

**MASTER COPY:** PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

**MEMORANDUM OF TRANSMITTAL**

U.S. Army Research Office  
ATTN: AMSRL-RO-BI (TR)  
P.O. Box 12211  
Research Triangle Park, NC 27709-2211

Reprint (Orig + 2 copies)

Technical Report (Orig + 2 copies)

Manuscript (1 copy)

Final Progress Report (Orig + 2 copies)

Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER:

REPORT TITLE:

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

**REPORT DOCUMENTATION PAGE**Form Approved  
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)		2. REPORT DATE <b>01/26/2006</b>	3. REPORT TYPE AND DATES COVERED <b>Manuscripts: 01 June 2005 - 31 January 2006</b>	
4. TITLE AND SUBTITLE <b>Erasure Insertion for Coded DUSTM-FHSS Systems without A Priori Knowledge</b>			5. FUNDING NUMBERS <b>W911NF0410224</b>	
6. AUTHOR(S) <b>Haichang Sui, James R. Zeidler</b>				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>University of California - San Diego Office of Contract &amp; Grant Administration 9500 Gilman Dr. Mail Code 0934, La Jolla, CA, 92093-0934</b>			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) <b>U. S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211</b>			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.				
12 a. DISTRIBUTION / AVAILABILITY STATEMENT <b>Approved for public release; distribution unlimited.</b>			12 b. DISTRIBUTION CODE <b>N/A</b>	
13. ABSTRACT (Maximum 200 words)  In this paper, we consider coded frequency-hopping spread spectrum (FHSS) systems with differential unitary space-time modulation (DUSTM). We investigate the erasure insertion (EI) techniques for improving the decoding error probability, especially in presence of strong partial-band interference (PBI). Conditional likelihood functions are derived for the decision-feedback demodulation and used for EI. The two proposed erasure rules do not require knowledge of the distribution of the noise plus interference power, in contrast to their counterparts in [14]. Instead of assuming the unknown parameters involved in decision-feedback demodulation and EI to be known, a joint estimation-demodulation scheme based on the adaptive recursive least square (RLS) algorithm is proposed. The sensitivity of the proposed EI rules to mismatched thresholds is studied by simulation. Numerical results show that EI can improve the decoding error probability significantly in presence of PBI, even when the required parameters have to be estimated and suboptimal thresholds are used.				
14. SUBJECT TERMS  <b>N/A</b>			15. NUMBER OF PAGES <b>7</b>	
			16. PRICE CODE <b>N/A</b>	
17. SECURITY CLASSIFICATION OR REPORT <b>UNCLASSIFIED</b>	18. SECURITY CLASSIFICATION ON THIS PAGE <b>UNCLASSIFIED</b>	19. SECURITY CLASSIFICATION OF ABSTRACT <b>UNCLASSIFIED</b>	20. LIMITATION OF ABSTRACT  <b>UL</b>	

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)  
Prescribed by ANSI Std. Z39-18  
298-102

# Erasure Insertion for Coded DUSTM-FHSS Systems without A Priori Knowledge

Haichang Sui and James R. Zeidler

Dept. of Electrical & Computer Engineering, Univ. of California, San Diego, La Jolla, CA 92093-0407

Emails: hsui@ucsd.edu, zeidler@ece.ucsd.edu

**Abstract**—In this paper, we consider coded frequency-hopping spread spectrum (FHSS) systems with differential unitary space-time modulation (DUSTM). We investigate the erasure insertion (EI) techniques for improving the decoding error probability, especially in presence of strong partial-band interference (PBI). Conditional likelihood functions are derived for the decision-feedback demodulation and used for EI. The two proposed erasure rules do not require knowledge of the distribution of the noise plus interference power, in contrast to their counterparts in [14]. Instead of assuming the unknown parameters involved in decision-feedback demodulation and EI to be known, a joint estimation-demodulation scheme based on the adaptive recursive least square (RLS) algorithm is proposed. The sensitivity of the proposed EI rules to mismatched thresholds is studied by simulation. Numerical results show that EI can improve the decoding error probability significantly in presence of PBI, even when the required parameters have to be estimated and suboptimal thresholds are used.

## I. INTRODUCTION

Frequency-hopping spread spectrum (FHSS) has been widely considered in packet radio networks [10] [7] due to the desirable properties inherent in the waveform. These properties include anti-jamming capability, low probability of intercept, operability in non-continuous spectrum, and reduced sensitivity to near-far problems. These properties also make FHSS a good candidate for multiple-access in ad hoc networks, where TDMA or FDMA may be difficult to implement due to the lack of central control while DS-CDMA may suffer from the lack of accurate power control and/or the non-continuity of the available spectrum.

Previous studies on FHSS systems mainly focus on single transmit antenna and frequency-shift-keying (FSK) modulation, e.g. [6] [5] and references therein. One advantage of FSK modulation is that it can be noncoherently demodulated. This is important because hopping introduces additional time-variation of the channel which complicates accurate channel estimation. In [14], differential unitary space-time modulation (DUSTM) is employed in FHSS systems. The DUSTM effectively exploits the time-correlation of the fading process within a dwell to allow for noncoherent demodulation, while FSK relies on the orthogonality of its modulated signals. In addition to the spatial diversity from multiple transmit

antennas, DUSTM can offer higher spectral efficiency than FSK [14].

In [14], it is shown that properly designed erasure insertion (EI) can achieve significantly better decoding error probability compared to the performance without EI. Such performance gain comes from the fact that RS codes with a simple bounded distance decoder can correct twice as many erasures as errors. The quality a symbol estimate can be accessed by comparing the relative magnitudes of the likelihood function under different hypotheses [2] [1] [3] [8] [13] [14]. However, due to hopping, the power of the noise plus interference ( $\sigma_w^2$ ) in a dwell is random. Evaluation of the likelihood function involves averaging over the distribution of  $\sigma_w$ , which is assumed perfectly known in [2] [3] [8] [13] [14] but may not be available at the receiver in practice. To circumvent such difficulties, we study EI schemes using the likelihood function conditioned on the realization of  $\sigma_w$ . To be distinctive, we use the term “unconditional likelihood function” and “conditional likelihood function” for the likelihood function with and without averaging over the distribution of  $\sigma_w$ , respectively. The use of conditional likelihood function for EI was studied by Pursley et al [1] for FSK-FHSS systems.

This paper makes major extensions to the previous work [14] in the following aspects. General DUSTM constellations and decision-feedback demodulation (DFD) [4] [15] are considered in this paper, of which the specific modulation and demodulation scheme in [14] are special cases. We study EI based on the conditional likelihood function. The resulting EI rules do not require knowledge of the distribution of  $\sigma_w$ , but depend on the realization of  $\sigma_w$  and the time-correlation coefficients of the fading process. In this paper, an adaptive algorithm is proposed to jointly estimate the required parameters and demodulate the received signal. Block-based estimation of the unknown parameters is possible [1] but may be less flexible when they are time-varying. Furthermore, the performance of EI techniques depends crucially on the choice of the threshold. In [14], optimal thresholds are found numerically and used for EI. Here we quantify the sensitivity of EI to the choice of suboptimal thresholds. We show here that the two EI techniques similar to those in [14] but using the conditional likelihood function and adaptively estimated parameters with suboptimal thresholds can still significantly outperform the non-erasure decoding.

The organization of the paper is as follows. The system

This work is supported by, or in part by, the U. S. Army Research Office under the Multi-University Research Initiative (MURI) grant # W911NF-04-1-0224 and the Office of Naval Research, Code 313.

model is presented in Section II. DFD of DUSTM signal is reviewed in Section III based on the results from [15]. The Bayesian and the likelihood ratio test (LRT) rule for EI based on the conditional likelihood function are discussed in Section IV. Section V addresses estimation of the parameters required by demodulation and EI. Simulation results are presented in Section VI and Section VII concludes the paper.

## II. SYSTEM MODEL

In a FHSS system, the carrier frequency on which the base-band signal is transmitted changes periodically. Each carrier occupies a constant bandwidth  $\Delta f$ . The period of time during which the carrier frequency remains constant is defined as the dwell duration. The transmission of each symbol occupies a  $T_s \times \Delta f$  time-frequency slot. We consider slow FHSS systems where a number of symbols are transmitted in each dwell (e.g. see Fig. 2 in [14]). The sequence of carrier frequencies occupied by a user along time is called the hopping pattern. We assume that the elements in the hopping pattern for the user of interest are i.i.d. randomly chosen from all available carrier frequencies with equal probability, as in [2] [14].

The model for the transmitter and receiver are depicted in Fig.1. A RS encoder with rate  $k/L$  is used. The codeword length  $L$  is also the size of the DUSTM constellation  $\mathcal{V} = \{\mathbf{V}_0, \dots, \mathbf{V}_{L-1}\}$ , where  $\mathbf{V}_l$  is  $M \times M$  unitary matrices for  $l = 0, \dots, L-1$  and  $M$  is the number of transmit antennas. The interleaver is the same as in [2] [14], which ensures the elements from the same codeword are transmitted in different dwells. After interleaving, the  $L$ -ary coded symbols are naturally mapped to the corresponding matrices in  $\mathcal{V}$ . Without loss of generality, we consider the dwell starting at time 0 and denote the coded symbol to be transmitted in the  $[(\tau-1)T_s, \tau T_s]$  as  $z_\tau$ . The actual transmitted space-time signal is given by  $\mathbf{S}_\tau = \mathbf{V}_{z_\tau} \mathbf{S}_{\tau-1}$ . The  $M$  columns of  $\mathbf{S}_\tau$  are fed to different transmit antennas while the elements in the same row of  $\mathbf{S}_\tau$  are transmitted at the same time. The initial transmission  $\mathbf{S}_0$  does not carry information and can be an arbitrary unitary matrix. It can be shown that DUSTM can offer higher spectral efficiency than FSK modulation [16]. In a FHSS multiple-access network with a total bandwidth constraint, lower spectral efficiency translates into fewer available frequency slots (or lower spreading gain) and consequently more significant MAI effect.

The receiver employs  $N$  antennas. We assume the coherent time and bandwidth of the fading are greater than  $T_s$  and  $\Delta f$ , respectively. Thus, the fading in each  $T_s \times \Delta f$  time-frequency slot can be viewed as constant and frequency-flat. We further assume the fading processes between different transmit-receive antenna pairs are statistically i.i.d. and circularly complex Gaussian with zero mean. The discrete channel gains between all transmit-receive antenna pairs during  $[(\tau-1)T_s, \tau T_s]$  are collected into an  $M \times N$  matrix  $\mathbf{H}_\tau$ . The time-correlation coefficients between the channel gains are defined as  $\varphi_{h,\xi} = E\{h_{\tau+\xi,i,j} \cdot h_{\tau,i,j}^*\}$  for  $\xi = 0, \pm 1, \dots$ . Without loss of generality, we normalize  $\varphi_{h,0}$  to unity. The received space-

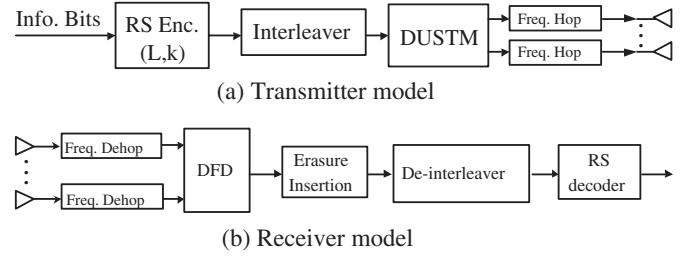


Fig. 1. System model

time signal in  $[(\tau-1)T_s, \tau T_s]$  can be written as

$$\mathbf{R}_\tau = \mathbf{S}_\tau \mathbf{H}_\tau + \mathbf{W}_\tau. \quad (1)$$

The entries in  $\mathbf{W}_\tau$  consist of interference in addition to the thermal noise. In this paper, we focus on the partial-band interference (PBI) and ignore other types of interference including the multiple-access interference. As assumed in [2] and [14], the entries in  $\mathbf{W}_\tau$  are i.i.d. zero-mean circularly complex Gaussian random variables (rv's) with variance  $\sigma_0$  ( $\mathcal{CN}(0, \sigma_0^2)$ ) in a fraction  $(1-p_I)$  of the band and are i.i.d.  $\mathcal{CN}(0, \sigma_0^2 + \sigma_I^2/p_I)$  rv's in the remainder of the band. Therefore, conditioned on the variance of the noise plus PBI  $\sigma_w$ ,  $\mathbf{W}_\tau$  contains  $\mathcal{CN}(0, \sigma_w^2)$  rv's. Entries corresponding to different  $\tau$ s are assumed to be uncorrelated. Due to the random hopping,  $\sigma_w$  can be viewed as a rv independent in different dwells and identically distributed as

$$f\sigma_w(\sigma) = p_I \delta(\sigma_0^2 + \sigma_I^2/p_I - \sigma^2) + (1-p_I) \delta(\sigma_0^2 - \sigma^2), \quad (2)$$

where  $\delta(x)$  equals 1 if  $x=0$  and 0 otherwise.

In general, the time-correlation coefficients  $\{\varphi_{h,\xi}\}_\xi$  depend on the normalized Doppler frequency shift  $f_d T_s$ , which may be different for different carrier frequencies. However, since the range of carrier frequencies is normally much smaller than their values, the variation in  $f_d T_s$  and the time-correlation coefficients can be ignored. For example, if the center frequency in the two ISM frequency bands 902–928MHz and 2.4–2.4835GHz corresponds to  $f_d T_s = 0.05$ , then the normalized Doppler shifts experienced by the carriers at the edges of the two bands will be  $0.05 \pm 0.0007$  and  $0.05 \pm 0.001$ , respectively. For simplicity, we regard  $\{\varphi_{h,\xi}\}_\xi$  as deterministic unknown parameters common for all dwells. The two EI techniques, however, can be easily extended to account for the variation in  $\{\varphi_{h,\xi}\}_\xi$  due to random hopping by treating them as a set of (correlated) rv's in addition to  $\sigma_w$  [16]. The EI techniques proposed in Section IV and the estimation scheme in Section V are independent of the distribution of  $\sigma_w$  and/or  $\{\varphi_{h,\xi}\}_\xi$ , which allows them to be applied in more general settings.

## III. DECISION-FEEDBACK DEMODULATION (DFD)

In this section, we review the DFD for the DUSTM signal based on the results from [15]. Without loss of generality, we consider the dwell starting from time 0. We assume  $\sigma_w$  and  $\{\varphi_{h,\xi}\}_\xi$  are known and time-invariant in deriving the DFD test statistics. Estimation of these parameters is discussed in

Section V and the resulting algorithm is applicable even when these parameters are time-varying.

We start from the likelihood function of receiving  $\mathbf{R}_0, \dots, \mathbf{R}_\tau$  given  $z_1, \dots, z_\tau$  are transmitted, that is,

$$f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_\tau^1) = \det^{-N}(\pi \mathbf{C}_{r,\tau}) \exp \left\{ -\text{tr} \left[ (\mathbf{R}_\tau^0)^H \mathbf{C}_{r,\tau}^{-1} \mathbf{R}_\tau^0 \right] \right\}. \quad (3)$$

Here we use the notation  $\mathbf{R}_\tau^0$  and  $z_\tau^1$  to denote  $[\mathbf{R}_0^H, \dots, \mathbf{R}_\tau^H]^H$  and  $(z_1, \dots, z_\tau)$ , where the superscript  $(\cdot)^H$  denotes matrix Hermitian. The subscript in  $f_{|\sigma_w}$  implies the dependence on  $\sigma_w$ , while the unconditional likelihood function can be obtained by taking expectation over the distribution of  $\sigma_w$ , that is,

$$f(\cdot | \cdot) = \mathbb{E} \{ f_{|\sigma_w}(\cdot | \cdot) \}. \quad (4)$$

The  $M(\tau+1) \times M(\tau+1)$  matrix  $\mathbf{C}_{r,\tau}$  in (3) is the covariance matrix of any column of the  $M(\tau+1) \times N$  matrix  $\mathbf{R}_\tau^0$ . It can be shown that  $c_{|\sigma_w}^{(1)} \triangleq \det^N(\pi \mathbf{C}_{r,\tau})$  does not depend on  $z_\tau^1$  and the exponent in  $f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_\tau^1)$  can be expressed as [15]

$$\gamma_{\tau|\sigma_w}(z_\tau^1) \triangleq -\text{tr} \left[ (\mathbf{R}_\tau^0)^H \mathbf{C}_{r,\tau}^{-1} \mathbf{R}_\tau^0 \right] \quad (5)$$

$$= \gamma_{\tau-1|\sigma_w}(z_{\tau-1}^1) + \Delta_{|\sigma_w}(z_\tau; z_{\tau-1}^1). \quad (6)$$

The second term in (6) can be further reduced to [15]

$$\Delta_{|\sigma_w}(z_\tau; z_{\tau-1}^1) = \frac{1}{\sigma_\tau^2} \left\| \left( \Lambda_1^\tau \right)^H \mathbf{R}_\tau - \sum_{j=1}^{\tau} a_j^{(\tau)} \left( \Lambda_1^{\tau-j} \right)^H \mathbf{R}_{\tau-j} \right\|^2, \quad (7)$$

where  $\|\cdot\|$  denotes the Frobenius norm of matrices and  $\Lambda_t^\tau = \mathbf{V}_{z_\tau} \cdot \mathbf{V}_{z_{\tau-1}} \cdots \mathbf{V}_{z_t}$  for  $\tau \geq t$ . The coefficients  $\{a_j^{(\tau)}\}_{j=1}^P$  and  $\sigma_\tau$  can be solved from  $\{\varphi_{h,\xi}\}_{\xi=0}^\tau$  and  $\sigma_w$  [15]. In (6),  $\gamma_{\tau|\sigma_w}(z_\tau^1)$  and  $\gamma_{\tau-1|\sigma_w}(z_{\tau-1}^1)$  can be viewed as the accumulated likelihood metrics for the sequence  $z_\tau^1$  and  $z_{\tau-1}^1$ , respectively, while  $\Delta_{|\sigma_w}(z_\tau; z_{\tau-1}^1)$  can be viewed as the branch metric. Unfortunately, their dependence on all previous transmitted symbols prevents efficient sequence estimation like the standard Viterbi algorithm. However, if we assume that the scalar processes  $\{h_{t,m,n} + w_{t,m,n}\}_t$  can be modelled by an AR( $P$ ) process, it is shown in [15] that for  $\tau \geq P$ , (7) reduces to

$$\Delta_{|\sigma_w}(z_\tau; z_{\tau-1}^1) = \frac{1}{\sigma_P^2} \left\| \mathbf{V}_{z_\tau}^H \mathbf{R}_\tau - \sum_{j=1}^P a_j^{(P)} \Lambda_{\tau-j+1}^{\tau-1} \mathbf{R}_{\tau-j} \right\|^2 \triangleq \Delta_{|\sigma_w}(z_\tau; z_{\tau-1}^{\tau-P+1}). \quad (8)$$

Thus, the branch metric  $\Delta_{|\sigma_w}(z_\tau; z_{\tau-1}^1)$  only depends on  $z_{\tau-1}^{\tau-P+1}$  instead of  $z_{\tau-1}^1$  and the standard Viterbi algorithm can be applied.

The Viterbi algorithm has a complexity which is exponential in  $L$  since the number of states in the trellis is  $L^{P-1}$ . In order to further reduce the complexity and facilitate symbol-by-symbol erasure insertion, we consider a DFD scheme which has an approximate complexity of  $O(L)$  per symbol. In DFD,  $P-1$  previous decisions  $\hat{z}_{\tau-P+1}, \dots, \hat{z}_{\tau-1}$  are fed back and

$P+1$  received blocks  $\mathbf{R}_{\tau-P}, \dots, \mathbf{R}_\tau$  are used to demodulate  $z_\tau$ . Specifically,  $z_{\tau-P+1}, \dots, z_{\tau-1}$  in (8) are replaced by  $\hat{z}_{\tau-P+1}, \dots, \hat{z}_{\tau-1}$  and  $\hat{z}_\tau$  is demodulated as

$$\hat{z}_\tau = \arg \max_{k=0, \dots, L-1} \Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1}). \quad (9)$$

The symbol error probability of  $\hat{z}_\tau$  in (9) is analyzed in [15] under the assumption that  $\sigma_w$ ,  $\{\varphi_{h,\xi}\}_\xi$  are known and no feedback errors. It is also shown there that the ML demodulation of  $z_\tau$  based on  $\mathbf{R}_\tau^{\tau-1}$  in [14] is a special case of the DFD if  $P$  is set to be 1.

#### IV. ERASURE INSERTION DECODING

A general EI rule is proposed in [2] based on the Bayesian framework to minimize a Chernoff bound on the decoded error probability. The resulting rule involves  $f(\mathbf{X}_\tau | k)$  ( $k = 0, \dots, L-1$ ), the likelihood of  $\mathbf{X}_\tau$  under all possible hypotheses for  $z_\tau$ , where  $\mathbf{X}_\tau$  consists the data used by the demodulator. Various EI rules have been proposed by specializing the general rule in [2] to demodulators in different settings [1] [3] [8] [13] [14]. In this section, we propose a Bayesian EI rule for the decision-feedback demodulator considered in III. We further propose a so-called Likelihood Ratio Test (LRT) rule which can be viewed as an approximation of the Bayesian rule but has implemental advantage.

##### A. Likelihood for erasure insertion

The choice of the data matrix  $\mathbf{X}_\tau$  in the likelihood function  $f(\mathbf{X}_\tau | k)$  depends on the demodulator under consideration. For the DFD discussed in III,  $\mathbf{R}_\tau^{\tau-P}$  is used and previous decisions  $\hat{z}_{\tau-1}^{\tau-P+1}$  are fed back for demodulating  $z_\tau$ . Therefore, it is natural to consider the likelihood function  $f(\mathbf{R}_\tau^{\tau-P} | k, \hat{z}_{\tau-1}^{\tau-P+1})$ . However, as can be seen from (4), the unconditional likelihood function  $f(\mathbf{R}_\tau^{\tau-P} | k, \hat{z}_{\tau-1}^{\tau-P+1})$  is hard to obtain because the distribution of  $\sigma_w$  may not be available at the receiver and the complication involved in evaluating the expectation in (4). It is therefore more desirable to consider the conditional likelihood function  $f_{|\sigma_w}(\mathbf{R}_\tau^{\tau-P} | k, \hat{z}_{\tau-1}^{\tau-P+1})$ . As pointed out in [1], the use of the conditional likelihood function can be justified by regarding  $\sigma_w$  as a received statistic in addition to  $\mathbf{R}_\tau^{\tau-P}$ . Specifically, we consider  $f(\mathbf{R}_\tau^{\tau-P}, \sigma_w | k, \hat{z}_{\tau-1}^{\tau-P+1})$  instead of  $f(\mathbf{R}_\tau^{\tau-P} | k, \hat{z}_{\tau-1}^{\tau-P+1})$ . It is straightforward to see that

$$f(\mathbf{R}_\tau^{\tau-P}, \sigma_w | k, \hat{z}_{\tau-1}^{\tau-P+1}) = f_{|\sigma_w}(\mathbf{R}_\tau^{\tau-P} | k, \hat{z}_{\tau-1}^{\tau-P+1}) f(\sigma_w), \quad (10)$$

where we have assumed  $f(\sigma_w | k, \hat{z}_{\tau-1}^{\tau-P+1}) = f(\sigma_w)$ . The conditional likelihood  $f_{|\sigma_w}(\mathbf{R}_\tau^{\tau-P} | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1})$  will be derived in the sequel.

From (3), (5), (6), and (8), we can easily obtain for  $\tau > P$

$$f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_{\tau-P}^1, \hat{z}_{\tau-1}^{\tau-P+1}, z_\tau) = \exp[\Delta_{|\sigma_w}(z_\tau; \hat{z}_{\tau-1}^{\tau-P+1})] \times \exp[\gamma_{\tau-1|\sigma_w}(z_{\tau-P}^1, \hat{z}_{\tau-1}^{\tau-P+1})] / c_{|\sigma_w}^{(1)} \quad (11)$$

under the assumption that the processes  $\{h_{t,m,n} + w_{t,m,n}\}_t$  can be statistically modelled by a common AR( $P$ ) process for all different  $(m, n)$  pairs. Given  $\hat{z}_{\tau-1}^{\tau-P+1}$ , the probability of receiving  $\mathbf{R}_\tau^0$  conditioned on  $z_\tau$  alone is given by

$$f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}) = \sum_{z_1, \dots, z_{\tau-P}} f_{|\sigma_w}(z_{\tau-P}^1 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}) \\ \times f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_{\tau-P}^1, \hat{z}_{\tau-1}^{\tau-P+1}, z_\tau), \quad (12)$$

where the summation is over all  $L^{\tau-P}$  possible combinations of  $(z_1, \dots, z_{\tau-P})$ . It is not hard to see that  $f_{|\sigma_w}(z_{\tau-P}^1 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}) = L^{P-\tau}$  since the transmitted symbols in a dwell can be treated as i.i.d. rv's due to interleaving. Consequently, we obtain

$$f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}) = c_{|\sigma_w}^{(2)}(\mathbf{R}_\tau^0, \hat{z}_{\tau-1}^{\tau-P+1}) \\ \times \exp[\Delta_{|\sigma_w}(z_\tau; \hat{z}_{\tau-1}^{\tau-P+1})] \quad (13)$$

where

$$c_{|\sigma_w}^{(2)}(\mathbf{R}_\tau^0, \hat{z}_{\tau-1}^{\tau-P+1}) \triangleq \frac{\sum_{z_1, \dots, z_{\tau-P}} \exp[\gamma_{\tau-1|\sigma_w}(z_{\tau-P}^1, \hat{z}_{\tau-1}^{\tau-P+1})]}{L^{\tau-P} \cdot c_{|\sigma_w}^{(1)}} \quad (14)$$

Note that  $c_{|\sigma_w}^{(2)}(\mathbf{R}_\tau^0, \hat{z}_{\tau-1}^{\tau-P+1})$  in (14) does not depend on  $z_\tau$ .

Alternatively, we can express  $f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1})$  as

$$f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}) = f_{|\sigma_w}(\mathbf{R}_\tau^{\tau-P} | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}) \\ \times f_{|\sigma_w}(\mathbf{R}_{\tau-P-1}^0 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}, \mathbf{R}_\tau^{\tau-P}). \quad (15)$$

Since  $z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}$  are transmitted after  $\mathbf{R}_{\tau-P-1}^0$  being received, the first conditional probability in (15) does not depend on  $z_\tau$  or  $\hat{z}_{\tau-1}^{\tau-P+1}$ . That is,

$$f_{|\sigma_w}(\mathbf{R}_{\tau-P-1}^0 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}, \mathbf{R}_\tau^{\tau-P}) = f_{|\sigma_w}(\mathbf{R}_{\tau-P-1}^0 | \mathbf{R}_\tau^{\tau-P}), \quad (16)$$

where the dependence in the right hand side of (16) is due to the channel's time-correlation. Consequently, if we compare the likelihood  $f_{|\sigma_w}(\mathbf{R}_\tau^0 | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1})$  obtained from (15) and (16) with (13), we get

$$f_{|\sigma_w}(\mathbf{R}_\tau^{\tau-P} | z_\tau, \hat{z}_{\tau-1}^{\tau-P+1}) = c_{|\sigma_w}^{(3)}(\mathbf{R}_\tau^0, \hat{z}_{\tau-1}^{\tau-P+1}) \\ \times \exp[\Delta_{|\sigma_w}(z_\tau; \hat{z}_{\tau-1}^{\tau-P+1})] \quad (17)$$

where

$$c_{|\sigma_w}^{(3)}(\mathbf{R}_\tau^0, \hat{z}_{\tau-1}^{\tau-P+1}) = c_{|\sigma_w}^{(2)}(\mathbf{R}_\tau^0, \hat{z}_{\tau-1}^{\tau-P+1}) f_{|\sigma_w}^{-1}(\mathbf{R}_{\tau-P-1}^0 | \mathbf{R}_\tau^{\tau-P}). \quad (18)$$

From (10) and (17), we obtain

$$f(\mathbf{R}_\tau^{\tau-P}, \sigma_w | k, \hat{z}_{\tau-1}^{\tau-P+1}) = f(\sigma_w) c_{|\sigma_w}^{(3)}(\mathbf{R}_\tau^0, \hat{z}_{\tau-1}^{\tau-P+1}) \\ \times \exp[\Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1})]. \quad (19)$$

The multiplicative factor  $f(\sigma_w) c_{|\sigma_w}^{(3)}(\mathbf{R}_\tau^0, \hat{z}_{\tau-1}^{\tau-P+1})$  in (19) is common for all  $k = 0, \dots, L-1$  and, as will be shown later, has no effect on EI.

## B. Bayesian erasure insertion

Assuming the transmitted symbols take values in  $\{0, \dots, L-1\}$  with equal *a priori* probabilities, the decision rule in [2] suggests that a symbol estimate  $\hat{z}_\tau$  should be erased if

$$\frac{f(\mathbf{X}_\tau | l_{ML})}{\sum_{k=0}^{L-1} f(\mathbf{X}_\tau | k)} \leq 1 - \theta, \quad (20)$$

and  $\hat{z}_\tau$  should be chosen as  $l_{ML} = \arg \max_{k=0}^{L-1} f(\mathbf{X}_\tau | k)$  otherwise. The quantity  $f(\mathbf{X}_\tau | k)$  is the likelihood of receiving  $\mathbf{X}_\tau$  provided  $z_\tau = k$  is transmitted. Such a rule is shown to minimize the quantity  $P_e + \theta P_a$  where  $P_e$  and  $P_a$  are the probabilities with which the receiver processing block outputs an error or erasure given  $\mathbf{X}_\tau$  is available, respectively. It is also shown in [2] that the decoded error probability  $P_E$  satisfies

$$P_E = \sum_{i=0}^n \sum_{j=j_0(i)}^{n-i} \frac{n! P_e^i P_a^j}{i! j! (n-i-j)!} (1 - P_e - P_a)^{n-i-j} \quad (21)$$

$$\leq \min_{0 < \theta < 1/2} \beta^{k-n} \sum_{i=0}^n \binom{n}{i} (\beta^2 - 1)^i (P_e + \theta P_a)^i \quad (22)$$

where  $\beta = \theta^{-1} - 1$ . Therefore, the decision rule (20) is optimal in the sense of minimizing the bound on  $P_E$ .

For the DFD of DUSTM signals, we may set  $f(\mathbf{X}_\tau | k)$  to  $f(\mathbf{R}_\tau^{\tau-P}, \sigma_w | k, \hat{z}_{\tau-1}^{\tau-P+1})$  from the discussion in Section IV-A. By replacing  $f(\mathbf{X}_\tau | k)$  with (19), equation (20) reduces to

$$\frac{\exp[\Delta_{|\sigma_w}(\hat{z}_\tau; \hat{z}_{\tau-1}^{\tau-P+1})]}{\sum_{k=0}^{L-1} \exp[\Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1})]} \leq 1 - \theta. \quad (23)$$

where

$$\hat{z}_\tau = \arg \max_k f(\mathbf{R}_\tau^{\tau-P}, \sigma_w | k, \hat{z}_{\tau-1}^{\tau-P+1}) \\ = \arg \max_k \Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1}) \quad (24)$$

is the same as in (9). The Bayesian rule inserts an erasure if (23) is satisfied; otherwise  $\hat{z}_\tau$  is sent to the decoder.

## C. Likelihood ratio test

The Bayesian rule (23) requires the exact value of  $\exp[\Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1})]$  for  $k = 0, \dots, L-1$ . In this section, we propose an approximation to (23) where only the value of  $\Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1})$  is needed. This is preferable in implementation because it may be hard to compute  $\exp[\Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1})]$  accurately due to its small magnitude. The proposed EI rule can be viewed as a threshold test for the ratio of two likelihood functions. We will refer it as the likelihood ratio test (LRT) rule in the following.

The LRT rule essentially approximates the sum in the denominator in (23) by the first two largest terms, that is, the two terms corresponding to  $k = \hat{z}_\tau, \hat{z}'_\tau$  where  $\hat{z}_\tau = \arg \max_k \Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1})$  and  $\hat{z}'_\tau = \arg \max_{k \neq \hat{z}_\tau} \Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1})$ . The LRT rule is to erase  $\hat{z}_\tau$  if

$$\Delta_{|\sigma_w}(\hat{z}_\tau; \hat{z}_{\tau-1}^{\tau-P+1}) - \Delta_{|\sigma_w}(\hat{z}'_\tau; \hat{z}_{\tau-1}^{\tau-P+1}) \leq \ln \mu^{-1}, \quad (25)$$

for a specific threshold  $\mu \in (0, 1)$ .

From (25) we can see that the LRT rule only requires the value of  $\Delta_{|\sigma_w}(z_\tau; \hat{z}_{\tau-1}^{\tau-P+1})$  while the Bayesian rule (23) requires  $\exp(\Delta_{|\sigma_w}(z_\tau; \hat{z}_{\tau-1}^{\tau-P+1}))$ , which could be hard to compute accurately due to its small magnitude.

## V. ESTIMATING THE STATISTICS BY RLS

We have so far assumed perfect knowledge of  $\sigma_w$  and  $\{\varphi_{h,\xi}\}_\xi$  in deriving the demodulation and EI. We observe from (9) (23), (25) that it suffices to estimate  $\{a_1^{(P)}, \dots, a_P^{(P)}, \sigma_P^2\}$  for the purpose of demodulation or EI. In this section, we propose a simple adaptive scheme based on the RLS algorithm for jointly estimating  $\{a_1^{(P)}, \dots, a_P^{(P)}, \sigma_P^2\}$  and demodulating the symbols. Details of derivations and discussions are not presented due to space limit but can be found in [16].

To facilitate the application of the RLS algorithm, we observe that the  $n$ th column of  $\mathbf{V}_k^H \mathbf{R}_\tau - \sum_{j=1}^P a_j^{(P)} \mathbf{\Lambda}_{\tau-j+1}^{\tau-1} \mathbf{R}_{\tau-j}$  can be written as

$$\mathbf{\Lambda}_{\tau-P+1}^{\tau-1} \left[ \mathbf{h}'_{\tau,n} + \mathbf{w}'_{\tau,n} - \sum_{j=1}^P a_j^{(P)} (\mathbf{h}'_{\tau-j,n} + \mathbf{w}'_{\tau-j,n}) \right] \quad (26)$$

where  $\mathbf{h}'_{\tau-j,n} = \mathbf{S}_{\tau-P} \mathbf{h}_{\tau-j,n}$  and  $\mathbf{w}'_{\tau-j,n} = (\mathbf{\Lambda}_{\tau-P+1}^{\tau-j})^H \mathbf{w}_{\tau-j,n}$  for  $j = 0, \dots, P$ . The new vector sequences  $\{\mathbf{h}'_{\tau-j,n}\}_{j=0}^P$  and  $\{\mathbf{w}'_{\tau-j,n}\}_{j=0}^P$  have the same time-correlation as the corresponding original sequences. The term in the bracket in (26) can be viewed as the  $P$ th order prediction error of the vector process  $\{\mathbf{h}'_{\tau-j,n} + \mathbf{w}'_{\tau-j,n}\}_{j=0}^P$ . By recalling  $\mathbf{\Lambda}_{\tau-P+1}^{\tau-1}$  is unitary, we know that the solution of the Yule-Walker equations for  $\{a_j^{(P)}\}_{j=1}^P$  will minimize the expectation of (26) and the resulting minimal mean-squared error will be equal to  $N\sigma_P^2$ .

In practice, we can feedback the previous estimates  $\hat{z}_{\tau-P+1}, \dots, \hat{z}_{\tau-1}$  and use them in place of  $z_{\tau-P+1}, \dots, z_{\tau-1}$ . The unavailability of  $z_\tau$  may be circumvented by joint demodulation and estimation. That is, we use the estimated values for  $\{a_1^{(P)}, \dots, a_P^{(P)}\}$  together with  $\mathbf{R}_\tau^{\tau-P}$  and  $\hat{z}_{\tau-1}^{\tau-P+1}$  to demodulate  $z_\tau$ . After obtaining  $\hat{z}_\tau$ , the estimates for  $\{a_1^{(P)}, \dots, a_P^{(P)}, \sigma_P^2\}$  are updated even if  $\hat{z}_\tau$  might be incorrect or erased later. We observe in simulation that the proposed joint estimation-demodulation algorithm works satisfactorily if it is initialized properly with a small number of training symbols.

We extend the standard RLS algorithm for scalars to vector processes and apply it for our problem. Specifically, we denote

$$\mathbf{c}(\tau) = \begin{bmatrix} a_P^{(P)}(\tau) & \dots & a_1^{(P)}(\tau) \end{bmatrix}^T, \quad (27)$$

$$\mathbf{y}(\tau) = \text{vec}(\mathbf{V}_{\hat{z}_\tau}^H \mathbf{R}_\tau) \in \mathbb{C}^{MN \times 1}, \quad (28)$$

and  $\mathbf{X}(\tau)$  as given in (29), where  $\text{vec}(\cdot)$  is the vectorization operator which transforms a  $M \times N$  matrix into a  $MN \times 1$  vector by concatenating its columns. The exponentially-weighted average MSE  $E(\tau)$  is defined by

$$E(\tau) = \frac{1-\lambda}{MN} \sum_{i=0}^{\tau} \lambda^{\tau-i} \|\mathbf{y}(i) - \mathbf{X}(i) \mathbf{c}(i)\|^2. \quad (30)$$

The RLS algorithm together with symbol demodulation is described as following.

(i) Initialization: Set  $\tau = 0$ ,  $\mathbf{c}(-1) = \mathbf{0}_P$ , and  $\mathbf{P}(-1) = \delta^{-1} \mathbf{I}_P$  for some small positive number  $\delta$ . We choose  $E(-1) = \delta$  for convenience but in general  $E(-1)$  can be arbitrarily chosen.

(ii) Demodulation: Compute the test statistics  $\Delta_{|\sigma_w}(k; \hat{z}_{\tau-1}^{\tau-P+1})$  as in (8) where  $\sigma_P^2$  and  $a_P^{(P)}, \dots, a_1^{(P)}$  are replaced by  $E(\tau-1)$  and entries in  $\mathbf{c}(\tau-1)$ , respectively.  $z_\tau$  is then chosen as in (9). When a training symbol is transmitted at time  $\tau$ ,  $z_\tau$  is perfectly known at the receiver and no demodulation is performed.

(iii) RLS updates: Compute  $\mathbf{y}(\tau)$ ,  $\mathbf{X}(\tau)$  and

$$\bar{\mathbf{G}}(\tau) = \frac{1}{\lambda} \mathbf{P}(\tau-1) \mathbf{X}^H(\tau),$$

$$\mathbf{G}(\tau) = \bar{\mathbf{G}}(\tau) [\mathbf{I}_M + \mathbf{X}(\tau) \bar{\mathbf{G}}(\tau)]^{-1},$$

$$\mathbf{P}(\tau) = \frac{1}{\lambda} \mathbf{P}(\tau-1) - \mathbf{G}(\tau) \bar{\mathbf{G}}^H(\tau),$$

$$\mathbf{e}(\tau) = \mathbf{y}(\tau) - \mathbf{X}(\tau) \mathbf{c}(\tau-1),$$

$$\mathbf{c}(\tau) = \mathbf{c}(\tau-1) + \mathbf{G}(\tau) \mathbf{e}(\tau),$$

$$\boldsymbol{\epsilon}(n) = \mathbf{y}(n) - \mathbf{X}(n) \mathbf{c}(n),$$

$$MSE(\tau) = \frac{1}{MN} \|\mathbf{e}(\tau)\|^2, \quad (31)$$

$$E(\tau) = \lambda E(\tau-1) + \frac{1-\lambda}{MN} \mathbf{e}^H(\tau) \mathbf{e}(\tau). \quad (32)$$

## VI. NUMERICAL RESULTS

In this section, numerical results are presented to quantify the gain from EI. The simulation settings are as follows. The channels between different transmit-receive antenna pairs are generated independently using Jakes' model [9] with the normalized Doppler frequency shift  $f_d T_s = 0.05$ . The transmitter has two antennas and the DUSTM constellation in [12] is used with constellation size  $L$  equal to 16. RS(16,4) codes are considered. The symbol error and erasure probability  $P_e(P, N, \theta, \sigma_w)$  and  $P_a(P, N, \theta, \sigma_w)$  are simulated for different combinations of  $P$ ,  $N$ ,  $\sigma_w$ , and  $\theta$  (for LRT rule, the notation for the threshold is replaced by  $\mu$ ). The average symbol error and erasure probability are obtained by taking expectation of  $P_e(P, N, \theta, \sigma_w)$  and  $P_a(P, N, \theta, \sigma_w)$  over the distribution of  $\sigma_w$  in (2), that is,  $P_e(P, N, \theta, \sigma_0, \sigma_I, p_I) = \mathbb{E}\{P_e(P, N, \theta, \sigma_w)\}$  and  $P_a(P, N, \theta, \sigma_0, \sigma_I, p_I) = \mathbb{E}\{P_a(P, N, \theta, \sigma_w)\}$ . The decoding error probability  $P_E(P, N, \theta, \sigma_0, \sigma_I, p_I)$  is evaluated by (21) where  $P_e$  is replaced by  $P_e(P, N, \theta, \sigma_0, \sigma_I, p_I)$  and similarly with  $P_a$ . The decoding error probability without EI can be obtained by either setting  $\theta$  in the Bayesian rule to 0 or  $\mu$  in the LRT rule to 1.

### A. Estimation Performance

In this section, we present numerical results of estimating  $\sigma_P^2$  and  $\{a_j^{(P)}\}_{j=1}^P$  by the proposed joint estimation-demodulation algorithm. We set  $P = 2$ ,  $\frac{1}{\sigma_w^2} = 0\text{dB}$  and plot the learning curves of the two RLS weights (real parts only) and  $E(\tau)$ ,  $MSE(\tau)$  in Fig.2. The learning curves in Fig.2 show good convergence to their theoretical values. It is expected that  $E(\tau)$  converges to  $\sigma_P^2$  more slowly than  $MSE(\tau)$  due to the exponentially weighting. In practice, a

combination of  $E(\tau)$  and  $MSE(\tau)$  or the local mean of  $MSE(\tau)$  can be used to better acquire and track  $\sigma_P^2$  in a time-varying environment.

As discussed in [16], training is mandatory for the proposed algorithm to converge correctly. Here we use five training symbols at the start of the algorithm. Detailed study of the convergence issue and the optimal training strategy is beyond the scope of this paper. In the following simulation, we set the dwell length large enough with sufficient training at the beginning so as to compare the performance of different approaches in the converged state of the adaptive algorithm. Consequently, the error probabilities presented in Fig.3 ~ Fig.5 can be viewed as lower bounds of what would be expected in the realistic setting. When the length of the dwell is not much larger than the convergence time, the performance loss due to the transient phase may be significant. In this case the symbol estimates may have to be re-computed after the algorithm converges, e.g., using the estimates for  $\{a_j^{(P)}\}_{j=1}^P$  and  $\sigma_P^2$  at the end of the dwell.

### B. Erasure insertion performance in presence of PBI

In Fig.3(a), the decoding error probabilities  $P_E(P, N, \theta, \sigma_0, \sigma_I, p_I)$  with and without EI are plotted versus the SINR  $1/\sigma_I^2$ . Optimal thresholds for both the Bayesian rule and the LRT rule are obtained numerically and used in generating the curves in Fig.3(a). Other parameters in  $P_E$  are fixed as the following:  $N = 2$ ,  $1/\sigma_0^2 = 14\text{dB}$ ,  $p_I = 0.2$ , and  $P = 1$  or  $5$ . The following observations can be made from Fig.3(a). Firstly, the gain from EI is significant, especially in the low SIR range. This makes EI very attractive for an environment with strong jammer or narrowband interferers without power control. Secondly, the LRT rule results in negligible performance loss when compared to the Bayesian rule, while the former is more preferable in implementation. Furthermore, noticeable gain can be achieved by increasing the observation window length from  $P = 1$  to  $P = 5$ . It is shown in [16] that the gain from further increasing  $P$  is negligible and two receive antennas significantly improve performance over single receive antenna.

To better understand the performance under AWGN plus PBI, we plot  $P_E(P, N, \theta, \sigma_0, \sigma_I, p_I)$  versus the loading factor  $p_I$  in Fig.3(b) while setting  $1/\sigma_I^2 = -4\text{dB}$  and other parameters same as in Fig.3(a). We can see that the decoding error probabilities degrades as  $p_I$  increases, or equivalent, as the PBI becomes more like AWGN. This is consistent with the fact that the Gaussian noise is the worst type of noise in an information-theoretic sense. The EI techniques exploit the bursty nature in the PBI and effectively suppresses it. When the PBI becomes more like AWGN as  $p_I \rightarrow 1$ , the gain from EI diminishes.

### C. Effect of the threshold

The choice of the threshold is crucial to the performance of EI decoding. Unfortunately, finding the optimal threshold analytically is highly complicated. We know the optimal threshold should depend on  $\sigma_0^2, \sigma_I^2, p_I, P, N$ , among which  $\sigma_0^2, \sigma_I^2, p_I$  are not known *a priori*. It is therefore desirable if a single fixed threshold would work well for a wide range of  $\sigma_0^2, \sigma_I^2, p_I$ . To illustrate that this is actually possible under certain conditions, we set  $1/\sigma_0^2 = 14\text{dB}$ ,  $P = 1$ ,  $N = 2$ , and find the optimal thresholds ( $\mu_{opt}$ ) for the LRT EI numerically. These optimal thresholds are plotted versus  $p_I$  and  $1/\sigma_I^2$  in Fig.4. We can see that the surface of  $\mu_{opt}$  is flat at small loading factor  $p_I$  or low SIR  $1/\sigma_I^2$ . Therefore, when the distribution of the PBI power falls into that range, a single fixed threshold may achieve close performance as the optimal threshold, which is different for different values of  $\sigma_I^2$  and  $p_I$  in general. This is verified in Fig.5, where decoding error probabilities of LRT using a fixed threshold ( $\mu = 0.0526$ ) are compared with LRT using the optimal threshold found numerically for every  $\sigma_I^2$  and  $p_I$ . Fig.5 shows that even when the optimal threshold is very different from the fixed threshold (e.g. in Fig.4 where  $p_I = 0.9$  or  $1/\sigma_I^2 = 8\text{dB}$ ), the performance loss by using the fixed threshold may be tolerable. Similar results as in Fig.4 and Fig.5 are also obtained for the Bayesian EI.

## VII. CONCLUSION

We have extended the modulation/demodulation in [14] to general DUSTM and DFD. Instead of using unconditional likelihood functions [2] [14], we propose to use the likelihood function conditioned on the variance of the noise plus interference. The resulting EI rules based on the conditional likelihood function do not require knowledge of the distribution of the noise plus interference power, in contrast to their counterparts in [14]. The DFD structure also facilitates the application of the RLS algorithm to adaptively estimate the required parameters in demodulation and EI. Specifically, a joint estimation-demodulation scheme is proposed and studied. Another problem of practical interest in applying EI is the performance loss due to mismatched thresholds, which is investigated numerically in this paper. The two proposed EI rules, the Bayesian rule and the LRT rule, are shown to achieve significantly lower decoding error probability, even when the required parameters have to be estimated and suboptimal thresholds are used.

## REFERENCES

- [1] C. Baum, A. Daraiseh, K. Balakrishnan, Adaptive erasure generation for coded frequency-hop communications with partial-band interference and fading, Proc. of IEEE Milcom, 1995
- [2] C. W. Baum, M. B. Pursley, Bayesian methods for erasure insertion in frequency-hop communications with partial-band interference, IEEE Trans. Comm., July 1992

$$\mathbf{X}(\tau) = \left[ \text{vec} \left( \overline{\Lambda}_{\tau-P+1}^{\tau-1} \mathbf{R}_{\tau-P} \right), \text{vec} \left( \overline{\Lambda}_{\tau-P+2}^{\tau-1} \mathbf{R}_{\tau-P+1} \right), \dots, \text{vec} \left( \overline{\Lambda}_{\tau-1}^{\tau-1} \mathbf{R}_{\tau-2} \right), \text{vec} \left( \mathbf{R}_{\tau-1} \right) \right] \in \mathbb{C}^{MN \times P} \quad (29)$$

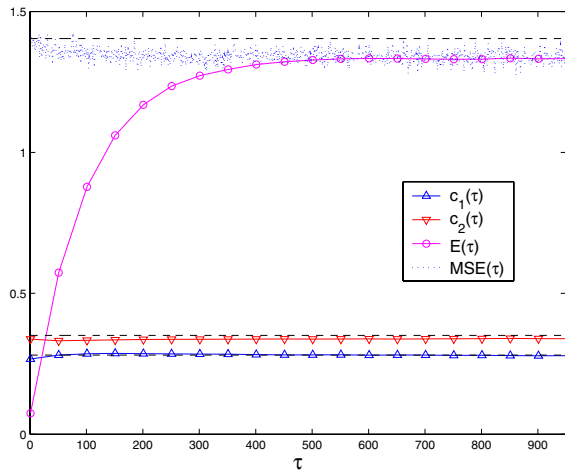


Fig. 2. Learning curves of the RLS weights and  $E(\tau)$ ,  $MSE(\tau)$  in the joint estimation-detection scheme ( $P = 2$ ,  $1/\sigma_w^2 = 0\text{dB}$ , averaged over 1023 realizations,  $\lambda = 0.99$ ,  $\delta = 0.01$ ), five training symbols)

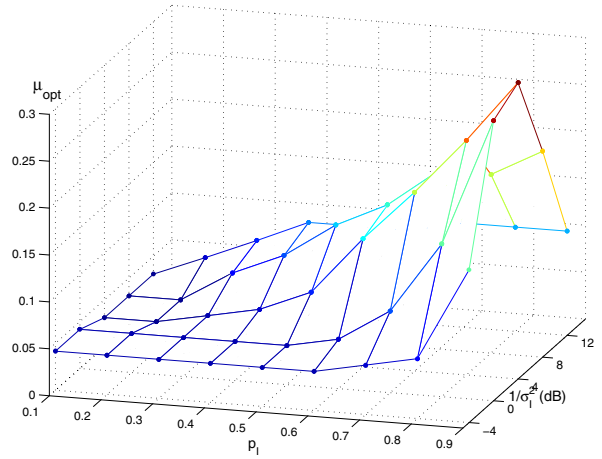


Fig. 4. Optimal thresholds for LRT EI ( $1/\sigma_0^2 = 14\text{dB}$ ,  $P = 1$ ,  $N = 2$ )

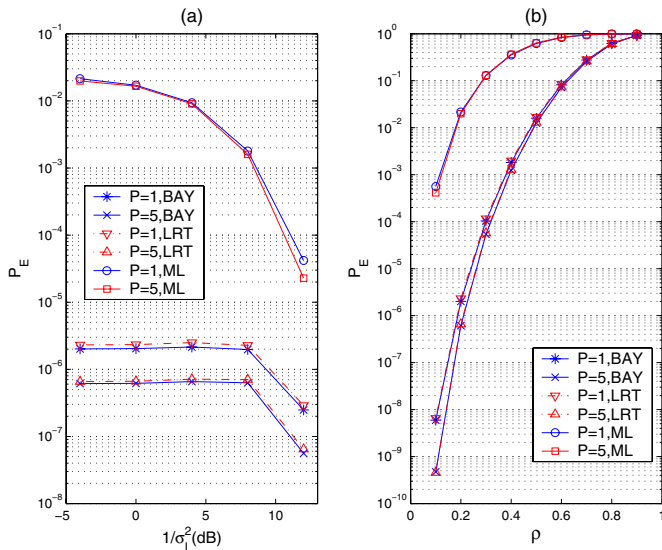


Fig. 3. decoding error probability with and without EI ( $1/\sigma_0^2 = 14\text{dB}$ ,  $N = 2$ , (a):  $p_1 = 0.2$ , (b):  $1/\sigma_1^2 = -4\text{dB}$ )

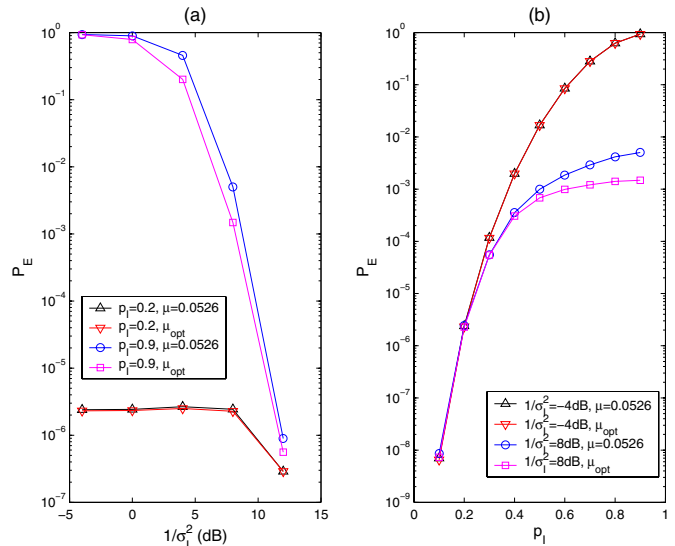


Fig. 5. Performance loss due to fixed, suboptimal threshold for LRT EI ( $1/\sigma_0^2 = 14\text{dB}$ ,  $P = 1$ ,  $N = 2$ )

[3] C. W. Baum, M. B. Pursley, Erasure insertion in frequency-hop communications with fading and partial-band interference, *IEEE Trans. Veh. Tech.*, Nov. 1997

[4] B. Bhukania, P. Schniter, Decision-feedback detection of differential unitary space-time modulation in fast Rayleigh-fading channels, *Proc. Allerton Conf. Comm., Control, and Computing*, 2002

[5] K. Choi, K. Cheun, Maximum Throughput of FHSS Multiple-Access Networks Using MFSK Modulation, *IEEE Trans. Comm.*, Mar. 2004

[6] J. D. Choi, D.-S. Yoo, W. E. Stark, Performance limits of M-FSK with Reed-Solomon coding and diversity combining, *IEEE Trans. Comm.*, Nov. 2002

[7] A. Ephremides, J. E. Wieselthier, D. J. Baker, A design concept for reliable mobile radio networks with frequency hopping signaling, *Proc. of the IEEE*, Jan. 1987

[8] S. D. Fina, Bayesian methods for erasure insertion in frequency-hopping multiple access with Rice selective fading, *IEEE Trans. Veh. Tech.*, Nov. 2003

[9] W. C. Jakes, *Microwave mobile communications*, IEEE Press, 1994

[10] M. B. Pursley, The role of spread spectrum in packet radio networks, *Proc. of the IEEE*, Jan. 1987

[12] V. Tarokh, H. Jafarkhani, A differential detection scheme for transmit diversity, *IEEE J. Sel. Areas Comm.*, July 2000

[13] L.-L. Yang, L. Hanzo, Low complexity erasure insertion in RS-coded SFH spread-spectrum communications with partial-band interference and Nakagami-m fading *IEEE Trans. Comm.*, Jun. 2002

[14] H. Sui, J. R. Zeidler, Erasure Insertion for Coded MIMO Slow Frequency-Hopping Systems in Presence of PBI, *Proc. of the 48th IEEE Globecom*, Nov. 2005

[15] H. Sui, J. R. Zeidler, An explicit and unified error probability analysis of two detection schemes for differential unitary space-time modulation, *Proc. of the 39th Asilomar Conference on Signals, Systems, and Computers*, Oct. 2005

[16] H. Sui, J. R. Zeidler, Erasure and non-erasure decoding in DUSTM-FHSS systems with DF demodulation, in preparation