University of California
Los Angeles

Universal trellis codes and concatenated trellis-coded modulations for the compound linear vector Gaussian channel

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

by

Cenk Kose

2004

The dissertation of Cenk Kose is approved.

Michael P. Fitz

Lieven Vandenberghe

Kirby A. Baker

Richard D. Wesel, Committee Chair

University of California, Los Angeles

2004

To my dad

Table of Contents

List of Figures

List of Tables

Vita

1975	Born, Ankara, Türkiye.
1998	B.S., Electrical Engineering, Honors, Boğaziçi University, İstanbul, Türkiye.
1998–2000	Graduate research assistant, Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA.
2000	M.S., Electrical Engineering, University of Massachusetts, Amherst, MA.
2000-2003	Graduate student researcher, Electrical Engineering Department, UCLA.
2003	Teaching assistant, Electrical Engineering Department, UCLA. Graduate Information Theory and undergraduate Probability.
2003	Summer intern, Conexant Systems, Inc., San Diego, CA.
2004	Summer intern, Conexant Systems, Inc., San Diego, CA.
2004	Ph.D., Electrical Engineering, UCLA.

Publications

- C. Köse, R. Wesel, "High-rate universal space-time codes from standard trellis codes", accepted for GLOBECOM 2004.

- C. Köse, R. Wesel, "Universal space-time trellis codes," *IEEE Trans. Inform. Theory, Special Issue on Space-Time Transmission, Reception, Coding and Signal Design*, Vol. 49, No. 3, pp. 2717-2727, Oct. 2003.

- W. Weng, C. Köse, R. Wesel, "Serially-concatenated trellis-coded modulation for the periodic compound erasures channel," submitted to *IEEE Trans. Commun.*, Dec. 2003.

- C. Köse, W. Weng, R. Wesel, "Serially concatenated trellis coded modulation for the compound periodic erasures channel,", IEEE International Conference on Communications, ICC 2003, Vol. 4, pp. 2953-2957, Anchorage, AL, May 2003.

- C. Köse, R. Wesel, "Universal space-time trellis codes," Proc. IEEE Global Telecommunications Conference, GLOBECOM 2002, Vol. 2, pp. 1108-1112, Taipei, Taiwan, Nov. 2002.

- C. Köse, R. Wesel, "Robustness of likelihood-ratio tests," Conf. Rec., 35th Asilomar Conference on Signals, Systems and Computers, Vol. 2, pp. 1738-1742, Asilomar, CA, Nov. 2001.

- C. Köse, R. Wesel, "Code design metrics for space-time systems under arbitrary fading," Conf. Rec., International Conference on Communications, ICC 2001, Vol. 4, pp. 1099-1103, Helsinki, Finland, June 2001.

- C. Köse, D. Goeckel, "Minimum complexity sequential multihypothesis detection: Theory and reduced complexity decoding," Conf. Rec., IEEE Int. Symp. on Information Theory, ISIT 2001, Washington, D.C., June 2001.

- C. Köse, D. Goeckel, "Minimum complexity sequential multihypothesis detection: Weak sequential tests," Proc. IEEE Wireless Networking and Communications Conference, Chicago, IL, Sept. 2000.

- C. Köse, D. Goeckel, "On power adaptation in adaptive signaling systems", *IEEE Trans. on Commun.*, Vol.48, No. 11, pp. 1769-1773, 1998.

- C. Köse, B. Sankur, V. Stravoitov, "Generalized distance based matching of nonbinary images," Proc. IEEE Conference on Image Processing, Vol. 1, pp. 803-807, Chicago, IL, October 4-7, 1998.

Abstract of the Dissertation

Universal trellis codes and concatenated trellis-coded modulations for the compound linear vector Gaussian channel

by

Cenk Kose

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2004

Professor Richard D. Wesel, Chair

In broadcast scenarios or in the absence of accurate channel probability distribution information, code design for consistent channel-by-channel performance, rather than average performance over a channel distribution may be desirable. Root and Varaiya's compound channel coding theorem for linear Gaussian channels promises the existence of codes that operate reliably whenever the channel mutual information is above the transmitted rate. Such codes are called universal codes for their respective classes. Many communication channels of interest, such as multiple-antenna (space-time) channels, frequency-hopped communication channels and multicarrier communication channels can be modeled as a family of linear vector Gaussian channels. The practical design goal of universal codes is to guarantee consistent performance on every channel that supports the information transmission rate with a mutual information gap that is similar to the capacity gap of a well-designed AWGN-specific code on the AWGN channel.

The first part of this dissertation illustrates the design of universal trellis codes for the most comprehensive compound linear Gaussian vector channel as

a low-latency universal coding solution for multiple-antenna wireless links with the assumption of reliable channel estimation at the receiver. Universal trellis codes presented here deliver comparable or in some cases superior error-rate performance under quasistatic fading channels to trellis codes of similar complexity that are designed specifically for the quasistatic fading scenario, while providing consistently good channel-by-channel performance that may not be provided for codes designed for the average fading scenario.

The design criteria for universal trellis codes is then extended to the case of capacity-approaching serially-concatenated turbo codes. The second part of the dissertation investigates the design of universal serially-concatenated trellis codes (serial turbo codes) first for the periodic erasures channel with period two, and then for space-time channels with two transmitters.

Among other contributions of this dissertation is an analysis of the effect on universality of the use of matrix constellations on the vector Gaussian channel and the design of robust phase-noncoherent trellis-coded modulations for multicarrier communication channels with phase errors.

CHAPTER 1

Universal codes for the compound linear vector Gaussian channel

1.1 Introduction

Root and Varaiya's compound channel theorem [64] applied to the linear Gaussian vector channel

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \qquad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, N_0\mathbf{I}) \tag{1.1}$$

indicates that a *single* code can achieve reliable information transmission at a rate R bits per symbol for every channel $\mathbf{H}$ that induces at least R bits per symbol of mutual information (MI). The immediate implication of this result is that good error performance over one particular linear Gaussian channel does not have to come at the expense of significant performance degradation over others. At a given decoding complexity, codes with consistently good proximity to capacity over a class of channels will be referred to as *universal* codes in this thesis.

Universal codes are particularly interesting for multiple antenna (space-time) systems in richly scattering environments [19] because of the wide range of possible channels. However, recent research on space-time coded systems has exclusively focused on designing for the average performance over quasistatic Rayleigh fading channels ([26, 18, 51], [27], [61], [53], [60], [39]). In the light of the aforementioned compound channel coding theorem, universal codes will have good

performance on every channel with sufficient mutual information and thus under any quasistatic fading statistics. In contrast, there exist good codes for quasistatic Rayleigh channels that experience a wide variation in performance over matrix channels with the same mutual information.

Section 1.2 reviews the capacity expressions for the vector Gaussian channel (1.2). This section also proves that in the absence of channel information at the transmitter, when the compound channel includes all the channels with a given set of eigenvalues, the capacity achieving input distribution on the compound channel employs uniform transmit power across the transmit antennas.

Section 1.3 introduces the concept of excess mutual information as a measure of how close a code is operating to the channel capacity on a particular channel. This section then poses the universal code design problem over a compound channel as a minimax excess mutual information game.

For robust performance across the compound channel, a universal trellis code has good minimum distance over all linear Gaussian channels with similar mutual information. Section 1.4 formulates the smallest (i.e., worst-case) minimum distance of an error event over all channels with a given set of eigenvalues. The minimum distances of a code under unitary channels and under singular channels yield a lower bound on the worst-case minimum distance over all channels that support the transmission rate.

Section 1.5 presents encoder rate and complexity requirements on space-time trellis coded modulations for universal performance. A minimax excess-MI search for universal space-time trellis coded modulation over two transmit antennas is detailed. An example universal space-time trellis-coded modulation (ST-TCM) has better compound channel performance than ST-TCMs designed for quasistatic Rayleigh fading and has identical or better bit error rate performance

under quasistatic Rayleigh fading than these other ST-TCMs. As an alternative to ST-TCMs, coded performance of Alamouti's space-time block (STB) scheme over the compound channel is analyzed.

Raw frame error rates in quasistatic Rayleigh fading for universal codes are slightly worse than their counterparts specifically designed for average performance in Rayleigh fading. However, the favorable distribution of bit errors among errored frames allows the frame error rate of universal codes to improve significantly with the addition of a high-rate outer code.

The success of universal codes in quasistatic fading channels motivates Section 1.6, which shows that the determinant criterion used for quasistatic Rayleigh fading channels is a minimax worst-case distance criterion over a strict subset channels that support the rate of transmission. This subset does not include singular channels.

1.2 The compound linear Gaussian channel

Consider the $N_t \times N_r$ linear Gaussian channel with unit average power constraint per transmit antenna

$$\begin{aligned} \mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \qquad & \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}), \\ \mathbb{E}||\mathbf{x}||^2 \qquad & \leq N_t \end{aligned} \tag{1.2}$$

where $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the channel gain matrix, and $\mathbf{w} \in \mathbb{C}^{N_r \times 1}$ is a vector of additive Gaussian noise with unit variance per complex dimension. The noise components at different receivers are assumed to be uncorrelated.

Suppose that the channel matrix belongs to a collection $\mathcal{H} = \{\mathbf{H}_\alpha, \alpha \in \mathcal{I}\}$, where $\mathcal{I}$ is an arbitrary index set. Root and Varaiya's compound channel theory [64] defines the rate R bits/symbol to be achievable for the compound channel

$\mathcal{H}$ if there exists a sequence of $(2^{NR}, N)$ codes and a sequence of real numbers $\{\epsilon_N > 0\}$, such that for each N, the word error probability is bounded above by ϵ_N uniformly over all channels, and $\epsilon_N \to 0$ as $N \to \infty$.

The capacity of the compound channel, defined to be the the supremum of all achievable rates, is attained by a codebook $\{\mathbf{X}\}$ with a vector Gaussian distribution of mean zero and covariance matrix $\mathbf{S}$ such that $\text{trace}(\mathbf{S}) \leq N_t$ in accordance with the transmit power constraint.

Theorem 1 (Compound channel capacity, [64]) *The compound channel capacity is*

$$C(\mathcal{H}) = \sup_{\mathbf{S} \in \mathcal{S}} \inf_{\mathbf{H} \in \mathcal{H}} \log_2 \det \left(\mathbf{I} + \mathbf{H}\mathbf{S}\mathbf{H}^\dagger \right) \tag{1.3}$$

bits per symbol where $\mathcal{S}$ is the set of $N_t \times N_t$ complex valued Hermitian matrices such that trace$(\mathbf{S}) \leq N_t$.

From this point on we will assume $N_t = N_r = N$, to simplify the analysis. For $N_r \neq N_t$, the analysis will be the same with $N = \min(N_r, N_t)$, since the smallest $|N_r - N_t|$ eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$ will be zero.

As shown in [64], this capacity may be achieved without a-priori knowledge of the channel at the receiver. However, since this theorem is for fixed channels and blocklengths that are growing to infinity, pilot assisted channel estimation affects throughput negligibly. For simplicity, we assume the receiver knows the channel in the rest of this chapter.

The capacity achieving covariance matrix of the input distribution depends on the collection of channels. The following corollary shows that when all channels with a given set of singular values are included in $\mathcal{H}$, the capacity-achieving input distribution is uncorrelated and has uniform average power over transmit

antennas. We will write

$$\mathbf{H} \cong \boldsymbol{\lambda} = (\lambda_1, ..., \lambda_N) \tag{1.4}$$

if $\mathbf{HH}^\dagger$ has eigenvalues $\lambda_1, ..., \lambda_N$.

Corollary 1.1 *Let $\mathcal{H}(\boldsymbol{\lambda}) \subset \mathbb{C}^{N \times N}$ be the class of $N \times N$ channel matrices $\mathbf{H} \cong \boldsymbol{\lambda}$. The compound channel capacity of $\mathcal{H}(\boldsymbol{\lambda})$ is achieved by an uncorrelated Gaussian input distribution, i.e. $\mathbf{S} = \mathbf{I}$. As a result, the compound channel capacity is $\sum_{i=1}^{N} \log_2(1 + \lambda_i)$ bits/symbol.*

Proof 1 *Pick a codeword covariance matrix $\mathbf{S}$, let $\mathbf{S} = \mathbf{Q}\mathrm{diag}(\mathbf{s})\mathbf{Q}^\dagger$ be the symmetric eigendecomposition with $\mathbf{s} = (s_1, ..., s_N)$, $\sum_{i=1}^{N} s_i \leq N$. Let π be a permutation on $\{1, ..., N\}$. Let $\mathbf{s}_\pi = (s_{\pi(1)}, ..., s_{\pi(N)})$ denote the vector obtained by permuting the elements of $\mathbf{s}$ by π. Let $\boldsymbol{\Lambda} = \mathrm{diag}(\boldsymbol{\lambda})$, $\boldsymbol{\Lambda}_\pi = \mathrm{diag}(\boldsymbol{\lambda}_\pi)$, and $\mathbf{H}_\pi = \boldsymbol{\Lambda}_\pi^{1/2}\mathbf{Q}^\dagger$. Note that $\mathbf{H}_\pi \in \mathcal{H}(\boldsymbol{\lambda})$. Then*

$$\begin{aligned}
&\min_\pi \log_2 \det\left(\mathbf{I} + \mathbf{H}_\pi \mathbf{S} \mathbf{H}_\pi^\dagger\right) \\
&= \min_\pi \log_2 \det\left(\mathbf{I} + \boldsymbol{\Lambda}^{1/2}\mathrm{diag}(\mathbf{s}_{\pi^{-1}})\boldsymbol{\Lambda}^{1/2}\right) \\
&\leq \sum_\pi \frac{1}{N!} \log_2 \det\left(\mathbf{I} + \boldsymbol{\Lambda}^{1/2}\mathrm{diag}(\mathbf{s}_{\pi^{-1}})\boldsymbol{\Lambda}^{1/2}\right) \\
&\overset{(a)}{\leq} \log_2 \det\left(\mathbf{I} + \boldsymbol{\Lambda}^{1/2}\left(\frac{1}{N!}\sum_\pi \mathrm{diag}(\mathbf{s}_{\pi^{-1}})\right)\boldsymbol{\Lambda}^{1/2}\right) \\
&= \sum_{i=1}^{N} \log_2(1 + \lambda_i).
\end{aligned}$$

where inequality (a) follows from the fact that $\log_2 \det\left(\mathbf{I} + \boldsymbol{\Lambda}^{1/2}\mathrm{diag}(\mathbf{s})\boldsymbol{\Lambda}^{1/2}\right)$ is a concave function of $\mathbf{s}$. Thus, the worst-case mutual information of any admissible $\mathbf{S}$ is upper bounded by that achieved by $\mathbf{S} = \mathbf{I}$. ■

For $N = 2$ at MI = 2 bits/symbol (= 1 bit/symbol/transmit antenna) level,

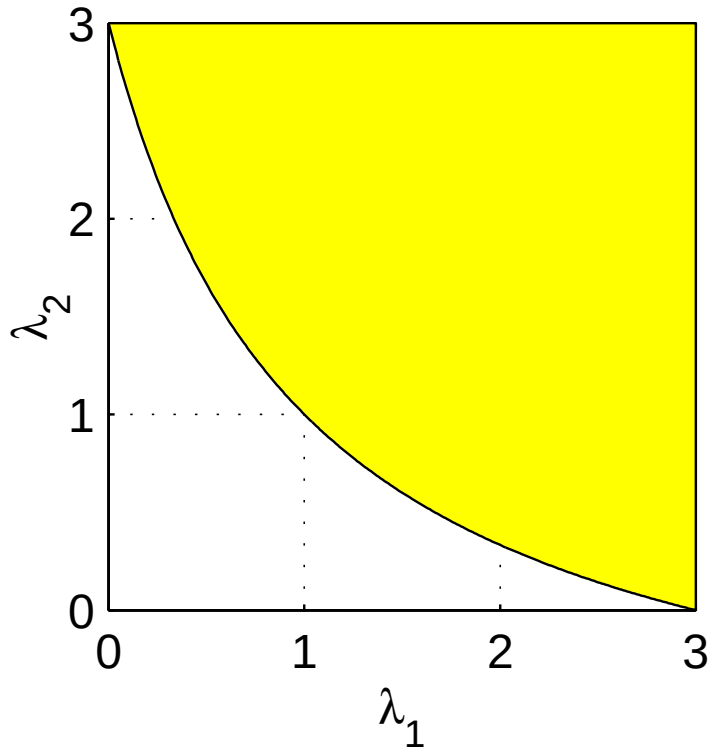

Figure 1.1: The equal-MI curve $\frac{1}{2}\log_2(1+\lambda_1)(1+\lambda_2) = R$ for $R = 1$ bit per symbol per antenna. The shaded region contains the pairs (λ_1, λ_2) such that MI ≥ 2 bits/symbol (1 bit/symbol/antenna).

Figure 1.1 shows the achievable region in the λ_1-λ_2 space with uniform power across antennas ($\mathbf{S} = \mathbf{I}$).

1.3 Excess mutual information criterion

With uniform average transmit signal energy $E_s = \mathbb{E}|x_i|^2$ across the antennas, the compound channel capacity theorem promises the existence of a code that can reliably transmit at R bits/symbol on every channel

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, N_0\mathbf{I}_N) \tag{1.5}$$

such that $\log_2 \det(\mathbf{I} + \frac{E_s}{N_0}\mathbf{H}) > R$. For a given rate and decoding complexity per information bit, a universal code delivers similar error performance over all channels with the same mutual information. Consider the set of channels that support a given decoding error probability. The level of universal performance over these channels is measured by the difference between the compound channel capacity and the information transmission rate. This difference is called the *excess mutual information.*

Consider a code that delivers R bits/symbol over N transmit antennas. Fix a target bit error rate BER_0. Associated with each channel matrix $\mathbf{H}$ is an average transmit energy requirement $E_s = E_s(\mathbf{H})$ to achieve BER_0. This transmit energy requirement translates to a mutual information figure

$$\mathrm{MI}\,(\mathbf{H}) = \log_2 \det\left(\mathbf{I} + \frac{E_s(\mathbf{H})}{N_0}\mathbf{H}\mathbf{H}^\dagger\right), \tag{1.6}$$

which is the maximum rate that can be reliably transmitted over the static channel $\mathbf{H}$ with uniform transmit power E_s over the antennas. The difference $\mathrm{MI}(\mathbf{H}) - R$ is the excess mutual information, which is a measure of how close the code is operating to channel capacity at $\mathrm{BER} = \mathrm{BER}_0$. In order to compare performance for various numbers of transmit antennas, the excess mutual information is normalized by the number of transmit antennas to obtain the excess mutual information per transmit antenna. The purpose of universal coding is to minimize the worst-case excess mutual information per transmit antenna

$$\sup_{\mathbf{H}\in\mathcal{H}} \frac{1}{N}\left(\mathrm{MI}\,(\mathbf{H}) - R\right) \tag{1.7}$$

at fixed target bit error rate (or frame error rate), latency and decoding complexity per bit.

Consider the standard additive white Gaussian noise (AWGN) channel over which complex valued symbols are transmitted with average transmit energy of

E_s:

$$y = \sqrt{E_s}x + w, \quad w \sim \mathcal{N}(0, N_0)$$
$$\mathbb{E}|x|^2 = 1$$

For $R = 1$ bit/symbol transmission, the best free-distance rate-1/2 64-state convolutional encoder driving a Gray-labeled QPSK constellation requires a signal-to-noise-ratio (SNR) of $E_s/N_0 = 4.3$ dB to achieve a bit error rate (BER) of 10^{-5}. At this SNR, the capacity of the AWGN channel is

$$\log_2\left(1 + \frac{E_s}{N_0}\right) = 1.88 \text{ bits/symbol} \tag{1.8}$$

therefore the code requires an *excess mutual information* (MI) of $1.88 - R = 0.88$ bits at this BER. Table 1.1 lists the excess MI performance under AWGN of trellis codes formed by the best free-distance rate-1/2 encoders driving Gray-labeled QPSK constellations.

Now consider the same code transmitted over two transmit antennas by means of a serial-to-parallel converter, delivering 2 bits/symbol (1 bit/symbol/antenna). Write

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \sqrt{E_s}\mathbf{H} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \tag{1.9}$$

where x_1 and x_2 are the odd and even-indexed symbols, respectively, from the original single dimensional trellis code, and the average transmit energy per antenna, E_s, is explicit. The AWGN scenario discussed above is $\mathbf{H} = \mathbf{I}$.

When

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \tag{1.10}$$

Table 1.1: Performance of best free-distance rate-1/2 convolutional codes driving Gray-labeled QPSK constellations. The first column is the number of encoder states, S. The second column shows the minimum distance and the bit multiplicity of the minimum-distance events. For each code the SNR value, SNR_0, required to achieve BER $= 1 \times 10^{-5}$ is displayed in the third column. The fourth column shows the corresponding excess-MI values.

S	$d_{\min}$, N_b	SNR_0	ΔMI
4	3.16, 8	5.85 dB	1.26
8	3.46, 8	5.4 dB	1.16
16	3.74, 2	4.9 dB	1.03
32	4, 16	4.65 dB	0.97
64	4.21, 3	4.3 dB	0.88

the rate-1/2 code with QPSK can at best deliver uncoded performance (BER=10^{-5} at SNR=12.5 dB, or equivalently, excess MI = 1.12 bits/antenna) since the code rate effectively raises to 1/1. In fact, the 64-state code in Table 1.1 completely fails under the channel (1.10) with BER $> 10^{-1}$ at SNR=12.5 dB.

However, as shown in [46], there exist 64 state trellis codes that require less than 0.88 bits of excess mutual information per antenna at BER $= 1 \times 10^{-5}$ on both $\mathbf{H} = \mathbf{I}$ and the singular $\mathbf{H}$ of (1.10). In order to deliver coded performance over all 2×2 matrix channels (even singular channels) with sufficient MI, the binary rate of the encoder should be less than 1/2. However, per-symbol rates may be maintained by using constellations larger than QPSK. The maximum rate possible for an encoder with universal performance is discussed in Section 1.5. This example shows that a typical good AWGN trellis code does not exhibit universal behavior. As another example, Ungerboeck type trellis-coded modulations

[22] fail under erasure scenarios such as (1.10) due to the presence of uncoded bits.

A universal trellis code should have consistently good minimum distance under linear transformations (channels) having the same nonzero mutual information. The following section characterizes the variation of minimum distance under transformations by channels with a fixed amount of MI.

1.4 Minimum distance over the compound channel

Consider a code $\mathcal{C}$ for the vector channel (1.2). A codeword $\mathbf{X} \in \mathcal{C}$ is represented here as a matrix of N rows and l columns where each column $\mathbf{x}_i \in \mathbb{C}^{N\times 1}$ is a codeword symbol at time i:

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_l]\,. \tag{1.11}$$

Let $\mathbf{X}$ and $\hat{\mathbf{X}}$ be two different codewords. Under maximum-likelihood decoding, the probability that the decoder mistakes $\hat{\mathbf{X}}$ for $\mathbf{X}$ conditioned on the perfect knowledge of the channel matrix $\mathbf{H}$ at the receiver is given by

$$P\left(\mathbf{X} \to \hat{\mathbf{X}}|\mathbf{H}\right) = Q\left(d\left(\mathbf{X} \to \hat{\mathbf{X}}|\mathbf{H}\right)/\sqrt{2N_0}\right) \tag{1.12}$$

where $d(\mathbf{X} \to \hat{\mathbf{X}}|\mathbf{H}) = \|\mathbf{H}(\mathbf{X}-\hat{\mathbf{X}})\|_2$ is the Euclidean distance of the event under $\mathbf{H}$, and $Q(\cdot)$ is the standard Gaussian tail integral.

Definition 1 *For a given* $\mathbf{H}$*, the smallest of pairwise codeword distances is called the* $d_{\min}(\mathbf{H})$ *of the code:*

$$d_{\min}(\mathbf{H}) = \min_{\mathbf{X}\neq\hat{\mathbf{X}}} d\left(\mathbf{X} \to \hat{\mathbf{X}}|\mathbf{H}\right). \tag{1.13}$$

The $d_{\min}(\mathbf{H})$ of a code is closely related to the eigenvalues of $\mathbf{H}\mathbf{H}^{\dagger}$. Throughout this section we will use the following ordering for the eigenvalues:

$$\mathbf{H} \cong \boldsymbol{\lambda}, \qquad \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_N \tag{1.14}$$

$$\mathbf{X} - \hat{\mathbf{X}} \cong \boldsymbol{\zeta}, \qquad \zeta_1 \leq \zeta_2 \leq ... \leq \zeta_N. \tag{1.15}$$

Lemma 1 (Universal codes have to be full rank [9]) *Let ζ_i, $i = 1, ..., N$ be the eigenvalues of $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^{\dagger}$ such that $\zeta_i \leq \zeta_{i+1}$. If $\zeta_1 = 0$, then for any $m > 0$, there exists $\mathbf{H}_0$ with $\lambda_1 = 0$, $\mathrm{M}I(\mathbf{H}) = m$ such that $d(\mathbf{X} \to \hat{\mathbf{X}}|\mathbf{H}_0) = 0$.*

Consider a system with two transmit antennas. One example of $\zeta_1 = 0$ arises under spatial repetition, i.e. when the two transmit antennas simply repeat the same single-dimensional coded signal, in which case channels of the form

$$\begin{bmatrix} h_1 & -h_1 \\ h_2 & -h_2 \end{bmatrix} \tag{1.16}$$

will cause destructive interference of the transmitted signals resulting in zero minimum distance. Simple spatial repetition cannot achieve robust performance over the matrix compound channel. The full-rank criterion also appears in code design optimizing average performance in quasistatic Rayleigh fading channels ([26],[61]).

For an $N \times N$ system, Theorem 2 below formulates the smallest $d_{\min}(\mathbf{H})$ of a given code over all channels $\mathbf{H} \cong \boldsymbol{\lambda}$ and finds those least-favorable channels under which $d_{\min}(\mathbf{H})$ is the smallest. This result can then be used to characterize the minimum distance over all channels $\mathbf{H} \cong \boldsymbol{\lambda}$ such that $\mathrm{MI} = \sum_{i=1}^{N} \log_2(1 + \lambda_i)$=constant.

Theorem 2 (Least favorable channels) *For a given codeword pair $(\mathbf{X}, \hat{\mathbf{X}})$, $\mathbf{X} \neq \hat{\mathbf{X}}$, let*

$$\mathbf{V} \cdot \mathrm{diag}(\boldsymbol{\zeta}) \cdot \mathbf{V}^{\dagger} \tag{1.17}$$

be the eigendecomposition of $(\mathbf{X}-\hat{\mathbf{X}})(\mathbf{X}-\hat{\mathbf{X}})^{\dagger}$ *such that* $\zeta_1 \leq \zeta_2 \leq ... \leq \zeta_N$. *Then,*

$$\min_{\mathbf{H}\cong\boldsymbol{\lambda}} d^2\left(\mathbf{X}\to\hat{\mathbf{X}}|\mathbf{H}\right) = \tilde{\boldsymbol{\lambda}}^T\boldsymbol{\zeta} = \sum_{i=1}^{n}\lambda_{N+1-i}\zeta_i \tag{1.18}$$

and the minimum is attained by any of the least favorable channels

$$\left\{\mathbf{H}=\mathbf{U}\tilde{\boldsymbol{\Lambda}}^{1/2}\mathbf{V}^{\dagger} : \mathbf{U}\ \textit{unitary}\right\} \tag{1.19}$$

where $\tilde{\boldsymbol{\lambda}} = (\lambda_N, ..., \lambda_1)$, *and* $\tilde{\boldsymbol{\Lambda}} = \mathrm{diag}(\tilde{\boldsymbol{\lambda}})$.

Proof 2 *Any* $\mathbf{H}\cong\boldsymbol{\lambda}$ *can be written as* $\mathbf{H}=\mathbf{U}\tilde{\boldsymbol{\Lambda}}^{1/2}\mathbf{R}^{\dagger}$ *where* $\mathbf{U}$, $\mathbf{R}$ *are unitary matrices over* $\mathbb{C}$. *Let* $\mathbf{Q}=\mathbf{R}^{\dagger}\mathbf{V}$. *Then* $\mathbf{Q}$ *is unitary. Since* $\mathbf{U}$ *does not effect the pairwise distance*

$$\begin{aligned}
&\min_{\mathbf{H}=\tilde{\boldsymbol{\Lambda}}^{1/2}\mathbf{R}} d^2\left(\mathbf{X}\to\hat{\mathbf{X}}|\mathbf{H}\right) &(1.20)\\
&= \min_{\mathbf{Q}\ unitary} tr\left(\tilde{\boldsymbol{\Lambda}}^{1/2}\mathbf{Q}\cdot\mathrm{diag}(\boldsymbol{\zeta})\cdot\mathbf{Q}^{\dagger}\tilde{\boldsymbol{\Lambda}}^{1/2}\right)\\
&= \min_{\mathbf{Q}\ unitary}\sum_{i=1}^{N}\sum_{j=1}^{N}|Q_{ij}|^2\lambda_{N+1-i}\zeta_j\\
&= \min\left\{\sum_{i=1}^{N}\lambda_i\zeta_{\pi(i)} :\ \pi\ \textit{permutation on}\ (1,...,n)\right\}\\
&= \sum_{i=1}^{N}\lambda_{N+1-i}\zeta_i &(1.21)
\end{aligned}$$

since we adopted $\zeta_i \leq \zeta_{i+1}$, $\lambda_i \leq \lambda_{i+1}$. *Note that (1.21) holds when* $\mathbf{Q}$ *is the identity matrix. Therefore, the least favorable channel has* $\mathbf{R}=\mathbf{V}$, *and the statement follows.* ■

For $N=2$, two cases for the eigenvalue pair (λ_1, λ_2) are of special interest: $\lambda_1=\lambda_2$, and $\lambda_1=0$. At MI = 2 bits/symbol level, these cases are $\lambda_1=\lambda_2=1$, and $\lambda_1=0$, $\lambda_2=3$. Note that when $\lambda_1=\lambda_2=1$, the channel matrix $\mathbf{H}$ is

necessarily unitary, and the transmitted codeword matrix can be recovered with no interference between the two antennas. The error performance on any unitary channel is identical to the error performance under the parallel AWGN channel $\mathbf{H} = \mathbf{I}$. Theorem 3(a) shows that the minimum distance of a code over these two cases at the same MI level yields a lower bound on the minimum distance over the compound 2×2 channel. Part (b) of the theorem extends this result to general N.

Theorem 3 (Unitary and rank-deficient channels) *For the $N \times N$ compound channel at* $\mathrm{MI} = R$ *bits/symbol level, let $d_{\boldsymbol{\lambda}}$ be the smallest $d_{min}(\mathbf{H})$ of a code over all channels $\mathbf{H} \cong \boldsymbol{\lambda}$.*

(a) *For $N = 2$, the minimum distance $d_{\boldsymbol{\lambda}} = d_{(\lambda_1,\lambda_2)}$, under the least favorable channel is lower bounded as*

$$\begin{aligned} &\min_{(\lambda_1,\lambda_2):\mathrm{MI}=R} d_{(\lambda_1,\lambda_2)} \\ &\quad \geq f_R \times \min\left\{d_{(2^{R/2}-1,2^{R/2}-1)}, d_{(0,2^R-1)}\right\}, \end{aligned} \tag{1.22}$$

where

$$f_R = \sqrt{\frac{2^{1+(R/2)} + 2^{R/4} + 1}{2^{3R/4} + 2^{R/2} + 2^{R/4} + 1}}. \tag{1.23}$$

f_R is decreasing for increasing R. The maximum *value of $d_{(\lambda_1,\lambda_2)}$ over the compound 2×2 channel is attained on either a rank-deficient channel or a unitary channel:*

$$\begin{aligned} &\max_{(\lambda_1,\lambda_2):\mathrm{MI}=R} d_{(\lambda_1,\lambda_2)} \\ &\quad = \max\left\{d_{(2^{R/2}-1,2^{R/2}-1)}, d_{(0,2^R-1)}\right\}. \end{aligned} \tag{1.24}$$

(b) *For $N > 2$,*

$$\min_{\boldsymbol{\lambda}:\mathrm{MI}(\boldsymbol{\lambda})=R} d_{\boldsymbol{\lambda}}$$
$$\geq f_R \min\{d_{(0,\ldots,0,2^{R/n}-1,2^{R/n}-1)}, d_{(0,\ldots,0,0,2^{2R/n}-1)}\}.$$

Proof 3 *Part (a): At R bits/symbol MI, we have $(\log(1+\lambda_1)+\log(1+\lambda_2)) = R$. For a given pair (λ_1, λ_2) of eigenvalues, let $(\mathbf{X} \to \hat{\mathbf{X}})$ be an error event that achieves the smallest $d_{min}(\mathbf{H})$ of the code over all channels $\mathbf{H} \cong (\lambda_1, \lambda_2)$. Let $(\mathbf{X} - \hat{\mathbf{X}}) \cong (\zeta_1, \zeta_2)$ with $\zeta_1 \leq \zeta_2$. Then*

$$d^2_{(\lambda_1,\lambda_2)} = \lambda_1\zeta_2 + \zeta_1\lambda_2 \tag{1.25}$$

Let $\lambda_1 = \lambda \in [0, 2^{R/2}-1]$, so that $\lambda_2 = (2^R - 1 - \lambda)/(1+\lambda)$. Using (1.25) and $\zeta_1 \leq \zeta_2$,

$$d^2_{(2^{R/2}-1,2^{R/2}-1)} \leq (2^{R/2}-1)(\zeta_1+\zeta_2) \tag{1.26}$$
$$d^2_{(0,2^R-1)} \leq (2^R-1)\zeta_1. \tag{1.27}$$

These inequalities arise because any particular (ζ_1, ζ_2) pair upper bounds the min-

imum d_{λ_1,λ_2} over all (ζ_1,ζ_2) pairs. Therefore

$$\begin{aligned}
d^2_{(\lambda_1,\lambda_2)} &= \lambda\zeta_2 + \left(\frac{2^R-1-\lambda}{1+\lambda}\right)\zeta_1 \\
&\overset{(a)}{\geq} \lambda\left(\frac{d^2_{(2^{R/2}-1,2^{R/2}-1)}}{2^{R/2}-1} - \zeta_1\right) + \frac{2^R-1-\lambda}{1+\lambda}\zeta_1 \\
&= \frac{\lambda}{2^{R/2}-1} d^2_{(2^{R/2}-1,2^{R/2}-1)} \\
&\quad + \left(\frac{2^R-1-\lambda}{1+\lambda} - \lambda\right)\zeta_1 \\
&\overset{(b)}{\geq} \frac{\lambda}{2^{R/2}-1} d^2_{(2^{R/2}-1,2^{R/2}-1)} \\
&\quad + \left(\frac{2^R-1-\lambda}{1+\lambda} - \lambda\right)\frac{d^2_{(0,2^R-1)}}{2^R-1} \\
&\overset{(c)}{\geq} \left(\frac{\lambda}{2^{R/2}-1} + \frac{1}{2^R-1}\left(\frac{2^R-1-\lambda}{1+\lambda} - \lambda\right)\right) \\
&\quad \times \min\left\{d^2_{(2^{R/2}-1,2^{R/2}-1)}, d^2_{(0,2^R-1)}\right\} \\
&= r_R(\lambda)\min\{d^2_{(2^{R/2}-1,2^{R/2}-1)}, d^2_{(0,2^R-1)}\}. \qquad (1.28)
\end{aligned}$$

Inequalities (a) and (b) follow from (1.26) and (1.27) respectively, and (c) follows when the larger of the two distances is replaced by the smaller. The function $r_R(\lambda)$ attains its minimum on $\lambda \in (0,1)$ at $\lambda_* = 2^{R/4}-1$, with $\sqrt{r_R(\lambda_*)} = f_R$ given in (1.23). Figure 1.2 shows $\sqrt{r_R}$ as a function of λ for several R values. For $R \leq 3$ bits/symbol, $f_R \geq 0.9$.

To prove the second statement, we use the fact that the function

$$g(\lambda) = d^2_{(\lambda,\frac{2^R-1-\lambda}{1+\lambda})}, \quad \lambda \in [0, 2^{R/2}-1] \qquad (1.29)$$

is continuous and convex on $(0, 2^{R/2}-1)$ therefore it has a maximum on $[0, 2^{R/2}-1]$ at either $\lambda = 0$ or $\lambda = 2^{R/2}-1$.

Part (b): Consider $N = 3$. Let $\mathbf{E} = \hat{\mathbf{X}} - \mathbf{X}$ be an error event of the code, let $\zeta_1 \leq \zeta_2 \leq \zeta_3$, be the eigenvalues of $\mathbf{E}\mathbf{E}^\dagger$. For $\lambda_1 \leq \lambda_2 \leq \lambda_3$, let $d^2_{(\lambda_1,\lambda_2,\lambda_3)}(\mathbf{E})$

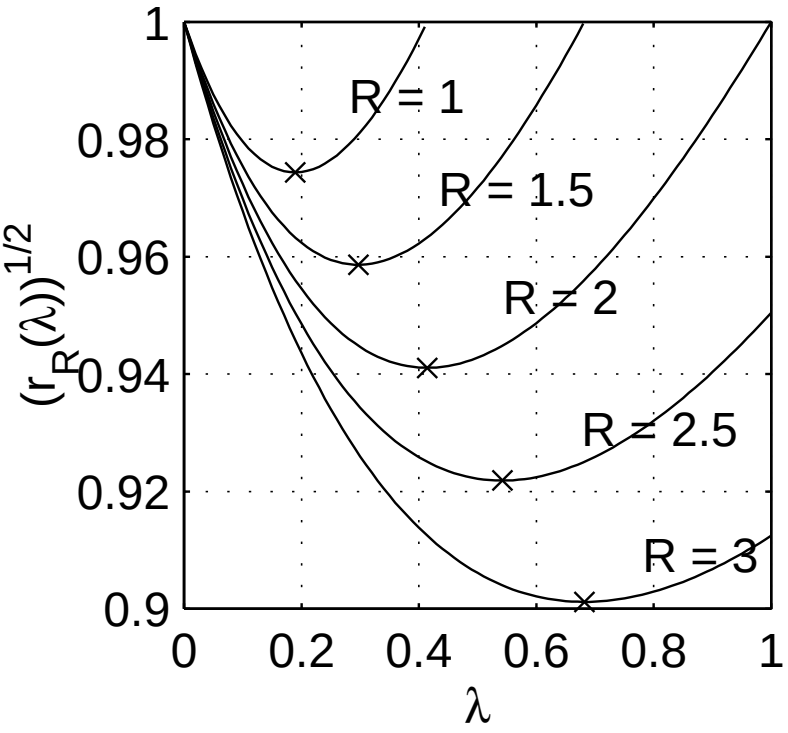

Figure 1.2: The function $\sqrt{r_R}$ for $R = 1$, 1.5, 2, 2.5, 3.

be the minimum squared Euclidean distance of this event under channels $\mathbf{H} \cong (\lambda_1, \lambda_2, \lambda_3)$. *By Theorem 2,*

$$d^2_{(\lambda_1,\lambda_2,\lambda_3)}(\mathbf{E}) = \lambda_1\zeta_3 + \lambda_2\zeta_2 + \lambda_3\zeta_1.$$

Fix $\lambda_1 \in [0, 2^{R/3}-1]$ *and minimize* $d^2_{(\lambda_1,\lambda_2,\lambda_3)}$ *under the constraint* $\mathrm{MI}(\lambda_1,\lambda_2,\lambda_3) = R$ *or* $\mathrm{MI}(\lambda_2,\lambda_3) = R - \log_2(1+\lambda_1)$*:*

$$\begin{aligned}
&\min_{\lambda_2,\lambda_3:\mathrm{MI}(\lambda_2,\lambda_3)=R-\log_2(1+\lambda_1)} d^2_{(\lambda_1,\lambda_2,\lambda_3)}(\mathbf{E}) \\
&= \lambda_1\zeta_3 + \min_{\lambda_2,\lambda_3:\mathrm{MI}(\lambda_2,\lambda_3)=R-\log_2(1+\lambda_1)} \lambda_2\zeta_2 + \lambda_3\zeta_1 \\
&\overset{(i)}{\geq} \lambda_1\zeta_3 \\
&\quad +(f_{R'})^2 \min\left\{d^2_{(0,2^{R'/2}-1,2^{R'/2}-1)}, d^2_{(0,0,2^{R'}-1)}\right\}
\end{aligned}$$

where $R' = R = \log_2(1+\lambda_1)$. *Inequality (i) follows from the proof of part (a) by noting that under the constraint* $\mathrm{MI}(\lambda_2,\lambda_3) = R$,

$$d^2_{(0,2^{R'/2}-1,2^{R'/2}-1)} \leq (2^{R'/2}-1)(\zeta_1+\zeta_2) \tag{1.30}$$

$$d^2_{(0,0,2^{R'}-1)} \leq (2^{R'}-1)\zeta_1. \tag{1.31}$$

Now, $R' \leq R$ *implies* $f_{R'} \geq f_R$ *and* $R' \geq 2R/3$ *implies*

$$d^2_{(0,2^{R'/2}-1,2^{R'/2}-1)} \geq d^2_{(0,2^{R/3}-1,2^{R/3}-1)} \tag{1.32}$$

$$d^2_{(0,0,2^{R'}-1)} \geq d^2_{(0,0,2^{2R/3}-1)} \tag{1.33}$$

therefore,

$$\begin{aligned}
&\min_{\lambda_2,\lambda_3:\mathrm{MI}(\lambda_2,\lambda_3)=R-\log_2(1+\lambda_1)} d_{(\lambda_1,\lambda_2,\lambda_3)}(\mathbf{E}) \\
&\quad \geq f_R \min\left\{d_{(0,2^{R/3}-1,2^{R/3}-1)}, d_{(0,0,2^{2R/3}-1)}\right\}
\end{aligned}$$

which leads to

$$\begin{aligned}
&\min_{\lambda_1,\lambda_2,\lambda_3:\mathrm{MI}(\lambda_1,\lambda_2,\lambda_3)=R} d_{(\lambda_1,\lambda_2,\lambda_3)} \\
&\quad \geq f_R \min\left\{d_{(0,2^{R/3}-1,2^{R/3}-1)}, d_{(0,0,2^{2R/3}-1)}\right\}.
\end{aligned}$$

Note that on channels with $\lambda_1 = 0$, the worst-case minimum distance is already upper bounded by $d^2_{(0,2^{R/2}-1,2^{R/2}-1)}$, similarly, on channels with $\lambda_1 = 0$, $\lambda_2 = 0$, an upper bound to the worst-case minimum distance is $d^2_{(0,0,2^R-1)}$.

The case for general N is proved by induction. To illustrate the induction step, take $N = 4$. For an error event $\mathbf{E} \cong (\zeta_1, \zeta_2, \zeta_3, \zeta_4)$, we have, by using the result for $N = 3$,

$$\begin{aligned} &\min_{\lambda_2,\lambda_3,\lambda_4:\mathrm{MI}(\lambda_2,\lambda_3,\lambda_4)=R'} d^2_{(\lambda_1,\lambda_2,\lambda_3,\lambda_4)}(\mathbf{E}) \\ &\quad \geq \quad f_{R'} \min\left\{ d_{(0,0,2^{R'/3}-1,2^{R'/3}-1)}, d_{(0,0,0,2^{2R'/3}-1)} \right\} \end{aligned}$$

where $R' = R - \log_2(1+\lambda_1) \in [3R/4, R]$. Therefore,

$$\begin{aligned} &\min_{\lambda_2,\lambda_3,\lambda_4:\mathrm{MI}(\lambda_2,\lambda_3,\lambda_4)=R'} d^2_{(\lambda_1,\lambda_2,\lambda_3,\lambda_4)}(\mathbf{E}) \\ &\quad \geq \quad f_{R} \min\left\{ d_{(0,0,2^{R/4}-1,2^{R/4}-1)}, d_{(0,0,0,2^{2R/4}-1)} \right\} \end{aligned}$$

and the theorem follows. ■

1.5 Minimax search for universal space-time trellis codes

This section describes the tools used to conduct an efficient search for universal codes over two transmit antennas. Section 1.5.1 discusses the binary encoder rate and complexity requirements for ST-TCMs to maintain coded performance across a family of matrix channels. Section 1.5.2 details a minimax excess-MI code search for one bit/symbol/antenna over two transmit antennas and describes the performance of the resulting codes over the compound 2×2 channel. Part 1.5.3 compares the performance of the universal codes with the performance of existing ST-TCMs as well as with Alamouti's orthogonal STB scheme in conjunction with a standard trellis code. Finally, Part 1.5.4 discusses the case for more than two transmit antennas.

1.5.1 Requirements for universal performance

Consider k bits/symbol transmission over N antennas using $N \times 2^m -$ point constellations. A consequence of Lemma 1 is that for robust performance over the compound matrix channel, m should be large enough to maintain redundancy under singular channels, when one or more eigenmodes of the channel is zero. In fact m should be larger than k to maintain an effective code rate strictly less than one under every channel $\mathbf{H}$ with rank 1 (when only one eigenvalue is nonzero). Equivalently, a given $N \times 2^m -$ PSK/QAM scheme can deliver at best uncoded performance over channels $\mathbf{H}$ with rank less than or equal to $\frac{k}{m}$.

Another consequence of Lemma 1 is a requirement on effective code length (ECL). The ECL of a code is defined ([49], [14]) to be minimum symbol-wise Hamming distance of the code, i.e., the smallest of the number of symbols by which any two codewords differ:

$$\text{ECL} = \min_{\mathbf{X} \neq \hat{\mathbf{X}}} | \{i : \mathbf{x}_i \neq \hat{\mathbf{x}}_i\} | \tag{1.34}$$

For trellis codes, effective code length is upper bounded by the the duration of the shortest error event, in symbols. If the length of the error event $\mathbf{X} \to \hat{\mathbf{X}}$ is l symbols, then $\text{rank}(\mathbf{X} - \hat{\mathbf{X}}) \leq \min(N, l)$. If $l < N$, then per Lemma 1, there exists a singular $N \times N$ channel with a prescribed MI such that $d(\mathbf{X} \to \hat{\mathbf{X}} | \mathbf{H}) = 0$. Therefore, to retain nonzero distance over the compound $N \times N$ matrix channel, the code should have ECL $\geq N$. Theorem 4, below, relaxes these constraints of rate and ECL by forcing a minimum rank requirement on the compound channel.

Theorem 4 *For $1 \leq r \leq N$, let $\mathcal{H}^r \subset \mathbb{C}^{N \times N}$ denote the the collection of $N \times N$ complex valued matrices $\mathbf{H}$ such that rank$(\mathbf{H}) \geq r$. A k-bits/symbol 2^ν-state $N \times 2^m -$ QAM/PSK universal trellis code for $\mathcal{H}^r$ satisfies*

$$\frac{k}{m} < r, \quad \left\lfloor \frac{\nu}{k} \right\rfloor \geq N + r. \tag{1.35}$$

Proof 4 *That $k/m < r$ is proved above. The effective code length of a k-bits/symbol trellis code with 2^ν states is bounded [66] as*

$$ECL \leq \left\lfloor \frac{\nu}{k} \right\rfloor + 1. \tag{1.36}$$

where $\lfloor z \rfloor$ is the largest integer not greater than z. For a trellis code with a particular ECL, there is at least one error event with at most ECL nonzero columns. Such an error event has at most ECL nonzero ζ's. If $ECL \leq N - r$ and channels may have $N - r$ zero eigenvalues, then we can find a channel $\mathbf{H}$ with $N - r$ zero eigenvalues that forces $d_{min}(\mathbf{H}) = 0$. ■

1.5.2 Code search for 1 bit/symbol/antenna transmission over two transmit antennas

Theorem 2 showed that a given error event $\mathbf{X} \to \hat{\mathbf{X}}$ attains its minimum distance over the compound 2×2 channel when the channel eigenvector corresponding to the larger eigenvalue (λ_2) of $\mathbf{H}\mathbf{H}^\dagger$ aligns with the eigenvector corresponding to the smaller eigenvalue (ζ_1) of $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^\dagger$. If the distance corresponding to the least-favorable channel can be evaluated for all the error events, then it is possible to conduct a search for the codes which have the highest worst-case pairwise Euclidean distance over all channels inducing a minimum mutual information. In general, forward trellis search with Viterbi elimination on the product-state trellis based on the squared distance $\lambda_1\zeta_2 + \zeta_1\lambda_2$ will not work since this distance measure is not additive over consecutive trellis branches.

However, for a fixed channel, the minimum-distance error event can be found by a forward trellis search. As a result of the continuity of $\lambda_1\zeta_2+\zeta_1\lambda_2$, a sufficiently fine sampling of the continuum of 2×2 channels will yield the worst-case minimum

distance event with the desired accuracy. For the worst-case minimum distance event at a given MI level, it is sufficient to consider 2×2 channels of the form

$$\mathbf{H} = \sqrt{\lambda_1} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\kappa^{-1}} \end{bmatrix} \begin{bmatrix} \cos(\phi) & \sin(\phi)e^{j\theta} \\ -\sin(\phi) & \cos(\phi)e^{j\theta} \end{bmatrix} \tag{1.37}$$

with $\log_2(1+\lambda_2)(1+\kappa\lambda_2)$=MI, $\phi \in [0, 2\pi)$, $\theta \in [0, 2\pi)$. The parameter $\kappa = \lambda_1/\lambda_2$ is the eigenvalue skew of matrix $\mathbf{HH}^\dagger$.

Theorem 3 proves that at fixed MI, good minimum distances on singular channels and on unitary channels guarantee good minimum distance under any channel with the same MI. For improved accuracy in guaranteeing bit error rate we have searched for codes with the following sampling of the channel space

$$\Delta\kappa = 0.5, \Delta\phi = \pi/4, \Delta\theta = \pi/8, \tag{1.38}$$

for 2×8-PSK transmission, which provides accuracy in the worst-case squared distance of more than two decimal digits. This sampling results in more than three hundred different channels at the same MI level, and the complexity of a simplistic exhaustive code search over these channels is prohibitive even for modest number of states. In order to alleviate this complexity burden, the brute-force search is pruned via the following observations.

1. For any channel instance, the distance of an error event is an upper bound on the worst-case distance of the code over all channels with the same MI. Therefore, a candidate universal code is discarded as soon as *an* error event for *any* channel sample produces a distance less than the highest worst-case distance for the best code found so far. This enables the search to accelerate as it progresses, since the rejection threshold of codes will increase with time.

2. An S-state code has a product-state trellis with S^2 states. A reduced-state description of the product-state trellis can be obtained by identifying a simpler equivalent state diagram ([30], [10]).

3. Codes with immediately apparent rank deficiency (for example, having the same generators for each antenna) are excluded from the search. Lemma 1 showed that universal codes have to be full rank. Similarly, catastrophic codes are eliminated.

4. Symmetry across antennas and time invariance of the generator functions is exploited to further reduce the search.

5. The sensitivity of the distance of an error event to the parameters ϕ, θ, κ is not uniform through the space of channels. For example, for channels with $\lambda_1 = \lambda_2$, the distance of an error event is independent of θ and ϕ. In fact,

$$\left| \frac{\mathrm{d}}{\mathrm{d}\theta} \left\| \mathbf{H}(\lambda_1, \lambda_2, \theta, \phi)(\mathbf{X} - \hat{\mathbf{X}}) \right\|^2 \right| \propto \lambda_2 - \lambda_1, \tag{1.39}$$

therefore a less dense sampling of the θ dimension could be sufficient for channels with $\lambda_1 \approx \lambda_2$.

In this fashion, the total number of channels to be considered can be reduced without sacrificing accuracy of the search.

A more accurate description of the error performance of a trellis code involves the multiplicities of the minimum-distance events. In particular, the bit error rate (BER) of a trellis code can be approximated using the truncated union bound (TUB) which takes into account the information bit errors caused by minimum-distance events as well as their distance. The truncated union bound yields an estimate (but not a bound) on the BER for high SNRs as

$$\mathrm{BER} \approx N_b \, Q\left(d_{\min}(\mathbf{H}) / \sqrt{2N_0} \right) \tag{1.40}$$

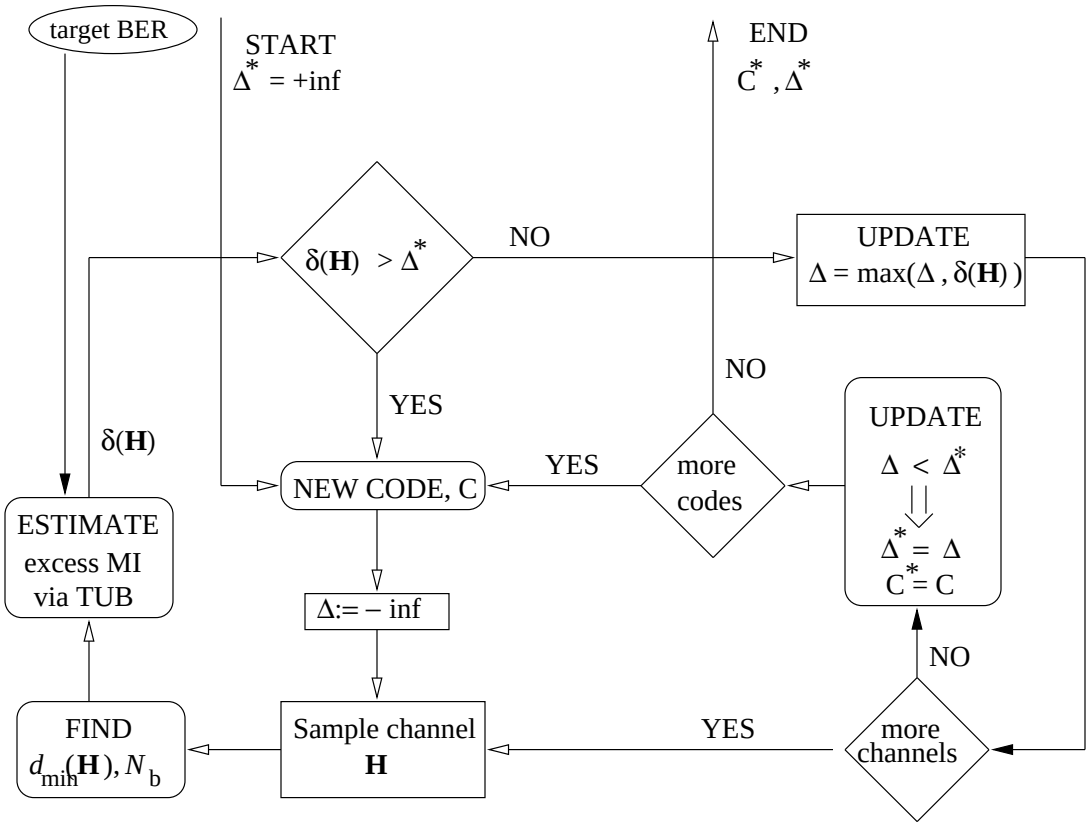

Figure 1.3: The minimax-excess MI search algorithm. At fixed BER, $\delta(\mathbf{H})$ is the TUB based estimate of the excess MI required to achieve BER on $\mathbf{H}$, Δ is the largest such excess MI estimate over all channels tested for the current code so far, and Δ^* is the smallest of Δ over all codes tested. The minimax algorithm stops testing a code as soon as $\delta(\mathbf{H}) > \Delta^*$.

where N_b is the total number of bit errors that are caused by minimum-distance events, averaged over all states. Note that in (1.40), the signal energy E_s per transmit antenna is absorbed in the channel matrix.

The truncated union bound approximation yields an estimate of the transmit SNR required to achieve a certain BER, say BER=10^{-5}, on any particular channel. The SNR figure is then used to compute the excess MI required on that particular channel, as discussed in Section III. Each channel in the sampled family of compound matrix channels (1.38) thus has an excess MI figure associated with it. The maximum of these excess MI figures gives the worst-case performance of the code over the family. Figure 1.3 illustrates the minimax excess-MI code-search algorithm based on TUB estimates of the bit error rate.

The excess MI computation based on the TUB is typically optimistic since error events that do not have minimum distance are ignored by the TUB. However, the TUB can be made very accurate by including a few more distances. Exact excess MI figures are obtained through simulation.

A linear feedforward encoder with $m-1$ memory elements which takes two input bits per cycle and generates six output bits (for two 8PSK symbols) per cycle can be compactly represented as

$$y_b^{(j)}(t) = \sum_{d=0}^{m} u(2t-d)\, g_b^{(j)}(d), \qquad j = 1,2, \;\; b = 0,1,2,$$

where two information bits, $u(2t)$, $u(2t-1)$, are input to the encoder at time t, $y_b^{(j)}$ is the bth bit of the 8PSK symbol from the jth transmit antenna and $g_b^{(j)}(d)$ are the corresponding binary generators. The least significant bit has index $b=0$. The number of states is $S = 2^{m-1}$. Figure 1.4 illustrates this encoder.

Table 1.2 lists 4-, 8-, 16-, 32- and 64-state linear encoders for $2 \times$ 8PSK transmission, found by exhaustive search to minimize the maximum TUB-estimated

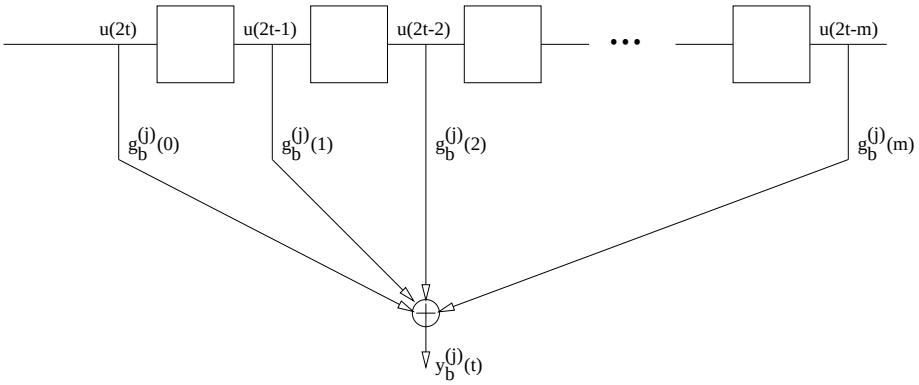

Figure 1.4: Two bits per symbol encoder using a single shift register. At time t, encoder takes two input bits, $u(2t)$, $u(2t-1)$ to produce bits $y_b^{(j)}(t)$, $j \in \{1, ..., N_t\}$, $b \in \{0, 1, ..., K-1\}$ where N_t is the number of transmit antennas and K is the number of output bits per antenna. With m memory elements, the number of states is 2^{m-1} since input bits are shifted two at a time.

excess MI at BER=10^{-5}. The actual maximum excess MI required to achieve BER=10^{-5} in simulation over the compound 2×2 matrix channel is denoted as ΔMI. The excess MI figures of equivalent complexity per bit standard (single dimensional) best rate-$\frac{1}{2}$ + QPSK schemes on the AWGN channel *only* are displayed in Table 1.1. Universal codes for the compound matrix channel are competitive with codes designed for those 2×2 diagonal channels with equal eigenvalues and with a single zero eigenvalue [46]. At 32 states, the universal code becomes competitive with the AWGN code performance on the AWGN channel.

For AWGN channels alone, the best free-distance rate-1/2 32-state linear encoder ($g_0 = 65, g_1 = 57$) driving a Gray-labeled QPSK constellation has $d_{\min}(\text{AWGN}) = 4$, and a corresponding $N_b = 1$. Using Viterbi decoding with trellis termination after 128 data symbols and full-traceback, this code achieves BER $= 10^{-5}$ on the AWGN channel at SNR = 4.65 dB, corresponding to an excess MI

Table 1.2: Minimax excess-MI 2×8-PSK codes found by exhaustive search. Encoder generator vectors in octal notation: $157 \rightarrow 1 + D + D^3 + D^4 + D^5 + D^6$, $035 \rightarrow D^2 + D^3 + D^4 + D^6$. Gray-labeled 8-PSK: $\{0, 2, 3, 1, 5, 7, 6, 4\}$ going around the circle. Third column: Maximum excess MI per transmit antenna required over the compound 2×2 channel at BER $= 1 \times 10^{-5}$, based on simulation. Fourth column: Percent excess MI loss under AWGN from the best Ungerboeck-code for AWGN performance at similar complexity (Table 1.1).

S	$(g_0^1, g_1^1, g_2^1), (g_0^2, g_1^2, g_2^2)$	ΔMI	%loss
4	$(013, 010, 001), (002, 004, 007)$	1.80	43
8	$(021, 017, 007), (002, 036, 035)$	1.40	21
16	$(061, 012, 071), (032, 017, 027)$	1.23	19
32	$(157, 066, 121), (050, 024, 070)$	1.06	8
64	$(316, 108, 272), (127, 218, 353)$	0.95	8

of 0.97 bits at 1 bit/symbol rate. The proposed universal 32-state 2 bits/symbol code has less minimum distance under the AWGN, $d_{\min}(\text{AWGN}) = 3.55$ but with better bit error multiplicity $N_b = 0.125$, and requires an excess MI of 1.06 bits per transmit antenna. However, this code handles *every* matrix channel with similar performance. In fact, the maximum excess MI required to achieve BER $= 1 \times 10^{-5}$ over the compound matrix channel is 1.06 bits. The parallel AWGN channel is therefore a least-favorable matrix channel for this code. Note that the two codes have similar complexity per decoded bit. The added consistent performance over the compound matrix channel costs less then 0.1 bits of excess MI for this case.

Figure 1.5 shows the eigenvalue pair scatter plot for the 32-state universal ST-TCM on the $\lambda_1 - \lambda_2$ plane. Every marker in the figure is a simulated op-

erating point of BER $= 10^{-5}$ on a channel with eigenvalues as the coordinates. The universal code requires at most 1.06 bits of excess MI per transmit antenna to achieve BER $= 10^{-5}$ on any channel, losing about 0.09 bits from the best AWGN performance achieved by the standard QPSK trellis code operating only in AWGN.

1.5.3 Comparison with other approaches

Orthogonal space-time block codes (STB) ([51], [60]) are attractive candidates for imposing robust error performance over the compound matrix channel since they transform matrix channels to single-antenna channels. For 2 bits/symbol transmission over two transmit antennas, the orthogonal scheme of Alamouti [51] can be combined with a 2 bits/symbol standard trellis code. This STB code effectively transforms the $2 \times N_r$ channel to a single-dimensional channel with quasistatic gain $||\mathbf{H}|| = (\sum_{1 \le i \le N_r} (|H_{i,1}|^2 + |H_{i,2}|^2))^{1/2}$. For two receivers, the effective channel is

$$\begin{aligned} y_k &= ||\mathbf{H}|| x_k + w_k, \quad w_k \sim \mathcal{N}(0, N_0) && (1.41) \\ &= \sqrt{\lambda_1 + \lambda_2} \cdot x_k + w_k. && (1.42) \end{aligned}$$

At constant mutual information level, a standard trellis code driving an Alamouti STB code will appear best under the singular channels since singular channels maximize the quantity $(\sum_{i,j} |H_{i,j}|^2)$ under a constant MI constraint. Similarly, the code will appear worst under the equal-eigenvalue (equivalently, unitary or AWGN) channels since these channels minimize $(\sum_{1 \le i,j \le 2} |H_{i,j}|^2)$ under constant MI. Note that the MI provided by the channel is

$$\mathrm{MI}(\mathbf{H}) = \log_2(1 + \lambda_1)(1 + \lambda_2) \quad \text{bits/symbol} \tag{1.43}$$

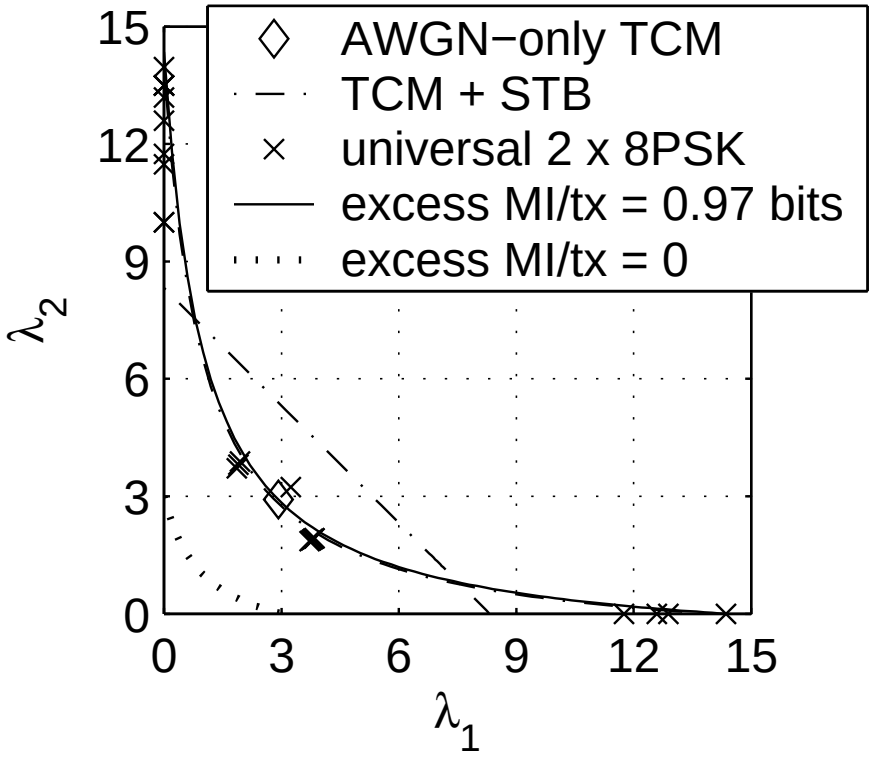

Figure 1.5: Eigenvalue pair scatter plot for the universal 32-state ST-TCM and the STB+TCM scheme at BER $= 10^{-5}$. For the ST-TCM scheme, the worst-case MI to achieve BER $= 10^{-5}$ is 4.12 bits/symbol, attained over unitary channels. The MI requirement of the TCM + STB scheme at BER $= 10^{-5}$ is a linear function of channel eigenvalues with worst-case total MI is 4.74 bits/symbol (excess-MI = 1.37 bits/symbol/antenna), attained on unitary channels. The standard trellis code operating only on AWGN requires MI = 3.94 bits/symbol at BER $= 1 \times 10^{-5}$.

whereas the MI provided by the Alamouti scheme is

$$\log_2(1 + \lambda_1 + \lambda_2) \quad \text{bits/symbol}, \tag{1.44}$$

the capacity of the effective channel in (1.42). As a consequence of the data processing inequality, the MI (1.44) provided by Alamouti's scheme could never exceed the original MI (1.43). Note that for nonsingular channels there is a loss in MI imposed by the orthogonal STB scheme. The worst loss occurs for parallel AWGN. For the operating point of interest, approximately $\lambda_1 = \lambda_2 = 4.17(6.2 \text{ dB})$, this loss is about 0.76 bits per symbol per antenna. For higher λ's associated with higher rates, this loss would be more severe. The 32-state 2 bits/symbol TCM of Ungerboeck [22] achieves BER $= 10^{-5}$ at $E_s/N_0 = 9.2$ dB. This corresponds to an excess MI of 1.22 bits per 2 bit symbol in AWGN. When driving the Alamouti STB code

$$\begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix} \tag{1.45}$$

over 2×2 singular channels, it requires 9.2 dB SNR per transmit antenna to achieve BER $= 10^{-5}$, yielding an excess MI of 0.61 bits per symbol per antenna. Over the diagonal equal-gain (AWGN) channel, it requires 6.2 dB SNR per transmit antenna to achieve BER $= 10^{-5}$, requiring an excess MI of 1.37 bits. This scheme requires at least 0.61 bits and at most 1.37 bits of excess MI to achieve BER $= 10^{-5}$ on any matrix channel. At the same complexity and BER, our universal space-time code requires at most 1.06 bits of excess MI. For 64 states, the TCM + STB scheme requires at least 0.58 bits (on singular channels) and at most 1.32 bits (on unitary channels) of excess MI per transmit antenna to achieve BER $= 10^{-5}$ on the 2×2 compound channel. At this complexity, the universal code requires a worst-case excess MI of 0.95 bits per transmit antenna.

Figure 1.5 also displays the MI requirement of the 32-state TCM + STB scheme over the unordered eigenvalues λ_1, λ_2 of $\mathbf{H}\mathbf{H}^\dagger$, $\mathbf{H} \in \mathbb{C}^{2\times 2}$. The operating MI curve of this scheme is a straight line since the BER performance depends only on the sum $\lambda_1 + \lambda_2$. Figure 1.6 shows the proximity to channel capacity for the 32-state universal code, the 32-state TCM + STB scheme and two recent codes designed for quasistatic Rayleigh fading, for a set of matrix channels with different eigenvalue skews. For each skew value, the performance results under 8 different channels sampled according to (1.38) are recorded, and the minimum and maximum values of the required excess MI at BER=1×10^{-5} are displayed. When the channel is singular, the universal code requires a maximum excess MI of 0.99 bits at BER=1×10^{-5}. The 32-state full-rate code of Yan and Blum [39], denoted YB, delivers worse than uncoded performance when the channel is singular ($\kappa = 0$). The expanded orthogonal construction of Siwamogsatham and Fitz [53], denoted SF, has minimum distance of $\sqrt{2}$ under every singular channel, thus delivering exactly uncoded performance with 1.12 bits of excess MI per antenna. The TCM + STB scheme is the best for singular channels, requiring only 0.61 bits of excess MI per antenna on all singular channels.

At the other end of the skew spectrum, the YB, SF, and the universal codes have similar performance under AWGN with excess MIs of 1.06-1.10 bits per antenna (SNR = 5-5.2dB/tx) whereas the TCM + STB code performs worst, requiring 1.37 bits per antenna (SNR = 6.2 dB/tx). The excess MI requirement for the SF code is 1.10 bits Figure 1.7 displays the BER performance of the four codes as well as the uncoded QPSK transmission over the channel

$$\mathbf{H}_0 = \begin{bmatrix} 0.5 & 0.5 \\ -0.5 & -0.5 \end{bmatrix}. \tag{1.46}$$

The YB code performs approximately 1 dB worse than the uncoded transmission

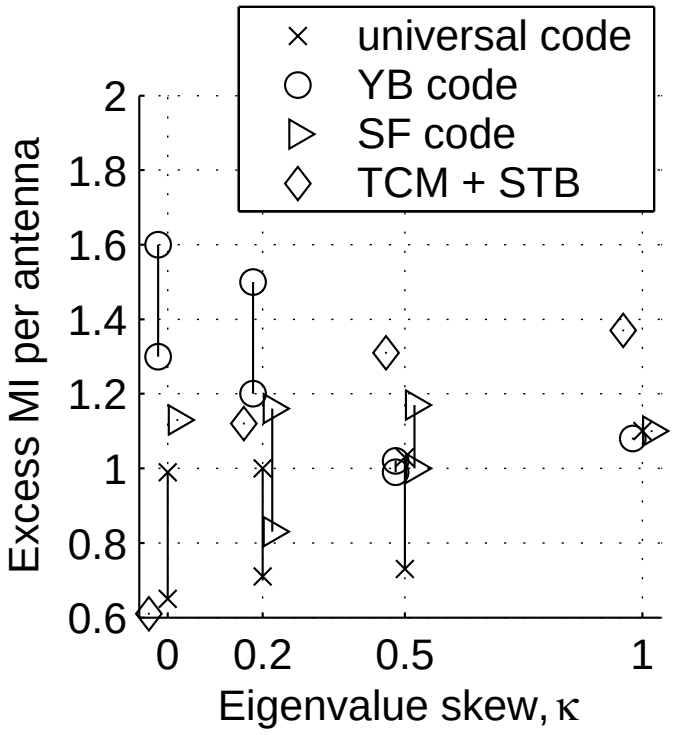

Figure 1.6: Excess MI profile of the universal code in comparison with the codes designed for quasistatic Rayleigh fading, based on BER $= 10^{-5}$. The excess MI profiles are consistent with the minimum distance analysis of Theorem 2: Worst case performance over singular channels ($\kappa = 0$) and worst-case performance over the AWGN channel ($\kappa = 1$) limit the worst-case performance over all eigenvalue skews.

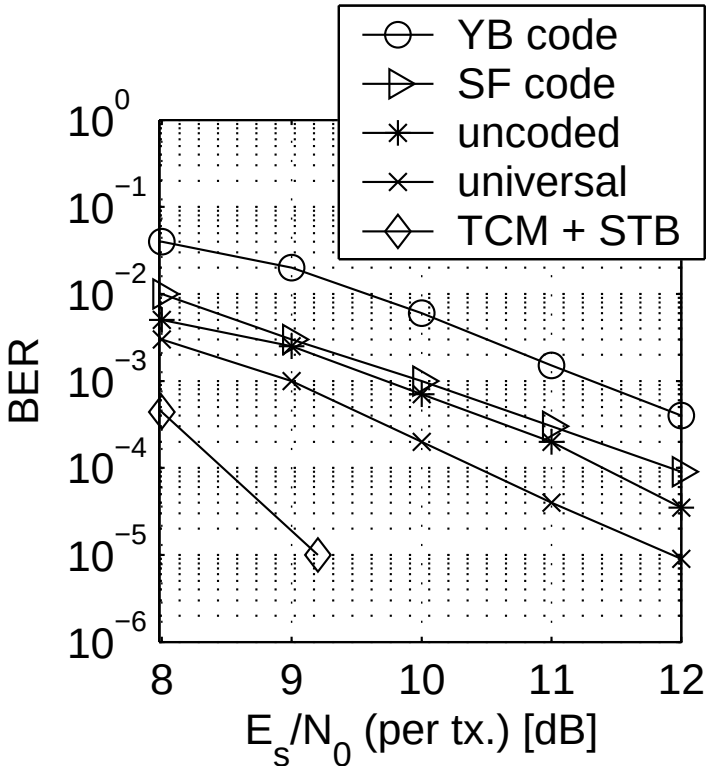

Figure 1.7: Bit error rate performance of the universal code in comparison with SF, YB and TCM + STB, under the singular channel $\mathbf{H}_0$ of (1.46). Uncoded 2 bits/symbol performance is provided for reference. Excess MI requirements for the universal, SF, YB and TCM + STB schemes are 1, 1.12, 1.55 and 0.61 bits per transmit antenna, respectively.

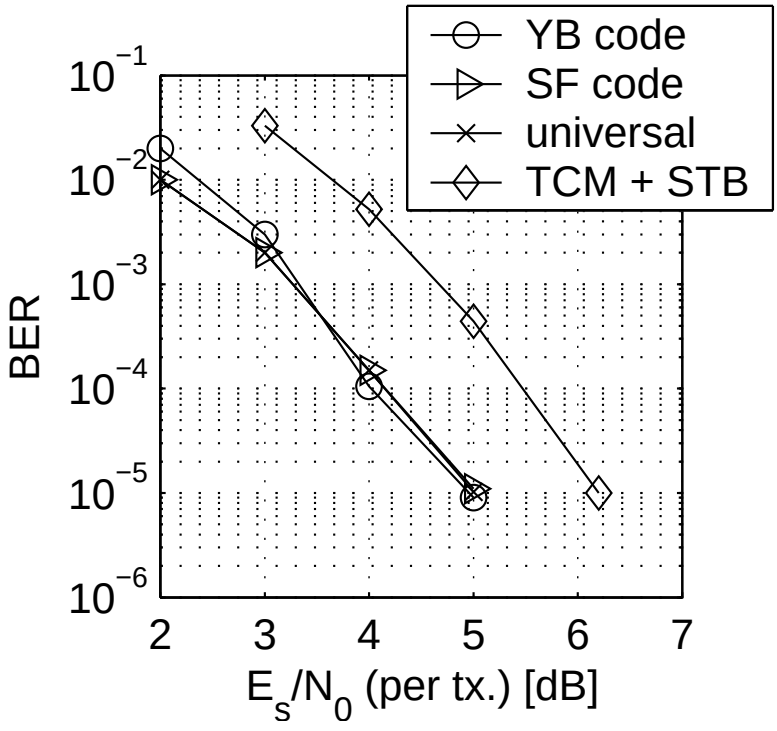

Figure 1.8: Bit error rate performance of the universal code in comparison with the SF, YB codes as well as the TCM + STB scheme under the unitary channel.

whereas the SF code delivers approximately uncoded performance. The universal code maintains coded performance with about 1 dB coding gain at BER = 10^{-5}. For this singular channel, the TCM + STB scheme outperforms the universal trellis code by 3 dB at BER=10^{-5}. Although the STB+TCM scheme requires 9.2 dB to achieve BER=10^{-5} on all singular channels, the SNR requirement for the ST-TCM scheme varies between 9.5 dB (excess-MI = 0.66 bits/tx.) and 11.6 dB (excess-MI = 0.97 bits/tx.).

Figure 1.8 displays the BER performance of these codes under the AWGN channel. The YB code, SF code and the universal code all perform similarly whereas the TCM + STB scheme is about 1 dB worse at BER=10^{-5}.

Figure 1.9 compares bit error rate (BER) performances under quasistatic

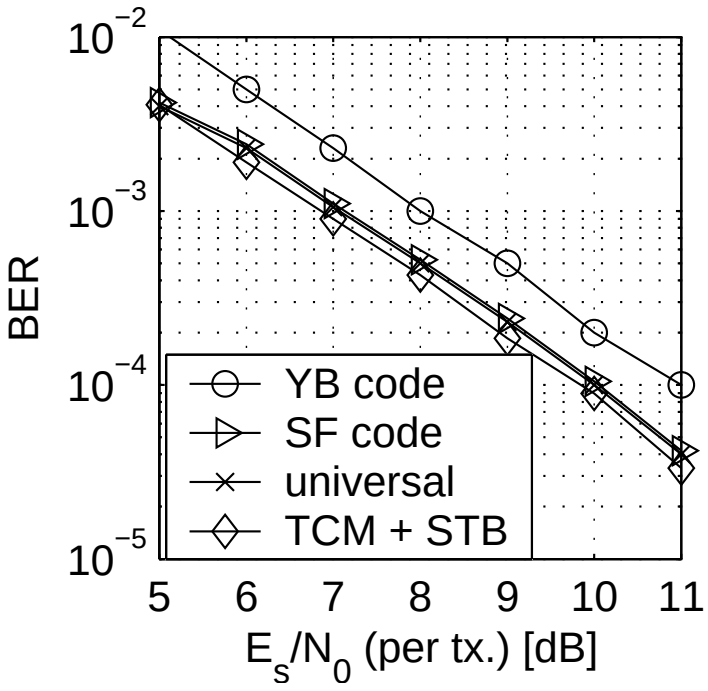

Figure 1.9: Bit error rate performance of the universal code under quasistatic Rayleigh fading in comparison with two full-rate codes designed for this channel model and a STB-coded TCM.

Rayleigh fading. For a frame duration of 128 2-bit data symbols with trellis termination, the BER performance of the 32-state universal code is 1 dB better than that of the YB code, and is same as the SF code. The TCM + STB scheme outperforms the universal ST-TCM by about 0.3dB. All codes have similar decoding complexity and rate.

For the same simulation, Figure 1.10 displays the distribution of bit errors among errored frames not in outage ($\mathrm{MI}(\mathbf{H}) \geq 2$ bits). For the universal code, approximately one quarter of these errored frames contain 3 or less bit errors. This ratio is around ten percent for the YB and SF codes. The histogram number of bit errors frame in the YB and SF code both peak around 6 bit errors

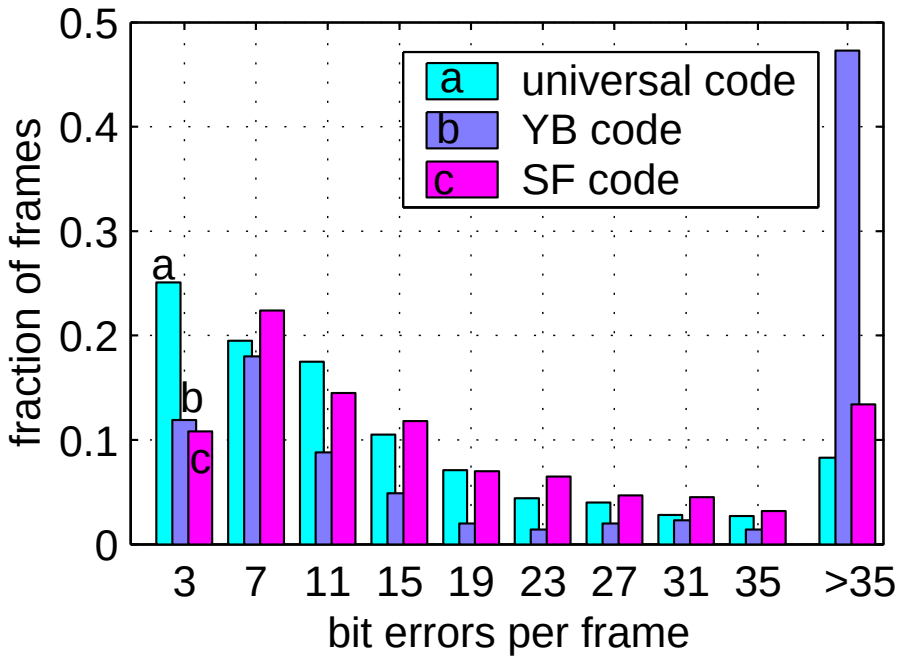

Figure 1.10: The distribution of bit errors over different non-outage frames ($\mathrm{MI}(\mathbf{H}) > R$) over ten thousand errored non-outage frames.

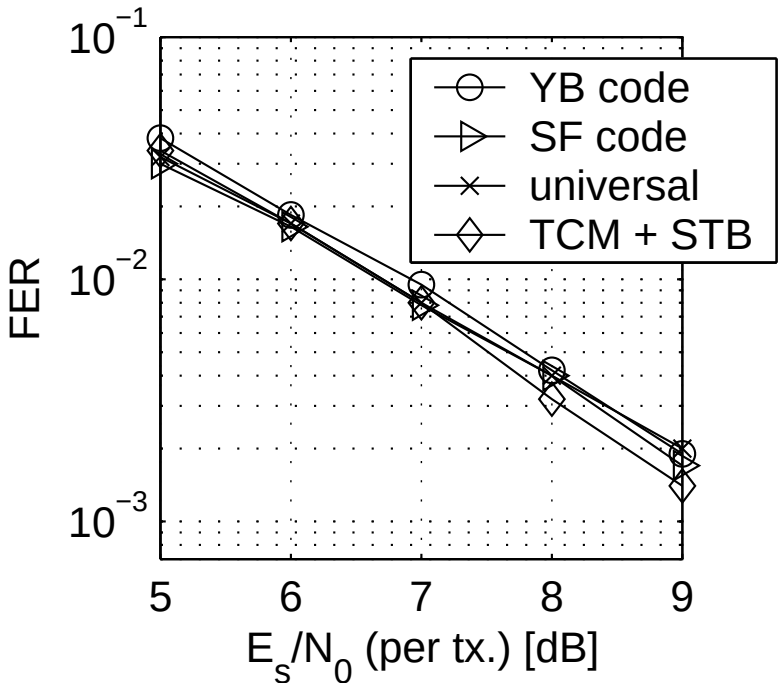

Figure 1.11: Frame error rate performance of the universal code under quasistatic Rayleigh fading in comparison with other coded schemes after a 3-bit-error correcting code.

per/frame. For the YB code nearly half of these errored frames contain 36 or more bit errors. The FER performance of the universal code benefits the most from a three-error-correcting outer code, by more than 0.6 dB. The FERs of other codes improve only by about 0.2 dB. The STB+TCM scheme has favorable FER performance over the ST-TCMs at high SNRs.

Figure 1.11 compares the frame error rate performances of the YB, SF, TCM + STB and universal codes under quasistatic Rayleigh fading after a three-error-correcting outer code.

For four and eight states, the work of Sandu, Heath and Paulraj [52] shows that the TCM + STB schemes have better FER performance than the ST-TCMs

under quasistatic Rayleigh fading. However, Sandu, Heath and Paulraj predicted that with increasing number of states the advantage of TCM + STB schemes diminishes. Results displayed in Figures 1.9 and 1.11 are in agreement with this prediction.

1.5.4 Search for universal codes over more antennas

For $N \times N$ ST-TCMs, the sampling of $N \times N$ complex-valued matrices becomes prohibitive. Instead, one can search over space of $N \times 1$ vector channels and minimize the maximum excess-MI required to achieve the target BER. For the single-receiver case, the channel matrix is a $1 \times N$ vector $\mathbf{h}$. Let $\mathbf{E}$ be an error event matrix of a code that is transmitting at R bits/symbol. For a single receiver, the minimum distance of this error event under a minimum MI-constraint of R bits/symbol is

$$\begin{aligned} \left(d^2_{(N_r=1)}(\mathbf{E})\right)_{\min} &= \min_{\mathbf{h}:\log(1+||\mathbf{h}||^2)\geq R} ||\mathbf{hE}||^2 && (1.47) \\ &= \min_{||\mathbf{h}||=(2^R-1)^{1/2}} ||\mathbf{hE}||^2 && (1.48) \\ &= (2^R-1)\zeta_1(\mathbf{EE}^\dagger). && (1.49) \end{aligned}$$

Therefore, a fine sampling of $N \times 1$ channels at constant MI will yield an accurate estimate of the worst-case minimum error event eigenvalue of the code.

Another possible way to decrease the complexity of an $N \times N$ search is to impose a minimum rank constraint on the matrix channels $\mathbf{H}$. If $\mathbf{H}$ has rows $\mathbf{h}_k$,

$k = 1, ..., N$, then

$$\begin{aligned}
&\left(d^2_{(N_r=N)}(\mathbf{E})\right)_{\min} \\
&\quad = \min_{\mathbf{H}:\mathrm{MI}(\mathbf{H})\geq R,\, \mathrm{rank}(\mathbf{H})\geq r} \mathrm{tr}(\mathbf{H}\mathbf{E}\mathbf{E}^\dagger\mathbf{H}^\dagger) && (1.50)\\
&\quad = \min_{\mathbf{H}:\mathrm{MI}(\mathbf{H})\geq R,\, \mathrm{rank}(\mathbf{H})\geq r} \sum_{k=1}^{N} ||\mathbf{h}_k\mathbf{E}||^2 && (1.51)\\
&\quad \geq r(2^{R/r}-1)\zeta_1(\mathbf{E}\mathbf{E}^\dagger) && (1.52)\\
&\quad = \frac{r(2^{R/r}-1)}{(2^R-1)} \left(d^2_{(N_r=1)}(\mathbf{E})\right)_{\min}. && (1.53)
\end{aligned}$$

1.6 Universal design vs. design for quasistatic Rayleigh channels

Consider a quasistatic channel with a coherence interval large compared to the constraint length of the code. When the entries of the channel matrix follow a complex Gaussian distribution ([26], [61]), the coding gain at high SNR (of a full-rank code) in terms of average performance is determined by the minimum determinant

$$\Delta_P^* = \min_{\mathbf{X}\to\hat{\mathbf{X}}} \det\left(\mathbf{X}-\hat{\mathbf{X}}\right)\left(\mathbf{X}-\hat{\mathbf{X}}\right)^\dagger \tag{1.54}$$

of the error events. From a universal coding perspective, the minimum squared distance under any channel $\mathbf{H} \cong \boldsymbol{\lambda}$ provides an upper bound on the minimum determinant of the code. In fact, using arithmetic-mean geometric-mean inequality, the minimum determinant is upper bounded as

$$\Delta_P^* \leq \frac{d^2_{\min}(\mathbf{H})}{2\det(\mathbf{H}\mathbf{H}^\dagger)}. \tag{1.55}$$

Theorem 5 shows that the determinant of an error event is proportional to the smallest squared distance of the event over a restricted set of channels that sup-

port the rate of transmission. The minimum determinant of the code is therefore proportional to the smallest $d_{\min}(\mathbf{H})$ of the code over the same set of channels.

Theorem 5 *The determinant of an error event* $\mathbf{X} \to \hat{\mathbf{X}}$ *is proportional to the minimum squared distance of the event over the set of channels*

$$\mathcal{H}_1 = \left\{ \mathbf{H} \in \mathbb{C}^{N \times N} : 1 + \log_2 \det(\mathbf{H}\mathbf{H}^\dagger) \geq R \right\}, \tag{1.56}$$

in fact, if $\mathbf{X} - \hat{\mathbf{X}} \cong \boldsymbol{\zeta}$*, then*

$$\left(\prod_{i=1}^{N} \zeta_i \right)^{1/N} = \frac{1}{N(2^{R-1})^{1/N}} \inf_{\mathbf{H} \in \mathcal{H}_1} d^2 \left(\mathbf{X} \to \hat{\mathbf{X}} | \mathbf{H} \right). \tag{1.57}$$

Proof 5 *Theorem 2 indicates that* $\inf_{\mathbf{H}:\mathbf{H} \cong \boldsymbol{\lambda}} d^2_{min}(\mathbf{X} \to \hat{\mathbf{X}} | \mathbf{H}) = \sum_{i=1}^{N} \lambda_{N+1-i} \zeta_i$ *with the convention* $\lambda_i \leq \lambda_{1+1}$ *and* $\zeta_i \leq \zeta_{i+1}$*. Over all channel eigenvalue vectors* $\boldsymbol{\lambda}$ *such that* $\prod_{i=1}^{N} \lambda_i \geq 2^{R-1}$*, this sum is minimized when*

$$\lambda_i = \frac{\beta}{\zeta_i}, \quad i = 1, ..., N \tag{1.58}$$

and $\beta > 0$ *is such that* $\prod_{i=1}^{N} \lambda_i = 2^R - 1$*, resulting in a minimum squared distance of*

$$d^2 = \sum_{i=1}^{N} \frac{\beta}{\zeta_i} \zeta_i = N\beta = N(2^{R-1})^{1/N} \left(\prod_{i=1}^{N} \zeta_i \right)^{1/N}. \tag{1.59}$$

■

Note that $\mathbf{H} \in \mathcal{H}_1$ implies

$$R \leq \log_2 \left(1 + \det(\mathbf{H}\mathbf{H}^\dagger) \right) \tag{1.60}$$

$$< \log_2 \left(\det(\mathbf{I} + \mathbf{H}\mathbf{H}^\dagger) \right) \tag{1.61}$$

$$= \mathrm{MI}(\mathbf{H}), \tag{1.62}$$

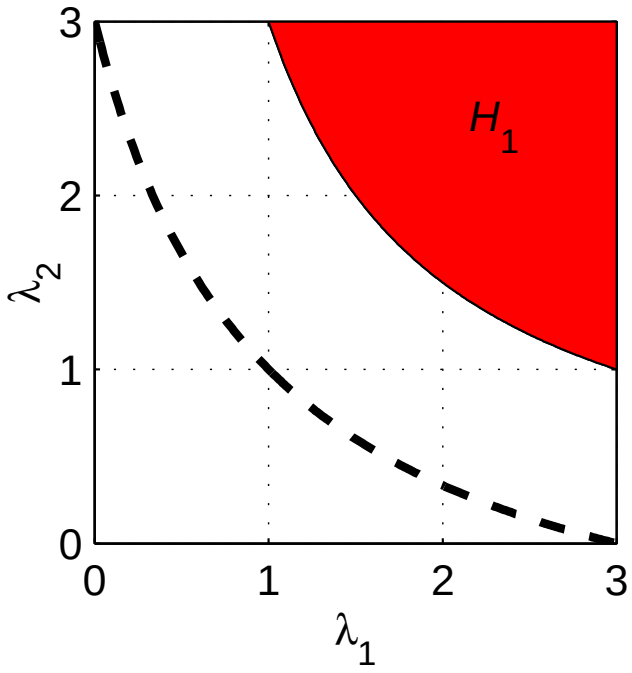

Figure 1.12: The subcollection of channels $\mathcal{H}_1$ of (1.56) at MI $= 1$ bit per symbol per antenna for two antennas, in the λ_1-λ_2 space. The collection of all the channels that have unit capacity per antenna, is the region above the dashed curve.

therefore $\mathcal{H}_1$ is a proper subset of all the $N \times N$ channels that can support rate R. In fact $\mathcal{H}_1$ does not include any singular channel. At $R = 2$ bit/symbol level, Figure 1.12 shows $\mathcal{H}_1$ in comparison to the achievable region of Figure 1.1.

Thus, searches based on the minimum-determinant of a code have some universality, but only for a subset of nonsingular channels that can support the rate of transmission. Perhaps more importantly, these searches often ignore the rate and ECL requirements for operation on singular channels.

Finally, the minimum distance of the error event over all channels with at least R bits of mutual information is given [65] by the waterfilling formula

$$\inf_{\mathbf{H}:\mathrm{MI}(\mathbf{H})\geq R} d^2\left(\mathbf{X} \to \hat{\mathbf{X}}|\mathbf{H}\right) = \sum_{i=1}^{N} (\zeta - \zeta_i)^+ \tag{1.63}$$

where

$$\zeta = 2^{R/q}\left(\prod_{i=1}^{q} \zeta_i\right)^{1/q} \tag{1.64}$$

$$q = \max\left\{j : 2^{R/j}\left(\prod_{i=1}^{j} \zeta_i\right)^{1/j} \geq \zeta_j\right\}. \tag{1.65}$$

However, this criterion is not amenable to code search.

1.7 Summary

This chapter characterized space-time trellis codes that deliver consistent performance on every channel matrix with similar mutual information. The minimum-distance based analysis led to a minimax search for space-time trellis codes having the best worst-case proximity to capacity over the entire collection of matrix channels. In particular, such universal codes maintain a nonzero coding gain over all channels. 32- and 64-state 2-bits/symbol trellis codes found by exhaustive search

have universal bit error rate performance across the channels that support the transmission rate, with worst-case excess mutual informations of 1.06 bits and 0.95 bits per antenna, respectively, guaranteeing good bit error rate performance under any quasistatic fading distribution.

The search space for general vector-labeled space-time trellis codes is prohibitively large for more than three transmit antennas to conduct an exhaustive search for universal trellis codes. Such a search is further complicated by the complex nature of the TUB-based search criterion. Chapter 2 focuses on the class of multiplexed two-dimensional trellis codes and a simplified universality criterion to search for universal codes for more than two transmit antennas.

CHAPTER 2

Universal space-time codes from two-dimensional trellis codes

2.1 Introduction

This chapter extends the ideas in Chapter 1 to propose universal space-time trellis codes formed by straightforward multiplexing of two-dimensional linear trellis codes over two, three and four transmit antennas. These multiplexed trellis codes have simpler maximum-likelihood decoding than general multidimensional (vector-labeled) trellis codes while still providing excellent performance.

Section 2.2 summarizes our results on the worst-case minimum-distance of a space-time code under linear transformations with equal mutual information and derives a simple approximate criterion for universal behavior. This section also formulates the encoder rate, constellation size and trellis complexity requirements for universal space-time trellis codes formed by multiplexing a two-dimensional dimensional trellis code.

Section 2.3 presents several linear trellis codes, found by exhaustive search over their respective encoder classes. These codes provide universal performance when multiplexed over two, three and four transmit antennas and received with maximum-likelihood decoding. When the trellis complexity required for universality over the rank-unconstrained compound channel exceeds practical limits,

reduced-rank schemes with favorable eigenvalue spreads provide universal performance in conjunction with algebraic outer codes that restore the lost diversity. The performance variation of universal codes as compared to other existing space-time codes over different channel instances is illustrated via extensive simulations. Our discussion ends with simulation results showing that the average error performance of the proposed universal codes over quasistatic Rayleigh fading is comparable, and in some cases superior, to existing space-time codes with similar decoding complexity designed specifically for the average error probability performance. A channel-by-channel look at the 2×2 and 3×3 quasistatic Rayleigh fading demonstrates why the proposed universal codes deliver excellent average Rayleigh fading error probability performance and why these universal codes provide a channel-by-channel reliability that codes designed only for good average performance may not provide.

2.2 Design guidelines for universal space-time trellis codes

2.2.1 An approximate criterion for universality

Let $\mathbf{X}$ and $\hat{\mathbf{X}}$ be two different codewords of a space-time code $\mathcal{C}$ for N_t transmit antennas, and let $\mathbf{E} = \hat{\mathbf{X}} - \mathbf{X}$ denote the codeword difference matrix. Recall that under maximum-likelihood (ML) decoding, the probability that the decoder mistakes $\hat{\mathbf{X}}$ for $\mathbf{X}$ conditioned on the perfect knowledge of the channel matrix $\mathbf{H}$ at the receiver is given by

$$P\left(\mathbf{X} \to \hat{\mathbf{X}}|\mathbf{H}\right) = Q\left(\sqrt{\frac{d^2\left(\mathbf{E}|\mathbf{H}\right)}{2N_0}}\right) \tag{2.1}$$

where $d^2(\mathbf{E}|\mathbf{H}) = ||\mathbf{HE}||^2 = \text{trace}\left(\mathbf{HEE}^\dagger\mathbf{H}^\dagger\right)$ is the squared Euclidean norm of the codeword difference matrix $\mathbf{E}$ when transformed by the channel $\mathbf{H}$, and

$Q(\cdot)$ is the standard Gaussian tail integral function. For a fixed channel $\mathbf{H}$, the minimum of the squared-distance $d^2(\mathbf{E}|\mathbf{H})$ over all $\mathbf{E}$ is the $d^2_{\min}(\mathbf{H})$ of the code.

With maximum-likelihood decoding, the universal performance of a space-time code results from its ability to sustain good $d^2_{\min}(\mathbf{H})$ over a family $\mathcal{H}$ of channel instances that support the transmission rate. The smallest value of $d^2_{\min}(\mathbf{H})$ over the compound channel is a function of the eigenvalues of the codeword difference matrices. For a given codeword difference matrix $\mathbf{E}$, $\boldsymbol{\eta} = (\eta_1, ..., \eta_{N_t})$ will be the vector of eigenvalues of $\mathbf{E}\mathbf{E}^\dagger$ with the ordering

$$\eta_1 \geq \eta_2 \geq ... \geq \eta_{N_t}, \tag{2.2}$$

and we will write $\mathbf{E} \cong \boldsymbol{\eta}$. The eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$, where $\mathbf{H}$ is an $N_r \times N_t$ channel gain matrix, are denoted by λ_i, $i = 1, ..., N_r$ with the ordering

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{N_r}, \tag{2.3}$$

and we write $\mathbf{H} \cong \boldsymbol{\lambda}$ where $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_{N_r})$. As in Chapter 1, we assume that $N_r \leq N_t$.

Consider the $N_r \times N_t$ compound channel with capacity R bits per symbol

$$\mathcal{H}(N_r, N_t, R) = \left\{\mathbf{H} \in \mathbb{C}^{N_r \times N_t} : \mathrm{MI}(\mathbf{H}) \geq R\right\}. \tag{2.4}$$

For a particular code, we are interested in the minimum of the squared Euclidean norm, $d^2(\mathbf{E}|\mathbf{H})$, of a codeword difference matrix $\mathbf{E} \cong \boldsymbol{\eta}$ over all instances $\mathbf{H}$ of the compound channel $\mathcal{H}$. In order to find the minimum of $d^2(\mathbf{E}|\mathbf{H})$ over all $\mathbf{H} \in \mathcal{H}(N_r, N_t, R)$, we first consider the minimum over all $\mathbf{H} \cong \boldsymbol{\lambda}$. Chapter 1 showed that for a given codeword difference matrix $\mathbf{E}$ and a given set of channel eigenvalues $\boldsymbol{\lambda}$, $d^2(\mathbf{E}|\mathbf{H}) = ||\mathbf{H}\mathbf{E}||^2$ is minimized when the channel aligns its strongest modes with the weakest eigendirections of $\mathbf{E}\mathbf{E}^\dagger$ and vice versa. In

particular,

$$\inf\left\{d^2(\mathbf{E}|\mathbf{H}) : \mathbf{H}\mathbf{H}^\dagger \cong \boldsymbol{\lambda}\right\} = \sum_{i=1}^{N_r} \lambda_i \eta_{N_t+1-i}, \tag{2.5}$$

therefore

$$\inf_{\mathbf{H}\in\mathcal{H}(N_r,N_t,R)} d^2(\mathbf{E}|\mathbf{H}) = \inf_{\boldsymbol{\lambda}:\sum_{i=1}^{N_r}\log_2(1+\lambda_i)\geq R} \sum_{i=1}^{N_r} \lambda_i \eta_{N_t+1-i}. \tag{2.6}$$

The solution to the optimization problem

$$\min_{\mathbf{H}} \sum_{i=1}^{N_r} \lambda_i \eta_{N_t+1-i}$$
$$\text{subject to} \quad \mathrm{MI}(\mathbf{H}) = \sum_{i=1}^{N_r} \log_2(1+\lambda_i) \geq R$$

has the waterpouring form [65] with the worst-case channel

$$\mathbf{H}^o = \sum_{i=1}^{N_r} \sqrt{\lambda_i^o}\,\mathbf{u}_i \mathbf{n}_{N_t+1-i}^\dagger, \tag{2.7}$$

where $\{\mathbf{u}_k \in \mathbb{C}^{N_r\times 1}\}_{k=1,\ldots,N_r}$ is any orthonormal set of vectors and $\mathbf{n}_i$ is the eigenvector of $\mathbf{E}\mathbf{E}^\dagger$ that corresponds to η_i. The worst-case channel eigenvalues are given by

$$\lambda_i^o = \left\{ \begin{array}{rl} 2^{R/N^*}\frac{G_{N^*}(\boldsymbol{\eta})}{\eta_{N_t+1-i}} - 1, & i = 1,\ldots,N^* \\ 0, & i = N^*+1,\ldots,N_r \end{array} \right\} \tag{2.8}$$

where $N^* = \max\left\{k : 2^{R/k}G_k(\boldsymbol{\eta}) > \eta_{N_t+1-k},\ \ k \leq N_t\right\}$ is the number of nonzero eigendirections that the worst-case channel uses and

$$G_k(\boldsymbol{\eta}) = \left(\prod_{i=N_t-k+1}^{N_t} \eta_k\right)^{1/k}, \quad k = 1,\ldots,N_t, \tag{2.9}$$

is the geometric-mean of the smallest k eigenvalues of $\mathbf{E}\mathbf{E}^\dagger$.

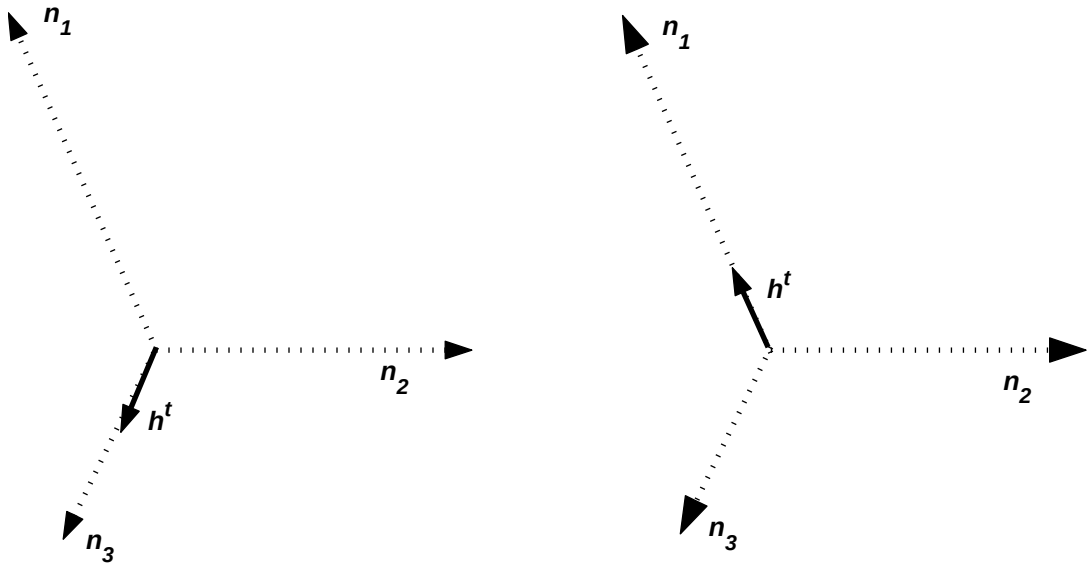

Figure 2.1: Interplay between the channel vector $\mathbf{h}$ and the eigenvectors of $\mathbf{E}\mathbf{E}^\dagger$ where $\mathbf{E}$ is a codeword difference matrix, $\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}}$, for $N_r = 1$. Among all channels $\mathbf{h}$, the probability of mistaking $\mathbf{X}$ for $\hat{\mathbf{X}}$ is the highest when $\mathbf{h}^\dagger$ aligns with the weakest eigenvector ($\mathbf{n}_3$ above) of $\mathbf{E}\mathbf{E}^\dagger$ (left). Among all channels $\mathbf{h}$, the most favorable channel for ML-detection of this error event is when $\mathbf{h}^\dagger$ aligns with the strongest eigenvector ($\mathbf{n}_1$ above) of $\mathbf{E}\mathbf{E}^\dagger$ (right).

This exact solution, however, does not yield a simple criterion for universality except for $N_r = 1$, in which case the worst-case channel is given by

$$\mathbf{h}^o = \left(2^R - 1\right)^{1/2} \mathbf{n}_{N_t}^\dagger \tag{2.10}$$

where $\mathbf{n}_{N_t}$ is the eigenvalue corresponding to the weakest eigenvector of $\mathbf{E}\mathbf{E}^\dagger$. Figure 2.1 illustrates this case.

A space-time code will perform identically over any $N_t \times 1$ channel with a fixed MI if and only if codeword difference matrices are scalar multiples of unitary matrices [31]. This implies that the eigenvalues of $\mathbf{E}\mathbf{E}^\dagger$ are equal, for any codeword difference $\mathbf{E}$. However, equality in the error event eigenvalues usually comes at

the expense of a reduced minimum-Euclidean distance (e.g., [51], [54]), which means that the same code will experience a performance penalty under unitary $N_t \times N_t$ channels, as compared to an AWGN code at the same complexity.

The purpose of universal coding is to deliver R bits per symbol with consistently good error probability across all instances of the $N_t \times N_t$ compound channel $\mathcal{H}(N_t, N_t, R)$. This would guarantee consistent performance for any $N_r \leq N_t$. In fact, for any $N_r \leq N_t$, $\mathcal{H}(N_r, N_t, R)$ is equivalent to the subset of $\mathcal{H}(N_t, N_t, R)$ consisting of square matrices $\mathbf{G}$ of rank less than or equal to N_r, i.e.,

$$\mathcal{H}(N_r, N_t, R) \equiv \{\mathbf{G} \in \mathcal{H}(N_t, N_t, R) : \text{rank}(\mathbf{G}) \leq N_r\} \tag{2.11}$$

where $\equiv$ is in the sense that for $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$, $\mathbf{G} \in \mathbb{C}^{N_t \times N_t}$,

$$\mathbf{H} \equiv \mathbf{G} \quad \text{if and only if} \quad \mathbf{UG} = \begin{bmatrix} \mathbf{H} \\ \mathbf{0} \end{bmatrix} \quad \text{for some unitary } \mathbf{U}. \tag{2.12}$$

For example, the subset of singular channels $\mathbf{G}$ in $\mathcal{H}(N_t, N_t, R)$ is equivalent to the compound channel $\mathcal{H}(1, N_t, R)$ with $N_r = 1$. The equivalence $\mathbf{H} \equiv \mathbf{G}$ implies that a space-time code would have identical error probability performance under $\mathbf{G}$ and $\mathbf{H}$ under ML decoding.

An approximate criterion for universality over $\mathcal{H}(N, N, R)$ is obtained by bounding the worst-case distance over the compound channel. First, an upper bound on the worst-case minimum-distance on the compound channel comes from the minimum-distance over the equal-eigenvalue ($\lambda_i = 2^{R/N} - 1$, $i = 1, ..., N$) channels. For a codeword difference $\mathbf{E}$,

$$\begin{aligned} &\inf_{\boldsymbol{\lambda}: \sum_{i=1}^{N} \log_2(1+\lambda_i) \geq R} \sum_{i=1}^{N} \lambda_i \eta_{N+1-i}(\mathbf{E}\mathbf{E}^\dagger) \\ &\qquad \leq \left(2^{\frac{R}{N}} - 1\right) \sum_{i=1}^{N} \eta_{N+1-i}(\mathbf{E}\mathbf{E}^\dagger) \qquad (2.13) \\ &\qquad = \left(2^{\frac{R}{N}} - 1\right) \Delta_E(\mathbf{E}) \qquad (2.14) \end{aligned}$$

where $\Delta_E(\mathbf{E}) = \text{trace}(\mathbf{E}\mathbf{E}^\dagger)$ is the squared Euclidean-norm of $\mathbf{E}$.

Now, a lower bound is obtained as

$$\begin{aligned}
&\sum_{i=1}^{N} \lambda_i \eta_{N+1-i}(\mathbf{E}\mathbf{E}^\dagger) \\
&= \left(\sum_{j=1}^{N} \eta_{N+1-j}\right) \left(\sum_{i=1}^{N} \frac{\eta_{N+1-i}}{\sum_{j=1}^{N} \eta_{N+1-j}}(1+\lambda_i) - 1\right) \\
&\overset{(a)}{\geq} \left(\sum_{j=1}^{N} \eta_{N+1-j}\right) \left(2^{\sum_{i=1}^{N} \frac{\eta_{N+1-i}}{\sum_{j=1}^{N} \eta_{N+1-j}} \log_2(1+\lambda_i)} - 1\right) \\
&\overset{(b)}{\geq} \left(\sum_{j=1}^{N} \eta_{N+1-j}\right) \left(2^{\left[R \frac{\eta_N}{\sum_{j=1}^{N} \eta_{N+1-j}}\right]} - 1\right) \\
&\overset{(c)}{=} \Delta_E(\mathbf{E}) \left(2^{R \frac{\eta_N}{\Delta_E(\mathbf{E})}} - 1\right) \qquad (2.15)
\end{aligned}$$

where (a) uses the concavity of the logarithm function, (b) follows from the fact that the minimum of $\sum_{i=1}^{N} \frac{\eta_{N+1-i}}{\sum_{j=1}^{N} \eta_{N+1-j}} \log_2(1+\lambda_i)$ subject to $\sum_{i=1}^{N} \log_2(1+\lambda_i) \geq R$ occurs at $\lambda_1^* = 2^R - 1$ and $\lambda_i^* = 0$ for $i > 1$, and (c) uses the shorthand notation $\Delta_E(\mathbf{E})$ for $\text{trace}(\mathbf{E}\mathbf{E}^\dagger) = \sum_{i=1}^{N} \eta_i(\mathbf{E}\mathbf{E}^\dagger)$. The following lemma summarizes these results.

Lemma 2 *The worst-case minimum distance of a space-time code over the compound $N \times N$ channel of capacity R bits per symbol is bounded as*

$$\begin{aligned}
\left(2^{\frac{R}{N}} - 1\right) \Delta_E^* &\geq \inf_{\mathbf{H} \in \mathcal{H}(N,N,R)} d_{\min}^2(\mathbf{H}) \\
&\geq \min_{\mathbf{E}} \left(2^{R \frac{\eta_N}{\Delta_E(\mathbf{E})}} - 1\right) \Delta_E(\mathbf{E}).
\end{aligned}$$

where $\Delta_E^* = \min_{\mathbf{E}} \text{trace}(\mathbf{E}\mathbf{E}^\dagger)$ is the minimum squared Euclidean-distance of the code.

The first inequality of Lemma 1 implies that a universal code should have good minimum Euclidean-distance. The second inequality of the lemma leads us to

choose, among good minimum Euclidean-distance codes, those codes with high minimum eigenvalue $\eta^*_{N_t} = \min_{\mathbf{E}} \eta_{N_t}(\mathbf{EE})$. Ultimately, the universal performance of a code over the compound channel will be measured by its excess mutual information requirement. Nevertheless, this criterion provides us with a basic rule to prune the search for universal space-time trellis codes.

2.2.2 Multiplexing a two-dimensional trellis code

The requirements for design of universal vector-labeled space-time trellis codes for two transmit antennas were considered in [8]. In this study, we focus on space-time trellis codes generated by straightforward multiplexing of a two-dimensional trellis code over the transmit antennas. This approach results in more manageable code searches. More importantly, the resulting space-time trellis codes have the same maximum-likelihood decoding complexity as the multiplexed code over a single-transmit antenna channel [12].

Consider a rate-k/n convolutional encoder with memory ν. For R-bits-per-symbol transmission over N_t transmit antennas using a 2^m-PSK/QAM constellation, we use this encoder $l = R/k$ times (assume for simplicity that k divides R evenly). Let $(b_0, ..., b_{n-1})$, ..., $(b_{(l-1)n}, ..., b_{ln-1})$ denote the ln codeword bits that the binary encoder would produce for lk successive input bits. If $ln = N_t m$, then we map $(b_{(i-1)m}, ..., b_{im-1})$ onto the 2^m-PSK/QAM constellation for the ith transmit antenna, $i = 1, ..., N_t$. If $ln > N_t m$, we puncture $ln - N_t m$ out of Rn bits and group the remaining $N_t m$ bits similarly, keeping the index order. For example, $R = 3$ information bits input to a rate-1/3 convolutional encoder produces 9 coded bits $b_0, b_1, ..., b_8$. For a 2×16-QAM space-time code delivering $R = 3$ bits per symbol, we puncture one bit, say b_5, and group the remaining eight bits as $(b_0 b_1 b_2 b_3)$, $(b_4 b_6 b_7 b_8)$ to map two 16-QAM symbols, one for each antenna. For

this case, there are nine different ways to puncture.

2.2.3 Encoder rate and constellation size requirements for universal space-time codes

Universal performance over the compound channel requires the following design rules for multiplexed two-dimensional trellis codes.

- Constellation size: Universal codes are designed to maintain redundancy over any instance of the compound channel. Therefore, the constellation size should be large enough to transmit R bits per symbol redundantly even when $N_t - 1$ of the transmit antennas fail, corresponding to a channel matrix with $N_t - 1$ zero rows[1]. Thus, the constellation size should satisfy

$$m > R. \tag{2.16}$$

  For example, for 2 bits/symbol transmission over $N_t = 2$ antennas, we propose rate-1/3 convolutional encoders mapping 8-PSK constellations. When one of the two transmit antennas fails, the effective code is then a rate-2/3 8-PSK trellis code.

- Trellis complexity: The effective code length (ECL) is the smallest symbol-wise Hamming weight of an error event. For a k-bits/symbol linear trellis code, ECL $\leq \lfloor \nu/k \rfloor + 1$ where ν is the memory of the encoder [46]. In order to have $\eta^*_{N_t} > 0$, the shortest error-event of the two-dimensional code should be long enough to occupy at least N_t channel symbols when multiplexed over N_t transmit antennas. Therefore,

$$\left\lfloor \frac{\nu}{k} \right\rfloor \geq N_t(N_t - 1). \tag{2.17}$$

[1]The channel matrix cannot be the zero matrix which has zero MI.

This was noted earlier in [28]. For example, 1 bit/symbol trellis code with $S = 2^{\nu}$ states multiplexed over three transmit antennas cannot provide full transmit diversity if $\nu < 6$. For $\nu = 5$, the shortest error-events of the space-time code are 3×2 matrices of the form

$$\mathbf{E} = \begin{bmatrix} e_0 & e_3 \\ e_1 & e_4 \\ e_2 & e_5 \end{bmatrix} \tag{2.18}$$

where $\mathbf{e} = (e_0, e_1, e_2, e_3, e_4, e_5)$ is an error event of the original 1 bit/symbol trellis code. The event in (2.18) has rank less than or equal to 2, i.e., $\eta_3\left(\mathbf{E}\mathbf{E}^{\dagger}\right) = 0$. The trellis complexity requirements can be relaxed by sacrificing one or more levels of transmit diversity. In general, for r-levels of transmit diversity, the necessary (but not sufficient) trellis complexity obeys

$$\left\lfloor \frac{\nu}{k} \right\rfloor \geq N_t(r-1). \tag{2.19}$$

With $r < N_t$ levels of transmit diversity, there is always a singular channel that supports the transmission rate, under which two codewords cannot be reliably (with error probability less than $1/2$) distinguished at the receiver. However, a diversity-r code with good η_r^* can provide universal performance over all channels that establish at least $N_t - r + 1$ equally strong spatial eigenmodes. Moreover, as will be illustrated in Section 2.3, missing levels of transmit diversity can be restored by an outer code.

2.3 Universal space-time codes from two-dimensional trellis codes

The search for universal codes is complicated by the non-additive nature of error-event eigenvalues over trellis branches. For $N_t = 2$, a finite sampling of channels can be used to approximate the worst-case minimum eigenvalue of the code over multiple forward trellis searches [8]. For $N_t = 3$ and $N_t = 4$, a small set of test channels (including the unitary and some rank-deficient diagonal channels) were used to prune the stack-based algorithm [17] to determine codes with the best worst-case eigenvalues.

Besides the worst-case minimum eigenvalue $\eta^*_{N_t}$, three other code parameters will be useful for interpreting the performance of space-time codes in different scenarios. The minimum squared Euclidean-distance of the code

$$\Delta^*_E = \min_{\mathbf{E}} \sum_{i=1}^{N_r} \eta_i(\mathbf{E}\mathbf{E}^\dagger), \tag{2.20}$$

together with its average bit multiplicity determines the high-SNR bit-error-rate performance under unitary channels. The diversity order of the code,

$$\Delta^*_H = \min_{\mathbf{E}} \operatorname{rank}(\mathbf{E}), \tag{2.21}$$

is the smallest number of non-vanishing eigenvalues that the error-event matrices have. A space-time code has nonzero worst-case minimum-distance on the rank-unconstrained compound channel if and only if it has full diversity. A final parameter of interest is the minimum product-distance

$$\Delta^*_P = \min_{\mathbf{E}} \prod_{i=1}^{\Delta^*_H} \eta_i(\mathbf{E}\mathbf{E}^\dagger), \tag{2.22}$$

which determines the coding gain under Rayleigh fading at high SNR. The worst-case ith eigenvalue over all $\mathbf{E}\mathbf{E}^\dagger$ is denoted by $\eta^*_i = \min_{\mathbf{E}} \eta_i(\mathbf{E}\mathbf{E}^\dagger)$.

All codes presented in this section are found by exhaustive search over their class of encoders to maximize the worst-case minimum-eigenvalue η_r^* under a transmit diversity constraint $\Delta_{H,\min} = r$ while sacrificing no more than twenty percent of the maximum squared Euclidean-distance achievable within the same class. In cases where $\eta_r^* > 0$ cannot be attained without less than twenty percent loss in squared Euclidean-distance, the search was carried out without the squared Euclidean-distance constraint. Ties are resolved by looking at the average bit multiplicity of worst-case eigenvalues.

We will compare our codes with the best space-time trellis codes with similar complexity that we have found in the literature: For $N_t = 2$ and $N_t = 3$, space-time codes of Aktas *et. al* [12] and for $N_t = 3$ and $N_t = 4$ the codes of Chen *et. al* [67] have excellent error probability performance quasistatic Rayleigh fading channels. Under quasistatic Rayleigh fading, comparison with other published codes (e.g., [36], [40]) should follow from the simulation results that are presented in Section 2.3.5. Due the complete loss of redundancy under at least one channel with rank one, the worst-case compound channel performances of these codes is similar to those of [12], [67] .

2.3.1 Universal codes for $N_t = 2$ transmit antennas

Table 2.1 lists universal trellis codes for $R = 1$, 2 and 3 bits per symbol transmission over two, three and four transmit antennas. We consider the $R = 2$ bits per symbol case in detail.

Table 1 2.1 lists rate-1/3 + 8PSK trellis-coded modulations for for 16, 32 and 64 states (codes #2, #3, #4, respectively) such that when multiplexed over two transmit antennas, deliver universal performance over the 2×2 compound channel.

Table 2.1: Multiplexed universal codes. R bits per symbol code trellis codes for N_t transmit antennas. A rate-k/n convolutional encoder outputs nR code bits which are mapped on $N_t \times 2^m$-PSK/QAM. 16QAM, QPSK: Gray labeling, 4PSK: Natural labeling. 8PSK: Gray labeling 0,2,3,1,5,7,6,4 in octal going around the circle.

Code	N_t	R	k/n	ν	g_i	Const.	$\eta^*_{\Delta_{H,\min}}$	Δ^*_P	Δ^*_E
1	2	1	1/4	4	03 22 36 04	QPSK	$\eta^*_2 = 5.53$	80	20
2	2	2	1/3	4	31 05 35	8PSK	$\eta^*_2 = 0.67$	16.32	12.6
3	2	2	1/3	5	71 31 61	8PSK	$\eta^*_2 = 0.89$	≤ 25.02	13.4
4	2	2	1/3	6	155 056 145	8PSK	$\eta^*_2 = 0.70$	17.7	17.2
5	2	3	1/3	3	06 16 13	16QAM	$\eta^*_2 = 0.4$	1.6	4
6	3	1	1/6	3	15 03 07 05 04 11	QPSK	$\eta^*_3 = 2.44$	248	26
7	3	2	1/5	4	31 06 02 31 02	8PSK	$\eta^*_3 = 0.14$	4.69	7.57
8	3	2	1/3	5	62 55 47	QPSK	$\eta^*_3 = 0.24$	32	22
9	3	3	1/2	5	75 62	QPSK	$\eta^*_2 = 2.0$	15.5	12.0
10	4	2	1/4	4	33 23 26 06	QPSK	$\eta^*_2 = 4.60$	78.6	24

Table 2.2: Multiplexed trellis codes with good Euclidean distance and/or good distance under periodic erasures. R bits per symbol code trellis codes for N_t transmit antennas. A rate-k/n convolutional encoder outputs nR code bits which are mapped on $N_t \times 2^m$-PSK/QAM. 16QAM, QPSK: Gray labeling, 8PSK: Gray labeling 0,2,3,1,5,7,6,4 in octal going around the circle.

Code	N_t	R	k/n	ν	g_i	Const.	$\eta^*_{\Delta_{H,\min}}$	Δ^*_P	Δ^*_E
11	4	2	1/4	5	75 71 67 53	QPSK	$\eta^*_3 = 0.29$	49.1	36
12	2	2	1/3	4	31 27	QPSK	$\eta^*_2 = 1.25$	16	14
13	2	2	1/2	5	53 75	QPSK	$\eta^*_2 = 0.46$	16	8
14	2	2	1/2	6	155 117	QPSK	$\eta^*_2 = 1.04$	18	24
15	2	1	1/4	4	25 27 33 37	QPSK	$\eta^*_2 = 5.53$	24	24
16	2	2	1/3	6	173 062 115	8PSK	$\eta^*_2 = 0.56$	32.3	17.6
17	3	3	1/4	6	117 155 145 137 [46]	16QAM	$\eta^*_2 = 0.28$	5.03	14.4
18	3	3	1/2	5	65 57	QPSK	$\eta^*_2 \leq 1.43$	19.7	16
19	4	2	1/4	4	25 27 33 37	QPSK	$\eta^*_2 = 2.62$	74.14	30
20	4	3	1/4	6	135 147 135 163	8PSK	$\eta^*_2 = 0.98$	45.01	24

Table 2.3: Space-time trellis codes for good average Rayleigh fading performance. 16QAM, QPSK: Gray labeling, 4PSK: Natural labeling. 8PSK: Gray labeling 0,2,3,1,5,7,6,4 in octal going around the circle.

Code	N_t	R	k/n	ν	g_i	Const.	$\eta^*_{\Delta_{H,\min}}$	Δ^*_P	Δ^*_E
21	2	2	2/4	4	from [12]	4PSK	$\eta^*_2 = 0.56$	8	16
22	2	2	2/4	5	from [12]	4PSK	$\eta^*_2 = 1.44$	32	16
23	2	2	2/4	6	from [12]	4PSK	$\eta^*_2 = 1.04$	28	18
24	3	2	2/6	4	[12]	4PSK	$\eta^*_2 = 0.91$	13	12
25	3	2	2/6	5	[12]	4PSK	$\eta^*_3 = 0.156$	8	14
26	4	2	2/8	5	[67]	4PSK	$\eta^*_3 = 0.144$	34.6	36
TCM + Alamouti STB									
27	2	2	2/3	6	[22],[51]	4PSK	$\eta^*_2 = 0.3$	6	9
Super-ortohogonal code									
28	2	2	2/4	6	[53]	4PSK	$\eta^*_2 = 0.4$	8	16
Universal STTCM [8]									
29	2	2	2/6	5	[8]	8PSK	$\eta^*_2 = 0.72$	13.0	12.0
30	2	2	2/6	6	[8]	8PSK	$\eta^*_2 = 0.80$	13.8	16.0

Figure 2.2 displays the simulated bit-error-rate performance of code #4 as well as two other 64-state transmit-diversity schemes over the 2×2 compound channel as a function of excess mutual information. The simulations are based on full ML-decoding on frames with 127 data symbols plus three symbols for trellis termination. With perfect channel state information at the receiver, the performance of the ML-decoder on unitary channels $\mathbf{H}$ does not depend on the particular instance of $\mathbf{H}$. However, on singular channels, the performance of the code will vary with the direction of the nonvanishing eigenvector of the channel matrix, except for the case of an orthogonal transmit design (such as Alamouti repetition [51]). For orthogonal transmit designs, the ML performance of the code depends only on the sum of the eigenvalues of the channel matrix.

At BER=10^{-5} code #4 requires no more than 0.88 bits of excess MI per transmit antenna on singular channels and requires 0.93 bits of excess MI on unitary channels. The rate-1/2 64-state maximal-free-distance convolutional code ([50]) with QPSK modulation (code #13 in Table 2.2) requires 0.84 bits of excess MI on unitary channels at BER=10^{-5}. At BER=10^{-5}, code #4 handles *every* instance of the 2×2 compound channel within 0.1 bits of excess MI per transmit antenna of the best rate-1/2 + QPSK AWGN-trellis code of similar complexity on the AWGN channel. The performance of the code over singular 2×2 channels *is* the performance of the code over 2×1 channels.

The 2 bits/symbol 64-state $\mathbb{Z}_4$-linear 4PSK-TCM of Aktas *et al.* [12](code #23 in Table 2.3) , designed to deliver good average error-probability under quasistatic Rayleigh fading channels, has $\eta_2^* = 1.0$. Over the singular channel

$$\mathbf{H} = \begin{bmatrix} 0.65 & -0.6 - 0.47i \\ 0 & 0 \end{bmatrix} \tag{2.23}$$

this code requires an excess MI of 1.10 bits per transmit antenna to achieve BER

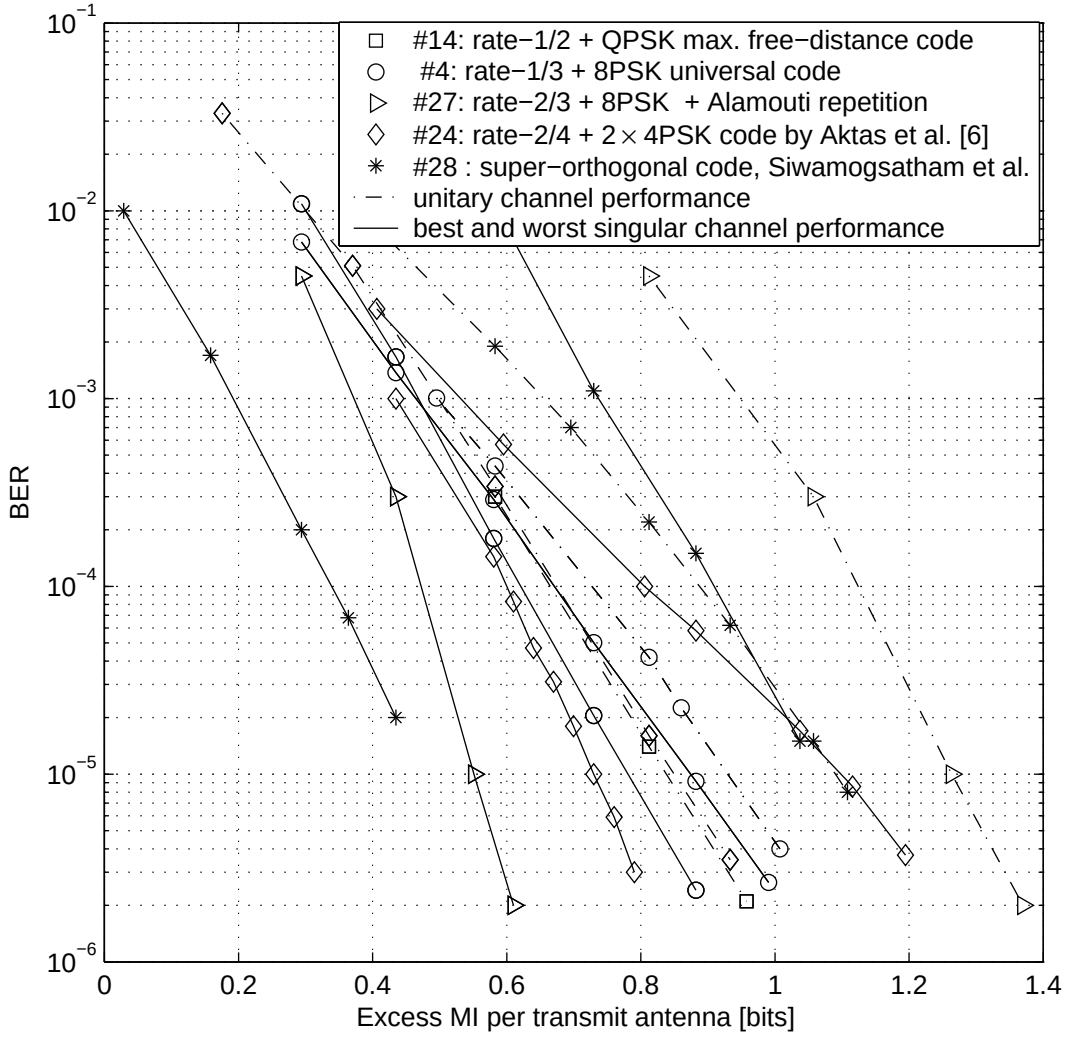

Figure 2.2: Channel-by-channel performance of the universal 64-state rate-1/3 + 8PSK code (#4) over the 2×2 compound channel. Best- and worst-case (for BER $= 10^{-5}$) singular channels are identified via extensive simulation. For comparison, the AWGN performance of the 64-state maximal free-distance rate-1/2 convolutional encoder (with QPSK) (code #14), as well as the compound channel performance of the 64-state rate-2/3 + 8PSK Ungerboeck-TCM + Alamouti block signaling (code #27) and that of the super-orthogonal code of Siwamogsatham and Fitz (code #28) are provided. Each frame consists of 127 data symbols and 3 symbols for trellis termination.

$= 10^{-5}$. The uncoded QPSK transmission on the AWGN channel has excess MI of 1.13 bits at BER $= 10^{-5}$. Code #23 has $\eta_2^* = 1.04$, with an average bit multiplicity of 22. In comparison, code #4 has $\eta_2^* = 0.73$, with average bit multiplicity of 1/32.

Another scheme that delivers 2 bits/symbol over two transmit antennas consists of a good AWGN-TCM followed by Alamouti repetition [51] (Code #27 in Table 2.3). The 64-state rate-2/3 Ungerboeck TCM [22] achieves BER $= 10^{-5}$ at SNR = 8.8 dB on the AWGN channel. On the compound 2×2 channel, the excess MI requirement of this scheme is a linear function of the sum of the channel eigenvalues [8]. On singular channels, this concatenated scheme requires only 0.55 bits of excess MI per antenna at BER $= 10^{-5}$, whereas on unitary channels the excess MI is requirement is 1.26 bits per transmit antenna. Among the three codes examined, code #4 has the most consistent channel-by-channel performance.

We make a final comparison of trellis code #4 with the 64-state universal space-time trellis code of [8] found by an exhaustive search over a larger class of encoders by using a channel-by-channel estimate of the excess MI requirement at the target bit error rate via the truncated union bound to the bit error probability. As shown in Fig. 2.3, over the 2×2 compound channel, the 64-state space-time code of [8] has a worst-case excess MI requirement of 0.95 bits per transmit antenna, 0.02 bit worse than code #4. The suboptimality of the more complex space-time trellis code was due to the search criterion employed in [8]. In particular, the truncated union bound failed to deliver a sufficiently accurate estimate of the excess mutual information requirement for singular channels under which worst-case error event multiplicities are typically low. The simplified encoder structure and search criterion used in this study, although approximate,

proved to produce better universal codes.

2.3.2 Universal codes for $N_t = 3$ and 4 transmit antennas

For $R = 2$ bits/symbol transmission over $N_t = 3$ transmit antennas, our search focused on rate-1/5 convolutional encoders. Two information bits produce ten codeword bits, one of which is punctured, and the remaining nine are used to map 3×8PSK constellations. An exhaustive search over 16-state encoders and all 1-out-of-10 puncturing patterns resulted in code #7 in Table 2.1 (b_9 is punctured). Code #7 has $\Delta_{H,\min} = 3$ and $\eta_3^* = 0.14$, whereas the 16-state 4PSK STTCM of [12] (code #24 in Table 2.3) has $\Delta_{H,\min} = 2$ and $\eta_2^* = 0.91$. Code #7 state achieves BER $= 10^{-5}$ with a worst-case excess MI of 1.3 bits per antenna. Code #24 cannot perform realiably under certain singular channels due to its limited diversity.

The Δ_E^*-constrained optimal-η_3^* search over 32-state encoders and all puncturing patterns resulted in an encoder that uses only four out of the eight points. This encoder may simply be represented as a 32-state rate-1/3 convolutional code driving QPSK (code #8).

Figure 2.4 displays the bit-error rate performances of code #8 and the 32-state code of [12] (code #25 in Table 2.3) on the compound 3×3 channel. Code #8 ($\Delta_H^* = 3$, $\eta_3^* = 0.24$, $\Delta_E^* = 22$) achieves BER $= 10^{-5}$ with a worst-case excess MI of 0.85 bits per antenna, code #25 requires about 0.95 bits of excess MI on the worst-case. Section 2.3.5 compares the *average* error performance of these codes under quasistatic Rayleigh fading.

A comparison the $R = 2$ bits per symbol schemes in Table 2.3 shows that the worst-case minimum eigenvalues of the full-diversity $N_t = 3$ schemes are much smaller than that of the full-diversity-schemes for $N_t = 2$ although the the Δ_E^*

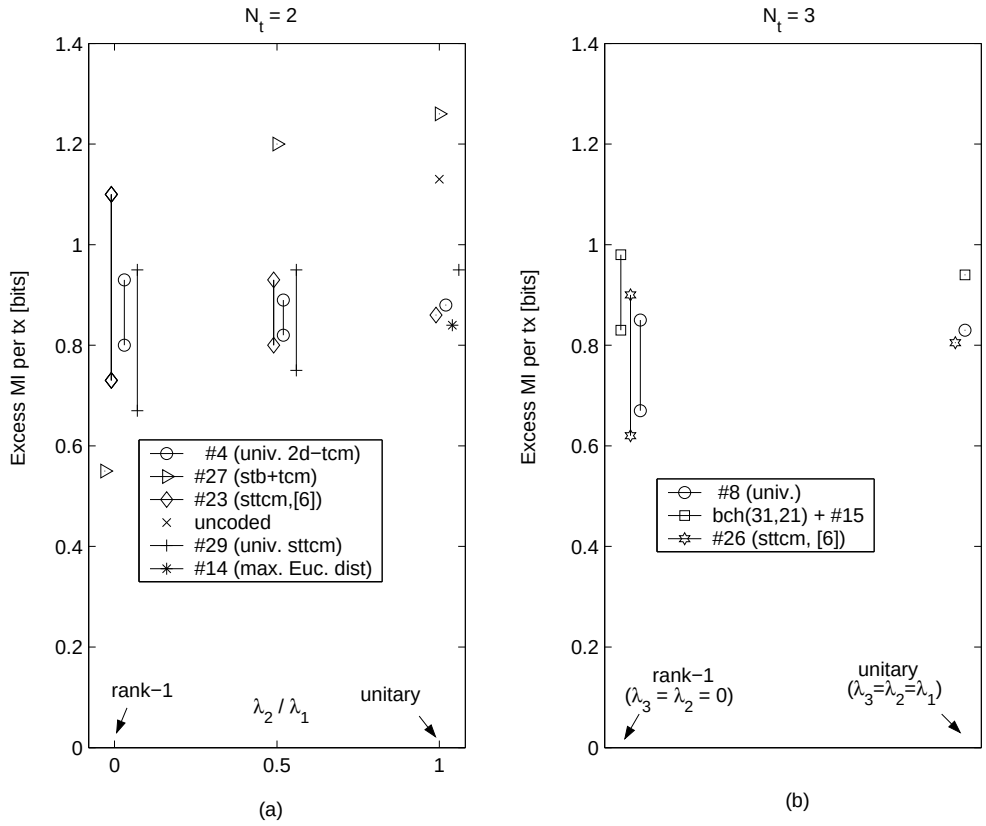

Figure 2.3: Excess mutual information requirement of $R = 2$ bits per symbol schemes $N_t = 2$ and $N_t = 3$ codes as a function of channel eigenvalue skew. Simulated target BER is 10^{-5} with maximum likelihood decoding on 127-data-symbol frames with trellis termination. Codes: *(a)* 64-state Rate-2/3 + 8-PSK trellis-coded modulation followed by Alamouti repetition (code #27) *(b)* 64-state 4-PSK code of Aktas *et al.* [6] (code #23), *(c)* 64-state 8-PSK multiplexed universal trellis code (code #4 of Table 1), *(d)* 64-state 8-PSK code from [20], *(e)* uncoded QPSK modulation, *(f)* 64-state rate-1/2 maximal free-distance convolutional encoder [26] + QPSK (code #14), *(g)* concatenation of the (31,21) BCH code and code #17 of Table 1 with soft-decision decoding via and max-log-APP, *(h)* 32-state 4-PSK code of Aktas *et al.* [6] (code #25), *(i)* 32-state universal 8-PSK trellis code (code #8 of Table 1).

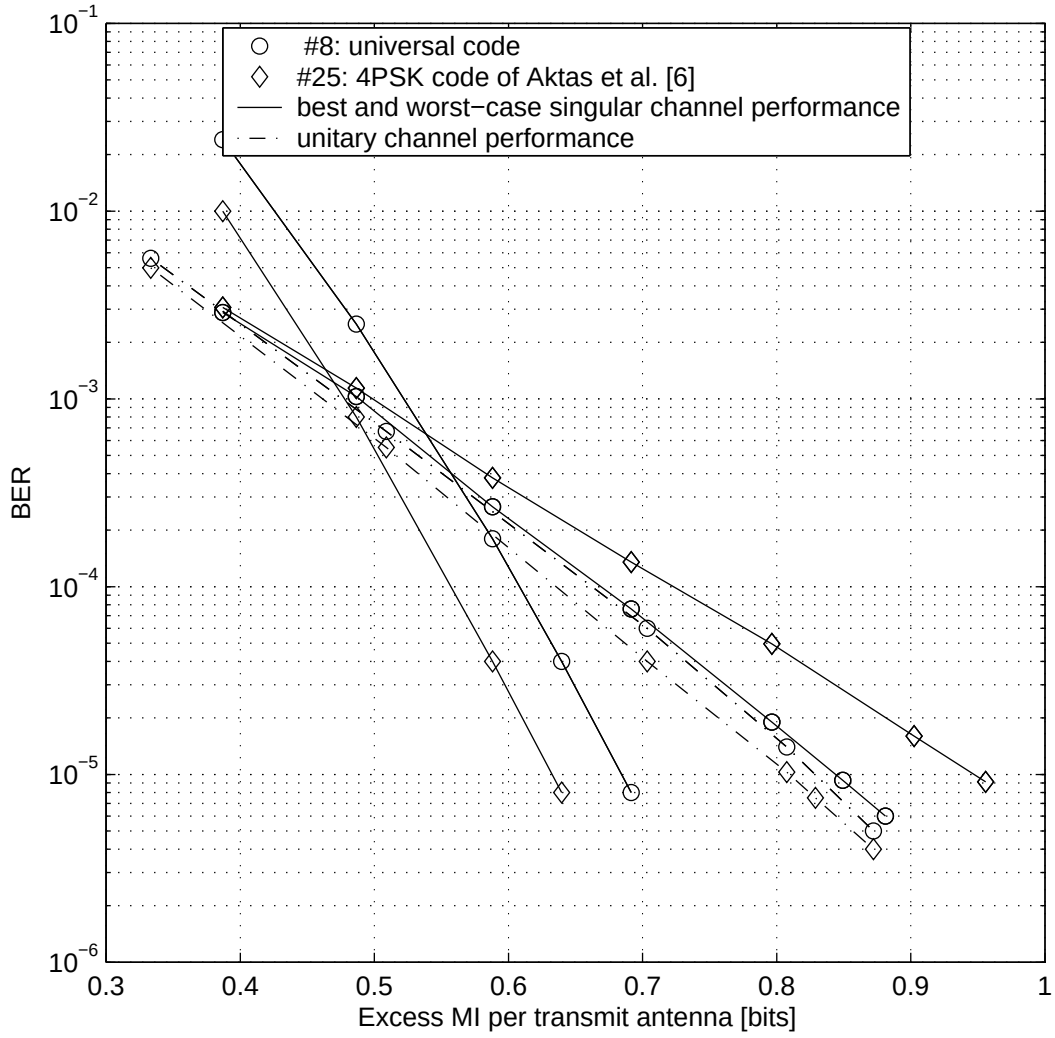

Figure 2.4: Channel-by-channel performance of the universal $R = 2$ bits per symbol QPSK code (#8) over the 3×3 compound channel. Best- and worst-case (for BER $= 10^{-5}$) singular channels are identified via extensive simulation. For comparison, the worst-case singular channel performance of the 32-state space-time code of Aktas *et al.* (code #25) is provided. The space-time code was proposed for average Rayleigh-fading performance. Each frame consists of 127 data symbols and 3 symbols for trellis termination.

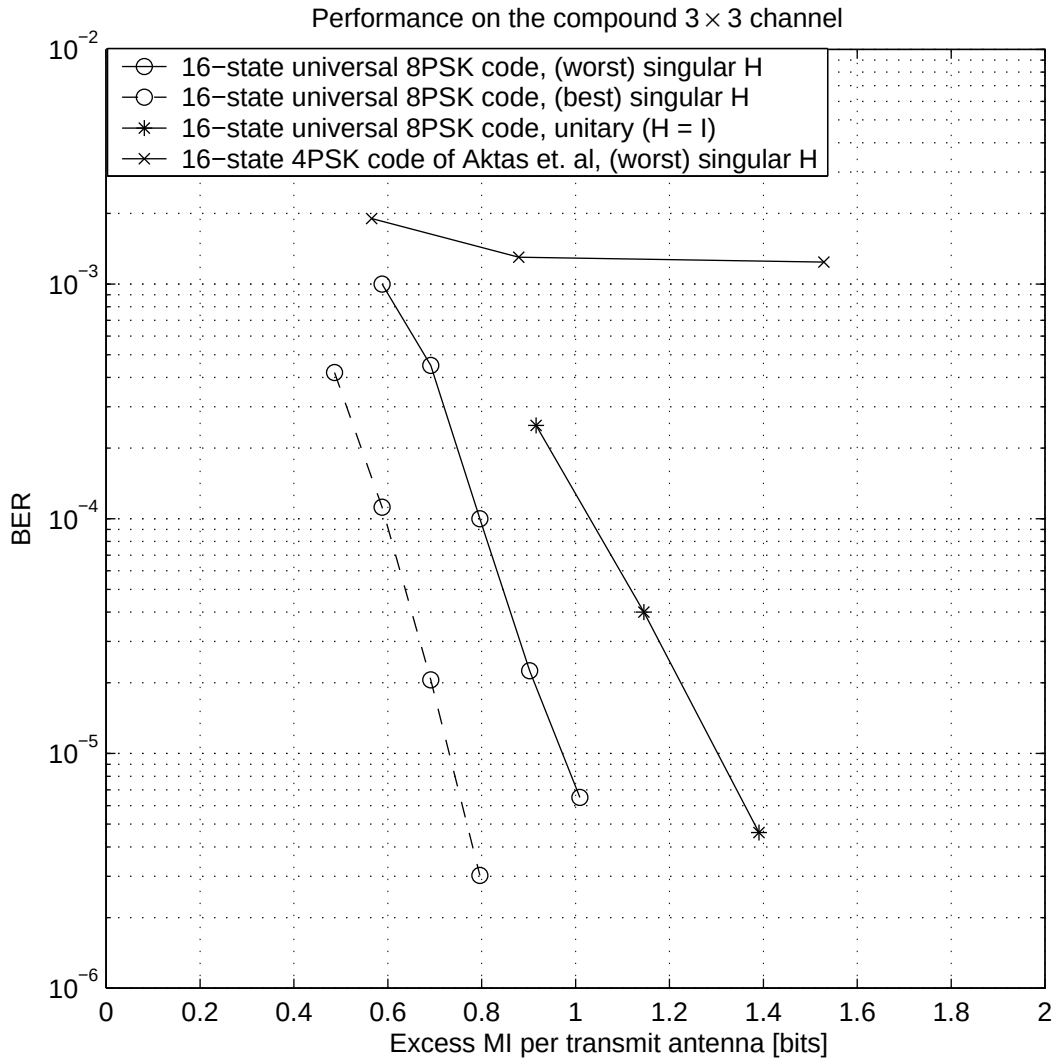

Figure 2.5: Channel-by-channel performance of the universal $R = 2$ bits per symbol QPSK code (#7) over the 3×3 compound channel. Best- and worst-case (for BER $= 10^{-5}$) singular channels are identified via extensive simulation. For comparison, the worst-case singular channel performance of the 16-state space-time code of Aktas *et al.* (code #24) is provided. The space-time code was proposed for average Rayleigh-fading performance. Each frame consists of 127 data symbols and 3 symbols for trellis termination.

figures are somewhat comparable. This is an inherent property of multiplexed linear trellis codes; at the same trellis complexity, the smallest eigenvalue decreases rapidly with increasing rates or with increasing N_t, but increases much more slowly with increasing trellis complexity. For this reason, trellis-code searches based on good-Δ_P^* for $N_t = 3$ and $N_t = 4$ often yield $\Delta_H^* = 2$ (which implies $\eta_i^* = 0$ for $i \geq 3$) for up to 64-states where typically $\eta_i^* < 1$ for $i \geq 3$ (for example, [67]).

For $R = 3$ bits per symbol transmission over $N_t = 3$ antennas, the binary encoder for the multiplexed trellis code needs to have at least 64-states for full diversity. Our search for a $R = 3$ bits per symbol universal TCM for $N_t = 3$ consisting of a 64-state rate-1/4 encoder mapping a 16QAM constellation failed to produce a full-diversity code. At 128 states, we were unable to complete the search. However, with $\Delta_H^* = 2$, a $R = 3$-bits per symbol maximal-η_2^* scheme was found as a result of an exhaustive search over 32-state rate-1/2 + QPSK codes (code #9). The parameters of the 32-state maximal-free-distance rate-1/2 encoder + QPSK code [50] (code #18) is provided for comparison. Code #9 is not universal over the 3×3 compound channel because it does not have full transmit diversity. However, it exhibits consistent performance over all channels with strong λ_2 presence.

For $N_t = 4$, code #10 provides maximal-η_2^* among its class of 16-state rate-1/4 encoders mapping two QPSK points for each information bit. Code #19 is the maximal-Δ_E^* code in this class [50]. For $R = 2$ bits per symbol over $N_t = 4$ transmit antennas, three levels of transmit diversity within this class of encoders becomes possible at 64-states. Code #11 is the maximal-η_3^* as well as the maximal-Δ_E^* among 32-state rate-1/4 encoders mapping two QPSK points for each information bit. This 32-state code has $\eta_3^* = 0.29$, $\Delta_P^* = 49.1$ and $\Delta_E^* = 36$.

We compare code #11 with the $R = 2$ bits per symbol 32-state vector-labeled 4PSK trellis-code of [67] (code #26 in Table 2.3) which was proposed for good average Rayleigh fading performance as a result of a maximum-Δ_E^*-search over a class of rate-2/8 encoders mapping $4 \times$ 4PSK constellations. This code, too, has $\Delta_E^* = 36$. Our analysis has indicated that it achieves 3 levels of transmit diversity with $\eta_3^* = 0.14$ and $\Delta_P^* = 34.6$. We were not able to simulate both codes extensively on 4×4 channels. However, code #10 has more favorable worst-case eigenvalues with $\eta_3^* = 0.28$, $(\eta_2\eta_3)^* = 2.17$ and $(\eta_2)^* = 4.27$, as compared to $\eta_3^* = 0.14$, $(\eta_2\eta_3)^* = 1.09$ and $(\eta_2)^* = 3.38$ of the 32-state code of [67]. Based on these figures, we might predict that at high SNR, code #10 will have a better worst-case performance on as compared to the latter code, over all reduced-rank channels that establish two strong spatial eigenmodes.

2.3.3 Universal codes by concatenation of algebraic block codes and rank-deficient TCMs with good eigenvalue spread

When the required trellis complexity for full-diversity exceeds practical limits, the concatenation of a rank-deficient trellis code with an outer code may restore the diversity order at the expense of reduced transmission rates.

Our first example is the (31,21) binary BCH code concatenated with the 64-state rate-1/4 + 16QAM scheme (code #17) to deliver $R = 2.03$ bits per symbol over three transmit antennas. The decoding is achieved by a Max-Log-APP BCJR [33] followed by a Max-Log-APP decoding of the BCH code. The interleaver is a 372-bit long block-interleaver hosting 12 codewords from the outer code in a 124-data-symbol packet with trellis termination. At BER $= 10^{-5}$, this scheme requires no more than 1 bit of excess MI on the collection of forty channels we simulated.

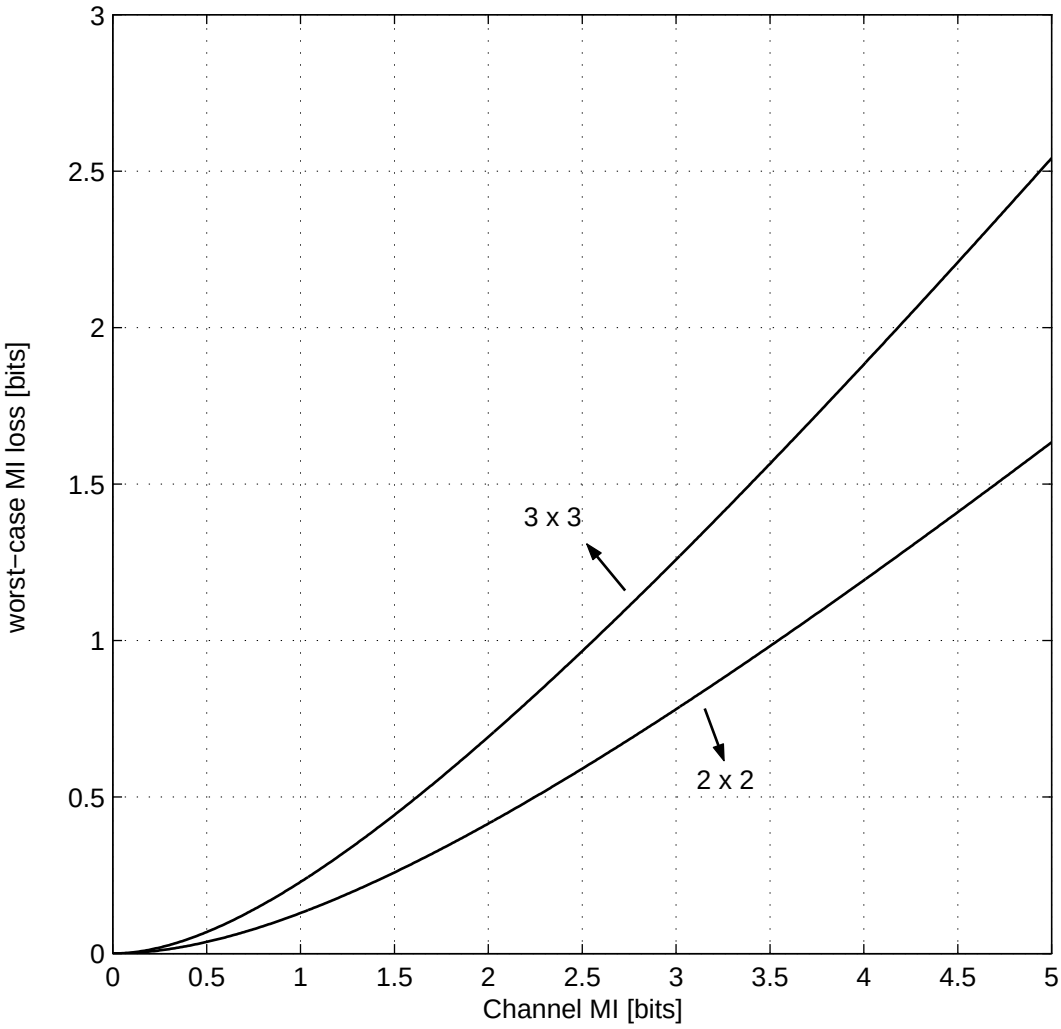

Figure 2.6: Worst-case mutual information loss of orthogonal space-time block codes over the compound channel as a function of channel mutual information. Alamouti repetition [51] for two antennas and the orthogonal block scheme of [6] for three transmit antennas.

Similar consistent channel-by-channel performance was observed with the concatenation of the (63,45) binary BCH code with code #17 to deliver $R = 2.75$ bits per symbol over $N_t = 4$ transmit antennas. For this scenario, we used a is a 512-bit block-interleaver interleaving 8 BCH codewords over a 129-data-symbol packet with trellis termination. This scheme required an excess MI of no more than 1.2 bits per transmit antenna to achieve BER $= 10^{-5}$ over a collection of thirty-two rank-1 channels and the 4×4 unitary channel.

We were unable to run a sufficient number (on the order of hundreds) of test channels to conclude that the concatenated approach yields universal codes for the compound channel. However, the promising simulation results so far encourage us to pursue an analytical understanding of the compound channel performance of concatenated schemes.

For the target bit error rate of BER=10^{-5}, the excess MI requirement of proposed trellis codes for two and three transmit antennas over the compound channel is illustrated in Fig. 2.3a and Fig. 2.3b respectively. Fig. 2.3a shows simulation-based excess MI figures for a dense sampling of 2×2 singular channels, unitary channels as well as 2×2 channels with eigenvalue skew equal to 1/2. The most consistent excess MI performance is obtained by the multiplexed universal code #4. The universal code of [8] performs better on some singular channels but requires more MI on unitary channels. The code of Aktas *et al.* performs well on unitary channels, but has a wide range of performance on singular channels. Its worst-case performance on a singular channel requires the most MI of any code studied here. In contrast, the rate-2/3 8-PSK TCM with Alamouti repetition requires the least MI on singular channels but performs noticeably worse than all the other codes on unitary channels. This poor performance on unitary channels is an unavoidable consequence of orthogonal space-time block designs

with multiple receive antennas. Subsection 2.3.4 explores this phenomenon more carefully.

Over singular channels, the TCM followed by an Alamouti repetition (marked with (a)) achieves the target BER with the smallest excess MI requirement among the codes displayed. The 4PSK code of [12] (marked with (b)) has a worst-case performance over singular channels that is similar to the performance of uncoded 4PSK over the AWGN channel (marked with (e)). At the other end of the eigenvalue skew spectrum, code #4 and the 4PSK code of [12] deliver good performance over unitary channels, code #4 requiring 0.02 bit more than the 4PSK code, but still 0.04 bits within that of the best 64-state rate-1/2 + QPSK AWGN trellis code (marked with (f)). For all the codes displayed, the excess MI requirement over channels with eigenvalue skew equal to 1/2 lies in between the cases for singular channels and unitary channels.

Figure 2.3b summarizes the excess MI requirement of several full-diversity schemes for $N_t = 3$ transmit antennas over a collection of 3×3 rank-1 channels as well as 3×3 unitary channels. Over the collection of channels simulated, code #8 required a maximum of 0.85 bits of excess MI per transmit antenna to achieve BER $= 10^{-5}$, whereas the 4PSK code of [12] required 0.95 bits of excess MI per transmit antenna in the worst-case.

2.3.4 Mutual information penalty of orthogonal space-time block codes

Now consider the MI penalty incurred by using an orthogonal space-time block (STB) scheme for transmit diversity. Racall that the channel MI is $\mathrm{MI}(\mathbf{H}) = \log_2 \det\left(\mathbf{I} + \mathbf{H}\mathbf{H}^\dagger\right)$ bits per symbol. For two transmit antennas, the constrained mutual information of the Alamouti scheme [51] is given by

$$\left(\mathrm{MI}(\mathbf{H})\right)_{\text{Alamouti}} = \log_2\left(1 + \mathrm{trace}(\mathbf{H}\mathbf{H}^\dagger)\right) \quad \text{bits.} \tag{2.24}$$

On 2×2 unitary channels $\mathbf{H}$ with $\mathrm{MI}(\mathbf{H})$, the MI loss of the Alamouti scheme is

$$R - \log_2\left(1 + 2(2^{R/2} - 1)\right) \quad \text{bits.} \tag{2.25}$$

For three transmit antennas, the orthogonal linear dispersion (LD) scheme of [6] achieves a constrained MI of

$$\left(\mathrm{MI}(\mathbf{H})\right)_{\text{Hassibi-LD}} = \frac{3}{4}\log_2\left(1 + \frac{4}{3}\mathrm{trace}(\mathbf{H}\mathbf{H}^\dagger)\right) \quad \text{bits.} \tag{2.26}$$

On 3×3 unitary channels, the MI loss of this scheme is therefore

$$R - \frac{3}{4}\log_2\left(1 + 4(2^{R/3} - 1)\right) \quad \text{bits.} \tag{2.27}$$

Figure 2.6 shows the worst-case mutual information loss over the compound channel of these two orthogonal STB schemes as a function of the channel mutual information. Both of these schemes experience heavy MI penalties for channels with equal eigenvalues that support high data rates. Universal trellis codes have superior worst-case compound channel performance as compared to orthogonal block schemes.

2.3.5 Universal space-time trellis codes under quasistatic Rayleigh fading

Universal space-time trellis codes deliver good average error performance under quasistatic Rayleigh fading as long as the quasistatic period is longer than several traceback depths of the codes. Figure 2.7 compares the frame-error-rate (FER) and the bit-error rate (BER) performances of code #4 the 64-state code of [12] (code #23) over 2×1 and 2×2 quasistatic Rayleigh fading with quasistatic duration of 130 channel symbols, three of which are used for trellis termination. The universal code has slightly better FER and BER over the SNR range displayed. The code performs 1.7 dB away from outage capacity.

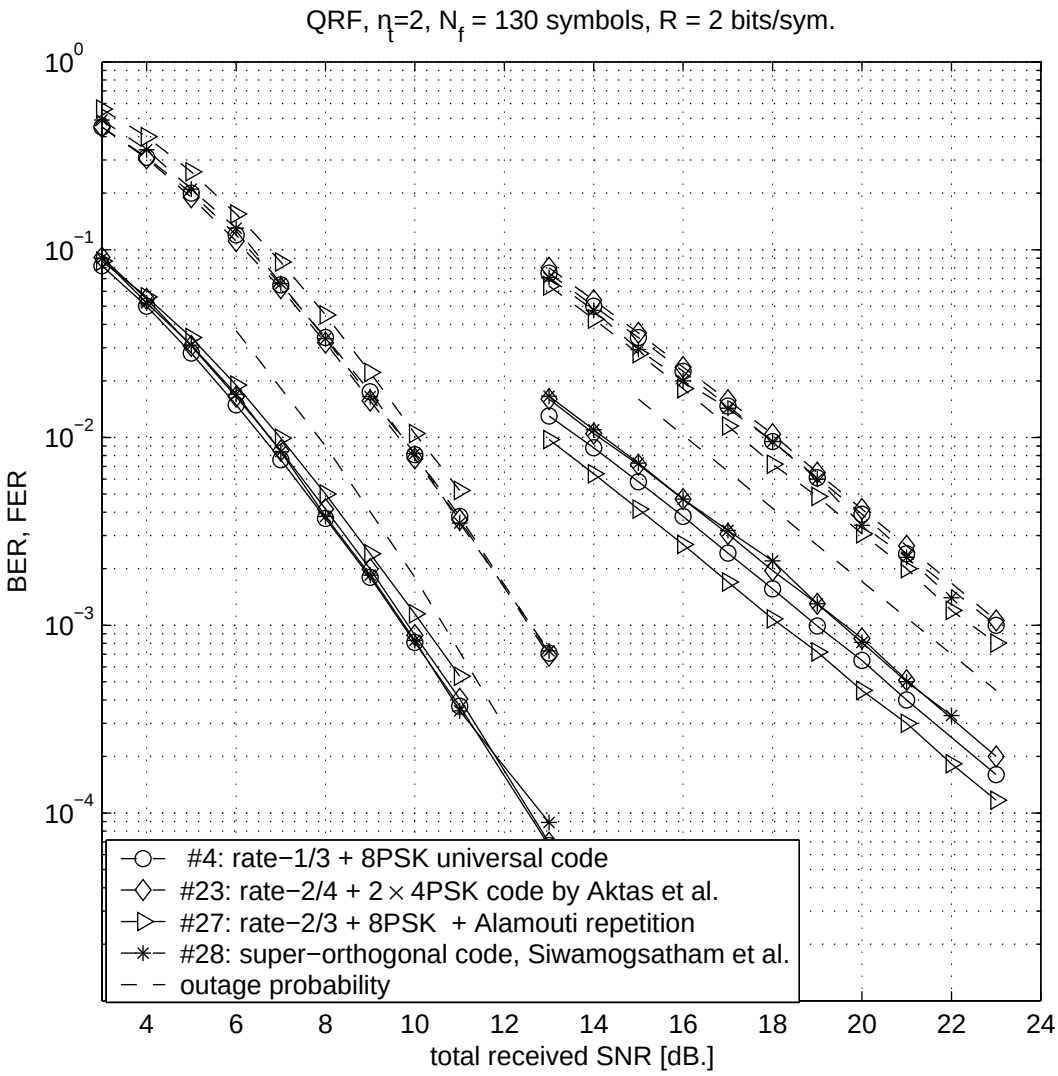

Figure 2.7: Bit-error-rate and frame-error-rate performance of the 64-state universal TCM (code #4) as compared to 64-state code of Aktas et. al, over the quasistatic Rayleigh fading channel, $N_t = 2$. Each frame consists of 127 data symbols and three symbols for trellis termination. Maximum-likelihood decoding on the entire frame.

Figure 2.8 compares the FER and BER performances of 32-state code #8 and the 32-state code of [12] (code #25) over 3×1 and 3×3 quasistatic Rayleigh fading with quasistatic duration of 130 channel symbols, three of which are used for trellis termination. The two codes have similar BER and FER performance over the range of SNRs displayed, with code #8 performing slightly worse in FER for the 3×3 scenario.

Finally, Figure 2.9 compares the FER and BER performances of code #11 and the 32-state code of [67]. Code #11 has approximately the same frame-error-rate and bit-error-rate performance with the 32-state code of [67] (code #26) over 4×2, 4×4 quasistatic Rayleigh fading channels, and slightly better bit-error-rate performance with similar frame-error-rate performance over 4×1 quasistatic Rayleigh fading.

Figure 2.10 displays the probability distribution of the ratio λ_1/λ_N in $N \times N$ Rayleigh fading for $N = 2, 3$. For two transmit antennas, the probability that the eigenvalues λ_1, λ_2 are more than 10 dB apart is 0.45. For $N = 3$, this probability is 0.83. Universal code design which takes into account the performance over singular channels (through high η_N) results in good codes for the average Rayleigh fading performance.

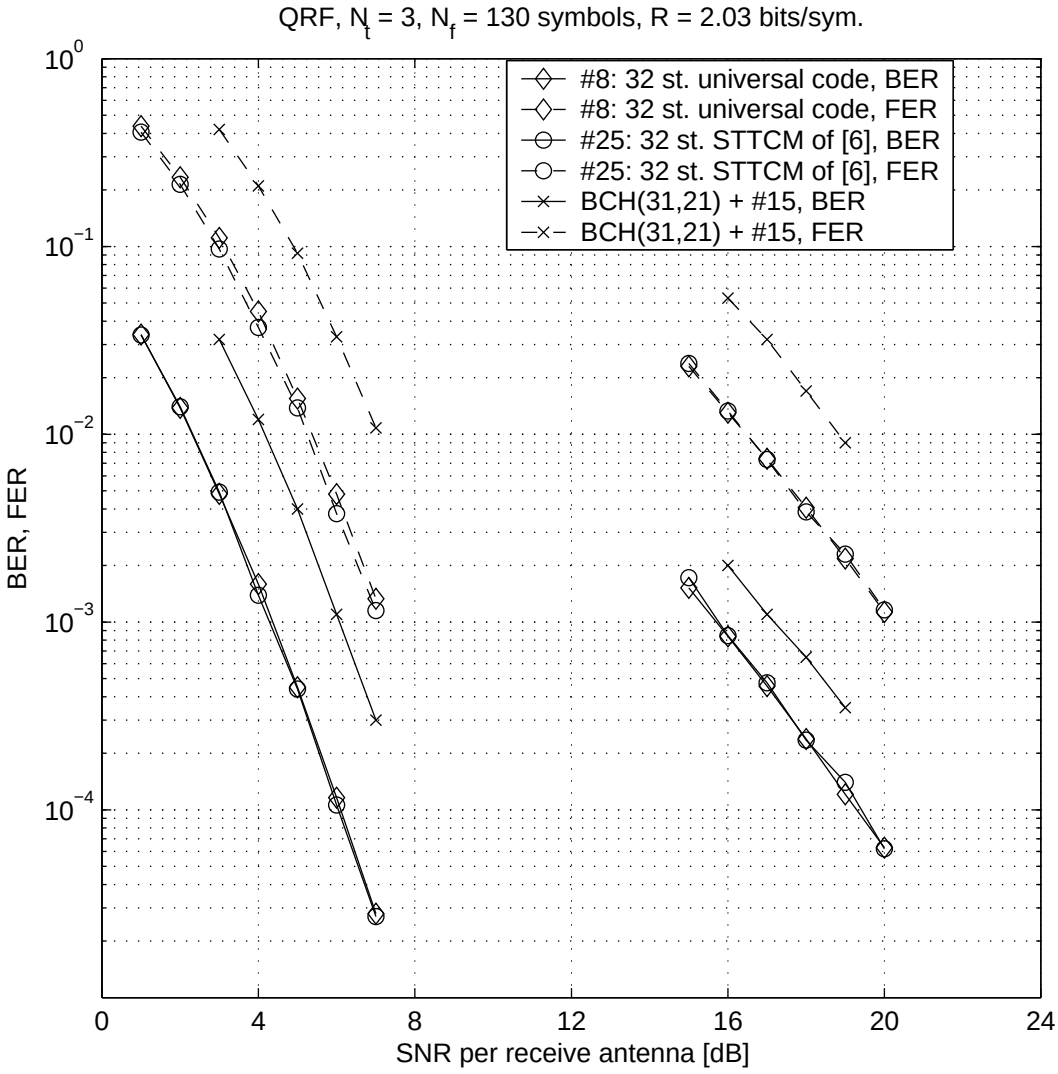

Figure 2.8: Bit-error-rate and frame-error-rate performance of the 32-state universal TCM (code #7) as compared to 32-state code of Aktas *et. al* (code #25), over the 3×1 and 3×3 quasistatic Rayleigh fading channels. Each frame consists of 127 data symbols and 3 symbols for trellis termination. Maximum-likelihood decoding on the entire frame.

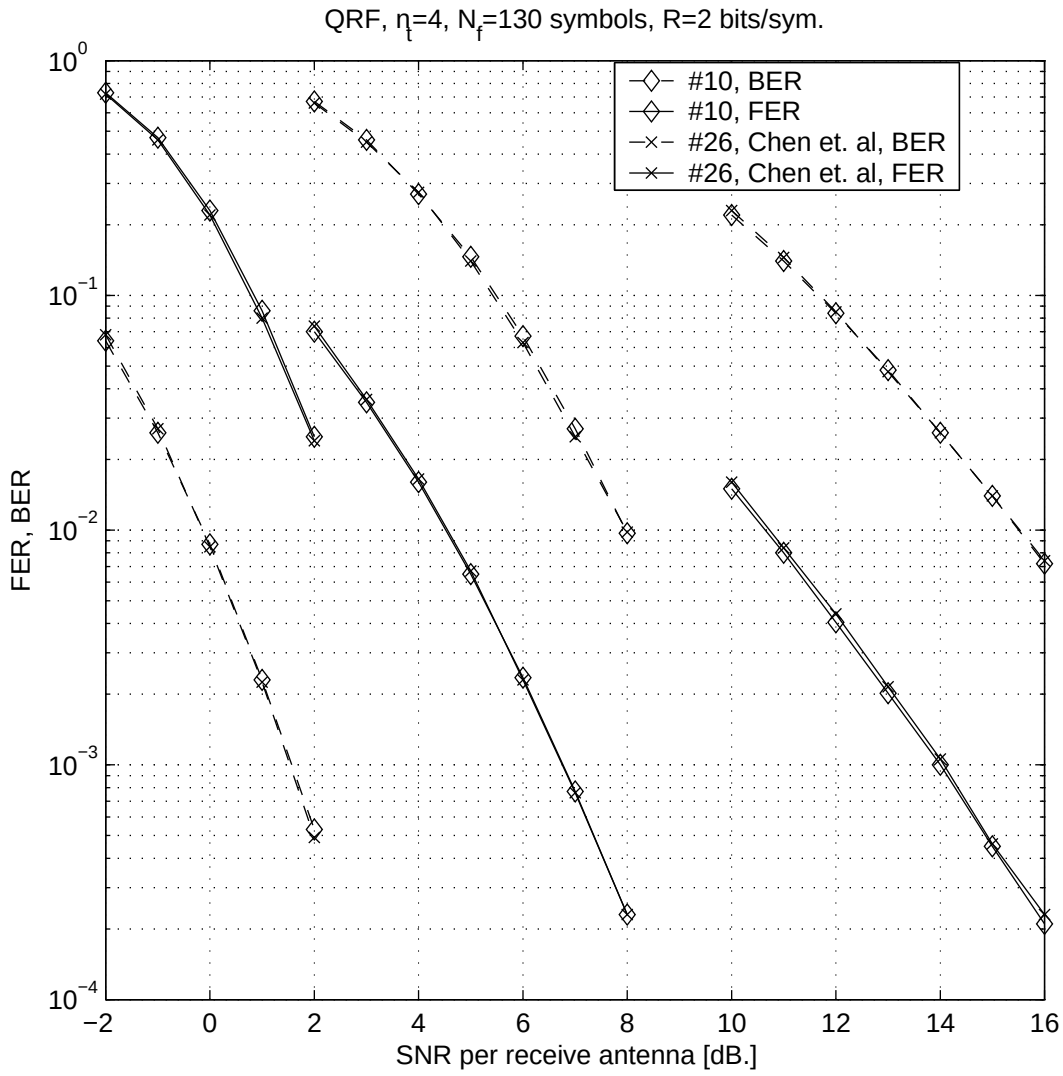

Figure 2.9: Bit-error-rate and frame-error-rate performance of the 32-state code (#11) as compared to 32-state code of Chen *et. al* (code #26), over the with 4×1, 4×2 and 4×4 quasistatic Rayleigh fading channel. Each frame consists of 127 data symbols and 3 symbols for trellis termination. Maximum-likelihood decoding on the entire frame.

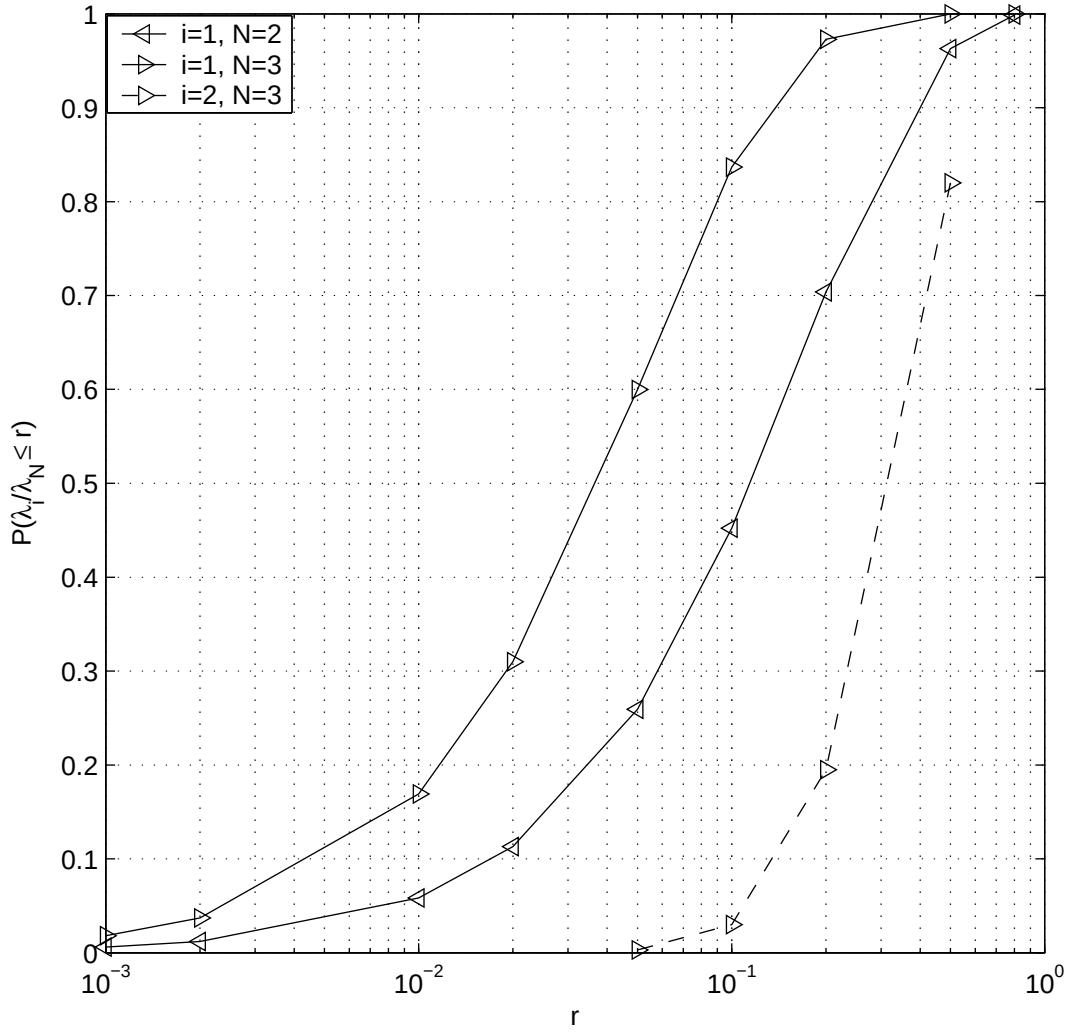

Figure 2.10: Distribution of the ratio of eigenvalues of $\mathbf{H}\mathbf{H}^\dagger$ where $\mathbf{H}$ is an $N \times N$ matrix of independent and identically ditributed complex Gaussian random variables. Top to bottom: Cumulative distribution function (CDF) of λ_3/λ_1 in 3×3 Rayleigh fading, CDF of λ_2/λ_1 in 2×2 Rayleigh fading, CDF of λ_2/λ_1 in 3×3 Rayleigh fading.

CHAPTER 3

Sercially concatenated trellis-coded modulations for periodic erasures

3.1 Introduction

The compound channel theorem [64] for linear Gaussian channels indicates that a single code can reliably transmit information at R bits/symbol on each channel in an ensemble of linear Gaussian channels with capacity $C > R$. The existence of codes that perform consistently close to channel capacity over a family of channels has become very significant with the introduction of turbo codes that operate within tenths of a dB of the capacity on the AWGN channel [7]. Codes that extend this performance to other channels of practical interest are highly desirable.

The design of turbo codes for the compound channel is motivated by the earlier work of Wesel, Liu and Shi [46] on robust trellis codes for channels with periodically varying signal-to-noise ratios (SNRs). Recently, McEliece [43] raised the question of how turbo-like codes would perform on nonstandard channels. This paper revisits ideas of [46] to design robust trellis codes for use as inner codes in a serially-concatenated scenario to provide consistent performance over the periodic-SNR channels. Such channels appear in frequency-hopped or multi-carrier transmissions as well as diagonally-layered space-time architectures.

A period-2 periodic-SNR channel is specified by

$$y_t = \tilde{a}_{(t \bmod 2)} \sqrt{E_s}\, x_t + z_t \tag{3.1}$$

where z_t is additive white Gaussian noise (AWGN) with variance $N_0/2$ per dimension, and the vector

$$\tilde{\mathbf{a}} = [\tilde{a}_0, \;\; \tilde{a}_1] \tag{3.2}$$

describes the nature of the time-varying attenuation behavior of the channel. The average transmitted signal energy is E_s.

With an equal allocation of power to each symbol in the period, it is possible to transmit R bits per two-dimensional symbol over all period-2 channels such that

$$\frac{1}{2} \sum_{i=0}^{1} \log_2 \left(1 + \frac{|\tilde{a}_i|^2 E_s}{N_0} \right) > R \tag{3.3}$$

where the left-hand side of (3.3) is the channel mutual information (MI).

For each channel in the ensemble, the required E_s/N_0 to achieve a certain bit-error rate (BER), say 10^{-5}, gives a mutual information figure. The difference between this MI and the transmitted rate, R, is the *excess MI*. The excess MI describes how far the code is operating from the capacity of the channel. At a fixed BER, the largest of the excess MI figures over the ensemble of channels determines the compound channel performance of the code.

The *compound period-2 periodic-erasures channel* consists of two instances of the attenuation vector: When $a_0 = 1$, $a_1 = 0$, every other symbol is erased, and when $a_0 = 1$, $a_1 = 1$, there are no erasures, the channel is AWGN.

For example, a 64-state code of [46] (Code 1) requires $E_s/N_0 = 4.3$ dB to achieve BER $= 10^{-5}$ on the AWGN channel. This corresponds to an excess MI

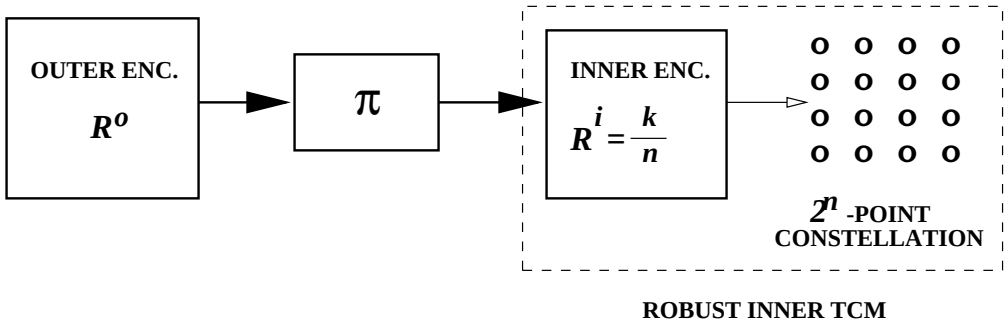

Figure 3.1: The SCTCM scheme.

of 0.88 bits. The same code requires $E_s/N_0 = 9.4$ dB at BER $= 10^{-5}$ when every other symbol is erased. The excess MI under periodic erasures is 0.64 bits. Therefore, the code requires at most 0.88 bits of excess MI to achieve BER $= 10^{-5}$ on the compound periodic-erasures channel.

With turbo codes it is possible to achieve lower excess MI figures at similar BERs. The serially-concatenated scheme is especially attractive in designing for the compound channel because robustness can be imposed on the inner code which directly interfaces the channel. Developing on this idea, Section II extends the existing AWGN design rules for SCTCM schemes to periodic-erasure channels with particular emphasis on the inner TCM. In Section III we propose 4- and 8-state inner TCM schemes which are robust under periodic erasures with period two. Section IV provides experimental results for 0.5 bits/symbol, 1 bit/symbol, 1.5 bit/symbol robust SCTCM schemes using the robust inner TCMs. Section V describes the ongoing work and concludes the paper.

3.2 Code Design

Figure 3.1 shows a generic SCTCM scheme consisting of a rate-R^o outer convolutional encoder whose output, scrambled by an N-bit block interleaver, drives

the input of a rate-R^i inner convolutional encoder mapping a two-dimensional 2^n-point constellation. The throughput of the overall scheme is

$$nR^oR^i \qquad \text{bits per symbol.} \tag{3.4}$$

The uniform interleaver analysis of [48] has shown that the bit error rate under maximum-likelihood (ML) decoding decreases exponentially with the effective free Euclidean distance of the inner code. For large interleaver sizes of N bits

$$-\log \text{BER} \propto \lfloor (d_{\text{f}}^o + 1)/2 \rfloor \log N + \delta^2 \frac{E_s}{4N_0} \tag{3.5}$$

where d_f^o is the free Hamming distance of the outer code, and

$$\delta^2 = \begin{cases} \frac{1}{2} d_{\text{f}}^o d_{\text{f, eff}}^2 & \text{for } d_{\text{f}}^o \text{ even} \\ \frac{1}{2}(d_{\text{f}}^o - 3) d_{\text{f, eff}}^2 + (h_m^{(3)})^2 & \text{for } d_{\text{f}}^o \text{ odd} \end{cases} \tag{3.6}$$

where $d_{\text{f,eff}}$ is the effective free Euclidean distance of the inner code, defined to be the minimum Euclidean distance of error events caused by two information bit errors, and $h_m^{(3)}$ is that distance of error events caused by three information bit errors.

The BER approximation (3.5) extends to the periodic erasures channel when the input to the channel is considered to be a vector of two consecutive symbols. Two necessary conditions for robust performance over the compound periodic erasures channel are:

1. The concatenated code should have positive redundancy under periodic erasures, i.e.,

$$2R^oR^i < 1. \tag{3.7}$$

   If the inequality in (3.7) is reversed transmission errors would be guaranteed when every other symbol is erased. If (3.7) were an equality, only uncoded performance would be possible under period-2 erasures.

2. The inner TCM should have nonzero effective free Euclidean distance under periodic erasures, in accordance with (3.5). For a given number of states, the rate of the inner encoder should be small enough to avoid parallel transitions:

   Lemma 3 *For $k > 1$, an S-state rate-k/n inner encoder mapping a 2^n-point constellation has* zero *minimum effective free Euclidean distance under periodic erasures if*

   $$\binom{k}{2} > S. \tag{3.8}$$

   Proof 6 *For $k > 1$, the encoder trellis has $\binom{k}{2}$ transitions per state to describe the inputs with Hamming weight 2. If $\binom{k}{2} > S$, there are parallel transitions between states and shortest error events are 1-symbol long.*

   For a 4-state encoder the maximum rate for nonzero effective free Euclidean distance under periodic erasures is bounded by 3/4, and for an 8-state encoder the bound is 4/5.

The design criteria consider both maximum likelihood (ML) decoding and iterative decoding. For ML decoding, the design criterion of SCTCM for large interleavers and very low bit error rates is to maximize the free Hamming distance of the outer code and to maximize the effective free Euclidean distance, $d_{\text{f,eff}}$, of the inner TCM. However, since iterative decoding is used in practice, Extrinsic Information Transfer (EXIT) charts proposed by ten Brink [56],[55], or similar techniques such as density evolution [15] are more useful in finding SCTCMs with low pinch-off thresholds. As will be demonstrated, the ML decoding criterion and iterative decoding criterion provide codes that perform very differently.

The proposed design method is based on the following steps:

- Design of the inner code: Although the inner TCM will have zero redundancy under periodic erasures for $R^i \geq 1/2$, it may still have nonzero effective free Euclidean distance. To achieve nonzero effective free Euclidean distance, the maximum rate of the inner encoder is limited by decoding complexity (Lemma 1).

  At a given decoding complexity, we search for the TCM schemes that have high $d_{\text{f, eff}}$ under both AWGN and periodic erasures. Among those codes having equal $d_{\text{f, eff}}$, we pick the ones with lowest number of error events caused by two information bit errors.

  From a ML decoding point of view, the aim is to find the highest rate inner encoders which have good effective free Euclidean distances under periodic erasures.

- Design of the outer encoder: The rate of the outer code is determined by the rate of the inner TCM and the desired bandwidth efficiency of the overall concatenated scheme. In accordance with (3.5), candidate outer encoders for a given rate are those with high free Hamming distances. For final selection, we use EXIT charts to estimate the pinch-off threshold under iterative decoding. The speed of convergence judged qualitatively from the distance between decoder characteristics (tunnel width) after the iterative decoding tunnel is also taken into consideration. Narrower tunnels require more iterations to achieve a given performance.

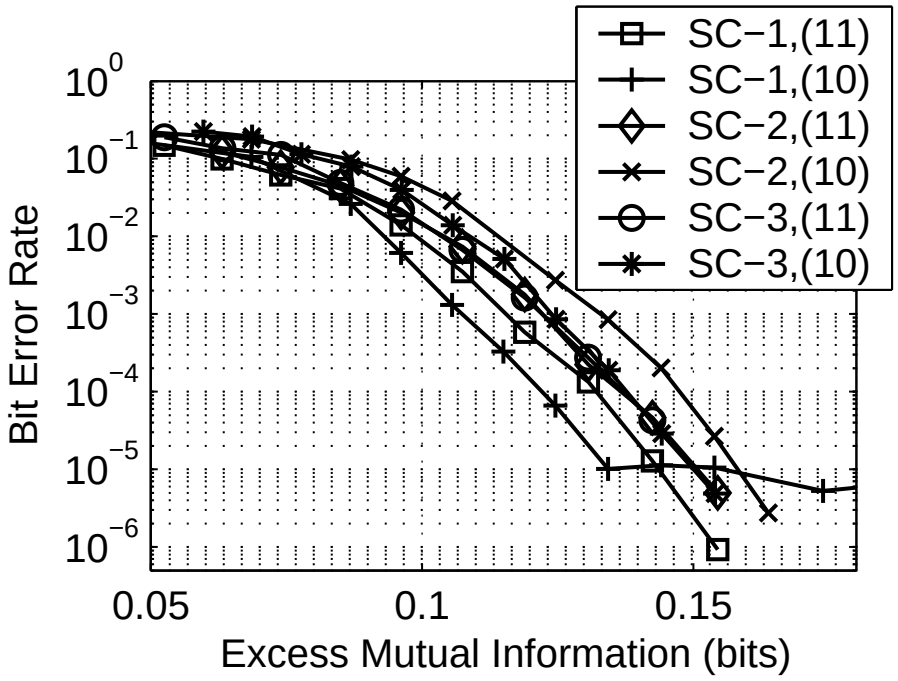

Figure 3.2: Simulations of BER versus Excess channel MI for 0.5 bits/symbol SCTCM's under both AWGN and period-2 erasures channel. See Table III for the SCTCM numbers. 11 and 10 represent period-2 erasures patterns. Block length=2,046, 12 iterations.

3.3 Example SCTCM Designs for Periodic Erasures

3.3.1 SCTCM Design of 0.5 bit per symbol

A rate-1/3 encoder mapping a Gray-labeled 8-PSK constellation provides 1 bit per symbol on the inner TCM. A rate-1/2 outer encoder then yields 0.5 bits/symbol overall throughput. The inner TCM still has redundancy under periodic erasures. The structure of a rate-1/3 linear systematic feedback encoder with 8 states is shown in Figure 3.4. An exhaustive search over the generator vectors yielded the maximal $d_{\text{f,eff}}$ encoder under period-2 erasures as well as under AWGN (C_1^i).

The scheme SC-1 (Table III) uses a 4-state maximum free-distance outer code with C_1^i. Figure 3.2 shows an error floor around BER=10^{-5} for SC-1, which is lowered down to BER=10^{-6} with 30 iterations. The long and narrow iterative decoding tunnel shown in Figure 3.3 requires an increased number of iterations.

A method to lower the error floor is to use an outer encoder with larger free distance. Using an 8-state maximal free-distance outer code (C_3^o) and 12 iterations, the error floor under periodic erasures is lowered (SC-3).

A better way to lower the error floor without increasing the overall complexity is to use EXIT chart analysis to find an 8-state inner code which has a shorter and wider iterative decoding tunnel than SC-1. Figure 3.2 shows that SC-2 which concatenates C_2^i and C_2^o is able to lower the error floor with only a slight increase in the pinch-off threshold. As a result, SC-2 requires at most 0.16 bits of excess MI at BER=10^{-5} to achieve BER=10^{-5} on AWGN as well as on periodic erasures.

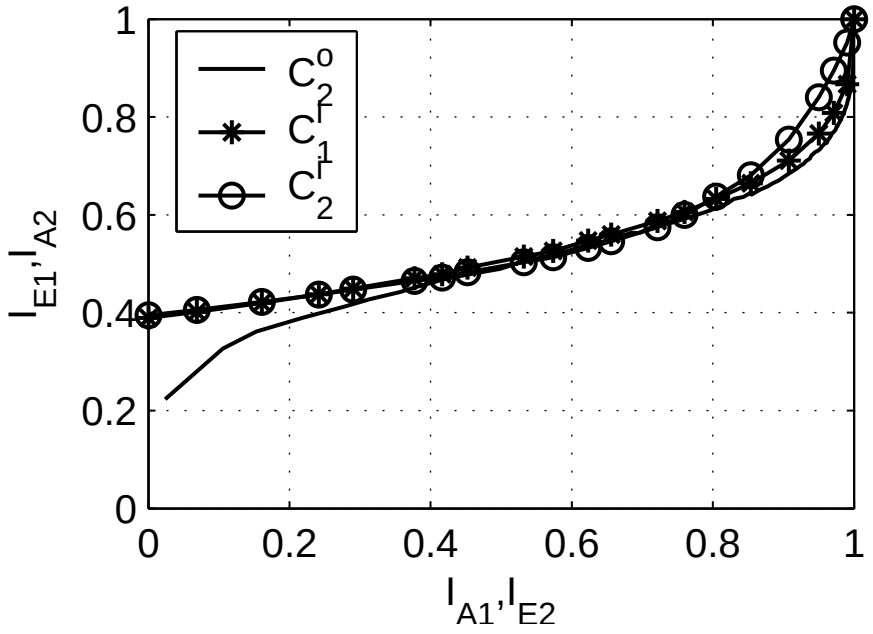

Figure 3.3: Extrinsic Information Transfer charts for some inner decoders at E_b/N_0=3.4 dB under period-2 erasures. I_{A1} is the a priori input to inner decoder and I_{E1} is the extrinsic and channel output of inner decoder. I_{A2} is the a priori input to outer decoder and I_{E2} is the extrinsic output of outer decoder.

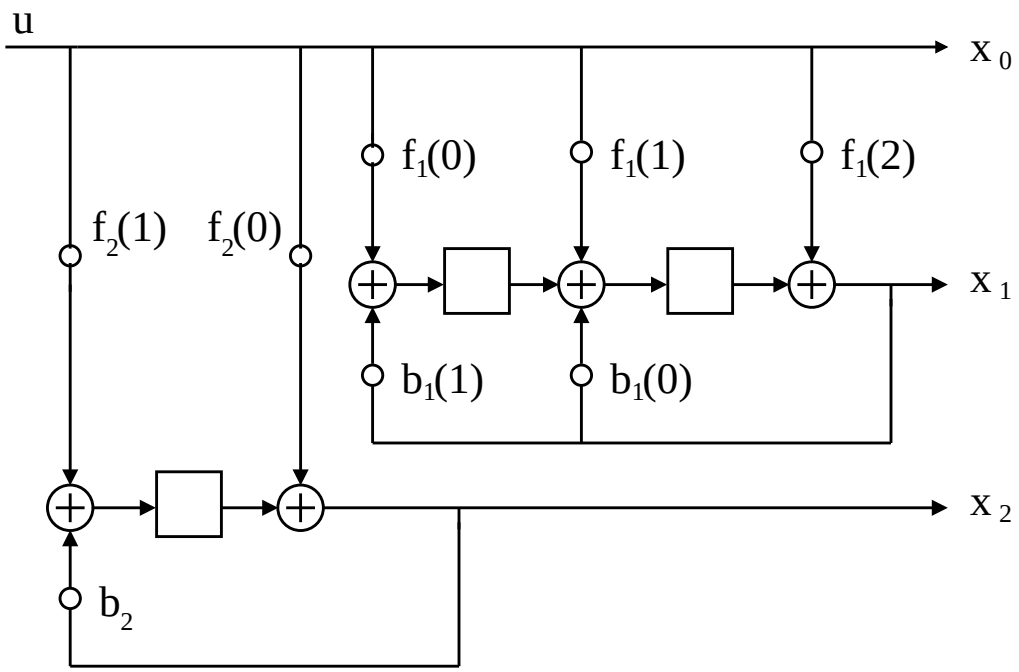

Figure 3.4: Structure of a rate-1/3 8-state recursive inner encoder with generator G(D)=[1 $\frac{f_1(2)+f_1(1)D+f_1(0)D^2}{1+b_1(0)D+b_1(1)D^2}$ $\frac{f_2(0)+f_2(1)D}{1+b_2D}$].

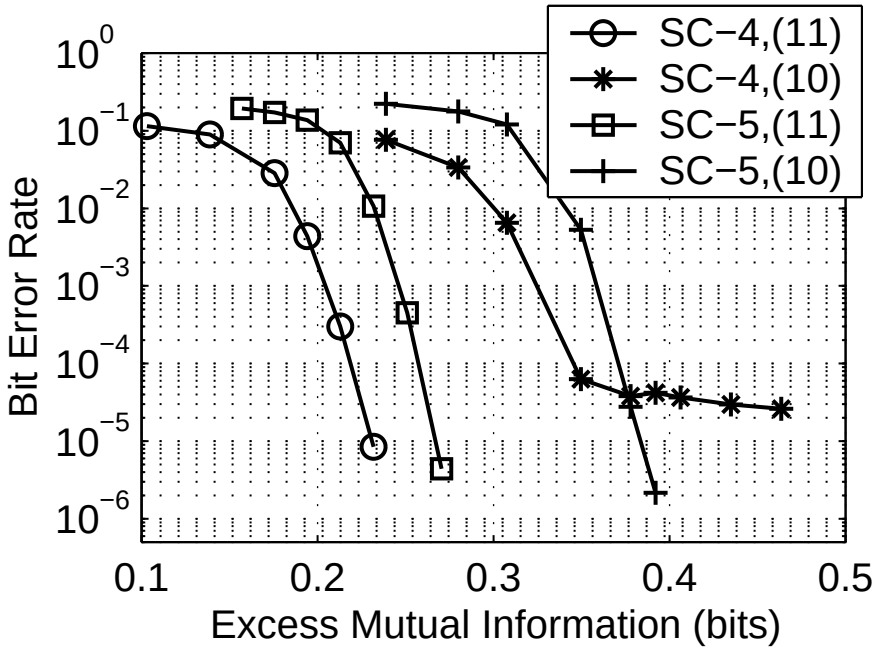

Figure 3.5: Simulations of BER versus Excess MI for two 1.0 bit/symbol SCTCMs which are only different in the complexity of the outer codes. Block length=10,000, 12 iterations.

3.3.2 SCTCM Design of 1.0 bit per symbol

The 1.0 bit per symbol SCTCM uses a rate-1/2 outer code and an inner TCM consisting of a rate-2/3 linear systematic recursive encoder driving an 8-PSK constellation. With a rate-2/3 inner encoder, the inner TCM has *negative redundancy* and therefore zero minimum distance under period-2 erasures. However, it is still possible to find an inner code with nonzero *effective* free distance.

Figure 3.5 shows the BER performance of two SCTCM schemes (SC-4, SC-5) using the same inner TCM (C_4^i driving a Gray-labeled 8-PSK) found by exhaustive search to maximize $d_{\mathrm{f,\,eff}}$ under periodic erasures. In comparison with SC-4, SC-5 trades pinch-off threshold for lower error floor under periodic erasures.

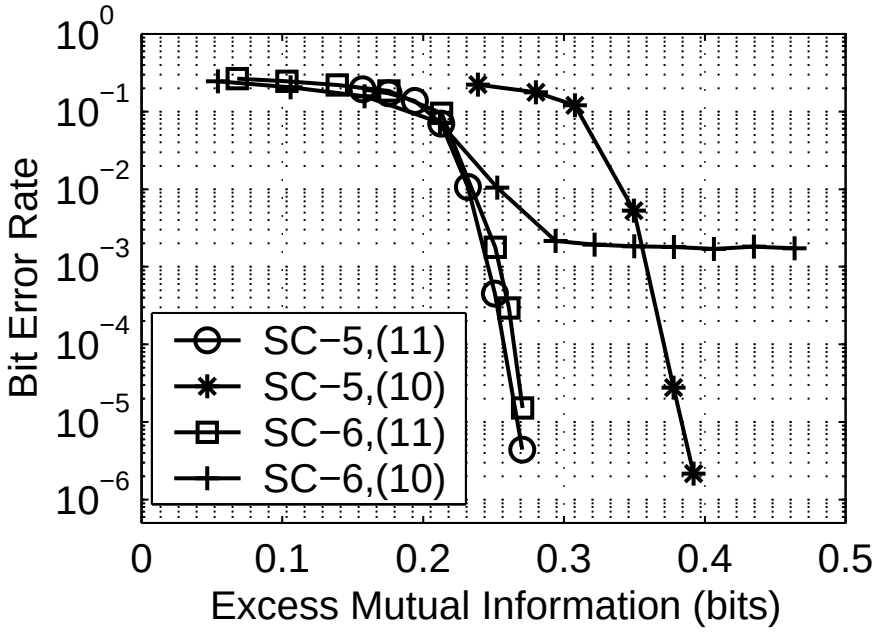

Figure 3.6: Simulations of BER versus Excess MI for two 1.0 bit/symbol SCTCMs with 4-state inner codes. SC-5 has maximal effective free distance while SC-6 has lowest pinch-off threshold. Block length=10,000, 12 iterations.

Figure 3.6 shows another set of simulations with a fixed outer code, C_2^o. SC-5, using a maximal-$d_{\text{f, eff}}$ inner TCM is compared to SC-6 using C_3^i, which has good pinch-off threshold based on EXIT analysis but has zero effective free distance. Although SC-5 and SC-6 have similar performance under AWGN, SC-6 achieves a slightly lower pinch-off threshold at the expense of a very high error floor, around BER=10^{-3}, due to its zero minimum distance, $d_{\text{f, eff}} = 0$. In contrast to the error floor of SC-1 (Figure 3.2), this error floor cannot be lowered down with more iterations.

When 8-state inner TCM's are used, the EXIT charts in Figure 3.7 (AWGN) and Figure 3.8 (period-2 erasures) show how the increased complexity does not help reduce the pinch-off thresholds via increased minimum distances. Divsalar

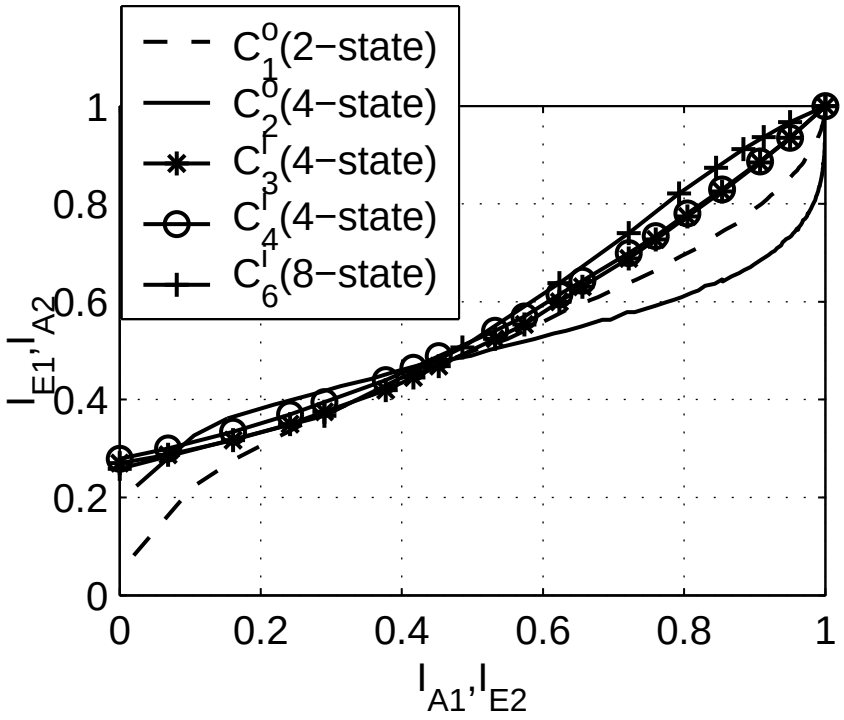

Figure 3.7: EXIT charts for some inner decoders at E_b/N_0=0.7 dB under AWGN.

[15] and ten Brink [55] also reported this kind of effect. This is more evident in the BER performance displayed in Figure 3.9 in comparison to Figures 3.5 and 3.6.

The search for robust SCTCMs over compound periodic erasure channel should start from low complexity constituent encoders, trying to avoid high error floors by using EXIT chart analysis and checking the effective free distance. Because of the tradeoff between pinch-off threshold and error floor, the code design should depend on the operating BER. For example, for 1.0 bit per symbol SCTCM at BER=10^{-5}, SC-5 is a better design as compared to SC-4 which has an error floor around BER = 4×10^{-5}. SC-5 does not exhibit an early floor, requiring an excess MI of 0.39 bits for BER = 10^{-5}.

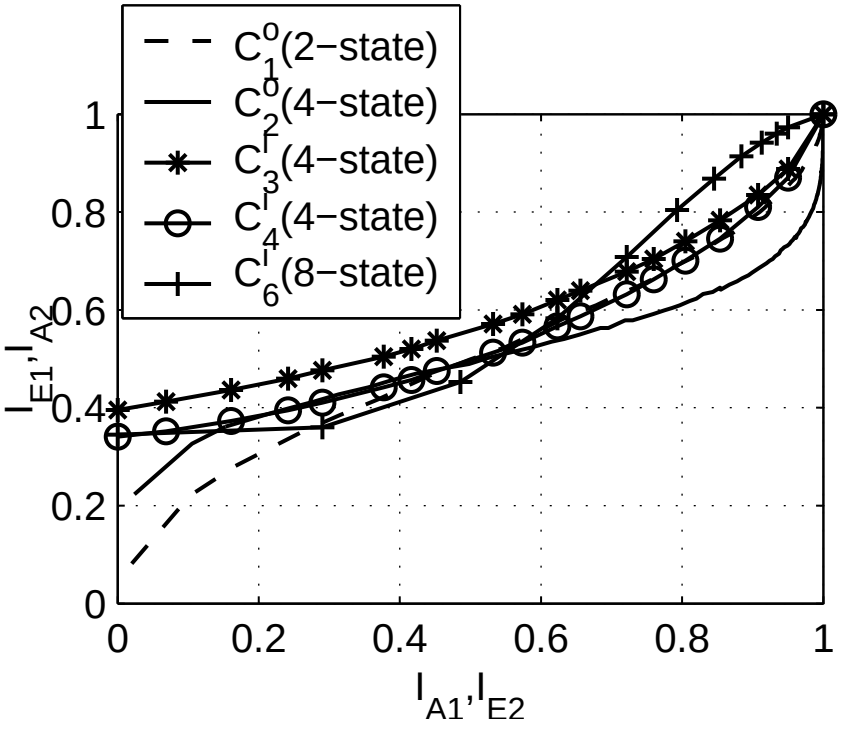

Figure 3.8: EXIT charts for some inner decoders at E_b/N_0=6.0 dB under period-2 erasures.

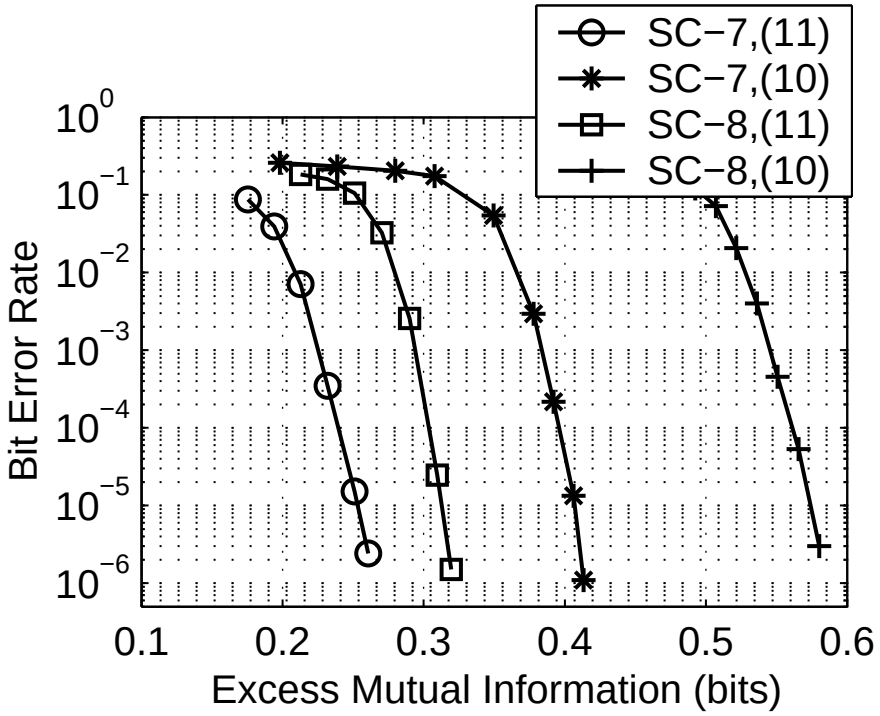

Figure 3.9: Simulations of BER versus Excess MI for two 1.0 bit/symbol SCTCMs with 8-state inner code. SC-7 has lowest pinch-off threshold while SC-8 has maximal effective free distance. Block length=10,000, 12 iterations.

Table 3.1: Rate-1/2 Outer encoders with memory ν

C^o	ν	G(D)
1	1	[1+D D]
2	2	[1+D^2 1+D+D^2]
3	3	[1+D^2 + D^3 1+D+D^2 + D^3]

3.3.3 SCTCM Design of 1.5 bits per symbol

Using the same criteria as in previous searches, we designed 1.5 bits/symbol SCTCMs by concatenating a rate-1/2 outer code with an inner TCM consisting of a linear recursive systematic rate-3/4 inner code driving a 16-QAM constellation. C_7^i has a low iterative decoding threshold and C_8^i has maximum $d_{\text{f,eff}}$ under period-2 erasures among 8-state rate-3/4 codes. As displayed in Fig. 3.10, we have not seen the error floor for SC-9 until BER=10^{-6} but SC-9 is expected to have a higher error floor than SC-10 because of its zero $d_{\text{f,eff}}$ under period-2 erasures. Unlike the case of 1.0 bit per symbol, at BER=10^{-5}, the SCTCM found by iterative decoding criterion, SC-9, is preferred. The excess MI requirement is 0.46 bits.

3.4 Conclusions

The design of robust SCTCMs in this paper aims to maintain turbo performance under periodic erasures with as little compromise as possible on the AWGN performance. At BER=10^{-5}, the proposed SCTCMs transmitting at 0.5 bits/symbol, 1 bit/symbol and 1.5 bit/symbol require 0.16, 0.39, 0.46 bits of excess MI, respectively. Typical 1 bit/symbol SCTCMs [23] at similar complexity require 0.35 bits of excess MI at BER = 10^{-5}. The proposed 1 bits/symbol

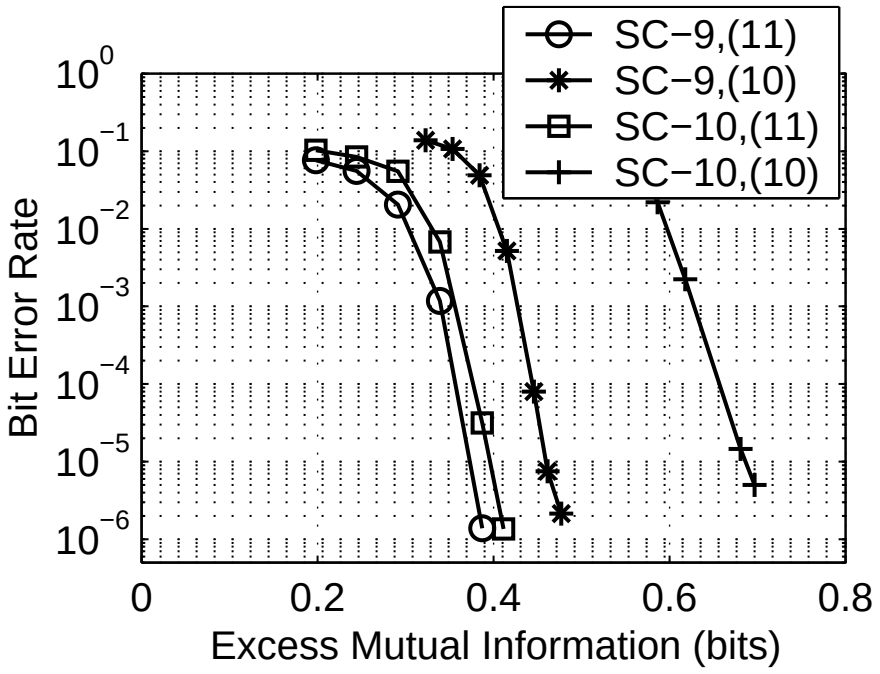

Figure 3.10: Simulations of BER versus Excess MI for two 1.5 bits/symbol SCTCMs. SC-9 has the lowest pinch-off threshold while SC-10 has maximal effective free distance. Block length=10,000,12 iterations.

Table 3.2: Inner encoders with memory ν

C^i	k	n	ν	G(D)
1	1	3	3	$[1 \ \frac{D}{1+D^2} \ 1+\mathrm{D}]$
2	1	3	3	$[1 \ \frac{D^2}{1+D+D^2} \ \frac{1}{1+D}]$
3	2	3	2	$\begin{bmatrix} 1 & 0 & \frac{1}{1+D^2} \\ 0 & 1 & \frac{D}{1+D^2} \end{bmatrix}$
4	2	3	2	$\begin{bmatrix} 1 & 0 & \frac{D}{1+D^2} \\ 0 & 1 & \frac{1+D}{1+D^2} \end{bmatrix}$
5	2	3	3	$\begin{bmatrix} 1 & 0 & \frac{1}{1+D^3} \\ 0 & 1 & \frac{D}{1+D^3} \end{bmatrix}$
6	2	3	3	$\begin{bmatrix} 1 & 0 & \frac{D}{1+D^2+D^3} \\ 0 & 1 & \frac{1+D+D^2}{1+D^2+D^3} \end{bmatrix}$
7	3	4	3	$\begin{bmatrix} 1 & 0 & 0 & \frac{1}{1+D^3} \\ 0 & 1 & 0 & \frac{D}{1+D^3} \\ 0 & 0 & 1 & \frac{D^2}{1+D^3} \end{bmatrix}$
8	3	4	3	$\begin{bmatrix} 1 & 0 & 0 & \frac{1}{1+D+D^3} \\ 0 & 1 & 0 & \frac{D^2}{1+D+D^3} \\ 0 & 0 & 1 & \frac{1+D+D^2}{1+D+D^3} \end{bmatrix}$

Table 3.3: SCTCM Schemes

SC	C^o	C^i	R	Constellation
1	2	1	1/6	8-PSK
2	2	2	1/6	8-PSK
3	3	1	1/6	8-PSK
4	1	4	1/3	8-PSK
5	2	4	1/3	8-PSK
6	2	3	1/3	8-PSK
7	2	5	1/3	8-PSK
8	2	6	1/3	8-PSK
9	2	7	3/8	16-QAM
10	2	8	3/8	16-QAM

SCTCM handles periodic erasures as well as AWGN with only 0.04 bits of excess MI loss.

An interesting outcome of our experiments is that it is possible to work under periodic erasures with an inner TCM that has negative redundancy when every other symbol is erased, as long as the effective free Euclidean distance under periodic erasures is nonzero. On the other hand, zero effective free distances on the inner code result in very high error floors that cannot be removed by simply increasing the number of turbo decoding iterations. This requirement on the effective free Euclidean distance places an upper bound on the rate of the inner encoder with a given number of states. The combined rate of the inner and outer encoders should be strictly lower than 1/2 to provide coded performance under period-2 erasures.

The maximum-likelihood decoding criterion and iterative decoding criterion

yield SCTCMs with very different performances. The ML-optimal SCTCMs have larger effective free distances and thus lower error floors but their pinch-off thresholds are usually higher. On the other hand, SCTCMs which are designed solely based on low pinch-off thresholds tend to have higher error floors. The trade-off between low pinch-off thresholds and low error floors is more evident under periodic erasures when most of the redundancy is removed. Such a tradeoff makes it necessary to consider both ML and iterative decoding criteria during the code search.

CHAPTER 4

Universal space-time block constellations

4.1 Introduction

4.2 System Model and notation

4.2.1 Multisymbol linear space-time modulations

Consider a space-time communication system with N_t transmit antennas. An (N_t, Q, L) linear space-time block signaling scheme is a space-time arrangement of Q data *threads* over L channel uses. Also known as linear dispersion (LD) [6] codes, an (N_t, Q, L) linear space-time block signaling scheme can be represented as

$$\mathbf{S} = \sum_{q=1}^{Q} x_q \mathbf{A}_q + j x_q^* \mathbf{B}_q, \tag{4.1}$$

$$\mathbf{A}_q, \mathbf{B}_q \in \mathbb{C}^{N_t \times L}, \ q = 1, ..., Q \tag{4.2}$$

where $\mathbf{A}_q$ and $\mathbf{B}_q$ are the linear dispersion matrices for the real part of complex symbol x_q and its conjugate x_q^*, respectively.

An equivalent representation of the block code (4.1) is through the set of $N_t \times Q$ matrices $\left\{\mathbf{T}_l^a, \mathbf{T}_l^b\right\}_{l=1}^{L}$, defined by

$$(\mathbf{T}_l^a)_{i,q} = (\mathbf{A}_q)_{i,l}, \quad (\mathbf{T}_l^b)_{i,q} = (\mathbf{B}_q)_{i,l}, \tag{4.3}$$

$$l = 1, ..., L, \ \ q = 1, ..., Q, \ \ i = 1, ..., N_t.$$

The qth column of the matrix $\mathbf{T}_l^a$ ($\mathbf{T}_l^b$) represents the the contribution of x_q (x_q^*) at time l to the transmitted vector. Since the horizontal dimension represents space and the vertical dimension represents threads, we call $\left\{\mathbf{T}_l^a, \mathbf{T}_l^b\right\}_{l=1}^L$ *space-thread* matrices.

Assuming the in-phase and quadrature components of each thread are independent random signals (although the threads may be dependent among themselves) with equal power 1/2, the total transmit power is

$$\begin{aligned} P &= \mathbb{E}[\mathrm{tr}(\mathbf{S}\mathbf{S}^\dagger)] \\ &= \sum_{i=1}^{Q} \mathrm{tr}\left(\mathbf{A}_{i,R}\mathbf{A}_{i,R}^\dagger + \mathbf{A}_{i,I}\mathbf{A}_{i,I}^\dagger + \mathbf{B}_{i,R}\mathbf{B}_{i,R}^\dagger + \mathbf{B}_{i,I}\mathbf{B}_{i,I}^\dagger\right). \end{aligned} \tag{4.4}$$

For $l = 1, ..., L$, let $\mathbf{s}^l$ denote the lth column of $\mathbf{S}$, i.e. the signal transmitted for the lth symbol. Then,

$$\begin{bmatrix} \mathbf{s}_R^l \\ \mathbf{s}_I^l \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{l,R}^a + \mathbf{T}_{l,R}^b & -\mathbf{T}_{l,I}^a + \mathbf{T}_{l,I}^b \\ \mathbf{T}_{l,I}^a + \mathbf{T}_{l,I}^b & \mathbf{T}_{l,R}^a - \mathbf{T}_{l,R}^b \end{bmatrix} \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix} \tag{4.5}$$

$$= \left(\begin{bmatrix} \mathbf{T}_{l,R}^a & -\mathbf{T}_{l,I}^a \\ \mathbf{T}_{l,I}^a & \mathbf{T}_{l,R}^a \end{bmatrix} + \begin{bmatrix} \mathbf{T}_{l,R}^b & \mathbf{T}_{l,I}^b \\ \mathbf{T}_{l,I}^b & -\mathbf{T}_{l,R}^b \end{bmatrix} \right) \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix} \tag{4.6}$$

$$= \left(\boldsymbol{\mathcal{T}}_l^a + \boldsymbol{\mathcal{T}}_l^b\right) \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix} \tag{4.7}$$

(note that $\boldsymbol{\mathcal{T}}_l^a$ and $\boldsymbol{\mathcal{T}}_l^b$ have different structure), and the space-time block-signaling scheme (4.1) can be represented as

$$\begin{bmatrix} \mathbf{s}_R^1 \\ \mathbf{s}_I^1 \\ \vdots \\ \mathbf{s}_R^L \\ \mathbf{s}_I^L \end{bmatrix} = \boldsymbol{\mathcal{T}} \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix} \tag{4.8}$$

where

$$\boldsymbol{\mathcal{T}} = \boldsymbol{\mathcal{T}}^a + \boldsymbol{\mathcal{T}}^b \tag{4.9}$$

and

$$\boldsymbol{\mathcal{T}}^a = \begin{bmatrix} \boldsymbol{\mathcal{T}}^a_1 \\ \boldsymbol{\mathcal{T}}^a_2 \\ \vdots \\ \boldsymbol{\mathcal{T}}^a_L \end{bmatrix}, \quad \boldsymbol{\mathcal{T}}^b = \begin{bmatrix} \boldsymbol{\mathcal{T}}^b_1 \\ \boldsymbol{\mathcal{T}}^b_2 \\ \vdots \\ \boldsymbol{\mathcal{T}}^b_L \end{bmatrix}. \tag{4.10}$$

4.2.2 The code-constrained signaling

Consider the linear Gaussian vector channel

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{4.11}$$

where $\mathbf{H}$ is the $N_r \times N_t$ channel matrix, and the additive noise vector $\mathbf{w}$ has zero-mean complex Gaussian distribution with unit covariance matrix $\mathbb{E}\mathbf{w}\mathbf{w}^\dagger = \mathbf{I}_{N_r}$.

The transmit structure of the linear space-time signaling scheme (4.8) on real coordinates can now be represented as

$$\tilde{\mathbf{y}} = (\mathbf{I}_L \otimes \mathcal{H}) \left(\boldsymbol{\mathcal{T}}^a + \boldsymbol{\mathcal{T}}^b\right) \tilde{\mathbf{x}} + \tilde{\mathbf{w}}. \tag{4.12}$$

where

$$\mathcal{H} = \begin{bmatrix} \mathbf{H}_R & -\mathbf{H}_I \\ \mathbf{H}_I & \mathbf{H}_R \end{bmatrix}. \tag{4.13}$$

and $\otimes$ is the Kronecker product.

Recall that the unconstrained Gaussian-input mutual-information is given by

$$C(\mathbf{H}) = \log_2 \det\left(\mathbf{I}_{N_r} + \mathbf{H}\mathbf{H}^\dagger\right) \quad \text{bits per channel use.} \tag{4.14}$$

Per equation 4.8, the code-constrained Gaussian-input mutual information

1. *Transmit power in terms of space-thread matrices:* The total transmit power, given in terms of the space-time dispersion matrices 4.4 is given in terms of the space-thread matrices as

$$P = \frac{1}{2}\sum_{l=1}^{L}\left(\text{trace}\left(\boldsymbol{\mathcal{T}}_l^a \boldsymbol{\mathcal{T}}_l^{aT}\right) + \text{trace}\left(\boldsymbol{\mathcal{T}}_l^b \boldsymbol{\mathcal{T}}_l^{bT}\right)\right). \tag{4.15}$$

2. *Singular value decompostion of the equivalent L-symbol real-valued channel matrix:* Let

$$\mathbf{H} = \mathbf{U}\mathbf{S}^{1/2}\mathbf{V}^{\dagger} \tag{4.16}$$

be a singular value decomposition of $\mathbf{H}$ where

$$\mathbf{S} = \begin{bmatrix}\boldsymbol{\Lambda}^{1/2} & \mathbf{I}_{N_t - N_r}\end{bmatrix} \tag{4.17}$$

and $\boldsymbol{\Lambda}$ is the $N_r \times N_r$ diagonal matrix of eigenvalues of $\mathbf{H}\mathbf{H}^{\dagger}$. Let $\mathbf{u}^i$ denote the ith column of $\mathbf{U}$, and let $\mathbf{v}^j$ denote the $j$the column of $\mathbf{V}$. Then, the real-valued matrix $\mathcal{H}$ has singular value decomposition

$$\mathcal{H} = \mathcal{U}\left(\mathbf{I}_2 \otimes \boldsymbol{\Lambda}^{1/2}\right)\mathcal{V}^{\dagger} \tag{4.18}$$

where

$$\mathcal{U} = \begin{bmatrix} \mathbf{u}_R^1 & -\mathbf{u}_I^1 & \cdots & \mathbf{u}_R^{N_r} & -\mathbf{u}_I^{N_r} \\ \mathbf{u}_I^1 & \mathbf{u}_R^1 & \cdots & \mathbf{u}_I^{N_r} & \mathbf{u}_R^{N_r} \end{bmatrix} \tag{4.19}$$

and

$$\mathcal{V} = \begin{bmatrix} \mathbf{v}_R^1 & -\mathbf{v}_I^1 & \cdots & \mathbf{v}_R^{N_t} & -\mathbf{v}_I^{N_t} \\ \mathbf{v}_I^1 & \mathbf{v}_R^1 & \cdots & \mathbf{v}_I^{N_t} & \mathbf{v}_R^{N_t} \end{bmatrix}. \tag{4.20}$$

Consequently, the equivalent channel matrix over L symbols will have the following singular value decomposition:

$$\begin{aligned} \mathbf{I}_L \otimes \mathcal{H} &= \mathbf{I}_L \otimes \left(\mathcal{U}\left(\mathbf{I}_2 \otimes \boldsymbol{\Lambda}^{1/2}\right)\mathcal{V}^{\dagger}\right) && (4.21) \\ &= \left(\mathbf{I}_L \otimes \mathcal{U}\right)\left(\mathbf{I}_{2L} \otimes \boldsymbol{\Lambda}^{1/2}\right)\left(\mathbf{I}_L \otimes \mathcal{V}\right)^{\dagger}. && (4.22) \end{aligned}$$

3. *Code-constrained mutual information and interference of threads:* Recall the constrained capacity expression:

$$C(\mathbf{H}, \boldsymbol{\mathcal{T}}) = \frac{1}{2L} \log_2 \left(\mathbf{I}_{2LN_r} + (\mathbf{I}_L \otimes \mathcal{H}) \, \boldsymbol{\mathcal{T}} \boldsymbol{\mathcal{T}}^T \left(\mathbf{I}_L \otimes \mathcal{H}^T \right) \right) \tag{4.23}$$

where

$$\boldsymbol{\mathcal{T}} \boldsymbol{\mathcal{T}}^T = \sum_{l=1}^{L} \left(\boldsymbol{\mathcal{T}}_l^a \boldsymbol{\mathcal{T}}_l^{aT} + \boldsymbol{\mathcal{T}}_l^b \boldsymbol{\mathcal{T}}_l^{bT} \right) + \sum_{l=1}^{L} \left(\boldsymbol{\mathcal{T}}_l^a \boldsymbol{\mathcal{T}}_l^{bT} + \boldsymbol{\mathcal{T}}_l^b \boldsymbol{\mathcal{T}}_l^{aT} \right). \tag{4.24}$$

The first summation is a symmetric positive semidefinite matrix whose trace determines the transmit power. The second summation is a symmetric matrix whose trace is *zero*:

$$\mathbf{T}_l^a \mathbf{T}_l^{bT} = \begin{bmatrix} \mathbf{T}_{l,R}^a & -\mathbf{T}_{l,I}^a \\ \mathbf{T}_{l,I}^a & \mathbf{T}_{l,R}^a \end{bmatrix} \begin{bmatrix} \mathbf{T}_{l,R}^b & \mathbf{T}_{l,I}^b \\ \mathbf{T}_{l,I}^b & -\mathbf{T}_{l,R}^b \end{bmatrix}^T \tag{4.25}$$

$$= \begin{bmatrix} \mathbf{T}_{l,R}^a \mathbf{T}_{l,R}^{b\,T} - \mathbf{T}_{l,I}^a \mathbf{T}_{l,I}^{b\,T} & \cdot \\ \cdot & \mathbf{T}_{l,I}^a \mathbf{T}_{l,I}^{b\,T} - \mathbf{T}_{l,R}^a \mathbf{T}_{l,R}^{b\,T} \end{bmatrix}. \tag{4.26}$$

The implication is that the matrix represented by the second summation has at least one *negative eigenvalue* unless it is identically the zero matrix. For orthogonal block *designs*, this matrix *is* identically zero, indicating complete thread independece.

4. Diversity: In terms of $\boldsymbol{\mathcal{T}}^a$, $\boldsymbol{\mathcal{T}}^b$?

4.3 Design of minimax-MI loss space-time block signaling schemes

For a given linear space-time block signaling $\boldsymbol{\mathcal{T}}$, the *worst-case channel* minimizes the code-constrained mutual information

$$C(\mathbf{H}, \boldsymbol{\mathcal{T}}) = \frac{1}{2L} \log_2 \left(\mathbf{I}_{2LN_r} + (\mathbf{I}_L \otimes \mathcal{H}) \, \boldsymbol{\mathcal{T}} \boldsymbol{\mathcal{T}}^T \left(\mathbf{I}_L \otimes \mathcal{H}^T \right) \right) \tag{4.27}$$

subject to

$$C(\mathbf{H}) = \frac{1}{2L}\log_2 \det\left(\mathbf{I}_{LN_r} + (\mathbf{I}_L \otimes \mathcal{H})(\mathbf{I}_L \otimes \mathcal{H})^T\right) \geq R. \tag{4.28}$$

where R is the rate of information transmission in bits per channel use.

We would like to find linear space-time block schemes $\boldsymbol{\mathcal{T}}$ that minimize the maximum mutual information loss

$$\sup\left\{R - C(\mathbf{H}, \boldsymbol{\mathcal{T}}) : C(\mathbf{H}) = R\right\}. \tag{4.29}$$

A space-time linear block signaling scheme is completely specified by the space-thread matrices $\{\boldsymbol{\mathcal{T}}_l^a, \boldsymbol{\mathcal{T}}_l^b\}$. The specifies the eigenvalues and eigenvectors of the symmetric positive semidefinite metrix

$$\boldsymbol{\mathcal{T}}\boldsymbol{\mathcal{T}}^T = \sum_{l=1}^{L}\left(\boldsymbol{\mathcal{T}}_l^a\boldsymbol{\mathcal{T}}_l^{aT} + \boldsymbol{\mathcal{T}}_l^b\boldsymbol{\mathcal{T}}_l^{bT}\right) + \sum_{l=1}^{L}\left(\boldsymbol{\mathcal{T}}_l^a\boldsymbol{\mathcal{T}}_l^{bT} + \boldsymbol{\mathcal{T}}_l^b\boldsymbol{\mathcal{T}}_l^{aT}\right) \tag{4.30}$$

where the first summationhas nonnegative eigenvalues, and the second summation has at least one negative eigenvalue unless not identically zero.

Recall that

$$\boldsymbol{\mathcal{T}} = \left[\mathcal{T}_i\mathcal{T}_j^T\right]_{i,j} = \begin{bmatrix} \mathcal{T}_1\mathcal{T}_1^T, & \mathcal{T}_1\mathcal{T}_2^T, & \cdots & \mathcal{T}_1\mathcal{T}_L^T \\ \mathcal{T}_2\mathcal{T}_1^T, & \mathcal{T}_2\mathcal{T}_2^T, & \cdots & \mathcal{T}_2\mathcal{T}_L^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{T}_L\mathcal{T}_1^T, & \mathcal{T}_L\mathcal{T}_2^T, & \cdots & \mathcal{T}_L\mathcal{T}_L^T \end{bmatrix} \tag{4.31}$$

therefore the matrix

$$(\mathbf{I}_L \otimes \mathcal{H})\,\boldsymbol{\mathcal{T}}\boldsymbol{\mathcal{T}}^T\left(\mathbf{I}_L \otimes \mathcal{H}^T\right) \tag{4.32}$$

can also be written as a block-partitioned matrix where the i,jth block is

$$\mathcal{H}\mathcal{T}_i\mathcal{T}_j^T\mathcal{H}^T \tag{4.33}$$

for $i, j \in \{1, ..., L\}$, i.e.,

$$(\mathbf{I}_L \otimes \mathcal{H})\, \boldsymbol{\mathcal{T}} \boldsymbol{\mathcal{T}}^T \left(\mathbf{I}_L \otimes \mathcal{H}^T\right) = \left[\mathcal{H}\mathcal{T}_i \mathcal{T}_j^T \mathcal{H}^T\right]_{i,j}. \tag{4.34}$$

The localization of eigenvalues of a block-partitioned matrix in terms of the eigenvalues of the partitioned matrices is a mathematically interesting problem. In the case where the partitions are scalars, the problem reduces to classical eigenvalue localization. A vast number of inequalities that describe the location of eigenvalues of a matrix in terms of its elements has been well-known. In particular, Gerschgorin circle theorem states that the eigenvalues $\{\lambda\}$ of an $n \times n$ matrix $A = [a_{ij}]$ obey

$$\min_{i=1,..,n} \left\{ -a_{ii} + \sum_{j \neq i} |a_{ij}| \right\} \leq \lambda \leq \max_{i=1,..,n} \left\{ a_{ii} + \sum_{j \neq i} |a_{ij}| \right\}. \tag{4.35}$$

Varga [44] extended this result to block-partitioned invertebrate matrices, partitiones matrices with vanishing main diagonal partitions. Varga's theory is explaind in Section 4.4. Write

$$\left[\mathcal{H}\mathcal{T}_i \mathcal{T}_j^T \mathcal{H}^T\right]_{i,j} = \mathcal{D} + \mathcal{E} \tag{4.36}$$

where $\mathcal{D}$ is the *vertebrae* of the partition, i.e.

$$\mathcal{D} = \mathrm{diag}\left(\mathcal{H}\mathcal{T}_1 \mathcal{T}_1^T \mathcal{H}^T, ..., \mathcal{H}\mathcal{T}_L \mathcal{T}_L^T \mathcal{H}^T\right). \tag{4.37}$$

Since $\mathcal{D}$ is symmetric, so is $\mathcal{E}$. The Gerschgorin spectral theory summarized in 4.4 is applied to $\mathcal{E}$ to upper-bound the spectral radius $\mu(\mathcal{E})$ of $\mathcal{E}$.

Theorem 6 (Code-constrained MI) *For a given linear space-time block code* $\{\tilde{\mathbf{A}}_t, \tilde{\mathbf{B}}_t\}$ *the code-constrained mutual information induced by the channel matrix*

H *is upper- and lower-bounded as*

$$C(\mathbf{H}, \boldsymbol{\mathcal{T}}) \leq \frac{1}{2L} \sum_{l=1}^{L} \sum_{i=1}^{N_r} \log_2 \left(1 + \lambda_i(\mathcal{H}\mathcal{T}_l\mathcal{T}_l^T\mathcal{H}^T) + \beta\right) \tag{4.38}$$

$$C(\mathbf{H}, \boldsymbol{\mathcal{T}}) \geq \frac{1}{2L} \sum_{l=1}^{L} \sum_{i=1}^{N_r} \log_2 \left(1 + \left(\lambda_i(\mathcal{H}\mathcal{T}_l\mathcal{T}_l^T\mathcal{H}^T) - \beta\right)^+\right) \tag{4.39}$$

bits per channel use, where

$$\beta = \max_i \sum_{k \neq i} s_{\max}\left(\mathcal{H}\mathcal{T}_i\mathcal{T}_k^T\mathcal{H}\right). \tag{4.40}$$

4.4 Appendix: Gerschgorin spectral theory for the eigenvalues of partitioned invertebrate matrices

Table 4.1: Frequently used symbols associated with linear space-time block signaling schemes

N_t	number of transmit antennas
N_r	number of receive antennas
$\mathbf{H}$	$N_r \times N_t$ channel matrix of complex numbers
Q	number of space-time threads
L	duration of the space-time block, channel uses
$\mu(\mathbf{E})$	spectral radius of $\mathbf{E}$
$(y)^+$	$\max(y, 0)$
$\lambda_i(\mathbf{E}\mathbf{E}^\dagger)$	ith largest eigenvalue of $\mathbf{E}\mathbf{E}^\dagger$
$s_{\max}(\mathbf{E})$	$\lambda_{\max}(\mathbf{E}\mathbf{E}^\dagger)$

In the following, we briefly review the Gerschgorin spectral theory for *partitioned* matrices and show how the spectral conjecture of Varga [44] can be used to estimate the eigenvalues of the equivalent channel correlation matrix.

Consider the nd-dimensional complex space $\mathbb{C}^{nd}$.

1. Define standard unit vectors $\mathbf{e}_k \in \mathbb{C}^d$ by $\mathbf{e}_k = (e_{k,1}, ..., e_{k,d})^T$, $e_{k,l} = \delta_{k,l}$ and define subspaces $\mathcal{C}_i = \text{span}\{\mathbf{e}_k, di \leq k \leq d(i+1)-1\}$, $i = 1, ..., n$. $\mathcal{C}_i$ is the projection of $\mathbb{C}^{nd}$ onto its n components with indices $di \leq k \leq d(i+1)-1$ with respect to the usual basis. Then $\mathbb{C}^{nd}$ is a direct sum of the subspaces $\mathcal{C}_i$: $\mathbb{C}^{nd} = \mathcal{C}_1 \oplus \mathcal{C}_2 \oplus ... \oplus \mathcal{C}_n$. Call this partition of the space $\mathbb{C}^{nd}$, π.

2. Let ϕ_i be a vector norm on $\mathcal{C}_i$, $i = 1, ..., n$. Define the operator norm n-tuple $\phi = (\phi_1, ..., \phi_n)$ by

$$\|\mathbf{z}\|_\phi = \max_{1 \leq i \leq n} \phi_i(P_i \mathbf{z}) \quad \mathbf{z} \in \mathbb{C}^{nd} \tag{4.41}$$

   where P_i is the projection operator from $\mathbb{C}^{nd}$ to $\mathcal{C}_i$. Let Φ_π be the collection of all such operator norm n-tuples associated with the partition π.

3. Consider the π-partitioned matrix

$$\mathcal{F} = \begin{bmatrix} \mathbf{F}_{1,1}, & \mathbf{F}_{1,2}, & \cdots & \mathbf{F}_{1,n} \\ \mathbf{F}_{2,1}, & \mathbf{F}_{2,2}, & \cdots & \mathbf{F}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{F}_{n,1}, & \mathbf{F}_{n,2}, & \cdots & \mathbf{F}_{n,n} \end{bmatrix} \tag{4.42}$$

   where $\mathbf{F}_{i,k} \in \mathbb{C}^{d \times d}$ represents a linear mapping from $\mathcal{C}_k$ to $\mathcal{C}_i$. Write $\mathcal{F} = [\mathbf{F}_{i,k}]_\pi$. $\mathcal{F}$ is called π-*invertebrate* if $\mathbf{F}_{i,i} = \mathbf{0}$, $i = 1, ..., n$. For a given operator norm n-tuple ϕ, let $\Gamma_\pi^\phi(\mathcal{F}) = \{\mathcal{E} = [\mathbf{E}_{i,k}]_\pi, \mathbf{E}_{i,i} = \mathbf{0}, \|\mathcal{E}\|_\phi = \|\mathcal{F}\|_\phi\}$ be the collection of π-invertebrate matrices whose ϕ-induced operator norms are equal to the operator norm $\|\mathcal{F}\|_\phi = \sup_{\mathbf{u} \in \mathbb{C}^{nd}: \|\mathbf{u}\|=1} \|\mathcal{F}\mathbf{u}\|_\phi$ of $\mathcal{F}$. Also let $\Gamma_\pi(\mathcal{F}) = \cap_\phi \Gamma_\pi^{\phi \in \Phi_\pi}(\mathcal{F})$.

4. Given $\mathcal{F} = [\mathbf{F}_{i,k}]_\pi$, define the $n \times n$ *comparison matrix* $\mathbf{M}^\phi = [M_{p,q}^\phi]$ by $M_{p,q}^\phi = \|\mathbf{F}_{p,q}\|_{\phi_p} = \sup_{\mathbf{z}_q \in \mathcal{C}_q, \phi_q(\mathbf{z}_q)=1} \phi_p(\mathbf{F}_{p,q}\mathbf{z}_q)$.

5. For any matrix $\mathbf{V}$, let $\mu(\mathbf{V})$ denote the spectral radius (largest magnitude of the eigenvalues) of $\mathbf{V}$. The largest eigenvalue of the positive semidefinite matrix $\mathbf{V}\mathbf{V}^\dagger$ is $(\mu(\mathbf{V}))^2$.

The spectral conjecture [44] states that $\sup_{\mathcal{E}\in\Gamma_\pi(\mathcal{F})}\mu(\mathcal{E}) = \inf_{\phi\in\Phi_\pi}\|\mathcal{E}\|_\phi$ for a given partition π and a π-invertebrate $\mathcal{F}$. Varga [44] went on to prove a stronger statement:

Theorem 7 (Spectral conjecture)

$$\sup_{\mathcal{E}\in\Gamma_\pi(\mathcal{F})}\mu(\mathcal{E}) = \inf_{\phi\in\Phi_\pi}\|\mathcal{E}\|_\phi = \inf_{\phi\in\Phi_\pi}\|\mathbf{M}^\phi(\mathcal{F})\|_\infty \tag{4.43}$$

where $\|\cdot\|_\infty$ is the maximum absolute row sum.

Corollary 7.1 *Let $\mathcal{E} = [\mathbf{E}_{i,k}]_\pi$ be π-invertebrate. Then,*

$$\mu_{max}(\mathcal{E}) \le \max_i \sum_{k\neq i}^{n} s_{max}(\mathbf{E}_{i,k}), \tag{4.44}$$

where s_{max} denotes the largest singular value (Table 4.1).

Proof 7 *Note that $\mathcal{E}\in\Gamma_\pi$ therefore (4.43) implies $\mu(\mathcal{E}) \le \|\mathbf{M}^\phi(\mathcal{E})\|_\infty$ for any norm n-tuple $\phi = (\phi_1, ..., \phi_n)$. Let ϕ_i be the L_2 norm norm on $\mathcal{C}_i$. Then $M_{i,k} = \sup_{\mathbf{z},\|\mathbf{z}\|_2=1}\|\mathbf{E}_{i,k}\mathbf{z}\|_2 = s_{max}(\mathbf{E}_{i,k})$ and the desired result follows.*

CHAPTER 5

A low-complexity approach to sequential multihypothesis detection

5.1 Introduction

Model selection and experiment design are essential to any hypothesis testing problem. For some problems, such as the transmission of one of a finite number of signals over a noisy channel whose parameters are known at the receiver, model selection is often trivial. However, in other cases there will be uncertainty about the model itself. In radar detection, the spectral characteristics of the interference and the target power spectra are often unknown [1]. In sonar detection, inaccurate characterization of the randomly fluctuating underwater channels results in uncertain models [24]. In environments with man-made interference, Gaussian assumptions in the estimation of signal parameters may result in significant performance degradation ([13], [42]). The issue of source and model mismatch also arises in classification [37], source coding [2], and decoding [3]. In addition to the limitations in modeling, the measurement of observables is often subject to unmodeled distortion, such as unmodeled noise and calibration errors. In such cases, none of the hypotheses in the model reflects the true distribution from which the observables are drawn. The first goal of this paper is to describe the performance limits of hypothesis testing when the experiment model is inaccu-

rate. To this end, the first part of the paper presents an upper bound on the best exponent with which the decision error probability under any hypothesis H_j decreases in the asymptote of small decision error probabilities in the possible presence of modeling errors.

Historically, the common starting point to address the problem of model uncertainty has been to incorporate the allowed deviations from the correct model by forming composite hypotheses, where each hypothesis is a *neighborhood* ([38], [47]) of a distribution. For two hypotheses, Huber [38] formulated a least favorable pair of distributions under various Bayesian and minimax constraints for likelihood-ratio tests. Hoeffding ([62], [63]) derived lower bounds on the expected sample size of a sequential test of binary *composite* hypotheses for which the decision error probabilities under both hypotheses are bounded above by predetermined numbers. For M-ary simple hypotheses, Simons [21] obtained lower bounds on the average sample size of a test when the observed data fits an $(M+1)$st hypothesis, however, there is no interpretation of the error probability under the new distribution. Other works related to hypothesis testing under model inaccuracy have concentrated on formulating the asymptotic (large sample-size) -distribution of the likelihood-ratio under an incorrect model [45], [32]. More recently, Baum and Veeravalli [11] raised the question of how their asymptotically optimum M-ary sequential probability ratio test (MSPRT) would perform under an incorrect model.

The first part of our analysis focuses on a scenario in which the experimenter is allowed to prune a subset of the original set of hypotheses based on the data observed so far. An example of hypothesis pruning is reduced-complexity decoding of trellis codes ([16], [25]). Section 5.3 introduces a two-step decision rule that applies the asymptotically optimum M-ary sequential probability ratio test

(MSPRT) to a subset of the original set of hypotheses after a pruning step. The pruning step, providing the complexity reduction for the remainder of the test, has the risk of rejecting the correct hypothesis in which case the final decision is delivered by a mismatched MSPRT. These issues found the bases of Section 5.4 and Section 5.5. Upon concluding, Section 5.3 summarizes the main results of the paper.

In order to analyze the feasibility of rejecting hypotheses, Section 5.4 introduces *weak* sequential decisions that as to what the correct state of the nature is *not*, rather than what it is based on threshold rules on the vector of posterior probabilities. Drawing a parallel to the sequential probability ratio tests, this section derives the asymptotic stopping-times and upper bounds to error probabilities for weak decisions. The results apply to the first part of the two-step test of Section 5.3.

Section 5.5 focuses on the MSPRT under incorrect models, a precalculated contingency in the proposed two-step test, also responding the authors' [11] concluding remarks. Under mild model mismatch, this section also derives a relationship between the decision error probabilities and the stopping-times of any sequential test in the asymptote of small decision error probabilities, based on Fano's inequality.

5.2 Sequential Tests: Previous Research

For binary hypothesis testing, Wald's [4] sequential probability ratio test (SPRT) has been known to be optimal [5] in the sense that it has the smallest expected sample size under either of the hypotheses, among all tests (sequential and non-sequential) for which the decision error probabilities do not exceed some pre-

defined values. The SPRT stops and delivers a decision whenever one of the posterior probabilities exceeds a threshold. The SPRT typically requires two to three times fewer observations on the average as compared to a fixed sample-size test operating at the same decision error probability levels.

A direct generalization of the binary SPRT to the case of multiple ($M \geq 3$) hypotheses does not guarantee optimality in the same sense as the binary case. In fact, the existence of an M-ary sequential test which minimizes the expected sample size under each of the hypotheses for given levels of decision error probabilities, has not been proven [59]. Nevertheless, the simplicity of the structure of the SPRT has inspired researchers to propose various M-ary sequential tests based on likelihood threshold rules (for example, [35],[29],[20],[11]). Aside from this structural simplification, the performance analysis have been facilitated by the use of an asmyptotic setting in which the decision error probabilities are forced to zero, either explicitly (e.g., [59]) or by allowing the experimentation cost per sample to go to zero in a Bayesian setting (e.g., [29], [11]).

5.3 Reduced-complexity tests based on pruning hypotheses

Let $\{X_t, t \in \mathbb{N}\}$ be an observed sequence of identically distributed random variables with probability density function f, in a probability space $(\Omega, F, \mathbb{P})$. Consider M simple hypotheses

$$H_i : f = f_i, \qquad i \in \mathcal{M} = \{0, ..., M-1\} \tag{5.1}$$

where all f_i are are defined with respect to the same $\mathbb{P}$. Let $\mathcal{M}$ denote the the set of indices $\{0, ..., M-1\}$. The nature selects index u with probability π_u. For $n \geq 1$, let F_n, denote the σ-algebra generated by the first n observations. A sequential

test is a pair $(N, \hat{u})$ where N is an $\{F_n\}$-stopping time and $\hat{u}(X_1, ..., X_N) \in \mathcal{M}$ is F_N measurable decision rule. We will use $\mathbb{P}_i$, $\mathbb{E}_i$ to denote probability and expectation conditioned on $u = i$. $\mathbb{E}$ will be shorthand for $\sum_{i \in \mathcal{M}} \pi_i \mathbb{E}_i$.

For a fixed error performance, the computational complexity of identifying the correct hypothesis often grows in proportion to M [11], which results from the contribution of each hypothesis to a decision metric update. To alleviate this complexity, the experimenter may choose to to prune the set of hypotheses based on the observed data.

For a sequential decision rule with stopping-time T, let M_t, $t \geq 1$ be the number of *active* hypotheses (ones contributing to decision metric computation) at time t. The computational effort of a sequential decision rule γ may be defined to be the total number of active hypotheses until the end of the test:

$$\Phi_\gamma = \left[\sum_{t=1}^{T} M_t \right] \tag{5.2}$$

For the M-ary sequential probability ratio test (MSPRT) [11], $M_t = M$ for $t = 1, ..., T$, therefore $\Phi_{\text{MSPRT}} = MT$. For $k \in \mathcal{M}$, let $p_n^{(k)}$ denote the conditional probability that $u = k$ given the first n observations:

$$p_n^{(k)} = \frac{f_k(\mathbf{x}^n)}{\sum_{l \in \mathcal{M}} f_l(\mathbf{x}^n)}, \tag{5.3}$$

where $\mathbf{x}^n = (x_1, ..., x_n)$. The sequential tests of concern operate on the vector of posterior probabilities

$$\boldsymbol{p}_n = \left(p_n^{(0)}, ..., p_n^{(M-1)} \right). \tag{5.4}$$

In order to assess the feasibility of complexity reduction we are first interested in identifying as quickly as possible (based on observed data) a subset of r indices that is likely to contain the correct index with high probability.

5.3.1 A two-step sequential decision rule: Rejection of unlikely hypotheses

Consider a two-step sequential decision rule which eliminates a high-probability r-subset, renormalizes the probability vector, and applies an MPSRT on the remaining $M - r$ hypotheses to deliver the final decision. The first step has the risk of rejecting the correct hypothesis, in which case the second part of the decision rule is an MSPRT under an incorrect model, which will be studied in Section 5.5.

Step 1: Define

$$L_r(a) = \min\left\{ n \geq 1 : \sum_{k\in\mathcal{R}} p_n^{(k)} \leq a(1+a)^{-1} \text{ for some } r\text{-subset } \mathcal{R} \right\}. \tag{5.5}$$

Let $\mathcal{R}^*$ be the set of indices that are rejected (those belonging to the smallest r entries in $\boldsymbol{p}_{L_r(a)}$).

$$\mathcal{R}^* = \{j \in \mathcal{M} : p_{L_r(a)}^{(j)} < a\} \tag{5.6}$$

$$\hat{\mathcal{M}} = \mathcal{M} - \mathcal{R}^*. \tag{5.7}$$

Renormalize the surviving posterior probabilities: For $j \in \hat{\mathcal{M}}$,

$$\hat{p}_n^{(j)} = \frac{\hat{\pi}_j f_j(\mathbf{x}^n)}{\sum_{l\in\hat{\mathcal{M}}} \hat{\pi}_l f_l(\mathbf{x}^n)} \tag{5.8}$$

where for $j \in \hat{\mathcal{M}}$, $\hat{\pi}_j = \pi_j / \sum_{l\in\hat{\mathcal{M}}} \pi_l$.

Step 2: MPSRT with parameter vector $\mathbf{b} = (b_j)_{j\in\hat{\mathcal{M}}}$ on $\left\{\hat{p}_n^{(j)},\ j \in \hat{\mathcal{M}}\right\}$ with optional truncation:

$$N(\mathbf{b}) = \min\left\{ n \in \left[L_r(a), -\frac{\log(b_j)}{\eta_j} \right] : \exists j \in \hat{\mathcal{M}} \ni \hat{p}_n^{(j)} > b_j(1+b_j)^{-1} \right\} \tag{5.9}$$

$$\hat{u} = \arg\max_{j\in\hat{\mathcal{M}}} \hat{p}_N^{(j)}. \tag{5.10}$$

There are three important criteria that jointly determine the performance of this two-step design in comparison to existing tests, such as the MSPRT: The average decision error probability, P_e, the average computational effort Φ, and the average (final) stopping-time, $\mathbb{E}N$. A reduction in computational effort may not be so desirable if the number of samples is increased significantly compared to the MSPRT at the same decision error probability. The design parameters are the number r of hypotheses rejected, rejection parameter a, and the MSPRT parameters $\{b_j\}$, along with the optional truncation parameter η. The two-step test will have

$$r \leq \frac{M-1}{2}, \tag{5.11}$$

$$\max_k b_k \leq a \tag{5.12}$$

so that upon error-free pruning, the second stage will behave as an MSPRT on the set of surviving hypotheses with modified prior probabilities: Suppose $u = j$.

$$\begin{aligned}
&\mathbb{P}_j \ \left(\exists n < L_r(a) \ni \hat{p}_n^{(j)} > (1+b_j)^{-1}\right) && (5.13)\\
&\overset{(a)}{\leq} \ \mathbb{P}_j \left(\exists n < L_r(a) \ni \sum_{k \in \hat{\mathcal{M}} - \{j\}} \hat{p}_n^{(k)} \leq b_j (1+b_j)^{-1}\right) \\
&\overset{(b)}{\leq} \ \mathbb{P}_j \left(\exists n < L_r(a), \mathcal{R} \ni |\mathcal{R}| = r, \ \sum_{k \in R} \hat{p}_n^{(k)} \leq a(1+a)^{-1}\right) \\
&\overset{(c)}{=} \ 0 && (5.14)
\end{aligned}$$

where (a) follows from the fact that $\sum_{k \in \hat{\mathcal{M}}} \hat{p}_n^{(k)} = 1$, (b) follows from (5.11) and (5.12), and finally (c) follows from the definition of $\hat{\mathcal{M}}$. Then it follows that N is the first time one of the estimated posterior probabilities $\hat{p}_n^{(j)}$ exceeds $1/(1+b_j)$.

5.3.2 Decision Errors

The two-step rule can fail in two ways: The correct hypothesis may be rejected at the first setp, or the correct hypothesis survives the first step, but is not the eventual MSPRT decision:

Suppose H_j is the correct hypothesis. The probability that H_j is falsely rejected in the first step was bounded as (refer to the equation):

$$\mathbb{P}_j(j \in \mathcal{R}^*) \leq \frac{a}{\pi_j}. \tag{5.15}$$

The second part of the error probability is bounded by the error probability of the embedded MSPRT on indices $\hat{\mathcal{M}}$:

$$\begin{aligned} \mathbb{P}_j(\hat{u} \neq j, j \in \hat{\mathcal{M}}) &\leq b_j \sum_{k \in \hat{\mathcal{M}}-j} \frac{\hat{\pi}_k}{\hat{\pi}_j} && (5.16) \\ &= b_j \sum_{k \in \hat{\mathcal{M}}-j} \frac{\pi_k}{\pi_j} && (5.17) \\ &\leq \frac{b_j}{\pi_j}, && (5.18) \end{aligned}$$

therefore

$$P_{e,j} \leq \frac{a + b_j}{\pi_j}. \tag{5.19}$$

For equally likely hypotheses, (5.18) can be tightened as $\mathbb{P}_j(\hat{u} \neq j, j \in \hat{\mathcal{M}}) \leq (M - r - 1)b_j$, yielding $P_{e,j} \leq Ma + (M - r - 1)b_j$.

1. Section 5.4 describes the number of samples necessary to *reject* a certain hypothesis with a given reliability via the likelihood ratios. The analysis develops to determine the asymptotics of the first time $L_r(a)$ that the combined likelihood of an r-subset of hypotheses falls below a. In particular, it is shown that

$$\frac{-\log(a)}{L_r(a)} \overset{\text{a.s.-}\mathbb{P}_j}{\longrightarrow} D(f_j, f_{[j]_{M-r}}) \quad \text{as} \quad a \to 0 \tag{5.20}$$

where $\{D(f_j, f_{[j]_m})\}_{m\in\mathcal{M}}$ is an ascending order of $\{D(f_j, f_m)\}_{m\in\mathcal{M}}$.

2. Under the separability assumptions (5.11), (5.12), with probability at least $1 - a/\pi_j$ under H_j, the stopping-time N is the first $n \geq 1$ such that $\hat{p}_n^{(k)}$ exceeds $(1 + b_k)^{-1}$ for some $k \in \hat{M}$, and the theory of Baum and Veeravalli [11] applies. However, with probability at most a/π_j under H_j, the second part of the test is an MSPRT under an incorrect model. Section 5.5 examines the behavior of the mismatched MSPRT. It is shown that

$$\frac{-\log b_j}{N(\mathbf{b})} \overset{\text{a.s.-}\mathbb{P}_j}{\leq} D(f_j, f_{[j]_1}) \quad \text{as} \quad \max_k b_k \to 0.. \tag{5.21}$$

However, with probability at most a/π_j, the MPSRT stage is under an incorrect model in which case the final-stopping time may not be finite, especially if the KL distances are arithmetic. Section 5.5 develops the theory of MSPRTs under model mismatch. The results are specialized to the two-step test. In any case the truncation at $t = -\log(b_j)/\eta_j$ with $\eta_j = D(f_j, f_{[j]_1})) + \epsilon$ may be used.

3. The asymptotically optimum MSPRT offers an average complexity of

$$\Phi_{(\text{MSPRT})} \overset{\mathbb{P}_j}{\leq} M \cdot \frac{-\log(\pi_j P_{e,j})}{D(f_j, f_{[j]_1})}. \tag{5.22}$$

whereas the average complexity of the two two-step test is bounded above as

$$\Phi \overset{\mathbb{P}_j}{\leq} (M - r)\frac{-\log(b_j)}{D(f_j, f_{[j]_1})} + r\frac{-\log(a)}{D(f_j, f_{[j]_{M-r}})}. \tag{5.23}$$

where $P_{e,j} \leq Ma + (M - r)\max_j b_j$ for equally likely hypotheses.

Example 1 *Consider the detection of one of eight PSK points in additive white Gaussian noise:*

$$H_k : s = \exp\left(j\frac{2\pi k}{8}\right) \tag{5.24}$$

The observables are $X_t = s + Z_t$, *where* $\{Z_t\}$ *are iid samples of complex zero-mean Gaussian noise with* $Var(Re\ Z_t) = Var(Im\ Z_t) = 1$. *The detection problem is symmetric in all hypotheses. The MSPRT(***a***) with* $a_k = 0.011$ *achieves* $P_{e,k} = 0.001$, $k = 0, ..., 7$. *with* $\mathbb{E}_k N = 45$. *The average computational effort of this test is* $\mathbb{E}_k \Phi = 360$. *At the same* $P_{e,k}$ *level, the two-step design with* $a = 0.006$, $b = 0.0055$ *prunes* $r = 3$ *hypotheses after an average of* $\mathbb{E}_k L_3 = 6.8$ *observations. The average final stopping-time (with no truncation) experiences an increase* $\mathbb{E}_k N = 50.2$*; with an overall average computational effort of* $\mathbb{E}_k \Phi = 237.4$, *a thirty-five percent saving over the effort of MSPRT.*

This simple experiment demonstrates that pruning hypotheses can lead do significant savings in computational effort with reasonable increase in final-stopping time. In the following, we analyze the risk of pruning hypotheses with simple likelihood-ratio rules. As in [11], we will assume the existence of positive numbers $D(f_j, f_i)$ such that

$$\frac{1}{n} \log \frac{f_j(\boldsymbol{X}^n)}{f_i(\boldsymbol{X}^n)} \overset{\text{a.s.-}\mathbb{P}_j}{\longrightarrow} D(f_j, f_i) \qquad \text{as} \quad n \to \infty \tag{5.25}$$

and also that the family

$$\left\{ \frac{1}{n} \log \frac{f_j(\boldsymbol{X}^n)}{f_i(\boldsymbol{X}^n)} \right\} \tag{5.26}$$

is uniformly integrable under $\mathbb{P}_j$ for any i, j.

5.4 Rejecting hypotheses by weak decisions

5.4.1 Stopping times for weak decisions

Probability ratio tests deliver decisions as to what the true hypothesis is based on whether the ratio of a linear combination of likelihoods of exceeds a predeter-

mined threshold. Analogously, a *weak* decision could be delivered against those hypotheses with likelihoods small compared to those of others.

Consider the first time that a posterior probability becomes smaller than a, a positive parameter less than 1. Define, for each $k \in \mathcal{M}$, the stopping times

$$L^{(k)}(a) = \inf\left\{n \geq 1 : p_n^{(k)} \leq a(1+a)^{-1}\right\}. \tag{5.27}$$

The following lemma establishes the asymptotics of $L^{(k)}(a)$ under various distributions.

Lemma 4 *For $k, j \in \mathcal{M}$, $k \neq j$,*

$$L^{(k)}(a) \to \infty \qquad a.s.\text{-}\mathbb{P}_j \ \ as \ \ a \to 0, \tag{5.28}$$

$$\frac{L^{(k)}(a)}{-\log a} \to \frac{1}{D(f_j, f_k)} \qquad a.s.\text{-}\mathbb{P}_j \ \ as \ \ a \to 0. \tag{5.29}$$

Moreover, $L^{(k)}(a) \to \infty$ $a.s.$-$\mathbb{P}_k$ as $a \to 0$.

Proof 8 *The proof follows the techniques of [11]. To prove the first statement, write*

$$\begin{aligned}
&\mathbb{P}_j\left(L^{(k)}(a) \leq n\right) \\
&= \mathbb{P}_j\left(\max_{1\leq m\leq n} \sum_{l\neq k} \frac{\pi_l}{\pi_k} \prod_{t=1}^{m} \frac{f_l(X_t)}{f_k(X_t)} \geq a^{-1}\right) \\
&\leq \sum_{l\neq k} \mathbb{P}_j\left(\max_{1\leq m\leq n} \sum_{t=1}^{m} \left[\log\frac{f_l(X_t)}{f_j(X_t)} + \log\frac{f_j(X_t)}{f_k(X_t)}\right] \geq -\log\left(\frac{aM\pi_l}{\pi_k}\right)\right) \\
&\leq \sum_{l\neq k} \mathbb{P}_j\left(\max_{1\leq m\leq n} \sum_{t=1}^{m} \log\frac{f_l(X_t)}{f_j(X_t)} \geq -\frac{1}{2}\log\left(\frac{aM\pi_l}{\pi_k}\right)\right) \\
&\quad + (M-1)\mathbb{P}_j\left(\max_{1\leq m\leq n} \sum_{t=1}^{m} \log\frac{f_j(X_t)}{f_k(X_t)} \geq -\frac{1}{2}\log\left(\frac{aM\pi_l}{\pi_k}\right)\right)
\end{aligned} \tag{5.30}$$

As $a \to 0$, every term in the right-hand side of (5.30) converges to zero.

To prove the second statement,we observe that

$$L^{(k)}(a) = \inf\left\{n : \zeta_n^{(k)} \geq -\log(a)/n\right\} \tag{5.31}$$

where

$$\zeta_n^{(k)} = \frac{1}{n}\log\frac{1-p_n^{(k)}}{p_n^{(k)}} \overset{a.s.\text{-}\mathbb{P}_j}{\longrightarrow} D(f_j, f_k) \quad as \;\; n \to \infty. \tag{5.32}$$

Since $L^{(k)}(a) \overset{a.s.\text{-}\mathbb{P}_j}{\longrightarrow} 0$ *as* $a \to 0$, $\zeta^{(k)}_{L^{(k)}(a)} \overset{a.s.\text{-}\mathbb{P}_j}{\longrightarrow} D(f_j, f_k)$ *as* $a \to 0$. *Also, by definition,*

$$\zeta^k_{L^{(k)}(a)} \geq \frac{-\log(a)}{L^{(k)}(a)}$$
$$\zeta^k_{L^{(k)}(a)-1} \leq \frac{-\log(a)}{L^{(k)}(a)-1}.$$

therefore the desired result follows from taking limsup and liminfs.

The last statement follows from the fact that for $l \neq k$,

$$\zeta_n^{(k)} \geq \frac{1}{n}\log\frac{f_k(\boldsymbol{X}^n)}{f_l(\boldsymbol{X}^n)} \to -D(f_k, f_l) < 0. \tag{5.33}$$

■

Now consider the first time *any* of the posterior probabilities falls below a. Define,

$$\begin{aligned} L(a) &= \inf\left\{n \geq 1 : \min_{k\in\mathcal{M}} p_n^{(k)} < a(1+a)^{-1}\right\} \\ &= \min_{k\in\mathcal{M}} L^{(k)}(a), \end{aligned} \tag{5.34}$$

so that for any j, $L(a) \overset{\text{a.s.-}\mathbb{P}_j}{\longrightarrow} \infty$ as $a \to 0$, and

$$\frac{L(a)}{-\log(a)} \overset{\text{a.s.-}\mathbb{P}_j}{\longrightarrow} \frac{1}{\max_{k\in\mathcal{M}-\{j\}} D(f_j, f_k)}. \tag{5.35}$$

For a fixed small threshold, the time it takes to deliver a weak decision is inversely proportional to the KL distance from the true distribution to the farthest one.

For strong (MSPRT) decisions, the decision time is inversely proportional to the KL distance from the true distribution to the nearest one [11].

It is also possible to reject a *subset* of the original set of hypotheses at once. Let $\mathcal{R}$ be a proper subset of $\mathcal{M}$. Define

$$L^{\mathcal{R}}(a) = \min\left\{ n : \sum_{k\in\mathcal{R}} p_n^{(k)} < a(1+a)^{-1} \right\}. \tag{5.36}$$

We have the following lemma.

Lemma 5 *If $j \notin \mathcal{R}$, as $a \to 0$,*

$$\frac{L^{\mathcal{R}}(a)}{-\log(a)} \overset{a.s.\text{-}\mathbb{P}_j}{\longrightarrow} \frac{1}{\min_{k\in\mathcal{R}} D(f_j, f_k)}. \tag{5.37}$$

If $j \in \mathcal{R}$, then $L^{\mathcal{R}}(a) \overset{a.s.\text{-}\mathbb{P}_j}{\longrightarrow} \infty$ as $a \to 0$.

Proof 9 *Fix $j \in \mathcal{M}$, $j \notin \mathcal{R}$. First,*

$$\begin{aligned} \frac{\pi_j f_j(\boldsymbol{x}^n)}{\pi_j f_j(\boldsymbol{x}^n) + \sum_{k\in\mathcal{R}} \pi_k f_k(\boldsymbol{x}^n)} &\geq (1+a)^{-1} \\ &\Leftrightarrow \frac{\sum_{k\in\mathcal{R}} \pi_k f_k(\boldsymbol{x}^n)}{\pi_j f_j(\boldsymbol{x}^n) + \sum_{k\in\mathcal{R}} \pi_k f_k(\boldsymbol{x}^n)} \leq a(1+a)^{-1} \\ &\Rightarrow \frac{\sum_{k\in\mathcal{R}} \pi_k f_k(\boldsymbol{x}^n)}{\sum_{l\in\mathcal{M}} \pi_l f_l(\boldsymbol{x}^n)} \leq a(1+a)^{-1} \\ &\Leftrightarrow \sum_{k\in\mathcal{R}} p_n^{(k)} \end{aligned} \tag{5.38}$$

therefore

$$L^{\mathcal{R}}(a) \leq \min\left\{ n \geq 1 : \frac{\pi_j f_j(\boldsymbol{x}^n)}{\pi_j f_j(\boldsymbol{x}^n) + \sum_{k\in\mathcal{R}} \pi_k f_k(\boldsymbol{x}^n)} \geq (1+a)^{-1} \right\} \tag{5.39}$$

where the right hand-side is an MSPRT stopping time whose asymptotics is inversely proportional to the minimum KL distance from f_j to the nearest f_k, where $k \in \mathcal{R}$. Therefore, as $a \to 0$

$$\frac{L^{\mathcal{R}}(a)}{-\log(a)} \overset{a.s.\text{-}\mathbb{P}_j}{\leq} \frac{1}{\min_{k\in\mathcal{R}} D(f_j, f_k)} \tag{5.40}$$

Also, for any $k \in \mathcal{R}$, $L^{\mathcal{R}}(a) \geq L^{(k)}(a)$ *and therefore* $L^{\mathcal{R}}(a) \geq \max_{k\in\mathcal{R}} L^{(k)}(a)$*:*

$$\frac{L^{\mathcal{R}}(a)}{-\log(a)} \geq \max_{k\in\mathcal{R}} \frac{L^{(k)}(a)}{-\log(a)} \tag{5.41}$$

$$\overset{a.s.\text{-}\mathbb{P}_j}{\geq} \max_{k\in\mathcal{R}} \frac{1}{D(f_j, f_k)} \tag{5.42}$$

$$= \frac{1}{\min_{k\in\mathcal{R}} D(f_j, f_k)} \tag{5.43}$$

using (5.35). The second statement follows from tha fact that if $j \in \mathcal{R}$*, then* $L^{\mathcal{R}}(a) \geq L^{(j)}(a)$*.*

This leads to first time when the combined posterior probability of any r of the M hypotheses falls below a predetermined level. Recall (Section 5.3):

$$L_r(a) = \min\left\{ n \geq 1 : \sum_{k\in\mathcal{R}} p_n^{(k)} \leq a(1+a)^{-1} \text{ for some } r\text{-subset } \mathcal{R} \right\}. \tag{5.44}$$

Theorem 8 *For any* $j \in \mathcal{M}$*, and any* $r < M$*,*

$$\frac{L_r(a)}{-\log(a)} \overset{a.s.\text{-}\mathbb{P}_j}{\longrightarrow} D(f_j, f_{[j]_{M-r}}) \tag{5.45}$$

where

$$D(f_j, f_{[j]_1}) \leq D(f_j, f_{[j]_2}) \leq ... \leq D(f_j, f_{[j]_{M-1}}) \tag{5.46}$$

is an ascending order of $\{D(f_j, f_k),\ k \in \mathcal{M} - \{j\}\}$*.*

Proof 10 *The proof follows from the fact that*

$$L_r(a) = \min_{\mathcal{R}\subset\mathcal{M}:|\mathcal{R}|=r} L^{\mathcal{R}}(a) \tag{5.47}$$

by using Lemma 5.

In the special case of $r = M - 1$, $L_{M-1}(a)$ is exactly the the first time one of posterior probabilities exceeds $(1+a)^{-1}$, i.e. an MSPRT final stopping time:

$$\frac{L_{M-1}(a)}{-\log(a)} \overset{\text{a.s.-}\mathbb{P}_j}{\longrightarrow} D(f_j, f_{[j]_1}). \tag{5.48}$$

5.4.2 Error probability associated with a weak decision

Consider the weak decision-rule (L, v)

$$\begin{aligned} L(a) &= \min\{n \geq 1 : \min_{k \in \mathcal{M}} p_n^{(k)} < a(1+a)^{-1}\} \\ \nu(a) &= \{k \in \mathcal{M} : p^{(k)}(\boldsymbol{X}^{L(a)}) < a(a+1)^{-1}\} \end{aligned}$$

which stops as soon as the posterior probability of one of the hypotheses falls below $a(1+a)^{-1}$, and rejects the corresponding hypothesis.

Under H_j, the weak decision rule makes an error is H_j is among the rejected. It is easily verified that

$$\pi_j \mathbb{P}_j(\nu(a) = j) \leq \frac{a}{1+a} \sum_{k \in \mathcal{M}} \pi_k \mathbb{P}_k(\nu = j) \tag{5.49}$$

therefore

$$\mathbb{P}_j(\nu(a) = j) \leq a/\pi_j. \tag{5.50}$$

Also, summing over j on both sides of (5.49) yields,

$$P_e = \text{Prob}(\nu(a) = u) \leq \frac{a}{1+a} \tag{5.51}$$

where u is the correct state.

For the first step of the two-step test of Section 5.3, the above steps apply:

$$\mathbb{P}_j(j \in \mathcal{R}^*) \leq \frac{a}{\pi_j} \tag{5.52}$$

$$P_{e,\text{I}} = \text{Prob}(u \in \mathcal{R}^*) \leq \frac{a}{1+a}. \tag{5.53}$$

5.5 Mismatched M-ary sequential probability ratio tests

When the correct hypothesis is mistakenly rejected, the second stage of the two-step reduced-complexity test introduced in Section 5.3 is a mismatched MSPRT.

For perfect models, the asymptotical optimality of the MSPRT with respect to the average sample size was first established by [11]. Dragalin, Tartakovsky and Veeravalli [59] strengthened this result by showing that the optimality is true for any moment of the sample-size. The authors asked how the MSPRT would perform under incorrect models in [11].

5.5.1 MSPRT under model mismatch

Definition 1 (MSPRT [11]) *For $j \in \mathcal{M}$, let*

$$N^{(j)}(a_j) = \inf\left\{n : \tilde{p}_n^{(j)} > (1+a_j)^{-1}\right\}. \tag{5.54}$$

For the model (5.1), the M-ary sequential probability ratio test

$$\gamma_{MPSRT}(\boldsymbol{a}) = (N, \hat{u}) \tag{5.55}$$

with parameters $\boldsymbol{a} = (a_0, a_1, ..., a_{M-1})$, $0 < a_j < 1$ has

$$N = \min_{j\in\mathcal{M}} N^{(j)}(a_j), \quad \hat{u} \;=\; \arg\max_{j\in\mathcal{M}} p_n^{(j)} \tag{5.56}$$

Now assume that when the nature selects index $u = 0$, the true probability density function of the observables is $\tilde{f}_0$, such that

$$\frac{1}{n}\log\frac{\tilde{f}_0(\boldsymbol{X}^n)}{f_k(\boldsymbol{X}^n)} \overset{\text{a.s.-}\mathbb{P}_{\tilde{f}_0}}{\longrightarrow} D(\tilde{f}_0, f_k) > 0, \tag{5.57}$$

and the model mismatch is mild enough that

$$D(\tilde{f}_0, f_0) < \min_{k\neq 0} D(\tilde{f}_0, f_k), \tag{5.58}$$

i.e., the true density resembles the hypothesized density for $u = 0$ more than any other model hypothesized density in the KL sense. Following the steps of Lemma 5.1 of [11], one can show that

$$N_j(a_j) \overset{\text{a.s.-}\mathbb{P}_{\tilde{f}_0}}{\longrightarrow} \infty \quad \text{as} \quad P_{e,0} \to 0 \tag{5.59}$$

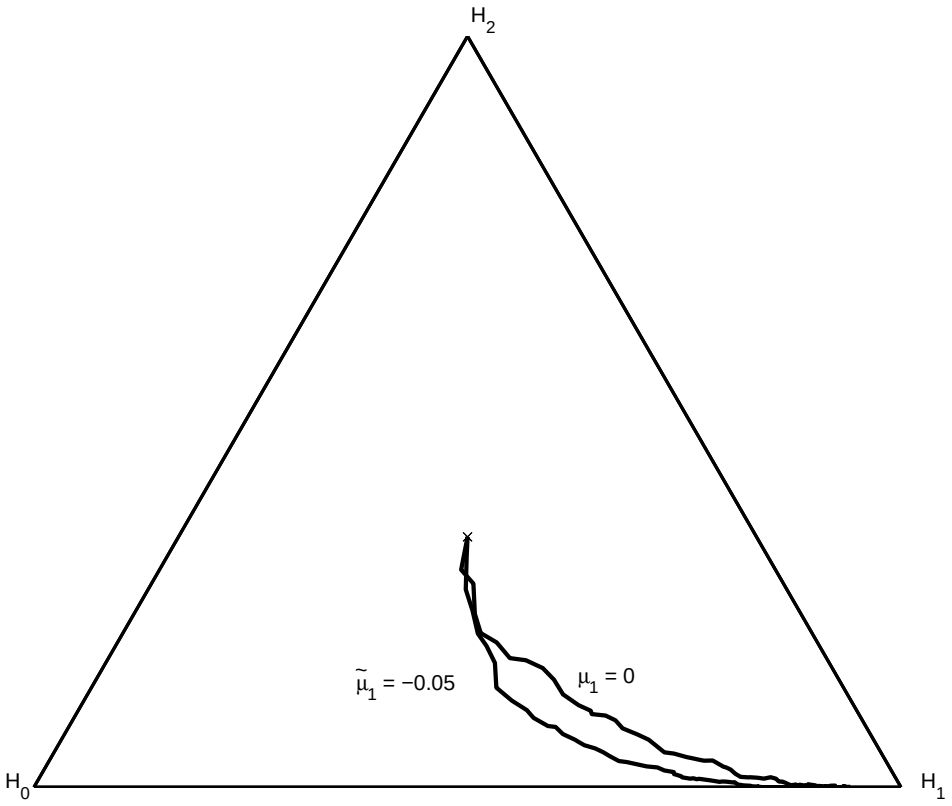

Figure 5.1: The *average* path of the posterior probability vector for the first 75 steps, averaged over 250 experiments, when $X_t \sim \Phi(\tilde{\mu}_1, 0.5)$, $\tilde{\mu}_1 = \delta = -0.05$. The average path is longer compared to the no-mismatch case.

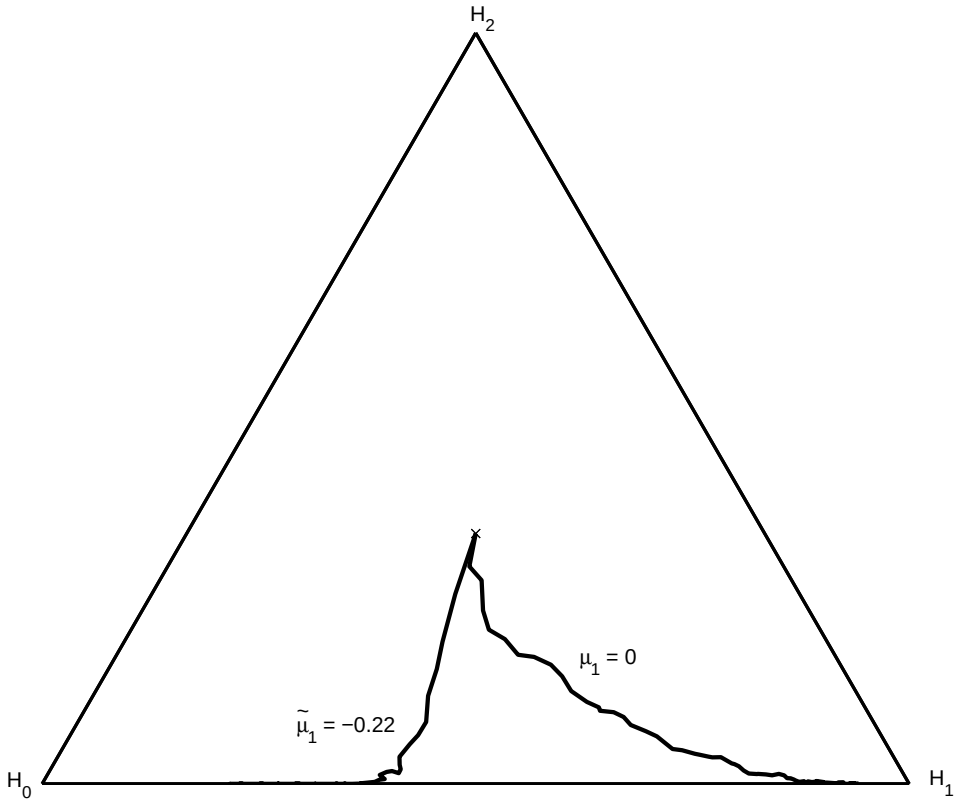

Figure 5.2: The *average* path of the posterior probability vector for the first 75 steps, averaged over 250 experiments, when $X_t \sim \Phi(\tilde{\mu}_1, 0.5)$, $\tilde{\mu}_1 = -0.22$.

where $P_{e,0} = \mathbb{P}_{\tilde{f}_0}(\hat{u} \neq 0)$.

With these above assumptions, the following theorem presents the behavior of the stopping-time and the error probability of the MSPRT in the asymptote of small decision error probability. Part (a) of the theorem is related to the final-stopping time of the two-step test described in Section 5.3.

Theorem 9 *Consider the family of M-ary sequential probability ratio tests*

$$\gamma_{MPSRT}(\boldsymbol{a}) = (N, \hat{u}) \tag{5.60}$$

with parameters $\boldsymbol{a} = (a_0, a_1, ..., a_{M-1})$, $0 < a_l < 1$, $\forall l$. Under the assumptions of Section 5.3,

(a)

$$\lim_{P_e \to 0} \frac{-\log a_0}{\mathbb{E}_{f_0} N} = \min_{l \in \mathcal{M}-\{0\}} D(\tilde{f}_0, f_l) - D(\tilde{f}_0, f_0), \tag{5.61}$$

(b)

$$\liminf_{P_e \to 0} \frac{-\log P_{e,0}}{-\log a_0} \geq 1 - \frac{D(\tilde{f}_0, f_0)}{\min_{l \in \mathcal{M}-\{0\}} D(\tilde{f}_0, f_l) - D(\tilde{f}_0, f_0)}. \tag{5.62}$$

Proof 11 *The proof is presented in Section 5.6.*

Example 2 (Three Gaussians) *Let $\phi(m, s^2)$ be the density corresponding to the Gaussian distribution with mean m and variance s^2. Consider the MSPRT(***a***) based on the three-hypotheses model*

$$H_i : f = \phi(\mu_i, \sigma^2), \quad i = 0, 1, 2$$

where $\mu_0 = -0.6$, $\mu_1 = 0$, $\mu_2 = 0.3$ and $\sigma^2 = 0.5$. The KL distances between the densities are

$$D(\phi(\mu_i, \sigma^2), \phi(\mu_j, \sigma^2)) = \frac{(\mu_i - \mu_j)^2}{2\sigma^2}. \tag{5.63}$$

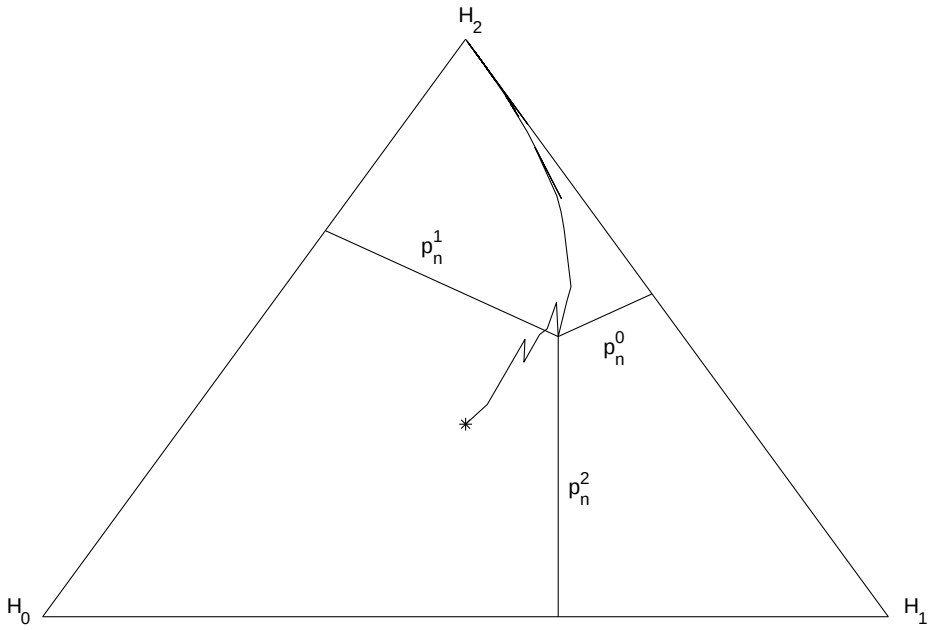

Figure 5.3: The barycentric coordinate system for three hypotheses. The posterior probabilities are represented as distances from a point inside the equilateral triangle to each of the three sides.

Table 5.1: Stopping-time under mismatch: $H_0 : f = \phi(-0.3, 0.5)$, $H_1 : f = \phi(0, 0.5)$, $H_2 : f = \phi(0.6, 0.5)$ where $\phi(m, s^2)$ is a Gaussian density, mean m, variance s^2. The true distribution is $\tilde{f}_1 = \phi(\delta, 0.5)$. For all the cases below $\tilde{f}_1$ is closest to $f_1 = \phi(0, 0.5)$ in the KL sense. MSPRT(**a**) with $a_1 = 0.001$.

δ	$(\mathbb{E}_{f_1} N)_{\text{sim}}$	$P_{e,1}$	$\left(\frac{-\log a_1}{\mathbb{E}_{\tilde{f}_1} N}\right)_{\text{sim}}$	$D(\tilde{f}_1, f_0) - D(\tilde{f}_1, f_1)$
0.01	77.4	1.38×10^{-3}	0.089	0.096
0.005	79.6	1.38×10^{-3}	0.087	0.093
0	82.0	1.46×10^{-3}	0.084	0.090
-0.005	84.5	1.56×10^{-3}	0.082	0.087
-0.01	87.2	1.77×10^{-3}	0.079	0.084
-0.02	93.4	2.43×10^{-3}	0.074	0.078
-0.05	118.6	8.72×10^{-3}	0.058	0.060

Figure 5.1 shows the *average* path of the posterior probability vector for the first 75 steps, averaged over 250 experiments, when $X_t \sim \phi(\tilde{\mu}_1, 0.5)$, $\tilde{\mu}_1 = \delta = -0.05$ as well as when $X_t \sim \phi(\mu_1, 0.5)$ in the probability simplex, where the posterior probability of each hypothesis is represented as the perpendicular distance from the edge of the triangle facing the vertex with the hypothesis label (Figure 5.3). Under the mean mismatch $\tilde{\mu}_1 = -0.05$, the average path is tilted toward the vertex H_0, indicating that the true probability density function $\tilde{f}_1$ is close to f_0 that the model (f_1) indicates.

Table 5.1 shows the stopping time of an MSPRT(a_0, a_1, a_2) with $a_0 = a_1 = a_2 = 0.001$ when the distribution of the observables has a density $\phi(\delta, \sigma^2)$, for various values of mean mismatch δ.

5.5.2 Performance limits

Fano's inequality provides a lower bound on the average probability of decision error when distinguishing between multiple hypotheses. Fano's inequality implies:

$$H_2\left(P_e\right) + P_e \log_2\left(M-1\right) \geq H\left(u|\boldsymbol{X}^N\right) \tag{5.64}$$

where $H(\cdot)$ is the Shannon entropy and $H_2(p) = -p\log(p) - (1-p)\log(1-p)$ is the binary entropy function. Theorem 10 presents a limit on the performance of hypothesis testing by relating the expected sample sizes to decision error probabilities in the asymptote of small decision error probabilities. The proof, detailed in Appendix, uses Fano's inequality (5.64).

Theorem 10 *Let $(N, \hat{u})$ be a sequential test such that $N \to \infty$ almost surely under any hypothesis as $P_e \to 0$. Then, for any given $\epsilon > 0$ and small enough P_e,*

$$\sum_{j=0}^{M-1} \pi_j \left(\min_{i \neq j} D(f_j, f_i) + \epsilon \right) \mathbb{E}_j N \geq -\log P_e - \epsilon. \tag{5.65}$$

Corollary 10.1 *For any sequential test $(N, \hat{u})$ such that $N \to \infty$ a.s. under any hypothesis as $P_e \to 0$,*

$$\limsup_{P_e \to 0} \frac{-\log P_e}{\mathbb{E}N} \leq \max_{j,k \in \mathcal{M}, k \neq j} D(f_j, f_k). \tag{5.66}$$

For the MSPRT with (possible) model mismatch $\{\tilde{f}_j, j \in \mathcal{M}\}$,

$$\liminf_{P_e \to 0} \frac{-\log P_e}{\mathbb{E}N} \geq \min_{j,k \in \mathcal{M}, k \neq j} D(\tilde{f}_j, f_k) - 2D(\tilde{f}_j, f_j). \tag{5.67}$$

5.6 Proof of the error exponents theorems

Lemma 6 *$\{X_n \in F_n, n \in \mathbb{N}\}$ be an identically distributed sequence of random variables with probability density function f_i. Let f_j be another probability density function defined with respect to the same dominating measure. T be an $\{F_n\}$-stopping time with $\mathbb{E}_i T < \infty$. For $b > 0$,*

$$\mathbb{P}_{f_i}\left(\frac{f_j(\boldsymbol{X}^T)}{f_i(\boldsymbol{X}^T)} \geq b^{-1}\right) \leq b.$$

Proof 12 *Since $\mathbb{E}T < \infty$,*

$$\begin{aligned}
&\mathbb{P}_{f_i}\left(\frac{f_j(\boldsymbol{X}^T)}{f_i(\boldsymbol{X}^T)} \geq b^{-1}\right) \\
&= \mathbb{E}_{T=t}\, \mathbb{P}_{f_i}\left(\frac{f_j(\boldsymbol{X}^T)}{f_i(\boldsymbol{X}^T)} \geq b^{-1} | T = t\right) \\
&\leq \mathbb{E}_{T=t}\, b \cdot \mathbb{E}_{f_i}\left(\frac{f_j(\boldsymbol{X}^T)}{f_i(\boldsymbol{X}^T)} | T = t\right) \\
&= b,
\end{aligned}$$

by using Markov's inequality [41].

Proof 13 (Proof of Theorem 9) (a) *Noting that*

$$N_0(a_0) = \min\left\{n : -\frac{1}{n}\log r_0(\boldsymbol{X}^n) > \frac{-\log a_0}{n}\right\}, \tag{5.68}$$

and that as $n \to \infty$

$$-\frac{1}{n}\log r_0(\boldsymbol{X}^n) \overset{a.s.\text{-}\mathbb{P}_{\tilde{f}_0}}{\longrightarrow} \min_{l \in \mathcal{M}-\{0\}} D(\tilde{f}_0, f_l) - D(\tilde{f}_0, f_0) \tag{5.69}$$

we have

$$\lim_{P_e \to 0} \frac{-\log a_0}{N_0(a_0)} = \min_{l \neq 0} D(\tilde{f}_0, f_l) - D(\tilde{f}_0, f_0) \qquad a.s.\text{-}\mathbb{P}_{\tilde{f}_0}. \tag{5.70}$$

Since $N_k(a_k) \to \infty$ as $P_e \to 0$, we have

$$\lim_{P_e \to 0} \frac{-\log a_0}{N(\mathbf{a})} = \min_{l \neq 0} D(\tilde{f}_0, f_l) - D(\tilde{f}_0, f_0) \qquad a.s.\text{-}\mathbb{P}_{\tilde{f}_0}. \tag{5.71}$$

The proof of Theorem 5.1 of [11] is valid under an incorrect model as long as $D(\tilde{f}_0, f_j) - D(\tilde{f}_0, f_0) > 0, \forall j \neq 0$ *the sequence*

$$\{-(1/N)\log(a_0), a_0 \to 0\} \tag{5.72}$$

is uniformly integrable under $\mathbb{P}_{f_j}$*, and therefore a.s.-convergence implies convergence in the mean:*

$$\lim_{P_e \to 0} \frac{-\log a_0}{\mathbb{E}_{\tilde{f}_0} N(\mathbf{a})} = \min_{l \neq 0} D(\tilde{f}_0, f_l) - D(\tilde{f}_0, f_0). \tag{5.73}$$

(b) *Let* $\delta_0 = \min_{l \neq 0} D(\tilde{f}_0, f_l)$ *and* $D_0^* = \min_{l \neq 0} D(\tilde{f}_0, f_l)$*. By part (a) of the theorem (a.s.-convergence and uniform integrability) we have*

$$\mathbb{P}_{\tilde{f}_0}\left(\left|\frac{-\log a_0}{N(D_0^* - \delta_0)} - 1\right| > \epsilon\right) \to 0 \qquad as \;\; a_0 \to 0. \tag{5.74}$$

By the same reasons,

$$\mathbb{P}_{\tilde{f}_0}\left(\left|\frac{1}{N\delta_0}\log\frac{\tilde{f}_0(\boldsymbol{X}^N)}{f_0(\boldsymbol{X}^N)} - 1\right| > \epsilon\right) \to 0 \qquad as \;\; a_0 \to 0. \tag{5.75}$$

Consequently, for small enough a_0*,*

$$\begin{aligned}
&\mathbb{P}_{\tilde{f}_0}\left(\left|\log\frac{\tilde{f}_0(\boldsymbol{X}^N)}{f_0(\boldsymbol{X}^N)} - (-\log a_0)\frac{\delta_0}{D_0^* - \delta_0}\right| > \epsilon\right) && (5.76)\\
&= \mathbb{P}_{\tilde{f}_0}\left(\left|\frac{1}{N\delta_0}\log\frac{\tilde{f}_0(\boldsymbol{X}^N)}{f_0(\boldsymbol{X}^N)} - \frac{-\log a_0}{N(D_0^* - \delta_0)}\right| > \frac{\epsilon}{\delta_0 N}\right) && (5.77)\\
&\leq \mathbb{P}_{\tilde{f}_0}\left(\left|\frac{1}{N\delta_0}\log\frac{\tilde{f}_0(\boldsymbol{X}^N)}{f_0(\boldsymbol{X}^N)} - 1\right| > \frac{\epsilon}{2\delta_0 N}\right)\\
&\quad + \mathbb{P}_{\tilde{f}_0}\left(\left|\frac{-\log a_0}{N(D_0^* - \delta_0)} - 1\right| > \frac{\epsilon}{2\delta_0 N}\right) && (5.78)\\
&\leq \epsilon. && (5.79)
\end{aligned}$$

Therefore, for given $\epsilon > 0$ and small enough a_0,

$$
\begin{aligned}
P_{e,0} = \mathbb{P}_{\tilde{f}_0}\left(\sum_{l\neq 0}\frac{\pi_l}{\pi_0}\frac{f_l(\boldsymbol{X}^N)}{f_0(\boldsymbol{X}^N)} \geq a_0^{-1}\right) & \\
\leq \sum_{l\neq 0}\mathbb{P}_{\tilde{f}_0}\left(\frac{f_l(\boldsymbol{X}^N)}{f_0(\boldsymbol{X}^N)} \geq \frac{a_0^{-1}}{M}\cdot\frac{\pi_0}{\pi_l}\right) & \\
= \sum_{l\neq 0}\mathbb{P}_{\tilde{f}_0}\left(\log\frac{f_l(\boldsymbol{X}^N)}{\tilde{f}_0(\boldsymbol{X}^N)} + \log\frac{\tilde{f}_0(\boldsymbol{X}^N)}{f_0(\boldsymbol{X}^N)} \geq -\log a_0 + \log\left(\frac{\pi_0}{M\pi_l}\right)\right) & \\
\leq \sum_{l\neq 0}\mathbb{P}_{\tilde{f}_0}\left(\log\frac{f_l(\boldsymbol{X}^N)}{\tilde{f}_0(\boldsymbol{X}^N)} - \log(a_0)\left(\frac{\delta_0}{D_0^* - \delta_0}\right) \geq -\log a_0 + \log\left(\frac{\pi_0}{M\pi_l}\right)\right) & \\
+\epsilon & \\
\leq \sum_{l\neq 0}\mathbb{P}_{\tilde{f}_0}\left(\log\frac{f_l(\boldsymbol{X}^N)}{\tilde{f}_0(\boldsymbol{X}^N)} \geq \left(\frac{D_0^* - 2\delta_0}{D_0^* - \delta_0}\right)(-\log a_0) + \log\left(\frac{\pi_0/(M-1)}{\pi_l}\right)\right) & \\
+\epsilon & \\
\leq \sum_{l\neq 0}\frac{M\pi_l}{\pi_0}(a_0)^{1-\frac{\delta_0}{d_0^*-\delta_0}} + M\epsilon = (M-1)\frac{1-\pi_0}{\pi_0}(a_0)^{1-\frac{\delta_0}{D_0^*-\delta_0}} + M\epsilon &
\end{aligned}
$$

that yields

$$
\frac{-\log P_{e,0}}{-\log a_0} \geq \left(1 - \frac{\delta_0}{D_0^* - \delta_0}\right) + \frac{const.}{-\log a_0}.
$$

■

The following lemmas are used to prove Theorem 10 of Section 5.5. For $k, j \in \{0, ..., M-1\}$, let

$$
r_k(\mathbf{x}^n) = \sum_{l=0, l\neq k}^{M-1}\frac{\pi_l}{\pi_k}\frac{f_l(\mathbf{x}^n)}{f_k(\mathbf{x}^n)}, \quad y_{k,j}(\mathbf{x}^n) = \frac{1}{N}\log\frac{\pi_k}{\pi_j} + \frac{1}{N}\log\frac{f_k(\mathbf{x}^n)}{f_j(\mathbf{x}^n)}. \tag{5.80}
$$

Lemma 7 *Under the assumptions of Section 5.3, for any k, j,*

$$
-\frac{1}{N}\log\left(r_k(\boldsymbol{X}^N)\right) \overset{a.s.\text{-}\mathbb{P}_j}{\longrightarrow} \min_{l\neq k} D(f_j, f_k) - D(f_j, f_l) \tag{5.81}
$$

as $P_e \to 0$.

Proof 14 *Fix j. Since $N \to \infty$ a.s.-$\mathbb{P}_j$ as $P_e \to 0$,*

$$-\frac{1}{N}\log\left(r_k(\boldsymbol{X}^N)\right) = -\frac{1}{N}\log\left(\sum_{l\neq k}\exp\left(-Ny_{k,l}(\boldsymbol{X}^N)\right)\right)$$

where

$$\begin{aligned} y_{k,l}(\boldsymbol{X}^N) &= \frac{1}{N}\log\frac{\pi_k}{\pi_l} + \frac{1}{N}\log\frac{f_j(\boldsymbol{X}^N)}{f_l(\boldsymbol{X}^N)} - \frac{1}{N}\log\frac{f_j(\boldsymbol{X}^N)}{f_k(\boldsymbol{X}^N)} \\ &\to D(f_j, f_l) - D(f_j, f_k) \end{aligned}$$

a.s.-$\mathbb{P}_j$ as $P_e \to 0$.

Lemma 8 *For any j,*

$$\lim_{P_e\to 0} r_j(\boldsymbol{X}^N) = 0 \qquad a.s.\text{-}\mathbb{P}_j. \tag{5.82}$$

Proof 15 *Follows from Lemma A.2, by the fact that*

$$-\frac{1}{N}\log\left(r_j(\boldsymbol{X}^N)\right) \to \min_{l\neq k} D(f_j, f_l) > 0 \tag{5.83}$$

and that $N \to \infty$ a.s.-$\mathbb{P}_j$, as $P_e \to 0$.

Theorem 11 *For any given $\epsilon > 0$ and small enough P_e,*

$$\sum_{j=0}^{M-1}\pi_j\left(\min_{l\neq j} D(f_j, f_l) + \epsilon\right)\mathbb{E}_j N \geq -\log P_e - \epsilon. \tag{5.84}$$

Proof 16

$$H\left(u|\boldsymbol{X}^N\right) = \mathbb{E}\left(\sum_{i=0}^{M-1} -Pr\left(u=i|\boldsymbol{X}^N\right)\log Pr\left(u=i|\boldsymbol{X}^N\right)\right)$$

where $Pr(u=i|\boldsymbol{X}^N)$ is the posterior probability of $u=i$ given that $\boldsymbol{X}^N$ is observed. Recognizing that

$$\begin{aligned} Pr(u=i|\boldsymbol{X}^N) &= \frac{\pi_i f_i\left(\boldsymbol{X}^N\right)}{\sum_{j=0}^{M-1}\pi_j f_j\left(\boldsymbol{X}^N\right)} \\ &= \left(1 + r_i(\boldsymbol{X}^N)\right)^{-1}, \qquad i = 0, ..., M-1, \end{aligned} \tag{5.85}$$

we have

$$
\begin{aligned}
H\left(u|\boldsymbol{X}^N\right) &= \mathbb{E}\left(\sum_{i=0}^{M-1}\frac{\log\left(1+r_i(\boldsymbol{X}^N)\right)}{1+r_i(\boldsymbol{X}^N)}\right) && (5.86)\\
&= \sum_{j=0}^{M-1}\pi_j\mathbb{E}_j\left(\sum_{i=0}^{M-1}\frac{\log\left(1+r_i(\boldsymbol{X}^N)\right)}{1+r_i(\boldsymbol{X}^N)}\right) && (5.87)\\
&= \sum_{j=0}^{M-1}\pi_j\mathbb{E}_j\left(\sum_{i=0}^{M-1}\frac{\log\left(1+r_i(\boldsymbol{X}^N)\right)}{1+r_i(\boldsymbol{X}^N)}\right) && (5.88)\\
&\geq \sum_{j=0}^{M-1}\pi_j\mathbb{E}_j\left(\frac{\log\left(1+r_j(\boldsymbol{X}^N)\right)}{1+r_j(\boldsymbol{X}^N)}\right). && (5.89)
\end{aligned}
$$

By Lemma A.2, for $\epsilon > 0$ and small enough P_e, there exists $F \in F_N$ such that $mP_j(F) > 1-\epsilon$ and

$$
\begin{aligned}
&\exp\left(-N(D(f_j, f_{[j]_1})+\epsilon)\right)\\
&\qquad \leq r(\boldsymbol{X}^N) \leq \exp\left(-N(D(f_j, f_{[j]_1})-\epsilon)\right) \quad \text{on } F. && (5.90)
\end{aligned}
$$

We have

$$
\frac{\log(1+r)}{1+r} = r - \frac{3}{2}r^2 + o(r^3) \tag{5.91}
$$

therefore for small enough P_e,

$$
\begin{aligned}
\mathbb{E}_j\left(\frac{\log\left(1+r_j(\boldsymbol{X}^N)\right)}{1+r_j(\boldsymbol{X}^N)}\right) &\geq \mathbb{E}_j\left(\frac{\log\left(1+r_j(\boldsymbol{X}^N)\right)}{1+r_j(\boldsymbol{X}^N)}; F\right) && (5.92)\\
&\geq \mathbb{E}_j\left(r_j(\boldsymbol{X}^N)-\epsilon; F\right) && (5.93)\\
&\geq \mathbb{E}_j\left(r_j(\boldsymbol{X}^N); F\right)-\epsilon && (5.94)\\
&\geq \mathbb{E}_j\left(\exp(-N(D(f_j, f_{[j]_1})+\epsilon)); F\right)-\epsilon && (5.95)\\
&\geq \mathbb{E}_j\left(\exp(-N(D(f_j, f_{[j]_1})+\epsilon))\right)-2\epsilon && (5.96)\\
&\geq \exp\left(-\left(D(f_j, f_{[j]_1})+\epsilon\right)\mathbb{E}_j N\right)-2\epsilon && (5.97)
\end{aligned}
$$

yielding

$$H\left(u|\boldsymbol{X}^N\right) \geq \sum_{j=0}^{M-1} \pi_j \exp\left(-\left(D(f_j, f_{[j]_1}) + \epsilon\right) \mathbb{E}_j N\right) - 2\epsilon \tag{5.98}$$

$$\geq \exp\left(-\sum_{j=0}^{M-1} \pi_j \left(D(f_j, f_{[j]_1}) + \epsilon\right) \mathbb{E}_j N\right) \tag{5.99}$$

Now, for small enough P_e*,*

$$\begin{aligned} &P_e(-\log(P_e) + 1 + \log_2(M-1)) \\ &\quad = -P_e \log(P_e) - (1-P_e)(-P_e) + P_e \log_2(M-1) \\ &\quad \geq H_2(P_e) + P_e \log_2(M-1) \qquad (5.100) \\ &\quad \geq H(u|\boldsymbol{X}^N) \qquad (5.101) \\ &\quad \geq \exp\left(-\sum_{j=0}^{M-1} \pi_j \left(D(f_j, f_{[j]_1}) + \epsilon\right) \mathbb{E}_j N\right) \qquad (5.102) \end{aligned}$$

where (5.100) follows from the fact that $P_e > -(1-P_e)\log(1-P_e)$ *for small enough* P_e*, and (5.101) follows from Fano's inequality. Operating with* $\log(\cdot)$ *on both sides and dividing by* $\log(P_e) < 0$*,*

$$1 + \frac{\log\left(-\log(P_e) + const.\right)}{\log P_e} \leq \left(\log(P_e)\right)^{-1} \sum_{j=0}^{M-1} \pi_j (D(f_j, f_{[j]_1}) + \epsilon)\mathbb{E}_j N. \tag{5.103}$$

Since $-\log P_e \to \infty$ *as* $P_e \to 0$*, we have, for small enough* P_e*,*

$$1 + \epsilon \leq \left(-\log(P_e)\right)^{-1} \sum_{j=0}^{M-1} \pi_j (D(f_j, f_{[j]_1}) + \epsilon)\mathbb{E}_j N, \tag{5.104}$$

and the proof is complete. ■

CHAPTER 6

A robust trellis coded-modulation solution for powerline communications

Powerline communications is a promising technology for next generation home-networking that promises to deliver high-rate, high-quality multiple simultaneous home entertainment streams and other digital content throughout the home using the powerline channel. The physical layer communication technology is rate-adaptive orthogonal frequency division multiplexing (OFDM), matching the transmit signal to the channel conditions that exist between the source and destination nodes.

In its ideal form, the OFDM medium is a collection of parallel (available for simultaneous use), non-interfering (ICI-free), independent, flat, stationary, additive white Gaussian noise (AWGN) channels with different channel gains. The carrier gains of a typical powerline communication channel is stationary for a duration that is much longer than the symbol duration for a megabyte-per-second communication scenario, therefore the transmitter side has access to reliable channel state information. The powerline communication channel, however, suffers from intermittent phase errors and ambiguities caused by the interference from many devices operating along the power line.

In this chapter we study an efficient forward error-correction coding strategy for powerline channels with periodic phase errors. A coding strategy that

can deliver the Avalanche data transmission goals is rate-adaptive trellis-coded modulation across tones. Trellis codes have low-latency, simple decoding and are amenable for rate adaptation. Trellis-coded modulation (TCM) was introduced in [22]. A rate-nonadaptive trellis-coded OFDM was proposed for the European (wireless) digital audio broadcasting (DAB) standard [58].

6.1 Structure of trellis-coded modulation across tones

An OFDM symbol of duration T may be regarded as a vector of Q available tones or subcarriers, received with no intercarrier interference (ICI) and no intersymb ol interference (ISI) under ideal channel conditions. A subset of the available tones provides the bandwidth for data transmission for a specific home entertainment application. Some tones are possibly reserved for channel estimation.

For uncoded transmission, the placement of modulation symbols over the available tones has no bearing on the performance as long as rate assignment over the tones is the same- since there is no memory in the modulation. However, coding introduces memory and the placement of symbols from a codeword sequence plays an important role in decoding delay and error performance if the channel is time-varying. At the heart of the envisioned TCM encoder is a fixed core TCM which consists of a rate-2/3 systematic feedback convolutional encoder driving an 8PSK constellation with natural labeling. Two information bits b_1 and b_2 cause a state transition and the convolutional encoder outputs a nonsystematic coded bit b_0. The three bit label $b_2b_1b_0$ describes the angle $2\pi(4b_2 + b_1 + b_0)/8$.

This core TCM delivers the minimum-rate transmission, 2 bits/symbol. For higher rates, uncoded bits are used to expand the basic constellation through set partitioning. Figure 6.1 illustrates the proposed TCM phase encoder. The

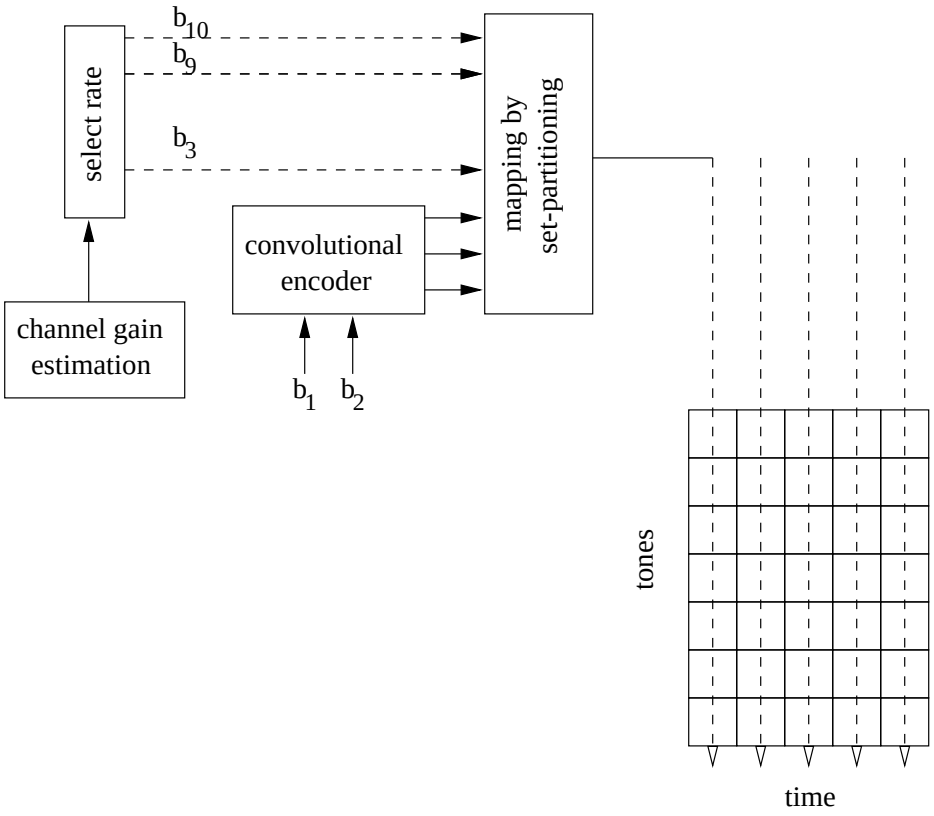

Figure 6.1: Trellis-coded modulation across OFDM tones.

exact structure of the TCM is shaped by many factors, including the nature of the variation of the channel transfer function, the accuracy and complexity of channel estimation, the need to minimize decoding delay.

6.2 Differential phase modulations: Uncoded performance

The phase components of the proposed coded streams are signaled differentially to avoid the need for locked carrier acquisition circuits, thereby simplifying the receiver design. The basic idea behind differential phase signaling is the cumulative encoding of the information-bearing phase [34]. For MPSK modulations,

the transmitted signal x_k ,at discrete symbol instant k , has the form

$$x_k = e^{j\sum_{t=1}^{k}\phi_t} \tag{6.1}$$

where ϕ_t is the uncoded phase at time t.

A discrete-time additive white Gaussian noise (AWGN) channel with an unknown channel carrier phase θ can be represented as

$$y_k = e^{j\theta_k}x_k + n_k \tag{6.2}$$

where k is the discrete time index and n_k is the equivalent complex valued i.i.d. AWGN sample at time k. The carrier phase is assumed to be constant and uniformly distributed over L symbols over which differential detection may be performed. The simplest case of differential detection is using $L = 2$ consecutive symbols. For decision on the uncoded phase ϕ_t, the differential detector compares

$$y_k y_{k-1}^{\dagger} = e^{j\phi_k} + \left(e^{j\theta_k}x_k + n_k\right) n_{k-1}^{\dagger} + \left(e^{-j\theta_{k-1}}x_{k-1}^{\dagger} + n_{k-1}\right) n_k \tag{6.3}$$

to candidate $\exp(j\phi)$ values. Once the differential detection is performed, the equivalent additive noise does not have Gaussian statistics. Moreover, it is correlated across samples and not accounting for this correlation results in a performance penalty: Uncoded 4DPSK requires 2.2 dB and uncoded 8DPSK requires 2.5 dB more to achieve BER = 1e-5 than their coherent counterparts. Noise-predictive filtering may gain a typical of 0.5-1 dB.

It is possible to involve as many symbols in the differential detector as there are in a period of constant channel phase. In fact, for a general N-sequence over which the phase is constant, the maximum-likelihood (ML) test for uncoded phases is given by

$$\left(\hat{\phi_1}, ..., \hat{\phi_N}\right)_{\mathrm{ML}} = \arg \min_{\phi_1,...,\phi_N} \sum_{1\le m<n\le n} \left| y_m y_n^{\dagger} - e^{j\sum_{t=n+1}^{m}\phi_t} \right|^2. \tag{6.4}$$

This decision consists of joint metrics from all pairs of received symbols in the observation intervals. There are $N(N-1)/2$ such pairs. As the observation interval (N) grows, the performance of the non-coherent N-symbol differential detector approaches that of a coherent detector with perfect knowledge of the phase, i.e., the differential detector *learns* the channel phase.

6.3 Trellis-coded differential phase modulations

Trellis-coding for differential phase modulations usually takes the form of a standard PSK-TCM concatenated with a cumulative phase encoder. Since memory is introduced in the phase modulation, codeword maximum-likelihood metrics are not additive over trellis branches, standard Viterbi decoding is suboptimal.

There are two main approaches to decoding a trellis-coded DPSK sequence: Trellis-based decoding and tree-based decoding [57]. Trellis-based decoding relies on the Viterbi algorithm to prune the set of hypothesized paths, either with an additive-suboptimal path metric or with a non-additive suboptimal metric. Tree-based algorithms use the optimum metric for an exhaustive or sparse search of the tree of hypothesized paths. The complex nature of the stack operations (such as sorting and backtracking) makes the tree-based algorithms better candidates for values of N higher than the constraint length of the TCM. For such values of N (e.g., N=10, 20), the simplicity of Viterbi pruning is clearly outweighed by the number of survivor paths retained per state. Trellis-based schemes are more suited to small values of N, typically up to 5. For this reason, a trellis-based low-delay decoding approach is proposed for the receiver.

For N-symbol differential detection of a k bits/symbol trellis, the Viterbi decoder considers for each state, k^{N-1} merging paths and uses a metric that is a

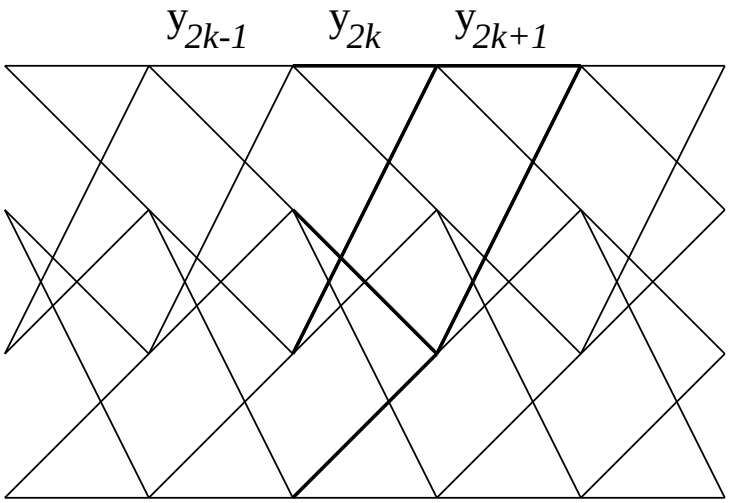

Figure 6.2: $N = 3$ decoding for a 4-state DPSK-TCM.

linear combination of joint metrics

$$\left| y_m y_n^\dagger - e^{j\sum_{t=n+1}^{m}\phi_t} \right|^2 \tag{6.5}$$

where m and n belong to any two symbols from the past N symbols. The full metric for an N-symbol update includes all $N(N-1)/2$ such terms:

$$\sum_{1\le m<n\le n} \left| y_m y_n^\dagger - e^{j\sum_{t=n+1}^{m}\phi_t} \right|^2. \tag{6.6}$$

With the full N-symbol metric, a k bits/symbol trellis with S states requires

$$\frac{1}{k(N-1)} \left(2^{k(N-1)} S \frac{N(N-1)}{2} \right) \tag{6.7}$$

individual metric computations. Figure 6.2 illustrates the trellis update of a 4-state code in the case of an N=3-symbol differential detection. Each of the four merging paths is updated with a metric of the form

$$\delta^2 = \left| y_{2k+1} y_{2k}^\dagger - x_{2k+1} \right|^2 + \left| y_{2k} y_{2k-1}^\dagger - x_{2k} \right|^2 + \left| y_{2k+1} y_{2k-1}^\dagger - x_{2k+1} x_{2k} \right|^2. \tag{6.8}$$

The 2-bits/symbol 8PSK TCMs of Ungerboeck [22] are selected as a baseline to asses the complexity-performance trade-off of multisymbol differential detection. Figure 6.3 displays the BER performance of TCMs with S=64, 16 and 8

state encoders with N=2, N=2 and 3, and N=4 symbol differential detection, respectively. Units of decoding complexity are indicated in parentheses. The pair (S=64, N=2) offers the best trade-off among the candidates. The 16-state TCM with N=2 differential decoding outperforms uncoded 4DPSK by 1dB at BER=10^{-3}, and by 1.5 dB at BER=10^{-4}.

6.4 Performance under channel-phase tracking errors

A second set of considerations on the design of the trellis code arises from the fact that the stationary assumption of channel transfer function is not true in practice. Consequently, channel gain estimation is not error-free. Of most importance to the TCM-based modem design is the phase of the channel.

Standard coherent trellis coded modulations are known to be more sensitive to phase errors as compared to their uncoded counterparts of the same rate as a result of the constellation expansion leading to a smaller angular separation. Experimental measurements indicate that the transfer function of powerline subchannels is subject to periodic state change, resulting from a superposition of multiple periodic behaviors with period 120 Hz, corresponding to several hundred OFDM symbols. Although channel gain amplitude changes are tracked easily and accurately, such is not the case for the phase, especially on a symbol-by-symbol basis. A channel amplitude change is almost certainly accompanied by a channel phase change.

The simplest model for the periodic behavior in the phase of the channel is a phase change, for a single symbol, every T=260 symbols, which is expected to degrade performance when overlooked at the receiver. The most severe phase change for a set-partitioned PSK-TCM is a phase reversal since cosets of such a

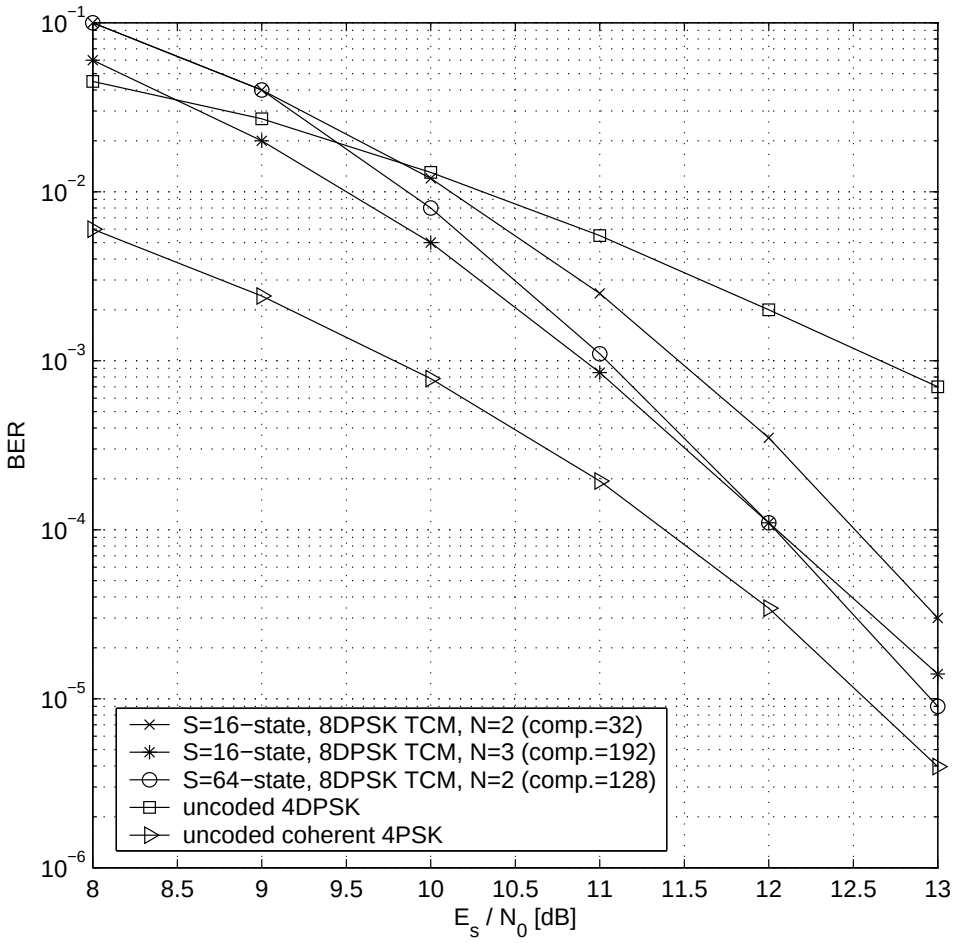

Figure 6.3: Bit error rate performance of 2 bits per symbol 16 and 64-state DPSK-TCMs for $N = 2, 3$ and $N = 3$-symbol differential detection, respectively. For reference, bit error rate performances of coherent and differential uncoded 4PSK are provided.

code have 180 degrees phase symmetry.

Apart from the periodic change of transfer function behavior, the phase of the channel has small fluctuations (typically 5-10 degrees) on a per symbol basis. In our simulations, these fluctuations are modeled as independent phase shifts drawn from a uniform distribution of a predetermined maximum magnitude.

Figure 6.4 displays the performance degradation in 2 bits/symbol trellis-coded and uncoded differential modulations when subjected to a phase change every 260 symbols. Both the coded and uncoded schemes experiences an error floor, the former at BER $= 10^{-2}$, and the latter around BER$=1/T = 3.8 \times 10^{-3}$, both of which are too high for practical purposes. The performance degradation for uncoded 4DPSK scheme is simply due to two unreliable symbols produced by differential detection every 260 symbols. For the coded TCM scheme, Figure 6.5 shows an error event that is due solely to the channel phase reversal that is unaccounted for.

A channel phase reversal produces two consecutive phase-difference reversals when differential-detection is used. With unit energy PSK constellations, a phase-delta reversal is a squared-Euclidean distance shift of $d^2 = 4$. Consequently, if the minimum squared Euclidean distance of the code, $d^2_{\min}$, is less than $2d^2 = 8$, the code may exhibit high error floor under differential detection. No 2 bits/symbol standard 8PSK TCM with 64-states (or less) has a minimum squared-distance greater than 8. With lowering the rate of transmission on the phase to 1 bit/symbol one can achieve greater minimum Euclidean distances thereby avoiding an error floor, however, the coding gain is entirely wiped out (Figure 8).

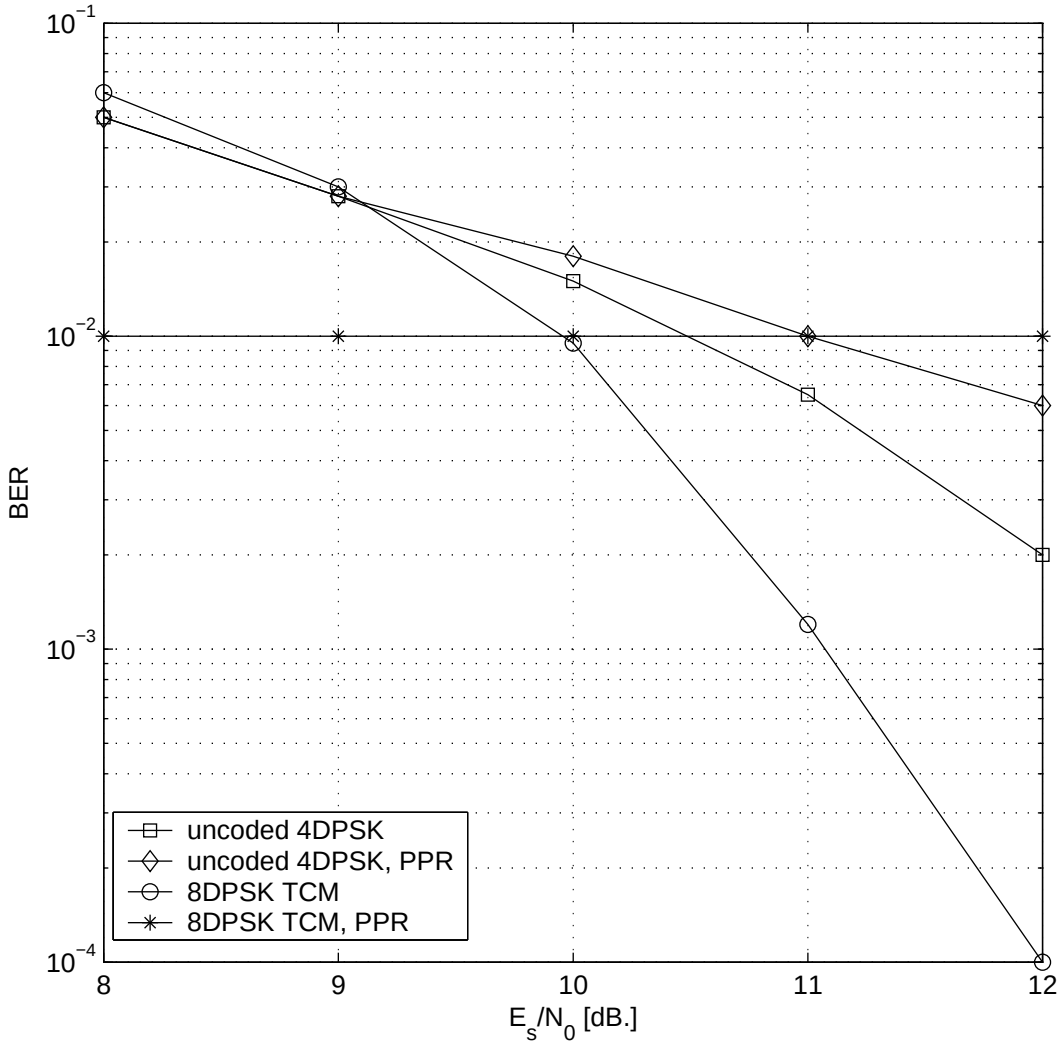

Figure 6.4: 16-state trellis coded and uncoded 2 bits per symbol DPSKs under periodic phase reversals (PPRs) with period 260 symbols.

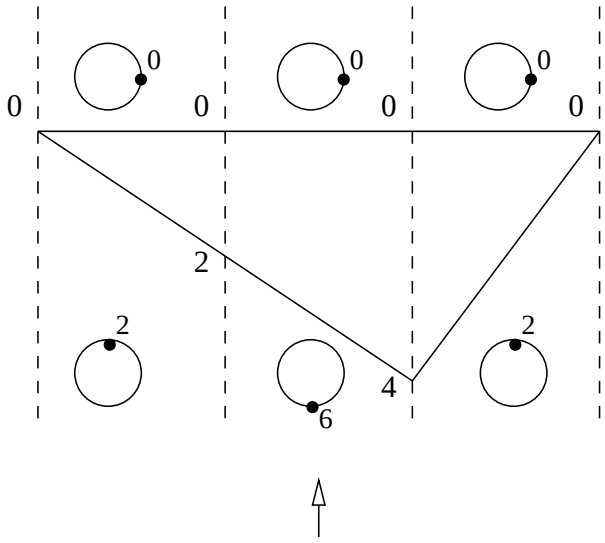

Figure 6.5: An error caused by the channel phase reversal. Even in the absence of receiver noise, the phase reversal causes an incorrect sequence appear closer to the received word.

6.4.1 Erasures Decoding

The sensitivity of trellis-coded phase modulations to sustained carrier-phase offsets has been their major drawback [22]. Similar performance degradation is experienced even with periodic phase errors with a period large compared to the constraint length of the code. However, the knowledge of the location of intermittent phase errors may help restore coded performance when erasures decoding is employed.

When the decoder is provided with information that a set of received symbols is unreliable, it may choose to declare an erasure, that is, simply ignore those received values when detecting a codeword. Necessary conditions to maintain coded performance when using erasures-decoding are:

- The erased rate of the encoder stays strictly less than 1/1. For example, the 2 bits/symbol 8PSK TCM that is being considered cannot deliver coded performance under erasures more often than (and including) one erasure every three symbols. In other words, if one out of three symbols is erased, then the decoder has 3+3+0=6 bits to rely on for 2+2+2=6 bits, which is uncoded transmission.

- The set of erased symbols does not wipe out a constraint length of the code. If as many symbols as the constraint length of the code are erased, those symbols cannot be recovered reliably. The symbol error rate performance will be above $K/2T$ where K is the constraint length of the code and T is the period of the K-erasures. Such a symbol error-rate is readily available by uncoded modulations under the same erasures pattern.

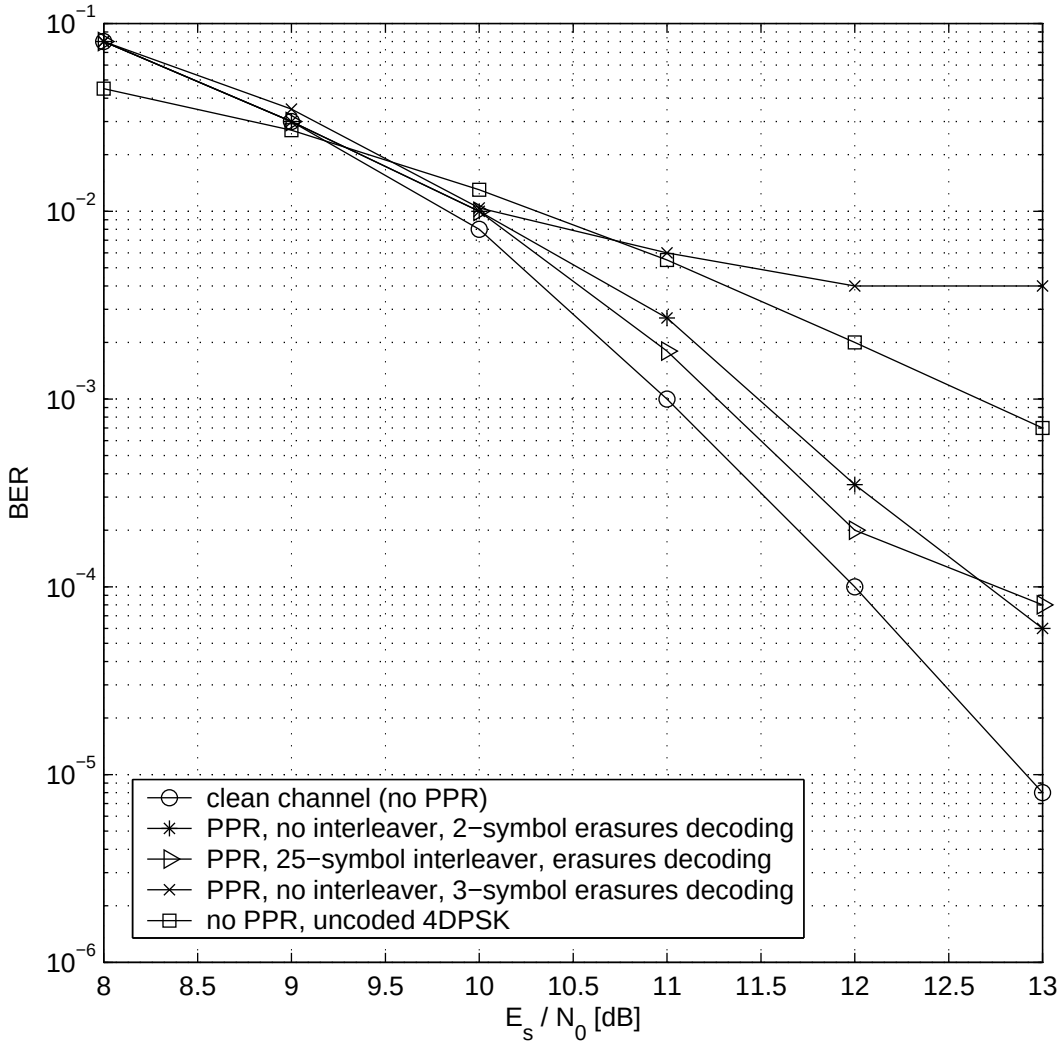

Figure 6.6: Performance of 32-state 2 bits/symbol 8DPSK-TCM under periodic phase reversals of period 260 symbols. Erasures decoding with and without interleaving on the traceback symbols.

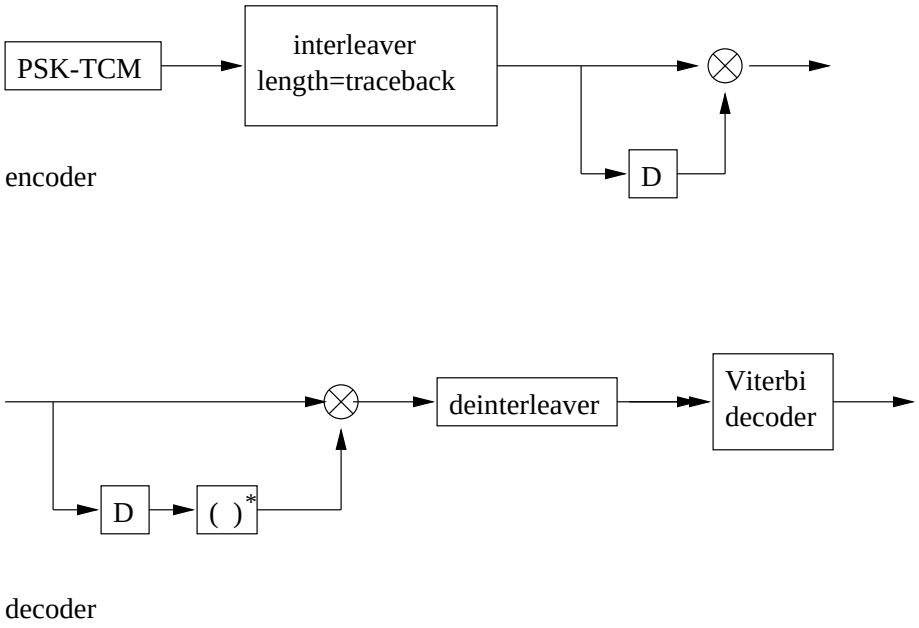

Figure 6.7: Trellis-coded interleaved differential MPSK (MDPSK) and its decoding.

6.4.2 Symbol interleaving

With N=2 differential decoding, a channel phase error of magnitude $\Delta\phi$ affects two consecutive symbols, producing angular mismatch of $(-\Delta\phi, \Delta\phi)$. Before an erasures-decoding at the receiver, interleaving may help to separate these consecutive phase difference errors (Figure 6.7). Figure 6.6 displays the performance of the 32-state 2 bits/symbol 8DPSK TCM under periodic phase reversals of period 260 symbols and with different strategies to mitigate the performance degradation. The interleaver for this application is chosen to have maximal circular spread for its blocklength. The *circular spread* of the interleaver is defined to be the smallest separation between adjacent points when interleaved, where adjacency is defined on a circle. Let π denote the interleaving permutation on integer indices between 0 and $N-1$ (inclusive). The circular spread is defined

as:

$$\text{circular spread } = \max\{|\pi(i) - \pi((i+1) \bmod N)|\} \tag{6.9}$$

A hand-crafted maximal circular-spread (6) interleaver of length N=26 is

$$\begin{aligned} \pi: \quad & 10, 18, 9, 24, 17, 11, 22, 13, 21, 1, 7, 16, \\ & 8, 19, 0, 6, 23, 2, 15, 5, 12, 3, 14, 20, 4 \end{aligned} \tag{6.10}$$

with the corresponding deinterleaver

$$\begin{aligned} \pi^{-1}: \quad & 14, 9, 17, 21, 24, 19, 15, 10, 12, 2, 0, 5, 20, \\ & 7, 22, 18, 11, 4, 1, 13, 23, 8, 6, 16, 3 \end{aligned} \tag{6.11}$$

It should be noted that the interleaver (6.11) and the deinterleaver (6.12) cannot be interchanged. The deinterleaver (6.12) has circular spread of only 2.

6.5 High-rate transmission using M-DAPSK constellations

For rates higher than 2 bits/symbol, not all information bits will be encoded on the phase dimension. The amplitude dimension of the MDPSK constellations is available for information transmission. The simplest 3 bits/symbol solution that builds on the 2 bits/symbol phase-coded system uses an uncoded ring-identifier bit. The 2x8DAPSK constellation is illustrated in Figure 6.8. The rate-2/3 8DSPK TCM selects angle $b_2b_1b_0$ and the uncoded bit b_3 selects the ring. The parameter $\lambda \in (0, 1)$ controls the tradeoff between error protection for the coded bits (b_1b_2) and that for the uncoded (b_3) bit. Larger values of λ_3 provide more protection for the uncoded bit and less protection for the coded bits. For the BER range of 10^{-4} to 10^{-3}, $\lambda = 0.83$ is approximately optimum in delivering the smallest average BER.

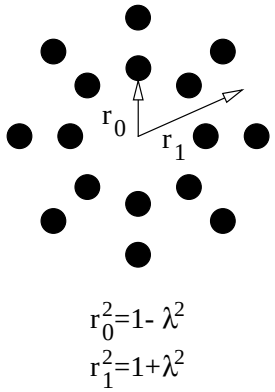

Figure 6.8: Unit energy 16-DAPSK constellation using two 8DPSK rings.

Figure 6.9 displays the BER performance of the 3 bits/symbol scheme that uses a 16-state and a 64-state encoder in comparison to uncoded 8DPSK as well as coherent 8PSK. The use of an uncoded dimension results in no coding gain. The desirable coherent-phase performance is 2 dB away. This simple experiment taught us the fact that amplitude dimension should be coded as well as the phase dimension to realize coding gain for transmission of rate higher than 2 bits/symbol.

The set-partitioning of the two-ring composite constellation depends on the relative magnitudes of the distance between the two rings and the distance between inner PSK points. For $\lambda^2 > 3/4$, the set-partitioning SP-I (Figure 6.10 does yield an uncoded ring bit, since the each of the final cosets consists of one point from the outer ring and one point from the inner ring. Note that SP-I is a valid set-partitioning that would produce a coding gain even with simple trellises using coherent phase modulation. However, differential phase encoding costs an immediate 1-2 dB. Moreover, this set-partitioning is correct for high values of λ, i.e. constellations that favor the amplitude dimension over the phase. For $\lambda = 0.83$, SP-I is not optimal. For $\lambda^2 < 3/4$, the uncoded bit produced by the set-partitioning is not the amplitude bit; it is b_1, which selects the half-plane on

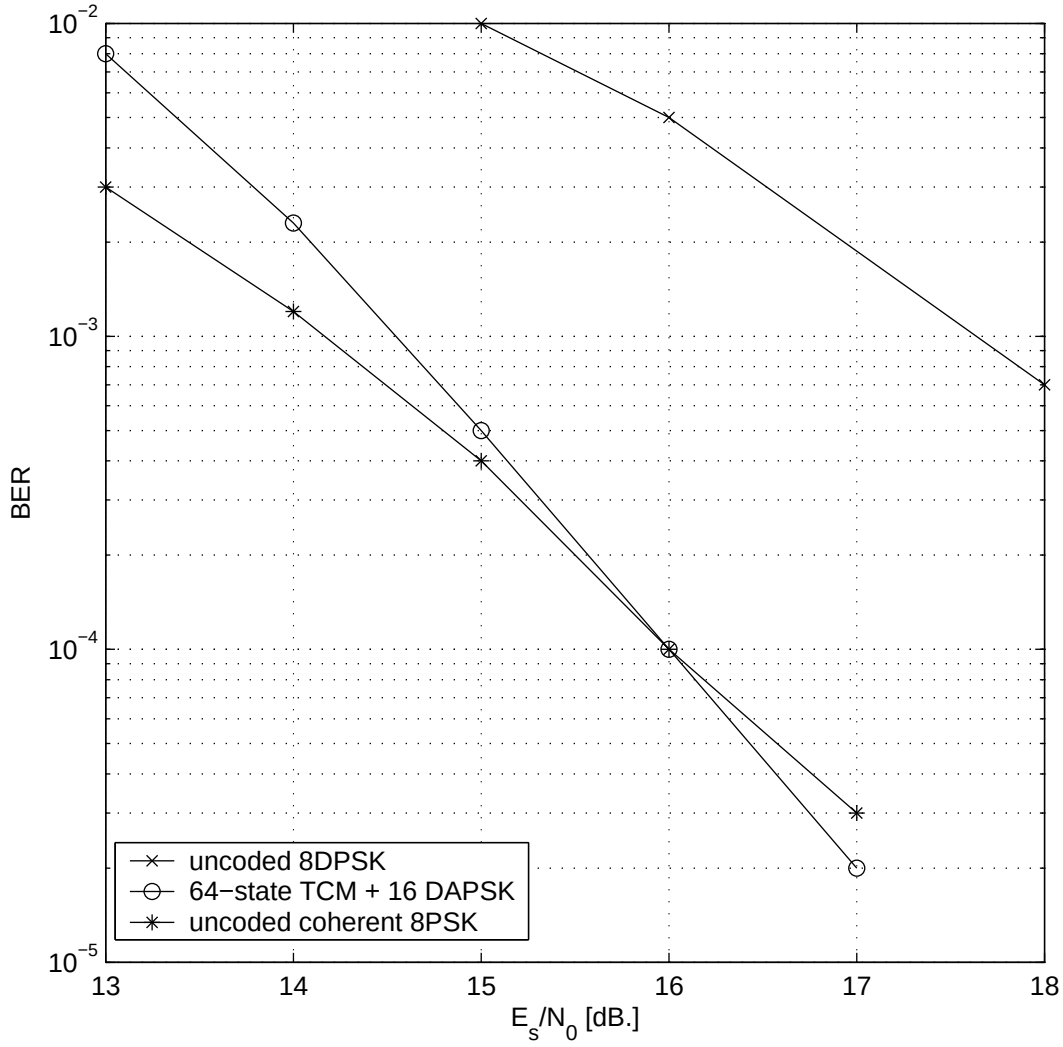

Figure 6.9: The performance of TCM + 16DAPSK with an uncoded amplitude bit.

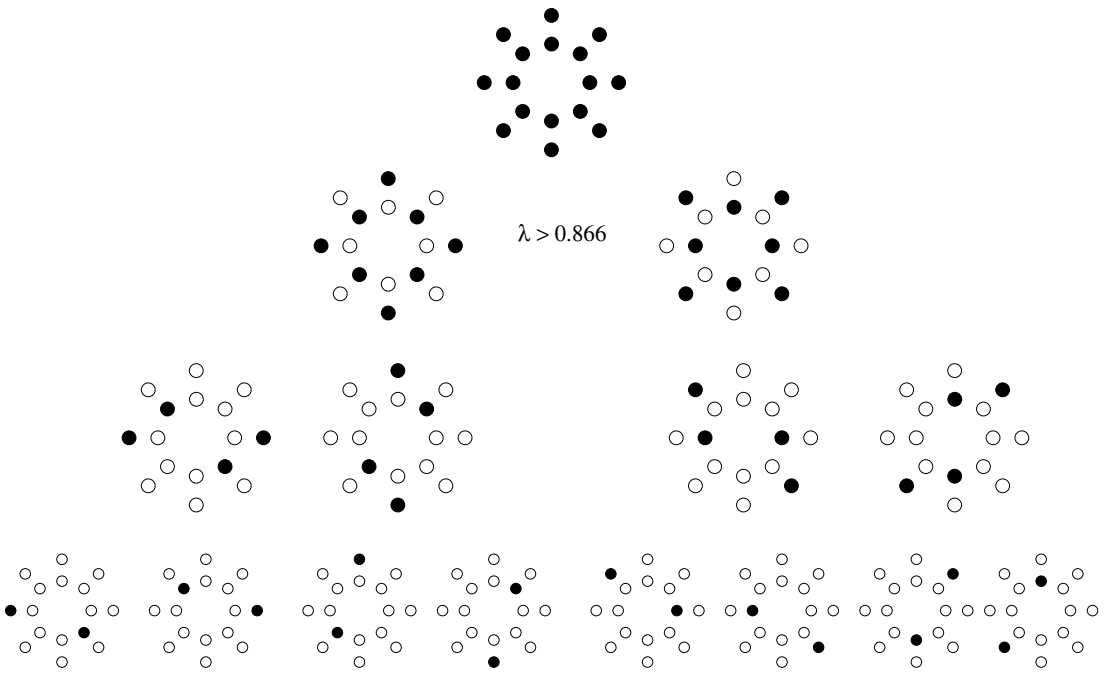

Figure 6.10: Set-partitioning of the two-ring 16DAPSK constellation for $\lambda > \sqrt{3/4} = 0.866$.

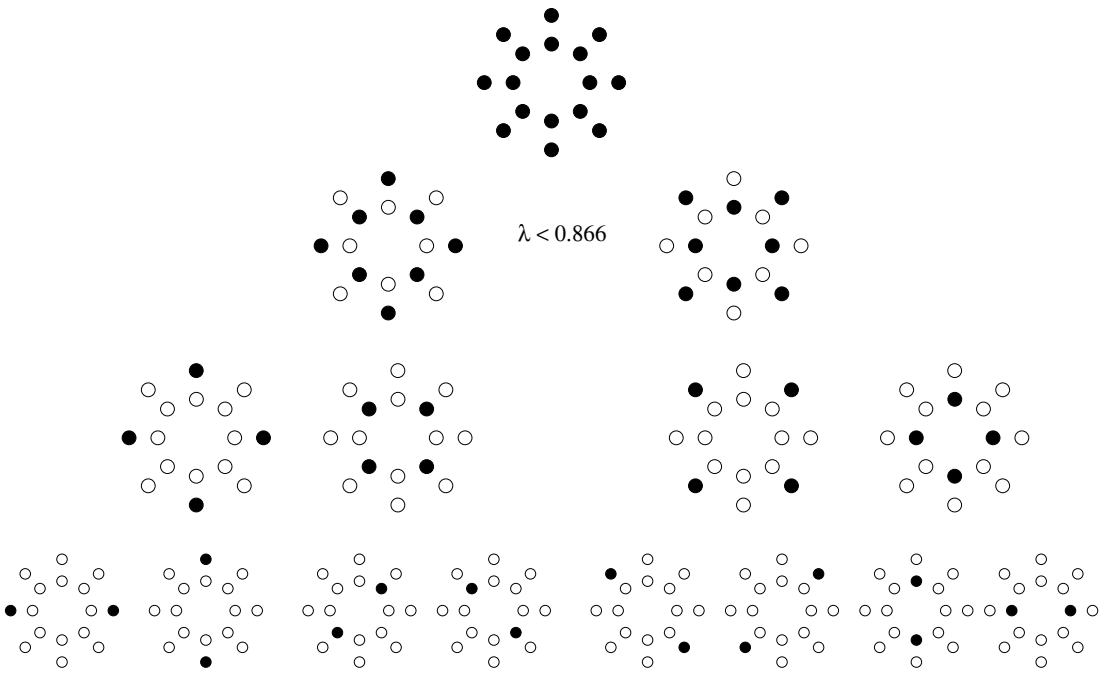

Figure 6.11: Set-partitioning of the two-ring 16DAPSK constellation for $\lambda < \sqrt{3/4} = 0.866$.

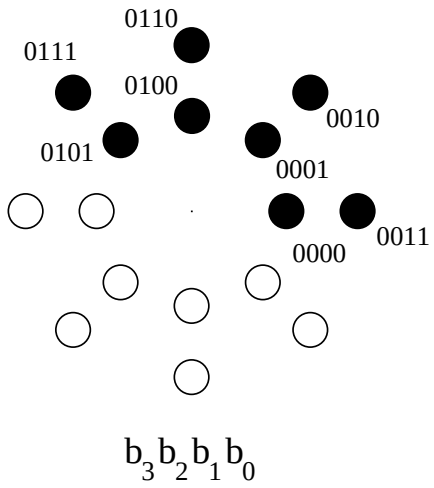

Figure 6.12: Set-partitioning (SP-II) labeling of the 16DAPSK constellation: b_1 is uncoded.

both circles. Figure 6.11 displays SP-II, the correct set-partitioning to be used for lower values of λ. Despite the fact that this would produce a more balanced coded and uncoded bit protection, the coded bits map a nonstandard 8-point constellation (Figure 6.12 that is not amenable to differential phase modulation. Having exhausted the scenarios based on joint optimal coding of amplitude and phase dimensions, we focus our attention to a separate-encoder approach for amplitude dimension. This approach is most attractive for its promise to develop the design independently of the already existing phase-dimension transmission.

6.5.1 Separate coding for amplitude and phase dimensions

The already existing phase-code provides integer rate transmission with two coded bits, and more with the standard set-partitioning around the circle. The amplitude lives on the positive real axis, therefore it may be encoded via standard PAM codes [22] that have one bit per symbol constellation expansion. This leads to, for example, using 8 different rings for 2 bits/T on the amplitude. With such

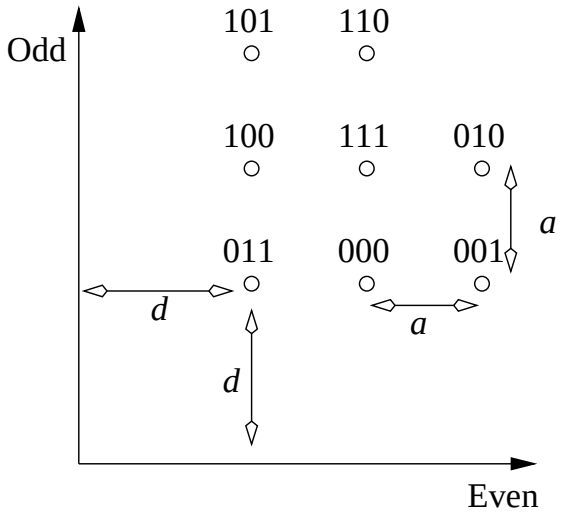

Figure 6.13: 2 bits/$2T$ amplitude tranmission using an 8point constellation over two symbols.

an expansion, the amplitude dimension gets very crowded very soon. Moreover, the amplitude code, being decoupled from any contingent phase errors, is more robust. We propose a half-a-bit-per-symbol expansion via one bit expansion over two symbols.

6.5.1.1 Coding the amplitude dimension over two symbol intervals

The concept of multiple-symbol coding is best explained through an example. With 1 bit/T of expansion (redundancy), 1 bits/T can be transmitted using 2 bits i.e. 4 different amplitudes. Instead, we propose to encode 2 bits over two symbols (which is the same rate) using 1bit/2T expansion, i.e. 8 different pairs of amplitude levels over two symbol intervals.

Figure 6.13 shows the constellation for amplitude pairs over two consecutive symbol intervals. The A-trellis encoder takes two information bits, $f_2 f_1$, makes a state-transition and outputs three bits, $f_2 f_1 f_0$, say 110. Based on the mapping

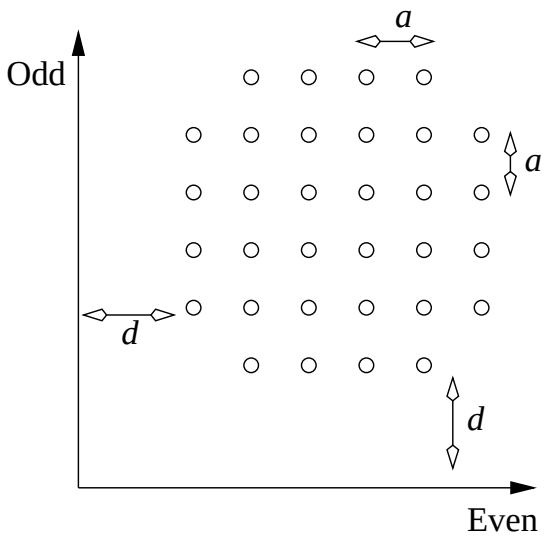

Figure 6.14: The 32-point constellation for 4 bits/$2T$ transmission on the A-dimension.

given in Figure 6.13, the radius of the current PSK symbol is scaled by $d + a$, and that of the next PSK symbol is scaled by $d + 2a$. The transmission rate on the A dimension is 2bits/2T, equivalently 1bit/T.

The 2 bits/2T transmission on the A-dimension is accomplished via 2-symbol trellis coding on the 8-point constellation given in Figure 6.13. For unit symbol energy (E_s) in the transmission, the parameters a and d for this 8-point constellation have to obey

$$8d^2 + 14ad + 11a^2 - 8 = 0. \tag{6.12}$$

For $R = 3$ bits/T overall transmission rate and target average BER of 5×10^{-4}, the parameters $(a, d) = (0.445, 0.548)$ have been found to be approximately optimal. For R = 4 bits/T, the A dimension still transmits 2 bits/$2T$. For this rate, the choice $(a, d) = (0.282, 0.729)$ is approximately optimal at target average BER of 5×10^{-4}.

The 4 bits/$2T$ transmission on the A-dimension is carried out by 2-symbol

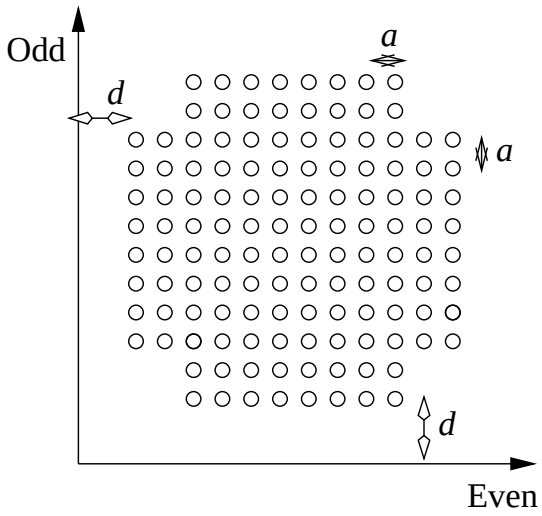

Figure 6.15: The 128-point constellation for 6 bits/$2T$ transmission on the A-dimension.

trellis coding on the 32-point constellation shown in Figure 6.14. The bit labels are detailed in the Appendix. For unit symbol energy (E_s) in the transmission, the parameters a and d for this 32-point constellation have to obey

$$d^2 + 5ad + 8.75a^2 - 1 = 0. \tag{6.13}$$

For $R = 5$ bits/T overall transmission rate, the A-dimension has 4 bits/$2T$ and the P-dimension has 3 bits/T. For target average BER of 5×10^{-4}, the parameters $(a, d) = (0.242, 0.375)$ have been found to be approximately optimal. For $R = 6$ bits/T, the A dimension still transmits 4 bits/$2T$. For this rate, the choice $(a, d) = (0.174, 0.626)$ is approximately optimal at target average BER of 5×10^{-4}.

The R= 7 bits/T and R=8 bits/T schemes use 6 bits/$2T$ transmission on A-dimension using the 128-point constellation shown in Figure 6.15 over two consecutive symbol intervals. For R =7 bits/T, the parameters $(a, d) = (0.117, 0.306)$ have been found to be approximately optimal for the target BER. For $R = 8$

bits/T, $(a, d) = (0.088, 0.491)$ is implemented.

6.5.1.2 Coding the phase dimension

The P-dimension is coded through a core 2 bits/T 8PSK TCM that is differentially signaled. The P-trellis encoder takes two information bits, b_2b_1, makes a state-transition and outputs three bits, $b_2b_1b_0$, which map to the angle $(4b_2 + 2b_1 + b_0)2\pi/8$ on the circle. For higher rates, uncoded bits are added as the most significant bits in the natural labeling. For example, 3 bits/T on the P-dimension is transmitted by the angle $((8b_3 + 4b_2 + 2b_1 + b_0)2\pi/16)$ (the 16PSK) where b_3 is uncoded.

6.5.2 Putting it all together: The multitone operation

The proposed coded-modulation scheme uses two independent encoders, a 32-state 2 bits/T encoder for the phase dimension, and a 16-state 2 bits/$2T$ encoder for the amplitude dimension. More bits are added by set partitioning expanded constellations as discussed before. Based on the standalone performance of individual rates, approximate SNR bins are developed for each rate at a target BER of 5×10^{-4}. Table 6.1 displays a collection of these SNR bins. With a minimum rate of 1 bit/T, the proposed scheme does not signal on tones with SNR less than 7 dB.

The multitone encoding of data starts with the selection of a set of tones that could support the transmission rate at a given error rate. Table 6.1 serves as a guideline for achievable rates at BER of order 5×10^{-4}.

Once the tones are selected, both encoders start running across tones and over again. The amplitude encoder determines the pair of amplitudes (for two symbol

Table 6.1: SNR bins for different rates.

SNR [dB.]	R
[7.0, 10.0)	1 bit/T
[12.0, 15.0)	2 bits/T
[15.0, 19.0)	3 bits/T
[19.0, 23.0)	4 bits/T
[23.0, 25.5)	5 bits/T
[25.5, 27.2)	6 bits/T
[27.2, 30.0)	7 bits/T
[30.0, 33.0)	8 bits/T
[33.0,36.0)	9 bits/T
[36.0, ∞)	10 bits/T

intervals) for all used tones one after another, and over again; whereas the phase encoder determines the phase on these rings, for all used tones one after another, and over again, cumulatively along the time dimension for each tone.

It is best to give an example: Consider only two tones, tone_1 with $\text{SNR}_0 =$ 16 dB tone_2 with $\text{SNR}_0 = 23$ dB. Tone-1 supports 3 bits/symbol transmission whereas tone-2 can signal at 5 bits/symbol. Suppose we are at the start of the transmission and both encoders are at state zero. Suppose also that the information sequence to be encoded is all ones, 1111111111, for simplicity. The amplitude on tone-1 will be encoded as 2 bits/$2T$ and on tone-2 as 4 bits/$2T$.

Starting with tone-1, the encoder takes 2 information bits (11), transits to state 12 with output label 7 of the 8-point constellation: For the next two symbols, the amplitudes on tone-0 will be $d_3 + a_3$ and $d_3 + a_3$ with proper d_3, a_3.

Tone-2 is next: The encoder takes 2 bits (11) to make a transition from

Table 6.2: Phase and amplitude encoding for two tones. Columns represent symbols in time.

Tone-1	$A_0 \exp(j\phi_0)$	$B_0 \exp(j(\phi_0 + \phi_2))$	$A_2 \exp(j(\phi_0 + \phi_2 + \phi_4))$	...
Tone-2	$A_1 \exp(j\phi_1)$	$B_1 \exp(j(\phi_1 + \phi_3))$	$A_3 \exp(j(\phi_1 + \phi_3 + \phi_5))$	...

state 12 to state 3, and outputs 6 ($f_2 f_1 f_0 = 110$), the constellation label is then $f_4 f_3 f_2 f_1 f_0 = 11110$ where $f_4 f_3 = 11$ are two more (uncoded) information bits, since tone-2 transmits 4 bits/2T. For the next two symbols, the amplitudes on tone-2 will be $d_5 + 3a_5$ and $d5 + 3_a 5$, with proper d_5, a_5, Back to tone-1, and the encoder will transit from state 3 to state 13 with output label 13.

On the phase dimension, the initial phase reference for differential signaling is set to zero. Tone-1 has 2 bits/T on the phase dimension, upon taking two information bits (11), the encoder transits to state 9 with label 6, which is the angle $6 \times 2\pi/8$. This is going to be the angle on the first symbol of tone-1. Next comes tone-2, which will have 3 bits/T (1 uncoded) on the phase: The phase encoder transits from state 9 to state 31 with label 7 (111), together with b_3=1 gives 1111 = 15 on the 16PSK, which is the angle $15 \times 2\pi/16$. Back to tone-1, the encoder transits from state 31 to state 20 with a label 7, and the angle on the second symbol of tone-1 will therefore be $6 \times 2\pi/8 + 7 \times 2\pi/8 = 5 \times 2\pi/8$ (modulo 2π), since it is added to the angle from the previous symbol. Now for tone-2, encoder transits from state 20 to state 3 with output 6 (110), together with an encoded bit 1 becomes 1110 = 14, and the angle for the second symbol on the second tone is $15 \times 2\pi/16 + 7 \times 2\pi/8 = 13 \times 2\pi/16$ (modulo 2π).

Table 6.2 illustrates the two-tone OFDM grid over four symbols. The amplitude encoder outputs the pairs of amplitudes (A_0, B_0), (A_1, B_1), (A_2, B_2), (A_3, B_3) in succession. The phase encoder outputs the phases ϕ_0, ϕ_1, ..., ϕ_7.

At the receiving end, two Viterbi-decoding operations are performed, one for the P-trellis, one for the A-trellis. The branch metric for the P-trellis is the squared-Euclidean distance between the unnormalized differential detector output and a unit vector with the angle on the branch:

$$\delta_{\mathrm{P}}^2 = \left| y_k y_{k-1}^{\dagger} - e^{j\phi_k} \right|^2 . \tag{6.14}$$

The A-trellis is decoded independently from P-decisions, where the branch metrics uses the most favorable angles possible for that tone over two consecutive symbols:

$$\delta_{\mathrm{A}}^2 = \min_{\phi_1 \in \Phi} \left| y_{2k+1} - e^{j\phi_1} x_{2k+1} \right|^2 + \min_{\phi_0 \in \Phi} \left| y_{2k} - e^{j\phi_0} x_{2k} \right|^2 \tag{6.15}$$

For phase constellations of 8 points and higher, the discrete minimization over the constellation angles can be approximated by a minimization over the continuum of angles $[0, 2\pi)$ with virtually no loss in performance:

$$\delta_{\mathrm{A}}^2 \sim \left| y_{2k+1} - e^{j\phi_1} x_{2k+1} \right|^2 + \left| y_{2k} - e^{j\phi_0} x_{2k} \right|^2 . \tag{6.16}$$

6.5.3 The proposed trellis codes

6.5.3.1 The 16-state A-encoder:

The 16-state basic A-encoder transmits at 2 bits/$2T$ (Table 6.3). Each row displays the set of four outgoing transitions, each in the form of (new-state, output-label) corresponding to inputs read from the column bits $(f_2 f_1)$. The encoder is systematic therefore the output label for input $(f_2 f_1)$ is $(f_2 f_1 f_0)$ where f_0 is the nonsystematic coded bit. For example, with the input 10, the code transits from state 12 to state 1 and outputs the constellation label 4, which is 100 in binary, and therefore the two consecutive ring radii are $(d, d+a)$ as seen in Figure 6.13.

Table 6.3: The 16-state A-trellis encoder.

state	$f_2 f_1$			
	00	01	10	11
0	(0, 0)	(2, 2)	(7, 4)	(5, 6)
1	(9, 1)	(11, 3)	(14, 5)	(12, 7)
2	(1, 0)	(3, 2)	(6, 4)	(4, 6)
3	(8, 1)	(10, 3)	(15, 5)	(13, 7)
4	(2, 0)	(0, 2)	(5, 4)	(7, 6)
5	(11,1)	(9, 3)	(12, 5)	(14, 7)
6	(3, 0)	(1, 2)	(4, 4)	(6, 6)
7	(10,1)	(8, 3)	(13, 5)	(15, 7)
8	(4, 0)	(6, 2)	(3, 4)	(1, 6)
9	(13, 1)	(15, 3)	(10, 5)	(8, 7)
10	(5, 0)	(7, 2)	(2, 4)	(0, 6)
11	(12,1)	(14, 3)	(11, 5)	(9, 7)
12	(6, 0)	(4, 2)	(1, 4)	(3, 6)
13	(15,1)	(13, 3)	(8, 5)	(10, 7)
14	(7, 0)	(5, 2)	(0, 4)	(2, 6)
15	(14, 1)	(12, 3)	(9, 5)	(11, 7)

For higher rates, uncoded bits are added as most significant positions of the binary representation in the label. For example, suppose the encoder is currently at state 7. For R=7 bits/T, the A-dimension will transmit 6 bits/$2T$. If the input is $f_5f_4f_3f_2f_1 = 10111$, then the code transits to state 15 with output $f_2f_1f_0$=111 (7) and the full constellation label is $f_5f_4f_3f_2f_1 = 101111$ (47).

6.5.3.2 The 32-state P-encoder

The P-dimension uses the standard 32-state Ungerboeck 2 bits/T 8PSK TCM ([22]). The generators in octal form for this encoder is $h_2 = 34, h_1 = 16, h_0 = 45$. Higher rates are achieved through uncoded bits that form the most significant positions on a naturally labeled MPSK.

6.6 Conclusions

This chapter described our efforts and solution towards developing an efficient coded-modulation solution for OFDM data streams. The proposed solution operates with two separate encoders that encode bits describing the phase and amplitude, respectively, of multi-ring PSK constellations in a fashion to ensure balanced error protection.

The observed powerline channel has the most uncertainty in the phase component, which could result in phase-estimation errors. The proposed solution is prepared for this contingency in the best sense that the symbol-interleaved encoding and erasures decoding for the phase component can restore performance under intermittent phase errors, with the receiver only requiring a localization of phase-errors within two- symbol accuracy. Phase signaling follows a differential rule to avoid carrier tracking at the receiver.

The amplitude dimension of the multi-ring constellations is encoded with a simpler encoder and decoded without any reference to phase decisions. This property ensures that even in the presence of complete loss of the channel phase, the amplitude bits can be recovered with no performance loss as long as the amplitude tracking is correct. This could add some benefit to overall system performance if the output of the RS block is carefully multiplexed in phase and amplitude encoders.

For a collection of data streams describing a specific application, the proposed scheme codes across tones, i.e. both encoder trellises run through the tones and over again, while the phase mapping follows a two-symbol differential signaling over time. This ensures a decoding delay of 1 OFDM symbol.

Even with a conservative differential modulation on the phase-component, the proposed solution achieves coding gains by achieving a faster transmission rate of 1 bit/symbol/tone over a similar uncoded scheme at the same error rates.

References

[1] A. di Vito and M. Naldi. Robustness of the likelihood ratio detector for moderately fluctuating radar targets. *IEE Proc.-Radar, Sonar Navig.*, 146(2):107–112, 1999.

[2] A. Lapidoth. On the role of mismatch in rate distortion theory. *IEEE Trans. on Inform. Theory*, 43(1):38–47, 1997.

[3] A. Lapidoth and S. Shamai. A lower bound on the mismatched viterbi decoding bit-error-rate. In *Proc. IEEE Int. Symp. on Information Theory, ISIT 1997*, 1997.

[4] A. Wald. Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, 16:117–186, 1945.

[5] A. Wald and J. Wolfowitz. Optimum character of the sequential probability ratio test. *Annals of Mathematical Statistics*, 19:326–339, 1948.

[6] B. Hassibi and B. M. Hochwald. High-rate codes that are linear in space and time. *IEEE Trans. Inform. Theory*, 48(7):1804–1824, July 2002.

[7] C. Berrou and A. Glavieux. Near optimum error correcting coding and decoding: Turbo-codes. *IEEE Trans. on Comm.*, 44(10):1261–1271, Oct. 1996.

[8] C. Köse and R. D. Wesel. Universal space-time trellis codes. *IEEE Trans. Inform. Theory*, 49(10):2717–2727, Oct. 2003.

[9] C. Köse and R. Wesel. Code design metrics for space-time systems under arbitrary fading. In *Proc. IEEE ICC–01*, pages 1099–1103, 2001.

[10] C. Schlegel. *Trellis Coding*. IEEE Press, 1997.

[11] C. W. Baum and V. V. Veeravalli. A sequential procedure for multihypothesis testing. *IEEE Trans. on Inform. Theory*, 40:1994–2007, 1994.

[12] D. Aktas, H. El Gamal, and M. P. Fitz. On the design and maximum-likelihood decoding of space-time trellis codes. *IEEE Trans. Commun.*, 51(6):854–859, June 2003.

[13] D. D. Lee and R. L. Kashyap. Robust maximum likelihood bearing estimation in contaminated Gaussian noise. *IEEE Trans. Sig. Processing*, 40(8):1983–1986, 1992.

[14] D. Divsalar and M. K. Simon. The design of trellis coded MPSK for fading channels: Performance criteria. *IEEE Trans. Commun.*, 36(9):1004–1012, Sept. 1988.

[15] D. Divsalar, S. Donilar, and F. Pollara. Iterative turbo decoder based on density evolution. *IEEE Journal on Sel. Areas in Comm.*, 19(5):891–907, May 2001.

[16] D. Haccoun and M. J. Ferguson. Generalized stack algorithms for decoding convolutional codes. *IEEE Trans. on Inform. Theory*, 21(6):638–651, 1975.

[17] D. K. Aktas and M. P. Fitz. Computing the distance spectrum of space-time trellis codes. In *Proc. IEEE WCNC–00*, pages 51–55, 2000.

[18] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, 1(3):41–59, Autumn 1996.

[19] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal. Commun.*, 6(3):311–335, Mar. 1998.

[20] G. Lorden. Likelihood-ratio tests for sequential k-decision problems. *Annals of Mathematical Statistics*, 43:1412–1427, 1972.

[21] G. Simons. Lower bounds for average sample number of sequential multihypothesis tests. *Annals of Mathematical Statistics*, 38:1343–1364, 1967.

[22] G. Ungerboeck. Channel coding with multilevel/phase signals. *IEEE Trans. Inform. Theory*, 28(1):55–67, Jan. 1982.

[23] H. M. Tullberg and P. H. Siegel. Serial concatenated trellis coded modulation with inner rate-1 accumulate code. In *Global Telecommunications Conference, GLOBECOM*, pages 936–940, 2001.

[24] I. M. G. Lourtie, G. C. Carter, and S. Basu. Signal to noise ratio improvement by likelihood ratio techniques under model and environmental mismatch. In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP-92*, volume 2, pages 457–460, 1992.

[25] J. B. Anderson and S. Mohan. Sequential coding algorithms: A survey and cost analysis. *IEEE Trans. Commun.*, 32(2):169–176, 1984.

[26] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo. Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels. In *Proc. IEEE VTC–96*, pages 136–140, 1996.

[27] J. Grimm, M. P. Fitz, and J. V. Krogmeier. Further results on space-time coding for Rayleigh fading. In *Proc. Allerton Conference*, pages 392–400, 1998.

[28] J. Grimm, M. P. Fitz, and J. V. Krogmeier. Further results on space-time coding for Rayleigh fading. In *Proc. Thirty-Sixth Annual Allerton Conference on Communication, Control, and Computing*, pages 391–400, Sept. 1998.

[29] J. Kiefer and J. Sacks. Asymptotically optimal sequential inference and design. *Annals of Mathematical Statistics*, 34:705–750, 1963.

[30] J. Shi and R. D. Wesel. Further error event diagram reduction using algorithmic techniques. *presented at IEEE ICC 2003, Alaska.*

[31] J. Shi and R. Wesel. Rotationally invariant space-time constellations. In *ISIT 2004.*

[32] J. T. Kent. Robust properties of the likelihood ratio test. *Biometrika*, 69:19–27, 1982.

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

[34] M. K. Simon and D. Divsalar. The performance of trellis coded multilevel DPSK on a fading mobile satellite channel. *IEEE Trans. on Veh. Tech.*, 37(2):78–91, May 1998.

[35] M. Sobel and A. Wald. A sequential design procedure for choosing one of three hypotheses concerning the unknown mean of normal distribution. *Annals of Mathematical Statistics*, 20:502–522, 1949.

[36] M. Tao and R. S. Cheng. Improved design criteria and new trellis codes for space-time coded modulation in slow flat fading channels. *IEEE Comm. Letters*, 5(7):313–315, July 2001.

[37] N. C. Warke and G. C. Orsak. Impact of statistical mismatch on both universal classifiers and likelihood ratio tests. In *Proc. IEEE Int. Symp. on Information Theory, ISIT 1998*, 1998.

[38] P. J. Huber. A robust version of the probability ratio test. *Annals of Mathematical Statistics*, 36:1753–1758, 1965.

[39] Q. Yan and R. S. Blum. Optimum space-time convolutional codes. In *Proc. IEEE WCNC–00*, pages 1351–1355, 2000.

[40] Q. Yan and R. S. Blum. Improved space-time convolutional codes for quasi-static slow fading channels. *IEEE Trans. Wireless Commun.*, 1(4):563–571, Oct. 2002.

[41] R. Durrett. *Probability: Theory and Examples.* Duxbury Press, 1996.

[42] R. J. Kozick and B. M. Sadler. Makimum-likelihood array processing in non-Gaussian noise with Gaussian Mixtures. *IEEE Trans. Sig. Processing*, 48(12):3520–3535, 2000.

[43] R. J. McEliece. Are turbo-like codes effective on nonstandard channels? *IEEE Information Theory Society Newsletter*, 51(4):1–8, Dec. 2001.

[44] R. S. Varga. Minimal Gerschgorin sets for partitioned matrices. *SIAM Journal on Numerical Analysis*, 7:493–507, 1970.

[45] R. V. Foutz and R. C. Srivastava. The performance of the likelihood ratio test when the model is incorrect. *Annals of Statistics*, 5:1183–1194, 1977.

[46] R. Wesel, X. Liu, and W. Shi. Trellis codes for periodic erasures. *IEEE Trans. Commun.*, 48(6):938–947, June 2000.

[47] S. A. Kassam. Robust hypothesis testing for bounded classes of probability densities. *IEEE Trans. on Inform. Theory*, 27:242–247, 1981.

[48] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara. Serial concatenation of interleaved codes: Performace analysis, design and iterative decoding. *IEEE Trans. on Inform. Theory*, 44(3):909–926, May 1998.

[49] S. G. Wilson and Y. S. Leung. Trellis coded phase modulation on Rayleigh channels. In *Proc. IEEE ICC–87*, pages 739–743, June 1987.

[50] S. Lin and D. J. Costello, Jr. *Error control coding:Fundamentals and applications.* Prentice-Hall, 1983.

[51] S. M. Alamouti. A simple transmit diversity tecnhique for wireless communications. *IEEE J. Select. Areas Commun.*, 16(8):1451–1458, Nov 1998.

[52] S. Sandhu, R. Heath, and A. Paulraj. Space-time block codes versus space-time trellis codes. In *Proc. IEEE ICC–01*, pages 1132–1136, 2001.

[53] S. Siwamogsatham and M. P. Fitz. Improved high-rate space-time codes via orthogonality and set partitioning. In *Proc. IEEE WCNC–02*, volume 1, pages 264–270, Mar 2001.

[54] S. Siwamogsatham and M. P. Fitz. Improved high-rate space-time codes via orthogonality and set partitioning. In *Proc. IEEE WCNC–02*, volume 1, pages 264–270, Mar. 2001.

[55] S. ten Brink. Design of serially concatenated codes based on iterative decoding convergence. In *Second International Symposium on Turbo Codes and Related Topics, Brest, France*, 2000.

[56] S. ten Brink. Convergence behavior of iterative decoded parallel concatenated codes. *IEEE Trans. on Comm.*, 49(10):1727–1737, Oct. 2001.

[57] T. R. Gialorrenzi and S. G. Wilson. Noncoherent demodulation techniques for trellis coded M-DPSK signals. *IEEE Trans. on Comm.*, 43(8):2370–2379, 1995.

[58] V. Engles and H. Rohling. Multi-resolution 64-DAPSK modulation in a hierarchical COFDM transmission system. *IEEE Trans. on Comm.*, 44(1):139–149, 1998.

[59] V. P. Dragalin, A. G. Tartakovsky, and V. V. Veeravalli. Multihypothesis sequential probability ratio tests-part I: Asymptotic optimality. *IEEE Trans. on Inform. Theory*, 45:2448–2461, 1999.

[60] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Trans. Inform. Theory*, 45(5):1456–1467, July 1999.

[61] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criteria and code construction. *IEEE Trans. Inform. Theory*, 44(2):744–765, March 1998.

[62] W. Hoeffding. A lower bound for the average sample number of a sequential test. *Annals of Mathematical Statistics*, 24:127–130, 1953.

[63] W. Hoeffding. Lower bounds for the expected sample size and the average risk of a sequential procedure. *Annals of Mathematical Statistics*, 31:351–368, 1960.

[64] W. L. Root and P. P. Varaiya. Capacity of classes of Gaussian channels. *SIAM J. of Applied Math.*, 16(6):1350–1393, 1968.

[65] W. Zhu and M. P. Fitz. through private communication.

[66] Wesel, R. D. *Trellis code design for correlated fading and achievable rates for Tomlinson-Harashima precoding.* PhD thesis, Stanford University, Aug. 1996.

[67] Z. Chen, B. S. Vucetic, J. Yuan, and K. L. Luo. Space-time trellis codes for 4-PSK with three and four transmit antennas in quasi-static flat fading channels. *IEEE Comm. Letters*, 6(2):67–69, Feb. 2002.