8-2013

Improved iterative detection techniques for slow-frequency-hop communications with Reed-Solomon codes

Madhabi Manandhar
Clemson University, mmanand@g.clemson.edu

Improved iterative detection techniques for slow-frequency-hop communications with Reed-Solomon codes

---

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Electrical Engineering

---

by
Madhabi Manandhar
August 2013

Accepted by:
Dr. Daniel L. Noneaker, Committee Chair
Dr. Michael B. Pursley
Dr. Harlan B. Russell

Abstract

The performance of a packet-level iterative detection technique is examined for a slow-frequency-hop packet radio system using interleaved Reed-Solomon codes and per-dwell differential encoding. A per-dwell soft-input-soft-output detector along with successive-erasures decoding results in a system that performs better than previously considered detection techniques in the presence of partial-band interference. The log-MAP algorithm and two forms of its max-log-MAP approximation are considered for the soft-input-soft-output detector along with different channel estimators. The performance and detection complexity of the systems is compared. A limit on the number of erasures allowed in successive-erasures decoding is also considered, and its effect on the system's performance and detection complexity is examined.

Acknowledgments

I would like to thank my advisor Dr. Noneaker for supporting me and guiding me patiently throughout my thesis. His immense knowledge and helpful insights have helped me in every step of my graduate school life. I would also like to thank Dr. Pursley and Dr. Russell for being on my committee and taking the time to review my thesis and provide feedbacks.

I also thank my colleagues of the Wireless Communications group at Clemson for their help and support. And last but not the least I thank my mom and dad without whose guidance and encouragement this work would not have been possible.

Table of Contents

List of Figures

Chapter 1

Introduction

Packet radio communication systems often operate in environments in which portions of the available frequency band may be occupied by intentional or unintentional interference, and a receiver in the system often lacks *a priori* knowledge of the portion of the frequency band in which interference is present or the strength of the interference. Slow-frequency-hop (SFH) spread spectrum provides good protection against such partial-band interference when used with proper channel coding and decoding. Reed-Solomon (R-S) coding with errors-and-erasures (EE) decoding is particularly well matched to counter partial-band interference in a SFH system [1]. As an example, the most widely deployed SFH tactical military packet radio system incorporates R-S coding with multiple code words per packet [2]. It uses a technique to erase unreliable code symbols at the receiver for use with one-shot EE decoding of each received word.

Concatenation of an inner R-S code with an outer channel code allows for effective exploitation of iterative decoding which can achieve much better performance than one-shot EE decoding. Iterative erasure insertion and decoding [3] results in better performance than the test-symbol method [4] in a SFH system with R-S coding. Iterative equalization-and-decoding significantly improves the performance of a SFH system with R-S coding if the system is subjected to multipath fading [5]. Furthermore, the use of a R-S outer code and a convolutional inner code with an erasure-insertion technique has been shown to

provide improved protection against partial-band interference [6]. An early form of iterative decoding [7] is shown to enhance the performance of a concatenated coding system with an outer R-S code.

Packet-level iterative detection techniques have been shown in [8] to provide better performance than one-shot decoding in the presences of partial-band interference for a SFH packet radio system with R-S coding and parity bits [4]. Differential encoding is exploited in packet-level iterative erasure insertion and bounded-distance EE decoding to obtain good performance with a modest increase in the average decoding time for a packet. Maximum-likelihood sequence detection [9] of the differentially encoded binary representation of each code symbol transmitted in a dwell interval is employed using the Viterbi algorithm [10], and parity bits [4] are used to determine erasures for bounded-distance EE decoding of each received word. The code words that have been detected through a given iteration are used as feedback to constrain the valid transitions in the two-state trellis of the Viterbi algorithm in each dwell interval for the next iteration. This step is referred to as *branch pruning* [8].

In this thesis, an alternative is considered for packet-level iterative detection in a SFH packet radio system with R-S encoding and differential encoding. Instead of inserting parity bits, soft measures of the reliability of the received symbols are utilized for erasure insertion prior to EE R-S decoding of each word. For this purpose, MLSE detection using the Viterbi algorithm is replaced by a soft-input, soft-output (SISO) detector using the BCJR (log-MAP) algorithm [11] or a variant of the algorithm. A branch-pruned version of the BCJR algorithm and its variants is used in analogy with iterations of the branch-pruned Viterbi algorithm used in [8].

Per-dwell SISO detection of the content of each dwell interval is considered in conjunction with bounded-distance successive-erasures (SE) decoding [12] (a form of generalized-minimum distance decoding [13]). The soft outputs from SISO detection are used to assign a reliability ranking to the hard code-symbol decisions for each received word; the ranking determines the order of erasure insertion in SE decoding. As with the previously investigated packet-level iterative detection method [8], successful decoding of some received words

in the received packet provide feedback which constrains the trellis for SISO detection in each dwell interval in the subsequent iterations.

In the investigation, we consider the log-MAP algorithm and two forms of its max-log-MAP approximation [14] as per-dwell SISO detectors. The performance of the system is considered with each SISO detector and compared with the performance of some previously considered SFH systems using R-S coding. Both the packet error probability and the computational complexity of packet detection are used as measures of performance.

We also consider several practical estimation algorithms for use with the SISO detectors and compare their effectiveness. Finally, we consider a limit on the number of code-symbol erasures in a received word for SE decoding, and we examine the effect that the choice of the limit has on the system performance.

The remainder of the thesis is organized as follows. Chapter 2 contains a description of the SFH system, including each receiver considered in the thesis. The SISO differential detectors are discussed in Chapter 3. The chapter also includes a comparison of the system performance and the detection complexity of the systems using each detection technique. In Chapter 4, various practical channel estimators are discussed, and differences in the system performance due to the use of the estimators are analyzed. Chapter 5 details how system performance and decoding complexities are affected by the changes in the maximum number of erasures allowed in SE decoding. Finally, a summary of the results is presented in Chapter 6.

Chapter 2

System Description

2.1 Transmitter

The block diagram of the transmitter for the SFH system is shown in Fig. 2.1. A singly extended (n, k) R-S code is used to encode the information message. (The code redundancy $n - k$ is even in all the examples in the thesis.) The code symbols of each R-S code word are elements of GF(n), $n = 2^m$.

Each packet of data contains N_s R-S code words $\left(s_0^{(p)} \dots s_{n-1}^{(p)}\right)$, $0 \le p \le N_s - 1$. The rectangular interleaver interleaves the N_s code words in an $n \times N_s$ matrix such that each row contains one code symbol from each code word in the packet. Each code symbol $s_i^{(p)}$ is expanded into an m-bit binary (*code bit*) representation $\left(b_{i,0}^{(p)} \dots b_{i,m-1}^{(p)}\right)$ by a code-symbol-to-bit mapper. The code-bit sequence forming the i^{th} row of the resulting $n \times (mN_s)$matrix is denoted $\left(\tilde{b}_0^{(i)} \dots \tilde{b}_{mN_s-1}^{(i)}\right)$, $0 \le i \le n - 1$.

A pseudo-random interleaver reorders the code-bit sequence in each row using a different interleaving pattern for each row. If $\pi_i(\cdot)$ denotes the interleaver for row i, the interleaved code-bit sequence is given by

$$\left(\overline{b}_0^{(i)} \dots \overline{b}_{mN_s-1}^{(i)}\right) = \left(\tilde{b}_{\pi_i(0)}^{(i)} \dots \tilde{b}_{\pi_i(mN_s-1)}^{(i)}\right).$$

The contents of each row are then passed through a differential encoder which is initialized to the zero state prior to encoding each row; the resulting contents of the i^{th} row are given by

$$\left(d_0^{(i)} \dots d_{mN_s-1}^{(i)}\right)$$

where $d_0^{(i)} = \overline{b}_0^{(i)}$, $d_j^{(i)} = d_{j-1}^{(i)} \oplus \overline{b}_j^{(i)}$, $1 \leq j \leq mN_s - 1$. The differentially encoded code bits of each row are then transmitted as a single dwell interval using BPSK modulation in a frequency slot with a pseudo-randomly selected carrier frequency.

In practice, the transmission in each dwell interval of a SFH system includes a preamble for symbol synchronization, and the preamble can be used to derive the reference phase for coherent demodulation of the received symbols in the dwell interval. In the thesis, we do not model the presence of the per-dwell preamble in the received signal but we assume that the correct reference phase is available at the receiver for coherent demodulation in each dwell interval. The baseband-equivalent signal transmitted at time $t = 0$ is thus given by

$$s(t) = \mathrm{Re}\left(\sqrt{P}\sum_{i=0}^{n-1}\sum_{j=0}^{mN_s-1}(-1)^{d_j^{(i)}}p_T(t-(i(mN_S)+j)T)e^{j2\pi f_i t}\right)$$

where T is the channel symbol duration, p_T is the unit amplitude pulse over $[0, T]$ and P is the power of the transmitted signal.

The center frequencies of the frequency slots used for the n dwell intervals are $\{f_0 = f_c + k_0\Delta f, \dots f_{n-1} = f_c + k_{n-1}\Delta f\}$ where f_c is the center frequency of the lowest-frequency slot, Δf is the offset between the center frequencies of adjacent frequency slots , $k_i \in \{0, \dots, S-1\}$ and S is the number of frequency slots available. The offset frequency $\Delta f = \frac{2}{T}$; thus, $B_t = \frac{2S}{T}$ is the total system bandwidth.

2.2 Channel

The channel considered is a static, single-path channel that is subjected to both full-band additive white Gaussian noise and partial-band additive Gaussian interference.

The double-sided power spectral density (p.s.d.) of the noise is $\frac{N_0}{2}$. The interference occupies a fraction ρ of the total frequency band, and within that portion of the band its p.s.d. is constant and equal to $\frac{N_I}{2\rho}$. Consequently, there is a probability ρ that a randomly selected frequency slot is subjected to the interference; in each such slot, the received signal encounters a total noise-plus-interference p.s.d. of $\frac{N_0}{2} + \frac{N_I}{2\rho}$. With a probability $1 - \rho$, the received signal encounters only a noise p.s.d. of $\frac{N_0}{2}$.

The (instantaneous) *signal-to-interference-plus noise ratio* (SINR) at the receiver during a dwell interval depends on the frequency slot that is used. It is given by

$$\text{SINR} = \begin{cases} \frac{E_b}{N_0}, & \text{if there is no interference in the frequency slot} \\ \frac{E_b}{N_0 + (N_I/\rho)}, & \text{if there is interference in the frequency slot} \end{cases} \tag{2.1}$$

where $E_b = (n/k)\,PT$ is the energy per bit of information. The *signal-to-noise ratio* (SNR) at the receiver is given by E_b/N_0, and the *signal-to-interference ratio* (SIR) (which does not depend on ρ) is given by E_b/N_I.

2.3 Receiver

The block diagram of the baseband-equivalent receiver is shown in Fig. 2.2 for each of the new systems considered in the thesis. Only some of the systems include the channel estimator (and its connections shown as dashed lines), whereas the remaining subsystems shown in the figure are included in each system. (The lines shown in bold represent feedback paths for iterative decoding in the receiver.)

The received signal in each dwell interval is passed through a coherent, matched-filter receiver that samples the output at the optimal instant for each received binary channel symbol. It is assumed that an automatic gain-control subsystem [15] normalizes the samples with respect to the received energy per channel symbol so that the j^{th} received symbol at

the demodulator output in dwell interval i is given by

$$r_j^{(i)} = (-1)^{d_j^{(i)}} + n_j^{(i)}.$$

The noise random variables $\left\{n_j^{(i)}\right\}$, $0 \le i \le n-1$, $0 \le j \le mN_s-1$, are conditionally mutually independent given the subset of dwell intervals that are subjected to interference. Under this condition, the zero-mean Gaussian random variable $n_j^{(i)}$ has a variance that depends on the frequency slot used in the i^{th} dwell interval. It is given by

$$\sigma_i^2 = \begin{cases} \frac{N_0}{2E_b(k/n)}, & \text{if there is no interference in the frequency slot} \\ \frac{N_0+(N_I/\rho)}{2E_b(k/n)}, & \text{if there is interference in the frequency slot.} \end{cases} \tag{2.2}$$

Consequently, the channel during the i^{th} dwell interval is characterized by a single parameter, σ_i^2, that is unknown *a priori* at the receiver. In the first iteration of differential detection and R-S decoding, the mN_s real-valued demodulator outputs for each dwell interval are provided as the input to a SISO code-bit detector.

The SISO algorithm is executed on the two-state trellis of the differential encoder. For dwell interval i, the state of the trellis at time j is equal to $d_{j-1}^{(i)}$ if $j \ge 1$ and it is equal to zero if $j = 0$. Each branch of the trellis connecting a state at time $j-1$ with a state at time j is labeled by both the code-bit polarity $\overline{b}_j^{(i)}$ that generates the corresponding state transition and the resulting channel symbol $(-1)^{d_j^{(i)}}$.

The SISO detector provides both a hard-decision output for each differentially encoded code bit in the dwell interval and a measure of the reliability of each bit decision (the *bit reliability*). The hard decisions on the m code bits corresponding to a code symbol yield the hard decision on the code symbol, and the m bit reliabilities are used to determine the reliability of the code-symbol decision (the *symbol reliability*). If $(r_0, r_1, \ldots r_{m-1})$ are the

bit reliabilities of the code bits forming a code symbol, the symbol reliability R is given by

$$R = \prod_{i=0}^{m-1} \frac{e^{r_i}}{1 + e^{r_i}}. \tag{2.3}$$

Each individual product term of equation (2.3) is the probability that the corresponding code bit is detected correctly if MAP detection with *a priori* knowledge of the noise variance is used in the SISO detector. The symbol reliability R is then the probability that the code symbol is detected correctly.

The hard code-symbol decisions and their reliabilities are employed in successive-erasures bounded-distance decoding of the received word for each R-S code word in the packet. A predetermined maximum allowable number of erasures is used in the first decoding attempt by erasing the least-reliable code symbols. The number of erasures is reduced by two in each subsequent decoding attempt until either successful decoding to a valid (correct or incorrect) code word occurs or a decoding failure is declared for the decoding attempt with zero erasures (which corresponds to errors-only decoding).

Bounded-distance R-S decoding often produces successful decoding to an erroneous code word (i.e., an *undetectable decoder error* occurs) if the number of erasures is close to $n - k$ [16]; thus, a maximum allowable number of erasures much less than $n - k$ is used. The packet-level iterative detector is terminated (and packet detection fails) if none of the received words is successfully decoded in the iteration. If all the received words are successfully decoded, packet detection is successful and the iterative decoder is terminated.

If, instead, one or more code words are successfully decoded in the iteration using successive-erasures decoding but others are not successfully decoded, another packet-level iteration of differential detection and R-S decoding occurs. In the new packet-level iteration, all code bits representing code symbols from received words successfully decoded in previous iterations are presumed to be known *a priori*; thus, the two-state trellis for each dwell interval is pruned to include only branches consistent with the known code-bit polarities, and the SISO algorithm is restricted to the remaining trellis branches when it is executed for

that dwell interval in the current iteration. This is illustrated in Fig. 2.3, which shows the pruned trellis for dwell intervals v and w in an iteration occurring later than the iteration in which received word 0 was successfully decoded. In the illustration, $n = 16$ and $N_s = 2$; the code bits of the i^{th} code symbol of the detected code word are denoted $\left(\hat{b}_{i,0}^{(0)} \ldots \hat{b}_{i,3}^{(0)}\right)$.

The output of the SISO algorithm for each dwell interval in the current packet-level iteration is used for successive-erasures decoding of the received words that have not yet been decoded successfully. Packet-level iterations continue until either all received words are decoded successfully or an iteration occurs in which no additional received words are decoded successfully.

Six variants are considered for the receiver using packet-level iterative detection with per-dwell SISO detection and successive-erasures R-S decoding. They are denoted by systems A-F, and they differ only in the SISO algorithm that is used in each iteration. The six systems are described below along with two previously introduced systems. (The performance of either of the previously introduced systems provides a benchmark against which the performance of each new system is compared.)

2.3.1 System A

System A uses the BCJR algorithm for log-MAP detection of the (differentially encoded) code bits in each dwell interval. The algorithm requires an estimate of the channel parameter σ_i^2 given in equation (2.2) for each dwell interval. The estimate of σ_i^2, denoted $\hat{\sigma}_i^2$, is referred to as the *channel estimate* for the i^{th} dwell interval.

We assume that system A obtains a perfect channel estimate for each dwell interval; that is, $\hat{\sigma}_i^2 = \sigma_i^2$ for $0 \leq i \leq n-1$. The bit reliabilities obtained from the BCJR algorithm are the log-likelihood ratios (LLRs) of the code bits; thus, the symbol reliability calculated by equation (2.3) is the probability that the code-symbol decision is correct at the output of the bit-to-code-symbol mapper. This is true for the first iteration of packet-level detection; it is also true for subsequent iterations under the condition that each successfully decoded received word is decoded correctly.

2.3.2 System B

In system B, the max-log-MAP algorithm is used instead of the log-MAP algorithm for the detection of the code bits and the determination of the bit reliabilities. The bit reliabilities are not LLRs in this instance, but they preserve the rank ordering of the code-bit LLRs. (The resulting symbol reliabilities do not necessarily preserve the rank ordering of code-symbol LLRs, however.) The correlator form of the max-log-MAP algorithm is considered in this system; it does not require channel estimates.

2.3.3 System C

System C is a modification of system B in which the outputs of the max-log-MAP algorithm for each dwell interval are weighted in inverse proportion to the channel estimate for the dwell interval. A perfect estimate is assumed for each dwell interval.

2.3.4 System D

System D operates in the same manner as system C, but it uses a practical channel estimator [17] based on the estimated moments of the demodulator outputs in the dwell interval. The channel estimate $\hat{\sigma_i}^2$ for the i^{th} dwell interval is obtained once per packet, and it is used for all iterations of the decoder. The channel estimator is described in Section 5.1.

2.3.5 System E

System E differs from System D in that it updates the channel estimate for each dwell interval in each packet-level iteration, using the code-bit polarities for received words that were successfully decoded in previous iterations. Conditioning on the hard code-bit decisions that are used as feedback alters the joint distribution of the remaining received symbols. A large-variance approximation to the maximum-likelihood estimate for the noise variance in the i^{th} dwell interval is used, as described in Section 5.2.

2.3.6 System F

System F is a hybrid of system D and system B. The decoder of system D is employed first. If the packet-level iterative detection algorithm is terminated with one or more received words that are not decoded successfully, packet-level iterative detection using System B is attempted with the successful decoding results from system D used as initial feedback to the SISO detector for each dwell interval.

2.3.7 System EO

Reference system EO uses the same transmitter as in Fig. 2.1, except that the code bits are not differentially encoded prior to transmission. The non-iterative receiver uses hard-decision detection of code bits, mapping of detected code bits to 2^m-ary code symbols, and one-shot errors-only decoding of the N_s received words in the received packet [6].

2.3.8 System PB

Reference system PB uses the same transmitter as in Fig. 2.1, except that the mapping of each 2^m-ary code-symbol to a binary representation includes the addition of an even-parity bit to the m code bits prior to bit interleaving [8]. The packet-level iterative receiver uses the Viterbi algorithm (instead of a SISO algorithm) with the (pruned) two-state trellis for each dwell interval. The Viterbi algorithm results in a hard decision for each differentially encoded code-bit and parity bit. The $m + 1$ bits corresponding to each code-symbol are then mapped to either a detected code-symbol (if even parity is satisfied) or an erasure symbol (if parity is not satisfied). The detected and erased symbols are used for a single instance of errors-and-erasures decoding in the packet-level iteration for each previously undecoded received word.

2.4 Measures of system performance

Each packet-detection attempt results in one of three outcomes: correct detection of the packet, a packet-detection error, or packet-detection failure. Correct detection of the packet occurs if all N_s received words in the received packet are decoded correctly at termination. A packet-detection error occurs if all received words are successfully decoded at termination, but one or more of them is decoded to an incorrect code word. Packet-detection failure occurs if one or more received words has not been successfully decoded at termination.

A packet-detection failure is known to the receiver. Furthermore, the use of an outer error-detection code and decoder allows the receiver to identify most occurrences of a packet-detection error. While we do not model the presence of an outer error-detection code and decoder in the systems considered in the thesis, we approximate the effect of their presence by classifying both a packet-detection error and a packet-detection failure as a *packet error* and using the probability of packet error as a measure of each system's performance. Specifically, the system performance is characterized by the SIR required to obtain a desired probability of packet error for a specified fractional interference bandwidth ρ and a specified SNR. A lower required SIR corresponds to better system performance.

The second measure of system performance is the computational cost or delay at the receiver to decode a packet. The variable computational burden of packet detection includes the work required for each decoding attempt for a received word and each instance of trellis-based code-bit detection for a dwell interval as well as each instance of channel estimation in the systems using it. The SISO algorithms operating on the two-state trellis used in each system impose only a small computational burden, however. Furthermore, we assume that the computation required for channel estimation is small in comparison with the computation required for R-S decoding. Consequently, we approximate the total detection complexity as the complexity of decoding.

One approach to SE decoding employs an underlying EE decoder that is used once

for each set of erasures applied to a received word. We assume that the computation required for EE decoding is similar regardless of the number of code-symbol erasures. Consequently, we approximate the total detection complexity in systems A-F using this SE decoder architecture as the number of EE decoding attempts per transmitted R-S code word. (The same measure of detection complexity is used for system EO and PB; the resulting detection complexity for system EO is thus constant and equal to one decoding attempt per transmitted R-S code word.)

Another approach to SE decoding employs an underlying decoder that shares much of the computation among all of the EE decoding attempts for a given received word [18], [19], [20]. We approximate the total detection complexity using this SE decoder architecture as the number of SE decoding attempts per transmitted R-S code word (regardless of how many EE decoding attempts any SE decoding attempt includes).

2.5 Examples used in thesis

All of the numerical results in the thesis are for packets of twelve code words from a $(32, 12)$ singly extended R-S encoder. Each packet is thus transmitted as 32 dwell intervals. Each dwell interval in each system includes 60 channel symbols representing twelve $32-$ary code symbols. System PB also includes a channel symbol for each of the twelve parity bits per dwell interval. The signal-to-noise ratio in each example is 20 dB.

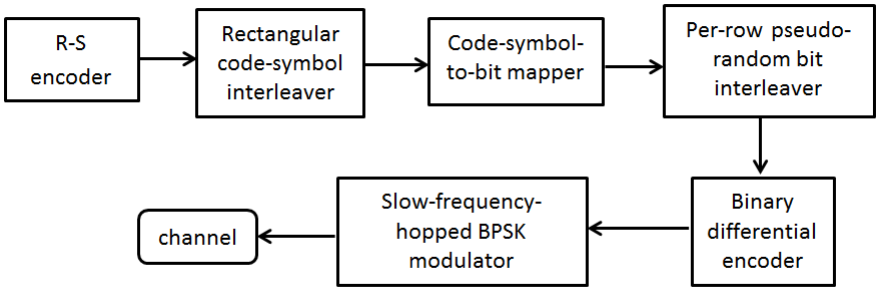

Figure 2.1: Transmitter for SFH system.

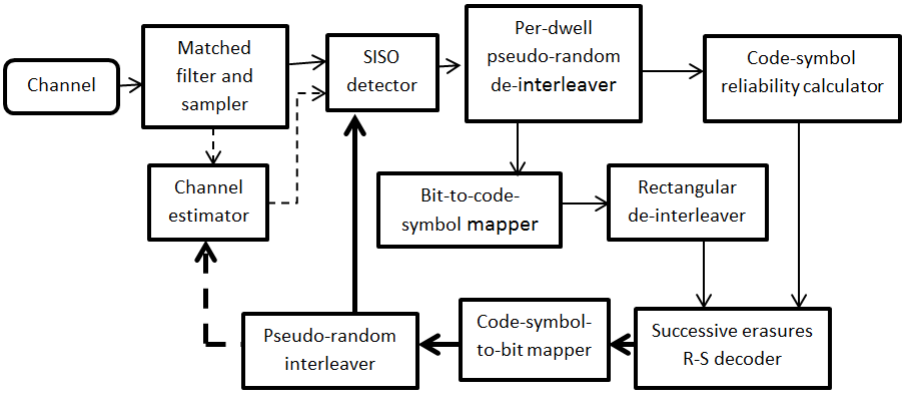

Figure 2.2: Baseband-equivalent receiver for SFH system.

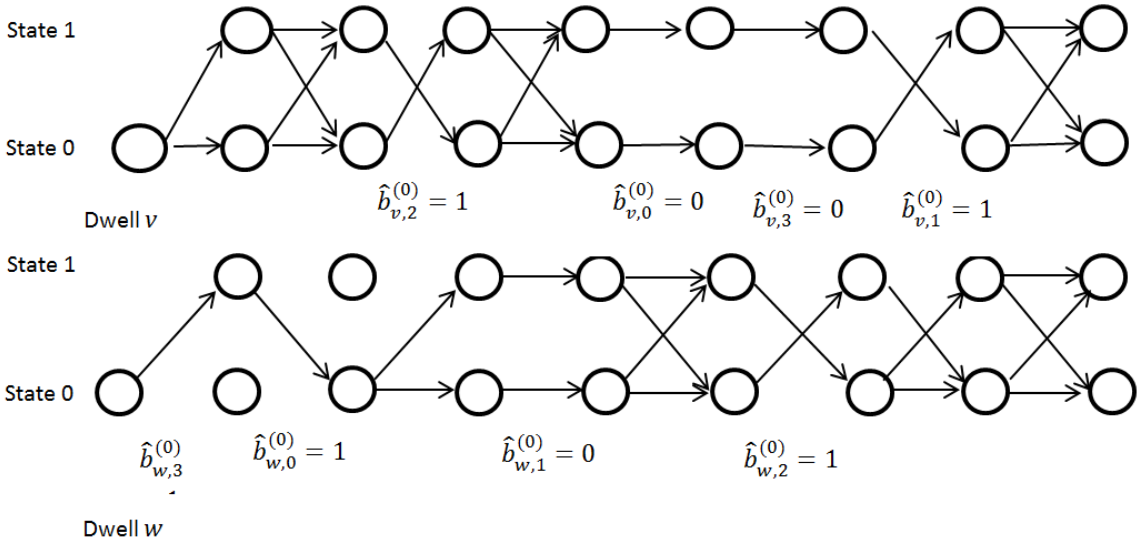

Figure 2.3: Examples of the branch-pruned trellis for two typical dwell intervals.

Chapter 3

Performance Using Different SISO Algorithms

3.1 MAP detection of code bits

The SISO detector uses the MAP algorithm for differential detection. The MAP algorithm decodes a code word by maximizing the *a posteriori* probability that an information bit is correctly decoded. In other words, the MAP decision for input information bit $b_j^{(i)}$ is the value $\tilde{b} \in \{0,1\}$ that maximizes $\Pr\left(b_j^{(i)} = \tilde{b} | \underline{r}_{0,mN_s-1}^{(i)}\right)$, where $\underline{r}_{0,mN_s-1}^{(i)} = \left(r_0^{(i)} \dots r_{mN_s-1}^{(i)}\right)$ is the received channel-symbol sequence in the i^{th} dwell interval. To obtain the MAP decision of a bit $b_j^{(i)}$, its log *a posteriori* probability (log APP) ratio is calculated as

$$L\left(b_j^{(i)}\right) = \log\left(\frac{\Pr\left(b_j^{(i)} = 0 | \underline{r}_{0,mN_s-1}^{(i)}\right)}{\Pr\left(b_j^{(i)} = 1 | \underline{r}_{0,mN_s-1}^{(i)}\right)}\right),$$

and the MAP decision for b_{k-1} is

$$\hat{b}_j^{(i)} = \begin{cases} 0, & L\left(b_j^{(i)}\right) > 0 \\ 1, & \text{otherwise.} \end{cases}$$

The differential encoder can be viewed as a rate-one convolutional encoder. An i.i.d. information source at the transmitter and the linearity of the R-S encoder results in a sequence of i.i.d. code bits within any single dwell interval. The sequence of code bits thus represents a special case of a Markov source; consequently, MAP detection of each code bit based on the sequence of channel outputs in the dwell interval (without consideration of the channel outputs for the other dwell intervals) is achieved using the BCJR algorithm [11]. This is expressed as

$$\Pr\left(b_j^{(i)}=0|\underline{r}_{0,mN_s-1}^{(i)}\right)=\sum_{(m',m)\in\tilde{A}}\frac{\Pr\left(S_{j-1}=m',S_j=m\right)f\left(\underline{r}_{0,mN_s-1}^{(i)}|S_{j-1}=m',S_j=m\right)}{f\left(\underline{r}_{0,mN_s-1}^{(i)}\right)},$$

$$\Pr\left(b_j^{(i)}=1|\underline{r}_{0,mN_s-1}^{(i)}\right)=\sum_{(m',m)\in\tilde{A}^c}\frac{\Pr\left(S_{j-1}=m',S_j=m\right)f\left(\underline{r}_{0,mN_s-1}^{(i)}|S_{j-1}=m',S_j=m\right)}{f\left(\underline{r}_{0,mN_s-1}^{(i)}\right)}.$$

The set $\tilde{A}=\left\{(m',m):b_{j-1}^{(i)}=0\right\}$ corresponds to all state transitions resulting from a channel input equal to zero, and $\tilde{A}^c=\left\{(m',m):b_{j-1}^{(i)}=1\right\}$ corresponds to all state transitions resulting from a channel input equal to one.

Since the source is Markov and the channel is memoryless,

$$\begin{aligned}
&\Pr\left(S_{j-1}=m',S_j=m\right)f\left(\underline{r}_{0,mN_s-1}^{(i)}|S_{j-1}=m',S_j=m\right)\\
&=\left[\Pr(S_{j-1}=m')f\left(\underline{r}_{0,j-2}^{(i)}|S_{j-1}=m'\right)\right]\\
&\quad\times\left[\Pr(S_j=m|S_{j-1=m'})f\left(r_{j-1}^{(i)}|S_{j-1}=m',S_j=m\right)\right]f\left(\underline{r}_{j,mN_s-1}^{(i)}|S_j=m\right)\\
&=\alpha_{j-1}\left(m'\right)\gamma_j\left(m',m\right)\beta_j(m).
\end{aligned}$$

For an AWGN channel with variance σ_i^2 and equally likely input bits, if the state transition from m' to m results in output $d(m',m)$,

$$\gamma_j\left(m',m\right)=\begin{cases}\frac{1}{2}\frac{1}{\sqrt{2\pi\sigma_i^2}}e^{-\frac{\left(r_{j-1}^{(i)}-(-1)^{d(m',m)}\right)^2}{2\sigma_i^2}}, & \Pr\left(S_j=m|S_{j-1}=m'\right)\neq 0\\ 0, & \text{otherwise,}\end{cases}$$

$$\begin{aligned}\alpha_j(m) &= \sum_{m'} \left[\Pr(S_{j-1}=m') f\left(\underline{r}_{0,j-2}^{(i)} | S_{j-1}=m'\right)\right] \\ &\quad \left[\Pr(S_j=m|S_{j-1}=m') f\left(r_{j-1}^{(i)} | S_{j-1}=m', S_j=m\right)\right] \\ &= \sum_{m'} \alpha_{j-1}\left(m'\right) \gamma_j\left(m', m\right),\end{aligned}$$

and

$$\begin{aligned}\beta_j(m) &= \sum_{m'} f\left(\underline{r}_{j+1,mN_s-1}^{(i)} | S_{j+1=m'}\right) \\ &\quad \left[\Pr(S_{j+1}=m'|S_j=m) f\left(r_j^{(i)} | S_j=m, S_{j+1}=m'\right)\right] \\ &= \sum_{m'} \beta_{j+1}\left(m'\right) \gamma_{j+1}\left(m, m'\right).\end{aligned}$$

The BCJR algorithm is used to efficiently determine the values of each auxiliary function $\gamma_j(m)$, $\alpha_j(m)$, and $\beta_j(m)$ for a given received channel-symbol sequence. Since the differential encoder begins at state 0 and terminates at either of the two states with equal probability, the initial values of $\alpha_j(m)$ and $\beta_j(m)$ are given as

$$\begin{aligned}\alpha_0(m) &= \begin{cases} 1, & m=0 \\ 0, & \text{otherwise} \end{cases} \\ \beta_{mN_s}(m) &= \frac{1}{2}\end{aligned}$$

Thus the log APP ratio is given by

$$L\left(b_j^{(i)}\right) = \log\left(\frac{\sum_{(m',m)\in\tilde{A}} \alpha_{j-1}(m')\gamma_j(m',m)\beta_j(m)}{\sum_{(m',m)\in\tilde{A}^c} \alpha_{j-1}(m')\gamma_j(m',m)\beta_j(m)}\right).$$

If a subset of the code bits in the dwell interval are known *a priori* (as occurs with feedback of successfully detected code words in packet-level iterative decoding), the same argument holds with the channel-symbol trellis for the dwell interval constrained to reflect known code-bit polarities.

3.1.1 Log-MAP form of the MAP algorithm

The large number of multiplications in the BCJR algorithm for MAP detection may cause numerical instability particularly when the block length is very large. A more stable form of the algorithm is obtained by converting all the calculations to the log domain; the resulting algorithm is referred to as the *log-MAP algorithm* [14].

In the log domain,

$$\begin{aligned}
\hat{\gamma}_j(m',m) &= \log\left(\gamma_j(m',m)\right) \\
&= \begin{cases} \log\left(\frac{1}{2\sqrt{2\pi\sigma_i^2}}\right) - \frac{\left(r_{j-1}^{(i)} - (-1)^{d(m',m)}\right)^2}{2\sigma_i^2}, & \Pr\left(S_j = m | S_{j-1} = m'\right) \neq 0 \\ 0, & \text{otherwise,} \end{cases} \\
\hat{\alpha}_j(m) &= \log\left(\alpha_j(m)\right) = \log\left(\sum_{m'} e^{\left(\hat{\alpha}_{j-1}(m) + \hat{\gamma}_j(m',m)\right)}\right), \\
\hat{\beta}_j(m) &= \log\left(\beta_j(m)\right) = \log\left(\sum_{m} e^{\left(\hat{\gamma}_{j+1}(m,m') + \hat{\beta}_{j+1}(m')\right)}\right),
\end{aligned}$$

$$\begin{aligned}
\alpha_0(m) &= \begin{cases} 0, & m = 0 \\ -\infty, & \text{otherwise,} \end{cases} \\
\beta_{mN_s}(m) &= \log\left(\frac{1}{2}\right)
\end{aligned}$$

and the log APP ratio is given as

$$\begin{aligned}
L\left(b_j^{(i)}\right) &= \log\left(\sum_{(m',m)\in\tilde{A}} e^{\left(\hat{\alpha}_{j-1}(m) + \hat{\gamma}_j(m',m) + \hat{\beta}_j(m)\right)}\right) \\
&\quad - \log\left(\sum_{(m',m)\in\tilde{A}^c} e^{\left(\hat{\alpha}_{j-1}(m) + \hat{\gamma}_j(m',m) + \hat{\beta}_j(m)\right)}\right).
\end{aligned}$$

The unstable multiplications are replaced by additions, and calculations of the form $\log(e^x + e^y)$ are simplified by using $\log(e^x + e^y) = \max(x, y) + \log\left(1 + e^{-|x-y|}\right)$ [14]. Outputs of the log-MAP algorithm are exactly the same as BCJR algorithm because only the method of computation is changed to simplify calculations.

3.2 Max-log-MAP detection of code bits

The calculations of the form $\log(e^x + e^y)$ in the log-MAP algorithm can be further simplified by using the approximation $\log(e^x + e^y) \approx \max(x, y)$ as the term $\log\left(1 + e^{-|x-y|}\right)$ ranges between 0 and 0.7 . The approximation results in the *max-log-MAP* approximation [21] to the log-MAP algorithm, which is implemented as in the BCJR algorithm, but with the modified auxiliary functions

$$\begin{aligned} \overline{\alpha}_0(m) &= \begin{cases} 0, & m = 0 \\ -\infty, & m \neq 0 \end{cases} \\ \overline{\beta}_{mN_s} &= \log\left(\frac{1}{2}\right) \end{aligned}$$

$$\begin{aligned} \overline{\alpha}_j(m) &= \max_{m'}\left(\overline{\alpha}_{j-1}(m) + \hat{\gamma}_j(m, m)\right) \\ \overline{\beta}_j(m) &= \max_{m'}\left(\hat{\gamma}_{j+1}(m, m') + \overline{\beta}_{j+1}(m')\right) \end{aligned}$$

and the approximate log-APP ratio

$$\begin{aligned} \overline{L}\left(b_j^{(i)}\right) = &\max_{(m',m)\in\tilde{A}}\left(\overline{\alpha}_{j-1}(m') + \hat{\gamma}_j(m', m) + \overline{\beta}_j(m)\right) \\ &- \max_{(m',m)\in\tilde{A}^c}\left(\overline{\alpha}_{j-1}(m') + \hat{\gamma}_j(m', m) + \overline{\beta}_j(m)\right). \end{aligned}$$

The value of $\hat{\gamma}_j(m', m)$ in the *max-log-MAP algorithm* is the same as that in the log-MAP algorithm, but $\overline{\alpha}_j(m)$ and $\overline{\beta}_j(m)$ are different than $\hat{\alpha}_j(m)$ and $\hat{\beta}_j(m)$, respectively,

of the log-MAP algorithm. Hence, even though it is easy to compute, $\overline{L}\left(b_j^{(i)}\right)$ is only an approximation to $L\left(b_j^{(i)}\right)$. Iterative decoders in which the algorithm is repeated a number of times, the approximation can have an observable effect on the probability of error at the decoder's output.

3.2.1 Correlator form of the max-log-MAP approximation

For a Gaussian noise channel with variance σ_i^2, if $r_0^{(i)}, r_2^{(i)}, \ldots r_{mN_s-1}^{(i)}$ are mN_s received channel symbols, we have

$$\begin{aligned}\hat{\gamma}_j(m',m) &= \log\left(\frac{1}{2\sqrt{2\pi\sigma_i^2}}\right) - \frac{\left(r_{j-1}^{(i)} - (-1)^{d(m',m)}\right)^2}{2\sigma_i^2} \\ &= -\log 2\sqrt{2\pi\sigma^2} - \left(\frac{\left(r_{j-1}^{(i)}\right)^2 + 1 - 2r_{j-1}^{(i)}(-1)^{d(m',m)}}{2\sigma_i^2}\right).\end{aligned}$$

Let $\tilde{\gamma}_j(m',m) = \sigma_i^2\left(\log\left(2\sqrt{2\pi\sigma_i^2}\right) + C_j + \hat{\gamma}_j(m',m)\right) = r_{j-1}^{(i)}(-1)^{d(m',m)}$

where, $C_j = \frac{\left(r_{j-1}^{(i)}\right)^2+1}{2\sigma_i^2}$. Then,

$$\tilde{\alpha}_j(m) = \sigma_i^2\left(j\log\left(2\sqrt{2\pi\sigma_i^2}\right) + \sum_{k=1}^{j} C_k + \overline{\alpha}_j(m)\right)$$

$$\tilde{\beta}_j(m) = \sigma_i^2\left((mN_s - j)\log\left(2\sqrt{2\pi\sigma_i^2}\right) + \sum_{k=j+1}^{mN_s} C_k + \overline{\beta}_j(m)\right).$$

And the *a posteriori* probability is given by

$$\begin{aligned}
\tilde{L}\left(b_j^{(i)}\right) &= \max_{(m',m)\in\tilde{A}}\left(\tilde{\alpha}_{j-1}(m')+\tilde{\gamma}_j(m',m)+\tilde{\beta}_j(m)\right) \\
&\quad - \max_{(m',m)\in\tilde{A}^c}\left(\tilde{\alpha}_{j-1}(m')+\tilde{\gamma}_j(m',m)+\tilde{\beta}_j(m)\right) \\
&= \sigma^2\left[\max_{(m',m)\in\tilde{A}}\left(mN_s\log\left(2\sqrt{2\pi\sigma_i^2}\right)+\sum_{k=1}^{mN_s}C_k+\overline{\alpha}_{j-1}(m')+\hat{\gamma}_j(m',m)+\overline{\beta}_j(m)\right)\right. \\
&\quad \left. - \max_{(m',m)\in\tilde{A}^c}\left(mN_s\log\left(2\sqrt{2\pi\sigma_i^2}\right)+\sum_{k=1}^{mN_s}C_k+\overline{\alpha}_{j-1}(m')+\hat{\gamma}_j(m',m)+\overline{\beta}_j(m)\right)\right] \\
&= \sigma_i^2\overline{L}\left(b_j^{(i)}\right).
\end{aligned}$$

This form of the max-log-MAP approximation is known as the *correlator form* of the max-log-MAP algorithm [22]. Here, $\tilde{\gamma}_j(m',m)$ in every step is obtained by the multiplication of $r_{j-1}^{(i)}$ and $(-1)^{d(m',m)}$. The evaluation of $\tilde{\gamma}\,(m',m)$ does not require noise variance estimation, and neither does the evaluation of $\tilde{\alpha}_j(m)$ and $\tilde{\beta}_j(m)$. The value $\tilde{L}\left(b_j^{(i)}\right)$ computed in this manner is a scaled version of $\overline{L}\left(b_j^{(i)}\right)$. But the scaling of the log APP ratio does not affect its polarity; hence, the bit decisions using either $\tilde{L}\left(b_j^{(i)}\right)$ or $\overline{L}\left(b_j^{(i)}\right)$ give the same results. The correlator form of the max-log-MAP algorithm is relatively easy to implement compared to all the previously discussed algorithms. If needed, its outputs $\tilde{L}\left(b_j^{(i)}\right)$ can be scaled with the estimated channel variance to recover $\overline{L}\left(b_j^{(i)}\right)$.

3.3 Bit and code-symbol reliability

Along with code-bit decisions, log-MAP detection also provides a measure of the reliability of each of the decisions. If $L\left(b_j^{(i)}\right)$ is the log APP ratio of $b_j^{(i)}$, its log reliability $R\left(b_j^{(i)}\right)$ is given by

$$R\left(b_j^{(i)}\right) = \left|L\left(b_j^{(i)}\right)\right|.$$

The *a posteriori* probability that $b_j^{(i)}$ has been detected correctly is given by

$$\Pr\left(b_j^{(i)}\text{detected correctly}\right) = \frac{e^{R\left(b_j^{(i)}\right)}}{1+e^{R\left(b_j^{(i)}\right)}}.$$

Thus the larger $R\left(b_j^{(i)}\right)$, the greater the probability that $b_j^{(i)}$ has been detected correctly.

If $\left(b_0^{(i)}, b_1^{(i)}, \ldots b_{m-1}^{(i)}\right)$ are the m bits forming a code-symbol, the *a posteriori* probability that the code-symbol has been detected correctly is

$$\Pr(\text{code-symbol detected correctly}) = \prod_{k=0}^{m-1} \frac{e^{R\left(b_k^{(i)}\right)}}{1+e^{R\left(b_k^{(i)}\right)}}. \tag{3.1}$$

The probability that a code-symbol has been detected correctly is described as the reliability of the code-symbol decision. The code-symbol reliabilities are used in SE decoding to determine erasure insertions.

If the max-log-MAP algorithm or its correlator form is used instead of the log-MAP algorithm, the respective log APP ratios $\overline{L}\left(b_j^{(i)}\right)$ or $\tilde{L}\left(b_j^{(i)}\right)$ are produced. Even though $\overline{L}\left(b_j^{(i)}\right)$ and $\tilde{L}\left(b_j^{(i)}\right)$ are not equal to the log APP ratio $L\left(b_j^{(i)}\right)$ produced by the log-MAP algorithm, with an AWGN channel each results in the same rank ordering of the bit reliabilities. The approximations do not necessarily result in the same rank ordering of code-symbol reliability as log-MAP detection, however. If different channel symbols are subjected to differing noise variance (as in the channels with partial-band Gaussian interference), even the rank ordering of code-bit reliabilities may differ among the log-MAP algorithm and the two forms of the max-log-MAP algorithm.

3.4 Comparison of systems with various SISO algorithms

In this section, the performance of systems A, B and C is compared with the performance of reference systems EO and PB and among systems A, B and C. Systems A and C use noise-variance estimation which is assumed to provide perfect estimates for each dwell

interval. The maximum number of erasures used for SE decoding in systems A, B and C is 10.

For a given SNR and value of ρ, the system with the smaller SIR required to achieve a specified packet error probability is the better performing of two systems. The robustness of a system to interference in an unknown fraction of the band is characterized by $\mathrm{SIR}_{\max}$ (corresponding to ENR^* in [1]), which is the largest value of the required SIR for the system over all values of ρ (for the given SNR). A smaller value of $\mathrm{SIR}_{\max}$ represents greater system robustness, and it represents better system performance in the presence of a hostile interferer that is able to adapt the fraction of the band it jams to maximize its harm to the system.

Many unintentional sources of interference (and some intentional jammers) are likely to have a much narrower bandwidth than the SFH system. Consequently, another parameter of interest is ρ^* [1], which measures the system's ability to perform well in the presence of very strong narrow-band interference. It is the largest fraction of the band jammed below which the system achieves the target packet error probability irrespective of the value of interference power. (That is, even if the interference power is infinite, acceptable system performance is achieved as long as the fraction of the band jammed is less than ρ^*.) Hence, a larger value of ρ^* represents a greater ability of the system to mitigate the effect of strong, narrowband interference.

The SIR required to achieve a packet error probability of 10^{-2} is shown as a function of ρ in Fig. 3.1 for system A and the two reference systems. System EO does not employ differential encoding and uses one-shot errors-only decoding; it results in $\rho^* = 0.15$ and $\mathrm{SIR}_{\max} = 8.83$ dB. System PB includes per-dwell differential encoding, parity bits for use with erasure insertion and packet-level iterative decoding. System PB results in $\rho^* = 0.19$ and $\mathrm{SIR}_{\max} = 7.53$ dB. The use of log-MAP detection in each dwell interval and SE decoding results in much better performance for system A, with $\rho^* = 0.28$ and $\mathrm{SIR}_{\max} = 6.12$ dB.

Fig. 3.2 compares the performance of systems A, B and C using the same performance criterion as in Fig. 3.1. System B employs the correlator form of the max-log-MAP algorithm without channel estimation, and it has a performance very similar to that of

system A for values of ρ greater than 0.33. Thus, $\mathrm{SIR}_{\max} = 6.12$ dB for system B as well as for system A. The disadvantage of not using channel estimation is demonstrated by the performance degradation of system B relative to system A for smaller values of ρ, however. Since the correlator form of the max-log-MAP algorithm used in system B does not scale the reliability based on the channel estimate for the dwell interval, there may be a substantial mismatch between the calculated code-symbol reliabilities and the ranking of the detected code symbols for a received word according to their probability of correct detection. Errors in the ranking are most common if there is a large difference in the noise variance among the dwell intervals, which is exactly the condition that determines ρ^*. System B thus results in $\rho^* = 0.16$; it can tolerate severe interference in only six-tenths as large a fraction of its system bandwidth as can be tolerated by system A.

Errors in ranking the detected code symbols by their reliability is largely eliminated by scaling the outputs of the max-log-MAP algorithm for each dwell interval by the inverse noise variance for the dwell interval, as done in system C. Hence, the performance of system C is almost identical to the performance of system A. The effect of using the max-log-MAP approximation in place of the log-MAP algorithm is not significant, if perfect channel estimates are used with both, as seen in Fig. 3.2.

The SIR required to achieve the more stringent packet error probability of 10^{-3} is shown in Fig. 3.3 for systems A, B and C. As the target probability of packet error is decreased, ρ^* decreases for all the three systems, as seen by comparing Fig 3.2 with Fig. 3.3. Systems A and C both result in $\rho* = 0.22$ whereas $\rho^* = 0.12$ for system B. For large values of ρ, use of the max-log-MAP approximation in systems B and C results in a small degradation in performance compared with system A; the degradation is greatest when $\rho = 1$. The performance of system A is about 0.36 dB better than that of system C and 0.22 dB better than that of system B for full-band noise. For system A, $\mathrm{SIR}_{\max} = 6.85$ dB, which is 0.08 dB and 0.04 dB lower than $\mathrm{SIR}_{\max}$ for systems C and B, respectively.

The detection complexity (as measured by EE decoding attempts) is shown as a function of ρ in Fig. 3.4 for systems A and PB under the condition that the SIR for each

values of ρ is equal to the SIR that results in a probability of packet error of 10^{-2}. Since the detection complexity is equal to one for system EO, it is not shown in the figure. If ρ is small, there is a greater disparity in the signal quality in different dwell intervals; consequently, erasure insertion is more effective and most received words are decoded in one of the first few EE decoding attempts of SE decoding in the first iteration of the packet-level iterative decoder. This is illustrated by the performance of System A. It has a detection complexity which increases with ρ — from an average of 1.14 EE decoding attempts per code word if $\rho = 0.3$ to an average of 5.41 EE decoding attempts per code word if $\rho = 1$. Since system PB does not use SE decoding in each iteration, the average EE decoding attempts per code word is small — between 1.01 to 1.15 over all values of ρ.

The average number of EE decoding attempts per code word for systems A, B and C is shown in Fig. 3.5 with a target probability of packet error of 10^{-2}. For system B, the average number of EE decoding attempts per code word decreases as ρ increases from 0.12 to 0.3; the average then increases with further increases in ρ. The detection complexity of system B and system C are very similar to the detection complexity of system A for $\rho \geq 0.3$. The average number of SE decoding attempts per code word for systems A, B and C is also shown in Fig. 3.5. The average number of SE decoding attempts increases steadily with increasing ρ for all three systems. However, the percentage increase is less than for the average number of EE decoding attempts. This indicates that more EE decoding attempts are required per SE decoding, on average, if ρ is large.

The average number of EE and SE decoding attempts per code word is shown in Fig. 3.6 for systems A, B and C and a target packet error probability of 10^{-3}. At the higher SIR required to achieve a lower probability of packet error, more code words are decoded in the first few iterations, decreasing the average number of EE and SE decoding attempts for all the systems. The dependence of detection complexity on ρ is the same for systems A and C for $0.3 \leq \rho \leq 0.8$. For larger values of ρ, there is little change in the detection complexity of system A as ρ varies, but the detection complexity of system C decreases slightly with an increase in ρ. On the other hand, the detection complexity of system B is

slightly less than that of system A for $\rho < 0.8$, but they are equal for $\rho > 0.8$. The average number of SE decoding attempts per code word demonstrates similar behavior among the three systems.

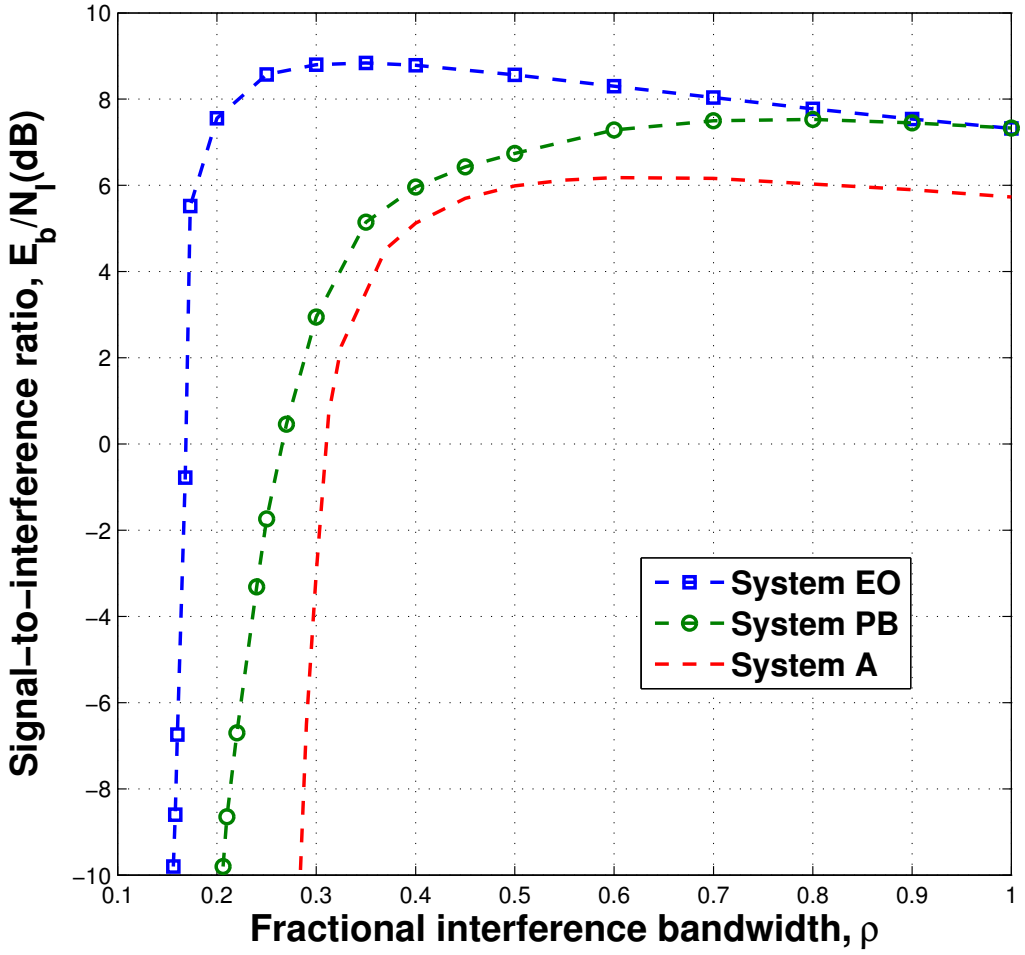

Figure 3.1: Performance of three SFH systems, $P_e = 10^{-2}$.

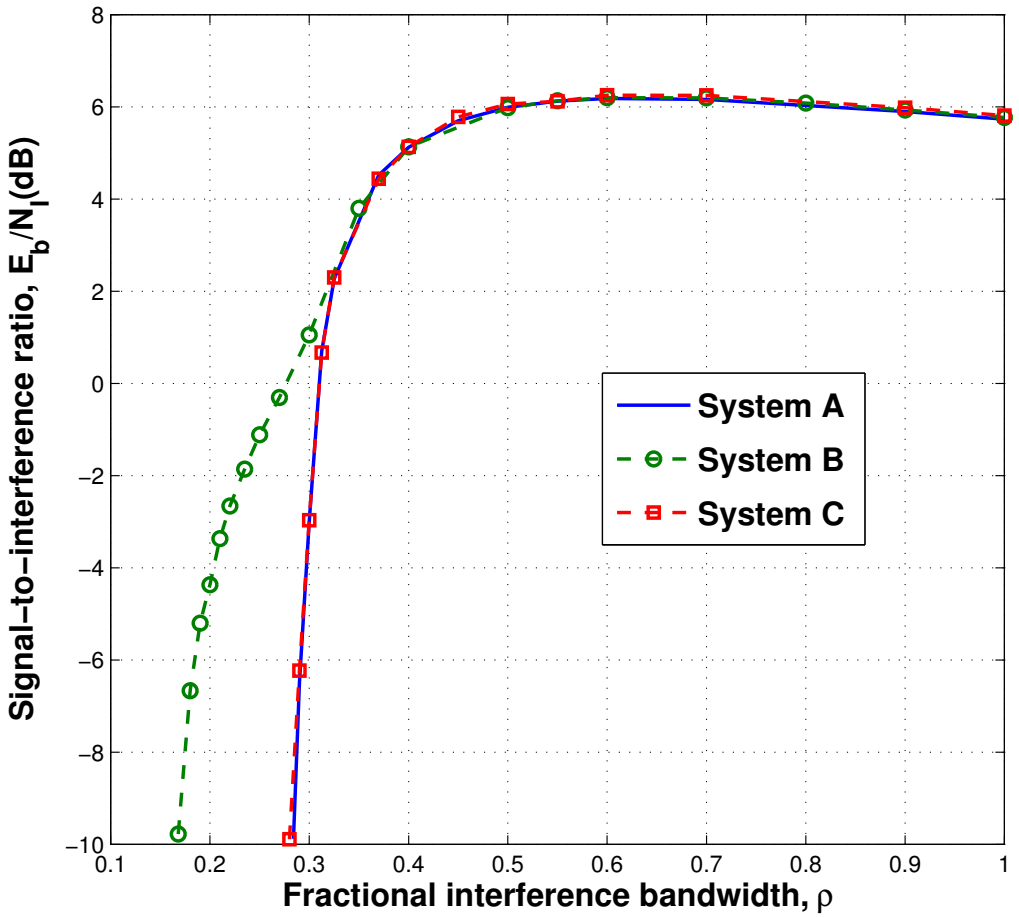

Figure 3.2: Performance of systems with reference SISO algorithms, $P_e = 10^{-2}$.

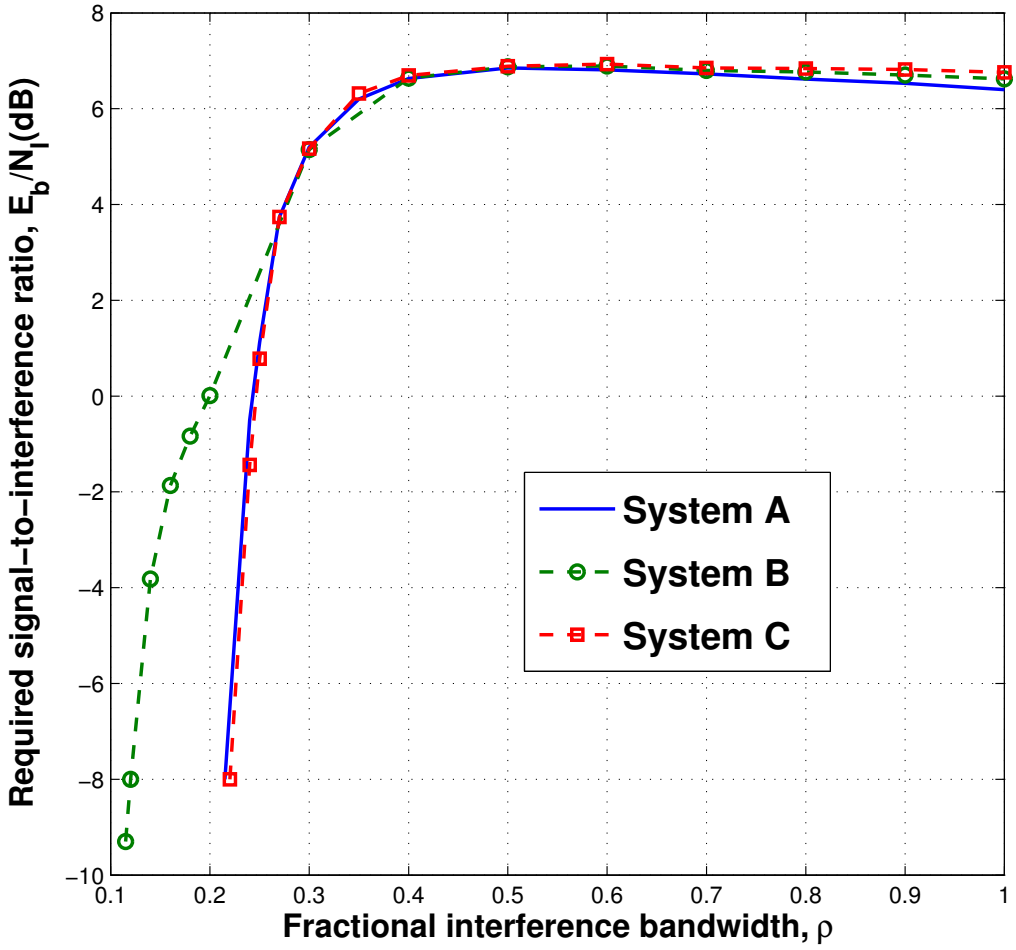

Figure 3.3: Performance of systems with reference SISO algorithms, $P_e = 10^{-3}$.

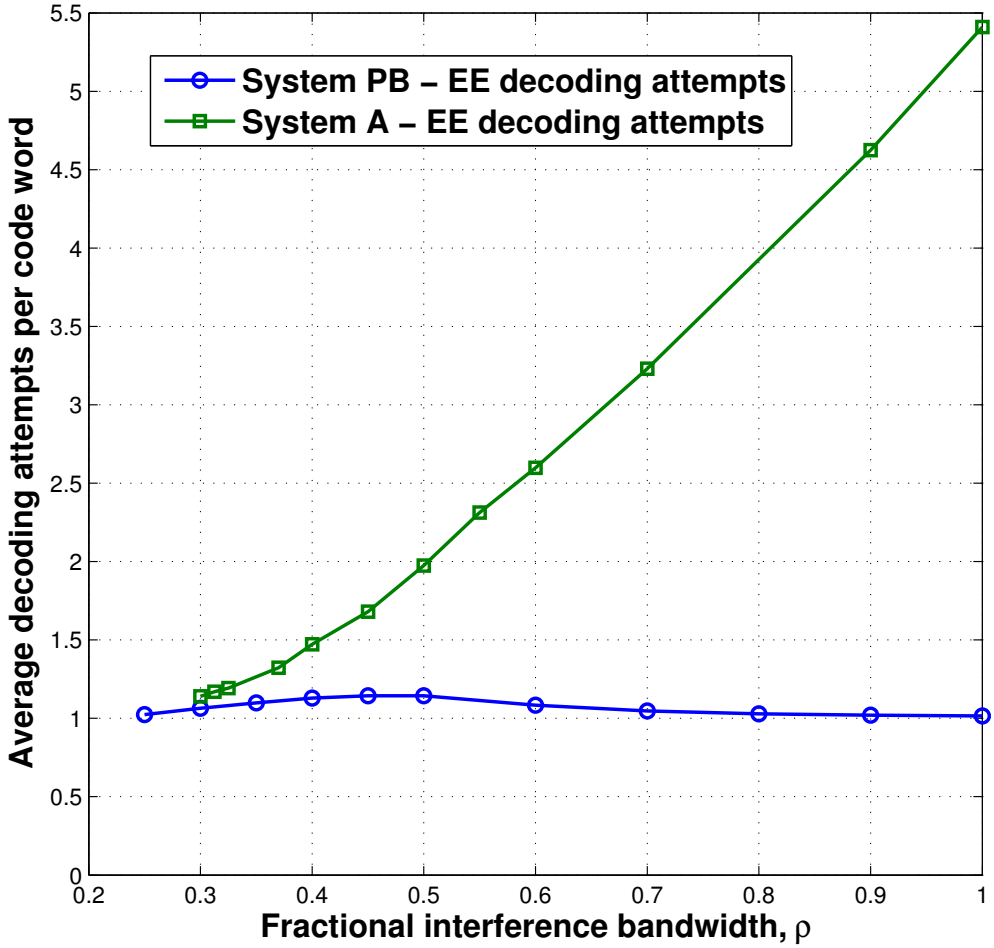

Figure 3.4: Detection complexity of two SFH systems, $P_e = 10^{-2}$.

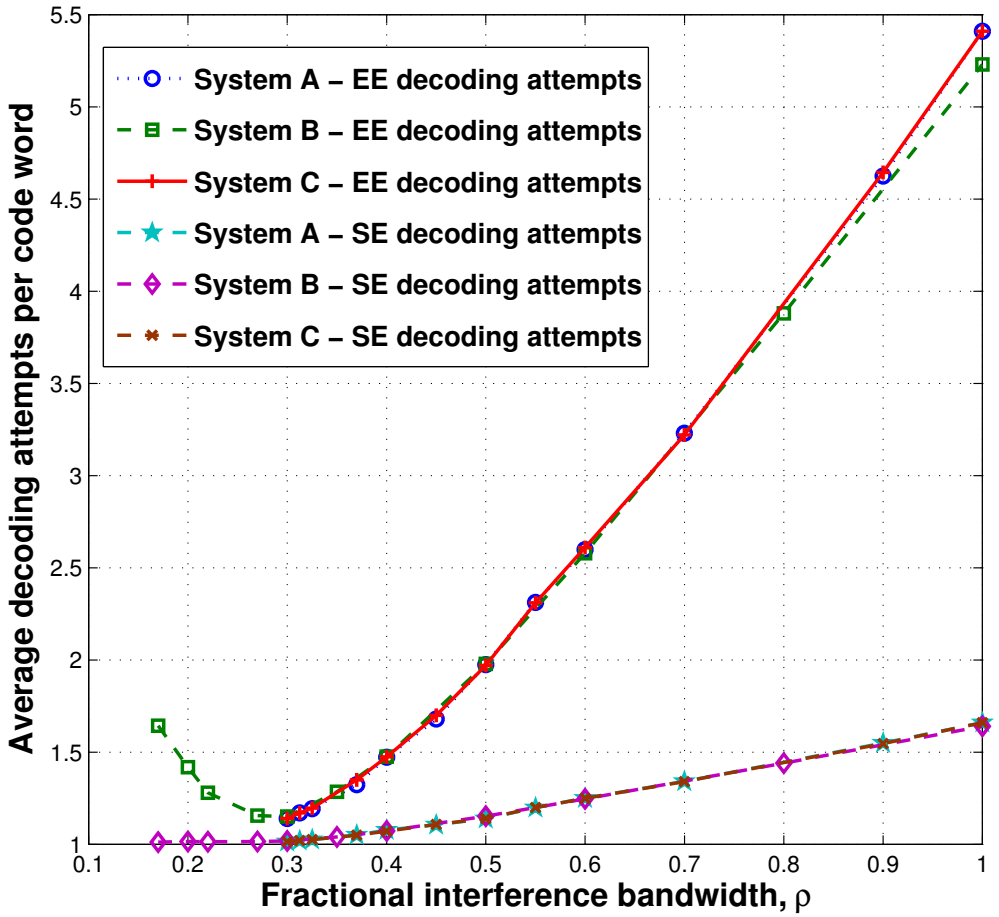

Figure 3.5: Detection complexity of systems with reference SISO algorithms, $P_e = 10^{-2}$.

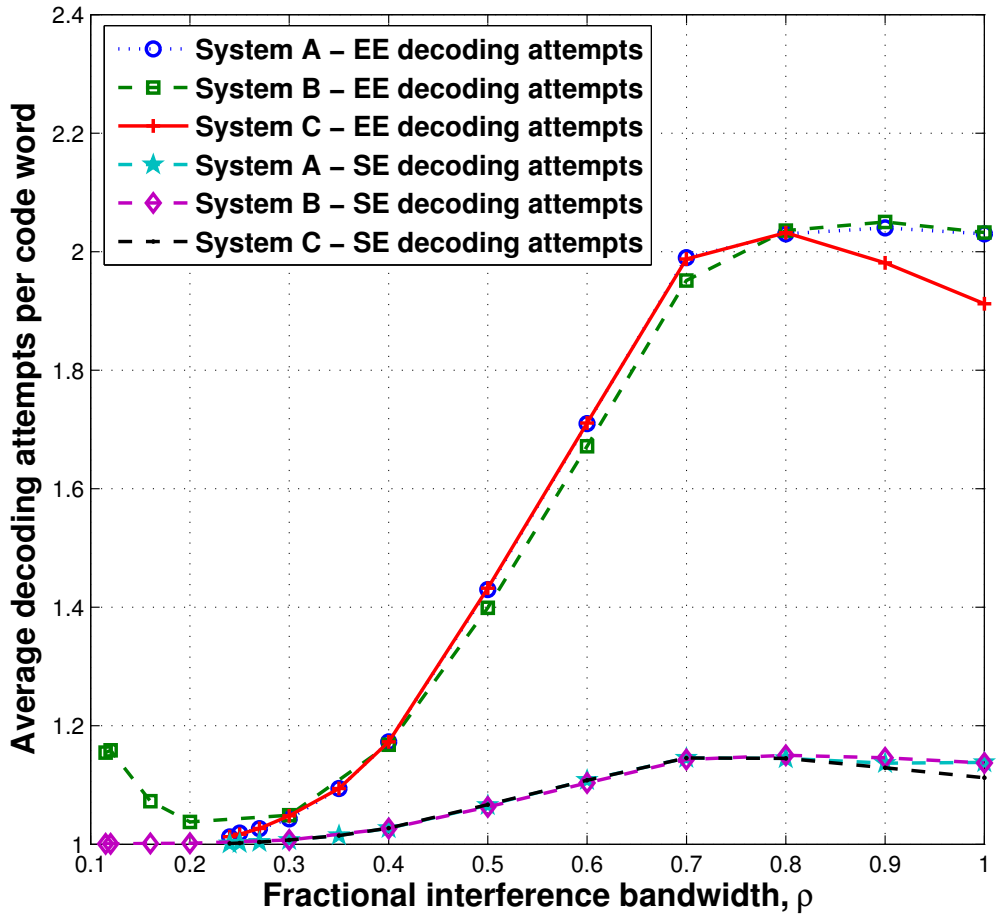

Figure 3.6: Detection complexity of systems with reference SISO algorithms, $P_e = 10^{-3}$.

Chapter 4

Performance Using Different Channel Estimators

The log-MAP algorithm and the form of the max-log-MAP algorithm that uses the noise variance both require channel estimation. The estimators discussed in this chapter estimate the per-dwell variance of the noise in the normalized received channel outputs. Two types of estimator are considered in this chapter: one is based on the method of moments, and the other is an approximate maximum-likelihood (ML) estimator.

4.1 Channel estimation using the method of moments

The estimator considered in this section is a previously reported estimator [17]. It does not require the transmission of training bits; instead, it uses the demodulator outputs for the data portion of the packet to estimate the noise variance in each dwell interval. The noise variance for each dwell interval is estimated from the demodulator outputs once for each packet, and it is used in all the iterations of packet-level detection.

If σ_i^2 is the noise variance of the i^{th} dwell interval, each demodulator output, $r_j^{(i)}$, in the dwell interval is a conditionally Gaussian random variable with variance σ_i^2 and mean

$\pm\sqrt{E_s}$ given the polarity of the corresponding channel symbol. Consequently,

$$\mathrm{E}\left[\left(r_j^{(i)}\right)^2\right] = E_s + \sigma_i^2$$

and

$$\mathrm{E}\left[\left|r_j^{(i)}\right|\right] = \sigma_i\sqrt{\frac{2}{\pi}}e^{-(E_s/2\sigma_i^2)} + \sqrt{E_s}\left[\mathrm{erf}\left(\sqrt{\frac{E_s}{2\sigma_i^2}}\right)\right].$$

Hence,

$$\frac{E\left[\left(r_j^{(i)}\right)^2\right]}{\left(E\left[\left|r_j^{(i)}\right|\right]\right)^2} = \frac{1+\frac{E_s}{\sigma_i^2}}{\left\{\sqrt{\frac{2}{\pi}}e^{-\left(\frac{E_s}{2\sigma_i^2}\right)} + \sqrt{\frac{E_s}{\sigma_i^2}}\left[\mathrm{erf}\left(\sqrt{\frac{E_s}{2\sigma_i^2}}\right)\right]\right\}^2}. \tag{4.1}$$

The estimate is determined from an approximation to the left side of equation (4.1) using empirical moments. Specifically, $\hat{\sigma}_i^2$ is the solution to the equation

$$\frac{\frac{1}{mN_s}\sum_{j=0}^{mN_s-1}\left(r_j^{(i)}\right)^2}{\left(\frac{1}{mN_s}\sum_{j=0}^{mN_s-1}\left|r_j^{(i)}\right|\right)^2} = \frac{1+\frac{E_s}{\hat{\sigma}_i^2}}{\left\{\sqrt{\frac{2}{\pi}}e^{-\left(\frac{E_s}{2\hat{\sigma}_i^2}\right)} + \sqrt{\frac{E_s}{\hat{\sigma}_i^2}}\left[\mathrm{erf}\left(\sqrt{\frac{E_s}{2\hat{\sigma}_i^2}}\right)\right]\right\}^2} \tag{4.2}$$

in the unknown $\hat{\sigma}_i^2$. There is a one-to-one relationship between the left side of equation (4.2) and the channel estimate $\hat{\sigma}_i^2$. There is no simple, closed-form explicit solution for $\hat{\sigma_i}^2$. Instead, in its implementation a lookup table is constructed and $\hat{\sigma}_i^2$ is computed by extrapolating between entries in the table.

A value exceeding $\pi/2$ for the ratio in the left side of equation (4.2) results in a negative value of $\hat{\sigma}_i^2$. The ratio is thus restricted to a maximum value somewhat less than $\pi/2$, which results in a maximum possible value for $\hat{\sigma}_i^2$, denoted $\hat{\sigma}_{max}^2$.

4.2 Approximate maximum-likelihood channel estimation

The ML estimator discussed in this section chooses as the estimated noise variance, $\hat{\sigma}_i^2$, for the i^{th} dwell interval the value that maximizes the joint density function of the demodulator outputs in the dwell interval if the additive, zero-mean, interference-plus-noise

random process is additive, white and Gaussian with a variance that is unknown *a priori*. The estimator uses decision feedback; specifically, it makes use of successfully decoded received words after each packet-level iteration to update the joint density function of the demodulator outputs and obtain a modified channel estimate in each dwell interval for use in the next packet-level iteration.

The ML channel estimator for the i^{th} dwell interval results in the noise-variance estimate given by the solution for σ_i^2 in equation (A.8) if the left side of the equation is set equal to zero. An approximation to $\hat{\sigma}_i^2$ under the assumption that σ_i^2 is small is given by

$$\begin{aligned}\hat{\sigma}_i^2 &\approx \frac{1}{mN_s}\left[\left(r_0^{(i)} - (-1)^{\bar{b}_0^{(i)}}\right)^2 + \sum_{k=1}^{p_0-1}\left(r_k^{(i)} - (-1)^{\bar{b}_0\oplus\ldots\oplus\bar{b}_k^{(i)}}\right)^2\right. \\ &+ \left.\sum_{(k=p_0)}^{(mN_s-1)}\left(\left(r_k^{(i)}\right)^2 + 1\right) - \sum_{l=0}^{\nu-1}\sum_{(k=j_l+p_l)}^{j_{(l+1)}-2} 2\left|r_k^{(i)}\right| - \sum_{l=1}^{\nu-1} 2|x_l| - 2\left|r_{j_\nu-1}^{(i)}\right|\right] \qquad (4.3)\end{aligned}$$

as developed in the appendix where ν, $(j_0,\ldots,j_{\nu-1})$, $(p_0,\ldots,p_{\nu-1})$ and $(x_1,\ldots,x_{\nu-1})$ are defined in the appendix. An approximation to $\hat{\sigma}_i^2$ under the assumption that σ_i^2 is large is given by

$$\begin{aligned}\hat{\sigma}_i^2 &\approx \frac{1}{mN_s}\left[\left(r_0^{(i)} - (-1)^{\bar{b}_0^{(i)}}\right)^2 + \sum_{k=1}^{p_0-1}\left(r_k^{(i)} - (-1)^{\bar{b}_0\oplus\ldots\oplus\bar{b}_k^{(i)}}\right)^2\right. \\ &+ \left.\sum_{(k=p_0)}^{(mN_s-1)}\left(\left(r_k^{(i)}\right)^2 + 1\right)\right] \qquad (4.4)\end{aligned}$$

which is also developed in the appendix.

Figures 4.1, 4.2 and 4.3 compare the normalized expected value of the channel estimate in decibels $\left(10\log\left[\mathrm{E}(\hat{\sigma}_i^2)/\sigma_i^2\right]\right)$ for the actual ML estimate ${\hat{\sigma}_i}^2$ and the approximations made to $\hat{\sigma}_i^2$ in equations (4.3) and (4.4) when the number of code words detected in previous iterations is zero, four and eight, respectively. Without loss of generality for our model, the channel affecting the dwell interval is assumed to contain thermal noise only. The

normalized expected value of the large-variance approximation approaches the normalized expected value of the ML estimate as the SNR approaches $-\infty$ dB, and the normalized expected value of the small-variance approximation approaches the normalized expected value of the ML estimate as SNR approaches $+\infty$ dB. (Note that the ML estimator is *not* an unbiased estimator.)

The normalized expected value of the ML estimate and the normalized expected value of the small-variance approximation are both small compared to the normalized expected value of the large-variance approximation in all examples. Furthermore, the variance of the ML estimate and the variance of its small-variance approximation are very similar throughout the observed region; they are also smaller than the variance of the large-variance approximation. In spite of this, the system using the large-variance approximation as its channel estimator (system E) results in system performance very close to the performance of the reference system (system C), as described in Section 4.3. The large-variance approximation, $\hat{\sigma}_i^2$ in equation (4.4), is used as the estimator in system E in the remainder of the thesis.

4.3 Comparison of systems with practical channel estimators

In this section, the performance of systems D, E and F (each using practical channel estimators) is compared with the performance of system C (with perfect channel estimation). In both examples, the parameter σ_{max}^2 in system D is set to 133 which corresponds to an SINR of -20 dB. The performance is illustrated for two values of the target probability of packet error: 10^{-2} and 10^{-3}. The maximum number of erasures allowed in SE decoding is 10.

Figure 4.4 illustrates the SIR required to achieve a packet error probability of 10^{-2} as a function of ρ for systems C, D, E and F. The estimators of systems D, E and F are able to discriminate effectively between dwell intervals subject to severe interference and dwell intervals that are free of interference; consequently, ρ^* is essentially the same for the

three systems as for system C (that is, $\rho^* = 0.28$). The penalty in performance due to imperfect channel estimation is negligible in systems D, E and F compared with system C for all values of ρ. For example, the value of $\text{SIR}_{\max}$ is 6.26 dB for systems D and F and 6.25 dB for system E, whereas $\text{SIR}_{\max} = 6.22$ dB for system C.

The relative performance of the four systems is also nearly identical if a more stringent performance requirement is considered. The SIR required to achieve a probability of packet error of 10^{-3} is shown in Fig. 4.5 for each system . Each system achieves comparable robustness in the presence of severe partial-band interference, with $\rho^* = 0.23$ for all four systems. Similarly, there is only a minimal difference in the worst-case performance among the four systems. The perfect channel estimates of system C result in $\text{SIR}_{\max} = 6.93$ dB. System E results in the same value of $\text{SIR}_{\max}$, whereas $\text{SIR}_{\max} = 6.94$ dB for system D and $\text{SIR}_{\max} = 6.98$ dB for system F.

The detection complexity of the four systems is also very similar. It is shown in Fig. 4.6 for systems C, D, E and F with a SIR for each value of ρ that is equal to the SIR required for the corresponding system to achieve a probability of packet error of 10^{-2}. The average number of EE decoding attempts and the average number of SE decoding attempts per transmitted code word for systems D, E and F exhibit similar dependence on ρ as for system C. Moreover, the difference in the detection complexity of the four systems is negligible.

The detection complexity of the four systems is shown in Fig. 4.7 for a target probability of packet error of 10^{-3}. Again, the detection complexity of all four systems is similar using either of the two measures of complexity. The detection complexity of any two of the systems differs by less than 5% for each value of ρ.

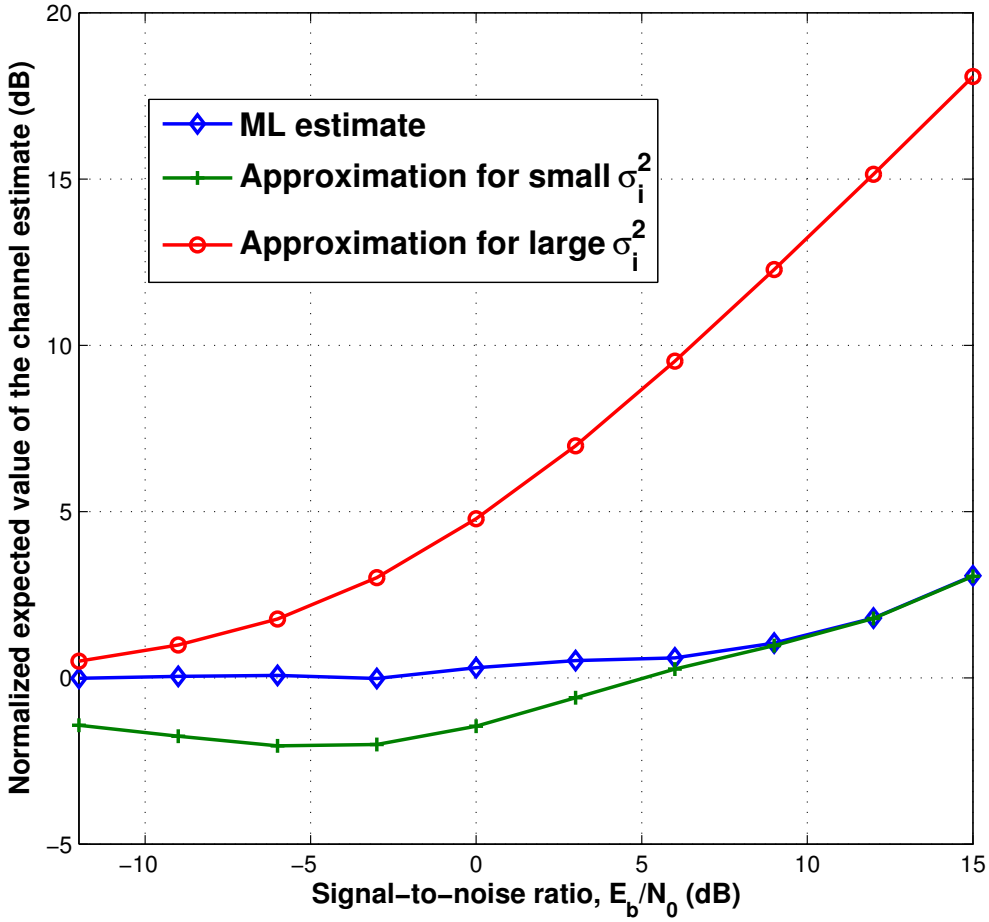

Figure 4.1: Normalized expected value of different channel estimates, zero code words previously detected.

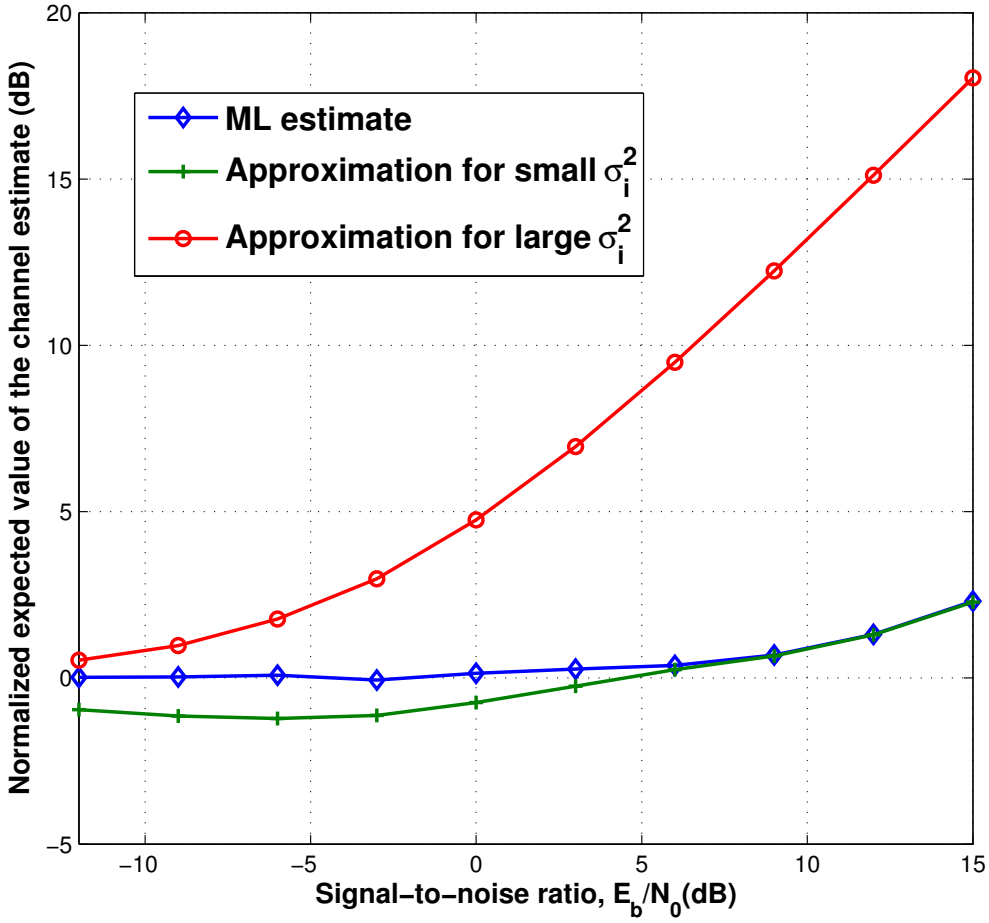

Figure 4.2: Normalized expected value of different channel estimates, four code words previously detected.

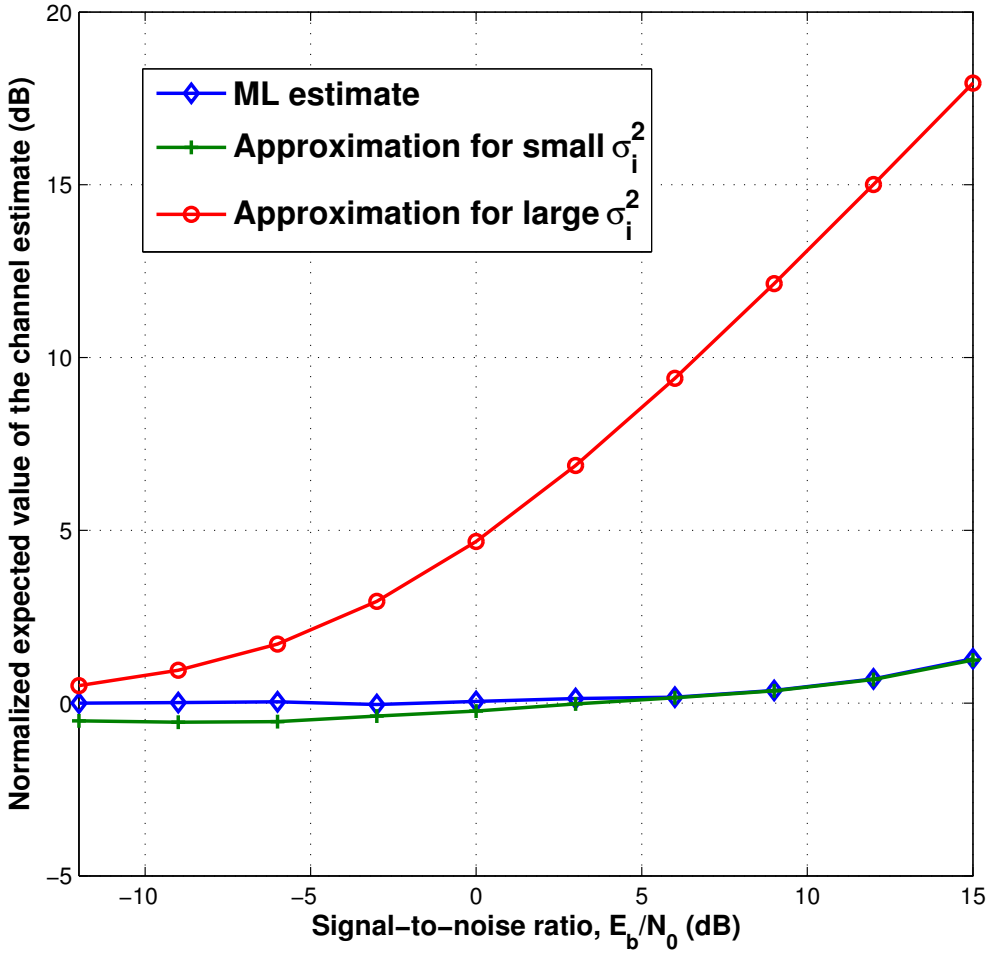

Figure 4.3: Normalized expected value of different channel estimates, eight code words previously detected.

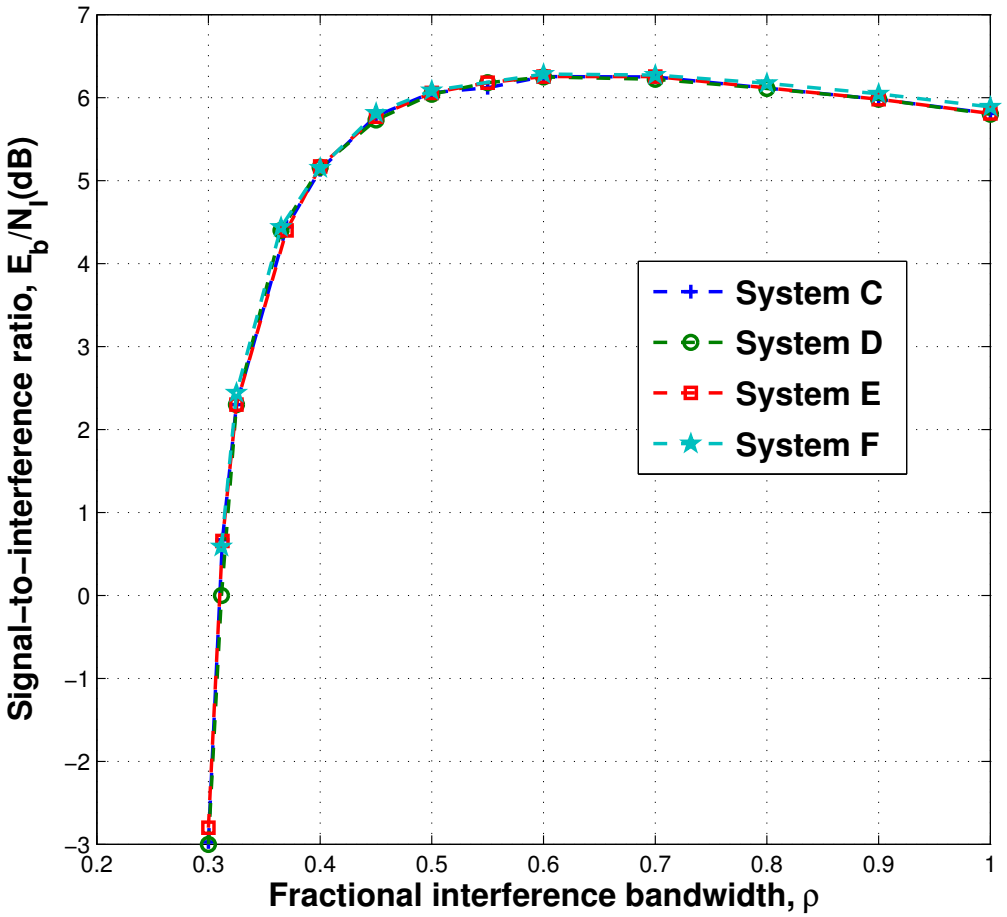

Figure 4.4: Performance of systems with different channel estimators, $P_e = 10^{-2}$.

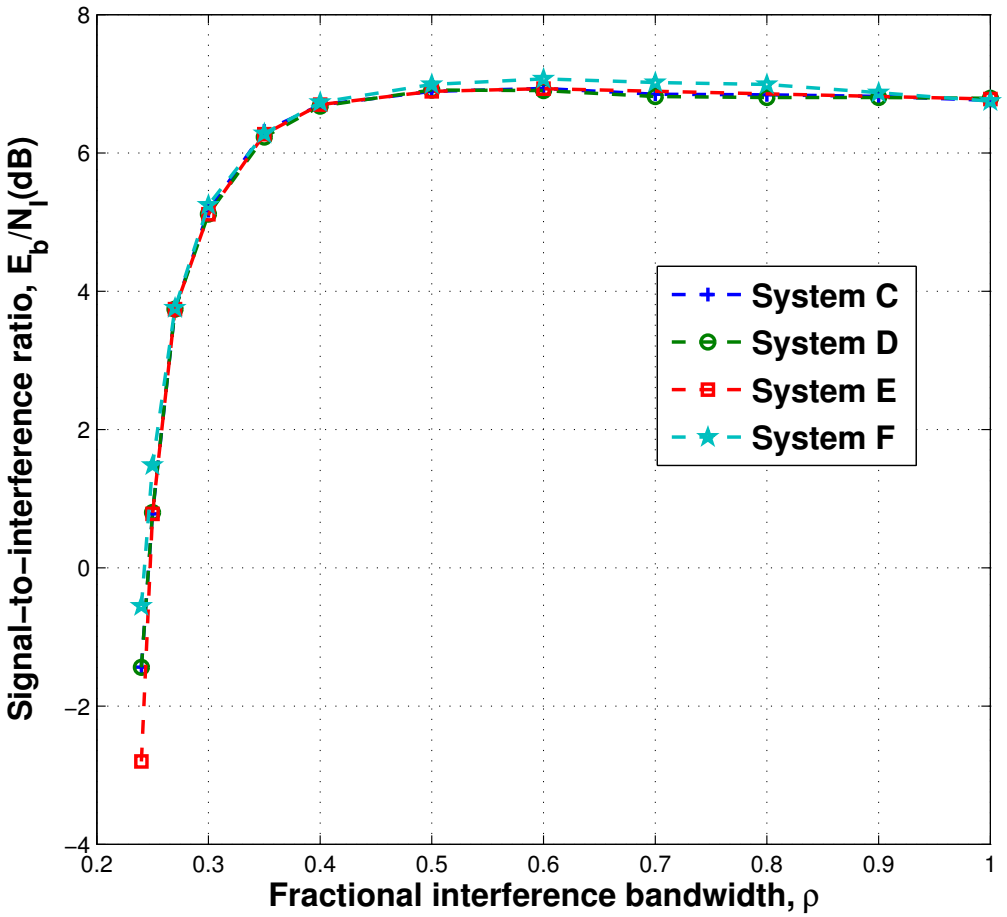

Figure 4.5: Performance of systems with different channel estimators, $P_e = 10^{-3}$.

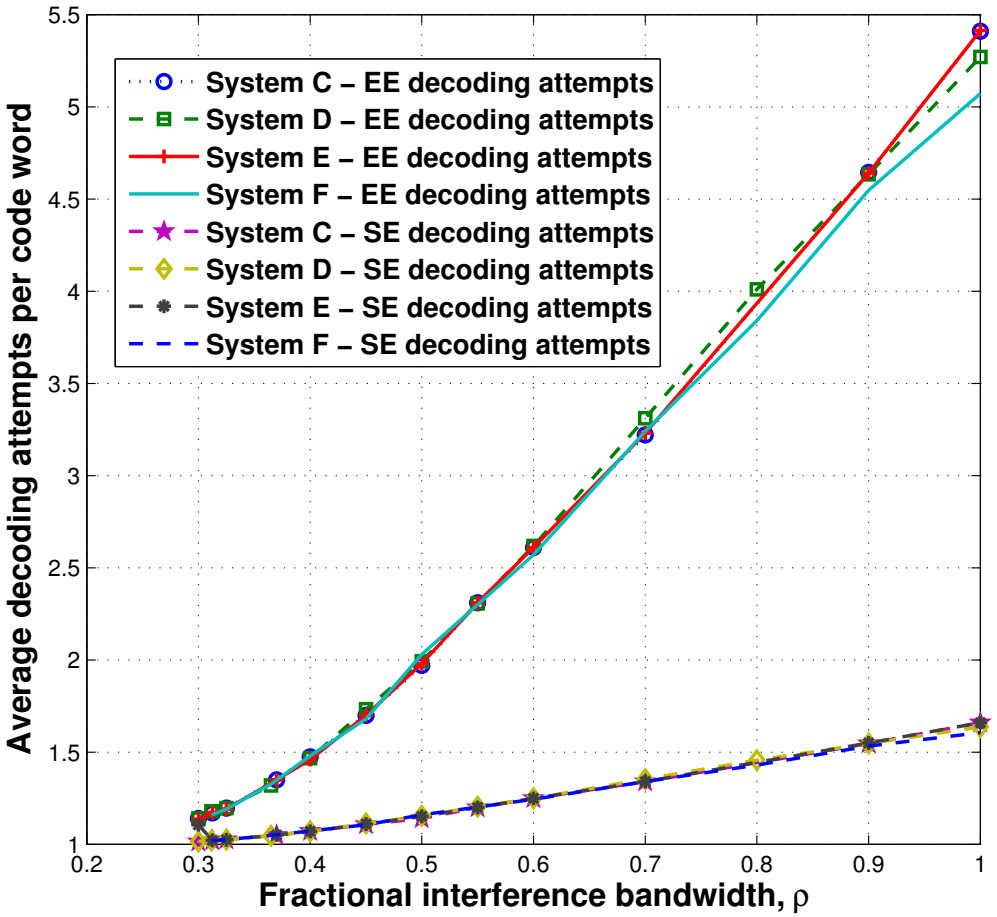

Figure 4.6: Detection complexity of systems with different channel estimators, $P_e = 10^{-2}$.

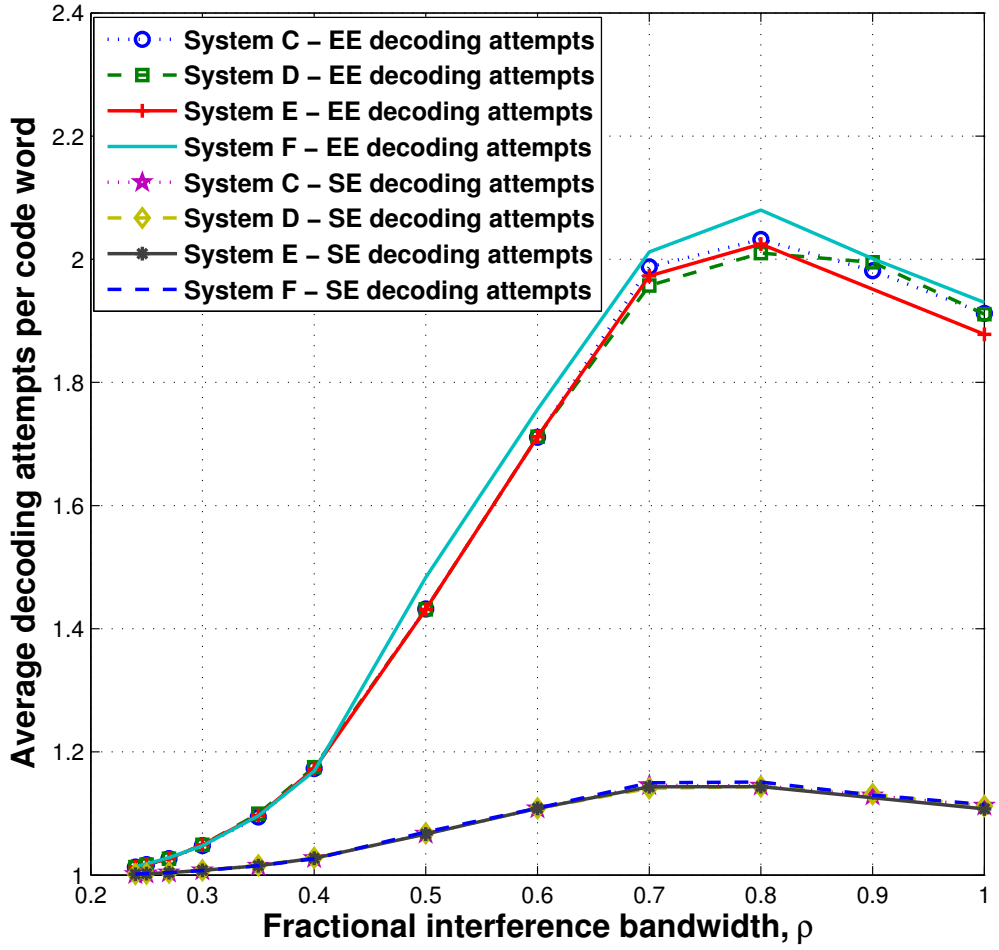

Figure 4.7: Detection complexity of systems with different channel estimators, $P_e = 10^{-3}$.

Chapter 5

Performance Using Different Erasure Constraints for SE Decoding

The form of SE decoding considered in the thesis first employs bounded-distance EE decoding of the received word using a maximum allowable number of code-symbol erasures, $e_{\max}$, where $e_{\max} \leq n-k$. The number of erasures is reduced by two in each subsequent EE decoding attempt, until either a decoding attempt results in a valid detected code word or the attempt with zero erasures (bounded-distance errors-only decoding) results in a decoder failure. SE decoding is most effective in circumstances in which the channel between the encoder and the input to the decoder produces code-symbol decisions of widely differing probabilities of correctness and those differences can be identified at the receiver prior to decoding.

Increasing the value of $e_{\max}$ in SE decoding increases the probability of successful decoding for a given channel between the encoder and the input to the decoder, but in some channels it also increases the probability that an incorrect code word will satisfy the decoder's distance bound, with a corresponding increase in the probability of undetected

code-word error at the output of the decoder. (For example, an increase in the value of $e_{\max}$ results in an increase in the probability of undetected code-word error in a single instance of SE decoding if there is an i.i.d. channel between the encoder and the input to the decoder [16].) In the context of the iterative packet-level detection considered in the thesis, the impact of the two counteracting effects on the probability of packet error is not apparent *a priori*.

Similarly, their impact on the average detection complexity is uncertain *a priori*. Since a change in $e_{\max}$ alters the occurrences of both correct and incorrect detected code words up to any given packet-level iteration in a manner that is difficult to predict, and those occurrences affect the subsequent decoding of the received words that have not yet been successfully decoded, it is not immediately apparent what effect the change will have on the success or failure of SE decoding in each remaining iteration for any of the received words. Furthermore, an increase in $e_{\max}$ can potentially increase or decrease the average number of EE decoding attempts in the SE decoding attempts that are *successful*. This chapter is focused on examining the effect of $e_{\max}$ on both the system's probability of error and its detection complexity.

5.1 Effect of $e_{\max}$ on the probability of packet error

The effect of $e_{\max}$ on the probability of packet error is illustrated in Fig. 5.1 which shows the SIR required to achieve a probability of packet error of 10^{-2} for system D. (Recall that system D uses the max-log-MAP SISO detector with one-shot channel estimation.) The performance is shown for even values of $e_{\max}$ between 8 and 14.

The performance of the system in full-band noise ($\rho = 1$) improves as $e_{\max}$ is decreased. If $e_{\max} = 14$, a SIR of 7.22 dB is required to achieve the desired probability of packet error. A reduction of $e_{\max}$ to 12 reduces the occurrences of undetected code-word errors significantly in the presence of full-band noise so that the required SIR is reduced by 0.58 dB to 6.64 dB. The required SIR decreases to 5.82 dB if $e_{\max}$ is decreased to 10, and

$e_{\max} = 8$ results in a required SIR of 5.8 dB in full-band noise.

The worst-case performance of the system occurs with a fairly wide interference bandwidth for each choice of $e_{\max}$; thus, the dependence of the worst-case performance on $e_{\max}$ is similar to its dependence with full-band noise. If $e_{\max} = 14$, the performance of the system is characterized by $\text{SIR}_{\max} = 7.45$ dB. A reduction of $e_{\max}$ to 12 reduces $\text{SIR}_{\max}$ by 0.73 dB to 6.72 dB. If $e_{\max}$ is equal to 8 or 10, $\text{SIR}_{\max} = 6.25$ dB.

The frequently occurring low reliability of a code-symbol decision in dwell intervals with severe interference is readily identified by the SISO detector. The ability of the receiver to compensate for severe interference by erasing the corresponding symbols is limited by the choice of $e_{\max}$, however, thus limiting the fraction of the frequency band in which the system can tolerate very strong interference. Consequently, a smaller value of e_{max} results in a smaller value of ρ^*. In Fig. 5.1, it is seen that a reduction in $e_{\max}$ from 14 to 12 results in a reduction in ρ^* from 0.34 to 0.32. Further reductions in $e_{\max}$ to 10 and 8 result in values of ρ^* of 0.29 and 0.26, respectively. System D can tolerate severe interference in a 31% wider bandwidth if the SE decoder uses $e_{\max} = 14$ than if it uses $e_{\max} = 8$. By comparing the results of Fig. 5.1 with those of Fig. 3.1, it is also seen that system D with $e_{\max} = 14$ tolerates severe interference in a 126% wider bandwidth than system EO and a 79% wider bandwidth than system PB.

The choice of $e_{\max}$ results in a clear tradeoff between the system's robustness against severe interference occupying a moderate fraction of the frequency band and its performance in the presence of wide-band interference (including a hostile interferer that selects its bandwidth to maximize the harm it causes). The highest limit on the number of erasures, $e_{\max} = 14$, results in the best performance in interference that occupies up to 42% of the frequency band, yet it results in the poorest performance among the four choices for $e_{\max}$ if the interference occupies more than 47% of the frequency band. Conversely, the system with $e_{\max} = 10$ is less effective against severe partial-band interference, but it achieves the best performance among all four choices for $e_{\max}$ if the interference occupies more than 51% of the frequency band and it matches the system with $e_{\max} = 8$ in achieving the lowest

value of $\mathrm{SIR}_{\max}$.

5.2 Effect of $e_{\max}$ on the detection complexity

The detection complexity of system D is shown in Figs. 5.2 and 5.3 as a function of ρ for each of the same four values of $e_{\max}$ considered in Section 5.1. The average number of SE decoding attempts per transmitted code word is shown in Fig. 5.2, and the average number of EE decoding attempts per transmitted code word is shown in Fig. 5.3. For each choice of $e_{\max}$ and a given value of ρ, the detection complexity of the system is shown for the SIR required to achieve a probability of packet error of 10^{-2} in a partial-band interference channel characterized by that value of ρ.

The average number of SE or EE decoding attempts is greater if a small value of $e_{\max}$ is used than if a large value is used when the system operates in a channel with wide-band interference. If the channel exhibits full-band noise only, for example, the average number of SE decoding attempts decreases as $e_{\max}$ is increased from 8 to 10, decreases markedly if $e_{\max}$ is increased from 10 to 12, and decreases further if $e_{\max}$ is increased from 12 to 14. Increasing $e_{\max}$ from 8 to 10 results in a reduction in the average number of SE decoding attempts per transmitted code word from 1.7 to 1.6. The increase in $e_{\max}$ increases the average number of EE decoding attempts per SE decoding from 2.84 to 3.25. however, so that the average number of EE decoding attempts per transmitted code word increases from 4.85 to 5.3. Further increases in $e_{\max}$ to 12 and 14 result in a decrease in the average number of SE decoding attempts to 1.14 and 1.03, respectively, and a decrease in the average number of EE decoding attempts to 2.4 and 1.7, respectively. The average number of SE decoding attempts with $e_{\max} = 12$ and $e_{\max} = 14$ are thus 29% less and 36% less, respectively, than with $e_{\max} = 8$ in full-band noise. The average number of EE decoding attempts with $e_{\max} = 12$ and $e_{\max} = 14$ are 51% less and 65% less, respectively, than with $e_{\max} = 8$ in full-band noise. The SIR required to achieve the target error probability is higher if $e_{\max} = 12$ and $e_{\max} = 14$ than if $e_{\max} = 8$ and $e_{\max} = 10$; hence, a typical SE

decoding attempt requires fewer EE decoding attempts with the larger values of $e_{\max}$ than with the smaller values of $e_{\max}$ at their respective values of the required SIR.

The average number of SE decoding attempts decreases with each increase in $e_{\max}$ for each value of ρ and those values of $e_{\max}$ for which the corresponding value of ρ^* is less than ρ. The same dependence of the average number of EE decoding attempts on $e_{\max}$ exhibited in full-band noise occurs in partial-band interference with $\rho \geq 0.4$. For lower values of ρ, the average number of EE decoding attempts is slightly greater if $e_{\max} = 8$ than if $e_{\max} = 10$. Both measures of complexity increase with increasing ρ if either of the two smaller values of $e_{\max}$ is used, whereas both measures exhibit a maximum at an intermediate value of ρ if $e_{\max} = 12$. The average number of SE decoding attempts is only slightly greater than one for all values of ρ if $e_{\max} = 14$, and the average number of EE decoding attempts never exceeds 1.7 in that instance. The detection complexity of system D with $e_{\max} = 14$ is still greater than the detection complexity of system PB for each value of ρ, however, as seen by comparing Fig. 5.3 with Fig. 3.4.

The discussion of the performance tradeoff in Section 5.1 can be expanded to account for the computational cost of the system's performance. A moderate value of the erasure constraint ($e_{\max} = 10$ in this example) provides good performance in full-band noise at the cost of high detection complexity and modest protection against severe partial-band interference. Larger values of the erasure constraint result in a lower detection complexity for all of the partial-band interference channels considered and improved protection against severe partial-band interference at the cost of poorer worst-case performance (and poorer performance in full-band noise) than with the smaller value of $e_{\max}$.

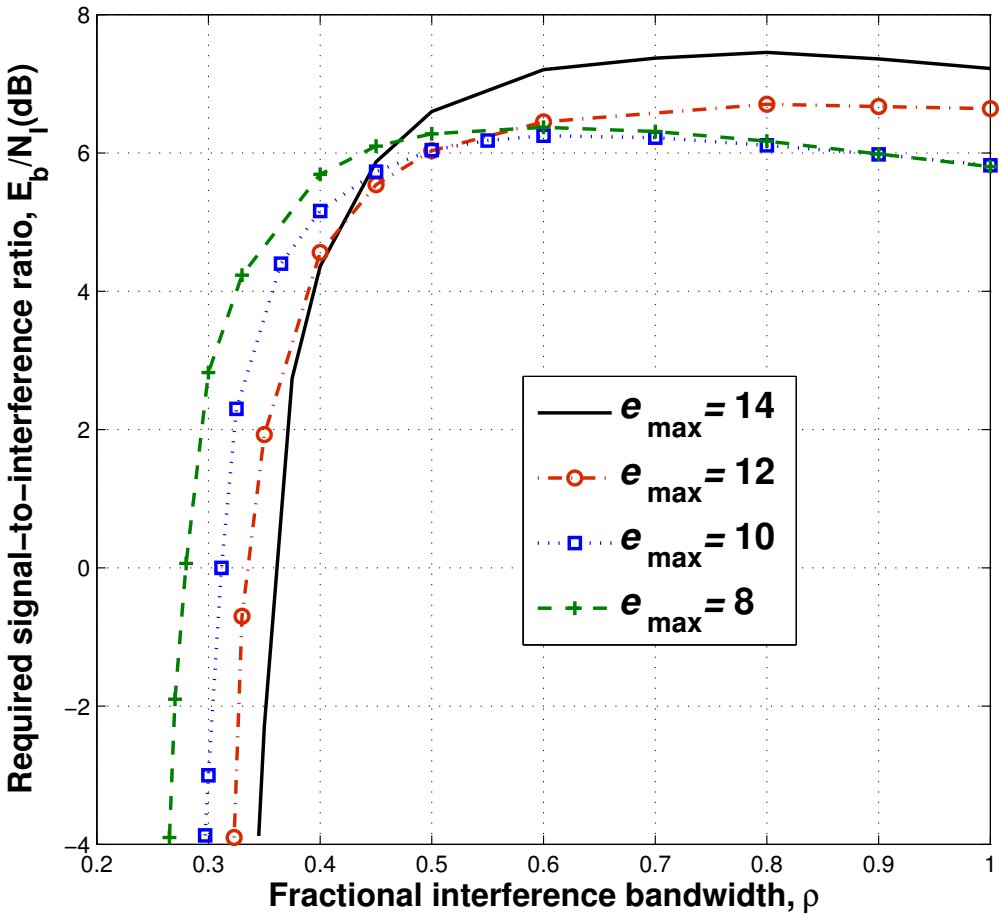

Figure 5.1: Required SINR of system D for several values of $e_{\max}$.

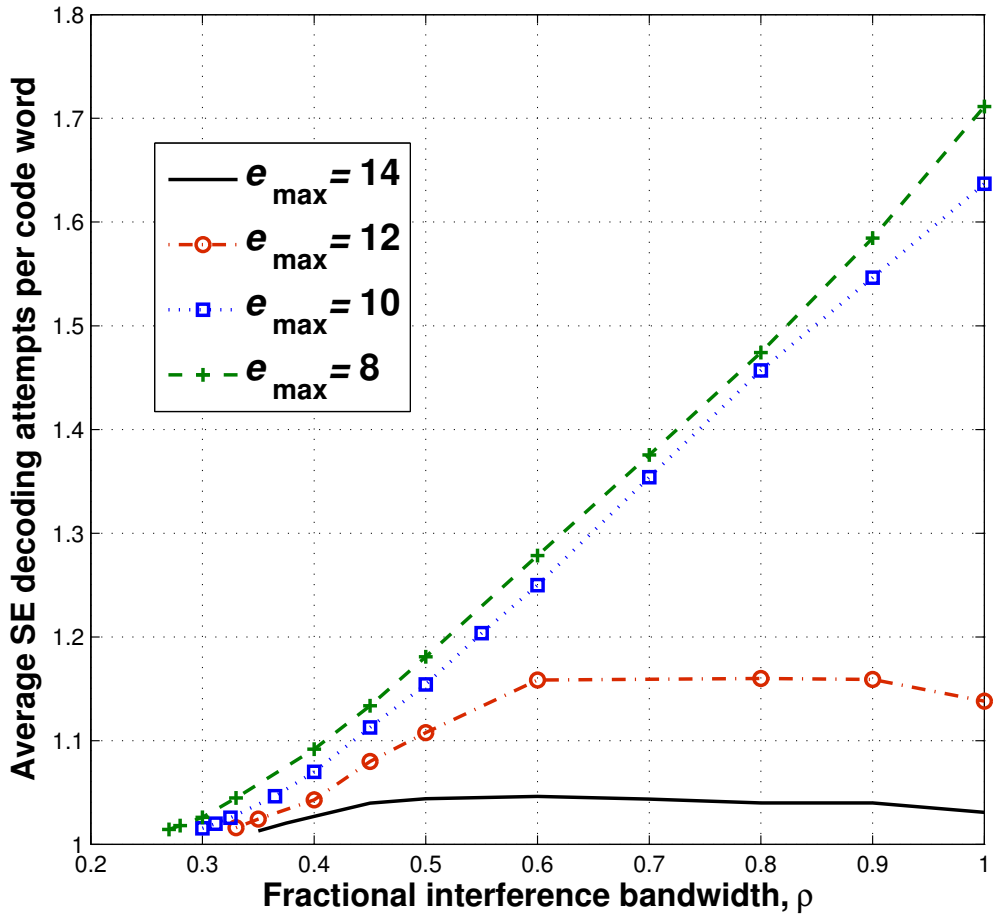

Figure 5.2: SE decoding complexity of system D for several values of $e_{\max}$.

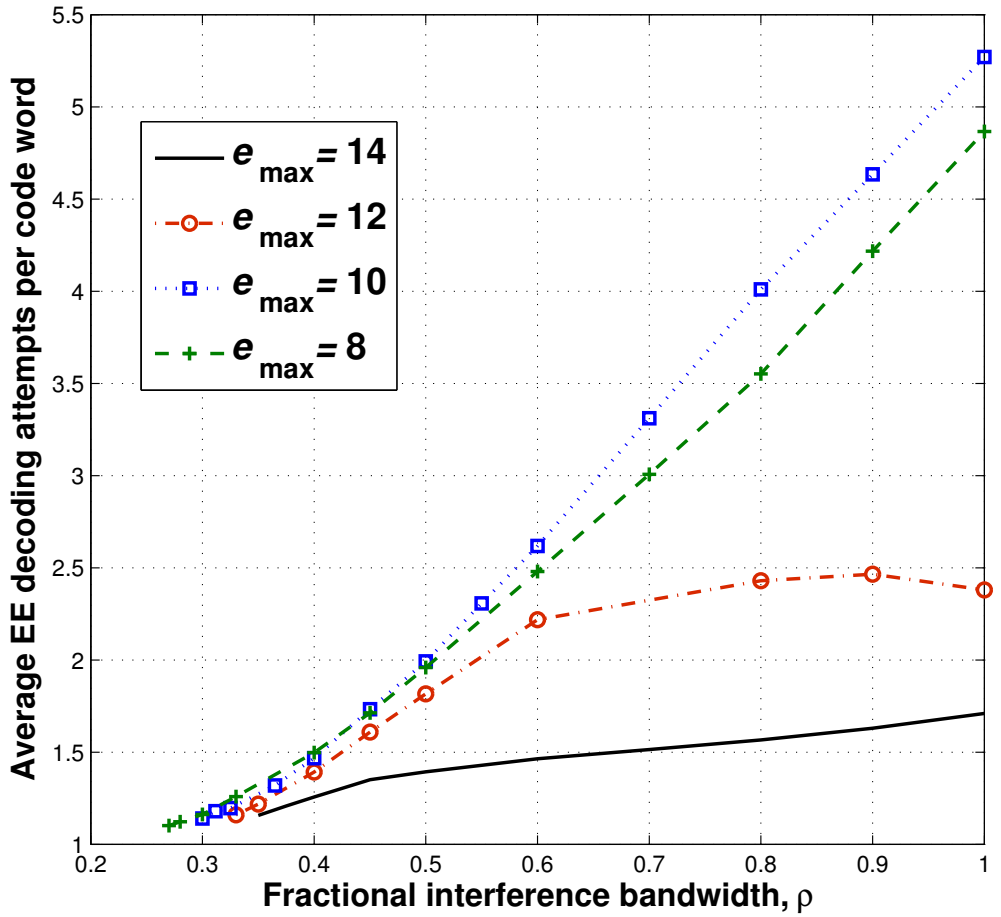

Figure 5.3: EE decoding complexity of system D for several values of $e_{\max}$.

Chapter 6

Conclusion

In this thesis we have considered a SFH packet radio system in which multiple R-S code words are interleaved across the dwell intervals of a packet transmission and the contents of each dwell interval are differentially encoded prior to transmission. Each dwell interval is transmitted in a randomly selected frequency slot within the system's frequency band. The received signal is subjected to both full-band white Gaussian noise and partial-band white Gaussian noise that is present in a fraction of the available frequency slots.

The use of SISO detection of the code bits within each dwell interval is exploited for SE decoding of each received word during each iteration of packet-level iterative detection. The performance of packet-level iterative detection using SISO per-dwell code-bit detection and SE decoding results in performance that is substantially better than the performance of either of two reference SFH systems: a system that uses a non-iterative receiver with one-shot EO decoding of each received word, and a system that uses a parity bit per code symbol and an iterative receiver with one EE decoding attempt per iteration for each received word. The performance is improved by a minimum of 2 dB and 1 dB over all the partial-band interference channels in comparison with the latter two systems, respectively, and severe interference is tolerated in a much larger fraction of the frequency band. The improvement in performance is achieved at the cost of up to a five-fold increase in the computation required at the receiver, however, which is most pronounced in the presence of wide-band

interference.

Three SISO detection algorithms are considered in the thesis: log-MAP detection (which requires an estimate of the noise variance of the samples in each dwell interval) and two forms of max-log-MAP detection, including one that requires an estimator and another that does not. The two SISO algorithms with noise-variance estimation result in essentially the same performance if perfect estimation occurs in each. Both exhibit greater robustness to severe partial-band interference than the system using max-log-MAP detection without estimation. The three choices of the SISO algorithm result in similar detection complexity at the receiver.

Two practical forms of noise-variance estimation in each dwell interval are considered in the thesis. A one-shot estimator uses the method of moments applied to the soft channel outputs in the dwell interval to produce a single estimate that is used throughput the detection of the packet. In contrast, an iterative decision-feedback ML estimator updates its estimate after each packet-level iteration based on the previous results of SE decoding. Two computationally efficient approximations to the ML estimator are also considered: a small-variance approximation, and a large-variance approximation. The fidelity of the ML estimator and its small-variance approximation are both very good, and both achieve better fidelity than the large-variance approximation. Yet the system using the large-variance approximation (as well as the one using the method-of-moments estimator) achieves performance and detection complexity almost identical to that of the system with access to perfect noise-variance estimates.

An erasure constraint in SE decoding (the maximum number of allowable code-symbol erasures in a received word) is also considered in the thesis, and its effect on the system performance is evaluated. An intermediate value of the constraint results in the best system performance in full-band noise. The performance is approximately 2 dB better than the performance if a large number of erasures is allowed or if no erasures are allowed (EO decoding). The worst-case performance over all partial-band channels is also optimized by the same choice of the erasure constraint. The tolerance of the system to severe partial-

band interference improves as the erasure constraint is increased, however, so the choice of the constraint provides a tradeoff between good performance in the presence of moderate-strength wide-band interference and good performance in the presence of severe interference of narrow bandwidth. The superior system performance in wide-band interference that is achieved at the cost of as much as a three-fold increase in detection complexity, however.

Appendix

Maximum-Likelihood Channel Estimator

In this appendix, we derive the ML estimator of the noise variance in the demodulator outputs within a single dwell interval of the SFH system. Two approximations to the ML estimator are developed, first in a form that provides some insight and analytical tractability and then in a form that is computationally feasible. The development is based on the assumption of independent, equally likely information bits at the source, so that the code bits within any single row of the interleaver (represented by the signal in the corresponding dwell interval) are independent and equally likely random variables. The circumstance is considered in which an arbitrary subset of the code bits are known *a priori* as a result of feedback of *correctly* detected code words from earlier iterations of packet-level detection. It is assumed that conditioning on the known code bits does not alter the joint distribution of the remaining unknown code bits. Only the contents of a single dwell interval are considered; thus, the notation introduced in Chapter 2 is simplified by omitting the notation that designates the dwell interval.

Derivation of the Estimator and Its Approximations

Suppose the interleaved code bits which correspond to received words decoded in a previous packet-level iteration have indices in $B_{1,0} \cup B_{1,1} \cup \ldots \cup B_{1,\nu-1}$ where

$$B_{1,l} = \{j_l, \ldots, j_l + p_l - 1\} \text{ for } 0 \le l \le \nu - 1. \tag{1}$$

Similarly, suppose the interleaved code bits which correspond to received words not decoded in any previous packet-level iteration have indices in $B_{0,0} \cup B_{0,1} \cup \ldots \cup B_{0,\nu-1}$ where

$$B_{0,l} = \{j_l + p_l, \ldots, j_{(l+1)} - 1\} \text{ for } 0 \le l \le \nu - 1.$$

The ordered sequence of interleaved code bits at the transmitter for the dwell interval can thus be expressed as

$$\left\{\overline{b}_{j_0}, \ldots, \overline{b}_{j_0+p_0-1}, \overline{b}_{j_0+p_0}, \ldots, \overline{b}_{j_1-1}, \ldots \overline{b}_{j_{(\nu-1)}}, \ldots, \overline{b}_{j_{(\nu-1)}+p_{(\nu-1)}-1}, \overline{b}_{j_{(\nu-1)}+p_{(\nu-1)}}, \ldots, \overline{b}_{j_\nu-1}\right\}.$$

Without loss of generality, $j_0 = 0$, $p_0 \geq 0$, $j_{l+1} > j_l + p_l$ for $0 \leq l \leq \nu - 2$, $p_l > 0$ for $1 \geq l \geq \nu - 2$, either $p_{\nu-1} > 0$ or $j_{\nu-1} + p_{\nu-1} = j_\nu$, and $j_\nu = mN_s$.

Let

$$\tilde{\mathbf{b}} = \left(\overline{b}_{j_0+p_0}, \ldots, \overline{b}_{j_1-1}, \overline{b}_{j_1+p_1}, \ldots, \overline{b}_{j_2-1}, \ldots, \overline{b}_{j_{\nu-1}+p_{\nu-1}}, \ldots, \overline{b}_{j_\nu-1}\right)$$

be the ordered sequence of the q unknown code bits, where

$$q = \sum_{l=0}^{\nu-1} (j_{l+1} - j_l - p_l).$$

Then the conditional joint distribution of the demodulator outputs in the dwell interval, given the known code bits can be expressed as

$$\begin{aligned} f(\mathbf{r}) &= \frac{1}{2^q}\left[f(\mathbf{r}|\tilde{\mathbf{b}} = (0,0,\ldots,0)) + f(\mathbf{r}|\tilde{\mathbf{b}} = (0,0,\ldots,1)) + \ldots + f(\mathbf{r}|\tilde{\mathbf{b}} = (1,1,\ldots,1))\right] \\ &= \frac{1}{2^q}\left[f_0(\mathbf{r}) + f_1(\mathbf{r}) + \ldots + f_{2^q-1}(\mathbf{r})\right]. \end{aligned} \tag{2}$$

Suppose $\{d_{j,0}, \ldots, d_{j,mN_s-1}\}$ is the sequence of differential encoder outputs in the dwell interval under the jth condition in equation (2), $0 \leq j \leq 2^q - 1$. Then

$$f_j(\mathbf{r}) = \frac{1}{(2\pi\sigma^2)^{\frac{mN_s}{2}}} e^{\frac{-A_j}{2\sigma^2}}$$

where

$$A_j = \sum_{k=0}^{mN_s-1} \left(r_k - (-1)^{d_{j,k}}\right)^2.$$

The joint conditional density function of the demodulator outputs given the known code

bits can thus be written as

$$f(\mathbf{r}) = \frac{1}{2^q} \frac{1}{(2\pi\sigma^2)^{\frac{mN_s}{2}}} \left[e^{\frac{-A_0}{2\sigma^2}} + e^{\frac{-A_1}{2\sigma^2}} + \ldots + e^{\frac{-A_{2^q-1}}{2\sigma^2}} \right].$$

The ML estimate, $\hat{\sigma}^2_{\mathrm{ML}}$, is the value of σ^2 that maximizes $f(\mathbf{r})$ for the given demodulator outputs. The derivative of $f(\mathbf{r})$ with respect to σ^2 is given by

$$\begin{aligned}
\frac{\mathrm{d}\left(f(\mathbf{r})\right)}{\mathrm{d}\sigma^2} &= \frac{1}{2^q} \frac{1}{(2\pi\sigma^2)^{\frac{mN_s}{2}}} \left[\frac{A_0}{2\sigma^4} e^{\frac{-A_0}{2\sigma^2}} + \frac{A_1}{2\sigma^4} e^{\frac{-A_1}{2\sigma^2}} + \ldots + \frac{A_{2^q-1}}{2\sigma^4} e^{\frac{-A_{2^q-1}}{2\sigma^2}} \right] \\
&\quad + \frac{1}{2^q (2\pi)^{\frac{mN_s}{2}}} \left(-\frac{mN_s}{2\sigma^{(mN_s+2)}} \right) \left[e^{\frac{-A_0}{2\sigma^2}} + e^{\frac{-A_1}{2\sigma^2}} + \ldots + e^{\frac{-A_{2^q-1}}{2\sigma^2}} \right] \\
&= \frac{1}{2^q (2\pi)^{\frac{mN_s}{2}}} \left[\frac{1}{2\sigma^{mN_s+4}} \left(A_0 e^{\frac{-A_0}{2\sigma^2}} + A_1 e^{\frac{-A_1}{2\sigma^2}} + \ldots + A_{2^q-1} e^{\frac{-A_{2^q-1}}{2\sigma^2}} \right) \right. \\
&\quad \left. + \left(-\frac{mN_s}{2\sigma^{(mN_s+2)}} \right) \left(e^{\frac{-A_0}{2\sigma^2}} + e^{\frac{-A_1}{2\sigma^2}} + \ldots + e^{\frac{-A_{2^q-1}}{2\sigma^2}} \right) \right].
\end{aligned}$$

Equating $\frac{\mathrm{d}(f(\mathbf{r}))}{\mathrm{d}\sigma^2}$ to 0, we get

$$\hat{\sigma}^2_{\mathrm{ML}} = \frac{1}{mN_s} \left(\frac{A_0 e^{\frac{-A_0}{2\sigma^2}} + A_1 e^{\frac{-A_1}{2\sigma^2}} + \ldots + A_{2^q-1} e^{\frac{-A_{2^q-1}}{2\sigma^2}}}{e^{\frac{-A_0}{2\sigma^2}} + e^{\frac{-A_1}{2\sigma^2}} + \ldots + e^{\frac{-A_{2^q-1}}{2\sigma^2}}} \right). \tag{3}$$

A closed-form expression for $\hat{\sigma}^2_{\mathrm{ML}}$ cannot be obtained from equation (3). Approximations for small values of σ^2 and for large values of σ^2 do lead to a closed-form expressions, however. From equation (3), for small values of σ^2,

$$\hat{\sigma}^2_{\mathrm{ML}} \approx \hat{\sigma}^2_{small} = \frac{1}{mN_s} \min(A_0, A_1, \ldots, A_{2^q-1}).,\tag{4}$$

and for large values of σ^2,

$$\hat{\sigma}^2_{\mathrm{ML}} \approx \hat{\sigma}^2_{large} = \frac{1}{mN_S} \frac{A_0 + A_1 + \ldots + A_{2^q-1}}{2^q}. \tag{5}$$

Computationally efficient approximations

The approximation of $\hat{\sigma}^2_{\mathrm{ML}}$ using either equation (4) or equation (5) involves calculation of a number of terms that increases exponentially in the number of unknown code bits (q). Hence, although the expressions in equations (4) and equations (5) are conceptually simple, they are not calculated efficiently in the forms written above. Therefore, we obtain a different (equivalent) expression for σ^2_{ML} which leads to computationally efficient forms for the same small-variance and large-variance approximations.

Assuming all detected code words are correct, the joint density function of the demodulator outputs is given as

$$
\begin{aligned}
f(\mathbf{r}) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(r_{j_0}-(-1)^{\overline{b}_{j_0}}\right)^2}{2\sigma^2}} \prod_{k=j_0+1}^{j_0+p_0-1} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(r_k-(-1)^{\overline{b}_{j_0}\oplus\ldots\oplus\overline{b}_k}\right)^2}{2\sigma^2}} \right) \\
&\times \left[\prod_{l=0}^{\nu-1} \prod_{(k=j_l+p_l)}^{j_{(l+1)}-2} \left(\frac{1}{2\sqrt{2\pi\sigma^2}} e^{-\frac{(r_k-1)^2}{2\sigma^2}} + \frac{1}{2\sqrt{2\pi\sigma^2}} e^{-\frac{(r_k+1)^2}{2\sigma^2}} \right) \right] \\
&\times \left(\frac{1}{2\sqrt{2\pi\sigma^2}} e^{-\frac{(r_{j_\nu-1}-1)^2}{2\sigma^2}} + \frac{1}{2\sqrt{2\pi\sigma^2}} e^{-\frac{(r_{j_\nu-1}+1)^2}{2\sigma^2}} \right) \\
&\times \prod_{l=1}^{\nu-1} \left(\frac{1}{2\times 2\pi\sigma^2} e^{-\frac{\left(r_{j_l-1}-1\right)^2}{2\sigma^2}} e^{-\frac{\left(r_{j_l}-(-1)^{\overline{b}_{j_l}}\right)^2}{2\sigma^2}} \prod_{q=j_l+1}^{(j_l+p_l-1)} \frac{1}{2\sqrt{2\pi\sigma^2}} e^{-\frac{\left(r_q-(-1)^{\overline{b}_{j_l}\oplus\ldots\oplus\overline{b}_q}\right)^2}{\sigma^2}} \right. \\
&\quad \left. + \frac{1}{2\times 2\pi\sigma^2} e^{-\frac{\left(r_{j_l-1}+1\right)^2}{2\sigma^2}} e^{-\frac{\left(r_{j_l}+(-1)^{\overline{b}_{j_l}}\right)^2}{2\sigma^2}} \prod_{q=j_l+1}^{(j_l+p_l-1)} \frac{1}{2\sqrt{2\pi\sigma^2}} e^{-\frac{\left(r_q+(-1)^{\overline{b}_{j_l}\oplus\ldots\oplus\overline{b}_q}\right)^2}{\sigma^2}} \right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(r_{j_0}-(-1)^{\overline{b}_{j_0}}\right)^2}{2\sigma^2}} \prod_{k=j_0+1}^{j_0+p_0-1} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\left(r_k-(-1)^{\overline{b}_{j_0}\oplus\ldots\oplus\overline{b}_k}\right)^2}{2\sigma^2}} \right) \\
&\times \prod_{(k=j_0+p_0)}^{j_\nu-1} \left(\frac{1}{2\sqrt{2\pi\sigma^2}} e^{-\frac{\left((r_k)^2+1\right)}{2\sigma^2}} \right) \left[\prod_{l=0}^{\nu-1} \prod_{(k=j_l+p_l)}^{j_{(l+1)}-2} \left(e^{\frac{r_k}{\sigma^2}} + e^{-\frac{r_k}{\sigma^2}} \right) \right] \\
&\times \prod_{l=1}^{\nu-1} \left(e^{\frac{r_{j_l-1}+r_{j_l}(-1)^{\overline{b}_{j_l}}+\sum_{q=j_l+1}^{(j_l+p_l-1)} r_q(-1)^{\overline{b}_{j_l}\oplus\ldots\oplus\overline{b}_q}}{\sigma^2}} \right. \\
&\quad \left. + e^{-\frac{r_{j_l-1}+r_{j_l}(-1)^{\overline{b}_{j_l}}+\sum_{l=j_l+1}^{(j_l+p_l-1)} r_q(-1)^{\overline{b}_{j_l}\oplus\ldots\oplus\overline{b}_q}}{\sigma^2}} \right) \times \left(e^{\frac{r_{j_\nu-1}}{\sigma^2}} + e^{-\frac{r_{j_\nu-1}}{\sigma^2}} \right) \qquad (6)
\end{aligned}
$$

Let

$$
x_l = r_{j_l-1} + r_{j_l}(-1)^{\overline{b}_{j_l}} + \sum_{l=j_l+1}^{(j_l+p_l-1)} r_q(-1)^{\overline{b}_{j_l}\oplus\ldots\oplus\overline{b}_q}.
$$

Then taking the natural log on both sides of equation (6),

$$\begin{aligned}
\ln(f(\mathbf{d})) &= \ln(c) - \frac{mN_s}{2}\ln(\sigma^2) - \frac{\left(r_{j_0} - (-1)^{\bar{b}_{j_0}}\right)^2}{2\sigma^2} - \sum_{k=j_0+1}^{j_0+p_0-1} \frac{\left(r_k - (-1)^{\bar{b}_{j_0}\oplus\ldots\oplus\bar{b}_k}\right)^2}{2\sigma^2} \\
&- \sum_{(k=j_0+p_0)}^{(j_\nu-1)} \frac{\left((r_k)^2+1\right)}{2\sigma^2} + \sum_{l=0}^{\nu-1} \sum_{(k=j_l+p_l)}^{j_{(l+1)}-2} \ln\left(e^{\frac{r_k}{\sigma^2}} + e^{-\frac{r_k}{\sigma^2}}\right) \\
&+ \sum_{l=0}^{\nu-1} \ln\left(e^{\frac{x_l}{\sigma^2}} + e^{\frac{x_l}{\sigma^2}}\right) + \ln\left(e^{\frac{r_{j_\nu-1}}{\sigma^2}} + e^{-\frac{r_{j_\nu-1}}{\sigma^2}}\right). \qquad (7)
\end{aligned}$$

Taking the derivative of equation (7) with respect to σ^2 results in

$$\begin{aligned}
\frac{\mathrm{d}(\ln(f(\mathbf{r})))}{\mathrm{d}\sigma^2} &= -\frac{mN_s}{2\sigma^2} + \frac{\left(r_{j_0} - (-1)^{\bar{b}_{j_0}}\right)^2}{2\sigma^4} + \sum_{k=j_0+1}^{j_0+p_0-1} \frac{\left(r_k - (-1)^{\bar{b}_{j_0}\oplus\ldots\oplus\bar{b}_k}\right)^2}{2\sigma^4} \\
&+ \sum_{(k=j_0+p_0)}^{(j_\nu-1)} \frac{\left((r_k)^2+1\right)}{2\sigma^4} + \sum_{l=0}^{\nu-1} \sum_{(k=j_l+p_l)}^{j_{(l+1)}-2} \left[\left(\frac{r_k}{\sigma^4}\right)\left(\frac{e^{-\frac{r_k}{\sigma^2}} - e^{\frac{r_k}{\sigma^2}}}{e^{\frac{r_k}{\sigma^2}} + e^{-\frac{r_k}{\sigma^2}}}\right)\right] \\
&+ \sum_{l=0}^{\nu-1} \left[\left(\frac{x_l}{\sigma^4}\right)\left(\frac{e^{-\frac{x_l}{\sigma^2}} - e^{\frac{x_l}{\sigma^2}}}{e^{\frac{x_l}{\sigma^2}} + e^{-\frac{x_l}{\sigma^2}}}\right)\right] + \left(\frac{r_{j_\nu-1}}{\sigma^4}\right)\left(\frac{e^{-\frac{r_{j_\nu-1}}{\sigma^2}} - e^{\frac{r_{j_\nu-1}}{\sigma^2}}}{e^{\frac{r_{j_\nu-1}}{\sigma^2}} + e^{-\frac{r_{j_\nu-1}}{\sigma^2}}}\right). \qquad (8)
\end{aligned}$$

The value of σ^2 that makes the expression in equation (8) equal to 0 is the ML estimate of the variance, $\hat{\sigma}^2_{\mathrm{ML}}$. But a closed-form solution for $\hat{\sigma}^2_{\mathrm{ML}}$ can not be obtained due to the complexity of equation (8); numerical methods must be applied instead. Approximations can be made for large and small values of σ^2 to simplify equation (8) and obtain a closed-form solution, however.

If σ^2 is small, we use the approximation $\pm e^y - e^{-y} \approx \pm e^y$ for $y >> 0$. Thus equation

(8) can be approximated as

$$\begin{aligned}\frac{\mathrm{d}(\ln(f(\mathbf{r})))}{\mathrm{d}\sigma^2} &\approx -\frac{mN_s}{2\sigma^2}+\frac{\left(r_{j_0}-(-1)^{\bar{b}_{j_0}}\right)^2}{2\sigma^4}+\sum_{k=j_0+1}^{j_0+p_0-1}\frac{\left(r_k-(-1)^{\bar{b}_{j_0}\oplus\ldots\oplus\bar{b}_k}\right)^2}{2\sigma^4}\\ &+\sum_{(k=j_0+p_0)}^{(j_\nu-1)}\frac{\left((r_k)^2+1\right)}{2\sigma^4}-\sum_{l=0}^{\nu-1}\sum_{(k=j_l+p_l)}^{j_{(l+1)}-2}\left[\left(\frac{|r_k|}{\sigma^4}\right)\right]\\ &-\sum_{l=0}^{\nu-1}\left[\left(\frac{|x_l|}{\sigma^4}\right)\right]+\left(\frac{|r_{j_\nu-1}|}{\sigma^4}\right).\end{aligned}$$

Solving for σ^2, the approximate solution of equation (8) for $\hat{\sigma}^2_{\mathrm{ML}}$ is given by

$$\begin{aligned}\hat{\sigma}^2_{\mathrm{ML}}\approx\hat{\sigma}^2_{small} &= \frac{1}{mN_s}\left[\left(r_{j_0}-(-1)^{\bar{b}_{j_0}}\right)^2+\sum_{k=j_0+1}^{j_0+p_0-1}\left(r_k-(-1)^{\bar{b}_{j_0}\oplus\ldots\oplus\bar{b}_k}\right)^2\right.\\ &+\left.\sum_{(k=j_0+p_0)}^{(j_\nu-1)}\left((r_k)^2+1\right)-\sum_{l=0}^{\nu-1}\sum_{(k=j_l+p_l)}^{j_{(l+1)}-2}2\,|r_k|-\sum_{l=1}^{\nu-1}2|x_l|-2\,|r_{j_\nu-1}|\right]. \quad (9)\end{aligned}$$

Similarly if σ^2 is large, we use the approximation $e^{-y}-e^{y}\approx 0$ for $y\approx 0$. Under this approximation, equation (8) is reduced to

$$\begin{aligned}\frac{\mathrm{d}(\ln(f(\mathbf{r})))}{\mathrm{d}\sigma^2} &\approx -\frac{mN_s}{2\sigma^2}+\frac{\left(r_{j_0}-(-1)^{\bar{b}_{j_0}}\right)^2}{2\sigma^4}+\sum_{k=j_0+1}^{j_0+p_0-1}\frac{\left(r_k-(-1)^{\bar{b}_{j_0}\oplus\ldots\oplus\bar{b}_k}\right)^2}{2\sigma^4}\\ &+\sum_{(k=j_0+p_0)}^{(j_\nu-1)}\frac{\left((r_k)^2+1\right)}{2\sigma^4}.\end{aligned}$$

Solution of the approximating equation yields

$$\begin{aligned}\hat{\sigma}^2_{\mathrm{ML}}\approx\hat{\sigma}^2_{large} &= \frac{1}{mN_s}\left[\left(r_{j_0}-(-1)^{\bar{b}_{j_0}}\right)^2+\sum_{k=j_0+1}^{j_0+p_0-1}\left(r_k-(-1)^{\bar{b}_{j_0}\oplus\ldots\oplus\bar{b}_k}\right)^2\right.\\ &+\left.\sum_{(k=j_0+p_0)}^{(j_\nu-1)}\left((r_k)^2+1\right)\right]. \quad (10)\end{aligned}$$

Bibliography

[1] M. B. Pursley and W. E. Stark, "Performance of Reed-Solomon coded frequency-hop spread-spectrum communications in partial-band interference," *IEEE Trans. Commun.*, vol. 33, no. 8, pp. 767–774, Aug 1985.

[2] B. J. Hamilton, "SINCGARS system improvement program (SIP) specific radio improvements," in *Proc. Tactical Commun. Conf.* (Ft. Wayne, IN), 30 Apr.-2 May 1996, pp. 397–406.

[3] T. G. Macdonald and M. B. Pursley, "Staggered interleaving and iterative errors- and-erasures decoding for frequency-hop packet radio," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 92–98, Jan. 2003.

[4] A. W. Lam and D. V. Sarwate, "A comparison of two methods for generation of side information in frequency-hop spread-spectrum multi access communications," in *Proc. 21st Annual Conf. Inform. Sci. Syst.* (Allerton, IL), Mar. 1987, pp. 869–877.

[5] H. Ramchandran and D. L. Noneaker, "Iterative equalization and decoding for SFH spread-spectrum communications using Reed-Solomon codes," *Intl. J. Wireless Inform. Networks*, vol. 15, no. 1, pp. 1–15, Mar. 2008.

[6] C. D. Frank and M. B. Pursley, "Concatenated coding for frequency-hop spread-spectrum with partial-band interference," *IEEE Trans. Commun.*, vol. 44, no. 3, pp. 377–387, Mar. 1996.

[7] E. Paaske, "Improved decoding for a concatenated coding system recommended by CCSDS," *IEEE Trans. Commun.*, vol. 38, no. 8, pp. 1138–1144, Aug. 1990.

[8] H. Ramchandran and D. L. Noneaker, "Packet-level iterative detection for SFH communications with Reed-Solomon coding in partial-band interference," in *Proc. 2006 IEEE Military Commun. Conf.* (Washington, DC), Oct. 2006, pp. 2610–2616.

[9] G. D. Forney Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, no. 3, pp. 363–378, May 1972.

[10] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun. Technol.*, vol. 19, no. 5, pp. 751–772, Oct. 1971.

[11] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[12] G. Einarsson and C. E. Sundberg, "A note on soft decision decoding with successive erasures," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 88–96, Jan. 1976.

[13] G. D. Forney Jr., "Generalized minimum distance decoding," *IEEE Trans. Inform. Theory*, vol. 12, no. 2, pp. 125–131, Apr. 1966.

[14] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Proc. 1995 IEEE Intl. Conf. Commun.* (Seattle, WA), Jun. 1995, pp. 1009–1013.

[15] D. V. Mercy, "A review of automatic gain control theory," *The Radio and Electronic Engineer*, vol. 51, no. 11/12, pp. 579–590, Nov./Dec. 1981.

[16] K.-M. Cheung and R. J. McEliece, "The undetected error probability for Reed-Solomon codes," in *Proc. 1988 IEEE Military Commun. Conf.* (San Diego, CA), Oct. 1988, pp. 163–167.

[17] T. A. Summers and S. G. Wilson, "SNR mismatch and online estimation in turbo decoding," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 421–423, Apr. 1998.

[18] D. V. Sarwate and N. R. Shanbhag, "High-speed architectures for Reed-Solomon decoders," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 9, no. 5, pp. 641–655, Oct. 2001.

[19] Z. Yan, D. V. Sarwate, and Z. Liu, "High-speed systolic architectures for finite field inversion," *Integration, the VLSI Journal*, vol. 38, no. 3, pp. 383 – 398, Jan. 2005.

[20] D. V. Sarwate and Z. Yan, "Modified Euclidean algorithms for decoding Reed-Solomon codes," in *Proc. IEEE ISIT* (Seoul, Korea), Jul. 2009, pp. 1398–1402.

[21] W. Koch and A. Baier, "Optimum and sub-optimum detection of coded data disturbed by time-varying intersymbol interference [applicable to digital mobile radio receivers]," in *Proc. IEEE Global Telecommum. Conf.* (San Diego, CA), Dec. 1990, pp. 1679–1684.

[22] A. Worm, P. Hoeher, and N. Wehn, "Turbo-decoding without SNR estimation," *IEEE Commun. Lett.*, vol. 4, no. 6, pp. 193–195, Jun. 2000.