Robust Frequency-Hopping System for Channels with Interference and Frequency-Selective Fading

Don Torrieri
U.S. Army Research Laboratory
Adelphi, MD
dtorr@arl.army.mil

Shi Cheng and Matthew C. Valenti
West Virginia University
Morgantown, WV
{shic,mvalenti}@csee.wvu.edu

Abstract—A robust frequency-hopping system with noncoherent detection, iterative turbo decoding and demodulation, and channel estimation is presented. The data modulation is the spectrally compact nonorthogonal continuous-phase frequency-shift keying, which strengthens the frequency-hopping system against multiple-access interference and multitone jamming. An analysis based on information theory provides the optimal values of the modulation index when there is a bandwidth constraint. The channel estimator, which is derived by applying the expectation-maximization algorithm, accommodates both frequency-selective fading and interference. Simulation experiments demonstrate the excellent system performance against partial-band interference.

I. INTRODUCTION

This paper describes and analyzes a robust frequency-hopping system with noncoherent detection, iterative turbo decoding and demodulation, and channel estimation. The system is designed to be effective not only when operating over the additive white Gaussian noise (AWGN) and fading channels but also in environments with multiple-access interference and multitone jamming.

Noncoherent or differentially coherent demodulation has practical advantages and is often necessary because of the difficulty of phase estimation after every frequency hop. A common choice of modulation is orthogonal frequency-shift keying (FSK). With FSK, there is a trade-off between the bandwidth B_u of each frequency channel and the energy efficiency. For the AWGN and fading channels, the energy efficiency can be improved by increasing the alphabet size q [1], which is equal to the number of possible transmit frequencies in the signal set. The problem is that a large B_u reduces the number of frequency channels available when the hopping is over a spectral region with fixed bandwidth W. This reduction makes the system more vulnerable to both multiple-access frequency-hopping signals and multitone jamming [2]. A reduction in B_u is obtained by using nonorthogonal continuous-phase frequency-shift keying (CPFSK).

As an example of the importance of B_u, consider multitone jamming of a frequency-hopping system with q-ary CPFSK in which the thermal noise is absent and each jamming tone has its carrier frequency within a distinct frequency channel. The uncoded symbol-error probability is approximately [2]

$$P_s = \left(\frac{q-1}{q}\right) B_u T_b \left(\frac{\mathcal{E}_b}{I_{t0}}\right)^{-1}, \quad B_u T_b \le \frac{\mathcal{E}_b}{I_{t0}} \le W T_b \quad (1)$$

where $\mathcal{E}_b$ is the energy per bit, T_b is the bit duration, $I_{t0}W$ is the total jamming power, and B_u is the uncoded bandwidth. This equation indicates the significant benefit of a small bandwidth in reducing the effect of multitone jamming.

Robust system performance is provided by using nonorthogonal CPFSK, a turbo code, *bit interleaved coded modulation* (BICM), iterative decoding and demodulation, and channel estimation. The bandwidth of q-ary CPFSK decreases with reductions in the modulation index h. Although the lack of orthogonality when $h < 1$ will cause a performance loss for the AWGN and fading channels, the turbo decoder makes this loss minor compared with the gain against multiple-access interference and multitone jamming. To implement BICM, which mitigates fading, a *binary* encoder is followed by a bitwise interleaver prior to performing the q-ary modulation [3]. The demodulator exchanges information with both the turbo decoder and the channel estimators. Frequency-selective fading changes the fading amplitude from hop to hop, and the partial-band and multiple-access interference change the interference and noise during some hop dwell intervals. Consequently, estimates of the fading amplitude and the spectral density of the interference and noise are computed for a block size N that is smaller than or equal to the number of symbols in the hop dwell interval. The decoder feeds *a priori* information (in the form of bit likelihoods) back to the demodulator and channel estimator, in accordance with the turbo principle.

Frequency hopping with binary orthogonal FSK, a turbo product code, and perfect channel information has been examined in [4]. Frequency hopping with differential q-ary phase-shift keying, iterative decoding, and channel estimation has been analyzed in [5], [6]. The proposed system with noncoherent, nonorthogonal CPFSK has the following primary advantages relative to the existing systems with differential detection or orthogonal modulation.

1. No extra reference symbol and no estimation of the phase offset in each dwell interval are required.

2. It is not necessary to assume that the phase offset is constant throughout a dwell interval.

3. The channel estimators are much more accurate and can estimate an arbitrary number of interference and noise spectral density levels.

4. The compact spectrum during each dwell interval allows more frequency channels and, hence, enhances performance
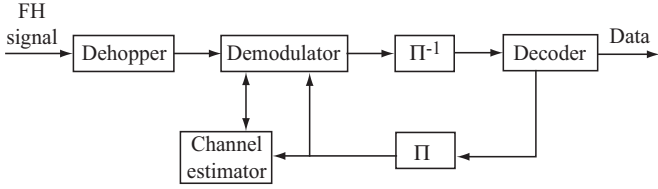
Fig. 1. Architecture of receiver for frequency-hopping system with turbo code. Π = interleaver; Π^{-1} = deinterleaver.

against multiple-access interference and multitone jamming.

II. SYSTEM MODEL

In the transmitter of the proposed system, encoded message bits are interleaved and then placed into a matrix $\mathbf{B}$ with $m = \log_2 q$ rows and N_d columns. $\mathbf{B}$ is then transformed into a length-N_d vector $\mathbf{d}$ with elements $d_i \in \{1, 2, ..., q\}$. Element d_i of $\mathbf{d}$ is found from column $\mathbf{b}_i$ of $\mathbf{B}$ through a Gray mapping. The vector $\mathbf{d}$ represents the sequence of tones that are frequency-translated by the carrier frequency of the frequency-hopping waveform during a signaling interval. The modulated signal passes through an AWGN or fading channel with partial-band interference.

As shown in Figure 1, the receiver front-end dehops the signal. The dehopped signal passes through a bank of q matched filters, each of which is implemented as a quadrature pair [1], [2]. The output of each matched filter is sampled at the symbol rate to produce a sequence of complex numbers. Assuming that symbol synchronization exists, the complex samples are then placed into an $q \times N_d$ matrix $\mathbf{Y}$ whose i^{th} column represents the outputs of the matched filters corresponding to the ith received symbol.

The channel estimator is derived under the assumption that blocks of N contiguous symbols within a dwell interval experience the same fading amplitude and the same interference and noise spectral density, either or both of which could vary from one block to the next in an arbitrary manner. If there are N symbols per block, then there will be $L = \lceil N_d/N \rceil$ blocks per codeword. The matrix $\mathbf{Y}$ can be partitioned according to $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, ..., \mathbf{Y}_L]$, where the $q \times N$ submatrix $\mathbf{Y}_\ell$ contains the received signal vectors corresponding to the ℓth fading block.

III. DEMODULATOR METRICS

The complex envelope of a unit-energy q-ary CPFSK symbol waveform with zero initial phase offset is

$$s_l(t) = \frac{1}{\sqrt{T_s}} e^{j2\pi lht/T_s}, \quad 0 \le t \le T_s, \quad l = 1, 2, \ldots, q \quad (2)$$

where T_s is the symbol duration and h is the modulation index. Suppose that symbol i of a codeword uses unit-energy waveform $s_{d_i}(t)$, where the integer d_i is a function of the codeword. If this codeword is transmitted over a channel with

fading and additive Gaussian noise, the received signal for symbol i can be expressed in complex notation as

$$r_i(t) = \text{Re}\left[a_i\sqrt{2\mathcal{E}_s}s_{d_i}(t)e^{j2\pi f_c t + \theta_i}\right] + n_i(t), \quad 0 \le t \le T_s$$
$$i = 1, 2, \ldots, N_d \quad (3)$$

where $n_i(t)$ is independent, zero-mean, white Gaussian noise with two-sided power spectral density $N_{0i}/2$, f_c is the carrier frequency, θ_i is the phase, $\mathcal{E}_s$ is the signal energy, and a_i accounts for the fading amplitude. The phase θ_i is the phase due to the contributions of the CPFSK constraint, the fading, and the frequency offset of the receiver. One might consider exploiting the inherent memory in the CPFSK when computing the metric transferred from the demodulator to a trellis decoder, as described in [7]. However, phase stability over several symbols is necessary, and the demodulator functions as a rate-one inner decoder. More design flexibility exists if the demodulator metrics are computed on a symbol-by-symbol basis, and the memory in the turbo code is exploited rather than the memory in the modulation.

Matched-filter k, which is matched to $s_k(t)$, produces the output samples

$$y_{k,i} = \sqrt{2}\int_0^{T_s} r_i(t)e^{-j2\pi f_c t}s_k^*(t)dt,$$
$$i = 1, 2, \ldots, N_d, \quad k = 1, 2, \ldots, q. \quad (4)$$

The substitution of (3) into (4) and the assumption that each of the $\{s_k(t)\}$ has a spectrum confined to $|f| < f_c$ yields

$$y_{k,i} = a_i\sqrt{\mathcal{E}_s}e^{j\theta_i}\rho_{d_i-k} + n_{k,i} \quad (5)$$

where

$$n_{k,i} = \sqrt{2}\int_0^{T_s} n_i(t)e^{-j2\pi f_c t}s_k^*(t)dt \quad (6)$$

and

$$\rho_l = \frac{\sin(\pi hl)}{\pi hl}e^{j\pi hl}. \quad (7)$$

Since $n_i(t)$ is zero-mean and white and the spectra of the $\{s_k(t)\}$ are confined, it follows that each $n_{k,i}$ is zero-mean,

$$E[n_{k,i}n_{l,i}^*] = N_{0i}\rho_{l-k} \quad (8)$$

and the $\{n_{k,i}\}$ have circular symmetry:

$$E[n_{k,i}n_{l,i}] = 0. \quad (9)$$

Since $n_i(t)$ is a Gaussian process, the real and imaginary components of $n_{k,i}$ are jointly Gaussian, and the set $\{n_{k,i}\}$ comprises *complex-valued jointly Gaussian random variables*.

Let $\mathbf{y}_i = [y_{1,i} \ldots y_{q,i}]^T$ denote the column vector of the matched-filter outputs corresponding to symbol i, and let $\mathbf{n} = [n_{1,i} \ldots n_{q,i}]^T$. Then given that the transmitted symbol is d_i, the symbol energy is $\mathcal{E}_s$, the fading amplitude is a_i, the noise spectral density is N_{0i}, and the phase is θ_i, $\mathbf{y}_i = \overline{\mathbf{y}}_i + \mathbf{n}$, where $\overline{\mathbf{y}}_i = E[\mathbf{y}_i|d_i, \mathcal{E}_s, a_i, N_{0i}, \theta_i]$. Equation (5) indicates that the *k*th component of $\overline{\mathbf{y}}_i$ is

$$\overline{y}_{k,i} = a_i\sqrt{\mathcal{E}_s}e^{j\theta_i}\rho_{d_i-k}. \quad (10)$$

The covariance matrix of $\mathbf{y}_i$ is

$$
\begin{aligned}
\mathbf{R} &= E[(\mathbf{y}_i - \overline{\mathbf{y}}_i)(\mathbf{y}_i - \overline{\mathbf{y}}_i)^H \mid d_i, a_i\sqrt{\mathcal{E}_s}, N_{0i}, \theta_i] \\
&= E[\mathbf{n}\mathbf{n}^H]
\end{aligned} \tag{11}
$$

and its elements are given by (8). It is convenient to define the matrix $\mathbf{K} = \mathbf{R}/N_{0i}$ with components

$$
K_{kl} = \rho_{l-k}. \tag{12}
$$

We can represent the conditional probability density function of $\mathbf{y}_i$ given that the transmitted symbol is d_i, the symbol energy is $\mathcal{E}_s$, the fading amplitude is a_i, the noise spectral density is N_{0i}, and the phase is θ_i as

$$
\begin{aligned}
&p(\mathbf{y}_i|d_i, a_i\sqrt{\mathcal{E}_s}, N_{0i}, \theta_i) \\
&= \frac{1}{\pi^q N_{0i}^q \det \mathbf{K}} \exp\left[-\frac{1}{N_{0i}}(\mathbf{y}_i - \overline{\mathbf{y}}_i)^H \mathbf{K}^{-1}(\mathbf{y}_i - \overline{\mathbf{y}}_i)\right]
\end{aligned} \tag{13}
$$

where $\mathbf{K}$ is independent of $(d_i, \mathcal{E}_s, a_i, N_{0i}, \theta_i)$.

An expansion yields

$$
\begin{aligned}
Q_i &= (\mathbf{y}_i - \overline{\mathbf{y}}_i)^H \mathbf{K}^{-1}(\mathbf{y}_i - \overline{\mathbf{y}}_i) \\
&= \mathbf{y}_i^H \mathbf{K}^{-1}\mathbf{y}_i + \overline{\mathbf{y}}_i^H \mathbf{K}^{-1}\overline{\mathbf{y}}_i - 2\operatorname{Re}(\mathbf{y}_i^H \mathbf{K}^{-1}\overline{\mathbf{y}}_i).
\end{aligned} \tag{14}
$$

Equations (10) and (12) indicate that $\overline{\mathbf{y}}_i$ is proportional to the $d_i th$ column of $\mathbf{K}$:

$$
\overline{\mathbf{y}}_i = a_i\sqrt{\mathcal{E}_s}e^{j\theta_i}\mathbf{K}_{:,d_i}. \tag{15}
$$

Since $\mathbf{K}^{-1}\mathbf{K} = \mathbf{I}$, only the $d_i th$ component of the column vector $\mathbf{K}^{-1}\overline{\mathbf{y}}_i$ is nonzero and

$$
Q_i = \mathbf{y}_i^H \mathbf{K}^{-1}\mathbf{y}_i + a_i^2\mathcal{E}_s - 2a_i\sqrt{\mathcal{E}_s}\operatorname{Re}(y_{d_i,i}e^{-j\theta_i}). \tag{16}
$$

For *noncoherent* signals, it is assumed that each θ_i is uniformly distributed over $[0, 2\pi)$. Substituting (16) into (13), expressing $y_{d_i,i}$ in polar form, and integrating over θ_i, we obtain the probability density function

$$
\begin{aligned}
p(\mathbf{y}_i|d_i, a_i\sqrt{\mathcal{E}_s}, N_{0i}) &= \frac{\exp\left(-\frac{\mathbf{y}_i^H \mathbf{K}^{-1}\mathbf{y}_i + a_i^2\mathcal{E}_s}{N_{0i}}\right)}{\pi^q N_{0i}^q \det \mathbf{K}} \\
&\quad \times I_0\left(\frac{2a_i\sqrt{\mathcal{E}_s}\,|y_{d_i,i}|}{N_{0i}}\right)
\end{aligned} \tag{17}
$$

where $I_0(\ )$ is the modified Bessel function of the first kind and order zero. Since the white noise $n_i(t)$ is independent from symbol to symbol, $\mathbf{y}_i$ with the density given by (17) is independent of $\mathbf{y}_l$, $i \neq l$.

Let $\widehat{A}$ and $\widehat{B}$ denote the estimates of $A = N_0$ and $B = 2a\sqrt{\mathcal{E}_s}$, respectively, for a dwell interval of N symbols during which $a_i = a$ and $N_{0i} = N_0$ are constants. Let $b_{k,i}$ denote bit k of symbol i. Let $\mathbf{Z}$ denote the $m \times N_d$ matrix whose element $z_{k,i}$ is the log-likelihood ratio for $b_{k,i}$ computed by the demodulator. The matrix $\mathbf{Z}$ is reshaped into a row vector and deinterleaved, and the resulting vector $\mathbf{z}'$ is fed into the turbo decoder. The extrinsic information $\mathbf{v}'$ at the output of the decoder is interleaved and reshaped into a $m \times N_d$ matrix $\mathbf{V}$ containing the *a priori* information:

$$
v_{k,i} = \log \frac{p(b_{k,i} = 1|\mathbf{Z}\backslash z_{k,i})}{p(b_{k,i} = 0|\mathbf{Z}\backslash z_{k,i})}. \tag{18}
$$

where conditioning on $\mathbf{Z}\backslash z_{k,i}$ means that the extrinsic information for bit $b_{k,i}$ is produced without using $z_{k,i}$. Since $\mathbf{V}$ is fed back to the demodulator,

$$
z_{k,i} = \log \frac{p(b_{k,i} = 1|\mathbf{y}_i, \gamma'_{\lceil i/N \rceil}, \mathbf{v}_i\backslash v_{k,i})}{p(b_{k,i} = 0|\mathbf{y}_i, \gamma'_{\lceil i/N \rceil}, \mathbf{v}_i\backslash v_{k,i})}. \tag{19}
$$

where $\gamma' = \{\hat{A}, \hat{B}\}$. To calculate (19) for bit k, first partition the set of symbols $\mathcal{D} = \{1, ..., q\}$ into two disjoint sets $\mathcal{D}_k^{(1)}$ and $\mathcal{D}_k^{(0)}$, where $\mathcal{D}_k^{(b)}$ contains all symbols labelled with $b_k = b$. The extrinsic information can then be expressed as

$$
z_{k,i} = \log \frac{\sum_{d\in\mathcal{D}_k^{(1)}} p(d|\mathbf{y}_i, \gamma'_{\lceil i/N \rceil}, \mathbf{v}_i\backslash v_{k,i})}{\sum_{d\in\mathcal{D}_k^{(0)}} p(d|\mathbf{y}_i, \gamma'_{\lceil i/N \rceil}, \mathbf{v}_i\backslash v_{k,i})}. \tag{20}
$$

From Bayes' rule,

$$
p(d|\mathbf{y}, \gamma', \mathbf{v}\backslash v_k) = \frac{p(\mathbf{y}|d, \gamma', \mathbf{v}\backslash v_k)p(d, \gamma', \mathbf{v}\backslash v_k)}{p(\mathbf{y}, \gamma', \mathbf{v}\backslash v_k)}. \tag{21}
$$

After conditioning on d and γ', $\mathbf{y}$ is independent of $\mathbf{v}$ and thus $p(\mathbf{y}|d, \gamma', \mathbf{v}\backslash v_k) = p(\mathbf{y}|d, \gamma')$. From the definition of conditional probability, $p(d, \gamma', \mathbf{v}\backslash v_k) = p(d|\gamma', \mathbf{v}\backslash v_k)p(\gamma', \mathbf{v}\backslash v_k)$. After conditioning on $\mathbf{v}\backslash v_k$, d is independent of γ', and thus $p(d|\gamma', \mathbf{v}\backslash v_k) = p(d|\mathbf{v}\backslash v_k)$. Gathering all these factors, inserting them into (20), and cancelling common factors yields

$$
z_{k,i} = \log \frac{\sum_{d\in\mathcal{D}_k^{(1)}} p(\mathbf{y}_i|d, \gamma'_{\lceil i/N \rceil})p(d|\mathbf{v}_i\backslash v_{k,i})}{\sum_{d\in\mathcal{D}_k^{(0)}} p(\mathbf{y}_i|d, \gamma'_{\lceil i/N \rceil})p(d|\mathbf{v}_i\backslash v_{k,i})}. \tag{22}
$$

This expression clearly delineates the contribution of the channel observation $\mathbf{y}$ and channel estimate γ', which influence only the $p(\mathbf{y}|d, \gamma')$ term, and the contribution of the *a priori* information passed to the demapper from the decoder, which affects only the $p(d|\mathbf{v}\backslash v_k)$ term.

The term $p(d|\mathbf{v}\backslash v_k)$ in (22) is only computed for those $d \in \mathcal{D}_k^{(b)}$, in which case $p(b_k = b|\mathbf{v}\backslash v_k) = p(b_k = b) = 1$. Thus, under the assumption of independent code bits (achieved by proper interleaving), the probability of d given the *a priori* input $\mathbf{v}$ is

$$
p(d|\mathbf{v}\backslash v_k) = \prod_{\substack{j=1 \\ j\neq k}}^{m} p(b_j(d)|v_j), \quad d \in \mathcal{D}_k^{(b)} \tag{23}
$$

where $b_j(d)$ is the value of the jth bit in the labelling of symbol d. The *a priori* input is interpreted by the demapper to be $v = \log[\hat{p}/(1 - \hat{p})]$, where $\hat{p}$ is the decoder's most recent estimate of the probability that the corresponding code bit is a *one*. Inverting the logarithm and solving for $\hat{p}$ yields $\hat{p} = e^v/(1 + e^v)$, which the demapper uses for $p(b = 1|v)$. Similarly, the

demapper uses $1 - \hat{p} = 1/(1 + e^v)$ for $p(b = 0|v)$. Since $b = \{0, 1\}$, we obtain

$$p(d|\mathbf{v} \backslash v_k) = \prod_{\substack{j=1 \\ j \neq k}}^{m} \frac{e^{v_j b_j(d)}}{1 + e^{v_j}}, \quad d \in \mathcal{D}_k^{(b)}. \tag{24}$$

Substituting (17) and (24) into (22) and cancelling common factors, we obtain

$$z_{k,i} = \log \frac{\sum_{d \in \mathcal{D}_k^{(1)}} I_0 \left(\gamma_{\lceil i/N \rceil} |y_{d_i,i}| \right) \prod_{\substack{j=1 \\ j \neq k}}^{m} \exp \left(b_j(d) v_{j,i} \right)}{\sum_{d \in \mathcal{D}_k^{(0)}} I_0 \left(\gamma_{\lceil i/N \rceil} |y_{d_i,i}| \right) \prod_{\substack{j=1 \\ j \neq k}}^{m} \exp \left(b_j(d) v_{j,i} \right)} \tag{25}$$

where only the ratio $\gamma = \hat{B}/\hat{A}$ is needed rather than the individual estimates.

IV. Channel Estimators

Since under block fading and time-varying interference, A and B can change on a block-by-block basis, each block is processed separately and in an identical fashion. To maintain robustness, the estimators make no assumptions regarding the distribution of the quantities to be estimated, nor do they make any assumptions regarding the correlation from block to block. The estimators directly use the channel observation for the ℓth block, $\mathbf{Y}_\ell$, while the observations of the other blocks are used indirectly through feedback of extrinsic information from the decoder. Since the form of the estimation algorithm is the same for each block, in the following discussion we assume that $\ell = 1$ and without ambiguity omit the dependence on ℓ. Thus in this section, $\mathbf{Y}$ is a generic $q \times N$ received block, $\mathbf{d} = [d_1, ... d_N]$ is the corresponding set of transmitted symbols, and $\{\hat{A}, \hat{B}\}$ is the corresponding set of channel estimators.

Rather than solving the ML problem directly, the expectation-maximization (EM) algorithm can be used as an iterative approach to solving this problem [9]. The EM algorithm consists of two steps per iteration, the *expectation* step (E-step) and the *maximization* step (M-step). Let $\{\mathbf{Y}, \mathbf{d}\}$ denote the *complete* data set. The log-likelihood of the complete data set is

$$\begin{aligned} L(A, B) &= \log p(\mathbf{Y}, \mathbf{d}|A, B) \\ &= \log \; p(\mathbf{Y}|\mathbf{d}, A, B) + \log p(\mathbf{d}) \\ &\sim \log \; p(\mathbf{Y}|\mathbf{d}, A, B). \end{aligned} \tag{26}$$

In the last step, the $\log p(\mathbf{d})$ have been removed from the likelihood because it is independent of A and B and therefore does not affect the maximization.

Since $\mathbf{y}_i$ and $\mathbf{y}_l$ are independent for $i \neq l$, (17) implies that

$$p(\mathbf{Y}|\mathbf{d}, A, B)$$
$$= \frac{\exp \left[-\frac{D}{A} - \frac{NB^2}{4A} + \sum_{i=1}^{N} \log I_0 \left(\frac{B|y_{d_i,i}|}{A} \right) \right]}{(\pi^q A^q \det \mathbf{K})^N} \tag{27}$$

where

$$D = \sum_{i=1}^{N} \mathbf{y}_i^H \mathbf{K}^{-1} \mathbf{y}_i. \tag{28}$$

After dropping irrelevant constants, (27) and (28) yield

$$L(A, B) \sim -qN \log A - \frac{D}{A} - \frac{NB^2}{4A} + \sum_{i=1}^{N} \log I_0 \left(\frac{B|y_{d_i,i}|}{A} \right). \tag{29}$$

The form of this equation indicates that the parameters A and B must both be estimated rather than just the ratio B/A.

Let r denote the EM iteration number, and $\hat{A}^{(r)}, \hat{B}^{(r)}$ the estimates of A, B during the rth iteration. Applying the E- and M-steps with (29) in a similar manner to that in [10], it is found that

$$\hat{A}^{(r)} = \frac{1}{qN} \left(D - \frac{N(\hat{B}^{(r)})^2}{4} \right) \tag{30}$$

$$\hat{B}^{(r)} = \frac{2}{N} \sum_{i=1}^{N} \sum_{k=1}^{q} p_{k,i}^{(r-1)} |y_{k,i}| F \left(\frac{4qN\hat{B}^{(r)}|y_{k,i}|}{4D - N(\hat{B}^{(r)})^2} \right). \tag{31}$$

where

$$p_{k,i}^{(r-1)} = \alpha_i^{(r-1)} I_0 \left(\frac{\hat{B}^{(r-1)}|y_{k,i}|}{\hat{A}^{(r-1)}} \right) p(d_i = k) \tag{32}$$

$\alpha_i^{(r-1)}$ is the normalization factor forcing $\sum_{k=1}^{q} p_{k,i}^{(r-1)} = 1$, i.e.

$$\alpha_i^{(r-1)} = \frac{1}{\sum_{k=1}^{q} I_0 \left(\frac{\hat{B}^{(r-1)}|y_{k,i}|}{\hat{A}^{(r-1)}} \right) p(d_i = k)} \tag{33}$$

and $p(d_i = k)$ is the probability that $d_i = k$ estimated by the decoder.

While a closed form solution to (31) is difficult to obtain, it can be found recursively [11]. The recursion involves initially replacing $\hat{B}^{(r)}$ on the right-hand side of (31) with $\hat{B}^{(r-1)}$ from the previous EM iteration. To select an initial estimate for B, consider what happens in the absence of noise. Without noise, (5) implies that either $|y_{k,i}| = a\sqrt{\mathcal{E}_s}$ (when $k = d_i$) or $|y_{k,i}| = 0$ (otherwise). Thus, an estimate for $a\sqrt{\mathcal{E}_s} = B/2$ can be achieved by taking the maximum $|y_{k,i}|$ over any column of $\mathbf{Y}$. To account for the possibility of noise, the average can be taken across all columns in the block, resulting in

$$\hat{B}^{(0)} = \frac{2}{N} \sum_{i=1}^{N} \max_k |y_{k,i}|. \tag{34}$$

The initial estimate of A is found from $\hat{B}^{(0)}$ by evaluating (30) for $r = 0$. After the initial values $\hat{A}^{(0)}$ and $\hat{B}^{(0)}$ are calculated, the initial probabilities $\{p_{k,i}^{(0)}\}$ are calculated from (32) and (33). The EM algorithm terminates when $\hat{B}^{(r)}$ converges to some fixed value, typically in fewer than 10 EM iterations.

V. Selection of Modulation Index

Let B_{max} denote the maximum bandwidth of the CPFSK modulation such that the hopping band accommodates enough frequency channels to ensure adequate performance against multiple-access interference and multitone jamming. We seek

to determine the values of h, q, and code-rate R of the turbo code that provide a good performance over the fading and AWGN channels in the presence of partial-band interference. For specific values of the modulation parameters h and q, the code rate is limited by the bandwidth requirement. Let $B_u T_b$ denote the normalized, uncoded, 99-percent power bandwidth of the CPFSK modulation. This value can be found for nonorthogonal CPFSK by numerically integrating the power-spectrum equations [1] and will be valid for frequency-hopping signals provided that the number of symbols per dwell interval is large [2]. When a code of rate R is used, the bandwidth becomes $B_c = B_u/R$. Since $B_c \leq B_{max}$ is required, the minimum code rate that achieves the bandwidth constraint is $R_{min} = B_u/B_{max}$.

Guidance in the selection of the best values of h, q, and $R \geq R_{\min}$ is provided by information theory [12]. For specific values of h and q, we evaluate the capacity $C(\gamma)$ as a function of $\gamma = \mathcal{E}_s/N_o$ under a bandwidth constraint for both the Rayleigh and AWGN channels. Since the noncoherent demodulator will include channel estimation, perfect channel-state information is assumed. Symbols are drawn from the signal set with equal probability. With these assumptions, a change of variables with $\mathbf{u} = \mathbf{y}_i/\sqrt{\mathcal{E}_s}$, and (17), the capacity for the fading channel may be expressed as [3]

$$C(\gamma) = \log_2 q - \frac{1}{q} \sum_{\nu=1}^{q} \int \int p(a)p(\mathbf{u}|\nu, a)$$
$$\times \log_2 \frac{\sum_{k=1}^{q} I_0\left(2a\gamma\left|u_k\right|\right)}{I_0\left(2a\gamma\left|u_\nu\right|\right)} d\mathbf{u}da \qquad (35)$$

where $p(a)$ is the density of the fading amplitude, the $(2q + 1)$−fold integration is over all values of a and the $2q$ real and imaginary components of $\mathbf{u}$, and

$$p(\mathbf{u}|\nu, a) = \frac{\gamma^q \exp[-\gamma(\mathbf{u}^H \mathbf{K}^{-1}\mathbf{u} + a^2)]}{\pi^q \det \mathbf{K}} I_0\left(2a\gamma\left|u_\nu\right|\right). \quad (36)$$

Equation (35) is numerically integrated by the Monte Carlo method. To determine the minimum $\mathcal{E}_b/N_o$ necessary to maintain $C(\gamma)$ above the code rate R, we use the relationship $\mathcal{E}_s = R\mathcal{E}_b \log_2 q$ and solve the equation

$$R = C(R\mathcal{E}_b \log_2 q/N_o) \qquad (37)$$

for all code rates such that $R_{min} \leq R \leq 1$. For noncoherent systems under severe bandwidth constraints, the R that minimizes $\mathcal{E}_b/N_o$ will typically be $R = R_{\min}$, but under loose bandwidth constraints the R that minimizes $\mathcal{E}_b/N_o$ could possibly be larger than R_{min} (in which case the actual bandwidth is less than B_{max}).

Figures 2 and 3 show plots of the minimum $\mathcal{E}_b/N_o$ versus h for $2 \leq q \leq 32$, $B_{\max}T_b = 2$, and $B_{\max}T_b = \infty$. When $B_{\max}T_b = 2$, the curves are truncated because there is a maximum value of h beyond which no code exists that satisfies the bandwidth constraint. Figure 2 is for the AWGN channel, and Figure 3 is for the Rayleigh fading channel. For each of the five values of q and two channel types, capacity curves were
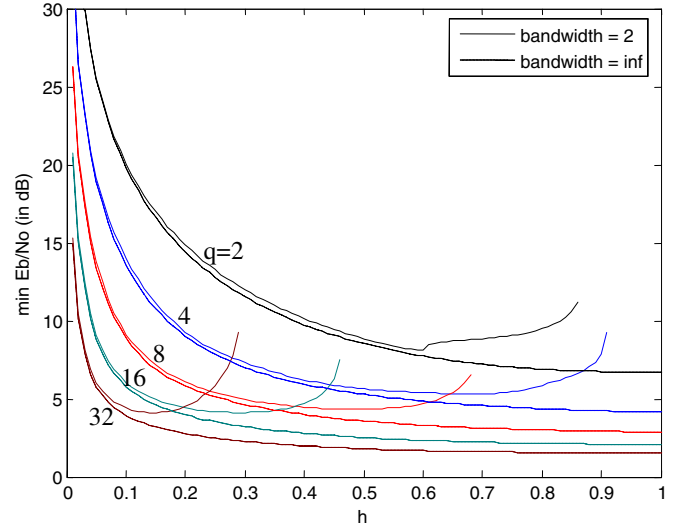

Fig. 2. Minimum $\mathcal{E}_b/N_o$ versus h for the AWGN channel, $2 \leq q \leq 32$, $B_{\max}T_b = 2$, and $B_{\max}T_b = \infty$.
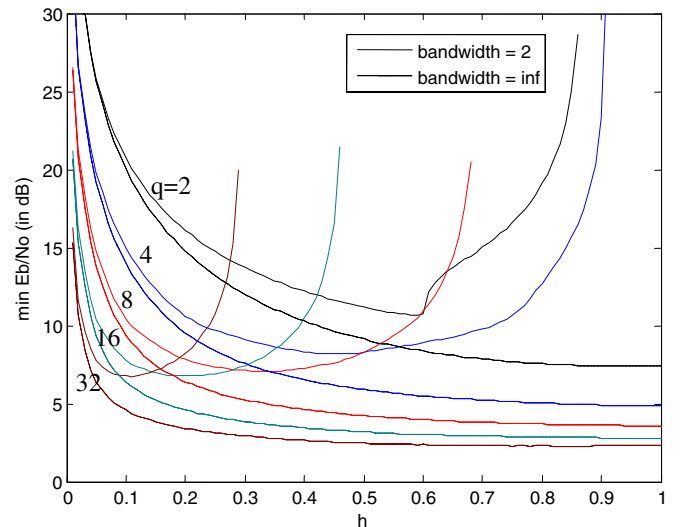

Fig. 3. Minimum $\mathcal{E}_b/N_o$ versus h for the Rayleigh channel, $2 \leq q \leq 32$, $B_{\max}T_b = 2$, and $B_{\max}T_b = \infty$.

generated for values of h ranging from $h = 0.01$ to $h = 1$ in increments of 0.01. Thus a total of $1,000$ capacity curves were generated and each curve was created using at least 2 million simulated symbols per SNR point in the range of interest. Altogether, nearly 1 trillion symbols were simulated, and it is estimated that this task would have taken about one year to complete on a single PC computer. To speed the run time, simulations were executed on a virtual private grid computer powered by the idle capacity of 30 workstations located in the teaching laboratories at the last author's institution[1]. The entire simulation scenario took just two weeks to complete on the grid computer.

[1]Job scheduling was performed online via the *Global Grid Exchange* (g2ex.com) which runs the *Frontier* Grid Platform developed by *Parabon Computation* (parabon.com).

For each value of q, in each figure there is an optimal value of h that gives the smallest value of the minimum $\mathcal{E}_b/N_o$. This smallest value decreases with q, but there are diminishing returns and the implementation complexity increases rapidly for $q > 8$. Let f_e denote the offset in the estimated carrier frequency at the receiver due to the Doppler shift and the frequency-synthesizer inaccuracy. The separation between adjacent frequencies in a CPFSK symbol is $hf_b/R\log_2 q$, where f_b denotes the information-bit rate. Since this separation must be much larger than f_e if the latter is to be negligible as assumed in (4),

$$f_e << \frac{hf_b}{R\log_2 q} \tag{38}$$

is required. Since the optimal h decreases while $R\log_2 q$ increases with q, (38) is another reason to prefer choose $q \leq 8$. For $q = 4$ in Figure 3, $h = 0.46$ is the approximate optimal value when $B_{\max}T_b = 2$, and the corresponding code rate is approximately $R = 16/27$. For $q = 8$, $h = 0.32$ is the approximate optimal value when $B_{\max}T_b = 2$, and the corresponding code rate is approximately $R = 8/15$. For both $q = 8$ and $q = 4$, (38) is satisfied if $f_e << 0.2f_b$. At the optimal values of h, the plots indicate that the loss is less than 1 dB for the AWGN channel and less than 2 dB for the Rayleigh channel relative to what could be attained with the same value of q, $h = 1$ (orthogonal CPFSK), and an unlimited bandwidth.

VI. SIMULATION RESULTS

Simulation experiments were conducted to assess the performance of frequency-hopping systems with quaternary CPFSK ($q = 4$) and octal CPFSK ($q = 8$) under the bandwidth constraint $B_{\max}T_b = 2$. The approximate optimal values of h and R determined from the bandwidth constraint and information theory are used. The interference signal is modeled as partial-band noise interference that introduces I_{t0}/μ additional interference spectral density in an interfered frequency channel, where μ is the fraction of the hopping band with interference and I_{t0} is the spectral density when $\mu = 1$. Thus, the total interference power is conserved as μ varies. The simulated system uses the turbo code from the UMTS specification [13] with 2048 information bits and the specified code-rate matching algorithm. The receiver executes no more than 20 iterations, as an early halting routine stops the iterations once the data is correctly decoded. The figures display the minimum value of $\mathcal{E}_b/N_o$ necessary to obtain a bit error probability equal to 10^{-3} versus μ for Rayleigh fading, Ricean fading with factor $K = 10$ dB, and the AWGN channel. A block coincides with a dwell interval, and the parameter A represents the spectral density due to the noise and the interference during a dwell interval. The symbols of a dwell interval undergo the same fading amplitude, and the fading amplitudes are independent from block to block, which models the frequency-selective fading that varies after each hop. The bandwidth is assumed to be sufficiently small that the fading is flat within each frequency channel.

Figures 4 and 5 plot the results for quaternary CPFSK and octal CPFSK, respectively, when there are 32 hops per
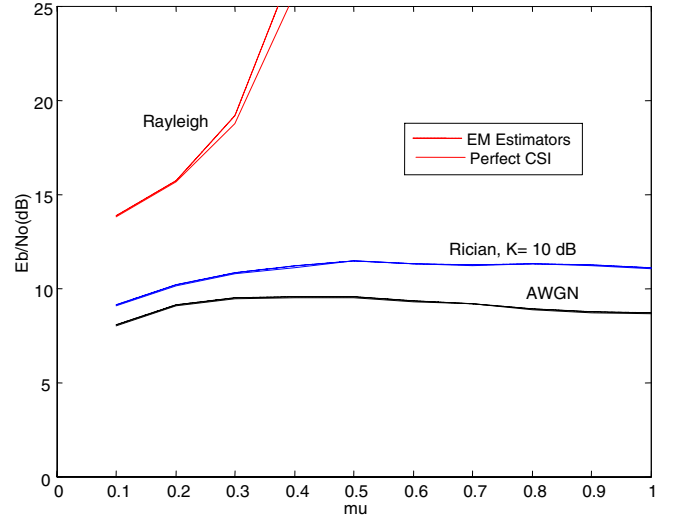

Fig. 4. Minimum required $\mathcal{E}_b/N_o$ of frequency-hopping system with quaternary CPFSK, $h = 0.46$, 32 hops per codeword, partial-band interference, $\mathcal{E}_b/I_{t0} = 10$ dB, and bit error probability $= 10^{-3}$.
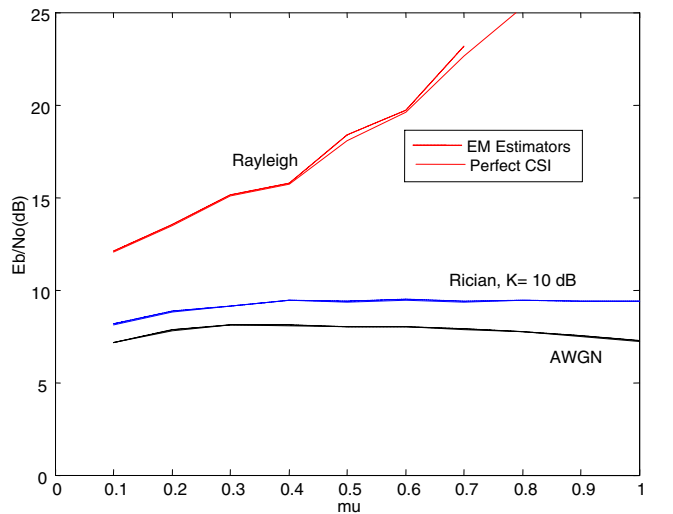

Fig. 5. Minimum required $\mathcal{E}_b/N_o$ of frequency-hopping system with octal CPFSK, $h = 0.32$, 32 hops per codeword, partial-band interference, $\mathcal{E}_b/I_{t0} = 10$ dB, and bit error probability $= 10^{-3}$.

codeword and 64 information bits per hop, and $\mathcal{E}_b/I_{t0} = 10$ dB. For quaternary CPFSK, there are 3456 code bits in a codeword, 108 code bits per hop, and 54 code symbols per hop. For octal CPFSK, there are 3840 code bits in a codeword, 120 code bits per hop, and 40 code symbols per hop. For quaternary CPFSK, $h = 0.46$ and $R = 16/27$, whereas for octal CPFSK, $h = 0.32$ and $R = 8/15$. Comparison of these two figures indicates that octal CPFSK has a nearly 2 dB advantage in $\mathcal{E}_b/N_o$ relative to quaternary CPFSK for Ricean fading and AWGN, and much more for Rayleigh fading. Both figures indicate that $\mu = 1$ or interference over the entire hopping band is the worst case for Rayleigh and Ricean fading. For the AWGN channel, a smaller value of μ is worst. The use of the EM channel estimators is
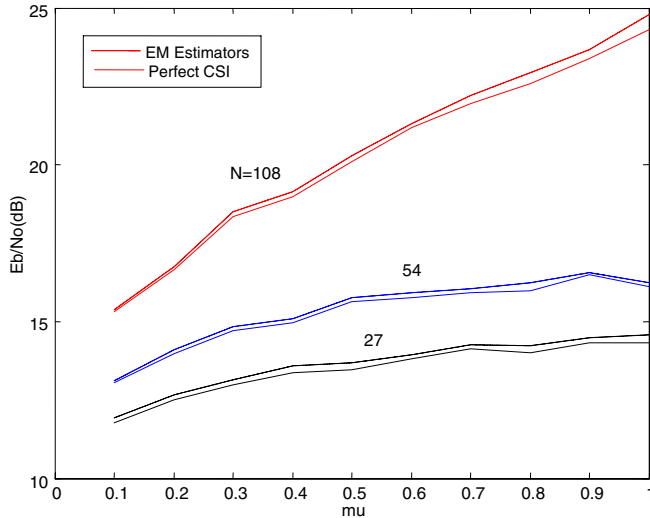
Fig. 6. Minimum required $\mathcal{E}_b/N_o$ of frequency-hopping system with quaternary CPFSK, $h = 0.46$, Rayleigh fading, partial-band interference, $\mathcal{E}_b/I_{t0} = 13$ dB, and bit error probability $= 10^{-3}$. N is the number of code symbols per hop.

Fig. 7. Minimum required $\mathcal{E}_b/N_o$ of frequency-hopping system with octal CPFSK, $h = 0.32$, Rayleigh fading, partial-band interference, $\mathcal{E}_b/I_{t0} = 13$ dB, and bit error probability $= 10^{-3}$. N is the number of code symbols per hop.

shown to produce a negligible loss relative to perfect channel-state information (CSI).

If the hop rate increases, the increase in the number of independently fading dwell intervals per codeword implies that more diversity is available in the processing of a codeword. However, the shortening of the dwell interval makes the channel estimation less reliable by providing the estimator with fewer samples. Figures 6 and 7 show the results for quaternary CPFSK and octal CPFSK, respectively, when the hop rate is varied so that there are 16, 32, or 64 hops per codeword. Independent Rayleigh fading occurs during each dwell interval, $\mathcal{E}_b/I_{t0} = 13$ dB, and the information-bit rate is maintained. For quaternary CPFSK, there are $N = 108$, 54, or 27 code symbols per hop. For octal CPFSK, there are $N = 80$, 40, or 20 code symbols per hop. For quaternary CPFSK, $h = 0.46$ and $R = 16/27$, whereas for octal CPFSK, $h = 0.32$ and $R = 8/15$. Comparison of these two figures indicates that for Ricean fading and AWGN, octal CPFSK maintains its nearly 2 dB advantage in $\mathcal{E}_b/N_o$ relative to quaternary CPFSK. Despite the slow decline in the accuracy of the EM channel estimates, the diversity improvement is sufficient to produce an improved performance as the number of code symbols per hop decreases. However, decreasing to fewer than 20 code symbols per hop will begin to broaden the spectrum significantly [2] unless the parameter values are changed.

Simulation experiments have confirmed the major performance advantage of the proposed frequency-hopping system against multiple-access interference. However, because of space limitations, an explicit demonstration is deferred to a subsequent paper.

VII. CONCLUSIONS

A noncoherent frequency-hopping system with nonorthogonal CPFSK has been designed to be highly robust in en-
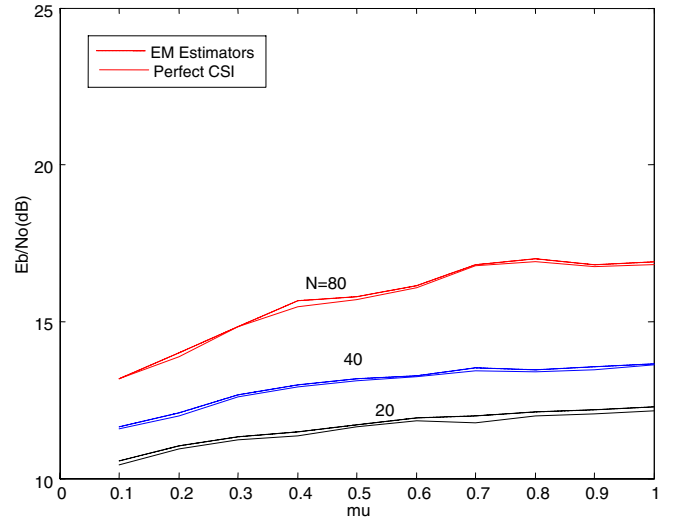
vironments including frequency-selective fading, partial-band interference, multitone jamming, and multiple-access interference. The robustness is due to the iterative turbo decoding and demodulation, the channel estimator based on the expectation-maximization algorithm, and the spectrally compact modulation.

REFERENCES

[1] J. G. Proakis, *Digital Communications, 4th ed*. New York: McGraw-Hill, 2001.
[2] D. Torrieri, *Principles of Spread-Spectrum Communication Systems*. Boston: Springer, 2005.
[3] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.
[4] Q. Zhang and T. Le-Ngoc, "Turbo product codes for FH-SS with partial-band interference," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 513-520, July 2002.
[5] W. G. Phoel, "Iterative demodulation and decoding of frequency-hopped PSK in partial-band jamming," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 1026-1033, May 2005.
[6] K. R. Narayanan and G. L. Stuber, "Performance of trellis-coded CPM with iterative demodulation and decoding," *IEEE Trans. Commun.*, vol. 49, pp. 676-687, April 2001.
[7] H. El Gamal and E. Geraniotis, "Iterative channel decoding and estimation for convolutionally coded anti-jam FH signals," *IEEE Trans. Commun.*, vol. 50, pp. 321-331, Feb. 2002.
[8] M. C. Valenti and S. Cheng, "Iterative demodulation and decoding of turbo coded M-ary noncoherent orthogonal modulation," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 1739-1747, Sept. 2005.
[9] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extentions*. Wiley, 1997.
[10] S. Cheng, M. C. Valenti, and D. Torrieri, "Turbo-NFSK: Iterative estimation, noncoherent demodulation, and decoding for fast fading channels," *IEEE Military Commun. Conf.*, 2005.
[11] S. C. Chapra and R. Canale, *Numerical Methods for Engineers, 4th ed*. New York: McGraw-Hill, 2002.
[12] S. Cheng, R. Iyer Seshadri, M. C. Valenti, and D. Torrieri, "The capacity of noncoherent continuous-phase frequency shift keying," *Conf. on Information Sciences and Systems (CISS)*, March 2007.
[13] *Third Generation Partnership Project (3GPP)*, "Universal mobile telecommunications system (UMTS): Multiplexing and channel coding (FDD)," 3GPP TS 25.212 Version 6.6.0 Release 6, Sept. 2005.