UNIVERSITY OF CALIFORNIA

Santa Barbara

Energy Efficient Wireless Communication using Distributed Beamforming

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy
in Electrical and Computer Engineering

by

Raghuraman Mudumbai

Committee in charge:

Professor Upamanyu Madhow, Chair

Professor Jerry Gibson

Professor Joao Hespanha

Professor Kannan Ramchandran

Professor Kenneth Rose

December 2007

The dissertation of Raghuraman Mudumbai is approved.

Jerry Gibson

Joao Hespanha

Kannan Ramchandran

Kenneth Rose

Upamanyu Madhow, Committee Chair

October 2007

Energy Efficient Wireless Communication using Distributed Beamforming

by

Raghuraman Mudumbai

Dedicated to my late grandmother Mrs. Janaki, who wanted us to go places she had never been.

VITA

RAGHURAMAN MUDUMBAI

October 2007

EDUCATION

June 1998	B.Tech Electrical Engineering Indian Insitute of Technology Madras India
December 2000	M.S. Electrical Engineering Polytechnic University Brooklyn, NY USA
December 2007	Ph.D. Electrical and Computer Engineering University of California, Santa Barbara, CA Advisor - Prof. Upamanyu Madhow

EXPERIENCE

2003-2007	*Research Assistant* Department of Electrical and Computer Engineering University of California, Santa Barbara, CA
2001-2002	*Systems Engineer* Division Multiservice Networks L. M. Ericsson Telephone Co., Goleta, CA
2000	*Intern* Applied Research Division Telcordia Technologies (formerly Bellcore), Morristown, NJ
1998-2000	*Research Assistant* Department of Electrical and Computer Engineering Polytechnic University, Brooklyn, NY

PUBLICATIONS

Journal Articles

- R. Mudumbai, G. Barriac, and U. Madhow, "On the feasibility of distributed beamforming in wireless networks," IEEE Trans. on Wireless Commun., vol. 6, no. 5, pp. 1754-1763, May 2007

Conference Proceedings

- N. Shrivastava, R. Mudumbai, U. Madhow and S. Suri, "Target Tracking with Binary Proximity Sensors: Fundamental Limits, Minimal Descriptions, and Algorithms," Proc. ACM SenSys, Nov. 2006, pp. 251-264.

- R. Mudumbai, B. Wild, U. Madhow and K. Ramchandran, "Distributed Beamforming using 1 Bit Feedback: from Concept to Realization," (invited paper) Proc. of 44'th Allerton Conference on Communication Control and Computing, Sept 2006.

- R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Scalable feedback control for distributed beamforming in sensor networks," in Proc. IEEE Intl. Symp. on Inform. Theory (ISIT'05), Sept 2005, pp. 137-141.

- K. Bruvold, R. Mudumbai, and U. Madhow, "A QoS framework for stabilized collision channels with multiuser detection," Proc. of the IEEE International Conference on Communications (ICC'05), May 16-20 2005, pp. 250-254 vol. 1.

- R. Mudumbai, G. Barriac, and U. Madhow, "Spread-spectrum techniques for distributed space-time communication in sensor networks," in Proc. of the 38th Asilomar Conference on Signals, Systems and Computers (invited paper), Pacific Grove, CA, Nov. 7–10, 2004, pp. 908-912 vol. 1.

- R. Mudumbai, G. Barriac, and U. Madhow, "Optimizing medium access control for rapid handoffs in pseudocellular networks," Proc. of the IEEE 60'th Vehicular Technology Conference, Sept. 2004, pp. 1098-1102, vol. 2.

- B. Ananthasubramaniam, G. Barriac, R. Mudumbai, and U. Madhow, "Distributed space-time communication for sensor networks," in Proc. of the First International Symposium on Control, Communications and Signal Processing, 2004, pp. 195-198.

- G. Barriac, R. Mudumbai, and U. Madhow, "Distributed beamforming for information transfer in sensor networks," in IPSN'04: Proc. of the Third International Symposium on Information Processing in Sensor Networks, 2004, pp. 81–88.

ABSTRACT

Energy Efficient Wireless Communication using Distributed Beamforming

by

Raghuraman Mudumbai

We consider the use of distributed beamforming to improve the energy efficiency and transmission range of wireless networks. Under distributed beamforming, a number of wireless transmitters collaboratively transmit a common message signal in such way that their individual transmissions combine *coherently* (i.e. in phase) at the intended receiver. In essence, a set of distributed wireless nodes organize themselves as a *virtual antenna array*. As in beamforming from a conventional antenna array, highly directional transmissions can be achieved using a virtual array, and therefore substantial SNR gains can be realized compared to a network in which each node transmits independently to the receiver.

Distributed beamforming arises naturally from information theoretic analyses of multi-user channels and is an essential ingredient of capacity-achieving coding strategies in several cases. However these analyses are based on baseband models of the channel and as such involve some implicit assumptions. The two most important such assumptions are (1) synchronized carrier signals, and (2) known phase relationship between the transmitters. The main contribution of this thesis is a detailed analysis of the feasibility of these assumptions, and a design for a practical wireless system based on distributed beamforming that explicitly addresses these issues. This design is based on a simple iterative procedure for

beamsteering using receiver feedback. Carrier synchronization is achieved by using phase locked-loops to lock all the transmitters to a common reference signal broadcast by a designated *master* transmitter. We show that the SNR gains from beamforming are sensitive to the choice of PLL parameters, and examine this dependence in detail.

The feedback procedure for beamsteering works as follows: each transmitter independently makes a small random adjustment to its phase at each iteration, while the receiver broadcasts one bit of feedback, indicating whether the signal-to-noise ratio is better or worse after the adjustment. The transmitters keep the 'good' phase adjustments and discard the 'bad' ones, thus implementing a distributed ascent algorithm. We show that, for a broad class of distributions for the random phase adjustments, this procedure leads to asymptotic phase coherence with probability one. A simple analytical model, borrowing ideas from statistical mechanics, is used to characterize the progress of the algorithm, and to provide guidance on parameter choices.

One of the motivating applications for this thesis is the problem of communication in WSNs. Therefore the scalability and robustness properties of our virtual array are important considerations. We show that the convergence time of the beamforming algorithm is linear in the number of collaborating nodes, and also that the SNR gains are robust to noise and fading effects in the wireless channel.

Contents

List of Figures

BLANK PAGE

Chapter 1

Introduction

This dissertation presents the design and analysis of a method for energy-efficient wireless communication from a network of transmitters to a distant receiver using distributed beamforming. Distributed beamforming refers to a form of cooperative communication in which a number of transmitters agree upon a common message, and then transmit it such that their contributions combine *coherently* at the receiver. In effect the transmitters organize themselves as the elements of a *virtual antenna array*; beamforming requires that the signals from each transmitter combine constructively at the receiver to give a strong combined signal. Physically the stronger signal is because of the fact that a large fraction of the transmitted power is focused in the direction of the receiver just like in centralized phased-array antennas.

As a result of this constructive interference, beamforming leads to a factor of N gain in power efficiency, where N is the number of collaborating transmitters. Thus, if the power of each transmitter is fixed, then distributed beamforming leads to an N^2 gain in received signal-to-noise ratio (SNR): a factor of N gain due to increase in total transmit power, and a factor of N gain in power efficiency due

to increased directivity. Such gains, if feasible, could have significant implications for the design of wireless ad-hoc and sensor networks, which typically suffer from severe power constraints.

1.1 Overview

Distributed transmit beamforming arises quite naturally as a key ingredient of several classical results in network information theory. However, some of the implicit assumptions in the traditional baseband model of the channel turn out to be difficult to realize in practice. Specifically the two related assumptions of synchronized carrier signals, and known phase relationships between the different transmitters are much more challenging for a virtual array of independent transmitters than for a centralized antenna array. This is because, unlike a centralized array, each transmitter in a virtual array obtains its RF carrier signal from a separate free-running local oscillator. As a result, the transmitters have RF carrier signals with unknown frequency and phase offsets between them. Unless these offsets are compensated for, methods for channel estimation and beamsteering that are designed for centralized antenna arrays (e.g. using reciprocity) cannot directly be used for distributed arrays. This is, perhaps, the most important difference between centralized and distributed beamforming.

Carrier *frequency* synchronization can be achieved by locking all of the transmitters to a common reference signal. This suggests a *master-slave* architecture, where a designated master transmitter broadcasts a reference carrier signal, and the slave transmitters lock to the reference signal using phase locked-loops (PLLs). However because of unknown propagation delays in the master-slave channel, there is still an unknown (but fixed) phase offset between the synchronized carrier sig-

nals. It is possible to use other methods for frequency synchronization (such as injection locking), but they are also subject to the same limitation of unknown phase offsets.

In principle, if the geometry of the virtual array is known accurately, the master-slave propagation delays of all the transmitters can be estimated, and therefore also the unknown phase offsets. Any residual errors in positioning can be considered as a source of phase noise [1], and along with other sources of noise such as jitter limit the achievable SNR gains from beamforming. The effect of these errors becomes larger when the frequency of the carrier signal is large. Fortunately, it turns out that the SNR gains are highly robust to phase errors; even with moderately large errors on the order of 60°, it is possible to achieve upto 75% of the available SNR gains from beamforming. This robustness to phase error is what makes distributed beamforming a potentially feasible technique for wireless networks.

However determining the geometry of the network with sufficient accuracy is a challenging task; for instance at a frequency of 1 GHz, in order to keep the phase errors less than 60°, the localization error needs to be below approximately 5 cm. Traditional localization methods such as GPS are not capable of this level of accuracy, indeed the physical dimensions of the antenna may be larger than 5 cm, which would make it impossible to regard the antenna as a point source. In order to avoid these difficulties, we consider a calibration procedure [2], where each transmitter measures its phase offset from the designated master-transmitter and compensates for it explicitly. If the phase drift in the local oscillators is a random process with memory the phase errors can build up rapidly. This requires that the calibration process be periodically repeated to keep the carrier signals synchronized. A detailed examination of this process reveals a fundamental tradeoff

between the synchronization overhead and the achievable beamforming gain.

While such a calibration process provides useful analytical insights, its complexity motivates the search for a simpler implementation of distributed beamforming. In this spirit, we investigate a simple iterative procedure [3], based on feedback from the receiver, for achieving phase coherence, and show that this procedure provides a powerful method to satisfy the requirements for distributed beamforming. The basic idea behind the feedback algorithm is described as follows. Each transmitter adjusts its phase randomly at each iteration, with the receiver broadcasting one bit of feedback per iteration as to whether its SNR is better or worse than before. If it is better, the transmitters keep the previous phase perturbation, and if it is worse, they undo the previous phase perturbation. This randomized ascent procedure is repeated until the transmitters converge to phase coherence. This procedure is especially appealing because it avoids the previously mentioned difficulties in channel estimation due to the unknown phase offsets; by using SNR measurements, it completely removes the need for any explicit channel estimation procedure.

It can be shown that under mild conditions, this feedback algorithm converges asymptotically to perfect coherence *with probability one*, assuming that the wireless channels are static, and there is no noise or SNR estimation errors. When these idealized assumptions are relaxed, the algorithm reaches a steady-state tracking mode. If the noise and time-variations are not too large, the SNR gains in the steady-state can be close to the ideal case. Thus, the algorithm is robust to channel impairments, and is also scalable in the sense that the convergence time increases no faster than linearly in the number of transmitters. When the number of transmitters becomes large, the algorithm also admits to a statistical mechanical analysis, which can be used to optimize the rate of convergence and to obtain

insight into its dynamics.

All these characteristics make the feedback algorithm an attractive method of realizing a virtual antenna array in practice. Indeed, it was shown in a proof-of-concept prototype [4] that significant SNR gains are achievable using a very simple implementation based on these ideas.

1.2 System Model

Our goals in this dissertation, can be summarized as follows:

1. To motivate the idea of a *virtual antenna array*, and to show how it relates to various aspects of the theory and practice of wireless communication systems.

2. To study the feasibility of distributed beamforming for wireless networks, to identify the important design parameters and determine the performance limits in terms of achievable SNR gains from beamforming.

3. To describe an implementation of distributed beamforming using a master-slave architecture for carrier frequency synchronization, and a 1-bit feedback algorithm for beamsteering.

4. To develop a theoretical analysis of the feedback algorithm in order to study its convergence and scalability, to optimize its performance, and to study its implications for other problems in wireless network design.

We consider an idealized wireless network to focus on the issues relevant to our goals, and avoid extraneous details. This idealized network is illustrated in Fig. 1.1, and its salient features are described as follows.

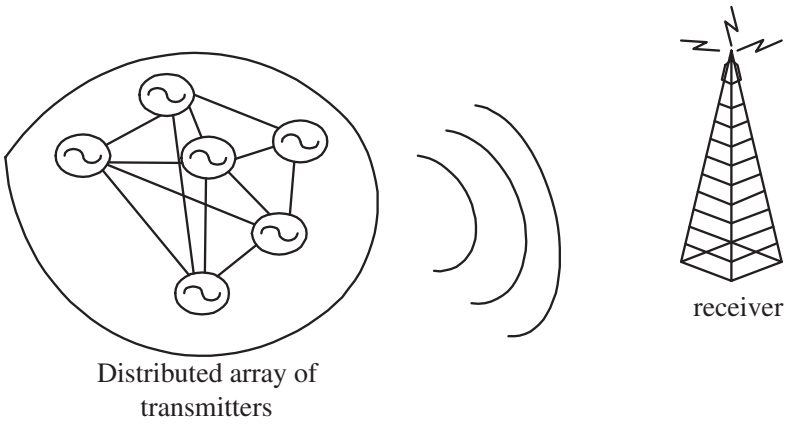


Figure 1.1. Idealized communication model for distributed beamforming.

1. The network consists of a set of N transmitters. The transmitters can communicate with each other wirelessly and also with the receiver (which is not considered part of the network). The locations of the transmitters is arbitrary, however the distance between any of the transmitters is small compared to the distance of the receiver from the network.

2. The transmitters seek to cooperatively communicate a common, baseband message signal $m(t)$ to the receiver. The message $m(t)$ is assumed to be known to all the transmitters. This may be arranged by a previous round of communication within the network.

3. The cost of "local" communication within the network, i.e. among the transmitters, is assumed to be negligible compared to the cost of communication with the receiver. Thus the cost of synchronizing the carrier signals of all the transmitters, and the cost of broadcasting the message signal $m(t)$ throughout the network are all assumed to be negligible. Each transmitter has a maximum power constraint for communication with the receiver.

This idealized network represents the most favorable conditions for distributed beamforming. In practice it approximates different networks to a greater or lesser degree, and it is necessary to extend our analysis to take our simplifying assumptions into account. The SNR gains from beamforming have to be balanced against the additional overheads that we have chosen to neglect here, such as the cost of synchronization and local communication. It is also necessary to compare our single-hop transmission scheme with a more general hybrid scheme that combines multi-hop routing with cooperative transmission. We briefly discuss some of these issues in Chapter 6.

1.3 Organization of this Dissertation

The rest of this dissertation is organized as followed. In Chapter 2, we present background information on topics related to cooperative wireless communications. Chapter 3 analyzes the requirements for distributed beamforming using a simple master-slave architecture for synchronization, and a reciprocity-based scheme for channel estimation and phase calibration. Chapter 4 presents a simple 1-bit feedback control algorithm for beamsteering, and derives an analytical model for its convergence. The effect of impairments such as phase jitter and time-variations on the feedback algorithm is examined, and a design for a practical implementation of distributed beamforming is presented in Chapter 5. Chapter 6 offers some remarks on the open issues and implications for future work.

Chapter 2

Background

Our work on collaborative beamforming was based on the idea of applying communication techniques designed for multi-antenna point-to-point wireless systems to wireless networks. The concept of *virtual antenna arrays* emerges naturally from traditional information theoretic models for the communication capacity of wireless networks. Yet these models suffer from serious limitations when applied to such virtual arrays, and indeed most of this thesis is devoted to finding explicit solutions to address these limitations. We now present a brief survey of areas relevant to our work: antenna arrays and wireless networking, and also more recent related work on cooperative communication in wireless networks.

2.1 Wireless Communication

Wireless communication can be properly said to have started with the invention of the wireless telegraph in the 1890s by Marconi and other pioneers. To be sure there were previous untethered communication systems. Even before Hertz's demonstration of electromagnetic waves, there were some attempts [5] to use near-

field induction effects for communication. Of course non-electric methods such as semaphores, and smoke-signals have been in use for a very long time. Nevertheless, it is clear that the telegraph represented a qualitatively new capability in terms of the amount of data, distances and the technical sophistication involved.

2.1.1 Early Wireless

Therefore wireless communication is a little more than a century old today. Within these 100 years or so, wireless communication started out as a discrete-time digital baseband system (the telegraph), however soon continuous-wave (CW) bandpass systems became dominant because it provided a way to avoid interference with other users, and also allowed continuous waveforms to be transmitted, audio being the most important. The "killer application" for this new technology turned out to be broadcast AM and FM radio, and later, television[1].

In contrast to these early applications, the ALOHA network at the University of Hawaii [7] represented the first implementation of a *packet radio* network with a decentralized and flat architecture i.e. mobile stations directly communicating with each other without the intervention of a Base Station. These ideas are extremely important, because they formed the basis for the ARPAnet [8], that has now grown to become the global Internet. For our purposes, this is especially important, because DARPA, the US government agency behind the ARPAnet, was encouraged by the success of the ARPAnet project to apply the same principles to pioneer a new type of network called the Distributed Sensor Network (DSN) [9].

[1]From reports from that period, it is clear that broadcast radio was an unprecedented social and cultural phenomenon, perhaps on a par with the Internet revolution of the 1990's in terms of its impact on society. According to an ad in *Scientific American* from 1922 [6]:

> The air is full of wireless messages every hour of the day. ... Famous people will talk to you, sing for you, amuse you. You don't have to buy a single ticket.

A DSN was conceived as a network of distributed, low-cost *sensing nodes* that collaborate with each other but are largely autonomous.

2.1.2 Wireless Sensor Networks

The early research efforts from the DSN project, today provides the intellectual foundation for research in sensor networks [9] but with some important modifications. Unlike the original conception of DSNs, wireless interconnection is now considered a defining feature for sensor networks. The sensors are visualized as miniature battery-driven computing devices; because of energy constraints, all computing, storage/retrieval, and especially communication operations are considered expensive [10], [11]. The sensors are deployed in large numbers; even though individual sensors are considered expendable, reliability is assured through *redundancy*. Therefore *scalability* and robustness (to sensor failures) are considered essential features for a sensor network. We use the term Wireless Sensor Network (WSN) to refer to a network with these characteristics.

The concept of a WSN as defined above motivates many interesting research problems in communication theory as well as in other related areas such as distributed computing and control. Traditionally communication networks have been designed with a focus on portability and rapid prototyping [12], using highly suboptimal models such as the collision model for interference in multiple-access channels. In recent years there have been a renewed effort to revisit these assumptions using information theoretical insights. The WSN model has proved to be well-suited for such theoretical inquiries, because in the idealized limit of large networks, details such as the network topologies can be statistically modeled.

In an influential contribution [13], Gupta et al showed that per-user bandwidth

in a non-hierarchical wireless network vanishes as network size increases i.e. flat wireless networks do not scale. Using an information theoretic analysis, they showed that as the network size increases[2], the interference from other transmitters also increases, leading to a decreasing per-node throughput. On the other hand, other authors have shown [14] that if multiple nodes cooperate to *control* the interference, the network becomes scalable. This can be interpreted as a form of distributed beamforming, and its impact on scalability was investigated in different contexts such as an ad-hoc network [15], and a relay network [16]. The idea of controlled interference comes from the early work on the capacity of multi-user channels (Section 2.3).

Separately, the idea of cooperative transmission [17] also emerged from the space-time coding techniques designed for multi-antenna systems which we survey next.

2.2 Antennas and Antenna-Arrays

From the early days of wireless, considerable effort has been spent on designing efficient structures for radiating electromagnetic waves i.e. antennas. Since all time-varying currents generate radiation, any arbitrary hunk of metal acts as an antenna when driven by an appropriate voltage or current source [18]. However a poorly designed antenna may dissipate most of the energy supplied to it as heat (high internal resistance), and/or it may have a large reactive impedance, so that high voltages are required to generate any appreciable amount of radiation (large Q-factor). In addition, for point-to-point (as opposed to broadcast) applications,

[2]There are actually two different but related models that have been used in the literature on network scaling; one model keeps the network area fixed while increasing the number of nodes i.e. increasing density, while the other model keeps density constant with increasing area. Both models are equivalent if an appropriate power-control mechanism is added.

it is usually desirable to have high directivity[3], so that the transmitted power is focused in the direction of the receiver. These three properties i.e. internal resistance, Q-factor[4] and directivity are the most important figures of merit for antennas.

It is also possible to view an antenna as a current distribution in a (3D) region of space. In this view, the overall electromagnetic fields radiated by the antenna can be considered as a superposition of the fields due to several infinitesimally small *current elements* that together constitute the antenna. It is important to note that electromagnetic influences also propagate within the antenna itself at the speed of light; if the driving signal varies slowly, the signal variations propagate themselves rapidly throughout the antenna, and the current distribution becomes almost uniform. In other words, if the spatial dimensions of an antenna are small compared to the wavelength of the signal, the current distribution is uniform throughout and therefore essentially trivial.

Putting these observations together leads to the concept of an antenna array: the radiation from any antenna can be emulated using a discrete grid of small current elements separated by distances on the order of a wavelength. This is analogous to replacing a bandlimited continuous-time signal by a sequence of its samples. The classical Yagi-Uda antenna [19] can be considered as the earliest example of an antenna array. However the term "antenna array" is usually reserved for an array of identical, regularly spaced elements, each of which can be *independently* driven by a voltage or current source (unlike in a Yagi antenna).

The greatest advantage of antenna arrays is their flexibility: by adjusting the

[3]We are mostly concerned with terrestrial-based communication, where (assuming the earth is flat!) all transmitters and receivers are coplanar, therefore we are interested in directivity in the azimuthal plane only, not in elevation.

[4]The Q-factor is also related to the bandwidth of the antenna, however our focus in this dissertation is on narrowband applications, where by definition, this is not an important factor.

gain and phase lag of the current in each element, a wide range of radiation patterns can be obtained, and the transmission can be steered electronically in any direction without physically moving any of its parts. Furthermore because the individual elements of the antenna are identical, the current distribution can be specified completely using only a discrete number of currents, one for each element; this allows us to rewrite the EM field equations as a system of linear equations of the element currents (the Method of Moments [18]) that can be efficiently solved on digital computers. Antenna arrays became popular during World War II, where they were used for radar applications. In the post-war period, they became popular for radio-astronomy, and were also adapted for problems such as Direction of Arrival estimation (DoA) and localization using acoustic signals [20]. Antenna arrays used for such applications are sometimes referred to as "smart antennas" or "adaptive antenna arrays" [21].

2.2.1 Distributed Antenna Arrays

So far it is not clear that there is any connection between antenna arrays and wireless networks, and indeed these subjects had little influence on each other until recently. Nevertheless, even though they were not thought of as networks, distributed arrays have been used in some applications for a long time. For instance, the so-called Very Large Array (VLA) is an antenna array used for radio-astronomy [22], that consists of 27 independent dish antennas each with a diameter of about 25 m, with the whole array mounted on railroad tracks that allows them to be deployed over a 23 mile distance. In essence, the array looks like a giant 23 mile wide antenna. This large aperture allows the array to achieve very high angular resolutions (upto 0.05 arcseconds or 0.000014°). Distributed acoustic arrays

i.e. an array of microphones [20] have also been used for localization applications. In all these applications, the distributed array acts as a receiver; it is possible to store the received signals throughout the array for offline processing. As we shall see later, using a distributed array for *transmitting* is much more difficult because the array has to be coordinated in real-time.

A key breakthrough in the use of antenna arrays came in the form of multi-input multi-output (MIMO) systems in the 1990's. Unlike previous multi-antenna applications, in MIMO, both transmitter and receiver are equipped with an multi-element antenna array, which allows them to simultaneously communicate *multiple streams* of data over the same frequency band. In essence there are multiple spatial degrees of freedom in MIMO systems over the same frequency spectrum; this not only increases the amount of data to be sent over the same bandwidth many times [23], but also provides *diversity* gains against stochastic fading i.e. when multiple spatial channels exist, the possibility of all of them simultaneously being in a deep fade becomes small.

Simple orthogonal "space-time" codes were developed [24] to realize these diversity gains in MIMO systems. Multiple spatially distributed nodes with single antennas could also achieve diversity gains by using the space-time coding techniques cooperatively. Under certain conditions, the benefits from cooperation were shown to outweigh the additional complexity involved in coordinating transmissions between multiple nodes [17], [25]. These analyses did not address synchronization issues in detail because they were based on the inherently baseband models of multi-user information theory. However it is well-known that carrier offsets cause some performance loss, and recent research [26] has focused on estimating these offsets and designing the receiver to degrade gracefully. The master-slave synchronization technique discussed in Chapter 3 offers an alterna-

tive solution to eliminate these offsets at the transmitters. We briefly comment on this possibility in Chapter 6.

2.2.2 MIMO Beamforming

Space-time diversity techniques exploit the statistical independence of the MIMO spatial channels; if the transmitter knows the channel gains to the receiver, it is possible to direct the transmissions from each antenna to minimize the interference between the MIMO spatial channels [27]. This is a generalization of the idea of beamsteering in phased-array antennas where the transmission is focused in a desired direction in space using the known geometry of the array. In the presence of multi-path fading, array geometry alone is no longer sufficient to determine the channel gains to a receiver. The most straightforward solution in this case, is to use training symbols to estimate the channel gains; this requires a feedback mechanism for the receiver to send its estimates of the channel gains back to the transmitter. This raises the important issue of the impact on capacity, of imperfections in the feedback channel. Such imperfections can be modeled in many different ways, such as quantized feedback where the channel information is conveyed using only a finite number of bits, or statistical feedback where only the fading averages such as the mean or covariance of the channel is known at the transmitter. It has been shown [28], [29] that gains from beamforming are substantial even with such limitations on channel knowledge. Similarly we find that significant SNR gains can be realized from distributed beamforming even in the presence of moderately large phase errors.

One ingenious method for channel estimation is to use the reciprocity property of electromagnetic channels [30] to estimate the forward-link channel indirectly us-

ing measurements on the reverse link and vice-versa. This requires that the same frequency band is used for both forward and reverse links i.e. time-division duplexing, however the same concept has also been extended to *wideband* frequency-division duplexed systems [31]. All of these techniques make it unnecessary to know the geometry of the antenna array, and are therefore of particular value in designing distributed arrays. We consider a reciprocity-based channel estimation procedure in Chapter 3, and an iterative feedback-based procedure in Chapter 4.

2.3 Information Theory of Multi-user Channels

In 1948, Shannon published a landmark paper [32] introducing the new mathematical area of *information theory*. Before this time, communication system design was mainly guided by intuition and rules of thumb and theory seriously lagged behind engineering practice. One of Shannon's subtler contributions was in his model of a communication system, which is divided into a source (modeled as a random process), a channel (a probabilistic medium) and a receiver (a decision device) (see Fig. 2.1). Today this model is considered so obvious that it is taken for granted. Very soon it was extended to other cases such as the multi-user [33] and stochastic channels [34] which are highly relevant to wireless communication systems with fading and interference.

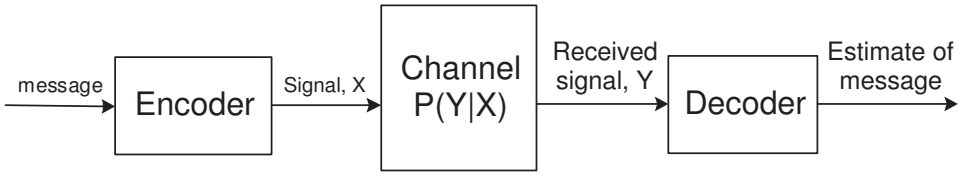


Figure 2.1. Shannon's model of a generic communication system.

The basic multi-user channel is illustrated in Fig. 2.2, and its relationship

to Fig. 2.1 is clear. While this model yields elegant theoretical insights, many objections can be raised against it. The main objection is that the above models are not reasonable representations of real multi-user channels, the most important one being the wireless channel. For instance, consider the Gaussian relay channel [35] shown in Fig. 2.3. This channel has played a very important role in the development of multi-user information theory. Its importance comes from the fact that it combines all three of the fundamental modes of communication in networks: direct transmission to receiver, multi-hop routing, and distributed coding.

The relay channel of Fig. 2.3 is fundamentally a *baseband* model, where the received signal is a fixed linear combination of the transmitted *baseband* signals from the source and the relay with additive noise as the only impairment. In fact this particular multiple-access scheme can be considered as a distributed beamforming scheme. The capacity-achieving schemes for WSNs surveyed in Section 2.1.2 are simply extensions of this relay channel to the whole network.

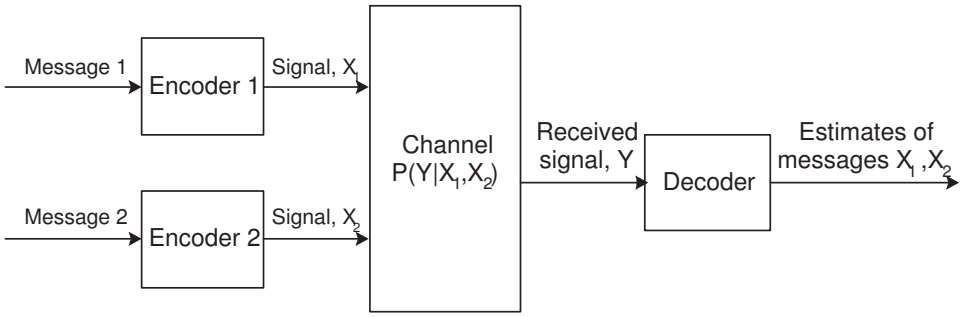


Figure 2.2. Multiple-access communication channel.

Unfortunately in a real multiple-access wireless network, the presence of additional inputs does not merely cause interference between the baseband signals, it can also introduce memory and time-variations into the channel. As we shall see later, the main source of these effects is the possibility of unsynchronized

carrier signals and unknown phase offsets between the different inputs. As we demonstrate in this dissertation, when carefully designed, however, a real wireless network can be made to mimic the theoretical multiple-access channels of Figs. 2.2 and 2.3.

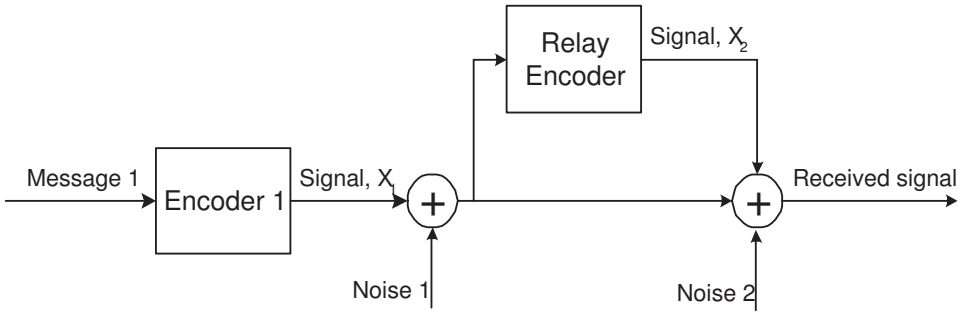


Figure 2.3. The Gaussian relay channel.

2.4 Recent Work

Following the pioneering work of [17], the idea of virtual antenna arrays and distributed beamforming has been studied extensively in recent years [36]. Beamforming continues to be the basis for several studies in network information theory [37], [38]. In [39], the statistics of the antenna pattern from a distributed array is examined; in particular their characterization of the *sidelobes* of the antenna pattern gives information about the interference from the distributed array, and thus complements the work described in this dissertation (which focuses exclusively on the mainlobe). A method of synchronization suitable for coherent communication was described in [40], where each transmitter separately synchronizes with the receiver so that it is phase aligned with all of the other transmitters. This is similar to the so-called "closed-loop" synchronization described in [39], and

can be considered as a special case of the *master-slave* synchronization presented in Chapter 3.

On another note, our feedback procedure for beamsteering can be considered as a distributed implementation of a stochastic approximation algorithm. This is a well-known class of mathematical algorithms [41], where a randomized search procedure is used to find the roots or extrema of a function that is difficult to characterize analytically. This technique was first used by Robbins-Monro [42] in 1952, and has since been recognized as a powerful technique with applications in many areas. In closely related work, a stochastic gradient algorithm was recently proposed [43] for downlink beamforming from a cellular Base Station. Other authors [44] have developed interesting variations on our algorithm for beamsteering to optimize its performance under specific conditions.

Chapter 3

A Master-Slave Architecture for Synchronization

We consider the communication system illustrated in Fig. 1.1, where a cluster of energy-constrained wireless transmitter nodes communicating with a distant receiver. The main assumption is that local communication among the cooperating transmitters is inexpensive compared to transmitting to the receiver.

In a traditional (centralized) multi-antenna transmitter, one way to perform beamforming is by exploiting reciprocity to estimate the complex channel gains to each antenna element. These channel gains are computed in a centralized manner with reference to a RF carrier signal supplied by a local oscillator. However, in a distributed setting, each transmitter has separate RF carrier signals supplied by separate local oscillator circuits. These carrier signals are not synchronized *a priori.* In the absence of carrier synchronization, it is not possible to estimate and pre-compensate the channel phase responses so as to assure phase coherence of all signals at the receiver. Accordingly, we consider a *master-slave* architecture, where a designated *master* transmitter coordinates the calibration and synchronization

of the carrier signals of the other *slave* transmitters, so that reciprocity can be used to estimate the channel gains to the receiver.

In this way, the transmitters use cheap local communication between the *master* and the *slave* transmitters to emulate a centralized antenna array, and to avoid the need for coordinating with the distant receiver. In Chapter 4, we show that by allowing a limited amount of coordination with the receiver in the form of 1-bit SNR feedback, we can use an alternative method of beamsteering that eliminates the need for explicit phase calibration and channel estimation, and is therefore simpler to implement in practical networks. In this chapter, we consider the more direct reciprocity-based method to gain insight into the synchronization process and to characterize the achievable gains from beamforming.

The achievable SNR gains are primarily limited by the residual phase errors from the synchronization process. We examine the different possible sources of phase error in detail. We observe that phase noise in practical oscillators causes them to drift out of synchronization, therefore, it is necessary for the master transmitter to resynchronize the slaves periodically. This, combined with the duplexing constraints of the wireless channel (i.e. it is not possible to transmit and receive on the same frequency simultaneously), reveals a fundamental tradeoff between synchronization overhead and beamforming gain. We quantify this tradeoff using a stochastic model for the internal phase noise of oscillators.

To get an idea about the effect of phase errors, consider the simple example of two equal amplitude signals from two transmitters combining at the receiver with relative phase error of δ. The resulting signal amplitude is given by $\left|1 + e^{j\delta}\right| = 2\cos\left(\frac{\delta}{2}\right)$. Even a significant phase error of $\delta = 30^\circ$ gives a signal amplitude of 1.93, which is 96% of the maximum possible amplitude of 2.0 corresponding to the zero phase error case. More generally, we show that it is possible to achieve

SNR of up to 70% of the maximum with moderately large phase errors on the order of 60°.

3.1 Analysis of Beamforming Gain

We consider a cluster of N transmitters, communicating a common (baseband) message signal $m(t)$ to a distant receiver,

We start by assuming that all of the transmitters modulate a common message signal $m(t)$, with a carrier signal at frequency f_c. Each transmitter derives its carrier signal from a separate local oscillator, therefore, the carrier signals of the different transmitters are not initially synchronized to each other. Therefore, an explicit synchronization process is necessary to get carrier signals at frequency f_c. Before we present our algorithm for carrier synchronization, we show using a simple analysis that beamforming gains are robust to moderately large phase errors. For this section we assume that the synchronization algorithm allows each transmitter to obtain synchronized carrier signals at frequency f_c and an estimate of their own channel gain to the receiver. Using this the transmitters can cooperatively transmit the message $m(t)$ by beamforming, just like a centralized antenna array. The resulting received signal $r(t)$ is the superposition of the channel-attenuated transmissions of all the transmitters and additive noise $n(t)$:

$$r(t) = \Re\Big(m(t)e^{j2\pi f_c t}\sum_i \big|g_i h_i\big|e^{j\phi_i(t)}\Big) + n(t) \tag{3.1}$$

where g_i is the pre-amplification and $\phi_i(t)$ is the cumulative phase error from the synchronization process for slave i. Under a constraint on the total transmit power, the optimum $|g_i| \equiv |h_i|$.

The phase errors have two effects on the received signal: a reduction in the

average SNR, and a time-dependent fluctuation of the received phase. The latter effect may cause limitations in the coherent demodulation of digital signals. However, there are several methods, e.g. differential modulation, available to deal with these fluctuations provided the time-variations are not too rapid. We concentrate on the first effect i.e. the reduction in average SNR. This is appropriate for power-limited sensor networks, where the feasible communication range is limited by SNR. For simplicity of notation, we suppress the time-dependence of $\phi_i(t)$ in this section.

We model the channel coefficients h_i, $i = 1...N$, as independent circularly symmetric complex normal random variables with zero mean and unit variance, as denoted by $h_i \sim CN(0,1)$. This can be considered as an extreme case of a non-LoS multipath wireless channel. We also assume that the phase errors ϕ_i are independent and identically distributed random variables for all the transmitters i.

Equation (3.1) motivates as our figure of merit, the beamforming gain defined as the normalized received power P_R, given that the total transmit power is $P_T = 1$:

$$P_R = \frac{1}{N}\left\|\sum_{i=1}^{N}|h_i|^2 e^{j\phi_i}\right\|^2 \tag{3.2}$$

Proposition 1: $\frac{1}{N}P_R \rightarrow \left(\beta_\phi\right)^2$ *a.s.* as $N \rightarrow \infty$, where $\beta_\phi = E[\cos\phi_i]$ and *a.s.* denotes almost sure convergence. In other words, when the total transmit power is kept a constant, the received signal power increases linearly with N as $N \rightarrow \infty$.

Note that when there are no phase errors, i.e. $f_\phi(\phi_i) = \delta(0)$, then $\frac{1}{N}P_R \rightarrow 1$ *a.s.*

Proposition 2: For finite N, $E[P_R] = 1 + (N-1)\left(\beta_\phi\right)^2$. Thus, even for finite

N, the expected value of the received signal power increases linearly with N. (β_ϕ is defined as in Proposition 1, i.e. $\beta_\phi = E[\cos\phi_i]$.)

In the absence of phase errors, Proposition 2 gives that $E[P_R] = N$.

Proposition 3: When N is large enough for the central limit theorem to apply,

$$P_R \approx X_c^2 + X_s^2 \tag{3.3}$$

where $X_c \sim N(m_c, \sigma_c^2)$, $X_s \sim N(0, \sigma_s^2)$, and the parameters m_c, σ_c^2, and σ_s^2, are given as follows:

$$\begin{aligned} m_c &= \sqrt{N}E[\cos(\phi_i)] \\ \sigma_c^2 &= 2E[\cos^2(\phi_i)] - E[\cos(\phi_i)]^2 \\ \sigma_s^2 &= 2E[\sin^2(\phi_i)] \end{aligned} \tag{3.4}$$

The variance of the received signal power is then

$$\mathrm{Var}[P_R] = 4\sigma_c^2 m_c^2 + 2\sigma_c^4 + 2\sigma_s^4 \tag{3.5}$$

which increases linearly with N.

When there are no phase errors, (3.5) reduces to $\mathrm{Var}[P_R] = 4N$.

(Refer to Appendix A for a proof of these results.)

Proposition 2 implies that as long as the distribution of phase errors is such that $\beta_\phi \equiv E\big(\cos\phi_i\big)$ is close to 1, large gains can still be realized using distributed beamforming.

We now present some numerical results comparing the above analytical model with Monte-Carlo simulations performed using SIMULINK. We assume that the

transmitters transmit a binary pulse train modulated by BPSK, with a bit-rate small compared to the carrier frequency:

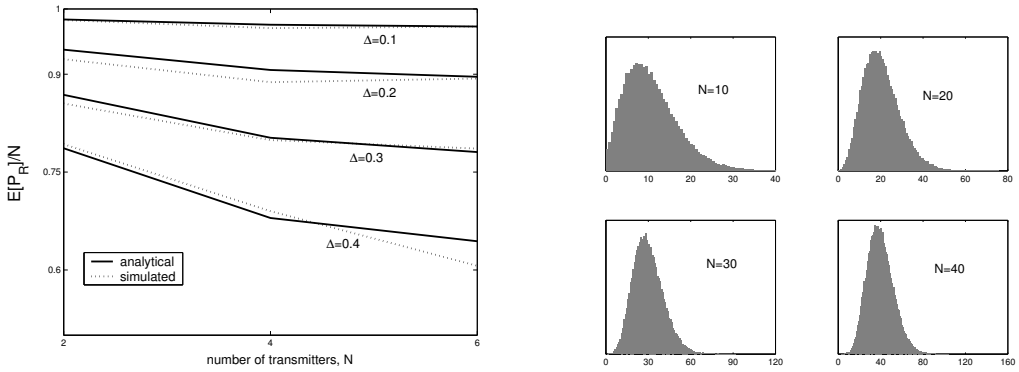
$$m(t) = \sum_k p(t - kT)s_k \tag{3.6}$$

where $\{s_k\}$ is the BPSK symbol stream, and $p(t)$ is the transmitted pulse. The average power of the pulse $p(t), t = 0..T$ is normalized to $\frac{1}{N}$ and $E[|s_k|^2] = 1$ so that the total power transmitted by all the transmitters is $P_T = 1$. Further we assume that the phases ϕ_i are distributed uniformly in the range $(-\Delta\pi, \Delta\pi)$.

Fig. 3.1(a) shows the variation of average beamforming gain normalized to the maximum possible: i.e. $\frac{E(P_R)}{N}$ against the phase error parameter Δ. We find that beamforming gains of more than 70% of the maximum are possible with phase errors as large as of 60°. In other words, the term β_ϕ decreases very slowly with the parameter Δ, which leads to the key conclusion that the beamforming gains are robust to moderately large phase errors.

While Fig. 3.1(a) shows the *average* beamforming gain, the actual beamforming gain is a random variable. We now look at the variation of the SNR with the phase errors uniformly distributed as above. Histograms of P_R, calculated using the Normal approximations as in Proposition 3, are shown in Fig. 3.1(b) where $\Delta = 0.1$ and $N = 10 : 10 : 40$.

The histograms in Fig. 3.1(b) show increased averaging for larger numbers of transmitters. This is expressed quantitatively in Proposition 3, which shows that while the mean of P_R is proportional to N, the standard deviation is proportional to $\sqrt{N}$ i.e. the fractional deviation $\frac{\sqrt{\mathrm{Var}(P_R)}}{E[P_R]}$ decreases with increasing N. This means that the probability of an outage event e.g. where the received SNR is smaller than 70% of its mean, decreases with increasing N, showing that beamforming has the effect of mitigating fading. This is true for perfect and imperfect



(a) $\frac{P_R}{N}$ vs N, empirical and analytical results.

(b) Histograms of P_R, $\Delta = 0.1$.

Figure 3.1. Variation of beamforming gain with number of transmitters.

synchronization. Of course, the existence of phase errors can only increase the variance over that of an ideal, error free system.

3.2 A Protocol for Synchronization

In this section, we present a protocol for achieving carrier phase synchronization based on a master-slave architecture. This is a multi-step process, and each step contributes to the overall phase error $\phi_i(t)$ that limits the beamforming gain. We now look at each step of the synchronization in detail.

The idea behind the protocol is illustrated in Fig. 3.2. The master transmitter has a local oscillator which generates a sinusoid $c_0(t)$:

$$c_0(t) = \Re\big(\tilde{c}_0(t)\big), \text{ where } \tilde{c}_0(t) = e^{j(2\pi f_c t + \gamma_0)} \tag{3.7}$$

that serves as the reference signal for the network. The master transmitter broadcasts $c_0(t)$ to all the slaves. Consistent with the assumption of inexpensive local communication, we ignore the receiver noise in this channel. After reception and

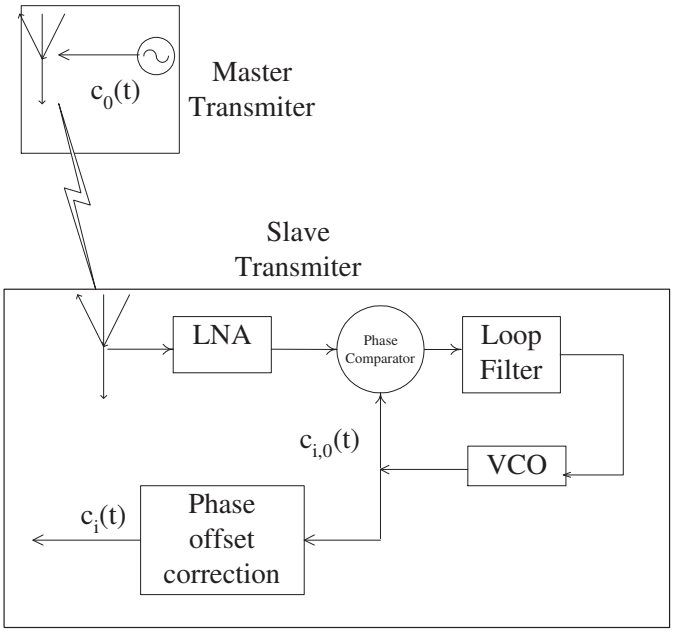


Figure 3.2. Master-Slave architecture for carrier synchronization.

amplification, the slave transmitter i receives the signal broadcast by the master as:

$$c_{i,0}(t) = \Re\big(\tilde{c}_{i,0}(t)\big), \text{ where } \tilde{c}_{i,0}(t) = A_{i,0}e^{j(2\pi f_c t + \gamma_0 - \gamma_i)} \tag{3.8}$$

where γ_i is the phase shift between the master and slave. $A_{i,0}$ is the amplitude of the received signal, its precise value is unimportant to the phase synchronization process (as the PLL is only sensitive to its phase). We set the term $A_{i,0}$ to unity, and the constant γ_0 to zero for simplicity.

The transmitter i uses this signal $c_{i,0}(t)$ from (3.8) as input to a second-order phase locked-loop, driven by a VCO with a quiescent frequency close to f_c. From PLL theory [45], we can show that the steady-state phase error between VCO output and $c_{i,0}(t)$ is zero, and therefore, the steady-state VCO output can be used as a carrier signal consistent across all transmitters - provided that the offset γ_i can be corrected for.

The phase offset γ_i is the total phase shift between the master transmitters' reference oscillator signal $c_0(t)$, and the input signal at the slave transmitters' PLL to which the slave VCO is synchronized in steady-state. One contribution to γ_i is from the phase response of the RF amplifiers at the master and slave transmitter. These offsets are fixed and precisely known, and therefore, can be corrected for. However, the propagation delay of the wireless channel between master and slave also contributes to γ_i. This contribution can be characterized by an effective channel length d_i as $\gamma_i = \frac{2\pi f_c d_i}{c}$.

Unfortunately, for the high-frequency RF carriers typical of wireless networks, even a small uncertainty in channel length d_i causes substantial phase uncertainty e.g. at $f_c = 1.0$ GHz, the wavelength of the transmission is 30 cm, and an uncertainty of 15 cm in the channel length causes an uncertainty of 180° in γ_i. If left uncorrected this is disastrous for distributed beamforming, because a 180° offset would change constructive interference between transmitters into destructive interference. In centralized antenna arrays, the array elements are arranged in a known geometry, and therefore, the offset for each element can be precisely computed. This is not a reasonable assumption for ad-hoc and sensor networks considered in this dissertation. Thus it is necessary to develop methods to explicitly measure and correct for this unknown offset. Fortunately, if the transmitters are not moving relative to each other, this offset stays roughly constant for significant time intervals, and therefore, frequent recalibration is not required. In Section 3.2.1, we describe a protocol for performing this calibration, based on each slave transmitter transmitting their frequency-locked carrier signal $c_{i,0}(t)$ back to the master transmitter. We now sketch the process of channel estimation, and the algorithm for distributed beamforming assuming that slave i has an estimate $\hat{\gamma}_i = \gamma_i + \phi_i^e$ of its phase offset, where ϕ_i^e is the estimation error in the phase

calibration. Slave i then has the calibrated carrier signal $c_i(t)$, which it uses to perform channel estimation and beamforming:

$$\begin{aligned} c_i(t) &= \Re\big(\tilde{c}_i(t)\big) \\ \text{where } \tilde{c}_i(t) &= \tilde{c}_{i,0}(t)e^{j\hat{\gamma}_i} \\ &= e^{j2\pi f_c t + j\phi_i^e} \end{aligned} \tag{3.9}$$

So far the synchronization process has been coordinated within the network by the master transmitter without requiring any interaction with the receiver. In order for the transmitters to beamform towards the receiver, some information about the direction of the receiver, or more precisely the channel response to the receiver is required. Using channel reciprocity allows us to achieve this with only a minimum interaction with the receiver. Specifically the receiver broadcasts an unmodulated carrier signal $g(t)$:

$$g(t) = \Re(\tilde{g}(t)) = \Re\big(e^{j2\pi f_{c,0}t+\phi_0}\big) \tag{3.10}$$

Each transmitter independently demodulates its received signal $g_i(t) = \Re\big(h_i\tilde{g}(t)\big)$ using $c_i(t)$ to obtain an estimate $\hat{h}_i$ of its own complex channel gain h_i to the receiver (for a narrowband message signal, the linear time-invariant channel to receiver is represented as a scalar complex gain). More precisely, the channel estimate $\hat{h}_i$ is obtained by the transmitter i by demodulating the received carrier signal $g_i(t)$ using $c_i(t)$, and sampling the result at some fixed time t_h. Note that while the transmitter nodes have a mutually consistent carrier signal, the receiver's carrier has not been explicitly synchronized to the master transmitter's reference carrier, and therefore, would not be at the same frequency as the transmitters. Letting $f_{c,0} = f_c + \Delta f$ we have:

$$\hat{h}_i = h_i \cdot e^{j(\phi_0 - \phi_i^e + \phi_h)} \tag{3.11}$$

where $\phi_h = 2\pi\Delta f\ t_h$. We observe that the term ϕ_0 is just a constant scaling term and adds no *relative* phase errors between transmitters. Similarly the term ϕ_h adds no relative phase error so long as the sampling term t_h is identical for all transmitters. If the sampling times are off due to timing errors τ_i, we get an effective phase noise: $\phi_i^h = 2\pi\Delta f\tau_i$. Therefore, we rewrite (3.11) as:

$$\hat{h}_i = C \cdot h_i \cdot e^{j(-\phi_i^e + \phi_i^h)} \tag{3.12}$$

where C is a (complex) scaling constant that has no impact on the beamforming process. For simplicity, we take $C = 1$.

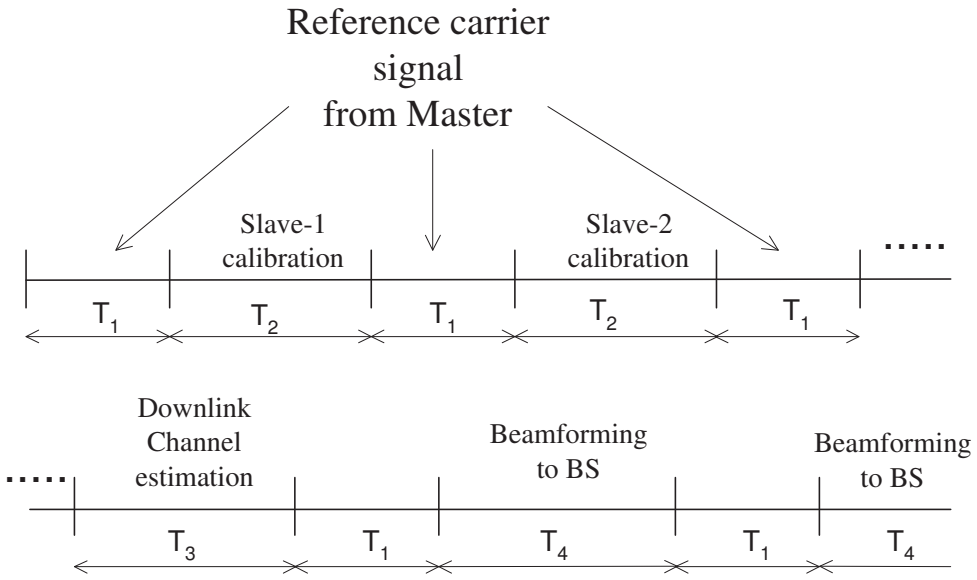


Figure 3.3. The Time-Division Duplexing constraint.

The transmitters now use the synchronized carrier signal $c_i(t)$, and the channel estimate $\hat{h}_i$ to modulate the message signal for beamforming. The slave transmitters obtain their carrier signal from the VCO that is synchronized to the reference signal from the master transmitter, however, it is not possible for the slave transmitters to receive a synchronization signal from the master transmitter, while they

are transmitting. Therefore, the VCOs of the slave transmitters need to operate in an open-loop mode (shown as switch position S_2 in Fig. 3.4), while the slave transmitters are transmitting. While in the open-loop mode, the slave's carrier signals obtained from the VCO undergoes uncompensated phase drift because of internal oscillator noise, and over time, the different slave carriers drift out of phase. This motivates the time-division duplexed mode of operation shown in Fig. 3.3, where the master transmitter periodically transmits a reference carrier signal to resynchronize the slave carriers, to keep the total phase error bounded. The phase noise can be considered as a cyclostationary random process with period $T = T_1 + T_2$, and we analyze it in detail in Section 3.3. The noisy carrier signal used by the slave transmitter i for modulation can be written as:

$$\begin{aligned} c_i^o(t) &= \Re\big(\tilde{c}_i^o(t)\big) \\ \text{where } \tilde{c}_i^o(t) &= \tilde{c}_i(t) e^{j\phi_i^d(t)} \\ &= e^{j2\pi f_c t + j\phi_i^e + j\phi_i^d(t)} \end{aligned} \tag{3.13}$$

$\phi_i^d(t)$ represents the uncompensated VCO drift when slave i is transmitting. After modulation by the carrier signal $c_i^o(t)$, slave transmitter i applies a complex amplification $\hat{h}_i^*$ to compensate for the channel, and transmits the signal:

$$\begin{aligned} s_i(t) &= \Re\big(\tilde{s}_i(t)\big) \\ \text{where } \tilde{s}_i(t) &= \hat{h}_i^* m(t) \tilde{c}_i^o(t) \end{aligned} \tag{3.14}$$

The received signal at the receiver is then given by:

$$\begin{aligned} r(t) &= \Re\big(\sum_i h_i \tilde{s}_i(t) + n(t)\big) \\ &= \Re\big(m(t) \sum_i h_i \hat{h}_i^* \tilde{c}_i^o(t)\big) \\ &= \Re\Big(m(t) \sum_i \big|h_i\big|^2 e^{j2\pi f_c t - j\phi_i^h + j2\phi_i^e + j\phi_i^d(t)}\Big) \end{aligned} \tag{3.15}$$

Comparing (3.15) with (3.1), we have for the total carrier phase error

$$\phi_i(t) = -\phi_i^h + 2\phi_i^e + \phi_i^d(t) \tag{3.16}$$

Equation (3.16) shows the different contributions to the total phase error in the received signal at the receiver. In Section 3.3 we look at the phase error in detail; we argue that the dominant component is the drift term $\phi_i^d(t)$, and show quantitatively how it affects the total beamforming gains.

3.2.1 Closed-Loop Method for Carrier Phase Calibration

In this section, we propose a flexible method for carrier phase calibration, where the master transmitter measures the round-trip phase offset, and uses it to estimate the unknown phase offset γ_i from (3.8) for each slave, assuming symmetry in the forward and reverse channels to the slave nodes. The flexibility of this method comes at the price of complexity, and the necessity of synchronizing each of the slaves individually. However, the calibration process has to be repeated only when the RF channel between the master and slave transmitter changes, therefore, the overhead from this process is small.

Remark: In the ideal case where the relative positions of the master and slave transmitters as well as any multi-path scatterers do not change, the calibration process has to be performed only once (at startup time). In practice, wireless channels are not perfectly static: mobile scatterers and physical changes in the medium may change the channel phase response even when the transmitters are stationary. Therefore, it makes sense to recalibrate the slave transmitters periodically to track the channel changes. Fortunately, the channel variations are slow compared to the channel transmission times, and the robustness benefits of this periodic recalibration (e.g. every 100 seconds) outweigh the small extra overhead.

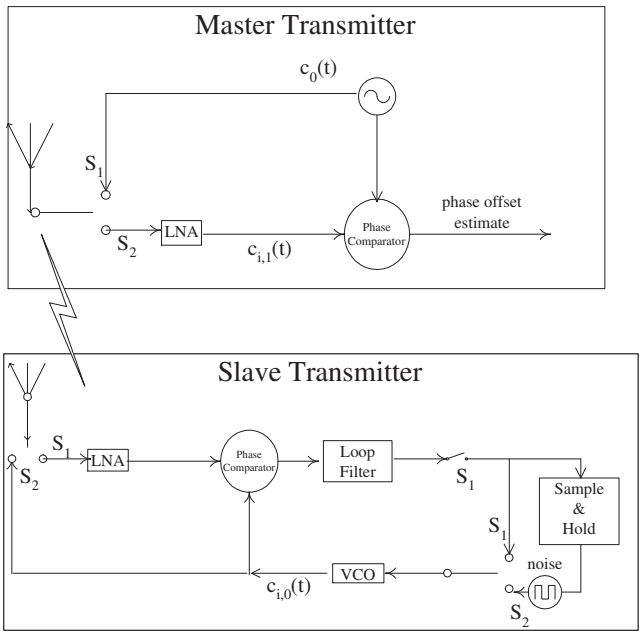


Figure 3.4. Round-trip phase calibration.

Fig. 3.4 illustrates the process of round-trip phase offset estimation. The basic idea is for the slave transmitter i to transmit back to the master transmitter the (uncompensated) VCO signal $c_{i,0}(t)$ represented in (3.8). The symmetry of the forward and reverse master-slave channels imply that the signal $c_{i,1}$ at the master transmitter can be written as:

$$c_{i,1}(t) = A_{i,1} \cdot \Re\left(e^{j(2\pi f_l t + \gamma_0 - 2\gamma_i)}\right) \tag{3.17}$$

where $A_{i,1}$ is the received signal amplitude at the master transmitter ($A_{i,1}$ is equal to $A_{i,0}$ by symmetry, but the actual value is not relevant to the phase noise, therefore, $A_{i,1}$ is set to unity for the discussion). Estimating the phase difference between $c_{i,1}(t)$ from (3.17) and $c_0(t)$ from (3.7) gives:

$$\Delta\phi_i = \left(2\gamma_i \, mod \, 2\pi\right) \tag{3.18}$$

Given a measured value of $\Delta\phi$, we have the estimated value of the offset γ_i:

$$\hat{\gamma}_i = \frac{\Delta\phi}{2} \tag{3.19}$$

Remark: There is one subtlety that needs to be noted here: the round-trip measurement of phase offset as in (3.19) leaves a 180° ambiguity in γ_i. In other words, by measuring $\Delta\phi_i$ we cannot distinguish between γ_i and $\gamma_i + 180°$. While it is possible to resolve this ambiguity by exchanging another set of messages between master and slave i, it turns out that a 180° phase difference does not affect the beamforming process. The reason is that the same carrier $c_i(t)$ is used by slave i for both channel estimation and distributed beamforming, and as (3.16) shows, the two ambiguities cancel each other.

3.2.2 Discussion

The time-division duplexing requirement for the master-slave link is the most important constraint of the synchronization protocol of Section 3.2. Other authors [46] have considered using two frequencies to avoid this problem, with one frequency f_1 reserved for the master-slave link and the slave transmitters beamforming to the receiver on a completely different frequency f_2. In such schemes, the slave transmitters use a frequency dividing PLL to obtain a carrier signal at frequency f_2 as $f_2 = \frac{m}{n} f_1$, where m and n are integers. Under this scheme the slave PLLs do not need to be open-loop while transmitting, therefore, an interesting question is whether such a frequency division duplexed (FDD) architecture can eliminate the problem of uncompensated carrier drift.

Unfortunately, the frequency divider introduces a phase ambiguity of integer multiples of $\frac{2\pi}{n}$ in the derived carrier signal. While it may appear that a constant phase ambiguity can be estimated and corrected for in a one-time calibration

process, closer analysis shows that such phase ambiguities may also occur during the dynamical operation of the PLL, e.g. due to cycle slips [45]. Therefore, periodic recalibration is still necessary even with a FDD architecture, and we conjecture that a fundamental tradeoff between the synchronization overhead and the achievable beamforming gain still applies in this case. One specific form of this tradeoff in a FDD synchronization scheme is examined in Chapter 5.

3.3 Analysis of Phase Error

So far, we have described a protocol for carrier synchronization and beamforming, while enumerating the different sources of phase errors ϕ_i^e, ϕ_i^h and $\phi_i^d(t)$. Of the three different sources of error, ϕ_i^e and ϕ_i^h are constant calibration errors, whereas $\phi_i^d(t)$ is a time-varying noise term that arises from oscillator drift. Theoretically, we could perform carrier phase calibration and channel estimation several times independently and reduce the error terms ϕ_i^e and ϕ_i^h to arbitrarily small levels. However, the drift term $\phi_i^d(t)$ represents an irreducible phase error. Therefore, we consider this as the dominant cause of performance degradation and we now develop a stochastic model to characterize it.

The previous discussion in Section 3.2 motivated the time-division duplexed (TDD) mode of operation as shown in Fig. 3.3, where the slave transmitters alternate between *sync* and *transmit* timeslots. The timeslots T_1 where the slave transmitter synchronize to the master is a synchronization overhead, therefore, it is desirable to keep it small relative to the useful timeslots T_2. T_1 is determined by the settling time of the slave PLL, and T_2 is determined by the maximum admissible phase error, and the statistics of oscillator phase noise. By tolerating larger phase error, we are able to make T_2 higher and thereby reduce the synchronization

overhead. We show in Section 3.1 that the SNR gains from beamforming are robust to moderately large phase errors. In the remainder of this section, we offer a quantitative analysis of this tradeoff using a stochastic model for oscillator phase noise.

Consider the PLL of the slave transmitter as shown in Fig. 3.4. We use a loop-filter with one pole to obtain a second-order PLL with the closed-loop transfer function [45]:

$$H(s) = \frac{s^2}{s^2 + 2\xi\omega_n s + \omega_n^2} \tag{3.20}$$

where ω_n is the *natural frequency* and ξ is the *damping ratio* of the loop. By standard PLL theory, the steady state phase error of a second-order PLL is zero, and if we require a 90° phase margin, then we need a damping ratio of at least $\xi = 1.0$, and the settling time (defined as the time required for the phase error to decrease to less than a given small fraction, say $\rho = 1\%$ of the initial error) is $T_s \approx \frac{4}{\omega_n}$. Since the synchronization timeslot $T_1 \geq T_s$, in order to minimize overhead we want to make T_s as small as possible. However, we observe that the loop has a low-pass frequency response with approximate bandwidth of ω_n, therefore, increasing ω_n also increases the phase noise. At the end of the *sync* timeslot, the loop-filter output is sampled and the VCO input is held to this value for the duration of the *transmit* timeslot. The phase error process in the transmit timeslot determines the achievable beamforming gain.

3.3.1 PLL Phase Noise as a Stochastic Process

In order to study this more quantitatively, we assume that the PLL input signal from the master transmitter (in the synchronization timeslot) is noiseless, and the only source of phase error is internal phase noise $\phi_i^d(t)$ in the slave transmitters'

local oscillator signal:

$$c_i(t) = \Re\left(e^{j2\pi f_c t + j\phi_i^d(t)}\right) \tag{3.21}$$

(We also assume that the PLL phase drift is always small enough to allow the use of a linearized model.)

The traditional way to measure phase noise is by specifying its root-mean squared frequency deviation and Allan variance [45]. However, these measures are most useful if the noise process is stationary in time. In our case the drift process $\phi_i^d(t)$ is not stationary; in the *transmit* timeslot, the dominant phase noise contribution is from a random residual frequency offset that causes the phase error to increase linearly in time until the next *sync* timeslot (see Fig. 3.5). Therefore, the statistics are more appropriately modeled as *cyclostationary* with the period $T = T_1 + T_2$. We use a more fundamental approach to model this process.

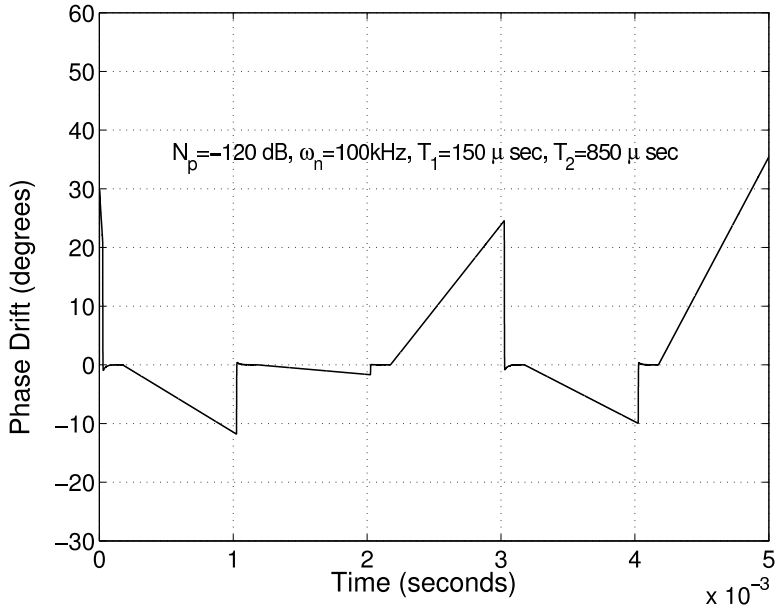


Figure 3.5. Simulation of oscillator phase drift.

In our model, the phase error in the oscillator in closed-loop (i.e. in the *sync* timeslot) consists of two components: a decaying transient of the initial phase

offset, and a phase noise internal to the oscillator. The phase error in the free-running oscillator (i.e. in the *transmit* timeslot) has those two components and an additonal linear phase drift. The linear drift arises because the VCO frequency set by the sample-and-hold (see Fig. 3.4) may have a small but non-zero offset from the reference frequency f_l. The oscillator internal phase noise is modeled as a wideband (white) Gaussian noise process with spectral density N_p. While phase noise in practical oscillators may also have other types of spectral densities e.g. flicker noise and random-walk noise, white Gaussian phase noise represents a worst case in terms of large instantaneous frequency deviations, because of the power in the high frequencies. Let N_p be the normalized spectral density defined such that the total power of the phase noise is $N_p\omega_n$. In other words, a white Gaussian phase noise with spectral density N_p will have the same power in a system of bandwidth ω_n as the oscillator's total internal phase noise.

Since the phase error process $\phi_i^d(t)$ is a zero-mean Gaussian process at all times, therefore, we characterize its statistics by computing its variance at the key time instants labelled A, B and C in Fig. 3.3. Let the random phase values at these instants be denoted as ϕ_A, ϕ_B and ϕ_C, and their standard deviations as σ_A, σ_B and σ_C respectively. By the cyclostationarity of $\phi_i^d(t)$, we have $\sigma_C \equiv \sigma_A$. Using the linearity of the PLL's phase response, we can write the phase at time B as the superposition of the deterministic decay of the initial error ϕ_A, to a small fraction ρ of its starting value (e.g. $\rho = 1\%$), and a noise term:

$$\phi_B = \rho\phi_A + \psi_1 \tag{3.22}$$

ϕ_B is small by design, and its variance can be written as:

$$\sigma_B^2 = \rho^2\sigma_A^2 + N_p\omega_n \tag{3.23}$$

In addition to the small phase error, at time instant B, the VCO input is sampled

to set the VCO frequency for the *transmit* timeslot. The sampled value has a random offset Δf from the reference carrier frequency, and this offset consists of a transient term and a noise term:

$$\Delta f = \rho\omega_n\phi_A + \omega_n\psi_3$$
$$\text{Therefore } \sigma_f^2 = \rho^2\omega_n^2\sigma_A^2 + \omega_n^3 N_p \tag{3.24}$$

We have for the evolution of the phase between time instants B and C:

$$\phi_C = \Delta f \ T_2 + \phi_B + \psi_2 \tag{3.25}$$

Of the three terms in (3.25), the frequency offset is the dominant term because it causes a phase drift that grows with time. The phase ϕ_B is small by design, and ψ_2 represents a stationary term, and we can safely neglect both terms compared to the linear drift. This is also illustrated in the simulation shown in Fig. 3.5. Thus we have:

$$\sigma_C^2 \equiv \sigma_A^2 = \sigma_f^2 T_2^2 \tag{3.26}$$

Combining (3.24), (3.23) and (3.26), we get:

$$\sigma_A^2 = \frac{N_p\omega_n^3 T_2^2}{1 - \rho^2\omega_n^2 T_2^2} \tag{3.27}$$

Fig. 3.5 shows a simulation of the phase error over time with $T_1 = 150\mu$sec, $T_2 = 0.85$ ms, $\omega_n = 100$ kHz, $\rho = 1\%$ and $N_p = 7 \times 10^{-11} Hz^{-1}$ or -101 dBc/Hz. The VCO in this simulation has a quiescent frequency that is 1 kHz offset from the reference carrier signal. The spectral density of phase noise is chosen conservatively compared to typical numbers reported e.g. -110 dBc/Hz in [47]. For these numbers, we get $\sigma_A \approx 24^\circ$ from (3.27). Since the phase error is a Gaussian variable with standard deviation smaller than 24° at all times, $\beta_\phi = E(\cos\phi_i) \geq 0.91$, and by Proposition 1, we can see that average beamforming gains of at least 91% are

achievable. This is an average number and occasionally, phase errors larger than this can occur as seen in Fig. 3.5, where phase error becomes almost 35° at one point. Even with this large phase error, the resulting beamforming gain is 81% of the maximum. This confirms the results of Section 3.1, that beamforming gain is robust to phase errors and demonstrates the basic feasibility of the distributed beamforming algorithm.

3.3.2 Cramer-Rao Bounds for Synchronization Error

So far in this analysis we have limited ourselves arbitrarily to a second-order PLL because it is the most commonly used device in practice. However, we can also derive fundamental limits on the size of the frequency and phase offsets, by viewing the PLL as a frequency and phase estimator. The PLL uses the (noisy) oscillator signal in the *sync* timeslot to form estimates $\hat{f}_l$ and $\hat{\phi}$. It uses the estimate $\hat{\phi}$ to drive the phase difference with the PLL input to zero, and $\hat{f}_l$ to tune the VCO's input to the frequency of the reference, and the sample-and-hold element keeps the VCO tuned to that estimate in the *transmit* timeslot. The Cramer-Rao lower bound for the variance of these offsets has been computed in previous work on frequency estimation [48]:

$$\begin{aligned} \hat{\sigma}_f^2 &\doteq \mathrm{Var}(\hat{f}_l) = \frac{3N_p}{\pi^2 T_1^3} \\ \hat{\sigma}_\phi^2 &\doteq \mathrm{Var}(\hat{\phi}) = \frac{2N_p}{T_1} \end{aligned} \tag{3.28}$$

Using the same values used in Fig. 3.5, we find $\hat{\sigma}_f = 2.5$ Hz, and $\hat{\sigma}_\phi \ll 1°$. Since $\hat{\sigma}_f$ is substantially smaller than the PLL's root-mean squared frequency offset $\sigma_f = 418$ Hz, we conclude that there is significant suboptimality in using an analog PLL. Thus, performance can be further improved by using optimal digital processing.

Chapter 4

A Feedback Control Algorithm for Beamsteering

In this chapter, we present a simple iterative procedure, based on feedback from the receiver, for achieving phase coherence, and show that this procedure provides a powerful method to satisfy the requirements for distributed beamforming. The basic idea is as follows. Each transmitter adjusts its phase randomly at each iteration, with the receiver broadcasting one bit of feedback per iteration as to whether its SNR is better or worse than before. If it is better, the transmitters keep the previous phase perturbation, and if it is worse, they undo the previous phase perturbation. This randomized ascent procedure is repeated until the transmitters converge to phase coherence.

This feedback-based approach presupposes carrier *frequency* synchronization i.e. all transmitters are assumed to modulate their transmissions with RF carriers at the same frequency. This can be achieved using a *master-slave* synchronization process similar to Section 3.2. However the feedback scheme automatically compensates for unknown transmitter phase offsets as well as unknown channel

gains, thereby obviating the need for explicit calibration and channel estimation procedures. This makes it substantially less complex than the reciprocity-based scheme of Chapter 3, and therefore easier to implement in practice. Any imperfections in the carrier synchronization process have an effect on the performance of the beamforming procedure. We postpone considerations of the synchronization process, as well as other impairments such as noise and channel time-variations to Chapter 5. In this chapter, we focus on gaining an understanding of the properties of the feedback algorithm under ideal conditions of frequency-synchronized carrier signals, constant channel gains and transmitter phase offsets, and error-free SNR estimation and feedback.

As shown in Fig. 4.1, we consider a system of N transmitters transmitting a RF carrier signal at frequency f_c modulated by a common narrowband message signal $m(t)$ to the receiver. Therefore the baseband signal of transmitter i can be written as $s_i(t) = Ae^{j\theta_i}m(t)$. Our goal is to adjust the complex gains $Ae^{j\theta_i}$ at the transmitters so as to achieve phase coherence at the receiver; we normalize the power constraint by setting $A = 1$.

The RF carrier signals of all the transmitters are *frequency* synchronized by assumption, however they still have unknown phase offsets between them, e.g. because of unknown propagation delays in the master-slave channels. The effect of these phase offsets is that the phase of the baseband signal transmitted from transmitter i gets rotated by an unknown amount γ_i.

We denote the complex channel gain of transmitter i to the receiver as $h_i = a_i e^{j\psi_i}$, where $a_i \geq 0$ represents the attenuation and ψ_i the phase response of the wireless channel. The received signal at the receiver is just a superposition of the signal received from each transmitter. The received signal due to transmitter i is given by $s_i(t)\ e^{j\gamma_i}\ h_i = a_i e^{j(\theta_i+\gamma_i+\psi_i)}\ m(t)$. We ignore distortions in the message

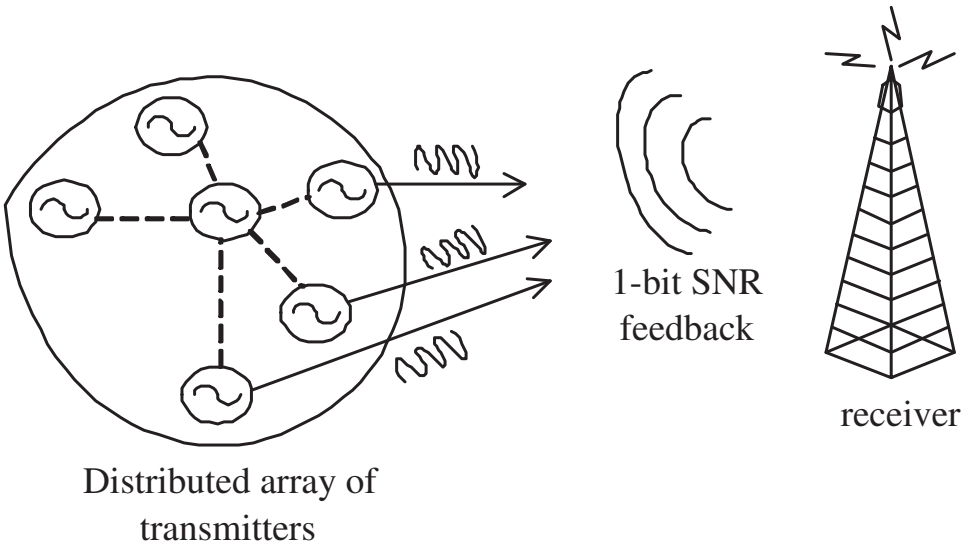


Figure 4.1. Beamsteering using receiver feedback.

due to small timing mismatches[1] between the transmitters, which allows us to ignore the presence of the message in what follows.

The net complex gain at the receiver is therefore given by

$$Y = \sum_{i=1}^{N} a_i e^{j(\theta_i+\gamma_i+\psi_i)} = \sum_{i=1}^{N} a_i e^{j\Phi_i} \tag{4.1}$$

where $Y \geq 0$ is the amplitude, or *received signal strength* (RSS), and $\Phi_i = \theta_i + \gamma_i + \psi_i$ is the phase at the receiver corresponding to the signal from transmitter i. Note that the RSS only depends on the unknown γ_i and ψ_i through the sum $\gamma_i+\psi_i$; nevertheless we write them separately to emphasize their different physical origins.

Our objective is to adapt the transmitter phases $\{\theta_i\}$ so as to maximize Y.

[1]Note that this requirement of time synchronization is unrelated to the phase synchronization required for beamforming; timing errors cause some ISI and message signal distortion, but do not affect the beamforming gain.

This happens if the received carrier phases Φ_i are all equal:

$$Y = \Big|\sum_{i=1}^{N} a_i e^{j\Phi_i}\Big| \leq Y_{opt} \equiv \Big(\sum_{i=1}^{N} a_i\Big), \text{ with equality if and only if } \Phi_i = \Phi_j \quad (4.2)$$

The purpose of the feedback control algorithm is to allow transmitter i to dynamically compute the optimal value of θ_i in (4.2), without requiring knowledge of either ψ_i or γ_i.

4.1 Description of Algorithm

The adaptation is performed in time-slotted fashion, with each transmitter adapting its phase in a time slot in response to feedback from the receiver. Such coarse time synchronization can be achieved by using the timing of the feedback broadcast by the receiver, assuming that the spread of propagation delays between the receiver and the transmitters is small compared to the time slot duration. At the beginning of slot n, let $\theta_i[n]$ denote the best known carrier phase at transmitter i. At each time-slot n, each transmitter i applies a random phase perturbation $\delta_i[n]$ to $\theta_i[n]$ in order to probe for a potentially better phase. The transmitted *probe* phase in slot n is then given by

$$\theta_i^{(p)}[n] = \theta_i[n] + \delta_i[n]$$

where $\delta_i[n]$ is a random phase perturbation. The corresponding RSS $Y[n] = \big|\sum_i a_i e^{j\Phi_i[n]}\big|$, where $\Phi_i[n] = \theta_i^{(p)}[n] + \gamma_i + \psi_i$. The receiver measures $Y[n]$, and broadcasts one bit of feedback indicating whether $Y[n]$ is bigger or smaller than its record of the highest observed signal strength so far, which we denote by

$$Y_{best}[n] = \max_{k \leq n} Y[k]$$

If the feedback from the receiver indicates an improvement in RSS, then the transmitters keep their random phase perturbations, otherwise they undo their perturbations. Thus, the best known phases at the transmitters are updated as follows:

$$\theta_i[n+1] = \begin{cases} \theta_i[n] + \delta_i[n] & Y[n] > Y_{best}[n] \\ \theta_i[n] & \text{otherwise.} \end{cases} \tag{4.3}$$

Simultaneously, the receiver also updates its record of the highest RSS so far as follows:

$$Y_{best}[n+1] = \max\Big(Y_{best}[n], Y[n]\Big) \tag{4.4}$$

The preceding procedure is repeated over multiple time slots. Equations (4.3) and (4.4) ensure that we retain phase perturbations that increase RSS, while discarding unfavorable ones. This distributed ascent procedure eventually converges to a set of transmit phases that satisfy (4.2), and achieve distributed beamforming. Fig. 4.2 shows the convergence to beamforming with $N = 10$ transmitters.

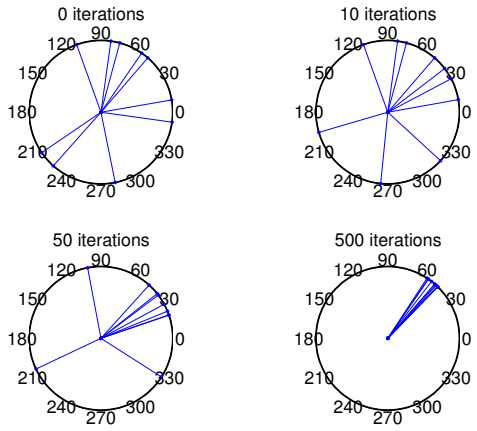


Figure 4.2. Convergence of feedback control algorithm.

The random perturbation $\delta_i[n]$ is chosen independently across transmitters from a probability distribution $\delta_i[n] \sim g_n(\delta_i)$, where the density function $g_n(\delta_i)$

is a parameter of the protocol. We show in Section 4.2 that the behavior of the algorithm is well-characterized by the variance of the distribution $g_n(\delta_i)$ and depends only weakly on the actual distribution. The two-point distribution $\Pr(\delta_i = \pm\delta_0) = 0.5$ permits a particularly simple baseband implementation of the phase shifts [4], and is therefore of great practical interest. In general, the distribution $g_n(\delta_i)$ can be adapted dynamically in time e.g. through the parameter δ_0 for the two-point distribution.

It follows from (4.3) that if the algorithm were to be terminated at time-slot n, the best achievable signal strength using the feedback information received so far, is equal to $Y_{best}[n]$, which corresponds to transmitter i transmitting with the phase $\theta_i[n]$.

$$Y_{best}[n] \equiv \Big|\sum_i a_i e^{\Phi_i[n]}\Big| \text{where } \Phi_i[n] = \theta_i[n] + \gamma_i + \psi_i \tag{4.5}$$

4.1.1 Asymptotic Coherence

We now show that the feedback control protocol outlined in Section 4.1 asymptotically achieves phase coherence for any initial values of the phases Φ_i. We define some notation first.

Let $\bar{\Phi}$ denote the vector of the received phase angles Φ_i. We define the function $\mathrm{RSS}(\bar{\Phi})$ to be the received signal strength corresponding to received phase $\bar{\Phi}$:

$$\mathrm{RSS}(\bar{\Phi}) \doteq \Big|\sum_i a_i e^{j\Phi_i}\Big| \tag{4.6}$$

Phase coherence means $\Phi_i = \Phi_j = \Phi_{const}$, where Φ_{const} is an arbitrary phase constant. In order to remove this ambiguity[2], it is convenient to work with the

[2]This is a form of degeneracy that can be eliminated by a slight modification of the feedback algorithm. Under this modification one previously designated transmitter acts as a reference and does not make any phase adjustments; this forces other transmitters to synchronize themselves

rotated phase values

$$\phi_i = \Phi_i - \Phi_0 \tag{4.7}$$

where Φ_0 is a constant chosen such that the phase of the total received signal is zero. This is just a convenient shift of the receiver's phase reference and as (4.6) shows, such a shift has no impact on the received signal strength, i.e. $\text{RSS}(\bar{\phi}) \equiv \text{RSS}(\bar{\Phi})$.

We interpret the feedback control algorithm as a discrete-time vector random process $\bar{\phi}[n]$ where $\bar{\phi}[n]$ is a N-dimensional vector of phases $\phi_i[n]$ constrained by the condition that the total phase of the received signal is zero as defined in (4.7). This random process is a Markov process because the phase perturbations $\bar{\delta}[n]$ are chosen independently at each time-slot n. However we will not need the Markov property, and therefore the following results are valid even when the $\bar{\delta}[n]$ are not independent across time-slots.

Let Ω be the sample space and let $\mathfrak{F}_1 \subset \mathfrak{F}_2 \subset \ldots$ be a sequence of σ-algebras, such that $\mathfrak{F}_n$ includes all the events upto time-slot n i.e. all events of the form $\{\omega : \omega \in \Omega, \bar{\delta}_1 \in \mathcal{A}_1, \bar{\delta}_2 \in \mathcal{A}_2, \ldots \bar{\delta}_n \in \mathcal{A}_n\}$, where $\mathcal{A}_i \subset (-\pi, \pi]$. The probability measure of the random process $\bar{\phi}$ is defined on the class of σ-algebras $\mathfrak{F}_i$. The sequence $Y_{best}[n]$ as defined in (4.5) is then a sequence of random variables with respect to this probability measure.

We now provide an argument that shows (under certain conditions on the probability density function $g_n(\delta_i)$), that $\{Y_{best}[n]\}$ converges *almost surely* to the constant Y_{opt} for arbitrary starting phases $\bar{\phi}$. It is easy to show that $Y_{best}[n] \to Y_{opt}$ is equivalent to $\bar{\phi}[n] \to 0$. The following proposition will be needed to establish the convergence.

to this reference. Simulation results indicate little effect on the convergence rates until the phases are near coherence, at which point the convergence of the modified algorithm is slightly slower.

Proposition 1: Consider a distribution $g_n(\delta_i)$ that has non-zero support (i.e. $g_n(\delta_i) > g_{lb} > 0$ for some g_{lb}) in an interval $(-\delta_0, \delta_0)$. Given any $\bar{\phi} \neq \bar{0}$, and $\text{RSS}(\bar{\phi}) < Y_{opt} - \nu$, where $\nu > 0$ is arbitrary, there exist constants $\epsilon > 0$ and $\rho > 0$ such that $\Pr\big(\text{RSS}(\bar{\phi} + \bar{\delta}) - \text{RSS}(\bar{\phi}) > \epsilon\big) > \rho$, where ϵ and ρ depend only on ν.

Proof. For the class of distributions $g_n(\delta_i)$ that we consider, the probability of choosing δ_i in any finite interval $I \subset (-\delta_0, \delta_0)$ is non-zero. One example of such a class of distributions is $g_n(\delta_i) \sim \text{uniform}[-\delta_0, \delta_0]$.

Recall that the phase reference is chosen such that the total received signal $\sum_i a_i e^{j\phi_i}$ has zero phase. First we sort all the phases ϕ_i in the vector $\bar{\phi}$ in the descending order of $|\phi_i|$ to get the sorted phases $\tilde{\phi}_i$ satisfying $|\tilde{\phi}_1| > |\tilde{\phi}_2| > ... > |\tilde{\phi}_N|$, and the corresponding sorted channel gains $\tilde{a}_i$. We use the condition $\text{RSS}(\bar{\phi}) < Y_{opt} - \nu$ to get:

$$\begin{aligned}
&\cos(\tilde{\phi}_1) \sum_i \tilde{a}_i < \sum_i \tilde{a}_i \cos(\tilde{\phi}_i) \leq Y_{opt} - \nu \\
&\tilde{\phi}_1 > \sum_i \tilde{a}_i \cos(\tilde{\phi}_i) \leq Y_{opt} - \nu \\
&\tilde{\phi}_1 > \phi_\nu \doteq \arccos\Big(\frac{Y_{opt} - \nu}{\sum_i \tilde{a}_i}\Big)
\end{aligned} \tag{4.8}$$

Now we choose a phase perturbation δ_1 that decreases $|\tilde{\phi}_1|$. This makes the most mis-aligned phase in $\bar{\phi}$ closer to the received signal phase, and thus increases the magnitude of the received signal. Without loss of generality we assume $\tilde{\phi}_1 > 0$, then we need to choose a $\delta_1 < 0$. Consider $\delta_1 \in (-\delta_0, -\frac{\delta_0}{2})$. This is an interval in which $g_n(\delta_1)$ is non-zero, therefore there is a non-zero probability $\rho_1 > 0$ of choosing such a δ_1. We have:

$$\tilde{a}_1 \cos(\tilde{\phi}_1 + \delta_1) - \tilde{a}_1 \cos(\tilde{\phi}_1) > 2\epsilon, \text{ where } \epsilon \doteq \frac{\tilde{a}_1 \delta_0}{4} \sin(\tilde{\phi}_\nu - \frac{\delta_0}{2}) \tag{4.9}$$

We observe that ϵ and ρ_1 depend only on ν and not on $\bar{\phi}$.

The perturbation δ_1 by itself will achieve a non-zero increase in total received signal, provided that the other phases $\tilde{\phi}_i$ do not get too mis-aligned by their respective δ_i:

$$\begin{aligned}\mathrm{RSS}(\bar{\phi}+\bar{\delta}) - \mathrm{RSS}(\bar{\phi}) &= \sum_i \tilde{a}_i\big(\cos(\tilde{\phi}_i+\delta_i) - \cos(\tilde{\phi}_i)\big) \\ &= \tilde{a}_1\big(\cos(\tilde{\phi}_1+\delta_1) - \cos(\tilde{\phi}_1)\big) + \sum_{i>1} \tilde{a}_i\big(\cos(\tilde{\phi}_i+\delta_i) - \cos(\tilde{\phi}_i)\big) \\ &> 2\epsilon + \sum_{i>1} \tilde{a}_i\big(\cos(\tilde{\phi}_i+\delta_i) - \cos(\tilde{\phi}_i)\big) \qquad (4.10)\end{aligned}$$

We note that since $\mathrm{RSS}(\bar{\phi})$ is continuous in each of the phases $\tilde{\phi}_i$, we can always find a $\epsilon_i > 0$ to satisfy:

$$\Big|\tilde{a}_i\big(\cos(\tilde{\phi}_i+\delta_i) - \cos(\tilde{\phi}_i)\big)\Big| < \frac{\epsilon}{N-1}, \forall |\delta_i| < \epsilon_i \qquad (4.11)$$

In particular the choice $\epsilon_i \doteq \frac{\epsilon}{\tilde{a}_i(N-1)}$, satisfies (4.11), and this choice of ϵ_i is independent of $\bar{\phi}$. With the δ_i's chosen to satisfy (4.11), we have:

$$-\epsilon < \sum_{i>1} \tilde{a}_i\big(\cos(\tilde{\phi}_i+\delta_i) - \cos(\tilde{\phi}_i)\big) < \epsilon \qquad (4.12)$$

Since $g_n(\delta_i)$ has non-zero support in each of the non-zero intervals $(-\epsilon_i, \epsilon_i)$, the probability ρ_i of choosing δ_i to satisfy (4.11) is non-zero, i.e. $\rho_i > 0$, which is independent of $\bar{\phi}$. Finally, we recall that each of the δ_i are chosen independently, and therefore with probability $\rho = \prod_i \rho_i > 0$, it is possible to find δ_1 to satisfy (4.9) and $\delta_i, i > 1$ to satisfy (4.11). For $\bar{\delta}$ chosen as above, $\mathrm{RSS}(\bar{\phi}+\bar{\delta}) - \mathrm{RSS}(\bar{\phi}) > \epsilon$, and therefore Proposition 1 follows. □

Theorem 1: For the class of distributions $g_n(\delta_i)$ considered in Proposition 1, starting from an arbitrary $\bar{\phi}$, the feedback algorithm converges to perfect coherence of the received signals *almost surely*, i.e. $Y_{best}[k] \to Y_{opt}$ or equivalently $\bar{\phi}[k] \to \bar{0}$ (i.e. $\phi_i[k] \to 0, \forall i$) *with probability 1.*

Remark: Proposition 1 provides a *sufficient* condition for asymptotic convergence. The conditions of Proposition 1 are satisfied, for instance, by using a fixed uniform distribution $\delta_i \sim \text{uniform}[-\delta_0, \delta_0]$. However the *fixed* two-valued distribution $\delta_i = \pm\delta_0$ does not satisfy these conditions. Convergence can still be assured by using a two-point distribution that varies in time.

Proof. We wish to show that the sequence $Y_{best}[k] = \text{RSS}(\bar{\phi}[n]) \to Y_{opt}$ given an arbitrary $\bar{\phi}[1] = \bar{\phi} \neq 0$. Let $G = \text{RSS}(\bar{\phi})$, and $\nu > 0$ be an arbitrary constant. We now show that with large enough k, $Y_{best}[k]$ exceeds $Y_{opt} - \nu$ for arbitrarily small ν.

Let $\mathcal{D}_{\nu,k} \doteq \{\omega \in \Omega : Y_{opt} - Y_{best}[k] \leq \nu\}$. Since the sequence $Y_{best}[k]$ is non-decreasing, we have $\mathcal{D}_{\nu,j} \subset \mathcal{D}_{\nu,k}$ whenever $j < k$. Therefore we have

$$\mathcal{E}_{\nu,k} \doteq \bigcap_{j=k}^{\infty} \mathcal{D}_{\nu,j} \equiv \mathcal{D}_{\nu,k} \tag{4.13}$$

We start with $Y_{best}[1] = G$. Assume $\nu_1 \doteq (Y_{opt} - G) - \nu > 0$ (otherwise the following proof is trivial). We write $Y_{best}[k] = Y_{best}[1] + \sum_{i=1}^{k-1} x_i$, where the random variable x_i represents the increment in Y_{best} in time-slot i. We then have:

$$\mathcal{D}_{\nu,k} \equiv \{\omega \in \Omega : \sum_{i=1}^{k-1} x_i > \nu_1\}. \tag{4.14}$$

Define $\mathcal{F}_{\epsilon,k} \equiv \{\omega \in \Omega : x_k > \epsilon\}$. From Proposition 1, we have $\Pr(\mathcal{F}_{\epsilon,k} | \mathcal{D}^c_{\nu,k}) > \rho$, for some $\rho, \epsilon > 0$. We note that $Y_{best}[k+1] > Y_{opt} - \nu$ if (but not only if) both $Y_{best}[k] > Y_{opt} - \nu - \epsilon$ and $x_k > \epsilon$ are satisfied. More specifically:

$$\mathcal{D}_{\nu,k+1} \supset \mathcal{D}_{\nu,k} \bigcup \big(\mathcal{D}_{\nu-\epsilon,k} \bigcap \mathcal{F}_{\epsilon,k} \bigcap \mathcal{D}^c_{\nu,k}\big) \tag{4.15}$$

In terms of probabilities, we have:

$$\Pr(\mathcal{D}_{\nu,k+1}) > \Pr(\mathcal{D}_{\nu,k}) + \rho \Pr\big(\mathcal{D}_{\nu-\epsilon,k} \bigcap \mathcal{D}^c_{\nu,k}\big) \tag{4.16}$$

Equation (4.16) can be extended recursively over m time-slots to give:

$$\Pr(\mathcal{D}_{\nu,k+m}) > \Pr(\mathcal{D}_{\nu,k}) + \rho^m \Pr\big(\mathcal{D}_{\nu-m\epsilon,k} \bigcap \mathcal{D}^c_{\nu,k}\big) \tag{4.17}$$

If we set $m \doteq M = \lceil \frac{\nu_1}{\epsilon} \rceil$, $\mathcal{D}_{\nu-M\epsilon,k}$ becomes the sure event and (4.17) simplifies to:

$$\Pr(\mathcal{D}_{\nu,k+M}) > \Pr(\mathcal{D}_{\nu,k}) + \rho^M \Pr(\mathcal{D}^c_{\nu,k}) \tag{4.18}$$

Let $p_j = \Pr(\mathcal{D}_{\nu,jM})$. From (4.18) we have $p_{j+1} > \rho^M + (1-\rho^M)p_j > 1 - (1-\rho^M)^j$, and therefore $\lim_{j\to\infty} p_j = 1$. Since $\Pr(\mathcal{D}_{\nu,k}) \geq \Pr(\mathcal{D}_{\nu,M\lceil \frac{k}{M} \rceil}) \equiv p_{\lceil \frac{k}{M} \rceil}$, we have:

$$\lim_{k\to\infty} \Pr\big(\mathcal{E}_{\nu,k}\big) = \lim_{k\to\infty} \Pr\big(\mathcal{D}_{\nu,k}\big) = 1 \tag{4.19}$$

Since (4.19) holds for arbitrary ν, by the definition of almost sure convergence (Ch. II in [49]), we have proved that $Y_{best}[k] \to Y_{opt}$ *almost surely.* □

4.2 Analytical Model for Algorithm Dynamics

The analysis in Section 4.1.1 shows that the feedback control algorithm of Section 4.1 asymptotically converges for a large class of distributions $g_n(\delta_i)$; however it provides no insight into the rates of convergence. We now derive an analytical model that accurately predicts the average convergence rate of $Y_{best}[n]$. We then use this analytical model, to optimize $g_n(\delta_i)$ for fast convergence.

Recall that $\phi_i[n]$ as defined in (4.7) satisfies $\angle\big(\sum_i a_i e^{j\phi_i}\big) = 0$. Initially all the received phases $\phi_i[1]$ are completely unknown, and therefore are randomly distributed in the interval $(-\pi, \pi]$. We now derive a statistical description of the phases for the case of equal channel gains, setting $a_i \equiv 1$. Physically this corresponds to Line-of-Sight channels to a distant receiver. It is possible to extend

the same analysis to channel gains that vary across transmitters, however to obtain insight into the algorithm dynamics, we restrict the statistical analysis to the Line-of-Sight case.

4.2.1 Statistical Model

The most complete description of the feedback algorithm is given by the statistics of the random vector process $\bar{\phi}[n]$. The update equation (4.3) can be thought of as a stochastic approximation procedure, therefore we might consider using the standard "mean ODE" approach that is used for such problems [41]. Our analytical model is based on similar ideas, however we depart from the standard "mean ODE" approach by focusing on the dynamics of $Y_{best}[n]$ rather than the full phase vector $\bar{\phi}[n]$. This is possible by taking advantage of some important regularities observed in the behavior of the feedback algorithm in large networks, which enables a *statistical* characterization of the phase angles $\phi_i[n]$, in analogy with methods employed in statistical mechanics. Specifically, we use a heuristic application of the *Gibbs conditioning principle* [50] to approximate the phases $\{\phi_i[n]\}$ as iid, with distribution parameterized by the received signal strength $Y_{best}[n] = \sum_{i=1}^{N} \cos(\phi_i[n])$.

For random variables $X_1, ..., X_N$ and a function f, let us denote the empirical average of $\{f(X)\}$ as follows: $\hat{E}[f(X)] = \frac{1}{N}\sum_{i=1}^{N} f(X_i)$. The Gibbs conditioning principle can now be stated as follows.

Gibbs conditioning principle (see [50] and references therein): Suppose that $\{X_i, i = 1, ..., N\}$ are iid random variables with marginal distribution $p(x)$. Then, conditioned on $\hat{E}[f(X)] = a$, the $\{X_i\}$ are approximately iid with marginal dis-

tribution $q(x)$, where

$$q = \arg\min_{p'} D(p'||p) \;\; \text{subject to} \;\; \mathsf{E}_{p'}\left[f(X)\right] = a \tag{4.20}$$

where $E_{p'}[\cdot]$ denotes expectation with respect to p'. That is, $q(x)$ is the distribution closest to $p(x)$ in terms of information-theoretic divergence, subject to the given constraint.

Remark: The "approximately iid" property can be elaborated as follows: subject to the conditioning, $X_1, ..., X_K$ can be modeled as iid with distribution $q(x)$, where $K = K(N)$ must grow slower than N [50].

Gibbs distribution: The variational problem (4.20) can be rewritten using Lagrange multipliers as:

$$q = \arg\min_{p'}\Big(\int p'(x)\log\Big(\frac{p'(x)}{p(x)}\Big)dx - \eta\Big(\int p'(x)f(x)dx - a\Big)\Big) \tag{4.21}$$

Solving (4.21), we obtain that the density for the conditional distribution q is given by

$$q(x) = \frac{e^{\eta f(x)}}{Z(\eta)}p(x) \tag{4.22}$$

where η is a Lagrange multiplier chosen so that the constraint in (4.20) is satisfied with equality and

$$Z(\eta) = \int e^{\eta f(x)}p(x)dx$$

is a normalization constant.

Using the statistical mechanics framework in our setting, the phases $\{\phi_i\}$ are approximated as iid at each iteration of the algorithm, and our goal is to estimate the empirical distribution $f_n(\phi)$ of $\phi_i[n]$ at the n'th step of the algorithm, as a function of the received signal strength at that step. It is convenient to define a normalized version of the received signal strength which can be interpreted as an

empirical average:

$$z_{best}[n] = \hat{E}[\cos\phi[n]] = \frac{1}{N}\sum_{i=1}^{N}\cos\phi_i[n] \tag{4.23}$$

We first show that a heuristic application of the Gibbs conditioning principle yields results that exhibit remarkably close agreement with simulations. This is what is used as a basis for algorithm optimization later in this section. A rigorous analysis based on deriving large deviations principles for the evolution of the algorithm is left as an open problem, but we offer some comment on the salient technical issues.

Iteration 1: The initial phases $\{\phi_i[1]\}$ are iid, uniform over $(-\pi, \pi]$. When we apply iid phase perturbations $\{\delta_i[1]\}$, the resulting phases remain iid, uniform over $(-\pi, \pi]$. When we hit a normalized RSS $z_{best}[2]$ higher than the initial value $z_{best}[1]$, we obtain the next set of phases $\{\phi_i[2]\}$ which satisfy $\hat{E}[\cos(\phi[2])] = z_{best}[2]$. Using the Gibbs conditioning principle, we obtain from (4.22) that $\{\phi_i[2]\}$ are approximately iid with marginal density

$$f_2(\phi) = \frac{e^{\eta\cos\phi}}{Z_2(\eta)} f_1(\phi)\ , \quad -\pi < \phi \leq \pi \tag{4.24}$$

where

$$f_1(\phi) = \frac{1}{2\pi}\ , \quad -\pi < \phi \leq \pi$$

and

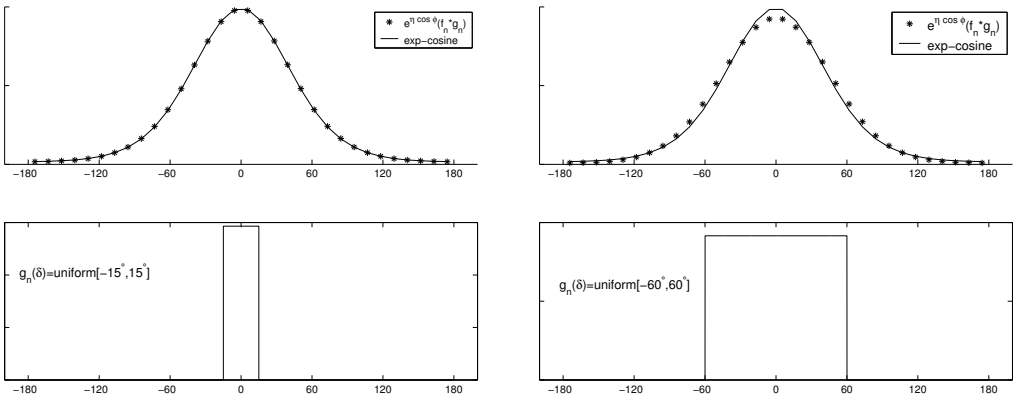
$$Z_2(\eta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{\eta\cos\phi}\ d\phi = I_0(\eta)$$

Recall that

$$I_k(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(k\phi)\ e^{x\cos\phi}\ d\phi \tag{4.25}$$

where $I_k(x)$ is the modified Bessel function of the first kind and order k.

Iteration n: Suppose that the phases $\{\phi_i[n]\}$ are approximately iid with marginal density f_n, and that the phase perturbations being applied are iid with marginal



(a) Small perturbations (b) Large perturbations

Figure 4.3. Approximation of f_{n+1} with exp-cosine distribution.

density g_n. When we hit a normalized RSS $z_{best}[n+1] > z_{best}[n]$, we obtain $\phi_i[n+1] = \phi_i[n] + \delta_i[n]$ which are approximately iid with density $f_n * g_n$ (the convolution is circular over $(-\pi, \pi]$), subject to the conditioning $\hat{E}[\cos\phi[n+1]] = z_{best}[n+1]$. Applying the Gibbs conditioning principle, we obtain that

$$f_{n+1}(\phi) = \frac{e^{\eta\cos\phi}}{Z_{n+1}(\eta)} \left(f_n * g_n\right)(\phi) \ , \quad -\pi < \phi \leq \pi \tag{4.26}$$

where

$$Z_{n+1}(\eta) = \int_{-\pi}^{\pi} e^{\eta\cos\phi} \ (f_n * g_n)(\phi) d\phi$$

When the phase perturbations are small, $f_n * g_n \approx f_n$. In this case, repeated applications of (4.26) preserve the "exp-cosine" form (4.24). For the other extreme case where the phase perturbations are arbitrarily large, $f_n * g_n$ is approximately uniform in $(-\pi, pi]$, and once again f_{n+1} follows the exp-cosine form. In between these two extremes, f_{n+1} is well-approximated by the exp-cosine as shown in Fig. 4.3. This leads us to the following convenient approximation which we use henceforth.

Approximation by exp-cosine density: At the nth step of the algorithm, the phases $\{\phi_i[n]\}$ are approximately iid with density given by

$$f_n(\phi) = \frac{e^{\eta_n \cos\phi}}{I_0(\eta_n)} \ , \quad -\pi < \phi \leq \pi \tag{4.27}$$

where η_n is chosen so that

$$E[\cos(\phi[n])] = \int_{-\pi}^{\pi} \cos\phi \ f_n(\phi) \ d\phi = z_{best}[n] \tag{4.28}$$

From (4.25), it follows that η_n satisfies

$$\frac{I_1(\eta_n)}{I_0(\eta_n)} = z_{best}[n] = \frac{Y_{best}[n]}{N} \tag{4.29}$$

Equation (4.27) is found to approximate the distribution of the $\{\phi_i\}$ extremely well for most reasonable distributions for the phase perturbations $\{\delta_i\}$ $g_n(\delta_i)$. This is illustrated in Fig. 4.4, where the exp-cosine distribution is shown to closely fit a histogram of phases ϕ_i obtained by a Monte-Carlo sstribution is shown to closely fit a histogram of phases ϕ_i obtained by a Monte-Carlo simulation of the beam-formingalgorithm. Also shown for comparison is a best-fit Gaussian distribution for the same data. The simulation used $N = 3000$, $g_n(\delta) = \text{uniform}[-12^\circ, 12^\circ]$ for all iterations, and the histogram corresponds to $z_{best} = 0.63$.

Discussion of open issues: The setting required for the Gibbs conditioning principle is satisfied exactly at iteration 1, when the phases are iid, uniform over $(-\pi, \pi]$. Moreover, the phases after the conditioning are also approximately iid, in that subsets of phases of size k, where k grows more slowly than N, are asymptotically iid. Does this weak notion of iid phases suffice for applying the Gibbs conditioning principle again at iteration 2, and so forth? We may be helped by the fact that, in any given iteration, the role of the random variables X_i is played not by ϕ_i, but by the perturbed phases $\phi_i + \delta_i$, where the $\{\delta_i\}$ are truly iid. Ad-

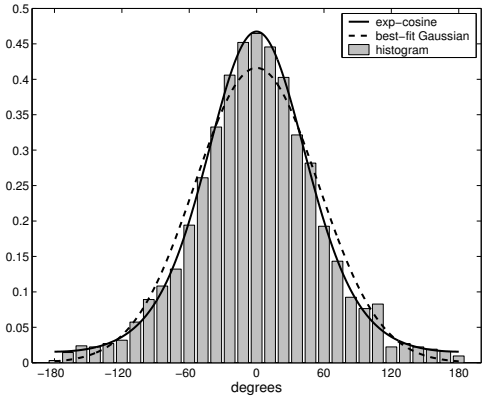


Figure 4.4. Exp-cosine distribution with histogram from simulation.

dressing these questions rigorously appears to require a detailed, problem-specific, large deviations analysis that is beyond the scope of this work.

4.2.2 Dynamical Evolution

We define the averaged sequence $y[n]$ recursively, as the conditional expectation of $Y_{best}[n]$ given $Y_{best}[n-1]$:

$$y[1] = 0 \tag{4.30}$$

$$y[n+1] = \mathsf{E}_{\bar{\delta}[n]}\Big[Y_{best}[n+1]\Big|Y_{best}[n] = y[n]\Big] = y[n] + h(y[n]) \tag{4.31}$$

where $h(y)$ is defined as the expected increase in RSS: $h(y) \doteq \mathsf{E}\big[Y_{best}[n+1]|Y_{best} = y\big] - y$. Equation (4.31) is a discrete time version of the mean ODE for $Y_{best}[n]$:

$$\frac{dy(t)}{dt} = h(y) \tag{4.32}$$

Note that $y[n]$ is a deterministic, monotonically increasing function of time.

From (4.31) we have:

$$Y_{best}[n+1] = \begin{cases} \mathrm{RSS}(\bar{\phi}[n] + \bar{\delta}[n]) & \text{if } \mathrm{RSS}(\bar{\phi}[n] + \bar{\delta}[n]) > y[n] \\ y[n] & \text{otherwise.} \end{cases} \tag{4.33}$$

We now express $\mathrm{RSS}(\bar{\phi}+\bar{\delta})$ as a sum of iid terms from each transmitter, and invoke the Central Limit Theorem (CLT).

$$\mathrm{RSS}(\bar{\phi}+\bar{\delta}) = \Big|\sum_i a_i e^{j\phi_i + j\delta_i}\Big| \tag{4.34}$$

$$= \Big|\sum_i a_i\big(\cos\phi_i\cos\delta_i - \sin\phi_i\sin\delta_i\big) + j\sum_i a_i\big(\cos\phi_i\sin\delta_i + \sin\phi_i\cos\delta_i\big)\Big|$$

$$= \Big|\big(C_\delta \mathrm{RSS}(\bar{\phi}) + x_1\big) + jx_2\Big| = \Big|\big(C_\delta y[n] + x_1\big) + jx_2\Big|, \tag{4.35}$$

$$\text{where } C_\delta \doteq E_\delta\big(\cos\delta_i\big), \tag{4.36}$$

$$x_1 = \sum_i a_i\Big(\cos\phi_i\big(\cos\delta_i - C_\delta\big) - \sin\phi_i\sin\delta_i\Big), \tag{4.37}$$

$$x_2 = \sum_i a_i\Big(\cos\phi_i\sin\delta_i + \sin\phi_i\cos\delta_i\Big) \tag{4.38}$$

where we omitted the time-index n from $\bar{\phi}[n]$ and $\bar{\delta}[n]$ for convenience. The random variables x_1, x_2 are illustrated in Fig. 4.5.

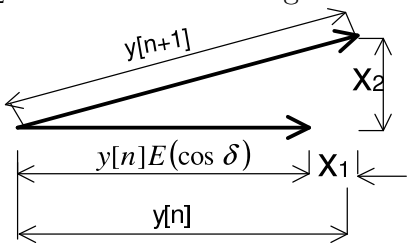


Figure 4.5. Effect of phase perturbations on the total received signal.

Both x_1 and x_2 as defined in (4.35) are linear combinations of iid random variables, $\sin\delta_i$ and $\cos\delta_i$. Therefore as the number of transmitters N increases, these random variables can be well-modeled as Gaussian, as per the CLT [51]. More precisely, the random variables $\frac{x_1}{\sqrt{N}}$ and $\frac{x_2}{\sqrt{N}}$ converge in distribution to zero mean, uncorrelated Gaussian random variables. The variances of $x_1,\ x_2$ are given

by:

$$\sigma_1^2 \doteq \mathrm{Var}(x_1) = \frac{1}{2}\sum_i a_i^2\Big((1-C_\delta^2) - \cos(2\phi_i)(C_\delta^2 - C_{2\delta})\Big)$$
$$\sigma_2^2 \doteq \mathrm{Var}(x_2) = \frac{1}{2}\sum_i a_i^2\Big((1-C_\delta^2) + \cos(2\phi_i)(C_\delta^2 - C_{2\delta})\Big)$$
$$\text{where } C_{2\delta} \doteq E_\delta\big(\cos(2\delta_i)\big) \tag{4.39}$$

With these simplifications, the statistics of $y[n+1]$ only depends on the density function $g_n(\delta_i)$ through C_δ and $C_{2\delta}$. We now specialize to the Line-of-Sight case of $a_i = 1$, and apply the statistical distribution to write

$$\sum_i \cos(2\phi_i) = N\mathsf{E}\left[\cos(2\phi_i)\right] = N\frac{I_2(\eta_n)}{I_0(\eta_n)} \tag{4.40}$$

where we used $k = 2$ in (4.25), and η_n is related to $y[n]$ as $\frac{I_1(\eta_n)}{I_0(\eta_n)} = \frac{y[n]}{N}$. Using (4.40) in (4.39), we have:

$$\sigma_1^2 = \frac{N}{2}\Big((1-C_\delta^2) - \frac{I_2(\eta_n)}{I_0(\eta_n)}(C_\delta^2 - C_{2\delta})\Big)$$
$$\sigma_2^2 = \frac{N}{2}\Big((1-C_\delta^2) + \frac{I_2(\eta_n)}{I_0(\eta_n)}(C_\delta^2 - C_{2\delta})\Big) \tag{4.41}$$

We have the following proposition.

Proposition 3: Assuming that the CLT applies for random variable x_1, the expected value of the received signal strength is given by:

$$y[n+1] \geq y[n]\Big(1 - p(1-C_\delta)\Big) + \frac{\sigma_1}{\sqrt{2\pi}}e^{-\frac{(y[n](1-C_\delta))^2}{2\sigma_1^2}} \tag{4.42}$$
$$\text{where } p = Q\Big(\frac{y[n](1-C_\delta)}{\sigma_1}\Big) \tag{4.43}$$

where the bound in (4.42) is asymptotically exact for large N.

Proof. First we observe that the small imaginary component x_2 of the perturbation mostly rotates the received signal, with most of the increase in $y[n+1]$

coming from x_1 (see Fig. 4.5).

$$\begin{aligned} \mathrm{RSS}(\bar{\phi}+\bar{\delta}) &= \left|C_\delta y[n] + x_1 + jx_2\right| \\ &\geq \left(C_\delta y[n] + x_1\right) \end{aligned} \tag{4.44}$$

Defining p as the probability that $Y_{best}[n+1] > y[n]$, (4.42), (4.43) readily follow from (4.44), (4.33) using Gaussian statistics. The bound in (4.44) is asymptotically exact in the sense that if $\frac{y[n]}{N}$ is kept fixed while N becomes large, then $\left|(C_\delta y[n] + x_1) + jx_2\right| = C_\delta + x_1(1 + \mathcal{O}(\frac{1}{\sqrt{N}}))$. To see this, consider

$$\left|(C_\delta y[n]+x_1)+jx_2\right| \leq C_\delta y[n]+x_1+\frac{x_2^2}{2C_\delta y[n]} = C_\delta y[n]+\sqrt{N}\left(\frac{x_1}{\sqrt{N}}+\frac{1}{\sqrt{N}}\frac{x_1^2+x_2^2}{2C_\delta y[n]}\right) \tag{4.45}$$

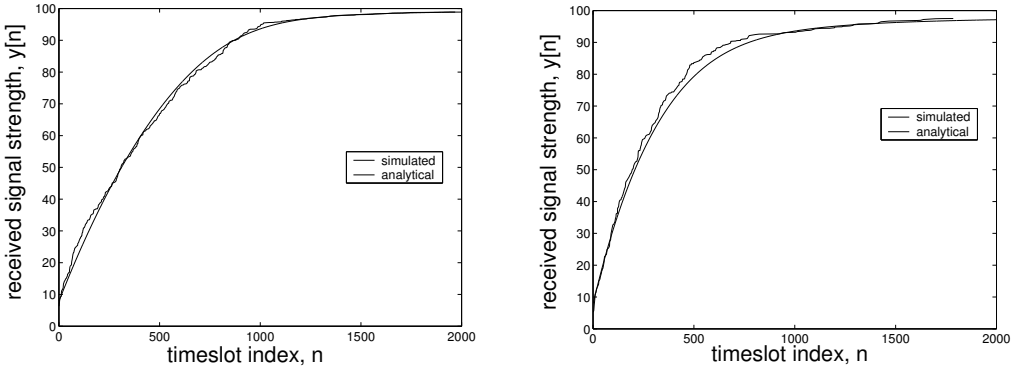
We note that the x_1 term in the parantheses in (4.45) is of the order of $\frac{\sigma_1}{\sqrt{N}}$ which is independent of N (for constant $\frac{y[n]}{N}$), whereas the last term decreases as $\frac{1}{\sqrt{N}}$. $\square$

Using (4.42) in (4.31), we have:

$$h(y) = \sigma_1 f(x), \text{ where } x = \left(\frac{y(1-C_\delta)}{\sigma_1}\right) \text{ and } f(x) \doteq \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} - xQ(x) \tag{4.46}$$

Proposition 3 allows us to iteratively compute the function $y[n]$ for all time-slots n, using the expression in (4.41) for σ_1. Next we use the analytical model to get insight into the convergence of the algorithm, and also to optimize the distribution $g_n(\delta_i)$ of the phase perturbations to maximize the convergence rate.

We now compare the averaged function $y[n]$ obtained from the analytical model with realizations of $Y_{best}[n]$ obtained by simulation of the feedback algorithm. Fig. 4.6 plots both $y[n]$ and an instance of $Y_{best}[n]$ for a network of $N = 100$ transmitters, for two different choices of the distribution $g_n(\delta_i)$: a uniform distribution in $[-\frac{\pi}{30}, \frac{\pi}{30}]$ and a two-point distribution choosing $\pm\frac{\pi}{30}$ with equal probability. The



(a) $g_n(\delta_i) \sim \text{uniform}(-6°, 6°)$ (b) $g_n(\delta_i) \sim \pm 6°$

Figure 4.6. Evolution of received RSS, from analytical model and simulation.

close match with the simulation data in both cases provides validation for the analytical model.

We observe from Fig. 4.6, that the received signal grows rapidly in the beginning, but after $y[n]$ gets to within about 25% of Y_{opt}, the rate of convergence becomes slower.

4.3 Performance Analysis and Optimization

We now use the analytical model to optimize the convergence rate of the feedback algorithm and to study its scalability properties with the number of transmitters N. We show that the convergence rate depends only weakly on the distribution $g_n(\delta_i)$, and near-optimal performance can be achieved with a simple one-parameter family of distributions. We also show the following scalability results:

- The expected received signal strength at any time, always increases when more transmitters are added.

- The number of timeslots required for the expected signal strength to reach

a given level of convergence (e.g. 80% of Y_{opt}) always increases with more transmitters, but increases no faster than linearly in the number of transmitters.

We note that beamforming with a centralized antenna array of N elements, requires $\mathcal{O}(N)$ bits or training symbols to learn N unknown channel gains. Surprisingly the average time to convergence of our simple 1-bit feedback algorithm also scales as $\mathcal{O}(N)$ provided that the distribution $g_n(\delta_i)$ of the phase perturbations is chosen optimally.

4.3.1 Optimizing the Random Perturbations

In Fig. 4.6, we used the same distribution for the perturbations for all iterations of the algorithm. However this choice is not optimal: intuition suggests that it is best to choose larger perturbations initially to speed up the convergence and make the distribution narrower when the phase angles are closer to coherence. We now use the analytical model to dynamically choose the distribution $g_n(\delta_i)$ as a function of $y[n]$.

The general problem of choosing a distribution is a problem in calculus of variations. Fortunately, it is possible to restrict ourselves to a family of distributions without losing optimality, because the analytical model only depends on the distribution through the two parameters $C_\delta, C_{2\delta}$. Furthermore the parameters C_δ, $C_{2\delta}$ are "highly correlated" i.e. determine each other to a large extent. To see this, recall from (4.36) and (4.39) the definitions of C_δ and $C_{2\delta}$ as the expected values of $\cos(\delta_i)$ and $\cos(2\delta_i)$ respectively. We are interested in δ_i corresponding to small random perturbations i.e. $\delta_i \ll \frac{\pi}{2}$. For such small values of δ_i, $\cos(\delta_i)$ and $\cos(2\delta_i)$

are very well approximated by the first two terms of the Taylor series:

$$\cos\delta \approx 1 - \frac{\delta^2}{2} \text{ and } \cos(2\delta) \approx 1 - 2\delta^2, \text{ if } |\delta| \ll \frac{\pi}{2} \tag{4.47}$$

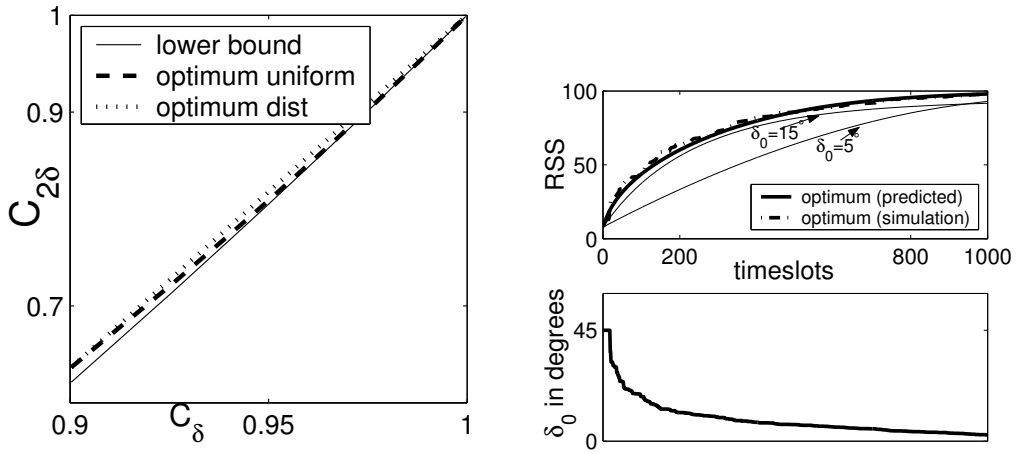
Equation (4.47) indicates that both C_δ and $C_{2\delta}$ are essentially determined by the second moment of δ_i. More precisely we can bound $C_{2\delta}$ as:

$$C_{2\delta} = \mathsf{E}\left[\cos(2\delta_i)\right] = \mathsf{E}\left[2\cos^2\delta_i - 1\right] \geq 2C_\delta^2 - 1 \tag{4.48}$$

where we used Jensen's Inequality to get $\mathsf{E}\left[\cos^2\delta_i\right] \geq \left(\mathsf{E}\left[\cos\delta_i\right]\right)^2 = C_\delta^2$. Equation (4.47) indicates that the bound in (4.48) is tight for distributions of interest. Therefore in practice we can get near-optimal results even if we restrict ourselves to a convenient one-parameter family of distributions e.g. the uniform distribution in $\delta_i \in (-\delta_0, \delta_0]$ or the two-point distribution $\delta_i \in \{\pm\delta_0\}$. The two-point distribution, in fact achieves the lower-bound in (4.48), and gives the slowest convergence rate of all distributions with the same C_δ.

Fig. 4.7(a) shows plots of the $(C_\delta, C_{2\delta})$ pair from a simulation with $N = 100$ over 1500 time-slots, where at each time-slot the expected convergence rate (i.e. $h(y[n])$) was optimized numerically (1) over the family of uniform distributions uniform$(-\delta_0, \delta_0)$ corresponding to different choices of δ_0, and (2) over a general distribution specified by two parameters p and δ_0 as $\Pr(\delta = 0) = 1 - p$, $\Pr(\delta = \pm\delta_0) = \frac{p}{2}$. The lower bound from (4.48) is also shown. It is clear from these plots that a simple one-parameter distribution is sufficient for near-optimal performance, as expected from (4.47).

Fig. 4.7(b) shows the convergence of $y[n]$ with $g_n(\delta_i) = \text{uniform}(-\delta_0, \delta_0)$ for different fixed values of δ_0. In addition Fig. 4.7(b) also shows $y[n]$ and δ_0 when the parameter δ_0 is chosen dynamically to optimize the convergence rate at each time-slot n as given by Proposition 3. This confirms our intuition that at the initial



(a) Optimum $C_{2\delta}$ as a function of C_δ

(b) Comparison of convergence rate with optimum and fixed $g_n(\delta_i)$

Figure 4.7. Optimizing the convergence rate of beamforming algorithm.

stages of the algorithm, it is preferable to use larger perturbations (corresponding to large δ_0), and when $y[n]$ gets closer to Y_{opt}, it is optimum to use narrower distributions (corresponding to smaller δ_0). In all cases, the convergence rate decreases with time, with near linear increase in RSS observed in the initial stage.

4.3.2 Scalability Results

We now establish the scalability of the algorithm for large N.

Theorem 2: We consider the convergence rate of the algorithm with N_1 and N_2 transmitters, where $N_2 > N_1$, and a common sequence of distributions $g_n(\delta_i[n])$ is used in both cases. If $y_1[n]$ and $y_2[n]$ represent the expected received signal magnitude at timeslot n with N_1 and N_2 transmitters, then the following

holds for all n:

$$y_2[n] \geq y_1[n] \tag{4.49}$$

$$\text{and } \frac{y_1[n]}{N_1} \geq \frac{y_2[n]}{N_2} \tag{4.50}$$

Proof. We offer a proof by induction. From (4.30), we have $y_2[1] = 0 \equiv y_1[1]$, and $\frac{y_1[1]}{N_1} = \frac{y_2[1]}{N_2}$. To prove (4.49), we need to show that $y_2[n+1] \geq y_1[n+1]$ given $y_2[n] \geq y_1[n]$.

We write $y_1[n+1] = F_1(y_1[n]) \doteq y_1[n] + h_1(y_1[n]),\ y_2[n+1] = F_2(y_2[n]) \doteq y_2[n] + h_2(y_2[n])$ where $F_1(y),\ h_1(y)$ and $F_2(y),\ h_2(y)$ are defined as in (4.46) for N_1 and N_2 transmitters respectively. Note that $F_1(y_1[n])$ $(F_2(y_2[n]))$ depends on the time index $n$espectively. Note that $F_1(y_1[n])$ $(F_2(y_2[n]))$ depends on the time index n not only through $y_1[n]$ $(y_2[n])$, but also through the distribution $g_n(\delta_i)$. We have suppressed this additional time-dependence to keep the notation simple. By their definition, the functions $F_1(y)$ and $F_2(y)$ are monotonically increasing in y; furthermore $F_1(y) > F_2(y),\ \forall y > 0$. To see this we observe from the definition of the variable x_1 in (4.37) that for the same value of y, the variance σ_1 of x_1 is larger for larger N, and $h(y)$ in (4.46) increases with σ_1.

We are now ready to complete the proof of (4.49) by induction. Given that $y_2[n] > y_1[n]$, we have:

$$y_2[n+1] \equiv F_2\big(y_2[n]\big) > F_1\big(y_2[n]\big) > F_1\big(y_1[n]\big) \equiv y_1[n+1] \tag{4.51}$$

This completes the proof of (4.49).

Now we assume $\frac{y_1[n]}{N_1} \geq \frac{y_2[n]}{N_2}$, or $y_1 \geq \frac{y_2 N_1}{N_2}$. Using Proposition 3 and the definition of $h(y)$ from (4.31), we have:

$$\frac{h(y)}{N} = \mathsf{E}\big[|C_\delta \frac{y}{N} + \frac{x_1}{N} + j\frac{x_2}{N}|\big] - \frac{y}{N} \tag{4.52}$$

In (4.41), we note that η_n is determined by $\frac{y}{N}$, and therefore if both $g_n(\delta_i)$ and $\frac{y}{N}$ are fixed, the variances of $\frac{x_1}{N}$, $\frac{x_2}{N}$ both decrease as $\frac{1}{N}$. Therefore $\frac{h(y)}{N}$ decreases with increasing N for a fixed $\frac{y}{N}$ i.e. $\frac{h_1(y)}{N_1} > \frac{h_2(yN_2/N_1)}{N_2}$. We have

$$\begin{aligned}
\frac{y_1[n+1]}{N_1} &\equiv \frac{F_1(y_1[n])}{N_1} \\
&\geq \frac{F_1(\frac{y_2[n]N_1}{N_2})}{N_1} = \frac{y_2[n]}{N_2} + \frac{h_1(\frac{y2[n]N_1}{N_2})}{N_1} \\
&> \frac{y_2[n]}{N_2} + \frac{h_2(y_2[n])}{N_2} \\
&\equiv \frac{y_2[n+1]}{N_2}
\end{aligned} \tag{4.53}$$

This proves (4.50). □

Corollary: The scalability relations (4.49) and (4.50) also hold when the transmitters use optimized distributions $g_n(\delta_i)$ in both cases.

Proof. Let $\tilde{y}_1[n]$ and $\tilde{y}_2[n]$ be the expected received signal magnitudes using the respective optimized distributions. We apply Theorem 2 to the case where we use the distribution $g_n(\delta_i)$ optimized for N_1 transmitters in both cases. By definition $\tilde{y}_2[n] \geq y_2[n]$, and $\tilde{y}_1[n] = y_1[n]$, therefore $\tilde{y}_2[n] \geq \tilde{y}_1[n], \forall n$. This proves (4.49). Using the same argument for the distribution $g_n(\delta_i)$ optimized for N_2 transmitters, we can prove (4.50). □

We now consider the time to convergence of the algorithm. We define $T_N(\alpha)$ as the number of timeslots required for $y[n]$ to exceed a given fraction, say $\alpha = 75\%$ of the maximum RSS, which equals $Y_{opt} \equiv N$ with N transmitters. Theorem 2 shows that $T_N(\alpha)$ monotonically increases with N. We now show that when the feedback algorithm is appropriately optimized, $T_N(\alpha)$ increases with N no faster than linearly. We need the following proposition.

Proposition 4: The expected RSS increment $h(y)$ as defined in (4.31), under an optimum choice of the distribution $g_n(\delta_i)$, can be lower-bounded as $h(y) > h_0(\frac{y}{N})$, where $h_0(.)$ is strictly positive and independent of N.

Proof. First we use (4.48) and (4.41) to get a lower-bound for the variance σ_1^2. We have:

$$C_\delta^2 - C_{2\delta} \leq (1 - C_\delta^2) \tag{4.54}$$

$$\text{and therefore } \sigma_1^2 \geq N\frac{(1-C_\delta^2)}{2}\Big(1 - \frac{I_2(\eta)}{I_0(\eta)}\Big) \tag{4.55}$$

$$\text{or } \sigma_1 \geq \frac{N(1-C_\delta^2)}{2\sigma_1}\Big(1 - \frac{I_2(\eta)}{I_0(\eta)}\Big) \tag{4.56}$$

We note that the above bounds are achievable; in particular, the two-point distribution $\Pr(\delta_i = \pm\delta_0) = 0.5$ achieves the bound in (4.56). We now rewrite (4.56) as $\sigma_1 \geq \frac{x}{2\alpha}(1 + C_\delta)(1 - \frac{I_2(\eta)}{I_0(\eta)})$, where we used the definition of x in (4.46). Therefore we have

$$\begin{aligned} h(y) = \sigma_1 f(x) &\geq \frac{1}{2\alpha}\Big(1 - \frac{I_2(\eta)}{I_0(\eta)}\Big)\big(xf(x)(1 + C_\delta)\big) \\ &= h_1(\alpha)(1 + C_\delta)h_2(x) \geq h_1(\alpha)h_2(x) \end{aligned} \tag{4.57}$$

$$\text{where } h_1(\alpha) \doteq \frac{1}{2\alpha}\big(1 - \frac{I_2(\eta)}{I_0(\eta)}\big) \text{ depends only on } \alpha, \tag{4.58}$$

$$\text{and } h_2(x) \doteq xf(x) = \frac{x}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} - x^2Q(x) \text{ depends only on x} \tag{4.59}$$

It is easily shown that the function $h_2(x)$ attains its maximum value h_m, at x_m that satisfies:

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{x_m^2}{2}} = 2x_mQ(x_m) \tag{4.60}$$

Solving (4.60) gives $x_m \approx 0.6120$ and $h_m = \max_{x\in\mathsf{R}^+} h_2(x) = h_2(x_m) \approx 0.1012$. We now show that a distribution $g_n(\delta_i)$ that achieves $x = \frac{y(1-C_\delta)}{\sigma_1} = x_m$ exists for any $\alpha = \frac{y}{N}$ and N. We show this by finding a suitable two-valued distribution for

any α and N. Using (4.56), we have:

$$x = \sqrt{\frac{1-C_\delta}{1+C_\delta}} \frac{\alpha\sqrt{N}}{\sqrt{1-\frac{I_2(\eta)}{I_0(\eta)}}} \tag{4.61}$$

We can show that the RHS of (4.61) equals x_m for the two-valued distribution with the parameter:

$$\delta_0 = 2\arctan\Big(\frac{x_m}{\alpha\sqrt{N}}\sqrt{1-\frac{I_2(\eta)}{I_0(\eta)}}\Big) \tag{4.62}$$

Thus we have shown that it is possible to choose a distribution $g_n(\delta_i)$ that satisfies $h(y) > h_0(\alpha) \equiv h_m h_1(\alpha)$, which proves the proposition. □

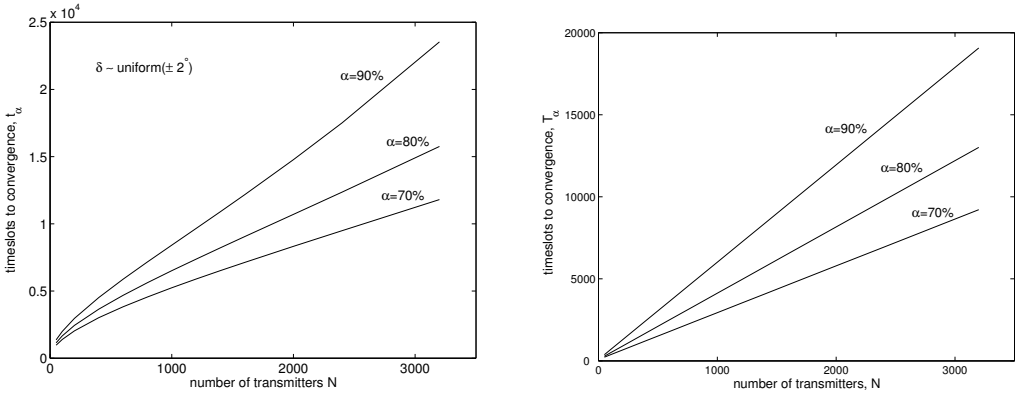
Theorem 3: The number of timeslots required for the feedback algorithm to achieve the fractional level of convergence α satisfies the following:

$$T_N(\alpha) < N \times t_\alpha, \text{ where } t_\alpha \text{ is a constant independent of } N \tag{4.63}$$

Proof. It is more convenient to work with continuous time functions, therefore we first define the function $f_c(t)$ as the piece-wise linear interpolation of $\frac{y[n]}{N}$ i.e. $f_c(t) = \frac{y[t]}{N},\ \forall\ t \in \mathsf{N}$. The function $f_c(t)$ is to be interpreted as representing the expected level of convergence over time.

We note that $f_c(t)$ is a monotonically increasing function of time t, with the range $[0, 1)$. Therefore the inverse function $\tilde{T}_N(\alpha) : [0, 1) \rightarrow \mathsf{R}^+$ of $f_c(t)$ exists. Also $\tilde{T}_N(\alpha)$ is piece-wise linear and continuous at all points in its domain. Although it is not differentiable in the countable set of points $\{y[n]\}$, the one-sided derivative $\frac{d\tilde{T}_N(\alpha)}{d\alpha^+} \doteq \lim_{\Delta\alpha\rightarrow 0^+} \frac{\tilde{T}_N(\alpha+\Delta\alpha)-\tilde{T}_N(\alpha)}{\Delta\alpha}$ exists everywhere in the domain. We can then use Proposition 4 to show:

$$\frac{d\tilde{T}_N(\alpha)}{d\alpha^+} < \frac{N}{h_0(\alpha)} \tag{4.64}$$



(a) Fixed Distribution (b) Optimized Distribution

Figure 4.8. Scalability of beamforming algorithm with number of transmitters.

Therefore we have:

$$\begin{aligned}\tilde{T}_N(\alpha) &= 1 + \int_0^{\alpha} \frac{d\tilde{T}_N(\alpha)}{d\alpha^+} d\alpha \\ &< 1 + N \int_0^{\alpha} \frac{1}{h_0(\alpha)} d\alpha \end{aligned} \tag{4.65}$$

The integral in (4.65) does not depend on N; we denote this integral by $\tilde{t}_\alpha$. Thus we have:

$$\begin{aligned} T_N(\alpha) \equiv \lceil \tilde{T}_N(\alpha) \rceil &< 1 + \tilde{T}_N(\alpha) \\ &< 2 + N\tilde{t}_\alpha < N(\tilde{t}_\alpha + 2) \end{aligned} \tag{4.66}$$

which proves the theorem with $t_\alpha = \tilde{t}_\alpha + 2$. □

Theorem 3 is illustrated by the results in Fig. 4.8, where the number of timeslots required to get within a certain fraction of convergence is plotted against number of transmitters N for a fixed distribution (Fig. 4.8(a)) as well as optimized distributions (Fig. 4.8(b)). These results show that the feedback algorithm is highly scalable with number of transmitters.

Chapter 5

Time-Varying Channels and Other Practical Considerations

The 1-bit feedback algorithm has many attractive features such as simplicity and scalability that make it suitable for real-world networks. In this chapter, we discuss some issues that are relevant for a practical implementation of the algorithm. One of the prerequisites for the algorithm is carrier frequency synchronization, and a stable phase relationship between the transmitters. In Section 5.1, we describe a master-slave architecture similar to Chapter 3 to achieve carrier frequency synchronization; since the feedback algorithm obviates the need for any calibration procedure to estimate the transmitter phase offsets, it is possible to use different frequencies for the reference carrier signal and for beamforming, which substantially simplifies the isolation issues.

The residual phase noise from the master-slave synchronization process, as well as other effects such as channel noise, errors in SNR estimation and channel time-variations all cause some degradation in the beamforming gains. In Section 5.2, we modify the feedback algorithm to dynamically track channel variations in

the presence of noise.

If the channel noise is small and time-variations slow enough that the phase gain can be considered static over several time-slots, the statistical model of Chapter 4 can be expected to provide a good approximation of the rate of convergence of the algorithm when the network first starts tracking. In a sense this is the initial *transient* behavior of the algorithm. To model the *steady-state* behavior we extend the statistical analysis of Chapter 4 to take time-variations and other channel impairments into account. We present the derivation of this model and some results on the steady-state beamforming gain in Section 5.3.

These analytical insights were confirmed in a recent experimental prototype for beamforming based on the feedback algorithm. Some results from this prototype are summarized in Section 5.4.

5.1 Carrier Frequency Synchronization

Fig. 5.1 shows the carrier frequency synchronization process for the slave transmitter. The elimination of the phase calibration step leads to two important simplifications compared to Fig. 3.4. First the synchronization process and the beamforming process are free to use different carrier frequencies i.e. they are frequency division duplexed rather than time-division duplexed. Second the use of the frequency dividers makes it possible to use much smaller frequencies at the phase comparator. This in turn allows the use of the cheaper charge-pump "digital" PLLs.

For instance, let the master transmitter broadcast a reference carrier at $f_1 =$ 1000 MHz. If we use frequency divider ratios of $M_1 = 200$ and $M_2 = 199$, then the synchronized VCO signal has the frequency of 995 MHz. This gives a comfort-

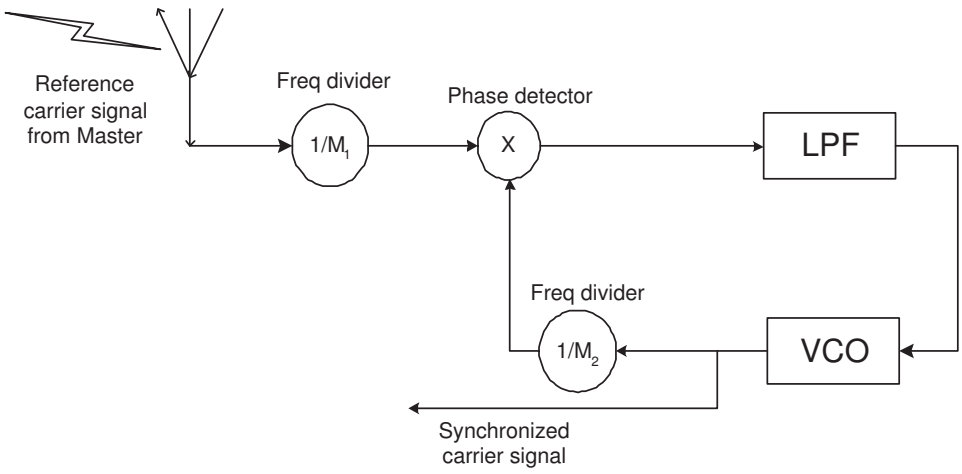


Figure 5.1. Carrier synchronization process in slave transmitter.

able guard interval of 5 MHz between the master-slave synchronization and the beamforming frequencies thereby avoiding isolation issues. The phase comparator frequency is 5 MHz, which is small enough to enable the use of inexpensive charge-pump type PLLs using digital comparators. The main design parameter in the synchronization process is the low-pass loop filter of the slave PLLs. In the case of a charge-pump PLL, the filtering function is provided by the impedance that is driven by the charge-pump, as illustrated in Fig. 5.2 [52]. Unlike the TDD system of Section 3.3, the VCO never operates in open-loop mode, therefore PLL phase error can be reduced to a very small value. Indeed since the PLL is only required to lock to a reference carrier signal i.e. a pure sinusoid, it is tempting to conclude that the phase error can be made arbitrarily small. This is however not the case.

We note that because of the frequency multiplier in Fig. 5.1, the loop phase noise gets multiplied by a factor of M_2 in the slave's carrier signal. Therefore noise suppression in the loop is of crucial importance. This suggests a small bandwidth for the loop filter.

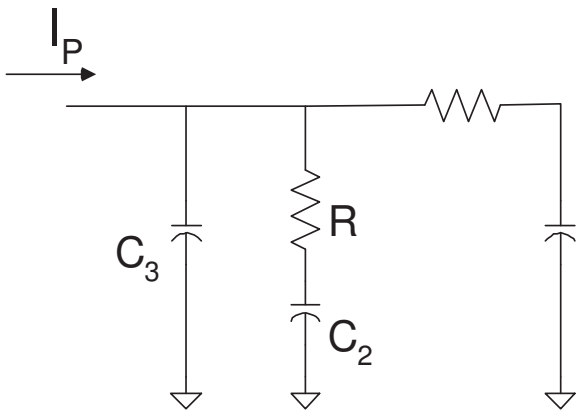


Figure 5.2. Impedance loading for a charge-pump PLL.

We consider a second-order PLL as in Section 3.3, with the closed-loop transfer function given by (3.20) i.e. $H(s) = \frac{s^2}{s^2+2\xi\omega_n s+\omega_n^2}$. The cross-over frequency of of this loop is given by $\omega_c = \frac{K_{VCO}I_P R}{2\pi}$, where K_{VCO} is the sensitivity of the VCO (Hz/V), and I_P is the charge-pump current. The resistance R cannot be made arbitrarily small without reducing the damping ratio ξ and therefore the phase margin. Furthermore the charge-pump undergoes a polarity change each cycle from a current of $-I_P$ to I_P and back [45], and this leads to a frequency excursion or ripple of size $\Delta\omega = \omega_c$. Supressing this ripple requires adding an extra pole (capacitor C_3 in Fig. 5.2) to the second-order loop, which causes a further loss of phase margin. We find that the phase noise in the synchronized carrier cannot be reduced below a certain point.

In general there appears to be a fundamental tradeoff between the loop bandwidth and the achievable SNR from beamforming. A more detailed analysis of the nature of this tradeoff is beyond our scope here.

5.2 Tracking a Time-Varying Channel

In Chapter 4, we derived an analytical model for the convergence of the synchronization when the wireless channels from each transmitter to the receiver is static i.e. constant in time. For such channels, the 1-bit algorithm can be shown to converge asymptotically to perfect coherence with *probability 1.* Once converged, the transmitters can use the optimal value $\theta_{best,i}$ obtained from the algorithm to maintain coherent transmission in subsequent timeslots. However in practical cases, the channel phase responses change in time e.g. due to Doppler effects from moving scatterers. For such channels, the channel variations cause the transmitted signals to lose coherence over time: even when the transmitters use the same phase rotation $\theta_{best,i}$, the received phase $\Phi_{best,i}[n] = \gamma_i + \theta_{best,i}[n] + \psi_i[n]$ will not remain the same, because of the change in the channel phase response $\psi_i[n]$. As a result, the received signal strength $Y_{best}[n] = \left|\sum_{i=1}^{N} a_i e^{j\Phi_{best,i}[n]}\right|$ decreases on average. Fortunately, the 1-bit algorithm can be easily adapted to dynamically adjust the transmitted phase $\theta_{best,i}[n]$. We now present this modified algorithm.

1. At each timeslot n, each transmitter keeps a record $\theta_{best,i}[n]$ of the best known value of its phase rotation, and the receiver keeps an estimate $Z_{best}[n]$ of the best achievable RSS. Unlike the static case, $Z_{best}[n]$ is only an estimate of the best achievable RSS $Y_{best}[n]$ that changes randomly because of channel variations.

2. At timeslot $n+1$, each transmitter generates a random phase perturbation δ_i from some probability distribution $f_\delta(\delta_i)$, and transmits its message signal with an incremental phase rotation δ_i: $\theta_i[n+1] = \theta_{best,i}[n] + \delta_i$. This results

in the received phase:

$$\Phi_i[n+1] = \Phi_{best,i}[n] + \delta_i + \Delta_i[n] \tag{5.1}$$

where $\Phi_{best,i}[n] \doteq \gamma_i + \theta_{best,i}[n] + \psi_i[n]$ and $\Delta_i[n]$ is the channel drift i.e. $\psi_i[n+1] = \psi_i[n] + \Delta_i[n]$.

3. The BS measures the received signal strength, $Y[n+1] = \left|\sum_{i=1}^{N} a_i e^{j\Phi_i[n+1]}\right|$ and generates a single bit of feedback that is set to '1' if the received signal strength in the current timeslot is better than the estimated best RSS $Z_{best}[n]$, and '0' otherwise. The BS then broadcasts this bit of feedback to all transmitters.

4. If the feedback bit is '1', the BS updates its value of $\mathcal{Z}_{best}[n+1]$ with the new measured RSS, and the transmitters update the phase rotations $\theta_{best,i}[n+1]$ to retain the perturbations δ_i; otherwise the BS *discounts* its estimated best RSS $Z_{best}[n]$ by a factor $\rho < 1$ to reflect the expected deterioration due to channel variation, and the transmitters discard the perturbations δ_i.

5. The process is repeated in the next timeslot.

The received phases change due to both the update process and the channel drifts. The update process can be written mathematically as:

$$Z_{best}[n+1] = \begin{cases} Y[n+1], & Y[n+1] > Z_{best}[n] \\ \rho\ Z_{best}[n], & \text{otherwise.} \end{cases} \tag{5.2}$$

$$\theta_{best,i}[n+1] = \begin{cases} \theta_{best,i}[n] + \delta_i[n], & Y[n+1] > Z_{best}[n] \\ \theta_{best,i}[n], & \text{otherwise.} \end{cases}$$

$$\Phi_{best,i}[n+1] = \begin{cases} \Phi_{best,i}[n] + \Delta_i[n] + \delta_i[n], & Y[n+1] > Z_{best}[n] \\ \Phi_{best,i}[n] + \Delta_i[n], & \text{otherwise.} \end{cases}$$

5.3 Modeling the Tracking Performance

Unlike the static case, this tracking version of the 1-bit feedback algorithm does not converge to a fixed $Y_{best}[n]$, but rather to a dynamic steady state. Intuitively, if at any time the received phases $\Phi_i[n]$ become more highly coherent, it becomes harder to find 'favorable' perturbations δ_i, and therefore, the overall tendency for the RSS is to decrease because of the channel drifts. The steady state balances the tendency of the channel drifts Δ_i to drive the phases away from coherence, and this is partly compensated by the random perturbations δ_i with feedback.

To quantitatively analyze this, we model the drift process $\Delta_i[n]$ as iid across transmitters, and stationary and uncorrelated in time with a distribution $f_\Delta(\Delta_i)$. As in Section 4.2.2, we restrict purselves to the Line-of-Sight case where all the transmitters have equal attenuations to the receiver i.e. $a_i = 1$. Much of the analysis of Section 4.2 can now be extended for the time-varying case. In particular, the typicality argument of Section 4.2.1 can be used in this case also, and therefore the empirical distribution of the phases $\phi_i[n]$ at any instant is still given

by (4.27) for large N. (Here $\phi_i \doteq \Phi_{best,i} - \Phi_0$, Φ_0 is defined as in (4.7) as the "centered" value of the phases $\Phi_{best,i}$.) As before we can write the aggregate effect of the phase perturbations, and channel drift as an increase or decrease in the magnitude $Y_{best}[n]$ of the received signal, and a rotation of its phase, and we can write an expression similar to (4.35):

$$Y[n+1] = \Big| C_\delta C_\Delta Y_{best}[n] + z_1 + j z_2 \Big| \tag{5.3}$$

where $C_\Delta \doteq E\big(\cos \Delta_i\big)$. This also suggests a natural choice for the discounting factor as $\rho = C_\Delta$. This choice would make $Z_{best}[n] = E\big(Y_{best}[n]\big)$ in the absence of the perturbations δ_i. As before z_1 and z_2 are uncorrelated, zero mean random variables whose distributions are approximately Gaussian because of the Central Limit Theorem, and their variances can be shown to be respectively:

$$\begin{aligned} \sigma_{11}^2 = \frac{N}{2}\Big(& \big(1 - C_\delta^2 C_\Delta^2\big) \\ & - \big(C_\delta^2 C_\Delta^2 - C_{2\delta} C_{2\Delta}\big) E\big(\cos(2\Phi_{best,i})\big)\Big) \end{aligned} \tag{5.4}$$

$$\begin{aligned} \sigma_{22}^2 = \frac{N}{2}\Big(& \big(1 - C_\delta^2 C_\Delta^2\big) \\ & + \big(C_\delta^2 C_\Delta^2 - C_{2\delta} C_{2\Delta}\big) E\big(\cos(2\Phi_{best,i})\big)\Big) \end{aligned} \tag{5.5}$$

where $C_{2\Delta} \doteq E\big(\cos(2\Delta_i)\big)$. From (5.2) and (5.3), we observe that $Y_{best}[n]$ is a Markov process, and its transition probability function is defined by:

$$f_M\big(y_2 \big| y_1\big) \doteq f\big(Y_{best}[n+1] = y_2 \big| Y_{best}[n] = y_1\big) \tag{5.6}$$

The transition probability function $f_M(y_2|y_1)$ can be expressed in terms of the known Gaussian densities of z_1 and z_2. From $f_M(y_2|y_1)$, we can calculate the steady-state probability density $f_{ss}(y)$ of the Markov chain as the solution to the eigenvalue problem:

$$f_{ss}(y) = \int_{y_1=-\infty}^{\infty} f_M(y|y_1) f_{ss}(y_1) dy_1 \tag{5.7}$$

Fig. 5.3 compares the steady-state distribution $f_{ss}(y)$ computed by solving (5.7), with a histogram of $\mathcal{Y}_{best}[n]$ obtained from a simulation of the 1-bit algorithm with channel time-variations (after discarding the initial "transient" samples). The simulation parameters are $N = 40$, channel drift $f_\Delta(\Delta_i) \sim$ uniform$[-\frac{\pi}{25}, \frac{\pi}{25}]$, phase perturbations $g_n(\delta_i) \sim$ uniform$[-\frac{\pi}{7}, \frac{\pi}{7}]$. The excellent agreement between $f_{ss}(y)$ and the histogram shows that the analytical model accurately predicts the behaviour of the algorithm.

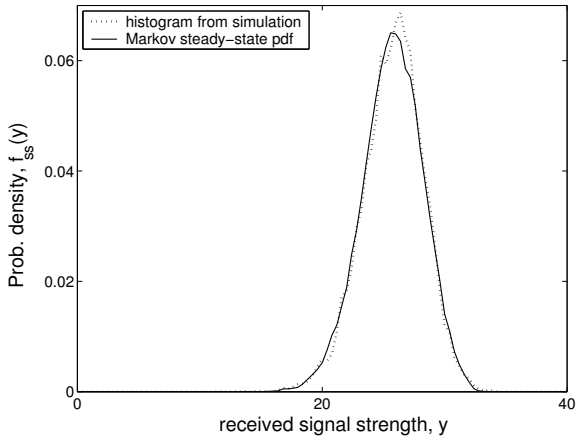
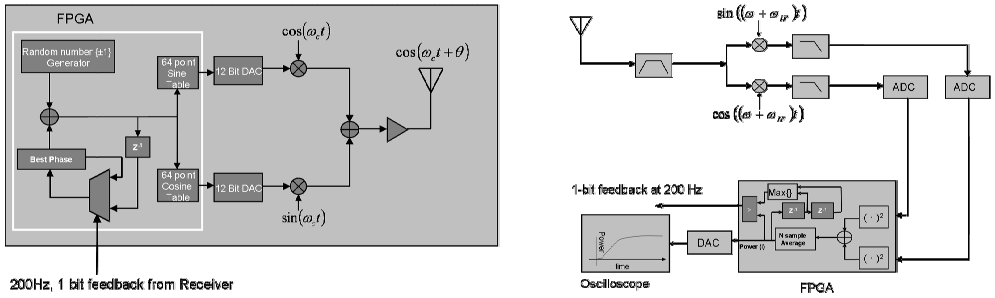


Figure 5.3. Steady-state distribution of RSS.

The above analysis of the tracking algorithm as a Markov process can be used to choose the distribution $g_n(\delta_i)$ optimally to maximize the average steady state RSS. This analysis is beyond the scope of the present analysis, but we can make a few general remarks.

1. In order to get good tracking performance, the perturbations δ_i need to be at least as large as the channel drifts Δ_i on average.

2. The perturbations δ_i should not be too large on average, to avoid large fluctuations in the RSS.



(a) Block diagram of a transmitter

(b) Block diagram of the receiver

Figure 5.4. Baseband functionality for 1-bit algorithm.

3. The effect of phase jitter (i.e. fluctuations in the phase γ_i) is similar to channel variations. If we apply the above steady state analysis to the experimental results of Section 5.4, the steady state beamforming gain correspond to a rms phase jitter of about 10°. This is consistent with a visual observation of the RF carrier signals on an oscilloscope.

5.4 Proof-of-Concept Prototype

We now summarize some results from a recent experimental prototype [4] developed to investigate the performance of the feedback algorithm in a practical situation. This prototype was designed by Ben Wild of UC Berkeley, and its design is illustrated in Fig. 5.4.

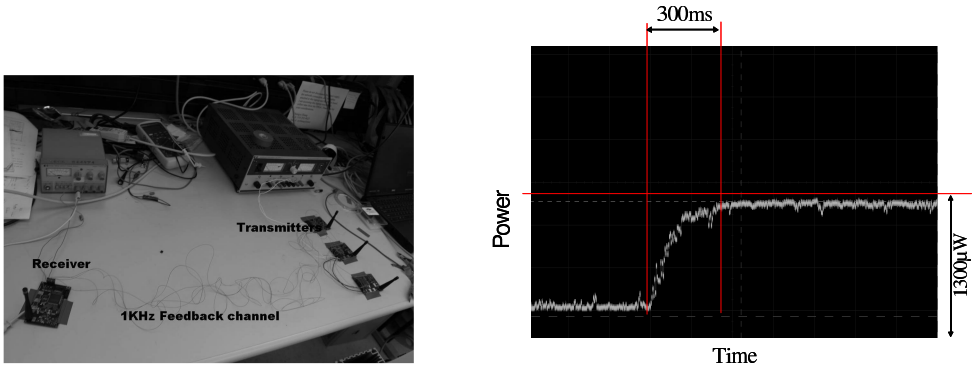
The experiment consisted of mounting $N = 3$ beamformer nodes across from a receiver node. Fig. 5.5(a) shows the setup. First, each transmitter was turned on and the other 2 were turned off, in order to measure their individual received power levels. After this calibration, all of the beamformer nodes were switched on, and the algorithm was started. After convergence we measured the received power and compared this with the theoretical value we were expecting. This allowed us

Transmitters ON	Received Power (μW)
1	120
2	85
3	280
1, 2 and 3	1230

Table 5.1. Received power measurements.

to measure the beamforming gain. Table 5.1 shows the results for one case.

If the beamforming was ideal, the theoretical received power would have been $1370\mu W$ while the actual received power was $1230\mu W$. Thus the results were within 90% of the theoretical limit. We also ran the algorithm while the beamformer nodes BPSK modulated the carrier with a known 10 kbps sequence. The modulation did not affect the convergence time or the beamforming gain as expected. The convergence of the algorithm took approximately 60 iterations. With a 200 Hz feedback rate that corresponded to approximately 300 ms. A typical realization of the received power is shown in Fig. 5.5(b).



(a) Photograph of experimental setup

(b) Received power over time

Figure 5.5. Experimental demonstration of distributed beamforming.

Chapter 6

Conclusions and Future Work

In previous chapters, we motivated the idea of communication using beamforming from a distributed antenna array. We showed that the problem of carrier synchronization was central to its implementation, and presented a solution based on a master-slave architecture. We also presented a feedback algorithm that allows a simple practical implementation and analyzed its properties by theory and simulations. Our main conclusion is that the idea of a distributed antenna array for wireless communication in general, and the technique of collaborative beamforming in particular, offer attractive performance gains compared to conventional methods; and furthermore these gains can be realized in practice using a master-slave architecture for synchronization combined with a iterative adaptation procedure. The adaptation procedure has the virtue of requiring minimal feedback from the receiver, and is also scalable and completely decentralized. Finally, the performance of the beamsteering algorithm can be accurately characterized partly by borrowing the elegant mathematical techniques of statistical physics.

We now offer some observations on other implications of this research for future work.

6.1 Open Technical Issues

Our analysis of the feedback algorithm for beamforming leaves several interesting open problems related to both analysis and design. We now summarize some of these problems.

1. The analytical model presented in Chapter 4 for the convergence of the feedback algorithm was based on the received phase angles following the "exp-cosine" distribution. We justified this distribution using the Gibbs conditioning principle. Unfortunately the version of the Gibbs principle from the literature requires the phases to be conditionally independent, which is not strictly satisfied under the feedback algorithm. An important open problem is making this argument rigorous, either by relaxing the independence assumption, or by generalizing the Gibbs principle.

2. Our feedback algorithm for beamforming is simple by design; it depends on the transmitters making random phase adjustments independent of each other, and also independently over time, and the receiver broadcasting just one bit of feedback. This simple design immediately suggests some possible improvements such as multiple bits of feedback from the receiver, coordinated phase adjustments across the transmitters, or using the past history of phase adjustments to pick the phase adjustments more optimally.

   The last mentioned possibility is especially appealing because it preserves the decentralized nature of the algorithm. A similar idea called *Polyak averaging* is well-known in the stochastic approximation literature (see [41], Ch. 11); such averaging is known to give near-optimum convergence rates in many cases, however it does not *exceed* the performance of a suitably

optimized un-averaged algorithm. A detailed investigation of this idea for the beamforming algorithm is an important subject for future work.

3. The master-slave carrier synchronization process has an important effect on performance of the beamforming algorithm. Our analysis of the tradeoffs involved in the PLL design in Sections 3.3 and 5.1 only scratches the surface of the underlying issues. One open question is the relative advantages of charge-pump and multiplier type PLLs for the overall phase error. On a more basic level, we would like to determine the *fundamental* limits on the achievable synchronization error using any PLL.

4. In Section 5.3, we extended the statistical analysis of Section 4.2 to the modified feedback algorithm for time-varying channels, and showed that it predicts the distribution of the steady-state fluctuations of the RSS accurately. Some authors [44] have suggested that the steady-state RSS distribution may be asymptotically Gaussian. Deriving these and other properties of the steady-state of the tracking algorithm is an interesting open problem. A related problem is generalizing the analysis to model impairments such as SNR estimation error, quantization and channel noise, errors in the feedback channel, and phase jitter from the master-slave synchronization process.

6.2 Connection with Stochastic Approximation

We observe that the feedback algorithm is a distributed version of the well-known class of stochastic approximation (SA) algorithms [41]. This class of algorithms is extremely versatile, and there exists a considerable literature on understanding its convergence properties. With the use of the statistical model of Section 4.2.1, a

mean-ODE can be obtained for the beamforming algorithm (as in (4.32)) similar to any other SA algorithm, and therefore results from the SA literature can be applied to the beamforming problem.

For instance in Section 5.3 we extended the statistical analysis of Section 4.2.1 to the case of time-varying channels. Other authors have exploited the similarity with SA, to reveal additional interesting properties, such as an asymptotic analytic form for the steady-state distribution [44] in a time-varying channel. This may also provide the analytical tools to explore interesting generalizations of the basic feedback algorithm such as the effect of multi-bit feedback, correlated adaptations over time and across transmitters, and the tradeoff between fast tracking and SNR fluctuations in the steady-state.

The connection with SA also suggests the possibility of using the technique of 1-bit feedback to other problems in distributed computing in WSNs. This is similar to the recent work [53] in using iterative techniques to compute medians and quantiles in WSNs.

6.3 Implications for Cooperative Communication

The feedback algorithm can also be generalized to other problems such as beamforming in the presence of interferers; rather than adapting the transmit phases to increase SNR at the receiver, we can also impose additional constraints on interference at other nodes. This would of course require feedback from other nodes as well as the intended receiver. The feedback algorithm may also lose some of its attractive properties in more complicated cases such as the property of *almost sure* convergence from arbitrary starting points.

An interesting extreme case of this generalization is the problem of commu-

nicating from one virtual array to another, as in a MIMO wireless system [27]; instead of focusing one message stream in the direction of a single receiver, this involves simultaneously directing multiple streams of data to multiple destinations. Of course local communication and carrier synchronization would be required separately in both the transmit and receive arrays. Also the transmitters would in general, need to adapt their amplification as well as their phase. The 1-bit feedback could be based on the combined mutual information over all the streams given a set of transmit weights. It would be of great interest to explore the convergence properties of such an algorithm, which could potentially bring some of the popular information theoretic relaying techniques [16] into practical wireless networks. A more modest goal may be to adapt the master-slave synchronization method for the implementation of simpler techniques such as cooperative space-time coding [17] for diversity in fading channels; it is well-known that the lack of carrier synchronization [26] reduces the benefits from cooperation offered by such techniques.

6.4 The Limits of Information Theory

We have seen that information theoretic models of wireless networks naturally give rise to cooperative communication techniques. However the inherently baseband character of the models neglects important aspects of wireless communication systems. In this dissertation, we have demonstrated that one specific form of cooperative communication i.e. collaborative beamforming can indeed be implemented in practice with large achievable SNR gains with a careful design of the RF carrier signal. However we focused on a highly simplified model (Fig. 1.1), that allows us to neglect the overheads of local communication, synchronization

and channel estimation.

6.4.1 Synchronization Overheads

There still remains the question of how these overheads impact the optimality of cooperative schemes. The existing literature on cooperative communication provides a partial answer; for instance Sendonaris et al [17] have shown that cooperative diversity schemes can improve outage rates in fading channels even including the costs of inter-user communication. However this only counts the overhead of message-sharing, and neglects carrier synchronization issues. Unfortunately this limitation is also shared by most of the other related literature [37].

As shown in Chapter 3, carrier *frequency* synchronization can be established with little overhead using a master-slave architecture; it only requires one master node to broadcast a reference tone for all collaborating transmitters to lock to[1]. The reference signal carries no information and therefore occupies zero bandwidth. Beamsteering using the 1-bit feedback algorithm is simple to implement as shown in Section 5.4. However it does require a finite number of time-slots for the SNR gains to be realized. This can be thought of as a training overhead, and becomes proportionally smaller as the length of the message to be transmitted becomes large. A more detailed accounting of the synchronization overheads is beyond our scope, and we leave it as an open problem for future work.

[1]There is some additional cost associated with the complexity of the PLL and other components required for locking to the reference signal. We confine our discussion here strictly to overheads in *communication resources* such as bandwidth and power.

6.4.2 Benefits for Multi-Hop Networks

So far we have focused on the benefits of the SNR gain from beamforming; in multi-hop wireless networks, beamforming offers other advantages that would further discount the cost of the overheads. One advantage follows directly from the SNR gain: because of increased communication range, shorter routes (i.e. fewer hops) may be possible with beamforming. Another advantage is interference mitigation: beamforming results in highly directive transmissions, which leaves a smaller proportion of the transmitted energy as interference.

From this discussion it is clear that a joint design of network routing protocols and the physical layer communication link is necessary to realize the full benefits of distributed beamforming. This is at odds with the layered protocol model that most modern communication networks are based on. However, there has been an increasing appreciation of the costs of "layering" [54], especially for wireless channels, and the idea of *cross-layer design* [55] has been extensively studied for wireless network design.

For example, a two-step strategy was recently proposed [15] that uses broadcast to a set of relays followed by beamforming to the destination. To fully evaluate the optimality of such a strategy, the overheads involved in synchronizing the relays and beamsteering to the receiver need to be estimated. These is already some work in the literature to address such questions; a cross-layer scheme was recently proposed [56] that simultaneously specifies a MAC protocol for propagation of the message signal, and a PHY transmission scheme based on distributed beamforming.

Any comprehensive information theory for wireless networks has to include the entire spectrum of possible communication methods such as direct transmis-

sion, multi-hop routing and cooperative relaying, including combinations of these methods. Existing multi-user information theory is inherently a baseband theory; explicit consideration of the overheads of cooperation may be one way to fix what has been called [12] an "unconsummated union" between information theory and networks.

Bibliography

[1] G. Barriac, R. Mudumbai, and U. Madhow, "Distributed beamforming for information transfer in sensor networks," in *IPSN'04: Proc. of the Third International Symposium on Information Processing in Sensor Networks*, 2004, pp. 81–88. 1.1

[2] R. Mudumbai, G. Barriac, and U. Madhow, "On the feasibility of distributed beamforming in wireless networks," *IEEE Trans. on Wireless Commun.*, vol. 6, no. 5, pp. 1754–1763, May 2007. 1.1

[3] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Scalable feedback control for distributed beamforming in sensor networks," in *Proc. IEEE Intl. Symp. on Inform. Theory, ISIT'05*, Sept 2005, pp. 137–141. 1.1

[4] R. Mudumbai, B. Wild, U. Madhow, and K. Ramchandran, "Distributed beamforming using 1 bit feedback: from concept to realization," in *Proc. of 44'th Allerton Conference on Communication, Control and Signal Processing*, Sept 2006. 1.1, 4.1, 5.4

[5] T. K. Sarkar, R. Mailloux, A. A. Oliner, M. Salazar-Palma, and D. L. Sengupta, *History of Wireless*. IEEE Press, 2006. 2.1

[6] P. J. Nahin, *The science of radio*. AIP, 2001. 1

[7] N. Abramson, "The ALOHA System - another alternative for computer communications," in *Proc. of AFIPS FJCC*, 1970, pp. 281–285. 2.1.1

[8] L. G. Roberts and B. D. Wessler, *"The ARPA Network,"*. Computer Commun. Networks, N. Abramson and F. F. Kuo (Eds.), 1973. 2.1.1

[9] C.-Y. Chong and S. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003. 2.1.1, 2.1.2

[10] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, 2001, pp. 2033–2036. 2.1.2

[11] J. M. Kahn, R. H. Katz, and K. S. J. Pister, "Next century challenges: mobile networking for smart dust," in *MobiCom 99: Proc. 5th ACM/IEEE Intl. Conf. on Mobile computing and networking*, 1999, pp. 271–278. 2.1.2

[12] A. Ephremides and B. Hajek, "Information theory and communication networks: an unconsummated union," *IEEE Trans. on Inform. Theory*, vol. 44, no. 6, pp. 2416–2434, 1998. 2.1.2, 6.4.2

[13] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. on Inform. Theory*, vol. 46, no. 2, pp. 388–404, 2000. 2.1.2

[14] A. Ozgur, O. Leveque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Trans. on Inform. Theory (submitted)*, June 2007. [Online]. Available: http://www.eecs.berkeley.edu/~dtse/hierarchical_0612.pdf 2.1.2

[15] A. Dana and B. Hassibi, "On the power efficiency of sensory and ad hoc wireless networks," *IEEE Trans. on Inform. Theory*, vol. 52, no. 7, pp. 2890–2914, 2006. 2.1.2, 6.4.2

[16] M. Gastpar and M. Vetterli, "On the capacity of large gaussian relay networks," *IEEE Trans. on Inform. Theory*, vol. 51, no. 3, pp. 765–779, 2005. 2.1.2, 6.3

[17] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part I. System description," *IEEE Trans. on Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov 2003. 2.1.2, 2.2.1, 2.4, 6.3, 6.4.1

[18] J. D. Kraus, *Antennas, Second Edition*. Mc-Graw Hill, 1988. 2.2

[19] S. Uda and Y. Mushiake, *Yagi-Uda Antennas*, 1954. 2.2

[20] K. Yao, R. Hudson, C. Reed, D. Chen, and F. Lorenzelli, "Blind beamforming on a randomly distributed sensor array system," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1555–1567, 1998. 2.2, 2.2.1

[21] J. Winters, "Smart antennas for wireless systems," *IEEE Personal Communications [see also IEEE Wireless Communications]*, vol. 5, no. 1, pp. 23–27, 1998. 2.2

[22] A. R. Thompson, B. G. Clark, C. M. Wade, and P. J. Napier, "The Very Large Array," *Astrophysical Journal Supplement Series*, vol. 44, pp. 151–167, Oct. 1980. 2.2.1

[23] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Oct. 1996. 2.2.1

[24] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998. 2.2.1

[25] J. Laneman and G. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. on Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct 2003. 2.2.1

[26] R. A. Iltis, S. Mirzaei, R. Kastner, R. E. Cagley, and B. T. Weals, "Carrier offset and channel estimation for cooperative MIMO sensor networks," in *IEEE Global Telecommunications Conference*, no. 1930-529X, 2006, pp. 1–5. 2.2.1, 6.3

[27] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions of Telecommunications*, vol. 10, pp. 585–595, December 1999. 2.2.2, 6.3

[28] D. J. Love, J. R. W. Heath, W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, 2004. 2.2.2

[29] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. on Inform. Theory*, vol. 47, no. 6, pp. 2632–2639, 2001. 2.2.2

[30] C. Coleman, "Application of the reciprocity theorem to complex propagation problems," in *IEEE Intl. Symp. Ant. and Prop. Soc.*, vol. 1, 2001, pp. 444–447. 2.2.2

[31] G. Barriac and U. Madhow, "Space-time communication for OFDM with implicit channel feedback," *IEEE Trans. on Inform. Theory*, vol. 50, no. 12, pp. 3111–3129, 2004. 2.2.2

[32] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication.* University of Illinois Press, 1949. 2.3

[33] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. on Inform. Theory*, vol. 20, no. 1, pp. 2–10, 1974. 2.3

[34] D. Blackwell, L. Breiman, and A. Thomasian, "The capacities of certain channel classes under random coding," *Ann. Math. Statist.*, vol. 31, pp. 558–567, 1960. 2.3

[35] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. on Inform. Theory*, vol. 25, no. 5, pp. 572–584, 1979. 2.3

[36] M. Dohler, J. Dominguez, and H. Aghvami, "Link capacity analysis for virtual antenna arrays," in *Proceedings 56th IEEE Vehicular Technology Conference VTC*, vol. 1, 2002, pp. 440–443. 2.4

[37] A. Ozgur, O. Leveque, and D. Tse, "How does the information capacity of ad hoc networks scale?" in *Proc. of the 44th Annual Allerton Conf. on Communication, Control and Signal Processing*, Sept 2006. 2.4, 6.4.1

[38] A. Sayeed and T. Sivanadyan, "Source-channel communication protocols and tradeoffs in active wireless sensing," in *Proc. of 44'th Allerton Conference on Communication, Control and Signal Processing*, Sept 2006. 2.4

[39] H. Ochiai, P. Mitran, H. Poor, and V. Tarokh, "Collaborative beamforming for distributed wireless ad hoc sensor networks," *IEEE Trans. on Signal Process. [see also IEEE Trans. on Acoustics, Speech, and Signal Processing]*, vol. 53, no. 1053-587X, pp. 4110–4124, 2005. 2.4

[40] Y.-S. Tu and G. Pottie, "Coherent cooperative transmission from multiple adjacent antennas to a distant stationary antenna through AWGN channels," in *Proc. 55th IEEE Vehicular Technology Conference*, vol. 1, Spring 2002, pp. 130–134. 2.4

[41] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications*. Springer-Verlag, NY, 1997. 2.4, 4.2.1, 2, 6.2

[42] H. Robbins and S. Monro, "A stochastic approximation method," *Annals of Mathematical Statistics*, vol. 22, pp. 400–407, 1954. 2.4

[43] B. Banister and J. Zeidler, "Feedback assisted stochastic gradient adaptation of multiantenna transmission," *IEEE Trans. on Wireless Commun.*, vol. 4, no. 3, pp. 1121–1135, 2005. 2.4

[44] J. A. Bucklew and W. A. Sethares, "Convergence of a class of decentralized beamforming algorithms," in *IEEE/SP 14th Workshop on Statistical Signal Processing*, 2007, pp. 123–125. 2.4, 4, 6.2

[45] H. Meyr and G. Ascheid, *Synchronization in Digital Communications*. Wiley and Sons, NY, 1990. 3.2, 3.2.2, 3.3, 3.3.1, 5.1

[46] D. Brown, G. Prince, and J. McNeill, "A method for carrier frequency and phase synchronization of two autonomous cooperative transmitters," in *IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, 2005, pp. 260–264. 3.2.2

[47] B. Razavi, "A study of phase noise in CMOS oscillators," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 3, pp. 331–343, 1996. 3.3.1

[48] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. on Inform. Theory*, vol. 20, no. 5, pp. 591–598, 1974. 3.3.2

[49] E. Lukacs, *Stochastic Convergence, Second Edition*. Academic Press, NY, 1975. 4.1.1

[50] A. Dembo and O. Zeitouni, "Refinements of the Gibbs conditioning principle," *Prob. Theory Rel. Fields 104*, pp. 1–14, 1996. 4.2.1, 4.2.1

[51] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 1984. 4.2.2

[52] ADIsimPLL. Analog Devices. [Online]. Available: http://www.analog.com/Analog_Root/static/pdf/RFComms/adisimpll.pdf 5.1

[53] R. Rajagopal, M. J. Wainwright, and P. Varaiya, "Universal quantile estimation with feedback in the communication-constrained setting," in *IEEE Intl. Symp. on Inform. Theory*, 2006, pp. 836–840. 6.2

[54] J. Crowcroft, I. Wakeman, Z. Wang, and D. Sirovica, "Is layering, harmful?" *IEEE Network*, vol. 6, no. 1, pp. 20–24, 1992. 6.4.2

[55] S. Shakkottai, T. Rappaport, and P. Karlsson, "Cross-layer design for wireless networks," *IEEE Communications Magazine*, vol. 41, no. 10, pp. 74–80, 2003. 6.4.2

[56] L. Dong, A. P. Petropulu, and H. V. Poor, "Cooperative beamforming for wireless ad hoc networks," 2007. 6.4.2

Appendix A

Proofs for Chapter 3:

Proof of Proposition 1: We can rewrite (3.2) as follows:

$$P_R = N \left\| \frac{1}{N} \sum_{i=1}^{N} \left|h_i\right|^2 e^{j\phi_i} \right\|^2 \tag{A.1}$$

Invoking the law of large numbers, and the fact that the $\{\left|h_i\right|^2\}$, are i.i.d. exponential random variables which are independent from the i.i.d $\{\phi_i\}$, we have

$$\frac{1}{N} \sum_{i=1}^{N} \left|h_i\right|^2 e^{j\phi_i} \rightarrow E\left[\left|h_i\right|^2 (\cos\phi_i + j\sin\phi_i)\right] \quad a.s. \tag{A.2}$$

The expectation on the RHS of (A.2) simplifies as follows

$$\begin{aligned} E\left[\left|h_i\right|^2 (\cos\phi_i + j\sin\phi_i)\right] &= E[\left|h_i\right|^2] E[\cos\phi_i] \\ &= E[\cos\phi_i] \end{aligned} \tag{A.3}$$

We have assumed that ϕ_i is symmetrically distributed around 0, and hence $E[\sin\phi_i] = 0$. Equation (A.3) results because $h_i \sim CN(0,1)$ and hence $\left|h_i\right|^2$ is exponential with unit mean. We thus have that

$$\left\| \frac{1}{N} \sum_{i=1}^{N} \left|h_i\right|^2 e^{j\phi_i} \right\|^2 \rightarrow (E[\cos(\phi_i)])^2 a.s. \tag{A.4}$$

since continuous functions of variables which are converging almost surely also converge almost surely, and the desired result follows. □

Proof of Proposition 2: The expected value of P_R can be written as

$$\begin{aligned} E[P_R] &= \frac{1}{N} E\left[\sum_{i=1}^{N} \left|h_i\right|^2 e^{j\phi_i} \sum_{l=1}^{N} \left|h_l\right|^2 e^{-j\phi_l}\right] \\ &= \frac{1}{N}(N + \frac{N(N-1)}{2} E[\left|h_1\right|^2 \left|h_2\right|^2 \\ &\quad .2\Re(e^{j(\phi_1-\phi_2)})] \end{aligned} \tag{A.5}$$

$$\begin{aligned} &= \frac{1}{N}\left(N + \frac{N(N-1)}{2} 2E[\cos(\phi_1 - \phi_2)]\right) \\ &= 1 + (N-1)E[\cos(\phi_1 - \phi_2)] \\ &= 1 + (N-1)E[\cos(\phi_1)\cos(\phi_2) \\ &\quad - \sin(\phi_1)\sin(\phi_2)] \\ &= 1 + (N-1)E[\cos(\phi_i)]^2 \end{aligned} \tag{A.6}$$

where we have used the fact that the $\{h_i\}, \{\phi_i\}$ are i.i.d. and that the $\{\phi_i\}$ are symmetrically distributed around 0. □

Proof of Proposition 3: We once again begin with the definition for P_R.

$$\begin{aligned} P_R &= \frac{1}{\sqrt{N}} \left\| \sum_{i=1}^{N} \left|h_i\right|^2 \cos(\phi_i) + j\left|h_i\right|^2 \sin(\phi_i) \right\|^2 \\ &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (\left|h_i\right|^2 \cos(\phi_i) - \alpha) \right. \\ &\quad \left. + j\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left|h_i\right|^2 \sin(\phi_i) \ + \sqrt{N}\alpha \right\|^2 \end{aligned} \tag{A.7}$$

where $\alpha = E[\left|h_i\right|^2 \cos(\phi_i)] = E[\cos(\phi_i)]$. Invoking the central limit theorem, as

N gets large, the first term in (A.7) tends to a Gaussian random variable with mean 0 and variance $\sigma_c^2 \equiv \mathrm{Var}[\left|h_i\right|^2 \cos(\phi_i)]$. Similarly, the second term tends to a Gaussian random variable with mean 0 and variance $\sigma_s^2 \equiv \mathrm{Var}[\left|h_i\right|^2 \sin(\phi_i)]$. Since the last term in (A.7) is real and constant, it only shifts the mean of the first Gaussian random variable, so we can write

$$P_R \approx \left|X_c + jX_s\right|^2 \tag{A.8}$$

where $X_c \sim N(\sqrt{N}\alpha, \sigma_c^2)$, and $X_s \sim N(0, \sigma_s^2)$. Making use of the fact that $\left|h_i\right|^2$ is a unit mean exponential random variable,

$$\begin{aligned}
\sigma_c^2 &= \mathrm{Var}[\left|h_i\right|^2 \cos(\phi_i)] \\
&= E[\left|h_i\right|^4 \cos^2(\phi_i)] - E[\left|h_i\right|^2 \cos(\phi_i)]^2 \\
&= 2E[\cos^2(\phi_i)] - E[\cos(\phi_i)]^2
\end{aligned}$$

$$\text{and similarly, } \sigma_s^2 = 2E[\sin^2(\phi_i)] \tag{A.9}$$

Letting $m_c \equiv \sqrt{N}\alpha$, we have that $P_R = X_c^2 + X_s^2$, as given. The variance of P_R follows from standard calculations for moments of Gaussian random variables. □