UPPSALA UNIVERSITY

Smart Antenna Implementation Issues for Wireless Communications

Mattias Wennström

October 1999

SIGNALS AND SYSTEMS
UPPSALA UNIVERSITY
UPPSALA, SWEDEN

Submitted to the Faculty of Science and Technology, Uppsala University in partial fulfillment of the requirements for the degree of Technical Licentiate in Signal Processing.

Printed in Sweden by Elanders Digitaltryck, Angered, 1999

Abstract

In this thesis, implementation issues for adaptive array antennas in wireless systems are treated. First, an overview of different implementation options are discussed and their drawbacks and benefits. We then discuss the performance degradation of the adaptive antenna array when it is implemented in hardware. The degradation is due to quantization of the received signal in the sampling process, the limited accuracy in the weighting units and the accuracy of the calibration. Due to temperature drift of active components in the hardware, the validity of the calibration will degrade over time.

We derive a theoretical model and a simulation model of an adaptive array antenna. The results are compared to measurements on an adaptive antenna testbed, developed at Signals and Systems Group, Uppsala University in cooperation with Ericsson Radio Access AB. From both the theory and the simulations, we concluded that the calibration accuracy limited the ability of the adaptive antenna testbed to suppress interferers. We also propose two algorithms for on-line calibration of the adaptive antenna array. The direct algorithm is successful in a slowly varying signal environment, which is typical in rural areas. The indirect calibration algorithm estimates the temperature drift and is shown to benefit from a fast varying signal environment, as typical in urban areas. Both algorithms are successful in maintaining the interferer suppression capability despite the temperature drift of the hardware parameters.

We also investigate, with measurements and simulations, the effect of an non-ideal multicarrier power amplifier in the transmitting downlink of an array antenna. We show how the CDF of the carrier to interference ratio of the mobiles in an cellular system depends on the basestation transmit amplifier back-off. We then define the total degradation function to find a power efficient choice of the amplifier back-off. Furthermore, the derived theory

predicts the direction of the radiated intermodulation products. These predictions were verified by measurements on an four element antenna array in an anechoic chamber. Finally, we discuss the impact of weight tapering on the radiated intermodulation power.

Acknowledgements

First, I wish to thank my supervisors Dr. Anders Rydberg and Dr. Tommy Öberg for all their help and encouragement. Our discussions are always intense and stimulating. I would also like to express my gratitude to Professor Anders Ahlén and Dr. Mikael Sternad for their advice and also for reading the thesis manuscript carefully and giving useful comments that improved the thesis.

I acknowledge all the people at Signals and Systems Group who contribute to the stimulating atmosphere, especially the often clarifying discussions at the coffee-table. I particularly would like to thank Jonas Strandell and Dr. Erik Lindskog for the co-authoring of some of my papers.

The co-operation with Dr. Leonard Rexberg, Olle Gladh and Erik Sandberg at Ericsson Radio Access AB in the work with the adaptive antenna testbed was very interesting, and gave me hands-on experience with smart antenna equipment. Bengt V. Andersson and especially Magnus Appelgren at Communicator have taught me the ABC of antenna measurements, to them I am mostly greatful.

The financial support from the National board for Industrial and Technical Development (NUTEK) is greatfully acknowledged.

Last, but no least, I would like to thank Mian for her patience, support and for her positive thinking during this thesis work. You made it possible.

Contents

Chapter 1

Introduction

1.1 Background

The number of subscribers to wireless services has increased enormously in the last few years. In some countries, the user penetration is more than 60% of the population[1], and in the predictions, there is no sign of saturation. The anticipation is to provide anyone, anywhere and anytime with mobile communication services at low cost, high quality and high data rates, with low emission of radiation and so forth. This will certainly lead to challenges for the system designer as these objectives are counteracting.

The adaptive antenna technology has been used in military applications for many years, but are now becoming economically viable in commercial systems. Certainly, the specifications for a wireless telecommunication system are completely different, but many basic problems were already solved by the military industry.

It has been verified, both theoretically and experimentally by many researchers that adaptive antenna array technology can contribute significantly to the solution of the problems mentioned above. The antenna array is useful for different purposes, depending on the signal environment it is situated in. It can be used at the base station in a cellular system, at the access point in a wireless local area network (WLAN) system, or at the users portable units, giving different benefits in different cases.

Some of the proven benefits of implementing adaptive antennas in cellular system basestations are increased spectrum efficiency and channel ca-

[1] July 1999

pacity by extending range coverage and reduction of co-channel interference. Also a reduction of multipath fading can be achieved. Perhaps the most important feature is the ability to reduce co-channel interference, by separating signals that arrive from different directions, or the ability to avoid transmission of signals in directions where another user or basestation is known to exist. It is this separation of signals that have different *spatial signatures*, or angles of arrival, that is the fundamental property that is utilized by the antenna arrays.

A major challenge in this technology is to realize the rigorously developed algorithms in hardware. By introducing an array with N antennas, the amount of hardware equipment is expected to increase N-fold, if we use the same technology as today. This motivates the need for increased integration. One example is the multi carrier power amplifier (MCPA) which replaces several single carrier power amplifiers (SCPA), thereby reducing the amount of hardware at the expense of an increased cost per unit. Another example is the advent of new technological challenges like finding robust linearisation techniques.

Another challenge is the calibration of the antenna arrays to maintain the performance at a high and reliable level. Several proposed adaptive antenna algorithms require a direction of arrival (DOA) estimate of the impinging signals and these estimates are crucially dependent on the calibration of the antenna array.

The Future

So where will the technology development lead us? The goal is to replace as much of the analog parts as possible with digital processing in a digital signal processor (DSP). Functions performed in software are normally cheaper to implement, they are more stable to environmental variations, such as temperature, and they add flexibility to the system. The system can be upgraded when new standards are available by just replacing a software program in the DSP. This (not yet existing) digital broadband radio is sometimes called a *software radio*.

The software radio concept implies that the sampling by the A/D and D/A converters (ADC,DAC) is moved closer to the antennas, which requires higher bandwidth and higher dynamic range of these digital to analog and analog to digital domain converters. Naturally, the bandwidth and dynamic range of ADC:s are counteracting, so a suitable tradeoff must be chosen. Another challenge is to limit the power dissipation which increases with performance. The ultimate goal is to place the ADC and DAC at the antenna, sampling at the radio frequency (RF) or intermediate frequency (IF), and perform filtering and down-conversion to the baseband in the DSP. The technology to perform this is not mature yet, and the ADC/DAC technology development which increases the number of ADC bits at a given sampling rate by 1.5 bits per every eight years has shown to stagnated somewhat[1]. The market need a application driver and software radios may provide the incentives for a break-trough in ADC performance advancement. Since adaptive antennas have high computational and calibration demands, the software radio technology will work in symbiosis with adaptive antennas and the beamformer will be only a part of the software in the DSP.

1.2 Array Antennas in Wireless Communication Systems

This section introduces some common terms used in the context of cellular systems and adaptive antennas. Although this thesis has its primary focus on cellular TDMA systems such as GSM and DCS-1800, some results are applicable to WLAN and CDMA systems as well.

1.2.1 Cellular systems

The first generation cellular systems was based on the analog modulation techniques. Examples are the Nordic Mobile Telephone Network (NMT) and Total Access Communication Systems (TACS) in Great Britain. In the United States the analog system is called Advanced Mobile Phone System (AMPS). These systems use frequency modulation and frequency division multiple access (FDMA) as access method.

The digital modulation format was introduced in the second generation systems. Some of the benefits of using digital modulation is increased capacity due to the introduction of channel coding and also the flexibility to introduce other services. Also lower manufacturing costs and power consumption was achieved. Some examples of systems that are used over the whole world are the Global System for Mobile Communications (GSM) and Digital-AMPS (D-AMPS). The access methods for these systems are time division multiple access (TDMA) and FDMA. Another system is the Digital Communication System (DCS-1800) which is essentially the GSM system moved to the 1.8 GHz band instead of the 0.9 GHz band, to increase capacity in dense traffic areas, so called “hot spots”.

Now consider a digital system, e.g. GSM. The system is allocated to a limited frequency band. Thus the total number of available channels is limited. In a TDMA / FDMA system, a *channel* is simply described as a certain timeslot and a certain frequency. When all channels in a system are filled, the system cannot provide service to any additional users. However, due to the propagation loss of the electromagnetic waves, the same channel can be reused in another geographical area. These geographical areas, using a unique set of channels are known as *cells*. In Figure 1.1 the concept of channel reuse is shown. All available channels are divided into three sets A,B, and C. As seen in Figure 1.1, the neighboring cells of a cell have all different channel sets than the cell itself.

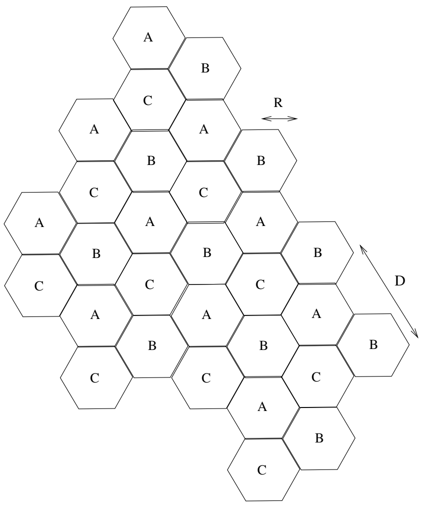


Figure 1.1: Channel assignment for cellular mobile telecommunications. Each cell is given a channel from the channel set $\{A, B, C\}$. D is the co-channel reuse distance and R is the cell radius.

There is a minimum distance between two cells using the same channel, known as the channel reuse distance D. Two cells that uses the same channel sets are known as co-channel cells and two mobiles in these cells are known as co-channel interferers, because they use the same channel and thus interfer with each other. If the reuse distance D is decreased more mobile users per square kilometer can be served if the reuse distance does not get too small so that the co-channel interferers block each others transmission.

The quality of the digital transmission can be measured by the bit error rate (BER), i.e. the percentage of the transmitted bits that are detected with errors by the receiver. The BER depends on the received carrier-to-interference ratio (CIR) which further depends on the frequency reuse distance D. Consequently, there is a tradeoff between capacity (decreased D) and quality of service (increased D).

By introducing adaptive antennas, it is possible to reduce the co-channel reuse distance D while maintaining the CIR for the mobiles at their required level for the desired quality of service (e.g. specified in BER). Thereby the system capacity is increased, without degrading the quality. Introducing adaptive antennas at all basestation sites might not be cost effective due to the additional cost of more hardware. A radio network planning tool was used in a simulation by Aurodaki and Bandelow [2], where adaptive antennas was deployed for different fractions of all the basestations in a GSM system. The simulations showed that with only 25% of the base stations deployed with adaptive antennas, the capacity increase was 50%. It was assumed that the adaptive antennas were introduced in cells with the highest traffic. Increasing the adaptive antenna penetration rate to 50% gave a 70% capacity increase compared to a conventional system. Thus the first deployed adaptive antennas gives the largest capacity increase and hence it might not be cost effective to deploy all base station sites with an adaptive antenna array.

1.2.2 Introducing adaptive antennas

This section presents the introduction of adaptive antennas at the basestation and the impacts this will have on the system capacity. The term downlink is introduced and denotes the basestation to mobile station transmission path, whereas the uplink is the opposite direction. In [3], three different strategies for deploying adaptive antennas in wireless communication are described:

HSR -High Sensitivity Reception By using an array of antennas at the basestation site, the horizontal antenna directivity is increased, and thus the range can be increased. Alternatively, the transmitted power of the users mobile can be reduced, thereby lowering the exposure of radiation to the user, or increasing the battery life in the mobile handset. The capacity increase using HSR is however small [4] and the benefits in the downlink is small with this technique. The same effect could in the downlink be obtained by increasing the total radiated power of a single basestation antenna.

SFIR -Spatial Filtering for Interference Reduction In contrast to HSR, SFIR uses beamforming in both the uplink and the downlink. The main purpose is to reduce the co-channel interference level, thereby enabling a shorter reuse distance and an increased capacity. The beamforming can be implemented in various ways:

Switched beamforming The beams are formed by a beamformer network, implemented in hardware or in software by some algorithm. Often a hardware Butler matrix [5] is used to make the antenna-space to beam-space transformation. The Butler matrix gives a number of fixed beams with narrow beam-widths. The Butler matrix output ports are connected to the receivers, and an algorithm in a DSP switches to the "best" beam, for reception. Similarly, the transmission is performed in a certain direction, by choosing the appropriate beam. Thereby the disturbance to other co-channel cells are reduced. The Butler matrix is described in Appendix A. Figure 1.2 shows the radiation pattern from a Butler matrix beamforming network, using 8 antennas in a uniform linear array. Beam number 3 (solid) and 4 (dashed-dotted) is shown. We assume that only one strong signal from each user is impinging onto the array. The desired signal is assumed to impinge from 55° and the interfering co-channel user from 110°, relative to broadside. The beamformer switches to beam number 3 (solid) which provides a reduction in desired signal power to co-channel interference power ratio of approximately 22 dB. Note that no "null" in the radiation pattern is towards the co-channel interferer, the interferer suppression has to rely on the low side-lobe level of the radiation pattern.

Full adaptive beamforming The disadvantages of the switched beamformer is that the "nulls" in the radiation pattern are fixed. Also the side-lobe level of the radiation pattern is fixed, and cannot be altered, if not a tapering of the antenna gains are introduced, which will on the other hand imply a lower antenna directivity. It is therefore desirable to introduce an individual magnitude and phase weight on each antenna channel. This will allow "null-steering", by placing nulls in the radiation pattern in the direction of known co-channel users. Compare with Figure 1.3 where an adaptive solution have been calculated for the same scenario as in Figure 1.2. The co-channel interference reduction is now more than 40 dB. However, the amount of signal processing to calculate these weights are increased, and the direction to co-channel mobiles must be estimated, to be used in the downlink algorithm. Also the requirements on the hardware increases when the adaptive solution is used. As seen in Figure 1.3, the null against the interfering signal source is very narrow, and if the calibration is inaccurate or the weights are to coarsely quantized, this null will be shifted, and the processing gain of spatial filtering is reduced. A benefit of the fully adaptive beamforming is the reduction of the number of inter-beam handovers among the beams, as the user can be tracked continuously as compared to the switched beam approach.

There might be a combination between the switched beam approach and the full adaptive approach. In the uplink, for instance, the channel between the mobile user and the basestation is known (estimated), and can be used to calculate beamforming weights and temporal equalizer coefficients jointly. However, in the downlink, the channel is unknown. Thus, it might be appealing to use the switched beamforming method, by just choosing the "best" beam according to some criterion, see Section 1.2.4.

A critical problem in SFIR is how to distinguish desired signals from interfering co-channel users. In GSM, the training sequences in the mid-amble of each transmitted data burst is used as a reference signal in many algorithms. By assigning different co-channel users to different training sequences, they can be distinguished in the receiver. The training sequence is originally used for tuning the temporal equalizer by the receivers to equalize the temporal channel. Another problem with some systems, such as GSM, is the use of frequency hopping, which makes the pattern of interferers change from one data burst to another. Thus, to implement SFIR in a GSM system requires a lot of network signalling to track the directions to future co-channel interferers. An alternative to this extensive network signalling, more intelligence can be put at the basestation. It is possible, at each data-burst to detect the interferering signals and their corresponding relevant parameters as power and direction of arrival. When these parameters are known, a bootstrap algorithm can be used to improve the signal to interference ratio in the receiver. Bootstrapping for interference rejection with array antennas was investigated by Tidestav and Lindskog[6].

SDMA -Spatial Division Multiple Access By introducing SDMA, the same channel is reused more than once in a single cell. Fully employed SDMA means that all available channels are reused in every cell. This requires a very complex channel assignment algorithm, because the user mobile must be sufficiently separated in angle to be able to use the same channel, or, in a multipath scenario, the two impulse responses of the channels between the base and the two mobiles must be uncorrelated. Otherwise the processing gain in the spatial filtering will not be sufficient to provide all users with an acceptable quality of service. So the channel assignment algorithm must keep track of the users location, and make intra-cell handovers when two users are insufficiently separated. Polarization diversity can be used to separate two users that are closely separated, but to use polarization diversity in the transmission in the downlink, the downlink channel must be completely known by the basestation. This is not possible in e.g. the GSM system, where the frequency duplex distance makes the uplink and downlink channels uncorrelated. An option is to use a dual polarization diversity antenna at the receiving mobile but one should note that no interference suppression is possible with this diversity method.

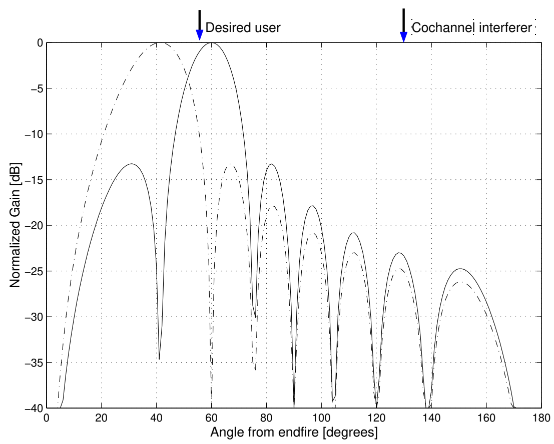


Figure 1.2: Radiation patterns from a Butler matrix beamformer. Beam 3 is solid and beam 4 is dash-dotted. These are used in the switched beam solution, note that no nulls are created in the radiation pattern in the direction of the co-channel interferer.

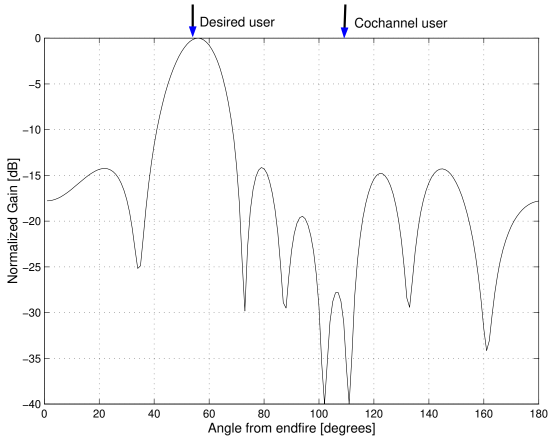


Figure 1.3: Radiation patterns from an adaptive beamformer using the SMI algorithm. The mainbeam centers in the direction of the desired user, and a sharp null is created in the radiation pattern in the direction of the interferer.

1.2.3 Uplink implementation options

The implementation of the receiving adaptive antenna array is divided into three cases, the analog, the digital, and the hybrid digital-analog beamformer. Each antenna in the array is in all three cases followed by a duplex filter to attenuate out-of-band interference. If the multi-standard software radio is implemented, then this filter and the following low noise amplifier has to be designed to cover all standards over all possible frequencies.

Analog beamformer

The analog beamformer, shown in Figure 1.4, consists of an analog beamforming network. The network can be passive, like the Butler matrix, or adaptive using analog circuits to track signals using, for example, the Howells-Appelbaum analog servo-control-loop processor [7]. Using analog circuits to implement adaptive algorithms are best suited for analog systems with a continuous transmission (no frequency hopping) as the NMT system. For a frequency hopping TDMA system as GSM this method becomes too complicated to use in practice. The passive analog beamformer is however a more feasible choice and can be seen as sectorization of the cell into sectors with a sector angle dependent on the array beam-width. By dividing the

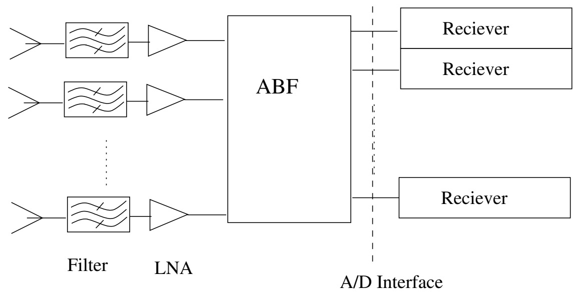


Figure 1.4: Uplink analog beamformer. After filtering and amplification, the signals are processed in an analog circuit (ABF) to create the beams. The beam-formed signals are then connected to the receivers.

cells into smaller sectors, an increased spectrum efficiency can be expected. The Butler matrix bandwidth is theoretically unlimited, so the beams cover all frequency channels. However the beam directions will change (squint) with frequency due to the change of the electrical dimensions of the array, see appendix A.

Hybrid analog-digital beamformer

In the hybrid analog-digital beamformer, the antenna signals are down-converted to baseband and digitized by A/D converters. A DSP calculates the uplink beamformer weights which are used to control an analog beamformer. The beamformer is implemented in hardware and a beamforming unit might consist of a phase shifter and an attenuator. Unfortunately, the attenuators introduce a phase shift that is a nonlinear function of the level of attenuation. Analog phase shifters suffer similar problems, and both require complex calibration schemes. One beamforming unit is required for each antenna for each user. Thus a large quantity of hardware is required. The analog beamformer can be implemented in ASIC to minimize the size and cost. The benefits of the hybrid-ABF is that it can be connected to conventional basestations as an add-on system to boost capacity in “hot spot” traffic areas. The output signals of the beamformer is at the RF frequency or might have been down-converted to the intermediate frequency (IF) if the beamforming has been performed on IF signals. The output is connected to the receivers which performs the possibly subsequent temporal equalization. Thus the spatial filtering and temporal equalization is decoupled in the hybrid analog-digital beamformer.

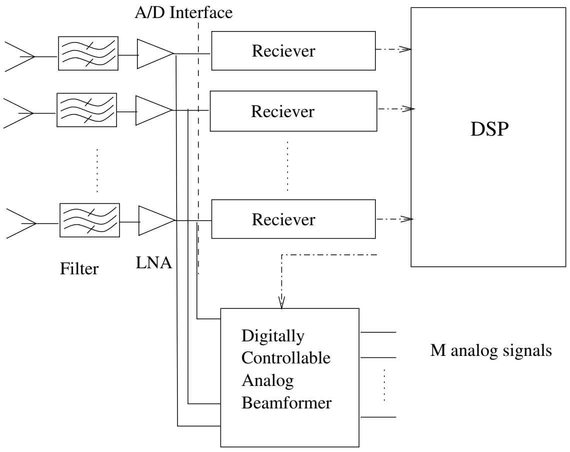


Figure 1.5: Uplink hybrid analog-digital beamformer. The received antenna signals are splitted and one part is connected to the receivers and the other is connected to the analog beamformer. The analog beamformer consists of digitally controllable phase shifters and attenuators.

Digital beamformer

In the digital beamformer, depicted in Figure 1.6, all antenna signals are down-converted to IF or baseband and then sampled by the ADC. To cover all frequency channels and to handle the near-far ratio, the ADC must be wideband and have a large dynamic range. The spatial beamforming is carried out in the DSP and can be combined with the temporal equalizer to give a spatio-temporal filter. The digital beamformer gives thus the largest flexibility but requires a high capacity DSP to carry out the computations and the key to this technology is an accurate translation of the analog signal into the digital domain. As the software radio concept develops, with wideband ADC:s and high performance DSP:s, it will be possible to implement an adaptive array antenna together with the multi-standard software radio.

1.2.4 Downlink implementation options

When implementing the downlink, several interesting design options arise. I will in this section try to systematically discuss the different choices of implementation and their benefits, drawbacks, and technical difficulties. The first choice for the system designers is to select an analog or digital beamforming system. The analog beamformer (ABF) can consist of digitally controlled

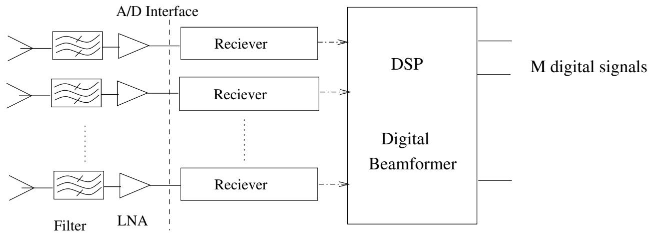


Figure 1.6: Uplink digital beamformer. The A/D converters are wide bandwidth and high dynamic range components. The signals are received and the beamforming is carried out in the DSP.

phase shifters and attenuators. This requires one phase shifter and attenuator (weighting unit) per frequency channel and antenna. Thus the required amount of hardware becomes large when the number of channels increases and becomes unrealizable large in a practical system. The phase dependent attenuation described in the previous section (up-link analog beamformer) applies here too, and thus a complex calibration scheme is required. A simpler implementation is to use a fixed beamforming network like the Butler Matrix Transformer (BMT) which generates a number of beams in specified directions, while keeping the sidelobes low [8]. See Appendix A. The algorithm switches between the beams is discussed in Section 1.2.2. I will in this section assume that the ABF is of the fixed beamformer type.

The second choice at hand is to use a digital beamforming (DBF) system where the weights are applied to the signals in the DSP. The weights can be tuned adaptively, or be of the fixed beamforming type. The fixed (or switched beam) beamformer is beneficial when the DSP processing load is high and a simple beamforming algorithm must be used. Measurements have shown that only an additional 2-3 dB can be gained in the carrier to interference ratio at a mobile station in an suburban area by using adaptive beams in the downlink as compared to fixed (switched) beams. These measurements were conducted by Andersson et.al. with an four element antenna array for the DCS-1800 system [9]. Similar results were obtained in the Tsunami project [4], with a few dB gain improvement by using adaptive compared switched beam antenna arrays. These results might be surprising, if one compares Figures 1.2 and 1.3 the interferer suppression difference is 18 dB. The difficulty lies in the non-stationary multipath scenario which makes the adaptive null-steering difficult due to the absence of good downlink channel estimates. So this motivates the switched beam or analog beamformer approach as a feasible choice to implement downlink beamforming at the basestation.

The different implementation options are shown in Figures 1.7-1.12. It is assumed that the notation "D/A interface" in the figures comprises both the D/A converter and the necessary up-conversion and filtering functions. The fixed beamforming ABF converts the inputs in the beam-space to N outputs in the antenna space by analog microwave circuits. The ABF is preceded by a switch that selects the beam to be used for transmission (i.e. selects an ABF input port). A beam selection algorithm must be based on the *average* of the received power and direction of arrival data to reduce fast fading effects and introduce hysteresis in the switching. The duplex filter combiner is omitted in the figures, as it is common for all implementation options.

The signals could be amplified in a single carrier amplifier (SCPA) and then combined using a combiner. A combiner consists of a tunable cavity filter that can handle high powers and is tuned to the carrier frequencies of the signals to be combined. A drawback with this combining method is if the frequency distribution in the network is replanned, then the filter has to be retuned manually. An alternative is a hybrid combiner which does not need this retuning but has higher losses. To avoid these difficulties, it is desirable to combine the signals in the digital domain and use a multicarrier power amplifier (MCPA) which jointly amplifies all signals together.

The MCPA needs however to be linearized to reduce the intermodulation products created when co-amplifying several signals in a single nonlinear amplifier. These different linearisation techniques are expensive or have moderate performance. The MCPA has large advantages by adding flexibility to the system and allows the multichannel software radio concept to be introduced.

ABF-SCPA

Figure 1.7 shows the Analog Beamformer with Single Carrier Power Amplifier (ABF-SCPA) setup. We assume that the ABF is of the switched beam type using fixed beams, thus there are switches to select the beam for transmission. This approach is straightforward and can easily be assembled by "off the shelves" components. Each of the M signals to be transmitted is converted to an analog signal in separate D/A converters and upconverted to the respective carrier frequency. A drawback of this technique is the high losses in the path between the SCPA and the antenna and the need for tunable cavity filters for combining. The ABF network must have capacity to handle high powers. A benefit is that the calibration is simple because

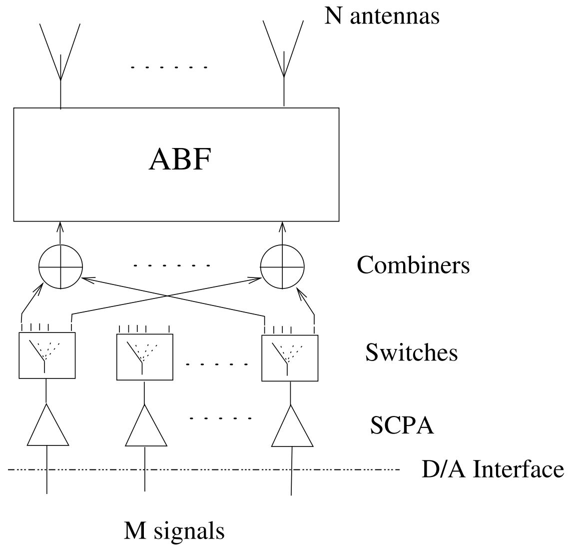


Figure 1.7: ABF-SCPA. Each of the M signals are D/A converted in a separate DAC and amplified in separate single carrier amplifiers. The combining are performed using cavity filter combiners. The beamforming is carried out using a hardware beamforming circuit.

only the path between the output of the ABF and the antenna needs to be calibrated, and it contains only passive parts which are stable under environmental changes as temperature, humidity etc.

DBF-SCPA

Figure 1.8 shows the Digital Beam Former with Single Carrier Power Amplifiers (DBF-SCPA) setup. Using an adaptive beamformer offers flexibility to the system. The downlink beamforming algorithm can be changed when better downlink channel estimates are available or when high-performance DSP:s or ASICS:s are available to do the calculations. The large drawback is that $N \times M$ SCPA:s are needed to amplify all signals to all antennas. This will be a bulky system that produces a lot of heat. There is however no need for linearisation if the signals have constant envelope modulation. The calibration is of moderate complexity, due to the SCPA amplifiers as compared to using MCPA amplifiers. Some type of combiners that can handle high power is required.

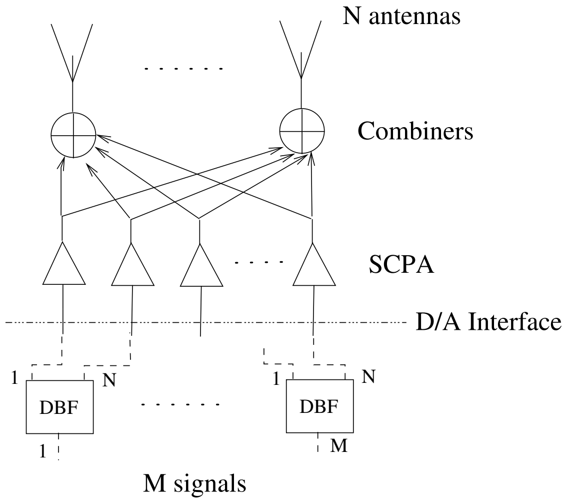


Figure 1.8: DBF-SCPA. The beamforming is performed in the DSP. M×N DAC and single carrier amplifiers are used. The amplified signals corresponding to the same antenna are combined in cavity filter combiners.

ABF-MCPA

In Figure 1.9 the ABF-MCPA option is shown (assuming switched beams). Here the MCPAs are placed prior to the ABF. Thus the power amplification is carried out in the beam-space as compared to ABF-MCPA option II described below. Only the signals to be transmitted in the same beam are amplified in the same MCPA. Thus, on average, only a fraction of all M signals are co-amplified in the same MCPA which could be designed for lower peak-to-average power ratio. Another nice property is that the calibration has to be performed on the antenna channel between the ABF output and the antennas, thus only over passive devices. The ABF will introduce losses which have to be taken into account when dimensioning the MCPA and the ABF must be capable of handling high power. Noticeable is that the intermodulation products created in the MCPAs are radiated in the same beam as the desired signals due to the location of the amplifiers in beam-space. No known investigation has been done on whether this is beneficial or if it is advantageous to scatter the intermodulation in other directions as in ABF-MCPA option II below.

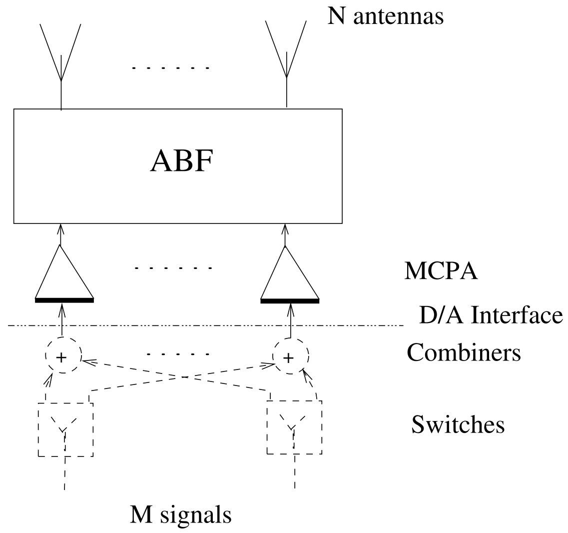


Figure 1.9: ABF-MCPA. Using wide bandwidth and high dynamic range DAC:s, the beam is chosen in the digital domain. The beam-space combined signals are amplified in the MCPAs. An analog beamformer performs the beamforming on the RF power signals.

ABF-MCPA option II

Another ABF-MCPA option is shown in Figure 1.10 where the MCPA is placed to amplify the signals in the antenna-space as compared to Figure 1.9. We decrease the losses by placing the MCPAs at the antennas, but now all M signals have to be co-amplified in each MCPA. Now, the ABF network can be designed for low power. The intermodulation products will now be radiated in another direction than the desired signals as derived in Chapter 5 in this thesis. Calibration is a more difficult task here as the channel between the ABF output and the antenna contains active components. The D/A interface could be placed prior to combining on each of the M signals (option B) or if high bandwidth D/A converters are available, prior to the ABF (option A).

DBF-MCPA

We now take the next step by moving the beamforming into the digital signal processing software, as depicted in Figure 1.11. The signals are combined after or prior to up-conversion in combiners at low power. This scheme needs some calibration algorithm to track changes in the antenna channels caused

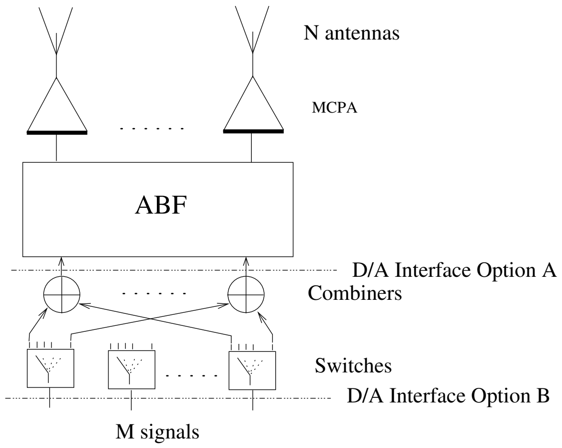


Figure 1.10: ABF-MCPA option II. The MCPAs are placed in the antenna-space, after the ABF. It is sufficient that the ABF is capable of handle pre-amplifier low power signals. The DAC interface can be placed prior or post the combining, dependent on the availability of high performing ADC:s.

by temperature drift, aging and so forth. However, these difficulties can be reduced by appropriate pre-distortion of the signals in the DSP. Note also that we need $N \times M$ D/A converters which, on the other hand, only need to cover the bandwidth of a single channel as compared to the software radio described below. No specific direction of the radiation of the intermodulation products that are generated in the MCPA can be predicted. It depends on the beamformer weights at a specific time, but as is concluded in Chapter 5, the mean power of the radiated intermodulation is less than in the ABF case. As in the ABF-MCPA option II case, the antenna branches have an active device, the MCPA, which makes the calibration more complicated. It has to be re-performed often to track the temperature drift and so forth.

Software radio (SWR)

The most versatile way to implement the downlink is by using the software radio concept shown in Figure 1.12. Here the flexibility is maximal and system and algorithms can be changed simply by changing the software in the DSP. Very high demands is put on the D/A converters which must handle a broad bandwidth and a large dynamic range. Calibration of the transmit channels are necessary and should be performed in parallel with normal operation. The major drawback with this solution is the expensive D/A con-

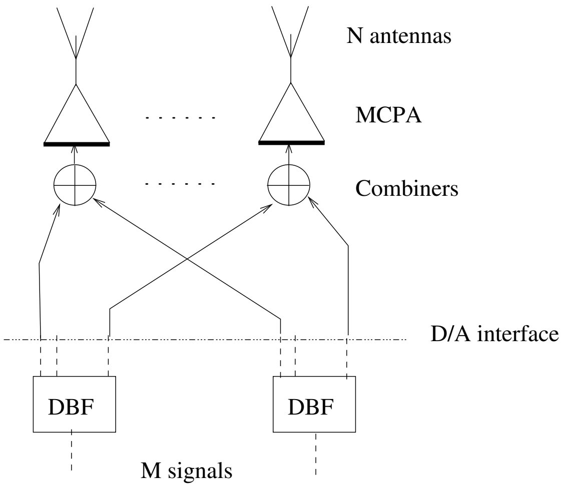


Figure 1.11: DBF-MCPA. The signals are beam-formed in the DSP. $M \times N$ ADC:s convert the signals and the combiners are in the analog RF domain. N MCPAs amplifies the signals prior to transmission.

verters, high speed DSP:s, and maybe, most costly of all, the requirements on very linear MCPAs.

The drawbacks and benefits of the different options described above are summarized in Table 1.1. The ABF-MCPA and ABF-MCPA option II are economically and technologically viable choices. If the improvement by using more intelligent beam-steering makes the DBF an option then the DBF-MCPA is to be preferred until the D/A converters makes the SWR possible.

1.3 Literature review

Many authors have given contributions to the field of adaptive antenna arrays for wireless communications. For an excellent review and introduction to the topic in a comprehensive way, with many references, refer to Godara's paper [10] or the book by Lo and Litva [11]. Also the textbooks by Compton [12], Monzingo/Miller [13] and Hudson [14] provide excellent introductions to the general theory of adaptive array antennas. This thesis is focusing on implementation issues, and the literature review given here, will be covering that area only.

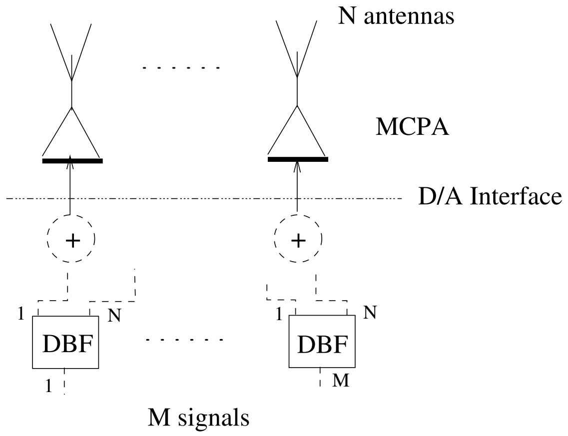


Figure 1.12: Software radio concept (SWR). High performance ADC:s convert the software beam-formed signals. N MCPAs amplifies the antenna signals to the power required for the radio interface. This solution is extremely flexible and easily reconfigurable, although expensive.

1.3.1 Implementation errors and calibration

The analog beamformer using digitally controlled analog weights was studied in the 1970's by several researchers, mainly for contributions in the field of military applications. This section covers some of the papers presented in the area of signal quantization by the ADC and weight quantization by the hardware weighting units.

Nitzberg made several contributions in the field of quantization effects, such as errors in adaptive weights [15] and the necessary precision in the DSP, when computing the inverse of the covariance matrix [16]. The effect of quantization of the signal by the ADC and also the quantization in weights when using hybrid analog-digital beamformers was examined by Hudson [17] and Takahashi *et.al.*[18]. Hudson showed how the quantization limited the adaptive antenna capability to reject interference. Similar results was presented by Takahashi's group, but applied to the field of mobile satellite communications.

Some more recent work of the effects of signal and weight quantization was presented in 1985 by Godara [19],[20], who studied the effect of phase shifter errors in the hybrid analog-digital beamformer, and how these errors affected the output signal-to-noise-ratio (SNR) of the optimal beamformer which maximizes the SNR. It was shown that the output desired signal will be suppressed in the presence of phase shifter errors, and that the suppres-

Type	Advantages	Disadvantages
ABF-SCPA	Off the shelf components and low calibration requirements	High losses Large amount of hardware
DBF-SCPA	DBF $\rightarrow$ flexible	Many SCPA:s required Losses in combiners
ABF-MCPA	Each MCPA co-amplifies only $k \leq M$ signals Simple calibration Direction of IM predictable	Losses in Analog Beam-Forming network (high power)
ABF-MCPA option II	low losses in combiner Direction of IM predictable	Losses in ABF Calibration difficulties
DBF-MCPA	Low loss combining low bandwidth D/A	Calibration difficulties
SWR	Extremely flexible and reconfigurable	Expensive and calibration difficulties

Table 1.1: Comparison of different downlink implementation options.

sion is proportional to the product of the phase shifter error variance and the total input power.

In his second paper [20], Godara extended the analysis to cover steering vector errors as well as weight errors and concluded that the steering vector errors affected the output SNR differently depending on how the covariance matrix of the received signals is estimated. Steering vector errors, or look-direction errors, might arise from a uncalibrated antenna array, or uncertainties in the estimation of the direction to the desired signal, due to angular spread. If the desired signal is included in the estimation of the covariance matrix, then the output SNR is extremely sensitive to errors in the steering vector. This is also called the hyper-sensitivity effect.

The hyper-sensitivity effect is an example of when the choice of the algorithm and the hardware errors combines to degrade the performance of the array antenna. The origin of the hyper-sensitivity effect has been studied by Bull *et.al.* [21], who showed how the steering vector error makes the desired signal to appear as an interferer, and thus being suppressed. Several authors have provided the remedy for this problem, by projecting the steering vector onto the signal subspace or by adding artificial noise to the diagonal of the covariance matrix [22],[23],[24],[25].

The effect of errors and inaccuracies in the calibration of the antenna array have been studied by Tsoulos *et.al.* [26],[27]. They implemented a calibration algorithm into the adaptive antenna used in the TSUNAMI project. They also studied the requirements of calibration accuracies to achieve a desirable performance. In [28], an auto-calibrating digital beamformer for the downlink was presented.

1.3.2 Effects from nonlinearities

Nonlinearities in the active stages of an array antenna can result in increased side-lobe levels, decrease the null depths in the radiation pattern and even change the positions of the nulls. This increases the interference level for the mobile terminals and the basestations in the system. The effects of nonlinearities for adaptive antennas has been studied by Litva [11], who introduced the concept of *phantom-interferers* when receiving with an array antenna. A phantom interferer appear when the intermodulation products from the nonlinearities in the different receiving antenna channels combines to make a virtual signal (not existing in the air interface) appear from some direction. When the receiving amplifiers are highly saturated, more phantom interferers occur. Litva also showed that with adaptive receiver beamforming it is possible to suppress the phantom interferers.

Nonlinearities in adaptive antennas for satellite communication was investigated by Johannsen [29]. The paper shows how the intermodulation distortion created in an adaptive array antenna system in some cases is directed from the earth and thus the intermodulation distortion power for receivers located on the ground is decreased. Hence, the array antenna improves the system capacity, or it enables a smaller back-off for the satellite amplifiers, which leads to an increase in power efficiency. The output back-off of an amplifier is defined as the ratio of the maximum output power when the amplifier is saturated to the actual mean output power. It is often expressed in decibel.

Nonlinearities in base-station adaptive array antennas in cellular systems was discussed by Tsoulous *et.al.* [26] and by Xue *et.al.* [30]. Although no measurements or simulation were carried out, the authors discussed how the intermodulation distortion affected the side lobe levels, the null depths, and the change in null direction for the transmitted and received radiation pattern.

Simulation studies on active array antennas and how the nonlinearities affect the achievable performance has been investigated by Loyka *et.al.* [31]. There a method of modeling and simulation of system-level nonlinear effects is presented. Loyka shows how the simulation must be divided into a multi-level analysis where electromagnetic tools are needed for the front end antenna parts, and time domain circuit simulators for the nonlinear parts. The results are then combined in a system (behavioral)-model for analysis of the entire system. In Coutiere's paper [32] regarding satellite transmitters, simulations show the direction of intermodulation distortion in the case of an array antenna transmitting towards northern Europe. It is in one case shown that the intermodulation falls onto southern Spain with a power which is 13 dB lower then in the main beam. It is thus concluded that a careful investigation of the directions and powers of the intermodulation products is important to avoid exceeding the allowed maximum in any direction.

1.4 Outline of the thesis

In Chapter 2, a presentation of the adaptive antenna testbed developed at Uppsala University in co-operation with Ericsson Radio Access AB is given. Outdoor field trials were commenced and the receiving array antenna is shown to successfully (bit error free) separate two signal sources transmitting in the same GSM time-slot and on the same frequency although the carrier to interference ratio (CIR) is -20 dB for one of the users. The improvement in CIR as a function of direction-of-arrival (DOA) separation was investigated by measurements. Regularization was introduced by means of diagonal loading in the received signal covariance matrix. In this way the algorithm was stabilized. Also a demonstration was commenced with voice quality test in a multipath signal environment. The adaptive antenna was able to separate the signals from two co-channel user mobiles placed 0.2 meters apart, so two simultaneous conversations were maintained during several minutes.

The measurements on the testbed, both in the outdoor field trials, and in the measurements in the laboratory raised some questions. What is limiting the performance of the testbed antenna array? Which parts of the hardware would degrade the performance most? Where should effort be put if we decided to improve the interferer suppression capability? A model of a receiving adaptive array antenna was developed. It is presented in Chapter 3. There an expression for the output CIR in the presence of hardware imperfections is derived. A/D quantization, calibration errors, weight errors, and limited dynamic range were considered. This work is useful for a system designer for understanding the impact of these errors and how to balance the accuracy in the implementation of e.g. the phase shifters with respect to the accuracy in the attenuators (weight magnitude). Thus, it helps the designer to avoid over- or under-dimensioning of the hardware parts.

The theoretical and simulated results coincided with the testbed measurements and we found that the performance bottleneck in the particular testbed was the coarse magnitude steps in the weighting units (1 dB) which also had effect on the calibration. If the calibration could be improved such that the channel errors after compensation become negligible, then the CIR on the adaptive antenna testbed output would be improved up to 8 dB.

This shows that the calibration of adaptive antenna arrays is an important issue. Due to seasonal changes in temperature, aging and so forth, the calibration is of out-most importance. A problem is that the calibration requires control over the weighting units so it has to be performed off-line, i.e. prior to "normal" antenna operation. Thus to calibrate the system, it has to be stopped and taken out of operation, which is undesirable. To avoid such undesirable action we propose in Chapter 4 two different on-line calibration algorithms that run in parallel to normal operation and continuously calibrate the antenna. The first algorithm uses a modified least mean square (LMS) algorithm to adaptively change the weights to minimize a mean square error criterion. The second algorithm estimates the changes in the channel transfer functions by solving a regression problem. The two algorithms are successful in maintaining the CIR, as measured on the output of the beamformer on a constant level when the calibration data becomes old.

In Chapter 5, the effect of using MCPAs in the downlink in the array antenna system is investigated. We simulated how the linearity and back-off are connected to the CIR sensed by a test-mobile in a seven cell cellular system. The spatial distribution of intermodulation products is derived and verified by measurements in an anechoic chamber on an four element linear antenna array. We also show how the nonlinear distortion can be approximated by an equivalent weight distortion.

In Chapter 6 conclusions are drawn and suggestions for future research work is discussed.

1.5 Contributions

Parts of the material in this thesis have been published or submitted to the following journals and conferences:

Chapter 2: Jonas Strandell, Mattias Wennström, Anders Rydberg, Tommy Öberg, Olle Gladh, Leonard Rexberg and Eric Sandberg, "Design and

Evaluation of a Fully Adaptive Antenna for Telecommunication Systems", In *Nordiskt Antennsymposium 1997*, Göteborg, Sweden, May 27-29 1997, pp.357-365

Jonas Strandell, Mattias Wennström, Anders Rydberg, Tommy Öberg, Olle Gladh, Leonard Rexberg, Eric Sandberg, Bengt V. Andersson and Magnus Appelgren, "Experimental evaluation of an Adaptive Antenna for a TDMA Mobile Telephony System", Proceedings of the *Eight International Symposium on Personal, Indoor and Mobile Radio Communications* (PIMRC), Helsinki, Finland, September 1-4 1997, pp.79-84

Chapter 3: Mattias Wennström, Tommy Öberg and Anders Rydberg, "Analysis of Quantization Effects in Adaptive Array Antennas", In *Radiovetenskap och Kommunikation 99*, Karlskrona, Sweden, June 15-17 1999, pp.451-455

Chapter 4: Mattias Wennström, Jonas Strandell, Tommy Öberg, Erik Lindskog and Anders Rydberg, "An Auto-Calibrating Adaptive Array for Mobile Telecommunications", Submitted to *IEEE Trans. on Aerospace and Electronic Systems.* Revised version submitted spring 1999.

Chapter 5: Mattias Wennström, Anders Rydberg and Tommy Öberg, "Implications of Nonlinear Amplifiers in the Downlink on Mobile Telephone Systems Utilizing Adaptive Antenna Arrays", Submitted to *IEE Proceedings Microwaves, Antennas and Propagation*

Mattias Wennström, Anders Rydberg and Tommy Öberg, "Effects on Nonlinear Transmit Amplifiers in Smart Antennas for Wireless Systems", Presented at *European Wireless'99*, October 6-8 1999, Munich, Germany

1.6 Abbreviations

The following abbreviations are used in this thesis.

ABF Analog Beam-Former

ADC Analog to Digital Converter

AGC Automatic Gain Control

AM/AM Amplitude Modulation to Amplitude Modulation

AM/PM Amplitude Modulation to Phase Modulation

AMPS Advanced Mobile Phone System

ASIC Application Specific Integrated Circuit

BER Bit Error Rate

BMT Butler Matrix Transformer

BPSK Binary Phase Shift Keying

BT Time Bandwidth product

CDF Cumulative Distribution Function

CDMA Code Division Multiple Access

CINR Carrier to Interferer plus Noise Ratio

CIR Carrier to Interference Ratio

CNR Carrier to Noise Ratio

CW Continuous Wave

DAC Digital to Analog Converter

D-AMPS Digital-AMPS

DBF Digital Beam-Former

DCS-1800 Digital Communication System at 1800 MHz

DOA Direction Of Arrival

DSP Digital Signal Processor

DTX Discontinuous Transmission

FFT Fast Fourier Transform

FDD Frequency Division Duplex

FDMA Frequency Division Multiple Access

FEC Forward Error Correction

FNBW First Null Beam-Width

FSRX Feedback Sampling Receiver

GAA Gaussian Angle of arrival

GSM Global System for Mobile communications

GMSK Gaussian Minimum Shift Keying

HSR High Sensitivity Reception

IBO Input Back-Off

IF Intermediate Frequency

IM_3 Third order Intermodulation product

IMD Intermodulation Distortion

INR Interference to Noise Ratio

ISRX Input Sampling Receiver

LMS Least Mean Square

MCPA Multi Carrier Power Amplifier

NMT Nordic Mobile Telephony network

OBO Output Back-Off

PAE Power Added Efficiency

PDF Probability Density Function

PRBS Pseudo Random Binary Sequence

RF Radio Frequency

SAF Shimbo Amplitude Function

SCPA Single Carrier Power Amplifier

SDMA Spatial Division Multiple Access

SFIR Spatial Filtering for Interference Reduction

SICR Summed inverse Interference to Carrier Ratio minimizing beamformer

SIR Signal to Interference Ratio

SLL Side Lobe Levels

SMI Sample Matrix Inversion

SRX Sampling Receivers

SWR Soft-Ware Radio

TACS Total Access Communication System

TDMA Time Division Multiple Access

TSUNAMI Technology in Smart antennas for Universal Advanced Mobile Infrastructure

ULA Uniform Linear Array

WLAN Wireless Local Area Network

Chapter 2

Testbed Description and Evaluation

2.1 Introduction

This chapter describes briefly the adaptive antenna testbed designed and realized at Signals and Systems Group, Uppsala University. For a more detailed description of the internal hardware parts and a discussion of the choice of algorithm, refer to Andersson and Landing [33]. The adaptive antenna testbed project commenced 1994 at Uppsala University in a cooperation with Ericsson Radio Access AB. The aim was to develop and evaluate adaptive antenna technologies for mobile communication systems and to obtain experience of using this technology in base stations. Due to limited resources, only the uplink was implemented, and analog beam-formers consisting of digitally controlled phase shifters and attenuators were used. This corresponds to the hybrid analog-digital beamformer described in Section 1.2.3. Some other testbed projects have been reported by Ericsson [9], and by the TSUNAMI project [34].

This chapter starts with a description of the adaptive antenna architecture. Section 2.4 presents results from measurements performed with the testbed in a laboratory, where a Butler matrix was used to emulate the front end. Section 2.5 describes measurements performed at an outdoor antenna measurement range and Section 2.6 describes a demonstration with voice transmission using the testbed. Finally, this chapter ends with some conclusions in Section 2.7.

2.2 Adaptive antenna architecture

The adaptive antenna is a test system which is intended to work in the uplink only. The antenna is designed for integration with an existing base station, using the DCS-1800 standards, thus the radio interface frequency is 1.8 GHz and only the uplink (receiving mode) is implemented. A photo of the antenna is shown in Figure 2.1. See Figure 2.2 for a schematic outline of the system. The front-end consists of ten antenna elements, mounted in a circular array configuration. The antennas are connected to directional couplers to allow injection of the calibration signal, low noise amplifiers, and cavity filters. Each of the ten signals from the front end is split, and one replica is connected to a sampling receiver, and the other two to the two beam-formers.

Two sets of hardware beamforming weights are implemented, to be able to use two parallel and beam-formed receiving channels simultaneously. Each set consists of ten weighting units. Thus, the two beam-formers enable trials with the SDMA access method, by assigning two signal sources (mobiles) to transmit different training sequences. A summary of some characteristics of the adaptive array system is shown in Table 2.1. The beam-forming is performed on the received RF signals to enable the use of an ordinary base station as a receiver. Thus, the beam-formed, or spatially filtered signal is connected to the base-station. The weights are calculated in the DSP, while they are applied to the signals via hardware phase shifters and attenuators. This is denoted a hybrid-analog beamformer due to the use of both digital and analog signal processing and is also described in Section 1.2.3, cf. Figure 1.5.

The base-station also generates the data for transmission, so it is possible to compare with the received data, to calculate the bit error rate (BER) and other parameters that characterize the transmission. Only traffic channel frames are transmitted. The basestation also provides synchronization signals, used by the adaptive antenna for correct timing of the sampling.

The sampling receivers (SRX) down-convert the RF signal to the baseband and separate it into I and Q channels using double-down-conversion receivers. The sampling in the ADC:s is performed at 270 kbit/s which is the symbol rate for the GSM signal. The receiver gain is set so the quantization noise is equal to the thermal noise for maximal use of the dynamic range. The ADC:s use 8 bits and the DSPLINK2 bus that connects to the DSP uses 32 bits. Thus four channels can be read from the ADC:s to the DSP simultaneously. The increase in performance by using ADC:s with more bits is investigated in Section 3.4.3.

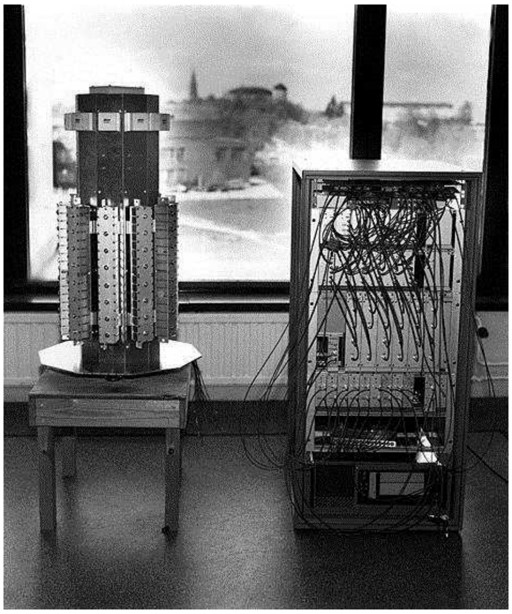


Figure 2.1: The adaptive antenna, to the left the front end and to the right the rack with receivers, DSP-system, and weighting units.

The digital signal processor (DSP) consists of seven TMS320C40 signal processors. The algorithm used is the sample matrix inversion (SMI) algorithm [12], described in Section 4.2.1. It was chosen for its rapid convergence compared to other algorithms [13], and to avoid the need for calibration of the channel between the antennas and the SRX. The training sequence of 26 bits in the mid-amble of each DCS-1800 traffic channel burst is used as a reference signal.

The signals that are used in the DSP to calculate the weights are not exactly the same as the signals used in the beam-formers. Rather they are a phase shifted and attenuated replica of them. The phase shift and attenuation is dependent on the length of the cables, their losses, and also the characteristics of the weighting units. The weighting units are built from active components, as phase shifters, attenuators and amplifiers. These active components suffer from temperature drift which will make the calibration invalid. It is therefore necessary to recalibrate often.

For calibration purposes, a feedback receiver is placed to down-convert and sample the summed beamformer output signal. The calibration has to take place off-line, i.e. prior to operating the antenna system by injecting a CW signal through directional couplers after the antenna elements. The calibration algorithm needs full access to the weight controller. Thus, the calibration must be performed off-line. The data from the calibration is stored as a look-up table in the DSPs memory. To avoid off-line calibration, which means that the base-station must be taken out of service, two different on-line calibration methods are presented in Chapter 4.

Item	**Specification**
Front end and radio parameters	
Center Frequency	1721 MHz
Modulation	GMSK (BT=0.3)
Number of antenna elements	10
Antenna configuration	Circular
Antenna element spacing	0.56λ @ 1721 MHz
Antenna polarization	Vertical
Antenna element beam-width	80 degrees
Receiver and DSP system	
DSP Processors	Seven TMS320C40
Sampling frequency per I- and Q-channel	270 kHz
ADC resolution	8 bit
ADC dynamic range	-32 dBm to -80 dBm
Algorithm for weight calculation	SMI
Weighting units	
Class	Analog
Phase shift resolution	1 degree
Amplitude attenuation resolution	1 dB
Calibration	
Phase accuracy	1 degree
Amplitude accuracy	0.75 dB

Table 2.1: Characteristic parameters of the adaptive antenna testbed

2.3 Optimal array antenna

In this section we derive the expression for the CINR of an optimal adaptive antenna to be able to compare with the measurements results. The minimum mean square error solution, using the training sequence at a reference signal leads to the Wiener-Hopf equations. The vectors $\mathbf{a}_i(\theta_i), \mathbf{a}_d(\theta_d)$ are the array response vectors towards the interferer and the desired mobile respectively. We assume a line of sight model with no angular spread

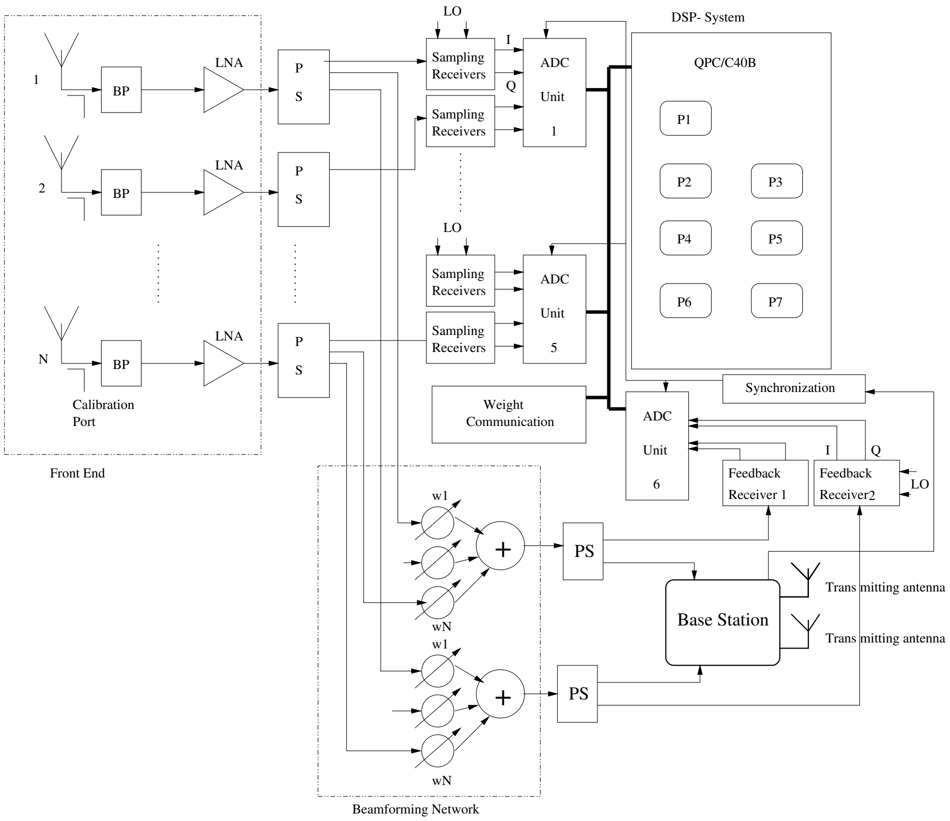


Figure 2.2: Adaptive antenna architecture

for simplicity. Furthermore, we assume that the array response vectors are normalized, i.e. $\mathbf{a}_d^H(\theta_d)\mathbf{a}_d(\theta_d) = 1$. If we assume that the training sequence has infinite length, to avoid finite sample size effects, then the covariance matrix and the cross-correlation estimates (as shown in equation 4.4) can be replaced by their true values, and the Wiener-Hopf solution for the weight vector $\mathbf{w}$ is

$$\mathbf{w} = \mathbf{R}_{xx}^{-1}\mathbf{r}_{xs} \quad . \tag{2.1}$$

Here $\mathbf{R}_{xx}$ is the signal covariance matrix and $\mathbf{r}_{xs}$ is the cross-correlation between the received signal and the training sequences. The covariance matrix of the received signal vector is assumed given by

$$\mathbf{R}_{xx} = \sigma_d^2\mathbf{a}_d(\theta_d)\mathbf{a}_d^H(\theta_d) + \sigma_i^2\mathbf{a}_i(\theta_i)\mathbf{a}_i^H(\theta_i) + \sigma_n^2\mathbf{I} \tag{2.2}$$

where σ_i^2, σ_d^2 and σ_n^2 is the interferer, the desired signal and the noise power respectively, as measured by a virtual omnidirectional antenna at the base-station.

By applying the matrix inversion lemma [14] twice in equation (2.2) and separating the interferer, the desired signal and the noise parts of the output signal, it is possible to show that the beamformer output carrier to interferer plus noise ratio can be written [14],[35] as

$$CINR = \frac{N\sigma_d^2}{\sigma_n^2}\left(1 - \frac{N^2\sigma_i^2|\rho|^2}{(\sigma_n^2 + N\sigma_i^2)N}\right) \tag{2.3}$$

where

$$\rho = \mathbf{a}_i^H(\theta_i)\mathbf{a}_d(\theta_d)/N \tag{2.4}$$

is the spatial correlation between the array response vectors. When the interferer to noise power (INR) is high and the spatial correlation is low, $\frac{\sigma_i^2}{\sigma_n^2} \gg \frac{1}{N(1-|\rho|^2)}$, we have

$$CINR \simeq \frac{N\sigma_d^2}{\sigma_n^2}\left(1 - |\rho|^2\right) \quad . \tag{2.5}$$

Equation (2.5) shows how the CINR at the output of the optimal adaptive array is the input signal to noise ratio (σ_d^2/σ_n^2), amplified N times by the array, as in the absence of an interferer, and then reduced by the factor $(1 - |\rho|^2)$. The loss is due to the pattern distortion when forming a null in the interference direction. Note that the interferer power is not present in the expression (2.5). The interfering signal is suppressed to the noise floor independent of the power.

If the desired signal and interferer are very close in azimuth angle, then the spatial correlation is $|\rho| \simeq 1$. In this case the noise appears to be "amplified" by the spatial correlation and $\frac{\sigma_i^2}{\sigma_n^2} \ll \frac{1}{N(1-|\rho|^2)}$. This will lead to the approximation

$$CINR \simeq \frac{N\sigma_d^2}{\sigma_n^2 + N\sigma_i^2} \tag{2.6}$$

which is the signal to interferer plus noise ratio of an N element omnidirectional receiver. Thus, no improvement by using the smart antenna is obtained. A more thorough derivation of the output CINR for the minimum variance beamformer is presented by Wax and Anu [36] which takes the finite sample size effect when estimating the covariances and also the correlation between the desired and the interfering signal into account.

2.4 Measurements in laboratory

To perform a measurement that verifies the adaptive antenna performance in a laboratory, without introducing complicated effects such as time delays and multipath propagation, the front end was replaced with an 8×8 Butler matrix. See appendix A for a description of the Butler matrix. Two signal generators were connected to two different input ports of the Butler matrix to emulate a scenario corresponding to a 8-element linear array of isotropic elements spaced $\lambda/2$ with two impinging signals with DOAs (direction of arrival) -14.5° and 14.5° respectively, relative to broadside. Here, λ is the wavelength of the RF signal. The 8 outputs of the Butler matrix are then connected to the adaptive antenna. Ideally, the array response vectors from the signals generated by a Butler matrix are orthogonal, i.e. $\mathbf{a}(\theta_k)^H\mathbf{a}(\theta_l) = 0$, where k and l represents different ports of the Butler matrix, generating signals from direction θ_k and θ_l. However, imperfections in manufacturing of the Butler matrix, peak errors of 8.5° in phase and 0.8 dB in magnitude makes the array response vectors non-orthogonal, which corresponds to a more general signal environment. The performance of the adaptive antenna was examined by measuring the output BER and output power of the desired signal, called the carrier (C_{out}) and output interferer signal (I_{out}) for different settings of input power (C_{in}, I_{in}). The input power of the two signals were measured at the input of the Butler matrix. With these parameters, it is possible to calculate interferer suppression, and the gain of the desired signal.

The desired signal is GMSK-modulated with actual DSC-1800 traffic channel data containing the 26 bit training sequence. The interferer is a GMSK-modulated signal transmitting pseudo random (PRBS) data or traffic data with another training sequence, depending on if both beamformers are used simultaneously. Ten measurements were performed and the diagrams presented show the mean of these series, see Figure 2.3, Figure 2.4 and Figure 2.5.

Figure 2.3 shows the suppression of the interferer through the adaptive antenna system. The standard deviation is 7 dB. This large standard deviation is due to the randomness in estimation of the steering vector which is calculated as the correlation between the training sequence and the received data. Due to the short training sequence, the correlation between

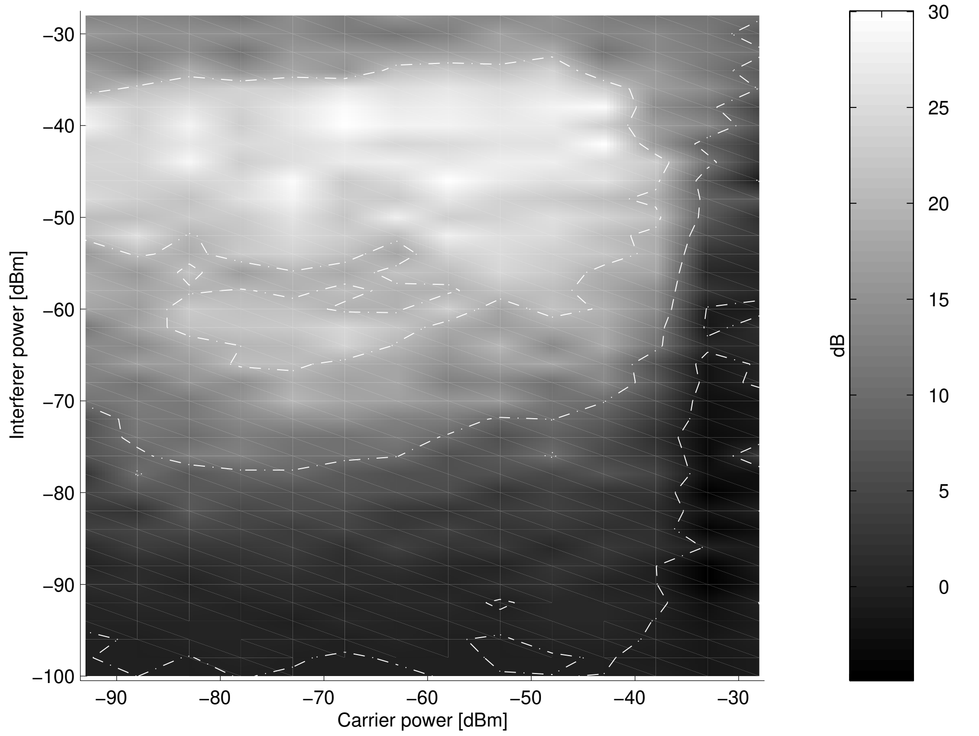


Figure 2.3: Suppression of interfering signal, in dB. Measurements performed in lab, on an 8 element array antenna, mean values over 10 measurements.

the training sequence and the PRBS data transmitted by the interferer will be nonzero and will vary randomly. If the correlation is small, then the interferer suppression becomes excellent and if it is large, then the suppression will be inferior. This is the reason of the large standard deviation in the measurements. This effect was also observed in the adaptive antenna testbed in the TSUNAMI project [4].

The adaptive antenna is able to suppress the interfering signal when the power of the interferer and the desired signal is in the ADC dynamic range. The ADC uses 8 bits for conversion and this gives a dynamic range of 48 dB. The receiver gain was adjusted so the quantization noise power and the thermal noise power was approximately equal. The limited dynamic range implies that for a received signal power lower than -80 dBm the signal is hidden in noise and if it is an interferer, no suppression is possible. If the interferer is present, then it is suppressed to the quantization noise floor,

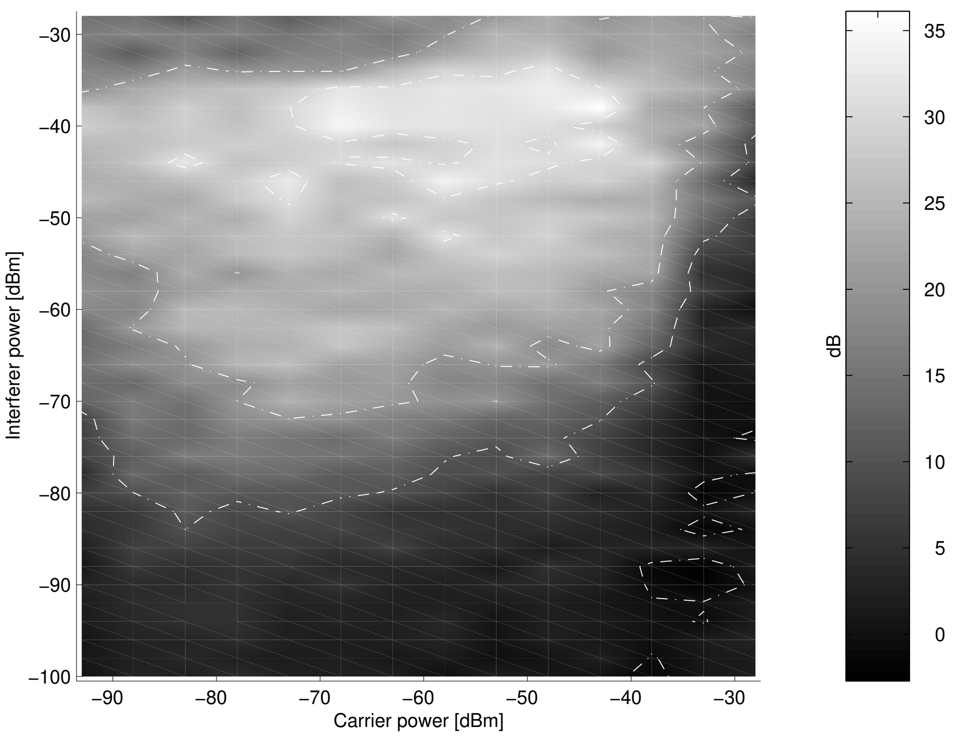


Figure 2.4: Improvement in CIR, in dB relative to one element antenna. Measurements performed in lab, on an 8 element array antenna, mean values over 10 measurements.

assuming that the signal sources are sufficiently separated in angle. Stated differently: If the spatial correlation is sufficiently low, which can be assumed here. Compare with equation (2.5) which expresses the output CINR for the optimal antenna. It is independent of the interferer power. Thus, the interferer suppression is proportional to I_{in}. The dashed line surrounding the brightest area in Figure 2.3 is the 20 dB iso-suppression line, and the maximum suppression is about 30 dB.

If the ratio of the output carrier to interference ratio (CIR_{out}) to the input CIR (CIR_{in}) is plotted, we obtain the CIR improvement, presented in Figure 2.4 and 2.6. The CIR_{in} was measured using one of the antennas in the array so the improvement will be referred to a single antenna. The main contribution to the CIR improvement is from the interferer suppression. The gain of the desired signal was approximately constant at about 2-6 dB for all carrier and interferer powers in the ADC dynamic range. With an interferer power at -40 dBm and the desired input signal power between -70 dBm and -40 dBm ($\mathrm{CIR}_{in} \leq 0$), the adaptive antenna is improving the CIR more than 30 dB. When either of the signal levels (I_{in},C_{in}) exceeds the ADC dynamic range the CIR improvement drops abruptly to 0 dB or less. The reason is that with saturated receivers, the calculation of correct weights is not successful, and the beamforming is not performed correctly.

Figure 2.6 shows a vertical cut of the plot in Figure 2.4 at C=-73 dBm. The carrier to interference improvement increases linearly (on dB scale) between interferer power between -90 and -70 dBm in accordance with the optimal antenna theory in Section 2.3. Above -70 dBm the improvement starts to saturate and levels out between 25 and 30 dB. The reason for this is investigated in Chapter 3.

Another illustrative way of presenting the measured data is to plot the improvement of the BER for different settings of C_{in} and I_{in}, see Figure 2.5. The improvement is measured relative to one of the Butler matrix outputs, i.e. using a single antenna element compared to the adaptive antenna. Clearly the coded BER improvement is zero for $\mathrm{CIR}_{in} \geq 9$ dB because the BER is 0% for an ordinary one-antenna base station receiver in this triangular area. This is the threshold level in DCS-1800 for 0 % BER. Above this triangle is the region where we benefit from using the adaptive antenna compared to using an ordinary base station antenna. Here the BER improvement is 50 percentage units, giving a BER out from the adaptive antenna of 0%. A CIR_{in} ratio above -20 dB is thus necessary to give a BER of 0% at this setup, using the Butler matrix. Clearly, the separation affects the minimum CIR_{in} to still maintain 0% BER at the beamformer output. This is further investigated in Section 2.5.3. The presented data is post-FEC (forward error correction) decoding of class Ib data in the base station.

A deeper theoretical insight into what limits the suppression of the interfering signals is given in Chapter 3. It is shown that the weight quantization, both in the phase and the magnitude, limits the maximum achievable interferer suppression. For this particular testbed, it is the 1 dB step size in weight magnitude that dominates the performance degradation. The quantization in the weights also limits the calibration accuracy. The number of bits of the ADC limits the range of received signal power where an improvement with the adaptive antenna is possible.

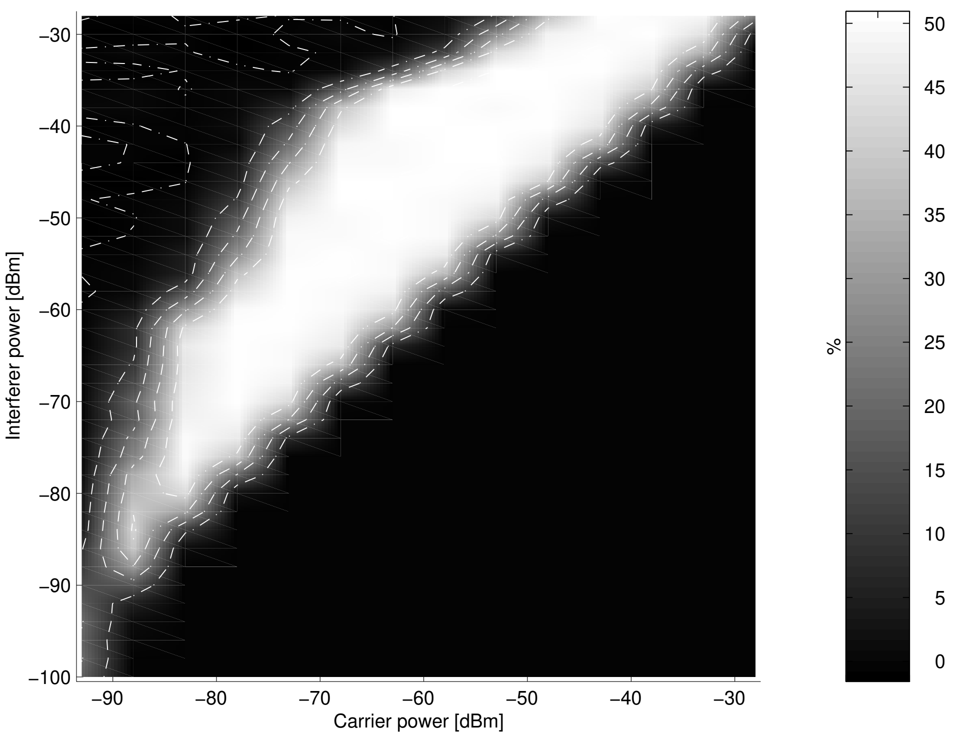


Figure 2.5: Improvement in BER using the adaptive antenna, in percentage units relative to one element antenna. Measurements performed in lab, on an 8 element array antenna, mean values over 10 measurements.

2.5 Outdoor measurements

2.5.1 The measurement range setup

The purpose of the outdoor field trial was first to characterize the system and its performance in a controlled environment with few variable parameters. In the next stage, measurements in a more complicated signal environment with problems of multipath propagation, fading, Doppler spread etc. were to be performed. The adaptive antenna performance was investigated at the outdoor antenna measurement range at FFV Aerotech in Arboga, Sweden.

The front-end was mounted on a turntable enabling 360 degrees azimuth rotation, see Figure 2.7. The circular array was then approximately three meters above the ground level. A five meter horizontal bar was mounted on the rotational turntable on which a horn antenna was placed. A second horn antenna was placed 30 cm above ground, 20 meters away from the turntable.

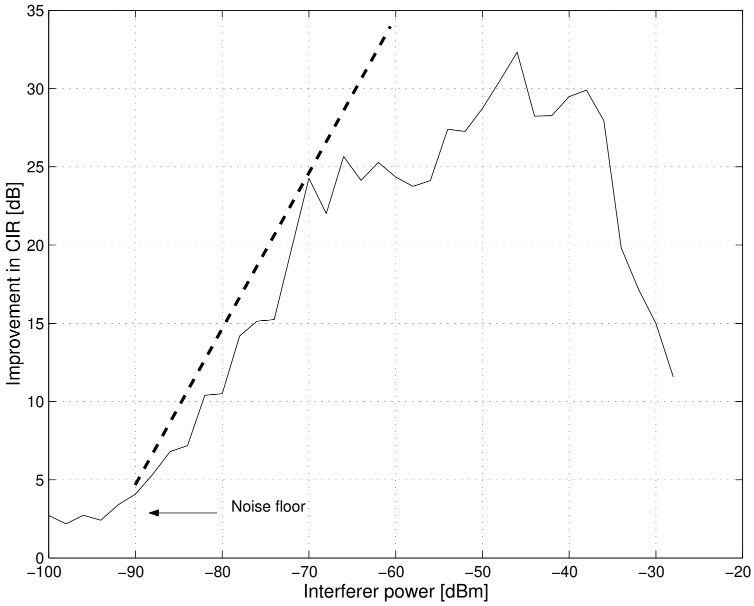


Figure 2.6: Improvement in CIR at carrier power -73 dBm. Note how the improvement levels out for high interferer powers and when the power is below the noise floor

The two horns were used as "mobiles" and illuminated the antenna from two directions. The separation is then adjustable by rotating the turntable. All equipment was stored in a room below the turntable. The distance to the closest obstacles (trees) is about 500 meters providing a rather echo free environment. All experiments were made at the frequency of 1721 GHz.

2.5.2 Measured radiation patterns

Measurement procedure

Two signals with GSM standard GMSK modulation were generated and connected to their respective horn antennas. The separation angle between the interfering and desired signal was set by rotating the turntable. The adaptive antenna was then allowed to adapt on the signal environment. The adaption was then stopped and the weights were held constant. The two signal sources were turned off and a CW signal at 1721 MHz was connected to the field antenna.

During rotation of the turntable, the output power from the adaptive antenna was measured at angles from 0 to 360 degrees in steps of half a de-

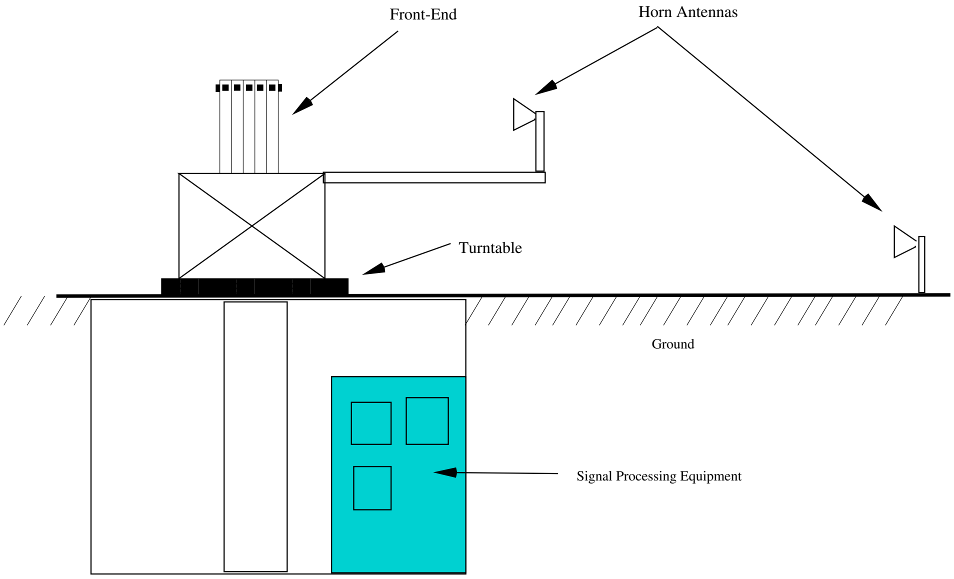


Figure 2.7: Antenna range setup. The front-end is placed on a turntable and is illuminated by two horn antennas. Below the turntable, a small room for the signal processing equipment.

gree, using a spectrum analyzer. The measurements were made for different angle separations and different power levels of the desired and interfering signals.

Results

Figure 2.8 shows the adapted radiation pattern of the second beamformer where the separation angle between the desired and interfering mobile is 90 degrees and $C_{in}=I_{in}$, thus $CIR_{in}=0$ dB. A null is directed towards the interfering signal and the interfering signal is suppressed 25 dB relative to the desired signal.

Diagonal loading

As seen in Figure 2.8, there is no well defined main-lobe against the desired signal. This is due to the short data sequence used to estimate the covariance matrix and the steering vector. The covariance matrix and steering vector is estimated based on the 26 samples in the training sequence. A study on how the number of samples affects the performance of the SMI algorithm can be

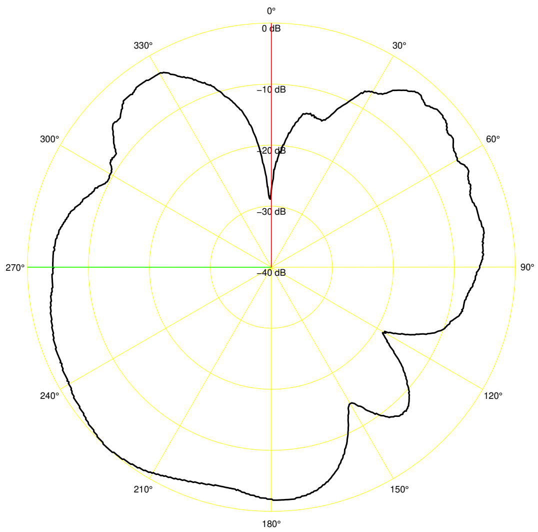


Figure 2.8: Adaptive antenna radiation pattern. Desired and interfering signals of equal strength with DOAs of 270° and 0° respectively

found in [37]. With the desired signal present in the received signal, as is the case in a communication systems, the SMI algorithm converges slower than without its presence. One way of dealing with the slow convergence problem is to use the technique of diagonal loading, i.e. the adding of a small value to the diagonal elements of the covariance matrix.

If we assume that the noise is spatially white, then the perfectly estimated (converged) covariance matrix would have all noise eigenvalues identical and equal to the noise variance. A poor (unconverged) estimate gives non-identical noise eigenvalues resulting in a radiation pattern with high sidelobes and the shape of the pattern changing from solution to solution. By choosing the loading value larger than the noise eigenvalues but smaller than the eigenvalues of the desired and interfering signal the overall equivalent noise level is increased. This results in almost identical noise eigenvalues [25]. The loading value L was therefore chosen so that $L/\sigma^2 \sim 100$, where σ^2 is the noise power. The diagonal loading will decrease the CIR$_{out}$ somewhat, but this is compensated for by an increase in the carrier to noise ratio CNR$_{out}$ because of the decrease in side-lobe levels [37]. Figure 2.9 shows the measured radiation pattern where the signal environment is the same as in Figure 2.8 but with a diagonally loaded covariance matrix. The radiation pattern retains its null in the direction of the interfering signal, but now the main lobe points towards the desired signal and the side-lobe level is

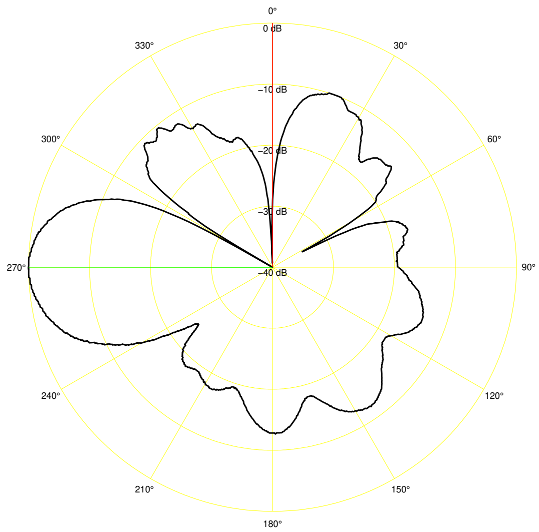


Figure 2.9: Effect of diagonal loading. The covariance matrix is regularized and the radiation patterns show lower sidelobes and a more stable radiation pattern between different realizations.

suppressed down to below -10 dB.

In Figure 2.11 and Figure 2.10 the corresponding weights for ten successive adaptions are presented. When no diagonal loading is used, a completely new weight solution is produced, although the signal environment is stable, with the same DOAs and power levels. This is due to the noise eigenvalues of the covariance matrix that are not equal, and differ from solution to solution. When diagonal loading is introduced, Figure 2.11, the solution becomes stable, because the noise eigenvalues are almost the same. In Figure 2.12, the effect on the radiation pattern is shown. It can be seen that almost the same weights makes successive radiation patterns very similar.

Effect of DOA separation

Figure 2.13(a) and 2.13(b) show radiation patterns of the two beam-formers where the desired signal and interferer are separated 2.5 degrees. Note that the two signals impinging on the array play opposite roles in the two beam-formers. Despite the small angle separation, the interfering signal is suppressed more than 20 dB relative to the desired signal in both plots. Note also the similarity of the two radiation patterns. This indicates that apart from the opposite treatment of the two signals the beam-formers experience

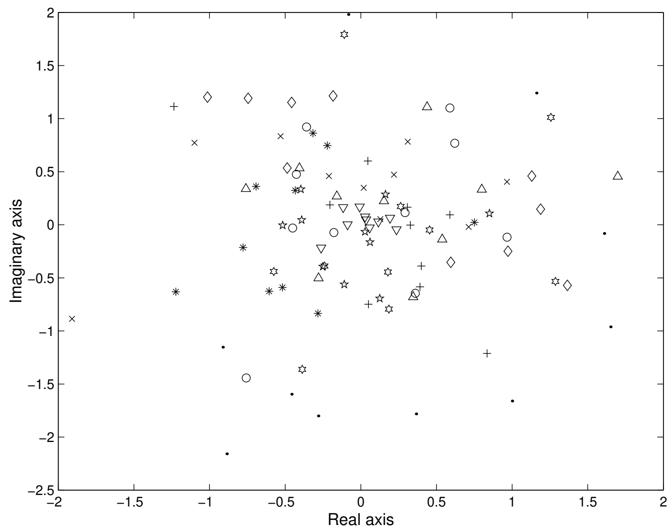


Figure 2.10: Weights in the complex plane. No diagonal loading and ten successive measurements on the same signal environment. Weights corresponding to the same channel is marked using the same symbol. Note how the solution changes completely between different adaptions, due to the short training sequence.

approximately the same noise environment.

For every radiation pattern measured, the suppression of the interferer and the amplification of the desired signal was measured. The results for CIR_{in}=-20 dB is presented in Table 2.2. It can be seen that the suppression of the interferer is, within the error margin, independent of the DOA separation whereas the amplification of the desired signal decreases with a decreasing DOA separation. This is a natural consequence of the limited beam-width of the array, i.e. the most narrow beam possible with the actual array configuration is not narrow enough to provide the main-lobe peak towards the desired signal when a null is directed in the DOA of the interferer. The results of Table 2.2 is in agreement with equation (2.5) and (2.6). When the DOA separation is large, the spatial correlation $|\rho|$ is small and the output CINR (CIR) is approximately constant, as is the improvement in CIR. When the separation angle becomes small, the spatial correlation approaches one and the performance degrades.

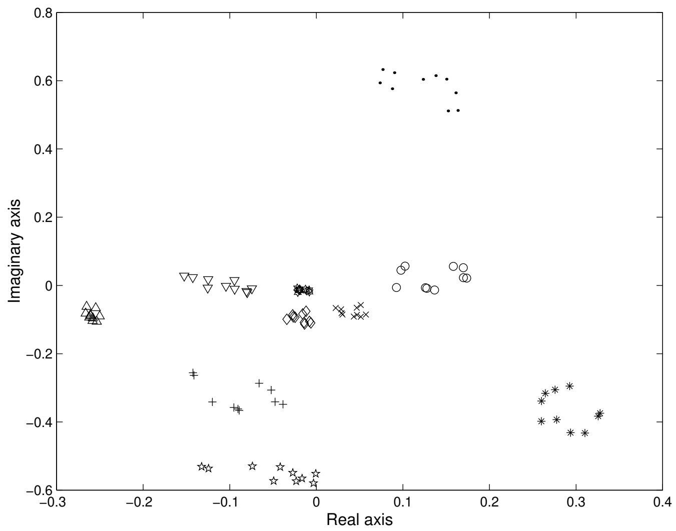


Figure 2.11: Weights in the complex plane when diagonal loading is used. Ten successive measurements are presented. Almost similar weight solution arise every time.

2.5.3 BER for different DOA separation between interferer and carrier

The BER was logged from the base station while rotating the antenna giving BER as a function of DOA separation between interfering and desired signal sources. These measurements were made for different settings of CIR_{in}, see Figure 2.14 for data after the decoding of class Ib bits by the forward error correcting (FEC) code. Figure 2.15 presents the BER of unprotected, class II data. The transition between a BER of 0% and BER of 50% is very sharp as is characteristic for a digital communication system. For CIR_{in}=-10 dB, the minimum angle separation for 0% BER in class Ib bits is less than 4 degrees and for decreasing CIR_{in} this minimum angle increases.

The class Ib bits are coded with a rate 1/2 convolutional code and class II bits have no error correcting coding. As seen from the Figures 2.14 and 2.15, the forward error correcting coding improves the adaptive array antenna ability to separate users in the spatial dimension.

The BER output is related to the CIR_{out} and depends on the base station receiver detection properties. The minimum angular separation allowed until the BER becomes nonzero is an important measure of the performance of the adaptive antenna. It determines when an intra-cell handover has to be

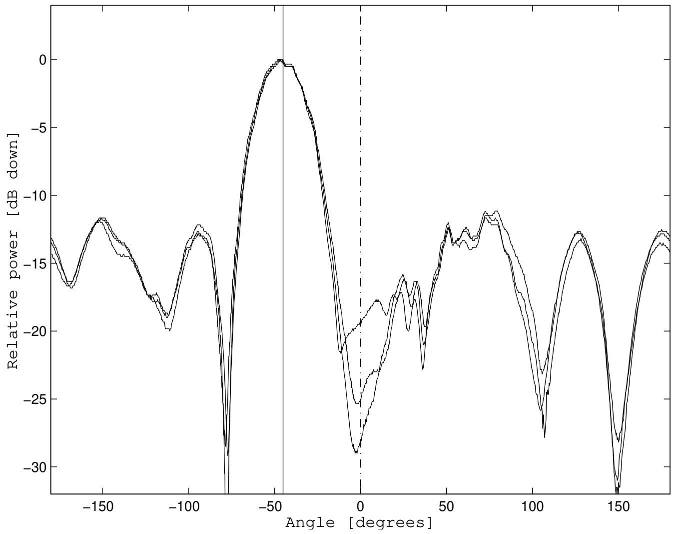


Figure 2.12: Three successive radiation patterns when diagonal loading is introduced. The DOAs and signal levels are identical for all three cases. Desired signal from -48° and interferer from 0°.

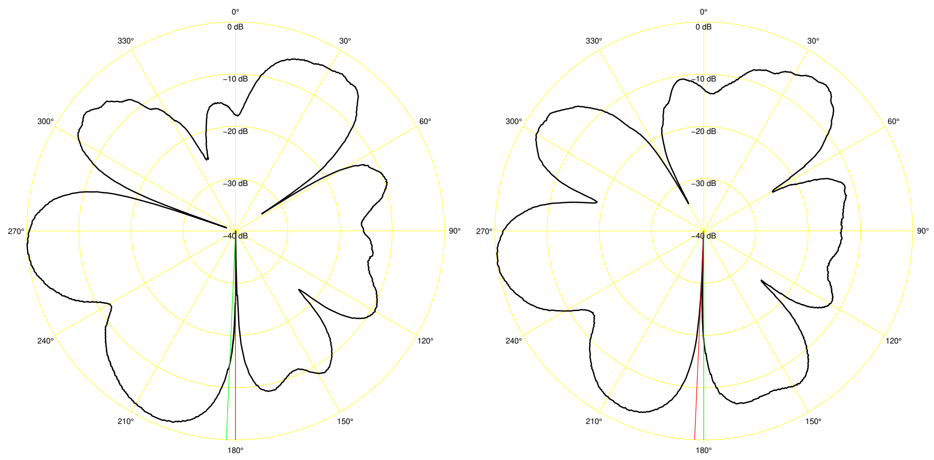
initiated in an SDMA system.

For a CIR_{in}=0 dB a BER of 0% was measured although the plane waves impinging on the array only differed in elevation angle (same azimuth angle). This can be explained by the vertical beamforming of the circular array, i.e. plane waves with the same azimuth angle but with different elevation angles give rise to different spatial signatures or array response vectors, which makes it possible to separate them.

For practical reasons the two illuminating antennas were positioned at different heights as can be seen in Figure 2.7, giving a separation in the elevation plane of approximately eight degrees. Different spatial signatures and low correlation between the training sequences of the two signal sources are enough for the adaptation algorithm to separate the signals, and thus provide a beamforming property in elevation angle also.

2.6 Qualitative test in a multipath environment

The testbed has also been used in a multipath scenario. The venue was a demonstration of the adaptive antenna using speech data transmission with “real” DCS-1800 cellular phones. The place was in central Kista out-



(a) Beamformer one

(b) Beamformer two

Figure 2.13: Desired and interfering signals of equal strength with DOAs of 182.5° and 180°, respectively

side of Stockholm and the occasion was a demonstration of the project in cooperation with Ericsson Radio Access AB. The demonstration gave qualitative results of the performance although no quantitative measurements was conducted.

In the demonstration, the adaptive antenna array was placed on the ground in a yard with high building surrounding it, thus a signal environment with severe multipath components due to many wall reflections. We successfully managed to connect the two DCS-1800 cellular phones transmitting on the same frequency channel and timeslot and used the testbed as a receiver. The voice transmission using the two cellular phones progressed perfectly although the physical separation between the two users was reduced to only 0.2 meters. This small separation distance corresponded to an angle separation at the basestation site of less than 1°, which is to small for bit error free transmission in the line of sight scenario.

Conclusively, the demonstration showed that with an array antenna, a multipath channel between the mobiles and the basestation actually helps the basestation to separate the users. The time delay spread for the multipath components must be considered small as the extent of the yard was not so large. It also shows that weight updating once per frame was sufficient for adapting to the changes in the channel.

DOA separation	Interferer suppression	Carrier gain	Improvement in CIR
180°	31 dB	2 dB	32 dB
135°	30 dB	4 dB	34 dB
90°	30 dB	4 dB	34 dB
45°	32 dB	2 dB	34 dB
10°	29 dB	2 dB	31 dB
5°	27 dB	-10 dB	17 dB
2.5°	26 dB	-8 dB	18 dB

Table 2.2: CIR_{out} improvement in dB for different DOA separation. CIR_{in}=-20 dB. The improvement in CIR is maintained above 30 dB until the DOA separation angle is smaller than 10°.

2.7 Conclusions

The laboratory tests show that the adaptive antenna presented here is capable of improving the CIR ratio more than 30 dB when the desired signal is weak and in the presence of a strong interferer. When the angle between the desired and the interfering signal is decreased the CIR improvement is reduced, mainly due to a loss in carrier gain. With a CIR=-20 dB the minimum separation between the desired and the interfering mobile for an bit-error-free transmission is 10 degrees.

Furthermore, it was verified that diagonal loading is a way to improve the estimation of the covariance matrix, giving lower sidelobes and a distinct main-lobe towards the desired signal.

This Chapter demonstrated the feasibility of using adaptive antennas at base stations for mobile communication networks. Although the quantitative measurements were conducted in the simplest signal scenario with line of sight communication between the mobile transmitters and the base station array antenna, the results are not optimal in the sense that the interferer is not fully suppressed to the noise floor for all interferer power levels. This was expected by the theory of optimal array antennas for received power levels well inside the dynamic range of the receivers. The reason is due to the limited accuracy in the weights of the analog beamformer and the calibration errors. This is more thoroughly studied in Chapter 3.

A demonstration of the testbed in a multipath scenario showed that the weight updating frequency was sufficient to track the changes in the multipath channel. Although no qualitative measurements were performed, the voice quality test was satisfactory. Hence, the SMI algorithm is useful

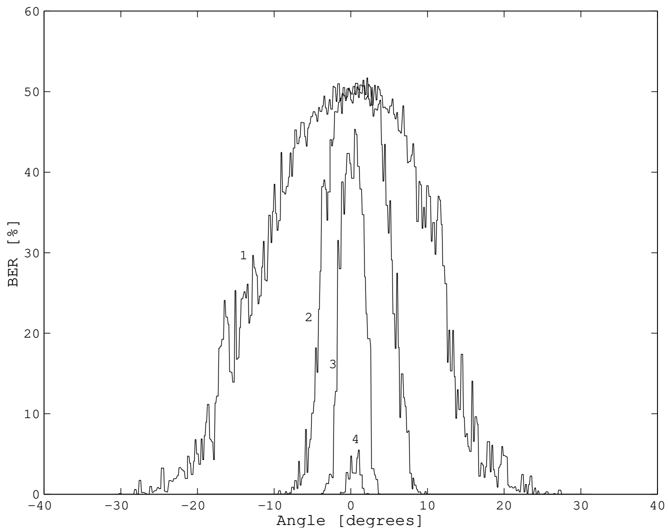


Figure 2.14: BER (after FEC decoding of class Ib data) as a function of interferer and desired signal DOA separation in the horizontal plane. CIR_{in}=-25 dB(curve 1), -20 dB(2), -15 dB(3), -10 dB(4) and 0 dB(5)

in this environment with multipath with short time delay spread.

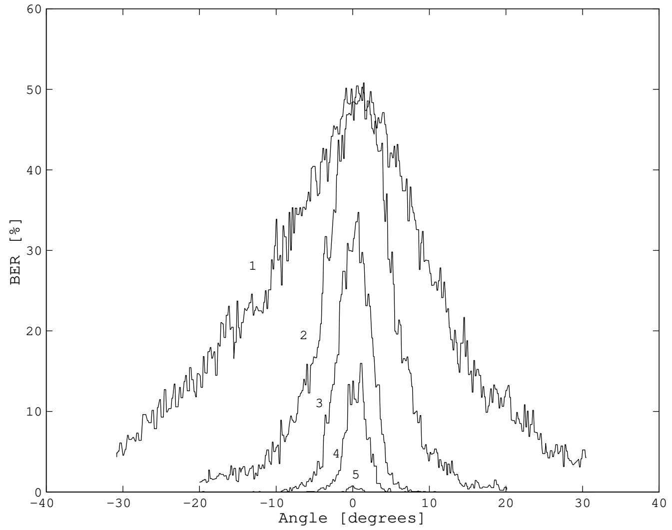


Figure 2.15: BER (Class II unprotected data) as a function of interferer and desired signal DOA separation in the horizontal plane. CIR_{in}=-25 dB(curve 1), -20 dB(2), -15 dB(3), -10 dB(4) and 0 dB(5)

Chapter 3

Performance of Analog Beamforming on Uplink

3.1 Introduction

The topic in this Chapter is to investigate how the implementation of the signal processing in hardware affects the performance of the adaptive array antenna. The investigation is confined to the uplink or receiving antenna array only. We consider the characteristics of the used sampling-receivers (SRX), the resolution of the adaptive weights that constitute the beamformers and the accuracy of the calibration. The goal is to understand how these implementation effects contribute to the total performance.

Furthermore, a hybrid-analog-digital beamformer (ABF) is considered as was described in Section 1.2.3, where the adaptive algorithm is implemented in a digital signal processor (DSP), but the beamformer weights are implemented in hardware. Thus, the beamforming is carried out on the RF or IF frequency, using digitally controlled analog phase-shifters and attenuators. This concept is shown in Figure 3.1. We derive an expression for the adaptive antenna output signal, and create a simulation model to make comparisons with measurements and theory. The simulation model includes the limited dynamic range of the receivers.

The novel aspect introduced in this chapter is the investigation of the implementation errors described below in the framework of mobile communication systems. Comparisons are made with an adaptive antenna testbed, developed for the DCS-1800 system to verify the simulations. Also, the dynamic range of the sampling receiver is included in the investigation. We develop a theoretical model and the theoretical results are compared to the simulations and testbed results.

The results can be used to find out the impact of different implementation techniques on the total performance, to find performance bottle-necks and to help the designer to avoid over-dimensioning of the antenna parts. Although an in-depth theoretical analysis of the implementation errors is not carried out, the discussion of how these errors affect the output CIR will be most helpful. The results from the measurements with the testbed are used to verify the simulation results and the qualitative theoretical expressions. With the simulator, different parameters can be varied to study the impact on the performance, which is measured as the CIR in the ABF output. The results will also explain the performance limit of the testbed presented in Chapter 2.

To experimentally verify the theoretical derivations and the simulations, measurements on the adaptive antenna testbed, described in Section 2.2, is used.

The model of the ABF must consider the limitations of the hardware and we study the following antenna parts:

Sampling receivers The ADCs will introduce quantization noise with a variance that should be in the same order of magnitude or less as the thermal noise variance for best utilization of the dynamic range. So by using more bits in the ADCs the dynamic range is increased and signals with larger amplitudes can be received without saturating the ADCs. Hudson [17] studied the quantization of received signals and in weights separately for an adaptive array and concluded that eight bits in the ADC was necessary to give 40 dB interference suppression with a ten element array. In practice, the dynamic range requirement (up to 90 dB in GSM), implies that a high-end ADC must be used with high resolution (many bits), so the generated quantization noise will not be a serious problem. If the dynamic range is insufficient, then the received signal will saturate the ADC and clipping will occur, which is a nonlinear effect on the signal that is much more severe than the introduction of extra noise power.

DSP The DSP has a finite word length that affects the numerical stability and accuracy of matrix inversions used by some algorithms. The required word-length to achieve desirable performance by using the optimal weights has been studied by Nitzberg [16]. It was shown how the required precision in the DSP (number of internal bits), to achieve the optimal performance, varies by orders of magnitude depending on number of interfering sources and their positions. Also errors from the choice of algorithm are introduced that affect performance, as the number of samples used in the estimation of the covariance matrix in SMI or the achievable precision in the DOA estimation in e.g. MUSIC. The former problem has been studied by Monzingo and Miller [13].

Weighting units The weighting units have finite accuracy which are determined by the type of weight used, and the number of control bits from the DSP. The effect of quantization of weights has been studied by Nitzberg [15] who showed how addition of quantization errors to optimal weights affected the output power. Godara [19] and Voloshina *et al.* [38] studied the effect of phase errors, whereas Davis and Sher [39] compared different weight implementations and their respective output quantization noise power.

Calibration To match the different antenna channels, a calibration must be performed. Quantization and time-varying calibration errors, caused by temperature, humidity variations or mechanical vibrations will affect the achievable accuracy. These effects will produce a residual calibration error that will degrade the performance. Calibration errors were studied by Tsoluous *et al.* [26],[27], where measurements on an adaptive array using a calibration algorithm was presented. In Chapter 4 an on-line calibration algorithm is presented that continuously tracks the changes in antenna channels during “normal” adaptive antenna operation.

The chapter is organized as follows. In Section 3.2 the simulation model is derived. Section 3.3 describes the adaptive antenna testbed and Section 3.4 the results from measurements and simulations. Finally, the chapter is concluded with a discussion, in Section 3.5.

3.2 The hardware model

A signal model of the assumed hardware is described in this section. Figure 3.2 shows the equivalent signal model of the ABF concept presented in Figure 3.1. We assume that all signals are represented by their complex baseband equivalents.

The noise generated in the low noise amplifiers in the receiver front-end and the noise received by the N antennas are modeled as an equivalent $N \times 1$

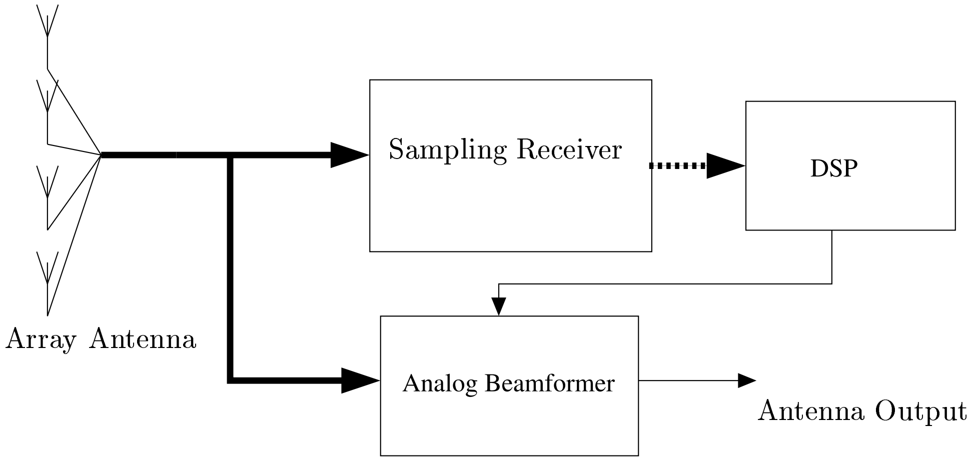


Figure 3.1: Concept of hybrid digital-analog beamformer

noise vector, $\mathbf{n}_t(t)$, which is assumed to be spatially and temporally white[1]. We assume that $\mathbf{n}_t(t)$ is a vector with zero mean variables and covariance matrix $\mathbf{R}_{\mathbf{n}_t} = \sigma_t^2\mathbf{I}$.

We assumed that two narrowband signals impinge on an uniform linear array (ULA) antenna from two distinct directions θ_d and θ_i.

The signal received by the N antennas can then be described by the $N \times 1$ vector $\mathbf{x}(t)$ as

$$\begin{aligned} \mathbf{x}(t) =& \mathbf{x}_d(t) + \mathbf{x}_i(t) + \mathbf{n}_t(t) = \\ =& \mathbf{a}(\theta_d)s_d(t) + \mathbf{a}(\theta_i)s_i(t) + \mathbf{n}_t(t) \end{aligned} \tag{3.1}$$

where $\mathbf{a}(\theta_d)$ and $\mathbf{a}(\theta_i)$ is the array response vector in direction θ_d and θ_i respectively, including antenna element gain and polarization. The signals $s_d(t)$ and $s_i(t)$ denote the desired and interferer signals respectively.

The receivers will introduce noise in the down-converting process and the limited dynamic range of the ADC will determine the maximum distortion-free amplitude before clipping of the received signal occurs. The ADC also introduces quantization noise which will set the minimum detectable signal

[1]This assumption means that we assume that no other interference sources are present which have a directional property. Furthermore, the white noise generated in each antenna branch in the front end are assumed mutually independent. Thus the noise is both temporally and spatially white.

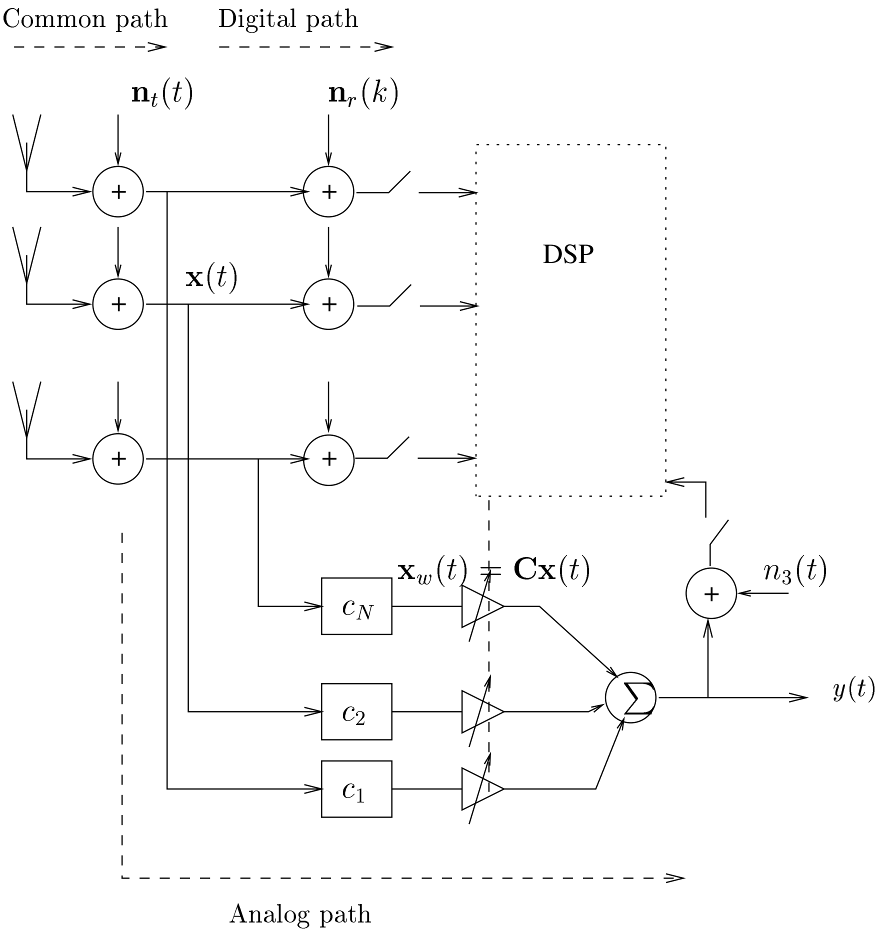


Figure 3.2: Signal model of the adaptive antenna

amplitude of the receiver. In the signal model, the sampling receivers (SRX) are replaced by limiters and a noise source $\mathbf{n}_r(t)$. This noise represents the internal noise generated in the receivers and the quantization noise generated in the sampling process. The corresponding signal vector in the DSP can thus be written as

$$\mathbf{x}_{DSP}(t_k) = \mathrm{sat}\left[\mathbf{a}(\theta_d)s_d(t_k) + \mathbf{a}(\theta_i)s_i(t_k) + \mathbf{n}(t_k)\right] \tag{3.2}$$

where t_k represents the sampling instants and $\mathbf{n}(t_k)$ is the sum of front end thermal noise $\mathbf{n}_t(t)_{t=t_k}$, quantization noise and receiver noise $\mathbf{n}_r(t)_{t=t_k}$.

$$\mathbf{n}(t_k) = \mathbf{n}_t(t_k) + \mathbf{n}_r(t_k) \quad . \tag{3.3}$$

The saturation operator, $sat[\cdot]$, hard-limits the signals, where the limiting amplitude level is dependent on the automatic gain control (AGC) in the receivers and the dynamic range of the ADC. If the signal voltage amplitude exceeds the maximal distortion-less amplitude, the signal is clipped.

3.2.1 Calibration

The aim of the calibration is to measure the transfer function of the channel between the SRX and the weighting units. The transfer functions are used to compensate the weights before they are applied in the weighting of the signals. This is important because the weights are calculated based on the signals from the SRX but applied to the signals entering the weighting units. We make the assumption that the channel between the SRX and the weighting units is wideband compared to the received narrowband signals and that the transfer function is flat over the passband of interest. This implies that the transfer function can be represented by a complex number representing the gain/attenuation and the phase shift of the channel.

By introducing the complex matrix $\mathbf{C} = \mathrm{diag}(c_1, \ldots, c_N)$, the relation of the signal at the weighting units, $\mathbf{x}_w(t)$, and the SRX, $\mathbf{x}(t)$, can be expressed as

$$\mathbf{x}_w(t) = \mathbf{C}\mathbf{x}(t) \quad . \tag{3.4}$$

The off-diagonal elements in $\mathbf{C}$, which here are assumed negligible[2], represent the mutual coupling between the channels in the beamformer. Assume that the algorithm calculates the weight vector $\mathbf{w}_{DSP}$, based on the received signals $\mathbf{x}_{DSP}(t_k)$. To compensate for the differences in the receiving channels, the weights are pre-adjusted to

$$\mathbf{w} = (\mathbf{C}^{-1})^H \mathbf{w}_{DSP} \tag{3.5}$$

where $(\cdot)^H$ denotes hermitian transpose. The analog beamformer output

[2]The coupling between two cables is extremely small and the main contribution to the mutual coupling comes from the antenna array, which here is neglected to simplify the discussion.

signal $y(t)$ will then be

$$y(t) = \mathbf{w}^H \mathbf{x}_w(t) = \mathbf{w}_{DSP}^H \mathbf{C}^{-1} \mathbf{C} \mathbf{x}(t) = \mathbf{w}_{DSP}^H \mathbf{x}(t) \quad . \tag{3.6}$$

Thus the effect of the transfer function $\mathbf{C}$ is canceled.

When the calibration is performed, the weights are adjusted to measure the transfer function for different weight settings. This implies that the accuracy in the weight settings will also have an impact on the accuracy of the calibration. Thus, the calibration has a finite accuracy, so $\mathbf{C}$ will not be known exactly. Also, due to temperature drift and aging, the calibration correction matrix will not describe the actual transfer function. Let the matrix $\mathbf{C}_0$ be the calibration correction matrix that is stored in the DSP. If the calibration is error-free, then $\mathbf{C}_0 = \mathbf{C}$. We can now write $\mathbf{C}_0^{-1}\mathbf{C} = \mathbf{I} + \delta\mathbf{C}$, where $\delta\mathbf{C}$ is a diagonal matrix with complex elements, representing the relative calibration errors. Writing the diagonal elements of $\delta\mathbf{C}$ as $\delta c_l = d_l e^{j\phi_l}$, we separate the relative calibration errors into a magnitude error d_l and a phase error ϕ_l. For further use, the magnitude d_l is assumed to be bounded in the range $[\pm\epsilon_a]$ and the phase error ϕ_l is in the range $[\pm\epsilon_\phi]$.

3.2.2 The SMI algorithm

The DSP calculates the weights by using the sample matrix inverse (SMI) algorithm [13]. The covariance matrix of the received signals and the cross-correlation between the received signal vector $\mathbf{x}(t_k)$ and the reference signal $s_d(t_k)$ is estimated using m samples. As the reference signal, the 26 bit training sequence in the mid-amble of a DCS-1800 timeslot is used. The signal covariance matrix estimate is then

$$\hat{\mathbf{R}}_{xx} = \frac{1}{m} \sum_{k=1}^{m} \mathbf{x}_{DSP}(t_k) \mathbf{x}_{DSP}^H(t_k) \quad . \tag{3.7}$$

The cross-correlation is estimated as

$$\hat{\mathbf{r}}_{xd} = \frac{1}{m} \sum_{k=1}^{m} \mathbf{x}_{DSP}(t_k) s_d^H(t_k) \quad . \tag{3.8}$$

Where$(\cdot)^H$ denotes the complex conjugate transpose. Thus, the reference signal, $s_d(t_k)$, is a modulated, pre-recorded copy of the training sequence of the desired signal, stored in the DSP memory.

The SMI weight vector is now calculated as

$$\mathbf{w} = \gamma \hat{\mathbf{R}}_{xx}^{-1} \hat{\mathbf{r}}_{xd} \tag{3.9}$$

where the factor γ scales the weights for full utilization of the dynamic range of the hardware weights. This weight vector is applied to the RF signals $\mathbf{x}_w(t)$, see Figure 3.2. Due to the finite step size in the hardware weighting units, the weights $\mathbf{w}$ will be quantized and an error vector $\delta_{\mathbf{w}}$ is introduced. The phase and magnitude of the weight quantization errors are assumed to be independent and zero mean, white, and uniformly distributed variables in the ranges $[\pm\epsilon_M]$ and $[\pm\epsilon_\theta]$ respectively. The total weight error covariance matrix is $\sigma_w^2 \mathbf{I}$, where σ_w^2 is derived in Section 3.2.4. The signal on the beamformer output can thus be written as

$$\begin{aligned} y(t) &= (\mathbf{w} + \delta_w)^H \mathbf{C}_0^{-1} \mathbf{C} \mathbf{x}(t) = (\mathbf{w} + \delta_w)^H (\mathbf{I} + \delta\mathbf{C}) \mathbf{x}(t) = \\ &= \mathbf{w}^H \mathbf{x}(t) + \text{error terms} \quad . \end{aligned} \tag{3.10}$$

3.2.3 The output signal power

As a performance measure for the adaptive antenna, we use the CINR on the beamformer output. Therefore an expression for the output power is derived and the different terms are identified as desired (carrier) terms and interferer and noise terms. These are then used to calculate the CINR. Some simplifying assumptions are made to make the analysis possible. The saturating effect of the SRX is not taken into consideration, and the two signals, $s_d(t)$ and $s_i(t)$, are assumed to be sufficiently separated to make the spatial correlation close to zero.

The power of the beamformer output signal can be written as

$$E\left\{|y(t)|^2\right\} = E\left\{((\mathbf{w} + \delta_w)^H \mathbf{C}_0^{-1} \mathbf{C} \mathbf{x}(t))((\mathbf{w} + \delta_w)^H \mathbf{C}_0^{-1} \mathbf{C} \mathbf{x}(t))^H\right\} \quad . \tag{3.11}$$

Using equation (3.10) and assuming that the elements of the stochastic vectors $\delta_{\mathbf{w}}$, $\mathbf{x}(t)$ and the matrix $\delta\mathbf{C}$ are mutually uncorrelated, equation (3.11) can be written as

$$\begin{aligned} E\left\{|y(t)|^2\right\} &= \mathbf{w}^H E\left\{\mathbf{x}(t)\mathbf{x}^H(t)\right\} \mathbf{w} + E\left\{\delta_{\mathbf{w}}^H \mathbf{x}(t)\mathbf{x}^H(t)\delta_{\mathbf{w}}\right\} + \\ &+ \mathbf{w}^H E\left\{\delta\mathbf{C}\mathbf{x}(t)\mathbf{x}^H(t)\delta\mathbf{C}^H\right\} \mathbf{w} + E\left\{\delta_{\mathbf{w}}^H \delta\mathbf{C}\mathbf{x}(t)\mathbf{x}^H(t)\delta\mathbf{C}^H \delta_{\mathbf{w}}\right\} \quad . \end{aligned} \tag{3.12}$$

The terms in (3.12) are referred to as the first, second, third and fourth term. The fourth term contains the product of the two error variables and is neglected in the sequel. The second and the third term are quadratic forms with the error terms $\delta_{\mathbf{w}}$ and $\delta\mathbf{C}$. We assume that the power from these "small" terms can be regarded as noise and interference. This is a pessimistic assumption, because these terms also include some signal which are correlated with the desired signal.

The first term

The first term in (3.12) is the output signal from a beamformer without calibration or weight quantization errors. This term is dominant in the digital beamforming type of antenna, where weighting of the signals is performed in the DSP itself. In the digital beamforming case there is no calibration errors of the type considered here, because the beamforming is performed on the signals in the DSP, $\mathbf{x}_{DSP}(t_k)$. However, a digital beamformer often requires channel matching calibration, between the antennas and the SRX. In the digital beamformer, the weight quantization is due to the finite word-length in the DSP, which usually gives a negligible weight quantization error.

The performance of the error-free SMI beamformer is dependent on the number of samples m used in the estimation (3.7) and (3.8). Due to the short length of the training sequence, the temporal correlation between the two signal sources will be nonzero. Also the separation angle between the two sources, i.e. the spatial correlation, will have impact on the performance of the SMI algorithm. The study by Monzingo and Miller [13] yielded that the SMI algorithm has a fast convergence rate as compared to other known algorithms, e.g. the LMS algorithm. The "convergence" rate of a block-algorithm like SMI can be compared to a recursive algorithm as the LMS algorithm by comparing the number of iterations with the number of data samples used in each block. Their conclusion were that for the SMI algorithm, approximately $2N$ samples in the block are required to obtain weights that give a carrier to interferer plus noise ratio CINR within 3 dB of the optimum achievable CINR at the beamformer output. In the following, to simplify the analysis, we assume that $m > 2N$ and proceed with the derivation, allowing the derived CINR to have an error within 3 dB of the correct value.

For the analog signal $\mathbf{x}(t)$ the following holds

$$\begin{aligned}\mathbf{R}_{xx} &= E\left\{\mathbf{x}(t)\mathbf{x}^H(t)\right\} = \sigma_d^2\mathbf{a}(\theta_d)\mathbf{a}(\theta_d)^H + \sigma_i^2\mathbf{a}(\theta_i)\mathbf{a}(\theta_i)^H + \sigma_t^2\mathbf{I} = \\ &= \mathbf{R}_d + \mathbf{R}_i + \sigma_t^2\mathbf{I} \quad .\end{aligned} \tag{3.13}$$

Where $\mathbf{R}_d$ and $\mathbf{R}_i$ are the desired and interfering signal covariance matrix respectively. Also, σ_d^2 and σ_i^2 is the desired and interfering signal power, respectively. So the first term in (3.12) can be written as

$$\begin{aligned} \mathbf{w}^H E\left\{\mathbf{x}(t)\mathbf{x}^H(t)\right\}\mathbf{w} &= \mathbf{w}^H\mathbf{R}_d\mathbf{w} + \mathbf{w}^H\mathbf{R}_i\mathbf{w} + \sigma_t^2\left|\mathbf{w}\right|^2 \\ &\leq \mathbf{w}^H\mathbf{R}_d\mathbf{w} + \mathbf{w}^H\mathbf{R}_i\mathbf{w} + N\sigma_t^2 w_{max}^2 \end{aligned} \tag{3.14}$$

where w_{max} is the maximum allowed weight magnitude, determined by the hardware and σ_t^2 is the thermal noise variance. The first term represents the power in the desired, or useful signal. The second term is the interferer output power, which also is dependent on the beamformer weights $\mathbf{w}$. Clearly, the adaptive beamformer tries to calculate the weight $\mathbf{w}$ so that $\mathbf{w}^H\mathbf{R}_i\mathbf{w}$ is small. The third term in (3.14) represents the sum of noise power from N uncorrelated noise sources.

The second term

The second term in (3.12) is due to weight quantization which is expanded by using the non-correlation assumption between $\mathbf{x}(t)$ and the weight errors $\delta_{\mathbf{w}}$. Thus,

$$\begin{aligned} E_{x,\delta_{\mathbf{w}}}\left\{\delta_{\mathbf{w}}^H\mathbf{x}(t)\mathbf{x}^H(t)\delta_{\mathbf{w}}\right\} &= E_{x,\delta_{\mathbf{w}}}\left\{\mathrm{Tr}\left\{\mathbf{x}(t)\mathbf{x}^H(t)\delta_{\mathbf{w}}\delta_{\mathbf{w}}^H\right\}\right\} \\ &= \mathrm{Tr}\left\{E_x\left\{\mathbf{x}(t)\mathbf{x}^H(t)\right\}E_{\delta_{\mathbf{w}}}\left\{\delta_{\mathbf{w}}\delta_{\mathbf{w}}^H\right\}\right\} \\ &= \sigma_w^2\,\mathrm{Tr}\left\{\mathbf{R}_{xx}\right\} \\ &= \sigma_w^2(\sigma_d^2|\mathbf{a}(\theta_d)|^2 + \sigma_i^2|\mathbf{a}(\theta_i)|^2 + N\sigma_t^2) \quad . \end{aligned} \tag{3.15}$$

Above, we used the property that the trace of the covariance matrix is equal to the sum of the received signal power. This term, (3.15), is interpreted as interfering or noise power. The weight errors, $\delta_{\mathbf{w}}$, causing signal power to leak through the beamformer, with a power proportional to total impinging power. The amount of "leaking" power is determined by weight error variance through the proportionality constant σ_w^2.

The third term

The third term in (3.12) includes the diagonal matrix $\delta\mathbf{C}$ that models the calibration errors. The structure is a Hermitian form and we find the upper bound as

$$\begin{aligned}\mathbf{w}^H E_{\mathbf{x}}\left\{\delta\mathbf{C}\mathbf{x}(t)\mathbf{x}^H(t)\delta\mathbf{C}^H\right\}\mathbf{w} &= \mathbf{w}^H\delta\mathbf{C}\mathbf{R}_{xx}\delta\mathbf{C}^H\mathbf{w} \\ &\leq \lambda_{max}|\mathbf{w}^H\delta\mathbf{C}|^2 \\ &\leq \mathrm{Tr}\left\{\mathbf{R}_{xx}\right\}|\mathbf{w}^H\delta\mathbf{C}|^2\end{aligned} \tag{3.16}$$

where λ_{max} is the largest eigenvalue of the Hermitian matrix $\mathbf{R}_{xx}$. The bound cannot be tighter, since the bound is attained when the error vector $\mathbf{w}^H\delta\mathbf{C}$ is co-linear with the eigenvector to $\mathbf{R}_{xx}$ associated with the largest eigenvalue. The largest eigenvalue has to be found on a case-by-case basis. Thus, we resort to examine the worst possible case. Since $\mathbf{R}_{xx}$ is Hermitian positive definite, all eigenvalues are real and positive. Thus the largest eigenvalue is less than the trace of $\mathbf{R}_{xx}$.

Furthermore, we cannot evaluate the vector norm expression, because it depends on the magnitude of the weights and the calibration errors. We resort to an upper bound, assuming that all weight magnitudes are smaller than w_{max}, which is the maximum possible weight that can be steered out by the hardware weighting units. Similarly, we use the upper bound of the calibration errors c_{max}. Thus by setting $|w_l|^2 = w_{max}^2$ for all weights l and $|c_l|^2 = c_{max}^2$ for all calibration paths, we can write the vector norm in expression (3.16) as

$$\begin{aligned}\left|\mathbf{w}^H\delta\mathbf{C}\right|^2 &= |w_1^* c_1|^2 + \cdots + |w_N^* c_N|^2 \leq |w_1|^2|c_1|^2 + \cdots + |w_N|^2|c_N|^2 \leq \\ &\leq N w_{max}^2 c_{max}^2\end{aligned} \tag{3.17}$$

where the Cauchy-Schwartz inequality was used. Now, we use that the trace of $\mathbf{R}_{xx}$ is equal to the total impinging power on the array and rewrite (3.16) as

$$\mathbf{w}^H E_x\left\{\delta\mathbf{C}\mathbf{x}(t)\mathbf{x}^H(t)\delta\mathbf{C}^H\right\}\mathbf{w} \leq N w_{max}^2 c_{max}^2 (\sigma_d^2|\mathbf{a}(\theta_d)|^2 + \sigma_i^2|\mathbf{a}(\theta_i)|^2 + N\sigma_t^2). \tag{3.18}$$

The value of c_{max}^2 is related to the assumed maximum magnitude and phase errors in Section 3.2.4. Similar to the weight error term (3.15), the total impinging power, described by the trace of $\mathbf{R}_{xx}$ is weighted by a proportionality constant. The constant in this case depends on the maximum calibration error c_{max}.

The output CINR

Using the equations above, we now derive an expression for the beamformer output CINR. The effect from using a finite number of samples when estimating the covariance matrix $\mathbf{R}_x x$ and the cross-correlation vector $\mathbf{r}_{xd}$ is neglected. Also, the saturation of the SRX discussed above is not taken into consideration. The power of the desired, or useful signal in the beamformer output is described by the first term in equation (3.14). The two other terms in equation (3.14) and the equations (3.15) and (3.18) are assumed to be interferer and noise power terms. Thus, the beamformer output can be expressed as:

$$\begin{aligned}\text{CINR} \geq &\left(\mathbf{w}^H \mathbf{R}_d \mathbf{w}\right) / \\ &\left(\mathbf{w}^H \mathbf{R}_i \mathbf{w} + N\sigma_t^2 w_{max}^2 + (\sigma_w^2 + Nc_{max}^2 w_{max}^2)(\sigma_d^2 |\mathbf{a}(\theta_d)| + \sigma_i^2 |\mathbf{a}(\theta_i)| + N\sigma_t^2)\right) .\end{aligned} \tag{3.19}$$

To maximize this expression, the weight calculation algorithm should make $\mathbf{w}^H \mathbf{R}_i \mathbf{w}$ as small as possible. Thus the weight vector $\mathbf{w}$ will yield an antenna radiation pattern[3] with “nulls” in the direction of the interfering sources.

The expression (3.19) was derived using several simplifications, but it illustrates how the desired and interfering signal power “leaks” through the beamformer, due to calibration errors and weight quantization. The system designer should choose the implementation of the weighting units so that the weight error variance σ_w^2 is equal in magnitude to the term $Nc_{max}^2 w_{max}^2$ that represents the calibration errors. In the limit when calibration errors and the weight quantization errors approaches zero, i.e. $c_{max} \to 0$ and $\sigma_w^2 \to 0$, the CINR approaches the well known expression for CINR of a digital beamformer,

$$CINR_{DBF} \simeq \frac{\mathbf{w}^H \mathbf{R}_d \mathbf{w}}{\mathbf{w}^H \mathbf{R}_i \mathbf{w} + N\sigma_t^2 w_{max}^2} \quad . \tag{3.20}$$

How the different terms in the denominator dominates in different cases is investigated by the simulations and the measurements presented in Section 3.4.

[3] The term *radiation pattern* is used in this thesis, irrespectively of antenna operation in the receiving mode or the transmitting mode. This is due to the reciprocity of the transmitting radiation pattern and receiving sensitivity pattern when the same beamformer weights are applied and the same frequency is used.

3.2.4 The weight error variance and calibration errors

The expressions derived in the previous section, and the CINR in (3.19) contains the weight error variance σ_w^2 and the maximal calibration error c_{max}. In this section, expressions for the weight error variance and the maximal calibration error are calculated, given the respective phase and magnitude accuracies. These expressions are later used to calculate the theoretically expected CIR_{out} from the beamformer.

The weight error variance σ_w^2

Since the testbed weighting units use attenuators and phase shifters as weights, we must find the relation between magnitude and phase accuracy and the resulting weight error variance. The probability density function of the weight quantization errors of different weight implementations was derived by Davis and Sher [39]. They considered the amplitude and phase weighting technique with linear weight magnitude steps. Here we derive an upper bound for the weight error variance and consider logarithmic attenuators in the weighting units.

An arbitrary weight with magnitude M and phase θ with magnitude accuracy $\pm\epsilon_M$ and phase accuracy $\pm\epsilon_\theta$ can be written as

$$w = Me^{j\theta} \quad . \tag{3.21}$$

Assuming that the magnitude and phase errors are small, and by differentiation of both sides of (3.21) and collecting terms gives

$$|dw|^2 = dM^2 + M^2 d\theta^2 \tag{3.22}$$

where dM is the magnitude error and $d\theta$ is the phase error. We assume that dM and $d\theta$ are independent, random variables and that dM is uniformly distributed in the interval $\pm\epsilon_M$ and $d\theta$ uniformly distributed in the interval $\pm\epsilon_\theta$. Taking the expectation value of (3.22) then gives

$$\sigma_w^2 = \sigma_M^2 + M^2\sigma_\theta^2 \tag{3.23}$$

where $\sigma_M^2 = \epsilon_M^2/3$ and $\sigma_\theta^2 = \epsilon_\theta^2/3$.

The total error variance (3.23) is dependent on the weight magnitude M. An upper bound on the weight error variance is given by replacing M

with the maximum weight magnitude w_{max}, thus

$$\sigma_w^2 \leq \sigma_M^2 + w_{max}^2 \sigma_\theta^2 = \frac{1}{3}\left(\epsilon_M^2 + w_{max}^2 \epsilon_\theta^2\right) \quad . \tag{3.24}$$

The contribution from the phase error to the sum in equation (3.24) is thus dependent on the maximal weight magnitude. This is intutievly explained by the increased distance between two points on a circle´s circumference when the radius of the circle is increased. Thus, the weight quantisation noise is smallest for small weight amplitudes.

The calibration error term c_{max}^2

The constant c_{max}^2, given in equation (3.17), which describes the maximum squared magnitude of the calibration errors is bounded by the squared sum of the magnitude and phase errors. See Figure 3.3. If the relative calibration

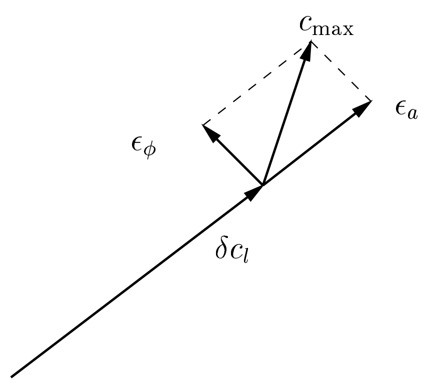


Figure 3.3: The squared calibration error is bounded by $c_{max}^2 = \epsilon_a^2 + \epsilon_\phi^2$

errors, defined in Section 3.2.1, is written as $\delta c_l = d_l e^{j\phi_l}$, then c_{max}^2 is given as

$$c_{max}^2 = \epsilon_a^2 + \epsilon_\phi^2 \tag{3.25}$$

where ϵ_a and ϵ_ϕ is the magnitude and phase ranges for the calibration errors. The expressions (3.25) and (3.24) can now be used when calculating the theoretical CINR in equation (3.19).

3.3 The adaptive antenna testbed

3.3.1 Hardware

In this section the adaptive antenna testbed is described briefly. The testbed is described in Section 2.2. For a detailed description, see [33].

The weighting units has a phase resolution of 1° and uses logarithmic attenuators with 1 dB magnitude step size and a 50 dB range between [-2 dB,-52 dB][4]. Thus, the constant w_{max}=-2 dB. The weight magnitude variance is then calculated using (3.24) as

$$\begin{aligned}\sigma_M^2 &= \frac{(0.5 \cdot (10^{\frac{1}{20}} - 1))^2}{3} = 1.24 \cdot 10^{-3} \\ \sigma_\theta^2 &= \frac{(\frac{0.5\cdot\pi}{180})^2}{3} = 0.025 \cdot 10^{-3} \\ w_{max}^2 &= 10^{-\frac{2}{10}} \\ \Rightarrow \sigma_w^2 &\leq 1.25 \cdot 10^{-3}\end{aligned} \tag{3.26}$$

where a conversion from dB scale to the linear scale is performed. It should be noted that the main contribution to the weight error variance of the testbed arise from magnitude errors (the step attenuators). The phase error contribution is two orders of magnitude smaller than the magnitude error contribution. Noticeable is that the weight magnitude variance is calculated for the worst case using logarithmic weighting units, i.e. at minimum attenuation. When the weight magnitude is smaller the steps are smaller, due to the logarithmic nature of the weights. A simple calculation yields that at a weight magnitude of -16dB, the phase error variance and the magnitude error variance are equal.

Calibration of the antenna array is performed prior to normal operation of the testbed. A continuous wave (CW) signal is injected at one antenna element at a time by directional couplers. The received signal is compared to the beamformer output signal and the phase and magnitude of the transfer functions $d_l e^{j\phi_l}$ can be calculated. The calibration have an accuracy of 1° in phase and 0.75 dB in magnitude [33]. Using equation (3.25), the relative calibration error constant c_{max}^2 can be calculated as $c_{max}^2 = 8.4 \cdot 10^{-3}$. This

[4] The reason for choosing logarithmic attenuators, is that in the design process, it was believed that a high dynamic range of the weights were beneficial. The logarithmic attenuator was thus chosen in favor of the linear one. The logarithmic attenuator has a range of attenuation of 50 dB but it was later discovered that a 50 dB range is overabundant. The beamformer seldom uses weight magnitudes below -15 dB.

implies that the proportionality constant $Nc_{max}^2 w_{max}^2$ in equation (3.19) is of an order of magnitude larger than the weight error constant σ_w^2. Thus, the calibration errors limits the performance for the testbed, which is verified by the simulations in Chapter 3.4.

3.3.2 Measurement setup

Measurements were performed in a laboratory, to validate the performance of the adaptive array antenna in an easy controllable signal environment. The front end array antenna is replaced by an 8×8 Butler matrix. The output ports of the Butler matrix are connected to the receivers. The eight input ports are used to emulate signals impinging on an uniform linear array from eight different directions. The use of a Butler matrix will ideally make the array response vectors orthogonal, i.e. $\mathbf{a}(\theta_i)^H \mathbf{a}(\theta_d) = 0$ and thus, the spatial correlation between the two signals is zero. But the hardware Butler matrix have imperfections from the manufacturing, in our case a peak error 0.8 dB in amplitude and 8.5° in phase. These errors will make the array response vectors $\mathbf{a}(\theta_d)$ and $\mathbf{a}(\theta_i)$ non-orthogonal which corresponds to a more realistic signal environment.

Two signal generators were connected to the Butler matrix to emulate signals impinging ideally from $-61°$ and $-7.2°$. The signal generators transmitted pseudo random binary data as traffic data. In DCS-1800, each traffic data burst contains 26 bits, located in the middle of the burst, known as the training sequence. The training sequence is a static bit pattern that is used for tuning the temporal equalizer in the receiver. In DCS-1800, eight different training sequences (bit patterns) are defined, and we used number 0 and number 4 in the DCS-1800 standard for the two users respectively. The training sequence is then used in the SMI algorithm to calculate the beamformer weights.

3.3.3 Comparison of measurements and theory

In practice, the CINR in equation (3.19) cannot be measured, due to the inability to separate the desired signal from the interfering and noise signal. Instead, we measured a modified carrier to interference ratio on the adaptive antenna output, denoted CIR_{out}, in fact, we measured the carrier plus noise to interferer plus noise ratio, where "noise" includes the thermal noise, calibration error noise and the weight quantization noise.

To perform measurements, the levels of the two signals were set and the adaption of the beamformer weights was started. To measure CIR_{out} for an arbitrary weight realization, the adaption was stopped at an arbitrary time instant. When the adaption was stopped, the weight update also stopped and the CIR in the beamformer output could be measured. This was made in a two step procedure, so the interferer power and the desired signal power was measured separately: The signal generator emulating the desired signal was turned off and the ABF interferer plus noise output power was measured using a spectrum analyzer. Then the desired signal's generator was turned on and the interfering signal's generator was turned off and the ABF desired signal plus noise output power was measured. By assuming that the system is linear and the superposition principle holds, the CIR_{out} can be expressed by the use of equation (3.11) to (3.18) as

$$\mathrm{CIR}_{out} \leq \frac{\mathbf{w}^H \mathbf{R}_d \mathbf{w} + N\sigma_t^2 w_{max}^2 + (\sigma_w^2 + Nc_{max}^2 w_{max}^2)(\sigma_d^2 |\mathbf{a}(\theta_d)|^2 + N\sigma_t^2)}{\mathbf{w}^H \mathbf{R}_i \mathbf{w} + N\sigma_t^2 w_{max}^2 + (\sigma_w^2 + Nc_{max}^2 w_{max}^2)(\sigma_i^2 |\mathbf{a}(\theta_i)|^2 + N\sigma_t^2)} . \tag{3.27}$$

The second and third term in the numerator of equation (3.27) makes the CIR_{out} estimate too optimistic. Therefore the expression (3.27) is larger or equal to the actual CIR_{out}. The term $\mathbf{w}^H \mathbf{R}_i \mathbf{w}$ is the interferer signal power that leaks through the correct beamformer weights, and is usually very small, if the DOA separation is large enough, which is the case in our measurements. It will therefore be neglected in the following. The term $\mathbf{w}^H \mathbf{R}_d \mathbf{w}$ expresses the received power from the desired signal. If the DOA separation is large enough it can be approximated as a constant K multiplied with the desired signal power, σ_d^2, where K is the array gain, including amplifiers in the front end and in the weighting units, degraded by the spatial and temporal correlation of the two signals. By making these simplifications, equation (3.27) can be rewritten as

$$\mathrm{CIR}_{out} \lesssim \frac{K\sigma_d^2 + N\sigma_t^2 w_{max}^2 + (\sigma_w^2 + Nc_{max}^2 w_{max}^2)(\sigma_d^2 |\mathbf{a}(\theta_d)|^2 + N\sigma_t^2)}{N\sigma_t^2 w_{max}^2 + (\sigma_w^2 + Nc_{max}^2 w_{max}^2)(\sigma_i^2 |\mathbf{a}(\theta_i)|^2 + N\sigma_t^2)} . \tag{3.28}$$

Clearly, if the desired signal power σ_d^2 is large compared to the interfering signal power σ_i^2 and noise power σ_t^2, then the CIR_{out} in equation (3.28) will differ from the CINR_{out}, due to the large term in the numerator. If the interferer to carrier ratio is high, say 20 dB, then equation (3.28) will resemble equation (3.19) well. Conclusive, we expect the theory and the measurements to correspond well for high interference to noise power ratios.

3.4 Simulation and measurement results

3.4.1 Simulation setup

The simulation was performed to imitate the testbed as much as possible. The simulation model is shown in Figure 3.2.Two GMSK-modulated signals were generated with 8 samples per symbol. The Gaussian filter had an impulse response length of 6 symbols and the normalized bandwidth was 0.3. As training sequences, the training sequences 0 and 4 given by the DCS-1800 standard were used. The received signal was quantized and the calculated weights were used with finite accuracy. Two impinging signals from the same directions as generated by the Butler matrix in the measurements were used in the simulations. Furthermore, calibration errors were also introduced, to emulate the testbed.

The ADCs used the sampling frequency of 270 kHz, the same as in the adaptive antenna testbed. The covariance matrix (3.7) and the cross-correlation vector (3.8) were estimated using the 26 complex samples. The direction of arrivals (DOA) for the two signals were equal to the DOA:s in the measurements, described in Section 3.3. The array response vectors were slightly distorted to emulate the 0.8 dB magnitude errors and 8.5° phase errors in the Butler matrix and to make the spatial correlation nonzero.

The weight vector was calculated and normalized using (3.9). The calculated weight vector was quantized to the desired accuracy in magnitude and phase and applied to the signals at the weighting units after multiplication with the matrix $(\mathbf{C}^{-1})^H$ as described by equation (3.5).

3.4.2 Validation of simulation model

After adaption, the same procedure to measure CIR_{out} as described in Section 3.3.3 for the measurements was used to measure the CIR_{out} in the simulation, to make the comparisons fair. A GMSK modulated PRBS signal using $m_l = 200$ samples was used to estimate the CIR_{out}. To verify the theoretical model and the simulation results, a comparison is made in Figure 3.4. The simulation parameters are set equal to the parameters used in the adaptive antenna testbed, e.g. 8 bit ADCs and 1° and 1 dB weight accuracy. Figure 3.4 shows the measured output CIR, denoted CIR_{out} as a function of interfering signal power, when the desired signal power was held constant at two different levels, -44 dBm and -54 dBm respectively. The figure also shows the corresponding simulation results and the theory, described by equation (3.28). The measured curves constitute mean values over ten measurements and the simulated are averaged over 100 simulations.

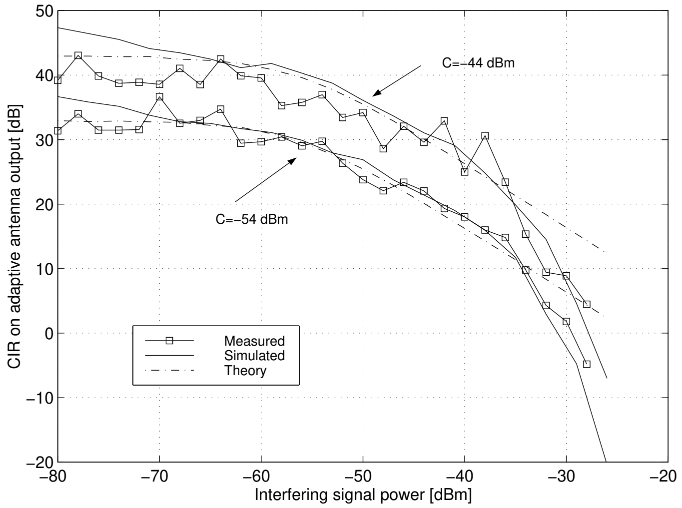


Figure 3.4: CIR_{out} as a function of interferer power. Comparisons between measurements, simulation and theory (equation (3.28)). Carrier power constant at -44 dBm and -54 dBm respectively. Noise level at -76 dBm.

The standard deviation is 7 dB and 0.8 dB in measurements and simulations respectively. The theoretical expression does not consider the limited dynamic range of the ADC, so the theoretical CIR_{out} is larger than the measured and simulated CIR_{out} when the interfering signal saturates the ADC.

The theory, equation (3.28), predicts the CIR_{out} to reach a constant level when the interferer is decreased below the thermal noise level, i.e. when the dominating term in the denominator of equation (3.28) is the noise term $N\sigma_t^2 w_{max}^2$. This is verified by the measurements and the level is determined by the desired signal power and the weight and quantization error variances.

Noticeable is also how well the theory and measurements agree for high carrier to noise ratios. It was in this region were the CIR_{out} approximation, equation (3.28), was less accurate.

3.4.3 Number of ADC bits

By increasing the number of bits used in the ADC, the dynamic range of the SRX is increased, as shown in Figure 3.5. Here the same calibration errors and weight errors as in the testbed is used. With a larger dynamic range, the SRX can receive a signal with higher power without saturating. Increasing the number of bits from 10 to 12 will not improve the output CIR in this scenario, because the SRX is not saturated for these input levels. The limiting factor in the 10 and 12 bit case is the weight and calibration errors.

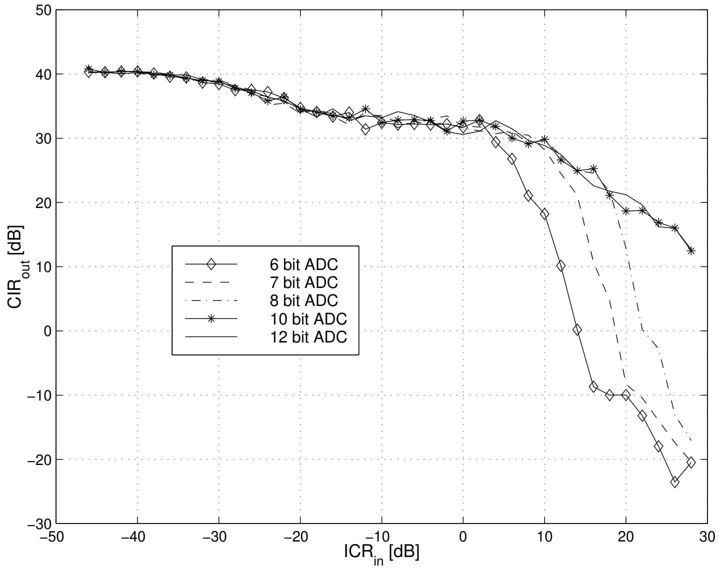


Figure 3.5: CIR$_{out}$ [dB] when the number of ADC bits is varied as a function of CIR. Carrier constant at CNR=42 dB and interferer power varied. Note that the abscissa is shown in interferer to carrier ratio.

3.4.4 Weight accuracy

To solely investigate the impact of weight accuracy on the performance of the adaptive antenna array, we assumed a perfect calibration, i.e. $\mathbf{C}_0^{-1}\mathbf{C} = \mathbf{I}$. Also, an 8 bit ADC was used to make comparisons with the adaptive antenna testbed possible. The CIR$_{out}$ was measured for different settings of the range of the relative weight errors ϵ_M and ϵ_θ.

The CIR$_{in}$ that a single virtual omnidirectional antenna would measure was set to -15 dB and the carrier to noise ratio (CNR) was 42 dB. The CIR$_{out}$ on the adaptive antenna output was estimated and the results are presented in Figure 3.6.

Figure 3.6 shows that a 30 dB CIR$_{out}$ can be achieved if the phase

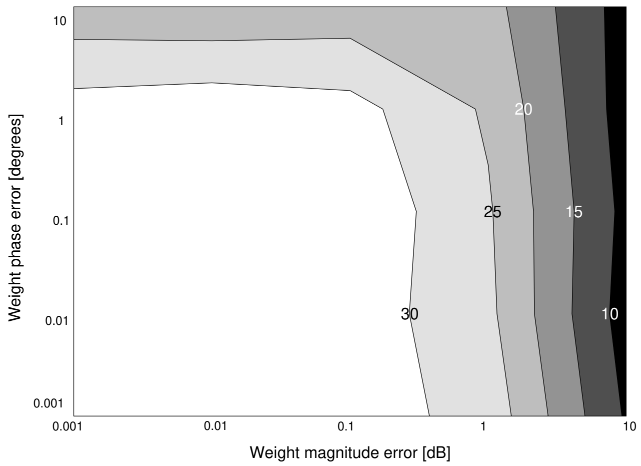


Figure 3.6: CIR_{out} [dB] as a function of quantization of phase and magnitude of weights, CIR_{in}=-15 dB, CNR=42 dB. The different shades refer to certain CIR_{out} levels with white and black being the largest and smallest respectively.

and magnitude quantization steps are less than 1° and 0.3 dB. Thus, the maximal CIR improvement in this scenario is 45 dB and decreasing the weight quantisation steps further will not improve the antenna performance for this particular DOA and signal levels. To achieve a larger improvement, more antennas can be used, which will make the spatial correlation smaller and the ability to suppress interferers larger.

3.4.5 Calibration errors

We investigate how the calibration accuracy affects the CIR_{out} in Figure 3.7. This figure displays CIR_{out} as a function of the weight error variance, for four different calibration accuracies. A weight error variance less than 10^{-5} will not further improve the CIR because the calibration errors limit the maximum achievable CIR[5]. To compare with the testbed, with a weight variance of $1.25 \cdot 10^{-3}$ the maximum achievable CIR_{out} in Figure 3.7 is approximately 24 dB if the calibration is performed without errors. This should be com-

[5]This weight error variance could for example be achieved with 1° phase step size and 0.1 dB weight amplitude step size.

pared to the measured 18 dB from the testbed in the same conditions. Thus, to improve the testbed performance, effort should be put to improve the calibration algorithm, in favor for improving the weight accuracy. Improving weight accuracy will raise CIR_{out} only a few dB.

When the weight errors are negligible, an amplitude calibration error of 1% gives a CIR_{out} degradation of 7-8 dB from the ideal case, with no errors. This is a severe degradation and the conclusion is that the calibration must have a high accuracy and also be performed frequently to maintain the adaptive antenna's interferer suppression performance over a long time of operation.

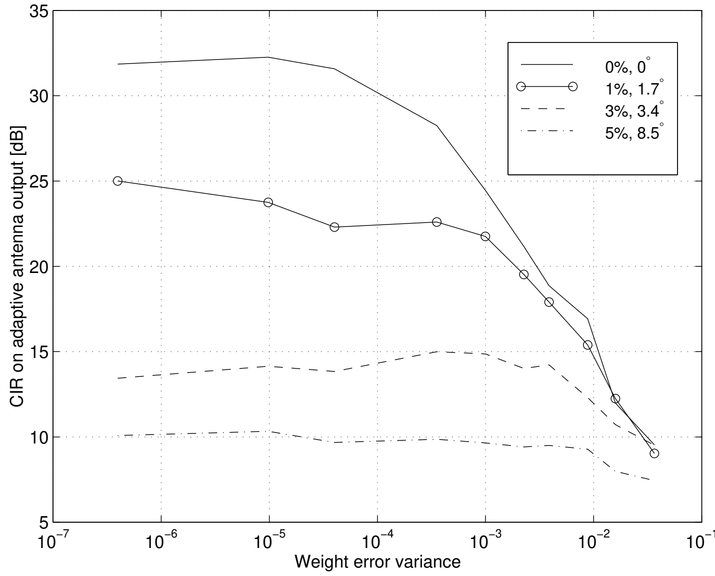


Figure 3.7: CIR_{out} as a function of the weight quantization error for different calibration errors, CIR_{in}=-15 dB, CNR=42 dB.

The effect of calibration errors when the CIR_{in} is varied by varying the interference to noise ratio (INR) is presented in Figure 3.8. The carrier to noise ratio (CNR) was held constant at 22dB. A comparison with the adaptive antenna testbed was made, and the CIR_{out} was measured for different calibration errors and different CIR_{in}. The weight quantization steps were set to 1 dB and 1°. The measured curve fits the curve with calibration error of 1% in relative magnitude and 1° in phase. Furthermore, when INR $<$ 22 dB, i.e. when CIR$> 0dB$, the calibration errors have an negligible effect on the antenna performance. When the interferer power gets large, the output

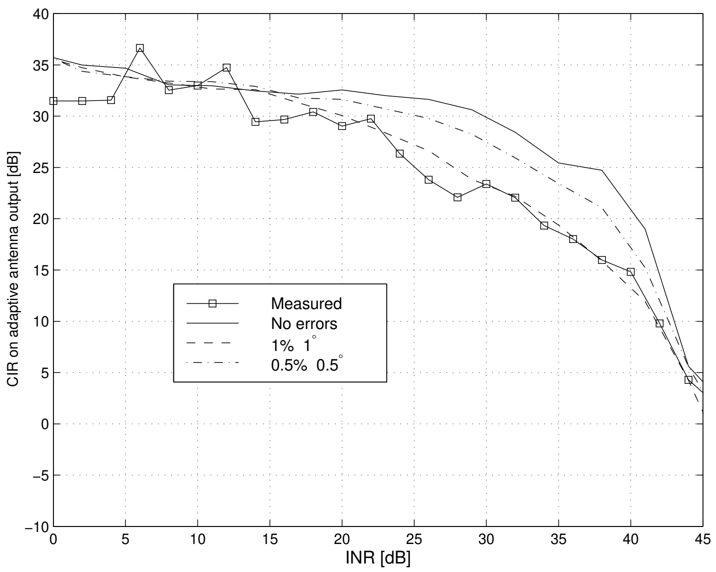


Figure 3.8: CIR_{out} as a function of interferer to noise ratio (INR) for different calibration errors. CNR=22 dB

noise, described in equation (3.15) and (3.16), is increasing and the CIR decreases. When the total input power saturates the ADC, the CIR decreases abruptly and the antenna cannot maintain a CIR above 0 dB on the output. Note that the RF analog signal on the output from the adaptive antenna is not saturated. But the signal used in the DSP to calculate the weights is saturated and the CIR_{out} drops.

3.5 Conclusions

A ten element adaptive antenna testbed have been built for the DCS-1800 system, working in uplink and uses the SMI-algorithm to calculate the weights. The testbed is of the hybrid analog-digital type, where the weights are calculated in a DSP, but the weighting of the received signals is performed by hardware implemented weighting units working at the RF frequency. To study the performance degradation when implementing the algorithm in hardware, a qualitative expression for the output CINR was derived. The weight quantization errors and the calibration errors increases the output noise power, thus decreasing the output CINR. The decrease is inversely proportional to the total impinging signal power on the array.

A simulation tool that model the hardware imperfection was developed, to study how the performance would improve if the calibration errors and weight errors was reduced. The simulation results were compared to a derived theoretical expression. The expression was derived under some simplifying assumptions, to obtain a quantitative expression for the adaptive antenna output CIR.

The simulations gave some useful results. First, it was concluded that in the adaptive antenna testbed, the magnitude steps in the weighting units (1 dB) were too coarse, as compared to the phase accuracy ($1°$). Thus the bottleneck in the testbed is identified as the coarse magnitude steps. The system designer should put effort into making the phase and magnitude error variances equal to minimize a large overhead in either the phase-shifter or attenuator accuracy.

Furthermore, the balance between weight accuracy and calibration accuracy was examined, and it was shown that with a certain calibration accuracy, the output CINR could not improve above a certain limit, regardless of weight accuracy. Thus, the system designer should balance these two sources of error. For the testbed, it can be concluded that improving the calibration accuracy would gain more in output CIR as compared to improving the weight accuracy. Note that it was assumed that the number of bits in the ADC:s were sufficiently large, i.e. chosen so that the dynamic range is not a limitation of the system performance.

Chapter 4

On-line Calibration Algorithms

4.1 Introduction

In this chapter an important issue in array technology is addressed, namely the calibration of the antenna array. As explained in Chapter 3, the performance of the adaptive antenna depends on the accuracy with which the calibration can be performed. During operation, the calibration data becomes old and a re-calibration is necessary to re-establish a high performing adaptive antenna.

Calibration of antenna arrays for mobile communication systems has earlier been discussed in [26]. There, the sensitivity of a digital beamforming system to calibration errors was studied and a method to calibrate the array prior to operation was presented. In [28] an auto-calibration algorithm for the transmitter part of a digital beamformer was presented.

In this chapter, the uplink, or receiving part of the array antenna in a TDMA system is considered. It is also assumed that we use the hybrid-analog beamformer (ABF), where the beamformer weights are calculated in a digital signal processor (DSP), but the weighting is performed by hardware weighting units on the RF signals, cf. Section 1.2.3.

When using the hybrid-ABF, it is important to know the transfer function between the input sampling receivers and the point in the signal path where the weights are applied to the signals, to be able to compensate for this difference. Prior to start-up, this transfer function is measured using an off-line calibration algorithm. However, there are active components in the weighting units and the receivers, which are sensitive to temperature variations and which characteristics also will change due to aging. This will soon make the off-line calibration data invalid.

In Figure 4.1, measurements of the temperature drift in amplitude and phase for the hardware weighting unit used in an ABF testbed described in Section 2.2 and in the papers [40],[41] is presented. The measurements were performed over a period of ten hours after a cold start-up. We can see that the drift is approximately 0.1 dB and 1° per hour of operation and it demonstrates the need for frequent re-calibrations. It is therefore desirable if the calibration can be performed simultaneous with normal antenna operation.

This chapter presents two on-line calibration algorithms, that are transparent to normal antenna operation and have low complexity. By transparent, we mean that the algorithm runs in parallel to the beamforming algorithm and does not degrade the beamforming capabilities. Thus, if we assume a steady-state condition, the performance of the beamformer algorithm and the beamformer with on-line calibration performance are indistinguishable. The first algorithm is derived from the least mean square (LMS) algorithm and is a non-parametric and direct solution. It uses the difference between expected and actual beamformer output to create an error signal to be used in the LMS algorithm. No estimation of the drift in the hardware is performed. In the second algorithm, we have parametrized the transfer function. The parameters are then estimated and are used to update the calibration look-up table. This is called the indirect approach.

To study the performance of the proposed algorithms, simulations were performed. We assumed a simple signal environment with two signal sources (mobiles) and no multipath propagation. The time variations in the hardware were simulated by using a parameter drift similar to the measured one, cf. Figure 4.1. Both algorithms succeed to maintain the output carrier to interference ratio (CIR) at the same level as with perfect knowledge of the transfer functions. Thus, both algorithms are transparent. The direct approach, using the LMS algorithm requires a slowly time varying or time-invariant signal environment, while the indirect algorithm actually benefits from a rapidly varying signal environment, which is characteristic for fading in a multipath scenario.

The chapter is organized as follows. In Section 4.2, the model of the hardware in the adaptive antenna is introduced and the sample matrix inversion (SMI) algorithm is described. In Section 4.3, the proposed auto-calibration algorithms are described and in Section 4.4 the simulation results are presented. Finally the results are summarized in Section 4.5.

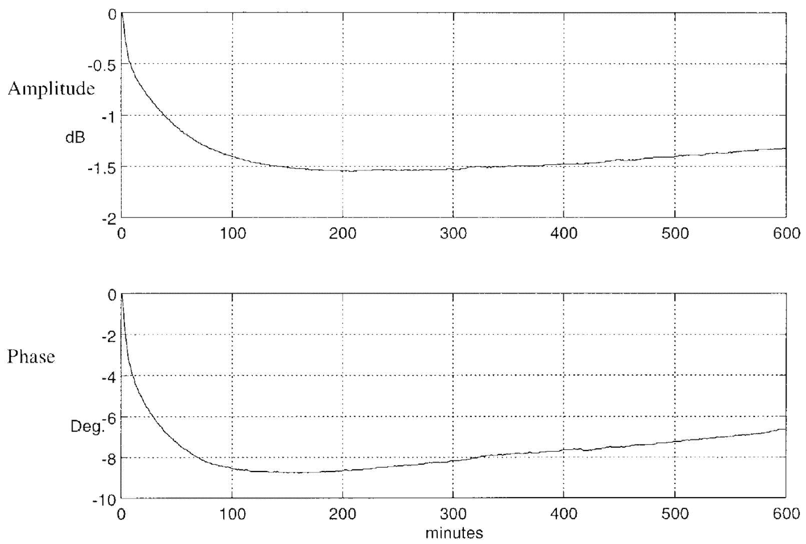


Figure 4.1: Temperature drift of the amplitude and phase of a hardware weight

4.2 Problem formulation

A block diagram of the adaptive antenna is shown in Figure 4.2. The signals arriving at the N antennas are described by the $N \times 1$ column vector $\mathbf{x}_a(t)$. The noise generation in the front end is modeled as a zero mean white Gaussian noise vector, $\mathbf{n}_1(t)$, with covariance matrix $\sigma_1^2\mathbf{I}$. The noise in the down-converting and sampling receivers, $\mathbf{n}_2(t)$, is also modeled as a zero mean white Gaussian noise vector, but with covariance matrix $\sigma_2^2\mathbf{I}$. The signals from the antennas are split into a digital path and an analog path. The analog path consists of cables and the hardware weighting units. The digital part consists of digital signal processing for weight computation and calibration.

Assume that all weights are set to unity, the transfer function measured between the input sampling receiver at point A in Figure 4.2 and the beamformer weighting units, point B, is described by the complex diagonal matrix $\mathbf{C} = \mathrm{diag}\,[c_1, \ldots, c_N]$. This assumes that the signal is narrowband in comparison to the bandwidth of the channel between A and B. Thus, the transfer function can be represented by a complex number. The signals

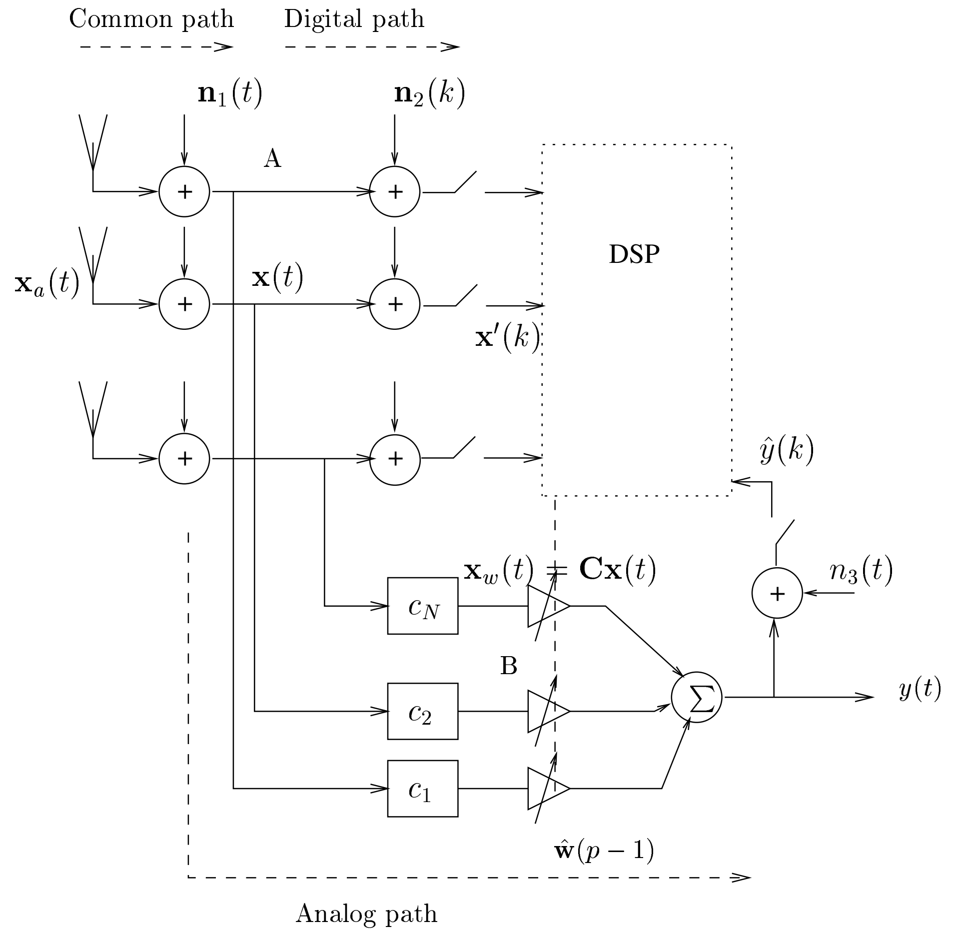


Figure 4.2: Block diagram of the adaptive antenna

at the beamformer is denoted $\mathbf{x}_w(t)$ and is defined as

$$\mathbf{x}_w(t) = \mathbf{C}\mathbf{x}(t) \quad . \tag{4.1}$$

Each factor c_l describes the attenuation and the phase shift of the analog path l relative to the corresponding digital path, except for the noise $\mathbf{n}_2(t)$.

The factors c_l will in generally be nonlinear functions of the calculated weights, $\mathbf{w}$, due to coupling between the phase shifters and the attenuators. The matrix $\mathbf{C}$ will thus in general be a nonlinear function of $\mathbf{w}$, i.e. $\mathbf{C} = \mathbf{C}(\mathbf{w})$and a look-up table for $\mathbf{C}$ must include values for all possible weight vectors $\mathbf{w}$. Furthermore, the elements in $\mathbf{C}$ are assumed to be constant over the receiver passband, but they will change in time during operation due to temperature drift in the weighting units, as seen experimentally in Figure 4.1.

In the on-line calibration algorithms, we use the beamformer output signal to generate an error signal. Thus, the output from the beamformer is down-converted to the baseband and sampled giving

$$\hat{y}(k) = y(t_k) + n_3(t_k) \quad . \tag{4.2}$$

The notation t_k will be used in this chapter to denote the continuous time signal sampled at the time instant $t = t_k$. The term $n_3(t)$ is the noise added by the feedback receiver. We will assume that $n_3(t)$ is zero mean complex Gaussian with variance σ_3^2. The signal $\hat{y}(k)$ will be used in the calibration algorithms, to be presented below.

4.2.1 The SMI-algorithm

We assume that the Sample Matrix Inversion (SMI)-algorithm is used to calculate the adaptive antenna weights. A reference signal, $s(k)$, present in each transmitted burst, e.g. the training sequence in a Time Division Multiple Access (TDMA)-systems like GSM and IS-136, is utilized in order to form a least-squares problem. The sampled baseband signals of the antenna elements form a column vector denoted $\mathbf{x}'(k)$. The mean square optimal weight vector is obtained by solving the Wiener-Hopf equations [12].

$$\hat{\mathbf{w}}_0 = \hat{\mathbf{R}}_{\mathbf{xx}}^{-1} \hat{\mathbf{r}}_{\mathbf{x}s} \tag{4.3}$$

where the sample-mean covariance matrix and cross-correlation vector are estimated by using M samples as

$$\begin{aligned} \hat{\mathbf{R}}_{\mathbf{xx}} &= \frac{1}{M} \sum_{k=1}^{M} \mathbf{x}'(k) \mathbf{x}'^H(k) \\ \hat{\mathbf{r}}_{\mathbf{x}s} &= \frac{1}{M} \sum_{k=1}^{M} \mathbf{x}'(k) s^H(k) \end{aligned} \tag{4.4}$$

Here, $(\cdot)^H$ denotes complex conjugate transpose. The SMI-algorithm tries to combine the N antenna signals in an optimum way, given the structure. There is no need to estimate the direction of arrival for the impinging signals, so a calibration of the array manifold (the common path) is not necessary.

4.2.2 Calibration

Compensation of phase and magnitude differences between the digital path and analog path must be performed for each antenna element, since the adaptive antenna weights are calculated based on the sampled signals of the digital paths, $\mathbf{x}'(k)$, whereas the beamforming is performed on the RF-signals of the analog paths $\mathbf{x}_w(t)$, cf. Figure 4.2.

Prior to operation, the matrix $\mathbf{C}$ is computed by performing a calibration. This calibration step is called an off-line calibration and the corresponding matrix is denoted $\mathbf{C}_0$. A continuous wave (CW) signal at the carrier frequency is injected at one antenna element at a time by directional couplers. The attenuation and the relative phase of that specific analog path compared to the digital path can be measured by comparing the beamformer output signal $\hat{y}(k)$ and the corresponding element in the vector $\mathbf{x}'(k)$. The calibration process is then repeated for all N antenna branches.

The calibration has to be repeated for all possible weight vectors also. The reason is that if the magnitude of a weight is adjusted, then the adjusted phase will also change, and vice versa, due to non-ideal isolation between the phase shifter and the attenuator in the hardware weights. Thus, the measurement of the matrix $\mathbf{C}$ has to be performed for every possible weight vector $\mathbf{w}$. Since full access to the control of the weighting units is required, the off-line calibration can only be performed prior to operation of the adaptive antenna.

During operation, temperature drift in the receivers and weighting units and other possible changes will degrade the performance of the adaptive antenna. The grade of degradation for a particular calibration phase and magnitude error was investigated in Chapter 3. It is therefore desirable to perform the calibration simultaneously with the main (beamforming) operation. This calibration must be of relatively low complexity and yet accurate enough to maintain a good performance of the adaptive antenna. Furthermore, the calibration should not interfere with the main operation of the antenna.

4.3 The auto-calibrating algorithms

In this section, the two proposed on-line algorithms are presented. They are called the direct and the indirect approach respectively.

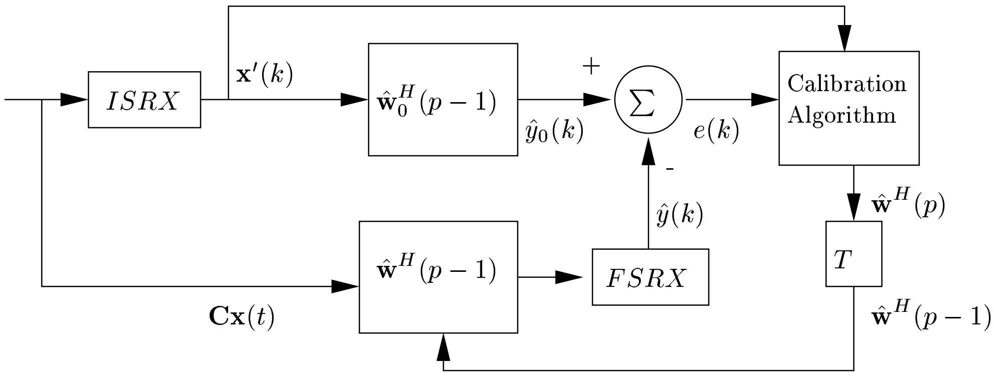


Figure 4.3: Block diagram of the LMS-approach. $FSRX$=Feedback Sampling Receiver, $ISRX$=Input Sampling Receiver. T represents one frame delay

4.3.1 The direct approach

This section outlines the auto-calibration algorithm using a direct approach and presents the assumptions made concerning the off-line calibration and the thermal drift of the involved hardware. The main idea of the direct approach is to create a reference signal in the DSP, by using the calculated beamformer weight vector and the received signal. An error signal is then created in the DSP by subtracting the reference signal from the sampled and down-converted beamformer output signal. This error signal and the received signal are then used to calculate a weight update to minimize the error. The update is performed by using the LMS algorithm.

The difficulties in this approach lies in the fact that the received signal in the DSP is not exactly the same as the signal at the weighting units. Also the error signal will be noisy due to the noise in the feedback receiver. The success of the algorithm lies in the robustness of the LMS algorithm. A drawback of this method is the delay introduced from calculating the beamformer weights to the actual use of the weights in the beamformer. First, the weights are calculated to create a reference signal in the DSP, then the weights have to be adjusted to minimize the error. Because of the analog weighting technique adopted, we cannot store data and apply the weights to the same data used to calculate the weights. Thus, the delay will be two TDMA frames and this will create a problem if the signal environment is changing rapidly, i.e. the weight vector $\mathbf{w}_0$ is not changing "slowly".

A signal-flow graph of the proposed algorithm is shown in Figure 4.3. The signals at the weighting units, $\mathbf{x}_w(t)$, are multiplied with the weight-vector $\hat{\mathbf{w}}^H(p-1)$, based on the calculations in the previous frame $p-1$, where p is the frame index, i.e.

$$y(t) = \hat{\mathbf{w}}^H(p-1)\mathbf{x}_w(t) \quad . \tag{4.5}$$

This is necessary since otherwise the weights have to be calculated and steered out during a time period that is much shorter than a frame. The analog beamformer makes it impossible to store the received block of data and apply the weights to the same data block that was used for calculating the beamformer weights. This also implies that the time step in the algorithm is one TDMA frame.

The sampled beamformer output signal, $\hat{y}(k)$, is compared with the corresponding beamformer output signal in the DSP, $y_0(k)$, and the error signal $e(k) = y_0(k) - \hat{y}(k)$ is formed. The signal $y_0(k)$ is calculated by using the SMI-weights from the previous frame as

$$y_0(k) = \hat{\mathbf{w}}_0^H(p-1)\mathbf{x}'(k). \tag{4.6}$$

Note that the signal $y_0(k)$ exists internally in the DSP only. Only one sample k is used in the calculation of $y_0(k)$ and it can be taken arbitrarily from the whole frame p. Alternatively, more samples can be used to improve the estimate of the gradient in the LMS algorithm.

For calculation of the necessary adjustments an LMS-like (least mean square) algorithm is proposed. The aim of the algorithm is to minimize the mean squared magnitude of the error signal $e(k)$.

We assume that the signal vector $\mathbf{x}_w(t_k)$ relates to $\mathbf{x}'(k)$ according to

$$\mathbf{x}_w(t_k) = \mathbf{C}\mathbf{x}(t_k) = \mathbf{C}\left\{\mathbf{x}'(k) - \mathbf{n}_2(k)\right\} \tag{4.7}$$

where $\mathbf{C}$ is a complex valued diagonal matrix. The relationship between $\mathbf{C}\mathbf{x}(t_k)$ and $\mathbf{x}'(k)$ in (4.7) is an approximation, since $\mathbf{C} = \mathbf{C}(\mathbf{w})$ and $\mathbf{w}$ in turn depends on $\mathbf{x}'(k)$. We thus assume in the calibration algorithm that the matrix $\mathbf{C}$ is independent of the weights, to simplify the algorithm and reduce the memory consumption in the DSP. In the LMS-approach, we will even assume that $\mathbf{C}$ in (4.7) is constant to obtain a relation between $\mathbf{x}_w(t_k)$ and $\mathbf{x}'(k)$.

To adjust the weights to compensate for the temperature drift, the well known LMS-approach would use $\mathbf{x}_w(t_k)$ and $e(k)$ to update the weight vector as:

Desired algorithm: (p≥1)

$$\begin{aligned}
y_0(k) &= \hat{\mathbf{w}}_0^H(p-1)\mathbf{x}'(k) \\
e(k) &= y_0(k) - y(t_k) \\
\hat{\mathbf{w}}(p) &= \hat{\mathbf{w}}(p-1) + \mu\mathbf{x}_w(t_k)e^H(k)
\end{aligned} \tag{4.8}$$

Where the initialization of the algorithm has been omitted.

The constant μ is the step-size parameter in the algorithm and $y(t)$ is the output of the analog beamformer. The algorithm (4.8) cannot be used since $\mathbf{x}_w(t_k)$ is not measurable and only a noisy estimate of $y(t_k)$ is available. The approach used here is therefore to use $\mathbf{x}'(k)$, i.e. the noisy measurement of $\mathbf{x}(k)$, and the matrix $\mathbf{C}_0$ obtained from the off-line calibration process to estimate $\mathbf{x}_w(t_k)$, using (4.7). The proposed algorithm can then be stated as follows:

Proposed algorithm:
Initialization: (p=0)

$$\hat{\mathbf{w}}(0) = (\mathbf{C}_0^{-1})^H\hat{\mathbf{w}}_0(0) \tag{4.9}$$

Algorithm: (p≥1)

$$\begin{aligned}
y_0(k) &= \hat{\mathbf{w}}_0^H(p-1)\mathbf{x}'(k) \\
\hat{e}(k) &= y_0(k) - \hat{y}(k) \\
\hat{\mathbf{x}}_w(k) &= \mathbf{C}_0\mathbf{x}'(k) \\
\hat{\mathbf{w}}(p) &= \hat{\mathbf{w}}(p-1) + \mu\hat{\mathbf{x}}_w(k)\hat{e}^H(k)
\end{aligned} \tag{4.10}$$

The fact that the calibration data from the off-line calibration is used will only slightly affect the performance of the gradient method since the LMS-approximation of the gradient is in itself very crude. The correct instantaneous gradient direction is given by $\mathbf{x}_w(t_k)e^H(k)$ but here the approximation $\hat{\mathbf{x}}_w(k)\hat{e}^H(k) = \mathbf{C}_0\mathbf{x}'(k)\hat{e}^H(k)$ is used instead. An exact analysis of the influence of the properties of this error in gradient estimate on the convergence of the algorithm remains to be performed. Simulation studies presented in Section 4.4 show that the performance of the proposed algorithm will be satisfactory in a relatively stationary signal environment, i.e. with a slowly varying SMI-weight vector. When the signal environment is non stationary, the long convergence time inherent in LMS-like algorithms can be expected to create problems. The error term $\hat{e}(k)$ will be large if the SMI-weight vector $\mathbf{w}_0$ is not changing slowly. The performance will then be degraded.

4.3.2 The indirect approach

The second proposal to an on-line calibration algorithm is to estimate the transfer functions c_l, $l = 1, \dots, N$ explicitly at each time step (frame) p. The estimated transfer functions, c_l, are then used to adjust the weights in the beamformer. Thus, the calibration look-up table is updated at each time step p. In this approach the delay in the weight update is only one frame as opposed to the LMS method. This method does not require "slow" changes in the weight vector $\mathbf{w}$ as was the case for the direct approach.

To derive the calibration algorithm we generate a model where the order of the weights and the transfer functions c_l has been reversed, see Figure 4.4. The weighted signals are denoted $\mathbf{z}(t_k)$. We now consider the sampled output signal from the beamformer, $y(t_k)$, as a linear combination of the signals $\mathbf{z}(t_k)$, after the weights, $\hat{\mathbf{w}}$, as in Figure 4.4. The estimated transfer functions at time step p, denoted $\hat{c}_l(p)$ are used to form the row-vector $\hat{\mathbf{c}}(p) = [\hat{c}_1(p)\hat{c}_2(p) \cdots \hat{c}_n(p)]$.

To find the coefficients c_l from $\mathbf{z}(t_k)$ and $y(t_k)$ is a simple regression problem. However, only the sampled received signal $\mathbf{x}'(k)$ and the sampled beamformer output signal, $\hat{y}(k)$, which is noisy can be measured and not $\mathbf{z}(t_k)$ and $y(t_k)$ which would have been desired. Thus, we resort to using the weights $\hat{\mathbf{w}}(p-1)$ and the identified transfer function vector from the previous frame, $\hat{\mathbf{c}}(p-1)$, to estimate $\mathbf{z}(t_k)$. The estimate is denoted $\mathbf{z}(k)$ and element l in the column vector $\mathbf{z}(k)$ is defined as:

$$\begin{aligned} z_l(k) =& x_l'(k)\hat{w}_l^H(p-1) \\ =& x_l'(k)\frac{\hat{w}_{0,l}^H(p-1)}{\hat{c}_l(p-1)} \end{aligned} \tag{4.11}$$

where in the second step we used the relation (3.5) between the weight $\hat{w}_{0,l}^H(p-1)$, calculated in the DSP and the weight used in the beamformer. Thus, the DSP weight is divided by the estimated transfer function $\hat{c}_l(p-1)$.

The estimated regressor vector $\mathbf{z}(k)$ is now used in the estimation of $\mathbf{c}$. We assume that the system is linear and form a least-squares problem utilizing the sampled output $\hat{y}(k)$. The output $y(t_k)$ is a linear combination of the regressor vector $\mathbf{z}(k)$:

$$y(t_k) = \mathbf{c}(p)\mathbf{z}(t_k) \quad . \tag{4.12}$$

By multiplying with $\mathbf{z}^H(t_k)$ and applying the expectation operator, we can

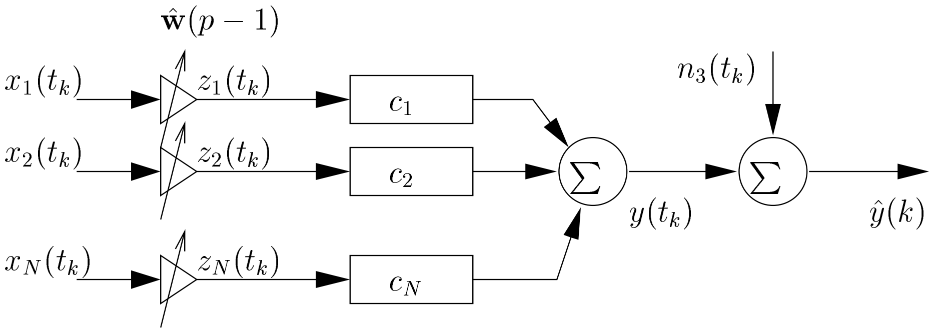


Figure 4.4: Regression model. Note that the order of the weights and the transfer functions c_l have been reversed. Using the inputs z_l and the output $y(t_k)$ we can find the least squares solution of the coefficients c_l.

write

$$\mathbf{r}_{y\mathbf{z}} = \mathbf{c}(p)\mathbf{R}_{\mathbf{zz}} \quad . \tag{4.13}$$

We can now calculate the least-squares estimate of the transfer functions $\hat{\mathbf{c}}(p)$ by using the estimated covariance and cross-correlation matrices

$$\begin{aligned}\hat{\mathbf{R}}_{\mathbf{zz}} &= \frac{1}{M}\sum_{k=1}^{M}\mathbf{z}(k)\mathbf{z}^H(k)\\ \hat{\mathbf{r}}_{\hat{y}\mathbf{z}} &= \frac{1}{M}\sum_{k=1}^{M}\hat{y}(k)\mathbf{z}^H(k)\end{aligned} \tag{4.14}$$

and calculate the diagonal elements of $\hat{\mathbf{C}}(p)$ as

$$\hat{\mathbf{c}}(p) = \hat{\mathbf{r}}_{\hat{y}\mathbf{z}}\hat{\mathbf{R}}_{\mathbf{zz}}^{-1} \tag{4.15}$$

The M samples used in (4.14) need not necessarily all be from the same frame p. The performance will be improved if samples from several frames are used, as discussed below.

The elements in $\hat{\mathbf{c}}(p)$ are low-pass filtered in order to introduce a memory in the algorithm. A first order filter with a pole at 0.9 is used. This will also reduce the variance of the regressor vector, because $\hat{\mathbf{c}}(p-1)$ appears in the calculation of the regressor, equation (4.11). Experiences from the simulations show that the low-pass filtering improves the ability to track $\mathbf{c}(p)$.

When using this indirect approach instead of the direct method described in Section 4.3.1, samples from the whole frame are used in the identification process, instead of the utilization of only one sample as in the LMS algorithm in the direct approach. Furthermore, mobile communication systems are interference limited and not noise limited. Thus the noise level in $\mathbf{x}(t)$ (and also in $\mathbf{z}(\mathrm{k})$) will be low compared to the signal levels. Therefore $\hat{\mathbf{R}}_{\mathbf{zz}}$ will be badly conditioned if the number of incoherent rays impinging on the array is less than N. The calculation of the inverse $\hat{\mathbf{R}}_{\mathbf{zz}}^{-1}$ will create problems when computing $\hat{\mathbf{c}}(p)$ in (4.15) and the tracking of the temperature drift may not succeed. This can however be mitigated by using a few number of samples from several frames to estimate the covariance matrices. This will in general make $\hat{\mathbf{R}}_{\mathbf{zz}}$ better conditioned. The idea is that the fading will cause the signal components in $\mathbf{z}(k)$ to vary over the frames, creating a sequence that is persistently exciting with a well conditioned covariance matrix $\hat{\mathbf{R}}_{\mathbf{zz}}$. The condition number can also be improved by diagonal loading.

This approach will thus benefit from a rapidly time-varying signal environment and as opposed to the direct approach presented in Section 4.3.1, which requires a stationary or slowly time-varying signal environment.

4.4 Simulation study

The aim of the simulation is to study the performance of the two autocalibration algorithms presented above. A simple signal environment with two mobiles transmitting binary phase shift keying (BPSK) modulated signals of equal power but with different training sequences was used. The two signals were transmitting on the same frequency and the wavelength was λ. The desired and the interfering signals are assumed to impinge on an eight element uniform linear array (ULA) from distinct directions θ_d and θ_i respectively. No multipath propagation or fading is assumed. The antenna array inter-element spacing was $\lambda/2$ and the number of antennas N=8. The length of the training sequence was 26 symbols to comply with the GSM/DCS-1800 standard.

The received signal, cf. Figure 4.3, can be expressed as

$$\mathbf{x}_a(t) = \mathbf{a}(\theta_d)s_d(t) + \mathbf{a}(\theta_i)s_i(t) \tag{4.16}$$

where $\mathbf{a}(\theta_d)$ and $\mathbf{a}(\theta_i)$ are the array response vectors in the direction θ_d and θ_i, while $s_d(t)$ and $s_i(t)$ are the desired and interfering signals respectively.

4.4.1 Generating the temperature drift

In the simulations the temperature drift in magnitude and phase was generated as independent integrated random-walk processes to obtain a "smooth" drift. The diagonal elements of the diagonal matrix $\mathbf{C}$ can be written as

$$c_l = A_l e^{j\varphi_l} \qquad l = 1, \ldots, N \quad . \tag{4.17}$$

Where A_l and φ_l are generated as integrated random walk processes:

$$\begin{aligned} A_l(p+1) &= A_l(p) + \frac{1}{1-q^{-1}} v_{A,l}(p) \\ \varphi_l(p+1) &= \varphi_l(p) + \frac{1}{1-q^{-1}} v_{\varphi,l}(p) \end{aligned} \tag{4.18}$$

with $v_{A,l}$(p) and $v_{\varphi,l}$(p) being mutually uncorrelated white noise sequences of appropriate variance. The variables A_l and φ_l are initialized by the values from the off-line calibration. This gives a temperature drift with statistical properties similar to the measurements presented in Figure 4.1.

4.4.2 Regularization of the covariance matrix

To stabilize the beamformer weights, we regularize the covariance matrix, by the method of diagonal loading. This is crucial to the operation of the direct approach as discussed below.

The conditional number (or eigenvalue spread) of the covariance matrix $\hat{\mathbf{R}}_{\mathbf{xx}}$ in (4.4) is reduced by employing regularization to obtain a condition number approximately equal to 10^2. More details about this regularization method can be found in [24]. The regularized covariance matrix can be inverted without numerical problems, and the weight vector calculated by the SMI-algorithm will be approximately the same from frame to frame, under steady state signal conditions. See also Section 2.5.2 for measurements on the effect of diagonal loading. Figure 2.11 shows how the weight vector is stabilized when diagonal loading is employed.

As the delay from calculation of weights to setting of weights is one frame, it is required that the signal environment and thus the beamformer weights are slowly changing. Furthermore, the proposed on-line LMS algorithm in the direct approach require that the weight vector $\mathbf{w}$ changes "slowly". So the regularization of the covariance matrix is crucial to make the direct approach work.

4.4.3 The simulation

Since the LMS-algorithm proposed in Section 4.3.1 is a closed loop algorithm, it is not possible to update the weights more frequently than once per frame; the beamformer output using the old weights is used to calculate the new weights. This means that the time step of the algorithm is one TDMA frame (or 4.615 ms in GSM/DCS-1800) and this is the time step used in the simulations for the update of the weights.

The noise variances of the integrated random walk processes (4.18), that model the temperature drift of the magnitude and the phase, sets the time scale of the simulation. In the simulations, we increased the noise variance to reduce the simulation time. Thus, the drift is more severe than in the measurements presented in Figure 4.1. It is therefore not necessary to simulate more than 500 frames[1].

4.4.4 Results

To measure the performance of the adaptive antenna, the CIR on the beamformer output was estimated as

$$\widehat{CIR} = \frac{E(|\hat{\mathbf{w}}^H \mathbf{a}(\theta_d) s_d(t)|^2)}{E(|\hat{\mathbf{w}}^H \mathbf{a}(\theta_i) s_i(t)|^2)} = \frac{|\hat{\mathbf{w}}^H \mathbf{a}(\theta_d)|^2}{|\hat{\mathbf{w}}^H \mathbf{a}(\theta_i)|^2} \tag{4.19}$$

where the weight vector is constant. Here we have assumed that the noise levels are well below the signal and interference levels.

The direct approach

The beamformer output CIR utilizing the auto-calibration algorithm presented in Section 4.3.1 is plotted as a function of frame number in Figure 4.5. In this particular simulation the angles of the two mobiles have been chosen to $\theta_d = 15°$ and $\theta_i = 43°$ relative to the broadside direction of the ULA. The step size μ used in the LMS-like algorithm was 0.005[2]. The variances of the noise sources σ_1^2, σ_2^2 and σ_3^2 were chosen to 20 dB, 40 dB and 30 dB below the signal level, respectively. The CIR of the adaptive array only utilizing the off-line calibration method is also plotted in Figure 4.5 for comparison. It is evident that the auto-calibration algorithm is able to

[1] This corresponds to 2.3 seconds in GSM/DCS-1800

[2] The step-size, μ, was found by trial and error, until a good performance was obtained with sufficiently fast convergence. No optimality is claimed in this choice.

maintain the CIR on a high level, whereas the performance of the traditional SMI adaptive array is severely degraded as the drift in the magnitude and phase of the weights is introduced. The variations of the CIR is also lower for the direct method compared to the indirect method. This is due to the recursion involved in the algorithm that smoothes the CIR.

The direct approach

Figure 4.6 shows the beamformer output CIR when using the indirect approach presented in Section 4.3.2, compared to the SMI algorithm with the off-line calibration data only. All simulation parameters are identical to those in the direct approach simulations. Also in this case the CIR of the auto-calibration algorithm is maintained on a high level. However, the variance of the beamformer output CIR is higher in this case compared to the direct approach, and more similar to the SMI without drift compensation. This can be explained by the batch oriented approach in the indirect method as opposed to the direct method where the feedback signal smoothes the weight adjustments and the variation of CIR is slower.

The performance of the indirect approach can also be studied by observing the tracking of the temperature drift. The generated temperature drift for the fifth antenna path is presented in Figure 4.7 together with the corresponding estimated phase and magnitude of the channel transfer function. It can be seen that the tracking of both magnitude and phase is successful. The large variance of the CIR in Figure 4.6 can be explained by the variations in the channel estimates, cf. Figure 4.7. If the N=8 channel estimates, similar to the one presented in Figure 4.7, have similar variance, we can explain the large variance in the beamformer output CIR. The combined effect from N=8 random variables, yields an output CIR with large variance.

4.5 Conclusions

Two algorithms are proposed for mitigating the temperature drift in adaptive antenna arrays using hardware weights. The two methods utilize a feedback signal from the beamformer output. The first method used the initially calculated SMI-weights to create a reference signal in the DSP. An LMS-like algorithm then adjusts the hardware weights to make the feedback signal follow the reference. The second method, which is an indirect method, attempts to estimate and follow the drift in the transfer functions, and use this tracking information to adjust the SMI-weights. Simulations shows that the output CIR is unaffected for both methods when a realistic

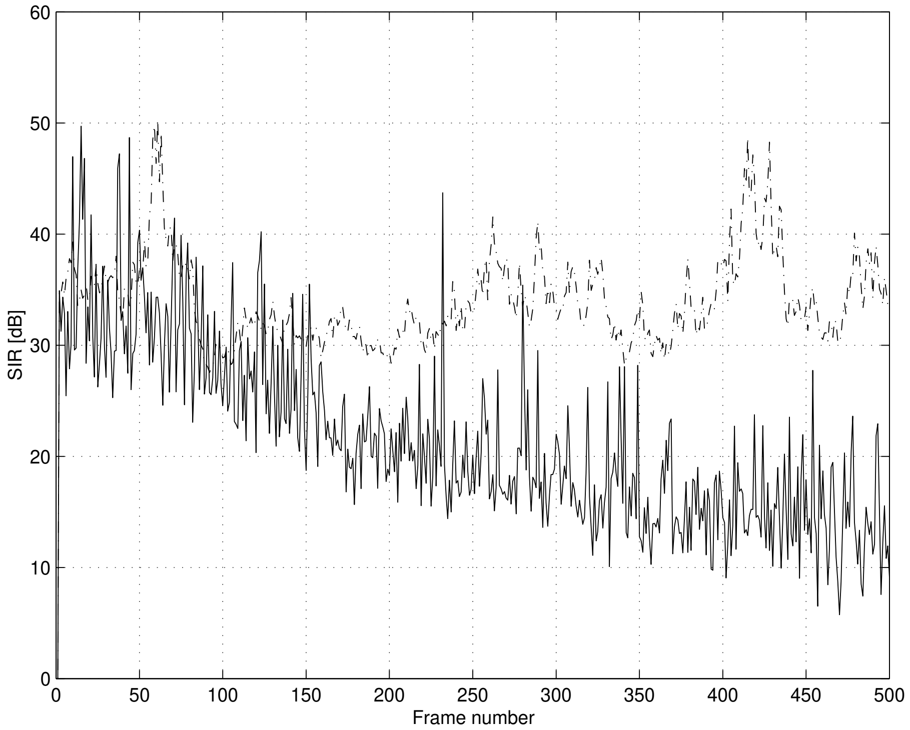


Figure 4.5: Beamformer output CIR (SIR) as a function of the frame number. Dash-dotted: auto-calibration using LMS, solid: SMI with off-line calibration data only

temperature drift generated as an integrated random walk process is introduced. The SMI-algorithm without compensation for the temperature drift however suffers from a considerable performance degradation, in terms of beamformer output CIR.

The variance of the beamformer output CIR of the direct approach is lower due to the smoothing effect introduced by the recursive LMS-algorithm. One drawback of the direct approach is that it requires that the beamformer weight vector is approximately equal during two frames. Thus, the signal environment most be slowly varying, otherwise the algorithm will lose track of the weight-vector $\mathbf{w}$. To reduce the effect of a particular noise realization on the calculated weight vector, we adopt a regularization of the covariance matrix by diagonal loading.

The indirect approach does not demand a slowly varying or quasi-static

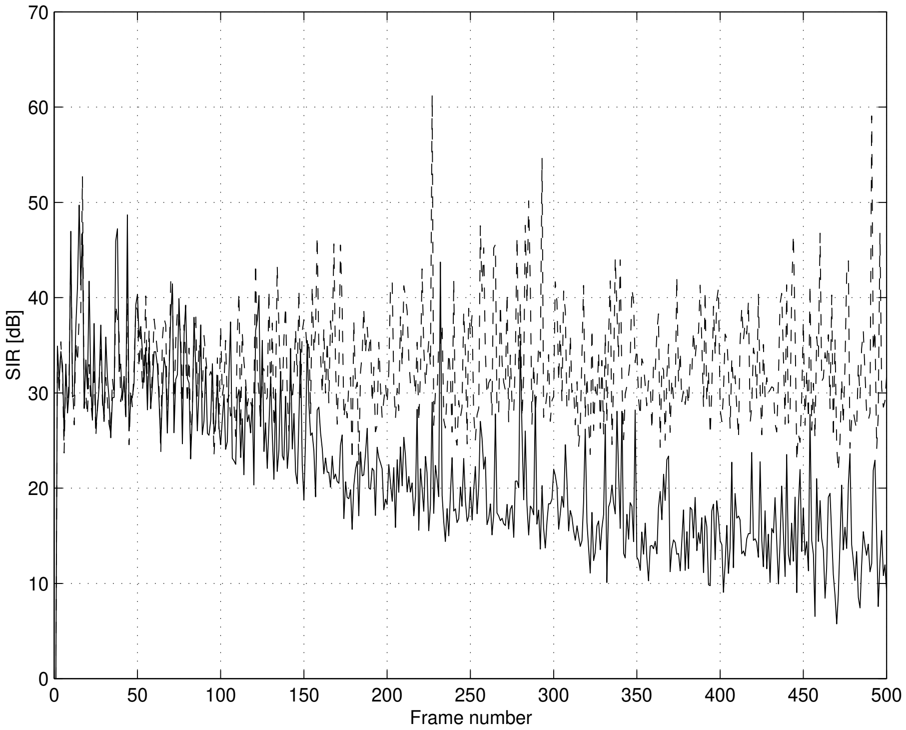


Figure 4.6: Beamformer output CIR (SIR) as a function of the frame number. Dashed: auto-calibration using the direct approach, solid: SMI with off-line calibration data only

signal environment as the direct approach. It actually benefits from rapid time variations as in a multipath environment. Such variations in the signal environment will tend to make the input signal to the algorithm persistently exciting. As long as the problem of a persistently exciting input signal can be handled, it will therefore be the more attractive approach. It is however more computationally complex and give a CIR on the beamformer output which has larger variance compared to the direct approach. The large variance in CIR is attributed to the variance in the estimation of the N antenna channels which contribute to the total variance in the beamformer output CIR. The performance of the indirect approach could be improved by applying a smoothing filter on the tracked channel parameters.

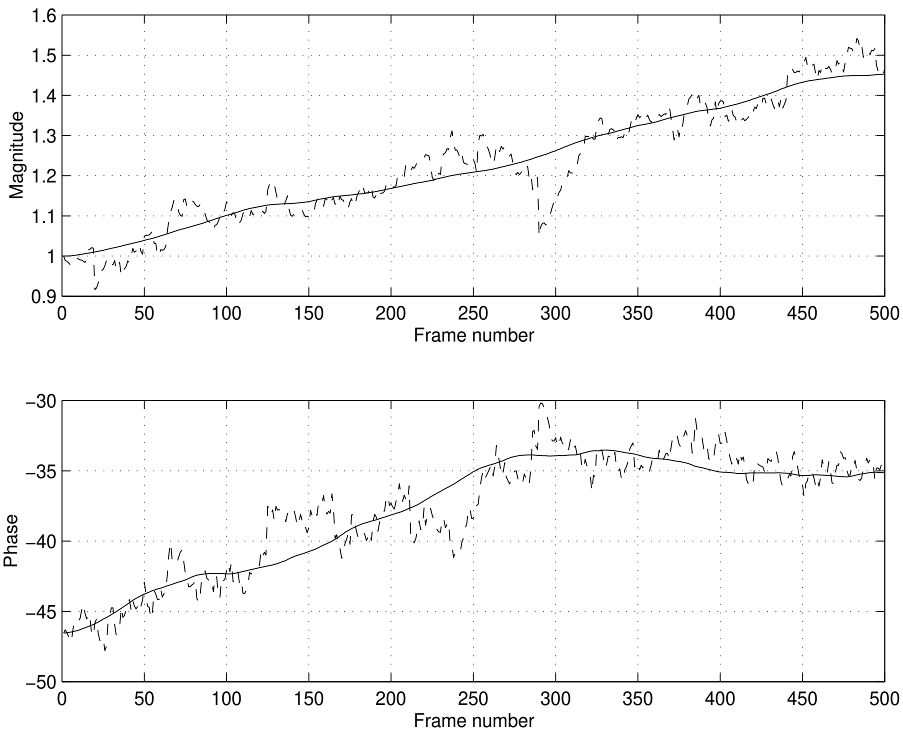


Figure 4.7: Tracking of variations in magnitude and phase for the fifth antenna path obtained from the direct method. Dashed: tracked drift, solid: true drift

Chapter 5

Transmit Amplifier Nonlinearities

5.1 Introduction

Introducing antenna arrays at the basestation site increases the complexity of the basestations by multiplying the number of radio modems, amplifiers and so on. Simply, if a basestation is operating on M frequency channels, it needs M radio modems and M amplifiers if the conventional single carrier amplifier (SCPA) technology is applied, as discussed in the downlink implementation options Section 1.2.4. By introducing an N element array antenna, the number of amplifiers increases to NM. See the DSP-SCPA configuration in Figure 1.8. Thus, it is highly desirable to integrate the hardware used at the base station site to reduce cost, power dissipation, and the space a basestation site occupies.

One proposed solution, applicable on the transmitting link is to combine signals prior to amplification and co-amplify the low power signals in a multicarrier amplifier (MCPA) (Figure 1.11). This eliminates the need for combining signals of high power using bulky cavity filters and the M SCPAs in a conventional basestation site is reduced to a single MCPA. By introducing MCPAs in the transmit channel of adaptive array antennas, the number of hardware units is reduced from NM to N, which is a considerable reduction if the number of used frequency channels M is large. This is denoted the DBF-MCPA or ABF-MCPA configuration in Chapter 1.

Another benefit of MCPAs is the possibility to move the combining of the M signals to the digital domain. This reduces the size of the basestation even further, but requires high dynamic range and wideband ADC. By moving more and more functionality into the digital domain, we approach the full software radio architecture (SWR), see Figure 1.12.

When several constant-envelope signals are combined in the MCPA, the envelope of the composite signal becomes non-constant. If the signals have different frequencies and a constant phase, the resulting combined signal becomes periodic. In digital communication systems, a phase modulation is often used, and the composite signal is the sum of several signals with independent phase information, and the resulting envelope will thus be non-constant and non-periodic. This implies that the amplifier cannot be operated in the high-efficient nonlinear region, unless some kind of linearisation is performed. The linearisation of MCPAs has been the subject of considerable research effort, using, for example, feed-forward [42] and pre-distortion [43] techniques. These results are promising, although still they are computationally expensive if high performance is to be achieved.

The combination of several independent signals give rise to a large peak-to-average ratio of the MCPA input signal, as indicated above. Even if it is possible to construct an MCPA which meets the stringent linearity requirements, its peak power requirement is much greater than the individual power of the M carriers. For example, a 16 channel basestation requiring 10 Watts per channel, would require a peak power in the MCPA as high as 2.56 kW [44]. The amplifier has to be biased with a large back-off from the saturating point to avoid that the peak power rating of the MCPA is exceeded. A large back-off implies a decreased power-added efficiency (PAE), so most of the DC power is lost as heat.

The input back-off (IBO) and the output back-off (OBO) of an amplifier is defined in dB as

$$IBO = 10 \log_{10} \left(\frac{P_{sat,in}}{P_{avg,in}} \right) \tag{5.1}$$

$$OBO = 10 \log_{10} \left(\frac{P_{sat,out}}{P_{avg,out}} \right) \tag{5.2}$$

where $P_{avg,in}, P_{avg,out}$ is the average input/output power and $P_{sat,in}$ is the input power when the output power is saturated and $P_{sat,out}$ is the corresponding output power. At maximal PAE the OBO and IBO is thus 0 dB. However, with decreasing back-off, the amount of generated intermodulation products increase and the link quality decreases. Thus, there is a trade-off between amplifier efficiency and link quality. In this chapter, we find a reasonable choice of the back-off, by defining the term total degradation, which is a function of the back-off and the carrier to interference degradation of a mobile user as compared to ideal amplifiers (linear at all input levels) in the basestation.

5.1.1 Background

Linearity considerations for adaptive antennas in mobile communications have previously been studied by Beach *et.al.* [30], where the intermodulation distortion (IMD) created in the nonlinear amplifier was shown to degrade the null depth and increase the side-lobe levels of the antenna radiation pattern. Another study by Litva and Lo [11], investigates the receive radiation patterns from a digital beamforming (DBF) array under nonlinear distortion. Fixed and adaptive beamforming was compared and it was shown that adaptive beamforming in the uplink can reduce the effect of nonlinearities, by suppressing the *phantom interferers* created by the nonlinearity. Recent research shows that it is possible to limit the peak-to-average ratio of the input signal (in software or hardware) without introducing too much intermodulation distortion that violates the overall system specifications. The signal is simply limited in amplitude prior to amplification and a reduction in peak power rating of the MCPA [44] is thereby allowed. The level of the amplitude limiter for a specific bit error rate is dependent on peak to mean ratio of the used modulation scheme [45].

Modeling of nonlinearities

Nonlinearities in communication systems are commonly characterized by the amplitude to amplitude modulation (AM/AM) distortion characteristics and amplitude to phase modulation (AM/PM) distortion characteristics, which is the input amplitude dependent gain and phase conversion of the amplifier.

When simulating multicarrier communication systems operating over nonlinear channels, the required sampling rate contributes significantly to long execution times. When the sub-carriers have constant envelopes and the number of carriers is large, the intermodulation distortion is often modeled by an additive Gaussian noise source [46]. In smart antennas, the signals that enter the MCPA has a time varying envelope, due to weight adaption and this will make the signals correlated with each other and with the intermodulation products. To simulate this, the sampling rate must be extremely high and the executing time becomes very long.

Schneider et.al. [46] presented a method to simulate a multicarrier digital communication system over a band limited channel with a nonlinear device. The method is based on Shimbo's work [47] and reduces the simulation time significantly by only considering the intermodulation products that falls within the bandwidth of the observed sub-carriers. The simulation method is applied in this chapter to evaluate the impact of nonlinearities on adaptive antenna system performance.

Effects of nonlinearities in array antenna systems

By introducing MCPAs in smart antenna systems and assuming that they are not perfectly linearized, we observe three effects:

1. Intermodulation distortion power will be emitted from the array antenna. As opposed to a conventional basestation with one antenna, the array makes the radiated intermodulation distortion to be dependent on the azimuthal angle. Thus in some direction the intermodulation products from all N antennas adds coherently and a mobile on that frequency and in that particular direction will experience an increased interference level. This has previously been studied by Loyka [48] for receiving arrays, where the concept *phantom-lobes* were introduced. In this chapter, we investigate the increased interference levels impact on the down-link link budget in a cellular system.

2. The AM/AM and AM/PM conversion in the MCPA will make the transmitting antenna branches unmatched, due to gain compression and phase shifts. An intended "null" in the radiation pattern towards a known co-channel user will be shifted, and the interference for that particular user will increase. For a two element array and a phase error of 1° and an amplitude error of 0.1 dB, the maximum null depth is limited to 30 dB for interference cancellation [26]. For example, the phase and amplitude shift through the amplifier feeding antenna 1 is not the same as through antenna amplifier 2, due to different input power levels, which in turn depends on the respective weight amplitude. Using MCPAs the situation gets more complicated, since the phase and amplitude shifts are then a function of every signal power entering the MCPA, due to the interaction between the input signals. Thus the superposition property of the linearly amplified signals is violated when the MCPA is nonlinear.

3. Due to the large signal envelope variations of the combined signal on the input of the MCPA, the peak-to-average ratio of the input signal is often limited prior to amplification, so the peak power rating of the amplifier can be reduced. The limiting is often performed in

software, prior to D/A conversion. This leads to considerable benefits in terms of power and cost savings in the design of the MCPA, on the other hand, the limiter will also generate intermodulation distortion that, unfortunately, will decrease the signal to interference ratio for the mobiles.

Although excellent linearisation techniques have been proposed, there is a trade-off between linearity and cost, and if many basestations are to be equipped with array antennas, it is important that the employed linearisation technique is inexpensive. Thus, a certain degree of nonlinearity must be tolerated.

In this chapter, we investigate the degradation or improvement in CIR for a mobile in a cellular system, when the back-off of the transmitting MCPA is decreased, the number of antennas is adjusted or if the transmitter amplifier is replaced by an amplifier with better linearity properties. The transmitting link is studied, because it handles higher power than the low noise amplifier in the receiving front end. It is essential to increase the transmit MCPAs efficiency to minimize the total base-station power consumption. We would like to know what parameters that affects the CIR most and by how much. It is also important to choose the amplifier back-off appropriately to maximize the amplifier efficiency whilst maintaining the average CIR (or link budget) at the mobiles at an acceptable level. The link budget specification can, for example, consist of the requirement that the mobile user CIR is larger than 5 dB, 90% of the time. We will thus, in this chapter estimate the cumulative distribution function (CDF) of the mobile CIR by means of simulations. From this, we can calculate the total degradation function and get the best choice of back-off for a certain amplifier and for a specific number of antennas. The downlink is assumed to be implemented as the DBF-MCPA and SWR configurations as defined in Section 1.2.4. We also assume that the adaptive antenna arrays are utilized as spatial filters for interference reduction (SFIR), cf. Section 1.2.2. Thus, the base-station transmits only to one mobile per frequency channel at a particular time instant. Of course, if a TDMA system is used, the frequency channel is further divided in time into timeslots.

Another issue studied in this chapter is the beam direction of the intermodulation distortion when using a transmitting array antenna of the ABF-MCPA type. We show that intermodulation distortion adds up in some direction, dependent on the current adaptive weights. This is also illustrated by measurements, performed in an anechoic chamber using a Butler matrix as a beamformer.

5.2 Basic considerations

We want to estimate the CDF of the CIR for the mobiles in a cellular system, when non-ideal MCPAs are used in the transmitting link. First, we define the basic parameters that characterize the downlink, and how the beamformer weights can be calculated. Then, the architecture of the DBF-MCPA is re-examined, and the necessary variables are introduced. We define the parameters in our representation of a multicarrier signal, to be used later in the calculations. Furthermore, to be able to determine the amplifier back-off, the average transmit power is required, and we discuss a way to estimate this.

5.2.1 Downlink transmission

The basic operation of an adaptive antenna array at a base station is to enhance the received signal by suppressing interferers transmitting on the same channel and optimally combine the signals to reduce multipath fading. The receiver estimates the radio channel and performs a spatio-temporal equalization. Due to the low correlation between the uplink and downlink channel in frequency division duplex (FDD) systems [49], downlink beamforming has to be based on other information than the estimated uplink channel. This study considers a FDD system, which is a very common duplex method (GSM, AMPS, IS54, IS136, IS95 are all FDD systems). If the angular spread of the signal is small, then the spatial signature of the channel will be approximately reciprocal. This information is used in calculating the downlink beamforming weights. One downlink algorithm is described in Appendix B.

The spatial signature is defined by denoting the direction of arrival (DOA) of user i, θ_i and its respective spreading angle σ_i as seen from the base station array. The spatial signature of user i is then the collection of all DOAs and spreading angles corresponding to user i. The receiver must estimate the spatial signature of all co-channel mobiles from the received signals. The received signal power Γ_i from user i is also estimated, where we assume that the delay spread is negligible. The received signal power has an inverse relationship to the square of the base to mobile distance, and also depends on the slow fading (shadowing). It is assumed that Γ_i is averaged over the fast Rayleigh-distributed fading.

The topology of the downlink transmission of the SWR scheme is shown in Figure 5.1. The signals to be transmitted are spatially multiplexed in the beamformer to give one weighted signal for each of the N antennas.

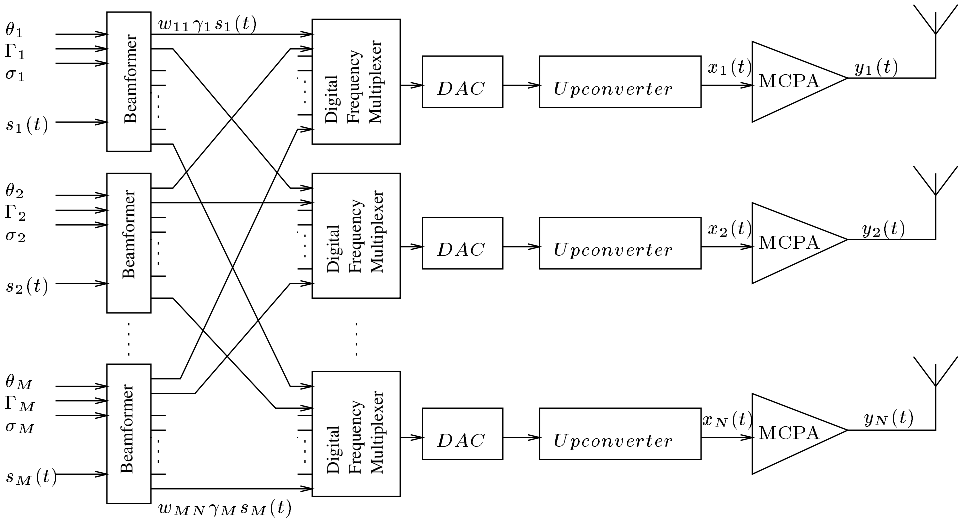


Figure 5.1: Downlink topology in the software radio adaptive antenna

The M signals to be transmitted by the same antenna are multiplexed in frequency, that is, signal i is digitally upconverted to frequency ω_i. Each antenna signal is then D/A converted and upconverted by mixing with the carrier frequency ω_c and further amplified in the MCPA.

The beamformers are using the spatial signatures of the received signals to calculate the weights for the downlink beamforming. If we assume that the multipath propagation is negligible, then there exist only one "principal" DOA per mobile user, the M complex baseband signals $\{s_1(t) \cdots s_M(t)\}$ should simply be transmitted in the directions $\{\theta_1 \cdots \theta_M\}$. The signals are D/A converted and up-converted to their respective RF frequencies $\{\omega_c + \omega_1 \cdots \omega_c + \omega_M\}$ where ω_c is the carrier frequency and ω_i is the offset from the carrier frequency of user i. We call these offset frequencies sub-carrier frequencies. It is assumed that $\omega_c \gg \omega_i$. Furthermore, a constant envelope modulation scheme is assumed. Setting the constant to unity, $|s_i(t)| = 1$, creates no loss of generality.

5.2.2 A multicarrier signal representation

Assume that the base-band signals $s_i(t)$, $i = 1, \ldots, M$ are to be transmitted at the RF frequencies $\omega_c + \omega_i$, $i = 1, \ldots, M$ respectively. Each signal is assumed to have a modulation bandwidth of W Hz. The M signals are first distributed over the baseband as in Figure 5.2. The complex baseband multicarrier signal is thus

$$s_{baseband}(t) = \sum_{i=1}^{M} s_i(t) e^{j\omega_i t} \tag{5.3}$$

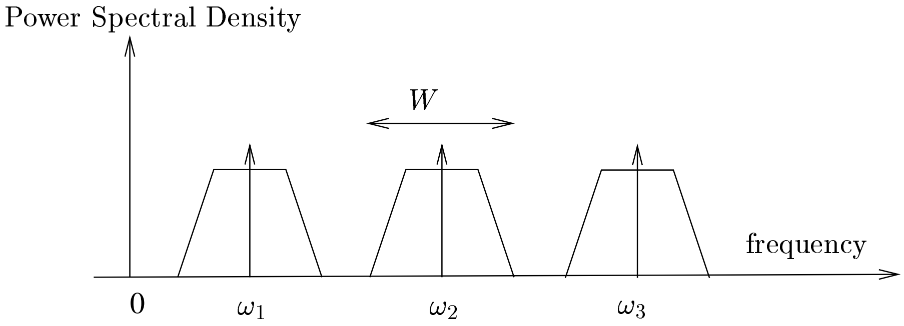


Figure 5.2: Placement of the signals $s_1(t), s_2(t)$ and $s_3(t)$ in the base-band at corresponding base-band (sub-carrier) frequencies ω_1, ω_2 and ω_3, prior to D/A conversion and up-conversion to the carrier frequency ω_c. Note that $\omega_c \gg \omega_i$ for $i = 1, 2, 3$ and W is the modulation bandwidth of each signal (sub-carrier).

This multicarrier signal is then D/A converted using a DAC with sufficient bandwidth (more than $2M$ times the modulation bandwidth of each sub-carrier), and upconverted to the RF frequency ω_c, using a mixer.

If, in addition, the signals $s_i(t)$ are weighted and power controlled, the upconverted multicarrier signal at the MCPA amplifier input at antenna branch n can be written as:

$$x_n(t) = \mathrm{Re}\left\{ \sum_{i=1}^{M} w_{ni} \gamma_i s_i(t) e^{j(\omega_c + \omega_i)t} \right\} \tag{5.4}$$

where w_{ni} is the complex weight for signal i at antenna n and $s_i(t)$ is the complex base-band information signal of user i. The gain factor γ_i, defined in equation (5.9), is the power control to user i[1]. Thus, if a mobile user is far away and/or is shadowed, the factor γ_i is large to compensate for the path loss and/or the slow fading. This issue is further discussed in Section 5.2.3. The weights and the gain factor are assumed to be constant during

[1]Later, this power control factor will be denoted γ_{ki}, as the gain factor from base-station k to user i. Here we have omitted the base-station index k, to simplify the expressions.

the time period of interest.

To make the expressions more compact, we combine the contributions to the signals magnitude and phase into the two variables V_{ni} and ϕ_{ni} respectively,

$$w_{ni}\gamma_i s_i(t) = |w_{ni}|\gamma_i c e^{j(\varphi_i(t)+\alpha_{ni}))} \triangleq V_{ni}e^{j\phi_{ni}(t)}, \tag{5.5}$$

where c is the constant modulus envelope amplitude, that is $s_i(t) = ce^{j\varphi_i(t)}$ and $\varphi_i(t)$ is the information-bearing phase of signal i. We assume that $c = 1$, without loss of generality. The phase argument α_{ni} is the argument of weight w_{ni}, that is, $w_{ni} = |w_{ni}|e^{j\alpha_{ni}}$ and in the last step of equation (5.5) we used the definition $\phi_{ni}(t) \triangleq \varphi_i(t) + \alpha_{ni}$. Thus, the right hand side of (5.5) is in polar form with V_{ni} real.

Using (5.5), we can rewrite expression (5.4) in the useful polar form as

$$x_n(t) = (I_n^2 + Q_n^2)^{1/2} \cos\left(\omega_c t + tan^{-1}\left(\frac{Q_n}{I_n}\right)\right) \tag{5.6}$$

where the quadrature components are

$$I_n(t) = \text{Re}\left\{\sum_{i=1}^{M} V_{ni}e^{j(\omega_i t+\phi_{ni}(t))}\right\} \tag{5.7}$$

$$Q_n(t) = \text{Im}\left\{\sum_{i=1}^{M} V_{ni}e^{j(\omega_i t+\phi_{ni}(t))}\right\} \tag{5.8}$$

Thus, the quadrature components $I_n(t)$ and $Q_n(t)$ contains the complex base-band signals, at their respective sub-carrier frequencies ω_i, $i = 1, \ldots, M$. They are used to represent the envelope and the phase of the RF signal, as seen in equation 5.6.

5.2.3 The dynamic range of the transmitted signal

The dynamic range of the transmitted signal depends on the dynamic range of the used weights and the power control. The dynamic range of the weights is usually negligible as compared to the power control that adapts to equalize the near-far ratio and the slow fading. The magnitude of the weights usually have a maximum dynamic range of 15 dB. This was measured on the antenna array testbed, described in Section 2.2 using the SMI algorithm.

The near-far ratio is due to the varying distance between the basestation and a mobile user i. The received power Γ_i has a large dynamic range, up to 100 dB [50]. The downlink power control dynamic range in GSM is limited to 30 dB in 15 discrete steps [51]. With an adaptive array the downlink power control does not need to be performed in discrete steps. We assume that the power control is designed to give each mobile the same received power as the basestation receives from the corresponding mobile . The aim is to achieve a transmitter dynamic range that is equal to the dynamic range of the received signal. Thus, if the path loss between basestation k and mobile i is denoted $G_{k,i}$, then the transmitted power is $G_{k,i}^{-1}$. This power is to be distributed over the N antennas. Therefore, the output power of each antenna must, for the power control criterion to hold, fulfill

$$\gamma_{ki} = \frac{G_{k,i}^{-1}}{N} \quad . \tag{5.9}$$

If a MCPA with M signals is used, then the dynamic range of the composite signal is even larger as compared to the single carrier case. The transmitted power γ_{ki} can be modeled as a stochastic variable with a probability distribution that depend on the size of the cell, the traffic, type of sectorization, type of environment (urban, rural etc.), and the used power control scheme of the mobile station. The probability density function of γ_{ki} has to be estimated from measurements or from simulations. When calculating the back-off of the MCPA, an average of the transmitted power can be obtained by setting

$$P_{avg,out} = E\left\{\gamma_{ki}\right\} \tag{5.10}$$

where E is the expectation operator, and then use equation (5.2) to calculate OBO.

5.3 Introducing the nonlinearity

In this section we introduce the Shimbo Amplitude Function (SAF) model for the nonlinear power amplifier. The SAF method significantly reduces the simulation time when simulating a multicarrier signal with non-constant envelopes in a narrowband channel with nonlinear devices [46]. In Section 5.3.2 the output signal from a nonlinear amplifier model is described using the SAF method. In Section 5.3.3 we show how the distortion from the nonlinearity can be approximated as an equivalent weight distortion and finally in 5.3.4 the distorted radiation pattern is calculated when the nonlinearity is present.

5.3.1 The power amplifier model

In this section we introduce a memoryless model for the nonlinear power amplifier. It is desirable that the model is in analytic in a closed form and that different amplifier characteristics can be generated by changing a single model parameter. If we express the complex envelope of the nonlinearity input signal as

$$z_{in}(t) = \varrho(t) \cdot e^{j\psi(t)}, \tag{5.11}$$

where the envelope $\varrho(t)$ is normalized such that $0 \leq \varrho(t) \leq 1$, then the complex envelope of the output signal can be expressed by

$$z_{out}(t) = g\left(\varrho(t)\right) e^{j(\psi(t)+f(\varrho(t)))} \tag{5.12}$$

where $g(\cdot)$ and $f(\cdot)$ is the AM/AM and AM/PM conversion of the nonlinear amplifier. For the envelope distortion, $g(\varrho(t))$, Cann suggested a limiter model that is linear for small signals, has an asymptotic output level at large input signals and is analytic in a closed form which is desirable. Also the amplifier characteristics can be changed easily by adjusting three parameters [52]. Cann's limiter model is given by

$$g(\varrho(t)) = \frac{\nu\varrho(t)}{\left(1 + \left(\frac{\nu\varrho(t)}{K_a}\right)^s\right)^{\frac{1}{s}}} \tag{5.13}$$

where ν is the small signal amplification (set as ν=1 in the following), K_a is the output saturation level as $|\varrho(t)| \rightarrow \infty$ and s controls the knee sharpness, or the transition smoothness between the linear and nonlinear region, as seen in Figure 5.3 for different s and with K_a=1. The figure shows the envelope output amplitude as a function of envelope input amplitude, i.e. $g(\varrho(t))/|\varrho(t)|$ defined for a continuous wave (CW) tone.

The AM/PM conversion $f(\varrho(t))$, for a solid state power amplifier is usually small as compared to the traveling wave tube (TWT) amplifiers. At large back-offs, the AM/PM conversion is negligible, but when the input amplitude approaches the saturation point, the relative input-output phase shift becomes noticeable. This can empirically be modeled by the following

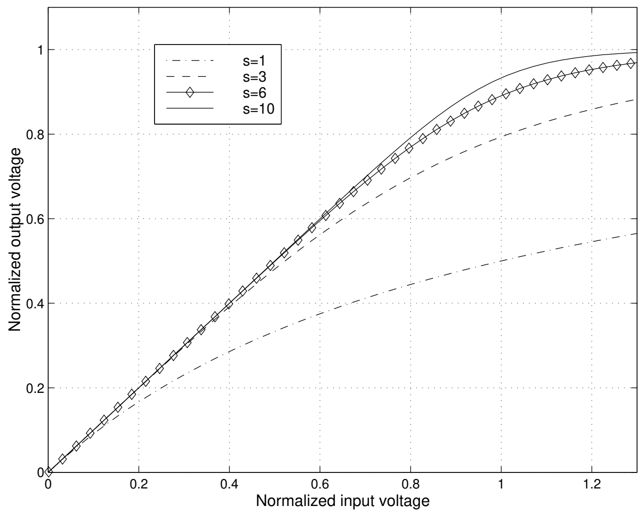


Figure 5.3: Parametric AM/AM conversion curve for $g(\varrho(t))$. When $s = 10$, the amplifier is approximately linear up to a normalized input voltage of 0.8, for the more nonlinear amplifier, $s = 3$, this breakpoint occurs at an input voltage of 0.4.

expression

$$f(\varrho(t)) = -\frac{K_\phi |\varrho(t)|^s}{1 + |\varrho(t)|^s} \tag{5.14}$$

where K_ϕ controls the maximum AM/PM conversion in the saturation region and the parameter s is the same as for the AM/AM curves above. This leads to the curve family in Figure 5.4, with K_ϕ=1 radian.

5.3.2 Output of a nonlinear bandpass memoryless amplifier with a multicarrier input signal

For the following simulations, we need the relative phase shift and gain compression for each of the M signals that are amplified in each MCPA. We are also interested in the power of the intermodulation products in the output signal that falls onto the sub-carrier frequencies in the input multicarrier signal. Thereby, the carrier to intermodulation ratio of each of the output sub-carriers can be calculated. It is assumed that the amplifier is memoryless, i.e. the AM/PM and AM/AM conversion curves are not functions of frequency. Stated differently, the bandwidth of the amplified signal is much

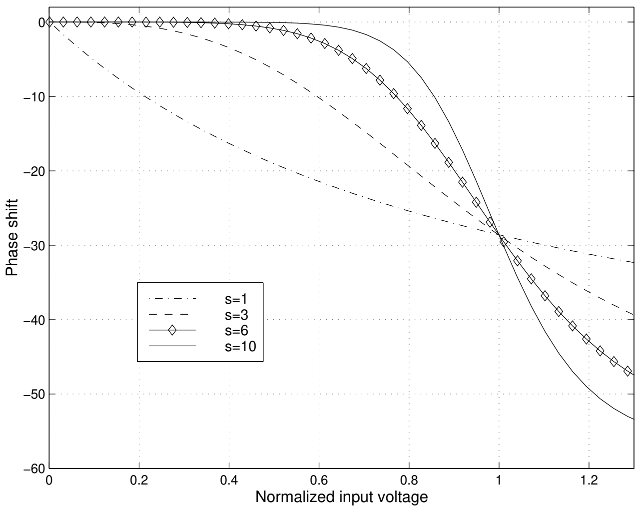


Figure 5.4: Parametric AM/PM conversion curves for $f(\varrho(t))$. When $s = 10$, the amplifier generates no phase shift up to a normalized input voltage of 0.6, compared to the $s = 3$ amplifier, which has a -10° phase shift at input voltage 0.8.

less than the inherent bandwidth of the amplifier. In this sense, we assume that the signal is narrowband.

We now define the first zone in the output spectrum as the region between $\omega_c - B$ and $\omega_c + B$, where $\omega_c \gg B > \omega_M > \omega_{M-1} > \ldots > \omega_1$. Correspondingly, the second zone is defined by the region $2\omega_c - B$ and $2\omega_c + B$, and so forth. The assumption that the memoryless amplifier is bandpass implies that only the spectral components in the first-zone, around the input frequencies need to be considered in the amplifier output signal. Higher order zones are easily attenuated by a filter, after the MCPA.

In communication systems, the carriers, $\omega_c + \omega_i$, $i = 1, \ldots, M$ are placed in a grid of equal frequency spacing. That is, $\omega_{i+1} - \omega_i = \Delta\omega$ is a constant and $\Delta\omega$ is the frequency distance in the grid. This implies, that the frequency of the generated intermodulation products falls exactly onto other carrier frequencies. This is inevitable, because the frequency distance is chosen as $\Delta\omega = W$, for maximal utilization of the frequency spectrum. Figure 5.5 shows the input and amplifier output signal spectrum of a frequency division multiplexed (FDM) signal in the first zone, consisting of three equally spaced carriers, each with a modulation bandwidth of W. The space between the carriers is exaggerated in the Figure 5.5 for clarity. Inter-

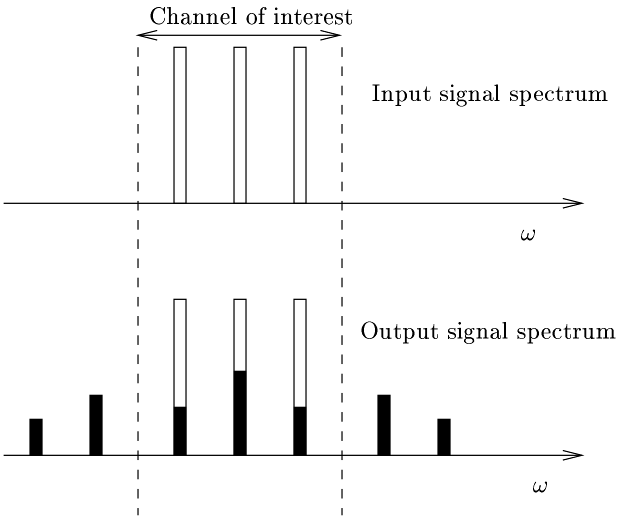


Figure 5.5: Spectrum of input FDM signals (upper) and output (lower). The white bars corresponds to the three sub-carriers in the multicarrier signal, each with a modulation bandwidth of W. The black bars in the output spectrum shows the intermodulation products. To reduce simulation time, it is only necessary to calculate the magnitude and phase of the intermodulation products that falls onto input sub-carrier frequencies, i.e. in the channel of interest.

modulation products are shown with black bars in the lower figure of Figure 5.5. The channel of interest in the output spectrum is shown and only those intermodulation products (black bars) in the output spectrum of the channel of interest need to be calculated in the simulation. This is utilized in the SAF method and it will lead to a reduction in simulation bandwidth and simulation time.

The instantaneous output from amplifier n can be written, by using (5.6) and (5.12) as

$$y_n(t) = g\left[(I_n^2 + Q_n^2)^{1/2}\right] \cos\left(\omega_c t + tan^{-1}\left(\frac{Q_n}{I_n}\right) + f\left[(I_n^2 + Q_n^2)^{1/2}\right]\right) \tag{5.15}$$

The nonlinearity will make the gain compression and phase shift of one carrier dependent on the amplitudes of all the other carriers that enter the amplifier. This is due to the gain compression function $g(\cdot)$ uses the total envelope of the input signal as an argument, see Equation (5.15). The gain compression for a specific sub-carrier is thus a nonlinear function of the amplitude and phase of each signal that enters the MCPA. To efficiently calculate this, the SAF method described below is useful. For array antennas, this is equivalent by saying that the beamformer weights corresponding to different beam-patterns (different signals) affect each other, which will create a beam-pattern distortion and intermodulation beams.

The work by Shimbo [47] and Schneider [46] gives an useful expression (5.16), which is a reformulation of (5.15). Equation (5.16) below provides the output of a bandpass memoryless nonlinear device given the nonlinearity functions $g(\cdot)$ and $f(\cdot)$, and a multicarrier input consisting of sub-carriers with arbitrary amplitudes. The output is given as a sum over all the frequencies that is represented in the output spectrum. By choosing the term that corresponds to the frequency of interest, we easily get the voltage and the phase of the corresponding output frequency component, by evaluating the complex function $S_n(k_1, k_2, \ldots, k_M; t)$.

$$y_n(t) = Re\left\{ \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \cdots \sum_{k_M=-\infty}^{\infty} S_n(k_1, k_2, \cdots, k_M; t) \right. \\ \left. \exp\left(j\omega_c t + j\sum_{l=1}^{M} k_l(\omega_l t + \phi_{nl}(t)) \right) \right\} \quad (5.16)$$

where ϕ_{nl} is the sum of the information-bearing phase of signal l and the phase of the weight for signal l at antenna n as:

$$\phi_{nl}(t) = \varphi_l(t) + \alpha_{nl} \quad (5.17)$$

The expression (5.16) describes the signals at the fundamental frequencies as well as those generated by the nonlinearity, e.g. harmonics and spurioses of all frequencies and all orders of intermodulation products. To restrict the sum only to the fundamental and third order intermodulation products in the first zone of the output, we add some constraints to the sums in (5.16). First, to observe only the frequency channel in the output spectrum corresponding to sub-carrier i in the multicarrier signal, we set a constraint as

$$k_1\omega_1 + k_2\omega_2 + \cdots + k_M\omega_M = \omega_i \quad . \quad (5.18)$$

This constraint on the integers k_i, will make the output signal in Equation (5.16) to only contain terms that falls between $\omega_c + \omega_i - W/2$ and $\omega_c + \omega_i +$

$W/2$. To further constrain the sums to the terms that are present in the input signal, we add

$$|k_1| + |k_2| + \cdots + |k_M| = 1 \quad . \tag{5.19}$$

This will give the output signal of each of the sub-carriers in the multicarrier signal. Furthermore, we use the constraint

$$|k_1| + |k_2| + \cdots + |k_M| = 3 \tag{5.20}$$

to consider the third order intermodulation products in the output signal. The order parameters k_i are integers. To clarify this, we present an example:

Example 1

Consider an input multicarrier signal consisting of three modulated sub-carriers, M=3. We assume that the sub-carriers are equally spaced, so that $\omega_1 + \Delta\omega = \omega_2$ and $\omega_1 + 2\Delta\omega = \omega_3$. Assume that we are interested in the output signal in the frequency channel $\omega_c + \omega_2$. Constraint (5.18) *gives*

$$k_1\omega_1 + k_2\omega_2 + k_3\omega_3 = \omega_2, \tag{5.21}$$

or, if we rewrite this

$$k_1\omega_1 + k_2(\omega_1 + \Delta\omega) + k_3(\omega_1 + 2\Delta\omega) = \omega_1 + \Delta\omega \quad . \tag{5.22}$$

Thus, the constraint in this case implies that

$$\begin{aligned} k_1 + k_2 + k_3 &= 1 \\ k_2 + 2k_3 &= 1 \end{aligned} \tag{5.23}$$

which can be fulfilled, if for example $(k_1, k_2, k_3) = (1, -1, 1)$. This is a mixing of all the three input sub-carriers, and the corresponding output amplitude is $S_n(1, -1, 1)$ where S_n is the Shimbo Amplitude Function as described below. Another solution to equation (5.23) *is $(k_1, k_2, k_3) = (2, -3, 2)$. This is an example of a seventh order intermodulation product term, and it can be verified, with the assumptions above, that $2\omega_1 - 3\omega_2 + 2\omega_3 = \omega_2$. These high order intermodulation products are usually small compared to the lower order products and we can thereby neglect them. If the constraint* (5.20) *is added to the example, to consider all third order intermodulation products that fall onto the frequency channel with center frequency $\omega_c + \omega_2$, only the choice $(k_1, k_2, k_3) = (1, -1, 1)$ solves Equation* (5.23). □

The function $S_n(k_1, k_2, \cdots, k_M; t)$, which first appeared in [47] is referred to as the *Shimbo Amplitude Function* (SAF) and is defined as

$$S_n(k_1, k_2, \cdots, k_M; t) = \int_0^\infty r \left\{ \prod_{i=1}^{M} J_{k_i} \left[V_{ni}(t) r \right] \right\} \int_0^\infty \rho g(\rho) e^{jf(\rho)} J_1(r\rho)\, d\rho\, dr \tag{5.24}$$

where $V_{ni}(t)$ is the envelope of the ith input signal component in amplifier n and $J_p(x)$ is the pth order Bessel function of the first kind. The SAF function S_n is a complex valued function and it gives the complex baseband amplitude of the signal component in the output signal corresponding to the set of indices $\{k_i\}_{i=1}^M$. The SAF is derived in [47] by calculating the autocorrelation of the multicarrier output signal and by using some theorems on integration of Bessel functions that allows rewriting integrals as sums over the indices k_i.

In general, the calculation of (5.24) is not simple, but Shimbo found that the Bessel function series expansion of the nonlinearity $g(\cdot)e^{jf(\cdot)}$, yields a simple and useful expression for the SAF [53]. Therefore, finding the coefficients b_p and α in (5.25) by solving

$$\{\alpha, b_1, \ldots, b_P\} = \arg \min_{\alpha, b_1, \ldots, b_P} \left\{ \left| g(\rho) e^{jf(\rho)} - \sum_{p=1}^{P} b_p J_1(\alpha p \rho) \right|^2 \right\} \tag{5.25}$$

in some sufficiently large range of the normalized input amplitude[2] ρ, results in a desirable closed form solution of SAF

$$S_n(k_1, k_2, \cdots, k_M; t) \simeq \sum_{p=1}^{P} b_p \prod_{i=1}^{M} J_{k_i}(\alpha p V_{ni}(t)) \tag{5.26}$$

which can be used in simulations to calculate the distortion of signals due to AM/AM and AM/PM conversion in the transmitting MCPA. This is shown in [53], by using properties of integration of Bessel-functions. In addition, the Bessel series expansion of the envelope nonlinearity corresponds to a Fourier series expansion of the instantaneous voltage transfer function.

[2] We used the range $0 \leq \rho \leq 1.5$ for the normalized input amplitude, and obtained satisfactory results.

5.3.3 Weight distortion interpretation

In this section, as a side result, it is shown how the calculated beamformer weights can be replaced by distorted weights, that take into account the third order intermodulation products. Thus the signal distortion, introduced by the nonlinearity, is transferred into an equivalent weight distortion, that affects the radiation pattern. Beamformer weights for different signals on different frequencies will thereby affect each other, due to the nonlinearity. The results presented in this section are not necessary for the following discussion and the reader can jump directly to Section 5.3.4 without loosing track of the argumentation.

The distorted weight i applied to a signal on frequency $\omega_c+\omega_i$ is denoted $\tilde{w}_{ni}$. The weight w_{ni} in equation (5.4) can be replaced by the expression[3]

$$\tilde{w}_{ni} = \frac{1}{\gamma_i}\left[\sum_{k_1,k_2,\ldots,k_M=-\infty}^{\infty} S_n(k_1,k_2,\ldots,k_M;t) \exp\left(j\sum_{l=1}^{M} k_l(\omega_l t+\phi_{nl}(t)) - j(\omega_i t+\varphi_i(t))\right)\right] \quad (5.27)$$

where the constraint $k_1\omega_1 + k_2\omega_2 + \cdots + k_M\omega_M = \omega_i$ must be fulfilled to assure outputs on frequency $\omega_c + \omega_i$ only. To shorten the notation, the M sums in (5.16) over indices $k_1,\ldots,k_M$ is denoted by a single summation symbol in equation (5.27). The weight expression 5.27 is normalized with the gain factor γ_i to cancel the gain factor in expression (5.4). The power control gain factor γ_i is included in S_n, and to get the weight magnitude correct, the SAF function must be divided by the corresponding gain factor. To demonstrate the weight distortion due to nonlinearities, we present an example:

Example 2

Reconsider Example 1 with three carriers, M=3, and with the three signals being uniformly spaced, so $\omega_2 = \omega_1 + \Delta$ and $\omega_3 = \omega_1 + 2\Delta$. Then equation

[3]As before, the base-station index k is left out of the gain factor γ_{ki}.

(5.27) *gives the distorted weight on frequency* $\omega_c + \omega_1$ *as*

$$\begin{aligned}\tilde{w}_{n1} = \gamma_1^{-1} [S_n(1,0,0)\exp(j\alpha_1)+ \\ +S_n(0,2,-1)\exp(j(2\alpha_2 - \alpha_3) + j(2\varphi_2(t) - \varphi_1(t) - \varphi_3(t)))]\end{aligned} \quad (5.28)$$

on frequency $\omega_c + \omega_2$ *as*

$$\begin{aligned}\tilde{w}_{n2} = \gamma_2^{-1} [S_n(0,1,0)\exp(j\alpha_2)+ \\ +S_n(1,-1,1)\exp(j(\alpha 1 - \alpha_2 + \alpha_3) + j(\varphi_1(t) - 2\varphi_2(t) + \varphi_3(t)))]\end{aligned} \quad (5.29)$$

and finally on frequency $\omega_c + \omega_3$ *as*

$$\begin{aligned}\tilde{w}_{n3} = \gamma_3^{-1} [S_n(0,0,1)\exp(j\alpha_3)+ \\ +S_n(-1,2,0)\exp(j(2\alpha_2 - \alpha_1) + j(2\varphi_2(t) - \varphi_1(t) - \varphi_3(t)))]\end{aligned} \quad . \quad (5.30)$$

The beamformer weight for the signal at the middle frequency $\omega_c + \omega_2$ *is distorted by the intermodulation product corresponding to* $(k_1, k_2, k_3) = (1, -1, 1)$ *in accordance with Example 1. Note that the SAF functions* $S_n(k_1, k_2, k_3)$ *are complex if the amplifier has AM/PM conversion. It can be verified that if these three equations are inserted into Equation* (5.4) *the result will yield:*

$$\begin{aligned}y_n(t) = Re\Big\{\Big(\Big[S_n(1,0,0)e^{j\phi_1(t)} + S_n(0,2,-1)e^{j(2\phi_2(t)-\phi_3(t))}\Big]\, e^{j\omega_1 t} \\ + \Big[S_n(0,1,0)e^{j\phi_2(t)} + S_n(1,-1,1)e^{j(\phi_1(t)-\phi_2(t)+\phi_3(t))}\Big]\, e^{j\omega_2 t}+ \\ + \Big[S_n(0,0,1)e^{j\phi_3(t)} + S_n(-1,2,0)e^{j(2\phi_2(t)-\phi_1(t))}\Big]\, e^{j\omega_3 t}\Big)\, e^{j\omega_c t}\Big\}\end{aligned} \quad (5.31)$$

Which is exactly the expression (5.16)*, describing the output of the amplifier. Thus the nonlinearity has been replaced by an equivalent weight distortion. Note that the weights are applied to signals on different frequencies, but affecting each other through the intermodulation terms.*

In each frequency channel the output signal consists of the desired term and the term due to third-order intermodulation. The third order intermodulation "signal" at frequency ω_1 *is* $\exp[j(2\varphi_2(t) - \varphi_3(t))]$. *Thus it is a function of the information signals transmitted on two other frequencies. This nonsense-signal is radiated with a radiation pattern determined by the equivalent "weights"* $S_n(0,2,-1)e^{j(2\alpha_{n2}-\alpha_{n3})}$ *for* $n = 1, \ldots, N$.

If the amplifier is ideal, with unity gain, then $S_n(1,0,0) = w_{n1}\gamma_1$, $S_n(0,2,-1) = 0$ *and so on. The ideal case will thus give the amplifier* *output signal as*

$$y_n(t) = Re\left\{\left(w_{n1}\gamma_1 e^{j(\omega_1 t+\varphi_1(t))} + w_{n2}\gamma_2 e^{j(\omega_2 t+\varphi_2(t))} + \right.\right.$$
$$\left.\left. + w_{n3}\gamma_3 e^{j(\omega_3 t+\varphi_3(t))}\right) e^{j\omega_c t}\right\} \quad (5.32)$$

as expected. □

The effective weight thus consist of the original, "correct" weight plus the distortion terms from the nonlinearity. Interpreted as radiation patterns, this is a superposition of the original "correct" radiation pattern and the intermodulation radiation patterns, as discussed in Section 5.3.4 below.

Carrier to intermodulation power ratio

Based on the theory and assuming three equally spaced signals $(\omega_1, \omega_2, \omega_3)$ as in the example above, it is possible to express the carrier to third order intermodulation power ratio on the output of antenna n on e.g. the frequency ω_1 as

$$\left(\frac{C}{IM_3}\right)_{\omega=\omega_1} = \left|\frac{S_n(1,0,0)}{S_n(0,2,-1)}\right|^2 \quad (5.33)$$

5.3.4 The radiation pattern

In this section we use the previously defined expressions of the output of a nonlinear MCPA in the context of array antennas. The expression becomes very complex and we then restrict the following discussion to the two signal case and a simple Butler matrix beamformer. Using these simplifying assumptions, we derive the direction of radiation of the intermodulation products. The generalization of the results are then discussed.

Assuming a transmitting linear equally spaced antenna array, the received complex baseband signal in the far-field at an angle θ from broadside is given by the antenna-space to beam-space transformation

$$U(\theta) = \sum_{n=1}^{N} \Omega_n(\theta) y_n(t) e^{-j(n-1)\frac{2\pi d}{\lambda}\sin(\theta)} \quad (5.34)$$

where d is the antenna element spacing, λ is the wavelength of the transmitted signal, $\Omega_n(\theta)$ is the antenna element gain for antenna n in direction θ and $y_n(t)$ is the signal transmitted from base-station antenna n. Here we assume that the antenna elements have identical radiation patterns, thus

$\Omega_n(\theta) = \Omega(\theta)$ and that we have line of sight propagation and no multipath propagation. Furthermore, the difference in λ for different carrier frequencies are neglected under the assumption that the carriers are closely spaced in frequency compared to the carrier frequency ω_c. We assume that $d = \lambda/2$ to satisfy the spatial Nyquist criterion.

We assume that a MCPA is used to amplify the multicarrier downlink signal, as in the ABF or DBF-MCPA configurations described in Chapter 1. We describe the nonlinearity with the SAF in equation (5.24). Now, by inserting the output signal from each antenna, that is, equation (5.16) in to equation (5.34) and consider a complex baseband signal representation, the signal in the far field of direction θ is

$$U(\theta) = \Omega(\theta) \sum_{n=1}^{N} \sum_{k_1,k_2,\ldots,k_M=-\infty}^{\infty} S_n(k_1, \ldots, k_M) e^{-j(n-1)\pi \sin(\theta)} \times \\ \times \prod_{l=1}^{M} e^{j(k_l(\phi_{nl}(t)+\omega_l t))} \quad . \tag{5.35}$$

This is a very general expression and to interpret it, it is useful to restrict the study to the two tone case together with the Butler matrix transformer (BMT), see appendix A for a description of the Butler matrix.

Two signal case with the Butler Matrix Transformer

The two signal case is studied by setting $M = 2$ in (5.35). The first zone output constraint (5.19) gives $k_2 = 1 - k_1$. One of the third order intermodulation products is given by (5.20) as $k_1 = 2, k_2 = -1$ and the radiation pattern for that component is

$$U_{2,-1}(\theta) = \Omega(\theta) \sum_{n=1}^{N} S_n(2, -1) e^{-j(n-1)\pi \sin(\theta)} e^{j((2\omega_1-\omega_2)t+2\phi_{n1}(t)-\phi_{n2}(t))} \tag{5.36}$$

A simple way to create beams in the downlink is to use the BMT. The BMT performs essentially a beam-space to antenna-space transformation. The BMT performs a linear transformation of the N input ports and has a structure similar to the Fast Fourier Transform [54],[5]. The N output ports are then connected to the N antennas in the array. The signal at an input port p will be present at all N output ports, but with a phase difference that is a multiple of $\Delta\phi_p$. The normalized input/output relationship for the

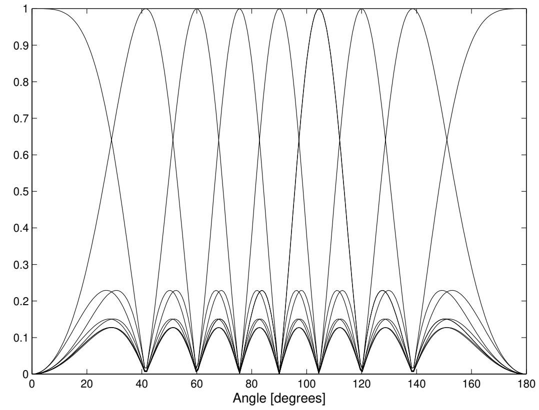


Figure 5.6: Plot of the eight beams of an 8×8 BMT

BMT can thus be written as

$$x_n(t) = \frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} s_p(t) e^{-jn\pi\left\{\frac{2p}{N} - \frac{(N-1)}{N}\right\}} \tag{5.37}$$

where $x_n(t)$ is the BMT output port n signal to be connected to the multicarrier amplifier at antenna n and $s_p(t)$ is the signal connected to input port p of the BMT.

The output phase gradient $\Delta\phi_p = \frac{\pi(2p-(N-1))}{N}$ depends on the input port p used and will give rise to a beam in a certain direction given by (5.43). For the 8×8 BMT the eight beams are plotted in Figure (5.6). The beams corresponding to different input ports are orthogonal in the sense that where one beam has a maxima, the others have nulls.

The output of the BMT for signal l has a constant phase gradient property, i.e. $\phi_{n,l} - \phi_{n-1,l} = \alpha_{n,1} - \alpha_{n-1,1}$ is not a function of n, where $\alpha_{n,l}$ is the phase of beamformer weight for signal l at antenna n. In words, a signal l appears on all BMT output ports with phase difference $\Delta\phi_p$, $2\Delta\phi_p$,…,$(N-1)\Delta\phi_p$ and so forth, compared to antenna 1. This leads to a simplification of (5.36) by using the relation

$$\phi_{nl} = \varphi_l + \alpha_{nl} = \varphi_l + (n-1)\Delta\phi_{p_l} \tag{5.38}$$

meaning that signal l is connected to port p_l of the BMT with output phase gradient $\Delta\phi_{p_l}$. Furthermore, the factor $S_n(2,-1)$ will not depend on n due to the uniform aperture distribution (the weight magnitude of the $M \times N$ weights are equal). So the average amplitudes of the multicarrier signals that enter the amplifiers are equal for all MCPAs. Equation (5.36) can now be expressed as

$$U_{2,-1}(\theta) = \Omega(\theta) S_n(2,-1) e^{j((2\omega_1-\omega_2)t+2\varphi_1(t)-\varphi_2(t))} \times \\ \times \sum_{n=1}^{N} e^{j\left\{(n-1)(2\Delta\phi_{p_1}-\Delta\phi_{p_2})-(n-1)\pi\sin(\theta)\right\}} \quad (5.39)$$

which has a maximum when

$$\pi\sin(\theta_{max}) = 2\Delta\phi_{p_1} - \Delta\phi_{p_2} \quad (5.40)$$

or

$$\theta_{max} = \arcsin\left(\frac{2\Delta\phi_{p_1} - \Delta\phi_{p_2}}{\pi}\right) \quad . \quad (5.41)$$

In the direction θ_{max} all the N signals are added coherently, giving an increase in intermodulation power by $10\log(N)$ dB. More generally, the direction of radiated power maxima for arbitrary M and arbitrary combination of $k_1, \ldots, k_M$ can be written as

$$\theta_{max} = \arcsin\left(\frac{\sum_{l=1}^{M} k_l \Delta\phi_{p_l}}{\pi}\right) \quad . \quad (5.42)$$

The direction of the main lobe corresponding to input port p is given by

$$\theta_{main,p} = \arcsin\left(\frac{\Delta\phi_p}{\pi}\right) \quad (5.43)$$

Two signals at different input ports will radiate the intermodulation in some other beam direction, different from the direction of the desired signals. Two signals using the same butler matrix input port p will create odd-order intermodulation products that radiates in the same beam direction as the transmitted signals. This can be shown by observing, for the BMT that

$$k_1\Delta\phi_{p_1} + k_2\Delta\phi_{p_2} = \Delta\phi_{p_r} \quad (5.44)$$

where

$$r = (k_1 p_1 + k_2 p_2) \,\text{modulo}\, N \tag{5.45}$$

When two signals are connected to the same port, $p_1 = p_2 = r$ for odd order intermodulation products, so the IM_3 is radiated in the same beam direction as the two signals. This increases the carrier to interference level of the desired users, but, on the other hand, reduces the carrier to interference ratio of mobiles in other directions.

General case

In the general case, a more "intelligent" weighting technique might be used to create the downlink beams as in the DBF-SCPA, DBF-MCPA and SWR cases. If we restrict ourselves to analog fixed/switched beamforming, then the next step in complexity from the standard BMT is to use tapered weights in the beamformer, but keep the constant phase difference property of the BMT. Different aperture weighting distributions have been studied in the field of radar systems. Using cosine, parabolic or Hamming weight magnitude distributions it is possible to decrease the side lobe levels (SLL) to more than 40 dB below the maximum intensity as compared to the uniform aperture distribution of the BMT which gives a theoretical SLL of 13 dB. The uniform aperture distribution however has the highest gain in the main lobe of any weight distribution. To maintain the same received power at the mobile when using tapered weights, the amplifier must therefore be driven at a smaller back-off. Thus, the C/IM_3 is decreased.

The BMT beamformer gives the worst possible amplification of intermodulation products of $10\log_{10}(N)$ dB in the direction θ_{max} determined by equation (5.42). All N intermodulation products adds up coherently in this direction for the BMT beamformer due to the constant phase gradient property. If the constant phase gradient criterion (5.38) is not fulfilled as in the DBF- and SWR cases, where the phase is used in the placement of "nulls" in the transmitted radiation pattern, the intermodulation products will be "smeared out" over the horizontal plane and no distinct maximum appears. The calculation of radiated intermodulation power must in that case be performed on a case-by-case basis for each weight and signal set using (5.35).

In this case, coherent intermodulation from all N antennas does not add coherently in any direction as in the case with uniform amplitude distribution and the BMT. The received intermodulation voltage is a vector sum of

N vectors and the triangle inequality gives that the received intermodulation voltage is less than the in-phase voltage addition from the BMT. Conclusively: The constant phase gradient property corresponds to the worst case of radiated intermodulation power, where all N antenna signals contribute, in-phase, to the intermodulation interference. In the more general case, with null-steering, the angle distribution of the intermodulation power has to be calculated on a case-by-case basis and has a peak value that is less or equal to the BMT case.

5.4 Simulations

The simulations aim to estimate the CDF of the CIR in a cellular system for different back-off and amplifier linearities. It is important for the system link budget analysis to know the probability of a certain CIR level for the mobiles in the system given a certain amplifier and at a given back-off. Comparisons will be made with the case of an ideal amplifier characteristics, that is an amplifier with no gain compression and constant phase conversion. The simulations follow the flowchart in Figure 5.7.

5.4.1 Simulation assumptions

A hexagonal cell structure with seven cells is used in the simulations. Each cell is further divided into three sectors of 120 degree coverage each, see Figure 5.8. We assume that every frequency channel is used in every sector. The signal carrier to interference ratio is measured for the mobile in the center cell, to decrease the "edge effect" from the use of a finite number of cells.

In Figure 5.8 the mobile under study is marked with a filled circle. The basestation array antennas that interfere with this mobile have an unfilled circle (mobile) in its sector. In the simulations the following assumptions are made:

1. Each sector uses frequency multiplexing with three frequency channels and the uplink interfering co-channel mobiles are assumed to be co-channel mobile in the downlink also.

2. Each sector has an array antenna with N=8,4 or 2 antenna elements, linearly spaced with an uniform $\lambda/2$ spacing, where λ is the wavelength of the carrier frequency. The individual antenna element patterns $\Omega(\theta)$ are given by

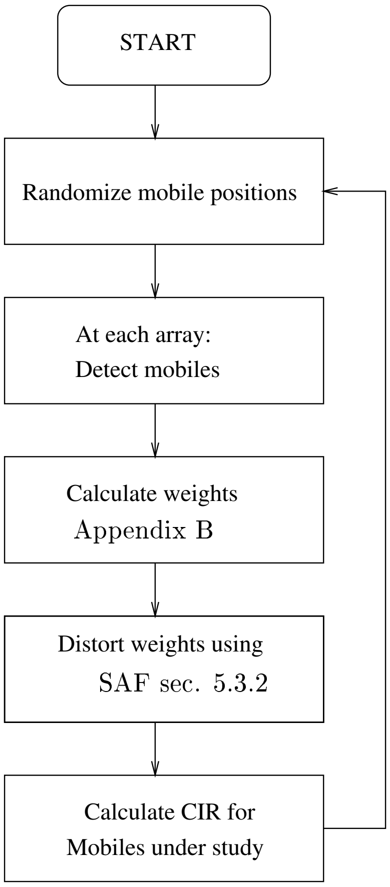


Figure 5.7: Simulation flowchart

$$|\Omega(\theta)|^2 = \begin{cases} cos^2(\theta) & \text{if } |\theta| \leq 90^\circ \\ 0 & otherwise \end{cases} \tag{5.46}$$

3. The path loss slope is assumed to be γ=3.8 and the angular spreading at the mobile as seen from the base is $\sigma = 4^\circ$, independent of the base to mobile distance. Only one cluster of scatterers is surrounding the mobile, thus no delay spread is assumed.

4. The received signal from the mobiles exhibit slow fading, log-normally distributed with standard deviation σ_L=6 dB. The log-normal fading

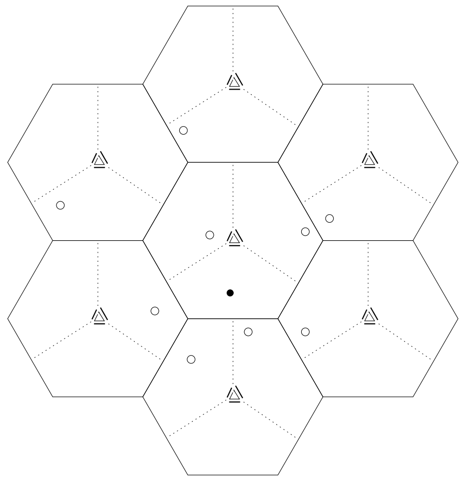


Figure 5.8: Cell configuration in simulations, $\bullet$=mobile under study, $\circ$= other mobiles

between a mobile and two different basestations are correlated with correlation coefficient 0.5

5. Each mobile is randomly placed in each sector m with equal probability in the area

$$0.1 \leq \left(\frac{cos(30^\circ)}{cos(\theta_m)}\right)^{\frac{2}{\gamma}} \left(\frac{r_m}{R}\right) \leq 1.3, \quad |\theta_m| \leq 60^\circ \qquad (5.47)$$

where R is the cell radius. This gives an uniform area distribution [49].

6. The mobile power control used makes the power received from the mobile equal for all mobiles in all sectors, implying that a received signal from a mobile interfering strongly in uplink will be strongly suppressed in the downlink transmission

7. No discontinuous transmission (DTX), frequency hopping or handover are assumed for simplicity

8. The mobile positions are completely re-randomized at each of the 800 iterations, and the CIRs of the three mobiles in the center sector are

calculated to give the cumulative distribution function (CDF) of the CIR for the test mobile.

The simulation was carried out for three different antenna array sizes, with $N = 8, 4, 2$. The received power at the mobile was held constant for the three antenna configurations by adjusting the transmitted power. Thus changing N from 4 to 2, implies that the IBO of the amplifiers has to be decreased by 3 dB

5.4.2 Determination of $\mathbf{P}_{avg,out}$

The value of $\mathrm{P}_{avg,out}$ is used in Equation (5.2) to set the output back-off in the simulations So, the probability density function of the transmitted power using perfect power control was estimated in a pre-run simulation by randomly placing the mobiles according to Figure 5.8 and equation (5.47). The received power was calculated and using the power control criterion (5.9) we obtain a sample from the distribution. The probability density function (PDF) of γ_{ki} was then estimated by a Monte Carlo simulation over 10000 iterations. The result is shown in Figure 5.9.

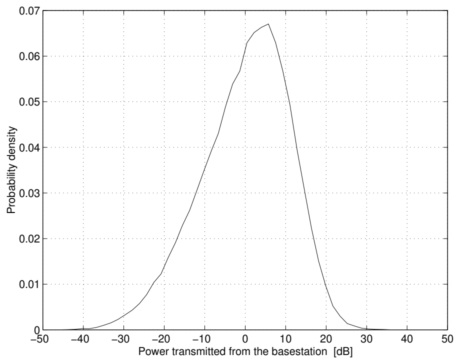


Figure 5.9: Estimated probability density function of the transmitted power γ_{ki} from simulations. The PDF is normalized around the mean, and it can be seen that the necessary dynamic range of the power control is approximately 60 dB.

The figure is normalized such that the mean transmitted power is 0 dB. It can be seen that the users should have a power control setting between -30dB to 30 dB, in accordance with the results in [49]. We neglect the dynamics in the weight magnitudes and the OBO is defined as the difference (in dB) between the mean of this distribution and the output saturation point of the amplifier used, see equations (5.10) and (5.2). Note that we defined OBO as the back-off for one of the M carriers in the input. Thus the back-off of the multicarrier signal is on average $10\log_{10}(M)$ dB lower than the single carrier case.

5.4.3 Calculation of CIR

It is assumed that the system is interference limited, thus the Carrier to Interferer ratio (CIR) is the quantity used for comparisons. If the received signal from a mobile is lower than the receiver sensitivity level, then the mobile is not detected as an interferer and will thus not be taken into consideration when the downlink radiation patterns are calculated.

The total interference experienced by the test mobile is due to direct interference and intermodulation from neighboring basestations. Here, we consider the third order intermodulation products only, which normally is the intermodulation product with the highest power. Thus, the CIR is calculated as

$$CIR = \frac{C}{\sum_m IM_3 + \sum_m I} \tag{5.48}$$

where C is the carrier power from the desired basestation. To the total interference power experienced by the mobile, there are contributions from the m basestations in the first tier, which contributes to the total interference power by third order intermodulation products (IM_3) and direct interference (I), weighted by their respective radiation patterns.

5.5 Simulation results

The cumulative distribution function (CDF) of the CIR for the test-mobile was estimated from the simulations. We investigated the impact of the number of antennas in the array, the OBO, and the linearity of the amplifiers. Also, the measured total degradation (TD) is defined and used to find a power efficient choice of OBO.

5.5.1 Impact of the number of antennas

One important option in the design of the antenna array is the choice of the number of antennas. Many antennas implies that narrower beams can be formed and more interferers can be nulled out as the degrees of freedom increase. However, the number of antennas is limited by space and economic factors. As the amount of transceivers must be multiplied with the number of antennas, there will be a trade-off between cost and performance.

We simulated the system with $N = 2, 4$ and 8 antennas and with a moderate linear amplifier ($s = 3$) and an amplifier with poor linearity ($s = 1$). The initial OBO[4] (for a single carrier) was chosen to 17 dB and 12 dB and three carriers were co-amplified in each MCPA. The OBO was then adjusted according to equation (5.9) to take the number of antennas into account. Thus with four antennas, each amplifier has to produce twice as much power as the eight antenna case, to yield the same received carrier to noise ratio at the mobile.

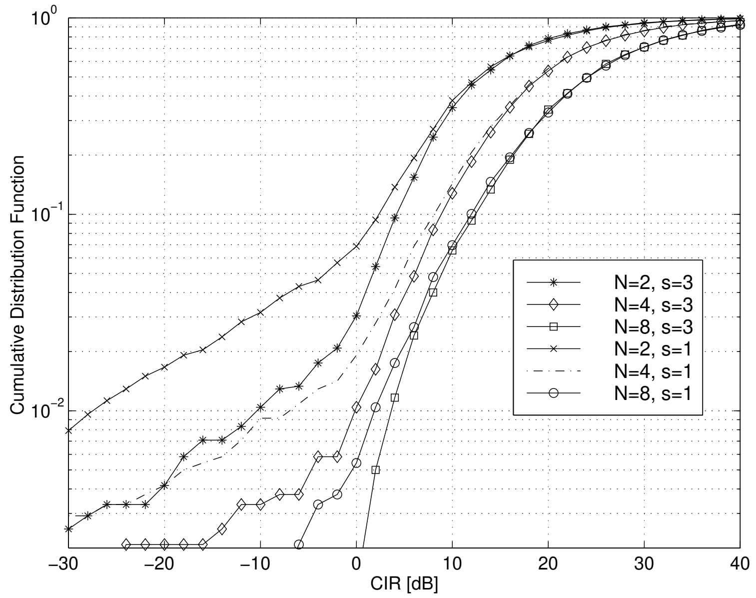


Figure 5.10: Cumulative distribution function for CIR using parametric amplifier models, Figure 5.3 and Figure 5.4 with 17 dB OBO

The results depicted in Figure 5.10 and Figure 5.11 shows that when the number of antennas is increased the impact of the nonlinearity is decreased. For OBO=17 dB and the 5th percentile (CDF=0.05), the CIR is -4 dB, 4 dB and 8 dB for the N=2,4,8 case respectively. This is due to the distribution of the amplification over several amplifiers, each of which can be used with a larger back-off, which in turn gives a higher linearity. When $N = 2$ the

[4]That is, the back-off from saturation of of the amplifier output signal.

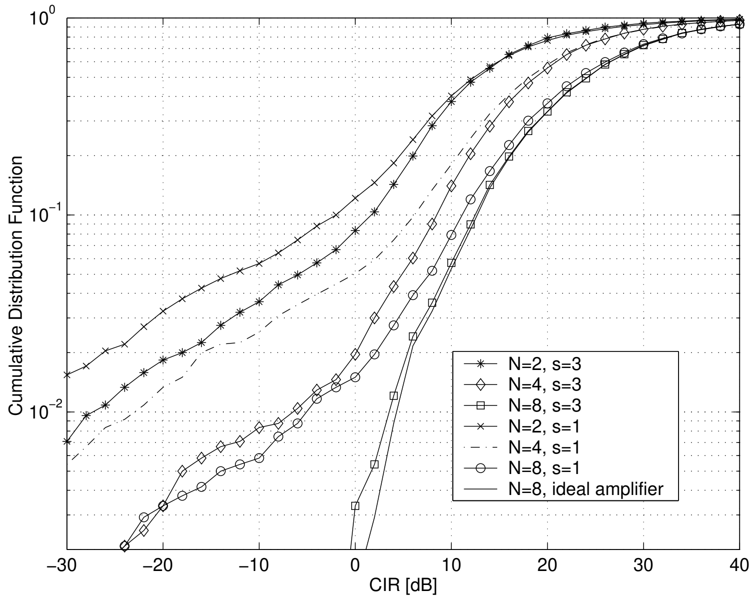


Figure 5.11: Cumulative distribution function for CIR using parametric amplifier models, Figure 5.3 and Figure 5.4 with 12 dB OBO

difference between the moderate and the poorly linear amplifier is strongly pronounced.

The difference between the curves using the same amplifier but with different number of antennas is due to the reduced ability to suppress interferers when the degrees of freedom in adaptive calculation of the weight are less than the number of interferers. The number of detected interferers was in the simulation 6 to 9, and with a two and a four element array it is impossible to fully suppress them all. Also the angle spread "uses up" degrees of freedom. Thus the curves in Figure 5.10 and 5.11 will not coincide for an array with ideal amplifiers.

When $N = 2$ the CIR is 6 dB lower for the $s = 1$ amplifier as compared to the $s = 3$ at the CDF=5% percentile for both OBO=12 dB and OBO=17 dB. When the number of antennas is increased to eight, the difference is 2 dB for the OBO=12 dB case, but for the large 17 dB back-off case, the difference is negligible. Thus with a very large back-off, the performance degradation due to the nonlinearity vanishes. So we can reduce the nonlinearity effects by adding more antennas, linearizing the amplifiers or increasing the OBO.

An interesting question is, of course, how large back-off should be used? This is investigated in the following section.

5.5.2 Choosing the Output Back-Off

The required amount of OBO that yields a certain quality of service was investigated and the results are shown in Figure 5.12 and Figure 5.13 for an $s = 3$ amplifier and an $s = 10$ amplifier respectively. The number of antennas were fixed to $N = 4$. When $s = 10$, the amplifier is very linear and

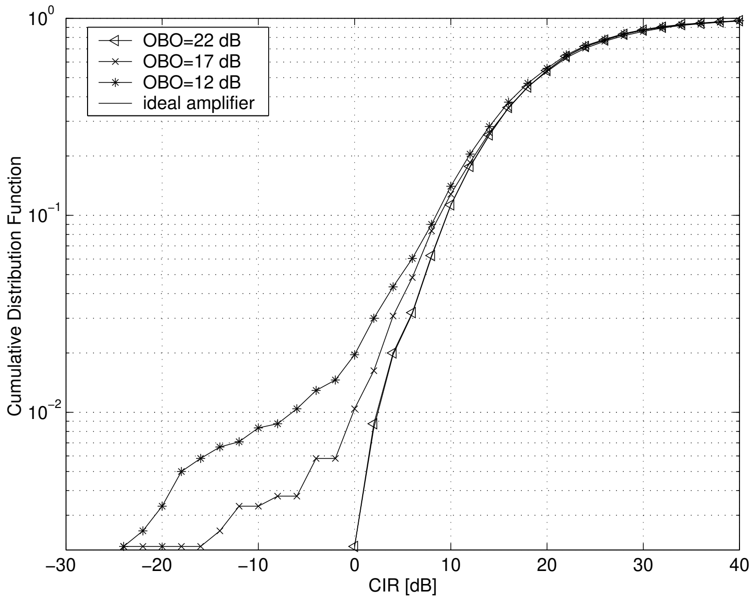


Figure 5.12: Cumulative distribution function for CIR using parametric amplifier model with $s = 3$ and $N = 4$ antennas.

the OBO=22 dB coincides with the ideal amplifier both for $s = 3$ and $s = 10$ in Figure 5.12 and Figure 5.13. With a smaller OBO, the amplifier has higher efficiency which is desirable and we see that the trade-off is between performance and efficiency. We propose a method of choosing the OBO for maximum power efficiency. The idea is to study the lower tail of the probability density function for the CIR of the mobiles, e.g. the 5th or the 10th percentile of the CIR distribution. Then the degradation in CIR between a system with an ideal amplifier (perfectly linear) and a nonlinear amplifier at the chosen percentile is calculated. As a reasonable choice of the OBO, we introduce the total degradation (TD) which is a useful performance measure for digital systems in the presence of nonlinearities. This parameter

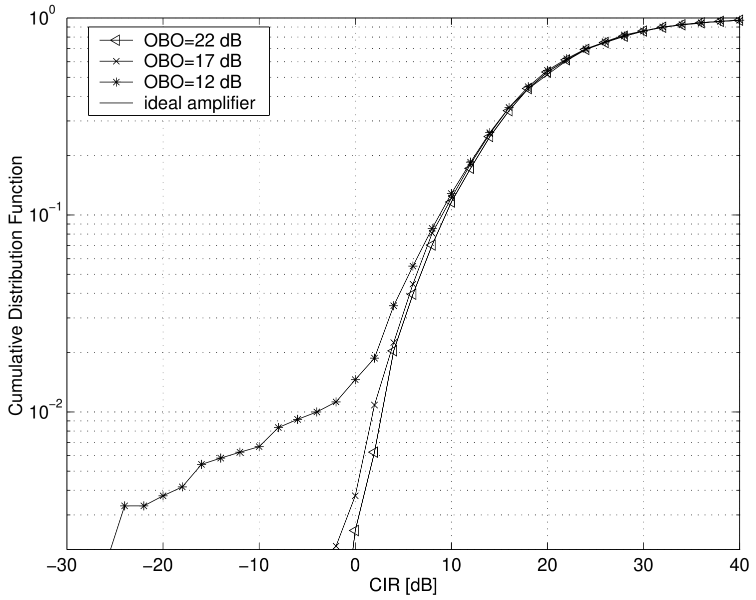


Figure 5.13: Cumulative distribution function for CIR using parametric amplifier model with $s = 10$ and $N = 4$ antennas

is defined as

$$\mathrm{TD}_{dB} = \widehat{CIR}_{dB} - \mathrm{CIR}_{dB} + \mathrm{OBO} \tag{5.49}$$

where CIR_{dB} is the carrier to interference ratio in dB for a given percentile with the nonlinear amplifier and $\widehat{CIR}_{dB}$ is the required carrier to interference ratio to obtain the same percentile with an perfectly linear amplifier. A good choice of operation point of the systems corresponds to the minimum of equation (5.49) and yields an intuitively power efficient choice although no properties of optimality is claimed.

The total degradation curves for the $s = 3$ amplifier is shown in Figure 5.14. At the 5th percentile the TD attains a minimum when OBO=14.5 dB, 9.5 dB and 7 dB for the $N = 2, 4, 8$ antenna case respectively. At larger back-off's the degradation in CIR for the mobiles is smaller, but the MCPA efficiency is lower and at smaller back-off's the CIR degradation becomes larger due to the increased interference levels in the system caused by intermodulation and distorted radiation patterns.

The results show, as expected, that the eight antenna has the smallest total degradation. When the OBO increases, the curves approaches asymp-

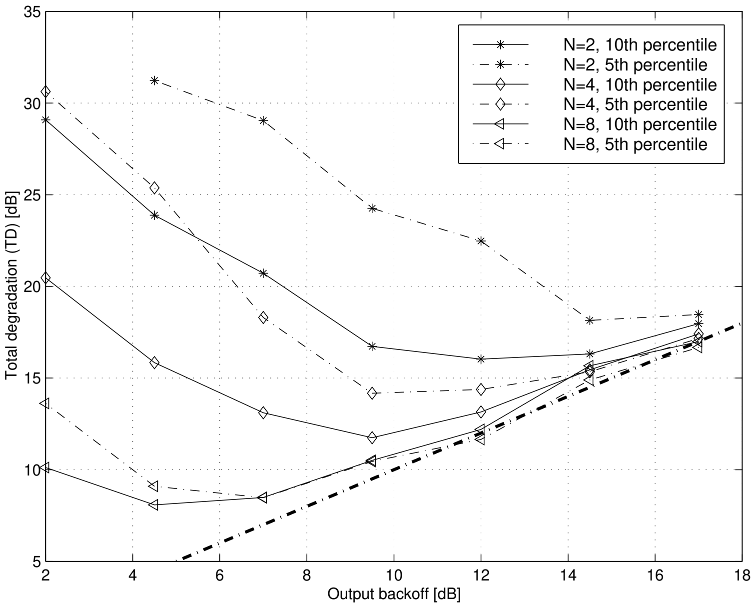


Figure 5.14: Total degradation using parametric amplifier model, Figure 5.3 and Figure 5.4 with $s = 3$. The curves has a minimum point which provides a suitable output back-off level for the tradeoff between OBO and CIR degradation.

totically the TD=OBO line as the CIR_{dB} approaches $\widehat{CIR}_{dB}$. From these curves it is possible to select the OBO level for each system configuration considered. One must, however, keep in mind, that there is restriction on the total radiated intermodulation power in adjacent channels, and it is necessary to calculate these to make sure that the specifications are not violated. The choice of OBO, that was proposed here, is only a good power efficient choice, maybe the OBO has to be increased further to meet specifications.

5.6 Measurements and calculations of intermodulation

To verify the results of Section 5.3.4 where a method of calculating the radiation pattern of the intermodulation distortion, measurements were performed in an anechoic chamber. We used a 4-element linear antenna array and an 8-by-8 analog BMT. Four standard microwave amplifiers were used and their respective AM/AM and AM/PM characteristics is shown in Figure 5.20. The setup corresponds to the ABF-MCPA option II described in Section 1.2.4 and is shown in Figure 5.15. The BMT was fed with two GSM

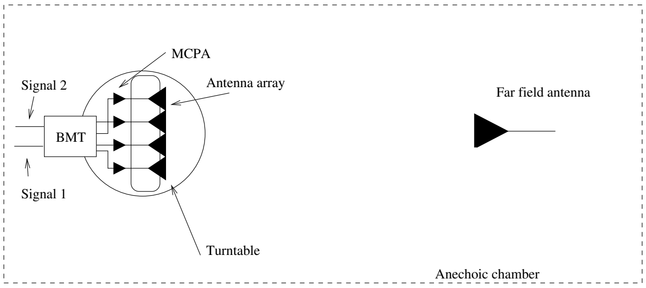


Figure 5.15: Measurement setup in the anechoic chamber for intermodulation measurements. A four element antenna array was used and a BMT for the beamforming.

signals with center frequencies 1.8000 GHz and 1.8004 GHz in two different BMT input ports. The input power of the two signals were adjusted to the appropriate input back-off and the radiation pattern was measured using the anechoic chamber with a turntable as shown in Figure 5.15. The angle was swept in 1° steps and for each angle the power density spectrum was measured with a spectrum analyzer.

5.6.1 Frequency-angle power spectral density measurements

The measurements, in Figure 5.17-5.19 give an intuitive picture of the spatial distribution of the intermodulation distortion. Input port number 6 was used for the f_1=1.800 0 GHz signal and port number 4 was used for the f_2=1.800 4 GHz signal. According to (5.42) and (5.43) the main-lobes will in this case be in the directions 22.0° and −7.2° respectively, while the third order intermodulation beams are in direction 61° for the $2f_1 - f_2$ product at 1.7996 GHz and $2f_2 - f_1$ product at 1.8008 GHz in −39°. This is easily verified by studying Figure 5.17 and Figure 5.16. These figures are normalized to 0dB maximum. In Figure 5.16, the fifth-order intermodulation product ($e.g. k_1 = 3, k_2 = -2$) is weakly visible at the DOA angles 61° and −39° as predicted by equation (5.42). Noteworthy is also the orthogonal array response vector properties of the beams from the BMT. In direction −7.2°, where beam 2 has its maxima, the other beams have nulls, giving a high level of isolation. This includes the intermodulation, which also has a null in that direction. So, by using a BMT, the spatial dimension can be used to reduce the intermodulation distortion at the mobile user. However, the IM_3

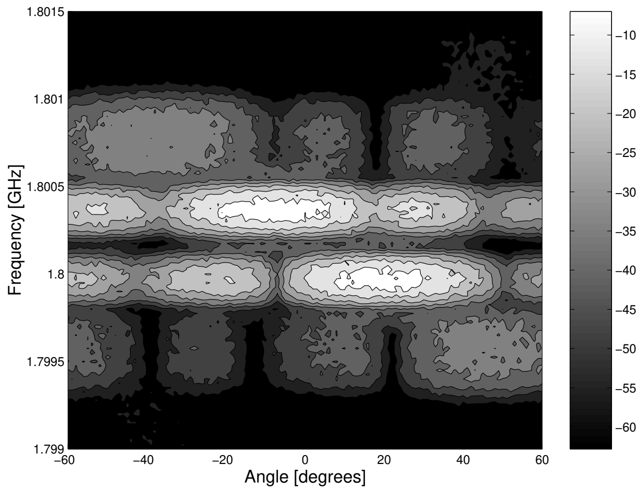


Figure 5.16: Normalized frequency-angle plot of power spectral density with saturated amplifiers (back-off decreased 8 dB compared to Figure 5.17). Here the third order intermodulation products are visible and also the fifth order intermodulation products are visible.

adds up coherently in another beam which increases the interference level of a potential co-channel user located in that direction.

This property of spatial intermodulation distortion filtering by the array antenna has earlier been described in [5] for satellite mounted array antennas. The analog BMT used here have a manufacturing error of $0.8dB$ and $8.5°$ which detoriates the orthogonality properties of the BMT array response vectors. These errors degrades the isolation and explains the not perfect positioned nulls in the measurements. This also demonstrate the need for calibration.

5.6.2 Measurements with tapered weights in ABF-MCPA systems

The BMT is a phase only beamforming method and thus all weight "magnitudes" are equal and equals the losses in in BMT. This corresponds to a uniform aperture distribution, or uniform tapering of the weight magnitude. By introducing a non-uniform tapering of the weight magnitude (e.g. a parabolic aperture distribution) the sidelobes of the radiation pattern can be reduced at the expense of a wider main-lobe and lower gain in the main-lobe as discussed in Section 5.3.4. The aperture distribution and the far

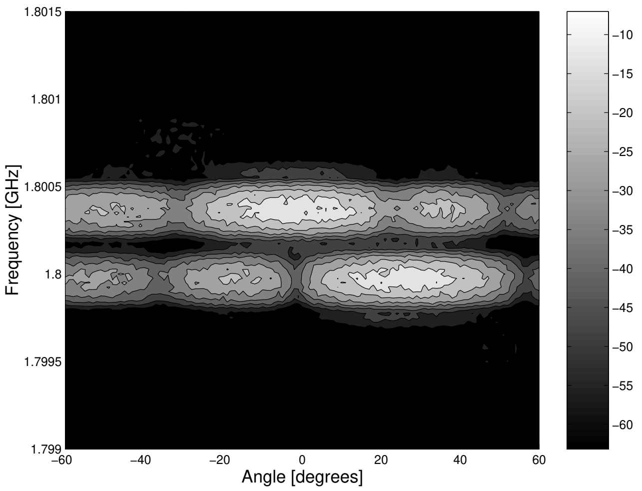


Figure 5.17: Normalized frequency-angle plot of power spectral density with large back off. The two beam-formed signals has main-lobes in 22° and $-7°$ respectively. The intermodulation products are weakly visible at -39° and 1.8004 GHz.

field radiation pattern are related by the Fourier transform [8].

A simple experiment with tapered weight magnitude was performed by inserting 3dB attenuators at the outermost antenna elements. This will degrade the antenna array gain in the main-lobe. The loss was measured to -1.84 dB using the tapered weights, but the side-lobe levels decreased from -9.7dB to -14.7 dB. See Figures 5.18 and 5.19. The loss in the mainbeam is not visible in the figures due to the normalization to 0 dB. The loss in mainbeam gain when using tapered weights must be compensated for by a higher amplifier power output, by using a more powerful amplifier or decreasing the IBO.

5.6.3 Calculations on weight tapering in ABF-MCPA systems

The AM/AM and AM/PM characteristics of the nonlinear amplifier were measured using a network analyzer and the results are given in Figure 5.20. The amplifier power characteristics were converted to instantaneous voltages and to which a Bessel function series expansion was fitted as described by equation (5.25). A BMT was simulated and different tapering and number of antennas were studied. The BMT creates main-lobes in direction $+22°$ and $-7.2°$ for the two signals respectively. The program generates two

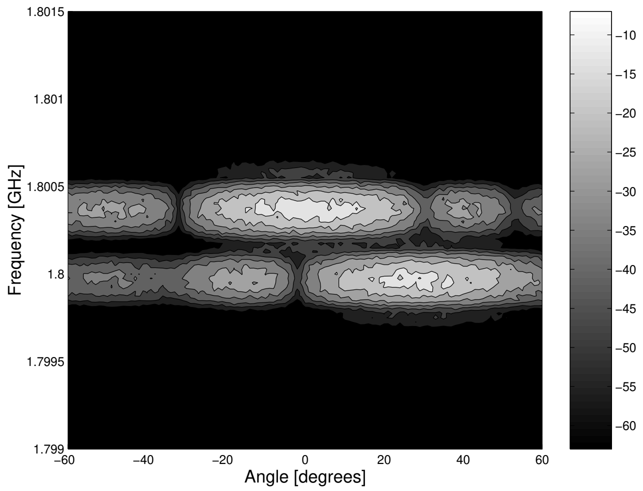


Figure 5.18: Normalized frequency-angle plot with back off and tapered weights. Note how the sidelobes are lower compared to Figure 5.17.

continuous wave (CW) baseband signals separated by 400kHz.

We adjusted the input amplitude of the two signals to give -15dBW output power in both main-lobe directions. The input amplitude is different for different tapering and different number of antennas to yield the same output power giving possibility to compare the results. Only third order intermodulation products were considered here and the output amplitude of the two signals plus the two intermodulation components was then calculated using (5.26). Finally, the total signal in direction θ could the be calculated using (5.35).

The weighting functions for $N = 8$ is shown in Figure 5.21. The use of more antennas implies that a larger back-off can be used at each amplifier, thus the C/IM_3 is larger for $N_{ant} = 8$ than for $N_{ant} = 4$. The use of tapering increases the SLL from 12.6 dB in the uniform case to 29.5 dB in the Hanning case, but it increases the first null beamwidth (FNBW). This tapering also reduces the gain in the main-lobe and the back-off of each antenna must be decreased to compensate for the loss in array gain. This increases the intermodulation power of the amplifiers, especially for the center amplifiers which are driven at the smallest back-off. Table 5.1 can be used to choose a proper tapering and as a means for trading off between SLL, FNBW and C/IM$_3$. With a single antenna, the required input to the amplifier is -12 dBW. Increasing the number of antennas to N=4, and maintaining the received power at the mobile, the input power

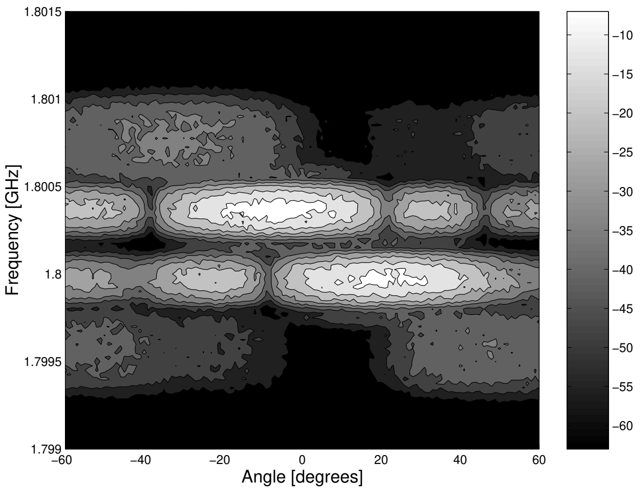


Figure 5.19: Normalized frequency-angle plot with saturated amplifiers and tapered weights. The intermodulation power is decreased by the tapering as compared to Figure 5.16.

can be decreased to -16.5 dBW, using the uniform taper distribution, or -15.5 dBW using the parabolic distribution. The parabolic distribution gives 8.6 dB lower side-lobes compared to the uniform distribution, at the expense of increased mainlobe beamwidth from 61° to 76°.

5.7 Conclusions

In this chapter, we investigated the usage of a not perfectly linearized MCPA in the downlink of a cellular system, by means of simulations and measurements. The conclusions can be tabulated as:

- By using more antennas in the antenna array at the base station the negative effects of nonlinearity is less pronounced. This is due to the distribution of the signal amplification over several amplifiers and combining of the signals in space. Each amplifier can thus operate at a lower output power.

- The nonlinearity will affect the radiation pattern in two ways. First, the nulls and the main beam pointing direction will change, which implies an increased interference level for mobiles using the same frequency channel in the vicinity of the basestation. We showed how

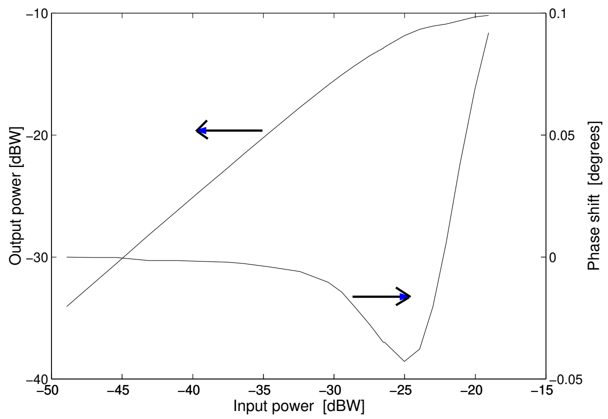


Figure 5.20: Measured AM/AM and AM/PM conversion characteristics for the amplifier used in the anechoic chamber measurements. The arrows points out the scale to be used with the respective curve.

this could be interpreted as a weight distortion. Second, intermodulation beams, or phantom beams, will be created that radiates the intermodulation products in undesired directions.

- The phantom lobes amplifies the radiated intermodulation products in some direction. We show that the amplification is at most $10\log_{10}(\mathrm{N})$ dB where N is the number of antennas in the array. This corresponds to coherent addition of the signal from all N antennas and we showed that this will be the case when a BMT is used. The BMT also gives a possibility to calculate the direction of the intermodulation beam. For a more general weighting, with null-steering, no specific direction of intermodulation can be specified, it depends on the particular beamformer weights, and the power amplification is less than $10\log_{10}(\mathrm{N})$ dB.

- We defined the total degradation to choose a power efficient back-off for the amplifier. With increased back-off, less intermodulation is generated and the CIR of the mobiles in the system increases. However, the power added efficiency of the MCPA decreases. We defined the total degradation function and found the function minimas corresponding to power efficient choices of the output back-off. We saw that with more antennas, a smaller back-off can be used and still maintain the CIR degradation within 2-3 dB. This choice of the OBO does not consider the maximum radiated intermodulation distortion power. So, after choosing the power efficient OBO, it must be verified that the radiated power in adjacent channels does not violate the specifications.

- Using an ABF as the Butler matrix, the beam directions and side-lobes are fixed. It is however possible to introduce a taper on the antenna weight magnitudes. This will make the side-lobes lower but some gain is lost in the main-beam, so the back-off of the amplifier must decrease, and an increased generation of intermodulation products follows.

Taper distr.	N_{ant}	FNBW	C/IM_3	SLL	P_{in}
Uniform	8	29°	52.8 dB	12.6 dB	-20 dBW
Cosine	8	40°	46.3 dB	22.2 dB	-17.7 dBW
Parabolic	8	35°	48.4 dB	17.9 dB	-19 dBW
Hanning	8	52°	39.7 dB	29.5 dB	-17.5 dBW
Uniform	4	61°	28.5 dB	11.0 dB	-16.5 dBW
Cosine	4	96°	23.2 dB	29.5 dB	-14 dBW
Parabolic	4	76°	28.6 dB	19.6 dB	-15.5 dBW
Hanning	4	87°	23.4 dB	24.5 dB	-14 dBW
Single	1	-	7.6 dB	-	-12 dBW

Table 5.1: Calculated effects of the tapering. Here, N_{ant}=number of antennas, FNBW=First Null Beam Width, C/IM_3=Carrier to third order intermodulation product ratio, SLL= highest side-lobe level, P_{in}=amplifier input power of each tone.

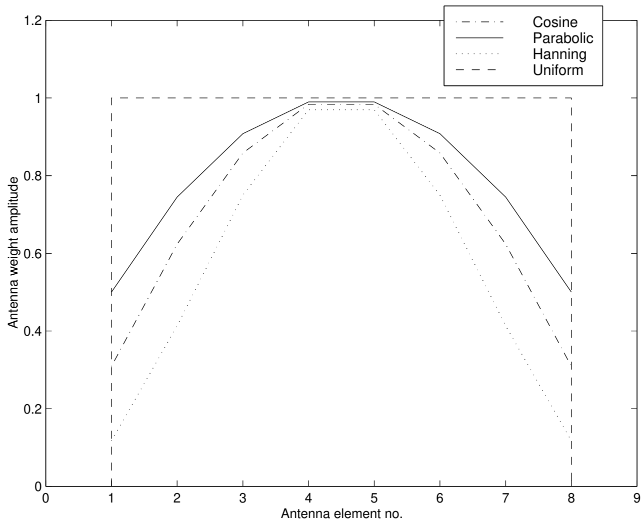


Figure 5.21: Different weight tapering functions

Chapter 6

Discussion and future work

The main results from the different chapters will be rewieved in this chapter and some general conclusions regarding implementation issues for smart antennas in wireless communications will be presented. Finally, some interesting future work will be presented in the last section.

6.1 Discussion

The measurements conducted on the adaptive antenna testbed shows promising results for the future of smart antennas in wireless systems. The testbed was able to suppress a 20 dB stronger co-channel interferer separated only 10 degrees in azimuth from the desired signal so effectively that the measured bit error rate was 0%. The scenario was however the simplest possible with two rays impinging on the array with zero angular spread and no multipath propagation. No quantitative measurements has been performed in a multipath scenario.

The testbed was evaluated in a multipath environment at a demonstration. The adaptive antenna array was placed on the ground in a yard with high building surrounding it, thus a signal environment with severe multipath components due to many wall reflections. The place was in central Kista outside of Stockholm and the occasion was a demonstration of the project in cooperation with Ericsson Radio Access AB. We successfully managed to connect two DCS-1800 mobile phones transmitting on the same frequency channel and timeslot and used the testbed as a receiver. The voice transmission over the two mobile phones progressed perfectly although the physical separation between the two users was reduced to only 0.2 meters.

This illustrates that with an array antenna, a multipath channel between the mobiles and the basestation actually helps the basestation to separate the users. It also shows that weight updating once per frame was sufficient for adapting to the changes in the channel.

In the testbed laboratory measurements we saw that the maximum achievable interferer suppression was limited at about 30 dB. To investigate what factors that affect this limit, we developed a model of an adaptive array antenna in Chapter 3 to consider the limitations in hardware. The limitations are the number of bits used by the ADC which also gives the dynamic range of the receiver. Also, as we use the hybrid-ABF with hardware weighting units, there is a finite accuracy in the weight settings. Another issue is the calibration performed prior to normal operation, which also has a finite accuracy. The calibration errors will change over time due to temperature drift in the receiving channels as was demonstrated with measurements.

The errors from finite accuracy in calibration and weight settings makes the interfering signal "leak" through the beamforming filter. Or, stated differently, the null depth in the radiation pattern becomes smaller, thus more power from the interfering signal appears in the beamformer output. The quantization noise from the ADC can be shown to be uncorrelated with the signal if the number of bits is larger than six and the signal is not saturating the ADC. Thus, the quantization noise power is added to the thermal noise power in the DSP. This could in fact be beneficial, acting as a regularizing noise in the algorithm, similar to the diagonal loading of the signal covariance matrix.

If however, the ADCs are saturated due to e.g. a strong transmitter close to the basestation, then the algorithm in the DSP collapses and the interference suppression capability falls rapidly. The dynamic range for the ADCs in an adaptive array system must be very large, up to 110 dB (in GSM) to be able to separate a user near the basestation from a distant user. So high dynamic range ADC is a key component in the adaptive antenna basestation.

In Chapter 4 we proposed two algorithms that operates in parallel with the adaptive antenna operation to mitigate performance degradation due to e.g. temperature drift in the receiving channels. We used the feedback signal from the output of the antenna beamformer to either track the temperature drift (direct method) and update the calibration data or to change the calculated weights in such a way that the effect of the drift is canceled. Both methods proved to work in simulations but it is argued that the direct approach is more robust to rapidly changing signal environments.

There exist uplink algorithms that don't need calibration. The SMI algorithm, used in the testbed is one example. The reason is that SMI uses only a temporal reference to calculate the weights, and no spatial direction to the arriving signals is calculated. The calibration we perform on the testbed is due to the analog beamformer which has to be matched to the weights calculated in the DSP. If the SMI algorithm is used in the uplink of a digital beamformer, then there would be no need for calibration of the antenna channels.

The downlink is however a more difficult story. The downlink transmission is of the spatial reference type, that is, we decide to transmit the signal in one or more specified directions. Thus the knowledge of the transfer function from the weighting unit to the antenna for each antenna channel is very important. If this antenna channel contains an active component such as an amplifier, then the calibration becomes very tricky. This is because the phase shift and gain of the amplifier depends on the amplitudes of the signals that enters the amplifier. Thus, constant gain and constant phase shift are very important characteristics for an amplifier to be used in the downlink channel of an adaptive antenna array.

In Chapter 5 the impact of using MCPAs in the transmit channel in a wireless cellular system was investigated. If the amplifier exhibits AM/AM and AM/PM conversion, then the problems discussed above will cause the nulls of the beamformer to be shifted. This will lead to an increased interference power for the co-channel mobiles. The second issue is the generation of intermodulation products in the MCPAs. These will also be radiated from the antenna array in a radiation pattern dependent on the beamformer weights and on the AM/AM and AM/PM conversion characteristics.

In Chapter 5 an expression for the radiation pattern of the intermodulation products is derived. The expression is verified by measurements in an anechoic chamber on a four element antenna array. To decrease the power in the generated intermodulation products, the back-off from amplifier saturation point can be increased. This will however lower the efficiency of the amplifier and most DC power is lost as heat. Thus there is a tradeoff between efficiency and generation of intermodulation products. We defined the total degradation as a function to be minimized to find a good and power efficient choice of the output back-off.

A seven cell system with antenna arrays at the base-station was simulated to investigate the effects of nonlinearities on the CIR of a mobile user in the center cell. We noted that by increasing the number of antennas in the array, the degradation due to the nonlinearity was decreased. With more amplifiers, the amplification is distributed which lower the requirements on each amplifier. With more antennas the system is also able to suppress more co-channel users, due to the increased degrees of freedom. It could also be seen from the total degradation curves that with two antennas the output back-off per carrier should be chosen to 12 dB but with eight antennas it could be decreased to 5 dB per carrier. This leads to a large increase in power amplifier efficiency.

6.2 Future work

We bring this thesis to an end by suggesting some issues of interest as future research topics:

- The calibration issues in smart antennas is a challenging area where not much work has been done. This will become increasingly important when the wideband-CDMA standard is implemented because it is important to come up with methods to calibrate the antenna front end over a wide bandwidth, to be able to fully utilize the capabilities of adaptive antenna arrays.

- A relating subject is how to model the nonlinearities of the receiving and transmitting front end, to use this in smart antenna simulation tools, and to better design the calibration method. The modeling can be seen as a nonlinear system identification problem and should incorporate mutual antenna coupling, instabilities and noise conversion in local oscillators, gain and phase sensitivity of amplifiers and mixers to power supply voltages, and so on.

- The simulations performed in Chapter 5 can be extended to incorporate multipath channel models. A specific modulation format can be specified and the bit error rate can then be used as a performance measure. The SAF method will be helpful in reducing the simulation time.

- The on-line calibration algorithms described in Chapter 4 can be implemented in the adaptive array antenna testbed described in Section 2.2 to verify its performance with experiments. The simulation model can also be extended to multipath scenarios, to verify the statement of robustness of the direct approach.

Appendix A

The Butler Matrix

In this appendix, the theory of the Butler matrix is explained more thoroughly. The Butler matrix was developed as a multiple beam feed network for a phased array antenna. It is attributed to Butler [55], but it was independently discovered by Shelton [56]. It uses 3-dB directional couplers, or hybrid junctions along with fixed phase shifters to form N beams from an N element array where N is an integer expressed as some power of 2, that is $N = 2^p$. The 3-dB directional couplers is a four port junction that divides the power equally on two ports with $+90°$ phase shift and no power is present at the fourth port. Similarly, a signal entering port four will be equal-power divided between the two ports with -90° phase difference and no signal at port one.

For a N-element antenna array, $(N/2)\log_2(N)$ directional couplers and $(N/2)(\log_2(N) - 1)$ phase shifters are required to create the Butler matrix. The Butler matrix is theoretically lossless, but there will always be a finite insertion loss, due to the inherent losses in directional couplers, phase shifters and transmission lines. There is no theoretical limit of the bandwidth, but due to the hardware, the bandwidth is usually limited, to about 30%.

The normalized input-output relationship can be written as

$$F_r = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} f_l \exp\left[-j\frac{2\pi lr}{N}\right] \tag{A.1}$$

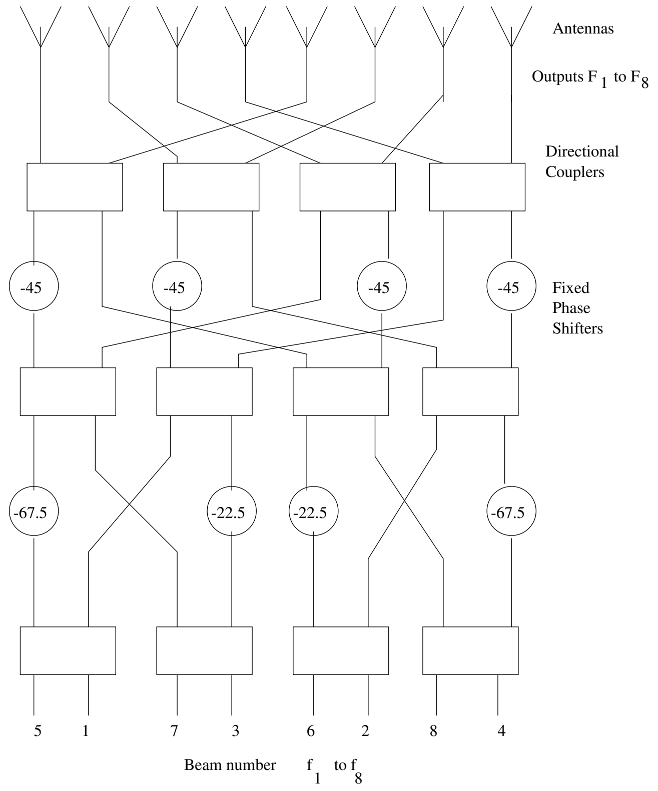


Figure A.1: Eight-element Butler matrix for beamforming

and

$$f_l = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} F_r \exp\left[j\frac{2\pi lr}{N}\right] \tag{A.2}$$

where f_l is the signal at the lth input port and F_r is the signal at the rth output port, connected to antenna r. Note the similarities with the fast Fourier transform (FFT). The fact is that the Butler matrix had existed a few years before the FFT appeared in 1965.

If the Butler matrix is connected to an antenna array, then the matrix will act so that the array will have a uniform amplitude distribution and constant phase difference between neighboring elements. This will then generate the beams depending on which input port is used. The direction of the beam depends on the wavelength of the transmitted signal, and will thus squint with frequency. If port one is used, then all other ports are grounded, the output of antenna r is

$$F_r = \frac{1}{\sqrt{N}} f_1 \exp\left[-j\frac{2\pi r}{N}\right] \tag{A.3}$$

and the radiation pattern in direction θ is

$$U(\theta) = \sum_{r=1}^{N} F_r(t) e^{-jr\frac{2\pi d}{\lambda}\sin(\theta) - j\frac{2\pi r}{N}} \tag{A.4}$$

where we assumed iso-tropical radiation patterns. This expression will have a maxima when the exponent is unity, given by

$$\frac{2\pi d}{\lambda}\sin(\theta_{max}) + \frac{2\pi r}{N} = 0 \tag{A.5}$$

Here d is the inter-element spacing (linear uniformly spaced array is assumed) and λ is the wavelength. To maintain the same direction of maximas when the frequency is changed, the distance d must be changed to maintain the quotient d/λ constant.

A plot of the radiation pattern from the Butler matrix are given in Figure 5.6 and in Figure 1.2.

Appendix B

Calculation of weights for downlink transmission

In this appendix we derive the downlink weight vector from the spatial signature of the uplink data. This method was presented by Zetterberg in his PhD thesis [49]. We assume that the Gaussian Angle of Arrival (GAA) model is valid for the received uplink data. The GAA method assumes that the received signal power is P and is Gaussian distributed with respect to the azimuth angle θ with mean θ_r and standard deviation σ_r.

If a uniform linear array with N antenna elements is used in the uplink, and if the Gaussian angle of arrival (GAA) model applies, then θ_r and σ_r can be estimated from the received data by the method in [57]. This information can then be used to calculate the downlink weights. If one assumes no multipath reflections or delay spread, for simplicity, then the channel between base station k and mobile i can be described by the N dimensional complex vector $\mathbf{v}_{k,i}$ with zero mean elements. The covariance matrix of $\mathbf{v}_{k,i}$ is denoted $\mathbf{R}_{k,i}$ and under the GAA assumption, $\mathbf{R}_{k,i}$ can be written

$$\mathbf{R}_{k,i} = G_{k,i}\mathbf{R}(\theta_{k,i}, \sigma_{k,i}) \tag{B.1}$$

where $\theta_{k,i}$ is the azimuth angle between basestation k and mobile i, $\sigma_{k,i}$ is the standard deviation of the Gaussian angular spread and $G_{k,i}$ is the path loss between basestation k and mobile i. The path loss includes the individual antenna element gain, slow fading and the $r_{k,i}^{-\gamma}$ dependence, where $r_{k,i}$ is the base-mobile distance and γ is the path loss slope.

The beamformer weights for basestation k are calculated by using the *Summed Inverse Interference to Carrier Ratio Minimizing Beamformer* (SICR) introduced in [49]. Assume that mobile k is connected to basestation k. The total number of detected mobiles, as seen from a basestation is Q. Transmitting from basestation k with the weights $\mathbf{w}_k$ yields a received signal power P at an arbitrarily mobile in the system, i as

$$P_{k,i} = \mathbf{w}_k^H \mathbf{R}_{k,i} \mathbf{w}_k \tag{B.2}$$

Here H denotes Hermitian transposition. By assuming that the array gain against the desired mobile is unity,

$$\mathbf{w}_k^H \mathbf{R}(\theta_{k,k}, \sigma_{k,k}) \mathbf{w}_k = 1 \tag{B.3}$$

and thus $P_{k,k} = G_{k,k}$. We now want to calculate the downlink weights for base-station k. Assume that only base-station k creates interference at mobile i. The mean carrier to interference ratio (CIR) at mobile i can then be written as

$$CIR_i = \frac{\mathbf{w}_i^H \mathbf{R}_{i,i} \mathbf{w}_i}{\mathbf{w}_k^H \mathbf{R}_{k,i} \mathbf{w}_k} = \frac{G_{i,i}}{G_{k,i} \mathbf{w}_k^H \mathbf{R}(\theta_{k,i}, \sigma_{k,i}) \mathbf{w}_k} \tag{B.4}$$

Then, a criterion for choosing $\mathbf{w}_k$ can be

$$\mathbf{w}_k = \arg\min_{\mathbf{w}} \left\{ \sum_{i=1, i \neq k}^{Q} CIR_i^{-1} \right\} \tag{B.5}$$

Thus $\mathbf{w}_k$ is chosen to minimize the sum of interferer to carrier ratio of the detected mobiles in the system. Defining the matrix $\mathbf{Z}$ as

$$\mathbf{Z} = \sum_{i=1, i \neq k}^{Q} \frac{G_{k,i}}{G_{i,i}} \mathbf{R}(\theta_{k,i}, \sigma_{k,i}) \tag{B.6}$$

yields

$$\mathbf{w}_k = \arg\min_{\mathbf{w}_k} \left\{ \mathbf{w}_k^H \mathbf{Z} \mathbf{w}_k \right\} \tag{B.7}$$

and it´s solution is given by the generalized eigenvalue problem

$$\mathbf{R}(\theta_{k,k}, \sigma_{k,k}) \mathbf{e} = \lambda \mathbf{Z} \mathbf{e} \tag{B.8}$$

Finally, the vector $\mathbf{w}_k$ is found by normalization to give unity array gain against the desired mobile as

$$\mathbf{w} = \frac{\mathbf{e}}{\sqrt{\mathbf{e}^H \mathbf{R}(\theta_{k,k}, \sigma_{k,k}) \mathbf{e}}} \tag{B.9}$$

where $\mathbf{e}$ is eigenvector corresponding to the largest eigenvalue λ_{max} in (B.8). For further details on the downlink algorithm refer to [49].

Note that this solution in not optimal in any sense. In practice, interference at the ith mobile will come from more than basestation k.

Bibliography

[1] Robert H. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, 1999.

[2] H.Aurodaki and K.Bandelow, "Effects of introducing adaptive antennas into existing gsm networks," in *VTC 99*, Houston, Texas, USA, 1999, pp. 670–674.

[3] TSUNAMI partners, "Algorithms and antenna array recommendations,", Technical Report Part 1, ERA, 1997.

[4] TSUNAMI partners, "Adaptive antenna performance in mobile systems,", Technical Report Part 2, ERA, 1998.

[5] W.A. Sandrin, "The Butler matrix transponder," *Comsat technical review*, vol. 4, no. 2, pp. 319–345, 1974.

[6] C.Tidestav and E.Lindskog, "Bootstrap equalization," in *IEEE International Conference on Universal Personal Communications (ICUPC'98)*, Florence, Italy, 1998.

[7] W.F. Gabriel, "Adaptive arrays- an introduction," *Proceedings of the IEEE*, vol. 64, pp. 239–273, 1976.

[8] Merril I. Skolnik, *Introduction to Radar Systems.* Auckland: McGraw Hill, 1980.

[9] S.Andersson et.al., "Ericsson/Mannesmann GSM field trials with adaptive antennas," in *Proceedings of Vehicular Technology Conference*, Phoenix,USA, 1997, pp. 1587–1591.

[10] L.C. Godara, "Applications of antenna arrays to mobile communications, Part I and II," *Proceedings of the IEEE*, vol. 85, no. 7,8, pp. 1031–1060,1195–1245, 1997.

[11] J.Litva and T.Lo, *Digital Beamforming in Wireless Communications.* Boston: Artech-House Publishers, 1996.

[12] JR. R.T. Compton, *Adaptive Antennas.* New Jersey: Prentice-Hall, Inc, 1988.

[13] R.Monzingo and T.Miller, *Introduction to Adaptive Arrays.* New York: John Wiley and Sons, 1980.

[14] Hudson J.E., *Adaptive Array Principles.* New York: Peter Peregrinus Ltd, 1981.

[15] R. Nitzberg, "Effect of errors in adaptive weights," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 12, no. 3, pp. 369–373, 1976.

[16] R. Nitzberg, "Computational precision requirements for optimal weights in adaptive processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 16, no. 4, pp. 418–425, 1980.

[17] J. Hudson, "The effects of signal and weight coefficient quantisation in adaptive array processors," in *Proceedings of the NATO Advanced Study Institute of Signal Processing*, Reidel, Dordrecht, Netherlands, 1977, pp. 423–428.

[18] Y.Karasawa T.Takahashi, I.Chiba, "The required resolution of A/D for null beamforming in a DBF antenna," in *IEEE International Symposium Digest Antennas and Propagation*, New York, USA, 1994, pp. 128–131.

[19] L. Godara, "The effect of phase-shifter errors on the performance of an adaptive antenna-array beamformer," *IEEE Journal of Oceanic Engineering*, vol. 10, pp. 278–284, 1985.

[20] L. Godara, "Error analysis of the optimal antenna array processors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, pp. 395–409, 1986.

[21] L.R. Burgess J. Bull, M.A. Arnao, "Hypersensitivity effects in adaptive antenna arrays," in *IEEE Symposium Digest. Antennas and Propagation*, 1990.

[22] L.J. Griffiths D. Feldman, "A constraint projection approach for robust adaptive beamforming," in *Procedings of International Conference on Acoustics, Speech and Signal Processing*, New York, USA, 1991, pp. 1381–1384.

[23] L. Pettersson, "Adaptive beamforming with imperfect arrays: pattern effects and their partial correlation,", Technical Report RL-TR-92-341, Rome Laboratory, 1992.

[24] B.D. Carlson, "Covariance estimation and diagonal loading in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 4, pp. 397–401, 1988.

[25] H.Steyskal, "Array error effects in adaptive beamforming," *Microwave journal*, vol. 34, pp. 101–112, 1991.

[26] M.Beach G.Tsoulus, J.McGeehan, "Space division multiple access (SDMA) field trials. Part 2: Calibration and linearity issues," *IEE Proceedings on Radar, Sonar and Navigation*, vol. 145, no. 1, pp. 79–84, 1998.

[27] G.Tsoulos and M.Beach, "Calibration and linearity issues for an adaptive antenna system," in *Proceedings of Vehicular Technology Conference*, Phoenix, USA, 1997, pp. 1597–1600.

[28] M.Beach S.M.Simmonds, "Downlink calibration requirements for the TSUNAMI(II) adaptive antenna testbed," in *Proceedings of the Ninth International Symposium on Personal, Indoor and Mobile Radio Communications*, Boston,USA, 8-11 September 1998.

[29] K.G. Johannsen, "Scan beam antenna intermodulation improvement due to spatial dispersion," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 23, pp. 543–557, 1987.

[30] M.Beach H. Xue, R. Davies and J. McGeehan, "Linearity considerations in adaptive antenna array applications," in *Proceedings of the Sixth International Symposium on Personal, Indoor and Mobile Radio Communications*, Toronto, Canada, 27-29 September 1995, pp. 682–686.

[31] J.R. Mosig S.L. Loyka, "Nonlinear modeling and simulation of active array systems," in *COST 260 MC and WG meeting*, Wroclaw, Poland, 1999.

[32] Sicsik-Pare' M. Coutiere, I, "Computed simulations of non-linearity in an satellite's active antenna system failures and degraded performance," in *COST 260 MC and WG meeting*, Dubrovnik, Croatia, 1997.

[33] M. Landing H. Andersson, "Adaptive antenna for mobile telephone systems,", Technical report, Signals and Systems, Uppsala University, Sweden, 1997, Technical Licentiate Thesis in Electronics.

[34] P.Mogensen et.al., "Preliminary results from an adaptive antenna array testbed for GSM/UMTS," in *Proceedings of Vehicular Technology Conference*, Phoenix,USA, 1997, pp. 1592–1596.

[35] H-C. Lin, "Spatial correlations in adaptive arrays," *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 2, pp. 212–223, 1982.

[36] M.Wax and Y.Anu, "Performance analysis of the minimum variance beamformer," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 928–937, 1996.

[37] S.L. Wilson M.W. Gantz, R.L. Moses, "Convergence of the SMI algorithms with weak interference," *IEEE Transactions on Antennas and Propagation*, vol. 38, no. 3, pp. 394–399, 1990.

[38] Shatskii Voloshina, Gabrielyan, "The discretization effect of phase control on the noise-to-signal ratio in adaptive arrays," *Radioelectronics and Communication Systems*, vol. 38, pp. 44–45, 1995, Translated from: Izvestiya Vysshikh Uchebnykh Zavedenii Radioelektronika vol.38, no.7; 1995; p.67-69.

[39] R. Davis and P. Sher, "Quantisation noise in adaptive weighting networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 20, pp. 547–559, 1984.

[40] J.Strandell, M. Wennström, A.Rydberg, and T.Öberg, "Experimental evaluation of an adaptive antenna for a TDMA mobile telephone system," in *Proceedings of the Eight International Symposium on Personal, Indoor and Mobile Radio Communications*, Helsinki, Finland, 1-4 September 1997, pp. 79–84.

[41] J.Strandell et.al., "Design and evaluation of a fully adaptive antenna for telecommunication system," in *Antenn 97*, Gothenburg,Sweden, 1997, pp. 357–366.

[42] D.W.Bennett P.B. Kenington, "Linear distortion correction using a feedforward system," *IEEE Transactions on vehicular technology*, vol. 45, no. 1, pp. 74–81, 1996.

[43] R.Reggiannini A.N.D'Andrea, V.Lottici, "RF power amplifier linearization through amplitude and phase predistortion," *IEEE Transactions on communications*, vol. 44, no. 11, pp. 1477–1484, 1996.

[44] P.B.Kenington D.W.Bennett, R.J.Wilkinson, "Determing the power rating of a multicannel linear amplifier," *IEE Proceedings on Communications*, vol. 142, no. 4, pp. 274–280, 1995.

[45] R.J.Wilkinson D.W.Bennett, P.B.Kenington, "Distortion effects of multicarrier envelope limiting," *IEE Proceedings on Communications*, vol. 144, no. 5, pp. 349–356, 1997.

[46] W.H.Tranter K.W. Schneider, "Efficient simulation of multicarrier digital communication systems in nonlinear channel environments," *IEEE Journal on selected areas in communications*, vol. 11, no. 3, pp. 328–339, 1993.

[47] O.Shimbo, "Effects of intermodulation, AM-PM conversion, and additive noise in multicarrier TWT systems," *Proceedings of the IEEE*, vol. 59, no. 2, pp. 230–239, 1971.

[48] S.L.Loyka, "Characteristics of a receiving intermodulation channel of active array antennas," *International journal of Electronics*, vol. 80, no. 4, pp. 595–602, 1996.

[49] P. Zetterberg. *Mobile Cellular Communications with Base Station Antenna Arrays: Spectrum Efficiency, Algorithms and Propagation Models*, PhD thesis, Royal Institute of Technology, Stockholm,Sweden, 1997.

[50] Version 4.2.1, "GSM/REC 03.03,", 1994.

[51] Version 4.22.1, "GSM/REC 05.08 (phase 2),", 1998.

[52] A.J.Cann, "Nonlinearity model with variable knee sharpness," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 16, pp. 874–878, 1980.

[53] J.P.De Almeida E.Alberquerque O.Shimbo, L.N. Nguyen, "Modulation-transfer noise effects among FM and digital signals in memoryless nonlinear devices," *Proceedings of the IEEE*, vol. 74, no. 4, pp. 580–597, 1986.

[54] S.Drabowitch, A.Papiernik, H.Griffiths, J.Encinas, and B.L.Smith, *Modern Antennas*: Chapman and Hall, 1998.

[55] J.L. Butler, *Digital matrix and intermediate frequency scanning, Scanning Antennas*, vol. 3: Academic Press, 1966.

[56] J.P Shelton and K.S. Kelleher, "Multiple beams from linear arrays," *IRE Transactions on antennas and propagation*, vol. 9, pp. 154–161, 1961.

[57] B Ottersten T Trump, "Maximum likelihood estimation of nominal direction of arrival and angular spread using an array of sensors," *Signal Processing*, vol. 50, no. 1-2, pp. 57–69, 1996.