Adaptive L_p–Norm Diversity Combining in Non–Gaussian Noise and Interference[1]

Amir Nasri, Ali Nezampour, and Robert Schober

The University of British Columbia

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

Abstract

In this paper, we introduce an adaptive L_p–norm metric for robust coherent, differential, and non-coherent diversity combining in non–Gaussian noise and interference. We consider the general case where all diversity branches may use different combining weights and different L_p–norms. We derive a general closed–form expression for the asymptotic bit error rate (BER) for L_p–norm combining in independent non–identically distributed Ricean fading and non–Gaussian noise and interference with finite moments. The asymptotic BER expression reveals that the diversity gain of L_p–norm combining is independent of the type of noise and the metric parameters. In contrast, the combining gain depends on both the type of noise and the metric parameters. Thus, the asymptotic BER can be minimized by optimizing the L_p–norm metric parameters for the underlying type of noise. For this purpose finite difference stochastic approximation (FDSA) and localized random search (LRS) algorithms are developed. Both adaptive algorithms do not require any *a priori* knowledge about the underlying noise and are able to track changes in the noise statistics. Simulation results confirm the validity of the derived asymptotic BER expressions, the effectiveness of the proposed adaptive algorithms, and the excellent performance of the proposed adaptive L_p–norm metric compared to other popular metrics.

[1]This paper has been submitted in part to the IEEE International Communications Conference (ICC), September 2008.

1 Introduction

Diversity combining is an efficient means for combating the detrimental effects of fading in wireless channels. For impairment by additive white Gaussian noise (AWGN) many different combining schemes have been proposed for coherent, differential, and noncoherent detection and their performances have been extensively studied [1]. Generally, these combining schemes are equivalent to the evaluation of a metric that involves an L_2–norm. Examples for such combining schemes include coherent maximal–ratio combining (MRC), differential equal gain combining (EGC), and noncoherent combining (NC).

In practice, wireless communication systems are not only impaired by AWGN but also by various forms of non–Gaussian noise and interference[2] such as man–made and natural impulsive noise [2], co–channel interference (CCI) [3, 4], partial–band interference [5], and ultra–wideband (UWB) interference [6, 7]. Unfortunately, diversity combining schemes optimized for AWGN do not perform well in non–Gaussian noise [7]–[9]. Of course, if the noise distribution is known in parametric form, the distribution parameters can be estimated first, and optimal maximum–likelihood (ML) combining can be applied subsequently, cf. [10] and references therein. However, in many cases, such knowledge is not available at the receiver and noise distribution may even change with time. This motivates the use of *robust* combining schemes and metrics, which perform well for a large class of noise distributions and possibly have a tunable parameter which can be adjusted to the underlying noise distribution. Prominent examples for such robust metrics include Huber's M–metric [11], Myriad and Meridian metrics [12], metrics involving hard and soft limiters [5], and the L_p–norm metric [7, 13]. Thereby, the L_p–norm metric is particularly interesting since it performs well in both noise with heavy–tailed distributions (e.g. impulsive noise) and noise with short–tailed distributions (e.g. CCI) if p is adjusted accordingly [13]. However, finding the optimum p for a particular type of noise is a formidable task, as appropriate optimization criteria are not known.

In this paper, we consider general L_p–norm metrics for coherent, differential, and noncoherent combining, where different diversity branches may use different L_p–norms and different combining weights. We derive analytical expressions for the asymptotic bit error rate (BER) of the considered combining schemes with L_p–norm metric, which are valid for any type of noise with finite moments.

[2] To simplify our notation, in the following, "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".

This analysis is similar in spirit to the asymptotic analysis of L_2–norm metrics for AWGN and non–Gaussian noise in [14, 15] and [16], respectively. The derived asymptotic BER expressions show that the diversity gain is independent of the L_p–norm used and the type of noise. In contrast, the combining gain depends on a generalized moment of the noise samples at the diversity branches, which enables the development of simple metric optimization criteria that directly minimize the asymptotic BER. We consider both off–line and on–line optimization of the metric parameters, and develop for the latter case adaptive multivariate finite difference stochastic approximation (FDSA) [17, 18] and localized random search (LRS) [18] algorithms. We note that the proposed adaptive diversity combining scheme is conceptually different from the asymptotically optimum space–diversity detector in [19]. Whereas the receiver in [19] was optimized for detecting the presence or absence of a weak signal in the *low* SNR regime, the receiver proposed in this paper is optimized for detecting a digitally modulated signal in the *high* SNR regime.

The remainder of this paper is organized as follows. In Section 2, we introduce the system model and the L_p–norm metric. Asymptotic BER expressions are derived in Section 3, and the calculation of generalized noise moments is discussed in Section 4. In Section 5, off–line and on–line optimization of the metric parameters is considered, and in Section 6, analytical and simulation results are presented. Conclusions are drawn in Section 7.

2 System Model and L_p–Norm Metric

In this paper, we consider coherent combining, differential combining, and noncoherent combining for coherent linear modulation formats (e.g. M–ary quadrature amplitude modulation (M–QAM), M–ary phase–shift keying (M–PSK)), differential M–PSK, and binary frequency–shift keying (BFSK), respectively.

2.1 Signal Model

Assuming L diversity branches, for coherent linear modulation and differential M–PSK the received signal in the lth branch and in the kth symbol interval can be modeled in equivalent complex baseband representation as

$$r_l[k] = \sqrt{\bar{\gamma}_l} h_l b[k] + n_l[k], \qquad 1 \leq l \leq L, \tag{1}$$

where $\bar{\gamma}_l$, h_l, and $n_l[k]$ denote the average signal–to–noise ratio (SNR), the fading gain, and the noise in the lth diversity branch, respectively. The transmitted symbols $b[k] \in \mathcal{A}$ are normalized to $\mathcal{E}\{|b[k]|^2\} = 1$ and taken from an M–ary alphabet $\mathcal{A}$. In case of differential M–PSK, $b[k]$ is obtained from $a[k] \in \mathcal{A}$ via differential encoding $b[k] = a[k]b[k-1]$.[3]

The noise is assumed to be independent of the fading gains but the noise samples[4] n_l, $1 \leq l \leq L$, may be statistically dependent and non–Gaussian. The noise variance is given by $\sigma_l^2 \triangleq \mathcal{E}\{|n_l|^2\}$, $1 \leq l \leq L$. The only restriction that we impose on the noise is that all joint moments of the n_l, $1 \leq l \leq L$, exist, i.e., $\mathcal{E}\{n_1^{\kappa_1}(n_1^*)^{\nu_1} n_2^{\kappa_2}(n_2^*)^{\nu_2} \cdots n_L^{\kappa_L}(n_L^*)^{\nu_L}\} < \infty$ for all $\kappa_l \geq 0$, $\nu_l \geq 0$, $1 \leq l \leq L$. Most practically relevant types of noise fulfill this condition (see next section). An exception is α–stable noise for which moments of order greater than α do not exist and which is sometimes used to model impulsive noise [20].

The fading gains h_l are modeled as independent, non–identically distributed (i.n.d.) Gaussian random variables with mean $\bar{h}_l \triangleq \mathcal{E}\{h_l\}$ and variance $\sigma_{h_l}^2 \triangleq \mathcal{E}\{|h_l - \bar{h}_l|^2\}$, i.e., i.n.d. Ricean fading is assumed. Note that for $\bar{\gamma}_l$ in (1) to be the SNR, the power of the fading gains has to be normalized to $\mathcal{E}\{|h_l|^2\} = \sigma_l^2$, $1 \leq l \leq L$. The Ricean factor is defined as $K_l \triangleq |\bar{h}_l|^2/\sigma_{h_l}^2$ and Rayleigh fading results as a special case for $K_l = 0$, $1 \leq l \leq L$.

For BFSK the signal model in (1) has to be augmented since, in this case, in each diversity branch the outputs of two matched filters (MFs) are processed. The first MF output is still given by (1) and the second MF output is modeled as

$$\bar{r}_l = \sqrt{\bar{\gamma}_l} h_l \bar{b} + \bar{n}_l, \qquad 1 \leq l \leq L, \tag{2}$$

where b, $\bar{b} \in \{0, 1\}$, $b \neq \bar{b}$, and $\bar{n}_l$ denotes the noise in the second MF output. While for AWGN n_l and $\bar{n}_l$ are statistically independent, this is not necessarily true for non–Gaussian noise. However, this does not affect the proposed asymptotic performance analysis and metric adaptation.

[3] In this paper, $\mathcal{E}\{\cdot\}$, $[\cdot]^T$, $[\cdot]^*$, $O(\cdot)$, and $I_0(\cdot)$ denote statistical expectation, transposition, complex conjugation, the big O notation, and the zeroth order modified Bessel function of the first kind, respectively. Furthermore, $A \doteq B$ means that A is asymptotically (i.e., for high SNR) equal to B and a function $f(x)$ is $o(x)$ if $\lim_{x \to 0} f(x)/x = 0$.

[4] To simplify our notation, we drop the time index k in variables such as $n_l[k]$ whenever possible.

2.2 Noise Models

In the following, we briefly discuss some important types of noise for which the analysis and metric optimization in this paper is applicable.

1) Gaussian Mixture Noise (GMN): For i.n.d. GMN the probability density function (pdf) of the noise in the lth diversity branch is given by

$$f_n(n_l) = \sum_{i=1}^{I} \frac{c_{i,l}}{\pi\sigma_{n,i,l}^2} \exp\left(-\frac{|n_l|^2}{\sigma_{n,i,l}^2}\right), \quad 1 \le l \le L, \tag{3}$$

where $c_{i,l} > 0$, $\sum_{i=1}^{I} c_{i,l} = 1$, and $\sigma_{n,i,l}^2$, $\sum_{i=1}^{I} c_{i,l}\sigma_{n,i,l}^2 = \sigma_l^2$, are constants. Special cases of GMN include ϵ–mixture noise ($I = 2$, $c_{1,l} = 1 - \epsilon_l$, $c_{2,l} = \epsilon_l$, $\sigma_{n,1,l}^2 = \sigma_l^2/(1 - \epsilon_l + \kappa_l\epsilon_l)$, $\sigma_{n,2,l}^2 = \kappa_l\sigma_{n,1,l}^2$, $0 \le \epsilon_l < 1$, and $\kappa_l > 1$) and Middleton's Class A noise ($I \to \infty$). GMN is a popular model for impulsive noise in systems with receive antenna diversity [9] and for partial band interference in frequency hopping (FH) systems with frequency diversity [5].

2) Co–Channel Interference I (CCI-I): The interference caused by I co–channel interferers in a system with receive antenna diversity can be modeled as [4]

$$n_l[k] = \sum_{i=1}^{I} g_{i,l} \sum_{\kappa=k_1}^{k_2} p_i[\kappa] b_i[k-\kappa], \quad 1 \le l \le L, \tag{4}$$

where $g_{i,l}$, $p_i[k]$, and $b_i[k]$ denote the fading gain at the lth receive antenna, the effective pulse shape, and the transmit symbols of the ith interferer, respectively. $p_i[k]$ depends on the transmit pulse shape of the interferer, the receiver input filter of the user, and the delay τ_i between the ith interferer and the user. The ith co–channel interferer is synchronous and asynchronous for $\tau_i = 0$ and $\tau_i \neq 0$, respectively. The limits k_1 and k_2 are chosen such that $p_i[k] \approx 0$ if $k < k_1$ or $k > k_2$. Here, we model the interference channel gains $g_{i,l}$ as (possibly correlated) Ricean fading gains with variances $\sigma_{g,i,l}^2$ and Ricean factors $K_{g,i,l}$. We note that CCI–I is spatially dependent even if the channel gains $g_{i,l}$ are independent because the term $\sum_{\kappa=k_1}^{k_2} p_i[\kappa]b_i[k-\kappa]$ is common to all diversity branches.

3) CCI-II: The CCI model for FH systems with frequency diversity is slightly different from CCI-I. Assuming the synchronous case and that at hopping frequency l, $1 \le l \le L$, co–channel interferer i, $1 \le i \le I$, is present with probability $\epsilon_{i,l}$, $0 \le \epsilon_{i,l} < 1$, the resulting interference can be modeled as

$$n_l = \sum_{i=1}^{I} X_{i,l} g_{i,l} b_{i,l}, \quad 1 \le l \le L, \tag{5}$$

where the $X_{i,l}$ are mutually independent, and $X_{i,l} = 1$ and $X_{i,l} = 0$ with probabilities $\epsilon_{i,l}$ and $1-\epsilon_{i,l}$, respectively. $b_{i,l}$ denotes the transmit symbols of the ith interferer at the lth hopping frequency and the interference gains $g_{i,l}$ are modeled as i.n.d. Ricean fading with variances $\sigma^2_{g,i,l}$ and Ricean factors $K_{g,i,l}$. CCI–II can be used to model the interference in systems that use FH for multiple access (e.g. Bluetooth).

4) Generalized Gaussian Noise (GGN): I.n.d. GGN is a popular model for non–Gaussian noise [7, 21]. The corresponding pdf for the lth diversity branch is given by

$$f_n(n_l) = \frac{\beta_l \Gamma(4/\beta_l)}{2\pi(\Gamma(2/\beta_l))^2} \exp\left(-\frac{|n_l|^{\beta_l}}{c_l}\right), \quad 1 \le l \le L, \tag{6}$$

where $c_l \triangleq (\Gamma(2/\beta_l)/\Gamma(4/\beta_l))^{\beta_l/2}$, and β_l, $0 < \beta_l < \infty$, denotes the shape parameter. GGN contains Laplacian ($\beta_l = 1$) and Gaussian ($\beta_l = 2$) noise as special cases. We note that the L_p–norm metric with appropriately chosen parameters is the ML metric for i.n.d. GGN [13].

5) UWB Interference: The interference from both multi–band orthogonal frequency division multiplexing (MB–OFDM) UWB and impulse–radio (IR) UWB is in general strongly non–Gaussian [6, 7]. We will test the theory and algorithms developed in this paper for interference caused by the MB–OFDM UWB and IR–UWB signal formats standardized by ECMA [22] and IEEE 802.15.4a [23], respectively.

We note that the proposed analysis is also applicable to any linear combination of the noises specified in 1)–5).

2.3 L_p–Norm Metric

In this subsection, we present the adopted L_p–norm metrics for the considered combining schemes.

Coherent Combining (CC): The L_p–norm metric for CC is given by

$$m_c(\tilde{b}) = \sum_{l=1}^{L} q_l |r_l - \sqrt{\bar{\gamma}_l} h_l \tilde{b}|^{p_l}, \tag{7}$$

where $\tilde{b} \in \mathcal{A}$ is a trial symbol, and $q_l > 0$ and $p_l > 0$, $1 \le l \le L$, are metric parameters that can be optimized for performance maximization for the underlying type of noise.[5] The decision $\hat{b}$ is that $\tilde{b}$ which minimizes $m_c(\tilde{b})$. For $q_l = 1$ and $p_l = 2$, $1 \le l \le L$, the L_p–norm metric $m_c(\tilde{b})$

[5]We note that, strictly speaking, $m_c(\tilde{b})$ is only a norm for $p_l \ge 1$, $1 \le l \le L$. However, whether or not $m_c(\tilde{b})$ is a norm is not important in our context.

is equivalent to MRC which is optimal in AWGN. For convenience we define the parameter vectors $\boldsymbol{q} \triangleq [q_1 \ \ldots \ q_L]^T$ and $\boldsymbol{p} \triangleq [p_1 \ \ldots \ p_L]^T$.

Differential Combining (DC): DC is applied for differential M–PSK modulation and the corresponding L_p–norm metric is

$$m_d(\tilde{a}) = \sum_{l=1}^{L} q_l |r_l[k] - \tilde{a} r_l[k-1]|^{p_l}, \tag{8}$$

where $\tilde{a} \in \mathcal{A}$ is an M–PSK trial symbol. For the special case $q_l = 1$ and $p_l = 2$, $1 \leq l \leq L$, the differential L_p–norm metric $m_d(\tilde{a})$ is equivalent to well–known differential EGC. The decision $\hat{a}$ is that $\tilde{a}$ which minimizes $m_d(\tilde{a})$.

Noncoherent Combining (NC): The considered NC metric for BFSK is

$$m_n = \sum_{l=1}^{L} q_l \left(|r_l|^{p_l} - |\bar{r}_l|^{p_l} \right), \tag{9}$$

where we decide for $b = 1$ if $m_n \geq 0$ and for $b = 0$ otherwise. For $q_l = 1$ and $p_l = 2$, $1 \leq l \leq L$, the L_p–norm metric in (9) is equivalent to conventional square–law combining for BFSK [1].

3 Asymptotic Analysis of L_p–Norm Combining

In this section, we develop asymptotic expressions for the pairwise error probability (PEP) of the combining schemes described in Section 2.3 and relate these PEPs to the respective asymptotic bit error rates (BERs).

3.1 Asymptotic PEP of CC

We show in Appendix A that for any type of noise with finite moments, the asymptotic PEP of CC for $\bar{\gamma}_l \to \infty$, $1 \leq l \leq L$, is given by

$$P_e(d) \doteq \frac{2^L \prod_{l=1}^{L} \left(\Gamma\left(\frac{2}{p_l}\right) \frac{1+K_l}{\sigma_l^2} \exp\left(-K_l\right) \right)}{d^{2L} \prod_{l=1}^{L=1} \left(\bar{\gamma}_l p_l q_l^{2/p_l} \right) \Gamma\left(\sum_{l=1}^{L} \frac{2}{p_l} + 1 \right)} M_{\boldsymbol{n}}(\boldsymbol{q}, \boldsymbol{p}), \tag{10}$$

where $M_{\boldsymbol{n}}(\boldsymbol{q}, \boldsymbol{p}) \triangleq \mathcal{E}\left\{ \left(\sum_{l=1}^{L} q_l |n_l|^{p_l} \right)^{\sum_{l=1}^{L} 2/p_l} \right\}$ can be interpreted as a *generalized moment* of the elements of noise vector $\boldsymbol{n} \triangleq [n_1 \ \ldots, n_L]^T$, and d denotes the Euclidean distance between the

alternative signal points considered for the PEP. The generalized noise moment $M_{\boldsymbol{n}}(\boldsymbol{q},\boldsymbol{p})$ in (10) can be calculated in closed form for special cases, cf. Section 4. Nevertheless, even if the generalized noise moment is not available in closed form, (10) can be used for fast evaluation of the asymptotic PEP since $M_{\boldsymbol{n}}(\boldsymbol{q},\boldsymbol{p})$ is independent of the SNR and has to be evaluated only once, which can be done e.g. by Monte–Carlo simulation. More importantly, (10) reveals how parameters q_l and p_l influence the asymptotic PEP, which will be exploited for metric optimization in Section 5.

For complexity reasons it may be desirable for some applications to limit the number of metric parameters to be optimized. For this purpose we may set $q_l = q$ and $p_l = p$, $1 \leq l \leq L$, and simplify (10) to

$$P_e(d) \doteq \frac{2^L \left(\Gamma\left(\frac{2}{p}\right)\right)^L \prod_{l=1}^{L} \left(\frac{1+K_l}{\sigma_l^2} \exp\left(-K_l\right)\right)}{d^{2L} \prod_{l=1}^{L} (\bar{\gamma}_l)\, p^L \Gamma\left(\frac{2L}{p}+1\right)} M_{\boldsymbol{n}}(p), \tag{11}$$

where $M_{\boldsymbol{n}}(p) \triangleq \mathcal{E}\left\{\left(\sum_{l=1}^{L} |n_l|^p\right)^{2L/p}\right\}$. Note that (11) depends on p but is independent of q. For the special case $p = 2$, (11) is equivalent to [16, Eq. (10)] for independent, identically distributed (i.i.d.) fading.

3.2 Asymptotic PEPs of DC and NC

The asymptotic PEPs of DC and NC can be derived in the same way as those for CC.

1) DC: The asymptotic PEP of M–PSK with DC is also given by (10) and (11) if the respective generalized noise moments of $\boldsymbol{n}$ are replaced with the generalized noise moments of the effective noise vector

$$\boldsymbol{z} = \boldsymbol{n}[k] - a[k]\boldsymbol{n}[k-1]. \tag{12}$$

If the $n_l[k]$ are rotational symmetric and $\boldsymbol{n}[k]$ and $\boldsymbol{n}[k-1]$ are statistically independent, $a[k]$ has no influence on the PEP and we may use $\boldsymbol{z} = \boldsymbol{n}[k] - \boldsymbol{n}[k-1]$ instead of (12).

2) NC: It can be shown that we formally obtain the PEP of BFSK with NC by letting $d = 1$ in (10) and (11), respectively.

3.3 Asymptotic BER

The asymptotic (average) BER can be obtained from the asymptotic PEP as [24]

$$\mathrm{BER} \doteq \frac{\xi_{\min}}{\log_2(M)} P_e(d_{\min}), \tag{13}$$

where $d_{\min}$ and $\xi_{\min}$ denote the minimum Euclidean distance of signal constellation $\mathcal{A}$ and the average number of minimum–distance neighbors, respectively. For example, for all binary modulation schemes $\xi_{\min} = 1$, for binary PSK (BPSK) $d_{\min} = 2$, and for M–QAM $\xi_{\min} = 4(1 - 1/\sqrt{M})$ and $d_{\min} = \sqrt{6/(M-1)}$.

3.4 Combining and Diversity Gain

It is convenient to express the asymptotic BER as $\mathrm{BER} \doteq (G_c \bar{\gamma})^{-G_d}$ [15, 24], where G_c and G_d denote the combining and the diversity gain, respectively, and $\bar{\gamma} = (\prod_{l=1}^{L} \bar{\gamma}_l)^{1/L}$, i.e., $\bar{\gamma}\,[\mathrm{dB}] = \frac{1}{L}\sum_{l=1}^{L} \bar{\gamma}_l\,[\mathrm{dB}]$. From (10) we observe that the diversity gain is given by $G_d = L$ independent of metric parameters $\boldsymbol{q}$ and $\boldsymbol{p}$, and independent of the type of noise. The combining gain for CC with L_p–norm metric can be expressed as

$$\begin{aligned} G_c\,[\mathrm{dB}] \;=\; & 10\,\log_{10}\left(\frac{d_{\min}^2 \log_2(M)^{1/L}}{2\xi_{\min}^{1/L}}\right) - \frac{10}{L}\sum_{l=1}^{L}\log_{10}\left(\frac{1+K_l}{\sigma_l^2}\exp\left(-K_l\right)\right) \\ & + \frac{10}{L}\sum_{l=1}^{L}\log_{10}\left(\frac{p_l q_l^{2/p_l}\left(\Gamma\left(\sum_{i=1}^{L}\frac{2}{p_i}+1\right)\right)^{1/L}}{\Gamma\left(\frac{2}{p_l}\right)}\right) - \frac{10}{L}\log_{10}\left(M_{\boldsymbol{n}}(\boldsymbol{q},\boldsymbol{p})\right). \quad (14) \end{aligned}$$

Eq. (14) reveals that the combining gain consists of four terms. The first and the second term on the right hand side (RHS) of (14) depend on the signal constellation and the fading parameters, respectively, but are independent of the metric parameters $\boldsymbol{q}$ and $\boldsymbol{p}$ and the properties of the noise. The third term on the RHS of (14) is a function of $\boldsymbol{q}$, $\boldsymbol{p}$, and L but is also independent of the noise. Only the last term on the RHS of (14) depends on the properties of the noise via the generalized moment $M_{\boldsymbol{n}}(\boldsymbol{q},\boldsymbol{p})$ of the noise samples. Eq. (14) reveals that the optimal parameters $\boldsymbol{q}_{\mathrm{opt}}$ and $\boldsymbol{p}_{\mathrm{opt}}$, which maximize G_c, only depend on L and the type of noise (via $M_{\boldsymbol{n}}(\boldsymbol{q},\boldsymbol{p})$) but are not influenced by the the modulation scheme and the Ricean factors K_l, $1 \le l \le L$.

For DC and NC similar observations as for CC can be made with respect to diversity gain and combining gain.

4 Generalized Noise Moments

In this section, we provide analytical results for the generalized noise moments defined in Section 3 for selected types of noise. To make the problem tractable, in this section, we consider not

necessarily independent but identically distributed (n.i.d.) noise and $M_{\boldsymbol{n}}(p)$, which depends only on p, instead of $M_{\boldsymbol{n}}(\boldsymbol{q},\boldsymbol{p})$. To simplify our notation, in the following, we will drop subscript l in all noise parameters (e.g. in $c_{i,l}$, ϵ_l, κ_l, $\sigma^2_{g,i,l}$, $K_{g,i,l}$, etc.) if the noise is n.i.d. (which includes i.i.d. as a special case).

4.1 Exact Noise Moments for $L = 2$

First, we consider the special case $L = 2$. Furthermore, for independent complex Gaussian random variables (RVs) x_1 and x_2 having variances $\sigma^2_{x_1}$ and $\sigma^2_{x_2}$ we define $M_G(p;\sigma^2_{x_1},\sigma^2_{x_2}) \triangleq \mathcal{E}\{(|x_1|^p + |x_2|^p)^{2L/p}\}$. Using the substitutions $|x_1| = r\sin^{2/p}\varphi$ and $|x_2| = r\cos^{2/p}\varphi$ with $0 \le r < \infty$ and $0 \le \varphi \le \pi/2$, we obtain

$$M_G(p;\sigma^2_{x_1},\sigma^2_{x_2}) = \frac{48\kappa\sigma^4_{x_1}}{2^{4/p}p}\, I_G(p,\kappa), \tag{15}$$

where $\kappa \triangleq \sigma^2_{x_1}/\sigma^2_{x_2}$, and the finite range integral $I_G(p,\kappa) \triangleq \int_0^{\pi/2}(\sin(2\varphi))^{4/p-1}/(\sin^{4/p}\varphi + \kappa\cos^{4/p}\varphi)^4\,\mathrm{d}\varphi$ depends only on p and κ and can be easily evaluated numerically.

Based on the result for $M_G(p;\sigma^2_{x_1},\sigma^2_{x_2})$, we provide analytical expressions for the generalized moments of AWGN, n.i.d. Rayleigh–faded CCI–I (i.e., $K_{g,i} = 0$, $1 \le i \le I$)[6], and i.i.d. Rayleigh–faded CCI–II (single interferer at each hopping frequency, i.e., $I = 1$) in Table 1. Furthermore, we also provide an expression for i.i.d. GGN in Table 1, which can be obtained in a similar fashion as the generalized moment in the Gaussian case and which contains the finite range integral $I_{GG}(p,\beta) \triangleq \int_0^{\pi/2}(\sin(2\varphi))^{4/p-1}/(\sin^{2\beta/p}\varphi + \cos^{2\beta/p}\varphi)^{8/\beta}\,\mathrm{d}\varphi$.

4.2 Noise Moments for General L

For general L a closed–form expression for the generalized moment can be calculated for several special cases. In particular, we will provide accurate approximations for $M_{\boldsymbol{n}}(p)$ for n.i.d. noise distributions that are based on the Gaussian distribution (i.e., i.i.d. GMN, n.i.d. Rayleigh–faded CCI–I, i.i.d. Rayleigh–faded CCI–II), and exact results for unfaded n.i.d. CCI–I and i.i.d. CCI–II with $I = 1$ and $K_{g,1} \to \infty$, $1 \le l \le L$.

1) Gaussian–based Noise Distributions: We first consider i.n.d. Gaussian RVs x_l with variances $\sigma^2_{x_l}$, $1 \le l \le L$, and our goal is to calculate $M_G(p;\sigma^2_{x_1},\ldots,\sigma^2_{x_L}) \triangleq \mathcal{E}\{(\sum_{l=1}^L |x_l|^p)^{2L/p}\}$. It

[6]We note that the fading gains $g_{l,i}$, $1 \le l \le L$, of n.i.d. CCI–I are i.i.d. RVs. However, the resulting CCI-I is n.i.d. since each interferer affects all receive antennas simultaneously, cf. (4).

can be shown that the pdf of $y_l = |x_l|^p$ is given by

$$f_{y_l}(y_l) = \frac{2}{p\sigma_{x_l}^2} y_l^{2/p-1} \exp\left(-\frac{y_l^{2/p}}{\sigma_{x_l}^2}\right), \tag{16}$$

which is a Weibull pdf. We are interested in the pdf of $z = \sum_{l=1}^{L} y_l$. Unfortunately, a closed–form expression for a sum of Weibull RVs is not known. However, an accurate approximation for the pdf of z is given by the α–μ pdf [25]

$$f_z(z) = \frac{\alpha\mu^\mu z^{\alpha\mu-1}}{\Omega^\mu \Gamma(\mu)} \exp\left(-\frac{\mu z^\alpha}{\Omega}\right), \tag{17}$$

where parameters α, μ, and Ω have to be optimized for the best possible agreement with the true pdf of z. For this purpose, the efficient moment–based method in [25, Eq. (5)–(9)] may be used. We note that in [25] only i.i.d. Weibull variables are considered, whereas we allow different variances $\sigma_{x_l}^2$. This small extension can be accommodated by replacing [25, Eq. (9)] by $\mathcal{E}\{y_l^n\} = \sigma_{x_l}^{pn}\Gamma(1+pn/2)$, $n \in \{0, 1, 2, \ldots\}$ (y_l is referred to as R_l in [25]), and we found the corresponding approximation to be still very accurate. Using (17) we obtain

$$M_G(p; \sigma_{x_1}^2, \ldots, \sigma_{x_L}^2) = \frac{\Gamma(\mu + 2L/(p\alpha))}{\Gamma(\mu)} \left(\frac{\Omega}{\mu}\right)^{2L/(p\alpha)}. \tag{18}$$

Based on the approximation for $M_G(p; \sigma_{x_1}^2, \ldots, \sigma_{x_L}^2)$ in (18), we can find the generalized moments of AWGN, i.i.d. GMN, n.i.d. Rayleigh–faded CCI–I, and i.i.d. Rayleigh–faded CCI–II ($I = 1$) given in Table 2 for general L.

2) Unfaded CCI: We first consider n.i.d. CCI–I. Assuming a single, unfaded interferer ($K_{g,1} \to \infty$), (4) simplifies to

$$n_l[k] = e^{j\Theta_{1,l}} \sum_{\kappa=k_1}^{k_2} p_1[\kappa] b_1[k-\kappa], \quad 1 \le l \le L, \tag{19}$$

with uniformly distributed, mutually independent phases $\Theta_{1,l} \in (-\pi, \pi]$, $1 \le l \le L$. Based on (19), the exact result for the generalized moment of unfaded CCI–I given in Table 2 can be obtained. Similarly, specializing (5) to $I = 1$ and $K_{g,1} \to \infty$, the exact expression for i.i.d. CCI–II in Table 2 can be derived.

The asymptotic PEP for CC and NC can be (approximately or exactly) obtained by combining the generalized moments in Tables 1 and 2 with the PEP formula in (11). Since the effective noise for DC is more complicated than the noise for CC and NC, cf. (12), a closed–form evaluation of the generalized moments does not seem possible for DC in most cases.

5 Metric Optimization

In this section, we optimize the metric parameters $\boldsymbol{p}$ and $\boldsymbol{q}$ for minimization of the asymptotic BER. In the following, we consider both off–line and on–line optimization.

5.1 Off–line Optimization

If the generalized noise moments are known, the metric parameters can be optimized off–line based on (10) or (11). If the underlying type of noise is *a priori* known, the generalized noise moments may obtained in closed–form, cf. Tables 1 and 2, or, if this is not possible, from Monte–Carlo simulation using locally generated noise samples.[7] Monte–Carlo simulation can also be applied to estimate the generalized noise moments from observed noise samples. To gain some insight and to make the problem tractable, we assume n.i.d. noise in this subsection. For n.i.d. noise we may set $q_l = q$ and $p_l = p$, $1 \leq l \leq L$, in metrics (7)–(9) without loss of optimality, i.e., we can base our off–line optimization on (11) and have to optimize only parameter p. Unfortunately, for most types of noise a closed–form optimization of p is not possible. An exception is n.i.d. unfaded CCI–I, where we can show based on (11) and Table 2 that the optimal p is given by $p_{\text{opt}} = \infty$ corresponding to metric $m_c(\tilde{b}) = \max_{l\in\{1,\ldots,L\}}\{|r_l - \sqrt{\overline{\gamma}_l}h_l\tilde{b}|\}$. Furthermore, exploiting (14) we obtain for the asymptotic SNR gain G_p of a metric using $p > 2$ over the L_2–norm metric

$$G_p\,[\text{dB}] = 10\log_{10}\left(\left(\frac{p}{2\Gamma(2/p)}\right)^{1-1/L}\frac{L^{1-2/p}}{L!^{1/L}}\right) \overset{p\to\infty}{=} 10\log_{10}\left(\frac{L}{L!^{1/L}}\right). \tag{20}$$

For example, for $L = 2$ we obtain $G_{20} = 1.3$ dB and $G_\infty = 1.5$ dB. Furthermore, using the Stirling formula [26] for $L \to \infty$ we can show that $G_\infty = 10\log_{10}(e) = 4.3$ dB. We note that it can be shown that $p_{\text{opt}} = \infty$ and (20) are also valid for DC and NC in n.i.d. unfaded CCI–I.

If the optimal p cannot be obtained in closed form, numerical optimization is necessary. To illustrate this, we show in Figs. 1 and 2 the BER of BPSK as a function of p for i.i.d. Rayleigh fading with $L = 2$ and $L = 3$, respectively. Details about the considered types of noise can be found in the captions of the figures. The solid lines represent analytical results generated based on (11), (13),

[7] Note that if the underlying noise model is known *a priori*, ML combining can be applied, of course. However, even in this case the proposed L_p–norm metric may be preferable if the ML metric is computationally complex or causes numerical problems. For example, the GMN pdf consists of a sum of exponential functions which may cause numerical problems for high SNRs.

and, respectively, Table 1 (Fig. 1) and Table 2 (Fig. 2). The markers indicate simulation results and the bold "+" markers denote the minima of the analytical BER. The agreement between analytical results and simulation results is excellent in both Figs. 1 and 2.

As expected, Figs. 1 and 2 show that $p = 2$ is optimal for AWGN and also for Rayleigh–faded CCI–I. In constrast, for heavy–tailed types of noise such as ϵ–mixture noise and Rayleigh–faded CCI–II $p_{\text{opt}} < 2$ holds. For short–tailed noise such as unfaded CCI–I $p_{\text{opt}} > 2$ is valid. For i.i.d. GGN with $\beta = 1$ we obtain $p_{\text{opt}} = 1$ from Fig. 1 as expected. While all other BER curves have a single minimum in the considered p range, the BER for unfaded CCI–II in Fig. 2 has two local minima. Figs. 1 and 2 clearly illustrate the benefits of optimizing p and confirm our analysis.

5.2 On–line Optimization

In practice, the statistical properties of the noise impairing a wireless communication system are often not known *a priori* and may change with time. Since multiplication of the L_p–norm metrics (7)–(9) with a positive constant does not change the decision, we can set $q_1 = 1$ without loss of optimality and optimize only the $2L - 1$ elements of vector $\boldsymbol{x} \triangleq [q_2 \ \ldots \ q_L \ \boldsymbol{p}^T]^T$.

Since the metric coefficients may not be updated in every symbol interval, we introduce a new time $t = N_m k$, where k is the symbol time and $N_m > 1$ can be used to specify how frequently the metric coefficients are updated. Furthermore, the proposed adaptive algorithms require an estimate of the cost function to be minimized. For CC we obtain based on (10) the cost function estimate

$$L_t(\boldsymbol{x}) \triangleq \frac{\prod_{l=1}^{L} \Gamma\left(\frac{2}{p_l}\right)}{\prod_{l=1}^{L=1} \left(p_l q_l^{2/p_l}\right) \Gamma\left(\sum_{l=1}^{L} \frac{2}{p_l} + 1\right)} \hat{M}_{\boldsymbol{n}}(\boldsymbol{x}, t) \tag{21}$$

$$\hat{M}_{\boldsymbol{n}}(\boldsymbol{x}, t) \triangleq \frac{1}{N_e} \sum_{\nu=0}^{N_e-1} \left(\sum_{l=1}^{L} q_l |\hat{n}_l[t-\nu]|^{p_l} \right)^{\sum_{l=1}^{L} 2/p_l}, \tag{22}$$

where we have neglected all irrelevant terms and N_e denotes the number of time steps used for estimation of the generalized moment $\hat{M}_{\boldsymbol{n}}(\boldsymbol{x}, t)$ at time t. Furthermore, $\hat{n}_l[t] \triangleq r_l - \sqrt{\overline{\gamma}_l} h_l b[t]$, where $b[t]$ may be a training symbol or a previous decision. A similar estimate for the cost function may be generated for DC and NC. In the following, two different algorithms for optimization of $\boldsymbol{x}$ are provided.

1) Multivariate Stochastic Approximation: The first algorithm is based on the finite difference stochastic approximation (FDSA) framework with gradient estimation [18]. This framework is

particularly well suited for the problem at hand since it employs a Kiefer–Wolfowitz type of gradient estimate $\hat{\boldsymbol{g}}_t(\boldsymbol{x}_t)$ avoiding cumbersome differentiation of $L_t(\boldsymbol{x})$ [17]. In the tth iteration the FDSA algorithm generates the estimate $\boldsymbol{x}_t$ for the optimal $\boldsymbol{x}$ as [18]

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - a_t\hat{\boldsymbol{g}}_t(\boldsymbol{x}_t), \tag{23}$$

$$\hat{\boldsymbol{g}}_t(\boldsymbol{x}_t) = \left[\frac{L_t(\boldsymbol{x}_t + c_t\boldsymbol{e}_1) - L_t(\boldsymbol{x}_t - c_t\boldsymbol{e}_1)}{2c_t} \cdots \frac{L_t(\boldsymbol{x}_t + c_t\boldsymbol{e}_{2L-1}) - L_t(\boldsymbol{x}_t - c_t\boldsymbol{e}_{2L-1})}{2c_t}\right]^T, \tag{24}$$

where $\boldsymbol{e}_n$ is a column vector of length $2L-1$ with a 1 in position n and 0's in all other positions. If $\boldsymbol{n}[k]$ is stationary and a_t and c_t fulfill $a_t > 0$, $c_t > 0$, $a_t \to 0$, $c_t \to 0$, $\sum_{t=0}^{\infty} a_t = \infty$, and $\sum_{t=0}^{\infty} a_t^2/c_t^2 < \infty$, the above algorithm will find the global minimum if the BER has only one minimum and at least a local minimum otherwise [18] (as long as the BER and $L_t(\boldsymbol{x})$ meet the mild conditions required for convergence outlined in [18]). However, since, in practice, $\boldsymbol{n}[k]$ will be non–stationary, we may set $a_t = a$ and $c_t = c$, $\forall t$, where a and c are small positive constants to give the algorithm some tracking capability. Furthermore, since the p_l may have a large dynamic range (e.g. $p_{\text{opt}} = \infty$ for unfaded CCI–I), the tracking ability of the algorithm can be improved by limiting the elements of $\boldsymbol{x}_t$ to some finite value $x_{\max}$ at the expense of some loss in performance if the optimal element of $\boldsymbol{x}$ exceeds $x_{\max}$. Note that for the problem at hand the FDSA algorithm may not find the global optimum as the cost function may have multiple local minimia, cf. Fig. 2. However, we did not find this to be a problem in practice as the BERs of most types of noise seem to have only a single minimum, and in case of multiple minima, all minima seem to result in similar performances. For initialization of the FDSA algorithm $q_l = 1$ and $p_l = 2$, $1 \leq l \leq L$, is a good choice since this guarantees that the solution found by the algorithm in combination with CC, DC, and NC will not perform worse than conventional MRC, EGC, and NC, respectively

2) Random Search Method: The second method that we consider is a localized random search (LRS) method. In contrast to FDSA algorithms, LRS algorithms do not get stuck in local minima and find the global minimum under mild conditions on the cost function [18]. Based on $\boldsymbol{x}_t$, the proposed LRS algorithm generates a new estimate [18]

$$\hat{\boldsymbol{x}}_{t+1} = \boldsymbol{x}_t + \boldsymbol{d}_t, \tag{25}$$

where $\boldsymbol{d}_t$ is a random vector whose elements are i.i.d. Gaussian random variables with variance σ_d^2. If at least one of the elements of $\hat{\boldsymbol{x}}_{t+1}$ lies outside the predefined search space $[0, x_{\max}]$, (25) is repeated until all elements of $\hat{\boldsymbol{x}}_{t+1}$ are inside the search space. Subsequently, if $L(\hat{\boldsymbol{x}}_{t+1}) < L(\boldsymbol{x}_t)$, we

let $\boldsymbol{x}_{t+1} = \hat{\boldsymbol{x}}_{t+1}$, otherwise $\boldsymbol{x}_{t+1} = \boldsymbol{x}_t$. In a non–stationary noise environment, the above algorithm is run continuously. The speed of convergence of the LRS algorithm depends crucially on the size of the search space (i.e., $x_{\max}$) and on σ_d^2 [18]. For initialization, the same initial vector as for the FDSA is appropriate. Since LRS algorithms suffer from performance degradation if the cost function estimate is noisy [18], comparatively large N_e may be advisable. Note, however, that the tracking capabilities of the algorithm decrease as N_e increases. We found $N_e = 100$ to give a good compromise between estimation noise suppression and tracking capabilities for the application at hand.

3) Complexity: From a practical point of view it is of interest to compare the complexity of the proposed adaptive algorithms assuming a fixed–point implementation with s digits. Denoting the complexity of one multiplication by $\mu(s)$, the complexity of evaluating the Gamma and the power functions is $O((\log(s))^2\mu(s))$ [27], whereas that of a division is $O(\mu(s))$. Taking this into account, neglecting the complexity of additions, and assuming that $2/p_l$, $1 \leq l \leq L$ is obtained form a look–up table, the complexities of one iteration of the FDSA and LRS algorithms are given by

$$C_{\mathrm{FDSA}} = O\left((2L-1)(N_e(L+1)+2L+1)(\log(s))^2\mu(s)\right) \tag{26}$$

and

$$C_{\mathrm{LRS}} = O\left((N_e(L+1)+2L+1)(\log(s))^2\mu(s)\right), \tag{27}$$

respectively. A comparison of (26) and (27) shows that, since typically N_e has to be chosen much larger for the LRS algorithm (e.g. $N_e = 100$) than for the FDSA algorithm (e.g. $N_e = 1$), the complexity per iteration of the LRS algorithm is larger than that of the FDSA algorithm for typical values of L (e.g. $L < 10$).

4) Performance: In Figs. 3 and 4, we show metric coefficients q_l, $2 \leq l \leq 4$, and p_l, $1 \leq l \leq 4$, vs. iteration t of, respectively, the FDSA and the LRS algorithms for i.i.d. Rayleigh fading with $L = 4$ and SNR = 16 dB. The corresponding BERs of BPSK with CC are shown in Fig. 5. Five different types of noise are considered which are specified in the caption of Fig. 3 and at $t = (\nu - 1) \cdot 10^6$ we switch abruptly to a new noise Nν, $1 \leq \nu \leq 5$. For the FDSA algorithm we used $a_t = a = 4 \cdot 10^{-4}$, $c_t = c = 10^{-5}$, $x_{\max} = 10$, $N_m = 1$, and $N_e = 1$. For the LRS algorithm we adopted $\sigma_d^2 = 0.1$, $x_{\max} = 10$, $N_m = 1$, and $N_e = 100$. For both algorithms $\boldsymbol{x}_t$ was initialized with $q_l = 1$, $2 \leq l \leq 4$, and $p_l = 2$, $1 \leq l \leq 4$, and previous decisions $\hat{b}[t]$ were used in the adaptation process. Figs. 3 and 4 show the results for one typical adaptation process and the corresponding BERs in Fig. 5 were

calculated with (10) and (13), where the generalized noise moments were obtained by Monte–Carlo simulation. Figs. 3–5 show that both algorithms work well and that after each switching to a new type of noise, steady state operation is achieved quickly. Thereby, with the chosen settings, the steady state error of the LRS algorithm is larger than that of the FDSA algorithm, but the LRS algorithm converges faster to the new steady state after the type of noise has changed. Note, however, that the trade–off between tracking capabilities and residual error strongly depends on how the parameters of the algorithms (e.g. a, c, N_e, and σ_d^2) are chosen [18]. Furthermore, as expected, Figs. 3 and 4 confirm that in steady state for the n.i.d. noises N1 and N5 all q_l and p_l are equal, respectively, whereas for the i.n.d. noises N2, N3, and N4 either the q_l or/and the p_l are not equal. For N5 $p_l = \infty$, $1 \leq l \leq 4$, is optimal and both algorithms yield $p_l = 10$, $1 \leq l \leq 4$, because we set $x_{\max} = 10$. Fig. 5 shows that the L_p–norm metric with FDSA and LRS adaptation substantially outperforms the L_2–norm metric (i.e., MRC).

6 Numerical Results and Discussions

In this section, we verify the analytical results derived in Sections 3 and 4 through simulations and compare the performance of the adaptive L_p–norm metric with that of other popular metrics. For convenience we consider n.i.d. noise throughout this section and drop subscript l in the noise parameters. The respective noise parameters are specified in the captions of the figures. The optimal metric parameter p_{opt} was obtained with the FDSA algorithm.

In Figs. 6 and 7, we show the BER of 16–QAM with CC in i.i.d. Rayleigh fading with $L = 2$ for the adaptive L_p–norm metric and several other popular robust metrics for, respectively, i.i.d. ϵ–mixture noise and n.i.d. unfaded QPSK CCI–I. To facilitate the definition of the various metrics, we introduce the notation $u_l \triangleq |r_l - \sqrt{\overline{\gamma}_l} h_l \tilde{b}|$. We consider the Huber metric $m(\tilde{b}) = \sum_{l=1}^{L} m_l(\tilde{b})$, $m_l(\tilde{b}) = u_l^2/2$ if $u_l \leq \delta$, and $m_l(\tilde{b}) = \delta u_l - \delta^2/2$ if $u_l > \delta$ [11], the Meridian metric $m(\tilde{b}) = \sum_{l=1}^{L} \log(u_l + \delta)$ [12], and the Myriad metric $m(\tilde{b}) = \sum_{l=1}^{L} \log(u_l^2 + \delta^2)$ [12]. Note that for all these robust metrics parameter δ has to be optimized by exhaustive search, which is quite tedious, since, unlike for the L_p–norm metric, a systematic optimization framework is not available. For Figs. 6 and 7 the robust metrics were optimized by simulation for SNR $=$ 30 dB. Fig. 6 shows that for the heavy–tailed ϵ–mixture noise the L_p–norm metric with $p_{\text{opt}} = 0.4$ outperforms the other robust metrics and the gap to the optimal ML metric is less than 1 dB. Fig. 7 shows that for short–tailed unfaded CCI–I the

Huber and Myriad metrics are essentially equivalent to the L_2–norm metric and are outperformed by 1.3 dB by the L_p–norm metric with $p = 20$ ($p_{\text{opt}} \to \infty$ holds in this case), as predicted in Section 5.1. We note that the ML metric does not seem tractable for unfaded CCI–I. Interestingly, while the L_p–norm metric was optimized based on the presented asymptotic analysis, Figs. 6 and 7 suggest that it also performs well for low SNRs.

In Fig. 8, we show the BER of BFSK with NC in i.i.d. Ricean fading with $L = 3$ and i.i.d. Rayleigh–faded QPSK CCI–II. Fig. 8 shows that the proposed L_p–norm combining also achieves considerable performance gains over L_2–norm combining for BFSK with NC and in Ricean fading. As expected from Section 3.4, the optimal value $p_{\text{opt}} = 0.5$ is independent of the Ricean factor K. Although for $K = 6$ dB the simulated BER approaches the asymptotic BER only for BER $< 10^{-10}$, the L_p–norm metric optimized for the asymptotic BER also results in large gains for higher BERs. For example, for BER $= 10^{-4}$, the L_p–norm metric achieves a gain of 3.5 dB over the L_2–norm metric.

Finally, in Fig. 9, we show the BER of 4–PSK in i.i.d. Rayleigh fading with $L = 3$ and impairment by MB–OFDM UWB and IR–UWB interference following the ECMA [22] and IEEE 802.15.4a [23] standards, respectively. The bandwidth of the receiver input filter of the 4–PSK system is assumed to be $B = 4$ MHz. Results for both CC and DC are shown in Fig. 9. For both combining schemes and both types of UWB, $p = 30$ was close to optimal for the L_p–norm metric. Fig. 9 shows that L_p–norm combining also achieves substantial gains over L_2–norm combining in UWB interference. Thereby, the performance gains are larger for CC than for DC. This can be explained by the fact that the effective noise for DC is the sum of two independent noise samples, cf. (12), and thus, according to the Central Limit Theorem [28], is closer to a Gaussian distribution than the noise relevant for CC (and NC).

7 Conclusions

In this paper, we have considered general L_p–norm coherent, differential, and noncoherent diversity combining in non–Gaussian noise and interference. For the asymptotic regime of high SNR we have derived closed–form expressions for the BER valid for i.n.d. Ricean fading and non–Gaussian noise and interference with finite moments. The asymptotic BER expressions reveal that while the diversity gain of L_p–norm combining is independent of the type of noise and the metric parameters, the combining gain depends on generalized noise moments and on the metric parameters. For on–

line metric optimization, we have developed two efficient adaptive algorithms which do not require any *a priori* knowledge about the noise statistics and can also cope with non–stationary noise. Simulation results have confirmed the analytical results presented in this paper and have shown that the proposed adaptive L_p–norm metric outperforms other robust metrics such as Huber's metric, the Myriad metric, and the Meridian metric in both heavy–tailed and short–tailed noise.

A Asymptotic PEP for CC

Assuming that b was transmitted and $\hat{b} \neq b$ was detected, the corresponding PEP can be expressed as

$$P_e(d) = \Pr\{m_c(b) > m_c(\hat{b})\}, \tag{28}$$

where $d \triangleq |e|$ and $e \triangleq b - \hat{b}$. In a first step, we calculate the PEP conditioned on the noise vector $\boldsymbol{n} \triangleq [n_1 \ \ldots \ n_L]^T$. With (7) and (28) this conditional PEP can be obtained as

$$P_e(d|\boldsymbol{n}) = \int_0^{m_c(b)} f_c(z)\,\mathrm{d}z, \tag{29}$$

where we have used the fact that due to the conditioning on $\boldsymbol{n}$, $m_c(b) = \sum_{l=1}^{L} q_l |n_l|^{p_l}$ is a constant, and $f_c(z)$ is the pdf of $m_c(\hat{b}) = \sum_{l=1}^{L} q_l |\sqrt{\bar{\gamma}_l} h_l e + n_l|^{p_l}$, which we calculate step–by–step in the following.

The conditional pdf of $x_l = |\sqrt{\bar{\gamma}_l} h_l e + n_l|$ is a Ricean pdf given by

$$f_{x_l}(x_l) = \frac{2x_l}{d^2 \bar{\gamma}_l \sigma_{h_l}^2} \exp\left(-\frac{x_l^2 + |\sqrt{\bar{\gamma}_l}\bar{h}_l e + n_l|^2}{d^2 \bar{\gamma}_l \sigma_{h_l}^2}\right) I_0\left(2\frac{x_l |\sqrt{\bar{\gamma}_l}\bar{h}_l e + n_l|}{d^2 \bar{\gamma}_l \sigma_{h_l}^2}\right). \tag{30}$$

The pdf of the transformed variable $y_l = x_l^{p_l}$ is given by $f_{y_l}(y_l) = \frac{1}{p_l} y_l^{1/p_l - 1} f_{x_l}\left(y_l^{1/p_l}\right)$ and the scaling with q_l leads to $z_l = q_l y_l$ with pdf $f_{z_l}(z_l) = \frac{1}{q_l} f_{y_l}(z_l/q_l)$. Taking into account these identities the pdf of $z_l = q_l |\sqrt{\bar{\gamma}_l} h_l e + n_l|^{p_l}$ is given by

$$f_{z_l}(z_l) = \frac{2 z_l^{2/p_l - 1}}{d^2 \bar{\gamma}_l \sigma_{h_l}^2 p_l q_l^{2/p_l}} \exp\left(-\frac{z_l^{2/p_l} + q_l^{2/p_l} |\sqrt{\bar{\gamma}_l}\bar{h}_l e + n_l|^2}{d^2 \bar{\gamma}_l \sigma_{h_l}^2 q_l^{2/p_l}}\right) I_0\left(2\frac{z_l^{1/p_l} |\sqrt{\bar{\gamma}_l}\bar{h}_l e + n_l|}{d^2 \bar{\gamma}_l \sigma_{h_l}^2 q_l^{1/p_l}}\right). \tag{31}$$

Considering the asymptotic case $\bar{\gamma}_l \to \infty$ and exploiting the Taylor series expansions of $\exp(\cdot)$ and $I_0(\cdot)$, $f_{z_l}(z_l)$ can be written as

$$f_{z_l}(z_l) = \frac{C_l}{\bar{\gamma}_l} z_l^{2/p_l - 1} + o(\bar{\gamma}_l^{-1}), \tag{32}$$

where $C_l \triangleq 2\exp\left(-|\bar{h}_l|^2/\sigma_{h_l}^2\right)/(d^2\sigma_{h_l}^2 p_l q_l^{2/p_l})$. Thus, the moment generating function (MGF) of z_l can be expanded as $\Phi_{z_l}(s) \triangleq \mathcal{E}\{e^{-sz_l}\} = C_l\Gamma(2/p_l)\bar{\gamma}_l^{-1}s^{-2/p_l} + o(\bar{\gamma}_l^{-1})$. Since conditioned on $\boldsymbol{n}$ the z_l are statistically independent, the MGF of $m_c(\hat{b})$ is given by $\Phi_c(s) = \prod_{l=1}^{L}\Phi_{z_l}(s)$, and the asymptotic expansion of the corresponding pdf is given by

$$f_c(z) = \frac{\prod_{l=1}^{L}\left(C_l\Gamma\left(\frac{2}{p_l}\right)\right)}{\Gamma\left(\sum_{l=1}^{L}\frac{2}{p_l}\right)\prod_{l=1}^{L}\bar{\gamma}_l} z^{\sum_{l=1}^{L}\frac{2}{p_l}-1} + o\left(\prod_{l=1}^{L}\bar{\gamma}_l^{-1}\right). \tag{33}$$

Using this result in (29) leads to

$$P_e(d|\boldsymbol{n}) = \frac{\prod_{l=1}^{L}\left(C_l\Gamma\left(\frac{2}{p_l}\right)\right)}{\Gamma\left(\sum_{l=1}^{L}\frac{2}{p_l}+1\right)\prod_{l=1}^{L}\bar{\gamma}_l}\left(\sum_{l=1}^{L}q_l|n_l|^{p_l}\right)^{\sum_{l=1}^{L}\frac{2}{p_l}} + o\left(\prod_{l=1}^{L}\bar{\gamma}_l^{-1}\right). \tag{34}$$

If all joint moments of the elements of $\boldsymbol{n}$ are finite, averaging $P_e(d|\boldsymbol{n})$ in (34) with respect to $\boldsymbol{n}$ yields (10). The assumption of finite joint noise moments is necessary, since the terms absorbed into $o(\prod_{l=1}^{L}\bar{\gamma}_l^{-1})$ in (34) involve sums of products of the elements of $\boldsymbol{n}$ which have to remain finite after expectation.

References

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

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

[3] C. Tellambura. Cochannel Interference Computation for Arbitrary Nakagami Fading. *IEEE Trans. Veh. Technol.*, 48:487–489, March 1999.

[4] A. Giorgetti and M. Chiani. Influence of Fading on the Gaussian Approximation for BPSK and QPSK with Asynchronous Cochannel Interference. *IEEE Trans. on Wireless Commun.*, 4:384–389, March 2005.

[5] C. Keller and M. Pursley. Clipped Diversity Combining for Channels with Partial-Band Interference - Part I: Clipped-Linear Combining. *IEEE Trans. Commun.*, 35:1320–1328, December 1987.

[6] A. Nasri, R. Schober, and L. Lampe. Analysis of Narrowband Communication Systems Impaired by MB–OFDM UWB Interference. *IEEE Trans. on Wireless Commun.*, 6:4090–4100, November 2007.

[7] H. Shao and N. Beaulieu. Analysis of a Novel P-Order Metric UWB Receiver Structure with Improved Performance in Multiple Access Interference. In *Proc. IEEE Global Telecommun. Conf. (Globecom)*, pages 4112–4117, November 2007.

[8] A. Shah and A. Haimovich. Performance Analysis of Maximal Ratio Combining and Comparison with Optimum Combining for Mobile Radio Communications with Cochannel Interference. *IEEE Trans. Veh. Technol.*, COM-49:1454–1463, July 2000.

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

[10] R. Blum, R. Kozick, and B. Sadler. An Adaptive Spatial Diversity Receiver for Non–Gaussian Interference and Noise. *IEEE Trans. Signal Processing*, 47:2100–2111, August 1999.

[11] P. Huber. *Robust Statistics*. Wiley, New York, 1981.

[12] T. Aysal and K. Barner. Meridian Filtering for Robust Signal Processing. *IEEE Trans. Signal Processing*, 55:3949–3962, August 2007.

[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. Abdel-Ghaffar and S. Pasupathy. Asymptotic Performance of M-ary and Binary Signals Over Multipath/Multichannel Rayleigh and Ricean Fading. *IEEE Trans. Commun.*, 43:2721–2731, November 1995.

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

[16] A. Nezampour, A. Nasri, R. Schober, and Y. Ma. Asymptotic BEP and SEP of Quadratic Diversity Combining Receivers in Correlated Ricean Fading, Non–Gaussian Noise, and Interference. *To appear in IEEE Trans. Commun. Available online:http://www.ece.ubc.ca/˜alinezam/TCOM-07.pdf*, 2008.

[17] J. Kiefer and J. Wolfowitz. Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, 23:462–466, September 1952.

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

[19] L. Izzo and L. Paura. Asymptotically Optimum Space-Diversity Detection in Non–Gaussian Noise. *IEEE Trans. Commun.*, 34:97–103, February 1986.

[20] 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.

[21] T. Pham and J. De Figueiredo. Maximum Likelihood Estimation of a Class of Non-Gaussian Densities with Application to l_p Deconvolution. *IEEE Trans. Signal Processing*, 37:73–82, January 1989.

[22] 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.

[23] 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.

[24] J.G. Proakis. *Digital Communications*. McGraw–Hill, New York, forth edition, 2001.

[25] J. Filho and M. Yacoub. Simple Precise Approximations to Weibull Sums. *IEEE Commun. Letters*, 10:614–616, August 2006.

[26] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. Dover Publications, Inc., New York, 1970.

[27] J. Borwein and P. Borwein. *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*. John Wiley & Sons, Inc., New York, 1987.

[28] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes*. McGraw–Hill, New York, 1984.

Figures and Tables:

Table 1: Generalized noise moments $M_{\boldsymbol{n}}(p)$ for $L = 2$ for various types of n.i.d. noise. In particular, we consider AWGN, i.i.d. GMN, n.i.d. Rayleigh–faded CCI–I ($\boldsymbol{s} \triangleq [s_1 \ldots s_I]^T$, $s_i \triangleq \sum_{\kappa=k_1}^{k_2} p_i[\kappa] b_i[\kappa]$, $\mathcal{S}$ contains all possible values of $\boldsymbol{s}$), i.i.d. Rayleigh–faded CCI–II ($I = 1$, $\boldsymbol{b}_I \triangleq [b_{1,1} \ldots b_{1,L}]^T$, $\mathcal{M}_I$ contains all possible values of $\boldsymbol{b}_I$, $c_1 \triangleq \epsilon_1$, $c_2 \triangleq 1 - \epsilon_1$, $\bar{\sigma}^2_{g,1} \triangleq \sigma^2_{g,1}$, $\bar{\sigma}^2_{g,2} \triangleq 0$), and i.i.d. GGN.

Noise Model	Moments $M_{\boldsymbol{n}}(p)$
AWGN	$M_G(p; 1, 1)$
GMN	$\sum_{i=1}^{I} \sum_{j=1}^{I} c_i c_j M_G(p, \sigma^2_{n,i}, \sigma^2_{n,j})$
CCI–I (Rayleigh)	$\frac{1}{\lvert\mathcal{S}\rvert} \sum_{\boldsymbol{s}\in\mathcal{S}} M_G(p, \sum_{i=1}^{I} \sigma^2_{g,i} \lvert s_i\rvert^2, \sum_{i=1}^{I} \sigma^2_{g,i} \lvert s_i\rvert^2)$
CCI–II (Rayleigh)	$\frac{1}{\lvert\mathcal{M}_I\rvert} \sum_{i=1}^{2} \sum_{j=1}^{2} c_i c_j \sum_{\boldsymbol{b}_I\in\mathcal{M}_I} M_G(p, \bar{\sigma}^2_{g,i} \lvert b_{1,1}\rvert^2, \bar{\sigma}^2_{g,j} \lvert b_{1,2}\rvert^2)$
GGN	$\frac{\Gamma(8/\beta)\beta}{(\Gamma(4/\beta))^2 2^{4/p-2} p} I_{GG}(p, \beta)$

Table 2: Approximations for the generalized noise moments $M_{\boldsymbol{n}}(p)$ for general L for the same types of n.i.d. noise considered in Table 1. Additionally, exact results for unfaded n.i.d. CCI–I ($I = 1$) and i.i.d. CCI–II ($I = 1$, $\xi_1 \triangleq 1$, $\xi_2 \triangleq 0$) are provided.

Noise Model	Moments $M_{\boldsymbol{n}}(p)$
AWGN	$M_G(p; 1, \ldots, 1)$
GMN	$\sum_{i_1=1}^{I} \cdots \sum_{i_L=1}^{I} c_{i_1} \cdots c_{i_L} \, M_G(p, \sigma^2_{n,i_1}, \ldots, \sigma^2_{n,i_L})$
CCI–I (Rayleigh)	$\frac{1}{\lvert\mathcal{S}\rvert} \sum_{\boldsymbol{s}\in\mathcal{S}} M_G(p, \sum_{i=1}^{I} \sigma^2_{g,i} \lvert s_i\rvert^2, \ldots, \sum_{i=1}^{I} \sigma^2_{g,i} \lvert s_i\rvert^2)$
CCI–II (Rayleigh)	$\frac{1}{\lvert\mathcal{M}_I\rvert} \sum_{i_1=1}^{2} \cdots \sum_{i_L=1}^{2} c_{i_1} \cdots c_{i_L} \sum_{\boldsymbol{b}_I\in\mathcal{M}_I} M_G(p, \bar{\sigma}^2_{g,i_1} \lvert b_{1,1}\rvert^2, \ldots, \bar{\sigma}^2_{g,i_L} \lvert b_{1,L}\rvert^2)$
CCI–I (Unfaded)	$L^{2L/p} \frac{1}{\lvert\mathcal{S}\rvert} \sum_{\boldsymbol{s}\in\mathcal{S}} \lvert s\rvert^{2L}$
CCI–II (Unfaded)	$\frac{1}{\lvert\mathcal{M}_I\rvert} \sum_{i_1=1}^{2} \cdots \sum_{i_L=1}^{2} c_{i_1} \cdots c_{i_L} \sum_{\boldsymbol{b}_I\in\mathcal{M}_I} \left(\sum_{l=1}^{L} \xi_{i_l} \lvert b_{1,l}\rvert^p \right)^{2L/p}$

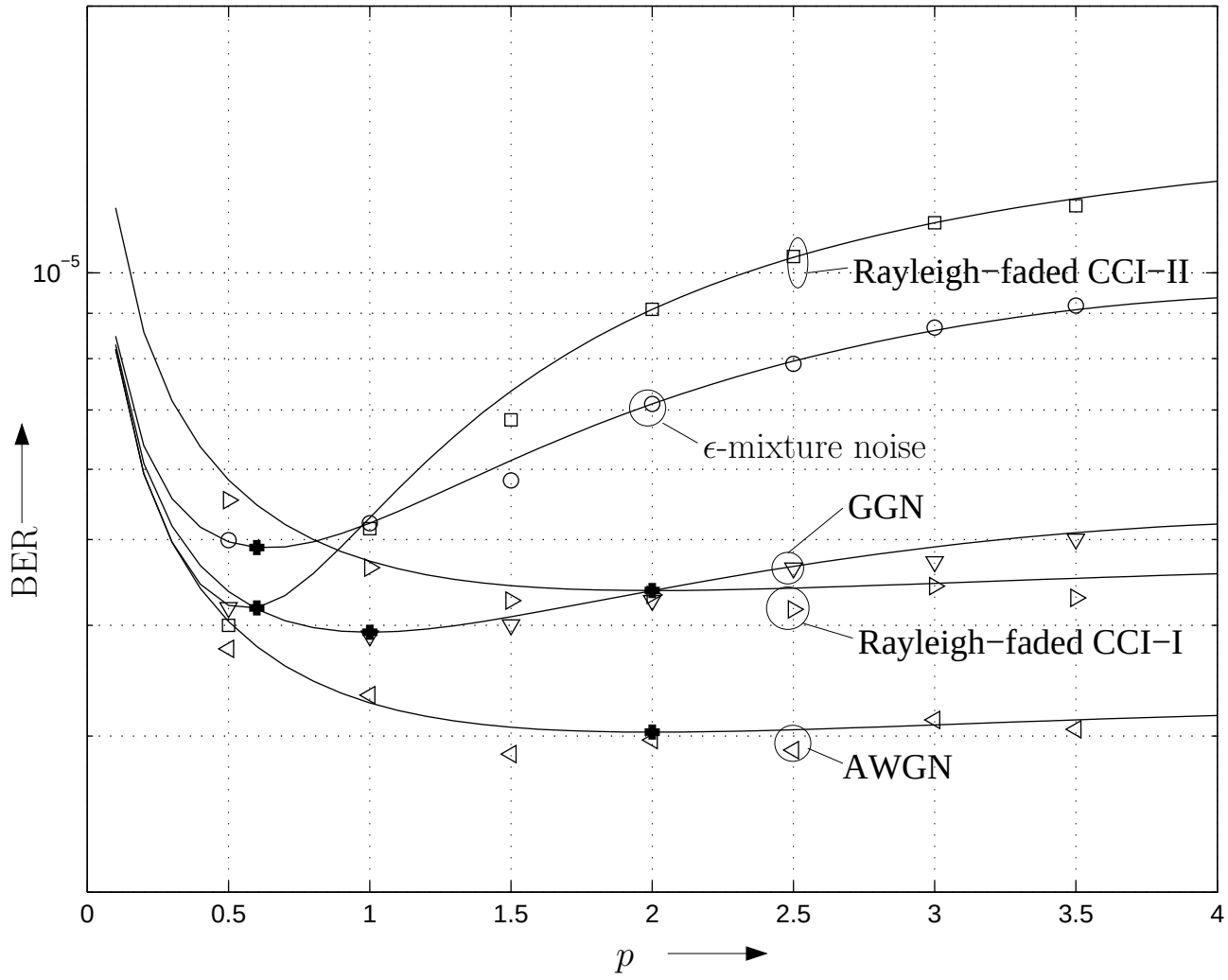

Figure 1: BER vs. p for BPSK, i.i.d. Rayleigh fading, $L = 2$, SNR = 24 dB, and different types of n.i.d. noise. Noise parameters: I.i.d. ϵ–mixture noise ($\epsilon = 0.1$, $\kappa = 10$), n.i.d. Rayleigh–faded QPSK CCI–I ($I = 1$, $\tau_1 = 0.25T$ with symbol duration T, raised cosine pulse shape with roll–off factor 0.22), i.i.d. Rayleigh–faded QPSK CCI–II ($I = 1$, $\epsilon_1 = 0.25$), and i.i.d. GGN ($\beta = 1$).

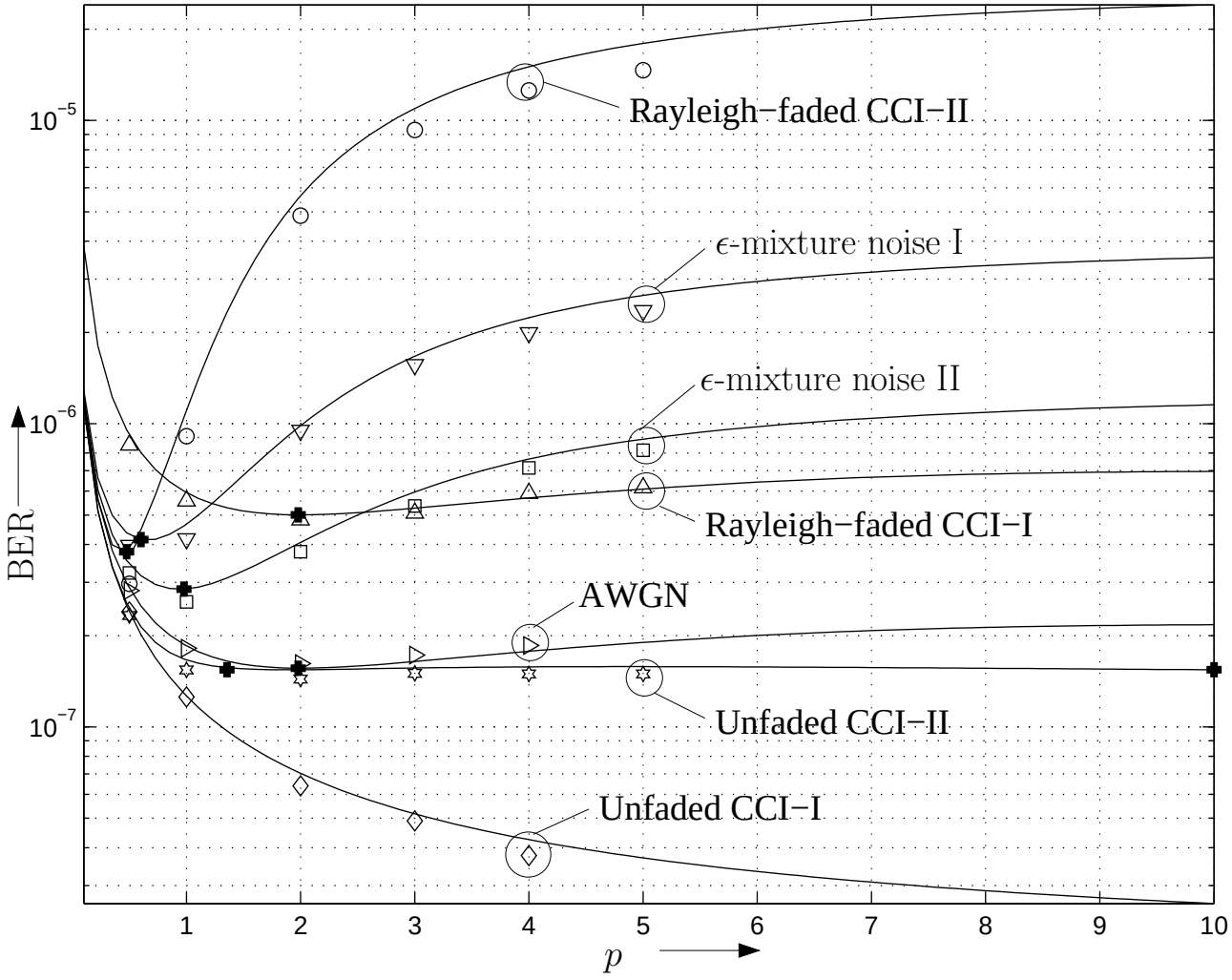

Figure 2: BER vs. p for BPSK, i.i.d. Rayleigh fading, $L = 3$, SNR = 20 dB, and different types of n.i.d. noise. Noise parameters: I.i.d. ϵ–mixture noise I ($\epsilon = 0.1$, $\kappa = 10$), i.i.d. ϵ–mixture noise II ($\epsilon = 0.1$, $\kappa = 5$), n.i.d. Rayleigh–faded and unfaded QPSK CCI–I ($I = 1$, $\tau_1 = 0.25T$, raised cosine pulse shape with roll–off factor 0.22), and i.i.d. Rayleigh–faded and unfaded QPSK CCI–II ($I = 1$, $\epsilon_1 = 0.41$).

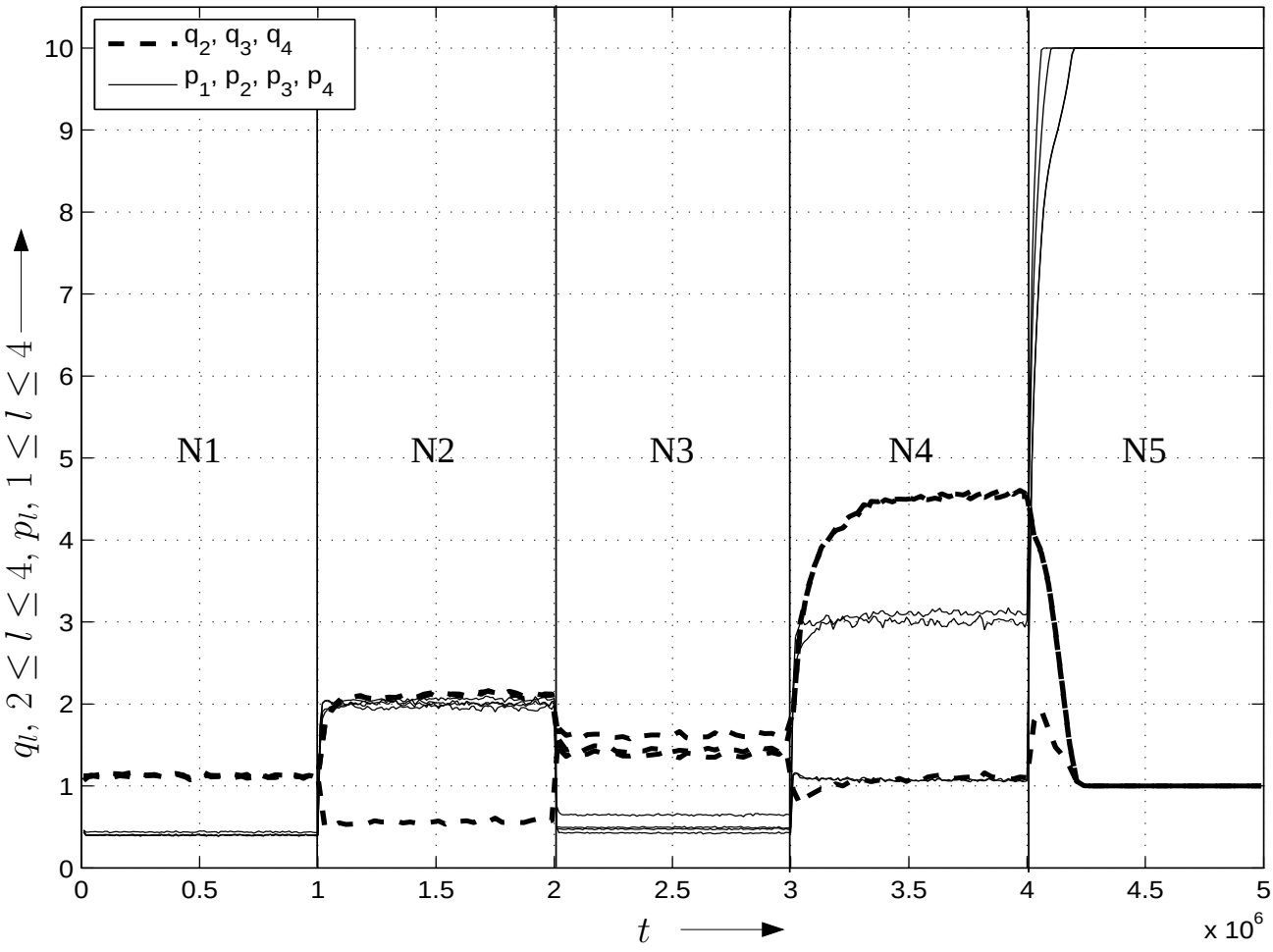

Figure 3: Metric coefficients q_l, $2 \leq l \leq 4$, and p_l, $1 \leq l \leq 4$, vs. iteration t of FDSA algorithm. N1: I.i.d. Rayleigh–faded QPSK CCI–II ($I = 1$, $\epsilon = 0.1$) and AWGN, where the CCI–II power is ten times larger than the AWGN variance; N2: I.n.d. Gaussian noise with variances $\sigma_1^2 = 1$, $\sigma_2^2 = 0.5$, $\sigma_3^2 = 0.5$, $\sigma_4^2 = 2$; N3: I.n.d. ϵ–mixture noise with $\epsilon_l = 0.1$, $1 \leq l \leq 4$, and $\kappa_1 = 20$, $\kappa_2 = 40$, $\kappa_3 = 50$, $\kappa_4 = 100$; N4: I.n.d. GGN with $\beta_1 = \beta_2 = 3$ and $\beta_3 = \beta_4 = 1$; N5: N.i.d. unfaded QPSK CCI–I ($I = 1$, $\tau_1 = 0.3T$, raised cosine pulse shape with roll–off factor 0.22).

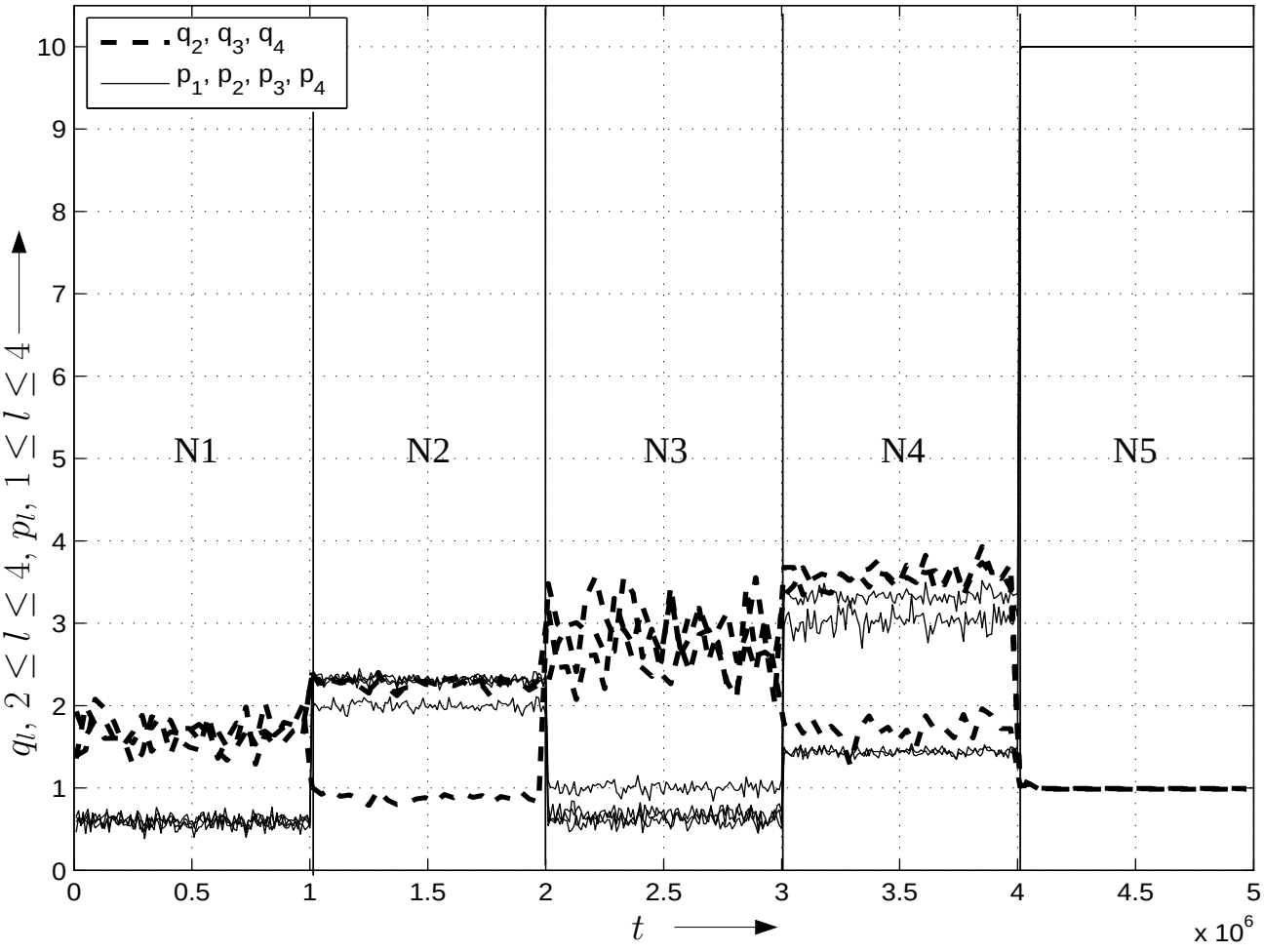

Figure 4: Metric coefficients q_l, $2 \leq l \leq 4$, and p_l, $1 \leq l \leq 4$, vs. iteration t of LRS algorithm. Noise types N1–N5 are specified in the caption of Fig. 3.

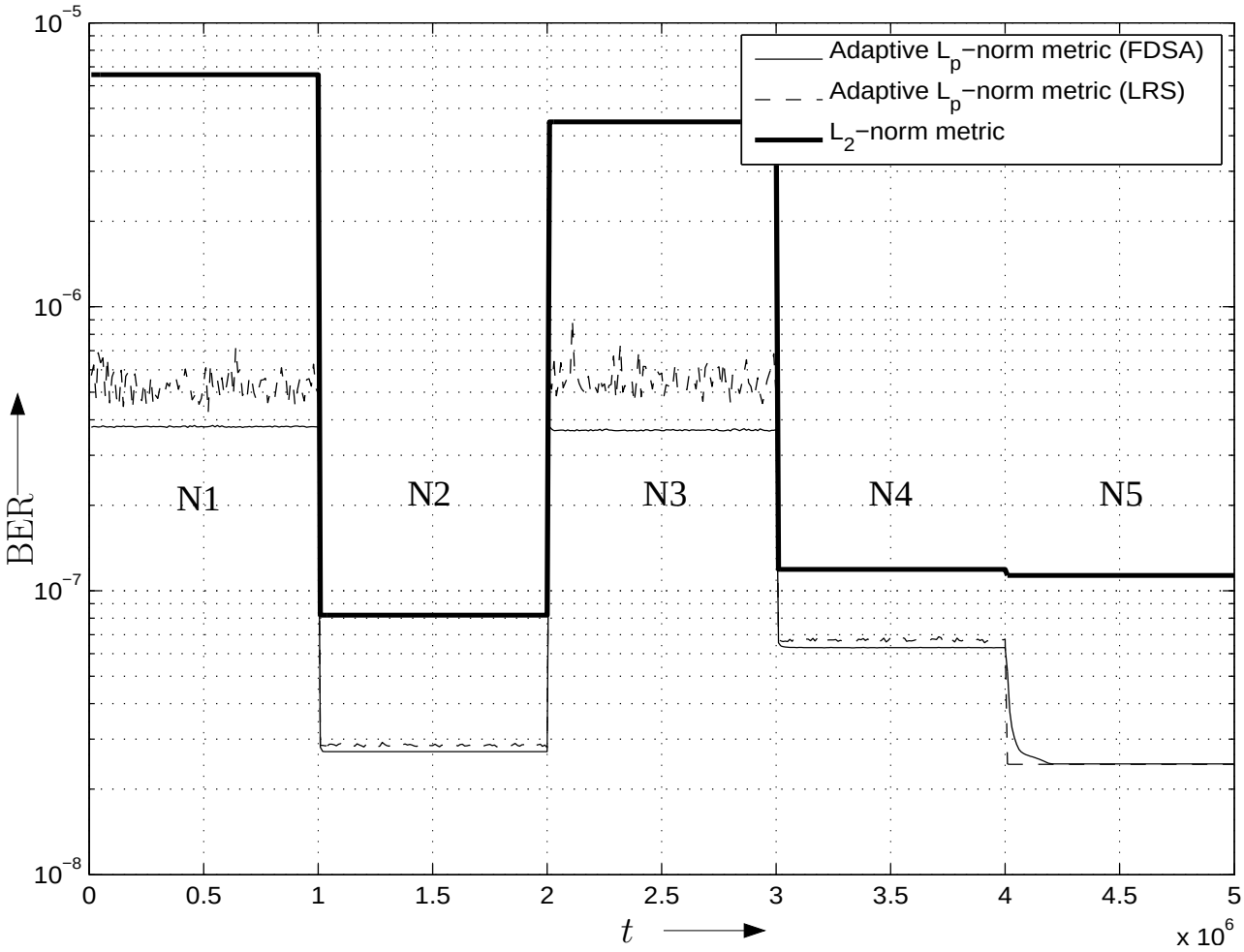

Figure 5: BER of BPSK with CC vs. iteration t for FDSA and LRS algorithms, respectively. For comparison BER of L_2–norm combining is also shown. Noise types N1–N5 are specified in the caption of Fig. 3.

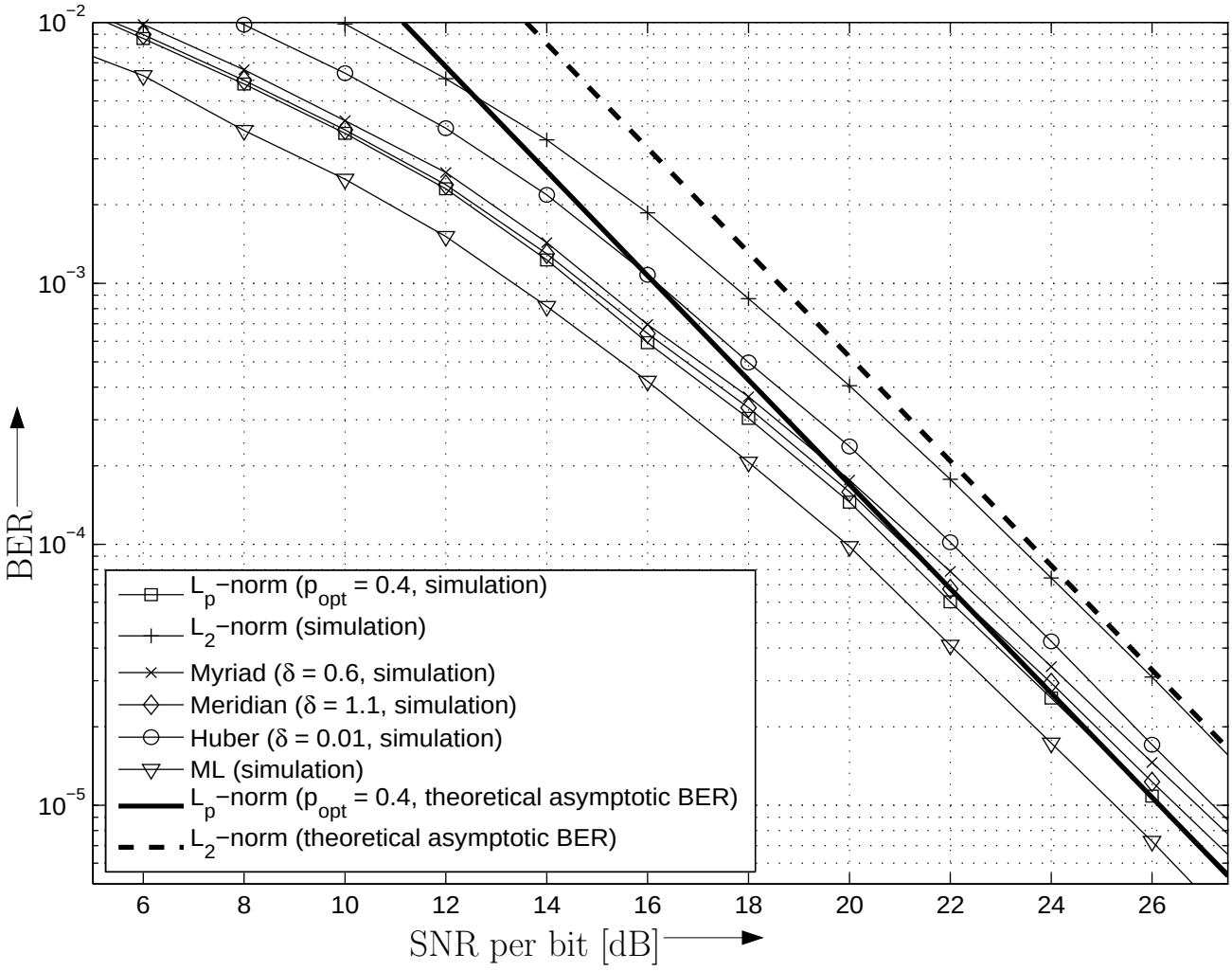

Figure 6: BER vs. SNR per bit per branch of 16–QAM with CC in i.i.d. Rayleigh fading ($L = 2$) and i.i.d. ϵ–mixture noise ($\epsilon = 0.1$, $\kappa = 100$). Solid lines with markers: Simulation results. Bold solid and dashed lines: Asymptotic BER based on (11), (13), and Table 1.

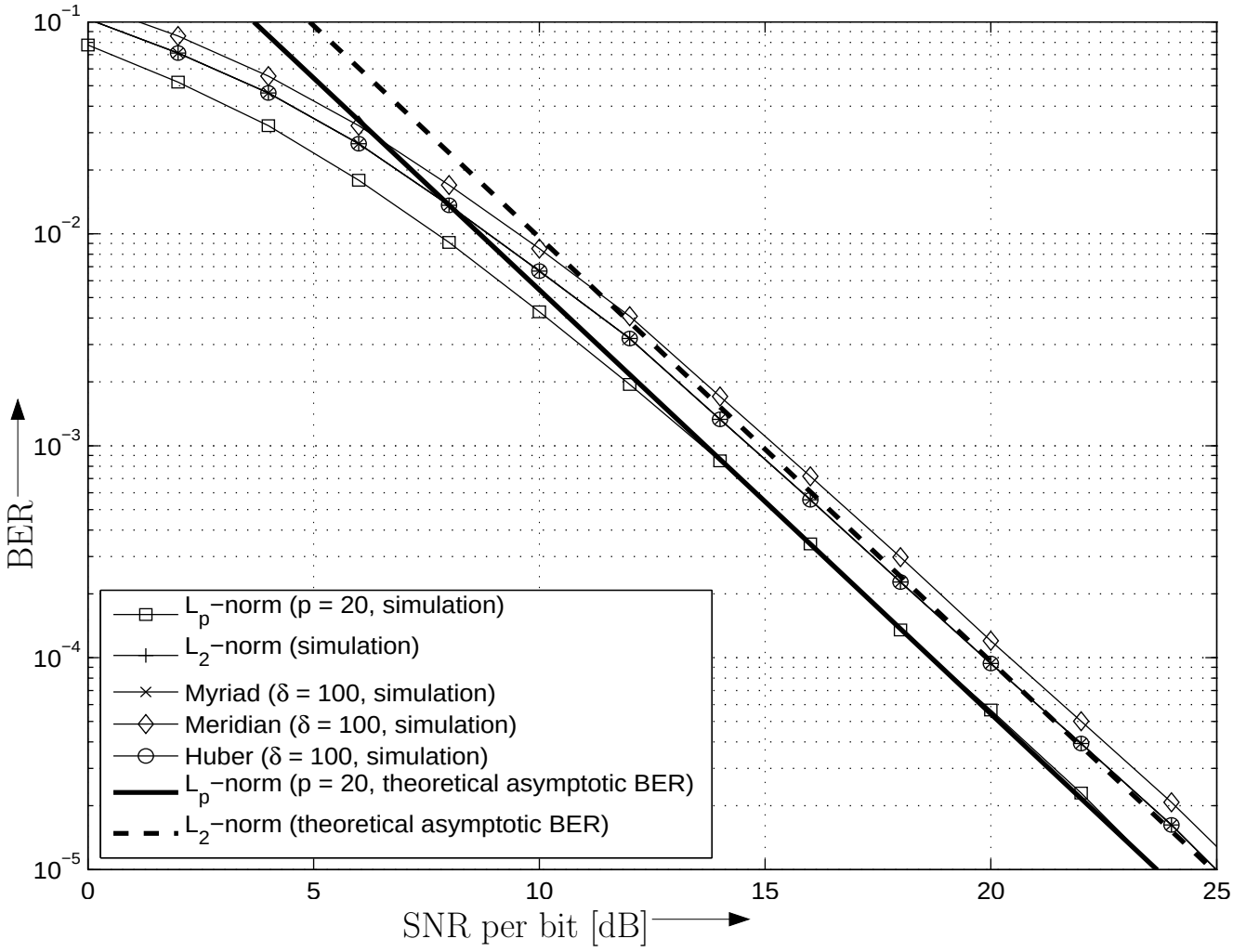

Figure 7: BER vs. SNR per bit per branch of 16–QAM in i.i.d. Rayleigh fading ($L = 2$) and n.i.d. unfaded QPSK CCI–I ($I = 1$, $\tau_1 = 0.3T$, raised cosine pulse shape with roll–off factor 0.22). Solid lines with markers: Simulation results. Bold solid and dashed lines: Asymptotic BER based on (11), (13), and Table 2.

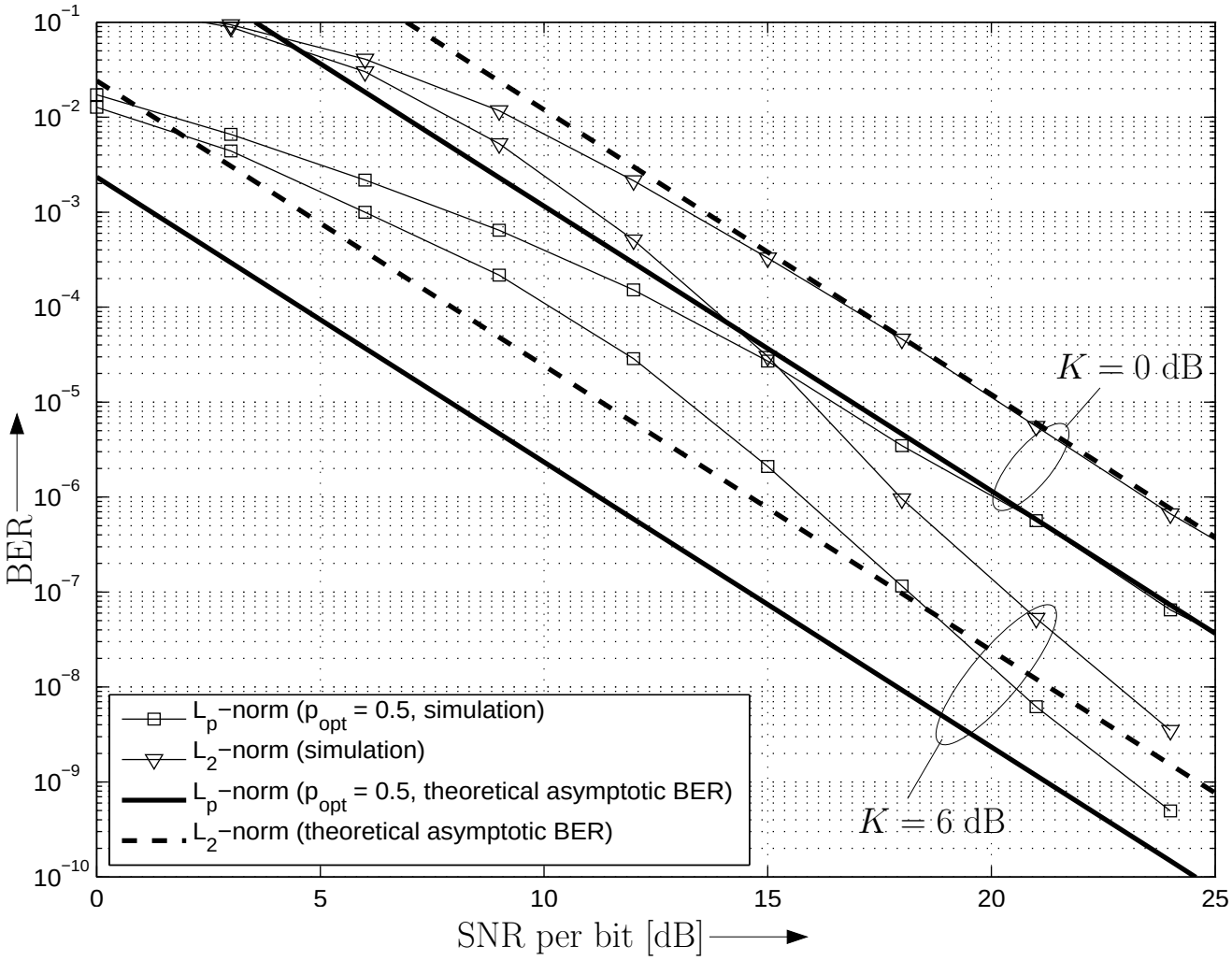

Figure 8: BER vs. SNR per branch of BFSK with NC in i.i.d. Ricean fading ($L = 3$) and i.i.d. Rayleigh–faded QPSK CCI–II ($I = 1$, $\epsilon_1 = 0.25$). Solid lines with markers: Simulation results. Bold solid and dashed lines: Asymptotic BER based on (11), (13), and Table 2.

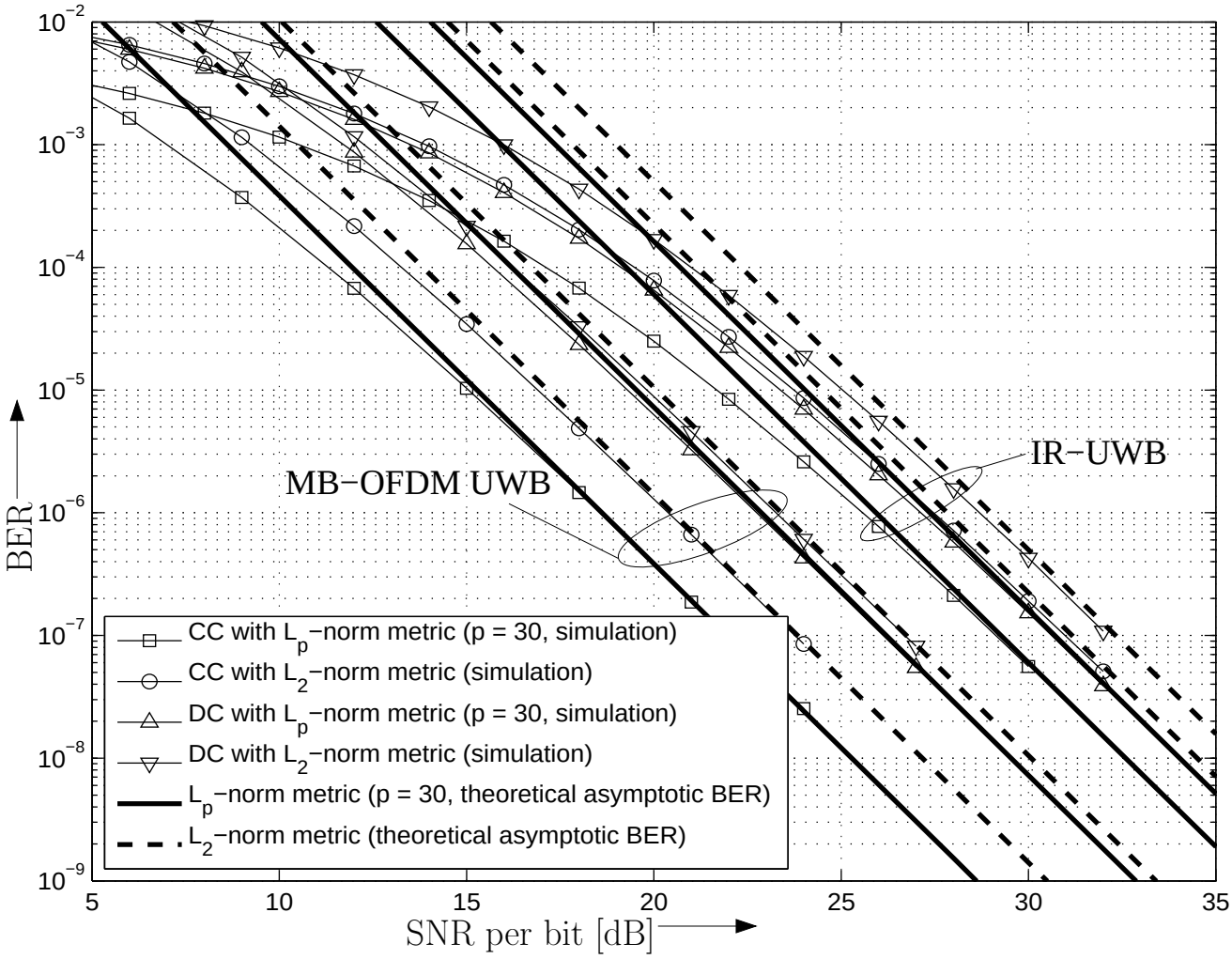

Figure 9: BER vs. SNR per bit per branch of 4–PSK system with bandwidth $B = 4$ MHz and CC or DC in i.i.d. Rayleigh fading ($L = 3$) and MB–OFDM UWB [22] and IR–UWB ($N_b = 32$ bursts per symbol and $L_c = 128$ chips per burst) [23] interference. Solid lines with markers: Simulation results. Bold solid and dashed lines: Asymptotic BER based on (11), (13), and Monte–Carlo simulation of generalized noise moments.