

Adaptive Spectrum Sharing for Secondary BICM Systems

¹ Amir Nasri and Robert Schober

Department of Electrical and Computer Engineering

The University of British Columbia

2356 Main Mall, Vancouver, BC, V6T 1Z4, Canada

Phone: +604 - 822 - 3515

Fax: +604 - 822 - 5949

E-mail: {amirn, rschober}@ece.ubc.ca

The frequency bands used by secondary systems such as cognitive radio and ultra-wideband (UWB) systems are expected to suffer from various forms of noise and interference with time-varying non-Gaussian distribution such as the co-channel interference caused by the primary user and other cognitive radios, UWB interference and man-made impulsive noise. To mitigate the harmful effect of noise and interference, we propose a robust L_p -norm metric for secondary systems that employ the popular bit-interleaved coded modulation (BICM) scheme. We propose two approaches for optimization of the L_p -norm metric based on BER performance analysis and Maximum-Likelihood parameter estimation principal. Based on both approaches we provide efficient adaptive algorithms that can be used for online metric optimization. Numerical and simulation results show that the proposed adaptive algorithms can effectively mitigate the adverse effects of noise and interference with time-varying statistics. Furthermore, the optimized L_p -norm metric is shown to significantly outperform other popular metrics in secondary system environments.

¹This work will be presented in part at the IEEE Global Telecommunications Conference (Globecom), New Orleans, 2008.

1 Introduction

The ever-increasing demand for high speed wireless access and inflexible methods of spectrum allocation have made the radio spectrum an increasingly scarce resource. On the other hand, recent studies have indicated that large portions of the frequency spectrum are rarely used in both space and time [1, 2]. This observation has spurred the development of secondary cognitive radio (CR) [3] and ultra-wideband (UWB) [4] systems that enable spectrum sharing by overlying and underlying the existing licensed (primary) systems, respectively. The idea behind both CR and UWB is to allow the secondary users to share the spectrum with primary users as long as the interference caused to the primary users is not noticeable. In the following, we will refer to CR and UWB systems collectively as *secondary* systems.

While most proposed secondary systems employ traditional methods of signal detection designed for additive white Gaussian noise (AWGN), various forms of non-Gaussian noise and interference with time-varying statistics can be expected to be present in practice. Examples include the narrowband and co-channel interference caused by the primary user and other secondary systems [5, 6], respectively, and man-made impulsive noise [7]. Therefore, the use of the L_2 -norm metric (also referred to as the Euclidean distance metric) for signal detection, which is optimal for AWGN, can result in significant performance losses in secondary user environments where non-Gaussian noise² is dominant. On the other hand, optimal maximum-likelihood (ML) detection requires exact knowledge of the noise distribution which may be difficult to obtain in practice and may change with time. This motivates the use of suboptimal robust metrics that perform well for a large class of noises and have a low computational complexity. These metrics should also have tunable parameters that can be adjusted to the time-variant noise statistics. Important examples of such robust metrics in the literature include Huber's M -metric [8], Myriad and Meridian metrics [9], the generalized Cauchy metric [10], and the L_p -norm metric [11]. Among these metrics, the L_p -norm metric is particularly interesting due to its low complexity and the ability to perform well in both heavy-tailed and short-tailed noise provided that the metric parameters are adjusted accordingly. However, due to the time-varying nature of noise statistics in secondary environments optimizing the metric parameters is a difficult task, as a suitable approach for metric optimization has not been proposed.

²To simplify our notation, in this paper, "noise" refers to any additive impairment of the received signal, i.e., our definition of noise also includes what is commonly referred to as "interference".

In this paper, we consider secondary systems employing bit-interleaved coded modulation (BICM) [12] in combination with either single-carrier (SC) or orthogonal frequency division multiplexing (OFDM) modulation and multiple receive antennas. The motivation for considering BICM-SC and BICM-OFDM systems is twofold. Firstly, BICM-SC and BICM-OFDM are very efficient in exploiting the time and frequency diversity of wireless fading channels, respectively. Secondly, these techniques have been adopted in a number of recent standards [13], such as the ECMA multi-band OFDM (MB-OFDM) UWB system [14], and are also prime candidates for the air interface of future CR and UWB systems [15]. We propose a robust L_p -norm decoding metric for secondary BICM systems to mitigate the harmful effects of non-Gaussian impairments. For metric optimization we propose two approaches which are based on minimizing the bit-error rate (BER) of the secondary system and ML parameter estimation, respectively. In the first approach we develop a general mathematical framework for BER performance analysis of the secondary systems that allows us to obtain an accurate approximate upper bound for the BER as well as closed-form expressions for the asymptotic BER. This framework is very general and applicable to arbitrary linear modulation formats, all commonly used fading models, and all practically relevant types of noise, and therefore can be used for accurate performance analysis of BICM systems employing L_p -norm metric. In the second approach we develop a ML estimator for the L_p -norm metric parameters based on the noise samples observed at the receiver. The latter approach, although suboptimal, has the advantage of being independent of various system parameters such as the channel type, code rate and modulation scheme, and also involves lower computational complexity. Based on both approaches, we develop efficient adaptive multivariate finite difference stochastic approximation (FDSA) algorithms that enable adaptive metric optimization.

Using numerical and simulation results we illustrate how the developed BER analysis and ML parameter estimation frameworks can be used for metric optimization. We study the performance of the proposed adaptive algorithms in a practical scenario with time-varying noise statistics and show that both algorithms can effectively adapt to the variations in the noise statistics. In addition, we show that using the optimized L_p -norm metric significant performance gains can be achieved compared to the conventional L_2 -norm metric for various types of noise present in secondary user environments.

Related Work: One approach to cope with interference in secondary BICM systems is to downscale those branch metrics which are affected by interference [16]. The disadvantage of this approach is that the location of the interference signals as well as their corresponding powers have to be accurately

estimated both in time and frequency, which not only involves high computational complexity but also is not possible if the interferers apply fast frequency/time hopping. Erasure marking was shown to be an effective technique for BICM affected by narrowband interference and impulsive noise but is computationally expensive as it requires joint erasure marking and Viterbi decoding [17, 18]. In contrast to [16, 17, 18], the approach proposed in this paper is not limited to specific types of noise and interference but leads to performance improvements for a large class of non-Gaussian impairments. We note that the performance of BICM with L_2 -norm decoding in non-Gaussian noise was studied in [19], and L_p -norm diversity combining was considered for uncoded transmission in [20]. However, the results in [19, 20] are not applicable to BICM with L_p -norm decoding.

Organization: The rest of this paper is organized as follows. In Section 2, the system model for the considered secondary communication systems is introduced. The BER analysis and ML metric parameter estimation frameworks are developed in Sections 3 and 4, respectively. In Section 5, adaptive metric optimization is considered, and analytical and simulation results are presented in Section 6. Finally, some conclusions in Section 7.

Notation: In this paper, $[\cdot]^T$, $(\cdot)^H$, $\Re\{\cdot\}$, $\|\cdot\|$, $\det(\cdot)$, and $\mathcal{E}_x\{\cdot\}$ denote transposition, Hermitian transposition, the real part of a complex number, the L_2 -norm of a vector, the determinant of a matrix, and statistical expectation with respect to x , respectively. Moreover, \mathbf{I}_M and $\mathbf{0}_M$ are the $M \times M$ identity matrix and the all-zero column vector of length M , respectively. Furthermore, we use the notation $u \stackrel{\circ}{=} v$ to indicate that u and v are asymptotically equivalent, and a function $f(x)$ is $o(g(x))$ if $\lim_{x \rightarrow 0} f(x)/g(x) = 0$.

2 System Model

We consider secondary BICM-SC and BICM-OFDM systems employing N_r receive antennas and L_p -norm decoding. In this section, we describe the corresponding signal model and the L_p -norm metric. We also present the models for several practically relevant types of noise affecting secondary user systems. For convenience, in this paper, all signals and systems are represented by their complex baseband equivalents.

2.1 Signal Model

The transmitter consists of a BICM encoder and either a SC or an OFDM modulator. The BICM encoder comprises a convolutional encoder of rate R_c , an interleaver, and a memoryless mapper [12]. The codeword $\mathbf{c} \triangleq [c_1, c_2, \dots, c_{m_c K_c}]$ of length $m_c K_c$ is generated by the convolutional encoder and interleaved. The interleaved bits are broken up into blocks of m_c bits each, which are subsequently mapped to symbols x_k from a constellation \mathcal{X} of size $|\mathcal{X}| \triangleq M = 2^{m_c}$ to form the transmit sequence $\mathbf{x} \triangleq [x_1, x_2, \dots, x_{K_c}]$ of length K_c . The transmit sequence is broken up into B frames containing N symbols each, i.e., $K_c = BN$. In BICM-SC, the symbols in each frame are transmitted in N consecutive symbol intervals, whereas in BICM-OFDM these symbols are transmitted over N consecutive sub-carriers. Furthermore, we assume that frequency hopping is applied such that each frame is transmitted with a different carrier frequency and the corresponding channels are statistically independent. For example, the ECMA MB-OFDM UWB system employs frequency hopping over $B = 3$ bands and future versions of the standard may use up to $B = 15$ bands [14].

Assuming perfect synchronization and demodulation, the observed signal for both BICM-SC and BICM-OFDM can be modeled as

$$\mathbf{r}_k = \sqrt{\gamma} \mathbf{h}_k x_k + \mathbf{n}_k, \quad 1 \leq k \leq K_c, \quad (1)$$

where γ , $\mathbf{h}_k \triangleq [h_{k,1} \dots h_{k,N_r}]^T$, and $\mathbf{n}_k \triangleq [n_{k,1} \dots n_{k,N_r}]^T$ denote the SNR per receive antenna, the fading vector, and the noise vector, respectively. Here, $h_{k,\nu}$ and $n_{k,\nu}$ are the fading gain and the noise at the ν th receive antenna, respectively. Without loss of generality we assume $\mathcal{E}\{||\mathbf{h}_k||^2\} = N_r$ and $\mathcal{E}\{||\mathbf{n}_k||^2\} = N_r$.

As customary in the literature, cf. e.g. [12, 21, 22], for our performance analysis we assume perfect interleaving, which means that \mathbf{h}_k and \mathbf{n}_k can be modeled as independent, identically distributed (i.i.d.) random vectors and only their first order probability density functions (pdfs) are relevant. We note, however, that the elements of \mathbf{h}_k and \mathbf{n}_k may in general be statistically dependent, respectively.

2.2 L_p -Norm Branch Metric

In this paper, we assume the secondary user employs an L_p -norm branch metric for Viterbi decoding. The employed branch metric for decoding bit i , $1 \leq i \leq m_c$, of symbol x_k is given by

$$\lambda_i(\mathbf{r}_k, b) \triangleq \min_{x_k \in \mathcal{X}_b^i} \{L_{q,p}(\mathbf{r}_k - \sqrt{\gamma} \mathbf{h}_k x_k)\} \quad (2)$$

where \mathcal{X}_b^i is the subset of all symbols in constellation \mathcal{X} whose label has value $b \in \{0, 1\}$ in position i . Furthermore, the generalized L_p -norm metric, $L_{\mathbf{q}, \mathbf{p}}(\mathbf{y})$, is defined as

$$L_{\mathbf{q}, \mathbf{p}}(\mathbf{y}) = \sum_{\nu=1}^{N_r} q_{\nu} |y_{\nu}|^{p_{\nu}}, \quad (3)$$

where $\mathbf{y} \triangleq [y_1 \dots y_{N_r}]^T$, and $\mathbf{q} \triangleq [q_1 \dots q_{N_r}]^T$ and $\mathbf{p} \triangleq [p_1 \dots p_{N_r}]^T$ denote the metric parameters. To achieve high performance \mathbf{q} and \mathbf{p} should be adapted to the underlying type of noise. For the special case $q_{\nu} = 1$ and $p_{\nu} = 2$, $1 \leq \nu \leq N_r$, (2) and (3) simplify to the well-known L_2 -norm branch metric, which is typically used in AWGN [12].

2.3 Fading Model

We assume that the fading gains can be expressed as $h_{k,l} \triangleq a_{k,l} e^{j\Theta_{k,l}}$, where $a_{k,l}$ and $\Theta_{k,l}$ are independent random variables (RVs). Specifically, $\Theta_{k,l}$ is uniformly distributed in $[-\pi, \pi)$ and $a_{k,l}$ is a positive real RV that follows the distribution $p_a(a_{k,l})$. In this paper, we consider both spatially independent and spatially dependent fading. For spatially independent fading the first order pdf $p_a(a_{k,l})$ is sufficient to fully describe the properties of random vector $\mathbf{a}_k \triangleq [a_{k,1} \dots a_{k,N_r}]^T$, while for spatially correlated fading the joint pdf $p_{\mathbf{a}}(\mathbf{a}_k)$ of the elements of \mathbf{a}_k is needed. However, as shown in [19], for infinitesimally small fading amplitudes, $\mathbf{a}_k \rightarrow \mathbf{0}_{N_r}$, spatially correlated fading can be modeled as spatially i.i.d. fading. Therefore, for $\mathbf{a}_k \rightarrow \mathbf{0}_{N_r}$ the joint pdf can be expressed as

$$p_{\mathbf{a}}(\mathbf{a}_k) \stackrel{\circ}{=} \prod_{\nu=1}^{N_r} p_a(a_{k,\nu}), \quad (4)$$

where

$$p_a(a_{k,\nu}) = 2\alpha_c a_{k,\nu}^{2\alpha_d-1} + o(a_{k,\nu}^{2\alpha_d-1}) \quad (5)$$

with fading distribution dependent constants α_c and α_d . For correlated Rayleigh ($\mathbf{C}_{hh} \triangleq \mathcal{E}\{\mathbf{h}_k \mathbf{h}_k^H\}$), Ricean ($\boldsymbol{\mu}_h \triangleq \mathcal{E}\{\mathbf{h}_k\}$, $\mathbf{C}_{hh} \triangleq \mathcal{E}\{(\mathbf{h}_k - \boldsymbol{\mu}_h)(\mathbf{h}_k - \boldsymbol{\mu}_h)^H\}$), and Nakagami- m ($\mathbf{C}_{aa} \triangleq \mathcal{E}\{\mathbf{a}_k \mathbf{a}_k^H\}$) fading as well as for spatially independent Nakagami- q and Weibull (with fading parameter c) fading, the fading pdf $p_a(a_{k,l})$ and parameters α_c and α_d are specified in Table 1.

2.4 Noise Model

The analysis and adaptive metric optimization presented in this paper are applicable to a large class of noises. The only restriction that we impose is that all joint moments of the elements of \mathbf{n}_k exist.

This condition is fulfilled by most noises of practical interest. An exception is α -stable noise, which is sometimes used to model impulsive phenomena [23]. However, other types of impulsive noise such as Gaussian mixture noise are included in our analysis. To illustrate the generality of our analysis and for future reference, we present in the sequel several practically relevant noise models that are frequency encountered in secondary user environments.

2.4.1 Noise Models for BICM-SC

For secondary BICM-SC systems we consider two important time-domain noise models, namely asynchronous co-channel interference (ACCI) and time-domain Gaussian-mixture noise (TD-GMN).

1) ACCI: In CR BICM-SC systems, ACCI may be caused by the primary user and/or other CRs [24, 25]. To describe this noise model, we consider a BICM-SC system with B different hopping frequencies and assume that at hopping frequency μ , $1 \leq \mu \leq B$, in addition to AWGN $\tilde{n}_{k,\nu,\mu}$, there are I_μ Ricean faded ACCI signals leading to time-domain noise

$$n_{k,\nu,\mu} = \sum_{i=1}^{I_\mu} \tilde{h}_{k,\nu,\mu}[i] \sum_{\kappa=\kappa_l}^{\kappa_u} g_{\kappa,\mu}[i] b_{\kappa,\mu}[i] + \tilde{n}_{k,\nu,\mu}, \quad 1 \leq \nu \leq N_r, \quad (6)$$

where $\tilde{h}_{k,\nu,\mu}[i]$ are temporally i.i.d. Gaussian random variables which model the Ricean interference channel gains with Ricean factor K_i and $b_{\kappa,\mu}[i]$ are the i.i.d. symbols of the i th interferer at the μ th hopping frequency. Furthermore, $g_{\kappa,\mu}[i] \triangleq g_{\mu,i}(\kappa T + \tau_{\mu,i})$, where $g_{\mu,i}(t)$, T , and $\tau_{\mu,i}$ are the effective pulse shape, the symbol duration, and the time offset of the i th interferer at the μ th hopping frequency, respectively, and we assume that $g_{\mu,i}(\kappa T + \tau_{\mu,i}) \approx 0$ for $\kappa < \kappa_l$ and $\kappa > \kappa_u$. Finally, we note that for $K_i \rightarrow \infty$, the interference channel gains $\tilde{h}_{k,\nu,\mu}[i]$ will be constant values. We refer to the resulting noise as unfaded ACCI (UF-ACCI).

2) TD-GMN: TD-GMN can be used to model the combined effects of Gaussian background noise and man-made or impulsive noise, cf. e.g. [7, 26, 27], which may affect e.g. CRs using unlicensed frequency bands. If the phenomenon causing the impulsive behavior affects the receive antennas independently, the TD-GMN is spatially i.i.d. [28] and the pdf of $n_{k,\nu}$ is given by [7]

$$p_n(n_{k,\nu}) = \sum_{i=1}^I \frac{c_i}{\pi \sigma_i^2} \exp\left(-\frac{|n_{k,\nu}|^2}{\sigma_i^2}\right), \quad 1 \leq \nu \leq N_r, \quad (7)$$

where $c_i > 0$ and $\sigma_i^2 > 0$ are parameters. In contrast, if all antennas are affected by the phenomenon causing the impulsive behavior concurrently, the TD-GMN is spatially dependent and the joint pdf

for the elements of the noise vector \mathbf{n} is given by [28]

$$p_{\mathbf{n}}(\mathbf{n}_k) = \sum_{i=1}^I \frac{c_i}{\pi^{N_r} \sigma_i^{2N_r}} \exp\left(-\frac{\|\mathbf{n}\|^2}{\sigma_i^2}\right). \quad (8)$$

Two popular special cases of both spatially independent and spatially dependent TD-GMN are Middleton's Class-A noise [7] and ϵ -mixture noise. For ϵ -mixture noise $I = 2$, $c_1 = 1 - \epsilon$, $c_2 = \epsilon$, and $\sigma_2^2 = \kappa \sigma_1^2$, where ϵ and κ denote the fraction of time when the impulsive noise is present and the ratio of the variances of the Gaussian background noise and the impulsive noise, respectively.

2.4.2 Noise Models for BICM-OFDM

For secondary BICM-OFDM systems, we consider impairment by narrowband interference (NBI) and frequency domain GMN (FD-GMN), as these two types of noise may impair both CR and UWB systems.

1) NBI: We consider a secondary BICM-OFDM system with coding over B different hopping frequencies. At hopping frequency μ , $1 \leq \mu \leq B$, the received frequency-domain signal is impaired by AWGN $\tilde{n}_{k,\nu,\mu}$ and I_μ Rayleigh faded NBI signals. The corresponding frequency-domain noise model is

$$n_{k,\nu,\mu} = \sum_{i=1}^{I_\mu} g_{k,\mu}[i] b_\mu[i] \tilde{h}_{k,\nu,\mu}[i] + \tilde{n}_{k,\nu,\mu}, \quad 1 \leq \nu \leq N_r, \quad (9)$$

where $\tilde{h}_{k,\nu,\mu}[i]$ are temporally i.i.d. Gaussian random variables which model the Ricean interference channel gains with Ricean factor K_i . Furthermore, $b_\mu[i]$ are the symbols of the i th interferer at the μ th hopping frequency and $g_{k,\mu}[i] \triangleq \exp[-j\pi(N-1)(k + f_{\mu,i}/\Delta f_s)/N + \phi_{\mu,i}] \sin[\pi(k + f_{\mu,i}/\Delta f_s)] / \sin[\pi(k + f_{\mu,i}/\Delta f_s)/N]$ [29]. Here, $f_{\mu,i}$ and $\phi_{\mu,i}$ denote the frequency and phase of the i th interferer at hopping frequency μ relative to the user, respectively, and Δf_s is the OFDM sub-carrier spacing. For future reference, we denote the ratio of the total NBI variance and the AWGN variance by κ , cf. Section 6. We note that for $K_i \rightarrow \infty$, the interference channel gains $\tilde{h}_{k,\mu,\nu}[i]$ are constant values. The resulting noise will be referred to as unfaded NBI (UF-NBI) in the rest of this paper.

2) FD-GMN: FD-GMN can be used to model the combined effects of frequency-domain Gaussian background noise and impulsive phenomena that only affect a small number of sub-carriers. For example, it can be used to model the effect of a Rayleigh faded NBI interferer or a tone interferer in a

BICM-OFDM UWB system. The mathematical model for FD-GMN is identical to that of TD-GMN in (7) and (8) except that now k denotes the sub-carrier index instead of the time.

2.4.3 UWB Interference

UWB interference may affect both secondary BICM-SC and secondary BICM-OFDM systems. In this paper, we test the proposed L_p -norm metric for impulse radio (IR) UWB and MB-OFDM UWB interference. Thereby, we generate the UWB interference signals following the respective ECMA and IEEE standards [14, 30].

3 BER Performance Analysis

In this section, we derive analytical expressions for the BER performance of the secondary BICM systems described in Section 2 in non-Gaussian noise environments. The derived expressions are obtained as a function of the metric parameters \mathbf{q} and \mathbf{p} , and therefore can serve as objective functions for metric optimization. In particular, in this section, we first provide an approximate upper bound for the BER based on the expurgated union-bound. We then analyze the behavior of the obtain BER bound for high SNR's to arrive at a closed-form expression for the asymptotic BER. The BER analysis provided in this section is based on the unified signal model presented in Section 2.1 and therefore the results are applicable to both BICM-SC and BICM-OFDM secondary systems.

3.1 Approximate Upper Bound for BER

Here, we provide an approximate upper bound for the BER performance of the considered BICM secondary systems impaired by non-Gaussian noise³. Assuming a secondary BICM system with code rate $R_c = k_c/n_c$ (k_c and n_c are integers) the union bound for the BER is given by [12]

$$P_b \leq \frac{1}{k_c} \sum_{d=d_f}^{\infty} w_c(d) P(\mathbf{c}, \hat{\mathbf{c}}), \quad (10)$$

³We note that since our derivation is based on the expurgated union bound in [12], we cannot prove that the proposed bound is a true upper bound, see discussion in [31, 32]. However, numerical evidence in e.g. [32, 19] suggests that the expurgated union bound does result in tight upper bounds if Gray labeling is applied. Our own results in Section 6 confirm this conjecture.

where \mathbf{c} and $\hat{\mathbf{c}}$ are two distinct code sequences with Hamming distance d that differ only in $l \geq 1$ consecutive trellis states, $w_c(d)$ denotes the total input weight of error events at Hamming distance d , and d_f is the free distance of the code. $P(\mathbf{c}, \hat{\mathbf{c}})$ is the pairwise error probability (PEP), i.e., the probability that the decoder chooses code sequence $\hat{\mathbf{c}}$ when code sequence $\mathbf{c} \neq \hat{\mathbf{c}}$ is transmitted. Adopting the expurgated bound from [12], the PEP can be expressed as

$$P(\mathbf{c}, \hat{\mathbf{c}}) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} (\Psi(s))^d \frac{ds}{s}, \quad (11)$$

with

$$\Psi(s) \triangleq \frac{1}{m_c 2^{m_c}} \sum_{i=1}^{m_c} \sum_{b=0}^1 \sum_{x_k \in \mathcal{X}_b^i} \Phi_{\Delta}(s) \quad (12)$$

where c is a small positive constant that lies in the region of convergence of the integrand. Furthermore, $\Phi_{\Delta}(s) \triangleq \mathcal{E}_{\mathbf{h}_k, \mathbf{n}_k} \{e^{-s\Delta}\}$ is the moment generating function (MGF) of the metric difference

$$\begin{aligned} \Delta &\triangleq L_{\mathbf{q}, \mathbf{p}}(\mathbf{r}_k - \sqrt{\gamma} \mathbf{h}_k z_k) - L_{\mathbf{q}, \mathbf{p}}(\mathbf{r}_k - \sqrt{\gamma} \mathbf{h}_k x_k) \\ &= \sum_{\nu=1}^{N_r} q_{\nu} |r_{k,\nu} - \sqrt{\gamma} h_{k,\nu} z_k|^{p_{\nu}} - \sum_{\nu=1}^{N_r} q_{\nu} |r_{k,\nu} - \sqrt{\gamma} h_{k,\nu} x_k|^{p_{\nu}} \end{aligned} \quad (13)$$

conditional on the transmission of $x_k \in \mathcal{X}_b^i$. Here, z_k is the nearest neighbor of x_k in $\mathcal{X}_{\bar{b}}^i$ with \bar{b} being the bit complement of b . Since conditional on the transmission of x_k , we have $r_{k,\nu} = \sqrt{\gamma} h_{k,\nu} x_k + n_{k,\nu}$, $1 \leq \nu \leq N_r$, we can rewrite (13) as

$$\Delta = \sum_{\nu=1}^{N_r} q_{\nu} |\sqrt{\gamma} h_{k,\nu} e_k + n_{k,\nu}|^{p_{\nu}} - \sum_{\nu=1}^{N_r} q_{\nu} |n_{k,\nu}|^{p_{\nu}} \quad (14)$$

where $e_k \triangleq x_k - z_k$. The MGF $\Phi_{\Delta}(s)$ can therefore be obtained as

$$\Phi_{\Delta}(s) = \mathcal{E}_{\mathbf{h}_k, \mathbf{n}_k} \left\{ \prod_{\nu=1}^{N_r} e^{-sq_{\nu} (|\sqrt{\gamma} h_{k,\nu} e_k + n_{k,\nu}|^{p_{\nu}} - |n_{k,\nu}|^{p_{\nu}})} \right\}. \quad (15)$$

The approximate union bound is obtained by numerically evaluating (10), (11), and (15). In particular, several different techniques can be used for evaluation of the complex integral in (11). One popular option is the application of Gauss–Chebyshev quadrature rules, cf. [31]. An even more efficient approach is the use of a saddle point approximation [31, 33]. In the latter case, the PEP is accurately approximated as

$$P(\mathbf{c}, \hat{\mathbf{c}}) \approx \frac{(\Psi(\hat{s}))^{(d+1/2)}}{\hat{s} \sqrt{2\pi d \Psi''(\hat{s})}} \quad (16)$$

where the saddlepoint \hat{s} is defined as the value for which $\Psi'(\hat{s}) = 0$ is valid. Here, $\Psi'(s)$ and $\Psi''(s)$ denote the first and the second derivative of $\Psi(s)$ with respect to s , respectively. $\Psi(s)$, $\Psi'(s)$, and $\Psi''(s)$ can be obtained efficiently by Monte–Carlo simulation. The saddle point is obtained by a simple one–dimensional search (is this true, please provide more details?????????).

The saddle point approximation is an efficient technique for evaluation of the PEP and consequently for calculation of the approximate BER bound in (10). Since the proposed BER bound depends on the metric parameters \mathbf{q} and \mathbf{p} , it can be used for offline metric optimization if the statistics of fading and noise are known *a priori* or a sufficient number of measurements for \mathbf{h}_k and \mathbf{n}_k are available to perform the averaging in (15). However, the required search for the saddle point and the fact that (15) requires averaging over both fading and noise, make the approximate upper bound unsuitable for adaptive metric optimization. In order to obtain analytical BER expressions that are suitable for adaptive metric optimization, we perform an asymptotic high SNR analysis in the next subsection.

3.2 Asymptotic BER

In this section, we analyze the asymptotic behavior of the BER bound provided in (10) for $\gamma \rightarrow \infty$, i.e., for asymptotically high SNR's. For this purpose, it is convenient to rewrite the PEP as

$$P(\mathbf{c}, \hat{\mathbf{c}}) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \mathcal{E}_{\mathbf{n}_k} \{ \Phi(s|\mathbf{n}_k) \} \frac{ds}{s}, \quad (17)$$

with

$$\Phi(s|\mathbf{n}_k) = \prod_{k=1}^d \left(\frac{1}{m_c 2^{m_c}} \sum_{i=1}^{m_c} \sum_{b=0}^1 \sum_{x_k \in \mathcal{X}_b^i} \Phi_{\Delta}(s|\mathbf{n}_k) \right), \quad (18)$$

where

$$\Phi_{\Delta}(s|\mathbf{n}_k) = \mathcal{E} \{ e^{-s\Delta} \} \triangleq e^{-s \sum_{\nu=1}^{N_r} q_{\nu} |n_{k,\nu}|^{p_{\nu}}} \prod_{\nu=1}^{N_r} \Phi_{y_{k,\nu}}(q_{\nu} s). \quad (19)$$

In obtaining the last equality, we have used (14), $y_{k,\nu} \triangleq |\sqrt{\gamma} h_{k,\nu} e_k + n_{k,\nu}|^{p_{\nu}}$, the asymptotic independence of the fading gains (4), and $\Phi_{y_{k,\nu}}(s) \triangleq \mathcal{E} \{ e^{-s y_{k,\nu}} \}$. The pdf $f_y(y_{k,\nu})$ of $y_{k,\nu}$ for $\gamma \rightarrow \infty$ is derived in the Appendix and given in (57). Based on this result, for $\gamma \rightarrow \infty$, the MGF $\Phi_{y_{k,\nu}}(s)$ can be expressed as

$$\Phi_{y_{k,\nu}}(s) = \frac{2\alpha_c}{p_{\nu} (\gamma d_{xz}^2)^{\alpha_d}} \sum_{i=0}^{\bar{\alpha}_d} \xi_{i,\nu} |n_{k,\nu}|^{2i} s^{-2(\alpha_d-i)/p_{\nu}} + o(\gamma^{-\alpha_d}), \quad (20)$$

where $d_{xz} \triangleq |e_k|$, $\bar{\alpha}_d \triangleq \lceil \alpha_d \rceil - 1$, $\xi_{i,\nu} \triangleq \frac{\Gamma(2(\alpha_d-i)/p_\nu) P_i}{\Gamma(\alpha_d-i)}$, and P_i is defined in (54). Substituting (20) in (19) yields

$$\Phi_\Delta(s|\mathbf{n}_k) = \frac{(2\alpha_c)^{N_r} e^{-s \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu}}}{(\gamma d_{xz}^2)^{N_r \alpha_d}} \prod_{\nu=1}^{N_r} \left(\sum_{i=0}^{\bar{\alpha}_d} \frac{\xi_{i,\nu} |n_{k,\nu}|^{2i}}{p_\nu q_\nu^{2(\alpha_d-i)/p_\nu}} s^{-2(\alpha_d-i)/p_\nu} \right) + o(\gamma^{-N_r \alpha_d}) \quad (21)$$

Using (21) in (18) we obtain for $\Phi(s|\mathbf{n}_k)$ the expression

$$\begin{aligned} \Phi(s|\mathbf{n}_k) &= X_m(\alpha_d, N_r, d) (2\alpha_c)^{dN_r} \gamma^{-dN_r \alpha_d} e^{-s \sum_{k=1}^d \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu}} \\ &\quad \times \prod_{k=1}^d \prod_{\nu=1}^{N_r} \left(\sum_{i=0}^{\bar{\alpha}_d} \frac{\xi_{i,\nu} |n_{k,\nu}|^{2i}}{p_\nu q_\nu^{2(\alpha_d-i)/p_\nu}} s^{-2(\alpha_d-i)/p_\nu} \right) + o(\gamma^{-dN_r \alpha_d}) \end{aligned} \quad (22)$$

with modulation and coding dependent constant

$$X_m(\alpha_d, N_r, d) \triangleq \left(\frac{1}{m_c 2^{m_c}} \sum_{i=1}^{m_c} \sum_{b=0}^1 \sum_{x_k \in \mathcal{X}_b^i} \frac{1}{(d_{xz}^2)^{N_r \alpha_d}} \right)^d. \quad (23)$$

After some manipulations, $\Phi(s|\mathbf{n}_k)$ can be rewritten as

$$\begin{aligned} \Phi(s|\mathbf{n}_k) &= X_m(\alpha_d, N_r, d) (2\alpha_c)^{dN_r} \gamma^{-dN_r \alpha_d} e^{-s \sum_{k=1}^d \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu}} \sum_{K=0}^{dN_r \bar{\alpha}_d} \sum_{i_1+\dots+i_{N_r}=K} s^{-2 \sum_{\nu=1}^{N_r} (d\alpha_d-i_\nu)/p_\nu} \\ &\quad \times \prod_{\nu=1}^{N_r} \sum_{j_1+\dots+j_d=i_\nu} p_\nu^{-d} \frac{\xi_{j_1,\nu} |n_{1,\nu}|^{2j_1}}{q_\nu^{2(\nu-j_1)/p_\nu}} \cdots \frac{\xi_{j_d,\nu} |n_{d,\nu}|^{2j_d}}{q_\nu^{2(\alpha_d-j_d)/p_\nu}} + o(\gamma^{-dN_r \alpha_d}) \end{aligned} \quad (24)$$

where $0 \leq j_k \leq \bar{\alpha}_d$, $1 \leq k \leq d$, and $0 \leq i_\nu \leq K$, $1 \leq \nu \leq N_r$. Applying (24) in (17) and using the Residue theorem, we obtain for the PEP for $\gamma \rightarrow \infty$ the expression

$$P(\mathbf{c}, \hat{\mathbf{c}}) = X_m(\alpha_d, N_r, d) (2\alpha_c)^{dN_r} \gamma^{-dN_r \alpha_d} \sum_{K=0}^{dN_r \bar{\alpha}_d} \sum_{i_1+\dots+i_{N_r}=K} M_{\mathbf{n}}(i_1, \dots, i_{N_r}) + o(\gamma^{-dN_r \alpha_d}), \quad (25)$$

where the *generalized* noise moment $M_{\mathbf{n}}(i_1, \dots, i_{N_r})$ is defined as

$$M_{\mathbf{n}}(i_1, \dots, i_{N_r}) \triangleq \frac{\mathcal{E}_{\mathbf{n}_k} \left\{ \prod_{\nu=1}^{N_r} \sum_{j_1+\dots+j_d=i_\nu} \prod_{k=1}^d \xi_{j_k,\nu} |n_{k,\nu}|^{2j_k} \left(\sum_{k=1}^d \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu} \right)^{2 \sum_{\nu=1}^{N_r} (d\alpha_d-i_\nu)/p_\nu} \right\}}{\Gamma \left(2 \sum_{\nu=1}^{N_r} \frac{d\alpha_d-i_\nu}{p_\nu+1} \right) \prod_{\nu=1}^{N_r} p_\nu^d q_\nu^{2(d\alpha_d-i_\nu)/p_\nu}}. \quad (26)$$

We note that $o(\gamma^{-dN_r \alpha_d})$ in (24) contains sums of products of the elements of the noise vector \mathbf{n}_k . For these term to remain finite after expectation with respect to \mathbf{n}_k , we have to assume that all joint

moments of the elements of \mathbf{n}_k exist. Based on (25) and (10) a closed-form approximation for the asymptotic BER can be obtained as

$$P_b \doteq \frac{w_c(d_f)}{k_c} X_m(\alpha_d, N_r, d_f) (2\alpha_c)^{d_f N_r} \gamma^{-d_f N_r \alpha_d} \sum_{K=0}^{d_f N_r \bar{\alpha}_d} \sum_{i_1 + \dots + i_{N_r} = K} M_{\mathbf{n}}(i_1, \dots, i_{N_r}). \quad (27)$$

In deriving (27), besides the assumption that all joint noise moments exist, we also have assumed that (a) the approximate BER bound in (10) is tight for high SNRs and (b) the first term with $d = d_f$ in (10) is dominant. Assumption (a) is confirmed by simulations in Section 6 and assumption (b) is justified for high SNR.

The asymptotic expression for the BER in (27) is very general since it is applicable to all types of fading whose pdf can be expressed as in (4), all types of noise whose joint moments exist, all modulation formats, and arbitrary code rates. Eq. (27) can be used for fast evaluation of the BER at high SNRs. For this purpose, it is most convenient to evaluate the generalized moments $M_{\mathbf{n}}(i_1, \dots, i_{N_r})$ by Monte-Carlo simulation. Note that this has to be done only once since these moments do not depend on the SNR. For special cases, $M_{\mathbf{n}}(i_1, \dots, i_{N_r})$ can be calculated in closed form. However, we do not derive corresponding expressions here because of space limitation and since the main goal of this paper is the exploitation of (27) for adaptive metric optimization, cf. Section 5. First, however, we consider some special cases and discuss some implications of (27).

3.2.1 Special Cases

In this subsection, we discuss some special cases that allow significant simplifications of the asymptotic result in (27).

Fading Channels with $\alpha_d = 1$: For $\alpha_d = 1$, which is valid for e.g. (possibly spatially correlated) Rayleigh, Ricean, and Nakagami- q fading, (25) can be significantly simplified. In particular, since we have $\bar{\alpha}_d = 0$, the double sum in (25) disappears and only the moment $M_{\mathbf{n}}(0, \dots, 0)$ is needed. This moment is given by

$$M_{\mathbf{n}}(0, \dots, 0) = \frac{\prod_{\nu=1}^{N_r} (\Gamma(2/p_\nu))^{d_f}}{\Gamma(\sum_{\nu=1}^{N_r} 2d_f/p_\nu + 1) p_\nu^{d_f} q_\nu^{2d_f/p_\nu}} \mathcal{E}_{\mathbf{n}_k} \left\{ \left(\sum_{k=1}^{d_f} \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu} \right)^{\sum_{\nu=1}^{N_r} 2d_f/p_\nu} \right\}. \quad (28)$$

Uncoded Transmission ($d_f = 1$): For uncoded transmission we have $d_f = 1$, $k_c = 1$, and $w_c(1) = 1$. Furthermore, $X(\alpha_d, N_r, 1) = N_{\min}/(m_c d_{\min}^{2\alpha_d N_r})$, where N_{\min} and d_{\min} are the average number of minimum distance neighbors and the minimum distance of signal constellation \mathcal{X} ,

respectively. Hence, the asymptotic BER can be expressed as

$$P_b \triangleq \frac{N_{\min}(2\alpha_c)^{N_r}}{m_c d_{\min}^{2\alpha_d N_r}} \gamma^{-N_r \alpha_d} \sum_{K=0}^{N_r \alpha_d} \sum_{i_1 + \dots + i_{N_r} = K} M_{\mathbf{n}}(i_1, \dots, i_{N_r}), \quad (29)$$

where now

$$M_{\mathbf{n}}(i_1, \dots, i_{N_r}) = \frac{\mathcal{E}_{\mathbf{n}_k} \left\{ \prod_{\nu=1}^{N_r} \xi_{i_{\nu}, \nu} |n_{k, \nu}|^{2i_{\nu}} \left(\sum_{\nu=1}^{N_r} q_{\nu} |n_{k, \nu}|^{p_{\nu}} \right)^{2 \sum_{\nu=1}^{N_r} (\alpha_d - i_{\nu}) / p_{\nu}} \right\}}{\Gamma \left(2 \sum_{\nu=1}^{N_r} (\alpha_d - i_{\nu}) / (p_{\nu} + 1) \right) \prod_{\nu=1}^{N_r} p_{\nu} q_{\nu}^{2(\alpha_d - i_{\nu}) / p_{\nu}}}. \quad (30)$$

If we assume additionally $\alpha_d = 1$, the double sum in (29) disappears and (30) simplifies in a similar manner as (28). For uncorrelated Rayleigh and Ricean fading, the corresponding asymptotic BER can be shown to be equivalent to [20, Eqs. (10),(13)]. Note, however, that even for uncoded transmission, the results in this paper, which are also valid for e.g. Nakagami- m and Weibull fading and correlated Rayleigh and Ricean fading, are much more general than those in [20].

Simplified Metric: The general L_p -norm metric in (3) requires the optimization of $2N_r - 1$ metric parameters (since the metric is invariant to a multiplication with a positive constant, we may choose $q_1 = 1$ without loss of generality). In practice, for complexity reasons, a metric with just one adjustable parameter may be more desirable. Therefore, we may choose $q_{\nu} = 1$ and $p_{\nu} = p$, $1 \leq \nu \leq N_r$. With this simplification, the relevant generalized noise moment can be rewritten as

$$M_{\mathbf{n}}(i_1, \dots, i_{N_r}) = \frac{\mathcal{E}_{\mathbf{n}_k} \left\{ \prod_{\nu=1}^{N_r} \sum_{j_1 + \dots + j_{d_f} = i_{\nu}} \prod_{k=1}^{d_f} \xi_{j_k, \nu} |n_{k, \nu}|^{2j_k} \left(\sum_{k=1}^{d_f} \sum_{\nu=1}^{N_r} |n_{k, \nu}|^p \right)^{2(d_f \alpha_d N_r - K) / p} \right\}}{\Gamma(2(d_f \alpha_d N_r - K) / p + 1) p^{d_f N_r}}, \quad (31)$$

which only depends on p . Thus, the metric optimization is greatly simplified. Furthermore, if we specialize (31) further to the L_2 -norm metric ($p = 2$), (29) and (31) can be shown to be equivalent to [19, Eq. (19), (20)]. On the other hand, if we make the additional assumption $\alpha_d = 1$, the double sum in (19) disappears again and only

$$M_{\mathbf{n}}(0, \dots, 0) = \frac{1}{\Gamma(2d_f \alpha_d N_r / p + 1) p^{d_f N_r}} \mathcal{E}_{\mathbf{n}_k} \left\{ \left(\sum_{k=1}^{d_f} \sum_{\nu=1}^{N_r} |n_{k, \nu}|^p \right)^{2d_f \alpha_d N_r / p} \right\} \quad (32)$$

is needed. This is of interest, since closed-form results for the expected value in (32) have been reported in [20] for various types of noise.

3.2.2 Diversity and Coding Gain

It is convenient to express the asymptotic BER in terms of the diversity gain G_d and coding gain G_c , i.e., $P_b \triangleq (G_c \gamma)^{-G_d}$ [34]. For the problem at hand, a comparison with (27) shows that diversity and coding gain are given by

$$G_d = d_f N_r \alpha_d \quad (33)$$

$$G_c [\text{dB}] = -\frac{10}{\alpha_d} \log_{10}(2\alpha_c) - \frac{10}{G_d} \log_{10} \left(\frac{w_c(d_f) X_m(\alpha_d, N_r, d_f)}{k_c} \right) - \frac{10}{G_d} \log_{10} \left(\sum_{K=0}^{d N_r \alpha_d} \sum_{i_1 + \dots + i_{N_r} = K} M_{\mathbf{n}}(i_1, \dots, i_{N_r}) \right) \quad (34)$$

From (33) we observe that the diversity gain is independent of the metric parameters, \mathbf{q} and \mathbf{p} , and of the type of noise. Therefore, the asymptotic BER curves for all types of noise are parallel independent of how \mathbf{q} and \mathbf{p} are chosen. Eq. (34) reveals that the coding gain consists of three terms. The first and the second term depend on the fading parameter α_c and on the modulation and coding scheme, respectively, but are independent of the metric parameters \mathbf{q} and \mathbf{p} . The third term is a function of \mathbf{q} and \mathbf{p} and also depends on the type of noise via the generalized moments $M_{\mathbf{n}}(i_1, \dots, i_{N_r})$. Thus, optimizing the metric parameters based on (27) will shift the asymptotic BER curve as far as possible to the left but has no influence on the slope of the BER curve. Another interesting observation from (33) and (34) is that coding and mapping schemes with maximum free distance d_f and minimum $w_c(d_f) X_m(\alpha_d, N_r, d_f)$ are not only asymptotically optimal for AWGN and L_2 -norm decoding but also for non-AWGN channels and L_p -norm decoding. In other words, transmitter and receiver can be independently optimized and the type of noise does not influence transmitter optimization, which simplifies the optimization of secondary BICM systems considerably.

4 Metric Parameter Estimation

In this section, we formulate the metric parameter optimization problem as an ML parameter estimation problem. For this purpose we first introduce the generalized Gaussian (GG) pdf. Using on this pdf, we then find the ML estimates for the L_p -norm metric parameters \mathbf{q} and \mathbf{p} based on the noise samples observed at the receiver. These estimates are then used as metric parameters for the L_p -norm metric employed in the secondary system. We note that this approach is suboptimal,

i.e., the metric parameter estimates obtained using this approach will not necessarily minimize the BER. However, the computational complexity of this approach is lower than the one presented in Section 3 as it leads to simpler expressions for the objective function for metric optimization.

4.1 GG Probability Density Function

The GG pdf encompasses a wide range of distributions and is also a popular model for non-Gaussian noise. This pdf can be expressed as

$$f_{GG}(z; \alpha, \beta) = \frac{\beta \alpha^{2/\beta}}{2\pi \Gamma(2/\beta)} \exp(-\alpha |z|^\beta), \quad (35)$$

with the corresponding variance σ_{GG}^2 given as

$$\sigma_{GG}^2 = \frac{\Gamma(4/\beta)}{\alpha^{2/\beta} \Gamma(2/\beta)}, \quad (36)$$

where α , $0 < \alpha < \infty$, is a scaling factor and β , $0 < \beta < \infty$, is a shape parameter. Smaller values of the shape parameter β ($0 \leq \beta < 2$) correspond to heavier-tailed and thus more impulsive distributions, whereas larger values of β ($\beta > 2$) result in shorter-tailed distributions. Well-known special cases of this density are Laplacian ($\beta = 1$) and Gaussian noise ($\beta = 2$). Furthermore, in the limiting case of $\beta \rightarrow \infty$ the GG pdf converges to a uniform density on the $|z| \leq 1/\alpha$ circle.

The motivation behind considering the GG pdf for parameter estimation is two fold. Firstly, this pdf is very flexible and therefore can be successfully used to approximate a wide range of distributions. Secondly, the L_p -norm metric employed in this paper is closely related to the GG pdf. In fact, it is not difficult to see that an optimized L_p -norm metric can achieve ML performance in the presence a non-Gaussian impairment distributed according to a GG pdf.

4.2 ML Parameter Estimation

The main idea behind the proposed ML metric parameter estimation method is to identify the GG pdf that best approximates the distribution of the underlying noise in an ML sense [35]. Based on the GG pdf given in (35), the ML parameter estimation problem can be formulated as follows. For the observed noise samples \mathbf{n}_k , $1 \leq k \leq K_m$, generated based on the pdf $p_{\mathbf{n}_k}(\mathbf{n}_k)$, we define the log-likelihood function (LLF) $L(\mathbf{n}; \mathbf{q}, \mathbf{p})$ as

$$L(\mathbf{n}; \mathbf{q}, \mathbf{p}) \triangleq \frac{1}{K_m} \log \left\{ \prod_{k=1}^{K_m} \prod_{\nu=1}^{N_r} p_{GG}(n_{k,\nu}; q_\nu, p_\nu) \right\} \quad (37)$$

with $\mathbf{n} \triangleq [\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_{K_m}]$. The ML parameter estimates $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$ are the values of \mathbf{q} and \mathbf{p} that maximize the LLF and therefore are given as

$$(\hat{\mathbf{q}}, \hat{\mathbf{p}}) = \arg \max_{\mathbf{q}, \mathbf{p}} \{L(\mathbf{n}; \mathbf{q}, \mathbf{p})\} \quad (38)$$

Using (35) in (37) yealds

$$\begin{aligned} L(\mathbf{n}; \mathbf{q}, \mathbf{p}) &= \frac{1}{K_m} \sum_{k=1}^{K_m} \sum_{\nu=1}^{N_r} \left[\log \left(\frac{p_\nu q_\nu^{2/p_\nu}}{2\pi \Gamma(2/p_\nu)} \right) - q_\nu |n_{k,\nu}|^{p_\nu} \right] \\ &= \sum_{\nu=1}^{N_r} \log \left(\frac{p_\nu q_\nu^{2/p_\nu}}{2\pi \Gamma(2/p_\nu)} \right) - \hat{m}_{\mathbf{n}_k}(\mathbf{q}, \mathbf{p}) \end{aligned} \quad (39)$$

with $\hat{m}_{\mathbf{n}_k}(\mathbf{q}, \mathbf{p}) = \frac{1}{K_m} \sum_{k=1}^{K_m} \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu}$. For $K_m \rightarrow \infty$, the strong law of large numbers can be invoked to concluded that $\hat{m}_{\mathbf{n}_k}(\mathbf{q}, \mathbf{p}) \rightarrow \mathcal{E}_{\mathbf{n}_k} \left\{ \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu} \right\}$. Therefore, (39) can be written as

$$L(\mathbf{n}; \mathbf{q}, \mathbf{p}) = \sum_{\nu=1}^{N_r} \log \left(\frac{p_\nu q_\nu^{2/p_\nu}}{2\pi \Gamma(2/p_\nu)} \right) - \mathcal{E}_{\mathbf{n}_k} \left\{ \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu} \right\}. \quad (40)$$

We further note that $\mathcal{E}_{\mathbf{n}_k} \left\{ \sum_{\nu=1}^{N_r} q_\nu |n_{k,\nu}|^{p_\nu} \right\} = \sum_{\nu=1}^{N_r} q_\nu m_n(p_\nu)$ where $m_n(p) \triangleq \mathcal{E}\{|n_{k,\nu}|^p\}$ is the p th moment of the underlying noise. Therefore we obtain

$$L(\mathbf{n}; \mathbf{q}, \mathbf{p}) = \sum_{\nu=1}^{N_r} \log \left(\frac{p_\nu q_\nu^{2/p_\nu}}{2\pi \Gamma(2/p_\nu)} \right) - \sum_{\nu=1}^{N_r} q_\nu m_n(p_\nu). \quad (41)$$

Close-form expressions for the noise moments $m_n(p)$ have been provided in [19] for different types of noise defined in Section 2. The corresponding noise moments can be used in (41) to arrive at a close-form expression for the LLF.

Eqs. (40) and (41) can be employed to calculate the LLF needed to solve the optimization problem formulated in (38). In particular, (41) can be used when the noise statistics are known (cf. Section 6.1) for offline LLF maximization whereas (40) cab be employed for adaptive minimization of the LFF in scenarios where only the noise samples are available (cf. Section 6.2).

5 Adaptive Metric Optimization

In practice, the type of noise impairing a secondary user system is not usually known *a priori* and change with time. In such scenarios, it is necessary to develop adaptive algorithms that solve the metric

optimization problem online only based on the noise samples observed at the receiver. Therefore, in this section, we present adaptive algorithms that enable adaptive BER minimization and adaptive parameter estimation based on the BER analysis and ML parameter estimation frameworks developed in Sections 3 and 4, respectively. For this purpose, due to the random nature of the optimization problem, a stochastic optimization algorithm has to be used. Although several types of stochastic optimization algorithms are available in the literature, numerical evidence shows that the finite-difference stochastic approximation (FDSA) algorithm [36] is the most suitable for the problem at hand. In essence, the FDSA algorithm iteratively optimizes an stochastic objective function by avoiding the computationally expensive calculation of the objective function's gradient.

For adaptive BER minimization, we use the asymptotic BER results obtained in Section 4 as objective function for the FDSA algorithm since the approximate upper bound derived in Section 3 is too cumbersome for adaptive optimization. In particular, based on (26) and (27) the objective function for the FDSA algorithm is given by

$$L_{k,\text{BER}}(\boldsymbol{\theta}) = \sum_{K=0}^{d_f N_r \bar{\alpha}_d} \sum_{i_1 + \dots + i_{N_r} = K} M_{\mathbf{n}_k}(i_1, \dots, i_{N_r}; \boldsymbol{\theta}) \quad (42)$$

with

$$M_{\mathbf{n}_k}(i_1, \dots, i_{N_r}; \boldsymbol{\theta}) \triangleq \frac{\frac{1}{N_e} \sum_{i=k}^{k+N_e-1} \prod_{\nu=1}^{N_r} \sum_{j_1 + \dots + j_d = i_\nu} \prod_{k=1}^{d_f} \xi_{j_k, \nu} |n_{k, \nu}|^{2j_k} \left(\sum_{k=1}^{d_f} \sum_{\nu=1}^{N_r} q_\nu |n_{k, \nu}|^{p_\nu} \right)^{2 \sum_{\nu=1}^{N_r} (d_f \alpha_d - i_\nu)/p_\nu}}{\Gamma \left(2 \sum_{\nu=1}^{N_r} (d \alpha_d - i_\nu)/p_\nu + 1 \right) \prod_{\nu=1}^{N_r} p_\nu^{d_f} q_\nu^{2(d_f \alpha_d - i_\nu)/p_\nu}} \quad (43)$$

where we have omitted all terms that do not affect the optimization and the vector $\boldsymbol{\theta} \triangleq [q_1 \dots q_{N_r} p_1 \dots p_{N_r}]^T$ contains all metric parameters to be optimized. Furthermore, $M_{\mathbf{n}_k}(i_1, \dots, i_{N_r}; \boldsymbol{\theta})$ is the instantaneous estimate for the generalized noise moment, and N_e denotes the length of the sliding window used to calculate this estimate. We note that multiplying the elements of \mathbf{q} with a constant does not affect the value of the objective function in (43), and therefore we can assume $q_1 = 1$ without loss of generality. We further note that $M_{\mathbf{n}_k}(i_1, \dots, i_{N_r}; \boldsymbol{\theta})$ can be significantly simplified if the fading or the metric parameters are constraint (e.g., $\alpha_d = 1$ or $q_\nu = 1$, $p_\nu = p$, $1 \leq \nu \leq N_r$), cf. Sections 3.2.1.

For adaptive ML parameter estimation, the objective function for the FDSA algorithm can be

formulated based on (40) as

$$L_{k,\text{MLE}}(\boldsymbol{\theta}) = \sum_{\nu=1}^{N_r} \log \left(\frac{p_\nu q_\nu^{2/p_\nu}}{2\pi \Gamma(2/p_\nu)} \right) - \frac{1}{N_e} \sum_{i=k}^{k+N_e-1} \sum_{\nu=1}^{N_r} q_\nu |n_{i,\nu}|^{p_\nu}. \quad (44)$$

Here, $q_1 = 1$ can not be assumed as it affects the value of the objective function.

With the objective functions at hand, the FDSA algorithm can be formulated as follows. The FDSA algorithm improves the parameter vector estimate at time k , $\boldsymbol{\theta}_k$, in the direction of the negative gradient vector estimate $\hat{\mathbf{g}}(\boldsymbol{\theta}_k)$ to obtain a new estimate $\boldsymbol{\theta}_{k+1}$ [36]:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \delta_k \hat{\mathbf{g}}(\boldsymbol{\theta}_k) \quad (45)$$

$$\hat{\mathbf{g}}(\boldsymbol{\theta}_k) = \left[\frac{L_{k,X}(\boldsymbol{\theta}_k + \zeta_k \mathbf{e}_1) - L_{k,X}(\boldsymbol{\theta}_k - \zeta_k \mathbf{e}_1)}{2\zeta_k} \dots \frac{L_{k,X}(\boldsymbol{\theta}_k + \zeta_k \mathbf{e}_{2N_r-1}) - L_{k,X}(\boldsymbol{\theta}_k - \zeta_k \mathbf{e}_{2N_r-1})}{2\zeta_k} \right]^T \quad (46)$$

where $X \in \{\text{BER}, \text{MLE}\}$, and $\delta_k > 0$ and $\zeta_k > 0$ are the gain sequences of the FDSA algorithm, and \mathbf{e}_i denotes a vector with a 1 in position i and 0s in all other positions. The convergence theory for the FDSA algorithm [36] states that if the gain sequence fulfills $\delta_k \rightarrow 0$, $\zeta_k \rightarrow 0$, $\sum_{k=0}^{\infty} \delta_k = \infty$, and $\sum_{k=0}^{\infty} \delta_k^2 / \zeta_k^2 < \infty$, under some mild conditions on the cost function, the algorithm is guaranteed to converge to a local minimum. However, in practice, it may be better to adopt $\delta_k = \delta$ and $\zeta_k = \zeta$, where δ and ζ are small constants, to give the algorithm some tracking capability.

In Section 6 we study and compare the performance of the proposed adaptive algorithms using a practical example (cf. Figs. 3, 4, and 5).

6 Numerical and Simulation Results

For numerical and simulation results presented in this section, the standard convolutional code with rate $R_c = 1/2$ and generator polynomials [133, 171] (octal representation) is adopted and higher code rates are obtained via puncturing. The considered type of fading, the assumed values for the SNR and number of antennas, and the employed code rate and modulation scheme are specified in the captions of the figures. The parameters for the considered types of noise are specified in the figure captions as well. Furthermore, the BER bound and asymptotic BER are calculated using (10) and (27), respectively.

6.1 Offline Metric Optimization

In Figs. 1 and 2 we consider a scenario where the noise statistics are known *a priori*, and therefore the task of metric optimization can be performed offline. For this scenario we illustrate how the BER analysis and ML parameter estimation techniques developed in Sections 3 and 4 can be used for offline BER minimization and offline ML parameter estimation, respectively. Although the assumption of known noise statistics is not realistic, it helps us study and compare the performance of the aforementioned approaches for metric optimization. To simplify the exposition, we assumed that $N_r = 1$ is valid and therefore the only metric parameters that have to be optimized are q and p , where we have dropped the indices for convenience. The more general case of $N_r \geq 1$ is considered in the next subsection.

We illustrate in Fig. 1 how the BER performance measures obtained in Section 3 for the secondary system can be used for offline metric optimization. Since for offline optimization computational complexity is not a major concern, we use the analytical BER bound for offline BER minimization. Thereby, assuming $q = 1$ without loss of generality, in Fig. 1 we have shown the BER bound for ACCI, NBI, and UF-ACCI noise vs. p for SNR = 15 dB. The value of p for which the BER bound is minimized is denoted by \hat{p}_{BER} and the corresponding point of the BER bound curve is marked by “*” markers. For comparison, we have also shown the BER obtained via Monte-Carlo simulation and the asymptotic BER. As seen, the BER bound, simulation and asymptotic results are in general in good agreement. The observed small differences between the curves are due to assuming a finite value for SNR in this figure. Nevertheless, Fig. 1 shows that for each type of noise the minimum BER happens at approximately the same value of p for all the three curves. Fig. 1 further shows that the BER of the secondary system strongly depends on the metric parameter p and therefore significant performance gains can be achieved by metric optimization.

The offline ML parameter estimation is illustrated in Fig. 2 for the same noise types as in Fig. 1. Here, $q = 1$ can not be assumed without loss of generality (cf. (40)). To avoid optimizing two parameters, we assume that q is determined based on p such that variance associated with the pdf $p_{GG}(z, q, p)$ is unity. Therefore, in this figure we have shown the LLF obtained using (40) for the considered noise types as function p . For each type of noise the value of p that maximized the LLF is denoted by \hat{p}_{MLE} and the corresponding point of the LLF curve is marked by “*” markers. We first note that the LFFs do not depend on the channel type, the value of SNR, or on the code

rate and modulation scheme. Furthermore, comparing Fig. 2 with Fig. 1 reveals that although the ML parameter estimation is a suboptimal approach for metric optimization, the parameter estimate \hat{p}_{MLE} is generally in good agreement with \hat{p}_{BER} and the incurred performance loss due to using \hat{p}_{MLE} instead of \hat{p}_{BER} is minimal.

6.2 Adaptive Metric Optimization

In Figs. 3, 4, and 5 we study and compare the performance of the adaptive BER minimization and adaptive ML parameter estimation algorithms proposed in Section 5 for a scenario where the noise statistics vary with time. In particular, we consider a BICM-OFDM secondary system impaired by NBI noise for which the number of interferers and their power change at time indexes $k = 5 \times 10^4$ and $k = 10 \times 10^4$, and are constant otherwise. The resulting impairments are denoted by N_1 , N_2 and N_3 , and details about corresponding noise parameters are provided in the caption of the figures. Furthermore, we have initialized both algorithms with $q_\nu = 1$, $1 \leq \nu \leq 2$, and $p_\nu = 2$, $1 \leq \nu \leq 2$ and shown the results for one typical adaptation process. In order to have fair comparison, for both algorithms we have adopted $\delta_k = \delta = 10^{-3}$, $\zeta_k = \zeta = 10^{-5}$ and $N_e = 1$, and have not used the normalization $q_1 = 1$ that is possible only for adaptive BER minimization.

In Figs. 3 and 4 we show the metric parameters estimates \mathbf{q} and \mathbf{p} resulting from adaptive BER minimization and adaptive parameter estimation, respectively. The corresponding asymptotic BERs as well as the asymptotic BER obtained for the L_2 -norm metric are shown and compared in Fig. 5. As seen, although similar gain sequences were chosen for both algorithms, the steady state values are reached relatively faster in Fig. 3 compared to Fig. 4. The reason can be explained by noting that the BER curves are relatively steeper than the LLF curves for similar values of \mathbf{q} and \mathbf{p} (cf. Section 6.1). Nevertheless, both algorithms can effectively adapt to variations in noise statistics and reach the steady state values fairly quickly. Fig. 5 reveals that both algorithms reach close to optimal BERs and achieve significant performance gains compared to L_2 -norm. Furthermore, this figure confirms the robustness of the proposed L_p -norm as the variations observed in the performance of the optimized L_p -norm are considerably smaller compared to the L_2 -norm.

6.3 Comparison with Other Metrics

We compare the performance of the optimized L_p -norm metric with that of L_2 -norm metric, optimized Huber's M-metric and ML metric in Fig. 6 for a BICM-SC secondary system impaired by TD-GMN noise. The L_p -norm metric was optimized using the offline approach described in Section 6 and the results are shown for both $p = \hat{p}_{\text{BER}}$ and $p = \hat{p}_{\text{MLE}}$. Huber's M-metric was optimized using simulation for SNR = 18 dB. For the L_2 -norm and optimized L_p -norm with $p = \hat{p}_{\text{BER}}$ the BER bound, asymptotic BER as well as simulation results as a function of the SNR. For the optimized L_p -norm with $p = \hat{p}_{\text{MLE}}$ only the asymptotic BER is included to avoid crowding the figure. For the Huber's M-metric only the simulation results are shown. Fig. 6 indicates that the performance loss due to using $p = \hat{p}_{\text{MLE}}$ instead of $p = \hat{p}_{\text{BER}}$ is negligible. Furthermore, it is observed that at BER = 10^{-6} the optimized L_p -norm metric outperforms the L_2 -norm metric and Huber's by 5 dB and 2 dB, respectively, and the gap to the optimal ML metric is less than 1 dB.

6.4 Effect of Different Channel Types

We consider the effect of different channel types on the performance of the proposed L_p -norm metric in Fig. 7. In particular, in Fig. 7, we show BER of a BICM-OFDM secondary system impaired by GMN-OFDM vs. SNR for Rayleigh, Rician, Nakagami-m and Weibul fading channels. For each fading type we show the BER bound, asymptotic BER, and simulation results for the optimized L_p -norm and L_2 -norm metrics. The L_p -norm was optimized using the adaptive BER minimization algorithm using a similar approach as explained in Section 6.2. Fig. 7 indicates that the optimized L_2 -norm metric achieves similar gains over the L_2 -norm metric for the different types of fading channels considered in this figure. This figure also confirms the validity of the asymptotic BER analysis presented in Section 3 as it shows that the simulations, the BER bound, and the asymptotic results are in great agreement for different types of fading, and in particular for Weibul fading for which $\alpha_d = 3/2$ is a non-integer. As expected, it is observed that the diversity gain is the same for L_2 -norm and L_p -norm metrics but differs for fading channels with different α_d .

6.5 Effect of UWB Interference

In Fig. 8 we study the performance of the proposed L_p -norm metric for a BICM-SC secondary system impaired by impulse radio (IR) UWB interference. We have shown the BER bound, asymptotic BER,

and simulated BER of the considered secondary system as a function of SNR for the L_2 -norm and optimized L_p -norm metrics. The results are shown for two values of B_s , where B_s denotes the total bandwidth of the BICM-SC system. The parameters for the IR UWB interference were adopted from the IEEE 802.15.4a standard [30]. Fig. 8 shows that the value of B_s which governs the impulsiveness of the noise [6], has a strong effect on the performance. This figure further suggests that the performance gain obtained using the optimized L_p -norm metric is larger when the underlying impairment is for more impulsive.

6.6 Comparison with Erasure Decoding

We compare the performance of erasure decoding with that of the optimized L_p -norm metric in Fig. 9 for a BICM-OFDM secondary system impaired by NBI noise. For the NBI noise we have assumed $B = 10$, $N_s = 64$ and a total of $N_I = 60$ interferers which are aligned with the center frequencies of the BICM-OFDM subcarriers. Erasure decoding is achieved by erasing the bit metrics associated with K_e affected subcarriers with largest interference powers. In this figure we have shown the simulation results for erasure decoding for different values of K_e as well as the BER bound and simulation results for L_2 -norm and the optimized L_p -norm metrics as a function of SNR. As seen, the performance of erasure decoding strongly depends on the number of erased subcarriers K_e and choosing the optimal K_e involves a tradeoff. In particular, increasing K_e improves the performance for low SNR's as the effect of interference from a larger number of subcarriers is suppressed. However, since erasure adversely affect the free distance of the underlying convolutional code, increasing K_e also results in the reduction of diversity gain and/or an error floor and therefore performance degradation at high SNR's. For the considered scenario the best tradeoff is achieved with $K_e = 30$. We note that for erasure decoding we have assumed ideal interference detection, i.e., we have assumed that the location of the affected BICM-OFDM subcarriers are perfectly known. In practice, the interference detection would be non-ideal and therefore the performance of erasure decoding would be worse than what shown in this figure. Nevertheless, Fig. 9 indicates that the optimized L_p -norm metric outperforms the best performance achieved by erasure decoding by more than 2.5 dB at $\text{BER} = 10^{-6}$.

7 Conclusions

In this paper, we have proposed an adaptive L_p -norm metric for secondary BICM systems operating in the presence of non-Gaussian noise and interference with time-varying statistics. For optimization of the proposed L_p -norm metric we have developed two different approaches based on minimizing the BER of the secondary system and ML parameter estimation. Based on both approaches we have derived FDSA algorithms that enable adaptive metric optimization. Our numerical and simulation results show that optimized L_p -norm decoding can substantially improve the performance of secondary systems in non-Gaussian noise environments and yields significant performance gains compared to other popular metrics. Furthermore we have shown that the proposed adaptive algorithms can effectively optimize the L_p -norm parameters in secondary environments with changing noise statistics.

A Pdf $f_y(y_{k,\nu})$ of $y_{k,\nu}$ for $\gamma \rightarrow \infty$

For convenience and without loss of generality, we drop all subscripts k and ν in this appendix. Furthermore, we first consider the pdf $f_X(X)$ of $X \triangleq |\sqrt{\gamma}he + n|^2$ and then calculate the pdf $f_y(y)$ of $y = X^{p/2}$ using

$$f_y(y) = 2/p f_X(y^{2/p}) y^{2/p-1}. \quad (47)$$

X can be expressed as

$$X = \gamma a^2 d_{xz}^2 + 2\sqrt{\gamma} d_{xz} a \Re\{\hat{n}\} + |n|^2, \quad (48)$$

where we introduced the definitions $\hat{n} \triangleq ne^{-j\Theta}$ and $d_{xz} \triangleq |e|$, and used the fact that $h = ae^{j\Theta}$ (cf. Section 2.2). The MGF of X can now be obtained as

$$\Phi_X(s) \triangleq \mathcal{E}_{a,\Theta}\{e^{-sX}\} = e^{-s|n|^2} \mathcal{E}_{a,\Theta}\{e^{-s\gamma a d_{xz}^2} e^{-s2\sqrt{\gamma} d_{xz} a \Re\{\hat{n}\}}\} \quad (49)$$

Using the Taylor series expansion $e^x = \sum_{i=0}^{\infty} x^i/i!$, we arrive at

$$\Phi_X(s) = e^{-s|n|^2} \mathcal{E}_{a,\Theta} \left\{ e^{-s\gamma a^2 d_{xz}^2} \sum_{i=0}^{\infty} \frac{(-2\sqrt{\gamma} d_{xz} a \Re\{\hat{n}\} s)^i}{i!} \right\}. \quad (50)$$

For $\gamma \rightarrow \infty$, (5) can be used along with the integral $\int_0^\infty x^{\mu-1} e^{-px^2} dx = p^{\mu/2} \Gamma(\mu/2)$ [37, 3.462] to rewrite (50) as

$$\begin{aligned}\Phi_X(s) &= \frac{\alpha_c e^{-s|n|^2}}{(\gamma d_{xz}^2 s)^{\alpha_d}} \sum_{i=0}^{\infty} 2^i \Gamma(\alpha_d + i/2) \mathcal{E}_\Theta\{\Re\{\hat{n}\}^i\} s^{i/2} + o(\gamma^{-\alpha_d}) \\ &= \frac{\alpha_c e^{-s|n|^2}}{(\gamma d_{xz}^2 s)^{\alpha_d}} \sum_{i=0}^{\infty} \beta_i |n|^{2i} s^i + o(\gamma^{-\alpha_d}).\end{aligned}\quad (51)$$

where $\beta_i \triangleq \Gamma(\alpha_d + i)/(i!)^2$. In deriving (51), we have used

$$\mathcal{E}_\Theta\{\Re\{\hat{n}\}^i\} = \begin{cases} \frac{i/2+1/2}{\sqrt{\pi}\Gamma(i/2+1)} |n|^i, & i \text{ even} \\ 0, & i \text{ odd} \end{cases}.\quad (52)$$

Using the Taylor series expansion $e^{-s|n|^2} = \sum_{i=0}^{\infty} (-1)^i |n|^{2i} s^i / i!$ in (51) yields

$$\Phi_X(s) = \frac{\alpha_c}{(\gamma d_{xz}^2 s)^{\alpha_d}} \sum_{i=0}^{\infty} P_i |n|^{2i} s^i + o(\gamma^{-\alpha_d}).\quad (53)$$

where

$$P_i \triangleq \sum_{\lambda=0}^i \frac{(-1)^{(i-\lambda)} \Gamma(\nu + \lambda)}{(\lambda!)^2 (i - \lambda)!}.\quad (54)$$

If α_d is an integer it can be shown that

$$P_i = \begin{cases} \binom{\nu-1}{i}^2 (\alpha_d - i - 1)!, & 0 \leq i \leq \alpha_d - 1 \\ 0, & \alpha_d - 1 < i \end{cases}.\quad (55)$$

Thus, for integer α_d , $\Phi_X(s)$ is a finite power series in s , whereas it is an infinite power series for non-integer α_d , cf. (53). However, truncating the power series for non-integer α_d after $\bar{\alpha}_d = \lceil \alpha_d \rceil - 1$ terms results in a good approximation for $\Phi_{X_{k,l}}(s)$. Applying this truncation, for $\gamma \rightarrow \infty$, based on (53), we obtain via the inverse Laplace transform

$$f_X(X) = \frac{\alpha_c}{(\gamma d_{xz}^2)^{\alpha_d}} \sum_{i=0}^{\bar{\alpha}_d} \frac{P_i}{\Gamma(\alpha_d - i)} |n|^{2i} X^{\alpha_d - i - 1} + o(\gamma^{-\alpha_d}).\quad (56)$$

Finally, $f_y(y)$ for $\gamma \rightarrow \infty$ can be found based on (56) and (47) as

$$f_y(y) = \frac{2\alpha_c}{p(\gamma d_{xz}^2)^{\alpha_d}} \sum_{i=0}^{\bar{\alpha}_d} \frac{P_i}{\Gamma(\alpha_d - i)} |n|^{2i} y^{2(\alpha_d - i)/p - 1} + o(\gamma^{-\alpha_d}).\quad (57)$$

Note that (57) is exact for integer α_d and an approximation for non-integer α_d .

References

- [1] F. C. C. (FCC). Spectrum Policy Task Force Report. *Tech. Rep. TR 02-155*, November 2002.
- [2] A. Petrin and P.G. Steffes. Analysis and comparison of spectrum measurements performed in urban and rural areas to determine the total amount of spectrum usage. *Proc. of the ISART*, March 2005.
- [3] J. Mitola. Cognitive Radio: an Integrated Agent Architecture for Software Defined Radio. *Ph.d. thesis, KTH Royal Inst. of Tech., Stockholm, Sweden*, 2000.
- [4] M. Di Benedetto, T. Kaiser, A. Molisch, I. Oppermann, C. Politano, and D. Porcino (Eds.). *UWB Communication Systems*. Hindawi, New York, 2006.
- [5] K. Watanabe, K. Ishibashi, and R. Kohno. Performance of Cognitive Radio Technologies in the Presence of Primary Radio Systems. In *Proc. IEEE Intern. Symp. Person., Indoor and Mobile Radio Commun. (PIMRC)*, pages 1–5, September 2007.
- [6] A. Nasri, R. Schober, and L. Lampe. Performance Evaluation of BICM-OFDM Systems Impaired by UWB Interference. In *Proceedings of IEEE Intern. Commun. Conf. (ICC)*, July 2008.
- [7] D. Middleton. Statistical-physical Models of Man-made Radio Noise – Parts I and II. *U.S. Dept. Commerce Office Telecommun.*, April 1974 and 1976.
- [8] P. Huber. *Robust Statistics*. Wiley, New York, 1981.
- [9] T. Aysal and K. Barner. Meridian Filtering for Robust Signal Processing. *IEEE Trans. Signal Processing*, 55:39493962, August 2007.
- [10] R.E. Carrillo, T.C. Aysaft, and K.E. Barner. Generalized Cauchy Distribution Based Robust Estimation. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3389–3392, March 2008.
- [11] G. Shevlyakov and K. Kim. Robust Minimax Detection of a Weak Signal in Noise With a Bounded Variance and Density Value at the Center of Symmetry. *IEEE Trans. Inform. Theory*, 52:12061211, March 2006.
- [12] G. Caire, G. Taricco, and E. Biglieri. Bit-Interleaved Coded Modulation. *IEEE Trans. Inform. Theory*, 44:927–946, May 1998.
- [13] H. Bölcskei. MIMO-OFDM Wireless Systems: Basics, Perspectives, and Challenges. *IEEE Wireless Commun.*, 13:31–37, August 2006.
- [14] ECMA. Standard ECMA-368: High Rate Ultra Wideband PHY and MAC Standard. [Online] <http://www.ecma-international.org/publications/standards/Ecma-368.htm>, December 2005.
- [15] E. Hossain and V. Bhargava (Eds.). *Cognitive Wireless Communications Networks*. Springer, New York, 2007.
- [16] C. Snow, L. Lampe, and R. Schober. Interference Mitigation for Coded MB-OFDM UWB. In *Proc. IEEE Radio and Wireless Symposium*, pages 17–20, January 2008.

- [17] T. Li, W. Mow, V. Lau, M. Siu, R. Cheng, and R. Murch. Robust Joint Interference Detection and Decoding for OFDM-Based Cognitive Radio Systems with Unknown Interference. *IEEE J. Select. Areas Commun.*, 25:566–575, April 2007.
- [18] T. Li, W. Mow, and M. Siu. Joint Erasure Marking and Viterbi Decoding Algorithm for Unknown Impulsive Noise Channels. *IEEE Trans. Wireless Commun.*, 7:3407–3416, September 2008.
- [19] A. Nasri and R. Schober. Performance of BICM-SC and BICM-OFDM Systems with Diversity Reception in Non-Gaussian Noise and Interference. *Submitted to IEEE Trans. on Commun.* [Online] <http://www.ece.ubc.ca/~amirn/TCOM-08.pdf>, 2008.
- [20] A. Nasri, A. Nezampoor, and R. Schober. Adaptive L_p -Norm Diversity Combining in Non-Gaussian Noise and Interference. *Submitted to IEEE Trans. on Wireless Commun.* [Online] <http://www.ece.ubc.ca/~amirn/TW-08.pdf>, 2008.
- [21] P.-C. Yeh, S. Zummo, and W. Stark. Error Probability of Bit-Interleaved Coded Modulation in Wireless Environments. *IEEE Trans. Veh. Technol.*, 55:722–728, March 2006.
- [22] D. Rende and T. Wong. Bit-Interleaved Space-Frequency Coded Modulation for OFDM Systems. *IEEE Trans. Wireless Commun.*, 4:2256–2266, September 2005.
- [23] G.A. Tsihrintzis and C.L. Nikias. Performance of Optimum and Suboptimum Receivers in the Presence of Impulsive Noise Modeled as an Alpha-Stable Process. *IEEE Trans. Commun.*, 43:904–914, Feb./Mar./Apr. 1995.
- [24] T.S. Rappaport. *Wireless Communications*. Prentice Hall, Upper Saddle River, NJ, 2002.
- [25] A. Giorgetti and M. Chiani. Influence of Fading on the Gaussian Approximation for BPSK and QPSK with Asynchronous Cochannel Interference. *IEEE Trans. Wireless Commun.*, 4:384–389, March 2005.
- [26] R. Prasad, A. Kegel, and A. de Vos. Performance of Microcellular Mobile Radio in a Cochannel Interference, Natural, and Man-Made Noise Environment. *IEEE Trans. Veh. Technol.*, 42:33–40, February 1993.
- [27] H. Nguyen and T. Bui. Bit-Interleaved Coded Modulation With Iterative Decoding in Impulsive Noise. *IEEE Trans. Power Delivery*, 22:151–160, January 2007.
- [28] C. Tepedelenlioglu and P. Gao. On Diversity Reception Over Fading Channels with Impulsive Noise. *IEEE Trans. Veh. Technol.*, 54:2037–2047, November 2005.
- [29] A. Coulson. Bit Error Rate Performance of OFDM in Narrowband Interference with Excision Filtering. *IEEE Trans. Wireless Commun.*, 5:2484–2492, September 2006.
- [30] IEEE P802.15.4a. Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs). January 2007.
- [31] E. Biglieri, G. Caire, G. Taricco, and J. Ventura-Traveset. Computing Error Probabilities over Fading Channels: a Unified Approach. *European Trans. Telecommun.*, 9:15–25, Jan./Feb. 1998.

- [32] V. Sethuraman and B. Hajek. Comments on "Bit-Interleaved Coded Modulation". *IEEE Trans. Inform. Theory*, 52:1795–1797, April 2006.
- [33] A. Martinez, A.G. Fabregas, and G. Caire. Error Probability Analysis of Bit-Interleaved Coded Modulation. *IEEE Trans. Inform. Theory*, 52:262–271, June 2006.
- [34] Z. Wang and G. Giannakis. A Simple and General Parameterization Quantifying Performance in Fading Channels. *IEEE Trans. Commun.*, 51:1389–1398, August 2003.
- [35] M.K. Varanasi. Parameter estimation of the generalize gaussian noise model. M.S. thesis, Rice University, Houston, TX, 1986.
- [36] J. Spall. *Introduction to Stochastic Search and Optimization*. Wiley & Sons, Inc., New Jersey, 2003.
- [37] I. Gradshteyn and I. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 2000.

Tables and Figures:

Table 1: Pdf $p_a(a)$ of fading amplitude a and fading parameters α_c and α_d for various fading models.

Channel type	$p_a(a)$ of the fading amplitude a	α_c	α_d
Rayleigh	$2 a e^{-a^2}$	$\det(\mathbf{C}_{hh})^{-1/N_R}$	1
Ricean	$2(K+1) a e^{-K-(1+K)a^2} I_0\left(2a\sqrt{K(K+1)}\right)$	$\left(\frac{\exp\left(-\boldsymbol{\mu}_h^H \mathbf{C}_{hh}^{-1} \boldsymbol{\mu}_h\right)}{\det(\mathbf{C}_{hh})}\right)^{1/N_R}$	1
Nakagami- m	$\frac{2}{\Gamma(m)} m^m a^{2m-1} e^{-ma^2}$	$\frac{m^m}{\Gamma(m)} \det(\mathbf{C}_{aa})^{-m/N_R}$	m
Nakagami- q	$\frac{(1+q^2)a}{q} \exp\left(-\frac{(1+q^2)^2 a^2}{4q^2}\right) I_0\left(\frac{(1-q^4)a^2}{4q^2}\right)$	$\frac{1+q^2}{2q}$	1
Weibull	$c \left(\Gamma(1 + \frac{2}{c})\right)^{\frac{c}{2}} a^{c-1} \exp\left(-\left(a^2 \Gamma(1 + \frac{2}{c})\right)^{\frac{c}{2}}\right)$	$\frac{c}{2} \left(\Gamma(1 + \frac{2}{c})\right)^{\frac{c}{2}}$	$\frac{c}{2}$

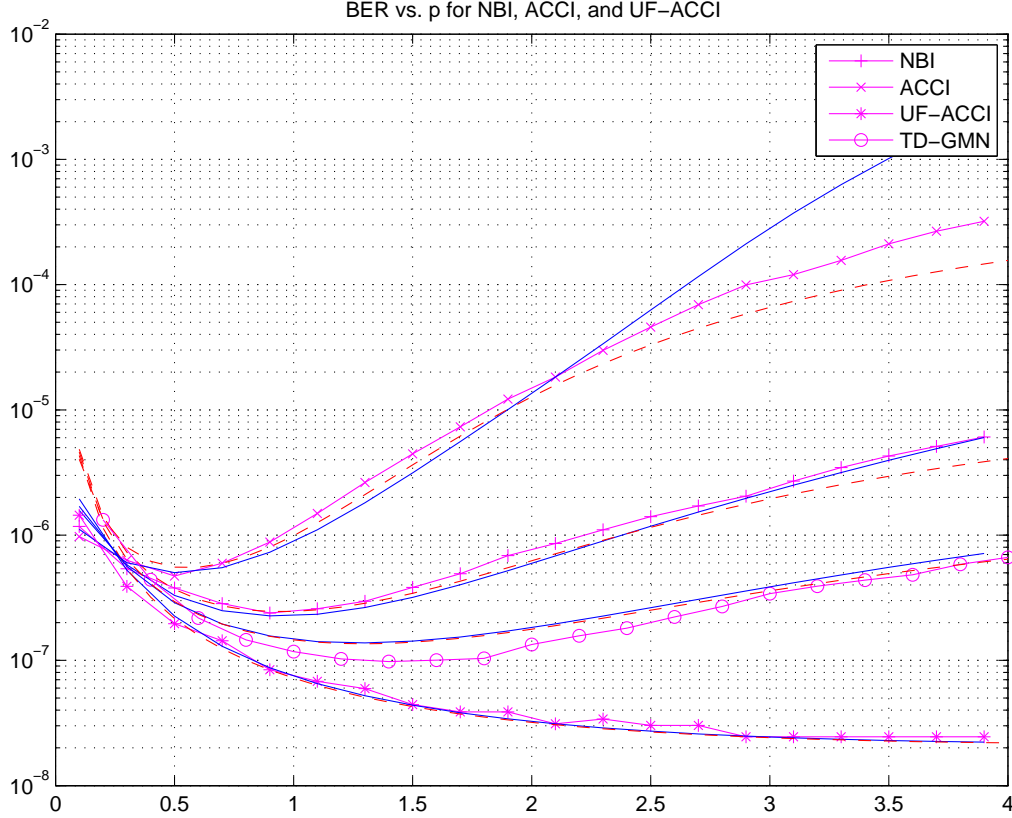


Figure 1: BER of BICM-SC and BICM-OFDM impaired by various types of noise vs. p . Rayleigh fading, $SNR = 16$ dB, $R_c = 3/4$, 4-PSK, and $N_r = 1$. GMN-TD: ϵ -mixture noise, $\epsilon = 0.1$, $\kappa = 3$. ACCI: One Rayleigh faded 4-PSK CCI signals, $B = 10$, $I_1 = 1$, $I_\mu = 0$, $2 \leq \mu \leq 10$, raised cosine pulses $g_{1,1}(t)$, with roll-off factor 0.3, $\tau_{1,1} = 0.3T$, $\kappa = 10$. NBI: Five equal power sub-carrier-centered Rayleigh faded NBI signals, $N = 64$, $B = 1$, $I_1 = 5$, $\kappa = 4$. UF-ACCI: One 4-PSK CCI signals, $B = 1$, $I_1 = 1$, raised cosine pulses $g_{1,1}(t)$, with roll-off factor 0.3, $\tau_{1,1} = 0.3T$, $\kappa = 10$. Solid lines with markers: Simulated BER. Solid lines without markers: BER bound (??). Dashed lines: Asymptotic BER (??).

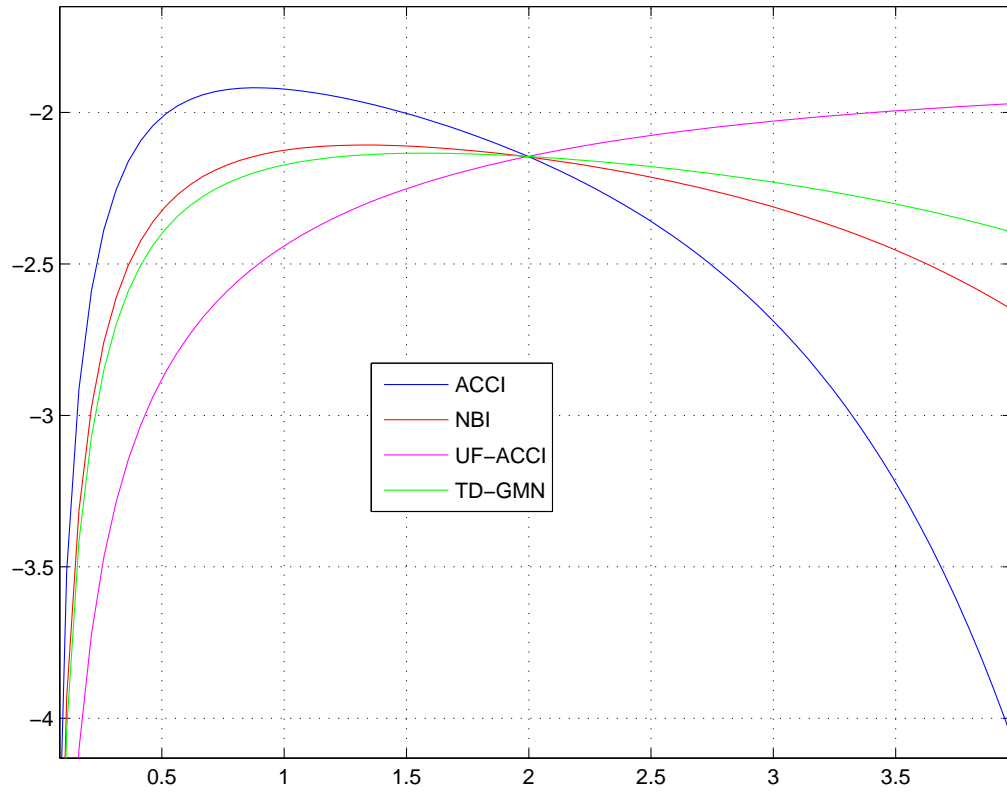


Figure 2: LLFs vs. p for the same types of noise as in Fig. 1

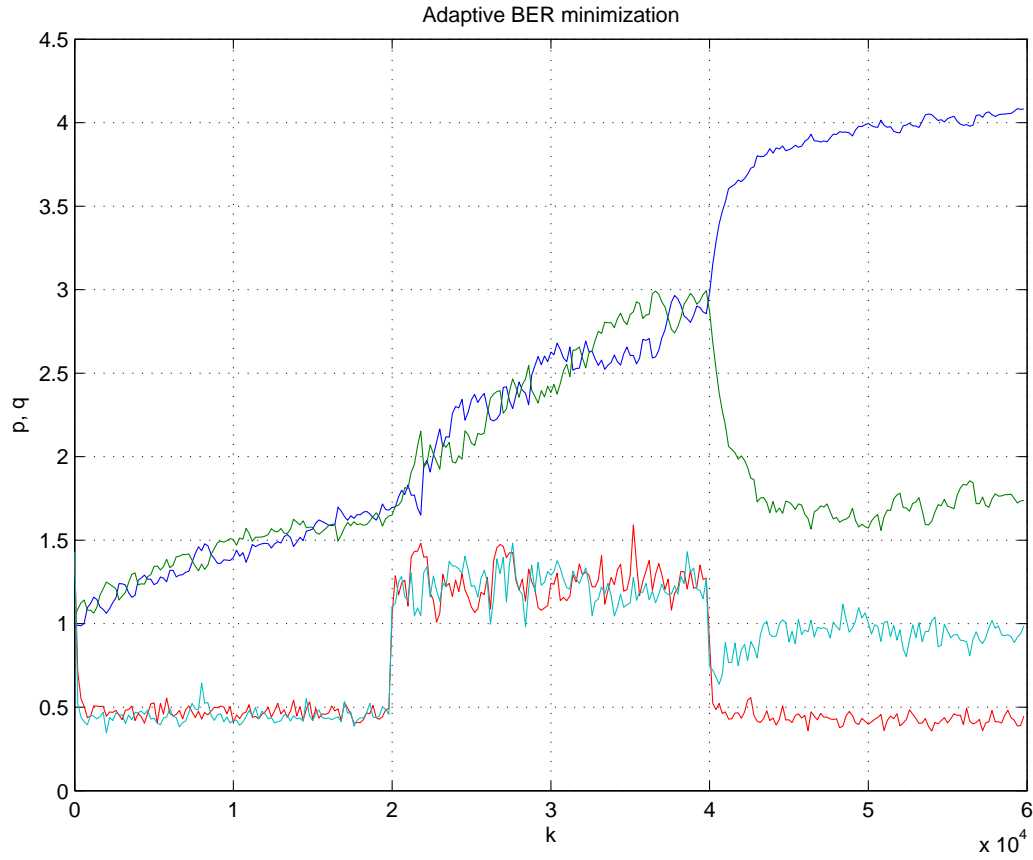


Figure 3: Metric coefficients q_ν , and p_ν , vs. iteration t of FDSA algorithm. Rayleigh fading, $SNR = 11$ dB, $R_c = 7/8$, 4-PSK, and $N_R = 2$. N1, N2 and N3, NBI with $N = 64$ and $B = 1$ with equal power sub-carrier-centered Rayleigh faded NBI signals. N1: $I_1 = 10$, $\kappa_1 = \kappa_2 = 40$. N2: $I_1 = 20$, $\kappa_1 = \kappa_2 = 4$. N3: $I_1 = 10$, $\kappa_1 = 40$, $\kappa_2 = 10$.

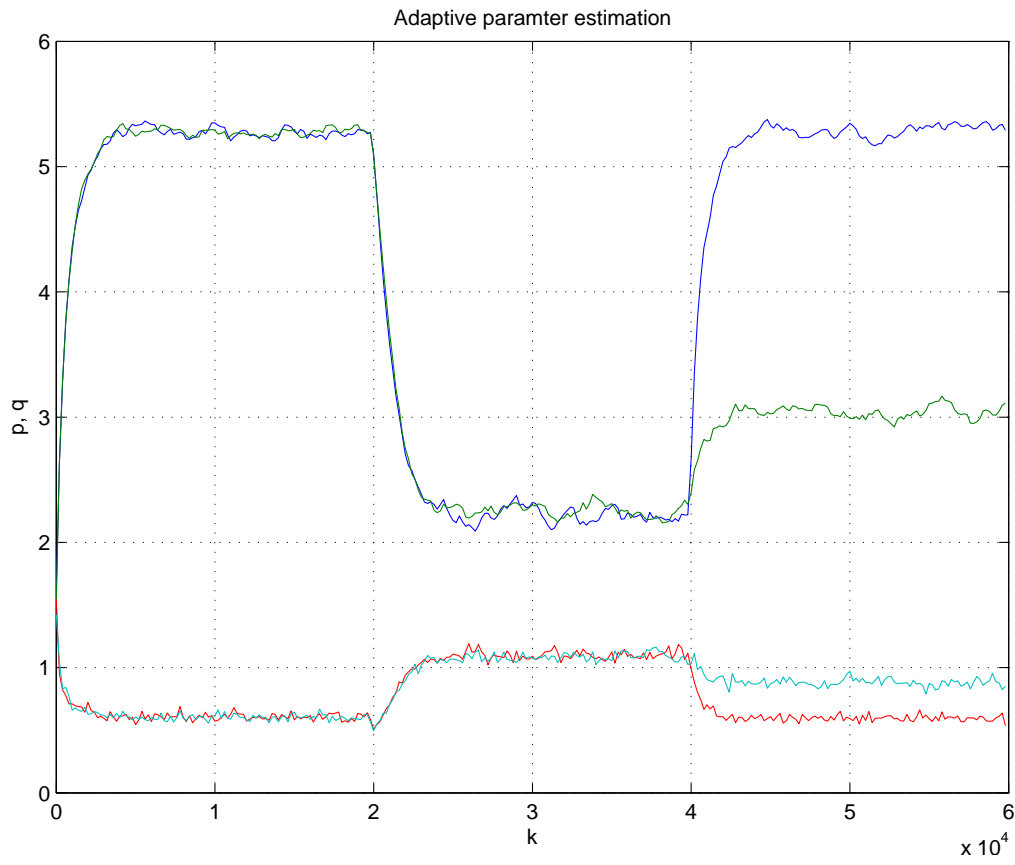


Figure 4: Parameters for N1–N3 are specified in the caption of Fig. 3

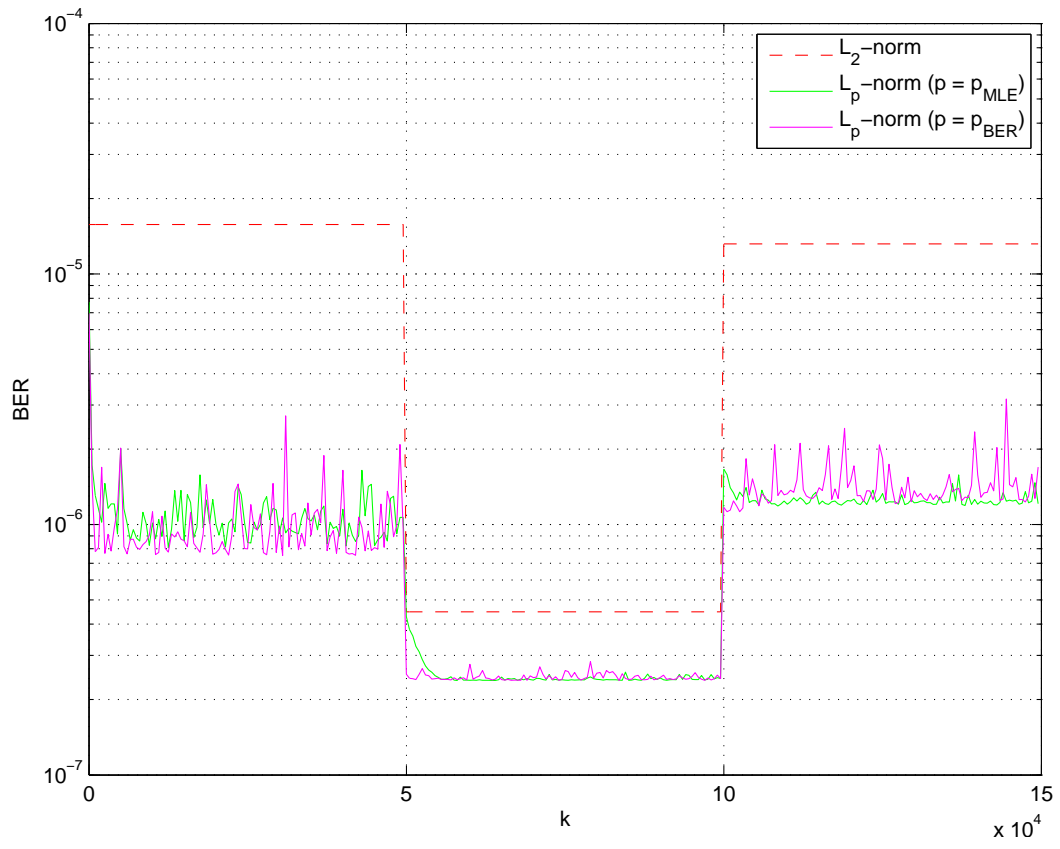


Figure 5: Parameters for N1–N3 are specified in the caption of Fig. 3

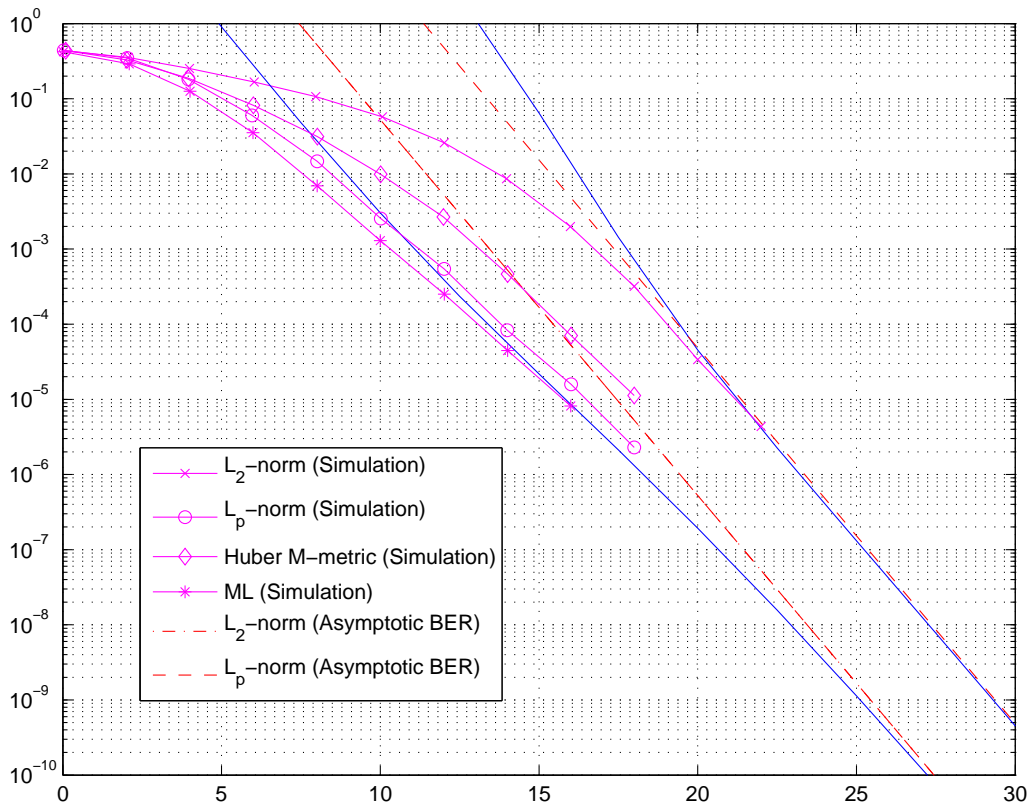


Figure 6: BER vs. SNR. Rayleigh fading, $R_c = 3/4$, 16-QAM, and $N_r = 1$. TD-GMN: ϵ -mixture noise with $\epsilon = 0.05$ and $\kappa = 50$.

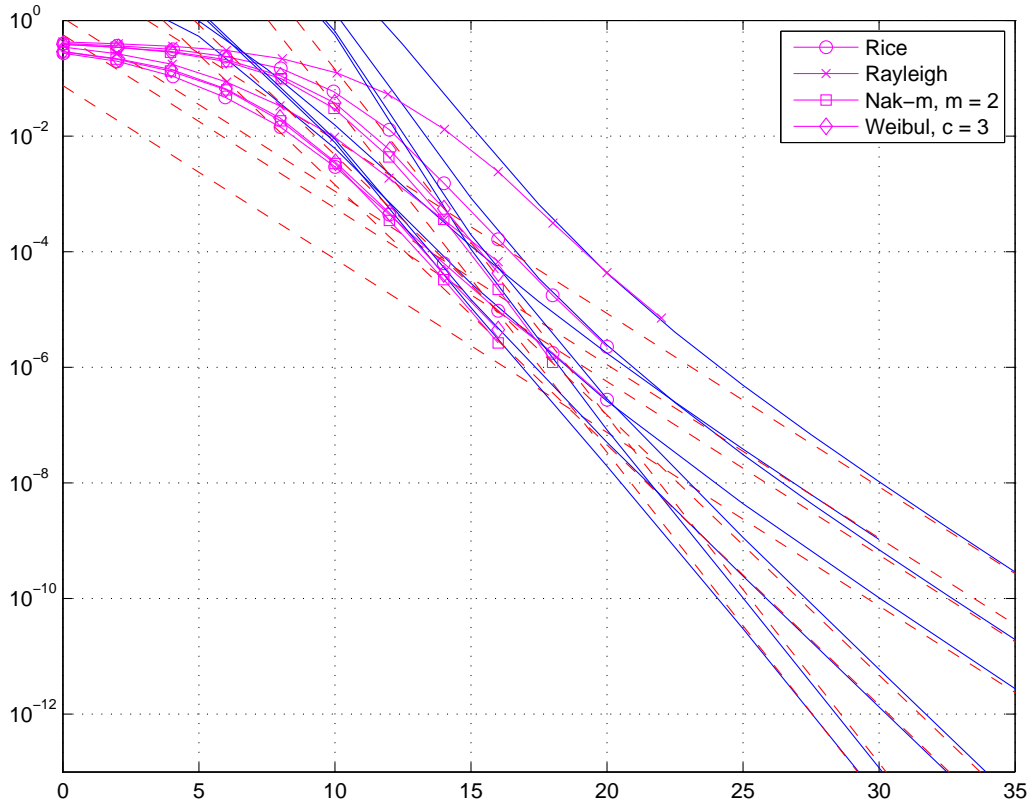


Figure 7: BER of BICM-SC impaired by TD-GMN vs. SNR for different types of fading channels. $R_c = 7/8$, 4-PSK, and $N_r = 1$. TD-GMN: ϵ -mixture noise with $\epsilon = 0.1$ and $\kappa = 100$.

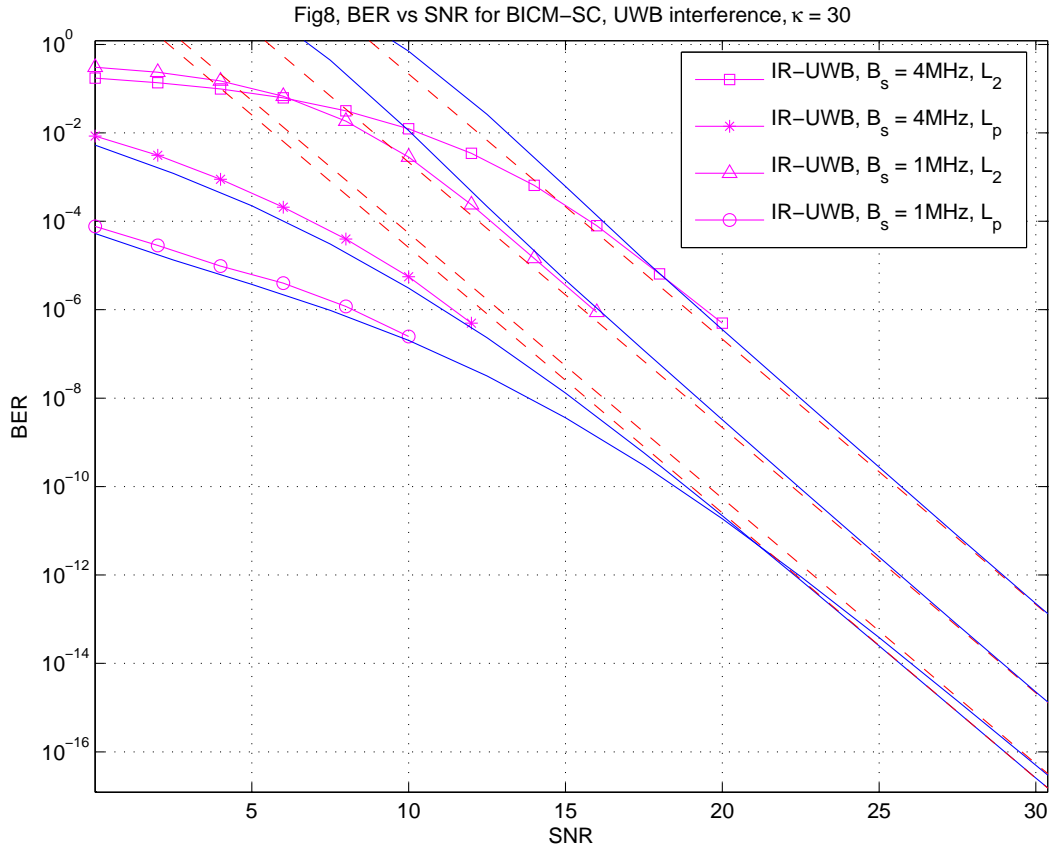


Figure 8: BER of BICM-OFDM system ($N = 64$) with sub-carrier spacing Δf_s impaired by IR-UWB [30] vs. SNR. IR-UWB: $N_b = 32$ bursts per symbol and $L_c = 128$ chips per burst. Rayleigh fading, $R_c = 2/3$, 4-PSK, and $N_r = 1$. Solid lines with markers: Simulated BER. Solid lines without markers: BER bound (??). Dashed lines: Asymptotic BER (??).

erasure decoding, BER vs SNR for BICM-OFDM, $R_c = 3/4$, NBI with $B = 10$, $N_s = 64$, $\varepsilon = 60/640 = 0.09$, $\kappa = 100$

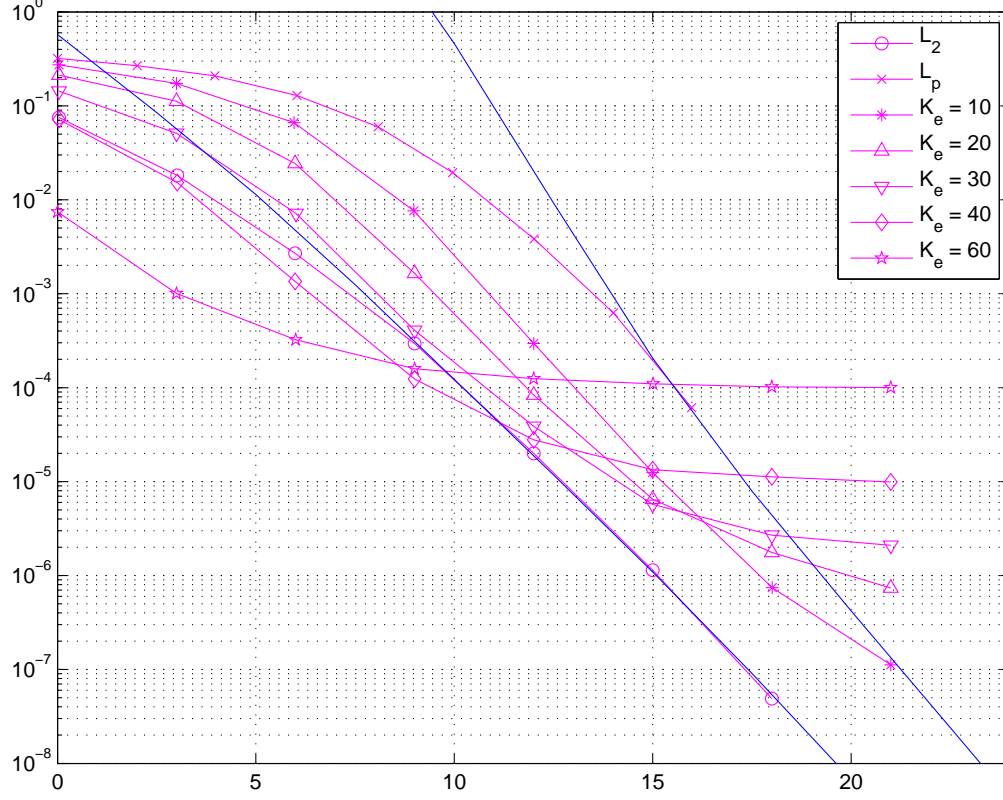


Figure 9: BER of BICM-OFDM system ($N = 64$) with K_e erasures vs. SNR. Rayleigh fading, $R_c = 3/4$, NBI: $B = 10$, $I_\mu = 6$ for $1 \leq \mu \leq 10$, $\kappa = 100$.