provide significant performance gains over the single-user receiver, especially in cases of unequal user power (near-far problem). However, the optimization of the criterion in (6.11) is exponential in the number of users and may be prohibitive in some cases. Moreover, this receiver requires exact knowledge of the channel for all users. The need for a simpler and more robust receiver led to the search for linear receivers.

B. Linear Receivers

Linear receivers for multiuser detection in CDMA were investigated by Lupas and Verdù in [26]. The general form of a linear CDMA detector is based on bit decisions given by

$$\hat{\mathbf{s}}(k) = \mathrm{dec}(\mathbf{T}\mathbf{r}(k)), \tag{6.12}$$

where $\mathbf{r}(k)$ is the vector-matched filter output, $\mathbf{r}(k) = [r_1(k) \dots r_Q(k)]^{\mathrm{T}}$, and $\mathbf{T}$ is a $Q \times Q$ matrix. Depending on the choice of $\mathbf{T}$, different receivers can be obtained among which the optimal receiver maximizes the asymptotic efficiency of the receiver. Note also that the conventional matched-filter receiver is obtained for $\mathbf{T} = \mathbf{I}$. The following two linear receivers have received some attention.

**Decorrelating Receiver**. Assuming no delay spread and perfect symbol synchronization, the output of the matched receivers can be written as

$$\mathbf{r} = \mathbf{P}\mathbf{s} + \mathbf{n}, \tag{6.13}$$

where $\mathbf{P}_{ij}$ is the cross correlation of the $i$-th channel with the $j$-th user signal. It is then natural to choose $\mathbf{T}$ to be

$$\mathbf{T} = \mathbf{P}^{-1}. \tag{6.14}$$

Since $\mathbf{P}$ may be singular, it is more appropriate to use a pseudo-inverse of $\mathbf{P}$. The bit decisions are given by

$$\hat{\mathbf{s}}(k) = \mathrm{dec}(\mathbf{P}^{+}\mathbf{r}(k)), \tag{6.15}$$

where $(\cdot)^{+}$ denotes the pseudo-inverse. Clearly, the receiver described in (6.15) is not optimal, since the premultiplication of the matched-filter output colors the noise. However, it was shown in [26] that this receiver is optimal in terms of near-far resistance. Its advantages are significant reduction of computational complexity and avoiding the need to estimate channel powers.

**MMSE Receiver**. Just as in temporal equalization, where the MMSE equalizer is used to reduce noise enhancement, and MMSE-type linear receiver can be proposed for multiuser detection instead of the decorrelating receiver. In this case, the bit decisions are given by

$$\hat{\mathbf{s}}_q(k) = \mathrm{dec}(\mathbf{T}\mathbf{r}(k)), \tag{6.16}$$

where each row, $\mathbf{t}_q$ of $\mathbf{T}$ is so chosen to minimize the mean-square error between the transmitted and training signal $\bar{s}_q(k)$

$$\mathbf{t}_q^{\mathrm{H}} = E\left\{\mathbf{r}(k)\mathbf{r}^{\mathrm{H}}(k)\right\}^{-1} E\left\{\mathbf{r}(k)\bar{s}_q^{*}(k)\right\}. \tag{6.17}$$

The MMSE multiuser detector is shown to be near-far resistant and also offers a significant performance improvement over a matched-filter receiver. Moreover, MMSE detector can be more easily structured to be adaptive. In the asymptotic case of background noise going to zero, the MMSE receiver reduces to the decorrelating detector. A drawback of the MMSE receiver is the need for training data. Blind methods can be used to avoid this problem.

### 6.3.2 Space-Time Processing for CDMA

The use of multiple antennas adds a new dimension to the CDMA receiver problem and, similar to TDMA, allows improved separation of user signals. Therefore, space-time processing can significantly increase the capacity of CDMA networks.

The received continuous-time $M \times 1$ signal vector in a multiple antenna CDMA system has the following form

$$\mathbf{x}(t) = \sum_{q=1}^{Q} \mathbf{x}_q(t) + \mathbf{n}(t), \tag{6.18}$$

where the vector signal contribution from a single user is given by

$$\mathbf{x}_q(t) = \sum_{k=-\infty}^{\infty} s_q(k)\mathbf{p}_q(t - kT), \tag{6.19}$$

where $\mathbf{p}_q(t)$ is the vector channel given by the convolution of the vector-channel impulse response, spreading code, and chip waveform, that is

$$\mathbf{p}_q(t) = \mathbf{h}_q(t) \otimes c_q(t), \tag{6.20}$$

where $\mathbf{h}_q(t)$ is the vector-channel impulse response. Once again, for a specular channel we can write

$$\mathbf{h}_q(t) = \sum_{l=1}^{L} \alpha_{lq}\mathbf{a}_{lq}\delta(t-\tau_{lq}) \tag{6.21}$$

where $\mathbf{a}_{lq}$ corresponds to the array-response vector for the $l$-th path from the $q$-th user.

The use of multiple antennas at the receiver therefore has merely converted a scalar channel $h_q(t)$ to a vector channel $\mathbf{h}_q(t)$. The output of a filter matched to the $i$-th sensor for user $q$ is given by

$$r_{iq} = \int_{kT}^{(k+1)T} x_i(t)p_{iq}^*(t-kT)\mathrm{d}t. \tag{6.22}$$

The collection of $M \times Q$ such outputs constitutes the sufficient statistics for ST processing. The techniques mentioned in the previous section for time-only processing can be adapted with care to space-time processing.

The first results for space-time processing in CDMA was proposed by Naguib and Paulaj [27] and was a space-time extension of a RAKE receiver. The algorithm was proposed as a blind approach in the sense that it did not need training sequences. The key ideas are summarized below.

Single-User Space-Time RAKE Receiver

This is an extension of the time-only RAKE receiver discussed earlier. In a space-time RAKE we use a space-time correlator for each path. This consists of a beamformer matched to the path array response vector $\mathbf{a}_{lq}$ followed by a conventional code correlator matched to the path delay $\tau_{lq}$. The output fingers from the code correlator can be summed in a RAKE combiner.

The output of the space-time RAKE for the $q$-th user is given by

$$\hat{s}_q = \mathrm{dec}\left\{\sum_{l=1}^{L}\alpha_{lq}^* \int_{kT+\tau_{lq}}^{(k+1)T+\tau_{lq}} \mathbf{a}_{lq}^{\mathrm{H}}x(t)c_q(t-kT-\tau_{lq})\mathrm{d}t\right\}. \tag{6.23}$$

A better receiver would be to use an MMSE beamformer followed by a conventional RAKE. The output vector of a matched filter $\mathbf{y}_{lq}(k)$, after correlating $\mathbf{x}(t)$ with a code synchronized to the $i$-th path of the $q$-th user is given by

$$\mathbf{y}_{lq}(k) = \int_{kT+\tau_{lq}}^{(k+1)T+\tau_{lq}} \mathbf{a}_{lq}^{\mathrm{T}}x(t)c_q(t-kT-\tau_{lq})\mathrm{d}t = P\alpha_{lq}\mathbf{a}_{lq}s_q(k) + \sqrt{P}\mathbf{q}_{lq}, \tag{6.24}$$

where $\mathbf{q}_{lq}$ is the contribution from paths other than the $l$-th path of user $q$ and from all other user signals. We can show that the covariance matrices of $\mathbf{y}_{lq}(k)$ and $\mathbf{x}(t)$ are given by

$$\mathbf{R}_{xx} = E\left\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\right\} = \|\alpha_{lq}\|^2\mathbf{a}_{lq}\mathbf{a}_{lq}^{\mathrm{H}} + \mathbf{Q}_{lq}, \tag{6.25a}$$

$$\mathbf{R}_{yy,\,lq} = \frac{1}{P}E\left\{\mathbf{y}_{lq}\mathbf{y}_{lq}^{\mathrm{H}}\right\} = P\|\alpha_{lq}\|^2\mathbf{a}_{lq}\mathbf{a}_{lq}^{\mathrm{H}} + \mathbf{Q}_{lq}, \tag{6.25b}$$

where $\mathbf{Q}_{lq}$ represents the covariance of the antenna outputs with the $l$-th path of user $q$ subtracted out. $\mathbf{a}_{lq}$ can be estimated as the generalized principal eigenvector of the matrix pair $(\mathbf{R}_{xx}, \mathbf{R}_{yy,\,lq})$. Similarly, the interference plus noise covariance can be estimated as

$$\mathbf{Q}_{lq} = \frac{P}{P-1}(\mathbf{R}_{xx} - \frac{1}{L}\mathbf{R}_{yy,lq}). \tag{6.26}$$

Using these estimates of $\mathbf{a}_{lq}$ and $\mathbf{Q}_{lq}$, the optimum beamforming weights are

$$\mathbf{w}_{lq} = \frac{\mathbf{Q}_{lq}^{-1}\mathbf{a}_{lq}}{\mathbf{a}_{lq}^{\mathrm{H}}\mathbf{Q}_{lq}^{-1}\mathbf{a}_{lq}}, \tag{6.27}$$

where $\mathbf{w}_{lq}$ is the beamformer for the $l$-th resolved path of the $q$-th user.

The space-time RAKE consists of a beamformer for each path with weights $\mathbf{w}_{lq}$ followed by a standard RAKE. The beamformer is an MMSE type and is followed by a RAKE combiner that does not represent an MMSE time receiver. Fig. 6.6 illustrates a space-time RAKE. Such a receiver leads to an improved figure of merit that can be traded for improved coverage or capacity. This improvement comes from reducing the intra- and intercell CCI through beamforming.

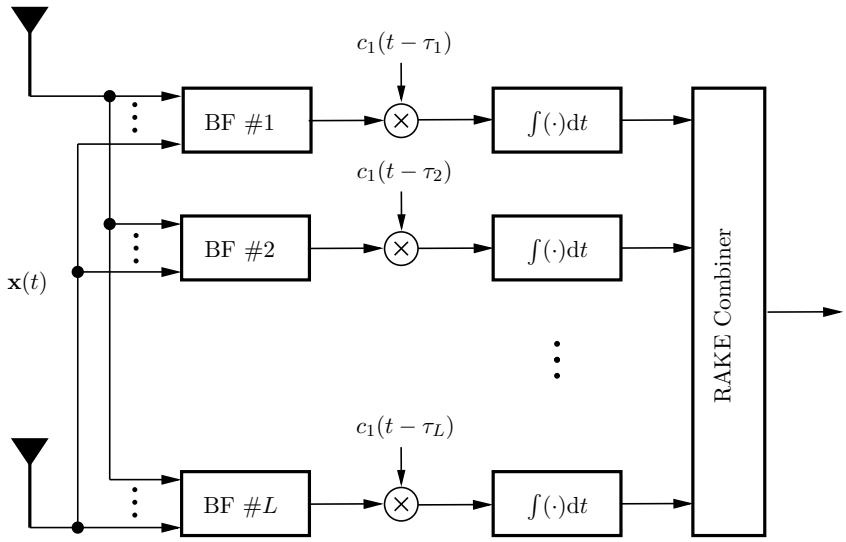



**Figure 6.6:** 2D RAKE receiver.

## 6.4 Forward Link

On forward link, the goal of space-time processing is again to minimize the interference generation to other users while maximizing the energy coupled to the intended user. As mentioned earlier, in CDMA paths separated by more than one chip period are quasi-orthogonal and therefore the individual path-channel estimates are isolated from each other. Hence, the problem of mapping the reverse channel to the forward channel is simpler. We only need to map the spatial channel for each path and in FDD systems the spatial channel path is nearly similar on both links. In the forward-link space-time processing, the signal for each path is transmitted with beamforming to couple into the desired path. Moreover, the individual path signals may be delayed, if necessary, to compensate for the path delays so as to arrive at the user with zero delay. To ensure coherent combining requires accurate estimation of the relative path phases, which may be difficult in practice. Therefore, it may be more pragmatic to transmit with zero-differential delays and use a RAKE receiver at the mobile end. Forward-link beamforming once again increases performance by reducing intra- and intercell interference. This improvement can be traded for improved coverage, capacity or quality of service.

# Chapter 7

# Smart Antennas on Mobile Handsets

## 7.1 Introduction

In this chapter we shall introduce the concept of using smart or intelligent antennas in a mobile handset. In a smart antenna the antenna beam is dynamically changed to enhance the system performance. In particular, by controlling the signal strengths at each element of an array antenna, by changing the weights of an adaptive antenna algorithm, the directivity of the antenna is dynamically controlled. We shall focus on the use of such an antenna on a mobile handset as a way of obtaining a higher capacity of the link.

First, we discuss the possibility of introducing adaptive antennas in a mobile handset, then we give the two examples of already built and tested mobile station adaptive arrays and we finish by discussing the possible research issues that still need to be undertaken.

## 7.2 Mobile Station Adaptive Beamforming

Most of the research effort on the use of adaptive antennas in mobile communications has been concentrated on the base station antennas. The objective of adaptive beamforming is to maximize the signal to interference and noise ratio of the received signal, by maximizing the strength of the desired signal while reducing the adverse effects of interference sources. Although it is a practical and effective technique for reducing the effect of multipath fading, it is rarely used at a mobile station due to cost, size and available power of the mobile station. For hand-held units that are small lightweight pocket communicators, any additional hardware or software must be minimized while maximizing its functionality. Furthermore, it must remain affordable for widespread market acceptance. A single base station often serves hundreds to thousands of mobile stations. Therefore it is more economical to add equipment to base stations rather than at each mobile stations. With the economy of scale on its side, it appears that base station complexity may be plausible trade space for achieving the requirements of the next generation wireless system [8].

However, with operators and manufacturers preparing and deploying the third generation systems, the increasing growth of mobile phone users has created a need for even higher capacity in cellular network. One way of overcoming the capacity problem could be by using multiple adaptive antennas on the handsets. In addition to the higher capacity benefit, this may offer improved efficiency in the following areas [28]

- reduction of multipath fading
- suppression of interference signals
- improvements of call reliability
- lowering the specific absorption rate
- mitigation against dead zones
- increased data rates
- spectral efficiency.

New technologies under development, i.e., small solid state antenna as manufactured by Antenova [29] and the Quadrifilar helix antenna (QHA) produced by Surrey University, have made the adaptive antenna on handsets a practical possibility. In the next section we briefly show some details of the manufactured examples.

## 7.3 Two Types of Mobile Handset Adaptive Antennas

### 7.3.1 The Quadrifilar Helix Antenna

The Intelligent Quadrifilar Helix Antenna was invented in the University of Surrey by A. A. Agius, S. M. Leach, P. Suvannapattana and S. R. Saunders (UK Patent Application No 9803273.3 February 1998). The schematic configuration of the antenna is shown in Fig. 7.1.

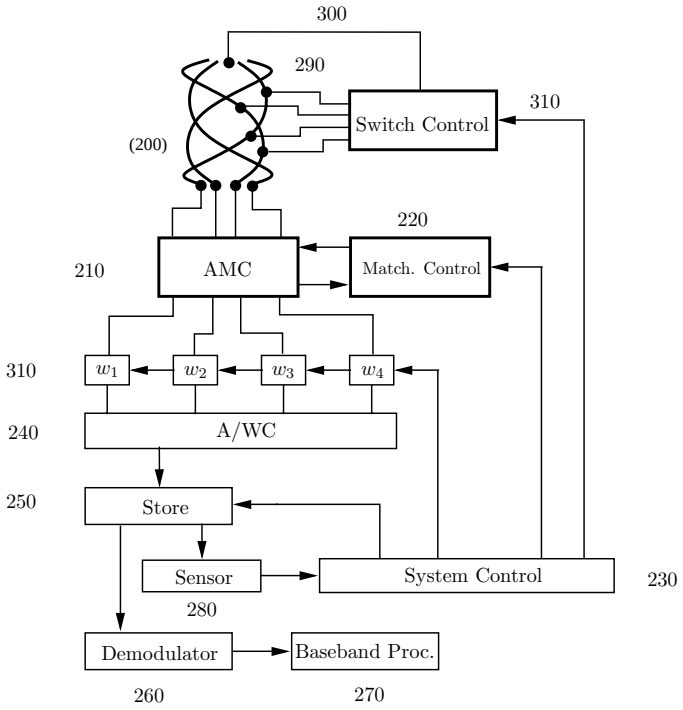
Specifications of the antenna are listed in the following [28]

- The QHA consists of four twisted wires in a form of a helix
- It permits reception inside the buildings where signals are too weak to be picked up by standard antennas
- Designed to improve performance and minimize radiation into the human head
- The signal is produced through use of a torch beam which means that instead of radiation going in all directions, it will use a narrow beam traveling towards the nearest base station or satellite.
- Reduces the required transmit power by a factor of 10, depending on the configuration of buildings and trees around the user
- It makes batteries last longer

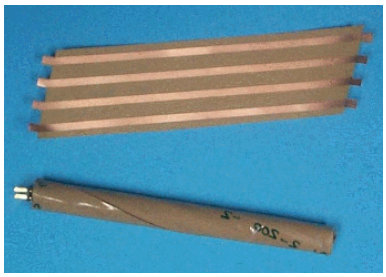
### 7.3.2 The solid state antenna

Specifications of the antenna are listed in the following [28, 29]

- Surface mounted to motherboard
- Approximately one tenth size of conventional copper antenna




**(a)** Schematic configuration.



**(b)** Basic Structure of printed QHA.

**Figure 7.1:** The Intelligent Quadrifilar Helix Antennas (I-QHA) configuration [30].

- The antenna is steerable (not multi-beam or switched beam)
- Life of the battery is increased as the handset only points in the direction of the base station so more efficient transmission is achieved
- Minimizes RF health issue by directing the RF emissions away from the head or body
- Antenna avoids detuning by the proximity of the user.

## 7.4 Research Issues

The first research issue is cost, including cost of power [7]. Researchers in Philips noted that 50% of the power in the handset is in the RF electronics. Therefore, multiple antennas in the handset not only increase the dollar cost of the handset, but also increase the power and thus decrease battery life. Research to reduce the power that each of these antennas requires needs to be undertaken.

The second key research issue is the size. Multiple external antennas on terminals are generally not practical. Therefore, some companies are researching putting antennas on the RF electronics IC in an "antenna-less" terminal (since the external antenna is not present). However, issues of gain and efficiency and the effect of hand placement on the terminal need further research.

The third issue is diversity, which is needed for multipath mitigation. Spatial diversity is difficult to achieve on a small handset. At Philips and other companies, researchers are using dual polarization diversity on handsets. Researchers at Nokia are studying the use of multiple antennas in the handset, where some of the antennas may be covered by the hand, and moving the hand around changes

the antenna pattern. These researchers believe that by adaptively combining the signals from such antennas, perhaps only using antennas not blocked by the hand or adjusting the antenna impedance to compensate for hand placement, it may be possible to obtain much better performance (including diversity) with multiple internal antennas as compared to an external antenna.

# Chapter 8

# Multiple Input - Multiple Output (MIMO) Communications Systems

## 8.1 Introduction

The adaptive antenna concept has now long been proved to bring a capacity increase of cellular networks [31]. Adaptive antennas were considered as a means to steer a beam towards the user, increasing the capacity by decreasing the interference level at the network level. The underlying property that was used is the *correlation* between the various sensors of the array.

A much older concept is that of diversity: two sensors receiving *decorrelated* signals. Space and/or polarization diversity at the base station receiver is widely deployed in present cellular networks, as a means to balance the link budget between a low power terminal and a strong power base station. However, this concept has recently received renewed attention.

The idea to separate signal transmission in a temporal and spatial domain is not new. This can be found in systems which use multiple antennas at the receive side (base station) and a single antenna at the transmit side (terminal) – so called single input - multiple output (SIMO) systems. Such technology is known as receive diversity and lead together with maximum ratio combining algorithm to higher capacities [32].

The throughput can be increased even further if an antenna array is used not only at the base station, but also at the terminal. Multiple antenna concepts are a further extension of diversity concepts [33]. They exploit independent multipath transmissions between the different antenna elements on the base station and the terminal side. Such MIMO wireless channels have a significantly higher capacity than single input - single output (SISO) wireless channels. They are particularly interesting for laptop type terminals that require high data rates and operate at higher carrier frequencies (e.g., in the 5 GHz range). Multiple antennas can easily be installed on such a terminal due to the fact that the antenna array requires a smaller physical size (aperture) at higher frequencies.

In this chapter we will present the signal model for MIMO wireless channel, review capacity expressions for SISO/MIMO frequency-nonselective/selective channels, and show some external simulations and measurements to prove the theoretical idea of capacity improvement.

## 8.2 Capacity for a Given Channel Realization

The increase in spectral efficiency offered by MIMO systems is based on the utilization of space (or antenna) diversity at both the transmitter and the receiver. With a MIMO system, the data stream from a single user is demultiplexed into $n_T$ separate substreams. The number $n_T$ equals the number of transmit antennas. Each sub-stream is then encoded into channel symbols. It is common to impose the same data rate on all transmitters, but adaptive modulation rate can also be utilized on each of the sub-streams (see Fig. 8.1).

With this transmission scheme, there is a linear increase in spectral efficiency compared to a logarithmic increase in more traditional systems utilizing receive-diversity or no diversity. The high spectral efficiencies attained by a MIMO system are enabled by the fact that in a rich scattering environment, the signals from each individual transmitter appear highly uncorrelated at each of the $n_R$ receive antennas. When the signals are conveyed through uncorrelated channels between the transmitter and receiver, the signals corresponding to each of the individual transmit antennas have attained different spatial signatures. The receiver can use these differences in spatial signature to simultaneously and at the same frequency separate the signals that originated from different transmit antennas.

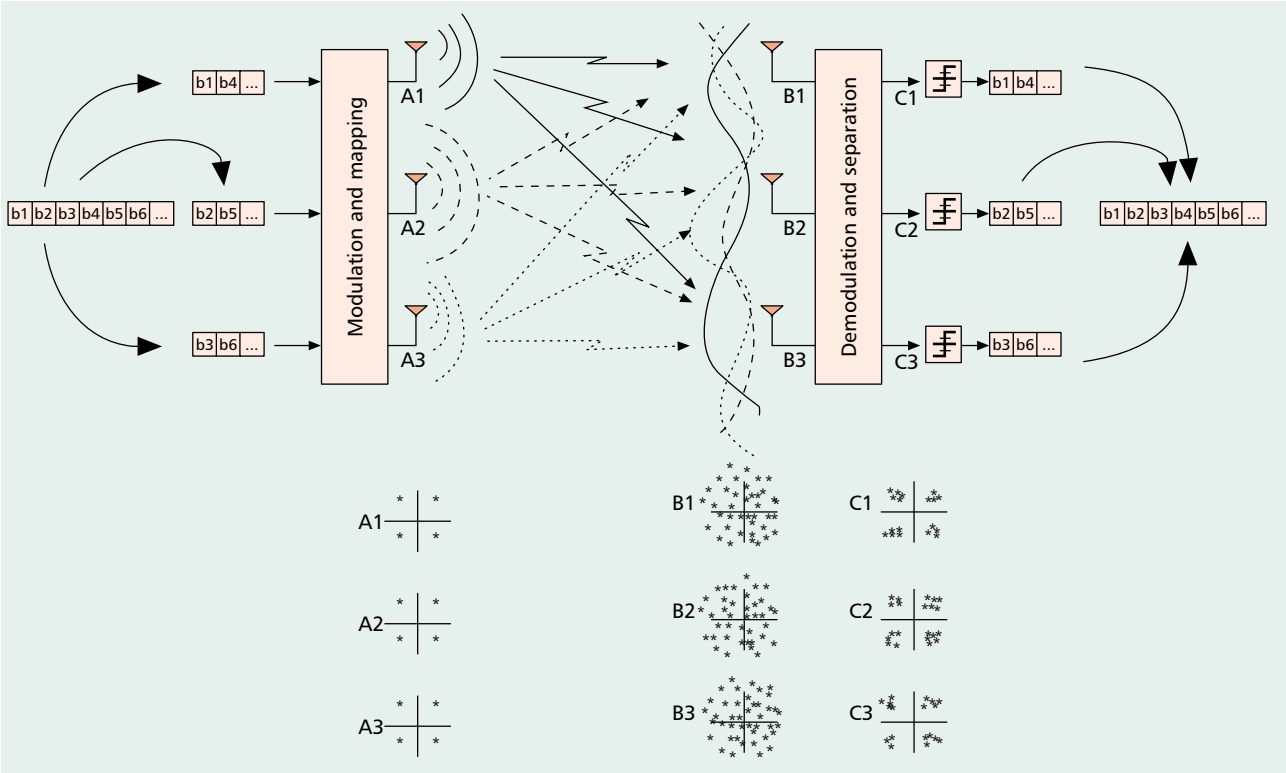



**Figure 8.1:** A basic MIMO scheme with three transmit and three receive antennas yielding three-fold improvement in system capacity [34].

A measure of how much information that can be transmitted and received with a negligible probability of error is called the channel capacity. The capacity of a channel depends completely on the channel realization, noise and transmitted signal power (along with its distribution). In this section, the expression of the capacity is reviewed for SISO/MIMO frequency-nonselective/selective channels [35].

### 8.2.1 Capacity of a SISO Channel

For a single-input single output (SISO) channel, the received signal model is given by

$$x(t) = \alpha u(t) + n(t), \tag{8.1}$$

where $u(t)$, $x(t)$ and $n(t)$ are the transmitted, received and noise signals respectively, and $\alpha$ is the channel fading. The capacity for such a model as was derived by Shannon in 1948. is given by

$$C = \log_2\left(1 + \frac{P}{\sigma_n^2}|\alpha|^2\right) \quad [\text{bits/sec/Hz}], \tag{8.2}$$

where $P$ is the transmitted power and $\sigma_n^2$ is the white Gaussian noise power within the channel band $B$.

In case the channel is frequency-selective, the received signal can be expressed as

$$x(t) = \sum_{l=1}^{L} \alpha_l u(t - \tau_l) + n(t), \tag{8.3}$$

where $L$ is the number of multipaths, $\alpha_l$ is the complex path fading and $\tau_l$ is the path delay.

After subdividing the signal into flat-fading frequency bands as

$$X(f) = \alpha(f)U(f) + N(f), \tag{8.4}$$

where now $X(f)$, $\alpha(f)$, $U(f)$ and $N(f)$ stand for the Fourier transformed $x(t)$, $\alpha_l$, $u(t)$ and $n(t)$, the capacity expression is given by integrating over the utilized bandwidth $B$

$$C = \frac{1}{B}\int_B \log_2\left(1 + \frac{P(f)}{\Phi(f)}|\alpha(f)|^2\right) \mathrm{d}f, \tag{8.5}$$

with $\Phi(f)$ being the noise power spectral density, $B$ being the transmitting bandwidth and $P(f) = E\left\{|U(f)|^2\right\}$ being the signal power spectral density. The transmitted power is constrained by $\int_B P(f)\mathrm{d}f \leq P_T$.

### 8.2.2 Capacity of a MIMO Channel

The received signal model for the case of a flat single-user MIMO channel is given by

$$\mathbf{x}(t) = \mathbf{H}\mathbf{u}(t) + \mathbf{n}(t), \tag{8.6}$$

where $\mathbf{u}(t)$ is an $n_T \times 1$ transmit vector, $\mathbf{x}(t)$ is an $n_R \times 1$ receive vector, $\mathbf{H}$ is an $n_R \times n_T$ channel matrix and $\mathbf{n}(t)$ is an $n_R \times 1$ additive white Gaussian noise (AWGN) vector at a given instant in time. In the following, it is assumed that the channel is memoryless (that is, for each use of the channel an independent realization of $\mathbf{H}$ is drawn) and that the receiver has perfect channel knowledge.

The capacity of a flat MIMO channel is given by the following expression

$$C = \log_2\left[\det\left(\mathbf{I} + \frac{1}{\sigma_n^2}\mathbf{H}\mathbf{Q}\mathbf{H}^{\mathrm{H}}\right)\right], \tag{8.7}$$

where $\mathbf{Q} = E\left\{\mathbf{u}(t)\mathbf{u}(t)^{\mathrm{H}}\right\}$ is the covariance matrix of the transmitted signal. The transmission power constraint can be expressed as $\mathrm{trace}(\mathbf{Q}) \leq P_T$.

For the case of frequency selective MIMO channels, the capacity is computed by integrating over the utilized bandwidth

$$C = \frac{1}{B}\int_B \log_2\left[\det\left(\mathbf{I} + \frac{1}{\Phi(f)}\mathbf{H}(f)\mathbf{Q}(f)\mathbf{H}^{\mathrm{H}}(f)\right)\right], \tag{8.8}$$

with the power constraint as $\int_B \mathrm{trace}(\mathbf{Q}(f))\mathrm{d}f \leq P_T$.

### 8.2.3 Capacity as a Random Variable

We presume that the communication is carried out using bursts. The burst duration is assumed to be short enough so that the channel can be regarded as essentially fixed during a burst, but long enough that the standard information-theoretic assumption of infinitely long code block lengths is a useful idealization. In this quasi-static scenario, it is meaningful to associate a channel capacity with a given realization of the channel matrix $\mathbf{H}$.

Since the channel capacity is a function of the random channel matrix, it can be regarded as a random quantity whose distribution is determined by the distribution of $\mathbf{H}$. In such cases, an important measure for the channel capacity is the channel capacity at a given outage probability $q$, denoted by $C_{out}$. It simply means that the channel capacity is less than $C_{out}$ with probability $q$ or, in other words, it is greater than $C_{out}$ with probability $(1-q)$. In the following, capacity results for different system configurations will be given by means of cumulative distribution functions (CDF) of the capacity, expressing it as the probability $(1-q)$ that the capacity is greater than $C_{out}$.

### 8.2.4 Power Allocation Strategies

Unlike the case of a flat SISO channel, where there is only one available channel, for the case of frequency-selective and/or MIMO channels, the available transmission power can be distributed over the antennas and/or frequency bands according to different strategies. The power allocation techniques will depend on the knowledge of the channel.

#### Uniform Distribution

The uniform distribution of the available transmission power has to be used when the channel is unknown for the transmitter. Note that the channel is always assumed to be known by the receiver. The capacity expression for the flat MIMO channel is then

$$C = \log_2\left[\det\left(\mathbf{I} + \frac{P/n_T}{\sigma_n^2}\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)\right] \tag{8.9}$$

and for a frequency-selective MIMO channel

$$C = \frac{1}{B}\int_B \log_2\left[\det\left(\mathbf{I} + \frac{P/(n_T B)}{\Phi(f)}\mathbf{H}(f)\mathbf{H}^{\mathrm{H}}(f)\right)\right]. \tag{8.10}$$

#### Water-filling

In some situations, such as when channel reciprocity between uplink and downlink can be applied (TDD) or explicit feedback information is used, the channel is known at the transmitter and, there-

fore, optimum distribution of power over the antennas and frequency bands can be performed. The maximization of the capacity gives a power distribution technique commonly referred to as "water-filling" or "water-pouring" because it resembles the act of filling a bowl.

The "water-filling" technique can be easily derived after performing the singular value decomposition of the channel matrix $\mathbf{H} = \mathbf{UDV}^{\mathrm{H}}$ and expressing the flat MIMO channel as a set of $L = \min(n_T, n_R)$ parallel channels,

$$\tilde{\mathbf{x}}(t) = \mathbf{D}\tilde{\mathbf{u}}(t) + \tilde{\mathbf{n}}(t), \tag{8.11}$$

or, equivalently, as

$$\tilde{y}_k(t) = \lambda_k^{1/2}\tilde{u}_k(t) + \tilde{n}_k(t), \qquad 1 \leq k \leq L, \tag{8.12}$$

where $\lambda_k$ is the gain of channel $k$. The optimum power allocation for the set of parallel channels is given by

$$p_k = K - \frac{\sigma_n^2}{\lambda_k}, \tag{8.13}$$

where $K$ is a constant to meet power constraints. Note that for frequency selective MIMO channels, the optimum power allocation has to be done simultaneously over the set of parallel channels resulting from the spatial channel parallelization and the orthogonal frequency bands.

## 8.3 Simulation Examples

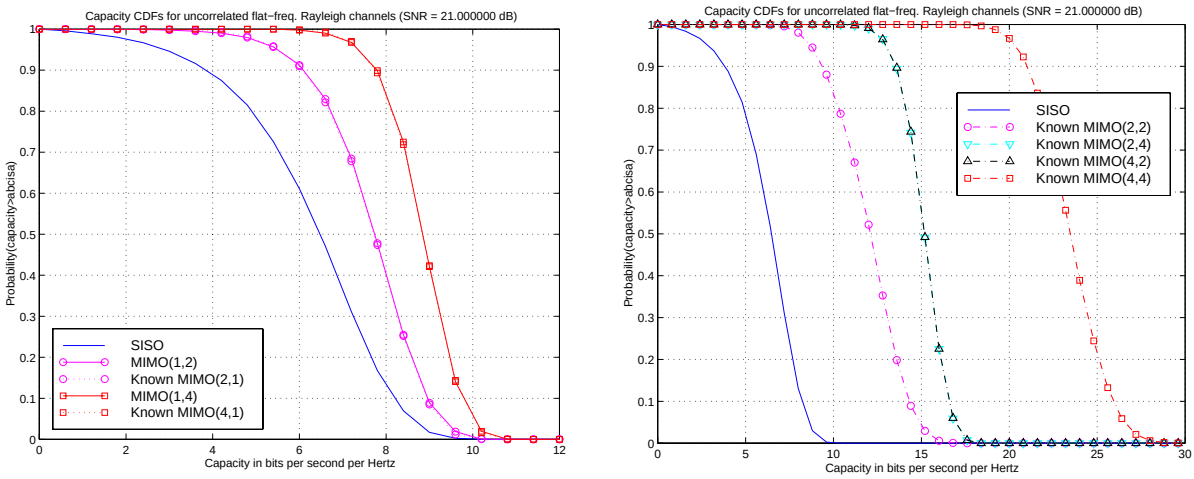
In this section capacity results for the different system configurations are given as curves of CDF or outage probability and can be found in [35]. They are computed using Monte-Carlo simulations based on 10000 random channel realizations. Different configurations will depend on many factors such as the number of transmit and receive antennas ($n_T$, $n_R$), whether the channel is known or not (it is always assumed known by the receiver), whether the fading is fully correlated or completely uncorrelated, the transmission power, and the frequency-selectivity of the channel (or its power delay profile). Note that, unless stated otherwise, the transmitted power is kept constant for fair comparisons (SNR = 21 dB at each receiving antenna independently of the number of transmit antennas).

### 8.3.1 Capacity of Flat SIMO/MISO vs. MIMO Channels

As illustrated in Fig. 8.2(a), where some capacity curves for different configurations of uncorrelated flat SIMO and MISO Rayleigh channels are plotted, the capacity increases with the number of either transmitting or receiving antennas. Nevertheless, when arrays of antennas are utilized both in transmission and reception simultaneously (see Fig. 8.2(b)), that is when MIMO channels are used, the capacity boosts (note that for the case of (4,4) the capacity is $C = 21$ bits/sec/Hz at $P_{out} = 0.1$). This significant increase of capacity is due to the existence of parallel channels, which do not exist in SIMO/MISO channel configurations.

### 8.3.2 Capacity as a Function of the Fading Correlation

The effect of the fading correlation can be seen comparing Fig. 8.2(b) with Fig. 8.3 (for (4,4) case, the capacity decreases from 21 to 8 bits/sec/Hz at $P_{out} = 0.1$). It can be seen how the huge potential capacity of the uncorrelated channels vanishes when the channel becomes fully correlated. The explanation for that substantial difference is the fact that, when the channel gets correlated, the number




**(a)** SIMO/MISO channels. **(b)** MIMO channels.

**Figure 8.2:** Flat uncorrelated channels.

of parallel channels decreases to the point of having just one single channel, which corresponds to the fully correlated case. In such cases, the capacity gain is obtained only by beamforming.

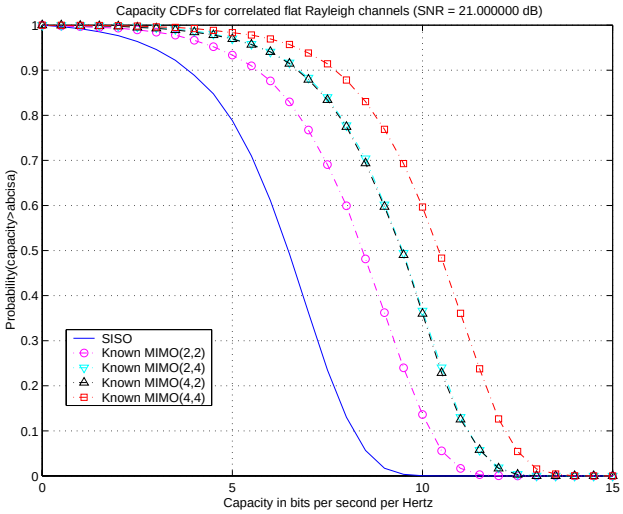



**Figure 8.3:** Flat correlated MIMO channels.

### 8.3.3 Capacity as a Function of the Transmitted Power

In Fig. 8.4, capacity CDF curves are plotted for the uncorrelated flat MIMO (4,4) case as a function of the transmitted power (or equivalently, the received average SNR at each antenna element). Note that for high SNR values, where the capacity of the SISO channel increases 1 bit per 3 dB, the capacity of an uncorrelated $(n, n)$ channel increases $n$ bits per 3 dB increase of SNR (see (8.9)).

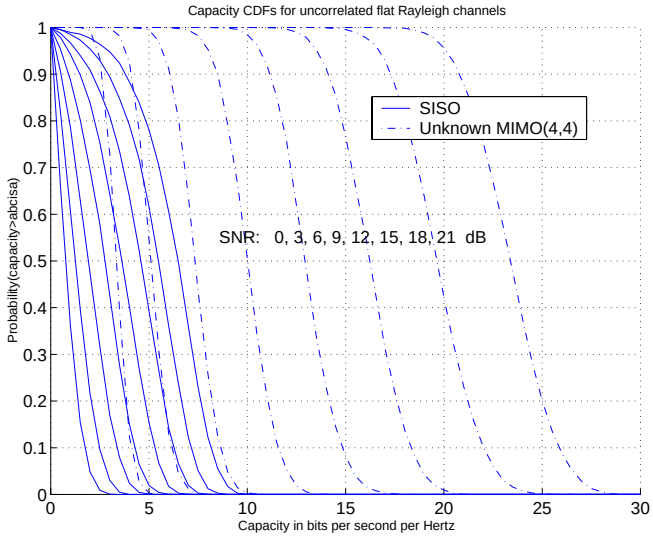



**Figure 8.4:** Capacity for different SNRs.

### 8.3.4 Capacity as a Function of the Number of Antenna Elements

In Fig. 8.5, the capacity CDF curves of flat MIMO $(n, n)$ channels are plotted. As predicted by (8.9), the capacity grows without limit as $n$ increases for the case of uncorrelated channel (actually, for large $n$ increases at least linearly).

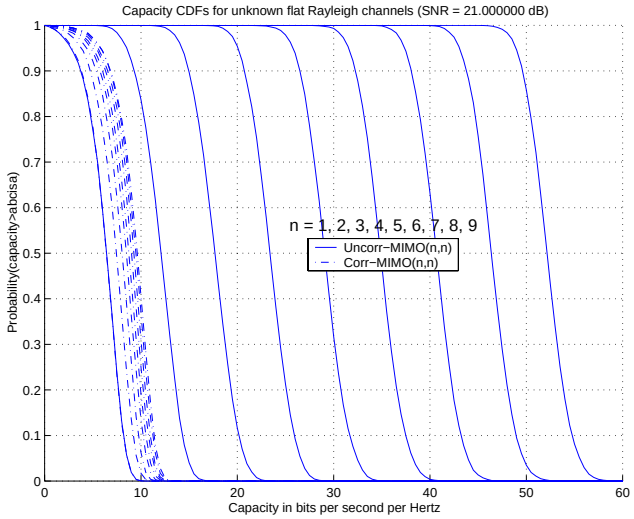
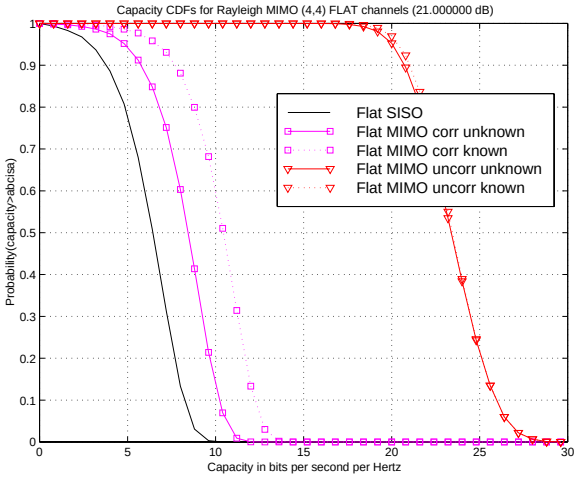



**Figure 8.5:** Capacity for different number of antennas.

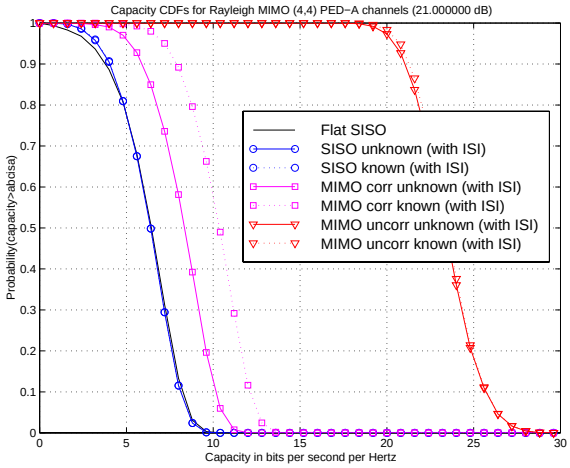
### 8.3.5 Capacity as a Function of the Frequency-Selectivity of the Channel

In Fig. 8.6 capacity CDF curves for a MIMO (4,4) configuration over a flat channel and two frequency-selective channels (with delay profiles PED-A and VEH-A, according to ETSI) are plotted. It can be

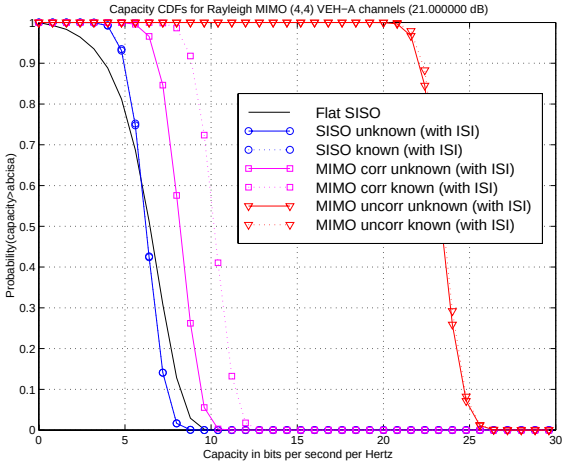
observed how the increase of frequency diversity increases the slope of the capacity curves, but does not shift it (as for the increase of $n$ in the $(n, n)$ case) improving the capacity at low probabilities of outage.




**(a)** Capacity of a Rayleigh flat MIMO channel.




**(b)** Capacity of a Rayleigh MIMO PED channel.




**(c)** Capacity of a Rayleigh MIMO VEH channel.

**Figure 8.6:** Capacity of Rayleigh MIMO channels.

### 8.3.6 Summary

The present analysis of the capacity of Rayleigh MIMO channels can be summarized as follows

- The capacity of a MIMO system generally decreases as the fades of the MIMO channel become more correlated or, in other words, as the angular spread decreases
- For the case of uncorrelated fading, there is a large amount of capacity available that increases linearly with $n$ in $(n, n)$ systems for large $n$ and with the transmitted power ($n$ bits per 3 dB increase in high SNR regime)

- The performance of a wireless communication system that has multiple transmit-receive antennas depends on how the transmitted power is distributed among the parallel channels. The difference between the capacities achieved by uniform and optimum power allocation is small when the fades associated with transmit-receive antenna pairs are independent, but can become very large when the fades are highly correlated. Therefore, the additional complexity of optimum power allocation over uniform power allocation is justified only in the correlated channel case

- The frequency-selectivity of MIMO channels increases the slope of the capacity CDF curves.

## 8.4 MIMO in Wireless Local Area Networks

The introduction of Wireless Local Area Networks (WLAN) motivates the use of multiple antennas on both transmitter and receiver sides (MIMO). The terminal in such a WLAN could be a laptop computer giving opportunity to carry multiple antennas. Examples of WLAN systems are IEEE 802.11 and HiperLAN/2 where standardization is ongoing. Versions of both systems today offer rates of more than 50 Mbits/s with a single terminal antenna. However, extremely good Signal-to-Noise-ratios are required. Multiple antennas on the terminal side are believed to increase the rate further and to relax the SNR requirements.

The mentioned WLAN systems both have versions for operation in the 5 GHz band. The propagation characteristics at this frequency are very appropriate for Radio LANs and it is possible to cover many users at a low cost. It is also possible to use such systems in both indoor and outdoor environments and thus provide coverage of a hot-spot area such as a campus or an airport. The results shown here are taken from [36] and are focused on the measured and simulated link capacity at 5.8 GHz (WLAN).

### 8.4.1 Channel Measurement Set-Up

The measurements were carried out by Telia Research at their office, located at the Scandinavian Center in Malmö, Sweden. The general planning of the floor consists of office rooms, open spaces and corridors. Most spaces are separated by walls, while glass is used in some of them. The transmitter was positioned in one of the offices, while the receiver was in an open area. The measured channel that was analyzed was typical non-line-of-sight (NLOS) situation and the distance from transmitter to receiver was 10-15 m.

The measurements were made at 5.8 GHz carrier frequency and the transmitter and receiver bandwidth was 400 MHz. By sending a pseudo-noise sequence at the transmitter and correlating with the same synchronous pseudo-noise sequence at the receiver, complex impulse responses were measured.

The data is measured with a synthetic array antenna, using one receive and one transmit antenna, both of monopole type. The transmit antenna is moved between seven different positions separated by 300 mm, that is, about $6\lambda$ (of these seven positions, three are used herein). For each of the seven transmitter positions, the receive antenna is moved between 21 different positions, using a step motor on a track (distance between two adjacent positions was about $\lambda/4$). This corresponds to spatial measurements over in total about $5\lambda$. At each combination of transmit and receive positions 20 samples are taken. All measurements have been performed during stationary conditions at night, and the measurement noise is assumed to be very low.

### 8.4.2 Results

The capacity for different signal-to-noise ratios, with three receive elements and one, two and three transmit elements is shown in Fig. 8.7. Note the substantial increase in capacity as the number of transmit elements increases from one to three. As a reference the simulated curve for the IID (Independently and identically distributed) Gaussian Channel is plotted as well, indicating low correlation among the elements of the channel matrix **H** under these conditions. The following results are given for the SNR fixed at 20 dB.

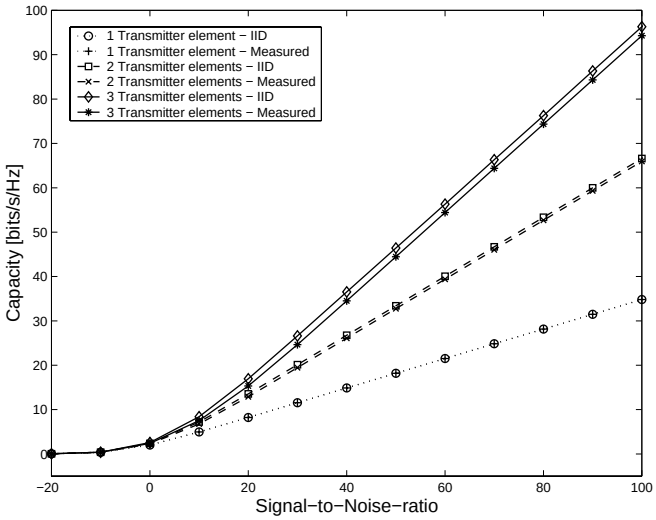



**Figure 8.7:** Capacity for different SNR, with 3 receive elements and one two and three transmit elements, respectively.

Increasing the number of receive and transmit antennas will also increase the capacity. This increase is shown in Fig. 8.8. The capacity increase is large when going from one to the number of elements in the transmit array $n_T = 3$, while the increase is less when further increasing the number of elements on the receive side. As expected, the increase for $n_R > n_T$ follows a logarithmic curve, i.e. it is due to an SNR increase from noise averaging over the channels.

By picking out three of the receive elements in the receive array, it is possible to see how the intra-cell distance influences the capacity. In Fig. 8.9 the capacity is shown for different number of elements as the intra-element distance is increasing. In Fig. 8.10 the capacity dependence of the intra-element distances is shown for two and three receive elements and three transmit elements. It is shown that the capacity increase is small when increasing the distance between the elements beyond $\Delta = \lambda$. After comparing the increase to simulations with random IID channels, which provides a statistical upper bound for the channel capacity, it has been concluded that for this experimental setup, the sub-channels are close to uncorrelated when the intra-element distance is about $2\lambda$.

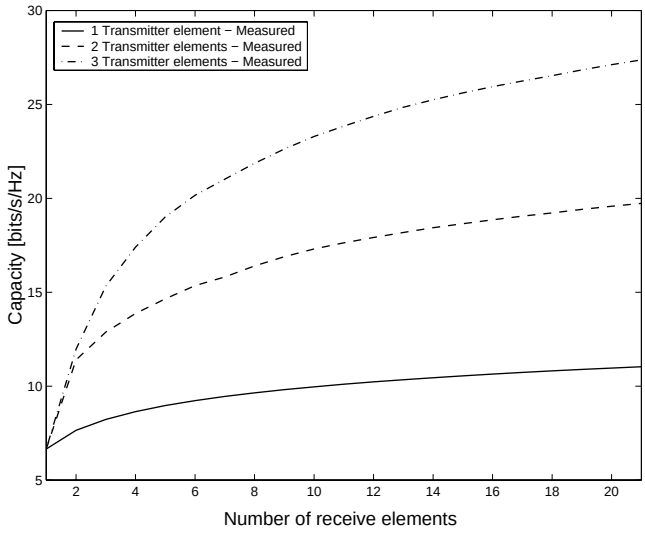



**Figure 8.8:** Capacity as a function of number of receive elements for one, two and three transmit elements, respectively, and an SNR of 20 dB.

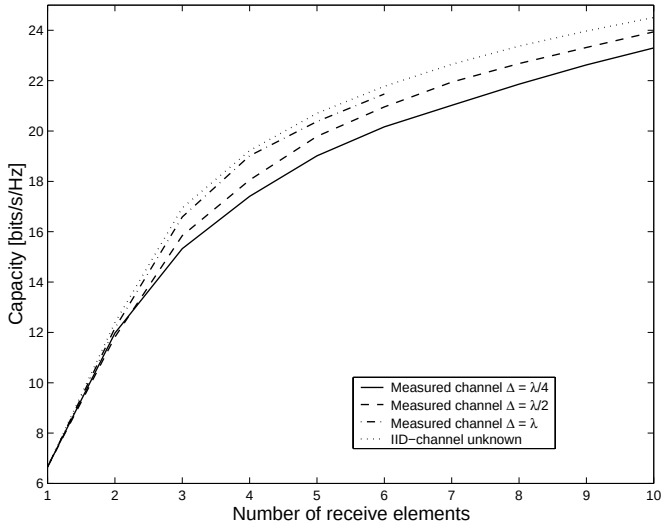



**Figure 8.9:** Capacity dependence on the number of elements in the receive array for different element distances. The SNR is 20 dB and three elements are used in the transmit array.

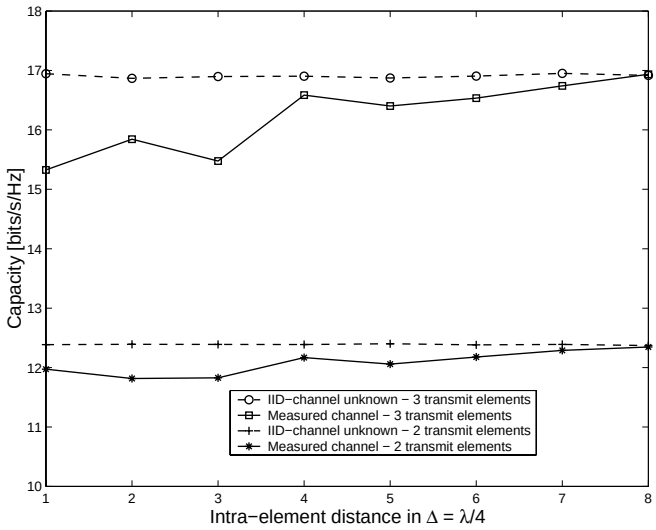



**Figure 8.10:** Capacity dependence on the intra-element distance at the receiver, for the measured channel compared to the simulated IID channel. Two and three receive elements are used together with three transmit elements. SNR=20 dB.

## 8.5 Conclusions

Smart antenna and MIMO technologies have emerged as the most promising area of research and development in wireless communications, promising to resolve the traffic capacity bottlenecks in future high-speed broadband wireless access networks [37].

In a multipath environment, the received power level is a random function of the user's location and, at times, fading occurs. When multiple antenna elements are used, the probability of losing the signal altogether decreases exponentially with the number of de-correlated signals (or antennas). The diversity scheme common in current SIMO (or MISO) wireless LAN (WLAN) systems uses a simple switching network to select the antenna that yields the highest SNR out of an array of two antennas.

MIMO systems can turn multipath propagation – usually a pitfall of wireless transmission – into an advantage for increasing the user's data rate.

In this chapter the information theoretical background for such new communication systems is investigated. The MIMO capacity formulas are shown and explained. Simulations and measurements taken from [35] and [36] are presented to show the capabilities of MIMO systems. It has been shown that the capacity of a MIMO system increases linearly with the $n = \min(n_T, n_R)$. However, the effect of spatial fading correlation on the capacity is significant where the angle spread and antenna spacing have a significant degrading influence if they decrease.

In a conclusion, it can be said that MIMO principle is able to provide future communication systems with dramatically increased capacity using the same bandwidth and transmit power as today. However, to date, a lot of questions and problems are open and need to be solved.

# Chapter 9

# Existing Smart Antenna Experimental Systems and Commercially Available Products

System-level field trials, which involve several GSM/DCS base stations with smart antennas, are in the focus of Ericsson - Mannesmann cooperation [38]. The system will experience a full commercial traffic load in the near future. In this experiment, SA receivers use eight elements, a dual-polarized array SA with DoA based beamformers at the uplink and downlink. Improvements in the carrier-to-noise (C/N) ratio in the order of 45 dB for uplink and downlink were reported. In rural and urban macrocells, the SA receiver provides an additional 10 and 6 dB, respectively, in the CIR. Based on the experiment, 100%-200% capacity gain is reached, and achievable range extension is determined by 4-5 dB C/N gain, which is equivalent to 50% fewer sites.

Another test bed was built by the international team of Ericsson, Stockholm, Sweden/Research Triangle Park, NC, for study of SA receiver performance for digital-advanced mobile phone service (D-AMPS) [39]. The uplink receiver uses space and polarization diversity. The antenna elements have 15 wavelength separations. Two types, a maximum ratio-combining (MRC) receiver and interference rejection-combining (IRC) receiver, were studied. Combined space and polarization approaches provided 3.5 dB gain in the C/N ratio and, additionally, 5 dB gain with IRC in an interference-limited scenario. The fixed-beam approach was used at the downlink.

A four-element adaptive antenna array (i.e., TRB) test bed with a DMI (Direct Matrix Inversion) algorithm was designed by AT&T, Holmdel, NJ [40] for evaluation of the SA concept in an IS-136 system operating at 850 MHz/1.9 GHz. A 5 dB higher gain was achieved at $10^{-2}$ BER in a Rayleigh fading environment compared to two-element antenna diversity. This corresponds to a 40% increase in range. It was shown that SA could maintain $10^{-2}$ BER when the interference level was near the level of desired signal with fading rates corresponded to 100 km/h. The power control performance was studied with a switched beam SA at the down-link.

NTT DoCoMo, Yokosukashi, Japan, is developing an SA experimental system for the third-generation UMTS wide-band CDMA (W-CDMA) network [41]. The 2-D RAKE receiver includes an MMSE beamformer, which tentatively will exploit user-dedicated pilot and recovered data symbols. There

are three cell sites in the experimental system, and it allows the evaluation of hand-over and other network functions. The first experimental results showed a substantial improvement in average BER with the SA compared to spatial diversity.

Participants of the SUNBEAM (formerly, TSUNAMI) ACTS Project are using an SA test bed designed by Era Technology Ltd., Leatherhead, Surrey, U.K. [2]. AoA is estimated by the MUSIC algorithm and Kalman filtering for tracking. The digital enhanced cordless technology (DECT) air interface was selected for trials since it can be easily integrated into an SA and allows networking aspect to be neglected. Two independent SDMA channels were supported. The uniform linear array (ULA) consists of eight elements.

An SA prototype of the SDMA system for a GSM/DCS1800 network was developed and tested by Thompson-CSF Communications, Cennevilliers, France, and CNET/France Telecom, Issyles Moulineaux, France [42]. The SA receiver consists of ten elements and a digital BF in the uplink and downlink. In test trials, three mobiles communicated in the same FDMA/TDMA channel. The MUSIC algorithm was used for AoA estimation. Such parameters as minimum angular separation, maximum dynamic signal separation, and achievable level of interference rejection were studied.

The Circuit and System Group, Uppsala University, Uppsala, Sweden, and Ericsson Radio Access AB, Stockholm, Sweden, built a ten-element experimental SA [42]. Real traffic data taken from a DCS 1800 network were used, and a spatial multiplexing concept was evaluated. 30 dB in CIR was obtained in a line of sight (LOS) propagation scenario. It was observed that different spatial signatures and low cross correlation between training are enough for separation, even for signals with the same angular position in the presence of CCI [43]. It was possible to maintain error-free transmission with minimum of a 10 degrees angular between desired and interfering signals when CIR $= -20$ dB.

The SPOTLIGHT Metawave Company, Redmond, WA, switched-beam system with 12 beams at the uplink and downlink is among the first commercially available products. SPOTLIGHT can be installed in CDMA IS-95 and AMPS networks and, according to Metawave, a 30% capacity improvements in IS-95 can be obtained [44].

Raytheon E-Systems, Fall Church, VA, introduced the "Fully Adaptive Digital Signal Processor System" based on an eight-element SA [45]. It is expected that the SA module can be directly connected to the RF input of the existing BS.

ArrayComm, San Jose, CA, offers a four-element SA for a wireless local loop (WLL) and personal handyphone system (PHS), which is similar to the DECT system in Europe [46]. During field trials that involved GSM protocols, interference mitigation of 20 dB was achieved.

Wireless Online, Santa Clara, CA is offering ClearBeam® 900 Smart Appliqué system for 900 MHz GSM Networks [47]. Global field deployments have demonstrated up to 18 dB improvement in Carrier-to-Interference ratio, two times capacity increase, three times coverage improvement, 60% reduction in dropped calls when compared to sectorized antenna system.

A listing of experimental SA systems and commercial products is presented in Table 9.1 and Table 9.2.

| Designer | Air Interface | Antenna (M) | SA Algorithm | Remarks | Ref. |
|---|---|---|---|---|---|
| **SA experimental systems** | | | | | |
| Ericsson Mannesmann Mobilfunk | GSM/DCS 1800 | 8 | **Uplink**: DoA based<br>**Downlink**: DoA SB and AA | Several BS equipped with SA integrated into network | [38] |
| Ericsson Research (SW/US) | IS-136 (D-AMPS) | spacing up-link $15\lambda$ & pol. div. | **Uplink**: MRC and IRC,<br>**Downlink**: fixed beams | - | [39] |
| AT&T Labs-Research (US) | IS-136 | 4 | **Uplink**: 4 branch adaptive TRB DMI algorithm,<br>**Downlink**: Switched beam with or without PC (up to 3 beams) | up and down links independent | [40] |
| NTT DoCoMo (Japan) | UMTS | 6 | **Uplink**: Decision directed MMSE (tentative data and pilot)<br>4 finger 2D-RAKE,<br>**Downlink**: calibration of weights generated for reverse link | - include 3 cell sites<br>- data transmission up to 2 Mbps | [41] |
| TSUNAMI (Sunbeam) Consortium (EU) | DECT SDMA DCS1800 | - | ULA-MUSIC for DoA estimation, Kalman filtering for tracking | SDMA based on DECT was studied | [2] |
| CNET & CSF-Thompson (F) | GSM/ DCS1800 SDMA | 10 circular | **Uplink**: DoA based Capon, MUSIC for DoA estim.<br>**Downlink**: DoA based | - | [42] |
| Uppsala University (SW) | DCS1800 SDMA | 10 circular | **Uplink only**: TRB with SMI | Data traffic from DCS-1800 was used | [43] |

**Table 9.1:** List of experimental SA systems [18].

| Designer | Air Interface | Antenna (M) | SA Algorithm | Remarks | Ref. |
|---|---|---|---|---|---|
| **Commercially available products** | | | | | |
| Metawave (US) Spotlight2000 | AMPS CDMA | 12 | **Up- and Downlink**: 12 Switched Beams | - | [44] |
| Raytheon (US) | Flexible upgraded by SW | 8 | **Uplink**: DoA based (?) | SA can be connected to RF input at the BS | [45] |
| ArrayComm "IntelliCell" (US) | WLL, PHS, GSM | 4 | **Uplink**: ESPRIT Adaptive interference cancellation | First mass market commercial product | [46] |
| Wireless Online "ClearBeam" (US) | GSM | - | **Uplink and Downlink**: Fixed Beams (7 narrow and 2 wide) | 2 × capacity increase, 3 × coverage improvement, up to 18 dB C/I improvement | [47] |

**Table 9.2:** List of commercially available products [18].

# Chapter 10

# Consequences of Introducing Smart Antennas

The introduction of smart antennas has a large impact on the performance of cellular networks. It also affects many aspects of both the planning and deployment of mobile systems. This chapter will discuss the potential benefits and cost factors.

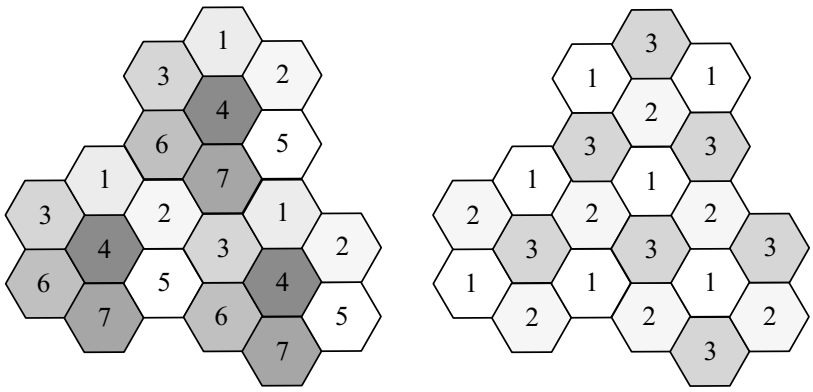
## 10.1 Improvements and Benefits

**Capacity Increase** - The principle reason for the growing interest in smart antennas is the capacity increase. In densely populated areas mobile systems are normally interference-limited, meaning that interference from other users is the main source of noise in the system. This means that the signal to interference ratio, SIR, is much larger than the signal to thermal noise ratio, SNR. Smart antennas will on average, by simultaneously increasing the useful received signal level and lowering the interference level, increase the SIR. Especially, the adaptive array will give a significant improvement. Experimental results report up to 10 dB increases in average SIR in urban areas [49].

In TDMA systems (GSM) the implication of the increased SIR is the possibility for reduced frequency re-use distance. An example is shown in Fig. 10.1, where the traditional seven-cell cluster has been reduced to a three-cell cluster. This will lead to a capacity increase of 7/3. Simulations performed on a FH-GSM network with 1/3 re-use distance utilizing SFIR report that a capacity increase of 300 percent can be expected.

CDMA systems, such as IS-95 or UMTS, are more inherently interference-limited than TDMA systems. The main source of noise in the system is the interference from other users due to the spreading codes being non-ideally orthogonal. This means that the expected capacity gain is even larger for CDMA than for TDMA. A fivefold capacity gain has been reported for CDMA in [50].

**Range Increase** - In rural and sparsely populated areas radio coverage rather than capacity will give the premises for base station deployment. Because smart antennas will be more directive than traditional sector or omnidirectional antennas, a range increase potential is available. This means that base stations can be placed further apart, potentially leading to a more cost-efficient deployment. The antenna gain compared to a single element antenna can be increased by an amount equal to the number of array elements, e.g. an eight-element array can provide a gain of eight (9 dB).



**(a)** Traditional 7-cell cluster.

**(b)** Cluster enabled by interference reduction like e.g. smart antennas.

**Figure 10.1:** Illustration of reduced frequency reuse distance.

**New Services** - When using smart antennas the network will have access to spatial information about users. This information can be used to estimate the positions of the users much more accurately than in existing networks. Positioning can be used in services such as emergency calls and location-specific billing.

**Security** - It is more difficult to tap a connection when smart antennas are used. To successfully tap a connection the intruder must be positioned in the same direction as the user as seen from the base station.

**Reduced Inter-Symbol-Interference (ISI)** - Multipath propagation in mobile radio environments leads to ISI. Using transmit and receive beams that are directed towards the mobile user of interest reduces the amount of multipaths and therefore the inter-symbol-interference.

## 10.2 Cost Factors

Although the benefits of using smart antennas are many, there are also drawbacks and cost factors. The gain should always be evaluated against the cost.

**Transceiver Complexity** - It is obvious that a smart antenna transceiver is much more complex than a traditional base station transceiver. The antenna will need separate transceiver chains for each of the array antenna elements and accurate real-time calibration of each of them. In addition, the antenna beamforming is a computationally intensive process, especially if adaptive arrays are to be used. This means that the smart antenna base station must include very powerful numeric processors and control systems. Smart antenna base station are no doubt much more expensive than conventional base stations.

**Resource Management** - Smart antennas are mainly a radio technology, but they also put new demands on network functions such as resource and mobility management. When a new connection is to be set up or the existing connection is to be handed over to a new base station, no angular information is available to the new base station and some means to "find" the mobile station is necessary. This can be handled by letting the base station continuously sweep through the cell with a "search" beam looking for candidates for a new connection or a handover. Another possibility is to

use an external system for positioning, e.g., GPS. As far as handover is concerned, a third possibility is available: directional information from the existing cell can be used by the network to provide an "educated guess" about which cell to hand the connection over to.

As was explained earlier, SDMA involves different users using the same physical communication channel in the same cell, separated only by angle. When angular collision between these users occurs, one of them must quickly switch to another channel so that the connection is not broken. This means that in systems providing full SDMA, there will be much more intra-cell handovers than in conventional TDMA or CDMA systems, and more monitoring by the network is necessary.

**Physical Size** - For the smart antenna to obtain a reasonable gain, an array antenna with several elements is necessary. Typically arrays consisting of six to ten horizontally separated elements have been suggested for outdoor mobile environments. The necessary element spacing is 0.4-0.5 wavelengths. This means that an eight element antenna would be approximately 1.2 meters wide at 900 MHz and 60 cm at 2 GHz. With a growing public demand for less visible base stations, this size although not excessive, could provide a problem. Fig. 10.2 shows a picture of an eight-element antenna array at 1.8 GHz.

**Figure 10.2:** Picture of an 8-element array antenna at 1.8 GHz. (Antenna property of Teila Research AB Sweden).

The exact cost-benefit relationship for upgrading systems with smart antennas is highly application specific. Factors such as: the diversity of services offered (e.g., asymmetrical data, low speed data, compressed voice, ATM pipe's, etc.), geographic dispersion of customers, uniformity of traffic demand, the available frequencies and regulatory constraints on their use, and topology will all impact the cost and revenue variables of the this calculation. However, certain general statements can be made based on early results of less sophisticated versions of smart antennas (e.g., switched beams used in cellular applications), which indicate that frequency utilization has been increased from anywhere between 3× to 20×, and geographic coverage increased by 20× to 200×, depending on the manufacturer making the claim. Since the adaptive array systems are more sophisticated, they could exceed these early results.

The following conservative examples, taken from [48], are provided for conceptualization purposes only since actual results will vary based on specific deployments.

Assumptions: Existing MMDS system licensed to operate on sixteen (16) channels, deployed in a full $360^0$ configuration, previously operating with 12 - fixed $30^0$ sector antennas, powered to cover the

maximum distance allowed by the FCC (35 mile radius, even though in practice more than one base station would be required due to interference without a smart antenna), with a fixed cost for the base station infrastructure (building, power, tower, heat, light, etc) of $1 MM, and incremental wireless equipment cost of $1 MM (excludes customer premises equipment which would be constant in either model), incremental cost to add the "Smart Antenna" application of $500 K, and an incremental increase in capacity of 5× (see examples in Table 10.1-Table 10.3).

| - | Without Smart Antenna (1-Base) | With Smart Antenna |
|---|---|---|
| Number of customers served | 7500 | 37500 |
| Total Cost | $2′000′000 | $2′500′000 |
| Cost/user | $270.00 | $67.00 |
| Facility related revenue/year(@ $40/year) | $300′000 | $1′500′000 |
| Simple payback period | 6.7 years | 1.7 years |

**Table 10.1:** Example #1: 100% Dedicated Internet type service (802.11 @ 144Kbps).

| - | Without Smart Antenna (1-Base) | With Smart Antenna |
|---|---|---|
| Number of customers served | 16'320 | 81'600 |
| Total Cost | $2′000′000 | $2′500′000 |
| Cost/user | $123.00 | $31.00 |
| Facility related revenue/year(@ $50/year) | $816′000 | $4′080′000 |
| Simple payback period | 2.5 years | 0.7 years |

**Table 10.2:** Example #2: 100% Shared DSL equivalent service (4.0 Mbps shared by 48 users).

| - | Without Smart Antenna (5-Base) | With Smart Antenna |
|---|---|---|
| Number of customers served | 81'600 | 81'600 |
| Total Cost | $10′000′000 | $2′500′000 |
| Cost/user | $123.00 | $31.00 |
| Facility related revenue/year(@ $50/year) | $4′080′000 | $4′080′000 |
| Simple payback period | 2.5 years | 0.7 years |

**Table 10.3:** Example #3: 100% Shared DSL equivalent service (4.0 Mbps shared by 48 users).

## 10.3 Research Issues

The first research issue is cost, including the cost of power [7]. For example, at Philips, researchers noted that 50% of the power in the handset is in the RF electronics. Therefore, multiple antennas in the handset not only increase the dollar cost of the handset, but also increase the power and thus decrease battery life. Research to reduce the power that each of these antennas requires needs to be undertaken. Similarly, the number of required receiver chains must be reduced because the RF electronics and the A/D converter required with each antenna are expensive. One method being considered is a low-cost phased array. At higher frequencies, some companies are considering using

large phased arrays to create very narrow beams to provide higher gain. But the issue is how to have, for example, hundreds of antenna elements and mass produce them at a reasonable cost. Thus, cost is limiting the number of antenna elements that can be used. Various solutions are being considered. For example, ATR is considering using optical beamforming for large phased arrays. Another solution being considered is integrating the antennas onto the RF electronics IC itself. Also, researchers at Ericsson are considering a limited introduction of smart antennas, because their research has shown that using smart antennas at just a small portion of the base stations, e.g., those having capacity problems or creating the most interference, can achieve most of the gain of complete deployment. In particular, Ericsson's results show that deploying smart antennas at only 10% of the base stations resulted in a 40% increase in capacity.

The second key research issue is size. Large base station arrays are difficult to deploy for aesthetic reasons, and multiple external antennas on terminals are generally not practical. For base stations, companies are using dual polarization, but at the terminal some companies are researching putting antennas on the RF electronics IC in an "antenna-less" terminal (since an external antenna is not present). However, issues of gain and efficiency and the effect of hand placement on the terminal need further research.

The third issue is diversity, which, as discussed above, is needed for multipath mitigation. For diversity, multiple antennas are needed on the base stations and/or terminals. As mentioned above there are three types of diversity: spatial, polarization, and angle (pattern) diversity. Spatial diversity (spatial separation of the antennas) is difficult on a small handset. Even though only a quarter wavelength separation is required for low correlation of the multipath fading between antennas on a handset, it also is difficult for base stations where the angular spread is small and large separation is required for low correlation. Spatial diversity is even more difficult to achieve in point-to-point systems where a near line-of-sight exists between the transmitter and receiver, and, further, at higher frequencies, sufficient spatial separation does not appear feasible. This problem can be partially avoided by the use of polarization diversity, where both vertical and horizontal polarizations are used to obtain dual diversity without spatial separation. For example, at Philips and other companies, researchers are using dual polarization diversity on handsets. Others are studying and implementing dual polarization on base station antennas. Polarization diversity provides only dual diversity, though polarization diversity can be used in combination with other forms of diversity to obtain higher orders of diversity.

Finally, companies are using angle diversity. That is, the signals from two or more beams (generally the beams with the highest signal powers) are used to obtain diversity. But performance depends on the angular spread. If the angular spread is small, then the received signal is mainly arriving on one beam and angle diversity will not provide a significant diversity gain. Also, some companies are studying pattern diversity, where antennas have different antenna patterns. In particular, researchers at Nokia are studying the use of multiple antennas in the handset, where some of the antennas may be covered by the hand, and moving the hand around changes the antenna pattern. These researchers believe that by adaptively combining the signals from such antennas, perhaps only using those antennas not blocked by the hand or adjusting the antenna impedance to compensate for hand placement, it may be possible to obtain much better performance (including diversity) with multiple internal antennas as compared to an external antenna.

A fourth issue is signal tracking, i.e., determining the angle-of-arrival of the desired signal with phased arrays to determine which beam to use and adjusting the weights with adaptive arrays to maximize

the desired signal-to-noise-plus-interference ratio in the output signal. At none of the sites the panel visited did the researchers feel that signal processing power was a significant issue for tracking in future systems. Instead, they felt that increases in signal processing power would permit new tracking algorithms to be implemented without substantial consideration of the processing requirements. Researchers at Ericsson, for example, noted that, although angle-of-arrival techniques for phased arrays use MUSIC or ESPRIT algorithms in 2000, improvements are needed to make these algorithms more robust with angular spread and to obtain higher resolution. For adaptive arrays, better subspace tracking methods are needed since higher data rates will require longer temporal equalizers, which require longer training sequences and greater overhead.

A fifth issue is spatial-temporal processing, i.e., equalization of intersymbol interference due to delay spread at high data rates, with cochannel interference suppression. During the WTEC panel's European site visits it was noted that better architectures are needed for spatial-temporal processing, as current architectures have room for significant improvement. However, the use of OFDM (orthogonal frequency division multiplexing) is being considered for fourth generation systems (as brought up during Japanese site visits), which may simplify spatial-temporal processing at high data rates, but further research is needed. Also, space-time coding is an area of significant research, primarily in the United States, but research on improved interference suppression and tracking with these codes is needed. Finally, multiple transmit/receive antenna systems (referred to as multiple input multiple output (MIMO) or BLAST for the Lucent Bell Labs version) are being touted mainly in the United States as a means for achieving very high capacities in wireless systems. With MIMO, different signals are transmitted from each antenna simultaneously in the same bandwidth and then are separated at the receiver, thus increasing the potential to provide an M-fold increase in capacity without an increase in transmit power or bandwidth. For example, Lucent has demonstrated 1.2 Mbps in a 30 kHz channel in an indoor environment using 8 transmit and 12 receive antennas. To be useful in a wider variety of wireless systems, however, research is needed to extend the technique to the outdoor environment, including determining the multipath richness of this environment, which is required for the technique to work properly, and to the cochannel interference environment of cellular systems.

A sixth issue involves putting the necessary hooks in the standards such that smart antenna technology can be used effectively. In second generation cellular systems, ANSI-136 and IS-95, implementing smart antennas had problems because the standards did not consider their use. In particular, ANSI-136 required a continuous downlink signal to all three users in a frequency channel, which precludes the use of different beams for each of these three users. In IS-95, there is a common downlink pilot, which also precludes the use of different beams for each user, as all users need to see the pilot. For third generation systems, smart antennas were taken into account in WCDMA, where downlink pilots are dedicated to each user, and therefore smart antennas can be effectively used on the downlink. In the EDGE system, the continuous downlink requirement is no longer present, but some signals from the base station still need to be broadcast to all users. Thus, further research is needed to ensure that smart antennas can be effectively used in this system. For fourth generation systems, therefore, smart antennas must be taken into account in standard development. Specifically, any packet or multimedia access to all users, as well as pilots, must be transmitted or done in such a way as to not preclude the use of smart antennas, if this technology is to be used to its full benefit. Since these standards are international, research in this area needs to be done globally.

The previous issue leads to the seventh and final issue: vertical integration or an interdisciplinary

approach. Research on smart antennas will require multiple factors/expertise to be considered-smart antennas cannot be studied in isolation. This issue was brought up repeatedly during the WTEC site visits. As discussed above, smart antennas must be considered in protocol development, i.e., expertise in both physical and media access layers is required. Also, smart antennas need to be considered in combination with other techniques, such as frequency hopping, power control, and adaptive channel assignment. Researchers at Nokia and Philips noted that smart antennas need to be considered in combination with RF matching, particularly with multiband antennas. At Nokia, the issue of adapting the antennas to hand position was noted. Ericsson has studied the limited introduction of smart antennas with nonuniform traffic. Another issue was the interaction when ad hoc networks are used. Furthermore, propagation measurements and channel modeling are needed to determine the performance of smart antennas in specific environments. Issues of base station versus terminal antenna (complexity) tradeoffs were also noted, as well as transmit diversity with space-time coding. From the above issues, it seems important that smart antenna research be multidisciplinary. However, few people have such a wide range of expertise, and it is often difficult for researchers with such different expertise to work together effectively. Thus, even though the critical need for such research was noted over and over again worldwide, there were few instances in any region where this was being done, or even planned in the future, as this type of research is different from the general method used in the past. Thus, this type of research appears to require a change of approach, but there was general agreement that the companies that can do this will make the greatest progress in smart antennas.

## 10.4 Conclusions

The following conclusions summarize our study [18]

- Proposed SA algorithms are becoming more complex and involve combinations with processing in time domain, multiuser detection, ST coding, and multiple antennas at MS.
- There are number of parameters such as the level of CCI reduction, diversity gain, and SNR, which can be improved with an SA. Some of these parameters can be interdependent and even conflicting. Their importance and tradeoff need to be decided on a cell-by-cell basis. The following parameters should be taken into consideration: propagation, interference environment, users mobility, and requirements for link quality.
- From the implementation point-of-view, there should always be a reasonable compromise between the amount of information about radio channels in different domains to be exploited at the SA receiver and the expected level of improvements. The possibility to exploit/obtain more detailed information related to the radio channel is restricted by the signal-processing algorithms and hardware, user mobility, and data transmission speed, and is highly dependent on the radio interface type and parameters. In complex (multipath) propagation environments, more complex SA algorithms should be used to maintain link quality requirements.
- Considerable improvements in the radio network performance with an SA can be achieved by combining different spatial-domain processing techniques like beam-forming, spatial diversity, sectorization with temporal-domain processing, and other diversity techniques. Correct and feasible combination can perhaps provide more improvements in system performance than implementation of very complex and sophisticated SA algorithms.

- A network planning concept with SA and site specific network planning tools are needed to be developed.
- Achievable capacity improvements with the SA depend on the penetration level of SA control functions into radio network control. The best performance will be obtained with an integrated approach to radio resource management and spatial processing. Jointly optimum resource management and spatial processing algorithms can be an interesting problem for future research and network design.
- The majority of the experimental SAs include a spatial diversity receiver as a reference model, which can be an economical solution. Many of the field trials show that SA receivers considerably outperform space diversity receivers.
- Today, experimental and commercially available SAs are mostly based on very simple algorithms.
- Network coverage and capacity in urban macrocells can at least be doubled with existing SA receivers. To achieve sensible capacity improvements in an urban microcell, more complex SA algorithms, discussed in this work, are required.
- A software radio will add flexibility to the SA receiver and network control, and perhaps will make them transparent to the air interface.

# Appendix A

# Acronyms

**AA** Adaptive Array

**ADC** Analog-to-Digital Converter

**AoA** Angle-of-Arrival

**ATM** Asynchronous Transfer Mode

**AWGN** Additive White Gaussian Noise

**BER** Bit-Error-Rate

**BPSK** Binary Phase-Shift Keying

**CCI** Co-Channel Interference

**CDF** Cumulative Distribution Function

**CDMA** Code-Division Multiple Access

**CIR** Carrier-to-Interference Ratio

**CM** Constant Modulus

**DAC** Digital-to-Analog Converter

**D-AMPS** Digital-Advanced Mobile Phone Service

**DCS1800** Digital Communications System 1800

**DECT** Digital Enhanced Cordless Telecommunications

**DMI** Direct Matrix Inversion

**DoA** Direction-of-Arrival

**DOCSIS** Data Over Cable Service Interface Specification

**DQPSK** Differential Quadrature Phase-Shift Keying

**DS-CDMA** Direct Sequence CDMA

**DSL** Digital Subscriber Line

**DSP** Digital Signal Processing

**ESPRIT** Estimation of Signal Parameters via Rotational Invariance Techniques

**ETSI** European Telecommunications Standards Institute

**FA** Finite Alphabet

**FCC** Federal Communications Commission (USA)

**FDD** Frequency-Division Duplex

**FDMA** Frequency-Division Multiple Access

**FIR** Finite Impulse Response

**GMSK** Gaussian Minimum Shift Keying

**GPS** Global Positioning System

**GSM** Global System for Mobile Communications

**HSR** High Sensitivity Receiver

**IC** Integrated Circuits

**IID** Independently and Identically Distributed

**IMT2000** International Mobile Telecommunications 2000

**IRC** Interference Rejection Combining

**ISI** Inter-Symbol Interference

**LAN** Local Area Network

**LMS** Least-Mean-Square Algorithm

**LOS** Line-of-Sight

**MAI** Multiple Access Interference

**MIMO** Multiple Input - Multiple Output

**MISO** Multiple Output - Single Input

**ML** Maximum Likelihood

**MLSE** Maximum Likelihood Sequence Estimation

**MMDS** Mulitipoint Microwave Distribution System

**MMSE** Minimum Mean-Square Error

**MNV** Maximum Noise Variance

**MRC** Maximum Ratio Combining

**MSE** Mean Square Error

**MU-CM** Multi-User Constant Modulus

**MU-MISO** Multiple User with Multiple antenna composite Input at the base station and Single antenna Output at each mobile

**MUSIC** MUltiple SIgnal Classification

**MU-SIMO** Multi User with Single antenna Input at each mobile and Multiple antenna Output at base station

**OFDM** Orthogonal Frequency-Division Multiplexing

**PA** Phased Array

**PBX** Private Branch Exchange

**PC** Power Control

**PCS** Personal Communication Service

**pdf** Probability density function

**PHS** Personal Handyphone System

**QHA** Quadrifilar Helix Antenna

**RDBF** Receive Digital Beam Former

**RF** Radio Frequency

**RLS** Recursive Least-Square

**SA** Smart Antenna

**SB** Switched Beam

**SDMA** Space-Division Multiple Access

**SFIR** Spatial Filtering for Interference Reduction

**SIMO** Single Input - Multiple Output

**SINR** Signal-to-Noise-and-Interference Ratio

**SISO** Single Input - Single Output

**SNR** Signal-to-Noise Ratio

**ST** Space-Time (coding)

**ST-MLSE** Space-Time MLSE

**ST-MMSE** Space-Time MMSE

**SU-MISO** Single User with Multiple antenna Input at the base station and Single antenna Output at the mobile

**SU-SIMO** Single User with Single antenna Input at mobile and Multiple antenna Output at base station

**TDBF** Transmit Digital Beam Former

**TDD** Time-Division Duplex

**TDMA** Time-Division Multiple Access

**TRB** Time Reference Beamforming (beamforming based on training signal)

**ULA** Uniform Linear Array

**UMTS** Universal Mobile Telecommunications System

**WLAN** Wireless LAN

**WLL** Wireless Local Loop

# Bibliography


[1] IoWave Inc. Smart Antenna. `http://www.iowave.com/`.

[2] G. Tsoulos, M. Beach, and J. McGeehan. Wireless Personal Communications for the 21st Century: European Technological Advances in Adaptive Antennas. *IEEE Communications Magazine*, September 1997.

[3] R. H. Ray. Application of Smart Antenna Technology in Wireless Communication Systems. ArrayComm Inc. `http://www.arraycomm.com/`.

[4] M. Cooper and M. Goldburg. Intellingent Antennas: Spatial Division Multiple Access. *Wireless, Annual Review of Communications*, 1996. ArrayComm Inc.

[5] Smart Antenna Systems. Web ProForum Tutorials, The International Engineering Consortium. `http://www.iec.org/`.

[6] J. H. Winters. Smart Antennas for Wireless Systems. *IEEE Personal Communications*, pages 23–27, February 1998.

[7] J. H. Winters. *WTEC Panel Report on Wireless Technologies and Information Networks*, chapter 6. Smart Antennas. International Technology Research Institute, Baltimore, July 2000.

[8] Ng K. Chong, O. K. Leong, P. R. P. Hoole, and E. Gunawan. *Smart Antennas and Signal Processing*, chapter 8. Smart Antennas: Mobile Station Antenna Beamforming, pages 245–267. WITPress, 2001.

[9] D. Nowicki and J. Roumeliotos. Smart Antenna Strategies. *Mobile Communications*, April 1995.

[10] A. Jacobsen. Smart Antennas for Dummies. Technical report, Telenor R&D, January 2001.

[11] `http//www.ececs.uc.edu/~radhakri/Research.htm/`.

[12] Per H. Lehne and Mangne Pettersen. An Overview of Smart Antenna Technology for Mobile Communication Systems . *IEEE Communications Surveys*, 2(4):2–13, Fourth Quarter 1999.

[13] Paul Petrus. *Novel Adaptive Array Algorithms and Their Impact on Cellular System Capacity.* PhD thesis, Faculty of the Virginia Polytechnic Institute and State University, March 1997.

[14] A. J. Paulraj and C. B. Papadias. Space-Time Processing for Wireless Communications. *IEEE Signal Processing Magazine*, pages 49–83, November 1997.

[15] A. J. Paulraj, D. Gesbert, and C. Papadias. *Encyclopedia for Electrical Engineering*, chapter Smart Antennas for Mobile Communications. John Wiley Publishing Co., 2000.

[16] J. Litva and T. K-Y. Lo. *Digital Beamforming in Wireless Communications*. Mobile Communications Series. Artech House Publishers, Boston - London, 1996.

[17] J. Baltersee. Smart antennas and space-time processing. Technical report, Institute for Integrated Signal Processing Systems, Aachen University of Technology, May 1998.

[18] B. O. Adrian and S-G. Häggman. System Aspects of Smart Antenna Technology in Cellular Wireless Communications – An overview. *IEEE Transactions on Microwave Theory and Techniques*, 48(6):919–929, June 2000.

[19] B. D. Van Veen and K. M. Buckley. Beamforming: A Versatile Approach to Spatial Filtering. *IEEE ASSP Magazine*, pages 4–24, April 1988.

[20] L. C. Godara. Application of Antenna Arrays to Mobile Communications, Part II: Beam-Forming and Direction-of-Arrival Considerations. *Proceedings of the IEEE*, 85(8):1195–1245, August 1997.

[21] H. Krim and M. Viberg. Two Decades of Array Signal Processing. *IEEE Signal Processing Magazine*, pages 67–94, July 1996.

[22] H. Liu and G. Xu. Multiuser blind channel estimation and spatial channel pre-equilization. *IEEE Proceedings ICASSP*, 3:1756–1759, 1995.

[23] A-J. van der Veen, S. Talwar, and A. Pulraj. Blind identification of fir channels carrying multiple finite alphabet signals. *IEEE Proceedings ICASSP*, pages 1213–1216, May 1995.

[24] A-J. van der Veen, S. Talwar, and A. Pulraj. Blind estimation multiple digital signals transmitted over fir channels. *IEEE Signal Processing Letters*, 2(5), May 1996.

[25] R. Price and Jr. P.E. Green. A communication technique for multipath channels. *Proc IRE*, 46:555–570, March 1958.

[26] R. Lupas and S. Verdù. Linear multiuser detectors for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, 35(1):123–136, January 1989.

[27] A. Naguib and A. Paulraj. Performance of CDMA cellular networks with base-station antenna arrays. *Proc. International Zurich Seminar on Digital Communications*, pages 87–100, March 1994.

[28] Malika Greene. Adaptive antennas on mobile handsets. Technical report, Radiocommunications Agency, June 2002.

[29] Antenova (2002). `http://www.antenova.com/`.

[30] `http://www.ee.surrey.ac.uk/Personal/A.Agius/index.html`.

[31] S. Mayrargue. Cluster on Adaptive Antennas - Report 2000. Technical report, Smart Antennas IST Cluster, 2000.

[32] C. Schneider. Multiple input - Multiple output (MIMO) Communications Systems. Technical report, Telenor R&D, May 2001.

[33] R. Becher, M. Dillinger, M. Haardt, and W. Mohr. Broad-band wireless access and future communication networks. *Proceedings of the IEEE*, 89(1):58–75, January 2001.

[34] D. Gesbert, L. Haumonté, H. Bölcskei, R. Krishnamoorthy, and A. J. Paulraj. Techonologies and Performance for Non-Line-of-Sight Broadband Wireless Access Networks. *IEEE Communications Magazine*, pages 86–95, April 2002.

[35] D. Pérez Palomar, J. R. Fonollosa, and M. A. Lagunas. Capacity results on frequency-selective Rayleigh MIMO channels. `http://www.ist-metra.org`, June 2000.

[36] R. Stridh and B. Ottersten. Spatial Characterization of Indoor Radio Channel Measurements at 5 GHz. Royal Institute of Technology, Departement of Signals, Sensors and Systems, Stockholm, March 2000.

[37] P. VanRooyen. Advances in space-time processing techniques open up mobile apps. *CommsDesign*, November 2002. `http://www.commsdesign.com/story/OEG20021107S0021`.

[38] S. Anderson, U. Forssen, J. Karlsson, T. Witzschel, P. Fisher, and A. Krug. Ericsson/Mannesmann GSM Field-Trials with Adaptive Antennas. *Proc. IEEE 47th VTC*, pages 1587–1591, 1997. USA.

[39] K. Molnar. Space-Time Processing in the Evolution of IS-136 System. *Fifth Stanford Workshop on Smart Antennas in Mobile Wireless Communications*, July 23-24 1998.

[40] J. H. Winters. Forward Link Smart Antennas and Power Control for IS-136. *Fifth Stanford Workshop on Smart Antennas in Mobile Wireless Communications*, July 23-24 1998.

[41] F. Adachi. Application of Adaptive Antenna Arrays to W-CDMA Mobile Radio. *Fifth Stanford Workshop on Smart Antennas in Mobile Wireless Communications*, July 23-24 1998.

[42] J. Monot, J. Thibault, P. Chevalier, F. Pippon, and S. Mayrague. Smart Antenna Prototype for the SDMA experimentation in UMTS and GSM/DCS1800 network. *IEEE PIMRC*, Septembar 1-4 1997. Helsinki.

[43] J. Strandell, M. Wennstrom, A. Tydberg, T. Oberg, and O. Gladh. Experimantal Evaluation of an Adaptive Antenna for TDMA Mobile Telephony System. *IEEE PIMRC*, Septembar 1-4 1997. Helsinki.

[44] `http//www.metawave.com/`.

[45] `http//www.raytheon.com/`.

[46] `http//www.arraycom.com/`.

[47] `http//www.wireless-online.com/`.

[48] `http//www.iowave.com/`.

[49] P. H. Lehne, O. Rostbakken, and M. Pettersen. Estimating Smart Antenna Performance from Directional Radio Channel Measurements. *Proc. 50th IEEE Vehic. Tech. Conf. - VTC 99 - Fall*, pages 57–61, September 1999. Amsterdam, Netherlands.

[50] G. V. Tsoulos, M. A. Beach, and S. C. Swales. DS-CDMA Capacity Enhancement with Adaptive Antennas. *Electronic Letters*, 13(16):1319–20, August 1995.