Institut Eurécom

Mobile Communications Department

2229, route des Crêtes

B.P. 193

06904 Sophia Antipolis

FRANCE

Technical Report no. 34

On Coding for Block Fading Channels[1][2]

(submitted to IEEE Transactions on Information Theory)

December 23, 1997

R. Knopp, P.A. Humblet

Tel: (+33) 4 93 00 26 26

Fax: (+33) 4 93 00 26 27

Email: knopp,humblet@eurecom.fr

[1]Eurecom's research is partially supported by its industrial partners: Ascom, Cegetel, Hitachi, IBM France, Motorola, Swisscom, and Thomson CSF

[2]R. Knopp's research was funded in part by a FCAR (Quebec, Canada) doctoral grant

On Coding for Block Fading Channels

R. Knopp, P.A. Humblet

Abstract

This work considers the achievable performance for coded systems adapted to a multipath *block–fading* channel model. This is a particularly useful model for analysing mobile–radio systems which employ techniques such as *slow frequency–hopping* under stringent time–delay or bandwidth constraints for slowly time–varying channels. In such systems, coded information is transmitted over a small number of fading channels in order to achieve *diversity*. The separation between the diversity effects of multipath resolution and coding are studied. Bounds on the achievable performance due to coding are derived using information–theoretic techniques. It is shown that high diversity can be achieved using relatively simple codes as long as very high spectral–efficiency is not required. Examples of simple block codes and carefully chosen trellis codes are given which yield, in some cases, performances approaching the information–theoretic bounds.

Key words: Block–Fading Channels, Diversity, Outage Probability, MDS Codes, Slow Frequency Hopping.

1 Introduction and Paper Outline

Consider the generic transmission scheme with diversity shown in figure 1. Information bits are coded/modulated into F blocks of length N symbols ($\cdots$), so that codewords have length NF symbols and are denoted as $\mathbf{c} = \begin{pmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,N-1} & c_{1,0} & \cdots & c_{F-1,N-1} \end{pmatrix}$. The coded symbols are formed by either a block or convolutional encoder and often passed to an interleaver for practical reasons. The coded symbols belong an arbitrary symbol set (constellation) $\mathcal{S}$ in the complex plane so that each occupies two dimensions. Each block is QAM modulated as

$$u_f(t) = \sum_{n=0}^{N-1} \sqrt{\mathcal{E}_s} c_{f,n} s(t - nT), f = 0, 1, \cdots, F-1 \tag{1}$$

where $s(t)$ is some unit energy signaling pulse shape, and $\mathcal{E}_s$ is the energy per coded symbol. The key feature of such a system is that the F blocks are transmitted over different time–varying channels, in order to partially average the performance over the different channel realizations. The complex baseband received signals before processing are given by

$$r_f(t) = u_f(t) * h_f(t, \tau) + z_f(t), t \in [0, NT], f = 0, 1, \cdots, F-1 \tag{2}$$

where $*$ denotes convolution, $h_f(t, \tau)$ is the channel response at time t to an impulse at time τ on the f^{th} channel, and $z(t)$ is complex white Gaussian noise with power spectral density N_0.

Coding across different channel realizations provides a certain amount diversity, which counters the effects of multipath fading. In what follows we will assume that the F channel realizations are correlated, although it may well be the case in some systems that they can be taken to be uncorrelated. In some cases with reasonable mobile speeds, the channel is virtually time-invariant during the block. We will assume this to be the case, and adopt the nomenclature of [MS84] who referred to this type of channel as a *block interference* or *block fading* channel.

In GSM [GSM90], which uses *slow frequency–hopping*, blocks modulate $F = 4$ (half–rate) or $F = 8$ (full–rate) carriers whose spacing is larger than the coherence bandwidth, resulting in virtually uncorrelated blocks. The practical advantages of such a system are firstly that reliable coherent communication is possible since the channel responses do not vary during the transmission of a block. Secondly and more importantly, the amount of diversity is independent of the rate of channel variation, since it is a result of exploiting frequency–selectivity. For wireless telephony, this is crucial since the majority of

calls are made at low speed. Another example is the IS54 standard[IS592] where coding is performed across $F = 2$ TDMA blocks separated in time so that the blocks start to become less correlated for high mobile speeds. The underlying system issues which force F to be small are usually imposed by either time–delay or bandwidth constraints, or even both. This model can also represent a multitone system, where F is the number of carrier frequencies. This more general problem is considered by Wesel and Cioffi [WC95] in the context of digital broadcasting using multitone signals.

Lapidoth [Lap94] considers a similar problem for convolutional codes with finite-depth interleaving over the Gilbert-Elliot erasure channel, which serves as a first–approximation for a fading channel. He compares different finite–interleaving strategies for rate $1/n$ binary convolutional codes. The reported results are applicable to slow frequency–hopped systems such as GSM. The main conclusion of this work is the that the code and interleaving strategy should be jointly optimized to maximize performance, and that coding complexity plays a rather unimportant role. Here we will draw similar conclusions regarding the relationship between the code and the number of blocks. Furthermore, we argue that the size of the underlying symbol alphabet also plays an important role with respect to the achievable performance. Furthermore, increasing code complexity is necessary to reduce error probabilities, up to a certain point.

1.1 Paper Outline

In section 2 we examine the achievable performance by determining the *pairwise error–probability (PEP)* between two arbitrary coded sequences. This analysis shows that the performance is characterized by two diversity effects which operate independently under certain conditions. The first is due to the degree of resolvability of different multipath components and depends on the relationship between the pulse shape and the delay spread of the channel. The second is due to the effect of coding across different (and hopefully independent) channel realizations.

In section 3 we restrict our attention to the effect of coding by considering a set of F discrete–time single–path channels, which completely describe a narrowband system without ISI. Because of the separation effect previously mentioned we lose nothing by considering this simplified scenario. We extend the results of Ozarow *et al* [OSSW94] and Kaplan and Shamai [KSS95] who characterized these types of systems by an *information outage probability* since, for finite F, a channel capacity does not exist. This was also termed *Outage Capacity* by Foschini and Gans [FG97]. Specifically we show the relationship between the information outage probability and the *frame* (FER) and bit error rates (BER). An important

conclusion is that for practical (i.e. small) signaling alphabets, the attainable diversity order due to coding (which in Rayleigh fading corresponds to the slope of the information outage probability curve versus the signal–to–noise ratio (SNR) on a log-log scale) is generally smaller than F. This warrants the use of larger constellations for achieving high diversity. We find that for practical spectral efficiencies ($<$1.5 bits/dim), a small increase in the size of the constellation can yield significant performance improvement.

Practical block and trellis codes are considered in section 4. We begin by showing that the achievable diversity for any coding scheme is given by a disguised version of the *Singleton bound* [Sin64]. This was also noted by Wesel and Cioffi [WC95] and stems from the fact that this type of coding can be interpreted as a non–binary coding problem with block–length constrained to F symbols. The Singleton bound predicts the same diversity order as the information outage probability analysis. Moreover, it shows that practical high diversity codes are difficult to construct when high spectral efficiency is required since very large constellations are required. We give many examples of coding schemes using standard modulation formats (AM and PSK) and spectral efficiencies in the range .25–1.5 bits/dim which meet the Singleton bound.

Finally, in section 5 we present computer simulations of some selected codes and show that their FER, with practical block sizes, is often very close to the information outage probability. This is perhaps the most important result of this work. The BER indicated by the information outage analysis is less indicative of practical performance.

2 Pairwise Error-Probability Analysis

Under the block–fading assumption, we assume that the time–variation of the channel is slow (i.e. that the coherence time is greater than the duration of a block) so that the channel attenuations and phases can be taken to be constant over blocks. We therefore express the complex baseband channel response as

$$h_f(t) = \sum_{l=0}^{L-1} \sigma_{f,l}\delta(t - d_{f,l}), \tag{3}$$

where $\sigma_{f,l}$ and $d_{f,l}$ are the complex attenuation and delay of the l^{th} path in the f^{th} block. These quantities are assumed to be random from block to block but known without error to the receiver. We consider a Gaussian fading model in this work so that the $\sigma_{f,l}$ are assumed to be circular symmetric complex

Gaussian random variables with mean $\overline{\sigma_l}$ and variance $\overline{|\sigma_l - \overline{\sigma_l}|^2}$ which are independent of f. We assume further a *wide–sense stationary uncorrelated scattering (WSSUS)* channel model so that

$$\mathrm{E}\left(\sigma_{f,l} - \overline{\sigma_{f,l}}\right)\left(\sigma_{f',l'} - \overline{\sigma_{f',l'}}\right)^* = \varrho_{f,f'}\overline{\sigma_l^2}\delta_{l,l'}, \tag{4}$$

where $\varrho_{f,f'}$ is the correlation coefficient between blocks f and f'. We have, therefore, that different paths are uncorrelated but that the strengths for a given path are correlated, in general, from block to block. Furthermore, we assume that the path strengths are normalized as $\sum_{l=0}^{L-1}\overline{\sigma_l^2} = 1$, so that the average attenuation is included in the transmitted signal strength. For convenience, we denote the column vector formed by the L path strengths of the f^{th} block by $\boldsymbol{\sigma}_f = \begin{pmatrix}\sigma_{f,0} & \sigma_{f,1} & \cdots & \sigma_{f,L-1}\end{pmatrix}^{\mathrm{T}}$.

The F received signals are processed by a maximum–likelihood decoding rule as

$$\hat{m} = \underset{m=0,\cdots,2^{FNR}-1}{\arg\min} \sum_{f=0}^{F-1}\int_0^{NT}\left|r_f(t) - u_f^{(m)}(t-NT) * h_f(t)\right|^2 dt. \tag{5}$$

Decoding in this fashion is too complex to be carried out in practice, and it is usually done in two steps, depending on the relationship between the coherence bandwidth of the channel and the bandwidth of $s(t)$. In medium-band systems like GSM where the multipath induces intersymbol interference (ISI), a sub–optimal approach is taken by first equalizing the F channels with a soft–output algorithm (e.g. soft–output Viterbi equalization [HH89]). These outputs are then deinterleaved and passed to a Viterbi decoder to retrieve the information bits. In narrow-band systems such as IS-54, the channel is almost ISI–free, and either a very simple equalizer or none at all is needed prior to deinterleaving/decoding. In spread–spectrum systems without ISI, equalization is also not required and some of the multipath can be exploited with a RAKE receiver prior to decoding [Pro95].

We now perform a standard Gaussian–fading performance analysis for the block–fading channel by determining the *pairwise error–probability (PEP)* between arbitrary code sequences. Denoting the codewords by $\mathbf{c} = \begin{pmatrix}\mathbf{c}_0 & \mathbf{c}_1 & \cdots & \mathbf{c}_{F-1}\end{pmatrix}$, the PEP conditioned on a particular set of channel realizations is given by

$$\Pr\left(\mathbf{c}^{(a)} \to \mathbf{c}^{(b)} \middle| \{h_f(t)\}\right) = \mathrm{Q}\left(\sqrt{d^2(a,b)\frac{\mathcal{E}_s}{2N_0}}\right), \tag{6}$$

where $d^2(a,b)$ is the squared Euclidean distance between the coded signals which in our case is

$$d^2(a,b) = \sum_{f=0}^{F-1} \int_0^{NT} \left| \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} \left(c_{f,n}^{(a)} - c_{f,n}^{(b)} \right) \sigma_{f,l} s(t - nT - d_{f,l}) \right|^2 dt, \tag{7}$$

and $\mathrm{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du$. This is a quadratic form in the path strengths which after some straightforward manipulation can be written as

$$\begin{aligned} d^2(a,b) &= \begin{bmatrix} \boldsymbol{\sigma}_0^* & \boldsymbol{\sigma}_1^* & \cdots & \boldsymbol{\sigma}_{F-1}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\Xi}_0(a,b) & 0 & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Xi}_1(a,b) & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \boldsymbol{\Xi}_{F-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\sigma}_0 \\ \boldsymbol{\sigma}_1 \\ \cdots \\ \boldsymbol{\sigma}_{F-1} \end{bmatrix} \\ &= \boldsymbol{\sigma}^* \boldsymbol{\Xi} \boldsymbol{\sigma}, \end{aligned} \tag{8}$$

where $\Xi_f^{(ll')}(a,b) = (\mathbf{c}_f^{(a)} - \mathbf{c}_f^{(b)}) \mathbf{P}_{f,l,l'} (\mathbf{c}_f^{(a)} - \mathbf{c}_f^{(b)})^*$, $P_{f,l,l'}^{(n,n')} = \rho_s \left((n-n')T + (d_{f,l} - d_{f,l'}) \right)$ and $\rho_s(\tau) = \int_{-\infty}^{\infty} s(t) s^*(t+\tau) dt$.

Since (8) is a quadratic form of circular symmetric Gaussian random variables, the moment generating function of the random variable $z = d^2(a,b) \frac{\mathcal{E}_s}{2N_0}$ is [SBS66, App. B]

$$\Phi_z(s) = \frac{\exp(s \frac{\mathcal{E}_s}{2N_0} \overline{\boldsymbol{\sigma}}^* (\boldsymbol{\Xi}^{-1} - s \frac{\mathcal{E}_s}{2N_0} \mathbf{K}_\sigma \boldsymbol{\Xi})^{-1} \overline{\boldsymbol{\sigma}}}{\det(\mathbf{I} - s \frac{\mathcal{E}_s}{2N_0} \mathbf{K}_\sigma \boldsymbol{\Xi})} \tag{9}$$

where $\mathbf{K}_\sigma$ is the covariance matrix of $\boldsymbol{\sigma}$.

When the path strengths are zero–mean (i.e. *Rayleigh fading*) (9) simplifies to

$$\Phi_z(s) = \prod_{i=0}^{FL} \frac{1}{\det(\mathbf{I} - s \frac{\mathcal{E}_s}{2N_0} \mathbf{K}_\sigma \boldsymbol{\Xi})} = \prod_{i=0}^{d_{\mathrm{H}}^F L - 1} \frac{1}{1 - s \lambda_i \mathcal{E}_s / 2N_0}, \tag{10}$$

where $\{\lambda_i\}$ are the *non–zero* eigenvalues of the matrix

$$\mathbf{K}_\sigma \mathbf{\Xi} = \begin{bmatrix} \mathbf{\Sigma\Xi}_0(a,b) & \varrho_{1,0}\mathbf{\Sigma\Xi}_1(a,b) & \cdots & \varrho_{F-1,0}\mathbf{\Sigma\Xi}_{F-1}(a,b) \\ \varrho_{0,1}\mathbf{\Sigma\Xi}_0(a,b) & \mathbf{\Sigma\Xi}_1(a,b) & \cdots & \varrho_{F-1,1}\mathbf{\Sigma\Xi}_{F-1}(a,b) \\ \cdots & \cdots & \cdots & \cdots \\ \varrho_{0,F-1}\mathbf{\Sigma\Xi}_0(a,b) & \varrho_{1,F-1}\mathbf{\Sigma\Xi}_1(a,b) & \cdots & \mathbf{\Sigma\Xi}_{F-1}(a,b) \end{bmatrix}, \tag{11}$$

and d_{H}^F is the number of non–zero $\mathbf{\Xi}_f(a,b)$ or equivalently the *Hamming distance* between $\mathbf{c}^{(a)}$ and $\mathbf{c}^{(b)}$ with the symbols taken as the sub–vectors $\mathbf{c}_f$ and $\mathbf{\Sigma} = \mathrm{diag}(\overline{\sigma_0^2}, \overline{\sigma_1^2}, \cdots, \overline{\sigma_{L-1}^2})$.

For the even simpler case where the blocks are uncorrelated (i.e. $\varrho_{f,f'} = \delta_{f,f'}$) (9) can be written as

$$\Phi_z(s) = \prod_{i=0}^{d_{\mathrm{H}}^F} \prod_{l=0}^{L-1} \frac{1}{1 - s\eta_{i,l}\mathcal{E}_s/2N_0}, \tag{12}$$

where $\eta_{i,l}$ is the l^{th} eigenvalue of the non–zero matrix $\mathbf{\Delta}_i = \mathbf{\Sigma\Xi}_i(a,b)$.

The PEP can be found analytically using the inverse Laplace transform of (9) to average (6). This amounts to performing a partial fraction expansion of (9) and yields simple closed–form expressions for the PEP. There are often numerical instabilities in the computation of the partial–fraction coefficients when there are repeated eigenvalues and another more numerically stable approach is considered in [BCTV96].

Alternately we may use the Chernov bound, $\mathrm{Q}(x) \leq \frac{1}{2}e^{-x^2/2}$ to upper–bound (12) as

$$\begin{aligned} \Pr\left(\mathbf{c}^{(a)} \to \mathbf{c}^{(b)}\right) &\leq \frac{1}{2}\mathrm{E}_z\left(e^{-\frac{z}{2}}\right) \\ &= \frac{1}{2}\Phi_z\left(-\frac{1}{2}\right) \qquad (13) \\ &< \frac{1}{2}\left(\frac{4N_0}{\chi^2\mathcal{E}_s}\right)^{d_{\mathrm{H}}^F L}, \qquad (14) \end{aligned}$$

where $\chi^2 = \left(\prod_{i=0}^{d_{\mathrm{H}}^F}\prod_{l=0}^{L-1}\eta_{i,l}\right)^{1/d_{\mathrm{H}}^F L}$. This geometric mean also surfaces in the study of antenna diversity systems [SBS66, Chap 10] and coded multitone systems [WC95]. Very similar performance measures also characterize systems with multiple transmit/receive antennas, and can be used to design coding schemes [TSC97a, TSC97b] We can obtain closer approximations by neglecting insignificant eigenvalues and reducing the power in the exponent of (14).

There are two limiting cases for the diversity offered by multipath. Either it cannot be resolved at all (i.e. frequency-flat fading) or when it can be completely resolved by using very wideband signals. In the special case of narrowband signals without ISI, (i.e. $|d_{f,l} - d_{f,l'}| \ll T$) $\mathbf{P}_{f,l,l'} \approx \mathbf{I}$ so that $\eta_{f,l} = d^2\left(\mathbf{c}_f^{(a)}, \mathbf{c}_f^{(b)}\right)\delta_l$. We see, therefore, that χ^2 is simply the d_{H}^F–root of the product distance measure described by Divsalar and Simon for trellis–code time–diversity schemes [DS88]. In very wide-band spread–spectrum systems without ISI (i.e. the bandwidth of $s(t)$ is much larger than the coherence bandwidth and the symbol rate), we have that $\mathbf{P}_{f,l,l'} \approx \delta_{ll'}\mathbf{I}$ so that $\eta_{f,l} = \overline{|\sigma_{f,l}|^2}d^2\left(\mathbf{c}_f^{(a)}, \mathbf{c}_f^{(b)}\right)$. The asymptotic slope of the PEP vs. $\mathcal{E}_s/N_0$ on a log–log scale is commonly referred to as the *diversity order*, and we see that it is the product of the code and multipath diversities.

The theoretical performance of a system will fall somewhere between the performance of these two limits which are straightforward to compute. Due to this separation, the goal of any coding system is therefore to maximize d_{H}^F, since it will affect the performance equally for any channel.

3 Outage Probability Analysis

For simplicity, let us now consider narrow–band signals so that ISI can be neglected. In addition, we assume the coded symbols belong to a real–valued symbol alphabet $\mathcal{S} \subseteq \mathbb{R}$ (i.e. each symbol uses 1 signaling dimension) since we wish to express our results on a per dimension basis. Extending this analysis to complex symbols is straightforward and brings no significant additional insight into the problem. We may write the continuous–time problem equivalently as

$$r_{f,k} = \sqrt{\alpha_f}c_{f,k} + z_{f,k} \tag{15}$$

where the $z_{f,k}$ are i.i.d. zero–mean Gaussian random variables with variance $N_0/2$. Under the Rayleigh fading model, α_f is an exponentially distributed random variable with unit mean (i.e. $f_{\alpha_f}(u) = e^{-u}, u \geq 0$). For unit–energy Ricean fading with a specular to diffuse power ratio K, α_f has the following density

$$f_{\alpha_f}^{\mathrm{Rice}}(a) = (1+K)\exp(-K(1+(1+1/K)a))\mathrm{I}_0\left(\sqrt{aK(1+K)}\right) \tag{16}$$

where $\mathrm{I}_0(\cdot)$ is the zero–order modified Bessel function of the first kind.

We define the NF–dimensional vectors, $\mathbf{r}$ and $\mathbf{c}$, representing the received and transmitted symbols over the F blocks, to which refer as a *frame* from this point onward. We take for granted that the transmitter

and receiver have agreed beforehand to use a codebook having M codewords so that the information rate is $R = (\log_2 M)/NF$ bits/dimension. We denote the F–dimensional vector of signal amplitudes by $\boldsymbol{\alpha}$, and assume that there is no feedback path so that the transmitter has no *a priori* knowledge of $\boldsymbol{\alpha}$. As a result, the transmitter and receiver agree beforehand on acceptable choices for R and the input source density $f_{\mathbf{C}}$ and do not modify them during the course of communication.

3.1 Frame Error Rates

We first recall an upper–bound on the ensemble average probability of codeword error (i.e. taken over all possible codes chosen at random) conditioned on the channel state $\mathbf{A} = \boldsymbol{\alpha}$. In our case this corresponds to the *frame error rate (FER)*. We denote this probability by $\overline{P_{\text{ens}|\mathbf{A}=\boldsymbol{\alpha}}}$. From [Gal68] we have that

$$\overline{P_{\text{ens}|\mathbf{A}=\boldsymbol{\alpha}}} \leq 2^{-NF(\mathrm{E}_0(\rho, f_{\mathbf{C}}, \mathbf{A}=\boldsymbol{\alpha}) - \rho R)} \tag{17}$$

where

$$\mathrm{E}_0(\rho, f_{\mathbf{C}}, \mathbf{A}=\boldsymbol{\alpha}) = -\frac{1}{NF}\log_2 \int_{\mathbf{r}} \cdots \int \left(\int_{\mathbf{c}} \cdots \int f_{\mathbf{C}}(\mathbf{c}) f_{\mathbf{R}|\mathbf{C},\mathbf{A}}(\mathbf{r}|\mathbf{c}, \alpha)^{\frac{1}{1+\rho}} d\mathbf{c} \right)^{1+\rho} d\mathbf{r}, \tag{18}$$

and ρ is arbitrary in $[0, 1]$. By maximizing over ρ, this can be expressed further as

$$\overline{P_{\text{ens}|\mathbf{A}=\boldsymbol{\alpha}}} \leq \begin{cases} 1 & I_{\mathbf{A}} < R \\ 2^{-NFE_r(R, f_{\mathbf{C}}, \mathbf{A}=\boldsymbol{\alpha})} & I_{\mathbf{A}} \geq R \end{cases} \tag{19}$$

where

$$\mathrm{E}_r(R, f_{\mathbf{C}}, \mathbf{A}=\boldsymbol{\alpha}) = \max_{0 \leq \rho \leq 1} \mathrm{E}_0(\rho, f_{\mathbf{C}}, \mathbf{A}=\boldsymbol{\alpha}) - \rho R, \tag{20}$$

and, under the assumption that the received signal in each frame is independent of the information transmitted in previous frames,

$$I_{\mathbf{A}} = \frac{1}{NF}\sum_{f=1}^{F-1}\int_{\mathbf{c}_f\in\mathcal{S}^N}\int_{-\infty}^{\infty} f_{\mathbf{R}_f,\mathbf{C}_f|\alpha_f}(\mathbf{r}_f;\mathbf{c}_f|\alpha_f)\cdot \log_2\frac{f_{\mathbf{R}_f|\mathbf{C}_f,\alpha_f}(\mathbf{r}_f|\mathbf{c}_f,\alpha_f)}{f_{\mathbf{R}_f|\mathbf{A}_f}(\mathbf{r}_f|\alpha_f)}d\mathbf{r}_f d\mathbf{c}_f \quad \text{bits/dim.} \tag{21}$$

We note that this is not a conditional mutual information functional and to avoid confusion we have used a slightly different notation. The conditional average mutual information between $\mathbf{R}$ and $\mathbf{C}$ is $I(\mathbf{R};\mathbf{C}|\mathbf{A}) = \mathrm{E}_{\mathbf{A}} I_{\mathbf{A}}$, which when maximized over the input distribution, is the capacity of an ergodic fading channel [Eri70].

We may bound the code–ensemble average probability of error as

$$\overline{P_{\text{ens}}} = \mathrm{E}_{\mathbf{A}}\overline{P_{e|\mathbf{A}=\boldsymbol{\alpha}}} \le P_{\text{out}}(R, f_{\mathbf{C}}) + \int_{\boldsymbol{\alpha}:I_{\mathbf{A}}\ge R} 2^{-NFE_r(R,f_{\mathbf{C}},\mathbf{A}=\boldsymbol{\alpha})} dF_{\mathbf{A}}(\boldsymbol{\alpha}) \tag{22}$$

where

$$P_{\text{out}}(R, f_{\mathbf{C}}) = \mathrm{Prob}(I_{\mathbf{A}} < R). \tag{23}$$

Unlike the time–invariant channel case, the irreducible term in (22) ($P_{\text{out}}(R, f_{\mathbf{C}})$) is independent of N which means that arbitrarily small error probabilities need not be achievable.

To get an idea of the achievable performance we now express the average codeword error probability for a particular code (i.e. not an ensemble average) as

$$\begin{aligned}\overline{P_e} &= \overline{P_{e|I_{\mathbf{A}}\ge R}}\,(1 - P_{\text{out}}(R, f_{\mathbf{C}})) + \overline{P_{e|I_{\mathbf{A}}<R}}P_{\text{out}}(R, f_{\mathbf{C}}) \\ &\ge \overline{P_{e|I_{\mathbf{A}}<R}}P_{\text{out}}(R, f_{\mathbf{C}})\end{aligned} \tag{24}$$

Practically speaking, this lower–bound on $\overline{P_e}$ is only meaningful if we consider the *strong converse* to the coding theorem which guarantees that $\overline{P_{e|I_{\mathbf{A}}\ge R}}$ tends to 1 with increasing N for *all* codes. This result can be extended to show [Gal68] that it must tend to 1 exponentially in N. In our context, this ensures that $\overline{P_{e|I_{\mathbf{A}}\ge R}} \approx 1$ if N is large so that $\overline{P_e} \gtrapprox P_{\text{out}}(R, f_{\mathbf{C}})$. In the limit of large N, the coding theorem gives us equality, since $P_{e|I_{\mathbf{A}}\ge R}$ is bounded by 1.

The reader may wonder whether why the previous result is meaningful, since the use of the strong converse says nothing about the error probability of the individual source bits or the *bit error rate*(BER). We note, however, that for many practical systems it is precisely the FER that is important. This is true for the transmission of some forms of digitized speech and in packet data communications. Typically, data is arranged into frames and then coded for transmission using both error correction and error detection techniques. At the receiver the frame is decoded and then checked for data integrity using the error detection scheme. If it is deemed intact, the data is passed on to the next level of the system. On the other hand, if the data is corrupted then the frame is often discarded or a retransmission is requested. Provided the number of symbols in the frame (NF) is large and a sophisticated coding scheme is used, $P_{\rm out}(R, f_{\mathbf{C}})$ will be a good indicator of the achievable error rate performance.

3.2 The Weak Converse and Bit–Error Rates

The weak converse (Fano's inequality) yields a less useful lower bound on the FER since it only shows that $P_{e|I_{\mathbf{A}}>R}$ is bounded away from zero when in an outage state. It is, however, more useful for obtaining a lower–bound to the BER, P_b. We have [Bla87] that the BER conditioned on $I_H > R$ satisfies

$$\mathcal{H}(P_{b|I_{\mathbf{A}}>R}) \geq 1 - \frac{I_{\mathbf{A}}}{R} \tag{25}$$

where $\mathcal{H}(\cdot)$ is the binary entropy function $\mathcal{H}(x) = -x\log_2(x) - (1-x)\log_2(1-x)$. The expression in (25) is only valid when the information source has maximum entropy. This yields the lower bound

$$P_b \geq \int_{\boldsymbol{\alpha}: I_{\mathbf{A}} \leq R} \mathcal{H}^{-1}\left(1 - \frac{I_{\mathbf{A}}}{R}\right) dF_{\mathbf{A}}(\boldsymbol{\alpha}), \tag{26}$$

where $\mathcal{H}^{-1}(\cdot)$ is taken to mean the smaller of the two roots of (25). We will see that the FER for many systems is quite close or even below $P_{\rm out}(R)$, but that the lower bound on the BER is rather optimistic for practical codes.

3.3 Discussion

In some cases the information outage probability may be zero, or equivalently a non–zero channel capacity exists. Any channel with $\min_{\mathbf{A}} I_{\mathbf{A}} > 0$ will exhibit this behaviour. It is also possible when the transmitter has *a priori* knowledge of the channel state and can adjust either R or $f_{\mathbf{C}}$ accordingly

[Gol94]. Another somewhat unrealistic case is when $F \to \infty$ and the $c_{f,k}$ are independent of each other and the channel state [OSSW94][Eri70], where by the law of large numbers $I_{\mathbf{A}} \to \frac{1}{N} I(R_{f,k}; C_{f,k}|A_f)$ bits/dim.

Kaplan and Shamai [KSS95] consider another outage probability based on the instantaneous cutoff rate in the place of the average mutual information. The behaviour of this measure is similar to (23) except that the outages are noticeably higher. It has less theoretical justification than (23) since it cannot be used to obtain a bound on the FER. Traditionally [Mas74],[Vit79] R_0 was taken to be the highest rate at which practical coding schemes can be implemented on ergodic channels. Humblet [Hum85] showed that on a direct detection optical channel there exist reasonably simple codes whose rates exceed those predicted by R_0 with acceptably low error probability. In recent years, the invention of *turbo codes* [BGT93] provides more evidence that R_0 is not a practical limit even on a Gaussian channel. We will soon see that some practical codes, which are not even as complex as turbo codes, can come very close to (23) when F is small, which shows that mutual information outage is sometimes more appropriate in our case as well. The main reason for this is that when the number of blocks (or, more generally, degrees of freedom of the fading process) is small, $P_{\text{out}}(R)$ is quite high and even fairly simple codes have FER on the order of $P_{\text{out}}(R)$ when $I_{\mathbf{A}} > R$ (i.e. when the system is not operating in an outage situation.) When this is the case, the average FER is dominated by the outage event.

3.4 AWGN Channels and Finite Symbol Alphabets

We now compute $P_{\text{out}}(R)$ for different symbol alphabets. Under an average power constraint $\sum_{n,f} \overline{c_{n,f}^2} < NF\mathcal{E}_s$, it well known [Gal68] that $I_{\mathbf{A}}$ is maximized when the $c_{n,f}$ are i.i.d. zero–mean continuous Gaussian random variables, yielding

$$I_{\mathbf{A}} = \frac{1}{F} \sum_{f=0}^{F-1} \frac{1}{2} \log_2 \left(1 + \frac{2\alpha_f \mathcal{E}_s}{N_0} \right) \text{bits/dim} \tag{27}$$

The corresponding information outage probability in this case is easily computed numerically, and a Chernov upper–bound in terms of Whittaker functions is discussed in [KSS95]. A tight lower–bound is found using the fact that

$$\log_2(1 + \frac{2\mathcal{E}_s}{N_0}\alpha_f) \leq \frac{1}{\ln 2} \left(\frac{2\mathcal{E}_s/N_0}{1 + 2\mathcal{E}_s\alpha_o/N_0} \right) (\alpha_f - \alpha_o) + \log_2 (1 + 2\mathcal{E}_s\alpha_o/N_0) \tag{28}$$

which for Ricean fading with specular–to–diffuse power ratio K yields

$$P_{\text{out}}(R) \geq \text{Prob}\left(\sum_{f=0}^{F-1} \alpha_i < \mu\right) = \text{Q}_F\left(\sqrt{2FK}, \sqrt{(2(K+1)\mu}\right) \tag{29}$$

where $\mu = 2F \ln 2 \left(2R - (\alpha_0 + N_0/2\mathcal{E}_s)\left(1 + \frac{2\mathcal{E}_s\alpha_0/N_0}{(1+2\mathcal{E}_s\alpha_0/N_0)\ln 2}\right) - \log_2(1 + 2\mathcal{E}_s\alpha_0/N_0)\right)$ and $\text{Q}_F(a,b)$ is the *Marcum Q–function of order* F. The bound can be tightened by maximizing (29) with respect to α_0. For the special case of Rayleigh fading ($K = 0$) we have

$$P_{\text{out}}(R) \geq e^{-\mu} \sum_{f=0}^{F-1} \frac{1}{f!}\mu^k. \tag{30}$$

A Gaussian input distribution is useful for assessing the potential performance of large signaling constellations. We show the outage probability for unit–energy Rayleigh fading and Ricean fading at a signal–to–noise ratio (SNR) of $\mathcal{E}_s/N_0$ =7 dB for $F = 1, 2, 4, 8$, which are reasonable choices for next generation mobile systems. The Ricean fading channel has a specular to diffuse ratio of K =6 dB which was measured for some typical indoor communication channels [Bul87]. The main conclusion to be drawn from these curves is that we cannot expect to transmit at spectral efficiencies much higher than 1 bit/dim if we require frame error rates on the order of 10^{-2}, even with as many as $F = 8$ independent blocks and a fairly strong specular signal component. In a recent study, Caire *et al* [CKH97] have applied these ideas to interference–limited FDM–TDMA cellular systems and have shown that comparable spectral efficiencies can be expected under certain assumptions regarding the system architecture (i.e. power control, frequency/time–hopping, frequency reuse and basestation assignment.) This motivates our search for practical codes operating in the range .25-1.5 bits/dim in the following sections.

We now examine the effect of using small constellations with equiprobable and independent symbols. In this case $P_{\text{out}}(R)$ can similarly be computed numerically using [Wil96]

$$I_{\mathbf{A}} = \log_2 |\mathcal{S}| - \frac{1}{F|\mathcal{S}|} \sum_{f=0}^{F-1} \sum_{s_i \in \mathcal{S}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{1}{N_0}(r - \sqrt{\alpha_f} s_i)^2\right) \cdot$$
$$\log_2 \sum_{s_j \in \mathcal{S}} \exp\left(-\frac{1}{N_0}\left[(r - \sqrt{\alpha_f}(s_j - s_i))^2 - r^2\right]\right) dr \ \ \text{bits/dim} \tag{31}$$

In Figs. 3(a),(b) we show $P_{\text{out}}(R)$ now as a function of the SNR per information bit $\mathcal{E}_b/N_0$ (where $\mathcal{E}_b = R\mathcal{E}_s$) for both small AM constellations and Gaussian signals in unit–mean Rayleigh fading and

spectral efficiencies of .5 and .75 bits/dim. The most important observation is that for a slight increase in the constellation size with respect to the minimum needed to achieve the target spectral efficiency, we approach the performance achievable with a continuous Gaussian input signal. We notice, however, that the diversity order (i.e. the slope of the error–rate curve) is low when the smallest constellation is used for transmitting at the target spectral efficiency (i.e. 2-AM for .5 bits/dim and .75 bits/dim). A slight constellation expansion (usually by a factor 2) can significantly increase the diversity order. In the following section we examine this observation more closely.

4 Maximum Code Diversity and the Singleton Bound

This section addresses practical block and convolutional codes which attain maximum code diversity (d_{H}^{F}) for a given number of uncorrelated blocks and information rate. If we consider, for example, binary modulation and binary convolutional codes, they need not exhibit maximum (free) Hamming distance, and, in general, $d_{\mathrm{H}}^{F} \leq d_{\mathrm{free}}$. A simple example is the rate 1/2 binary convolutional code with binary modulation employed in the full–rate GSM standard shown in Fig.4. The output bits are interleaved over 8 blocks transmitted on widely spaced carriers. The minimum free Hamming distance path ($d_{\mathrm{free}} = 7$) (after deinterleaving) is shown along with the blocks over which each bit were transmitted. As the first two bits of the error event are in the same blocks as the last two, it is clear that this path achieves $d_{\mathrm{H}}^{8} = 5$. It turns out that this is also the minimum diversity path for this code and, as we shall soon see, that there is no other code which achieves a larger diversity with binary modulation and $R = 1/2$ bits/dim, even with arbitrarily many states.

The important conclusion to be drawn from this simple example is that traditional codes cannot necessarily be used effectively on non–ergodic fading channels. For the case when $F \to \infty$, however, Caire *et al.* have recently shown [CTB97] that *off-the-shelf* binary codes can be used with arbitrary signaling alphabets for achieving high diversity. This stems from the fact that Hamming distance is the dominant performance indicator, and a coded–modulation approach is not necessarily warranted for these types of channels. We will see that this is not really the case in the problem at hand, although Hamming distance, in a non–binary sense, is still the primary performance indicator.

4.1 Maximum Diversity Bound

In order to determine the minimum pairwise d_{H}^F, it is convenient to group together the N symbols which are transmitted in the same block, and view them as a super-symbol over $\mathcal{S}^N$. The codeword is then a vector of length F super-symbols. This is the same view taken by McEliece and Stark in [MS84], except now that F is fixed. Under this interpretation, d_{H}^F is simply the Hamming distance in $\mathcal{S}^N$. This reduces the analysis to one of non–binary block codes with a fixed block length F, and therefore all traditional bounding techniques apply.

An important first observation is that the highest rate code which achieves $d_{\mathrm{H}}^F = F$ has $R = \frac{1}{F}\log_2|\mathcal{S}|$ bits/symbol, which was also noted by Leung and Wilson in [LWK93]. This follows directly from the fact that no two codewords can have identical symbols in the same position if $d_{\mathrm{H}}^F = F$, and therefore the number of codewords cannot exceed $|\mathcal{S}|$. We can achieve this, for example, using a repetition code over $\mathcal{S}^N$. The question, therefore, is one of determining how close we can get to $d_{\mathrm{H}}^F = F$ with high–rate codes and simple constellations. The answer lies in the Singleton bound [Sin64] which is proven in this context, for the sake of completeness.

Theorem 1 (Singleton Bound)

Any code $\mathcal{C}$ of rate R bits/symbol with M codewords consisting of F blocks of length N symbols from an alphabet $\mathcal{S}$ has d_{H}^F satisfying

$$d_{\mathrm{H}}^F \leq 1 + \left\lfloor F\left(1 - \frac{R}{\log_2|\mathcal{S}|}\right)\right\rfloor. \tag{32}$$

Proof: Let k ($0 < k \leq F-1$) denote the integer value satisfying $|\mathcal{S}|^{N(k-1)} < M \leq |\mathcal{S}|^{Nk}$, where $M = 2^{NFR}$. Consider any set $\mathcal{I}_{k-1}$ of $k-1$ coordinates, for instance $\mathcal{I}_{k-1} = 0, 1, \cdots, k-2$. Since $M > |\mathcal{S}|^{N(k-1)}$ there are necessarily at least two codewords, $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ such that $\mathbf{x}_i = \mathbf{y}_i, \forall i \in \mathcal{I}_{k-1}$. It follows that $d_{\mathrm{H}}^F \leq F - k + 1$ and therefore that

$$M \leq |\mathcal{S}|^{N(F - d_{\mathrm{H}}^F + 1)}. \tag{33}$$

Using the fact that d_{H}^F must be an integer yields (32). This bound for binary codes was also given in the context of convolutional codes with finite interleaving by Wesel and Cioffi [WC95].

The first interesting result of this analysis is that the shape of the constellation is not important with regard to the code diversity since it is a completely algebraic measure of the performance. The class

of *maximum distance separable codes* (MDS) therefore plays a large role in this context. There is a downside, however, which is that the block length of the code is constrained to be F which means that many existing codes, such as the *Reed-Solomon (RS) codes* cannot necessarily be used effectively. Shortly, however, we give some examples of codes which can be used with practical choices for F and guarantee maximum diversity.

Secondly, and more importantly, we see what was remarked earlier in the outage probability analysis concerning constellation expansion. Take for example transmission at $R = .5$ bits/dimension over $F = 8$ blocks as in full–rate GSM. With binary modulation ($|\mathcal{S}| = 2$), the maximum pairwise diversity is 5, which is what is achieved by the coding scheme used in GSM. With quaternary modulation we see that it can be increased to 7. Examining the slopes of the information outage curves in Fig. 3 we see that both results agree. On the downside, for high code rates (> 2 bits/dimension) very large symbol alphabets are required to achieve high asymptotic diversity. For example, with $F = 8$ and $R = 3$ bits/dimension, a 16-point constellation can only achieve a diversity of $d_{\mathrm{H}}^8 = 3$. To achieve $d_{\mathrm{H}}^8 = 7$ a constellation with 4096 points is needed. Since d_{H}^F is only an asymptotic indicator, it may be somewhat pessimistic at low SNR ratios.

4.2 Block Codes

Let us first consider some examples of linear block codes of codeword length F with k information symbols, so that the rates of the codes are $R = \frac{k}{F} \log_2 |\mathcal{S}|$ bits/dimension. For this case the Singleton bound assures that $d_{\mathrm{H}}^F \leq F - k + 1$.

4.2.1 Simple codes

As we already pointed out, the simplest possible coding scheme for achieving diversity F is repetition coding. The number of codewords is $M = |\mathcal{S}|$ and the spectral efficiency is $\log_2 |\mathcal{S}|/F$ bits/dim. The parameter χ^2 for these codes is the minimum Euclidean distance of $\mathcal{S}$. The receiver for this coding scheme simply performs a maximal ratio combining of the F received symbols.

There are equally simple codes which outperform repetition codes. An example for $F = 2$ and $R = 1.5$ bits/2 dim is the free $\mathbb{Z}_8$–module generated by the matrix $(1, 3)$ mapped to 8–PSK. The minimum χ^2 is easily shown to be 1.41 and whereas for the trivial repetition code χ^2 would be .59. This is an important advantage since we gain a power savings greater than 3dB. The code for $F = 4$ for 4–AM at $R = .5$

bits/dim, $\{0202, 2020, 1133, 1331\}$ has $\chi^2 = 1.39$ whereas the trivial repetition code has $\chi^2 = .8$. For $R = .75$ and $R = 1.0$ bits/2 dim the free $\mathbb{Z}_8$ and $\mathbb{Z}_{16}$ modules generated by $(1, 3, 5, 7)$ yield χ^2 of 1.41 and 1.18 for 8 and 16–PSK respectively. Repetition codes would have χ^2=.59 and .15. In what follows, we will use these codes with simple trellis structures to show how χ^2 can be increased.

4.2.2 Multidimensional Constellations

The multidimensional lattice codes considered by Giraud and Belfiore [GB96] and Boutros *et al.* [BVRB96] are perfectly suited for the block–fading channel, since they consider constellations over a finite and small number of dimensions. Each dimension has an independent signal attenuation, and therefore in the context of the block–fading model, this is equivalent to letting F be the number of dimensions with $N = 1$. In [GB96] the constructed codes have dimensionality $2 \leq F \leq 8$ and $M = 2^{2F}$ points (codewords) which have diversity F so that the code rate is 2 bits/dim. In general, when the multidimensional constellations are projected onto the coordinate axes, they produce non–uniformly spaced AM constellations, with the minimum number of points necessary to satisfy the Singleton bound with $d_{\mathrm{H}}^F = F$.

The parameter χ^2 for these constellations is small because of the fact that $N = 1$, as is the case for the simple codes mentioned previously. In order to achieve higher coding gain but keep diversity F, therefore, it may be worthwhile to perform a coset decomposition (based on χ^2) of the constellations to be used in conjunction with trellis codes. We have not attempted this.

4.2.3 Other MDS Codes

We now consider MDS code families for systems having $F = 4, 6, 8$. They are formed by either shortening or lengthening RS codes. Shortening RS codes by removing information symbols results in a code with the same d_{H}^F as the base code. Similarly, it is shown in [Wol69] that up to 2 information symbols can be added to an RS code without changing d_{H}^F. For the case $F = 6$ we also consider a particular less complex extended Hamming code which is also MDS. The combination of the constraints imposed by the structure of the codes and the number of blocks in the system does not assure minimal complexity, nor the flexibility of choosing arbitrary symbol alphabets. Another negative aspect is that the purely algebraic structure of the codes pays no attention to the other less critical performance indicator, χ^2.

Example A : $F = 4$

Consider a family of codes with rate $R = k/4$ bits/dimension for use with binary modulation. Assuming we form symbols over GF(4) by forming pairs of bits from the same block, we start with the $(3, k-1)$ RS code over GF(4) with $d_{\mathrm{H}}^4 = 5 - k$ and lengthen it to $(4, k)$. The resulting parity check matrix for this code family is

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & \alpha & \alpha^2 \\ 0 & 1 & \alpha^2 & (\alpha^2)^2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \alpha^{3-k} & (\alpha^{3-k})^2 \end{pmatrix}. \tag{34}$$

These codes achieve maximum diversity for $k/4$ bits/dimension with binary modulation. Clearly, we could also use the same code with a quaternary symbol alphabet to achieve $R = k/2$ bits/dimension and keep the same diversity. Here we see the first example of the effect of constellation expansion; if we take $k = 2$ and binary modulation we have $R = .5$ bits/dimension and $d_{\mathrm{H}}^4 = 3$. With $k = 1$ and quaternary modulation the information rate is still $.5$ bits/dimension but $d_{\mathrm{H}}^4 = 4$.

Example B : $F = 6$

We now examine another family of codes with binary modulation and $R = k/6$ bits/dimension for the case when $F = 6$. Consider the $(7, k+1)$ family of RS codes over GF(8), having $d_{\mathrm{H}}^6 = 7 - k$. The parity check matrix for a shortened code family $(6, k)$ is given by

$$\mathbf{H} = \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^5 \\ 1 & \alpha^2 & (\alpha^2)^2 & \cdots & (\alpha^2)^5 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \alpha^{7-k} & (\alpha^{7-k})^2 & \cdots & (\alpha^k)^5 \end{pmatrix}. \tag{35}$$

This shortened family achieves maximum diversity for binary modulation and $R = k/6$ bits/dimension. We can also use this family with 8-ary modulation to yield $R = k/2$ bits/dimension and the same diversity level.

It is interesting to point out that although the codes are optimal in an MDS (maximum diversity) sense, there may be other less complex codes which are also MDS. For example, the (6,3) extended Hamming

code over GF(4) with generator matrix

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & \alpha & \alpha^2 \\ 0 & 0 & 1 & 1 & \alpha^2 & \alpha \end{pmatrix}, \tag{36}$$

is also MDS with $d_{\mathrm{H}}^6 = 4$ for $k = 3$. It is much less complex than the (6,3) shortened RS code outlined above (64 codewords instead of 512). Moreover, it can be used with a quaternary signal set.

Example C : $F = 8$

As a final example we consider the case of a code family with $R = k/8$ bits/dimension when $F = 8$ and $N = 3$. Similarly to when $F = 4$, we look at the $(7, k-1)$ family of Reed–Solomon codes over GF(8), having $d_{\mathrm{H}}^8 = 9 - k$. The parity check matrix for the lengthened code family $(8, k)$ is given by

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & \alpha & \alpha^2 & \cdots & \alpha^6 \\ 0 & 1 & \alpha^2 & (\alpha^2)^2 & \cdots & (\alpha^2)^6 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \alpha^{7-k} & (\alpha^{7-k})^2 & \cdots & (\alpha^k)^6 \end{pmatrix}. \tag{37}$$

This family achieves maximum diversity for binary modulation and $R = k/8$ bits/dimension. As before, we can also use this family with 8–ary modulation yield $R = 3k/8$ bits/dimension and the same diversity level.

4.3 Trellis Codes

In the GSM system today, as previously mentioned, rate 1/2 binary convolutional codes are used. This is mainly due to the computational simplicity of implementing the Viterbi algorithm with soft decisions. The Singleton bound is also applicable to arbitrary trellis codes, since they can always be interpreted as very long block codes. In fact, in systems like GSM the convolutional codes are used in a block fashion by appending trailing zeros to the information sequence, and a one–shot decoding of the entire block is performed.

We will consider two approaches for designing trellis codes for these types of channels. We first give three examples of simple 4–dimensional trellis codes based on AM and PSK constellations for $F = 4$.

This is the approach taken by Divsalar and Simon in [DS88] who considered TCM schemes for *perfectly interleaved* fading channels. It relies on using partitioning rules applied to F–dimensional Cartesian products of M–PSK constellations. The rules are chosen such each set has a specified amount of diversity and maximum product distance. Each set corresponds to the parallel transitions in the trellis, and by properly choosing the trellis structure, the diversity of the code can be made equal to the minimum Hamming distance of the sets. Wei considers similar codes for M–DPSK systems [Wei93]. In both cases the diversity order of the codes is quite small (≤ 4) considering the perfect interleaving assumption. A similar approach for multiple-antenna systems are the *space-time codes* introduced by Tarokh *et al.* [TSC97a, TSC97b].

None of these codes are automatically applicable in our case since we no longer have a system with perfect interleaving. As previously mentioned, Leung and Wilson [LWK93] designed simple 1.5 bit/2 dimensions 8–PSK trellis codes for systems with $F = 2$ blocks. For illustration purposes, we will do similarly for .5,.75 bit/dim for $F = 4$ blocks in section 4.3.1.

The second approach considered in section 4.3.2 is to search for linear MDS convolutional codes which maximize χ^2. Malkamäki and Leib [ML97] recently considered conventional rate $1/F$ binary convolutional codes interleaved across F blocks. Here we present the results of code searches for binary convolutional codes applied to BPSK/QPSK and 4–AM/16–QAM modulation with an appropriate mapping from the output bits to modulation symbols. Similar coding schemes for channels with correlated fading were described by Wesel and Cioffi [WC95]. We also consider convolutional codes over $\mathbb{Z}_8$ for 8–PSK and 8–AM modulation which achieve maximum diversity. Coded modulation schemes for the AWGN channel using ring convolutional codes were introduced by Massey *et al.* in [MM89].

4.3.1 4–dimensional Trellis Codes for $F = 4$

We now illustrate several codes for $F = 4$, having spectral efficiencies of .5,.75 and 1.0 bit/dim, which are based on the simple block codes described previously. From the Singleton bound we require constellations of size 4, 8 and 16 respectively in order to attain diversity 4. We consider the two and four state trellis diagrams and the constellations shown in 5. Each transition has an assigned set of 4–dimensional outputs, $\mathcal{S}_i$, whose cardinalities depend on the desired spectral efficiency. The branch outputs are chosen such that

1. The sub-codes comprising the state outputs when leaving and entering each state has diversity 4

2. The parallel transitions have as large a χ^2 as possible.

These heuristic guidelines do not guarantee an optimal code, but assure maximum diversity and a large χ^2.

4–AM and QPSK codes (.5 bits/ (2) dim)

For codes with .5 bit/dim, the number of input bits per 4-dimensional output is 2 so that there are 4 branches leaving and entering each state. We assume parallel transitions with two 4–dimensional symbols on each. Examining first the case of a 4–AM constellation. We choose the 4 sets as $\mathcal{S}_0 = \{0202, 2020\}, \mathcal{S}_1 = \{1133, 3311\}, \mathcal{S}_2 = \{3113, 1331\}, \mathcal{S}_3 = \{0220, 2002\}$. With this assignment, the χ^2 for the parallel transitions is maximum (3.2) under the constraint of satisfying the first design rule above. For the two–state code, the overall χ^2 is 2.77, whereas the 4–state code achieves $\chi^2 = 3.20$. With QPSK the 2–state code achieves $\chi^2 = 4.00$ which is maximum for any code with parallel transitions.

8–AM and 8–PSK codes (.75 bits/ (2) dim)

Now consider the case $R = .75$ bits/dim with 8–AM and 8–PSK modulation. We now have 3 bits per 4–dimensional output so that there are 8 branches leaving each state. For 8–AM we choose the 4 sets which obey the guidelines above

$$\mathcal{S}_0 = \{0246, 6024, 4602, 2460\}$$

$$\mathcal{S}_1 = \{7531, 1753, 3175, 5317\}$$

$$\mathcal{S}_2 = \{0426, 6024, 4062, 2640\}$$

$$\mathcal{S}_3 = \{7351, 1573, 3715, 5137\}.$$

With two states, the code has $\chi^2 = .9867$ and with four states $\chi^2 = 1.28$. With 8–PSK, we choose the sets as the free submodules $\mathcal{S}_i = (2j + i)(1, 3, 5, 7), i = 0, 1, j = 0, 1, 2, 3$ and $\mathcal{S}_i = \mathcal{S}_{i-2} + (0, 2, 6, 0), i = 2, 3$. Recall that this code was mentioned previously, and is optimal with respect to χ^2. With two states, the minimum $\chi^2 = 2.00$ is achieved by the parallel transitions so that there is no need to consider the 4-state code.

16–PSK codes (1 bit/2 dim)

Finally, we consider a 16–PSK code based on the simple block code mentioned earlier for $R = 1$ bit/2dim with $\mathcal{S}_i = (2j+i)(1,3,5,7), i = 0,1, j = 0,1,\cdots,7$ and $\mathcal{S}_i = \mathcal{S}_{i-2} + (0,4,12,0)$. Again for a 2–state code we achieve the minimum $\chi^2 = 1.41$ with the parallel transitions.

These techniques could be used to design codes for $F > 4$ and higher spectral efficiencies. It becomes a problem of finding MDS block codes with large χ^2 which can be *set partitioned* and assigned to the transitions in the trellis. These block codes quickly become very large, and as a result, the decoding complexity increases quickly. For example, a 2–state code for $F = 8$ and 1 bit/dim would require branches with 128 parallel transitions.

4.3.2 Convolutional Codes

Binary Convolutional Codes with BPSK/QPSK and 4–AM/16–QAM

Since d_{H}^F is a purely algebraic measure of the performance we have performed a code search for rate 1/4,1/2 and 3/4 binary convolutional codes that are MDS for $F = 2,4,8$. The results are summarized in Tables 1–8, and the MDS codes are highlighted in bold type. In addition, the codes listed maximize χ^2. The rate 1/4 and 1/2 codes are conventional feedforward convolutional codes with generators listed in octal notation following [LC83]. The rate 3/4 codes are systematic codes with feedback shown in figure 7. We have chosen the recursive form because of the reduced size of the search space. The generators are listed in hexadecimal form representing $(h_{i,4}, h_{i,3}, h_{i,2}, h_{i,1})$, where the leftmost bit is the most significant bit.

We have used a *Gray mapping* for adjacent bits out of the encoder for both the 4AM/16QAM and QPSK codes. This is shown in figure 6 and greatly simplifies the code search (which is already more computationally intensive than for computing d_{free}). If we write the modulation symbols as $s_i \in \{00, 01, 10, 11\}$ and denote the Euclidean distance between symbols as $d^2(s_i, s_j)$ then under the Gray mapping shown in the figure we have with QPSK $d^2(s_i, s_j) = d^2(00, s_i \oplus s_j)$ and with 4–AM/16–QAM $d^2(s_i, s_j) \geq d^2(00, s_i \oplus s_j)$. For 4–AM/16QAM we guarantee that the χ^2 between any two paths in the trellis is at least as large as the χ^2 between the all–zero path and their component-wise difference modulo 2.

As a general rule, we need very few states to yield a maximum diversity code, and χ^2 can be made substantially larger than those of the codes we constructed using the multidimensional approach in Section 4.3.1. We have also found that a candidate code must often be scanned to a great depth in order to find

the minimum diversity path.

As a first example, consider the case of $R = .5$ bits/dim with $F = 8$. We can achieve maximum diversity with an eight–state code, and moreover, it turns out that it does not exhibit maximum free Hamming distance ($d_{\text{free}} = 5$, not 6). It is the only such code, so that it is a perfect example of the danger of using selection rules appropriate only for ergodic channels. It is interesting to note that the GSM standard uses a 16–state maximum free Hamming distance code, which offers a slightly larger χ^2 than its 8–state counterpart. The 16–state code listed in the table has a slightly larger χ^2 than the GSM code, but we have found that the performance improvement is negligible. For the case of $F = 4$, maximum diversity can be obtained with a 4–state code, whereas in the GSM standard a 64–state code is used.

There are other important issues requiring the use of more complex codes. For instance, the 16-state code used in full–rate GSM achieves maximum diversity with $F = 2, 4, 6$ and 8, whereas the 8–state code achieves maximum diversity only with $F = 2, 4, 8$. This is important since in a frequency–hopping system, the number of hopping frequencies is left up to the operator. Although we have not considered this issue, it would be interesting to determine *universally* good codes which achieve acceptable performance for many different values of F. The more important reason for increasing complexity, as we will see in section 5, is that larger values of χ^2 can yield significant coding gain in the FER performance.

Convolutional Codes over $\mathbb{Z}_8$ for 8–PSK and 8–AM

For 8–ary modulation we have considered convolutional codes over the ring $\mathbb{Z}_8$ shown in figure 8. These linear codes were introduced by Massey *et al* in [MM89] and are naturally suited for phase modulated signals since the codewords form a multiplicative group in the signal space yielding a geometrically–uniform code [For91][BGMM93]. Moreover, it was found that they perform at least as well as any M–PSK code designed by set partitioning and are fairly easily made rotationally invariant. These codes have some peculiar algebraic properties due to presence of zero–divisors when using rings which are important to rule out catastrophic behaviour. We note that there are other configurations for achieving the same number of states but we have found that they yield less powerful codes in our case.

With the mapping shown in figure 6 for 8–AM we do not have geometric–uniformity but still assure, as was the case with 4AM/16QAM with binary codes and a Gray mapping, that $d^2(s_i, s_j) \geq d^2(0, s_i + s_j)$ where addition is now modulo 8. Again we need only consider each path with respect to the all–zero path to determine the minimum χ^2 for any code. This would not be the case for a code over a finite field.

5 Performance Comparison of Various Codes

In order to assess the performance of some of the codes reported in this work, we resort to computer simulations of a subset of codes. We have found that a union–bound approach for assessing the performance analytically yields quite unfruitful results for trellis codes. This was also remarked by Leung and Wilson[LWK93] and Malkamäki and Leib[ML97]. The main reason is that as we progress through the trellis the contributions of the long paths cannot be discarded since their diversity order is limited to F. For codes with a high diversity order, all paths have roughly parallel PEPs vs. SNR curves and therefore contribute to the total error probability. The number of paths to be considered in the union bound is very large and the bound is quite loose (depending on the point where we stop including paths.)

In our simulations shown in figures 9–16 we assumed a block length of 100 uncoded bits and a single–path Rayleigh fading channel with an independent realization in each block and soft–decision decoding with perfect channel state information. As a general rule, we find that with practical codes, we can often achieve FER close or, in a few cases, lower than $P_{\text{out}}(R)$, when the diversity order is low (e.g. $F = 2, 4$). This is the case since, for low diversity codes, $P_{\text{out}}(R)$ is quite high and even fairly simple codes operate on the order of or less than $P_{\text{out}}(R)$ when $I_{\text{A}} > R$, so that their performance is dominated by the outage event. Furthermore, for finite N, the FER is only approximately lower bounded by $P_{\text{out}}(R)$, so that it is possible for some codes to have an FER below this indicator. The FER is, of course, highly dependent on N, and for larger values of N than were investigated here, we would expect to require more complex codes in order to approach $P_{\text{out}}(R)$.

The binary convolutional code chosen for half-rate GSM ($F = 4, R = .5$ bit/dim, 64 states) does not maximize χ^2 but its FER and BER performance is very close to the code shown in figures 15,16. Both fall within .25 dB of $P_{\text{out}}(.5)$ with binary modulation so that the use of more than 64 states would be unnecessary. This assesment would be more difficult to make based only on d_{H}^F (or with an erasure model), since even a 4–state code is MDS. For the full-rate case ($F = 8$, $R = .5$ bit/dim, 16 states) we have not included the simulation for neither the code chosen in the GSM standard nor the one which maximizes χ^2. Again both have virtually identical BER and FER. They offer around 1dB gain in FER over the 8-state code shown in figures 15,16 and only a fraction of a dB in BER. Both these simulation results can be found in [KH97]. The code for $F = 8$ with 64 states falls within .5dB of $P_{\text{out}}(.5)$.

Some of the RS-based codes mentioned in (4.2) were simulated and have BER comparable with the simple convolutional codes although these results are not shown here. The performance enhancement

due to constellation expansion can be very significant, most notably for the examples at .75 bit/dim and 1.5 bit/2 dim. Note that we have not considered rate 3/8 codes for quaternary alphabets which would achieve diversity 3 for $F = 4$ and $R =$.75 bit/(2) dim. These would require rather inconvenient encoder structures, but would provide gains in between the binary and octal examples shown here.

We also remark that increased complexity has a much more significant effect on the FER than on the BER, especially for low diversity codes. We also notice the peculiar result that simple codes can have lower BER than more complex codes, when χ^2 is the selection criterion (e.g. figure 16, $F = 2$), even if their FER is significantly higher. In addition, the strict lower bound on the BER in (26) gives much less indication of practical performance than does $P_{\rm out}(R)$ for the FER.

6 Conclusion

This work considered coding for block–fading channels with small number of blocks. This channel model has significant practical importance for block–oriented communications where the fading process is characterized by a small number of degrees of freedom during the decoding interval. The slow frequency–hopping scheme used in the current GSM mobile radio system is a prime example. It is reasonable to assume that next generation wireless systems will also use similar, and perhaps more complex techniques.

We described the separation between the diversity effects due to multipath resolvability and coding. We then turned our attention to the attainable diversity due to coding. We showed that there is a upper–limit to the diversity which depends on the number of blocks, the code rate and the size of the signaling constellation. This was shown in two ways; the first was based on the computation of the information outage probability for various constellations. We then showed that the maximum diversity for a code of a given rate is given by the Singleton bound, so that appropriately chosen MDS codes play a very important role for these types of channels. Both methods indicate that diversity is limited and that it can be increased by constellation expansion. A rather unfortunate result is that for high spectral–efficiency systems, in order to achieve a high asymptotic diversity order, very large constellations are required.

We gave examples of block and trellis codes, with more of an emphasis on the latter, which achieve maximum diversity. An important result is that maximum diversity can be achieved with rather simple codes and that, in terms of BER performance, increased complexity does not always yield significant gains. This is not true, however, for the FER performance, which is often important in both speech and

data applications. We showed that the information outage probability is a good indicator for practical FER values, when the diversity order is rather low (<8). This result should also apply to cellular systems where coding is used to combat intercell interference [PC95][CKH97] as well as coded multitone systems [WC95].

References

[BCTV96] E. Biglieri, G. Caire, G. Taricco, and J. Ventura. "Simple Method for Evaluating Error Probabilities". *Electronic Letters*, 32(3):191–192, feb 1996.

[BGMM93] S. Benedetto, R. Garello, M. Mondin, and G. Montorsi. "Geometrically Uniform Partitions of $L\times$ MPSK Constellations and Related Binary Trellis Codes". *IEEE Transactions on Information Theory*, 39(6):1773–1798, November 1993.

[BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima. "Near Shannon Limit Error–Correcting Coding and Decoding: Turbo–Codes". In *Proc. IEEE ICC '93*, May 1993.

[Bla87] R.E. Blahut. *Principles and Practice of Information Theory.* Addison-Wesley, Reading, Massachusetts, 1987.

[Bul87] R.J.C. Bultitude. "Measurement, Characterization and Modeling of Indoor 800/900 MHz Radio Channels for Digital Communications". *IEEE Commun. Mag.*, 25(6):5–12, 1987.

[BVRB96] J. Boutros, E. Viterbo, C. Rastello, and J.C. Belfiore. "Good Lattice Constellations for Both Rayleigh Fading and Gaussian Channels". *IEEE Transactions on Information Theory*, 42(2):502–518, March 1996.

[CKH97] G. Caire, R. Knopp, and P.A. Humblet. "System Capacity of F-TDMA Cellular Systems". *accepted for publication in IEEE Trans. on Comm.*, 1997.

[CTB97] G. Caire, G. Taricco, and E. Biglieri. "Bit–Interleaved Coded Modulation". *IEEE Transactions on Information Theory*, 1997.

[DS88] D. Divsalar and M.K. Simon. "The Design of Trellis Coded MPSK for Fading Channels: Performance Criteria". *IEEE Transactions on Communications*, 36:1004–1012, 1988.

[Eri70] T. Ericson. “A Gaussian Channel with Slow Fading”. *IEEE Transactions on Information Theory*, 1970.

[FG97] G. Foschini Jr. and M. Gans. “On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas”. *submitted to Wireless Personal Communications*, 1997.

[For91] G.D. Forney Jr. “Geometrically Uniform Codes”. *IEEE Transactions on Information Theory*, 37:223–236, September 1991.

[Gal68] R.G. Gallager. *Information Theory and Reliable Communication*. John Wiley and Sons, 1968.

[GB96] X. Giraud and J.C. Belfiore. “Constellations Matched to the Rayleigh Fading Channel”. *IEEE Transactions on Information Theory*, 42(1):106–115, January 1996.

[Gol94] A. Goldsmith. *Design and Performance of High-Speed Communication Systems over Time–Varying Radio Channels*. PhD thesis, University of California at Berkeley, 1994.

[GSM90] European Telecommunications Standards Institute. *European Digital Cellular Telecommunications System : Physical Layer on the Radio Path (GSM 05.02)*, 1990.

[HH89] J. Hagenauer and P. Höher. “A Viterbi Algorithm with Soft–Decision Outputs and its Applications”. In *Proceedings of IEEE Globecom*, pages 47.1.1–47.1.7, Dallas, Tx, feb 1989. IEEE.

[Hum85] P.A. Humblet. “Error Exponents for a Direct Detection Optical Channel”. In *Proc. 23rd Allerton Conference on Communication, Control, and Computing*, October 1985.

[IS592] Telecommunications Industry Association. *EIA/TIA Interim Standard, Cellular System Dual mode Mobile-Station Base-Station Compatibility Standard IS-54B*, 1992.

[KH97] R. Knopp and P.A Humblet. “Maximizing Diversity on Block–Fading Channels”. In *Proc. IEEE ICC’97, Montreal, Canada*, June 1997.

[KSS95] G. Kaplan and S. Shamai (Shitz). “Error Probabilities for the Block–Fading Gaussian Channel”. *Archiv für Elektronik und Übertragungstechnik*, 49(4):192–205, 1995.

[Lap94] A. Lapidoth. "The Performance of Convolutional codes over the Block Erasure Channel Using Various Finite Interleaving Techniques". *IEEE Transactions on Information Theory*, 40(5), 1994.

[LC83] S. Lin and D.J. Costello Jr. *Error Control Coding: Fundamentals and Applications*. Prentice Hall, Englewood Cliffs, New Jersey, 1983.

[LWK93] Y.S. Leung, S.G. Wilson, and J.W. Ketchum. "Multi-frequency Trellis Coding with Low-Delay for Fading Channels". *COM*, 41(10):1450–1459, October 1993.

[Mas74] J.L. Massey. "Coding and Modulation in Digital Communications". In *Proc. Int. Zürich Sem. Digital Commun.*, Zürich, Switzerland, March 1974.

[ML97] E. Malkamäki and H. Leib. "Rate $1/n$ Convolutional Codes with Interleaving Depth of n over a Block-Fading Ricean Channel". In *Proc. IEEE Vehicular Technology Conference*, pages 2002–2006, May 1997.

[MM89] J.L. Massey and T. Mittelholzer. "Convolutional Codes Over Rings". In *Fourth Joint Swedish–USSR Int. Workshop on Information Theory*, 1989.

[MS84] R. McEliece and W.E. Stark. "Channels with Block Interference". *IEEE Transactions on Information Theory*, 30(1):44–53, January 1984.

[OSSW94] L. Ozarow, S. Shamai (Shitz), and A.D. Wyner. "Information Theoretic Considerations for Cellular Mobile Radio". *IEEE Transactions on Vehicular Technology*, 43(2):359–378, May 1994.

[PC95] G. Pottie and R. Calderbank. "Channel Coding Strategies for Cellular Mobile Radio". *IEEE Transactions on Vehicular Technology*, 44(4):763–769, November 1995.

[Pro95] J.G. Proakis. *Digital Communications*. McGraw Hill, New York, third edition, 1995.

[SBS66] M. Schwartz, W.R. Bennett, and Stein S. *Communication Systems and Techniques*. McGraw Hill, New York, 1966.

[Sin64] R.C. Singleton. "Maximum Distance q–Nary Codes". *IEEE Transactions on Information Theory*, 10:116–118, 1964.

[TSC97a] V. Tarokh, N. Seshadri, and A.R. Calderbank. “Space-Time Codes for High Data Rate Wireless Communications:Performance Criteria”. In *IEEE Internation Conference on Communications (ICC ’97)*, 1997.

[TSC97b] V. Tarokh, N. Seshadri, and A.R. Calderbank. “Space-Time Codes for High Data Rate Wireless Communications:Code Construction”. In *IEEE Vehicular Technology Conference (VTC ’97)*, 1997.

[Vit79] A.J. Viterbi. “Spread–Spectrum Communications- Myths and Realities”. *IEEE Comm. Mag.*, 4:11–18, May 1979.

[WC95] R. Wesel and J. Cioffi. “Fundamentals of Broadcast OFDM”. In *Proc. of the 29th Asilomar Conference on Signals Systems and Computers*, 1995.

[Wei93] L.-F. Wei. “Coded M–DPSK with Built–in Time Diversity for Fading Channels”. *IEEE Transactions on Information Theory*, 39(6):1820–1839, November 1993.

[Wil96] S.G. Wilson. *Digital Modulation and Coding*. Prentice Hall, Englewood Cliffs, New Jersey, 1996.

[Wol69] J.K. Wolf. “Adding Two Information Symbols to Certain Non–Binary BCH Codes and Some Applications”. *Bell Systems Technical Journal*, 48:2405–2424, 1969.

	$F=4$			$F=8$		
States	d^4_{H}	$\chi^2_{\min}$	gen.	d^8_{H}	$\chi^2_{\min}$	gen.
4	**4**	**9.80**	**5,5,7,7**	6	**7.27**	**5,3,7,7**
8	**4**	**12.90**	**64,64,54,74**	**7**	**7.81**	**44,64,54,34**
16	**4**	**14.89**	**52,62,66,76**	**7**	**12.29**	**46,26,64,76**
32	**4**	**17.71**	**71,55,75,57**			

Table 1 Rate 1/4 convolutional codes for binary modulation (.25 bit/dim)

	$F=2$			$F=4$			$F=8$		
States	d^2_{H}	$\chi^2_{\min}$	gen.	d^4_{H}	$\chi^2_{\min}$	gen.	d^8_{H}	$\chi^2_{\min}$	gen.
4	**2**	**9.80**	**5,7**	**3**	**6.35**	**5,7**	4	5.66	5,7
8	**2**	**12.00**	**64,54**	**3**	**10.08**	**44,54**	**5**	**4.00**	**44,64**
16	**2**	**12.65**	**62,72**	**3**	**13.21**	**62,46**	**5**	**5.28**	**62,72**
32	**2**	**16.00**	**71,73**	**3**	**14.54**	**75,57**	**5**	**10.56**	**51,65**
64	**2**	**17.89**	**704,564**	**3**	**17.93**	**724,564**	**5**	**18.47**	**414,354**

Table 2 Rate 1/2 binary convolutional codes for binary modulation (.5 bit/dim)

	$F=4$			$F=8$		
States	d^4_{H}	$\chi^2_{\min}$	gen.	d^8_{H}	$\chi^2_{\min}$	gen.
4	**4**	**2.58**	**5,7,3,7**	6	2.02	5,7,3,7
8	**4**	**3.76**	**44,64,54,34**	**7**	**1.76**	**44,64,54,34**
16	**4**	**4.63**	**72,76,44,54**	**7**	**2.55**	**64,56,50,66**
32	**4**	**5.71**	**61,75,53,57**	**7**	**3.63**	**41,75,45,33**
64	**4**	**6.60**	**624,634,564,564**	**7**	**5.00**	**644,370,424,354**

Table 3 Rate 1/4 binary convolutional codes for 4-AM (.5 bit/dim)

	$F=4$			$F=8$		
States	d^4_{H}	$\chi^2_{\min}$	gen.	d^8_{H}	$\chi^2_{\min}$	gen.
4	**2**	**5.66**	**9,A,F**	**3**	**4.00**	**9,A,F**
8	**2**	**11.31**	**9,A,3,F**	**3**	**8.00**	**F,9,A,F**
16	**2**	**13.85**	**F,9,6,5,A**	**3**	**12.00**	**F,9,C,6,F**
32	**2**	**16.00**	**9,A,3,9,5,F**			

Table 4 Rate 3/4 bits/dim convolutional codes for binary modulation (.75 bit/dim)

	$F=2$			$F=4$			$F=8$		
States	d^2_{H}	$\chi^2_{\min}$	gen.	d^4_{H}	$\chi^2_{\min}$	gen.	d^8_{H}	$\chi^2_{\min}$	gen.
4	**2**	**4.00**	**5,7**	**3**	**3.17**	**5,7**	3	3.17	5,7
8	**2**	**6.00**	**64,54**	**3**	**4.00**	**20,54**	4	2.83	64,54
16	**2**	**6.32**	**62,66**	**3**	**5.04**	**62,54**	**5**	**2.64**	**26,74**
32	**2**	**8.00**	**31,57**	**3**	**6.60**	**51,17**	**5**	**3.48**	**25,73**
64	**2**	**9.80**	**664,474**	**3**	**8.00**	**664,774**	**5**	**4.34**	**604,564**

Table 5 Rate 1/2 binary convolutional codes for QPSK (1 bit/2 dim)

	$F=4$		
States	d^4_{H}	χ^2_{min}	gen.
4	3	2.54	5,5,7,7
8	**4**	**1.61**	**44,64,50,74**
16	**4**	**2.08**	**52,56,66,76**
32	**4**	**2.54**	**51,55,66,76**

Table 6 Rate 1/4 binary convolutional codes for 16-QAM (1 bit/2 dim)

	$F=4$			$F=8$		
States	d^4_{H}	χ^2_{min}	gen.	d^8_{H}	χ^2_{min}	gen.
4	**2**	**2.82**	**9,A,7**	**3**	**4.0**	**9,A,7**
8	**2**	**4.00**	**8,5,7,1**	**3**	**8.0**	**F,8,9**
16	**2**	**6.00**	**9,F,3,6,5**	**3**	**12.0**	**F,9,C,6,7**
32	**2**	**7.78**	**9,A,3,9,5,7**			

Table 7 Rate 3/4 binary convolutional codes for QPSK (1.5 bit/2 dim)

	$F=2$			$F=4$		
States	d^2_{H}	χ^2_{min}	gen.	d^4_{H}	χ^2_{min}	gen.
4	**2**	**2.00**	**10,32**	2	2.00	10,12
8	**2**	**2.82**	**13,34**	**3**	**2.18**	**11,16**
16	-	-	-	**3**	**2.33**	**210,354**
32	**2**	**3.42**	**112,230**	**3**	**2.88**	**112,250**
64	**2**	**5.10**	**113,361**	**3**	**4.56**	**116,311**

Table 8 Rate 1/2 convolutional codes over $\mathbb{Z}_8$ for 8-PSK (1.5 bit/2 dim)

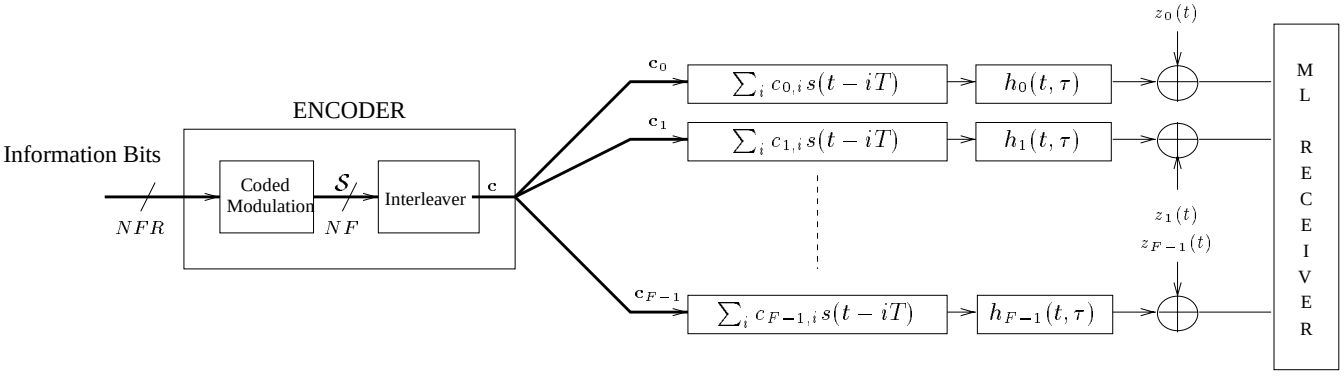

Fig. 1: System model

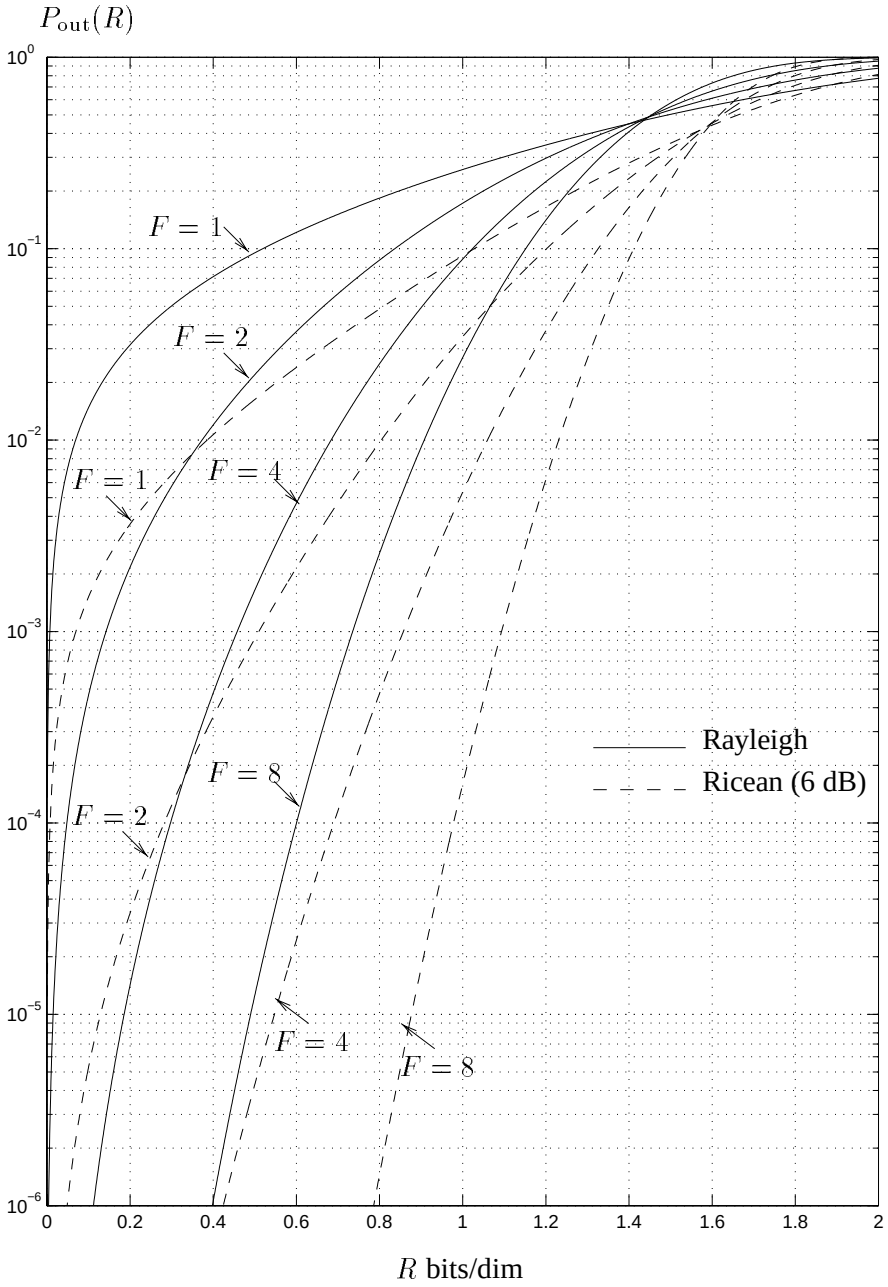

Fig. 2: Information Outage Probabilities for Rayleigh and Ricean Fading ($\mathcal{E}_s/N_0 = 7$ dB)

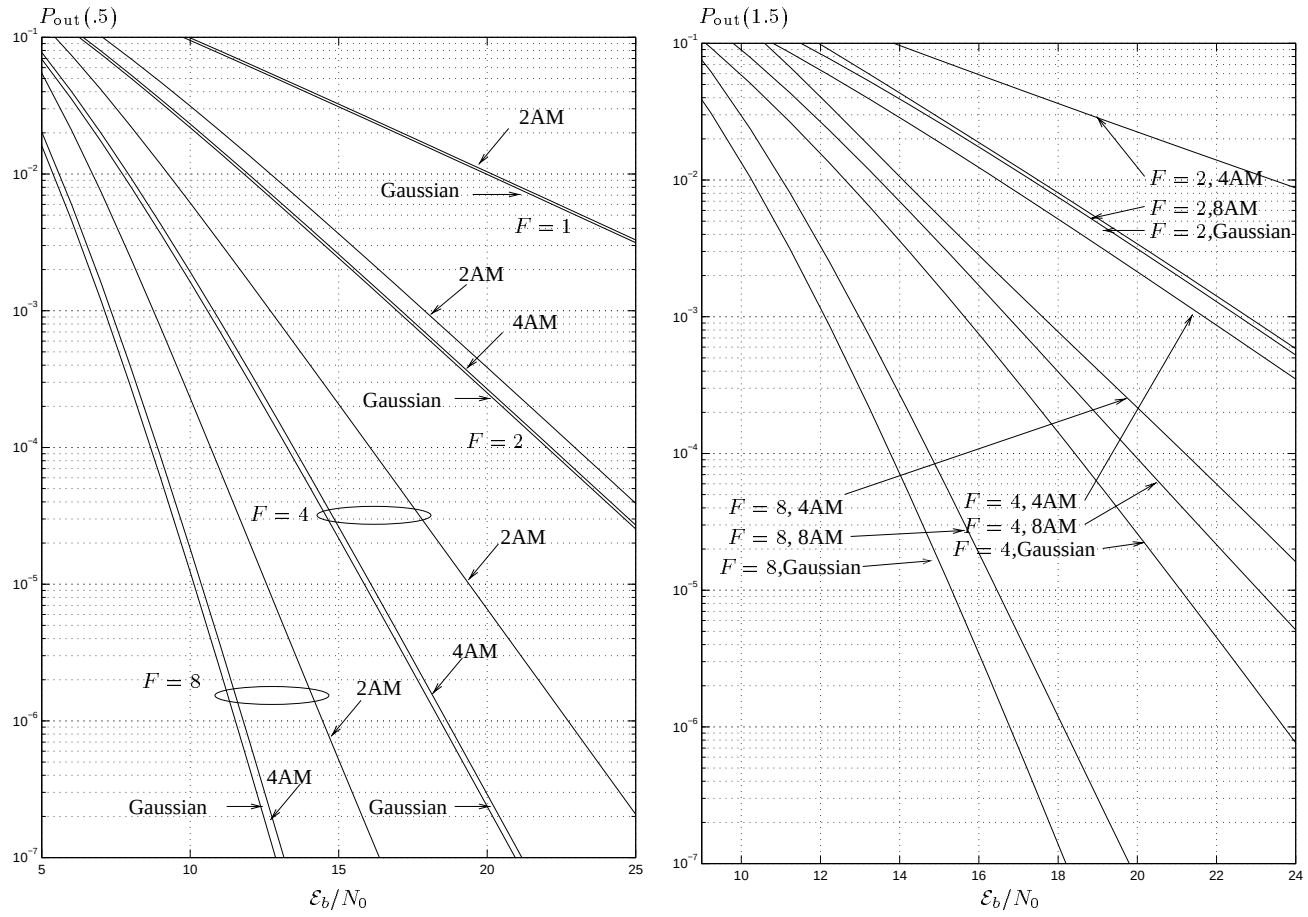

Fig. 3: Information Outage Probabilities vs. SNR for different size constellations ($R = .5, 1.5$ bits/dim)

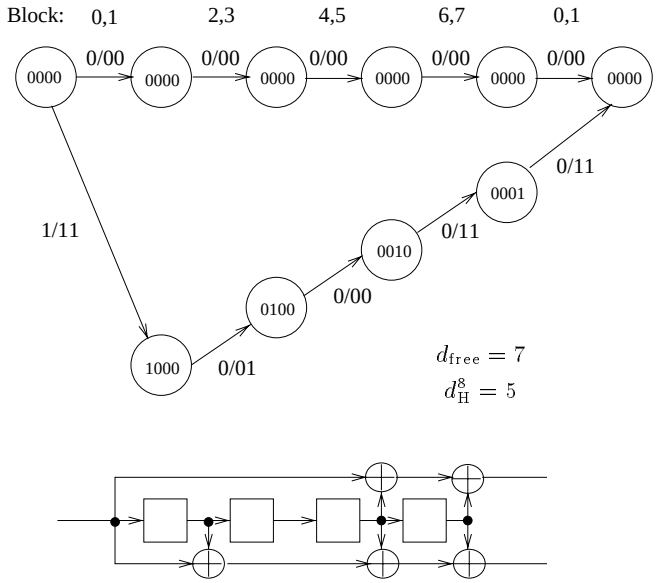

Fig. 4: Minimum diversity error event for full–rate GSM,

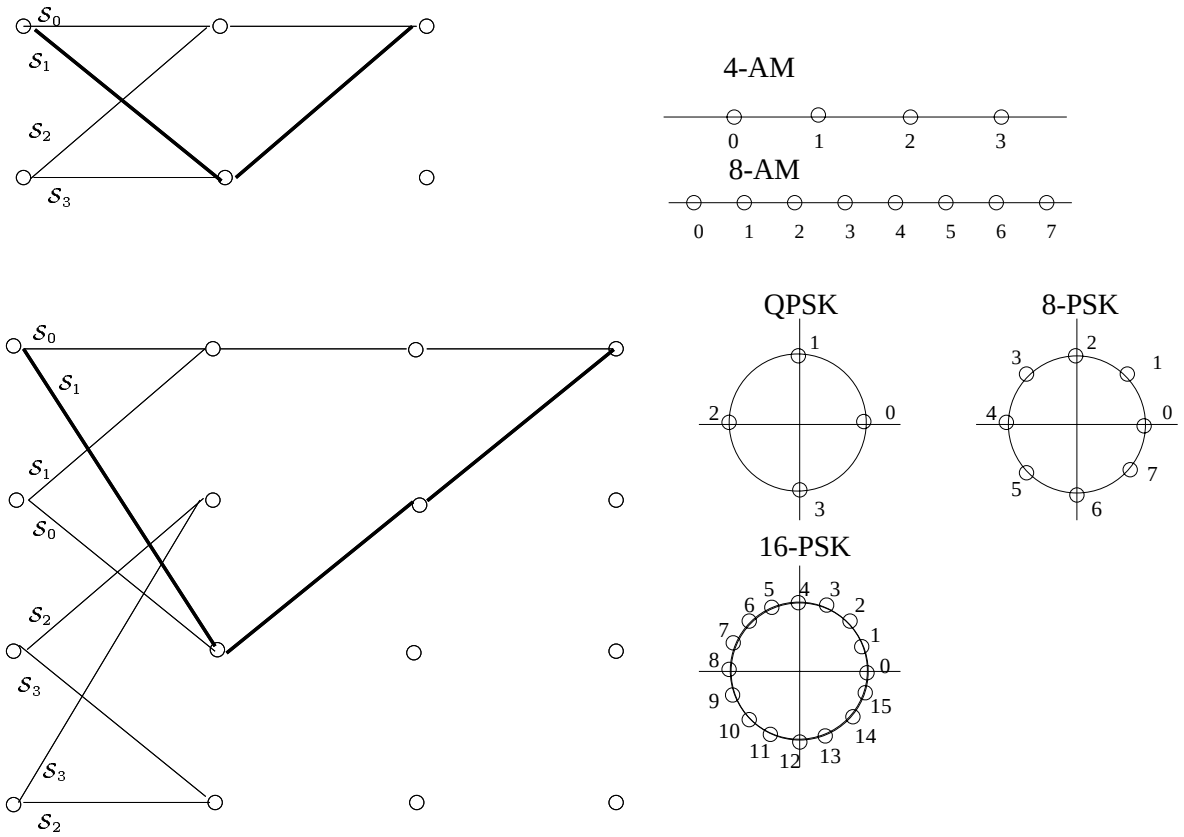

Fig. 5: Two and Four State Trellis Code Examples

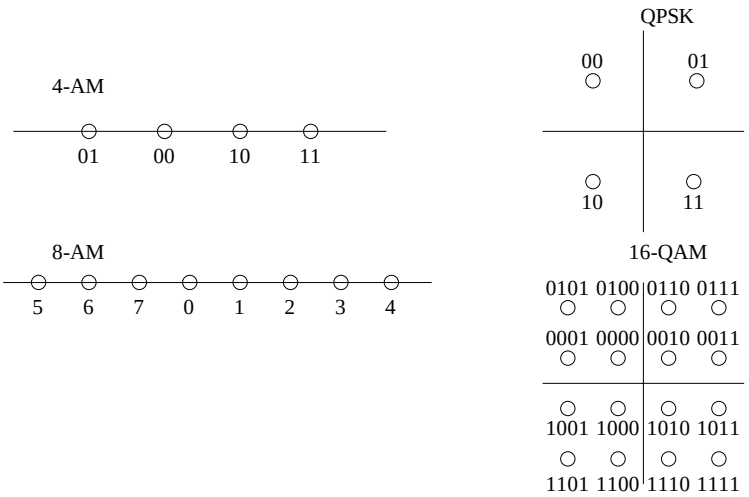

Fig. 6: Different Mappings

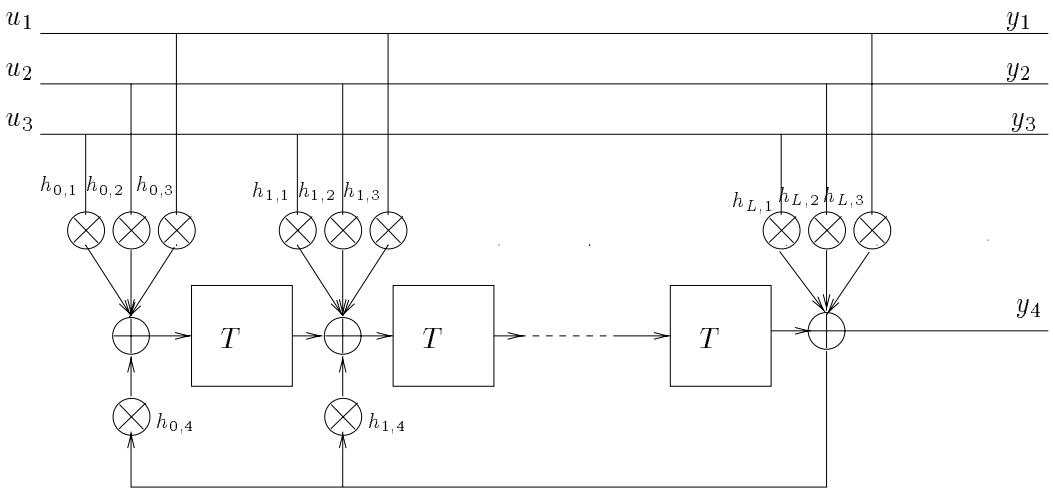

Fig. 7: Rate 3/4 Systematic Convolutional Code with Feedback

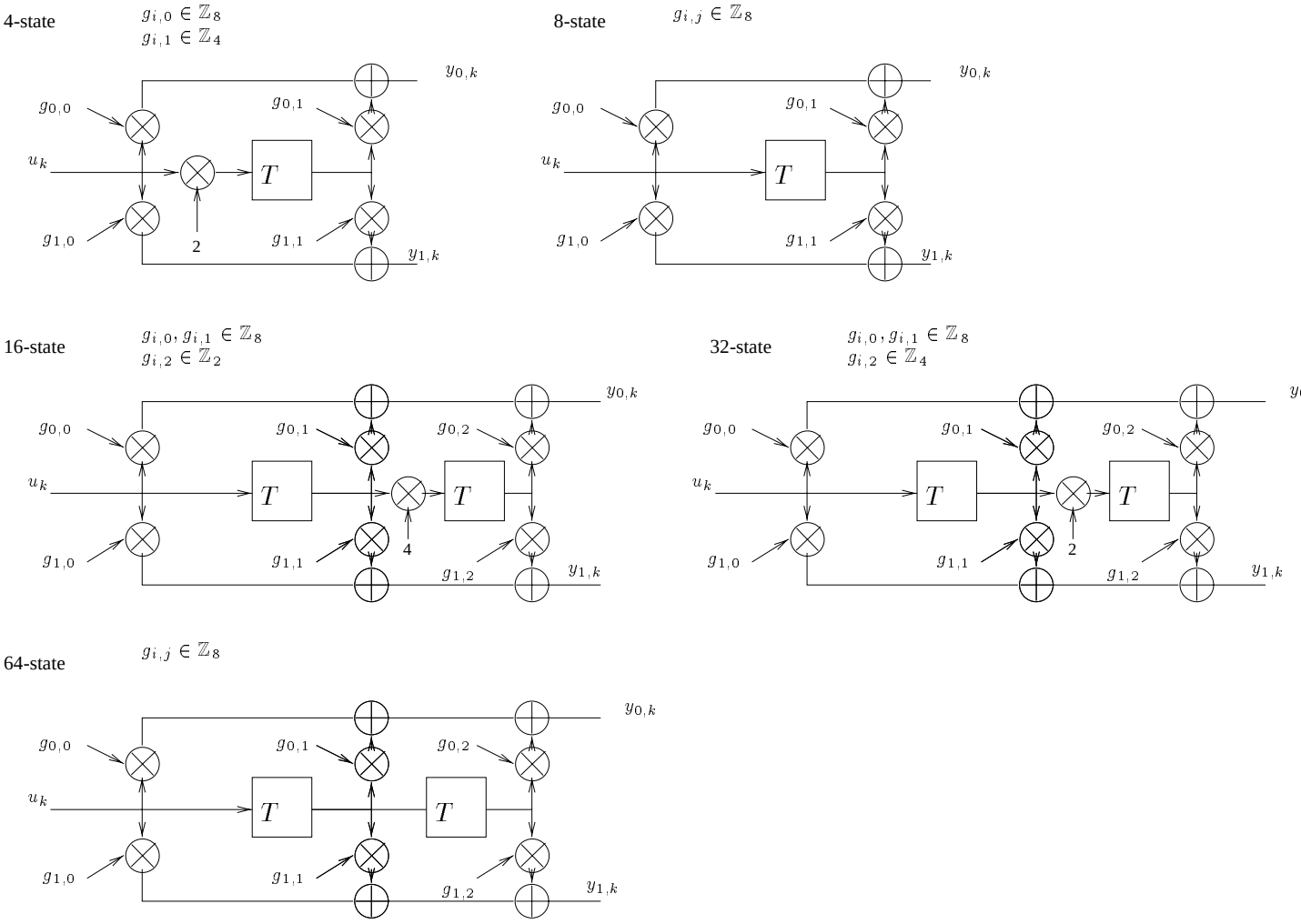

Fig. 8: Convolutional Codes over $\mathbb{Z}_8$

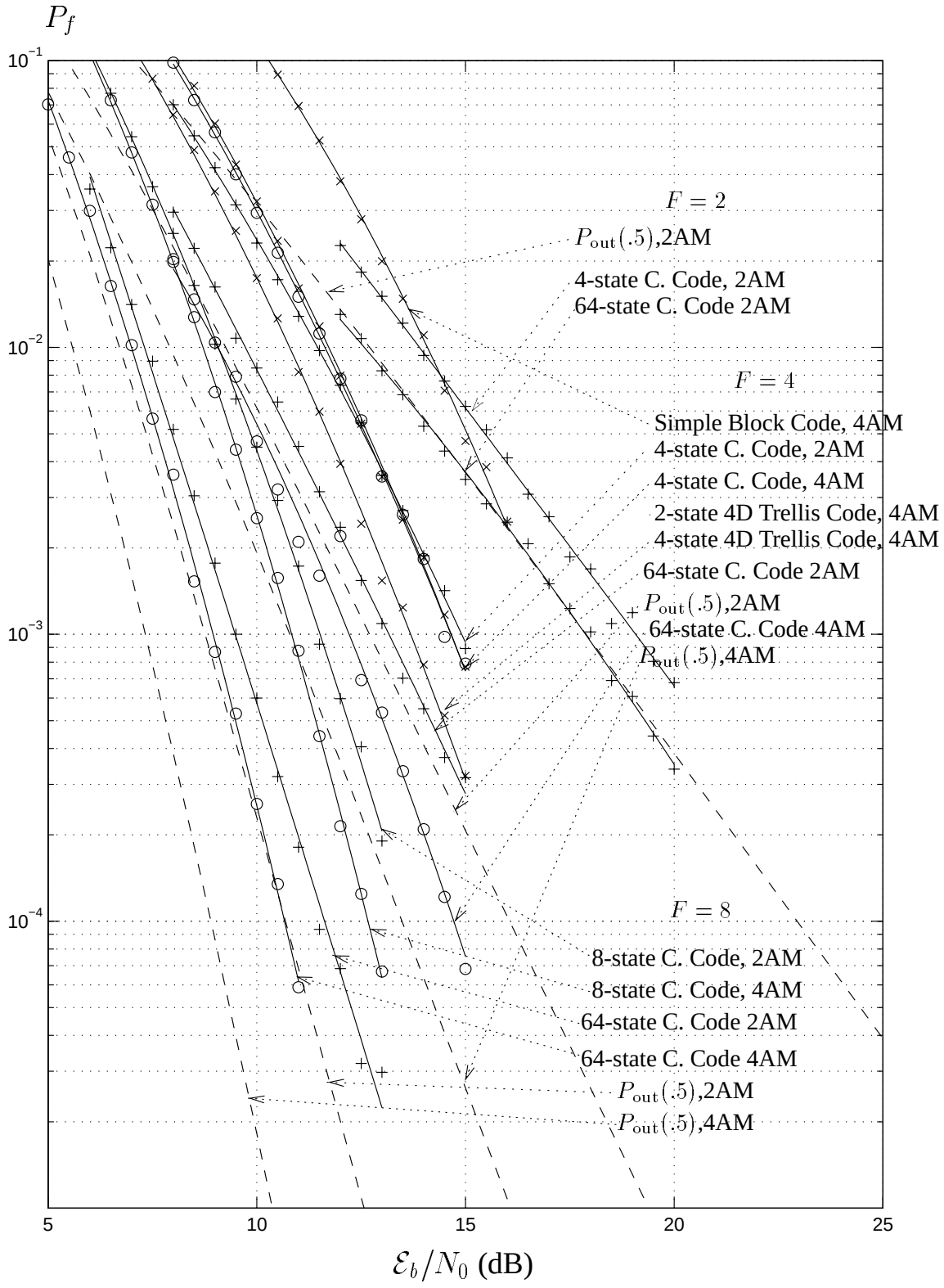

Fig. 9: Frame Error Rates for Codes with .5 bit/dim

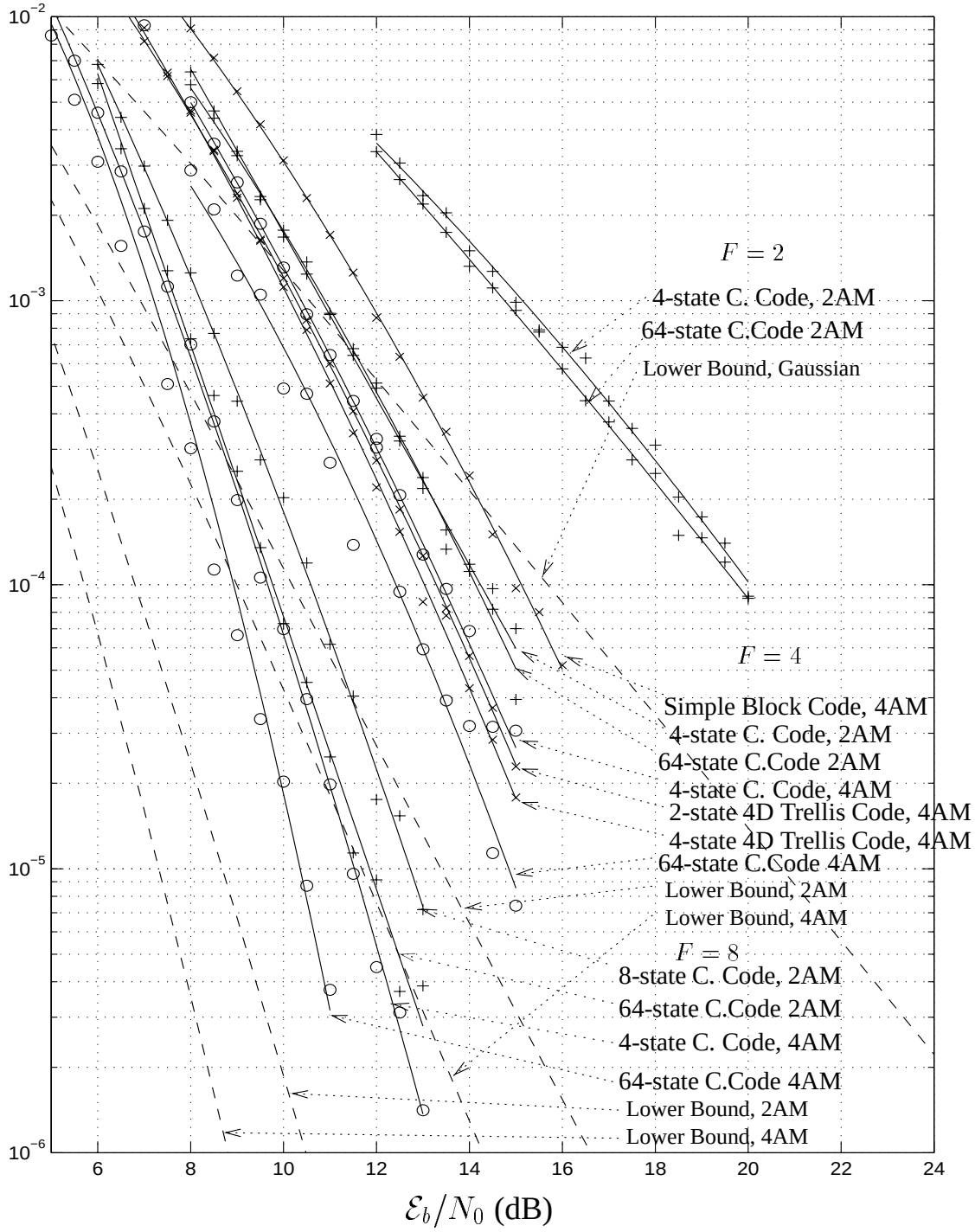

Fig. 10: Bit Error Rates for Codes with .5 bit/dim

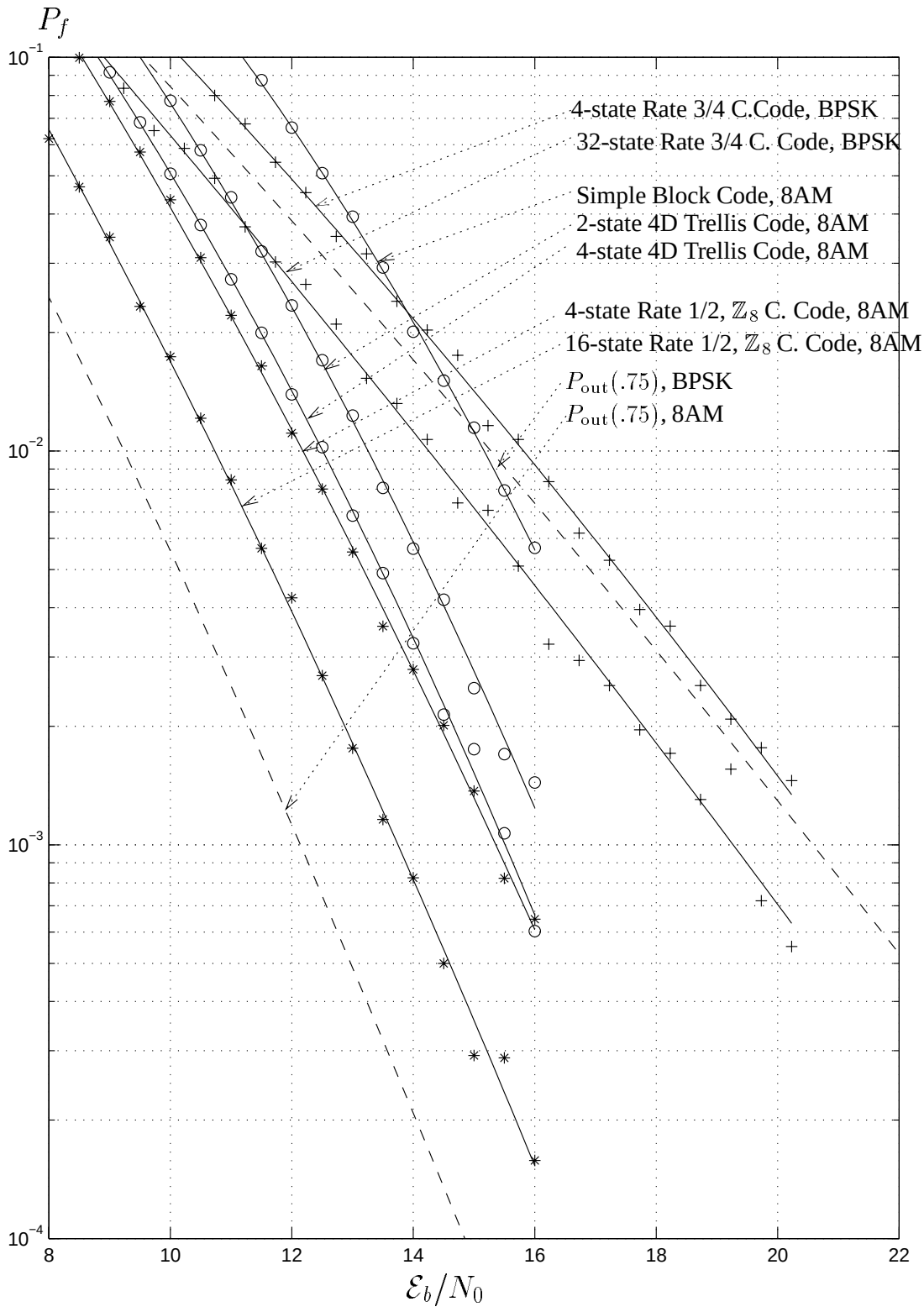

Fig. 11: Frame Error Rates for Codes with .75 bit/dim $F = 4$

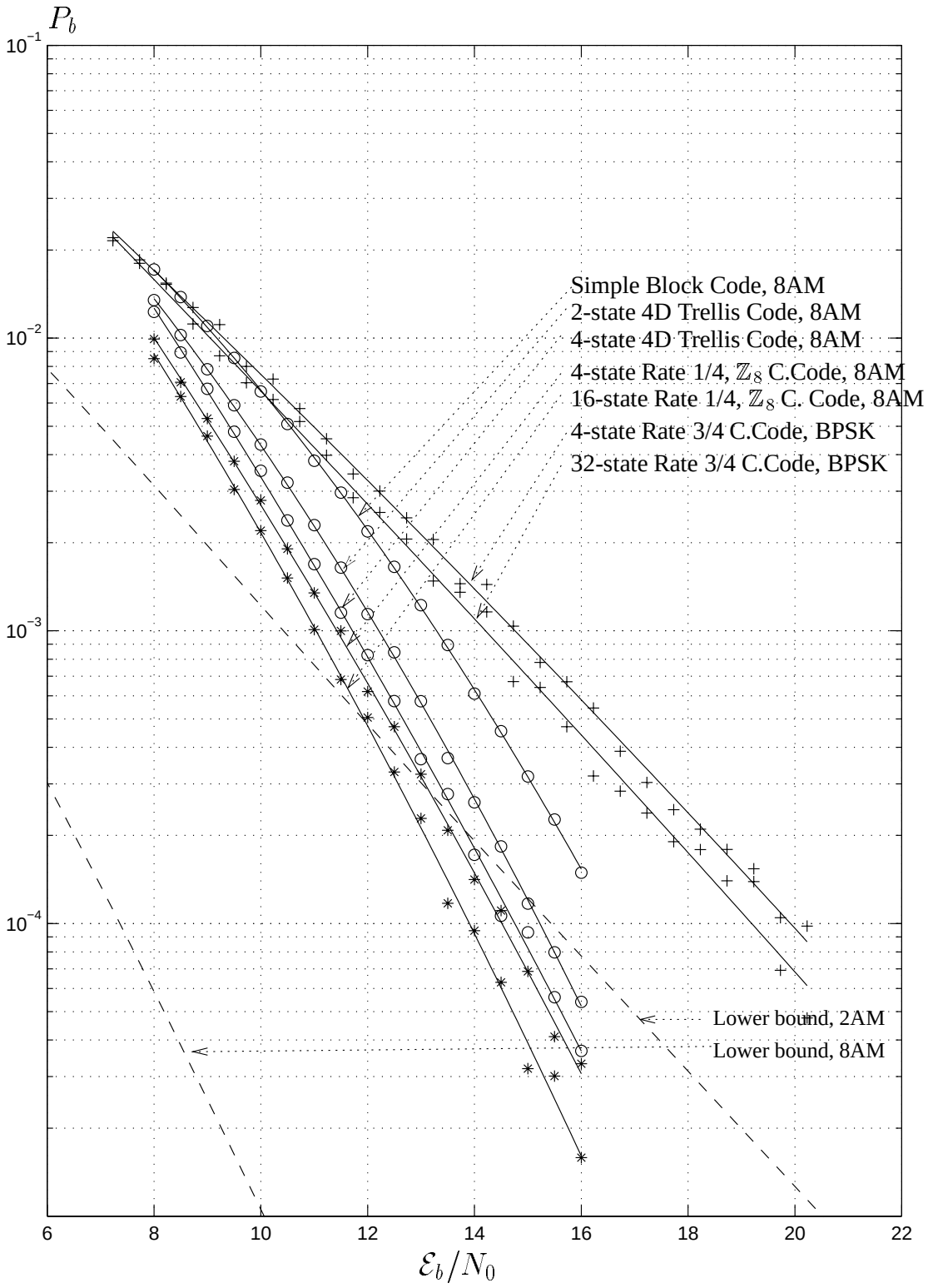

Fig. 12: Bit Error Rates for Codes with .75 bit/dim and $F = 4$

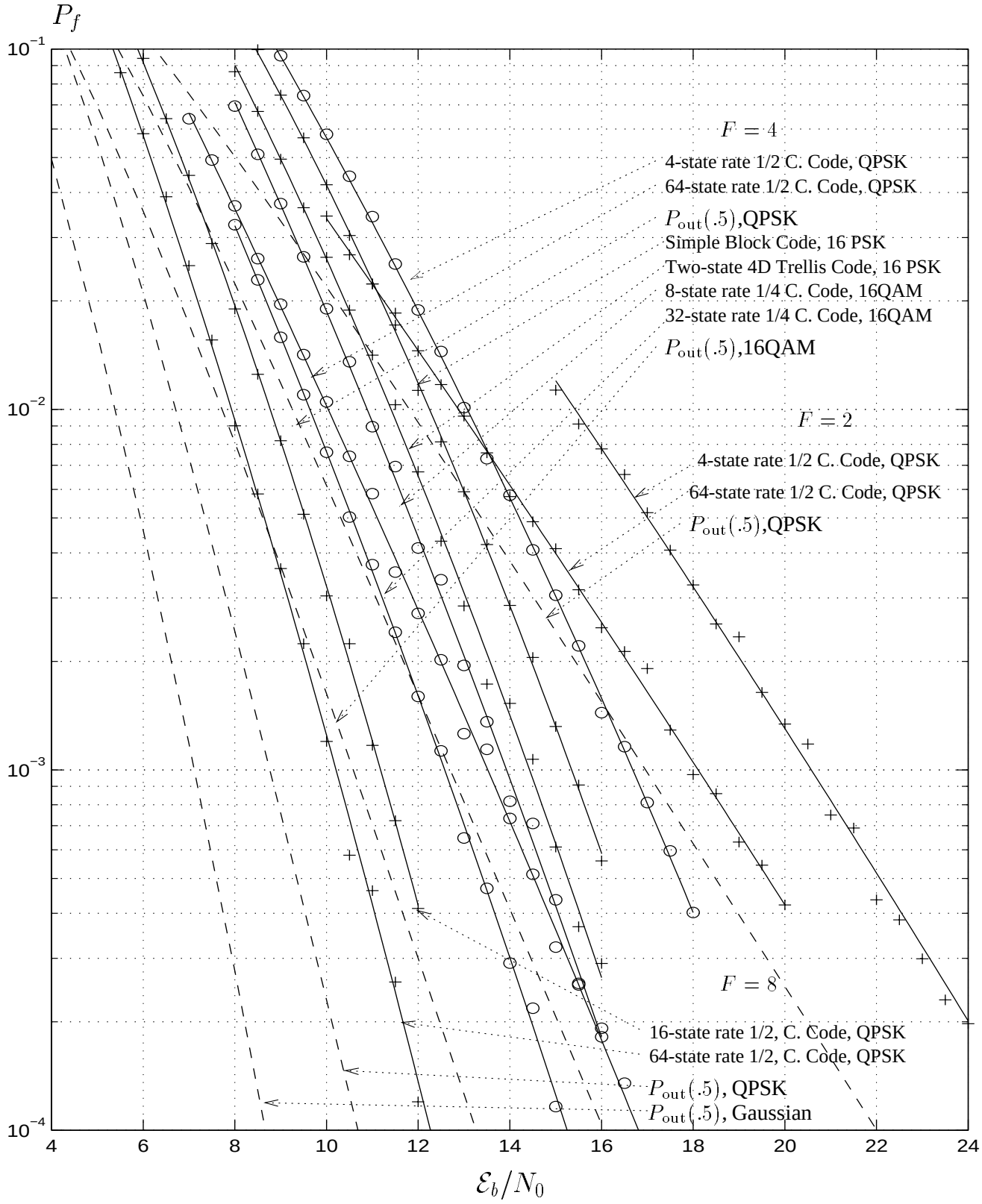

Fig. 13: Frame Error Rates for Codes with 1.0 bit/2 dim

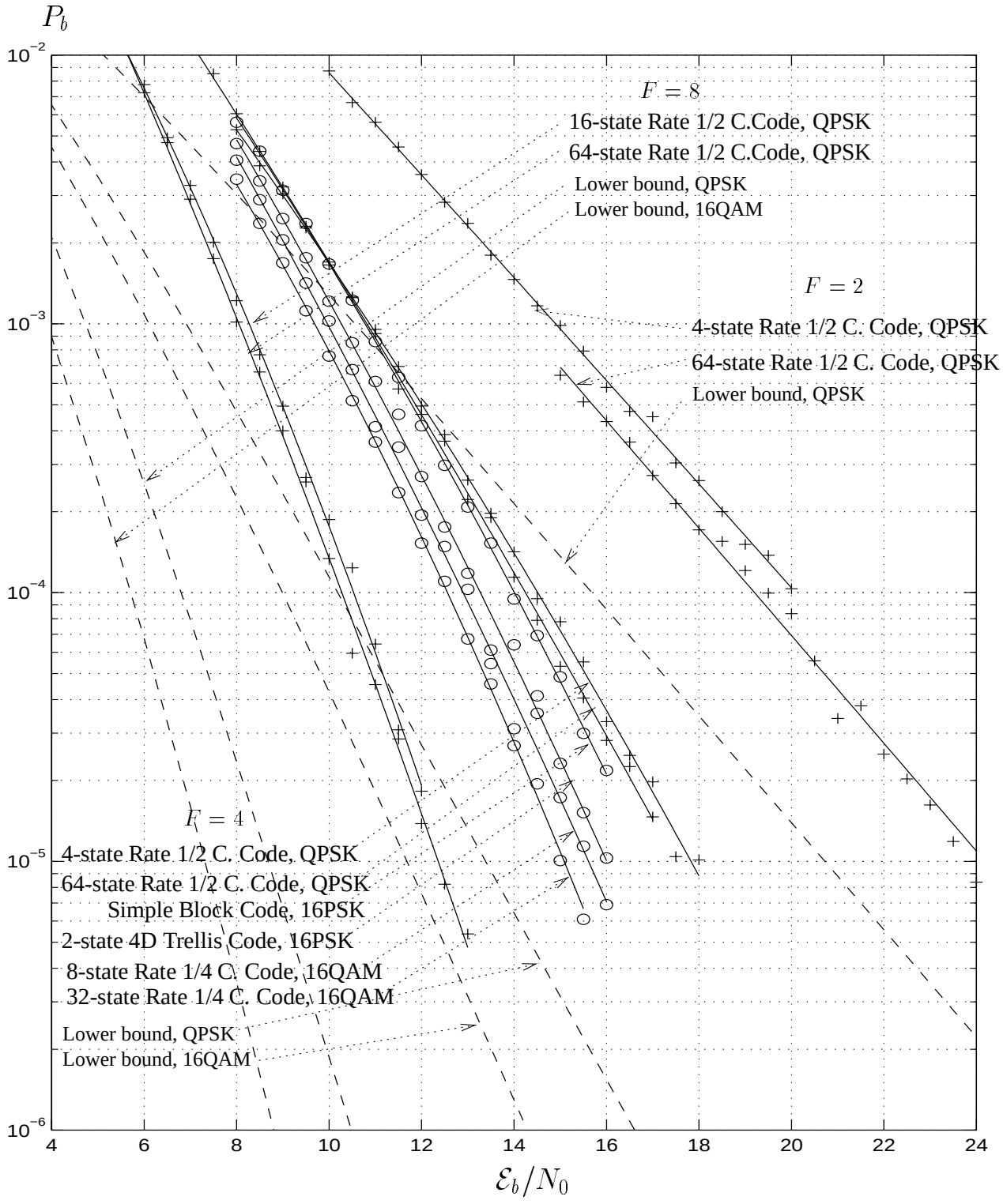

Fig. 14: Bit Error Rates for Codes with 1.0 bit/2 dim

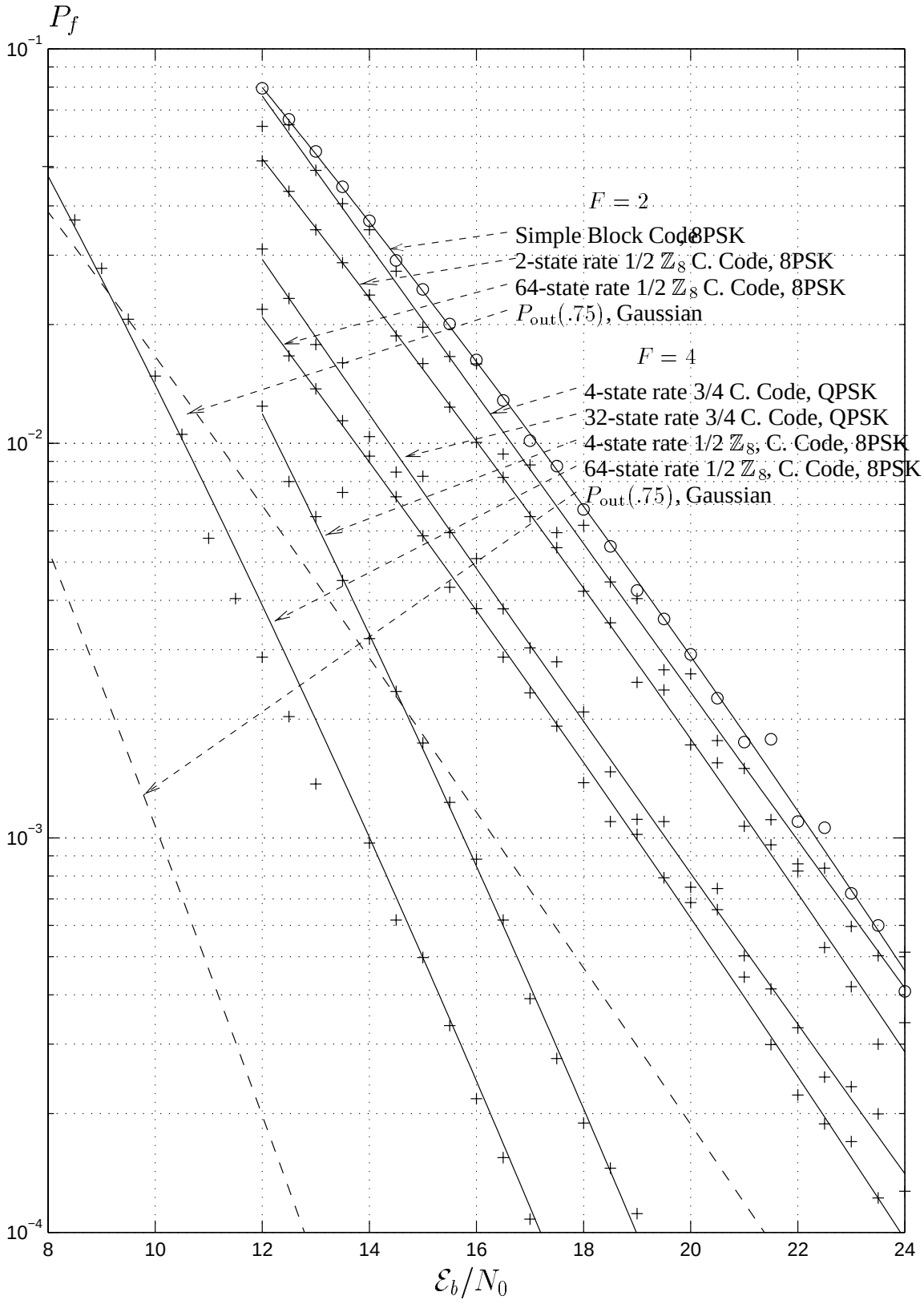

Fig. 15: Frame Error Rates for Codes with 1.5 bit/2 dim

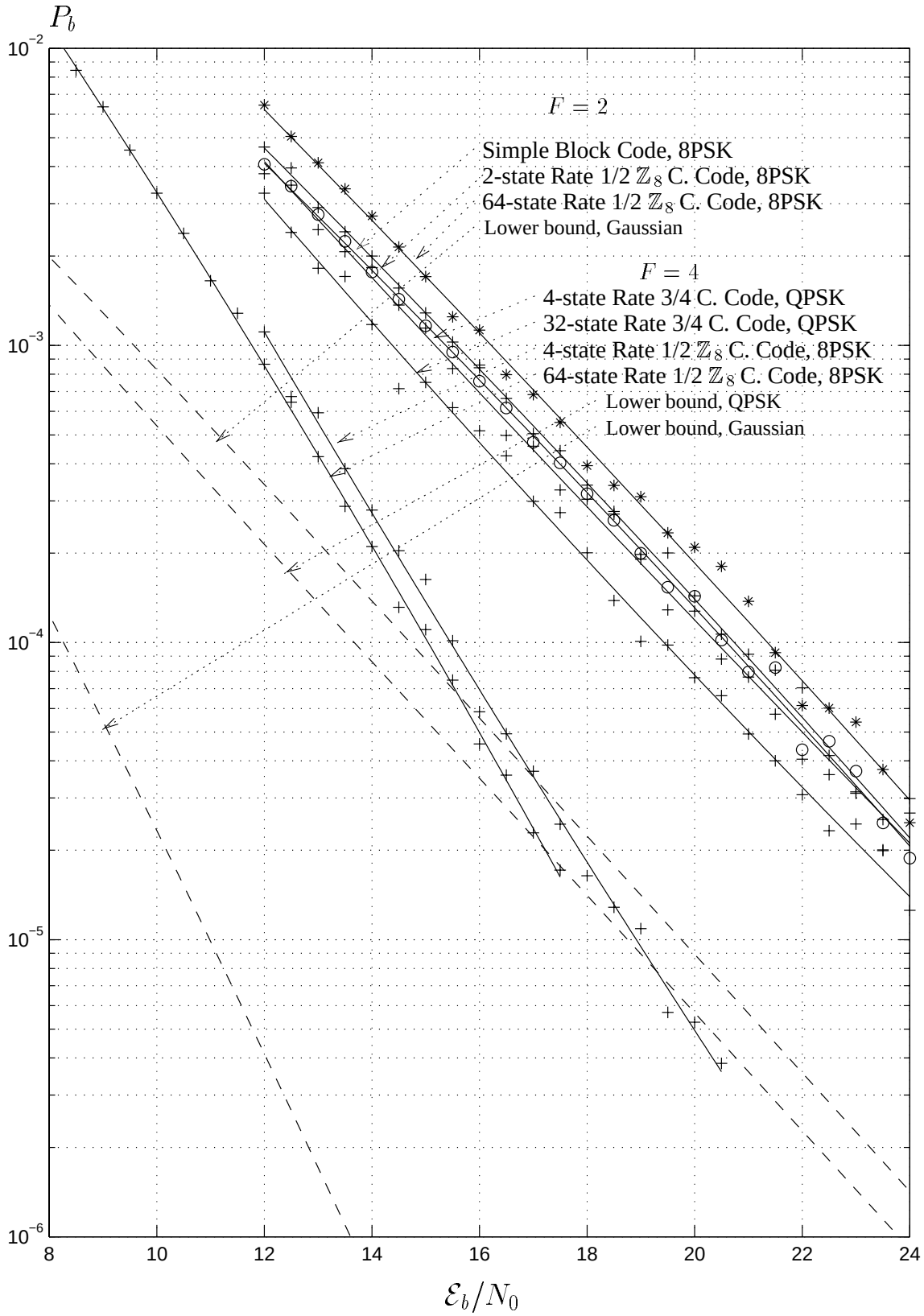

Fig. 16: Bit Error Rates for Codes with 1.5 bit/2 dim