UNIVERSITY OF CALIFORNIA
Santa Barbara

Bayesian approaches to noncoherent communication: from Shannon theory to practical architectures

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

by

Noah Jacobsen

Committee in Charge:

Professor Upamanyu Madhow, Chair

Professor Jerry D. Gibson

Doctor Roger Peterson

Professor Kenneth Rose

September 2005

The Dissertation of
Noah Jacobsen is approved:

Professor Jerry D. Gibson

Doctor Roger Peterson

Professor Kenneth Rose

Professor Upamanyu Madhow, Committee Chair

July 2005

Bayesian approaches to noncoherent communication: from Shannon theory to practical architectures

by

Noah Jacobsen

Curriculum Vitæ

Noah Jacobsen

Education

Sept. 2005	Ph.D. Electrical and Computer Engineering University of California, Santa Barbara, CA
April 2002	M.S. Electrical and Computer Engineering University of California, Santa Barbara, CA
June 2000	B.S. Electrical Engineering Cornell University, Ithaca, NY

Experience

Sept. 2000–present	Graduate Student Researcher, UCSB Advisor: Prof. Upamanyu Madhow
Winter 2004	Teaching Assistant, Optimal Estimation and Filtering, ECE 240A
Summer 2003	Research Assistant, Yokohama National University, Japan
Summer 2002	Research Intern, Motorola Labs, Schaumburg, IL
Sept. '98–Aug. '99	Research Assistant, Newman Laboratory of Nuclear Studies, Cornell University

Journal Publications

N. Jacobsen, G. Barriac, and U. Madhow, Noncoherent receive beamforming for frequency selective time-varying channels, in preparation.

N. Jacobsen and U. Madhow, Coded noncoherent communication with amplitude/phase modulation: from Shannon theory to practical turbo architectures, under review, *IEEE Trans. Communications*, November 2004

K. Solanki, N. Jacobsen, U. Madhow, B.S. Manjunath, and S. Chandrasekaran, Robust image-adaptive data hiding using erasure and error correction, in *IEEE Trans. Image Processing*, 13(12):1627–1639, December 2004

Abstract

Bayesian approaches to noncoherent communication: from Shannon theory to practical architectures

Noah Jacobsen

Forthcoming wireless cellular systems, such as Wireless Metropolitan Area Networks (WMANs), will have to deliver on promised "wired" bandwidths while overcoming hurdles such as severe channel dispersion, no line of sight between transmitter and receiver, mobility at vehicular speeds, and inter-cell interference arising from shared spectrum scheduling. Moreover, the fading rates that result in time and frequency may prohibit the use of conventional coherent transceiver designs. These issues motivate us to consider spectrally efficient noncoherent communication systems that do not rely on pilot-symbol based estimation of continuously varying channels. Rather, the channel is estimated implicitly, based on the statistics of the received signal and probabilistic models of the fading process. Given the success of belief propagation decoding in a variety of fields, a Bayesian framework for iterative demodulation and decoding is employed to approach the capacity of such channels. With the goal of practical transceiver designs, several reduced complexity implementations, which maintain near Shannon theoretic performance, are proposed. We further introduce a multi-antenna noncoherent *eigenbeamforming* receiver that adaptively learns the spatial channel to a given mobile with little or no pilot overhead. With the key observation that outdoor channels are characterized by relatively few dominant spatial modes, eigenbeamforming receivers enjoy beamforming gains in Signal-to-Noise Ratio (SNR) that result from scaling up the number of receive elements, while simultaneously reducing the complexity of noncoherent demodulation and decoding, which scales with the number of dominant modes. Finally, a side-by-side comparison of noncoherent and coherent transceivers is performed in the context of a packetized Orthogonal Frequency Division Multiplexing (OFDM) system, such as that in development for the IEEE 802.16 WMAN standard.

Contents

List of Figures

Chapter 1

Introduction

Bayesian approaches have been successful in a variety of fields [22, 29, 18], specifically in approaching Shannon-theoretic limits of communication channels using turbo-like codes [32]. The main focus of this dissertation is on using such approaches to attain capacity for wireless time-varying channels.

Time variations in wireless channels arise due to relative motion between the transmitter and receiver, resulting in a well-known apparent frequency shift, referred to as the Doppler shift. A measure of the time period over which the channel is expected to remain roughly constant is given by the inverse of the maximum Doppler shift. This is referred to as the *coherence time* of the channel. The other major contributor to fading in wireless channels is multipath interference. Multipath, or channel dispersion, is the noncoherent superposition of multiple waves at the receiver resulting from reflections, scattering, and diffraction of the transmitted radio wave. Channel variation in frequency results from the various delays associated with multipath components. The delay spread of the channel is given by the maximum delay of any multipath component relative to the first arriving component. Analogous to the definition of coherence time, the *coherence bandwidth* is a measure of the range in frequency over which the channel is roughly constant, and is defined as the inverse of the delay spread.

When the coherence bandwidth is less than the symbol rate, frequency selective fading results, and well-studied techniques of channel equalization are necessary

to decode the received signal (see for example [3]). An alternative approach, that is favored in developing standards, is the use of Orthogonal Frequency Division Multiplexing (OFDM) for modulating the transmitted data symbols. OFDM is a multi-carrier modulation technique that effectively divides a wideband frequency selective fading channel into many frequency non-selective (flat) fading subchannels, with coherence bandwidth greater than the symbol rate. For this reason, and for its suitability to noncoherent demodulation and decoding, we adopt a flat fading channel model, in which the channel is constant over a block of transmitted symbols. Moreover, with an eye towards application to OFDM systems, it is convenient to define a generalized notion of channel coherence that is the product of the coherence time and bandwidth, termed the *coherence length* of the channel.

In particular, we employ the Rayleigh block fading channel model, in which the summation of multiple unresolvable components is modelled with a complex Gaussian random variable that multiplies a block of T consecutive symbols, the coherence length, and is independent and identically distributed (i.i.d.) from block to block. The block fading model is further applicable to certain Time Division Multiple Access (TDMA), frequency hopping, and block interleaved systems. A further advantage of the Rayleigh block fading channel model is that the Shannon capacity was recently solved [27], thus enabling the comparison of constructive coding and modulation schemes to information theoretic limits.

The conventional approach to transceiver design is to estimate the channel using pilots, and then to employ coherent demodulation assuming that the channel estimates are perfect. There are two main drawbacks of this approach: the overhead required for pilots to accurately track rapid channel variations potentially requires a significant fraction of the available bandwidth; and channel estimates based solely on the pilots are suboptimal, since they do not exploit the bulk of the transmitted energy, which is in the data. A number of recent papers [8, 31, 15, 10] consider the alternative of turbo noncoherent communication, with iterative joint

estimation of the channel and data (which does not need pilots, but can incorporate them if available). This body of work is the starting point of this dissertation, which adopts the same basic transceiver architecture: an outer binary code, serially concatenated with a modulation code amenable to noncoherent demodulation [19, 18].

Noncoherent approaches to wireless transceiver design hold the potential for dramatic improvements in the throughput of forthcoming communication systems by eliminating the overhead required to estimate and track a time-varying channel. This gain is most significant for the moderately fast fading, dispersive channels considered here, where the channel coherence length is on the order of a dozen symbols. Such channels are expected of forthcoming wireless metropolitan area networks (WMANs), where pilot overhead can be as much as twenty percent. Indeed, from a more existential point view, pilot based coherent systems are seen to be a suboptimal implementation falling within the noncoherent paradigm, since the channel estimates are inherently noisy and channel estimation ignores the greater part of received signal energy, which is in the data.

Classical approaches to noncoherent communication include orthogonal modulations such as Frequency Shift Keying (FSK) and On/Off Keying (OOK), as well as Differential Phase Shift Keying (DPSK). Orthogonal modulations are bandwidth inefficient, allowing only one signal point per degree of freedom, and thus insufficient for use in modern commercial systems. DPSK, in which the channel is assumed constant over two symbols and information is encoded in symbol phase transitions, is somewhat more bandwidth efficient, but suffers from noise enhancement with classical two-symbol differential demodulation. Warrier and Madhow [38] have generalized differential modulation techniques to classical amplitude/phase constellations, thereby admitting truly bandwidth efficient signals to noncoherent systems. However, many of the design prescriptions obtained in their work, based on maximizing distances in Euclidean signal space with a non-

coherent metric, are found to be inappropriate for the coded systems considered here.

Divsalar and Simon [12] showed that when the channel is constant for a block of symbols, block differential demodulation significantly improves the performance of differential signaling techniques in uncoded systems. Yet such methods would have to be integrated with sophisticated coding techniques to obtain efficient noncoherent communication systems. Peleg and Shamai [31] were the first to do so, with a turbo-architecture for iterative block demodulation and decoding of DPSK signals on the phase-noisy channel. Chen *et al.* [10] further considered constructive coding and modulation schemes for the noncoherent Rayleigh block fading channel, in which near capacity performance was shown to be achievable with certain combinations of QPSK modulation codes and outer channel codes.

We also adopt a turbo-architecture for capacity approaching, spectrally efficient noncoherent communication with large amplitude/phase constellations (or Quadrature Amplitude Modulation (QAM)). Design prescriptions for choice of modulation code and channel code are provided with techniques for reducing complexity in implementations of practical systems. Furthermore, noncoherent demodulation and decoding are generalized to multi-antenna receivers, enabling beamforming gains, traditionally defined in the context of coherent systems, while also maintaining reasonable levels of complexity. The main results of this research are outlined in the following sections.

1.1 Efficient noncoherent communication

We consider the problem of bandwidth-efficient communication over time-varying single-antenna channels with memory, such as those encountered in high data rate outdoor wireless mobile communication. Motivated by the recent success of turbo-like codes and iterative techniques in a variety of classical settings,

we explore methods for the design and analysis of practical coded modulation schemes which approach the information-theoretic limits for such channels.

While the design techniques developed here are quite general, we consider a block fading frequency-nonselective channel model in our performance evaluations: the channel is a scalar complex gain modeled as constant over a block of symbols, with the gain chosen independently from block to block. This model allows for low-complexity noncoherent block demodulation techniques which implicitly estimate the channel gain and phase on each block, and is amenable to information-theoretic computations with which to compare the performance of practical coded modulation schemes.

More importantly, however, the nonselective block fading model is an excellent approximation for existing and projected cellular systems. The slow variation of the channel gain is valid for any system in which the symbol rate is significantly larger than the Doppler spread. Frequency nonselectivity applies, of course, to narrowband systems with bandwidth smaller than the channel coherence bandwidth, but it also applies to each subcarrier in wideband OFDM systems. Moreover, it is easy to efficiently adapt transceiver designs for the block fading model to continuously varying fading channels, as described in Chapter 4.

Specifically, the modulation code considered in our results is a simple generalization of standard differential modulation to QAM alphabets. No pilot symbols are employed. Iterative decoding with soft information exchange between the outer binary decoder and inner noncoherent block demodulator is employed. Our contributions are outlined, in the context of prior work, in the following.

Complexity Reduction: Maximum Likelihood (ML) or Maximum *a posteriori* Probability (MAP) block noncoherent demodulation has complexity that is exponential in the block length, in contrast to the linear complexity of coherent demodulation. One approach to reducing the complexity is to implicitly estimate the channel gain jointly with the data, on a block by block basis. In past

work on block noncoherent demodulation with PSK alphabets [10, 38, 26], this is accomplished simply by quantizing the channel phase into bins, in conjunction with a simple energy-based amplitude estimator. For a coded system as in [10], parallel coherent MAP decoders can be employed, one for each bin, followed by soft-combining of the outputs.[1] However, the simple amplitude estimator in [10] does not work when the signal amplitude varies due to the use of a QAM constellations. Furthermore, maintaining a large number of phase bins implies that the complexity of block noncoherent demodulation is still significantly larger (Q times larger, where Q is the number of phase bins) than that of coherent demodulation.

We address these shortcomings as follows. First, we provide an amplitude estimator that is bootstrapped with conventional two-symbol differential detection, yielding symbol amplitude-level estimates. The estimator is computed only once per block, with sufficient accuracy for noncoherent QAM demodulation that incurs only 0.4 dB loss compared to a *genie-based* system with perfectly known channel amplitude. The *bootstrap stage* also yields initial soft decisions to be passed up to the outer decoder. As for the channel phase, we do quantize it as in [10, 38, 26], and run parallel MAP decoders, but in contrast to prior work, we employ a GLRT-based phase arbitration mechanism based on feedback from the outer decoder to reduce the number of phase bins to two after the first iteration. These simplifications are crucial to enabling efficient noncoherent communication with large QAM alphabets, with the overall system complexity now approaching that of an idealized turbo coded coherent system.

Shannon Theoretic Computations: The capacity of the block fading channel was computed by Marzetta and Hochwald [27]. Roughly speaking, their result indicates that, for moderate and low SNRs, and reasonable channel coherence lengths, independent and identically distributed (i.i.d.) Gaussian inputs are near-

[1]For uncoded noncoherent systems [38], the best phase bin can be jointly estimated, along with the data, based on a Generalized Likelihood Ratio Test (GLRT) criterion.

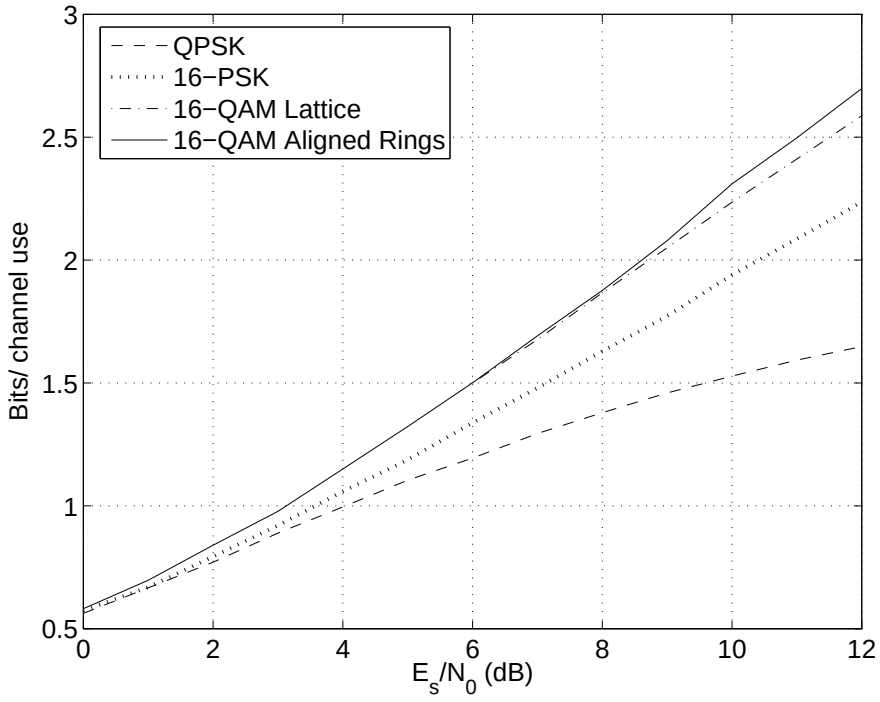


Figure 1.1: 16-QAM constellations significantly improve the noncoherent mutual information for large-SNR.

optimal. Chen *et al.* [10] provide information-theoretic computations showing that this capacity can be approached by the use of standard PSK and QAM constellations. Figure 3.3, computed using the techniques in [10, 27], shows the mutual information versus SNR for 16-ary constellations and QPSK. Evidently large constellations, and moreover, generalized amplitude/phase constellations are required to approach capacity at moderate to large SNR. The Figure further reveals that mutual information is relatively insensitive to constellation shape for QAM constellations for a given number of signal points. For example, the mutual information of the lattice 16-QAM constellation and 16-QAM based on aligned PSK rings (see Figures 2.6 and 2.7) is approximately the same. We therefore need a tool other than Shannon theory for constellation and bit map design in coded noncoherent systems, and for that we turn to a modified form of EXIT analysis.

Modified EXIT Analysis and its Implications: Extrinsic Information Transfer (EXIT) charts [36, 37] are a popular means of obtaining insight into the behavior of systems with a turbo structure, given that they incur far less computational complexity than density evolution techniques [33]. A key tool for simplifying EXIT computations is a Gaussian approximation [37] for the information transferred back and forth between the decoder blocks within a turbo-like structure. One possible intuitive justification is the addition of many contributions in the Log Likelihood Ratio (LLR) domain for a code with long block length. We have modified this methodology to understand the behavior of noncoherent block demodulation, with iterative information exchange with an outer binary decoder. Since the demodulator has a relatively small block length, its output is not well approximated as Gaussian, and is therefore modeled in detail. However, the Gaussian approximation does apply to the output of the outer binary decoder, which operates on a large block length. The resulting EXIT technique allows us to characterize the performance of noncoherent block demodulation for a given signal constellation and bit map, independent of the choice of the outer binary code. The results are employed to provide recommendations for 8- and 16-QAM constellations that are well matched to differential bit maps in phase and amplitude. In particular, for such a bit map, aligned concentric PSK rings have an EXIT curve that completely dominates that of offset concentric PSK rings, regardless of the outer binary code. This is confirmed by simulation results of a coded system, for which aligned concentric rings give the best performance.

Overall Design Summary: In our serially concatenated structure, we use unit rate differential amplitude/phase modulation as the inner modulation code, so that the interleaved bits from the outer binary encoder govern the amplitude and phase transitions between successive transmitted symbols. This inner code has a convolutional structure, so we expect to obtain a turbo effect when it is concatenated with an outer convolutional code. Indeed, the modified EXIT analysis

indicates that, for this choice of inner code, a convolutional outer code is close to optimal, based on a folk theorem [1] that the area under an EXIT chart equals the outer code rate. An alternative configuration which also yields good performance (see [10] for performance results with PSK constellations) is to concatenate a turbo-like binary outer code with *block* differential modulation. In the latter, information is encoded in the amplitude/phase transition of each symbol in the block relative to the amplitude/phase of a single reference symbol. As discussed in Section 2.1.3, however, the complexity of approximate noncoherent demodulation is higher for block differential modulation than for standard differential modulation, hence we focus for the most part on a system in which a convolutional outer binary code is concatenated with standard differential modulation. The signal constellations that we recommend are concentric aligned PSK rings with a differential bit map which is a straightforward extension of DPSK: a subset of bits index the phase transitions, while the other subset indexes amplitude transitions. Such constellations provide significant and realizable gains over DPSK for SNRs greater than 6.5 dB, and for constellation sizes of 16 or larger.

Contrast with prior design recommendations: Our recommendation of aligned PSK rings differs from the standard recommendation for coherent systems without differential modulation, for which offset PSK rings (as well as conventional rectangular QAM) perform better than aligned PSK rings. This also implies that our recommendation differs from prior recommendations for *uncoded* noncoherent communication [38], which can be paraphrased as follows: as the channel coherence length gets large, the best constellation choice (as determined using a noncoherent signal space metric based on minimum distance style arguments) for noncoherent systems is the same as that for coherent systems. The noncoherent signal space metric in [38] also leads to a recommendation for blockwise energy shaping for QAM alphabets, to ensure that the energy per block does not become too low due to a series of low amplitude symbols. However, we find

that such shaping is unnecessary for coded systems, for which it actually leads to performance loss. Overall, we conclude that the recommendations in [38], which are based on minimum distance style concepts relevant for systems operating at very low raw bit error rate (BER), are not applicable for heavily coded systems such as ours designed to operate at relatively high raw BER (1-10%).

1.2 Noncoherent eigenbeamforming

We investigate wideband space-time communication on the uplink of an outdoor cellular system, in which the base station is equipped with N antennas and the mobile has a single antenna. We assume noncoherent reception at the base station, which potentially incurs significantly less overhead than pilot-based estimation of the space-time channel from each mobile to the base station. As is common in outdoor cellular systems, we assume little or no scattering around the base station, so that, from the viewpoint of the base station antenna array, the incoming signal from a given mobile has a narrow Power Angle Profile (PAP). Thus, the spatial channel covariance matrix is typically highly colored, having one or two dominant eigenmodes. The space-time channel for each subcarrier can be modeled as identically distributed complex Gaussian random vectors that decorrelate across frequency [5].

Emerging wideband systems predominantly adopt OFDM techniques for dealing with channel dispersion that may or may not include a line of sight component, as well as logical channelization for multiple access scheduling. For such systems, with large delay spreads and subscribers moving at vehicular speeds, the channel quickly decorrelates in time and frequency. Thus, the amount of pilot overhead that is necessary to maintain accurate channel estimates can be as much as twenty percent. On the other hand, such rich channels implicitly provide feedback regarding the spatial modes of the received signal, that does not require *any* transmitted

pilots. This motivates the consideration of a generalized notion of channel coherence, the coherence length, T, that is given by the product of the coherence time and coherence bandwidth, which for the outdoor channels considered here is on the order of a dozen symbols. We thus employ a block fading channel model in time and frequency to characterize the information theoretic limits of such channels, and block noncoherent demodulation and decoding techniques to provide near capacity performance.

The main contribution of this work is a beamforming receiver that learns the dominant spatial modes to a given user based on the statistics of the received signal, with no pilot overhead (when there is no interference). This receiver offers the major advantage of complexity reduction, especially for outdoor channels where the number of dominant modes is typically one or two [6], and beamforming gain, that is typically expected of coherent reception with channel state information at the transmitter. For the case of an interference component in the received signal, pilots *are* employed to differentiate between the desired and interference signals' *second order* statistics. Note, however, that the pilot overhead required for this purpose is significantly less than that required to estimate the channel realization (first order statistics), as required for coherent reception.

The proposed noncoherent base station receiver architecture is outlined by the following [17]:

(a) Estimate the spatial channel covariance matrix from the covariance matrix of the received signal, averaged across subcarriers. This exploits the observation [5] that the space-time channel for different subcarriers can be modeled as identically distributed random vectors that decorrelate across frequency.

(b) Project the received signal in each subcarrier along the L dominant eigenmodes of the estimated spatial covariance matrix (L is typically much smaller than the number of receive elements N for a typical outdoor channel). This *eigenbeamforming* operation creates L parallel, independently fading, channels for the same

transmitted data.

(c) For each of the L eigenmodes, use noncoherent coded modulation strategies with turbo-like joint data and channel estimation, as in prior work on single antenna channels [31, 10, 19, 20]. In this paper, we propose a suboptimal but effective technique for combining the outputs of the L noncoherent demodulators.

Thus, by appropriately exploiting the covariance estimate available in a wideband system, the preceding noncoherent receiver architecture provides many of the benefits of explicit space-time channel estimation without incurring its overhead. For example, beamforming gains in received SNR (relative to a single antenna system) are realized, while incurring reasonable complexity by using a small number of dominant eigenmodes for demodulation and decoding.

For numerical evaluation of the proposed architecture, we approximate the fading gains for each eigenmode by an independent block fading channel [10, 20]. In an OFDM system, such an approximation might be applied to blocks of contiguous time-frequency bins in which the channel may be approximated as constant. We focus attention on operation at relatively low SNR, using Quadrature Phase Shift Keying (QPSK) constellations. Our main results are as follows:

• We compute the capacity with QPSK signaling of L parallel block fading channels of possibly unequal strengths. This is done in a manner analogous to prior work on symmetric block fading models. Capacity plots showing the diversity gain as a function of the number of dominant eigenmodes are provided.

• We further provide numerical results for iterative joint data and channel estimation for a constructive coded modulation scheme consisting of a convolutional code, serially concatenated with differential QPSK. For multiple dominant eigenmodes, optimal noncoherent processing is excessively complex, so that our numerical results compare the performance of a suboptimal diversity combining scheme with the information-theoretic benchmarks we have obtained.

Finally, a Minimum Mean Squared Error (MMSE) framework for interference suppression is developed and evaluated. The simulation results confirm that relatively few pilots (less than 3%) are required to train MMSE correlators to the dominant modes of the desired user's spatial channel, while performance approaches that of ideal MMSE interference suppression (which assumes infinite training). Thus, when the dominant modes of the signal and interference channels are roughly orthogonal, arising for example when the mean angle of arrival for the two channels is separated by 45°, performance approaches that with no interference, less approximately 0.15 dB of pilot overhead required to train the MMSE correlators.

1.3 Noncoherent OFDM

Finally, we perform a side-by-side comparison of noncoherent and coherent transceivers. Thus, it is worth commenting on the relation between the block fading model employed in our work, and continuously varying fading channels such as the OFDM channel considered in the comparison. For stationary fading, Lapidoth *et al.* [23, 24] show, in a rather general setting, that as SNR tends to infinity, capacity grows only as log log SNR, rather than the well-known log SNR growth for a classical AWGN channel. Thus, this extremely power-inefficient operating regime, which results when both the SNR and the Doppler are extremely high, is to be avoided if at all possible. Also, achieving this double logarithmic growth requires the use of constellations whose shape is very different [9] from the Gaussian, PSK or QAM inputs that work well over the AWGN channel. The double logarithmic growth occurs, roughly speaking, because the effect of the errors in (implicit or explicit) estimation of the channel dominates the effect of noise at high SNR. Fortunately, for outdoor cellular applications, the typical combination of SNRs and fading rates does not fall in this regime [13]; that is, the effect of

errors in implicit or explicit channel estimation are small compared to the effect of the channel noise. The block fading model corresponds to approximating the time-varying channel gain over a block of symbols by a constant scalar; intuitively, we expect the error in this approximation to be small in desirable operating regimes in which the channel gain is varying slowly enough that it can be accurately estimated. Note that the block fading channel has been shown to exhibit log SNR growth at high SNR [25], which implies that conventional power-bandwidth tradeoffs are applicable in this desired operating regime.

We then directly compare the performance of noncoherent and coherent wireless transceiver designs. For this purpose, a packetized OFDM system is employed, such as that of the IEEE 802.16 standard for WMANs. Noncoherent systems typically incur significantly more computational complexity than coherent systems, resulting from Bayesian processing techniques employed for approaching capacity. Thus to facilitate a fair comparison, a detailed analysis of the computational requirements of each approach is performed [21] and system parameters are chosen to yield similar complexity counts. Two main findings from the comparison result:
(1) The raw frame error rate (FER) performance of the coherent system is 0.5-1 dB better than that of the noncoherent approach. Part of this gap clearly stems from limiting the complexity of noncoherent receiver processing. Moreover, the block fading model does not leverage inter-block channel continuity as well as pilot-based channel estimation techniques, and we conjecture that more sophisticated modelling of channel frequency variation will eliminate this gap. Nonetheless, the block fading model works very well in capturing local channel memory within a block and should serve as a good starting point for further transceiver designs.
(2) As expected, the overall throughput of noncoherent systems is superior to coherent systems. This is directly attributable to the twenty percent pilot over head employed by the coherent system. Since throughput is defined with respect to the FER, this advantage will only broaden as techniques for solving (1) are realized.

Chapter 2

Coded noncoherent amplitude/phase modulation

We consider bandwidth-efficient communication using amplitude/phase modulation over a frequency nonselective channel whose time variations model the fading exhibited in outdoor wireless mobile communication. The system is *noncoherent,* not requiring pilots for channel estimation and tracking, and not assuming prior channel knowledge on the part of the receiver. Serial concatenation of a binary outer code with an inner differential modulator provides a turbo structure that, along with the channel memory, is exploited for joint iterative channel and data estimation at the receiver. While prior work on noncoherent systems mainly focuses on PSK alphabets, we consider here a moderate to high SNR regime in which amplitude/phase constellations are significantly more efficient. We first reduce the computational complexity of block noncoherent demodulation to a level comparable with that of standard coherent demodulation. We observe that channel capacity for a block fading channel is relatively insensitive to constellation shape, so that Shannon theory is not adequate for optimizing the choice of constellation and bit map. We provide a tool for such choices, independent of the choice of outer code, by modifying EXIT analysis for noncoherent demodulation. The results are consistent with simulations, and the recommended constellation shapes differ significantly from standard coherent designs, and from prior rec-

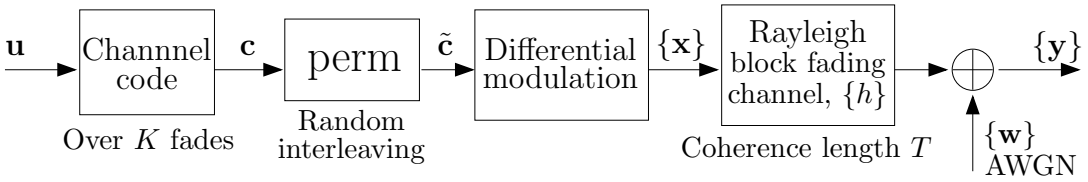


Figure 2.1: Baseband transmitter and channel model.

ommendations for *uncoded* noncoherent communication. The EXIT analysis also indicates that a convolutional outer code is nearly optimal for an inner differential modulator. The overall system is within about 1.8 dB of Shannon capacity for the block fading channel at moderate to large SNR, demonstrating that bandwidth-efficient noncoherent communication systems with reasonable complexity are now within reach.

2.1 Noncoherent transceiver processing

In this section, we describe the channel model, concatenated code and modulation structure, and turbo noncoherent processing of amplitude/phase constellations with complexity reducing techniques.

2.1.1 Channel model

We consider systems in which coding is performed on a scale that is much larger than the channel coherence length, thus leveraging the inherent diversity of a fading channel. Figure 2.1 depicts the complex baseband transmitter and channel model. The binary information sequence $\mathbf{u}$ is mapped to codeword $\mathbf{c}$ of the binary channel code $\mathcal{C}$ and pseudo-randomly permuted to the code-symbol sequence $\tilde{\mathbf{c}} = \{c[n]\}$. With the cardinality of the modulation alphabet, $\mathcal{A}$, equal to M, we adopt the convention that $c[n]$ denotes $m = \log_2(M)$ permuted code-bits which modulate the nth channel symbol, $x[n]$. Codewords in the modulation code, $\mathbf{x} \in \mathcal{M}$, belong in the T-fold product of the symbol alphabet, $\mathcal{A}^T$.

Block fading model: The channel is assumed to be constant over disjoint blocks of T symbol intervals, where T is the coherence length. Channel gains for different blocks are modeled as i.i.d. Letting $\mathbf{x}$ denote a block of T transmitted symbols, the block of received symbols is given by

$$\mathbf{y} = h\mathbf{x} + \mathbf{w}, \tag{2.1}$$

where the channel gain $h = ae^{j\theta}$ is a zero-mean, unit-variance proper complex Gaussian random variable, denoted, $h \sim \mathcal{CN}(0, 1)$. This is a classical Rayleigh fading model, with channel amplitude, a, Rayleigh, channel phase, θ, uniform over $[0, 2\pi]$, and a and θ independent. The additive noise vector, $\mathbf{w}$, is Gaussian, $\mathcal{CN}(0, 2\sigma^2\mathbf{I}_T)$, where $\mathbf{I}_T$ stands for the $T \times T$ identity matrix. The Rayleigh fading model is equivalently defined by the conditional Probability Density Function (PDF) of the received symbols given the transmitted symbols [27]:

$$P(\mathbf{y}|\mathbf{x}) = \frac{\exp\{-\operatorname{tr}([2\sigma^2\mathbf{I}_T + \mathbf{x}\mathbf{x}^{\mathrm{H}}]^{-1}\mathbf{y}\mathbf{y}^{\mathrm{H}})\}}{\pi^T \det(2\sigma^2\mathbf{I}_T + \mathbf{x}\mathbf{x}^{\mathrm{H}})}. \tag{2.2}$$

Since the channel is i.i.d. over blocks, ergodic analysis of a single block of symbols suffices to design systems with near capacity performance. Thus (2.2) completely specifies the information theoretic behavior of the block fading channel.

Block fading approximation to continuously varying channel: Since there is no absolute amplitude and phase reference within a block, the signals over a block of length T live in a $(T-1)$-dimensional manifold [40], which costs a rate penalty of $1/T$. This can be intuitively interpreted as resulting from the use of one symbol in the block as a amplitude/phase reference, or pilot (whether or not this is explicitly done). However, in practice, this rate loss can be avoided when applying the block fading model to a continuously varying channel, by overlapping successive blocks by one symbol. Thus, by including the last symbol of the previous block as the first symbol of the current block, we have $T-1$ new channel uses required for signaling in a $(T-1)$-dimensional manifold. Of course, when applying the block

fading approximation to a continuously varying model, there are two sources of performance loss: first, the approximation error in modeling the channel gain as constant over a block, and second, the loss due to not exploiting the continuity of the channel in adjacent blocks explicitly for channel estimation. However, these losses are expected to be small if the block length T is chosen appropriately, and the operating SNR is not extremely high.

Standard differential modulation (generalized to amplitude/phase constellations) enables demodulation of the transmitted symbol sequence, despite lack of absolute amplitude and phase reference. In an uncoded system, block-wise MAP demodulation of differentially modulated channel data minimizes the symbol error rate, and is thus the optimal bit-level demodulation strategy. In the coded system considered here, block noncoherent demodulation is used to compute the extrinsic *A Posteriori* Probability (APP) of the code-symbols, to be passed back to the outer channel decoder. We also consider classical two-symbol differential demodulation in a bootstrap phase for amplitude estimation. The transmitted symbol sequence is generated from the code-symbol sequence, $\{c[n]\}$, and a reference channel symbol, $x[0]$, with the differential bit-to-symbol mapping, $\nu : \{0,1\}^m \times \mathcal{A} \to \mathcal{A}$, according to

$$x[n] = \nu(c[n], x[n-1]). \tag{2.3}$$

The APP of the code-symbol $c[n]$ may thus be expressed as

$$P(c[n] = c|\mathbf{y}) = \sum_{\mathbf{x}: x[n]=\nu(c,x[n-1])} P(\mathbf{x}|\mathbf{y}). \tag{2.4}$$

Letting $P(\mathbf{x})$ denote the prior distribution on the block of transmitted symbols, the probability that $\mathbf{x}$ was transmitted given the observed symbols, $\mathbf{y}$, is given by Bayes' Law,

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}, \tag{2.5}$$

in terms of the conditional PDF of the received symbols (2.2). In practice, direct evaluation of $P(\mathbf{y}|\mathbf{x})$ is computationally infeasible, with complexity $O(M^T)$, that

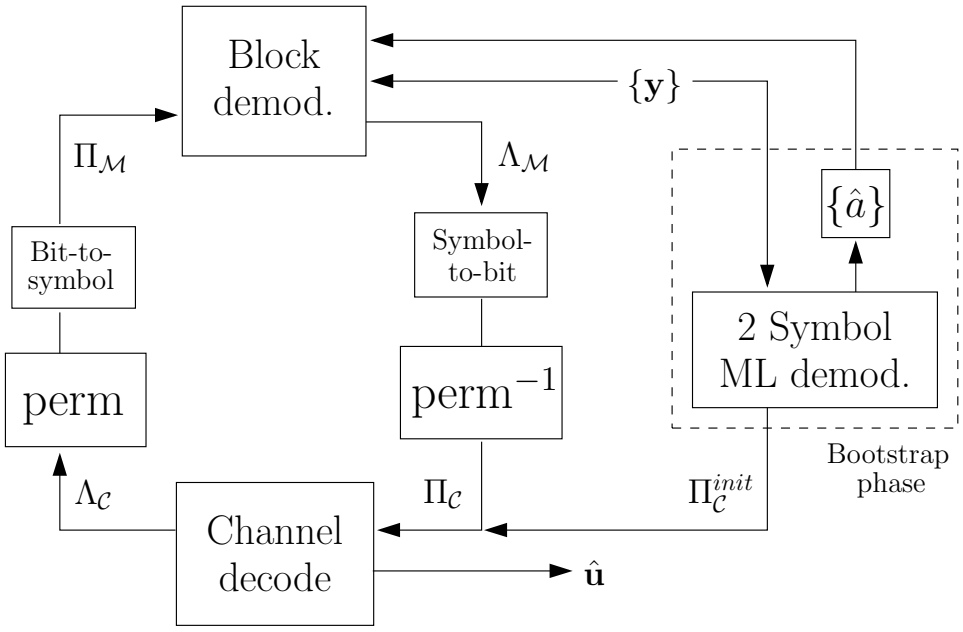


Figure 2.2: Iterative noncoherent demodulation and decoding.

scales exponentially in the coherence length. Following [10, 31], we employ an approximate noncoherent APP demodulator with channel amplitude estimation, phase quantization, and parallel *coherent* BCJR processing. The complexity is polynomial in M.

2.1.2 Turbo noncoherent demodulation and decoding

The data, $\mathbf{u}$, and channel, $\{h\}$, are estimated jointly with "turbo" iterative demodulation and decoding of the received symbol sequence, $\{\mathbf{y}\}$, as illustrated in Figure 2.2. Block demodulation consists of running parallel coherent demodulators, one for each quantized phase bin of the unknown channel phase, over the range of relevant channel rotation, with estimated channel amplitudes, one for each fading block. Note that the composition of unit-rate differential modulation with rotationally invariant constellations comprises a (relatively simple) rotationally invariant modulation code, for which the number quantizer bins required is usually much less than for non-rotationally invariant codes, which require quan-

tization over the full range of $[0, 2\pi]$; see [10]. Thus, the demodulator computes extrinsic APPs, $\Lambda_{\mathcal{M}}$, of the transmitted symbol sequence, based on the observed symbol sequence, $\{\mathbf{y}\}$, and prior probabilities on the transmitted symbols, $\Pi_{\mathcal{M}}$. Soft-Input, Soft-Output (SISO) noncoherent demodulation is described in detail the next section for the general case of amplitude/phase modulation constellations. As in the turbo-decoding algorithm for concatenated binary codes, the channel decoder operates at the code-bit level, producing the extrinsic probabilities $\Lambda_{\mathcal{C}}$ with de-interleaved priors, $\Pi_{\mathcal{C}}$, that are computed from demodulator symbol-level posteriors. Channel decoder posteriors are then re-interleaved and converted back to symbol-level probabilities, for use as priors in the next round of noncoherent demodulation. Random code-bit permutation justifies the independence of prior probabilities assumption for so-called belief propagation decoding of concatenated codes. Demodulation and decoding are thus performed until an accurate estimate of the transmitted data is attained, or complexity constraints are met.

Classical two-symbol Maximum Likelihood (ML) detection of differentially modulated symbols, not requiring any prior channel knowledge, serves to bootstrap the receiver, providing (i) the initial priors to the outer channel decoder and (ii) the probabilities of symbol amplitude levels to the channel amplitude estimator. This *bootstrap phase* is noted in Figure 2.2 with a dashed box. In the following proposed reduced-complexity receiver, the channel decoder is initialized with the bootstrap phase and computes the demodulator's first set of symbol priors. Next, parallel block demodulators, for *all* phase branches, compute code symbol APPs conditional on quantized phase bins. A GLRT criterion, described in Section 2.1.4, is then used to prune all but the two phase branches producing the highest quality of soft information per block. Thus, full phase quantization is employed only during the first iteration, and all subsequent iterations demodulate only the selected subsets of two phase branches. The resulting reduced-complexity receiver

requires only twice as many demodulation computations (after the first iteration) as a genie-aided demodulator that has access to the channel realizations.

2.1.3 Phase quantization

The unknown channel phase is implicitly estimated with Bayesian combining of coherent APPs calculated with quantized phase bins. This approach relies on an energy based estimate of the channel amplitude (described in Appendix B), and an approximation of the code symbol posteriors, $P(c[n]|\mathbf{y}) \approx P(c[n]|\mathbf{y}, a)$, that is conditional on the channel amplitude. We shall see that bit-to-symbol mappings satisfying the condition, $\nu(c[n], e^{j\phi}x[n-1]) = e^{j\phi}\nu(c[n], x[n-1])$, referred to as *$\phi$-rotational invariance*, are desirable for the method of phase quantization. Unit-rate differential modulation with rotationally invariant constellations satisfy the rotational invariance condition and are well-suited to noncoherent processing and amplitude estimation. Thus, let $\phi \in (0, 2\pi]$ denote the smallest angle for which rotation of $\mathcal{A}$ returns $\mathcal{A}$, i.e. $\mathcal{A} = e^{jk\phi}\mathcal{A}, \forall k \in \mathbb{Z}$. Further define $L \triangleq 2\pi/\phi$. The conditional APPs of the transmitted symbols, given the channel amplitude, are computed via Q-fold Riemann approximation (2.7) of the total probability expansion

$$P(c[n]|\mathbf{y}, a) = \frac{L}{2\pi}\int_0^{\phi} d\theta \; P(c[n]|\mathbf{y}, a\exp(j\theta)) \tag{2.6}$$

$$\approx \frac{1}{Q}\sum_{q=0}^{Q-1} P(c[n]|\mathbf{y}, a\exp(j\phi q/Q)). \tag{2.7}$$

Note the limits of integration (2.6) reflect ϕ-rotational invariance of the modulation code. Thus for each quantization phase bin, indexed with q, APP computation is efficiently performed with the standard coherent BCJR algorithm. Maps that do not have the rotational invariance property (e.g. block DPSK [10]) require quantization of the full $[0, 2\pi]$ interval, and thus L times as many demodulation computations.

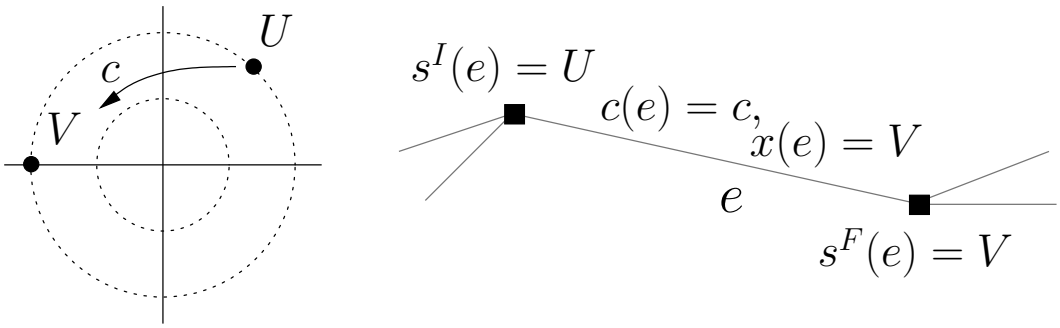


Figure 2.3: An example state transition in the trellis representation of differential modulation.

Viewing ν as a unit rate/memory recursive convolutional code, an equivalent trellis representation is depicted Figure 2.3. To each trellis edge, e, corresponds the initial and final states, input code bits, $c(e)$, and output channel symbol, $x(e)$. The coherent posteriori probability, $\lambda_n(c|h) \triangleq \log P(c[n] = c|\mathbf{y}, h)$, of the code symbol, $c[n]$, is computed with the logarithmic BCJR algorithm:

$$\lambda_n(c|h) = \max_{e:c(e)=c}{}^{\star} \left\{ \alpha_{n-1}(s^I(e)) + \gamma_n(e|h) + \beta_n(s^F(e)) \right\}, \tag{2.8}$$

where the forwards/backwards recursions for α_n and β_n are defined in Appendix A, [2, 7]. The *max-star* function, $\max^{\star}_{Z\in\mathcal{F}}\{Z\} = \log(\sum_{Z\in\mathcal{F}} e^Z)$, corresponds to logarithmic summation. The branch metric $\gamma_n(e)$ of edge e is given by

$$\gamma_n(e|h) = \Pi_n(c(e)) + \sigma^{-2}\, \Re\langle y[n], hx(e)\rangle, \tag{2.9}$$

where prior probabilities of the code-symbols, $\Pi_{\mathcal{M}} = \{\Pi_n\}$, $\Pi_n(c) = \log P(c[n] = c)$, are initially uniform and then set by the outer decoder through turbo processing.

Thus, noncoherent APP demodulation works with the coherent BCJR algorithm as a building block, the amplitude estimate of Appendix B, and Q-level quantization of the unknown channel phase θ in $[0, \phi]$. For each quantization level, $q \in \mathcal{Q} = \{0, 1, \ldots, Q-1\}$, (2.8)$-$(2.9) are evaluated with respect to $h = \hat{a}\exp(j\phi q/Q)$. The resulting (coherent) symbol likelihoods are then aver-

aged, (2.7), to yield extrinsic (noncoherent) APPs of the code symbols,

$$\Lambda_n(c) \triangleq \log\{P(c[n] = c|\mathbf{y}, a = \hat{a})\} - \Pi_n(c) \tag{2.10}$$

$$\approx \max_{q\in\mathcal{Q}}{}^{\star}\{\ \lambda_n(c|h = \hat{a}\exp(j\phi q/Q)\ \} - \Pi_n(c) \tag{2.11}$$

This method of SISO noncoherent APP demodulation enables turbo processing with the outer channel decoder in a serially concatenated system.

2.1.4 Phase selection

The method of phase quantization exhibits near capacity performance on the noncoherent block fading channel, yet each phase branch requires its own BCJR computation per iteration. Such a receiver requires Q times as many demodulation computations as coherent reception of the noncoherent code with the same number of iterations. Our study has shown that a genie-based system, which demodulates only the quantization branch with phase closest to the true channel phase, provides excellent performance. This motivates the development of a criterion for ranking and pruning parallel phase branches as iterative demodulation and decoding are performed.

We propose a Generalized Likelihood Ratio Test (GLRT) for phase branch selection where the observation is the received signal and extrinsic information from the decoder, and the parameters to be estimated are the channel, h, and the transmitted data, $\mathbf{x}$. Thus, the GLRT operates with the joint likelihood function, $P(\Gamma|\mathbf{x}, h)$, of the "observation" $\Gamma \triangleq \{\mathbf{y}, \Pi_{\mathcal{M}}\}$, given $\mathbf{x}$ and h. GLRT based phase estimation involves maximization of the likelihood function first over transmitted symbol vector and then over quantized channel phase (2.12), and may be viewed as joint maximum likelihood estimation of θ and $\mathbf{x}$ based on the observation Γ.

$$\hat{\theta}_{GLRT}(\gamma) = \arg\max_{q\in\mathcal{Q}}\max_{\mathbf{x}\in\mathcal{M}} P(\gamma|\mathbf{x}, \hat{a}\exp(j\phi q/Q)) \tag{2.12}$$

The inner maximization, $P(\gamma|\hat{\mathbf{x}}_q, \hat{a}\exp(j\phi q/Q))$, represents the conditional likelihood of the Maximum Likelihood Sequence Estimate (MLSE), $\hat{\mathbf{x}}_q$, of the trans-

mitted symbols on the qth phase trellis. We propose to estimate the likelihood of the MLSE sequence, typically computed with the Viterbi algorithm, with the forward recursion of BCJR algorithm, with $\max_{s\in\mathcal{A}}\{\alpha_T(s)|h = \hat{a}\exp(j\phi q/Q)\}$. This approximation is found to be very accurate. Thus, the MLSE statistic is a natural choice for measuring reliability of soft decisions output by each phase branch. The metric, when used to choose the best two phase branches after the first receiver iteration, yields performance within 0.1 dB to that of averaging over all phase bins. And, after the first iteration, noncoherent demodulation requires only twice as many BCJR computations as coherent demodulation of the same code.

2.2 EXIT functions of noncoherent codes

We study the convergence behavior of iterative noncoherent demodulation and decoding via Extrinsic Information Transfer (EXIT) charts [36, 37]. The EXIT chart of a noncoherent code is a graphical description of iterative noncoherent demodulation and decoding, portraying the mutual information between decoder messages communicated and code-symbols estimated, as evolved through turbo-processing. Consider first the inner modulation code, $\mathcal{M}$, that maps the code bits, $\tilde{\mathbf{c}} = \{c[n]\}$, $c[n] = [c_1[n]\cdots c_m[n]]$, $m = \log_2 M$, to channel symbols, $\{x[n]\}$. The noncoherent block demodulator computes *a posteriori* probabilities, $\Lambda_{\mathcal{M}} = \{\Lambda_n\}$, as described in section 2.1.3, defined here with respect to the code-bits, $\{c_n\}$, as LLRs,

$$\Lambda_n = \log\frac{Pr(c_n = 0|\Pi_{\mathcal{M}}, \{\mathbf{y}\})}{Pr(c_n = 1|\Pi_{\mathcal{M}}, \{\mathbf{y}\})} - \Pi_n, \tag{2.13}$$

where $\Pi_{\mathcal{M}} = \{\Pi_n\}$ denotes the code-bit priors, $\Pi_n = \log\frac{Pr(c_n=0)}{Pr(c_n=1)}$. The *EXIT function*, A, for $\mathcal{M}$ describes the mutual information of the code-bits and APPs, $a^{out} \triangleq I(\mathbf{c}; \Lambda_{\mathcal{M}})$, as a function of the *input* mutual information of the code bits and priors, $a^{in} \triangleq I(\mathbf{c}; \Pi_{\mathcal{M}})$, and channel SNR according to $a^{out} = A(a^{in}, \mathrm{SNR})$.

Conditional probability density functions of decoder priors are often well-modeled as i.i.d., with a single-parameter family of Gaussian densities [37],

$$\Pi_n \sim \mathcal{N}(\pm 2\gamma, 4\gamma),\ \gamma \in [0, \infty), \tag{2.14}$$

interpreted as the conditional likelihoods of the code-bits transmitted with BPSK over an AWGN channel of SNR γ. In this case, a^{in} has a simplified form, computed with the estimate (see [37])

$$a^{in}(\gamma) = 1 - E\left[\log\big(1 - \exp(\Pi_n)\big)|c_n = 0\right]. \tag{2.15}$$

With $\mathcal{M}$ discrete, we have $A : [0,1] \rightarrow [0,1]$, and the parameter γ is varied to generate a^{in} over the support of A. The *output* mutual information, a^{out}, is computed by measuring conditional probability density functions of Λ_n that are generated by the decoder fed with $\Pi_{\mathcal{M}}$ as in (2.14).

We next consider the EXIT function, B, of the outer channel code, $\mathcal{C}$. The APP decoder for $\mathcal{C}$ computes *a posteriori* probabilities of the code-bits, $\Lambda_{\mathcal{C}}$, with the priors $\Pi_{\mathcal{C}} = \text{perm}^{-1}(\Lambda_{\mathcal{M}})$ (permuted extrinsics from APP demodulation). Letting $b^{in} \triangleq I(\mathbf{c}, \Pi_{\mathcal{C}})$ and $b^{out} \triangleq I(\mathbf{c}, \Lambda_{\mathcal{C}})$ denote input and output decoder mutual information, respectively, the decoder EXIT function is given by $b^{out} = B(b^{in})$. In many cases, log-APPs produced by the outer channel decoder, e.g. the convolutional decoder, are well modeled as Gaussian. Then, a^{out} is accurate when computed empirically from demodulator APPs, $\Lambda_{\mathcal{M}}$, resulting from the priors (2.14). However, we find that the extrinsic information, $\Lambda_{\mathcal{M}}$, produced by the demodulator is non-Gaussian, resulting from the relatively short block length, so that estimates of B based on Gaussian priors at the decoder do not accurately model density evolution for noncoherent processing.

Since decoder priors are de-interleaved code-symbol posteriors from the demodulator, $\Pi_{\mathcal{C}} = \text{perm}^{-1}(\Lambda_{\mathcal{M}})$, we propose the following approach for measuring decoder output mutual information, b^{out}. First, Gaussian code-bit priors are noncoherently demodulated; demodulator input mutual information, $a^{in}(\gamma)$, is com-

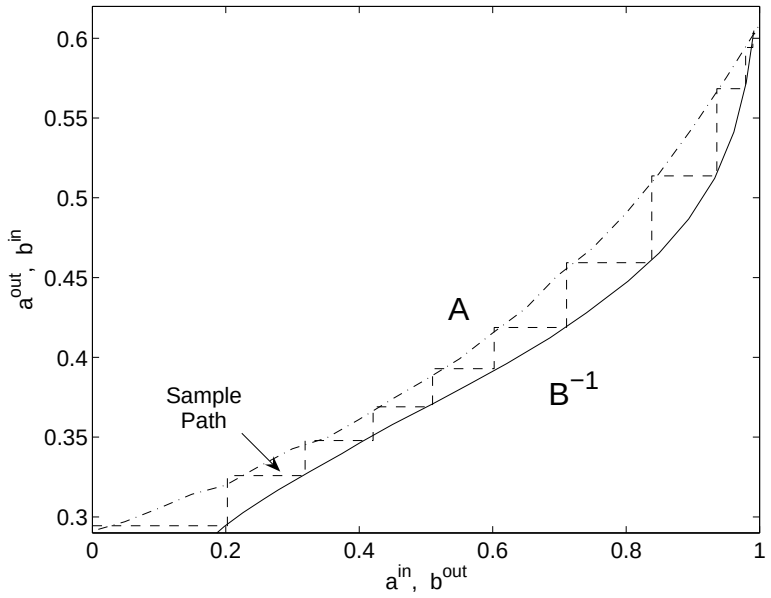


Figure 2.4: Rate 1.35 bits/channel symbol, noncoherent modulation code (A) and outer convolutional code (B), EXIT functions.

puted with (2.15) and output mutual information, a^{out}, is estimated empirically. The resulting extrinsic code-bit APPs are then de-interleaved and sent to the decoder as priors, $\Pi_{\mathcal{C}}$, now accurately modeling the priors observed in noncoherent processing. Decoder output mutual information (2.16) is computed empirically from the resulting decoder extrinsics $\Lambda_{\mathcal{C}}$.

$$b^{out} = BA(a^{in}(\gamma)) \tag{2.16}$$

Figure 2.4 is an EXIT chart of the mutual information (2.16) of a rate-3/8 convolutional code with aligned PSK rings 16-QAM at an SNR of 5.4 dB, near the SNR threshold for this code combination. The inverse decoder transfer function, B^{-1}, is plotted since $b^{in} = a^{out}$. A sample path, corresponding to one channel realization and the resulting mutual information sequences, $\{a_k^{out}\}$, $\{b_k^{out}\}$, that arise from iterative noncoherent demodulation and decoding of a transmitted codeword, is depicted.

We note some properties of EXIT charts and their implications:

Property 1. A given code combination *converges* when $\lim_{k\to\infty} b_k^{out} = 1$, if and only if the information bit error rate (BER) approaches zero. An equivalent condition for code convergence is that $B^{-1} < A$. In practice, we do not require the final decoder output mutual information to exactly equal one. In general, we have $B^{-1}(1) = 1$, and it may only be possible to arbitrarily approximate this condition. Note that a final decoder output mutual information less than one will give rise to an error floor.

Property 2. Channel SNR induces an ordering on demodulator EXIT functions such that if A and A' are measured at SNRs τ and τ', respectively, with $\tau < \tau'$, then $A \leq A'$. The *convergence threshold* of a code is the SNR threshold, τ, for which the code converges if and only if SNR$> \tau$.

Property 3 *(Conjecture).* The area property of trellis decoders: $\int_0^1 B^{-1} = r_{\mathcal{C}}$, where $r_{\mathcal{C}}$ denotes the code rate. This property, proved only for erasures channels [1], has consistently been observed in the literature, and in our own study, for a wide variety of channels. We show later that this property, if true, would imply that convolutional outer codes are near-optimal when the inner code is unit rate differential modulation.

As an empirical rule of thumb, the operating SNR does not affect the shape of the demodulator transfer function, A, but rather its vertical position. Since the channel decoder does not directly observe the channel output, its transfer function is unaffected by SNR. Figure 2.5 illustrates Property 2 for 16-ary constellations. EXIT chart analysis of noncoherent codes provides a quantitative framework for comparing signal alphabets or complexity reducing techniques without having to simulate BER performance. For example, Figure 2.5 shows aligned PSK rings 16-QAM are clearly superior to lattice 16-QAM over a wide range of SNR.

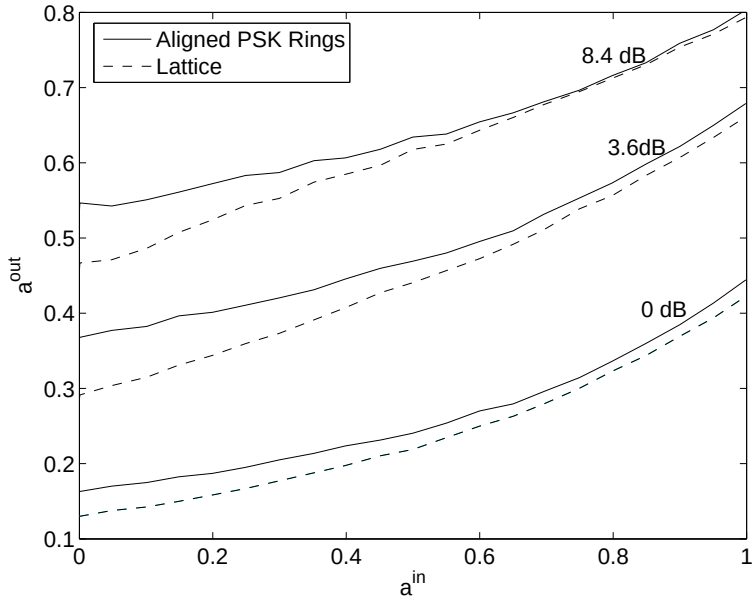
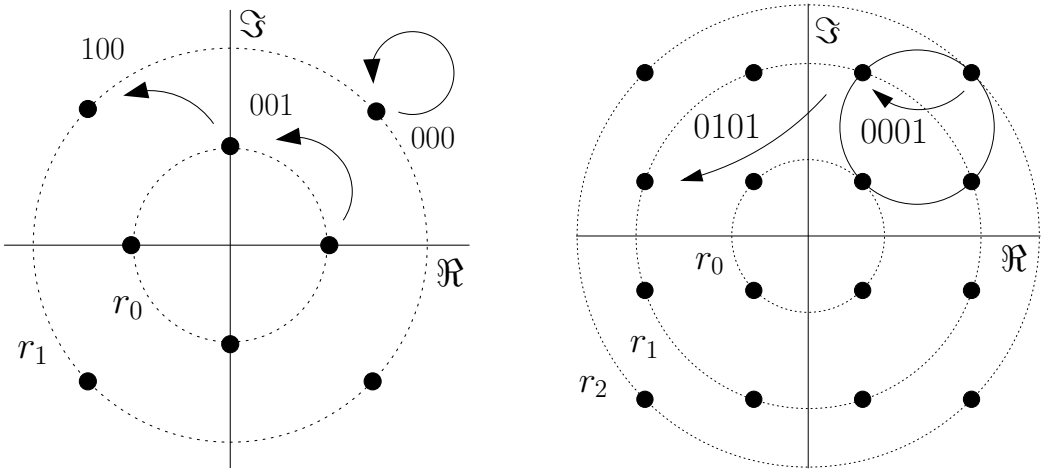


Figure 2.5: Transfer functions of 16-ary amplitude/phase constellations.

2.3 Constellation design for noncoherent communication

We consider the design of amplitude/phase constellations and bit-to-symbol maps well-suited for turbo noncoherent communication over the block Rayleigh fading channel. The main tools for design and optimization of the modulation codes are noncoherent capacity and modified EXIT chart analysis, where we use the term "noncoherent capacity" for the mutual information attained by various input distributions for noncoherent communication over the block fading channel. Marzetta and Hochwald [27] have shown that the unconstrained noncoherent capacity is achieved by T-dimensional isotropically distributed vectors, with $E[\mathbf{x}^{\mathrm{H}}\mathbf{x}] = T$. This is equivalent to sending i.i.d. Gaussian symbols, with the energy over a block of T symbols normalized to a constant. Intuitively, therefore, suitable modifications of designs similar to those employed for the AWGN channel, for which i.i.d. Gaussian input is optimal, are expected to work for the

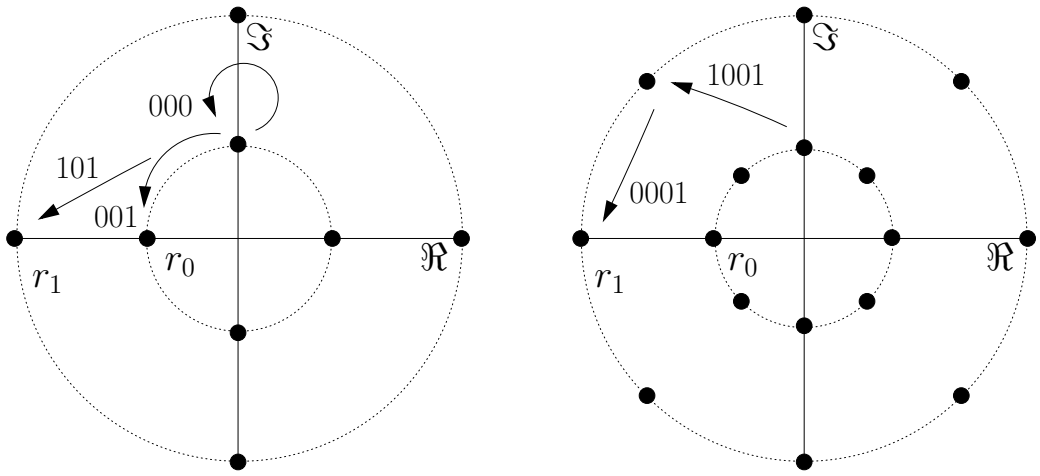


(a) 8-QAM: Offset PSK rings

(b) 16-QAM: Lattice

Figure 2.6: Standard AWGN constellations with Gray-like bit-mappings.

noncoherent Rayleigh block fading channel. In particular, QAM constellations approximate Gaussian input distributions more closely than PSK, especially for a large number of points. We first consider noncoherent communication with lattice QAM constellations with differential Gray-like bit maps, as in Figure 2.6. The bit maps in this case index transitions within and between QPSK sub-constellations within the QAM constellations. However, we find that, in simulations of a coded noncoherent system, such lattice QAM constellations, at least in conjunction with the bit maps we have considered, perform poorly, not delivering on the promised gains over PSK. We therefore consider an alternative class of QAM constellations, in the form of aligned PSK rings. These constellations, along with Gray-like bit maps for encoding data in the amplitude and phase transitions, are depicted in Figure 2.7. The ratio of the ring radii is chosen to optimize noncoherent capacity. As discussed in detail below, these constellations are found to perform much better than lattice QAM. We observe, Figure 2.8, that the noncoherent capacity is virtually identical for aligned PSK rings and lattice QAM, and turn to EXIT analysis for providing more precise guidance on constellation and bit map choice.



(a) 8-QAM (b) 16-QAM

Figure 2.7: Noncoherent signaling based on aligned, concentric PSK rings.

Figure 2.9 compares the simulated information BER of aligned PSK rings with lattice 16-QAM and (lattice-like) offset rings 8-QAM. Standard convolutional codes are employed for an overall data rate of 1.35 bits/channel symbol. The figure demonstrates the gain in using aligned PSK rings over rectangular lattices: 1 dB for 16-ary constellations, and a less drastic 0.2 dB for 8-ary constellations. It also displays the advantage of constellation expansion with heavier coding: for the same information rate, 16-QAM aligned PSK rings outperform 8-QAM aligned PSK rings by 0.5 dB. Of course, this advantage is not realized for poor constellation and bit map choices: the lattice 16-QAM performs significantly worse than the 8-ary constellations at the same information rate.

The advantage of aligned PSK rings is clearly brought out by EXIT analysis. Figure 2.5 shows that, at the same SNR, that the EXIT chart for aligned PSK rings 16-QAM lies strictly above that of lattice 16-QAM with noncoherent block demodulation and Gray-like differential bit maps. This implies that the convergence threshold of any outer code will be strictly larger for lattice 16-QAM. A possible intuitive explanation for the superior performance of aligned rings is as

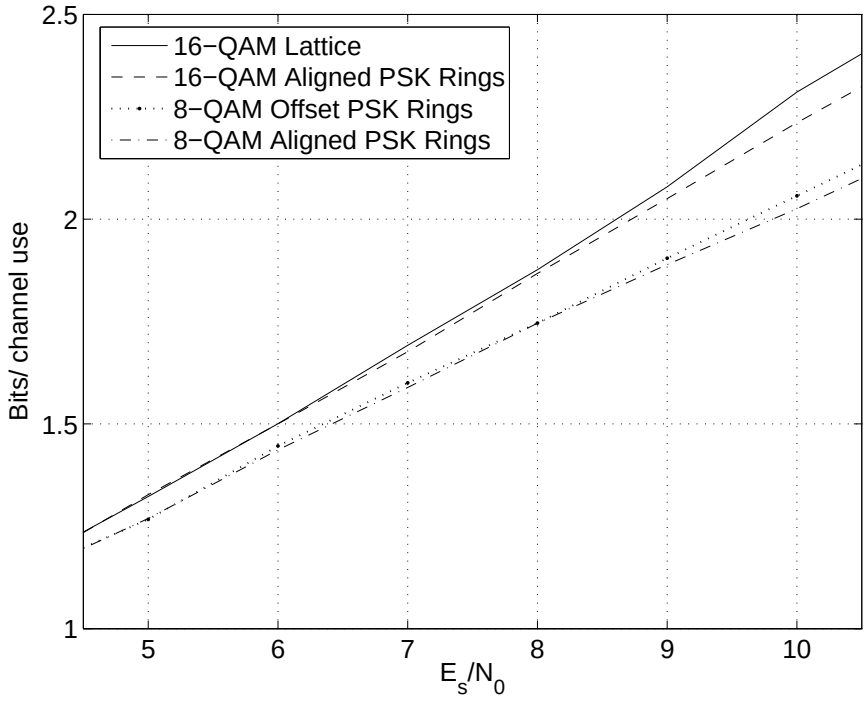


Figure 2.8: Noncoherent capacity of 8- and 16-QAM constellations.

follows. For blocks suffering from poor SNR due to fading, the aligned PSK rings effectively collapse to a more robust PSK constellation. Thus, while the bits encoded in amplitude transitions are difficult to recover, the bits encoded in phase transitions are relatively better preserved. For lattice constellations, on the other hand, all bits are affected adversely in a faded block. As SNR and the coherence interval increases, AWGN-like QAM constellations are preferable, for their better nearest neighbor distance characteristics. However, this is not the operating regime for the turbo-like system considered here, where we expect a relatively high uncoded BER. Thus, artful coupling of the constellation shape and bit-to-symbol map is key to the design of bandwidth efficient symbol alphabets for noncoherent communication.

Thus far, we have exclusively considered unit-rate rotationally-invariant differential modulation. These simple modulation codes are well-suited to block noncoherent processing for their low-complexity demodulation and bootstrap function-

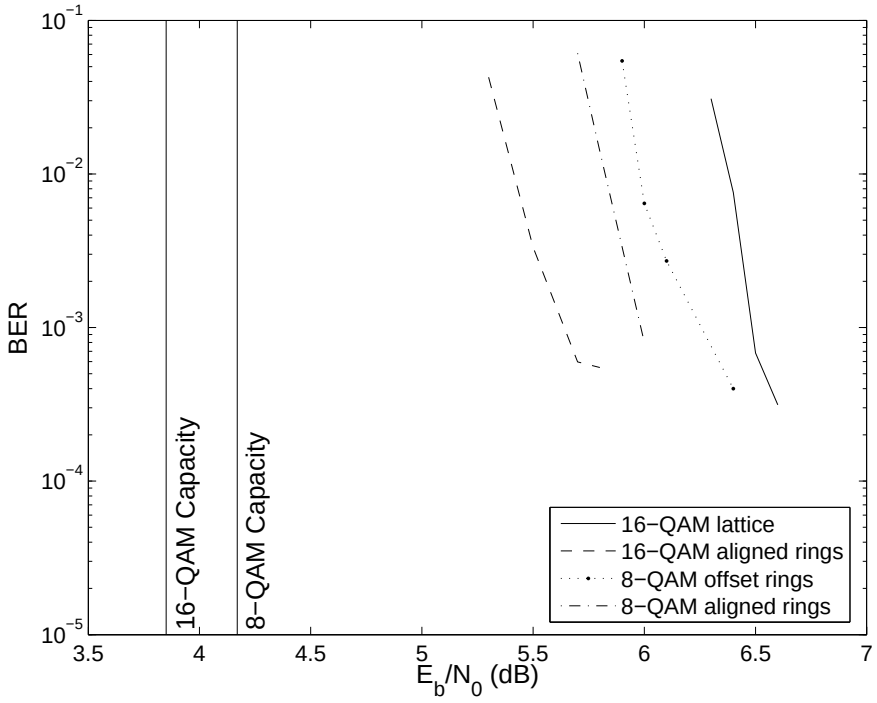


Figure 2.9: Coded performance of 8- and 16-QAM constellations.

ality. We now wish to to quantify the performance penalty associated with this restriction (e.g., as opposed to using more sophisticated trellis-based rotationally invariant modulation codes) for a serially concatenated system with an outer binary code. To this end, we invoke the conjectured area property, which states that the rate of the outer decoder equals the area under its exit curve, B^{-1}. The best possible (typically unrealizable) choice of outer code is when the decoder curve perfectly matches the inner demodulator curve at convergence. Thus, the highest possible rate of the outer code for convergence at a given SNR is the area under the demodulator curve at that SNR. This provides the following upper bound on the achievable rate (with serial concatenation) as a function of SNR for a given inner modulation code, as a function of its exit curve A:

$$I_{\mathcal{M}}(SNR) = \log_2(M)\frac{T-1}{T}\int A(u, SNR)du. \tag{2.17}$$

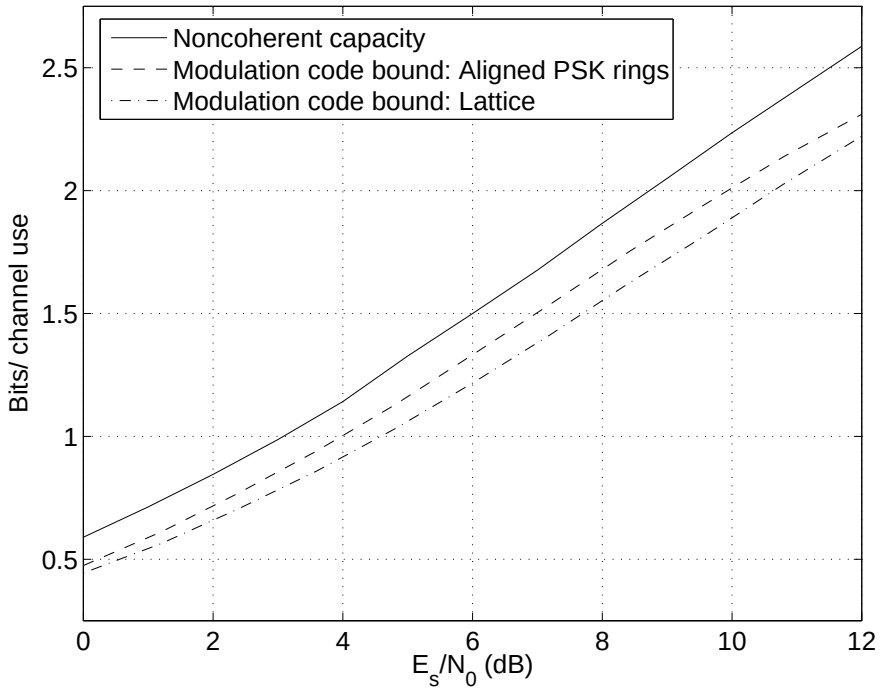


Figure 2.10: The modulation code bound for unit rate differential modulation with 16-QAM aligned PSK rings is about 1 dB away from the noncoherent capacity for that alphabet.

We term this bound the *modulation code bound.* Figure 11 compares the noncoherent capacity for aligned PSK rings 16-QAM with the preceding upper bound on the achievable rate, using the same constellation, when restricted to using serial concatenation with a unit rate differentially modulated inner code with Gray-like bit maps.

Finally, we comment on block differential modulation, a unit-rate modulation scheme that is an alternative to standard differential modulation. In block differential modulation, information is encoded in transitions in amplitude and phase relative to a fixed symbol: in practice, the reference symbol might be the first symbol of the current block, which would be the same as the last symbol of the previous block, if successive blocks overlap by a symbol. For turbo noncoherent communication with QPSK alphabets [10], block differential modulation was

found to yield the same convergence threshold when paired with a turbo-like code as standard differential modulation paired with a convolutional code. However, as mentioned earlier, demodulation for block differential modulation is significantly more complex, requiring phase quantization over $[0, 2\pi]$, compared to the much smaller interval required for standard differential modulation. Furthermore, the bootstrap mechanism that we employ for amplitude estimation using standard differential modulation is found not to work well for block differential modulation. A possible explanation is that there is less averaging in the amplitude estimator for two-symbol demodulation for block differential modulation, since the reference symbol must always be involved.

2.4 Channel coding for noncoherent modulation

Given the constellation and bit-mapping, EXIT chart analysis also guides the appropriate choice of outer channel code. Figure 2.11 compares the EXIT functions of aligned PSK rings 16-QAM at (a) 6.5 dB when paired with a standard rate-1/2 convolutional code and (b) 7.9 dB when paired with an irregular LDPC code optimized for the AWGN channel [32]. The number of LDPC decoder iterations is 30.

In order to estimate the gap to the modulation code bound, $I_{\mathcal{M}}$, a lower bound on the SNR convergence threshold of a given channel code is obtained by appealing to the waterfall behavior of concatenated codes and considering the first iteration of demodulation and decoding as a function of channel SNR. By Property 1, a necessary condition for code convergence is that the first iteration produces a net increase in demodulator output mutual information, i.e.

$$a_0^{out} < a_1^{out}. \tag{2.18}$$

Note that if the demodulator fails to yield a net increase in output mutual information in any one iteration, then the turbo demodulation and decoding algorithm

has reached a fixed point solution, and no further increase (or decrease) in mutual information is possible. The *Channel Threshold Bound* (CTB) for code convergence is the smallest SNR for which (2.18) holds. In general, this bound will be tight, given the waterfall characteristic of turbo-processing. In the figure, the CTB corresponds to the SNR at which the dotted curves, $A(0, SNR)$ (upper) and $B(A(0, SNR))$ (lower, inverse plotted), diverge. For standard rate-1/2 convolutional coding the bound is 5.7 dB, which is 2 dB better than the corresponding bound for the LDPC code. Note, however, that the channel threshold bound is much looser for the convolutional code since the EXIT charts of the demodulator and decoder are so well-matched.

We find that the convolutional code EXIT function is well-matched to unit-rate aligned-rings 16-QAM, yielding at least 1 dB performance improvement from the irregular LDPC code. Optimization of the LDPC degree sequence for the specific context of noncoherent demodulation in block fading can potentially close this gap, but such optimization is beyond the scope of this paper. We claim, however, that convolutional coding is near-optimal for noncoherent amplitude/phase modulation. Taking the CTB as the threshold for the code combination in Figure 2.11(a) and comparing to the modulation code bound of Figure 2.10, $I_{\mathcal{M}}^{-1}(1.35) = 4.8$ dB (E_b/N_0), idealized convolutional coding could improve the convergence threshold by only 0.9 dB. Since achieving the modulation code bound requires infinitely many demodulation and decoding iterations (by definition the decoder transfer function is perfectly matched and coincident with the demodulator function), we conjecture that only 0-0.4 dB could be gained by optimizing the convolutional code. Moreover, the gain for optimizing the LDPC code would be the same (with respect to the depicted convolutional code). Thus, we infer that standard convolutional coding is near-optimal for a unit-rate differential modulation code.

2.5 Results and discussion

We first consider simulated BER results of the aligned-rings 16-QAM constellation with unit-rate Gray-like differential modulation and a standard rate-1/2 convolutional code. The angle of rotational invariance, ϕ, is $\pi/4$ for aligned PSK rings 16-QAM constellation. The full complexity noncoherent receiver and the first iteration of the reduced complexity receiver, employ $Q = 5$ quantization phase bins, $\{\phi q/Q\}_{q=0}^{Q-1}$, sufficiently many to closely approximate performance with an arbitrary number of quantization levels. The overall codeword length is 64,000 bits. Accounting for the $1/T$ loss in rate for differential demodulation with i.i.d. block fading, at a modest coherence interval, $T = 10$, the corresponding data rate is 1.8 bits per channel symbol. Noncoherent capacity computations, Figure 1.1, give a 1.5 dB advantage to 16-ary amplitude/phase constellations over 16-PSK at this rate. Indeed, the aligned PSK rings constellation realizes almost all of the of the predicted gain, demonstrating a surprising robustness to amplitude distortion. We note that the system is operating around 1.7 dB from capacity, which agrees with the EXIT analysis, Sections 2.3 and 2.4, that predicts a loss of 1 dB for unit-rate modulation, 0.9 dB for non-ideal channel coding (recall the convergence threshold bound for the convolutional code is optimistic). Finally, there is a negligible 0.1 dB loss for GLRT selection of the best two phase branches after the first iteration, the reduced-complexity receiver.

Table 2.5 summarizes computer simulation results, showing serial concatenation of a convolutional code and differential amplitude/phase modulation approaches Shannon capacity for a block fading channel model, and performs significantly better than DPSK for moderately high SNRs and constellation sizes of 16 or larger. We have developed modified EXIT analysis tools for constellation and bit mapping choice, and for matching outer and inner codes. An important potential application is to the design of OFDM-based fourth generation wireless cellular systems: the complexity of our turbo noncoherent system is comparable

Table 2.1: Information rates attained using turbo designs based on aligned PSK rings with gap to capacity and improvement over PSK.

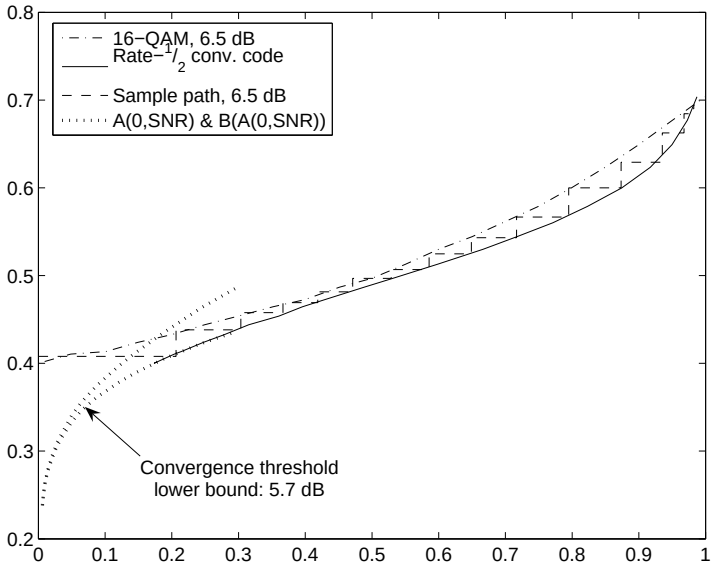
Bits/channel use	Constellation size	Outer code rate	Gap to capacity (dB)	Improvement over PSK (dB)
0.675	8	1/4	1.8	-0.2
1.35	8	1/2	1.8	0.4
1.35	16	3/8	1.8	1
1.8	16	1/2	1.7	1.5

to that of coherent systems with turbo-like coded modulation, so that noncoherent architectures are now implementable. However, we are still about 1.7 dB from capacity for an information rate of 1.8 bits/symbol, using a 16-ary amplitude/phase constellation. Since the convolutional outer code appears to be near-optimal if unit rate differential modulation is used as the inner code, one possible approach to close the gap may be to employ a lower rate inner code, alleviating the 1 dB loss for unit-rate modulation predicted by the modulation code bound (2.17). Other approaches include suitably optimizing the degree distribution of an LDPC outer code, with different possible choices of inner modulation code. It is also important to quantify how much of the gap to capacity can be closed simply by increasing the code length, or by increasing the constellation size and decreasing the outer code rate, while still employing a unit rate inner differential modulator.

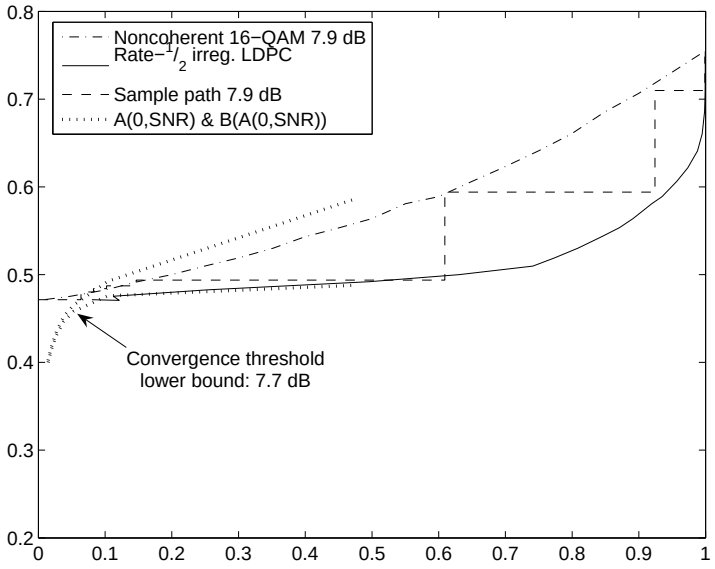
In practice, the fading gain for a mobile channel varies continuously with time. However, our block-wise constant approximation for the channel gain works well for the settings found in current and projected commercial digital cellular systems, in which, for typical Doppler shifts, the operating SNRs do not reach the extremely high levels [13] required for falling into the undesirable $\log \log \mathrm{SNR}$ regime of capacity growth predicted by Lapidoth and Moser [23]. Indeed, we conjecture that the capacity of the block fading model should be close to that of a continuously varying channel that it approximates in the desirable regime

of $\log \mathrm{SNR}$ growth. It is of interest to give precise shape to this intuition in the context of a specific continuously varying fading channel model.

While we have considered relatively small normalized Doppler frequencies (where the normalization is relative to the symbol rate) typical of outdoor cellular wireless systems, there are other situations in which the normalized Doppler may be large enough that the block fading approximation breaks down even at low SNR. It is of interest to explore alternative structures for turbo noncoherent communication in such settings, with preliminary work in this direction reported in [11, 28].



(a) Standard convolutional code.



(b) Irregular LDPC code.

Figure 2.11: Rate-1/2 coding for noncoherent 16-QAM.

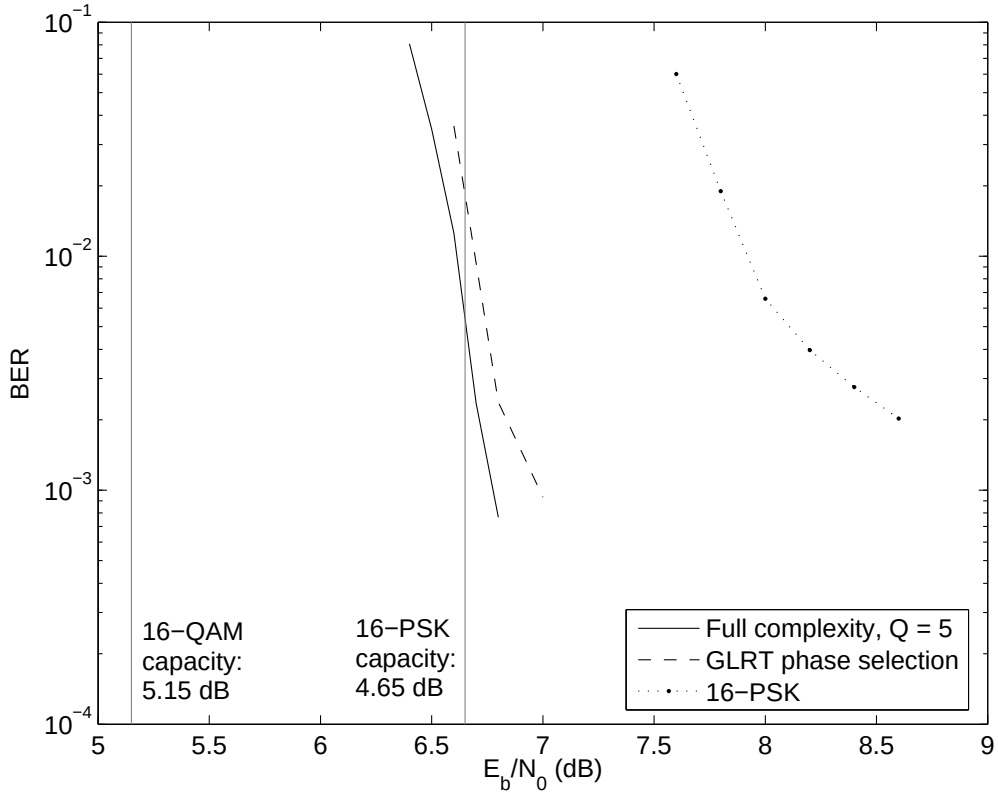


Figure 2.12: Rayleigh block fading results: 16-QAM.

Chapter 3

Noncoherent eigenbeamforming

Much of the Multi-Input Multi-Output (MIMO) communication literature assumes channel coherence, see [14] and the references therein, and classical receive beamforming is based on explicit channel estimation [35]. We demonstrate that beamforming gains are still realizable for multi-antenna receivers when the channel is *a priori* unknown to the transmitter or the receiver and no pilot-symbols are transmitted. Moreover, the complexity remains practical even as the receiver benefits from an SNR increase by scaling up the number of antennas. Noncoherent communication is particularly well-suited to the uplink of a cellular system in which the base-station must estimate the time-varying channel to each mobile. Pilot-symbol based channel estimation is more efficient on the downlink, since the mobiles share a common pilot channel. Accurate estimates of the spatial covariance matrix, available through averaging in wideband systems, allow eigenbeamforming [16] at the receiver along the dominant channel modes. For a typical outdoor channel, where the number of dominant modes is small, this allows the receiver to increase its SNR by scaling up the number of antennas, while limiting the demodulation and decoding complexity (which scales with the number of channel modes used by the receiver).

This research investigates receive beamforming at the base station of an outdoor wireless cellular system, when no prior channel realization information is assumed at the transmitter or receiver [17]. Rather, empirical estimates of the

received spatial covariance matrix, generated from uplink measurements, in conjunction with well-established techniques of coded noncoherent communication, are employed to approach capacity. We show that typical outdoor cellular channels are characterized by a relatively few number of dominant spatial modes. Thus, by projecting the received signal along the dominant modes, the receiver enjoys a SNR enhancement, or *beamforming gain*, that scales up with the number of receive elements, while simultaneously obtaining reasonable levels of complexity. Shannon theoretic analysis further reveals a diversity tradeoff for channels with multiple spatial modes. Key to these results is the property that second order channel statistics (the spatial covariance) vary slowly with respect to the mean (or realization) of the channel. This assumption holds true for emerging Wireless Metropolitan Area Networks (WMANs), for example the IEEE 802.16 standard, where user mobility at vehicular speeds and severe non-line-of-sight multipath delay profiles give rise to fast fading in both time and frequency. For such channels, the coherence length, or number of symbols for which the channel is expected to be roughly constant, is expected to be on the order of a dozen symbols. Such rich time-frequency diversity enables extremely accurate beamforming estimates, without any pilot overhead. For the case of non-white interference contributions to the overall received spatial covariance matrix, arising for example in a frequency reuse-1 system, a Minimum Mean Squared Error (MMSE) framework for interference suppression is proposed. Simulation results demonstrate that less than 3% pilot overhead is required to mitigate consistently strong spatial interference that is of equal power to the desired user.

We first review material from our earlier work [5, 4], which abstracts simple statistical models from the literature on outdoor channel measurements. We then show that, for typical outdoor channels, most of the beamforming gains relative to a single antenna system can be obtained by using a small number of modes.

3.1 OFDM System Model

As in the classical Saleh-Valenzuela model [34], the channel response is decomposed into clusters. Experimental measurements of outdoor channels [30] indicate that the number of clusters is small, usually one or two, and that the power delay profile and power angle profile for each cluster can be modeled as exponential and Laplacian, respectively.

We consider an OFDM system in which a mobile with one antenna communicates to an N antenna receive array. There are K subcarriers. The received signal vector on the kth subcarrier is

$$\mathbf{y}[k] = \mathbf{h}[k]x[k] + \mathbf{n}[k]$$

where $\mathbf{h}[k]$ is an $N \times 1$ channel frequency response, $\mathbf{n}[k]$ is Additive White Gaussian Noise (AWGN) where $E[\mathbf{n}[j]\mathbf{n}[k]^{\mathrm{H}}] = 2\sigma^2\delta_{jk}\mathbf{I}_N$, and $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. It follows from our earlier work [5] that the $\{\mathbf{h}[k]\}$ are well-modeled as identically distributed zero-mean proper complex Gaussian random vectors

$$\mathbf{h}[k] \sim \mathcal{CN}(0, \mathbf{C})$$

with covariance

$$\mathbf{C} = E[\mathbf{a}(\Omega)\mathbf{a}(\Omega)^{\mathrm{H}}] \tag{3.1}$$

where $\mathbf{a}(\Omega)$ is the base station array response for the angle of arrival, Ω. We model a standard linear array for which

$$\mathbf{a}(\Omega) = [a_1(\Omega) \cdots a_N(\Omega)]^{\mathrm{T}}, \;\; a_n(\Omega) = e^{j(n-1)2\pi\frac{d}{\lambda}\sin(\Omega)},$$

when d is the antenna array spacing, and λ the carrier wavelength. The expectation (3.1) is taken over the distribution of Ω, the PAP.

A spectral decomposition of the channel covariance yields

$$\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}} \tag{3.2}$$

where the eigenvector matrix $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_N]$ is unitary, and $\mathbf{\Lambda}$ is diagonal with eigenvalues $\{\lambda_l\}$ arranged in decreasing order. The eigenvalue λ_l represents the energy of the channel on lth eigenmode $\mathbf{u}_l$.

3.2 Covariance Estimation

For the large delay spreads typical of outdoor environments, the coherence bandwidth is small, and the correlation between the channel responses at different frequencies dies out quickly with their separation. Thus, the base station can accurately estimate $\mathbf{C}$ by measuring the channel over a rich enough set of frequencies on the uplink [4].

Averaging over frequency bins, the base station forms an empirical autocorrelation matrix

$$\mathbf{R} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{y}[k]\mathbf{y}[k]^{\mathrm{H}}.$$

With $E[|x[k]|^2] = 1$, it is easy to show that $\mathbf{R}$ is an estimate of $\mathbf{C} + 2\sigma^2\mathbf{I}_N$, where σ^2 is the noise variance per dimension. Thus, if λ_l are the eigenvalues of $\mathbf{C}$, the eigenvalues of $\mathbf{R}$ are $\lambda_l + 2\sigma^2$. The eigenvectors of the two matrices are the same.

An eigendecomposition of $\mathbf{R}$ thus yields the dominant channel eigenmodes. Typically, the number of dominant eigenmodes is small for an outdoor channel because of the narrow PAP corresponding to signals received from a given mobile.

3.3 Eigenbeamforming

We now perform receive *eigenbeamforming* along the principal eigenmodes of $\mathbf{R}$, namely $\{\mathbf{u}_1, \cdots, \mathbf{u}_L\}$. The received vectors in each frequency bin are projected along these eigenvectors to get L parallel scalar OFDM signals (L typically much smaller than the number of antenna elements N). The lth signal is given by

$$z_l[k] = \langle \mathbf{y}[k], \mathbf{u}_l \rangle = \alpha_l[k]x[k] + w_l[k],$$

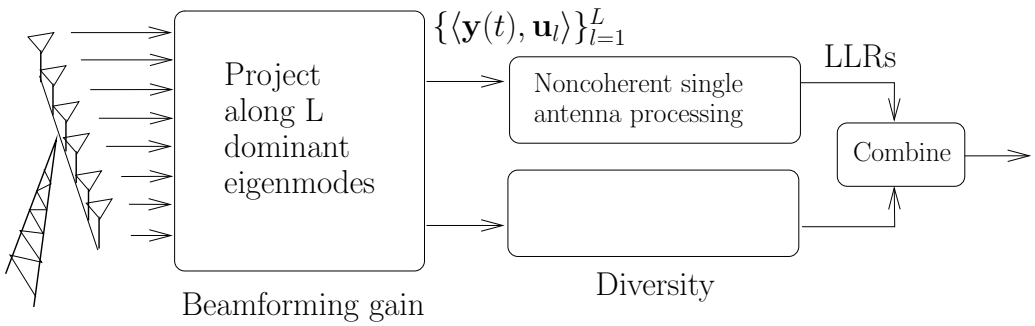


Figure 3.1: Beamforming gains are achieved by projecting onto the L dominant eigenmodes of the N-dimensional received signal $\mathbf{y}(t) = \sum_k \mathbf{y}[k] \exp(j2\pi(k-1)t/K)$.

where $\alpha_l[k] = \langle \mathbf{h}[k], \mathbf{u}_l \rangle$, $l = 1, \ldots, L$, are independent Gaussian channels, $\alpha_l[k] \sim \mathcal{CN}(0, \lambda_l)$, and $w_l[k] = \langle \mathbf{n}[k], \mathbf{u}_l \rangle$ is AWGN with $E[w[j]w[k]^*] = 2\sigma^2\delta_{jk}$.

3.3.1 Eigenbeamforming gain

As a rough measure of the performance gain relative to a single antenna system, we define the beamforming gain as the SNR if the signal power is summed over the L chosen eigenmodes, relative to the SNR for a single antenna element. This yields the following formula for the beamforming gain as a function of L:

$$G(L) = \frac{N}{\mathrm{tr}(\mathbf{C})} \sum_{l=1}^{L} \lambda_l \tag{3.3}$$

Figure 3.2 shows the beamforming gain as a function of the number of eigenmodes used for a 10 antenna system. The upper curve is for a single cluster channel whose PAP is Laplacian, $\mathcal{L}(\theta, \Delta\theta)$, with mean $\theta = 0$, and angular spread $\Delta\theta = 10^\circ$, where the variance is given by $(\Delta\theta)^2/2$. The middle plot is for a two cluster channel where the first cluster's PAP is as above, and the second cluster's PAP is also Laplacian with angular spread 10°, but has its mean at 45° (both clusters with the same power). Finally, the lower plot considers a third additional equal strength $\mathcal{L}(-45^\circ, 10^\circ)$ cluster. The total receive power is normalized to be the same in all three cases. Note that the beamforming gain quickly plateaus as

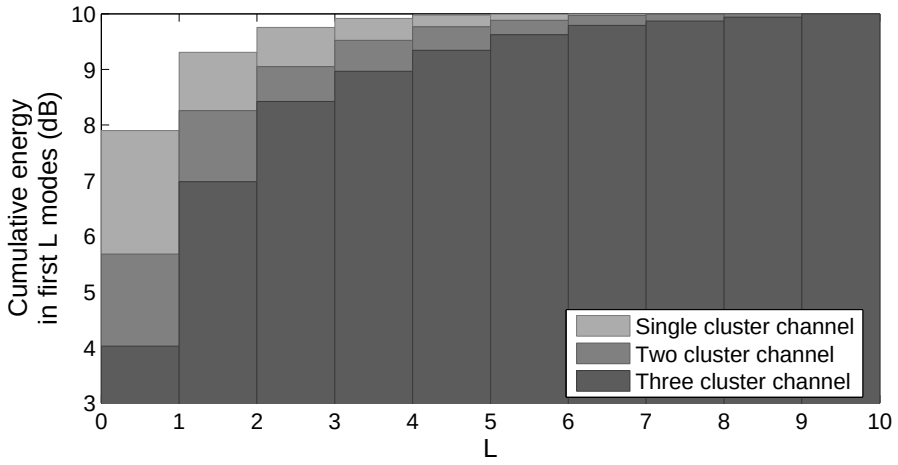


Figure 3.2: Eigenbeamforming gain over a single antenna receiver.

a function of L: thus, beamforming along the dominant eigenmode captures most of the channel energy for the one cluster channel, while using the first two (three) eigenmodes captures roughly the same amount of energy in the two (three) cluster system. Thus, for typical outdoor channels, estimation of the channel covariance enables the use of a small number of eigenmodes by the demodulator and decoder, limiting complexity while preserving the SNR advantage from scaling up the number of receive elements.

The signals for the L eigenmodes can be combined in a number of ways. One possibility is to explicitly estimate the scalar channels $\{\alpha_l[k]\}$ using pilots, and to then perform coherent diversity combining of the L branches to obtain an estimate of $x[k]$. The advantage that this may have over estimation of the original $N \times 1$ channel vector $\mathbf{h}[k]$ is that fewer gains may need to be explicitly estimated. In this paper, however, we consider noncoherent diversity combining, which is consistent with our goal of reducing overhead in uplink transmission. In particular, we consider serial concatenation of an outer code with differential modulation. Joint noncoherent processing of all modes is complex. Instead, we employ a suboptimal combining strategy for iterative noncoherent processing: parallel noncoherent demodulators are employed for each mode (each employing extrinsic information

from the outer decoder), the soft outputs of the demodulators are combined and sent up to the outer decoder, which then sends extrinsic information back to each of the parallel noncoherent demodulators. The details are in Section 3.6.

3.4 Noncoherent capacity of an L mode channel

In this section we compute the noncoherent capacity of the $L \times 1$ block fading channel, where the L modes can have unequal strengths. Ideas analogous to those used for isotropic [27] or scalar channels [10] can be used to simplify capacity calculations. Our numerical results are for a QPSK alphabet.

Consider M-ary modulation with an $L \times 1$ block fading channel of coherence time T. Letting $\mathcal{A}$ denote the symbol alphabet, with $M = 2^k$ for some $k \in \{1, 2, 3, ...\}$, the signal vector $\mathbf{x} = [x[1]\ x[2]\ \cdots\ x[T]]^{\mathrm{T}}$ is drawn randomly from $\mathcal{X} = \mathcal{A}^T$ according to $P(\mathbf{x}) = 1/M^T$. In the *parallel block fading model*, an $L \times 1$ Gaussian channel, $\mathbf{h} \sim \mathcal{CN}(0, \mathbf{C})$, communicates the signal $\mathbf{x}$,

$$\mathbf{Y} = [\mathbf{y}_1\ \mathbf{y}_2\ \cdots\ \mathbf{y}_L] = \mathbf{x}\mathbf{h}^{\mathrm{T}} + \mathbf{N}, \tag{3.4}$$

in AWGN, $\mathbf{N} = [\mathbf{n}_1\ \mathbf{n}_2\ \cdots\ \mathbf{n}_L]$, with $E[\mathbf{n}_j\mathbf{n}_k^{\mathrm{H}}] = 2\sigma^2\delta_{jk}\mathbf{I}_T$.

In general, the channel is completely defined by the conditional probability density function (PDF), $P(\mathbf{Y}|\mathbf{x})$, of the received symbols $\mathbf{Y}$ given the transmitted symbols $\mathbf{x}$. The corresponding *capacity* is the supremum of the mutual information,

$$\mathcal{C} = \sup I(\mathbf{x}; \mathbf{Y}) = \sup\{H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{x})\}, \tag{3.5}$$

over space of input alphabets $(\mathcal{X}, P(\mathbf{x}))$, satisfying the average power constraint $E[\|\mathbf{x}\|^2] = T$. The mutual information is independent of the choice of basis for the received signal (since it is a one-to-one transformation) [27]; in particular, $I(\mathbf{x}; \mathbf{U}^{\mathrm{H}}\mathbf{Y}\mathbf{U}) = I(\mathbf{x}; \mathbf{Y})$, for any unitary matrix $\mathbf{U}$. Thus, without loss in generality, let $\mathbf{C} = \mathrm{diag}([\lambda_1\ \lambda_2\ \cdots\ \lambda_L])$. Then, the vector channel, $\mathbf{h} = [h_1\ h_2\ \cdots\ h_L]^{\mathrm{T}}$,

is composed of L independent scalar fading channels $h_l \sim \mathcal{CN}(0, \lambda_l)$, and the conditional PDF of the received symbols given the transmitted symbols is

$$P(\mathbf{Y}|\mathbf{x}) = \prod_{l=1}^{L} P(\mathbf{y}_l|\mathbf{x}), \tag{3.6}$$

where

$$P(\mathbf{y}_l|\mathbf{x}) = \frac{\exp\{-\operatorname{tr}([2\sigma^2\mathbf{I}_T + \lambda_l\mathbf{x}\mathbf{x}^{\mathrm{H}}]^{-1}\mathbf{y}_l\mathbf{y}_l^{\mathrm{H}})\}}{\pi^T \det(2\sigma^2\mathbf{I}_T + \lambda_l\mathbf{x}\mathbf{x}^{\mathrm{H}})}$$

is the conditional PDF of a scalar block fading channel [27].

Direct evaluation of (3.5) with respect to (3.6) is prohibitively complex, and Appendix C discusses analytical simplifications that make the computation tractable. The techniques employed are analogous to those used in [10] for the case of a scalar fading channel, and [27] for isotropic channels. The general approach is to compute $H(\mathbf{Y}) = -E \log P(\mathbf{Y}|\mathbf{x} = \mathbf{1})$ with Monte-Carlo integration over the space of $\mathbf{h}$ and $\mathbf{n}$. The conditional entropy term, $H(\mathbf{Y}|\mathbf{x})$, is evaluated in closed form.

3.5 Capacity Plots

The information theoretic properties of multiple mode channels are now quantified for the case of QPSK signaling with the parallel block fading channel of coherence length $T = 10$ and L equal power modes. In order to compare capacity at different values of L, the noise variance per dimension is fixed and the strength of the modes (of the L mode channel) is set to $\lambda_l = 1/L$, $l = 1, ..., L$, so that the signal energy when summed over all modes is always one. Figure 3.3 shows noncoherent capacity versus SNR for $L = 1, 2, 3, 4$. There is a moderate diversity gain of up to 2 dB at high SNR, and no gain at low SNR. This indicates that, as long as there is enough averaging across frequency and time (which is implicitly assumed when we compute ergodic capacity), there is little penalty in performance due to the relatively few spatial eigenmodes present in a typical outdoor channel.

Note also that, since accurate estimates of the spatial covariance matrix are implicitly available in wideband systems, the number of receive elements can be

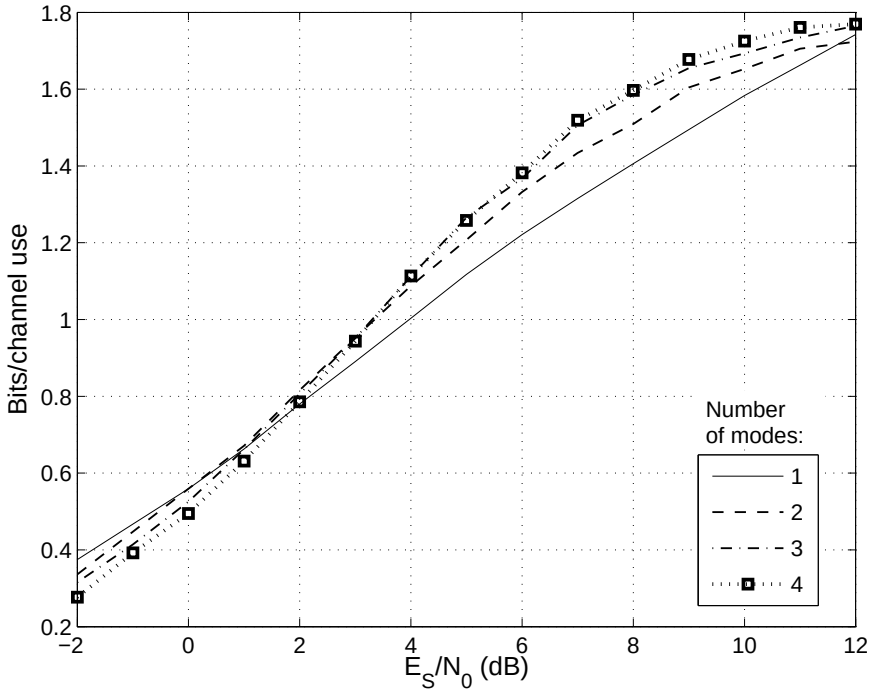


Figure 3.3: Block fading capacity with varying number of dominant eigenmodes.

scaled up to obtain large beamforming gains relative to a single antenna receiver, even when there is little spatial diversity gain available. This gain is a function of the number of antenna elements and array geometry, and is not shown in the plot above.

3.6 Noncoherent processing for the parallel block fading channel

Given the success of iterative noncoherent demodulation and decoding with a single antenna block fading channel [10, 19], we propose a diversity combining scheme for parallel block fading channels that falls within this framework. Optimal

computation of the *a posteriori* probabilities (APPs) of the transmitted symbols,

$$P(x[k] = x|\mathbf{Y}, \mathbf{\Pi}) = \frac{1}{P(\mathbf{Y})} \sum_{\mathbf{x}:x[k]=x} \prod_{t=1}^{T} P(x[t]) \prod_{l=1}^{L} P(\mathbf{y}_l|\mathbf{x}), \tag{3.7}$$

given the priors, $\mathbf{\Pi} = \{P(x[t])\}$, involves joint quantization of the L block fading channels, incurring a complexity that is Q^{L-1} times that of the single antenna demodulator, where Q denotes the number of quantizer bins per mode.

We instead propose a computationally efficient, suboptimal demodulator that is motivated by the conditional independence of the received signals $\{\mathbf{y}_l\}$ given the transmitted symbols $\mathbf{x}$ (3.6). The approach consists of running parallel noncoherent demodulators, one for each dominant mode of the channel, and linearly combining the resulting log-domain APPs. This approach has the advantage of leveraging well-known techniques for efficient noncoherent demodulation of the scalar block fading channel [10, 19], and entails phase quantization and amplitude estimation of the unknown channel for each mode. Moreover, the complexity of parallel noncoherent demodulation is linear (rather than exponential) in the number of modes. Thus, we simply sum the LLRs output from each noncoherent demodulator to get an estimate of the APPs of the transmitted symbols. The decoder then computes its own posteriors for the transmitted symbols, based on the aggregated output of the demodulators, and sends them back to each demodulator for use as priors. In this fashion, joint noncoherent demodulation and decoding are performed iteratively until the codeword is decoded. Note that our diversity combining differs significantly from conventional combining of differential demodulation statistics and is able to exploit channel continuity over the coherence length.

We have simulated the preceding linear diversity combining scheme with rate-3/4 convolutional coding and QPSK signaling. Figure 3.4 compares the performance of diversity combining to capacity for an idealized parallel block fading channel with one, two, and three equal strength modes. The results demonstrate

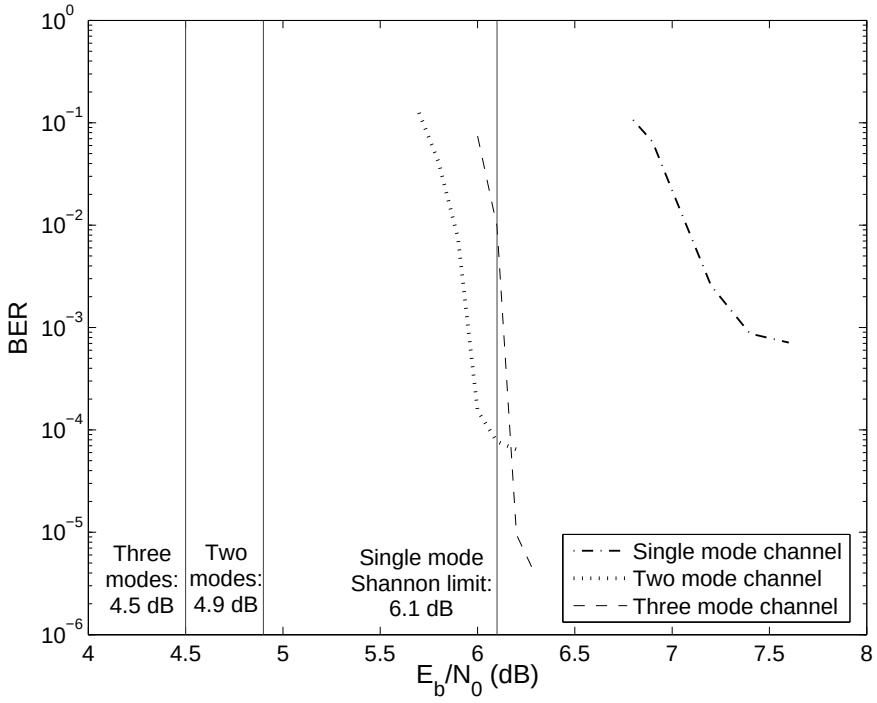


Figure 3.4: Performance of noncoherent diversity combining receivers with the parallel block fading channel, $T = 10$, at 1.35 bits/channel use.

near capacity performance for one and two mode channels. For the case of a three mode channel, the suboptimal combining scheme suffers from a per mode noise enhancement resulting from dividing the signal energy amongst the modes and running parallel independent demodulators. Nonetheless, the results are very encouraging, since the focus is outdoor channels with relatively few dominant modes, for which beamforming gains and complexity reductions increase significantly with the lesser number of modes. Further, there are larger potential diversity gains for the two mode channel over a one mode channel than there are by increasing the number of dominant modes beyond two.

We next consider more realistic fading channels whose spatial modes are generated explicitly according to outdoor channel models, with Laplacian PAPs for one and two cluster channels. As in the eigenbeamforming gain example, the single mode channel consists of a $\mathcal{L}(0°, 10°)$ PAP, and the two mode channel con-

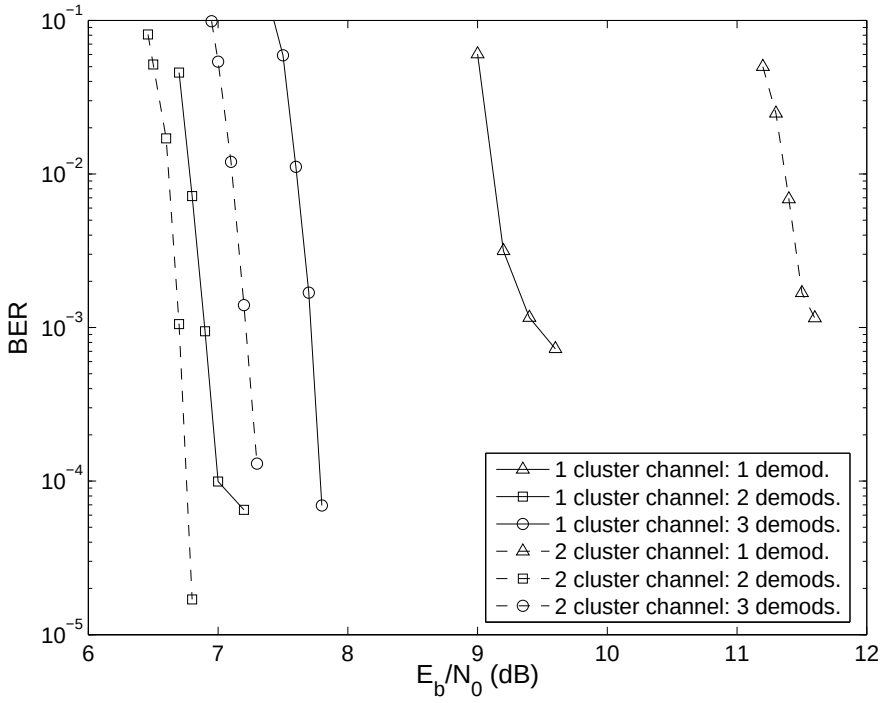


Figure 3.5: Performance of noncoherent diversity combining of multiple spatial modes in one and two cluster channels.

sists of the superposition of two equal strength $\mathcal{L}(0°, 10°)$ and $\mathcal{L}(45°, 10°)$ PAPs. The results are given in Figure 3.5. Performance clearly improves by considering two modes rather than one. Of course, in the model employed here, the improvement consists of an increase in SNR due to employing an additional mode (which is lesser for the single cluster channel), as well as a diversity advantage expected at this rate of 1.35 bits/channel use (which is greater for the two cluster channel). Again, parallel processing of three modes exhibits a slight performance degradation due to the suboptimal combining technique employed. Thus, more sophisticated demodulation techniques are required for channels with more than two dominant modes, in order to exploit the SNR advantage for considering the third mode. This is a topic of future research. Note, however, that for the one/two cluster channel simulated here the major gains in SNR and complexity reduction are achieved by the two mode demodulator.

3.7 Spatial interference

The issue of spatial interference from neighboring cells in frequency reuse-1 OFDM systems is largely mitigated by random-like scheduling in time and frequency. Inter-cell interference is further avoided in synchronous systems via coordinated Time Division Multiplexing (TDM) and Frequency Division Multiplexing (FDM) scheduling and power control techniques. Estimates of the desired user's spatial covariance matrix,

$$\mathbf{C}_{\mathrm{S}} = E[\mathbf{a}(\Omega)\mathbf{a}(\Omega)^{\mathrm{H}}], \tag{3.8}$$

rely on averaging over many subcarriers. If the interference in different subcarriers is spatially uncorrelated, then their contribution to the averaged spatial covariance matrix is approximately spatially white. In this case, the dominant eigenmodes of the overall spatial covariance matrix provide a good estimate of the spatial modes for the desired user. Interference resulting in non-white contributions to the overall spatial covariance matrix of the received signal, illustrated in Figure 3.6, is defined as *consistent spatial interference*. Such interference would arise, for example, if the uplink schedule of a radio in a nearby cell overlaps significantly with that of the desired user for a block of symbols that is much longer than the channel coherence length. Certainly, spatially non-white interference contributions will affect the eigen-structure of the received signal and thus estimated beamforming directions to the desired user. These issues lead us to consider interference learning receivers that estimate the spatial covariance of the interference based on randomized insertion of pilot blocks into the transmitted data sequence.

For typical outdoor scenarios, we would expect a small number of dominant spatial modes, since the received energy from a given radio is typically concentrated within a narrow angular spread, with contributions from only one or two dominant reflecting bodies. Thus, the overall spatial covariance matrix consists of a "mixture" of dominant modes from the desired user and interferer. Thus,

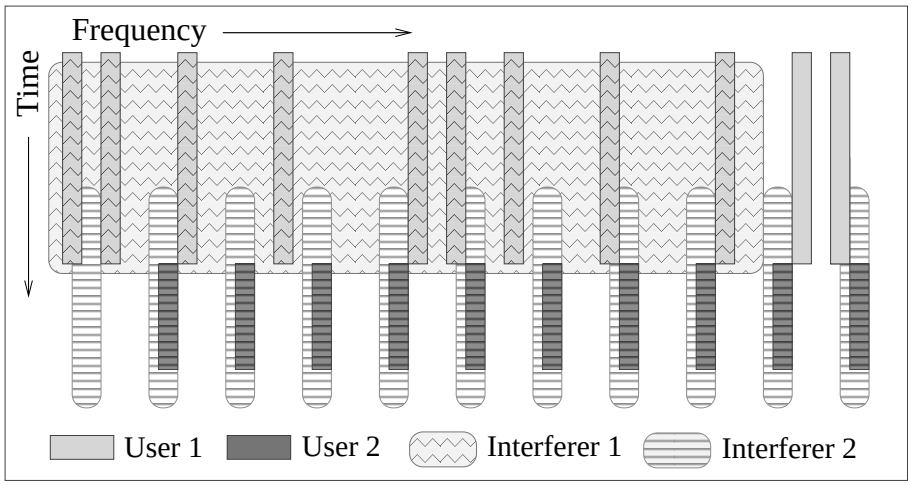


Figure 3.6: Examples of severe inter-cell interference on the uplink of a reuse-1 OFDMA system.

we might expect twice as many dominant modes as a single cluster channel, or the same number of dominant modes as a multi-cluster channel, in the absence of interference. The spatial covariance matrix can be estimated without overhead by computing second order statistics from uplink measurements. If the dominant spatial modes can be assumed to be due to the signal alone, then noncoherent eigenbeamforming leads to significant gains, as shown in [6] for downlink and [16] for uplink. This corresponds to projecting down to the low-dimensional signal subspace corresponding to the dominant eigenmodes. However, when there is significant contribution to the spatial eigenmodes due to the interference, then use of second order statistics is not enough to distinguish the desired signal from interference. We take the following approach for dealing with consistent spatial interference: (a) project down to the signal subspace corresponding to the dominant eigenmodes, (b) compute linear MMSE type interference-suppressing correlators using pilot symbols for a selected number of subcarriers, and (c) find a single interference-suppressing correlator that works well over the entire frequency band, exploiting the fact that the signal and interference spatial modes are the same across frequency. The projection step (a) reduces the dimension of

the interference-suppressing correlator, and reduces the number of pilot symbols needed in step (b).

3.8 Interference model

We consider the extreme case of a 0 dB interference process $\{x_{\mathrm{I}}[k]\}$, that is uncorellated with the desired signal $\{x_{\mathrm{S}}[k]\}$. This in particular means that the signal and interference are of equal energy, and we adopt the convention $E[|x_{\mathrm{S}}[k]|^2] = E[|x_{\mathrm{I}}[k]|^2] = 1$. Analogous to the desired user's own spatial channel (3.8), the $N \times 1$ spatial channel to the interferer is well modeled as zero-mean Gaussian with covariance $\mathbf{C}_{\mathrm{I}} = E[\mathbf{a}_{\mathrm{I}}(\Omega)\mathbf{a}_{\mathrm{I}}(\Omega)^{\mathrm{H}}]$. The interference array response $\mathbf{a}_{\mathrm{I}}(\Omega)$ varies according to the PAP of the angle of arrival, Ω. The PAPs of the interference and data channels are modelled independent Laplacians, with the same variance, but differing mean angles of arrival. This corresponds to the extreme scenario in which the interference process is virtually indistinguishable from an additional mode of the desired channel.

The baseband received symbol vector at the kth frequency channel is now

$$\mathbf{y}[k] = \mathbf{h}_{\mathrm{S}}[k]x_{\mathrm{S}}[k] + \mathbf{h}_{\mathrm{I}}[k]x_{\mathrm{I}}[k] + \mathbf{n}[k],$$

with $\mathbf{h}_{\mathrm{S}}[k] \sim \mathcal{CN}(0, \mathbf{C}_{\mathrm{S}})$, $\mathbf{h}_{\mathrm{I}}[k] \sim \mathcal{CN}(0, \mathbf{C}_{\mathrm{I}})$, and $\mathbf{n}[k] \sim \mathcal{CN}(0, 2\sigma^2\mathbf{I}_N)$. Letting $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_L]$ denote the dominant eigenvectors of $E[\mathbf{y}[k]\mathbf{y}[k]^{\mathrm{H}}]$, eigenbeamforming consists of projecting the received signal $\mathbf{y}(t)$ down to the reduced ($L \ll N$) complex vector subspace spanned by $\mathbf{V}$. This operation yields the following sufficient statistic for the transmitted symbol sequence:

$$\mathbf{z}[k] = \mathbf{V}^{\mathrm{H}}\mathbf{y}[k] = \boldsymbol{\alpha}_{\mathrm{S}}[k]x_{\mathrm{S}}[k] + \boldsymbol{\alpha}_{\mathrm{I}}[k]x_{\mathrm{I}}[k] + \mathbf{w}[k], \tag{3.9}$$

with $\boldsymbol{\alpha}_{\mathrm{S}}[k] \sim \mathcal{CN}(0, \boldsymbol{\Sigma}_{\mathrm{S}})$, $\boldsymbol{\alpha}_{\mathrm{I}}[k] \sim \mathcal{CN}(0, \boldsymbol{\Sigma}_{\mathrm{I}})$, and $\mathbf{w}[k] \sim \mathcal{CN}(0, 2\sigma^2\mathbf{I}_L)$. Where we have defined $\boldsymbol{\Sigma}_{\mathrm{S}} = \mathbf{V}^{\mathrm{H}}\mathbf{C}_{\mathrm{S}}\mathbf{V}$ and $\boldsymbol{\Sigma}_{\mathrm{I}} = \mathbf{V}^{\mathrm{H}}\mathbf{C}_{\mathrm{I}}\mathbf{V}$ to be the signal and interference

covariance in the reduced-space. The spatial covariance matrix of the received signal, (3.9), is thus given by

$$\mathbf{R} = E[\mathbf{z}[k]\mathbf{z}[k]^{\mathrm{H}}] = \mathbf{\Sigma}_{\mathrm{S}} + \mathbf{\Sigma}_{\mathrm{I}} + 2\sigma^2\mathbf{I}_L.$$

3.9 MMSE framework

We next develop a MMSE framework for interference suppression with the received signal model (3.9). The MMSE criterion is defined as

$$\mathbf{c} = \arg\min E[|\langle \mathbf{z}[k], \mathbf{c}\rangle - x_{\mathrm{S}}[k]|^2],$$

and the well known least squares solution: $\mathbf{c} = \mathbf{R}^{-1}\mathbf{p}$, where $\mathbf{p} = E[x_{\mathrm{S}}^*[k]\mathbf{z}[k]]$. As an unrealizable benchmark, we first consider a genie aided receiver, which does have access to the data and interference channel realizations, and thus ideal MMSE filter,

$$\mathbf{c}_{\mathrm{MMSE}}[k] = (\boldsymbol{\alpha}_{\mathrm{S}}[k]\boldsymbol{\alpha}_{\mathrm{S}}[k]^{\mathrm{H}} + \boldsymbol{\alpha}_{\mathrm{I}}[k]\boldsymbol{\alpha}_{\mathrm{I}}[k]^{\mathrm{H}} + 2\sigma^2\mathbf{I}_L)^{-1}\boldsymbol{\alpha}_{\mathrm{S}}[k]. \tag{3.10}$$

For continuously fading channels, this solution holds for the entire coherence length of the channel. In the wideband systems considered here, spatial covariances $\mathbf{\Sigma}_{\mathrm{S}}$ and $\mathbf{\Sigma}_{\mathrm{I}}$ are assumed to vary slowly with respect to the channel coherence length. Thus, accurate estimates of the desired user's spatial modes are achievable, at little cost, via pilot-based uplink measurements. Once the receiver has been trained to the data channel's spatial modes, the computed filters are applicable for the entire duration of the interference region. In simulation results, for example, less than 3% (or roughly 0.15 dB) of pilot overhead is required to adequately train the MMSE filters.

As a measure of performance of ideal MMSE interference suppression, we evaluate the Signal-to-Interference Ratio (SIR) at the correlator output, defined as,

$$\frac{E[\ |\langle \boldsymbol{\alpha}_{\mathrm{S}}[k], \mathbf{c}_{\mathrm{MMSE}}[k]\rangle|^2\ ]}{E[\ |\langle \boldsymbol{\alpha}_{\mathrm{I}}[k], \mathbf{c}_{\mathrm{MMSE}}[k]\rangle|^2\ ]}.$$

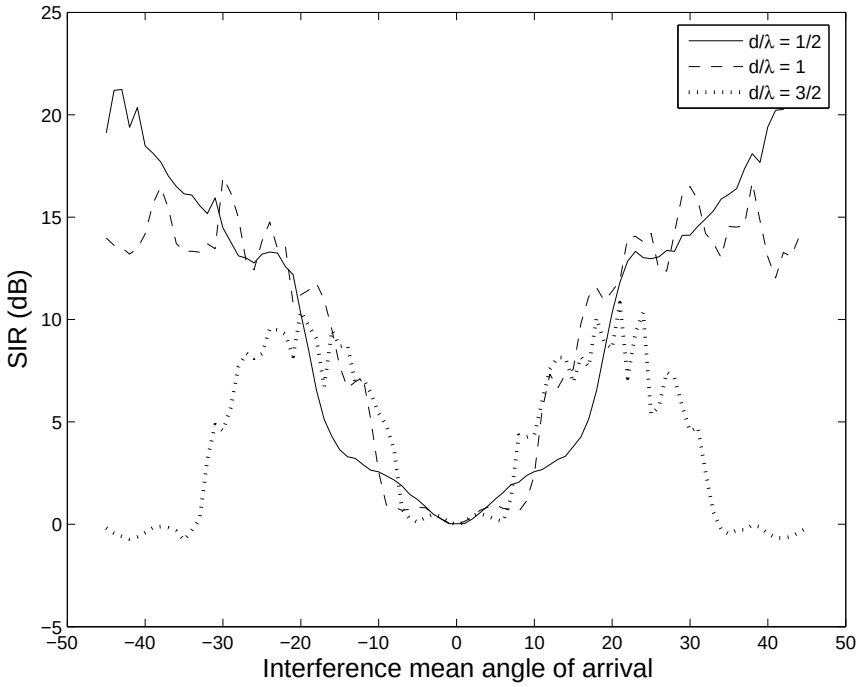


Figure 3.7: Ideal MMSE interference suppression with a 10 element receive array.

For this purpose, the signal and interference PAPs are modelled as Laplacian, $\mathcal{L}(\theta, \Delta\theta)$, with means $\{\theta\}$, and angular spread $\Delta\theta = 10°$. Figure 3.7 displays MMSE SIR as a function of the mean angle of arrival of the interference PAP, where the desired user's PAP is taken to be zero-mean. Observed, for example, is that half-wavelength spaced antenna arrays are in principle better able to resolve interference at mean angles of arrival between 30 and 45°.

In practice, the receiver employs the same, *averaged* MMSE filters, $\{\mathbf{c}_{avg}\}$, whose number depends on the number of dominant modes of the data channel, for the entire interference duration. They are computed from regular pilot symbols transmitted by the desired user. The pilot symbols are ideally transmitted in bursts of length chosen to match the coherence length of the channel, or sufficiently long to obtain accurate estimates of the instantaneous channel realization. By assumption, the channel $\boldsymbol{\alpha}(k) \triangleq \boldsymbol{\alpha}_{\mathrm{S}}[t]$ is constant over the kth coherence block

and is thus well estimated by

$$\mathbf{p}(k) = \frac{1}{T} \sum_{t \in \mathcal{T}(k)} x_{\mathrm{S}}^*[t]\mathbf{z}[t], \tag{3.11}$$

where $\mathcal{T}(k)$ denotes the set of symbol indices corresponding to the kth coherence block, and T is the coherence length of the channel. On the other hand, estimates of the spatial covariance, $\mathbf{R}$, do not require pilots, and are formed by averaging the received signal over the whole of the interference duration. Thus, the instantaneous MMSE filter estimate is given by $\mathbf{c}(k) = \mathbf{R}^{-1}\mathbf{p}(k)$. Due to the high degree of variability in fading channels, multiple pilot estimates are required to adequately train the receiver to the desired user's spatial covariance matrix. Thus, letting $\mathcal{P}$ denote the index set of P pilot blocks, the receiver is equipped with the correlator estimates $\{\mathbf{c}(k)\}$, $k \in \mathcal{P}$. Note that although each of these correlators is trained to the dominant mode of the data channel, they cannot be combined directly, since the phase of the data channel is i.i.d. uniform across pilot blocks. We thus propose the following solution for combining multiple pilot blocks. Letting $\mathbf{A} = \sum_{k \in \mathcal{P}} \mathbf{p}(k)\mathbf{p}(k)^{\mathrm{H}}$,

$$\mathbf{c}_{avg} = \mathbf{R}^{-1} \underset{\text{eigenvector}}{\text{dominant}} \{\mathbf{A}\} \tag{3.12}$$

is an estimate of the ideal MMSE filter trained to the dominant spatial mode of the data channel. Furthermore, employing the second dominant eigenvector of $\mathbf{A}$ in (3.12) gives an estimate of the MMSE filter trained to the second dominant mode of the data channel. In simulation results of coded systems, we have found this estimate to perform extremely well; for the details of its derivation see Appendix D.

3.10 Numerical results

We consider the performance of noncoherent eigenbeamforming with 0 dB interference, in which pilot blocks are employed to estimate the second order statis-

tics of the desired user's spatial channel. The model (3.9) is employed for simulation. That is, $\boldsymbol{\alpha}_{\mathrm{S}}[k] \sim \mathcal{CN}(0, \boldsymbol{\Sigma}_{\mathrm{S}})$, $\boldsymbol{\alpha}_{\mathrm{I}}[k] \sim \mathcal{CN}(0, \boldsymbol{\Sigma}_{\mathrm{I}})$, are generated explicitly for the simulation results, where Laplacian PAPs, and a 10 element half-wavelength spaced linear array, are considered for $\mathbf{C}_{\mathrm{S}}$ and $\mathbf{C}_{\mathrm{I}}$ computation. The PAPs of the data and interference channel are taken to be $\mathcal{L}(0°, 10°)$ and $\mathcal{L}(45°, 10°)$, respectively. These PAPs model single cluster channels, and are known to be characterized by one or two dominant spatial modes. The dimension of the reduced-space, span$\{\mathbf{V}\}$, is $L = 2$ ($L = 4$) for the one (two) mode demodulator. Moreover, the accuracy of the MMSE filter estimates (3.12) is expected to be superior to estimates computed in the original 10-dimensional complex vector space.

Figure 3.8 compares the performance of convolutional coded differential QPSK modulation at rate-1.35 bits/symbol with ideal and averaged MMSE interference suppression. The codeword length is chosen to be 64,000 bits, and the coherence length of the channel is $T = 10$ symbols. We first note that averaged MMSE performance is within 0.25 dB of ideal MMSE interference suppression, where approximately 0.15 dB is attributed to pilot overhead. Given the performance improvement for diversity combining with the top two modes of a one cluster channel with no interference, we employ the top two MMSE correlators (3.12) for diversity combining in the presence of interference. The results are also displayed in Figure 3.8, demonstrating comparable performance to that of the ideal single mode block fading channel with no interference.

3.11 Conclusions

This chapter focuses on transceiver design for increasing the capacity of wireless cellular uplinks. Forthcoming cellular systems, for example WMANs, will have to overcome hurdles such as severe channel dispersion, no line of sight between transmitter and receiver, mobility at vehicular speeds, and inter-cell interference

arising from shared spectrum scheduling. The fading rates that result in time and frequency may prohibit the use of conventional coherent transceiver designs. These issues motivate us to consider spectrally efficient noncoherent communication that does not rely on pilot-symbol based estimation of the continuously varying channel. Rather, the spatial covariance matrix is estimated with "feedback" implicitly available in wideband systems, while techniques of coded noncoherent communication are employed to approach noncoherent capacity, where no prior channel realization information is assumed at the transmitter or receiver. The outdoor channels that are considered here are typically characterized by a relatively few number of dominant spatial modes. Thus, by projecting along, and processing only, the dominant modes of the channel, the receiver is able to dramatically improve its per mode SNR while simultaneously obtaining reasonable levels of complexity. We further investigate a scenario in which inter-cell interference gives rise to non-white contributions to the received spatial covariance matrix. We propose an MMSE framework based on regularly transmitted pilot blocks to distinguish between the data and interference channels' dominant eigenmodes. Simulation results verify our hypothesis that significantly less pilot overhead (less than 3%) is required to measure the channel covariance than is required to measure its realization (which is expected to be as much as 20% [21]).

These concepts are demonstrated with simulation results, and capacity comparisons, for a proposed linear diversity combining receiver with convolutionally coded differential QPSK modulation. The diversity combining receiver, although suboptimal, performs extremely well for the case of two mode combining, and offers the advantage of leveraging well-established techniques of coding and modulation for single antenna noncoherent channels. Further avenues of research include more sophisticated receiver algorithms for joint processing of three or more spatial modes (at the expense of increased complexity). However, for the tradeoffs in beamforming gain and diversity level described in this paper, we argue that

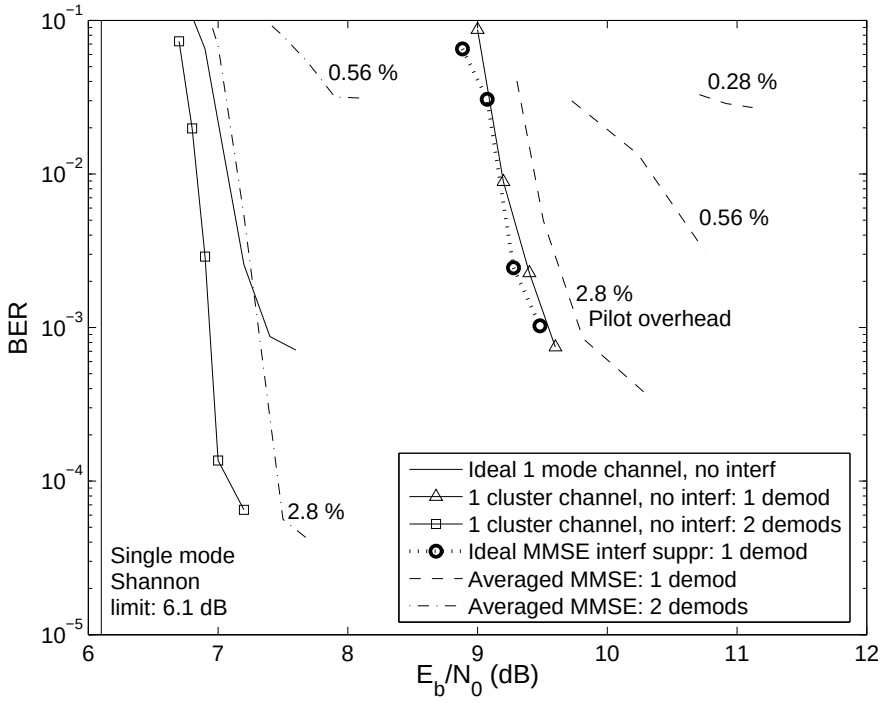


Figure 3.8: Performance of MMSE filters with 0 dB interference at mean angle of arrival 45°.

the proposed transceiver achieves the most gains, in terms of SNR increase and complexity reduction, for two mode channels. Further extensions to the case of multiple transmit antennas are also a topic of future research.

Chapter 4

Comparison with coherent transceiver

In this chapter, a detailed comparison of the performance of noncoherent and coherent transceiver designs in an OFDM system is conducted. Noncoherent transceivers hold the potential to significantly improve the throughput of wireless communication systems, as they eliminate the overhead of pilot symbol transmissions. However, to achieve a level of performance that is comparable to that of coherent systems, noncoherent systems typically require significantly more complex processing at the receiver. Thus, this chapter is devoted to a side-by-side comparison of the two approaches to wireless transceiver design in the context of a packetized OFDM system, exploring complexity and throughput tradeoffs. For the noncoherent transceiver, differential QPSK modulation with convolutional coding is employed. Noncoherent demodulation and decoding are performed as described in Chapter 2, with time-frequency symbol mapping and demodulator block size tailored to the OFDM channel. The coherent transceiver is composed of turbo-coded QPSK modulation with time-frequency channel estimation based on regularly transmitted pilot symbols. We assume 20% pilot overhead, as is expected in forthcoming wireless metropolitan area networks, with users moving at vehicular speeds and severe multipath environments. The coherent receiver also employs iterative demodulation and decoding, where feedback from the code is

used to refine channel estimates (referred to as decision directed channel estimation) in the subsequent demodulation iteration.

For the comparison, we consider a 4th Generation (4G) OFDM concept system that is under development at Motorola Labs [21]. Specifications of the OFDM system, such as number of subcarriers, subcarrier spacing, cyclic prefix length, and so forth, are the same for both systems. The proposed noncoherent transceiver (henceforth referred to as the differential system) is evaluated via direct throughput comparison with the coherent system. An effort is made to keep the decoding/demodulation complexity of the two systems roughly equivalent, so as to facilitate a fair comparison. In particular, as both the coherent and differential systems rely on APP computation as the basis for SISO demodulation and decoding, the complexity of both systems is estimated in terms of the number of max-star (log-APP) kernel operations required for complete decoding of a code word. Detailed descriptions of the OFDM system, channel model, and implementations of the differential and coherent systems are provided in the sequel.

4.1 Noncoherent OFDM

This section describes the application of noncoherent transceiver designs to OFDM systems. By definition, noncoherent systems do not rely on the transmission of pilot symbols for channel estimation and tracking. Rather, channel continuity in time and frequency is leveraged by a turbo architecture that iteratively demodulates and decodes the received data sequence. At the demodulator, channel uncertainty is handled with a block fading model with (1) signal energy based amplitude estimation and (2) Bayesian combination of coherent symbol probabilities that result from quantizing the unknown channel phase.

System model: The turbo architecture described in Section 2.1 of Chapter 2 is employed with convolutional channel coding and differential QPSK modulation.

For numerical evaluation, the channel is modeled with the Wide Sense Stationary with Uncorrelated Scattering (WSSUS) channel model, widely employed for modelling outdoor channels. Block noncoherent demodulation is based on approximating the channel as constant over an appropriate number of time and frequency bins. Thus, the transmitted symbol sequence is ideally mapped to time and frequency in a fashion that maximizes the correlation between channels seen by successive symbols. For this purpose, the OFDM channel is nicely represented with a time-frequency grid, illustrated in Figure 4.1, in which columns correspond to subcarriers and rows correspond to OFDM symbol intervals. The depicted arrows represent a mapping of QPSK symbols onto subcarriers of successive OFDM symbols. Thus, the depicted mapping is well suited to a time-frequency channel that does not vary significantly over four OFDM symbols. Note that the receiver is free to choose the number of subcarriers for which the block fading approximation makes sense, but is constrained by the transmitter as for the block length in time. For example, the Typical Urban (TU) WSSUS channel, which models severe channel dispersion and is employed here for numerical simulations, is best approximated as constant over two subcarriers. Thus, a demodulator block size of nine symbols, where one symbol overlaps adjacent blocks (see Section 2.1.1), best suits the time-frequency symbol mapping of Figure 4.1. Note that in contrast to the block fading channel model, block noncoherent demodulation of continuously modulated differential data with a continuously fading channel does not incur a one symbol per block rate penalty.

The received symbol vector for one block of transmitted symbols may be expressed as

$$\mathbf{y} = h'\mathbf{x}' + \mathbf{w} = h\mathbf{x} + \mathbf{w}, \tag{4.1}$$

where $h = h'x[0]$ and $\mathbf{x} = x^*[0]\mathbf{x}'$. Note that, for the QPSK modulation considered here, h and h' are identically distributed. In general, the received symbol block may be de-rotated by the phase of the reference symbol $x[0]$ (first symbol of

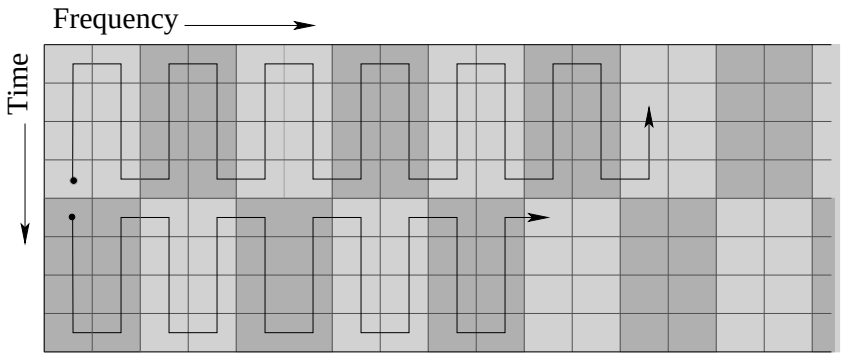


Figure 4.1: Mapping the transmitted symbol sequence to a time/frequency channel, for block noncoherent demodulation.

the block) without any affect on the channel model or symbol estimates. Thus, conditional on the channel phase, the demodulator may estimate $\mathbf{x}$ assuming the path traversed on the trellis of the modulation code begins from the zero state.

4.2 Coherent system

For the coherent system, we assume that one pilot symbol is transmitted for every four data bearing symbols. The transmitter accepts a frame of information bits, performs turbo-encoding and interleaving and Gray-maps the resulting coded bits to QPSK symbols. As in the differential system, the transmitted codeword extends over four OFDM symbols, with one QPSK symbol transmitted on each OFDM subcarrier and a pilot symbol preceding and succeeding each set of four QPSK symbols in time. The receiver employs the pilot symbols to estimate the channel and thus the transmitted codeword. Initially, iterations of the standard turbo-decoding algorithm are performed based on channel estimates obtained from the pilot symbols. The resulting bit-level APPs are then converted to probabilities on the transmitted symbols for refining the channel estimates in the next demodulation iteration. Thus, coherent systems also benefit from iterative demodulation and decoding, though, based on the results of the comparison performed here, to a lesser extent than noncoherent systems.

In the first iteration, channel estimation is performed solely based on the pilot symbols. In particular, the demodulated pilot symbols provide raw estimates of the channel gain for each subcarrier in the OFDM symbol in which they appear. Leveraging channel continuity in frequency, the raw channel estimates are then smoothed across subcarriers with an FIR filter corresponding to the coherence bandwidth of the channel. Channel estimates for the data symbols of a given subcarrier are then obtained by linearly interpolating the smoothed channel estimates preceding and succeeding the data symbols.

In subsequent iterations, channel estimates are refined according to feedback from the outer code. The per-symbol raw channel estimates for the data symbols are obtained as the solution to the following MMSE formulation:

$$\hat{h} = \arg\min_{h} E[|y[n] - hx[n]|^2], \tag{4.2}$$

where expectation is taken with respect to the PDF of the transmitted symbols, obtained by converting code-bit APPs yielded by the turbo-decoder to symbol-level probabilities. The solution is given by $y[n]E[x^*[n]]/\left|E[x[n]]\right|$. As before, the raw channel estimates (now for each OFDM symbol) are smoothed in frequency with an FIR filter. Thus, at this point of the decoder algorithm, there are six channel estimates obtained for each frequency bin, corresponding to the preceding and succeeding pilot symbols, and four data-bearing symbols. They are viewed as estimates of time samples of the channel fading process in each frequency bin. As a first order approximation, the channel fading process is assumed to be linear across two pilot symbols. Thus, the final channel estimate for the data symbols of each subcarrier is obtained by sampling a line fitted to the six channel estimates at the appropriate time instance.

4.3 Numerical results

In this section, results of the numerical comparison between the differential and coherent transceivers in the context a packetized OFDM system are presented. The comparison is performed in terms of the Frame Error Rate (FER) and throughput attained by the receivers. The WSSUS channel model with a TU power delay profile is employed for the comparison. Each ray of the power delay profile of the WSSUS channel constitutes an independent fading channel at the Doppler rate of 160 Hz, roughly equivalent to a speed of 30 MPH, at a carrier frequency of 3.676 GHz. The parameters of each system are chosen to maintain equivalent levels of complexity, defined in terms of the number of log-APP binary kernel operations per second (bkops) [21].

The parameters of the OFDM system are chosen to model a packetized system with relatively short codeword length, corresponding to one transmitted frame. The FFT size is 1024, sample rate 25.6 MHz, and length of the cyclic prefix is 256. The corresponding OFDM symbol duration is 50 μsec. The number of subcarriers used for data transmission is 768, corresponding to a system bandwidth of 19.2 MHz. In both the differential and coherent systems, coding is performed over four contiguous OFDM symbols, resulting in a codeword length of 6144 bits. For the coherent system, pilot symbols are inserted between each codeword frame, resulting in a throughput loss of 4/5. Both systems employ four iterations of demodulation and decoding. In the differential system, a standard rate-1/2 convolutional code of memory 6 is chosen, resulting in a maximum throughput (corresponding to zero frame errors) of 15.4 Mbps, at the computational complexity of 28.6 giga-bkops. The turbo-code employed by the coherent receiver consists of two rate-1/3, memory 3, convolutional codes, for an overall rate of 1/2. The corresponding maximum throughput is 12.28 Mbps, at the complexity of 24.6 giga-bkops.

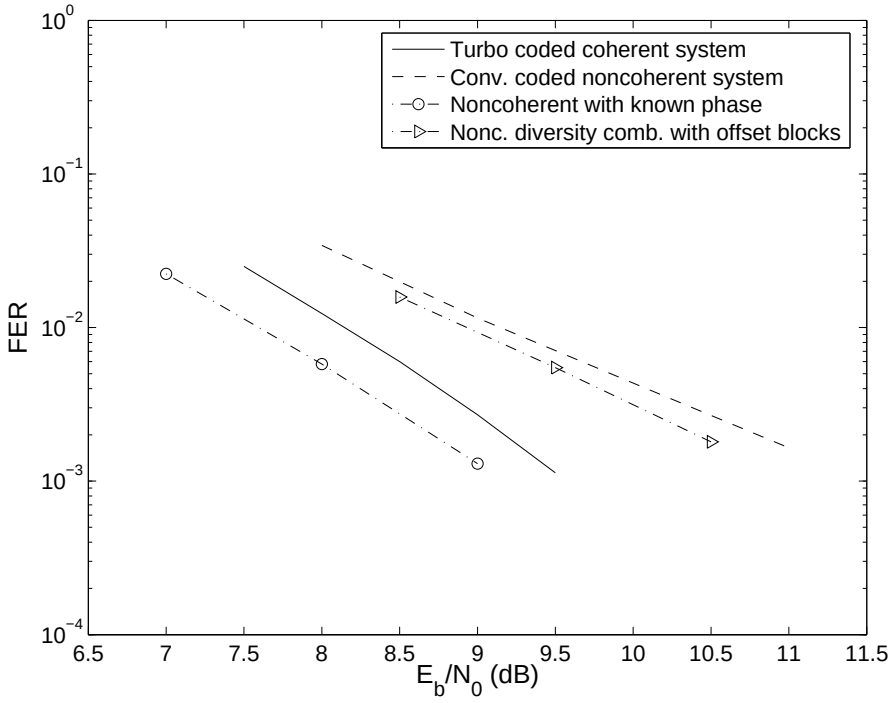


Figure 4.2: Raw FER of coherent and noncoherent transceivers.

The numerical results are displayed in Figures 4.2 and 4.3. The FER of the two systems is shown in Figure 4.2. The results demonstrate that the pilot-based system is better able to leverage channel continuity in frequency than the noncoherent system in its current form. This arises from the fact that block noncoherent demodulation assumes independent channel realizations from block to block. Also depicted is the performance of a genie-aided noncoherent receiver which has access to the true channel phase. Since there is a significant gap to the genie-aided receiver, we conjecture that there is room for improvement in the noncoherent system with more detailed modelling of channel frequency variation than is provided by the i.i.d. block fading model. However, design of such a receiver is beyond the scope of this paper and will likely incur additional complexity cost. Nonetheless, also reported in Figure 4.2 is an initial foray into this interesting topic, consisting of a diversity combining receiver similar to that employed for the parallel block fading channel. The diversity combining receiver demodulates the received symbol

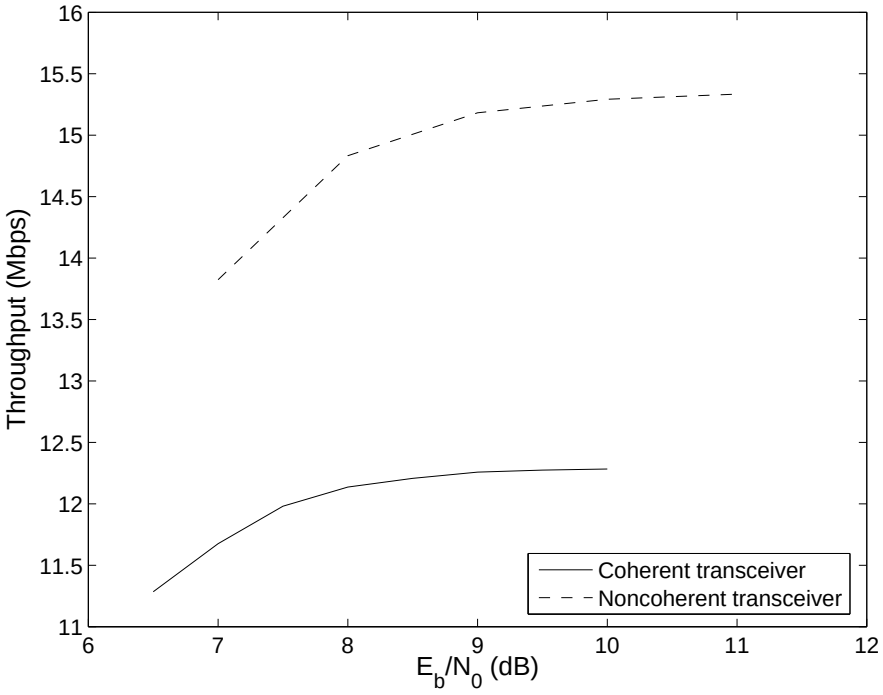


Figure 4.3: Throughput advantage of noncoherent transceiver designs.

sequence twice, with offset blocks in frequency, and linearly combines the resulting symbol posteriors. This approach essentially mitigates edge effects of noncoherent block demodulation for symbols at block boundaries, yielding a modest gain of roughly 0.4 dB.

Figure 4.3 demonstrates the throughput advantage of noncoherent communication, where throughput is defined as

$$T(SNR) = (1 - FER(SNR))R_{max}, \tag{4.3}$$

and R_{max} is the maximum bit rate supported by a given combination of modulation and channel code.

4.4 Discussion

Noncoherent iterative demodulation and decoding is demonstrated to be a viable alternative to standard receivers that rely on pilot symbols for channel estimation, demodulation and decoding. As demonstrated in this chapter, omission of pilot transmissions results in higher throughputs for the differential system relative to the coherent system. As noncoherent iterative schemes require Bayesian processing of the unknown channel phase, they are inherently more complex than standard coherent receivers. On the other hand, there is more to be gained in terms of the performance of coded differential systems by allowing for more iterations, and thus higher levels of complexity. This follows from the convolutional structure of differential modulation, which introduces memory into the modulated data and thus better exploits the channel's memory with SISO demodulation and decoding in a turbo-architecture. However, since the performance results reported in this Chapter were obtained at roughly equivalent levels of complexity, we find that this complexity increase may not be prohibitive. Moreover, there is clearly room for improvement in the noncoherent system by leveraging channel continuity between blocks. Since the throughput is defined in terms of the FER attained, the advantage of noncoherent systems will only increase as inter-block continuity is better leveraged by the receiver. This is a subject of future research.

Chapter 5

Conclusions

This body of work is primarily focused on attaining capacity of wireless time-varying channels. We have examined three main subjects:

Design of modulation and channel codes for efficient noncoherent communication: We have considered a single antenna transmitter and receiver pair in which no prior knowledge of channel realizations is assumed. Information theoretic computations demonstrate the advantage of large amplitude/phase constellations for the noncoherent block fading channel, as is well-known to be the case for a classical AWGN channel. Yet, heretofore, the research community has not considered the appropriate choice of QAM modulation code for turbo-architectures designed to approach capacity of the block fading channel.

Our research (i) finds that QAM constellations classically employed for signalling over AWGN channels perform poorly over noncoherent fading channels, and (ii) yields new constellations well-matched to noncoherent fading channels that realize the capacity advantage of amplitude/phase signalling. These results are facilitated by the development of EXIT analysis techniques for coded noncoherent communication. The analysis techniques developed here are useful in their own right, for guidance in the design of high performance coded modulation systems. For example, EXIT analysis of coded noncoherent communication provides a rigorous explanation for the near optimality of simple convolutional

codes in concatenation with unit-rate differential modulation. This fact, observed previously in the literature, had only been justified with heuristics.

Throughout this work we have consistently focused on practical implementations of the noncoherent transceivers developed. The proposed reduced complexity receiver loses only 0.1 dB to the full complexity implementation in performance simulations. Moreover, after the first iteration of demodulation and decoding, the complexity is only twice that of genie-aided demodulation of the noncoherent code, in which the channel is assumed to be known. In summary, a practical coded noncoherent transceiver is developed that realizes the capacity advantage of large amplitude/phase constellations over PSK for alphabets of size 16 or larger and SNR greater than 6.5 dB.

Eigenbeamforming receivers: Accurate estimates of the spatial covariance matrix, available through averaging in wideband systems, allow eigenbeamforming at the base station along the dominant channel modes. For a typical outdoor channel, where the number of dominant modes is small, this allows the receiver to increase its SNR by scaling up the number of antennas, while limiting the demodulation and decoding complexity (which scales with the number of channel modes used by the receiver). Averaging over time and frequency compensates for the reduced level of spatial diversity in outdoor systems. The proposed eigenbeamforming receiver is consistent with the approach of noncoherent transceiver design, in contrast to classical coherent receive beamforming based on explicit channel estimates. This can potentially eliminate the significant pilot overhead required for estimating the channel to each mobile at the base station.

The case of spatial interference at the receive array is handled with MMSE interference suppression techniques, yielding excellent performance with little pilot overhead. Thus, noncoherent eigenbeamforming is demonstrated to be a practical and realizable transceiver strategy that is well-suited to emerging wireless cellular systems.

Comparison with coherent transceiver: Although noncoherent transceivers are not yet able to achieve the same raw FER as their pilot-based coherent counterparts, the claimed throughput advantage is verified in the context of next generation cellular systems. Clearly, the performance of noncoherent transceivers benefits from allowing additional complexity, much more so than coherent systems. More importantly, however, noncoherent wireless transceivers will have to better leverage channel continuity in order to attain indisputable levels of performance gain. We conjecture that this goal is fully realizable, and will perhaps be facilitated as the fundamental limits of wireless time-varying channels are better understood. Progress in this direction is reported in [11], where the Gauss-Markov channel model for continuous variations is employed to derive capacity bounds and design near capacity performing receivers.

5.1 Open issues

Although we consider the design of capacity approaching wireless communication systems for time-varying channels, information theoretic guidance is based on the block fading model which fails to account for channel continuity between blocks, and thus falls sort in describing truly continuous fading channels. The Shannon limits of continuously varying noncoherent channels is still very much an open research topic.

Initial results in this direction [23] show that, for a general ergodic and stationary fading process, the MIMO capacity is dominated by a $\log \log$ SNR term in the limit as SNR tends to infinity, regardless of the number of antenna elements. This is in stark contrast to coherent MIMO capacity [35], which scales linearly in the minimum of the number of transmit and receive elements (the so-called degrees of freedom). The model employed in [23] essentially assumes a fixed symbol-to-

symbol channel variation, which dictates performance when the noise power is small.

However, more recently, it is argued [13] that this operating regime does not describe most practical cases of interest, in which the channel is underspread (with fading rates much slower than the system bandwidth) and thus both channel variation and noise determine capacity. Their result indicates that for a Gauss-Markov fading process, with channel variation and noise power tending to zero simultaneously, capacity again exhibits a linear growth in the degrees of freedom, and thus communication systems benefit from the use of multiple antennas.

It is certainly evident from both of these works that the SNR regime of operation drastically affects the capacity scaling laws of continuous fading channels, and thus plays a pivotal role in system design. Moreover, emerging results [13, 39] indicate that neither the SNR nor fading rate alone determines the behavior of capacity, but that these parameters must be analyzed jointly to determine the Shannon theoretic limits of real world channels.

Bibliography

[1] A. Ashikhmin, G. Kramer, and S. ten Brink. Extrinsic information transfer functions: model and erasure channel properties. *IEEE Trans. Information Theory*, to appear.

[2] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Information Theory*, 20:284–287, Mar. 1974.

[3] G. Barriac, N. Jacobsen, and U. Madhow. Beyond BAD: A parallel arbitration framework for low-complexity equalization. *Proc. 39th Annual Allerton Conf. on Communication, Control, and Computing*, October 2001.

[4] G. Barriac and U. Madhow. Wideband space-time communication with implicit channel feedback. In *Seventh International Symposium on Signal Processing and its Applications*, 2003.

[5] G. Barriac and U. Madhow. Characterizing outage rates for space-time communication over wideband channels. 52(12):2198–2208, Dec. 2004.

[6] G. Barriac and U. Madhow. Space-time communication for OFDM with implicit channel feedback. 50(12):3111–3129, Dec. 2004.

[7] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. A soft-input soft-output maximum a posteriori (MAP) module to decode parallel and serial

concatenated codes. *Jet Propulsion Laboratory, Pasadena, CA, PL TDA Prog. Rep. 42-127*, Nov. 1996.

[8] R.-R. Chen, D. Agrawal, and U. Madhow. Noncoherent detection of factor-graph codes over fading channels. In *Proc. Conf. on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, Mar. 2000.

[9] R.-R. Chen, B. Hajek, R. Koetter, and U. Madhow. On fixed input distributions for noncoherent communication over high SNR fading channels. *IEEE Trans. Information Theory*, to appear.

[10] R.-R. Chen, R. Koetter, D. Agrawal, and U. Madhow. Joint demodulation and decoding for the noncoherent block fading channel: a practical framework for approaching channel capacity. *IEEE Trans. Communications*, 51(10):1676–1689, Oct. 2003.

[11] R.-R. Chen, R. Koetter, and U. Madhow. Joint noncoherent demodulation and decoding for fast Rayleigh fading channels. In *Proc. Conf. on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, Mar. 2003.

[12] D. Divsalar and M. Simon. Multiple-symbol differential detection of MPSK. *IEEE Trans. Communications*, 38(3):300–308, Mar. 1990.

[13] R. Etkin and D. Tse. Degrees of freedom in underspread MIMO fading channels. *IEEE Trans. Information Theory*, submitted for publication.

[14] D. Gesbert, M. Shafi, S. Da-shan, P. Smith, and A. Naguib. From theory to practice: an overview of MIMO space-time coded wireless systems. *IEEE Trans. Communications*, 21(3):281–302, Apr. 2003.

[15] P. Hoeher and J. Lodge. Turbo DPSK: Iterative differential PSK demodulation and channel decoding. *IEEE Trans. Communications*, 47(6):837–842, June 1999.

[16] N. Jacobsen, G. Barriac, and U. Madhow. Noncoherent eigenbeamforming for a wideband cellular uplink. In *Proc. IEEE International Symp. on Information Theory (ISIT)*, Chicago, IL, USA, June 2004.

[17] N. Jacobsen, G. Barriac, and U. Madhow. Noncoherent receive beamforming for frequency selective time-varying channels, Sept. 2005. In preparation.

[18] N. Jacobsen, G. Barriac, U. Madhow, R.-R. Chen, and R. Koetter. Towards Shannon-theoretic limits on wireless time-varying channels. In *Proc. International Symp. on Control, Communications, and Signal Processing (ISCCSP)*, Hammamet, Tunisia, Mar. 2004.

[19] N. Jacobsen and U. Madhow. Reduced-complexity noncoherent communication with differential QAM and iterative receiver processing. In *Proc. Conf. on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, Mar. 2003.

[20] N. Jacobsen and U. Madhow. Coded noncoherent communication with amplitude/phase modulation: from Shannon theory to practical turbo architectures, Oct. 2004. Submitted for publication.

[21] N. Jacobsen, E. Visotsky, and R. Peterson. Noncoherent iterative demodulation and decoding for OFDM systems. *Motorola Labs Tech. Report*, Feb. 2005.

[22] F. Kschischang, B. Frey, and H. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Trans. Information Theory*, 47(2):498–519, Feb. 2001.

[23] A. Lapidoth and S. Moser. Capacity bounds via duality with applications to multi-antenna systems on flat fading channels. *IEEE Trans. Information Theory*, 49(10):2426–2467, Oct. 2003.

[24] A. Lapidoth and S. Shamai. Fading channels: how perfect need "perfect side information" be? *IEEE Trans. Information Theory*, 48(5):1639–1667, June 1999.

[25] Y. Liang and V. Veeravalli. Capacity of noncoherent time-selective block Rayleigh flat-fading channel. In *Proc. IEEE International Symp. on Information Theory (ISIT)*, Lausanne, Switzerland, June 2002.

[26] O. Macchi and L. Scharf. A dynamic programming algorithm for phase estimation and data decoding on random phase channels. *IEEE Trans. Information Theory*, Sept. 1981.

[27] T. Marzetta and B. Hochwald. Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading. *IEEE Trans. Information Theory*, 45(1):139–157, Jan. 1999.

[28] R. Nuriyev and A. Anastasopoulos. Capacity characterization for the noncoherent block-independent AWGN channel. In *Proc. IEEE International Symp. on Information Theory (ISIT)*, Yokohama, Japan, July 2003.

[29] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, San Mateo, CA, 1988.

[30] K. Pedersen, P. Mogensen, and B. Fleury. Dual-polarized model of outdoor propagation environments for adaptive antennas. In *IEEE 49th Vehicular Technology Conference*, volume 2, pages 990–5, May 1999.

[31] M. Peleg, S. Shamai, and S. Galán. Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channel. *IEE Proc. Communications*, 147(2):87–95, Apr. 2000.

[32] T. Richardson, M. Shokrollahi, and R. Urbanke. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. Information Theory*, 47(2):619–637, Feb. 2001.

[33] T. Richardson and R. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans. Information Theory*, 47(2):599–618, Feb. 2001.

[34] A. Saleh and R. Valenzuela. A statistical model for indoor multi-path propagation. *IEEE Journal on Selected Areas in Communications*, 5(2):128–137, February 1987.

[35] E. Telatar. Capacity of multi-antenna Gaussian channels. *AT&T Bell Labs Internal Tech. Memo # BL0112170-950615-07TM*, June 1995.

[36] S. ten Brink. Convergence behavior of iteratively decoded parallel concatenated codes. *IEEE Trans. Communications*, 49(10):1727–1737, Oct. 2001.

[37] M. Tüchler and J. Hagenauer. EXIT charts of irregular codes. In *Proc. Conf. on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, Mar. 2002.

[38] D. Warrier and U. Madhow. Spectrally efficient noncoherent communication. *IEEE Trans. Information Theory*, 48(3):651–668, Mar. 2002.

[39] L. Zheng, M. Medard, D. Tse, and C. Luo. On channel coherence in the low SNR regime. In *Proc. IEEE International Symp. on Information Theory (ISIT)*, Chicago, IL, USA, June 2004.

[40] L. Zheng and D. Tse. Communicating on the Grassmann manifold: A geometric approach to the non-coherent multiple antenna channel. *IEEE Trans. Information Theory*, 48(2):359–383, Feb. 2002.

Appendices

Appendix A

BCJR recursions

The forwards/backwards recursions employed for block noncoherent APP demodulation are, conditional on the channel, h,

$$\alpha_n(s) = \max_{e:s^F(e)=s}{}^{\star} \left\{ \gamma_n(e|h) + \alpha_{n-1}(s^I(e)) \right\}, \quad \alpha_1(s) = \sigma^{-2}\, \Re\langle y[1], hs \rangle, \quad s \in \mathcal{A}, \tag{A.1}$$

$$\beta_n(s) = \max_{e:s^I(e)=s}{}^{\star} \left\{ \gamma_{n+1}(e|h) + \beta_{n+1}(s^F(e)) \right\}, \quad \beta_T(s) = -\log_2(M), \quad s \in \mathcal{A}. \tag{A.2}$$

Appendix B

Channel amplitude estimation

We present an amplitude estimate that is well matched QAM constellations based on concentric rings. Developed here for the case of two ring-level, M-ary alphabets, the estimate generalizes easily to larger constellations. The complexity is linear in the product of the number of amplitude levels and size of the modulation alphabet. The polar representation of the nth transmitted symbol is defined as $x[n] = r[n]\exp(j\phi[n])$, where $r[n] \in \mathcal{R} = \{r_0, r_1\}$ denotes the amplitude level of the nth symbol, and $\phi[n]$ the phase. Conventional two-symbol differential detection statistics do not require channel knowledge, and are used to obtain the distribution of $r[n]$. The *amplitude likelihoods* (B.3) are then used to compute an energy based estimate of a. Note that this bootstrap is not required of constant-amplitude, PSK signaling [10]. The vector of two received symbols is

$$\mathbf{y}[n] \triangleq \begin{bmatrix} y[n-1] \\ y[n] \end{bmatrix} = h \begin{bmatrix} r[n-1] \\ r[n]\exp(j\phi[n]) \end{bmatrix} + \begin{bmatrix} w[n-1] \\ w[n] \end{bmatrix} = h\mathbf{x}[n] + \mathbf{w}[n]. \quad \text{(B.1)}$$

We note that the phase of the $(n-1)$th symbol may be factored into the channel, without changing its density. The log likelihood of the received symbol given the transmitted symbol is

$$\log P(\mathbf{y}[n]|\mathbf{x}[n]) = \frac{1}{2\sigma^2}\frac{\left|\mathbf{y}[n]^{\mathrm{H}}\mathbf{x}[n]\right|^2}{\|\mathbf{x}[n]\|^2 + 2\sigma^2} - \log(\|\mathbf{x}[n]\|^2 + 2\sigma^2) + const. \quad \text{(B.2)}$$

Defining the probability of the nth symbol amplitude, $r[n]$, as

$$\mu_n(r) \triangleq P(r[n] = r|\mathbf{y}[n]) \propto \sum_{\mathbf{x}\in\mathcal{R}\times\mathcal{A}:r[n]=r} P(\mathbf{y}[n]|\mathbf{x}[n] = \mathbf{x}), \tag{B.3}$$

the energy based channel amplitude estimate, $\hat{a}$, is computed with (B.4), where T denotes the coherence length of the channel,

$$\hat{a}^2 = \max\left\{0, \frac{\|\mathbf{y}\|^2 - 2T\sigma^2}{\sum_{n=1}^{T}\sum_{r\in\mathcal{R}}\mu_n(r)r^2}\right\}. \tag{B.4}$$

Appendix C

Parallel block fading capacity calculation

Property 4. Any MPSK signal vector can be written as a unitary matrix times the all ones vector $\mathbf{x} = \mathbf{\Phi}(\mathbf{x})\mathbf{1}$, where $\mathbf{\Phi}(\mathbf{x}) = \text{diag}(x_1, x_2, ..., x_T)$. We shall refer to $\mathbf{\Phi}$ of this form as a "discrete rotation," since the signal alphabet, $\mathcal{X} = \mathbf{\Phi}\mathcal{X}$, is the same with all of the elements multiplied by $\mathbf{\Phi}$.

The following lemma shows that the conditional PDF of the parallel block fading channel is consistent with respect to rotations of the underlying signal space. An immediate consequence is that the PDF of Y is unaffected by discrete rotations.

Lemma 1. $P(\mathbf{UY}|\mathbf{x}) = P(\mathbf{Y}|\mathbf{U}^{\text{H}}\mathbf{x})$ for any unitary transformation $\mathbf{U}$.

Proof. Let $f_N(\mathbf{n})$ denote the noise PDF. By the circular symmetry of Gaussian noise:

$$\begin{aligned} P(\mathbf{UY}|\mathbf{x}) &= \prod_{l=1}^{L} P(\mathbf{U}\mathbf{y}_l|\mathbf{x}) \\ &= \prod_{l=1}^{L} \int dh_l P(h_l) f_N(\mathbf{U}\mathbf{y}_l - h_l\mathbf{x}) \\ &= \prod_{l=1}^{L} \int dh_l P(h_l) f_N(\mathbf{y}_l - h_l\mathbf{U}^{\text{H}}\mathbf{x}) \\ &= P(\mathbf{Y}|\mathbf{U}^{\text{H}}\mathbf{x}) \end{aligned}$$

□

Corollary 1. For any discrete rotation, $\mathbf{\Phi}$, $P(\mathbf{\Phi Y}) = P(\mathbf{Y})$.

Proof. Given the above Lemma, note,

$$\begin{aligned} P(\mathbf{\Phi Y}) &= \frac{1}{M^T} \sum_{\mathbf{x} \in \mathcal{X}} P(\mathbf{\Phi Y}|\mathbf{x}) \\ &= \frac{1}{M^T} \sum_{\mathbf{x} \in \mathcal{X}} P(\mathbf{Y}|\mathbf{\Phi}^{\mathrm{H}}\mathbf{x}) \\ &= \frac{1}{M^T} \sum_{\mathbf{x} \in \mathbf{\Phi}^{\mathrm{H}}\mathcal{X}} P(\mathbf{Y}|\mathbf{x}) \\ &= P(\mathbf{Y}) \end{aligned}$$

□

Proposition 1. $H(\mathbf{Y}) = -E[\log P(\mathbf{Y})|\mathbf{x} = \mathbf{1}]$

Proof.

$$\begin{aligned} H(\mathbf{Y}) &= -\int d\mathbf{Y} P(\mathbf{Y}) \log P(\mathbf{Y}) \\ &= -\frac{1}{M^T} \int d\mathbf{Y} \big(\sum_{\mathbf{x} \in \mathcal{X}} P(\mathbf{Y}|\mathbf{x}) \big) \log P(\mathbf{Y}) \\ &= -\frac{1}{M^T} \sum_{\mathbf{x} \in \mathcal{X}} \int d\mathbf{Y} P(\Phi(\mathbf{x})^{\mathrm{H}}\mathbf{Y}|\mathbf{1}) \log P(\mathbf{Y}) \\ &= -\frac{1}{M^T} \sum_{\mathbf{x} \in \mathcal{X}} \int d\Phi(\mathbf{x})\mathbf{Y} P(\mathbf{Y}|\mathbf{1}) \log P(\Phi(\mathbf{x})\mathbf{Y}) \\ &= -\int d\mathbf{Y} P(\mathbf{Y}|\mathbf{1}) \log P(\mathbf{Y}) \\ &= -E[\log P(\mathbf{Y})|\mathbf{x} = \mathbf{1}], \end{aligned}$$

since the Jacobian determinant of a unitary matrix is one. □

Finally, regarding the conditional entropy of $\mathbf{Y}$ given $\mathbf{x}$, note that, conditioned on $\mathbf{x}$, $\{\mathbf{y}_l\}$ are independent Gaussian random vectors of entropy $H(\mathbf{y}_l|\mathbf{x}) = \log\det(\pi e E[\mathbf{y}_l \mathbf{y}_l^{\mathrm{H}}])$. Thus,

$$H(\mathbf{Y}|\mathbf{x}) = \sum_{l=1}^{L} H(\mathbf{y}_l|\mathbf{x}) = LT \log(2\pi e\sigma^2) + \log\Big(\prod_{l=1}^{L} (1 + \frac{\lambda_l T}{2\sigma^2}) \Big).$$

Using (C.1) and Proposition 1 with Monte-Carlo integration, we can efficiently compute the noncoherent capacity of the $L \times 1$ block fading channel.

Appendix D

MMSE interference suppression correlators

The following proposition pertains to the combining of multiple pilot measurements for computing interference suppression correlators. The proposition yields an estimates of the MMSE correlators trained to the dominant spatial modes of the desired signal.

Proposition 2. Let $\mathbf{\Sigma}_{\mathrm{S}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}$, $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_L]$, $\mathbf{\Lambda} = \mathrm{diag}([\lambda_1 \cdots \lambda_L])$, denote the eigendecomposition of the desired user's spatial covariance matrix. The data channel may thus be written as $\boldsymbol{\alpha}(k) = \mathbf{U}[\alpha_1(k) \cdots \alpha_L(k)]^{\mathrm{T}}$, where $\{\alpha_l(k)\}$ are independent Gaussian random variables, $\alpha_l(k) \sim \mathcal{CN}(0, \lambda_l)$. Then, with $\mathbf{A} = \sum_{k\in\mathcal{P}} \mathbf{p}(k)\mathbf{p}(k)^{\mathrm{H}}$,

$$\mathbf{c}_{avg} = \mathbf{R}^{-1} \underset{\text{eigenvector}}{\text{dominant}} \{\mathbf{A}\} \tag{D.1}$$

is an estimate of the MMSE filter, $\mathbf{c}_1 = \arg\min E[|\langle \mathbf{z}[k], \mathbf{c}\rangle - \alpha_1(p)x_{\mathrm{S}}[k]|^2]$, that is trained to the dominant mode of the data channel.

Proof. First, by the law of large numbers,

$$\lim_{T\to\infty} \mathbf{p}(k) = E[x_{\mathrm{S}}^*[k]\mathbf{z}[k]] = \boldsymbol{\alpha}(k) \tag{D.2}$$

Now, letting $\mathbf{\Psi}(P,T) = \frac{1}{P}\mathbf{A}$,

$$\lim_{P,T\to\infty} \mathbf{\Psi}(P,T) = \lim_{P\to\infty} \frac{1}{P} \sum_{k\in\mathcal{P}} \boldsymbol{\alpha}(k)\boldsymbol{\alpha}(k)^{\mathrm{H}} = \mathbf{\Sigma}_{\mathrm{S}}. \tag{D.3}$$

The eigendecomposition of $\mathbf{\Psi}(P,T)$ converges to $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}$ with probability one. Thus, the dominant eigenvectors of $\mathbf{A}$ yield estimates of the dominant spatial modes of the data channel, $\{\mathbf{u}_l\}$. Finally, the MMSE correlator trained to the lth dominant spatial mode of the desired user, defined as

$$\arg\min E[|\langle \mathbf{z}[t], \mathbf{c}\rangle - \alpha_l(k) x_{\mathrm{S}}[t]|^2], \tag{D.4}$$

is given by $\mathbf{c}_l = \mathbf{R}^{-1}\mathbf{u}_l$. □