Robust L_p–Norm Metric for BICM Systems

Amir Nasri and Robert Schober
University of British Columbia, $\{$amirn, rschober$\}$@ece.ubc.ca

***Abstract*— Cognitive radio (CR) system are capable of using the frequency spectrum more effectively by utilizing unoccupied or under–utilized frequency bands. The frequency bands used by CR systems however, are expected to suffer from various forms of noise and interference with non–Gaussian distribution such as the co–channel interference caused by the primary user and other cognitive radios, ultra–wideband (UWB) interference and man–made impulsive noise. To mitigate the harmful effect of non–Gaussian noise and interference, we propose a robust L_p–norm metric for CR systems that employ the popular combination of bit–interleaved coded modulation (BICM) and orthogonal frequency division multiplexing (OFDM). We propose two approaches for metric optimization based on BER performance analysis and Maximum–Likelihood parameter estimation principal, respectively. In both cases we provide effcient adaptive algorithms that can be used for online metric optimization. We show that the proposed adaptive algorithm can effectively mitigate the adverse effects of non–Gaussian noise in practical scenarios where noise statistics vary with time.**

I. 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. At the same time, 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 which are overlaid and underlaid on existing licensed (primary) systems, respectively. The idea behind both CR and UWB is to give secondary users opportunistic spectrum access as long as they do not cause noticeable interference to primary users. 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 can 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 noise[1] is dominant. This motivates the use of robust metrics that perform well for a large class of noises with (possibly) non–Gaussian distribution. These metrics may also have tunable parameters that can be adjusted to the possibly 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 parameter p is adjusted accordingly.

In this paper, we consider secondary systems employing bit–interleaved coded modulation (BICM) [12] in combination with orthogonal frequency division multiplexing (OFDM) modulation. The motivation for considering BICM–OFDM systems is twofold. Firstly, BICM–OFDM is very efficient in exploiting the frequency diversity of wireless fading channels. Secondly, this technique has 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]. Here, we propose a robust L_p–norm decoding metric for secondary BICM–OFDM systems to mitigate the harmful effects of non–Gaussian impairments. The L_p–norm metric has a tunable paramter p and therefore can be optimized by properly adjusting this parameter. For metric optimization we propose two approaches. In the first approach, we develop a general mathematical framework for bit–error rate (BER) performance analysis of the secondary systemsthat allows us to obtain an accurate approximate upper bound for the BER as well as a closed–form expression for the asymptotic BER. The aforementioned performance measures are both obtained as a function of the metric parameter p and thus can be employed for metric optimization. In the second approach we develop a maximum–likelihood (ML) estimator for the metric parameter p based on the noise samples observed at the receiver. In both cases we provide effcient adaptive algorithms that enable online metric optimization. Using numerical and simulation results, we show that the optimized L_p–norm metric can achieve significant performance gains compared to the conventional L_2–norm metric in secondary user environments with non–Gaussian noise. Furthermore, we study the performance of the proposed adaptive agroithms in a practical scenario with time–varying noise statistics and show that these algorithms are very effective in dealing with the harmful effects of non–Gaussian noise.

The rest of this paper is organized as follows. In Section II, the system model for the considered secondary communication systems is introduced. The BER analysis framework is provided in Section III, and the metric parametere estimation is developed in Section IV. In Section V, the adaptive metric optimization is considered, and analytical and simulation results are presented in Section VI. Finally, some conclusions

This work was supported in part by an NSERC Strategic Project Grant (STPGP 350451) and in part by Bell University Laboratories, Canada.

[1]To 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".

are drawn in Section VII.

II. System Model

In this section, we consider BICM–OFDM secondary systems employing L_p–norm decoding and describe the corresponding signal model and the L_p–norm metric. We also present the models for 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.

A. Signal Model

The transmitter for the considered secondary system consists of a BICM encoder and an OFDM modulator with N sub–carriers. The BICM encoder comprises a convolutional encoder of rate R_c, an interleaver, and a memoryless mapper [12]. The codeword $\boldsymbol{c} \triangleq [c_1, c_2, \ldots, c_{m_c K_c}]$ of length $m_c K_c$ is generated by the convolutional encoder and is interleaved by the interleaver. The interleaved bits are broken up into blocks of m_c bits, 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 frame $\boldsymbol{x} \triangleq [x_1, x_2, \ldots, x_{K_c}]$ of length K_c. We assume that one codeword spans B OFDM symbols, i.e., $K_c = BN$, and that the length of the OFDM cyclic prefix exceeds the length of the channel impulse response. Furthermore, we assume that the channel changes independently from OFDM symbol to OFDM symbol, which can be achieved by frequency hopping. For example, the ECMA multi–band OFDM (MB–OFDM) ultra–wideband (UWB) system employs interleaving and coding over $B = 3$ (future versions of the standard may use up to $B = 15$) frequency–hopped OFDM symbols [14]. Assuming perfect synchronization and OFDM demodulation, the received signal can be written as

$$r_k = \sqrt{\gamma}\, h_k\, x_k + n_k, \qquad 1 \leq k \leq K_c, \tag{1}$$

where h_k and n_k with $\mathcal{E}\{|h_k|^2\} = \mathcal{E}\{|n_k|^2\} = 1$ are the fading gains and the noise variables, respectively, and γ denotes the SNR.[2] In this paper, we consider Rayleigh fading which implies that h_k is a zero–mean complex Gaussian random variable (RV). Therefore, the fading gains h_k can be expressed as $h_k \triangleq a_k e^{j\Theta_k}$, where a_k and Θ_k are independent RVs. Specifically, Θ_k is uniformly distributed in $[-\pi, \pi)$ and a_k is a positive real RV that follows a Rayleigh distibution. As customary in the literature, cf. e.g. [12], [16], [17], for our performance analysis we assume perfect interleaving, which means that h_k and n_k can be modeled as independent, identically distributed (i.i.d.) RVs and only their first order probability density functions (pdfs) are relevant. The i.i.d. assumption is justified for severely frequency selective channels and/or sufficiently large B.

[2]In this paper, $\mathcal{E}_x\{\cdot\}$ denotes statistical expectation with respect to x. Furthermore, we use the notation $u \doteq v$ to indicate that u and v are asymptotically equivalent, and a function $f(x)$ is $o(g(x))$ if $\lim_{x\to 0} f(x)/g(x) = 0$.

B. 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

$$\lambda_i(r_k, b) \triangleq \min_{x_k \in \mathcal{X}_b^i} \{f_m(u_k)\} \tag{2}$$

where $u_k \triangleq |r_k - \sqrt{\gamma}\, h_k\, x_k|$, and $\mathcal{X}_b^i$ is the subset of all symbols in constellation $\mathcal{X}$ whose label has value $b \in \{0, 1\}$ in position i, and $f_m(\cdot)$ is a suitably chosen function that depends on the considered metric. For L_p–norm metric considered in this paper $f_m(u) = u^p$ is valid. To achieve high performance the parameter p should be adapted to the underlying type of noise. We note that for the special case $p = 2$, (2) is the well–known L_2–norm branch metric which is typically used in AWGN channels [12].

C. 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 $\boldsymbol{n}_k$ 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 [18]. 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 practically relevant noise models that are frequency encountered in secondary user environments. In particular, we consider narrowband interference (NBI) and frequency–domain Gaussian mixture noise (GMN). These noise models are used in Section V for performance valuation of the proposed L_p–norm metric.

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

$$n_{k,\mu} = \sum_{i=1}^{I_\mu} g_{k,\mu}[i] b_\mu[i] \tilde{h}_{k,\mu}[i] + \tilde{n}_{k,\mu}, \tag{3}$$

where $\tilde{h}_{k,\mu}[i]$ 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 ith interferer at the μth hopping frequency and $g_{k,\mu}[i] \triangleq \exp[-j\pi(N-1)(k+f_{\mu,i}/\Delta f_s)/N + \phi_{\mu,i}] \sin[\pi(k + f_{\mu,i}/\Delta f_s)]/\sin[\pi(k+f_{\mu,i}/\Delta f_s)/N]$ [19]. Here, $f_{\mu,i}$ and $\phi_{\mu,i}$ denote the frequency and phase of the ith 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 VI. We note that for $K_i \to \infty$, the interference channel gains $\tilde{h}_{k,\mu}[i]$ are constant values. The resulting noise will be referred to as unfaded NBI (UF–NBI) in the rest of this paper.

2) GMN: 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. For GMN noise the pdf of n_k is given by [7]

$$p_n(n_k) = \sum_{i=1}^{I} \frac{c_i}{\pi\sigma_i^2} \exp\left(-\frac{|n_k|^2}{\sigma_i^2}\right), \quad (4)$$

where $c_i > 0$ and $\sigma_i^2 > 0$ are parameters. 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 denotes the fraction of subcarriers affected by the impulsive noise and κ is the ratio of the variances of the Gaussian background noise and the impulsive noise.

III. BER Analysis

In this section, we analyze the BER performance of secondary BICM–OFDM systems employing L_p–norm decoding in non–Gaussian noise environments. 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 arive at a closed–form expression for the asymptotic BER. The BER bound and asymptotic BER are both derived as a function of p and therefore can serve as objective functions for metric optimization.

A. Approximate Upper Bound for BER

Here, we provide an approximate upper bound for the BER performance of the cosidered secondary BICM–OFDM systems. We note that our derivation is based on the expurgated union bound in [12], and therefore we cannot prove that the obtained bound is a true upper bound (see discussion in [20], [21]). However, numerical evidence in e.g. [21], [22] suggests that the expurgated union bound does result in tight upper bounds if Gray labeling is applied. Our own results in Section VI confirm this conjecture.

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]

$$P_b \le \frac{1}{k_c} \sum_{d=d_f}^{\infty} w_c(d)\, P(\boldsymbol{c}, \hat{\boldsymbol{c}}), \quad (5)$$

where $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ are two distinct code sequences with Hamming distance d that differ only in $l \ge 1$ 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. $P(\boldsymbol{c}, \hat{\boldsymbol{c}})$ is the pairwise error probability (PEP), i.e., the probability that the decoder chooses code sequence $\hat{\boldsymbol{c}}$ when code sequence $\boldsymbol{c} \ne \hat{\boldsymbol{c}}$ is transmitted. Adopting the expurgated bound from [12], the PEP can be expressed as

$$P(\boldsymbol{c}, \hat{\boldsymbol{c}}) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \prod_{k=1}^{d} \left(\frac{1}{m_c 2^{m_c}} \sum_{i=1}^{m_c} \sum_{b=0}^{1} \sum_{x_k \in \mathcal{X}_b^i} \Phi_\Delta(s) \right) \frac{\mathrm{d}s}{s}, \quad (6)$$

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

$$\Delta(x_k, z_k) \;\triangleq\; |r_k - \sqrt{\gamma}\, h_k\, z_k|^p - |r_k - \sqrt{\gamma}\, h_k\, x_k|^p \quad (7)$$

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 $r_k = \sqrt{\gamma}\, h_k\, x_k + n_k$, we can rewrite (7) as

$$\Delta(x_k, z_k) = y_k - |n_k|^p \quad (8)$$

where $y_k \triangleq |\sqrt{\gamma}\, h_k\, x_k + n_k|^p$ and $e_k \triangleq x_k - z_k$. Based on the MGF $\Phi_\Delta(s) = \mathcal{E}_{h_k,n_k}\{\mathrm{e}^{-s\,\Delta(x_k,z_k)}\}$ can be evaluated efficiently 8 using e.g. a Gauss–Chebyshev quadrature rule, cf. [14]. The result can subsequently be used along with (5) and (6) to calculate the approximate upper bound.

We note that although the method described above provides for an efficient means of calculating the approximation upper bound, the calculation of $\Phi_\Delta(s)$ required for the BER bound still involves numerical averaging over both fading and noise. Therefore the provided bound is only suitable for offline metric optimization where computational complexity is not a concern. In order to obtain analytical expressions suitable for online metric optimization, in the following subsection we provide an asymptotic analysis that results in simple–to–evaluate expressions for the BER.

B. Asymptotic BER

In this subsection we analyze the asymptotic behavior of the approximate upper bound in (5) for $\gamma \to \infty$. For this purpose, it is convenient to first rewrite the PEP as

$$P_e(\boldsymbol{c}, \hat{\boldsymbol{c}}) \;=\; \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \mathcal{E}_{n_k}\left\{\Phi(s|n_k)\right\} \frac{\mathrm{d}s}{s}, \quad (9)$$

with

$$\Phi(s|n_k) = \prod_{k=1}^{d} \left(\frac{1}{m_c 2^{m_c}} \sum_{i=1}^{m_c} \sum_{b=0}^{1} \sum_{x_k \in \mathcal{X}_b^i} \Phi_\Delta(s|n_k) \right), \quad (10)$$

where

$$\Phi_\Delta(s|n_k) \triangleq \mathcal{E}\{e^{-s\Delta(x_k,z_k)}|n_k\} = e^{-s|n_k|^p}\, \Phi_y(s), \quad (11)$$

For the last equality we have used (**??**) and the definition $\Phi_y(s) \triangleq \mathcal{E}\{e^{-sy_k}\}$. For $\gamma \to \infty$, the pdf $f_y(y_k)$ of y_k can be expressed as [**?**]

$$f_y(y_k) = \frac{2}{p(\gamma d_{xz}^2)}\, y_k^{\frac{2}{p}-1} + o\left(\gamma^{-1}\right). \quad (12)$$

Therefore, the asymptotic MGF $\Phi_{y_k}(s)$ can be obtained as the Laplace transform of $f_y(y_k)$ as

$$\Phi_y(s) = \frac{2}{p\left(\gamma d_{xz}^2\right)} \Gamma(2/p)\, s^{-\frac{2}{p}} + o\left(\gamma^{-1}\right). \tag{13}$$

Applying (13) in (11) yields

$$\Phi_\Delta(s|n_k) = \frac{2\, e^{-s|n_k|^p}}{\left(\gamma d_{xz}^2\right)} \Gamma(2/p)\, s^{-\frac{2}{p}} + o\left(\gamma^{-1}\right). \tag{14}$$

We can now obtain $\Phi(s|n_k)$ from (10) in (14) as

$$\begin{aligned}\Phi(s|n_k) = {} & 2^d X(d)\, \gamma^{-d}\, e^{-s\sum_{k=1}^{d}|n_k|^p} \left(\Gamma(2/p)\right)^d s^{-2d/p} \\ & + o\left(\gamma^{-d}\right), \end{aligned} \tag{15}$$

where

$$X(d) \triangleq \left(\frac{1}{m_c 2^{m_c}} \sum_{i=1}^{m_c} \sum_{b=0}^{1} \sum_{x_k \in \mathcal{X}_b^i} \frac{1}{d_{xz}^2} \right)^d. \tag{16}$$

The PEP can be calculated from (15) and (9) as

$$P_e(\boldsymbol{c},\, \hat{\boldsymbol{c}}\,|\, n_k) = 2^d X(d)\, \gamma^{-d} \frac{\left(\Gamma(2/p)\right)^d}{p^d\, \Gamma(2d/p+1)} M_n(d,p), \tag{17}$$

where the *generalized* noise moments $M_n(d,p)$ are defined as

$$M_n(d,p) \triangleq \mathcal{E}_{n_k}\left\{ \left(\sum_{k=1}^{d} |n_k|^p \right)^{2d/p} \right\}. \tag{18}$$

The generalized noise moments $M_n(d,p)$ can be obtained in closed–form using a similar approach as in [23] or can be efficiently calculated using Monte–Carlo simulation.

Based on (17) and (5) a closed–form approximation for the asymptotic BER $P_b \,\dot{=}\, \frac{w_c(d_f)}{k_c}\, P_e(\boldsymbol{c},\, \hat{\boldsymbol{c}})$ can be obtained as

$$P_b \,\dot{=}\, \frac{w_c(d_f) X(d_f)\, 2^{d_f} \left(\Gamma(2/p)\right)^{d_f}}{k_c\, p^{d_f}\, \Gamma(2d_f/p+1)} M_n(d_f,p)\, \gamma^{-d_f} \tag{19}$$

where d_f denotes the minimum free distance of the convolutional code. In deriving (19), besides the assumption that all joint noise moments exist, we also have assumed that (a) the approximate BER bound in (5) is tight for high SNRs and (b) the first term with $d = d_f$ in (5) is dominant. Assumption (a) is confirmed by simulations in Section VI and assumption (b) is justified for high SNR.

C. BER Minimization

Here, we study the depence of the BER bound and asymptotic BER on p and show how these performance measures can be employed for minizing the BER and thus for metric optimization. We consider a scenario where the noise statistics are known *a priori*, and therefore it is possible to perform the task of metric optimization offline. We postpone the discuss of online metric optimization until Section VI.

The offline optimization of the metric parameter p using the BER bound and the asymptotic BER is illustrated in Fig. 1. In this figure we have shown these performance measures vs. p for NBI, UF–NBI and GMN defined in Subsection II-C. The BER bound and the asymptotic BER obtained using

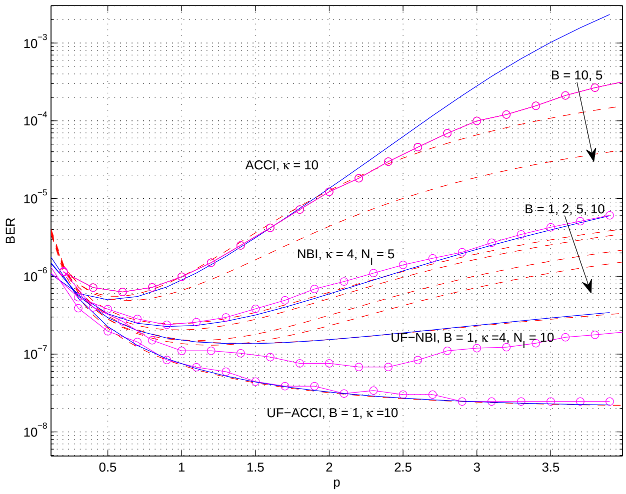

Fig. 1. BER of BICM–OFDM CR system with $B = 5$ and $N = 128$ impaired by NBI (5 equal power, sub–carrier–centered NBI signals, $I_\mu = 1$, $1 \leq \mu \leq 5$, $\kappa = 40$) vs. SNR γ.

(19) and (19), respectively. This value of p at with the BER bound is minimized denoted with p_{opt} and the corresponding point in the BER bound curve is marked using “X” markers. For comparison and to confirm our analysis, we have also shown the BER obtained via Monte–Carlo simulation. As seen, the BER bound and asymptotic BER are generally in good agreement with the simulation results. The observed small differences between the BER bound, simulation and asymptotic results are due to assuming a finite value for SNR (SNR = 15 dB) in this figure. Nevertheless, Fig. 1 shows that for each type of noise the minimum BER happens at virtually the same value of p for all the three curves. Fig. 1 further shows that the BER of the considered secondary sysmtems strongly depends on the metric parameter p and as a result, large performance improvements can be obtained using metric optimization.

IV. Metric Parameter Estimation

In this section, we derive a Maximum Likelihood (ML) paramter estimtor for the metric parameter p based on the noise samples observed at the receiver. For this purpose we first introduce a family of densities called the generalized Gaussian density (GGD) family which is parameterized by the parameter p. We then aim at finding the best estimate for p for which the GGD most accurately approximates the distribution of the underlying noise in an ML sense [24], [25]. This estimate is then used as metric paramter for the L_p–norm metric. We note that this approach is suboptimal i.e. the metric paramter estimate obtained using this approach will not necessarily minimize the BER. However, the computatinal compelexity of this appraoch is lower than the one presented in Section III as it leads to simpler expressions for the objective function used for metric optimization. We note that this approach is also amenable to both offline and online metric optimization.

A. GGD Family

The GGD family encompasses a wide range of distributions and is also a popular model for non–Gaussian noise. The corresponding pdf for GGD family is give as [25]

$$p_{GG}(z;p,\sigma) = \frac{c(p)}{\sigma^2} \exp\left(-b(p)\left(\frac{|z|}{c(p)}\right)^p\right), \qquad (20)$$

where we have defined $b(p) \triangleq \left(\frac{\Gamma(4/p)}{\Gamma(2/p)}\right)^{p/2}$ and $c(p) \triangleq \frac{p}{2\pi}\frac{\Gamma(4/p)}{\Gamma(2/p)^2}$. Furthermore, σ and p, $0 < p < \infty$, denote the standard deviation and the shape parameter, respectively. Smaller values of the shape parameter p ($0 \leq p < 2$) correspond to heavier–tailed and thus more implusive distributions, whereas larger values of p ($p > 2$) result in shorter–tailed distributions. Well-known special cases of this family are Laplacian ($p = 1$) and Gaussian noise ($p = 2$).

The motivation behind considering the GGD family for parameter estimation is two fold. Firstly, the GGD is very felexible and therefore can be succusfully used to approximate a wide range of distributions. Secondly, the L_p–norm metric employed in this paper is closely related to the GGD family. In fact, it is easy to see that an optimized L_p–norm metric can achieve ML performance in the presence a non–Gassian impairment with GGD.

B. ML Parameter Estimation

Here, we first assume that the standard deviation σ is known for the moment. The GGD family given in (20) can therefore be parametrized using parameter p leading to following formulation for the ML parameter estimation problem. For i.i.d. noise samples n_k, $1 \leq k \leq K_e$ [3], generated based on the pdf $p_n(n_k)$, we wish to find the parameter estimate $\hat{p}$ for which the GGD best approximates $p_n(n_k)$. The solution to this problem in an ML sense is the the estimate $\hat{p}$ given by

$$\hat{p} = \arg\max_{p} \{L(\boldsymbol{n};p)\} \qquad (21)$$

where $\boldsymbol{n} = [n_1, \dots n_k]^T$and the log–likelihood function (LLF) $L(\boldsymbol{n};p)$ is defined as

$$L(\boldsymbol{n};p) \triangleq \frac{1}{K_e} \log\left\{\prod_{k=1}^{Kc} p_{GG}(n_k;p,\sigma)\right\} \qquad (22)$$

Using (20) in (22) yealds

$$L(\boldsymbol{n};p) = \log(c(p)) - 2\log(\sigma) - \frac{b(p)}{(c(p))^p} T(p) \qquad (23)$$

with $T(p) \triangleq \frac{1}{K_e}\sum_{k=1}^{Kc} |n_k|^p$. For large enough values of K_e, the strong law of large number can be invoked to accurately approximate $T(p)$ as $T(p) \approx m_n(p)$ where $m_n(p) \triangleq \mathcal{E}\{|n_k|^p\}$ is the pth moment of the underlying noise. Therefore, from (23) it follows that

$$L(\boldsymbol{n};p) \approx \log(c(p)) - 2\log(\sigma) - \frac{b(p)}{(c(p))^p} m_n(p) \qquad (24)$$

[3] Assuming a frame length of K_c, these K_e noise samples are taken from $\lceil K_e/K_c \rceil$ frames (cf. Eq. (1)).

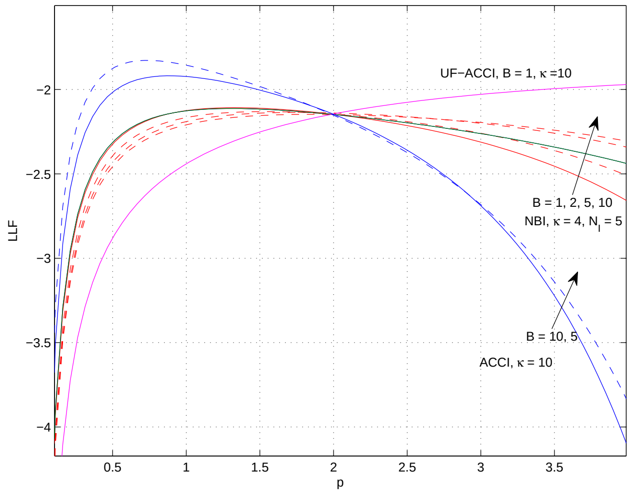

Fig. 2. BER of BICM–OFDM CR system with $B = 5$ and $N = 128$ impaired by NBI (5 equal power, sub–carrier–centered NBI signals, $I_\mu = 1$, $1 \leq \mu \leq 5$, $\kappa = 40$) vs. SNR γ.

Close–form expressions for the noise mements $m_n(p)$ have been provided in [22] for different types of noise defined in Section II. The corresponding noise moments can be used in (24) to arrive at a close–form and accurate approximation for the LLF.

In the end, we note that if the standard deviation σ happens to be unknown, we first obtain the following ML estimate $\hat{\sigma} = b(p)(p/2T(p))^{1/p}$ assuming p is known. We then use $\hat{\sigma}$ is Eqs. (22) and (21) to obtain the ML estimate $\hat{p}$.

C. LLF Maximization

In this subsection, assuming $\sigma = 1$, we study the depence of the LLF on p for different types of noise and find the estimate $\hat{p}$ that miximizes the LLF. Considering offline optimization for a scenario with known noise statistics, we postpone the discussion of online parameter estimation until Section VI.

The offline paremter estimation is illustrated in Fig. 2 for the same noise types as in Fig. 1. In this figure we have shown LLF obtained using (19) for the considered noise types vs. p. For each type of noise we have marked the estimate $\hat{p}$ that maximizes the LLF by "X" markers. Comparing Fig. 2 with Fig. 1 reveals that although parameter estimation is a suboptimal approach, the parameter estimate $\hat{p}$ is generally very close to p_{opt} and the incurred performance loss due to using $\hat{p}$ instead of p_{opt} is minimal.

V. Adaptive Metric Optimization

In practice, the type of noise impairing a secondary user system is usually not known *a priori* and changes with time. Therefore, in this section, we present efficient adaptive algorithms for optimization of the metric parameter p. We first develop an adaptive algorithm for BER minimization based on the general BER analysis framework described in Section III. Then we propose an adaptive parameter estimation algorithm

$$\frac{\partial L(\boldsymbol{n};p)}{\partial p} = \frac{1}{p} - \frac{4}{p^2}\psi(4/p) + \frac{4}{p^2}\psi(2/p) - \frac{1}{K_e}\sum_{k=1}^{K_e}\frac{|n_k|^p}{A^p}\left(\log\left(\frac{|n_k|}{A}\right) + \frac{1}{2p}\left[2\psi(2/p) - 4\psi(4/p)\right]\right) \tag{25}$$

$$\begin{aligned}\frac{\partial^2 L(\boldsymbol{n};p)}{\partial p^2} = & -\frac{1}{p^2} + \frac{8}{p^3}\psi(4/p) - \frac{8}{p^3}\psi(2/p) + \frac{16}{p^4}\psi'(4/p) - \frac{1}{K_e}\sum_{k=1}^{K_e}\frac{|n_k|^p}{A^p}\left(\log\left(\frac{|n_k|}{A}\right) + \frac{1}{2p}\left[2\psi(2/p) - 4\psi(4/p)\right]\right)^2 \\ & -\frac{8}{p^4}\psi'(2/p) - \sum_{k=1}^{K_e}\frac{|n_k|^p}{A^p}\left(\frac{1}{2p^2}\left[2\psi(2/p) - 4\psi(4/p)\right] - \frac{2}{p^3}\psi'(2/p) + \frac{8}{p^3}\psi'(4/p) - \frac{1}{2p^2}\left[2\psi(2/p) - 4\psi(4/p)\right]\right)\end{aligned} \tag{26}$$

based on the parameter estimation techniques developed in Section V.

A. Adaptive BER Minimization

Since the approximate upper bound derived in Section III-A is too cumbersome for real–time optimization, here we propose an adaptive algorithm based on the asymptotic BER results obtained in Section III-B. Due to the random nature of the optimization problem, a stochastic optimization algorithm has to be used. Although several types of stochastic optimization methods are available in the literature, our experience has shown that the Kiefer–Wolfowitz (KW) algorithm [26] is the most suitable for the problem at hand.

Based on (18) and (19) the cost function for the KW algorithm is given by

$$L_k(p) = \frac{w_c(d_f)X(d_f)\,2^{d_f}(\Gamma(2/p))^{d_f}}{k_c\,p^{d_f}\,\Gamma(2d_f/p+1)}M_{n_k}(d_f,p) \tag{27}$$

where

$$\hat{M}_n(d,p) \triangleq \left(\sum_{k=1}^{d}|n_k|^p\right)^{2d/p}. \tag{28}$$

where $\hat{M}_n(d,p)$ is the instantaneous estimate for the generalized noise moment, and we have omitted all terms that do not affect the optimization. The proposed KW algorithm recursively updates the estimates of the optimal p, i.e., it generates the parameter estimate p_k at the kth iteration as [26]:

$$p_{k+1} \;=\; p_k + \delta_k\,\frac{L_k(p_k+\zeta_k) - L_k(p_k-\zeta_k)}{2\zeta_k} \tag{29}$$

where $\delta_k > 0$ and $\zeta_k > 0$ are the gain sequences of the KW algorithm. The convergence theory for the KW algorithm [26] states that if the gain sequence fulfills $\delta_k \to 0$, $\zeta_k \to 0$, $\sum_{k=0}^{\infty}\delta_k = \infty$, 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.

B. Adaptive Parameter Estimation

In this subsection, we provide an adaptive algorithm that allows us to obtain an ML estimate for p by online maximization of LLF given in (23). For this porpose we use a Newton–Raphson (NR) based iterative algorithm that is widely used in the literature for adaptive parameter estimation [27]. We however note that a KW algorithm can also be constructed based on the approximate LLF in (24). We only present the NR based algorithm since this algorithm is more widely used in the litrature for parameter estimation. The proposed NR based iterative algorithm can be formulated as follows. At the kth iteration the algorithm generates the new estimate p_{k+1} as [27]

$$p_{k+1} = p_k - \left[\frac{\partial^2 L(\boldsymbol{n};p)}{\partial p^2}\right]^{-1}\frac{\partial L(\boldsymbol{n};p)}{\partial p}\Big|_{p=p_k} \tag{30}$$

The partial derivatives used in the above equation are obtained in (25) and (26) where we have defined $A \triangleq \sqrt{\frac{\Gamma(2/p)}{\Gamma(4/p)}}$, and $\psi(\cdot)$ and $\psi'(\cdot)$ denote the digamma and trigamma functions, respectively, for which efficient algorithms exist for numerical evaluation []. The convergence depends on the intial guess and the algorithm normally converges to a local minimum provided that the initial guess is not too far from the that minimum. The convergence rate is generally quadratice, i.e., the error is squared at each iteration.

C. Example

We now present an practical example where the noise statistics vary with time. For this example we illustrate the versatiliy of the proposed adaptive algorithms in solving the metric optimization problem and compare their performances.

In Fig. 3

VI. Numerical and Simulation Results

In this section, we verify the analytical results presented in Sections III and IV with computer simulations and compare the adaptive L_p–norm metric with several other popular robust metrics. We also demonstrate how the off–line and on–line optimization techniques described in Section V can be used to optimize the L_p–norm metric parameter p. For all results shown we consider a BICM–OFDM secondary system with $N = 64$ or $N = 128$, 4–PSK, and a code with rate

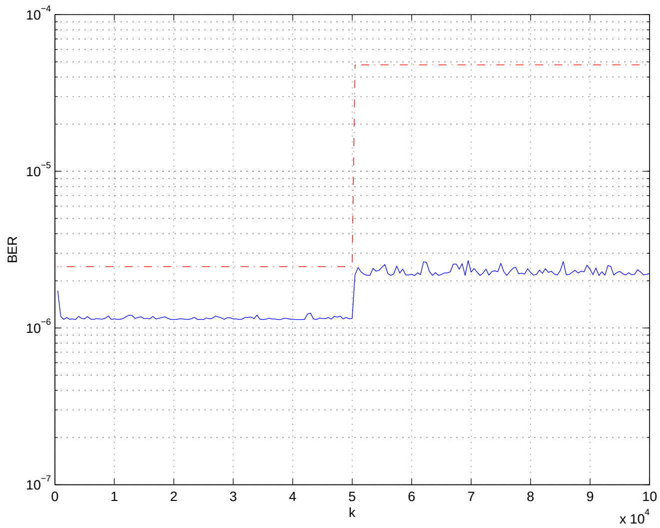

Fig. 3. BER of BICM–OFDM CR system with $B = 5$ and $N = 128$ impaired by NBI (5 equal power, sub–carrier–centered NBI signals, $I_\mu = 1$, $1 \leq \mu \leq 5$, $\kappa = 40$) vs. SNR γ.

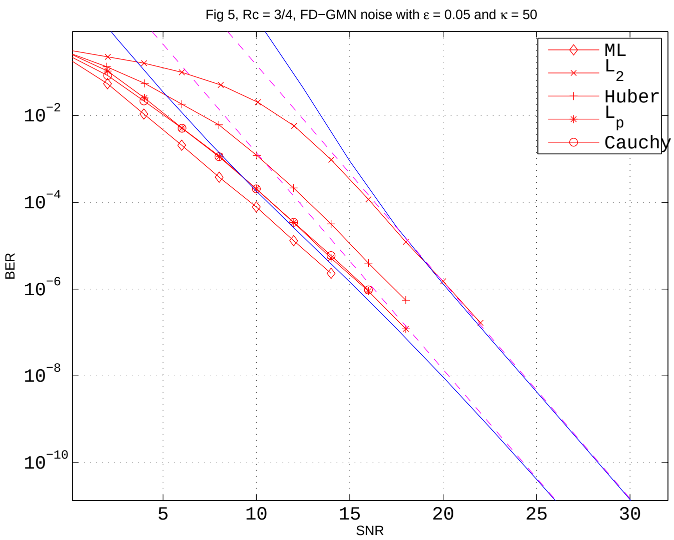

Fig. 4. BER of BICM–OFDM CR system with $B = 5$ and $N = 128$ impaired by NBI (5 equal power, sub–carrier–centered NBI signals, $I_\mu = 1$, $1 \leq \mu \leq 5$, $\kappa = 40$) vs. SNR γ.

$R_c = 3/4$ obtained through puncturing of the standard rate–1/2 convolutional code with generator polynomials $[133, 171]$ (octal representation). The parameters for the considered types of noise are specified in the captions of the figures.

In Fig. 4, we compare the performance of the proposed L_p–norm metric with that of the conventional L_2–norm metric and several other popular robust metrics for ϵ–mixture noise. We consider the Huber metric $f_m(u) = u^2/2$ if $u \leq \delta$, and $f_m(u) = \delta u - \delta^2/2$ if $u > \delta$ [8], and the Cauchy metric $f_m(u) = \log(u^2 + \delta^2)$ [**?**]. For the robust metrics δ was optimized by simulation for SNR = 18 dB. In contrast, the L_p–norm metric was optimized with the KW algorithm. The BER bound and aysmptotic BER are shown for L_p–norm metric and the L_2–norm metric using solid and dashed lines, respectively. Also shown are the simulation results for all considered metrics using solid lines with markers. The great agreements between the simulation results, the BER bound and asymptotic BER for the L_p–norm and the L_2–norm metrics again corroborate our analysis. Fig. 4 further shows that for the ϵ–mixture noise the L_p–norm metric outperforms the other robust metrics and the gap to the optimum ML–metric is less than 1 dB. Finally, Fig. 4 suggests that although the L_p–norm metric was optimized based on the presented asymptotic analysis for high SNR, it also performs well for small SNRs.

VII. Conclusions

In this paper, we have proposed an L_p–norm metric for BICM–OFDM secondary systems operating in the presence of non– Gaussian noise. We have derived an approximate upper bound and an accurate asymptotic approximation for the BER of the considered secondary system. These analytical results can be used for optimization of the metric parameter p. Simulation results have confirmed the validity of the provided analytical results and have shown the effectiveness of the proposed L_p–norm metric in mitigating the harmful effects of non–Gaussian noise in secondary systems.

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] 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.
[17] D. Rende and T. Wong. Bit–Interleaved Space–Frequency Coded Modulation for OFDM Systems. *IEEE Trans. Wireless Commun.*, 4:2256–2266, September 2005.
[18] 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.
[19] A. Coulson. Bit Error Rate Performance of OFDM in Narrowband Interference with Excission Filtering. *IEEE Trans. Wireless Commun.*, 5:2484–2492, September 2006.
[20] 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.
[21] V. Sethuraman and B. Hajek. Comments on "Bit–Interleaved Coded Modulation". *IEEE Trans. Inform. Theory*, 52:1795–1797, April 2006.
[22] 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.
[23] 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.
[24] M.K. Varanasi. Parameter estimation of the generalize gaussian noise model. M.S. thesis, Rice University, Houston, TX, 1986.
[25] M.S. Davis, P. Bidigare, and D. Chang. Statistical Modeling and ML Parameter Estimation of Complex SAR Imagery. In *Proceedings of IEEE ACSSC*, pages 500–502, 2007.
[26] J. Spall. *Introduction to Stochastic Search and Optimization*. Wiley & Sons, Inc., New Jersey, 2003.
[27] S.M. Kay. *Fundamentals of Statistical Signal Processing*. Prentice Hall PTR, second edition, 1998.