Robust L_p–Norm Metric for Cognitive Radio 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

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 bit–interleaved coded modulation (BICM) scheme. For the considered CR system we provide an approximate upper bound on the BER as well as simple and easy–to–evaluate analytical expressions for the asymptotic BER. The BER bound and the asymptotic BER are both obtained as functions of the metric parameters and therefore can be employed for metric optimization. While the BER bound is used for offline metric optimization, an effective adaptive algorithm is provided for online metric optimization based on the asymptotic BER results. Numerical and simulation results show that using the optimized L_p–norm metric significant performance gains can be achieved compared to the conventional L_2–norm metric.

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

1 Introduction

While demand for high speed wireless access and inflexible methods of spectrum allocation have made the spectral resources increasingly scarce, recent studies [1, 2] have indicated that wide spectral ranges are rarely used in both space and time. This observation has motivated the new idea of Cognitive Radios [3]. In a broad sense a cognitive radio is a smart radio capable of sensing the environment in order to provide efficient utilization of spectrum bands that temporally or spatially unoccupied or under–utilized by their licensed (primary) users. While primary users have priority in accessing the spectrum, CR (secondary) users can use the available spectrum opportunistically without interfering with the primary users. As a results, the spectrum is used more efficiently and the deployment of new applications is simplified.

The aforementioned definition of CR implies that the CR users have to be able to maintain their required quality–of–service (QoS) in the presence of the interference from the primary user. The CR users also should be able to cope with the interference caused by other CR user who aim at exploiting the same under–utilized frequency band [4]. Furthermore, other sources of interference such as UWB interference and man–made impulsive noise [5, 6] can be presence in CR environments.

While conventional CR systems are design for Gaussian noise, various forms of noise and interference present in CR environments can have non–Gaussian distribution. In particular, the interference caused by the primary user and other CR systems are forms of co–channel interference and therefore can potentially be non–Gaussian distributed [7]. Furthermore, recent studies have shown that other types of noise and interference present in CR environments such as UWB interference [8, 9] and man–made impulsive noise [10] can have pronounced non–Gaussian characteristics. Therefore, the use of an L_2–norm metric (also referred to as the Euclidean distance (ED) metric) for signal detection which is only optimal in the case of Gaussian noise can result in significant performance losses in CR environments where non–Gaussian noise[2] is present.

Fortunately, these harmful effects and be effectively reduced if the non–Gaussian behavior of the noise is carefully taken into account. In particular, a maximum–likelihood (ML) metric with optimal performance in the presence of non–Gaussian noise can be designed based on the noise statistics. We note however that depending on the noise distribution the implementation of the ML metric at the

[2]Unless stated otherwise, in the rest of this thesis, by "noise" we refer to any additive impairment of the received signal, i.e., our definition of noise also includes what is commonly referred to as "interference".

receiver may not be feasible due to high computational complexity. Furthermore, the design of ML metric requires the knowledge of the noise statistics which may not be available in many practical scenarios.

To overcome this problem, suboptimal robust metrics should be used that offer low complexity and perform well in a large class of noise distributions. Since the noise statistics may be time variant in practice, these metrics should also be able to adjust to variations in noise statistics. Important examples of such robust metrics available in the literature include Huber metric [11], generalized Cauchy metric [12], and L_p–norm metric [13]. 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 [13].

In this paper, we propose a robust CR system that exploits an L_p–norm metric to mitigate the harmful effects of non–Gaussian noise present in CR environments. The considered CR system employs a combination of bit–interleaved coded modulation (BICM) with single–carrier (SC) modulation or orthogonal frequency division multiplexing (OFDM). Furthermore, the receiver of the considered CR system is equipped with multiple receiver antennas. The motivation behind considering BICM–SC and BICM–OFDM is two fold. Firstly, it is well known that BICM–SC and BICM–OFDM are very efficient in exploiting time diversity and frequency diversity in wireless fading channels, respectively. The inclusion of the multiple receiver antennas allows the CR system to exploit the space diversity as well and therefore achieve more robustness against fading. Secondly, these techniques have been adopted by a number of recent standards [14] and are also expected to play a major role in future wireless standards.

For the considered CR system we provide a general mathematical framework for analyzing the BER performance in non–Gaussian environments. This framework is very general and applicable to arbitrary linear modulation formats, all commonly used fading models, and all practically relevant types of noise. Based on the developed framework we provide an approximate upper bound on the bit error rate (BER) performance based on the expurgated union bound of [15]. This bound is obtained as a function of the metric parameters and therefore can be used for the purpose of metric optimization. Furthermore, the obtained bound is very accurate in the BER range of practical interest (i.e. $\mathrm{BER} < 10^{-5}$) and computationally much more efficient compared to direct Monte–Carlo simulation of the communication system. However, since this bound is not obtained in closed–form, its computational complexity makes it unsuitable for scenarios where the noise statistics vary with time

and therefore online optimization is required. As a result, we use the approximate upper bound only in scenarios where the noise statistics are known *a priori* and the optimization can be performed offline. For online optimization we provide closed–formed asymptotic expressions by analyzing the asymptotic behavior of the BER bound for high signal–to–noise ratios (SNR's). Based on the obtained asymptotic results we develop an efficient adaptive multivariate finite difference stochastic approximation (FDSA) algorithm for online metric optimization. Simulation results confirm the validity of our analysis and show that the proposed L_p–norm metric is very effective in dealing with the harmful effects of nonGaussian noise.

The rest of this paper is organized as follows. In Section 2, the system model for the considered CR system is introduced. The approximate upper bound for the BER is derived in Section 3 and asymptotic BER expressions are obtained in Section 4. In Section 5, off–line and on–line optimization of the metric parameters are discussed and analytical and simulation results are presented in Section 6. Finally conclusions are drawn in Section 7.

Notations: 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, $\boldsymbol{I}_M$ 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 $\lim_{x\to 0} f(x)/g(x) = 0$.

2 System Model

We consider a CR system employing one of the popular BICM–SC or BICM–OFDM schemes combined with N_r receive antennas. In the rest of this paper, we refer to the aforementioned CR systems as CR–BS and CR–BO, respectively. In the following we provide unified signal and fading models that are applicable to both CR–BS and CR–BO systems. We also present several practically relevant noise models for the considered CR–BS and CR–BO systems. For convenience, in this paper, all signals and systems are represented by their complex baseband equivalents.

2.1 Signal Model

The CR transmitter employs a BICM transmission scheme, i.e., it consists of a convolutional encoder of rate R_c, an interleaver, and a memoryless mapper [15]. Specifically, the codeword $\boldsymbol{c} \triangleq [c_1, c_2, \ldots, c_{m_c K_c}]$ of length $m_c K_c$ is generated by a 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 $\boldsymbol{x} \triangleq [x_1, x_2, \ldots, x_{K_c}]$. The K_c transmit symbols are broken up into B frames of N symbols each implying $K_c = BN$. In CR–BS systems, all the N symbols in a frame are transmitted over a single carrier. In contrast, these symbols are modulated onto N OFDM sub–carriers in CR–BO systems.

Assuming perfect synchronization and demodulation, for both CR–BS and CR–BO systems the observed signal at the N_r receive antennas for the Lth transmitted frame can be modeled as

$$\boldsymbol{r}_k = \sqrt{\gamma}\, \boldsymbol{h}_k\, x_k + \boldsymbol{n}_k, \qquad (L-1)K_c + 1 \leq k \leq LK_c, \tag{1}$$

where k denotes the time index in CR–BS systems and the frequency index in CR–BO systems, respectively. Furthermore, in the above equation γ denotes the SNR per receive antenna as we have assumed without loss of generality $\mathcal{E}\{||\boldsymbol{h}_k||^2\} = N_r$ and $\mathcal{E}\{||\boldsymbol{n}_k||^2\} = N_r$. In (1) we have also defined $\boldsymbol{h}_k \triangleq [h_{k,1} \ \ldots \ h_{k,N_r}]^T$ and $\boldsymbol{n}_k \triangleq [n_{k,1} \ \ldots \ n_{k,N_r}]^T$ where $h_{k,l}$ and $n_{k,l}$, $1 \leq l \leq N_r$, denote the fading gains and the noise variables, respectively.

As customary in the literature, cf. e.g. [15, 16, 17], for our performance analysis we assume perfect interleaving, which means that $\boldsymbol{h}_k$ and $\boldsymbol{n}_k$ can be modeled as independent, identically distributed (i.i.d.) random vectors and only their first order probability density functions (pdfs) are relevant.

2.1.1 The L_p–Norm Branch Metric

Traditional BICM–based system employ an L_2–norm branch metric for Viterbi decoding at the receiver. In this paper, we assume that the Viterbi decoder at the CR system receiver employs an L_p–norm branch metric that allows it to adapt to the ambient noise and interference by adjusting the metric parameters. The employed branch metric for decoding bit i, $1 \leq i \leq m_c$, of symbol x_k is given as

$$\lambda_i(\boldsymbol{r}_k, b) \triangleq \min_{x_k \in \mathcal{X}_b^i} \{f_M(\boldsymbol{r}_k - \sqrt{\gamma}\, \boldsymbol{h}_k\, x_k)\} \tag{2}$$

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, $f_M(.)$ denotes a non–linearity that depends on the adopted branch metric for Viterbi

decoding. For the L_p–norm branch metric assumed in this paper we have

$$f_M(\boldsymbol{r}) = \sum_{l=1}^{N_r} q_l \, |r_l|^{p_l}, \tag{3}$$

where $\boldsymbol{r} \triangleq [r_1, \ldots, r_{N_r}]^T$, and $\boldsymbol{q} \triangleq [q_1, \ldots, q_{N_r}]^T$ and $\boldsymbol{p} \triangleq [p_1, \ldots, p_{N_r}]^T$ are the adjustable metric parameters.

2.2 Fading Model

We assume that the fading gains can be expressed as $h_{k,l} \triangleq a_{k,l} \mathrm{e}^{j\Theta_{k,l}}$, 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 $p_{a_{k,l}}(a_{k,l})$. In this paper, we consider spatially i.i.d. as well as spatially correlated fading channels. For spatially i.i.d. fading the first order pdf $p_a(a_{k,l})$ is sufficient to describe the properties of the RV $a_{k,l}$, while for spatially correlated fading the joint pdf $p_{\boldsymbol{a}_k}(\boldsymbol{a}_k)$ of the elements of $\boldsymbol{a}_k \triangleq [a_1 \ \ldots \ a_{N_r}]^T$ (cf. e.g. [18]–[20]) is also required. However, as shown in [21], for asymptotically high SNR's spatially correlated fading can be regarded as asymptotically spatially i.i.d. fading. Therefore for $\gamma \to \infty$ the joint pdf can be expressed as

$$p_{\boldsymbol{a}_k}(\boldsymbol{a}_k) \mathrel{\dot{=}} \prod_{l=1}^{N_r} p_a(a_{k,l}), \tag{4}$$

where

$$p_a(a_{k,l}) = 2A \, a_{k,l}^{2\nu-1} + o(a_{k,l}^{2\nu-1}) \tag{5}$$

with fading distribution dependent constants A and ν. The fading pdf $p_a(a_{k,l})$ and the parameters A and ν are specified in Table 1 for correlated Rayleigh, Ricean and Nakagami-m fading as well as for spatially independent Nakagami–q and Weibull fading channels.

2.3 Noise Model

In this subsection, we present several practically relevant noise models that are frequency encountered in CR environments. We use these noise models in Section 5 to show the ability of our proposed CR system to mitigate the harmful non–Gaussian noise effects, and to adjust to changes in the statistics of the ambient noise. Since the receivers of CR–BS and CR–BO systems effect the noise differently, in the following we discuss the noise models for these two systems separately.

2.3.1 Noise Models for CR–BS

Here, we consider two important time–domain noise models for CR–BS systems. In particular, we consider asynchronous CCI (ACCI) ad time–domain Gaussian–mixture noise (TD–GMN).

ACCI: In CR–BS systems, ACCI [22, 7] can be used to model the CCI caused by the primary user and other CR systems. To describe this noise model, we consider a CR–BS system with B different hopping frequencies and assume that at hopping frequency μ, $1 \leq \mu \leq B$, in addition to AWGN $\tilde{n}_{k,l,\mu}$, there are I_μ Ricean faded asynchronous CCI signals leading to time–domain noise

$$n_{k,l,\mu} = \sum_{i=1}^{I_\mu} \tilde{h}_{k,l,\mu}[i] \sum_{\upsilon=\upsilon_l}^{\upsilon_u} g_{\mu,i}[\upsilon] b_{\mu,i}[\upsilon] + \tilde{n}_{k,l,\mu} \tag{6}$$

where $\tilde{h}_{k,l,\mu}[i]$ are temporally i.i.d. Gaussian random variables which model the Ricean interference channel gains with Ricean factor $K_{\mu,i}$ and $b_{\mu,i}[\upsilon] \in \mathcal{M}_{\mu,i}$ ($\mathcal{M}_{\mu,i}$: $\tilde{M}_{\mu,i}$–ary symbol alphabet) are the i.i.d. symbols of the ith interferer at the μth hopping frequency, respectively. Furthermore, $g_{\mu,i}[\upsilon] \triangleq g_{\mu,i}(\upsilon T + \tau_{\mu,i})$, where $g_{\mu,i}(t)$, T, and $\tau_{\mu,i}$ are the effective pulse shape, the symbol duration, and the time offset of the ith interferer at the μth hopping frequency, respectively, and we assume that $g_{\mu,i}(\upsilon T + \tau_{\mu,i}) \approx 0$ for $\upsilon < \upsilon_l$ and $\upsilon > \upsilon_u$. In addition, for future reference we denote the set of all possible values of $\xi_{\mu,i} \triangleq \sum_{\upsilon=\upsilon_l}^{\upsilon_u} g_{\mu,i}[\upsilon] b_{\mu,i}[\upsilon]$ by $\mathcal{S}_{\mu,i}$, define $\mathcal{S}_\mu \triangleq \mathcal{S}_{\mu,1} \times \ldots \times \mathcal{S}_{\mu,I_\mu}$ (if $I_\mu = 0$, we formally set $\mathcal{S}_\mu = \{0\}$ with $|\mathcal{S}_\mu| = 1$) and denote the ratio of the total CCI variance and the total AWGN variance by κ, cf. Section 6. Finally, we note that for $K_{\mu,i} \to \infty$, the interference channel gains $\tilde{h}_{k,l,\mu}[i]$ will be constant values. The refer to the resulting noise as unfaded ACCI (UF–ACCI).

TD–GMN: TD–GMN can be used to model the combined effect of Gaussian background noise and man–made or impulsive noise, cf. e.g. [23, 10, 24] present in CR environments. If the phenomenon causing the impulsive behavior affects the receive antennas independently, the GMN is spatially i.i.d. [25] and the pdf of $n_{k,l}$ is given by [10]

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

where $c_i > 0$ and $\sigma_i^2 > 0$ are parameters, and $\sum_{i=1}^{I} c_i \sigma_i^2 = 1$. Two popular special cases of Gaussian mixture noise are Middleton's Class–A noise [10] and ϵ–mixture noise. For ϵ–mixture noise $I = 2$, $c_1 = 1 - \epsilon$, $c_2 = \epsilon$, $\sigma_1^2 = \sigma_g^2$, and $\sigma_2^2 = \kappa \sigma_g^2$, where ϵ is the fraction of time when the impulsive noise is present, κ is the ratio of the variances of the Gaussian background noise and the impulsive noise, and $\sigma_g^2 = 1/(1 - \epsilon + \kappa\epsilon) = 1$.

2.4 Noise Models for CR–BO

For CR–BO systems, we consider two practically relevant frequency–domain (FD) noise models, i.e., narrowband interference (NBI) and FD–GMN.

NBI: We consider a CR 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 $\tilde{n}_{k,\mu}$ and I_μ Rayleigh faded PSK NBI signals. The corresponding frequency–domain noise model is

$$n_{k,l,\mu} = \sum_{i=1}^{I_\mu} g_{k,\mu}[i] b_\mu[i] \tilde{h}_{k,l,\mu}[i] + \tilde{n}_{k,l,\mu}, \tag{8}$$

where $\tilde{h}_{k,l,\mu}[i]$ are temporally i.i.d. Gaussian random variables which model the Ricean interference channel gains with Ricean factor $K_{\mu,i}$. Furthermore, $b_\mu[i]$ is the PSK symbol of the ith interferer at the μth hopping frequency affecting the set $\mathcal{N}_{\mu,i}$ of sub–carriers via $g_{k,\mu}[i] \triangleq \exp[-j\pi(N-1)(k + f_{\mu,i}/\Delta f_s)/N + \phi_{\mu,i}] \sin[\pi(k + f_{\mu,i}/\Delta f_s)]/\sin[\pi(k + f_{\mu,i}/\Delta f_s)/N]$ [26]. 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 $f_{\mu,i} = \upsilon \Delta f$, the NBI affects only sub–carrier υ, i.e., $\mathcal{N}_{\mu,i} = \upsilon$, while, in theory, for $f_{\mu,i} \neq \upsilon \Delta f$ the NBI affects all sub–carriers. However, $g_{k,\mu}[i]$ decays quickly and we limit $\mathcal{N}_{\mu,i}$ such that $|g_{k,\mu}[i]| \approx 0$ for $k \notin \mathcal{N}_{\mu,i}$. Finally, we assume that no sub–carrier is affected by two narrowband interferers at a given hopping frequency, i.e., $\mathcal{N}_{\mu,i_1} \cap \mathcal{N}_{\mu,i_2} = \emptyset$, $i_1 \neq i_2$. 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 $K_{\mu,i} \to \infty$, the interference channel gains $\tilde{h}_{k,\mu,l}[i]$ will be constant values. The resulting noise will be referred to as unfaded NBI (UF–NBI) in the rest of this paper.

FD–GMN: FD–GMN can be used to model the combined effect of the frequency–domain Gaussian background noise and the interference caused by a Rayleigh faded NBI that employs frequency hopping technique. Denoting the by ϵ the probability that the interferer hops into the band used by the CR system, we can express the pdf of the corresponding frequency–domain noise as

$$p_n(n_{k,l}) = \frac{(1-\epsilon)}{\pi\sigma_g^2} \exp\left(-\frac{|n_{k,l}|^2}{\sigma_g^2}\right) + \frac{\epsilon}{\pi\kappa\sigma_g^2} \exp\left(-\frac{|n_{k,l}|^2}{\kappa\sigma_g^2}\right), \tag{9}$$

where σ_g is variance of the frequency–domain background noise and κ is the ratio of the variances of the background noise and the interference.

3 Approximate Upper Bound for BER

In this section, we provide an approximate upper bound on the BER performance of the considered CR system in non–Gaussian noise environments based on the expurgated union bound of [15].[3] The results obtained in this section are based on the unified system model presented in Section 2 and therefore they are applicable to both CR–BS and CR–BO systems. Furthermore, the BER bounds are obtained as a function of the metric parameters $\boldsymbol{q}$ and $\boldsymbol{p}$, and consequently can be used for optimizing these parameters.

For a CR system employing convolutional code of rate $R_c = k_c/n_c$ (k_c and n_c are integers) the union bound for the BER is given by [15]

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

where $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ 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. $P(\boldsymbol{c},\, \hat{\boldsymbol{c}})$ is the pairwise error probability (PEP), i.e., the probability that the decoder chooses code sequence $\hat{\boldsymbol{c}}$ when code sequence $\boldsymbol{c} \neq \hat{\boldsymbol{c}}$ is transmitted. Invoking the expurgated bound from [15], the PEP can be expressed as

$$P(\boldsymbol{c},\, \hat{\boldsymbol{c}}) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \prod_{k=1}^{d} \Psi_k(s)\, \frac{\mathrm{d}s}{s}, \tag{11}$$

with

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

where c is a small positive constant that lies in the region of convergence of the integrand. Furthermore, $\Phi_{\Delta(x_k,z_k)}(s) \triangleq \mathcal{E}_{\boldsymbol{h}_k, \boldsymbol{n}_k}\{\mathrm{e}^{-s\,\Delta(x_k,z_k)}\}$ is the MGF of the following metric different conditional on the transmission of symbol x_k at the transmitter

$$\begin{aligned} \Delta(x_k, z_k) &\triangleq f_M(\boldsymbol{r}_k - \sqrt{\gamma}\, \boldsymbol{h}_k\, z_k) - f_M(\boldsymbol{r}_k - \sqrt{\gamma}\, \boldsymbol{h}_k\, x_k) \\ &= \sum_{l=1}^{N_r} q_l\, |r_{k,l} - \sqrt{\gamma}\, h_{k,l}\, z_k|^{p_l} - \sum_{l=1}^{N_r} q_l\, |r_{k,l} - \sqrt{\gamma}\, h_{k,l}\, x_k|^{p_l} \end{aligned} \tag{13}$$

[3] As pointed out in [27, 28], the expurgated union bound found in [15] is not a true upper bound and constitutes only an approximation to the BER performance. However, numerical evidence [21] has taught us that for Gray labeling this bound is very accurate for the BER range of practical interest even in non–Gaussian environments (also cf. Section 6).

where z_k is the nearest neighbor of x_k in $\mathcal{X}_{\bar{b}}^i$ with $\bar{b}$ being the bit complement of b. Taking to account the fact that conditional on the transmission of x_k we have $r_{k,l} = \sqrt{\gamma}\, h_{k,l}\, x_k + n_{k,l}$, $1 \leq l \leq N_r$, we can write the above equation as

$$\Delta(x_k, z_k) = \sum_{l=1}^{N_r} q_l \, |\sqrt{\gamma}\, h_{k,l}\, e_k + n_{k,l}|^{p_l} - \sum_{l=1}^{N_r} q_l \, |n_{k,l}|^{p_l} \tag{14}$$

where we have defined $x_k - z_k \triangleq d_{xz} \mathrm{e}^{j\Theta_d}$ with ED $d_{xz} > 0$.The MGF $\Phi_{\Delta(x_k,z_k)}(s)$ can therefore be obtained as

$$\Phi_{\Delta(x_k,z_k)}(s) = \mathcal{E}_{\boldsymbol{h}_k, \boldsymbol{n}_k} \left\{ \prod_{l=1}^{N_r} \mathrm{e}^{-s(q_l \, |\sqrt{\gamma}\, h_{k,l}\, e_k + n_{k,l}|^{p_l} - q_l \, |n_{k,l}|^{p_l})} \right\} \tag{15}$$

The MGF $\Phi_{\Delta(x_k,z_k)}(s)$ obtained in the above equation can be used in (11) to calculate the PEP. To evaluate the complex integral in (11) efficiently, we use the saddlepoint approximation technique advocated in [27, 29]. This leads to

$$P(\boldsymbol{c},\, \hat{\boldsymbol{c}}) \approx \frac{(\Psi_k(\hat{s}))^{(d+1/2)}}{\hat{s}\sqrt{2\pi d \Psi_k''(\hat{s})}} \tag{16}$$

where $\Psi_k''(s)$ is the second derivative of $\Psi_k(s)$ and the saddlepoint $\hat{s}$ is defined as the value for which $\Psi_k'(\hat{s}) = 0$ is valid.

Once the PEP is obtained, it can be used along with (10) to obtain an approximate upper bound on the BER. We note that although the saddlepoint technique provides for an efficient means of calculating the PEP, evaluation of the PEP and therefore the BER bound according to (10) involves several integrations which have to be evaluated numerically. 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 next section we provide an asymptotic analysis which results in simple–to–evaluate expressions for the BER.

4 Asymptotic Analysis

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

$$P(\boldsymbol{c},\, \hat{\boldsymbol{c}}) \;=\; \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} \mathcal{E}_{\boldsymbol{n}_k} \{\Phi(s|\boldsymbol{n}_k)\} \, \frac{\mathrm{d}s}{s}, \tag{17}$$

with

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

where

$$\Phi_{\Delta(x_k,z_k)}(s|\boldsymbol{n}_k)=\mathcal{E}\{\mathrm{e}^{-s\Delta(x_k,z_k)}\}=\mathrm{e}^{-s\sum_{l=1}^{N_r}q_l|n_{k,l}|^{p_l}}\prod_{l=1}^{N_r}\Phi_{y_{k,l}}(q_l\,s). \tag{19}$$

In obtaining the last equality, we have used (15), and defined $y_{k,l}\triangleq|\sqrt{\gamma}\,h_{k,l}\,e_k+n_{k,l}|^{p_l}$ and $\Phi_{y_{k,l}}(s)\triangleq\mathcal{E}\{\mathrm{e}^{-sy_{k,l}}\}$. The pdf $f_y(y_{k,l})$ of $y_{k,l}$ has been calculated in Appendix A (cf. Eq. (53)). Therefore, the MGF $\Phi_{y_{k,l}}(s)$ can be found by calculating the Laplace transform of $f_y(y_{k,l})$ as

$$\Phi_{y_{k,l}}(s)\approx\frac{2A}{p_l\,(\gamma d_{xz}^2)^{\nu}}\sum_{i=0}^{\bar{\nu}}\xi_{i,l}\,|n_{k,l}|^{2i}\,s^{-2(\nu-i)/p_l}+o\left(\gamma^{-\nu}\right) \tag{20}$$

where $\xi_{i,l}\triangleq\frac{\Gamma(2(\nu-i)/p_l)\,P_i}{\Gamma(\nu-i)}$. Substituting (20) in (19) yields

$$\Phi_{\Delta(x_k,z_k)}(s|\boldsymbol{n}_k)\approx\frac{2^{N_r}A^{N_r}\,\mathrm{e}^{-s\sum_{l=1}^{N_r}q_l|n_{k,l}|^{p_l}}}{(\gamma d_{xz}^2)^{N_r\nu}}\prod_{l=1}^{N_r}\left(\sum_{i=0}^{\bar{\nu}}\xi_{i,l}\,|n_{k,l}|^{2i}\,p_l^{-1}q_l^{-2(\nu-i)/p_l}\,s^{-2(\nu-i)/p_l}\right)+o\left(\gamma^{-N_r\nu}\right). \tag{21}$$

We can now obtain $\Phi(s|\boldsymbol{n}_k)$ by using (21) in (18) as

$$\begin{aligned}\Phi(s|\boldsymbol{n}_k)\approx X_m(\nu,N_r,d)\,(2A)^{dN_r}\gamma^{-dN_r\nu}\,\mathrm{e}^{-s\sum_{k=1}^{d}\sum_{l=1}^{N_r}q_l|n_{k,l}|^{p_l}}\prod_{k=1}^{d}\prod_{l=1}^{N_r}\\ \left(\sum_{i=0}^{\bar{\nu}}\xi_{i,l}\,|n_{k,l}|^{2i}\,p_l^{-1}q_l^{-2(\nu-i)/p_l}\,s^{-2(\nu-i)/p_l}\right)+o\left(\gamma^{-dN_r\nu}\right)\end{aligned} \tag{22}$$

where

$$X_m(\nu,N_r,d)\triangleq\left(\frac{1}{m_c2^{m_c}}\sum_{i=1}^{m_c}\sum_{b=0}^{1}\sum_{x_k\in\mathcal{X}_b^i}\frac{1}{(d_{xz}^2)^{N_r\nu}}\right)^{d}. \tag{23}$$

After some manipulations $\Phi(s|\boldsymbol{n}_k)$ can be written as

$$\begin{aligned}\Phi(s|\boldsymbol{n}_k)\approx X_m(\nu,N_r,d)\,(2A)^{dN_r}\gamma^{-dN_r\nu}\,\mathrm{e}^{-s\sum_{k=1}^{d}\sum_{l=1}^{N_r}q_l|n_{k,l}|^{p_l}}\sum_{K=0}^{dN_r\bar{\nu}}\sum_{i_1+\cdots+i_{N_r}=K}s^{-2\sum_{l=1}^{N_r}(d\nu-i_l)/p_l}\\ \prod_{l=1}^{N_r}\sum_{j_1+\cdots+j_d=i_l}p_l^{-d}\,\xi_{j_1,l}\,|n_{1,l}|^{2j_1}\,q_l^{-2(\nu-j_1)/p_l}\ldots\,\xi_{j_d,l}\,|n_{d,l}|^{2j_d}\,q_l^{-2(\nu-j_d)/p_l}+o\left(\gamma^{-dN_r\nu}\right)\end{aligned} \tag{24}$$

where we have $0 \leq j_k \leq \bar{\nu}$ for $1 \leq k \leq d$, and $0 \leq i_l \leq K$ for $1 \leq l \leq N_r$. The PEP can be calculated from (24) and (17) as

$$P(\boldsymbol{c},\, \hat{\boldsymbol{c}}\,|\,\boldsymbol{n}_k) \approx X_m(\nu, N_r, d)\,(2A)^{dN_r}\gamma^{-dN_r\nu} \sum_{K=0}^{dN_r\bar{\nu}} \sum_{i_1+\cdots+i_{N_r}=K} M_{\boldsymbol{n}}(i_1,\ldots,i_{N_r}) \tag{25}$$

where we have defined the *generalized* noise moments $M_{\boldsymbol{n}}(i_1,\ldots,i_{N_r})$ as

$$M_{\boldsymbol{n}}(i_1,\ldots,i_{N_r}) \triangleq \mathcal{E}_{\boldsymbol{n}_k}\left\{ \frac{(\sum_{k=1}^{d}\sum_{l=1}^{N_r} q_l|n_{k,l}|^{p_l})^{2\sum_{l=1}^{N_r}(d\nu-i_l)/p_l}}{\Gamma(2\sum_{l=1}^{N_r}(d\nu-i_l)/p_l+1)\,\prod_{l=1}^{N_r} p_l^d q_l^{2(d\nu-i_l)/p_l}} \prod_{l=1}^{N_r} \sum_{j_1+\cdots+j_d=i_l} \prod_{k=1}^{d} \xi_{j_k,l}\,|n_{k,l}|^{2j_k} \right\}. \tag{26}$$

Based on (25) and (10) a closed–form approximation for the asymptotic BER can be obtained as

$$P_b \doteq \frac{w_c(d_f)}{k_c}\, X_m(\nu, N_r, d_f)\,(2A)^{d_f N_r}\gamma^{-d_f N_r\nu} \sum_{K=0}^{d_f N_r\bar{\nu}} \sum_{i_1+\cdots+i_{N_r}=K} M_{\boldsymbol{n}}(i_1,\ldots,i_{N_r}) \tag{27}$$

In deriving the above equation we have used the fact that the BER bound in (10) is tight for high SNR's (cf. Section 6). Furthermore, we have taken into account the fact that the first term in (10) for which $d = d_f$ is asymptotically dominant. We note that for non–integer ν, due to the approximation made in Appendix A the asymptotic BER given in (27) constitutes an approximation to the behavior of the BER bound for $\gamma \to \infty$. However, as will be shown in Section 6 the incurred approximation error is negligible for all practical purposes and therefore the obtained asymptotic BER is still a valid criterion for metric optimization.

The asymptotic expression for BER given in (27) is very general since it is applicable to different types of fading channels, a large class of noise and different code rates used in the CR system. Furthermore, (27) reveals the dependance of the asymptotic BER on the metric parameters $\boldsymbol{q}$ and $\boldsymbol{p}$, and therefore can be used for metric optimization. Although (27) is simple enough to be directly used for online metric optimization (cf. Section 5), in several practically relevant special cases less general, yet simpler expressions can be obtain for the asymptotic performance that can facilitate the task of online metric optimization. In the following, we consider important especial cases where such simplifications are possible in (27). Since the main difficulty in obtaining the BER performance arises from the calculation of generalized noise moments $M_{\boldsymbol{n}}(i_1,\ldots,i_{N_r})$, we will mainly focus on obtaining simplified expressions for these noise moments.

4.1 Case 1, Fading Channels with $\nu = 1$:

If $\nu = 1$ is valid for the fading channel (e.g. (possibly spatially correlated) Rayleigh, Ricean, and Nakagami–q fading), (26) can be significantly simplified. In this case we have $\bar{\nu} = \lceil \nu \rceil - 1 = 0$ which results in $i_l = 0$, $1 \leq l \leq N_r$ and $j_k = 0$, $1 \leq k \leq d_f$. Therefore, (26) simplifies to

$$M_{\boldsymbol{n}}(0, \ldots, 0) = \frac{\prod_{l=1}^{N_r} (\Gamma(2/p_l))^{d_f} p_l^{-d_f} q_l^{-2d_f/p_l}}{\Gamma(\sum_{l=1}^{N_r} 2d_f/p_l + 1)} \mathcal{E}_{\boldsymbol{n}_k} \left\{ (\sum_{k=1}^{d_f} \sum_{l=1}^{N_r} q_l |n_{k,l}|^{p_l})^{\sum_{l=1}^{N_r} 2d_f/p_l} \right\}. \tag{28}$$

and the asymptotic BER is therefore given by

$$P_b \doteq \frac{w_c(d_f)}{k_c} X_m(1, N_r, d_f) \, (2A)^{d_f N_r} \gamma^{-d_f N_r} M_{\boldsymbol{n}}(0, \ldots, 0) \tag{29}$$

Based on (29) it is possible to obtain the asymptotic BER performance for uncoded transmission with maximum–ratio combining (MRC) at the receiver by allowing $d_f = 1$, $k_c = 1$, and $w_c(1) = 1$. This results in

$$P_b \doteq X_m(1, N_r, 1) \, (2A)^{N_r} \gamma^{-N_r} \frac{\prod_{l=1}^{N_r} \Gamma(2/p_l) p_l^{-1} q_l^{-2/p_l}}{\Gamma(\sum_{l=1}^{N_r} 2/p_l + 1)} \mathcal{E}_{\boldsymbol{n}_k} \left\{ (\sum_{l=1}^{N_r} q_l |n_{k,l}|^{p_l})^{\sum_{l=1}^{N_r} 2/p_l} \right\} \tag{30}$$

It is easy to see that (30) is in agreement with [30, Eq. (13)] which is obtained assuming Rayleigh and Rician fading channels. Eq. (30) is however more general than the results obtained in [30] as it is not limited to Rayleigh and Rician fading channels and is also applicable to all other fading types with $\nu = 1$ such as Nagakami–q fading.

4.2 Case 2, Spatially i.i.d. Noise:

In many practical scenarios the elements of $\boldsymbol{n}_k$ are i.i.d. (e.g. due to sufficiently large antenna spacing at the receiver). When the noise is specially i.i.d. without loss of optimality we can set $q_l = 1$ and $p_l = p$, $1 \leq l \leq N_r$. Thus, (26) can be written as

$$M_{\boldsymbol{n}}(i_1, \ldots, i_{N_r}) = \mathcal{E}_{\boldsymbol{n}_k} \left\{ \frac{(\sum_{k=1}^{d_f} \sum_{l=1}^{N_r} |n_{k,l}|^p)^{2(d_f \nu N_r - K)/p}}{\Gamma(2(d_f \nu N_r - K)/p + 1) p^{d_f N_r}} \prod_{l=1}^{N_r} \sum_{j_1 + \cdots + j_{d_f} = i_l} \prod_{k=1}^{d} \xi_{j_{k,l}} |n_{k,l}|^{2j_k} \right\}. \tag{31}$$

Evidently, the number of metric parameters to be optimized reduces to one in this case and therefore the task of metric optimization is greatly simplified. In the following we consider two cases where even more simplifications are possible.

4.2.1 Case of $d_f N_r \gg 1$:

For $d_f N_r \gg 1$ and spatially i.i.d. noise we can invoke the strong low of large numbers to approximate the term $\sum_{k=1}^{d_f}\sum_{l=1}^{N_r}|n_{k,l}|^p$ as

$$\sum_{k=1}^{d_f}\sum_{l=1}^{N_r}|n_{k,l}|^p \approx N_r d_f\, m_n(p) \tag{32}$$

where $m_n(p) \triangleq \mathcal{E}\{|n_{k,l}|^p\}$ are the scalar moment of the noise. Using (32) in (26) leads to

$$M_{\boldsymbol{n}}(i_1,\ldots,i_{N_r}) \approx \frac{(N_r d_f\, m_n(p))^{2(d_f\nu N_r-K)/p}}{\Gamma(2(d_f\nu N_r-K)/p+1)\,p^{d_f N_r}}\prod_{l=1}^{N_r}\sum_{j_1+\cdots+j_{d_f}=i_l}\prod_{k=1}^{d_f}\xi_{j_k,l}\,m_n(2j_k). \tag{33}$$

4.2.2 Case of L_2–Norm Branch Metric:

The joint noise moments in this case can be found by allowing $p=2$ in (31). Furthermore, the asymptotic BER performance can be obtained from (27) as

$$P_b \doteq \frac{w_c(d_f)}{k_c}\,X_m(\nu,N_r,d_f)\,(2A)^{d_f N_r}\gamma^{-d_f N_r\nu}\sum_{K=0}^{d_f N_r\bar{\nu}}\sum_{i_1+\cdots+i_{N_r}=K}M_{\boldsymbol{n}}(i_1,\ldots,i_{N_r}). \tag{34}$$

It can be shown that (34) is in complete agreement with the results reported in [21] which are obtained for BICM–based systems employing L_2–norm branch metric in non–Gaussian noise environments.

4.3 Diversity Gain and Coding Gain

For completeness, in this subsection we calculate the diversity gain G_d (i.e., the asymptotic slope of the asymptotic BER curve on a log–log scale) and the coding gain G_c (i.e., a relative horizontal shift of the asymptotic BER curve). Thereby, the diversity gain and the coding gain can be obtain by comparing the general asymptotic results in (27) with $P_b \doteq (G_c\gamma)^{-G_d}$ [31] as

$$G_d = d_f N_r\nu \tag{35}$$

$$G_c\,[\mathrm{dB}] = -\frac{10}{G_d}\log_{10}\left(\frac{w_c(d_f)X_m(\nu,N_r,d)\,(2A)^{dN_r}}{k_c}\right)-\frac{10}{G_d}\log_{10}\sum_{K=0}^{dN_r\bar{\nu}}\sum_{i_1+\cdots+i_{N_r}=K}M_{\boldsymbol{n}}(i_1,\ldots,i_{N_r}) \tag{36}$$

From (35) we observe that the diversity gain is independent of the metric parameters $\boldsymbol{q}$ and $\boldsymbol{p}$ and is also independent of type of noise. Therefore the asymptotic BER curves for all noise types are parallel

for different choices of $\boldsymbol{q}$ and $\boldsymbol{p}$. Eq. (36) reveals that the coding gain consists of two terms. The first term depends on the types of convolutional code, signal constellation and the fading channel but is independent of the metric parameters $\boldsymbol{q}$ and $\boldsymbol{p}$ and the statistics of the noise. The second term is a function of $\boldsymbol{q}$ and $\boldsymbol{p}$, as well as the properties of the noise via the generalized moments $M_{\boldsymbol{n}}(i_1, \ldots, i_{N_r})$ of the noise. Eqs. (35) and (36) show that minimum BER can be achieved for a given SNR by optimizing the L_p–norm metric parameters, which results in shifting the asymptotic BER curves to the left as far as possible in a log–log scale.

5 Metric Optimization

In this section, we use the analytical results obtained in Sections 3 and 4 to optimize the parameters of the L_p–norm metric employed in the CR system. We consider both offline and online metric optimization depending on the availability of the noise statistics at the CR receiver.

5.1 Offline Metric Optimization

In scenarios where the noise statistics are known *a priori*, the task of metric optimization can be performed offline. Since in such cases computational complexity is not a major concern, we use the analytical BER bound obtained in (10) for metric optimization. We illustrate in Fig. 1 how this BER bound can be used for offline metric optimization. To simplify the exposition, in this figure we have assumed spatially i.i.d. noise for which only a single metric parameter p has to be optimized. Thereby, in Fig. 1 we have shown the BER bound for different types noise defined in Subsection 2.3 vs. p. For comparison, we have also shown the BER obtained via Monte–Carlo simulation and the asymptotic BER obtained in (27). The observed differences between the BER bound, simulation and asymptotic results are due to assuming a finite value for SNR (15 dB) in this figure. Nevertheless, Figure 1 shows that for each type of noise the minimum BER happens at approximately the same value of p for all the three curves.

5.2 Online Metric Optimization

Online metric optimization becomes necessary when the noise statistics are not known at the receiver or vary quickly with time. In such scenarios, a cost function simple enough to be evaluated in real–

time is needed to enable online and adaptive optimization of the metric parameters. We therefore base our online metric optimization on the simple–to–evaluate asymptotic BER results obtained in Section 4. In particular, in this subsection, we consider online metric optimization based on the asymptotic results obtained in (27) and (26), and provide adaptive algorithms that can be use to perform the optimization effectively. Due to random nature of the optimization problem, we propose an stochastic optimization algorithm to perform the online optimization efficiently. Although several types of stochastic optimization algorithms are available in the literature [32], numerical evidence has taught us that among them the finite–difference (FD) stochastic approximation (SA) algorithm is the most suitable for the problem at hand.

The cost function for FDSA algorithm can be obtained from (27) and (26) as

$$L_k(\boldsymbol{\theta}) = \sum_{K=0}^{d_f N_r \bar{\nu}} \sum_{i_1+\cdots+i_{N_r}=K} M_{\boldsymbol{n}_k}(i_1,\ldots,i_{N_r};\boldsymbol{\theta}) \tag{37}$$

with

$$M_{\boldsymbol{n}_k}(i_1,\ldots,i_{N_r};\boldsymbol{\theta}) \triangleq \frac{(\sum_{k=1}^{d}\sum_{l=1}^{N_r} q_l |n_{k,l}|^{p_l})^{2\sum_{l=1}^{N_r}(d\nu-i_l)/p_l}}{\Gamma(2\sum_{l=1}^{N_r}(d\nu-i_l)/p_l+1)\prod_{l=1}^{N_r} p_l^d q_l^{2(d\nu-i_l)/p_l}} \prod_{l=1}^{N_r} \sum_{j_1+\cdots+j_d=i_l} \prod_{k=1}^{d} \xi_{j_k,l}\,|n_{k,l}|^{2j_k} \tag{38}$$

where $M_{\boldsymbol{n}_k}(i_1,\ldots,i_{N_r};\boldsymbol{\theta})$ is the instantaneous estimate for the generalized noise moments, and we have omitted terms and parameters that do not effect the optimization. Furthermore, $\boldsymbol{\theta}$ is a vector containing all the metric parameters to be optimized, i.e., we have defined $\boldsymbol{\theta} \triangleq [q_2,\ldots,q_{N_r},p_1,\ldots,p_{N_r}]^T$ where without loss on optimality we have assumed $q_1 = 1$. We note that simplified cost functions can be obtained in the special cases described in Subsections 4.1 and 4.2.

The proposed FDSA algorithm employed to minimize the cost function in (37) can be explained as follows: The algorithm recursively updates the estimate $\boldsymbol{\theta}_k$ of the optimal $\boldsymbol{\theta}$, i.e., in the kth iteration, the the estimate $\boldsymbol{\theta}_{k+1}$ is obtained as [32]

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + a_k \hat{\boldsymbol{g}}(\boldsymbol{\theta}_k) \tag{39}$$

$$\hat{\boldsymbol{g}}(\boldsymbol{\theta}_k) = \left[\frac{L_k(\boldsymbol{\theta}_k + c_k\boldsymbol{e}_1) - L_k(\boldsymbol{\theta}_k - c_k\boldsymbol{e}_1)}{2c_k} \cdots \frac{L_k(\boldsymbol{\theta}_k + c_k\boldsymbol{e}_{2N_r-1}) - L_k(\boldsymbol{\theta}_k - c_k\boldsymbol{e}_{2N_r-1})}{2c_k}\right] \tag{40}$$

where $a_k > 0$ and $c_k > 0$ are the gain sequences of the FDSA algorithm, and $\boldsymbol{e}_i$ denotes a vector with a 1 in the ith place and 0's elsewhere. The convergence theory for the FDSA algorithm [32] states that if the standard conditions $a_k \to 0$, $c_k \to 0$, $\sum_{k=0}^{\infty} a_k = \infty$, and $\sum_{k=0}^{\infty} a_k^2/c_k^2 < \infty$ on the gain sequences are met, the algorithm converges to a local or global minimum of BER.

6 Numerical and Simulation Results

7 Conclusions

A The pdf $f_y(y_{k,l})$ of $y_{k,l}$

To obtained $f_y(y_{k,l})$, we first obtain the pdf $f_X(X_{k,l})$ of $X_{k,l} \triangleq |\sqrt{\gamma}h_{k,l}e_k + n_{k,l}|^2$ and then calculate the pdf $f_y(y_{k,l})$ of $y_{k,l} = X_{k,l}^{p_l/2}$ using

$$f_y(y_{k,l}) = 2/p_l \, f_X(y_{k,l}^{2/p_l}) \, y_{k,l}^{2/p_l - 1}. \tag{41}$$

We start the calculation of $f_X(X_{k,l})$ by performing the following reformulations

$$\begin{aligned} X_{k,l} &= |\sqrt{\gamma}h_{k,l}e + n_{k,l}|^2 = |\sqrt{\gamma}a_{k,l}e + \hat{n}_{k,l}|^2 \\ &= \gamma a_{k,l}^2 d_{xz}^2 + 2\sqrt{\gamma} d_{xz} a_{k,l} \Re\{\hat{n}_{k,l}\} + |n_{k,l}|^2, \end{aligned} \tag{42}$$

where we have defined $\hat{n}_{k,l} \triangleq n_{k,l}\mathrm{e}^{-j\Theta_{k,l}}$ and $e \triangleq |d_{xz}|$, and used the fact that $h_{k,l} = a_{k,l}\mathrm{e}^{j\Theta_{k,l}}$ (cf. Section 2). The MGF of $X_{k,l}$ can now be obtained as

$$\Phi_{X_{k,l}}(s) \triangleq \mathcal{E}_{a_{k,l},\Theta_{k,l}}\{\mathrm{e}^{-sX_{k,l}}\} = \mathrm{e}^{-s|n_{k,l}|^2}\mathcal{E}_{a_{k,l},\Theta_{k,l}}\{\mathrm{e}^{-s\gamma a_{k,l}d_{xz}^2}\,\mathrm{e}^{-s2\sqrt{\gamma}d_{xz}a_{k,l}\Re\{\hat{n}_{k,l}\}}\} \tag{43}$$

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

$$\Phi_{X_{k,l}}(s) = \mathrm{e}^{-s|n_{k,l}|^2}\mathcal{E}_{a_{k,l},\Theta_{k,l}}\left\{\mathrm{e}^{-s\,\gamma\,a_{k,l}^2\,d_{xz}^2}\,\sum_{i=0}^{\infty}\frac{\left(-2\sqrt{\gamma}\,d\,a_{k,l}\,\Re\{\hat{n}_{k,l}\}s\right)^i}{i!}\right\}. \tag{44}$$

For $\gamma \to \infty$, (5) can be used along with the integral $\int_0^\infty x^{\mu-1}e^{-px^2}\,\mathrm{d}x = p^{\mu/2}\Gamma(\mu/2)$ [33, 3.462] to rewrite the above equation as

$$\begin{aligned} \Phi_{X_{k,l}}(s) &= \frac{A\,\mathrm{e}^{-s|n_{k,l}|^2}}{(\gamma d_{xz}^2 s)^\nu}\sum_{i=0}^{\infty}2^i\Gamma(\nu+i/2)\mathcal{E}_{\Theta_{k,l}}\{\Re\{\hat{n}_{k,l}\}^i\}s^{i/2} + o\left(\gamma^{-\nu}\right) &(45)\\ &= \frac{A\,\mathrm{e}^{-s|n_{k,l}|^2}}{(\gamma d_{xz}^2 s)^\nu}\sum_{i=0}^{\infty}\beta_i|n_{k,l}|^{2i}s^i + o\left(\gamma^{-\nu}\right). &(46)\end{aligned}$$

where $\beta_i \triangleq \frac{\Gamma(\nu+i)}{(i!)^2}$. In deriving (46) we have used

$$\mathcal{E}_{\Theta_{k,l}}\{\Re\{\hat{n}_{k,l}\}^i\} = \begin{cases} \frac{i/2+1/2}{\sqrt{\pi}\Gamma(i/2+1)}|n_{k,l}|^i & i \text{ even} \\ 0 & i \text{ odd}. \end{cases} \tag{47}$$

Using the Taylor series expansion $\mathrm{e}^{-s|n_{k,l}|^2} = \sum_{i=0}^{\infty}(-1)^i|n_{k,l}|^{2i}s^i/i!$ in (46) yields

$$\Phi_{X_{k,l}}(s) = \frac{A}{(\gamma d_{xz}^2 s)^\nu}\sum_{i=0}^{\infty} P_i\,|n_{k,l}|^{2i}\,s^i + o\left(\gamma^{-\nu}\right). \tag{48}$$

where

$$P_i \triangleq \{(-1)^i/i!\}_{i=0}^{\infty} \otimes \{\Gamma(\nu+i)/(i!)^2\}_{i=0}^{\infty} = \sum_{\lambda=0}^{i}\frac{(-1)^{(i-\lambda)}\,\Gamma(\nu+\lambda)}{(\lambda!)^2\,(i-\lambda)!}. \tag{49}$$

If ν is an integer it can be shown that

$$P_i = \begin{cases} \binom{\nu-1}{i}^2(\nu-i-1)! & 0 \le i \le \nu-1 \\ 0 & \nu-1 < i \end{cases} \tag{50}$$

Therefore for integer ν, $\Phi_{X_{k,l}}(s)$ is given by a finite power series in s. For non–integer ν however, $\Phi_{X_{k,l}}(s)$ is given by a infinite power series as shown in (48). It can be shown that truncating this power series after $\bar{\nu} = \lceil\nu\rceil - 1$ terms will result in a close approximation for $\Phi_{X_{k,l}}(s)$. This approximation can therefore be written as

$$\Phi_{X_{k,l}}(s) \approx \frac{A}{(\gamma d_{xz}^2 s)^\nu}\sum_{i=0}^{\bar{\nu}} P_i\,|n_{k,l}|^{2i}\,s^i + o\left(\gamma^{-\nu}\right). \tag{51}$$

The pdf $f_X(X_{k,l})$ of $X_{k,l}$ can now be calculated by obtaining the inverse Laplace transform of $\Phi_{X_{k,l}}(s)$ as

$$f_X(X_{k,l}) \approx \frac{A}{(\gamma d_{xz}^2)^\nu}\sum_{i=0}^{\bar{\nu}}\frac{P_i}{\Gamma(\nu-i)}\,|n_{k,l}|^{2i}\,X_{k,l}^{\nu-i-1} + o\left(\gamma^{-\nu}\right). \tag{52}$$

Finally, the pdf $f_y(y_{k,l})$ of $y_{k,l}$ can be found based on (52) and (41) as

$$f_y(y_{k,l}) \approx \frac{2A}{p_l(\gamma d_{xz}^2)^\nu}\sum_{i=0}^{\bar{\nu}}\frac{P_i}{\Gamma(\nu-i)}\,|n_{k,l}|^{2i}\,y_{k,l}^{2(\nu-i)/p_l-1} + o\left(\gamma^{-\nu}\right). \tag{53}$$

The above equation gives the exact pdf $f_y(y_{k,l})$ for integer ν, but involves an approximation to this pdf for non–integer ν. This approbation has been found to be reasonably accurate for the purpose of metric optimization (cf. Section 6).

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] T. Li, W.H. Mow, V.K.N. Lau, M. Siu, R.S. Cheng, and R.D. 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.

[5] S. Haykin. Cognitive Radio: BrainEmpowered Wireless Communications. *IEEE J. Select. Areas Commun.*, 23:201–220, February 2005.

[6] J. Lunden, S.A. Kassam, and V. Koivunen. Nonparametric Cyclic Correlation Based Detection for Cognitive Radio Systems. In *Proceedings of the International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, pages 1–6, May 2008.

[7] 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.

[8] A. Nasri, R. Schober, and L. Lampe. Performance of a BPSK NB Receiver in MB-OFDM UWB Interference. *In Proceedings of the International Conference on Communications (ICC)*, June 2006.

[9] 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.

[10] D. Middleton. Statistical-physical Models of Man–made Radio Noise – Parts I and II. *U.S. Dept. Commerce Office Telecommun.*, April 1974 and 1976.

[11] S.A. Kassam and H.V. Poor. Robust Techniques for Signal Processing: A Survey. In *Proceedings of the IEEE*, volume 73, pages 433–481, 1985.

[12] 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.

[13] 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*, 52:1206–1211, March 2006.

[14] H. Bölcskei. MIMO–OFDM Wireless Systems: Basics, Perspectives, and Challenges. *IEEE Wireless Commun.*, 13:31–37, August 2006.

[15] G. Caire, G. Taricco, and E. Biglieri. Bit–Interleaved Coded Modulation. *IEEE Trans. Inform. Theory*, 44:927–946, May 1998.

[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] M.K. Simon and M.-S. Alouini. *Digital Communication over Fading Channels*. Wiley, Hoboken, New Jersey, 2005.

[19] O.C. Ugweje and V.A. Aalo. Performance of Selection Diversity System in Correlated Nakagami Fading. In *Proceedings of IEEE Veh. Techn. Conf. (VTC)*, pages 1488–1492, May 1997.

[20] C. Tan and N. Beaulieu. Infinite Series Representations of the Bivariate Rayleigh and Nakagamim Distributions. *IEEE Trans. Commun.*, 45:1159–1161, October 1997.

[21] A. Nasri and R. Schober. Performance of BICM-SC and BICM-OFDM Systems with Diversity Reception in NonGaussian Noise and Interference. *Submitted to IEEE Trans. on Commun. [Online] http://www.ece.ubc.ca/∼amirn/TCOM-08.pdf*, 2008.

[22] T.S. Rappaport. *Wireless Communications*. Prentice Hall, Upper Saddle River, NJ, 2002.

[23] 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.

[24] H. Nguyen and T. Bui. Bit–Interleaved Coded Modulation With Iterative Decoding in Impulsive Noise. *IEEE Trans. Power Delivery*, 22:151–160, January 2007.

[25] C. Tepedelenlioglu and P. Gao. On Diversity Reception Over Fading Channels with Impulsive Noise. *IEEE Trans. Veh. Technol.*, 54:2037–2047, November 2005.

[26] A. Coulson. Bit Error Rate Performance of OFDM in Narrowband Interference with Excission Filtering. *IEEE Trans. Wireless Commun.*, 5:2484–2492, September 2006.

[27] 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.

[28] V. Sethuraman and B. Hajek. Comments on "Bit–Interleaved Coded Modulation". *IEEE Trans. Inform. Theory*, 52:1795–1797, April 2006.

[29] 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.

[30] A. Nasri, A. Nezampoor, and R. Schober. Adaptive Lp-Norm Diversity Combining in Non-Gaussian Noise and Interference. *Submitted to IEEE Trans. on Wireless Commun. [Online] http://www.ece.ubc.ca/∼amin/TW-08.pdf*, 2008.

[31] Z. Wang and G. Giannakis. A Simple and General Parameterization Quantifying Performance in Fading Channels. *IEEE Trans. Commun.*, 51:1389–1398, August 2003.

[32] J. Spall. *Introduction to Stochastic Search and Optimization*. Wiley & Sons, Inc., New Jersey, 2003.

[33] I. Gradshteyn and I. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 2000.

[34] 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.

[35] 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.

Tables and Figures:

Table 1: Pdf $p_a(a)$ of fading amplitude a for popular fading models and corresponding values for α_c and α_d. We have omitted subscript l for convenience. The parameters for Rayleigh ($\boldsymbol{C}_{hh}$), Ricean ($\boldsymbol{\mu}_h$, $\boldsymbol{C}_{hh}$), and Nakagami–m (m, $\boldsymbol{C}_{aa}$) fading are defined in Appendix **??**. The parameters for Nakagami–q (q, b) and Weibull (c) fading are defined as in [18].

Channel type	$p_a(a)$ of the fading amplitude a	α_c	α_d
Rayleigh	$2\,a\,e^{-a^2}$	$\det(\boldsymbol{C}_{hh})^{-1/N_R}$	1
Ricean	$2(K+1)\,a\,e^{-K-(1+K)a^2}\,\mathrm{I}_0\left(2a\sqrt{K(K+1)}\right)$	$\left(\frac{\exp\left(-\boldsymbol{\mu}_h^H\boldsymbol{C}_{hh}^{-1}\boldsymbol{\mu}_h\right)}{\det(\boldsymbol{C}_{hh})}\right)^{1/N_R}$	1
Nakagami–m	$\frac{2}{\Gamma(m)}\,m^m\,a^{2m-1}\,e^{-ma^2}$	$\frac{m^m}{\Gamma(m)}\,\det(\boldsymbol{C}_{aa})^{-m/N_R}$	m
Nakagami–q	$\frac{2a}{\sqrt{1-b^2}}\,\exp\left(-\frac{a^2}{(1-b^2)}\right)\,\mathrm{I}_0\left(\frac{ba^2}{(1-b^2)}\right)$	$\frac{1+q^2}{2q}$	1
Weibull	$c\left(\Gamma(1+\frac{2}{c})\right)^{\frac{c}{2}}\,a^{c-1}\exp\left(-\left(a^2\Gamma(1+\frac{2}{c})\right)^{\frac{c}{2}}\right)$	$\frac{c}{2}\,(\Gamma(1+\frac{2}{c}))^{\frac{c}{2}}$	$\frac{c}{2}$

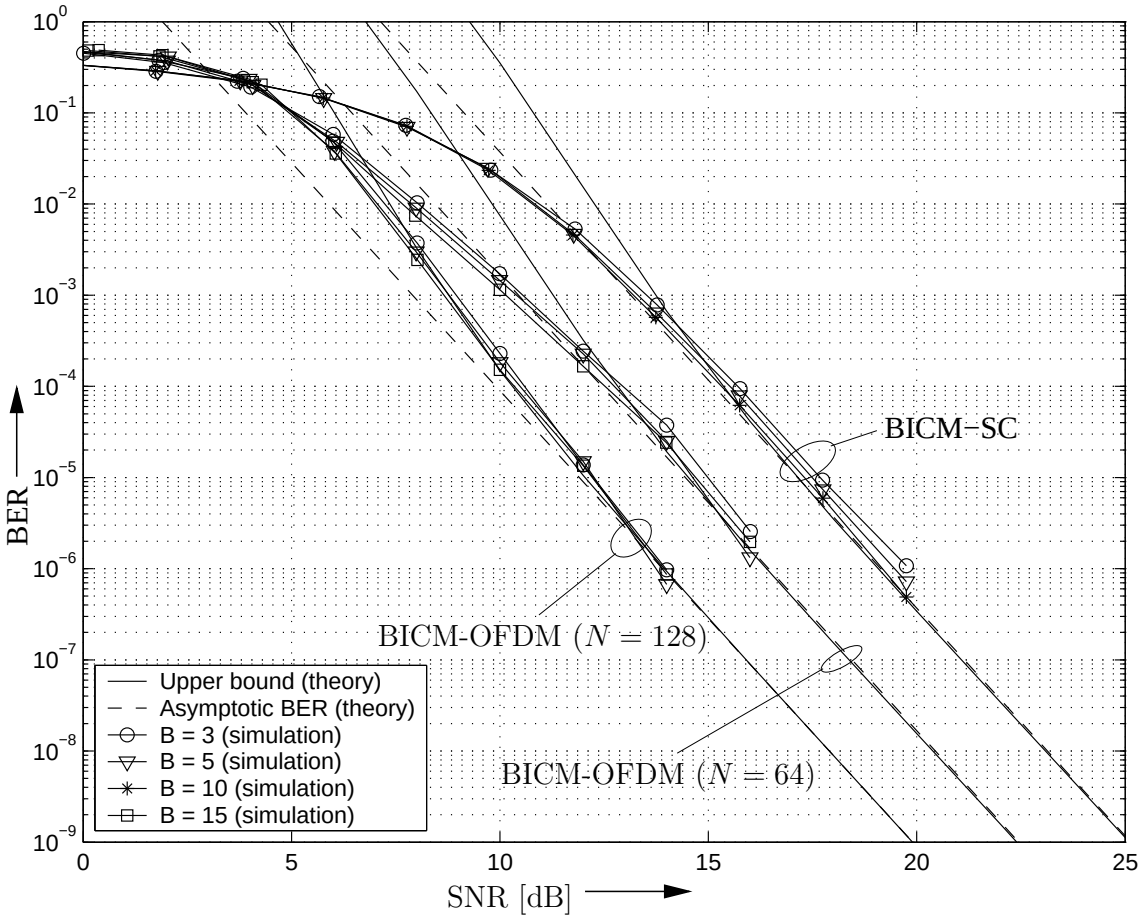

Figure 1: BER of BICM–SC and BICM–OFDM impaired by GMN (ϵ–mixture noise, $\epsilon = 0.1$, $\kappa = 100$) and NBI, respectively, vs. SNR γ. $R_c = 3/4$, Rayleigh fading, 4–PSK, and $N_R = 1$. BICM–SC: Frequency–flat time–selective fading, $N = 972$, and $B_f T = 0.007$. BICM–OFDM ($N = 64$): Frequency–selective Rayleigh fading with $L = 10$ and B equal power, sub–carrier centered NBI signals with $I_\mu = 1$, $1 \leq \mu \leq B$, $\kappa = 7$. BICM–OFDM ($N = 128$): Frequency–selective Rayleigh fading with $L = 20$ and B equal power, sub–carrier centered NBI signals with $I_\mu = 1$, $1 \leq \mu \leq B$, $\kappa = 2$. Solid lines with markers: Simulated BER. Solid lines without markers: BER bound (**??**). Dashed lines: Asymptotic BER (**??**).

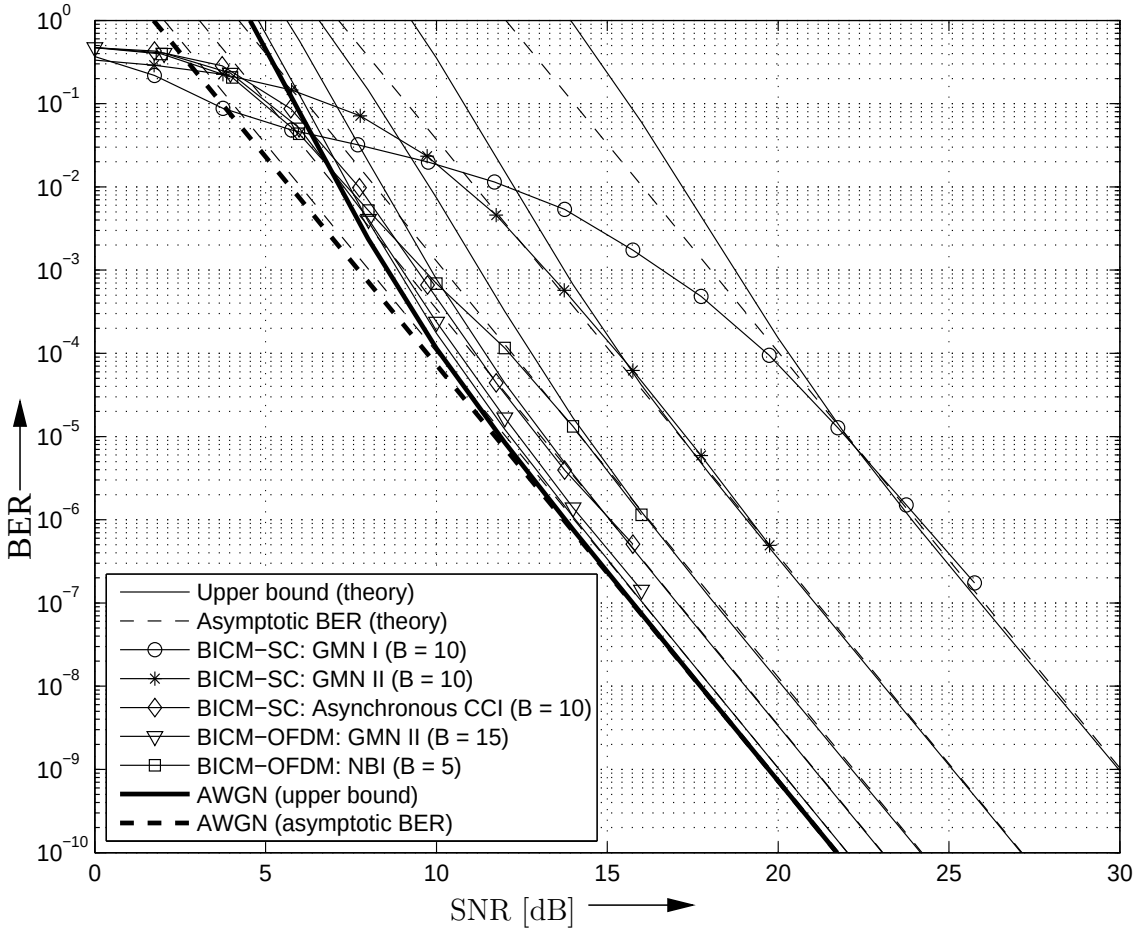

Figure 2: BER of BICM–SC and BICM–OFDM impaired by various types of noise vs. SNR γ. Rayleigh fading, $R_c = 3/4$, 4–PSK, and $N_R = 1$. BICM–SC: $N = 972$ and $B_f T = 0.007$. BICM–OFDM: $N = 64$ and $L = 10$. GMN I: ϵ–mixture noise, $\epsilon = 0.01$, $\kappa = 100$. GMN II: ϵ–mixture noise, $\epsilon = 0.1$, $\kappa = 100$. Asynchronous CCI: Two asynchronous equal power 4–PSK CCI signals, $I_\mu = 1$, $\mu \in \{1, 2\}$, $I_\mu = 0$, $3 \leq \mu \leq 10$, raised cosine pulses $g_{1,\mu}(t)$, $\mu \in \{1, 2\}$, with roll–off factor 0.3, $\tau_{1,\mu} = 0.3T$, $\mu \in \{1, 2\}$, $\kappa = 2$. NBI: One sub–carrier–centered NBI signal, $I_1 = 1$, $I_2 = I_3 = I_4 = I_5 = 0$, $\kappa = 9$. Solid lines with markers: Simulated BER. Solid lines without markers: BER bound (**??**). Dashed lines: Asymptotic BER (**??**).

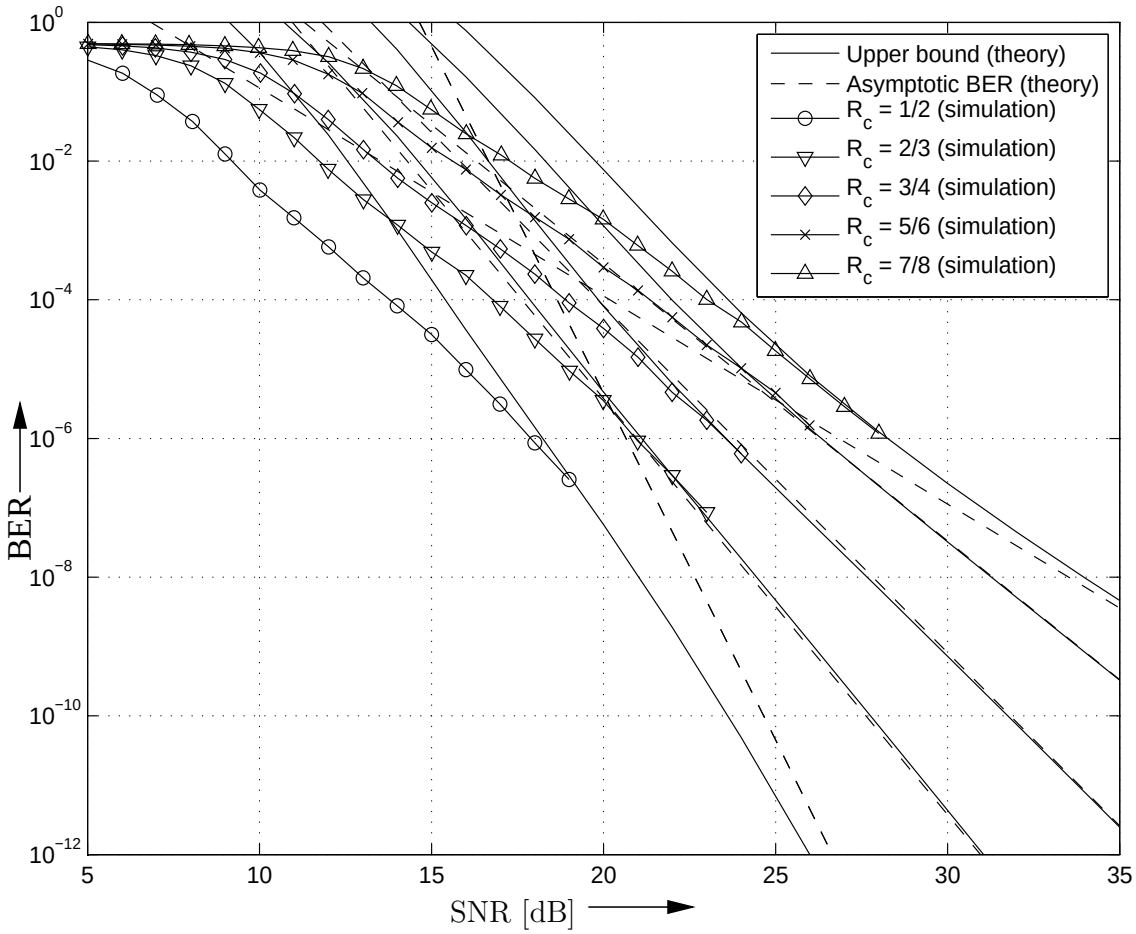

Figure 3: BER of BICM–OFDM impaired by NBI (3 equal power, sub–carrier–centered NBI signals, $I_1 = I_2 = I_3 = 1$, $\kappa = 10$) vs. SNR γ. I.i.d. Rayleigh fading, 64–QAM, $N = 128$, $B = 3$, and $N_R = 1$. Solid lines with markers: Simulated BER. Solid lines without markers: BER bound (**??**). Dashed lines: Asymptotic BER (**??**).

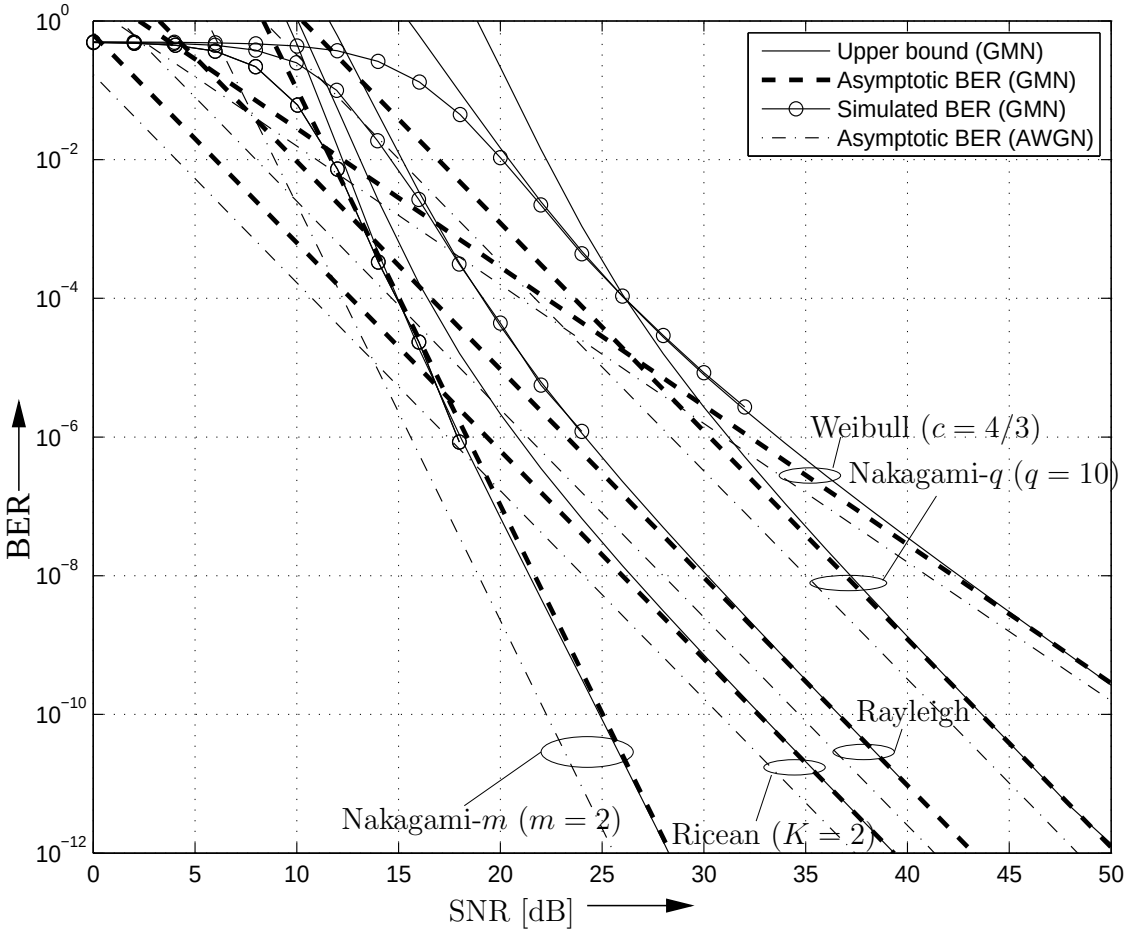

Figure 4: BER of BICM–SC impaired by GMN (ϵ–mixture noise, $\epsilon = 0.25$, $\kappa = 10$) and AWGN, respectively, vs. SNR γ. Ideal i.i.d. fading, $R_c = 7/8$, 16–QAM, and $N_R = 1$.

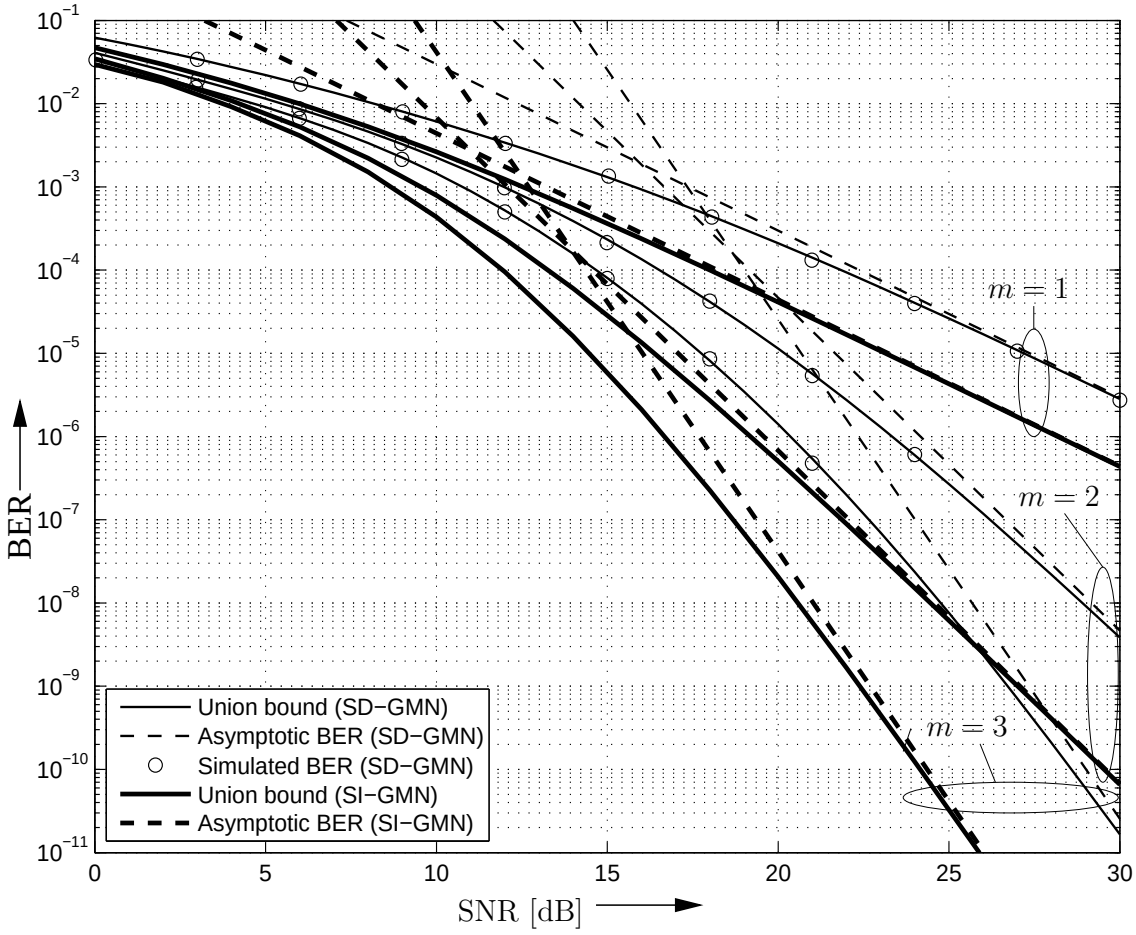

Figure 5: BER of uncoded SC transmission impaired by SD- and SI–GMN (ϵ–mixture noise, $\epsilon = 0.1$, $\kappa = 10$), respectively, vs. SNR γ. $N_R = 2$, Nakagami–m fading spatial correlation $\rho_a = 0.9$, and 4–PSK.

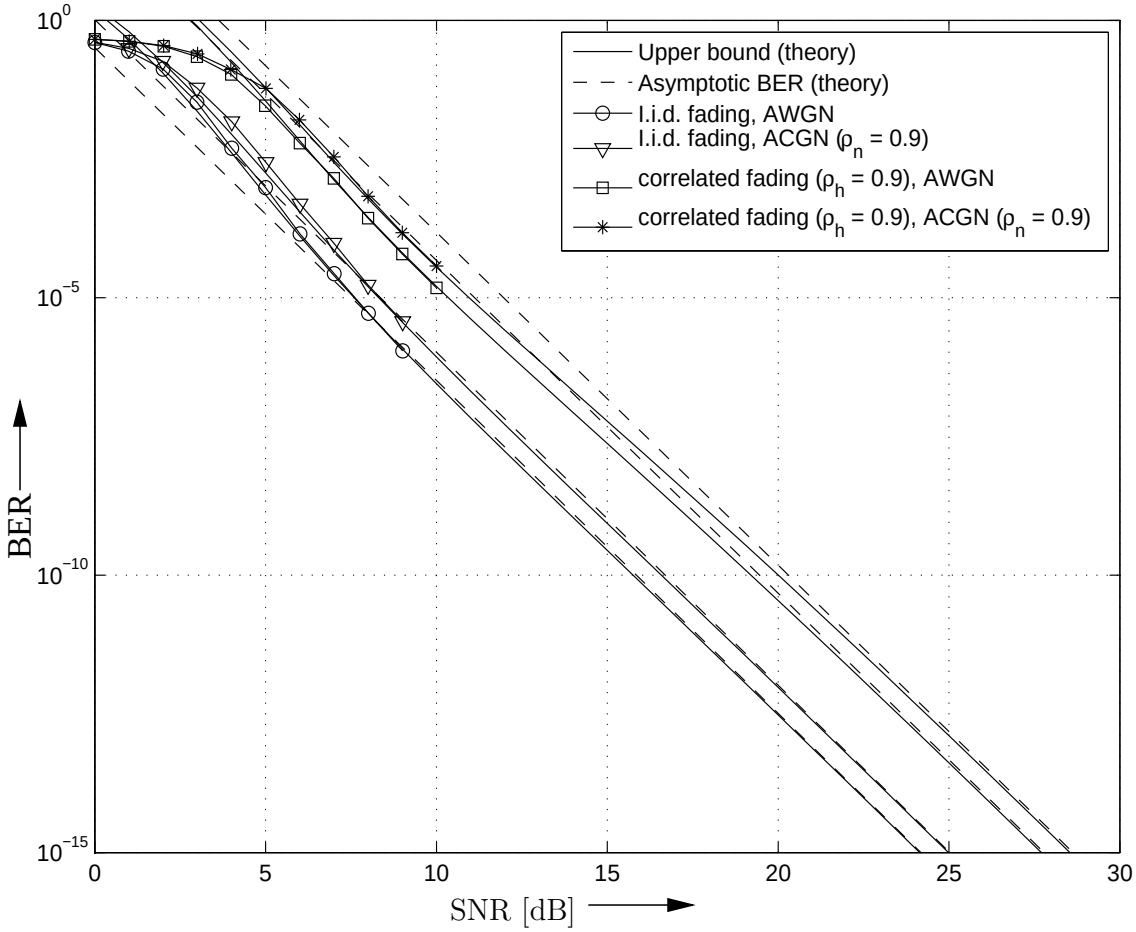

Figure 6: BER of BICM–SC impaired by AWGN/ACGN vs. SNR γ. Spatially i.i.d./spatially correlated, temporally i.i.d. Rayleigh fading, $R_c = 7/8$, 4–PSK, and $N_R = 2$. Solid lines with markers: Simulated BER. Solid lines without markers: BER bound (**??**). Dashed lines: Asymptotic BER (**??**).

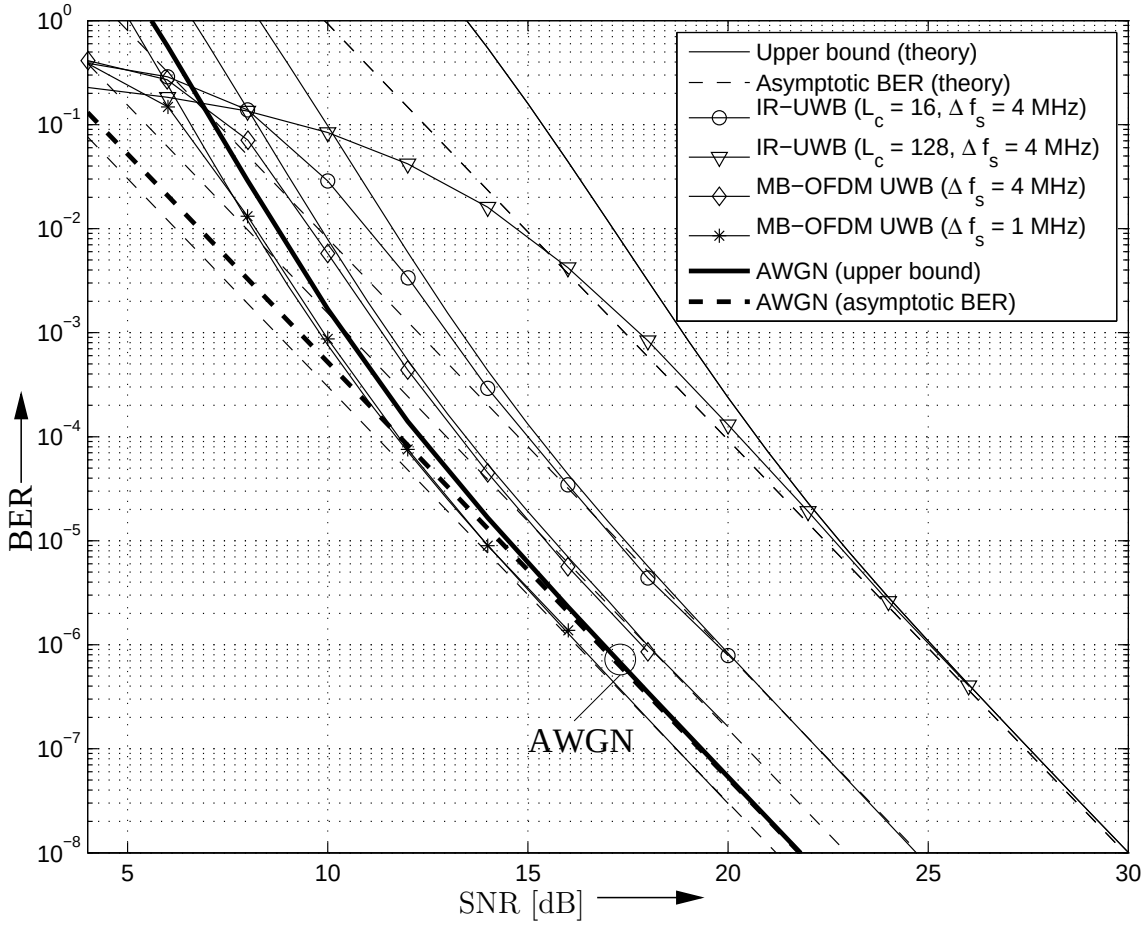

Figure 7: BER of BICM–OFDM system with sub–carrier spacing Δf_s impaired by IR–UWB [34] ($N_b = 8$ bursts per symbol and L_c chips per burst) and MB–OFDM UWB [35], respectively, vs. SNR γ. Ideal i.i.d. Rayleigh fading, $R_c = 5/6$, 4–PSK, and $N_R = 1$. Solid lines with markers: Simulated BER. Solid lines without markers: BER bound (**??**). Dashed lines: Asymptotic BER (**??**). For comparison the bound and the asymptotic BER for AWGN are also shown.