Adaptive Coherent L_p–Norm Combining

Amir Nasri, Ali Nezampour, and Robert Schober
University of British Columbia, {amirn, alinezam, rschober}@ece.ubc.ca

***Abstract*— In this paper, we introduce an adaptive L_p–norm metric for robust coherent diversity combining in non–Gaussian noise and interference. 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. Based on this asymptotic BER expression, the metric parameters can be adapted to the underlying type of noise and interference using a finite difference stochastic approximation (FDSA) algorithm. Simulation results confirm the validity of the derived asymptotic BER expression and the excellent performance of the proposed adaptive L_p–norm metric.**

I. 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) it is well known that maximal ratio combining (MRC) is optimal [1].

In practice, wireless communication systems are not only impaired by AWGN but also by various forms of non–Gaussian noise and interference[1] such as man–made and impulsive noise [2], co–channel interference (CCI) [1], partial–band interference [3], and ultra–wideband (UWB) interference [4]. Unfortunately, diversity combining schemes optimized for AWGN do not perform well in non–Gaussian noise [2]–[5]. Of course, if the distribution of the noise is *a priori* known, the optimum combining scheme can be derived based on the maximum–likelihood (ML) concept. However, in many cases, the noise distribution is not known at the receiver and may even change with time. This motivates the use of *robust* combining schemes, 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 robust metrics include Huber's M–metric [6], Myriad and Meridian metrics [7], and the L_p–norm metric [8], [4]. 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 parameter p is adjusted accordingly [8]. 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 coherent L_p–norm combining, where different diversity branches may use different L_p–norms and different combining weights. We derive an analytical expression for the asymptotic bit error rate (BER) for L_p–norm combining, which is valid for Ricean fading and any type of noise with finite moments. This analysis is similar in spirit to the asymptotic analysis of MRC for AWGN and non–Gaussian noise in [9] and [10], respectively. However, the tools developed in [9], [10] cannot be applied in the more general L_p–norm case. The derived analytical BER expression enables the optimization of the metric parameters. Since closed–form expressions for the optimal metric parameters cannot be obtained in general and the type of noise is not known in practice, we develop an efficient adaptive algorithm for on–line metric optimization.

The remainder of this paper is organized as follows. In Section II, we introduce the system model and the L_p–norm metric. The asymptotic BER expression is derived in Section III, and metric optimization is discussed in Section IV. In Section V, analytical and simulation results are presented, and conclusions are drawn in Section VI.

II. System Model

In this section, we present the considered signal and noise models and the L_p–norm metric.

A. Signal Model

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

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

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 powers of both fading gain and noise are normalized to $\sigma_l^2 \triangleq \mathcal{E}\{|h_l|^2\} = \mathcal{E}\{|n_l[k]|^2\}$, $1 \le l \le L$.[2] Furthermore, the transmitted symbols $b[k] \in \mathcal{A}$ are normalized to $\mathcal{E}\{|b[k]|^2\} = 1$ and taken from an M–ary alphabet $\mathcal{A}$ such as M–ary quadrature amplitude modulation (M–QAM) and M–ary phase–shift keying (M–PSK).

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. 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 $K_l = 0$, $1 \le l \le L$.

The noise is assumed to be independent of the fading gains but the noise samples[3] n_l, $1 \le l \le L$, may be statistically dependent and non–Gaussian. The only restriction that we impose on the noise is that all joint moments of the n_l, $1 \le l \le L$, 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 $\kappa_l \ge 0$, $\nu_l \ge 0$, $1 \le l \le L$. Most practically relevant types of noise fulfill this condition. An exception is α–stable noise for which moments of order greater than α do not exist and

[1]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".

[2]In this paper, $\mathcal{E}\{\cdot\}$, $[\cdot]^T$, $[\cdot]^*$, and $I_0(\cdot)$ denote statistical expectation, transposition, complex conjugation, 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 $\lim_{x\to 0} f(x)/x = 0$.

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

which is sometimes used to model impulsive noise. However, other models for impulsive noise such as Middleton's Class A model [2] are included in our considerations.

For diversity combining we adopt the L_p–norm metric

$$m(\tilde{b}) = \sum_{l=1}^{L} q_l |r_l - \sqrt{\bar{\gamma}_l} h_l \tilde{b}|^{p_l}, \quad (2)$$

where $\tilde{b} \in \mathcal{A}$ is a trial symbol, and $q_l > 0$ and $p_l > 0$, $1 \leq l \leq L$, are metric parameters that can be optimized for performance maximization for the underlying type of noise[4]. The decision $\hat{b}$ is that $\tilde{b}$ which minimizes $m(\tilde{b})$. For $q_l = 1$ and $p_l = 2$, $1 \leq l \leq L$, the L_p–norm metric $m(\tilde{b})$ is equivalent to MRC which is optimal in AWGN. For convenience we define $\boldsymbol{q} \triangleq [q_1 \, \ldots \, q_L]^T$ and $\boldsymbol{p} \triangleq [p_1 \, \ldots \, p_L]^T$.

B. Noise Models

For future reference and to demonstrate the versatility of the proposed approach, 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

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

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 ($I = 2$, $c_{1,l} = 1 - \epsilon_l$, $c_2 = \epsilon_l$, $\sigma_{n,1,l}^2 = \sigma_l^2/(1 - \epsilon_l + \kappa_l \epsilon_l)$, $\sigma_{n,2,l}^2 = \kappa_l \sigma_{n,1,l}^2$, $0 \leq \epsilon_l < 1$, 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 [5] and for partial band interference in frequency hopping (FH) systems with frequency diversity [3].

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

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

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 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.

[4]We note that strictly speaking the metric defined in (2) is only a norm if $p_l \geq 1$, $1 \leq l \leq L$. However, whether or not $m(\tilde{b})$ is a norm is not important in our context.

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 each hopping frequency co–channel interferer i, $1 \leq i \leq I$, is present at the lth hopping frequency with probability $\epsilon_{i,l}$, $0 \leq \epsilon_{i,l} < 1$, the resulting interference can be modeled as

$$n_l = \sum_{i=1}^{I} X_{i,l} g_{i,l} b_{i,l}, \quad 1 \leq l \leq L, \quad (5)$$

where the $X_{i,l}$ are mutually independent, and $X_{i,l} = 1$ and $X_{i,l} = 0$ with probabilities $\epsilon_{i,l}$ and $1 - \epsilon_{i,l}$, $1 \leq l \leq L$, $1 \leq i \leq I$, respectively. $b_{i,l}$ denotes the transmit symbols of the ith interferer at the lth hopping frequency and the Ricean fading interference gains $g_{i,l}$ are i.n.d. with Ricean factors and variances as defined in the CCI–I case. CCI–II can be used to model the interference in systems that use FH for multiple access (e.g. Bluetooth) and different users are assigned random, not necessarily orthogonal hopping patterns.

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

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

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 noise contains Laplacian ($\beta_l = 1$) and Gaussian ($\beta_l = 2$) noise as special cases. We note that the L_p–norm metric with properly chosen metric parameters $\boldsymbol{q}$ and $\boldsymbol{p}$ is the ML metric for GGN [8].

The proposed analysis is also applicable to any linear combination of the noises specified in 1)–4).

III. Asymptotic Analysis of L_p–Norm Combining

In this section, we develop an asymptotic expression for the pairwise error probability (PEP) of coherent L_p–norm combining and relate this PEP to the asymptotic BER.

A. Asymptotic PEP

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

$$P_e(d) \doteq \frac{2^L \prod_{l=1}^{L} \left(\Gamma\left(\frac{2}{p_l}\right) \frac{1+K_l}{\sigma_l^2} \exp\left(-K_l\right)\right)}{d^{2L} \prod_{l=1}^{L=1} \left(\bar{\gamma}_l p_l q_l^{2/p_l}\right) \Gamma\left(\sum_{l=1}^{L} \frac{2}{p_l} + 1\right)} M_{\boldsymbol{n}}(\boldsymbol{q}, \boldsymbol{p}), \quad (7)$$

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 (7) can be calculated in closed form for special cases, cf. Section III-C. Nevertheless, even if the generalized noise moment is not available in closed form, (7) 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, (7) reveals how parameters q_l and p_l influence the asymptotic PEP, which will be exploited for metric optimization in Section IV.

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 $p_l = p$, $1 \le l \le L$, and simplify (7) to

$$P_e(d) \doteq \frac{2^L\left(\Gamma\left(\frac{2}{p}\right)\right)^L \prod_{l=1}^L\left(\frac{1+K_l}{\sigma_l^2}\exp\left(-K_l\right)\right)}{d^{2L}\prod_{l=1}^L(\bar\gamma_l)\, p^L\Gamma\left(\frac{2L}{p}+1\right)} M_{\boldsymbol{n}}(p), \tag{8}$$

where $M_{\boldsymbol{n}}(p) \triangleq \mathcal{E}\left\{\left(\sum_{l=1}^L |n_l|^p\right)^{2L/p}\right\}$. Note that the PEP in (8) depends on p but is independent of q.

B. Asymptotic BER

The asymptotic BER can be obtained from the asymptotic PEP as [9]

$$\mathrm{BER} \doteq \frac{\xi_{\min}}{\log_2(M)} P_e(d_{\min}), \tag{9}$$

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 binary PSK (BPSK) $\xi_{\min} = 1$, $d_{\min} = 2$ and for M–QAM $\xi_{\min} = 4(1 - 1/\sqrt{M})$, $d_{\min} = \sqrt{6/(M-1)}$.

It is often convenient to express the asymptotic BER as $\mathrm{BER} \doteq (G_c\bar\gamma)^{-G_d}$ [9], 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 (7) 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. In contrast, (7) shows that the combining gain G_c does depend on the type of noise and on $\boldsymbol{p}$ and $\boldsymbol{q}$.

C. Generalized Noise Moments

For evaluation of the PEPs in (7) and (8) the generalized noise moments have to be calculated, which is possible in closed form for special cases. In particular, to make the problem tractable, in this section, we consider not necessarily independent but identically distributed (n.i.d.) noise and $M_{\boldsymbol{n}}(p)$ instead of $M_{\boldsymbol{n}}(\boldsymbol{q}, \boldsymbol{p})$. To simplify our notation, for n.i.d. noise (which includes i.i.d. noise as a special case), we drop subscript l in all noise parameters (e.g. in $c_{i,l}$, ϵ_l, κ_l, $K_{g,i,l}$, etc.) in the following.

In the following, 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., independent, identically distributed (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) Gaussian–based Noise Distributions: We first consider i.n.d. Gaussian RVs x_l with variances $\sigma_{x_l}^2$, $1 \le l \le L$, and our goal is to calculate $M_G(p; \sigma_{x_1}^2, \ldots, \sigma_{x_L}^2) \triangleq \mathcal{E}\{(\sum_{l=1}^L |x_l|^p)^{2L/p}\}$. It can be shown that the pdf of $y_l = |x_l|^p$ is given by

$$f_{y_l}(y_l) = \frac{2}{p\sigma_{x_l}^2} y_l^{2/p-1}\exp\left(-\frac{y_l^{2/p}}{\sigma_{x_l}^2}\right), \tag{10}$$

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 [11]

$$f_z(z) = \frac{\alpha\mu^\mu z^{\alpha\mu-1}}{\Omega^\mu\Gamma(\mu)}\exp\left(-\frac{\mu z^\alpha}{\Omega}\right), \tag{11}$$

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 [11, Eq. (5)–(9)] may be used. We note that in [11] 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 [11, 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 [11]), and we found the corresponding approximation to be still very accurate. Using (11) we obtain

$$M_G(p; \sigma_{x_1}^2, \ldots, \sigma_{x_L}^2) = \frac{\Gamma(\mu + 2L/(p\alpha))}{\Gamma(\mu)}\left(\frac{\Omega}{\mu}\right)^{2L/(p\alpha)}. \tag{12}$$

Based on the approximation for $M_G(p; \sigma_{x_1}^2, \ldots, \sigma_{x_L}^2)$ in (12), we can find the generalized moments for AWGN, GMN, Rayleigh–faded CCI–I, and Rayleigh–faded CCI–II ($I = 1$) given in Table I.

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

$$n_l[k] = e^{j\Theta_{1,l}}\sum_{\kappa=0}^{P_1-1} p_1[\kappa]b_1[k-\kappa], \quad 1 \le l \le L, \tag{13}$$

with uniformly distributed, mutually independent phases $\Theta_{1,l} \in (-\pi, \pi]$, $1 \le l \le L$. Based on (13), the exact result for the generalized moment of unfaded CCI–I given in Table I can be easily obtained. Similarly, specializing (5) to $I = 1$ and $K_{g,1} \to \infty$, the exact expression for i.i.d. CCI–II in Table I can be derived.

TABLE I

GENERALIZED NOISE MOMENTS $M_{\boldsymbol{n}}(p)$ FOR AWGN, I.I.D. GMN, RAYLEIGH–FADED AND UNFADED (UNF.) N.I.D. CCI–I AND I.I.D. CCI–II ($I = 1$). THE FOLLOWING DEFINITIONS ARE USED. CCI–I: $\mathcal{S}$ IS THE SET OF ALL POSSIBLE VALUES OF $\sum_{\kappa=k_1}^{k_2} p_1[\kappa]b_1[\kappa]$; CCI–II: $c_1 \triangleq 1-\epsilon$, $c_2 \triangleq \epsilon$, $\xi_1 \triangleq 1$, $\xi_2 \triangleq 0$, $\boldsymbol{b}_I \triangleq [b_{1,1}, \ldots, b_{1,L}]^T$, $\mathcal{M}_I$ CONTAINS ALL POSSIBLE VALUES OF $\boldsymbol{b}_I$.

Noise Model	Moments $M_{\boldsymbol{n}}(p)$
AWGN	$M_G(p; 1, \ldots, 1)$
GMN	$\sum_{i_1=1}^{I}\cdots\sum_{i_L=1}^{I} c_{i_1}\cdots c_{i_L} M_G(p, \sigma_{n,i_1}^2, \ldots, \sigma_{n,i_L}^2)$
CCI–I (Ray.)	$\frac{1}{\lvert\mathcal{S}\rvert}\sum_{\boldsymbol{s}\in\mathcal{S}} M_G(p, \sum_{i=1}^I \sigma_{g,i}^2\lvert s_i\rvert^2, \ldots, \sum_{i=1}^I \sigma_{g,i}^2\lvert s_i\rvert^2)$
CCI–II (Ray.)	$\frac{1}{\lvert\mathcal{M}_I\rvert}\sum_{i_1=1}^2\cdots\sum_{i_L=1}^2 c_{i_1}\cdots c_{i_L}\sum_{\boldsymbol{b}_I\in\mathcal{M}_I}$ $M_G(p, \xi_{i_1}\sigma_{g,i_1}^2\lvert b_1\rvert^2, \ldots, \xi_{i_L}\sigma_{g,i_L}^2\lvert b_{1,L}\rvert^2)$
CCI–I (Unf.)	$L^{2L/p}\frac{1}{\lvert\mathcal{S}\rvert}\sum_{\boldsymbol{s}\in\mathcal{S}}\lvert s\rvert^{2L}$
CCI–II (Unf.)	$\frac{1}{\lvert\mathcal{M}_I\rvert}\sum_{i_1=1}^2\cdots\sum_{i_L=1}^2 c_{i_1}\cdots c_{i_L}$ $\sum_{\boldsymbol{b}_I\in\mathcal{M}_I}\left(\sum_{l=1}^L \xi_{i_l}\lvert b_{1,l}\rvert^p\right)^{2L/p}$

IV. 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.

A. Off–line Optimization

If the underlying type of noise is known *a priori*, the metric parameters can be optimized off–line based on (7) or (8). To gain some insight and to make the problem tractable, we assume n.i.d. noise in this subsection. The more general case of non–identically distributed noise will be considered in the next subsection. For n.i.d. noise we may set $q_l = q$ and $p_l = p$, $1 \leq l \leq L$, in metric (2) without loss of optimality, i.e., we can base our off–line optimization on (8) and have to optimize only parameter p. The generalized noise moments required in (8) can be taken from Table I or be obtained by Monte–Carlo simulation. Unfortunately, even for those cases where analytical expressions for the generalized moments are available, a closed–form optimization of p is not possible in general. An exception is n.i.d. unfaded CCI–I, where we can show based on (8) and Table I that the optimum p is given by $p_{\mathrm{opt}} = \infty$ corresponding to metric $m(\tilde{b}) = \max_{l\in\{1,\dots,L\}}\{|r_l - \sqrt{\bar{\gamma}}h_l\tilde{b}|\}$.

If the optimum p cannot be obtained in closed form, numerical optimization is necessary. To illustrate this, we show in Fig. 1 the asymptotic BER (solid lines) calculated based on (8), (9), and the generalized moments in Table I for BPSK, i.i.d. Rayleigh fading, $L = 3$, SNR = 20 dB, and various types of n.i.d. noise as a function of p. The markers indicate simulation results and confirm the analytical results. The bold "+" markers denote the minima of the analytical BER. As expected, Fig. 1 shows that $p = 2$ is optimum for AWGN. In constrast, for heavy–tailed types of noise such as Rayleigh–faded CCI–II and ϵ–mixture noise $p_{\mathrm{opt}} < 2$ holds for the optimum p. For short–tailed noise such as unfaded CCI–I $p_{\mathrm{opt}} > 2$ holds. Note that for unfaded CCI–II there are two local minima. Fig. 1 clearly illustrates the benefits of optimizing p and confirms our analysis.

B. On–line Optimization

In practice, the statistical properties of the noise impairing a wireless communication system are often not known *a priori*. Therefore, in this section, we provide an adaptive algorithm for optimization of the L_p–norm metric parameters $\boldsymbol{q}$ and $\boldsymbol{p}$ that does not require any prior knowledge regarding the noise statistics. Since the outcome of the detection process with L_p–norm combining, is invariant to multiplication with a positive constant, we can fix $q_1 = 1$ and optimize only the elements of vector $\boldsymbol{x} \triangleq [q_2 \ \dots q_L \ \boldsymbol{p}^T]^T$ without loss of optimality. Furthermore, in each iteration t the proposed adaptive algorithm requires an estimate of the cost function to be minimized. Based on (7), we obtain the instantaneous cost function estimate

$$L_t(\boldsymbol{x}) \triangleq \frac{\prod_{l=1}^{L}\Gamma\left(\frac{2}{p_l}\right)\left(\sum_{l=1}^{L} q_l|\hat{n}_l[t]|^{p_l}\right)^{\sum_{l=1}^{L}2/p_l}}{\prod_{l=1}^{L=1}\left(p_l q_l^{2/p_l}\right)\Gamma\left(\sum_{l=1}^{L}\frac{2}{p_l}+1\right)} \qquad (14)$$

where we have neglected all irrelevant terms and $\hat{n}_l[t] \triangleq r_l - \sqrt{\bar{\gamma}_l}h_l b[t]$. Here, $b[t]$ may be a training symbol or a previous decision.

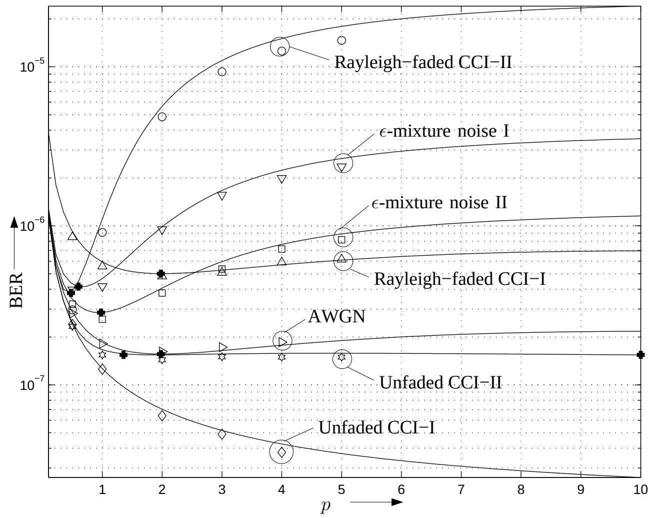

Fig. 1. BER vs. p for BPSK, i.i.d. Rayleigh fading, $L = 3$, SNR = 20 dB, and different types of n.i.d. noise. Noise parameters: I.i.d. ϵ–mixture noise I ($\epsilon = 0.1$, $\kappa = 10$), i.i.d. ϵ–mixture noise II ($\epsilon = 0.1$, $\kappa = 5$), n.i.d. Rayleigh–faded and unfaded QPSK CCI–I ($I = 1$, $\tau_1 = 0.25T$ with symbol duration T, raised cosine pulse shape with roll–off factor 0.22), and i.i.d. Rayleigh–faded and unfaded QPSK CCI–II ($I = 1$, $\epsilon_1 = 0.41$). Solid lines: Asymptotic BER based on (8), (9), and Table I. Markers: Simulation results. Bold "+"–markers: Minimum of asymptotic BER.

Adaptive Algorithm: The proposed adaptive algorithm is based on the multivariate finite difference stochastic approximation (FDSA) framework with gradient approximation [12]. 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 $L_t(\boldsymbol{x})$ [13]. In the tth iteration the FDSA algorithm generates the estimate $\boldsymbol{x}_t$ for the optimum $\boldsymbol{x}$ as [12]

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - a_t\hat{\boldsymbol{g}}_t(\boldsymbol{x}_t) \qquad (15)$$

$$\hat{\boldsymbol{g}}_t(\boldsymbol{x}_t) = \left[\frac{L_t(\boldsymbol{x}_t + c_t\boldsymbol{e}_1) - L_t(\boldsymbol{x}_t - c_t\boldsymbol{e}_1)}{2c_t} \ \dots \ \frac{L_t(\boldsymbol{x}_t + c_t\boldsymbol{e}_{2L-1}) - L_t(\boldsymbol{x}_t - c_t\boldsymbol{e}_{2L-1})}{2c_t}\right]^T \qquad (16)$$

where $\boldsymbol{e}_n$ is a column vector of length $2L-1$ with a one in position n and zeros 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 finds the global minimum if the BER has only one minimum and at least a local minimum otherwise (as long as $L_t(\boldsymbol{x})$ and the BER meet some mild conditions, see [12] for more details). 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 constants to give the algorithm some tracking capability. Furthermore, since p_l may assume very large values, it is advisable to limit the maximum of the elements of $\boldsymbol{x}_t$ to some finite value $x_{\max}$ to improve the tracking capabilities of the algorithm. For initialization, $q_l = 1$, $2 \leq l \leq L$, and $p_l = 2$, $1 \leq l \leq L$, are a good choice since this guarantees that the solution found by the algorithm will not perform worse than conventional MRC.

Example: In Figs. 2 and 3 we show metric coefficients q_l, p_l and the corresponding BER for BPSK as a function of the number of iterations t of the FDSA algorithm, respectively. For this example, we have chosen $a_t = a = 4 \cdot 10^{-4}$, $c_t =$

$c = 10^{-5}$, $x_{\max} = 10$, SNR = 16 dB, and i.i.d. Rayleigh fading with $L = 4$. At $t = (\nu - 1) \cdot 10^6$ we switch abruptly to a new noise Nν, $1 \le \nu \le 5$, which is defined as follows. 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 $\epsilon_l = 0.1$, $1 \le l \le 4$, 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). $\boldsymbol{x}_t$ is initialized with $q_l = 1$, $2 \le l \le 4$, and $p_l = 2$, $1 \le l \le 4$, and previous decisions $\hat{b}[t]$ are used in the adaptation process. For the metric coefficients in Fig. 2 the results of one adaptation process are shown and the BER in Fig. 3 is calculated with (7) and (9), where the generalized moments were obtained by Monte–Carlo simulation. Figs. 2 and 3 show that the FDSA algorithm works well and that after each switch to a new type of noise steady state operation is achieved quickly. Fig. 2 reveals 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 as expected. For N5 $p_l \to \infty$, $1 \le l \le 4$, is optimum and the FDSA yields $p_l = 10$, $1 \le l \le 4$, because we set $x_{\max} = 10$. Fig. 3 shows that the L_p–norm metric with FDSA adaptation substantially outperforms the L_2–norm metric (i.e., MRC).

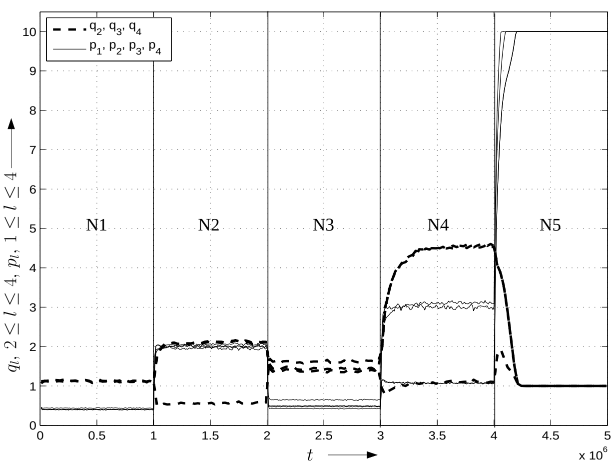

Fig. 2. Metric coefficients p_l, $2 \le l \le 4$, and p_l, $1 \le l \le 4$, vs. iteration t.

V. Results and Discussion

In this section, we compare the adaptive L_p–norm metric with the conventional L_2–norm metric and several other popular robust metrics. For convenience we define $u_l = |r_l - \sqrt{\gamma_l} h_l \tilde{b}|$. 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 $u_l \le \delta$, and $m_l(\tilde{b}) = \delta u_l - \delta^2/2$ if $u_l > \delta$ [6], the Meridian metric $m(\tilde{b}) = \sum_{l=1}^{L} \log(u_l + \delta)$ [7], and the Myriad metric $m(\tilde{b}) = \sum_{l=1}^{L} \log(u_l^2 + \delta^2)$ [7]. Note that for all these robust metrics parameter δ has to be optimized by hand, which is quite tedious, since, unlike for the L_p–norm metric, a systematic optimization framework is not available. Figs. 4 and 5 show the BER of 16–QAM for

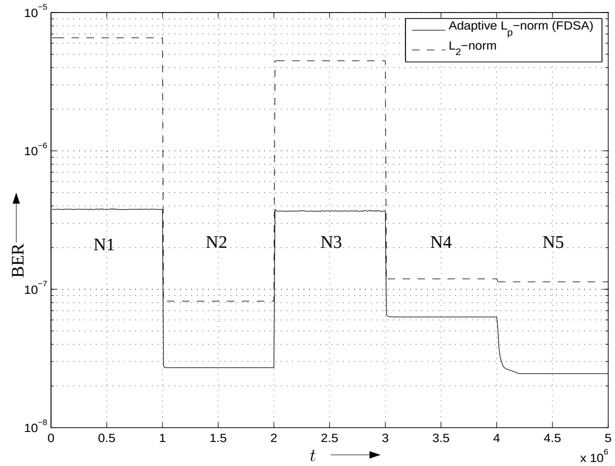

Fig. 3. BER vs. iteration t.

i.i.d. Rayleigh fading with $L = 2$ and impairment by i.i.d. ϵ–mixture noise and n.i.d. unfaded QPSK CCI–I, respectively. Simulation results for all metrics are shown (solid lines with markers). In addition, for the L_p–norm metric and the L_2–norm metric the simulation results are confirmed by the analytical asymptotic BER (bold solid and dashed lines). For the robust metrics δ was optimized by simulation for SNR = 30 dB. In contrast, the L_p–norm metric was optimized with the FDSA algorithm. Since the considered noises are n.i.d., $p = p_1 = p_2$ and $q = q_1 = q_2$ is valid for the L_p–norm metric. Fig. 4 shows that for the heavy–tailed ϵ–mixture noise the L_p–norm metric with $p_{\text{opt}} = 0.39$ outperforms the other robust metrics and the gap to the optimum ML–metric is less than 1 dB. Fig. 5 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 more than 1 dB by the L_p–norm metric with $p = 20$ ($p_{\text{opt}} \to \infty$ holds in this case). Interestingly, while the L_p–norm metric was optimized based on the presented asymptotic analysis, Figs. 4 and 5 suggest that it also performs well for small SNRs.

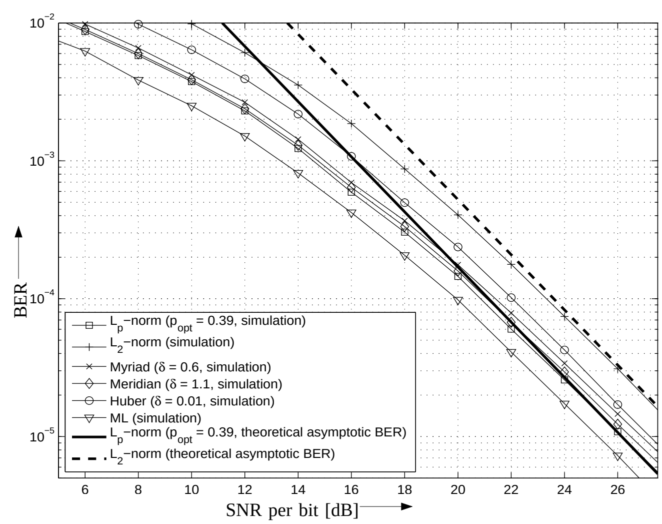

Fig. 4. BER vs. SNR per bit for 16–QAM, i.i.d. Rayleigh fading, $L = 2$, and i.i.d. ϵ–mixture noise ($\epsilon = 0.1$, $\kappa = 100$).

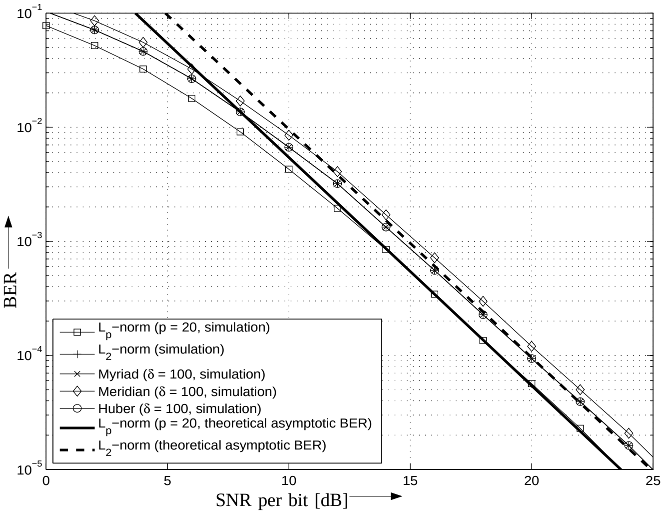

Fig. 5. BER vs. SNR per bit for 16–QAM, i.i.d. Rayleigh fading, $L = 2$, and n.i.d. unfaded QPSK CCI–I ($I = 1$, $\tau_1 = 0.3T$, raised–cosine pulse shape with roll–off $\alpha = 0.22$).

VI. Conclusions

In this paper, we have derived a closed–form expression for the asymptotic BER of coherent L_p–norm combining in Ricean fading and non–Gaussian noise and interference. Based on the asymptotic BER, we have developed an efficient FDSA algorithm for on–line adaptation of the metric parameters. The proposed adaptive L_p–norm metric was shown to outperform other robust metrics in both heavy–tailed and short–tailed non–Gaussian noise.

Appendix

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

$$P_e(d) = \Pr\{m(b) > m(\hat{b})\}, \tag{17}$$

where $d \triangleq |e|$ and $e \triangleq b - \hat{b}$. In a first step, we calculate the PEP conditioned on the noise vector $\boldsymbol{n} \triangleq [n_1 \, \dots \, n_L]^T$. With (2) and (17) this conditional PEP can be obtained as

$$P_e(d|\boldsymbol{n}) = \int_0^{m(b)} f(z)\,\mathrm{d}z, \tag{18}$$

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

$$\begin{aligned} f_{z_l}(z_l) = & \frac{2 z_l^{2/p_l - 1}}{d^2 \bar{\gamma}_l \sigma_{h_l}^2 p_l q_l^{2/p_l}} \exp\!\left(- \frac{z_l^{2/p_l} + q_l^{2/p_l} |\sqrt{\bar{\gamma}_l} \bar{h}_l e + n_l|^2}{d^2 \bar{\gamma}_l \sigma_{h_l}^2 q_l^{2/p_l}} \right) \\ & \times I_0 \left(2 \frac{z_l^{1/p_l} |\sqrt{\bar{\gamma}_l} \bar{h}_l e + n_l|}{d^2 \bar{\gamma}_l \sigma_{h_l}^2 q_l^{1/p_l}} \right). \end{aligned} \tag{19}$$

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

$$f_{z_l}(z_l) = \frac{C_l}{\bar{\gamma}_l} z_l^{2/p_l - 1} + o(\bar{\gamma}_l^{-1}), \tag{20}$$

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 $\Phi_{z_l}(s) \triangleq \mathcal{E}\{e^{-s z_l}\} = C_l \Gamma(2/p_l) \bar{\gamma}_l^{-1} s^{-2/p_l} + o(\bar{\gamma}_l^{-1})$. Since, conditioned on $\boldsymbol{n}$ the z_l are statistically independent, the MGF of $m(\hat{b})$ is given by $\Phi(s) = \prod_{l=1}^{L} \Phi_{z_l}(s)$, and the asymptotic expansion of the corresponding pdf is given by

$$f(z) = \frac{\prod_{l=1}^{L} \left(C_l \Gamma\left(\frac{2}{p_l}\right) \right)}{\Gamma\left(\sum_{l=1}^{L} \frac{2}{p_l} \right) \prod_{l=1}^{L} \bar{\gamma}_l} z^{\sum_{l=1}^{L} \frac{2}{p_l} - 1} + o\left(\prod_{l=1}^{L} \bar{\gamma}_l^{-1} \right). \tag{21}$$

Using this result in (18) leads to

$$\begin{aligned} P_e(d|\boldsymbol{n}) = & \frac{\prod_{l=1}^{L} \left(C_l \Gamma\left(\frac{2}{p_l}\right) \right)}{\Gamma\left(\sum_{l=1}^{L} \frac{2}{p_l} + 1 \right) \prod_{l=1}^{L} \bar{\gamma}_l} \left(\sum_{l=1}^{L} q_l |n_l|^{p_l} \right)^{\sum_{l=1}^{L} \frac{2}{p_l}} \\ & + o\left(\prod_{l=1}^{L} \bar{\gamma}_l^{-1} \right). \end{aligned} \tag{22}$$

If all joint moments of the elements of $\boldsymbol{n}$ are finite, averaging $P_e(d|\boldsymbol{n})$ in (22) with respect to $\boldsymbol{n}$ yields (7). The assumption of finite joint noise moments is necessary, since the terms absorbed into $o(\prod_{l=1}^{L} \bar{\gamma}_l^{-1})$ in (22) 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. 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.

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

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

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

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

[8] 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.

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

[10] 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 on-line:http://www.ece.ubc.ca/~alinezam/TCOM-07.pdf*, 2008.

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

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

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