Performance Analysis of Linear Modulation Schemes With Generalized Diversity Combining on Rayleigh Fading Channels With Noisy Channel Estimates

Ramesh Annavajjala, *Member, IEEE*, Pamela C. Cosman, *Senior Member, IEEE*, and Laurence B. Milstein, *Fellow, IEEE*

***Abstract*—Generalized diversity combining (GDC), also known as hybrid selection/maximal ratio combining or generalized selection combining, is a low-complexity diversity combining technique by which a fixed subset of a large number of available diversity channels is chosen and then combined using the rules of maximal ratio combining. In this paper, we analyze the performance of GDC on time-correlated Rayleigh fading channels with noisy channel estimates. We derive expressions for the probability of error for various linear modulation schemes with coherent detection, and discuss the conditions under which the analysis can be extended to noncoherent and differentially coherent receiver structures. Throughout the paper, using a fundamental approach to obtain the decision statistic at the combiner output, a number of new expressions for the error probabilities are obtained in a rigorous way, along with a presentation of their performance with channel estimation errors. The final expressions have roughly the same complexity of evaluation as that for the channel with only additive Gaussian noise. Our results correct various inaccuracies in the literature, and show that coherent receivers based on imperfectly estimated channel knowledge incur a significant performance loss.**

***Index Terms*—Generalized diversity combining, imperfect channel estimation, pilot symbol-assisted modulation (PSAM), Rayleigh processes, two-dimensional signal constellation.**

I. Introduction

WIDEBAND wireless channels are capable of resolving a large number of multipath components which can be combined constructively to improve communication reliability [1, Ch.29]. This can result in a low signal-to-noise ratio (SNR) on a per-resolvable path basis, which exacerbates the system's ability to obtain accurate channel estimates, as do the effects of a large Doppler spread and/or a low-rate coding scheme. Further, systems designed to have disparate users share a common spectrum, such as cognitive radio and ultra wideband, are dependent upon accurate channel estimation techniques to ensure efficient operation. In practice, due to implementation constraints, only a subset of the available paths are typically combined. Generalized diversity combining (GDC), also referred to as hybrid-selection/maximal ratio combining or generalized selection combining, is a technique to choose a fixed subset (of size K) of a large number of available diversity channels (of size L) and then combine them using the rules of maximal ratio combining (MRC) [2]. With perfect channel state information (CSI) at the receiver, and for large values of the average received SNR, a GDC(L, K) receiver with $K < L$ can achieve the same diversity order, L, as that of MRC [1]. In practice, the receiver has to estimate the channel and the CSI is not perfect. An information-theoretic approach to the effect of imperfect CSI on the channel capacity can be found, for example, in [3] and [4], whereas the main goal of this paper is an exact quantification of the effect of noisy channel estimates on the error probability performance of linear modulation schemes with GDC.

We now summarize the relevant research work dealing with the error performance of digital modulation schemes on fading channels with imperfect CSI, and contrast them with the results we derive in this paper. In [5], the authors analyze the performance of L-branch diversity[1] for independent and identically distributed (i.i.d) Rayleigh fading channels with a separate pilot channel for estimating the fade in the data channel. They consider both coherent binary phase-shift keying (BPSK) and noncoherent binary frequency-shift keying (BFSK) signaling schemes, derive the probability density function (pdf) of the instantaneous SNR random variable (r.v.) at the output of the combiner, and use it to average the conditional error probability expressions [5, Eqns. (16), (31), (34), and (40)] to obtain the average error rates. As shown in [6], with a completely decorrelated pilot channel, while the average error rate for BPSK signaling is 0.5, noncoherent BFSK is unaffected by estimation errors (see, also, Section IV-A of this paper). However, an inaccurate conclusion in [5] is that, with an uncorrelated pilot

Manuscript received June 10, 2005; revised May 15, 2007. This work was supported in part by the Office of Naval Research under Grant N00014-03-1-0280, the National Science Foundation under Grant CCF-0635165, the Center for Wireless Communications at the University of California, San Diego (UCSD), and LG Electronics. The material in this paper was presented in part at the IEEE Information Theory and Applications (ITA) Workshop, San Diego, CA, January 2007.

R. Annavajjala was with the Electrical and Computer Engineering Department, University of California, San Diego, La Jolla, CA 92093 USA. He is now with ArrayComm LLC, San Jose, CA 95131 USA (e-mail: ramesh.annavajjala@gmail.com).

P. C. Cosman and L. B. Milstein are with the Electrical and Computer Engineering Department, University of California, San Diego, La Jolla, CA 92093 USA (e-mail: pcosman@ucsd.edu; milstein@ece.ucsd.edu).

Communicated by A. Høst-Madsen, Associate Editor for Detection and Estimation.

Color versions of Figures 1 and 4–11 in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2007.909130

[1]Throughout this paper, *L-branch diversity* is to be interpreted as combining all the available diversity branches (i.e., GDC(L, L)).

channel, the error probability varies inverse linearly with the average received SNR.[2]

To analyze the performance of MRC with Gaussian weighting errors, [10], in a novel way, models the channel fade as a function of the channel estimate (since both the channel gain and its estimate are assumed to be jointly complex Gaussian) and obtains the pdf of the SNR at the output of the combiner to analyze the outage behavior. However, a conclusion of [10, eq. (50)], shows that, with a completely decorrelated fade estimate, the outage probability approaches that of a no-diversity system, whereas, in reality, outage occurs with probability one (see (14) in this paper and the discussion below it). In [11], the authors analyze the error performance of a binary differentially coherent PSK (BDPSK) receiver for L i.i.d. Rayleigh fading channels with imperfect channel estimates. A BDPSK receiver uses the signal received in the previous symbol interval as a channel estimate for the current symbol, and hence channel estimation is not a requirement. However, [11, eq. (20)] shows that with a completely decorrelated fade estimate the error probability approaches that of a system with no diversity, whereas we show that BDPSK is insensitive to channel estimation errors (see (119) in this paper). References [12], and [13] extend the results of [11] for selection combining (SC, i.e., GDC$(L, 1)$) and GDC schemes, respectively, and for various modulation/demodulation formats. We address the following main limitations of [12] and [13].

1) With coherent detection, [12] and [13] do not account for the effect of *crosstalk* and *signal-dependent noise*, due to imperfect estimates, on the quadrature branches of the modulation signals, and show that imperfect channel estimates reduce the diversity order without causing any error floor. Specifically, with completely decorrelated channel estimates, [13] shows that the outage performance of a coherent GDC(L, K) receiver approaches that of a no-diversity system.
   In this paper, we present a new analysis on the outage probability (see Appendix I), and show that the diversity order of a coherent GDC(L, K) scheme is preserved even with noisy CSI, whereas the error floor limits the receiver performance (also see [14] for a related study on multiple-input multiple-output channels). We re-examine the average symbol error probability (SEP) expressions for coherent phase-shift keying (PSK) [13, eq. (17)], quadrature amplitude modulation (QAM) [13, eq. (19)], and general two-dimensional (2-D) modulations with polygonal decision boundaries [13, eq. (12)], and derive versions of these expressions that take into account signal-dependent noise and crosstalk between the in-phase and quadrature branches.
2) The authors in [12] and [13] show that M-ary noncoherent and differentially coherent receivers are severely impacted by imperfect CSI (see, (18) in [13], for M-ary noncoherent FSK (NCFSK) and (20) for M-ary differentially encoded PSK (DPSK) (M-DPSK)). In particular, with completely decorrelated channel estimates, [13] shows that the average error performances of FSK and DPSK receivers with GDC(L, K) vary inverse linearly with the average received SNR.
   We show that these receivers are insensitive to channel estimation errors (see (123) for M-ary NCFSK and (130) for M-ary DPSK in this paper). Specifically, we use the channel estimates only for choosing the K diversity channels from the L available ones, but not for the actual demodulation/detection process (see Figs. 2 and 3). With completely decorrelated channel estimates, we show that the error rates of both noncoherent and differentially coherent receivers with GDC(L, K) coincide with that of an ideal GDC(K, K) receiver.

In addition to the above, the expressions derived in this paper also extend various published results on coherent modulation with imperfect channel estimates. To this end, we first briefly review some of these published results. An upper bound on the SEP for QAM with pilot-symbol assisted modulation (PSAM) [15] is presented in [16], whereas performance of M-PSK and M-QAM with minimum mean-square error (MMSE) channel estimation on Rayleigh and Rician fading channels without diversity is presented in [17] and [18], respectively. With an assumption that the amplitude and phase estimation errors are independent of each other, approximate bit error probability (BEP) performances of 16- and 64-QAM modulation schemes are analyzed in [19] for a Rayleigh fading channel. An exact expression, in terms of a complicated double-integral, for the average BEP of 16-QAM is obtained in [20] with MRC diversity and channel estimation errors. Using the results on Gaussian quadratic forms [21, Appendix B], [22] presents closed-form expressions for the average BEP of M-QAM with MRC on Rayleigh fading channels, whereas a Rician-fading channel is considered in [23]. An approximate analysis of BEP for M-QAM with GDC is conducted in [24] for Rayleigh, Rician, and Nakagami fading channels. SEP analysis for general 2-D modulation schemes is investigated in [25] for Rayleigh fading channels with channel estimation errors and no diversity. Probability density functions, with channel estimation errors, for analyzing the performance of PAM and QAM signals on Rayleigh fading channels with MRC diversity, and for Rician-fading channels without diversity, are developed in [26].

As described in the previous paragraph, most of the reported results are limited to either constellations with restricted alphabet sizes (such as 16-QAM/64-QAM), or a particular choice of diversity scheme (such as MRC or no diversity). In particular, for QAM constellations, the analytical framework with estimation errors, so far, is limited to BEP performance only. The following contributions in our paper extend various results summarized in the previous paragraph.

1) For M-PSK modulation, we derive the conditional (conditioned on the channel estimates) distribution of the phase angle of the received signal at the output of the GDC receiver with imperfect channel estimates. This result is a generalization of [27], wherein Proakis derives the distribution of the phase angle of the received signal for MRC with channel estimation errors. Using this distribution, we extend the BEP expressions of [28] to account for fading, GDC, and noisy CSI. Furthermore, our results are exact (whereas [29] presents an approximate analysis), and are in a simple closed form.

[2]This conclusion has also appeared in some classic textbooks (see [7, pp. 549–555] and [8, pp. 325–329, 371–375]), and in a recent work [9, eqs. (39), (40), (42), and (56).]

2) For an M-PAM (pulse-amplitude modulation) signal set, we derive new expressions for SEP and BEP (with Gray mapping) which extend [26] to GDC(L, K) reception based on noisy channel estimates.
3) With QAM, our average BEP expressions with GDC and channel estimation errors are valid for arbitrary rectangular constellation sizes with Gray code mapping. Our BEP results generalize the 16-QAM and MRC results of [20] to M-QAM and GDC, and M-QAM and MRC results of [22] to GDC. A closed-form analysis on the average SEP for M-QAM is also presented, which, to the best of our knowledge, has not been reported in the literature.
4) Analogous to the single-antenna results of [25], our analysis on the average SEP performance of 2-D constellations allows us to express the final results in terms of a single integral [30]. Our results are nontrivial generalizations of [25] to GDC and imperfect CSI.

The rest of this paper is organized as follows. In Section II, we describe the system and the channel estimation error models. Analysis of average SEP and BEP of various coherent signaling schemes is presented in Section III. In particular, M-PSK signaling is considered in Section III-A, M-PAM and M-QAM are considered in Sections III-B and III-C, respectively, and an analysis is presented for arbitrary 2-D constellations in Section III-D. Extensions to noncoherent and differentially coherent schemes are studied in Section IV. Numerical results and discussions are provided in Section V, and we conclude our work in Section VI.

II. System Model

We assume that the information bits are mapped onto a general 2-D constellation $\mathcal{S}$ with the r.v. $S \in \mathcal{S}$ denoting the transmitted signal point. The signal points are normalized to have an average energy of E_s (i.e., $E[|S|^2] = E_s$). We assume that the channel is frequency nonselective and slowly fading over the duration of the transmitted symbol, and the receiver employs L antennas for diversity reception. Assuming perfect recovery of symbol timing, the low-pass equivalent representation of the received signal at the output of a matched filter on the lth antenna path is given by

$$r_l = g_l S + n_l, \qquad l = 1, 2, \ldots, L \tag{1}$$

where g_l is the complex channel gain whose real and imaginary parts are assumed to be uncorrelated and are Gaussian distributed each with zero mean and variance of σ_g^2. The noise r.v. n_l is complex Gaussian with independent components each with zero mean and variance $\sigma^2 = N_0/2$. The channel gains, g_l and g_m, at two different diversity branches l and m, are assumed to be i.i.d. We also assume that g_l is independent of $\{n_l\}_{l=1}^{L}$. Note that this model is chosen because it has often been used in the past (see, e.g., [19] and [31]). The implicit assumption we are making is that various physical effects, such as path loss and multipath fading, as well as all normalizations from gains at the receiver, are embodied in the variance of g_l, $2\sigma_g^2$.

Let p_l be the estimate of the complex fade g_l on the lth diversity path, which is also assumed to be a complex Gaussian r.v. with zero mean and variance of $E\left[|p_l|^2\right] = 2\sigma_p^2$. Since p_l and g_l are jointly Gaussian, the conditional distribution of g_l, conditioned on p_l, is also Gaussian with mean proportional to p_l and variance independent of p_l. That is, conditioned on p_l, we can express g_l as [32]

$$g_l = (x_l + j y_l) + \varrho \frac{\sigma_g}{\sigma_p} p_l \tag{2}$$

where ϱ is the complex correlation coefficient between p_l and g_l, and x_l and y_l are independent Gaussian random variables (r.v.'s) each with zero mean and variance $\sigma_e^2 = \sigma_g^2(1 - |\varrho|^2)$, and are independent of p_l. The parameter σ_e^2 is defined as the channel estimation error variance (per dimension). The complex correlation coefficient $\varrho = \rho e^{\phi_\rho}$ between g_l and p_l is defined as

$$\varrho \triangleq \frac{E[g_l p_l^*]}{\sqrt{E[p_l p_l^*]\, E[g_l g_l^*]}} = \frac{R_c + j R_{cs}}{\sigma_p \sigma_g} \tag{3}$$

where $E[g_l p_l^*] \triangleq 2(R_c + j R_{cs})$. Then

$$\rho = \frac{\sqrt{R_c^2 + R_{cs}^2}}{\sigma_p \sigma_g} = \sqrt{\rho_c^2 + \rho_{cs}^2} \tag{4}$$

$$\text{and} \quad \phi_\rho = \tan^{-1}\left(\frac{\rho_{cs}}{\rho_c}\right) = \tan^{-1}\left(\frac{R_{cs}}{R_c}\right). \tag{5}$$

In (4), we defined $\rho_c = \frac{R_c}{\sigma_p \sigma_g}$ and $\rho_{cs} = \frac{R_{cs}}{\sigma_p \sigma_g}$.

A. Practical Channel Estimation Schemes

The previously described channel estimation error model can be specialized to a variety of practical channel estimation schemes. In this subsection, we illustrate this for three popular channel estimation schemes.

1) Additive Channel Estimation Errors: If a channel estimation scheme results in an additive error, then the estimate p_l can be written as $p_l = g_l + w_l$. With the assumption that w_l is a complex Gaussian r.v. with zero mean and variance $E\left[|w_l|^2\right] = 2\sigma_w^2$, and is independent of g_l, by using (3) we directly obtain $\varrho = \rho = \frac{\sigma_g}{\sqrt{\sigma_g^2 + \sigma_w^2}}$ and $\phi_\rho = 0$. Clearly, the channel estimation error variance is given by $\sigma_e^2 = \sigma_g^2(1-\rho^2) = \sigma_g^2 \sigma_w^2 / \left(\sigma_g^2 + \sigma_w^2\right)$. We point out that the *clairvoyant* and *pilot signal* estimates, as discussed by Proakis in [27], can be viewed as particular instances of the general additive estimation error model.

2) MMSE Channel Estimation: With an MMSE channel estimation scheme, the channel estimate is chosen in such a way that the mean square error between the estimate p_l and the fade g_l is minimized. From [33], it is well known that, with MMSE estimation, the estimation error $g_l - p_l$ is uncorrelated with the estimate p_l. Since both p_l and g_l are complex Gaussian, it follows that $g_l - p_l$ is independent of p_l. Upon setting $E[(g_l - p_l)p_l^*] = 0$, we arrive at $R_c = \sigma_p^2$, $R_{cs} = 0$, $\rho_{cs}^2 = R_{cs}^2 / \left(\sigma_p^2 \sigma_g^2\right) = 0$, and $\rho^2 = \rho_c^2 = R_c^2 / \left(\sigma_p^2 \sigma_g^2\right) = \sigma_p^2 / \sigma_g^2$. Finally, the estimation error σ_e^2 is given in terms of σ_g^2 and σ_p^2 as $\sigma_e^2 = \sigma_g^2(1-\rho^2) = \sigma_g^2 - \sigma_p^2$.

3) Pilot Symbol Assisted Modulation: In a PSAM system, as detailed in [15] and [19], information symbols are packed into N-length frames containing one pilot symbol followed by $N-1$ information symbols. The channel estimate is derived from the pilot symbols of N_1 past, the present, and N_2 future frames.

If $g_l^{\text{pilot}}(j)$ and $\eta_l^{\text{pilot}}(j)$ denote, respectively, the complex fade and the additive noise on the pilot symbol corresponding to the jth frame on the lth branch, and if E_p denotes the energy-per-branch invested on the pilot symbol, then the estimate $p_l(n)$ on the nth symbol corresponding to the current frame can be written as

$$p_l(n) = \sum_{j=-N_1}^{N_2} f_{j,n}\left(g_l^{\text{pilot}}(j) + \frac{\eta_l^{\text{pilot}}(j)}{\sqrt{E_p}}\right) \tag{6}$$

where $\{f_{j,n}\}$ is the set of real filter coefficients and the pilot symbols are assumed to be BPSK modulated. Clearly, $p_l(n)$ is zero mean complex Gaussian with variance [19]

$$E\left[|p_l(n)|^2\right] = 2\sigma_p^2 = 2\sigma_g^2 \sum_{j=-N_1}^{N_2}\sum_{k=-N_1}^{N_2} f_{j,n}f_{k,n} \times J_0(2\pi f_d|k-j|NT_s) + \frac{N_0}{E_p}\sum_{i=-N_1}^{N_2} f_{i,n}^2 \tag{7}$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind [8], f_d is Doppler frequency, and T_s is the symbol duration. Again, the estimate $p_l(n)$ of (6) and the fade $g_l(n)$ are jointly Gaussian, so that from (3), we have

$$R_c = E\left[\frac{p_l(n)g_l^*(n)}{2}\right] = \sigma_g^2 \sum_{j=-N_1}^{N_2} f_{j,n}J_0(2\pi f_d(jN-n)T_s) \tag{8}$$

$$\text{and} \quad R_{cs} = 0. \tag{9}$$

Using (7)—(9) in (4), ρ^2 and σ_e^2 are simply given by (10)–(11) at the bottom of the page, where, in (10), $\overline{\gamma}_{\text{pilot}} = 2\sigma_g^2E_p/N_0$ is the average SNR per branch for the pilot signal.

III. Error Probability Analysis

With the received signal of (1) and the corresponding channel estimates $\{p_l\}_{l=1}^L$, the output of the diversity combiner for a linear modulation scheme is given by

$$r = \frac{\sum_{l=1}^{K} p_{(l)}^* r_{(l)}}{\sum_{l=1}^{K}|p_{(l)}^*|^2}$$

$$= S\left(\frac{R_c + jR_{cs}}{\sigma_p^2}\right) + \frac{\sum_{l=1}^{K}\left\{Sp_{(l)}^*(x_l+jy_l) + n_l p_{(l)}^*\right\}}{\sum_{l=1}^{K}|p_{(l)}^*|^2}$$

$$\triangleq S\left(\frac{R_c + jR_{cs}}{\sigma_p^2}\right) + \eta_1 + j\eta_2 \tag{12}$$

where $p_{(1)}, p_{(2)}, \ldots, p_{(L)}$ are the order statistics of $p_1, p_2, \ldots, p_L$ such that $|p_{(1)}| \geq |p_{(2)}| \geq \cdots \geq |p_{(L)}|$, and $r_{(l)}$ is the received signal on the diversity branch for which $p_{(l)}$ is the corresponding channel estimate. For simplicity, let us define $\beta^2 = \left(\sum_{l=1}^{K}|p_{(l)}|^2\right)/\left(2\sigma_p^2\right)$. Conditioned on β^2, η_1 and η_2 are zero mean independent Gaussian r.v.'s each having a variance

$$\sigma_{|S|}^2 = \frac{|S|^2\sigma_e^2 + \sigma^2}{\sum_{l=1}^{K}|p_{(l)}|^2} = \frac{|S|^2\sigma_e^2 + \sigma^2}{2\sigma_p^2\beta^2}. \tag{13}$$

Recognize that β^2 is the normalized SNR r.v. at the output of a genie-aided GDC receiver [34], [35]. We also notice that, unlike the case of ideal channel estimation, the variance $\sigma_{|S|}^2$, conditioned on β^2, depends on the transmitted signal point S. However, for an M-PSK signal set, $\sigma_{|S|}^2$ is not a function of S. Interestingly, the r.v.'s η_1 and η_2 are correlated and are non-Gaussian distributed. Thus, we conclude from (12) and (13) that the effect of imperfect channel estimation at the output of a linear diversity combiner is scaling the transmitted signal S by *an unknown* (to the receiver) complex constant $(R_c + jR_{cs})/\sigma_p^2$ and then corruption by a complex, correlated, non-Gaussian noise whose variance is proportional to the transmitted signal energy.

An interesting observation can be made from (12) when the channel estimate is completely decorrelated from the actual channel gain. In this scenario, we have $\varrho \to 0$. That is, $(R_c + jR_{cs}) \to 0$ and (12) reduces to

$$r = \eta_1 + j\eta_2 \quad (\text{for } \varrho = 0). \tag{14}$$

That is, there is no signal component at the output of the combiner. As a result, with no further computation, we conclude that *with a completely decorrelated channel estimate, the outage probability, the probability that the received SNR at the output of a coherent diversity combiner falls below a predetermined threshold, is always unity irrespective of the modulation type, the number of paths L, and the parameter K of the diversity combiner.* For the sake of completeness, an

$$\rho^2 = \frac{R_c^2}{\sigma_p^2\sigma_g^2} = \frac{\left(\sum_{j=-N_1}^{N_2} f_{j,n}J_0(2\pi f_d(jN-n)T_s)\right)^2}{\sum_{j=-N_1}^{N_2}\sum_{k=-N_1}^{N_2} f_{j,n}f_{k,n}J_0(2\pi f_d|k-j|NT_s) + \frac{1}{\overline{\gamma}_{\text{pilot}}}\sum_{i=-N_1}^{N_2} f_{i,n}^2} \tag{10}$$

$$\sigma_e^2 = \sigma_g^2(1-\rho^2) \tag{11}$$

analysis of the outage probability for an arbitrary value of ϱ is provided in Appendix I, and is contrasted with the results in the literature.

A. M-PSK Constellation

For coherent M-PSK signaling, we have

$$\mathcal{S} = \left\{ \sqrt{E_s} \exp(j\theta_m), \theta_m = 2m\pi/M \right\}_{m=0}^{M-1}.$$

When $S_m = \sqrt{E_s}\exp(j\theta_m)$ is transmitted, (12) can be conveniently written as

$$r = Z_I + jZ_Q \tag{15}$$

where

$$Z_I = \sqrt{E_s}\frac{R_c\cos\theta_m - R_{cs}\sin\theta_m}{\sigma_p^2} + \eta_1 \tag{16}$$

$$\text{and } Z_Q = \sqrt{E_s}\frac{R_c\sin\theta_m + R_{cs}\cos\theta_m}{\sigma_p^2} + \eta_2. \tag{17}$$

The decision statistic that we are interested in is the phase Φ of the received complex variable r, which is defined as $\Phi = \tan^{-1}\left(\frac{Z_Q}{Z_I}\right)$. Note from (16) and (17) that, conditioned on β^2 and θ_m, Z_I and Z_Q are independent real Gaussian r.v.'s with the following means and variances:

$$m_{Z_I} = E[Z_I] = \sqrt{E_s}\frac{R_c\cos\theta_m - R_{cs}\sin\theta_m}{\sigma_p^2} \tag{18}$$

$$m_{Z_Q} = E[Z_Q] = \sqrt{E_s}\frac{R_c\sin\theta_m + R_{cs}\cos\theta_m}{\sigma_p^2} \tag{19}$$

$$\sigma_{Z_I}^2 = \sigma_{Z_Q}^2 = \frac{(\sigma^2 + E_s\sigma_e^2)}{2\sigma_p^2\beta^2}. \tag{20}$$

The following result will be useful for obtaining the pdf of Φ:

Lemma 1: If X and Y are two independent real Gaussian r.v.'s with mean values μ_X and μ_Y, respectively, and have variances of Σ^2 each, then the pdf of $\Theta = \tan^{-1}\left(\frac{Y}{X}\right)$ is given by

$$f_\Theta\left(\theta; \frac{s^2}{2\Sigma^2}, \psi\right) = \frac{e^{-\frac{s^2}{2\Sigma^2}}}{2\pi} + \frac{1}{\sqrt{\pi}}\sqrt{\frac{s^2}{2\Sigma^2}}\cos(\theta-\psi) \times e^{-\sin^2(\theta-\psi)\frac{s^2}{2\Sigma^2}} \times \left(1 - Q\left(\sqrt{\frac{s^2}{2\Sigma^2}}\sqrt{2}\cos(\theta-\psi)\right)\right) \tag{21}$$

where $f_\Theta(\theta; a, b)$ is the pdf of Θ evaluated at θ with parameters a and b, $s^2 = \mu_X^2 + \mu_Y^2$, $\psi = \tan^{-1}\left(\frac{\mu_Y}{\mu_X}\right)$, and

$$Q(u) = \frac{1}{\sqrt{2\pi}}\int_u^\infty \exp(-t^2/2)dt.$$

Proof: Refer to [36, Sec. 5A.5]. □

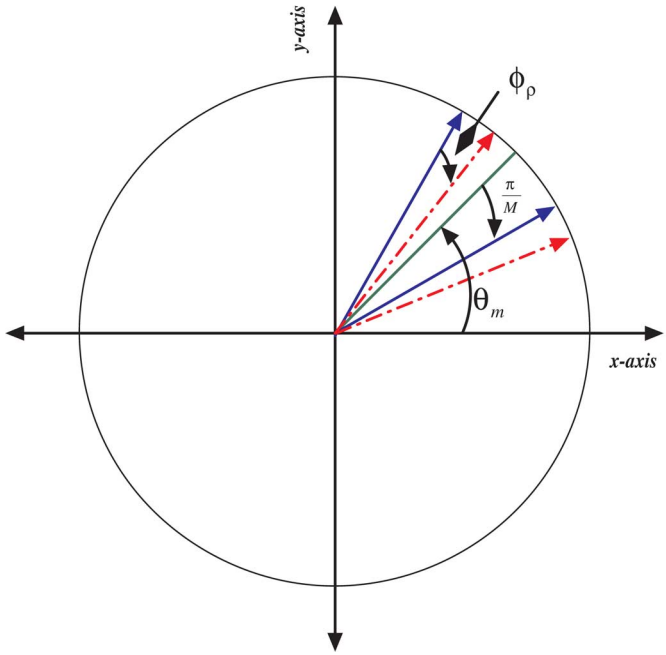

Fig. 1. M-PSK signal constellation with decision boundary in the presence of channel estimation errors. Notice that the angle ϕ_ρ is due to phase estimation errors. Due to this, the decision region when θ_m is transmitted is given by the wedge between $\theta_m - \frac{\pi}{M} - \phi_\rho$ and $\theta_m + \frac{\pi}{M} - \phi_\rho$.

Upon using (20) together with Lemma 1, and after some simplification,[3] we arrive at the following expression for the conditional pdf of Φ:

$$f_\Phi\left(\phi; \overline{\gamma}_{\text{eff}}\beta^2, \phi_m\right) = \frac{e^{-\overline{\gamma}_{\text{eff}}\beta^2}}{2\pi} + \sqrt{\frac{\overline{\gamma}_{\text{eff}}}{\pi}\beta^2}\cos(\phi-\phi_m) \times e^{-\sin^2(\phi-\phi_m)\overline{\gamma}_{\text{eff}}\beta^2} \times \left(1 - Q\left(\sqrt{2\overline{\gamma}_{\text{eff}}\beta^2}\cos(\phi-\phi_m)\right)\right) \tag{22}$$

where $\overline{\gamma} = E\left[|g_l|^2\right]E_s/N_0 = 2\sigma_g^2 E_s/N_0$,

$$\overline{\gamma}_{\text{eff}} = \frac{\overline{\gamma}\rho^2}{1+\overline{\gamma}(1-\rho^2)} \tag{23}$$

$$\text{and } \phi_m = \tan^{-1}\left(\frac{R_c\sin\theta_m + R_{cs}\cos\theta_m}{R_c\cos\theta_m - R_{cs}\sin\theta_m}\right) = \theta_m + \phi_\rho. \tag{24}$$

The probability of symbol error when θ_m is the transmitted phase, from Fig. 1, is

$$P_S(\beta^2) = 1 - \int_{\phi=\theta_m-\frac{\pi}{M}}^{\theta_m+\frac{\pi}{M}} f_\Phi\left(\phi; \overline{\gamma}_{\text{eff}}\beta^2, \theta_m+\phi_\rho\right)d\phi = 1 - \int_{\phi=-\frac{\pi}{M}-\phi_\rho}^{\frac{\pi}{M}-\phi_\rho} f_\Phi\left(\phi; \overline{\gamma}_{\text{eff}}\beta^2, 0\right)d\phi$$

[3]Detailed in Appendix II.

$$= \int_{\phi=-\pi}^{-\frac{\pi}{M}-\phi_\rho} f_\Phi\left(\phi;\overline{\gamma}_{\text{eff}}\beta^2,0\right)d\phi + \int_{\phi=\frac{\pi}{M}-\phi_\rho}^{\pi} f_\Phi\left(\phi;\overline{\gamma}_{\text{eff}}\beta^2,0\right)d\phi \quad (25)$$

where the second equality in (25) is due to a change of integration variable. Observe that $f_\Phi\left(\phi;\overline{\gamma}_{\text{eff}}\beta^2,0\right)$ is just the pdf of the phase angle when $\theta_0 = 0$ is transmitted on a fading channel with perfect CSI and with an instantaneous channel SNR of $\overline{\gamma}_{\text{eff}}\beta^2$ [37]. An important result is that the cumulative distribution function (cdf) of $f_\Phi\left(\phi;\overline{\gamma}_{\text{eff}}\beta^2,0\right)$ is obtained by Pawula *et al.* in [37] in a simplified form which is given in (26) at the bottom of the page, where $\psi_1 < \psi_2$ and

$$F_\Phi(\psi;\overline{\gamma}_{\text{eff}}\beta^2) = -\frac{\text{sgn}(\psi)}{2\pi} \times \int_0^{\pi-|\psi|} e^{-\overline{\gamma}_{\text{eff}}\beta^2\frac{\sin^2\psi}{\sin^2\theta}}d\theta, \quad -\pi < \psi < \pi. \quad (27)$$

In (27), $\text{sgn}(x) = 1$ for $x \geq 0$ and is equal to -1 otherwise. Due to the discontinuity of $F_\Phi(\psi;\cdot)$ of (27) at $\psi = 0$, for evaluating (26) either at $\psi_1 = 0$ or $\psi_2 = 0$,we have to use $F_\Phi(\psi_1 = 0;\cdot) = -1/2$ and $F_\Phi(\psi_2 = 0;\cdot) = 1/2$. For details please refer to [37].

Using (27) in (25), and using the fact that $\text{sgn}(-x) = -\text{sgn}(x)$ for $x \neq 0$, we obtain

$$P_S(\beta^2) = \frac{\text{sgn}\left(\frac{\pi}{M}-\phi_\rho\right)}{2\pi}\int_0^{\pi-|\frac{\pi}{M}-\phi_\rho|} e^{-\frac{\overline{\gamma}_{\text{eff}}\beta^2\sin^2\left(\frac{\pi}{M}-\phi_\rho\right)}{\sin^2\theta}}d\theta + \frac{\text{sgn}\left(\frac{\pi}{M}+\phi_\rho\right)}{2\pi}\int_0^{\pi-|\frac{\pi}{M}+\phi_\rho|} e^{-\frac{\overline{\gamma}_{\text{eff}}\beta^2\sin^2\left(\frac{\pi}{M}+\phi_\rho\right)}{\sin^2\theta}}d\theta. \quad (28)$$

It is to be noted that, due to the definition of the cdf of Φ in (26), (28) is valid only when $|\phi_\rho| < \pi/M$. Expressions similar to (28) can be readily obtained, using (25) and (26), even for the case of $|\phi_\rho| > \pi/M$ or $\phi_\rho = \pi/M$. In practice, $|\phi_\rho|$ is very small, and in what follows, we assume that $|\phi_\rho| < \pi/M$.

Notice that, fortunately, in (28) the r.v. β^2 appears in the exponent of the integrand. By recalling that $E[\exp(-s\beta^2)] = \mathcal{L}_{\beta^2}(s)$, where $\mathcal{L}_X(s)$ is the Laplace transform of the pdf of the r.v. X, the average SEP for M-PSK signaling, with GDC reception and noisy CSI, can be obtained by taking the expectation of (28) over β^2. The result is

$$\begin{aligned}\overline{P}_S &= E\left[P_S(\beta^2)\right] \\ &= \frac{\text{sgn}\left(\frac{\pi}{M}-\phi_\rho\right)}{2\pi} \times \int_0^{\pi-\left|\frac{\pi}{M}-\phi_\rho\right|} \mathcal{L}_{\beta^2}\left(\overline{\gamma}_{\text{eff}}\frac{\sin^2\left(\frac{\pi}{M}-\phi_\rho\right)}{\sin^2\theta}\right)d\theta \\ &\quad + \frac{\text{sgn}\left(\frac{\pi}{M}+\phi_\rho\right)}{2\pi} \times \int_0^{\pi-\left|\frac{\pi}{M}+\phi_\rho\right|} \mathcal{L}_{\beta^2}\left(\overline{\gamma}_{\text{eff}}\frac{\sin^2\left(\frac{\pi}{M}+\phi_\rho\right)}{\sin^2\theta}\right)d\theta. \quad (29)\end{aligned}$$

For the practical channel estimation schemes described in Section II-A, we have $R_{cs} = 0$. This implies that $\phi_\rho = 0$ and $\rho^2 = \rho_c^2$. With this, (29) reduces to

$$\overline{P}_S(\phi_\rho = 0) = \frac{1}{\pi}\int_{\theta=0}^{\frac{\pi(M-1)}{M}} \mathcal{L}_{\beta^2}\left(\overline{\gamma}_{\text{eff}}\frac{\sin^2\left(\frac{\pi}{M}\right)}{\sin^2\theta}\right)d\theta. \quad (30)$$

Equation (30) shows that the average SEP for M-PSK is similar to the ideal SEP, with the ideal average SNR $\overline{\gamma}$ replaced by the effective average SNR $\overline{\gamma}_{\text{eff}}$. When $\rho = 0$, (30) shows that the average SEP of M-PSK modulation is equal to $(M-1)/M$ (i.e., randomly choosing one of M signal points), whereas [13] shows that the average SEP varies inverse linearly with $\overline{\gamma}$.

To obtain expressions for the average SEP, averaged over β^2, the following expression for the Laplace transform of the pdf of β^2 is needed [38]:

$$\begin{aligned}\mathcal{L}_{\beta^2}(s) &= E[\exp(-s\beta^2)] = \left(\frac{1}{1+s}\right)^K \prod_{l=K+1}^{L}\frac{1}{1+\frac{sK}{l}} \\ &= \sum_{l=K+1}^{L}\frac{\chi(l)}{(s+1)^K\left(s+\frac{l}{K}\right)} \quad (1 \leq K < L) \quad (31) \\ &= \left(\frac{1}{1+s}\right)^L \quad (K = L) \quad (32)\end{aligned}$$

where (31) is due to partial fractions techniques and

$$\chi(l) = \frac{l}{K} \times \prod_{m=K+1, m\neq l}^{L}\frac{m}{m-l}. \quad (33)$$

We also need the following trigonometric identity [39]:

$$I_n(\theta;c_1,c_2) = \int_{\phi=0}^{\theta}\left(\frac{\sin^2\phi}{\sin^2\phi+c_1}\right)^n\left(\frac{\sin^2\phi}{\sin^2\phi+c_2}\right)\frac{d\phi}{\pi} \quad (34)$$

$$\text{Prob}\left(\psi_1 \leq \Phi \leq \psi_2;\overline{\gamma}_{\text{eff}}\beta^2\right) = \begin{cases} F_\Phi\left(\psi_2;\overline{\gamma}_{\text{eff}}\beta^2\right) - F_\Phi\left(\psi_1;\overline{\gamma}_{\text{eff}}\beta^2\right) + 1, & \text{if } \psi_1 < 0 < \psi_2 \\ F_\Phi\left(\psi_2;\overline{\gamma}_{\text{eff}}\beta^2\right) - F_\Phi\left(\psi_1;\overline{\gamma}_{\text{eff}}\beta^2\right), & \text{if } \psi_1 > 0 \text{ or } \psi_2 < 0 \end{cases} \quad (26)$$

$$\overline{P}_S = \sum_{l=K+1}^{L} \frac{K \times \chi(l)}{2l} \times \left\{ \operatorname{sgn}\left(\frac{\pi}{M} - \phi_\rho\right) I_K\left(\pi - \left|\frac{\pi}{M} - \phi_\rho\right|; \overline{\gamma}_{\text{eff}} \sin^2\left(\frac{\pi}{M} - \phi_\rho\right), \frac{K}{l}\overline{\gamma}_{\text{eff}} \sin^2\left(\frac{\pi}{M} - \phi_\rho\right)\right) + \operatorname{sgn}\left(\frac{\pi}{M} + \phi_\rho\right) I_K\left(\pi - \left|\frac{\pi}{M} + \phi_\rho\right|; \overline{\gamma}_{\text{eff}} \sin^2\left(\frac{\pi}{M} + \phi_\rho\right), \frac{K}{l}\overline{\gamma}_{\text{eff}} \sin^2\left(\frac{\pi}{M} + \phi_\rho\right)\right)\right\}. \quad (35)$$

where $I_n(\theta; c_1, c_2)$ is derived in closed form in [39, Appendix 5].

Upon using (31)–(34), (29) can be expressed in closed form, given in (35) at the top of the page. Equation (35) is a simple extension of the results developed in [27] for GDC and noisy CSI. The results of [27] are valid only for $\phi_\rho = 0$ (i.e., for the channel estimation schemes of Sections II-AI–III) and for $K = L$. With perfect CSI, we have $\rho = 1$ and $\phi_\rho = 0$, and (35) reduces to the well-known average SEP expression with GDC on Rayleigh fading channels [38]. As will be discussed in Section V, and illustrated in Fig. 6, by not considering the signal dependency on the noise variance, the average SEP expression of [13, eq. (17)] does not agree with (35), and is overly optimistic by not exhibiting any error floor.

1) Average BEP With Gray Mapping: We now derive the average BEP with Gray code mapping. Our approach is due to [40] (also see [28] for a correction to [40]). Similar to [28], we define $\mathcal{P}(k, m, \beta^2)$ as the probability of the received signal falling in a wedge of width $2\pi/M$ centered around the $(k+m)$th symbol point $k = 1, \ldots, M$, $k \neq m$, conditioned on β^2, when $S_m = \sqrt{E_s}\exp(j2\pi m/M)$ is the transmitted signal. That is, we have

$$\mathcal{P}(k, m, \beta^2) = \int_{\phi=\theta_m+\theta_k-\frac{\pi}{M}}^{\theta_m+\theta_k+\frac{\pi}{M}} f_\Phi\left(\phi; \overline{\gamma}_{\text{eff}}\beta^2, \theta_m + \phi_\rho\right) d\phi = \int_{\phi=\theta_k-\frac{\pi}{M}-\phi_\rho}^{\theta_k+\frac{\pi}{M}-\phi_\rho} f_\Phi(\phi; \overline{\gamma}_{\text{eff}}\beta^2, 0) d\phi \quad (36)$$

where the second equality in (36) is due to a change of integration variable. Note that, similar to the case of perfect channel knowledge, $\mathcal{P}(k, m, \beta^2)$ is not a function of the transmitted signal phase θ_m. As a result, we use $\mathcal{P}(k, \beta^2)$ instead of $\mathcal{P}(k, m, \beta^2)$. To proceed further, as done previously to arrive at (28) from (25), we employ (26) and (27), and simplify (36) to

$$\mathcal{P}(k, \beta^2) = \frac{\operatorname{sgn}\left(\theta_k - \frac{\pi}{M} - \phi_\rho\right)}{2\pi} \times \int_0^{\pi - \left|\theta_k - \frac{\pi}{M} - \phi_\rho\right|} e^{-\frac{\overline{\gamma}_{\text{eff}}\beta^2 \sin^2\left(\theta_k - \frac{\pi}{M} - \phi_\rho\right)}{\sin^2\theta}} d\theta - \frac{\operatorname{sgn}\left(\theta_k + \frac{\pi}{M} - \phi_\rho\right)}{2\pi} \times \int_0^{\pi - \left|\theta_k + \frac{\pi}{M} - \phi_\rho\right|} e^{-\frac{\overline{\gamma}_{\text{eff}}\beta^2 \sin^2\left(\theta_k + \frac{\pi}{M} - \phi_\rho\right)}{\sin^2\theta}} d\theta. \quad (37)$$

Note that (37) is valid only for $|\phi_\rho - \theta_k| > \pi/M$. The cases $|\phi_\rho - \theta_k| \leq \pi/M$ can also be treated in a similar manner. As a sanity check, with $\beta^2 = 1$, $\overline{\gamma}_{\text{eff}} = E_s/N_0$, and $\phi_\rho = 0$, (37) reduces to the expression derived in [39, eq. (8.29)] for the additive white Gaussian noise (AWGN) channel. Following the steps of (29) and (35), a closed-form expression for $\overline{\mathcal{P}}(k) = E[\mathcal{P}(k, \beta^2)]$ is given by (38) at the bottom of the page. Using (38), the average BEP for the Gray coded M-PSK signal set is

$$\overline{P}_{\text{bit},M\text{-PSK}} = \frac{1}{\log_2(M)} \sum_{k=1}^{M-1} \overline{d}(k)\overline{\mathcal{P}}(k) \quad (39)$$

where $\overline{d}(k)$ is the weight spectrum of Gray code, derived in [28], which is reproduced here:

$$\overline{d}(k) = 2\left|\frac{k}{M} - \left\lfloor\frac{k}{M}\right\rceil\right| + 2\sum_{i=2}^{\log_2(M)}\left|\frac{k}{2^i} - \left\lfloor\frac{k}{2^i}\right\rceil\right|. \quad (40)$$

In (40), $\lfloor x \rceil$ rounds x to the closest integer.

$$\overline{\mathcal{P}}(k) = E[\mathcal{P}(k, \beta^2)] = \sum_{l=K+1}^{L} \frac{K \times \chi(l)}{2l} \times \left\{ \operatorname{sgn}\left(\theta_k - \frac{\pi}{M} - \phi_\rho\right) \times I_K\left(\pi - \left|\theta_k - \frac{\pi}{M} - \phi_\rho\right|; \overline{\gamma}_{\text{eff}} \sin^2\left(\theta_k - \frac{\pi}{M} - \phi_\rho\right), \frac{K}{l}\overline{\gamma}_{\text{eff}} \sin^2\left(\theta_k - \frac{\pi}{M} - \phi_\rho\right)\right) - \operatorname{sgn}\left(\theta_k + \frac{\pi}{M} - \phi_\rho\right) \times I_K\left(\pi - \left|\theta_k + \frac{\pi}{M} - \phi_\rho\right|; \overline{\gamma}_{\text{eff}} \sin^2\left(\theta_k + \frac{\pi}{M} - \phi_\rho\right), \frac{K}{l}\overline{\gamma}_{\text{eff}} \sin^2\left(\theta_k + \frac{\pi}{M} - \phi_\rho\right)\right)\right\}. \quad (38)$$

We note that (38) and (39) extend, in a closed form, the results of [28] to the case of fading, GDC, and imperfect CSI. When $M = 2$, (39) reduces to the average BEP performance with BPSK and GDC, as reported in [6].

2) Remarks and Discussion: Recently, [31] analyzed the performance of M-PSK with MRC diversity and channel estimation errors. Specifically, for Rayleigh fading with i.i.d. branches (using [31, eqs. (4), (17), and (18)] and simplifying using our notation), the following expression for the average BEP was obtained:

$$\overline{P}_b([31]) = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \mathcal{L}_{\beta^2}\left(\frac{\overline{\gamma}\rho^2\cos^2(\phi_\rho)}{(1+\overline{\gamma}(1-\rho^2))\sin^2\theta}\right)d\theta \quad (41)$$

where $\mathcal{L}_{\beta^2}(s)$ is given in (32). We now show the limitations of (41). For simplicity, we set $\phi_\rho = \pi$. Using these parameters, (41) shows that the average BEP is unaffected by a phase rotation of π. However, with the help of (25), (26), (27), and (29), derived in this paper, the average BEP is given by

$$\begin{aligned}\overline{P}_b &= 1 - \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \mathcal{L}_{\beta^2}\left(\frac{\overline{\gamma}\rho^2}{(1+\overline{\gamma}(1-\rho^2))\sin^2\theta}\right)d\theta \\ &= 1 - \overline{P}_b([31])\Big|_{\phi_\rho=\pi} \end{aligned} \quad (42)$$

which is attributed to the fact that the decision region is flipped for bits "0" and "1" due to a phase rotation of 180°. The reason for this discrepancy is as follows: From (25), we observe that imperfect channel estimation affects a PSK system in two ways: a) the average SNR per branch $\overline{\gamma}$ is reduced to $\overline{\gamma}_{\text{eff}}$ and b) the decision region for symbol s_m shifts from $(\theta_m - \pi/M, \theta_m + \pi/M]$ to $(\theta_m - \pi/M - \phi_\rho, \theta_m + \pi/M - \phi_\rho]$, whereas the analysis of [31] did not take into account the effect of the phase offset ϕ_ρ on the demodulator's decision region.

In [41], the authors analyzed the average BEP performance of generalized hierarchical PSK constellations (i.e., embedded PSK constellations), with perfect CSI, using Pawula's F-function. By modifying Pawula's original F-function to incorporate the effects of noisy CSI, as done in this paper, we are extending the effects of channel estimation errors to the signal constellations of [41].

B. M-PAM Constellation

For an M-ary PAM constellation, S is a real-valued signal point. The mth signal point $s_m \in \mathcal{S}$ is represented as $s_m = (-M + 2m - 1)d$ for $m = 1, \ldots, M$, where $2d = \sqrt{12E_s/(M^2-1)}$ is the minimum distance between two signal points so that $E[s_m^2] = E_s$. From (12), the relevant decision statistic is the real part of r, which is given by

$$Z_{PAM} = \frac{R_c}{\sigma_p^2}s_m + \eta_1. \quad (43)$$

To proceed further, let us define the following:

$$\kappa \triangleq \frac{E\left[|p_l|^2\right]}{E\left[|g_l|^2\right]} = \frac{\sigma_p^2}{\sigma_g^2} \quad (44)$$

$$\begin{aligned}\overline{\gamma}_m &\triangleq \frac{1}{\beta^2}\frac{d^2}{\sigma^2_{|s_m|}} = \frac{2\sigma_p^2 d^2}{\sigma^2 + \sigma_g^2(1-\rho^2)|s_m|^2} \\ &= \frac{6\kappa\overline{\gamma}}{M^2 - 1 + 3(1-\rho^2)(M+1-2m)^2\overline{\gamma}} \end{aligned} \quad (45)$$

$$\kappa_c \triangleq \frac{R_c}{\sigma_p^2} = \frac{\rho_c}{\sqrt{\kappa}} \quad (46)$$

$$\text{and} \quad \kappa_{cs} \triangleq \frac{R_{cs}}{\sigma_p^2} = \frac{\rho_{cs}}{\sqrt{\kappa}} \quad (47)$$

where κ in (44) denotes the average power imbalance between the channel fade and its estimate, $\overline{\gamma}_m$ in (45) is the effective SNR due to s_m, and (45) is obtained by substituting (44), making use of the fact that $E_s = (M^2-1)d^2/3$, and the definition $\overline{\gamma} = 2\sigma_g^2 E_s/N_0$.

Due to the signal-dependent noise variance, as given by (13), to analyze the error performance of M-PAM, one has to consider each signal point separately. For signals s_1 and s_M, the probabilities of correct decision, conditioned on β^2, are given by

$$\begin{aligned}P_c(s_1,\beta^2) &= \text{Prob}\left(Z_{PAM} < -(M-2)d|\beta^2, s_1\right) \\ &= Q\left(\frac{(M-2)d + \left(R_c/\sigma_p^2\right)s_1}{\sigma_{|s_1|}}\right) \\ P_c\left(s_M,\beta^2\right) &= \text{Prob}\left(Z_{PAM} \geq (M-2)d|\beta^2, s_M\right) \\ &= Q\left(\frac{(M-2)d - \left(R_c/\sigma_p^2\right)s_M}{\sigma_{|s_M|}}\right). \end{aligned} \quad (48)$$

Note that since $s_1 = -s_M$, we have $P_c\left(s_1,\beta^2\right) = P_c\left(s_M,\beta^2\right)$. For s_m, $m \in \{2,\ldots,M-1\}$, $P_c\left(s_m,\beta^2\right)$ can be expressed as

$$\begin{aligned}P_c\left(s_m,\beta^2\right) &= \text{Prob}\left(s_m - d \leq Z_{PAM} < s_m + d|\beta^2, s_m\right) \\ &= Q\left(\frac{s_m\left(1 - R_c/\sigma_p^2\right) - d}{\sigma_{|s_m|}}\right) \\ &\quad - Q\left(\frac{s_m\left(1 - R_c/\sigma_p^2\right) + d}{\sigma_{|s_m|}}\right) \\ &= P_c\left(-s_m,\beta^2\right). \end{aligned} \quad (49)$$

Since the M-PAM signal set is symmetric about the origin, and $s_m = -s_{M+1-m}$ for $m = 1,\ldots,M/2$, the average probability of error, conditioned on β^2 and using (44)–(47) in (48) and (49), can be written as in (50) at the top of the following page.

We need the following definitions:

$$\begin{aligned}&\mathcal{H}_{\text{GDC}}(a,L,K) \\ &\quad \triangleq E\left[Q\left(a\sqrt{\beta^2}\right)\right] \\ &\quad = \begin{cases} \mathcal{G}_{\text{GDC}}(|a|,L,K), & \text{if } a > 0 \\ 1 - \mathcal{G}_{\text{GDC}}(|a|,L,K), & \text{if } a < 0 \\ \frac{1}{2}, & \text{if } a = 0 \end{cases} \end{aligned} \quad (51)$$

where, from [39], we have

$$\begin{aligned}\mathcal{G}_{\text{GDC}}(|a|,L,K) &\triangleq \int_{x=0}^{\infty} Q\left(|a|\sqrt{x}\right) f_{\beta^2}(x)dx \\ &= \binom{L}{K}\sum_{l=0}^{L-K}\frac{(-1)^l\binom{L-K}{l}}{1+\frac{l}{K}} \\ &\quad \times I_{K-1}\left(\frac{\pi}{2}; \frac{|a|^2}{2}, \frac{|a|^2/2}{1+\frac{l}{K}}\right) \end{aligned} \quad (52)$$

and where $I_n(\theta; c_1, c_2)$ is given by (34), and is derived in closed form in [39, Appendix 5A]. Using (51) and (52) to average (50) over β^2, we obtain the closed-form solution shown in (53) also

$$\begin{aligned} P_{e,M\text{-PAM}}(\beta^2) &= \frac{1}{M}\sum_{m=1}^{M} P_e\left(s_m,\beta^2\right) \\ &= \frac{1}{M}\sum_{m=1}^{M}\left(1-P_c\left(s_m,\beta^2\right)\right) \\ &= \frac{2}{M}\sum_{m=1}^{M/2}\left(1-P_c\left(s_m,\beta^2\right)\right) \\ &= \frac{2}{M}Q\left(\left[1-(M-1)(1-\kappa_c)\right]\sqrt{\overline{\gamma}_1\beta^2}\right) \\ &\quad + \frac{2}{M}\sum_{m=2}^{M/2}\Bigg\{Q\left(\left[1+(M+1-2m)(1-\kappa_c)\right]\sqrt{\overline{\gamma}_m\beta^2}\right) \\ &\qquad + Q\left(\left[1-(M+1-2m)(1-\kappa_c)\right]\sqrt{\overline{\gamma}_m\beta^2}\right)\Bigg\}. \end{aligned} \tag{50}$$

$$\begin{aligned} \overline{P}_{e,M\text{-PAM}} &= \frac{2}{M}\mathcal{H}_{\text{GDC}}\left(\left[1-(M-1)(1-\kappa_c)\right]\sqrt{\overline{\gamma}_1},L,K\right) \\ &\quad + \frac{2}{M}\sum_{m=2}^{M/2}\Bigg\{\mathcal{H}_{\text{GDC}}\left(\left[1+(M+1-2m)(1-\kappa_c)\right]\sqrt{\overline{\gamma}_m},L,K\right) \\ &\qquad + \mathcal{H}_{\text{GDC}}\left(\left[1-(M+1-2m)(1-\kappa_c)\right]\sqrt{\overline{\gamma}_m},L,K\right)\Bigg\}. \end{aligned} \tag{53}$$

at the top of the page. Note that (53) generalizes [26] to the case of GDC. Also, with $K=L$, our results are simpler than [26].

1) Average BEP With Gray Mapping: We now derive the average BEP for Gray coded M-PAM. Let $\mathcal{X}=\{0,1,\ldots,M-1\}$ denote the index set of the PAM signal points. For any $x\in\mathcal{X}$, let $(u_{k-1},u_{k-2},\ldots,u_0)$ denote the binary representation of x (i.e., $x=\sum_{j=0}^{k-1}u_j2^j$), where $k=\log_2(M)$. Let us also denote by $(a_{k-1},a_{k-2},\ldots,a_0)$, the Gray mapping of $(u_{k-1},u_{k-2},\ldots,u_0)$. For $j=0,\ldots,k-1$, let us define the following sets: $X_1(j)=\{x:(x \bmod 2^{j+2})=2^j+l,l=0,\ldots,2^j-1\}\cup\{x:(x \bmod 2^{j+2})=2^{j+1}+l,l=0,\ldots,2^j-1\}$ and $X_0(j)=\{x:(x \bmod 2^{j+2})=l,l=0,\ldots,2^j-1\}\cup\{x:(x \bmod 2^{j+2})=3\times 2^j+l,l=0,\ldots,2^j-1\}$. The sets $X_0(j)$ and $X_1(j)$, for various values of the constellation size M, were presented in [42]. For completeness, we tabulate these sets in Table I. It was shown in [42] that the decision statistic for bit a_j, $j=0,1,\ldots,k-1$, can be expressed as the following[4] disjoint union of intervals on the x-axis shown in (54) at the bottom of the page, where $\mathbf{1}_{\{\mathsf{A}\}}$ is the indicator function that evaluates to 1 when A is true. Otherwise, it evaluates to 0. As an example, consider 8-ary PAM and bit a_0. Table I gives us $X_1(0)=\{1,2,5,6\}$. With the help of Table I and (54), we can express the decision region for bit a_0 as $\widehat{a}_0=1$ if $Z_{PAM}\in[-6d,-4d)\cup[-4d,-2d)\cup[2d,4d)\cup[4d,6d)$ (i.e., $Z_{PAM}\in[-6d,-2d)\cup[2d,6d)$) and $\widehat{a}_0=0$ otherwise.

The average probability of bit error for bit a_j, conditioned on β^2, can be expressed as equation (55) at the bottom of the following page. Notice that the r.v. β^2 appears in the $Q(\cdot)$ functions of (55) only in the form of $Q(c\sqrt{\beta^2})$, where c is *real*. Using (51) to average (55) over β^2, the average probability of error for bit a_j, $\overline{P}_b(a_j)=E\left[P_b\left(a_j,\beta^2\right)\right]$, can be obtained. This task can be accomplished trivially by replacing each $Q(\cdot)$ function in (55) by $\mathcal{H}_{\text{GDC}}(\cdot,L,K)$ of (51). The resulting average BEP is obtained in closed form as shown in (56) also at the bottom of the following page. Finally, the average BEP can be obtained as

$$\overline{P}_{\text{bit},M-PAM}=\frac{1}{\log_2(M)}\sum_{j=0}^{k-1}\overline{P}_b(a_j). \tag{57}$$

Equations (56) and (57) provide a novel expression for the average BEP of M-PAM with Gray code mapping, GDC, and imperfect CSI.

[4]We note that [42] does not employ indicator functions for the decision boundaries of the end points.

$$\widehat{a}_j=\begin{cases}1, & \text{if } Z_{PAM}\in\cup_{x\in X_1(j)}\left[-\boldsymbol{\infty}\times\mathbf{1}_{\{x=0\}}+s_x-d,\boldsymbol{\infty}\times\mathbf{1}_{\{x=M-1\}}+s_x+d\right)\\ 0, & \text{otherwise}\end{cases} \tag{54}$$

TABLE I
TABULATION OF THE SETS $X_1(j)$ AND $X_0(j)$ FOR EACH VALUE OF THE PAM CONSTELLATION SIZE M. NOTE THAT FOR A GIVEN M, FOR ANY j, $X_1(j) \cup X_0(j) = \{0, 1, \ldots, M-1\}$

M	j	$X_1(j)$	$X_0(j)$
2	0	{1}	{0}
4	0	{1,2}	{0,3}
	1	{2,3}	{0,1}
8	0	{1,2,5,6}	{0,3,4,7}
	1	{2,3,4,5}	{0,1,6,7}
	2	{4,5,6,7}	{0,1,2,3}
16	0	{1,2,5,6,9,10,13,14}	{0,3,4,7,8,11,12,15}
	1	{2,3,4,5,10,11,12,13}	{0,1,6,7,8,9,14,15}
	2	{4,5,6,7,8,9,10,11}	{0,1,2,3,12,13,14,15}
	3	{8,9,10,11,12,13,14,15}	{0,1,2,3,4,5,6,7}
32	0	{1,2,5,6,9,10,13,14,17,18,21,22,25,26,29,30}	{0,3,4,7,8,11,12,15,16,19,20,23,24,27,28,31}
	1	{2,3,4,5,10,11,12,13,18,19,20,21,26,27,28,29}	{0,1,6,7,8,9,14,15,16,17,22,23,24,25,30,31}
	2	{4,5,6,7,8,9,10,11,20,21,22,23,24,25,26,27}	{0,1,2,3,12,13,14,15,16,17,18,19,28,29,30,31}
	3	{8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23}	{0,1,2,3,4,5,6,7,24,25,26,27,28,29,30,31}
	4	{16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31}	{0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15}

C. Rectangular M-QAM Constellations

From (12), we observe that the effects of channel estimation error on a QAM signal constellations are threefold: to scale the transmitted signal point by a factor of $\sqrt{R_c^2 + R_{cs}^2}/\sigma_p^2 = \rho/\sqrt{\kappa}$, to rotate the constellation by ϕ_ρ, and to add a signal-dependent noise term. We let $s_m = s_x + js_y$, $m = 0, 1, \ldots, M-1$, $x = 0, 1, \ldots, M_1 - 1$, $y = 0, 1, \ldots, M_2 - 1$, where the M-QAM constellation is of size $M = M_1 M_2$. Here $s_x = a_x d$, $s_y = a_y d$, where $a_x = -(M_1 - 1) + 2x$, $a_y = -(M_2 - 1) + 2y$, M_1 is the size of the in-phase PAM constellation, and M_2 is the size of the quadrature-phase PAM constellation. To proceed further, we define, for $x = 0, 1, \ldots, M_1 - 1$ and $y = 0, 1, \ldots, M_2 - 1$, the parameter $\overline{\gamma}_{x,y}$, shown in (58) at the bottom of the following page. In (58), we have used the fact that, for $M_1 \times M_2$-QAM, $E_s = \left(M_1^2 + M_2^2 - 2\right) d^2/3$ [43].

Let us denote by $P_{C,x,y}(\beta^2)$ the probability of correctly receiving $s_x + js_y$, conditioned on β^2. It is now straightforward to compute $P_{c,x,y}(\beta^2)$. For $x = 1, 2 \ldots, M_1 - 2$, $y = 1, 2, \ldots, M_2 - 2$, we have (59), at the bottom of the following page. For convenience, $P_{C,x,y}(\beta^2)$ for other values of x and y are expressed as (60)–(67) at the bottom of the following page, and simplified final expressions for (60)–(67) are tabulated in Table II.

Let us define by $\overline{P}_{C,x,y} \triangleq E\left[P_{C,x,y}(\beta^2)\right]$ the probability of correct reception of $s_x + js_y$, averaged over β^2. Each of the $P_{C,x,y}(\beta^2)$ expressions in Table II can be expressed as $Q(a\sqrt{\beta^2}) \times Q(b\sqrt{\beta^2})$. To derive $\overline{P}_{C,x,y}$, we need to evaluate $E[Q(a\sqrt{\beta^2}) \times Q(b\sqrt{\beta^2})]$. To this end, we define (68) at the

$$
\begin{aligned}
P_b\left(a_j, \beta^2\right) &= \frac{1}{2}\mathrm{Prob}\left(\widehat{a}_j = 1 | a_j = 0, \beta^2\right) + \frac{1}{2}\mathrm{Prob}\left(\widehat{a}_j = 0 | a_j = 1, \beta^2\right) \\
&= \frac{1}{M} \sum_{y \in X_0(j)} \sum_{x \in X_1(j)} \left\{ Q\left(\left[-\infty \times \mathbf{1}_{\{x=0\}} + 2x - M + \kappa_c(M - 1 - 2y)\right] \sqrt{\overline{\gamma}_y \beta^2}\right) \right. \\
&\qquad \left. - Q\left(\left[\infty \times \mathbf{1}_{\{x=M-1\}} + 2x + 2 - M + \kappa_c(M - 1 - 2y)\right] \sqrt{\overline{\gamma}_y \beta^2}\right) \right\} \\
&\quad + \frac{1}{M} \sum_{y \in X_1(j)} \sum_{x \in X_0(j)} \left\{ Q\left(\left[-\infty \times \mathbf{1}_{\{x=0\}} + 2x - M + \kappa_c(M - 1 - 2y)\right] \sqrt{\overline{\gamma}_y \beta^2}\right) \right. \\
&\qquad \left. - Q\left(\left[\infty \times \mathbf{1}_{\{x=M-1\}} + 2x + 2 - M + \kappa_c(M - 1 - 2y)\right] \sqrt{\overline{\gamma}_y \beta^2}\right) \right\}. \qquad (55)
\end{aligned}
$$

$$
\begin{aligned}
\overline{P}_b(a_j) = \frac{1}{M} &\sum_{y \in X_0(j)} \sum_{x \in X_1(j)} \left\{ \mathcal{H}_{\mathrm{GDC}}\left(\left[-\infty \times \mathbf{1}_{\{x=0\}} + 2x - M + \kappa_c(M - 1 - 2y)\right] \sqrt{\overline{\gamma}_y}, L, K\right) \right. \\
&\left. - \mathcal{H}_{\mathrm{GDC}}\left(\left[\infty \times \mathbf{1}_{\{x=M-1\}} + 2x + 2 - M + \kappa_c(M - 1 - 2y)\right] \sqrt{\overline{\gamma}_y}, L, K\right) \right\} \\
+ \frac{1}{M} &\sum_{y \in X_1(j)} \sum_{x \in X_0(j)} \left\{ \mathcal{H}_{\mathrm{GDC}}\left(\left[-\infty \times \mathbf{1}_{\{x=0\}} + 2x - M + \kappa_c(M - 1 - 2y)\right] \sqrt{\overline{\gamma}_y}, L, K\right) \right. \\
&\left. - \mathcal{H}_{\mathrm{GDC}}\left(\left[\infty \times \mathbf{1}_{\{x=M-1\}} + 2x + 2 - M + \kappa_c(M - 1 - 2y)\right] \sqrt{\overline{\gamma}_y}, L, K\right) \right\}. \qquad (56)
\end{aligned}
$$

bottom of the page, where $\mathcal{G}_{\mathrm{GDC}}(|a|, L, K)$ is given in (52), and $\mathcal{G}(|a|, |b|, L, K) \triangleq E[Q(|a|\sqrt{\beta^2})Q(|b|\sqrt{\beta^2})]$ is derived in closed form in (148), Appendix III.

For simplicity, let us define the following scalar variables:

$$\mathrm{t}_1(x,y) = (a_x - 1 - a_x k_c + a_y k_{cs})\sqrt{\overline{\gamma}_{x,y}} \quad (69)$$

$$sft_2(x,y) = (a_x + 1 - a_x k_c + a_y k_{cs})\sqrt{\overline{\gamma}_{x,y}} \quad (70)$$

$$\mathrm{t}_3(x,y) = (a_y - 1 - a_x k_{cs} - a_y k_c)\sqrt{\overline{\gamma}_{x,y}} \quad (71)$$

$$\mathrm{t}_4(x,y) = (a_y + 1 - a_x k_{cs} - a_y k_c)\sqrt{\overline{\gamma}_{x,y}}. \quad (72)$$

Notice that (69)–(72) appear as the arguments of $Q(\cdot)$ functions in Table II. Using (68) below and (69)–(72), each row in Table II can be averaged over β^2 to obtain closed-form expressions for $\overline{P}_{C,x,y}$, $x = 0, 1, \ldots, M_1 - 1$, $y = 0, 1, \ldots, M_2 - 1$. These expressions are tabulated in Table III. Using them, the average SEP can be written as

$$\overline{P}_{e,M-QAM} = \frac{1}{M}\sum_{x=0}^{M_1-1}\sum_{y=0}^{M_2-1}(1 - \overline{P}_{C,x,y}) = 1 - \frac{1}{M}\sum_{x=0}^{M_1-1}\sum_{y=0}^{M_2-1}\overline{P}_{C,x,y}. \quad (73)$$

It can be numerically shown (see Fig. 7 and the discussion in Section V) that this equation does not agree with the average SEP expression of [13, eq. (19)].

1) Average BEP With Gray Mapping: Similar to the sets $X_1(j)$ and $X_0(j)$, $j = 0, \ldots, k-1$, $k = \log_2(M)$, as in Section III-BI, we now introduce the following sets. We define $k_1 = \log_2(M_1)$, $k_2 = \log_2(M_2)$, and the sets $\mathcal{X} = \{0, 1, \ldots, M_1 - 1\}$ and $\mathcal{Y} = \{0, 1, \ldots, M_2 - 1\}$. The vector $(a_{k_1-1}, a_{k_1-2}, \ldots, a_0)$ is the Gray code mapping for the in-phase signal s_x, and $(b_{k_2-1}, b_{k_2-2}, \ldots, b_0)$

$$\overline{\gamma}_{x,y} \triangleq \frac{1}{\beta^2}\frac{d^2}{\sigma^2_{|s_m|}} = \frac{d^2 2\sigma_p^2}{\sigma^2 + \sigma_e^2|s_m|^2} = \frac{6\frac{\sigma_p^2}{\sigma_g^2}2\sigma_g^2\frac{E_s}{N_0}}{M_1^2 + M_2^2 - 2 + 3(1-\rho^2)2\sigma_g^2\frac{E_s}{N_0}\left((M_1 - 1 - 2x)^2 + (M_2 - 1 - 2y)^2\right)} = \frac{6\kappa\overline{\gamma}}{M_1^2 + M_2^2 - 2 + 3(1-\rho^2)\overline{\gamma}\left((M_1 - 1 - 2x)^2 + (M_2 - 1 - 2y)^2\right)}. \quad (58)$$

$$\begin{aligned} P_{C,x,y}(\beta^2) &= \mathrm{Prob}\left(s_x - d \le Z_I < s_x + d|\beta^2\right) \times \mathrm{Prob}\left(s_y - d \le Z_Q < s_y + d|\beta^2\right) \\ &= \left\{Q\left(\left[a_x - 1 - a_x\kappa_c + a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right) - Q\left(\left[a_x + 1 - a_x\kappa_c + a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\right\} \\ &\quad \times \left\{Q\left(\left[a_y - 1 - a_x\kappa_{cs} - a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right) - Q\left(\left[a_y + 1 - a_x\kappa_{cs} - a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\right\}. \end{aligned} \quad (59)$$

$$\begin{aligned} P_{C,x,y}(\beta^2) = {} & \mathrm{Prob}\left(s_x - d \le Z_I < s_x + d|\beta^2\right) \times \mathrm{Prob}\left(-\infty < Z_Q < s_y + d|\beta^2\right) \quad x = 1, 2 \ldots, M_1 - 2,\ y = 0 && (60) \\ & \mathrm{Prob}\left(s_x - d \le Z_I < s_x + d|\beta^2\right) \times \mathrm{Prob}\left(s_y - d \le Z_Q < \infty|\beta^2\right) \quad x = 1, 2 \ldots, M_1 - 2,\ y = M_2 - 1 && (61) \\ & \mathrm{Prob}\left(-\infty < Z_I < s_x + d|\beta^2\right) \times \mathrm{Prob}\left(s_y - d \le Z_Q < s_y + d|\beta^2\right) \quad x = 0,\ y = 1, 2 \ldots, M_2 - 2 && (62) \\ & \mathrm{Prob}\left(s_x - d \le Z_I < \infty|\beta^2\right) \times \mathrm{Prob}\left(s_y - d \le Z_Q < s_y + d|\beta^2\right) \quad x = M_1 - 1,\ y = 1, 2 \ldots, M_2 - 2 && (63) \\ & \mathrm{Prob}\left(-\infty < Z_I < s_x + d|\beta^2\right) \times \mathrm{Prob}\left(-\infty < Z_Q < s_y + d|\beta^2\right) \quad x = 0,\ y = 0 && (64) \\ & \mathrm{Prob}\left(s_x - d \le Z_I < \infty|\beta^2\right) \times \mathrm{Prob}\left(-\infty < Z_Q < s_y + d|\beta^2\right) \quad x = M_1 - 1,\ y = 0 && (65) \\ & \mathrm{Prob}\left(-\infty < Z_I < s_x + d|\beta^2\right) \times \mathrm{Prob}\left(s_y - d \le Z_Q < \infty|\beta^2\right) \quad x = 0,\ y = M_2 - 1 && (66) \\ & \mathrm{Prob}\left(s_x - d \le Z_I < \infty|\beta^2\right) \times \mathrm{Prob}\left(s_y - d \le Z_Q < \infty|\beta^2\right) \quad x = M_1 - 1,\ y = M_2 - 1. && (67) \end{aligned}$$

$$\begin{aligned} \mathcal{H}(a,b,L,K) &\triangleq E\left[Q\left(a\sqrt{\beta^2}\right)Q\left(b\sqrt{\beta^2}\right)\right] \\ &= \begin{cases} \mathcal{G}(|a|,|b|,L,K) & (a>0, b>0) \\ \mathcal{G}_{\mathrm{GDC}}(|a|,L,K) - \mathcal{G}(|a|,|b|,L,K) & (a>0, b<0) \\ \mathcal{G}_{\mathrm{GDC}}(|b|,L,K) - \mathcal{G}(|a|,|b|,L,K) & (a<0, b>0) \\ 1 - \mathcal{G}_{\mathrm{GDC}}(|a|,L,K) - \mathcal{G}_{\mathrm{GDC}}(|b|,L,K) + \mathcal{G}(|a|,|b|,L,K) & (a<0, b<0) \\ (1/2)\mathcal{H}_{\mathrm{GDC}}(b,L,K) & (a=0, b\neq 0) \\ (1/2)\mathcal{H}_{\mathrm{GDC}}(a,L,K) & (b=0, a\neq 0) \\ \frac{1}{4} & (a=0, b=0) \end{cases} \end{aligned} \quad (68)$$

TABLE II
FOR EACH $x \in \{0,1,\ldots,M_1-1\}$ AND $y \in \{0,1,\ldots,M_2-1\}$, CONDITIONED ON β^2, THE PROBABILITY OF CORRECT RECEPTION OF THE SYMBOL $s_x + j s_y$ IS THE THIRD COLUMN FOR AN $M_1 \times M_2$ RECTANGULAR QAM CONSTELLATION

x	y	$P_{C,x,y}(\beta^2) = \text{Prob}\left(s_x + js_y \text{ received successfully} \middle| \beta^2\right)$
$\{1,2,\ldots,M_1-2\}$	$\{1,2,\ldots,M_2-2\}$	$\left\{Q\left(\left[a_x-1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)-Q\left(\left[a_x+1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\right\}\times$ $\left\{Q\left(\left[a_y-1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)-Q\left(\left[a_y+1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\right\}$
$\{1,2,\ldots,M_1-2\}$	$\{0\}$	$\left\{Q\left(\left[a_x-1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)-Q\left(\left[a_x+1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\right\}\times$ $Q\left(-\left[a_y+1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)$
$\{1,2,\ldots,M_1-2\}$	$\{M_2-1\}$	$\left\{Q\left(\left[a_x-1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)-Q\left(\left[a_x+1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\right\}\times$ $Q\left(\left[a_y-1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)$
$\{0\}$	$\{1,2,\ldots,M_2-2\}$	$Q\left(-\left[a_x+1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\times$ $\left\{Q\left(\left[a_y-1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)-Q\left(\left[a_y+1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\right\}$
$\{M_1-1\}$	$\{1,2,\ldots,M_2-2\}$	$Q\left(\left[a_x-1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\times$ $\left\{Q\left(\left[a_y-1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)-Q\left(\left[a_y+1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)\right\}$
$\{0\}$	$\{0\}$	$Q\left(-\left[a_x+1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)Q\left(-\left[a_y+1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)$
$\{M_1-1\}$	$\{0\}$	$Q\left(\left[a_x-1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)Q\left(-\left[a_y+1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)$
$\{0\}$	$\{M_2-1\}$	$Q\left(-\left[a_x+1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)Q\left(\left[a_y-1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)$
$\{M_1-1\}$	$\{M_2-1\}$	$Q\left(\left[a_x-1-a_x\kappa_c+a_y\kappa_{cs}\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)Q\left(\left[a_y-1-a_x\kappa_{cs}-a_y\kappa_c\right]\sqrt{\overline{\gamma}_{x,y}\beta^2}\right)$

TABLE III
FOR EACH $x \in \{0,1,\ldots,M_1-1\}$ AND $y \in \{0,1,\ldots,M_2-1\}$ THE AVERAGE PROBABILITY OF CORRECT RECEPTION OF THE SYMBOL $s_x + j s_y$ IS THE THIRD COLUMN FOR AN $M_1 \times M_2$ RECTANGULAR QAM CONSTELLATION. THE FUNCTIONS $\mathsf{t}_1(x,y), \mathsf{t}_2(x,y), \mathsf{t}_3(x,y), \mathsf{t}_4(x,y)$ ARE DEFINED IN (69)–(72), RESPECTIVELY. THE FUNCTION $\mathcal{H}_{\text{GDC}}(a,L,K)$ IS DEFINED IN (51), WHEREAS THE FUNCTION $\mathcal{H}(a,b,L,K)$ IS DEFINED IN (68)

x	y	$\overline{P}_{C,x,y} = E\left[P_{C,x,y}(\beta^2)\right]$
$\{1,2,\ldots,M_1-2\}$	$\{1,2,\ldots,M_2-2\}$	$\mathcal{H}(\mathsf{t}_1(x,y),\mathsf{t}_3(x,y),L,K) - \mathcal{H}(\mathsf{t}_1(x,y),\mathsf{t}_4(x,y),L,K) -$ $\mathcal{H}(\mathsf{t}_2(x,y),\mathsf{t}_3(x,y),L,K) + \mathcal{H}(\mathsf{t}_2(x,y),\mathsf{t}_4(x,y),L,K)$
$\{1,2,\ldots,M_1-2\}$	$\{0\}$	$\mathcal{H}_{\text{GDC}}(\mathsf{t}_1(x,y),L,K) - \mathcal{H}(\mathsf{t}_1(x,y),\mathsf{t}_4(x,y),L,K) -$ $\mathcal{H}_{\text{GDC}}(\mathsf{t}_2(x,y),L,K) + \mathcal{H}(\mathsf{t}_2(x,y),\mathsf{t}_4(x,y),L,K)$
$\{1,2,\ldots,M_1-2\}$	$\{M_2-1\}$	$\mathcal{H}(\mathsf{t}_1(x,y),\mathsf{t}_3(x,y),L,K) - \mathcal{H}(\mathsf{t}_2(x,y),\mathsf{t}_3(x,y),L,K)$
$\{0\}$	$\{1,2,\ldots,M_2-2\}$	$\mathcal{H}_{\text{GDC}}(\mathsf{t}_3(x,y),L,K) - \mathcal{H}_{\text{GDC}}(\mathsf{t}_4(x,y),L,K) -$ $\mathcal{H}(\mathsf{t}_2(x,y),\mathsf{t}_3(x,y),L,K) + \mathcal{H}(\mathsf{t}_2(x,y),\mathsf{t}_4(x,y),L,K)$
$\{M_1-1\}$	$\{1,2,\ldots,M_2-2\}$	$\mathcal{H}(\mathsf{t}_1(x,y),\mathsf{t}_3(x,y),L,K) - \mathcal{H}(\mathsf{t}_1(x,y),\mathsf{t}_4(x,y),L,K)$
$\{0\}$	$\{0\}$	$1 - \mathcal{H}_{\text{GDC}}(\mathsf{t}_4(x,y),L,K) -$ $\mathcal{H}_{\text{GDC}}(\mathsf{t}_2(x,y),L,K) + \mathcal{H}(\mathsf{t}_2(x,y),\mathsf{t}_4(x,y),L,K)$
$\{M_1-1\}$	$\{0\}$	$\mathcal{H}_{\text{GDC}}(\mathsf{t}_1(x,y),L,K) - \mathcal{H}(\mathsf{t}_1(x,y),\mathsf{t}_4(x,y),L,K)$
$\{0\}$	$\{M_2-1\}$	$\mathcal{H}_{\text{GDC}}(\mathsf{t}_3(x,y),L,K) - \mathcal{H}(\mathsf{t}_2(x,y),\mathsf{t}_3(x,y),L,K)$
$\{M_1-1\}$	$\{M_2-1\}$	$\mathcal{H}(\mathsf{t}_1(x,y),\mathsf{t}_3(x,y),L,K)$

is the Gray code mapping for the quadrature-phase signal s_y. For $i = 0,\ldots,k_1-1$, let us define the following sets: $X_1(i) = \{x : (x \bmod 2^{i+2}) = 2^i + l, l = 0,\ldots,2^i-1\} \cup \{x : (x \bmod 2^{i+2}) = 2^{i+1} + l, l = 0,\ldots,2^i-1\}$ and $X_0(i) = \{x : (x \bmod 2^{i+2}) = l, l = 0,\ldots,2^i-1\} \cup \{x : (x \bmod 2^{i+2}) = 3\times 2^i + l, l = 0,\ldots,2^i-1\}$. For $j = 0,\ldots,k_2-1$, let us define the following sets: $Y_1(j) = \{y : (y \bmod 2^{j+2}) = 2^j + l, l = 0,\ldots,2^j-1\} \cup \{y : (y \bmod 2^{j+2}) = 2^{j+1} + l, l = 0,\ldots,2^j-1\}$ and $Y_0(j) = \{y : (y \bmod 2^{j+2}) = l, l = 0,\ldots,2^j-1\} \cup \{y : (y \bmod 2^{j+2}) = 3\times 2^j + l, l = 0,\ldots,2^j-1\}$. Using these sets, the decision statistic for each bit a_i, $i = 0,\ldots,k_1-1$, is given by the following disjoint union of intervals on the x-axis in (74) at the bottom of the page; whereas for bit b_j, $j = 0,\ldots,k_2-1$, it is given by (75) also at the bottom of the page. Following the steps of (55) and (56), we obtain closed-form expressions for the average probability of bit error $\overline{P}_b(a_j) = E\left[P_b\left(a_j,\beta^2\right)\right]$, $j=0,\ldots,k_1-1$, and $\overline{P}_b(b_j) = E\left[P_b\left(b_j,\beta^2\right)\right]$, $j=0,\ldots,k_2-1$,

$$\widehat{a}_i = \begin{cases} 1, & \text{if } z_I \in \cup_{x\in X_1(i)}\left[-\infty\times\mathbf{1}_{\{x=0\}} + s_x - d, \infty\times\mathbf{1}_{\{x=M_1-1\}} + s_x + d\right) \\ 0, & \text{otherwise} \end{cases} \tag{74}$$

$$\widehat{b}_j = \begin{cases} 1, & \text{if } z_Q \in \cup_{y\in Y_1(j)}\left[-\infty\times\mathbf{1}_{\{y=0\}} + s_y - d, \infty\times\mathbf{1}_{\{y=M_2-1\}} + s_y + d\right) \\ 0, & \text{otherwise.} \end{cases} \tag{75}$$

$$\overline{P}_b(a_j) = \frac{1}{M} \sum_{x_0 \in X_0(j)} \sum_{x_1 \in X_1(j)} \sum_{y \in \mathcal{Y}} \Bigg\{ \mathcal{H}_{\mathrm{GDC}} \Bigg(\Bigg[-\infty \times \mathbf{1}_{\{x_1=0\}} + 2x_1 - M_1 + \kappa_c(M_1 - 1 - 2x_0) - \kappa_{cs}(M_2 - 1 - 2y) \Bigg] \sqrt{\overline{\gamma}_{x_0,y}}, L, K \Bigg) - \mathcal{H}_{\mathrm{GDC}} \Bigg(\Bigg[\infty \times \mathbf{1}_{\{x_1=M_1-1\}} + 2x_1 + 2 - M_1 + \kappa_c(M_1 - 1 - 2x_0) - \kappa_{cs}(M_2 - 1 - 2y) \Bigg] \sqrt{\overline{\gamma}_{x_0,y}}, L, K \Bigg) \Bigg\} + \frac{1}{M} \sum_{x_1 \in X_1(j)} \sum_{x_0 \in X_0(j)} \sum_{y \in \mathcal{Y}} \Bigg\{ \mathcal{H}_{\mathrm{GDC}} \Bigg(\Bigg[-\infty \times \mathbf{1}_{\{x_0=0\}} + 2x_0 - M_1 + \kappa_c(M_1 - 1 - 2x_1) - \kappa_{cs}(M_2 - 1 - 2y) \Bigg] \sqrt{\overline{\gamma}_{x_1,y}}, L, K \Bigg) - \mathcal{H}_{\mathrm{GDC}} \Bigg(\Bigg[\infty \times \mathbf{1}_{\{x_0=M_1-1\}} + 2x_0 + 2 - M_1 + \kappa_c(M_1 - 1 - 2x_1) - \kappa_{cs}(M_2 - 1 - 2y) \Bigg] \sqrt{\overline{\gamma}_{x_1,y}}, L, K \Bigg) \Bigg\} \tag{76}$$

$$\overline{P}_b(b_j) = \frac{1}{M} \sum_{y_0 \in Y_0(j)} \sum_{y_1 \in Y_1(j)} \sum_{x \in \mathcal{X}} \Bigg\{ \mathcal{H}_{\mathrm{GDC}} \Bigg(\Bigg[-\infty \times \mathbf{1}_{\{y_1=0\}} + 2y_1 - M_2 + \kappa_{cs}(M_1 - 1 - 2x) + \kappa_c(M_2 - 1 - 2y_0) \Bigg] \sqrt{\overline{\gamma}_{x,y_0}}, L, K \Bigg) - \mathcal{H}_{\mathrm{GDC}} \Bigg(\Bigg[\infty \times \mathbf{1}_{\{y_1=M_2-1\}} + 2y_1 + 2 - M_2 + \kappa_{cs}(M_1 - 1 - 2x) + \kappa_c(M_2 - 1 - 2y_0) \Bigg] \sqrt{\overline{\gamma}_{x,y_0}}, L, K \Bigg) \Bigg\} + \frac{1}{M} \sum_{y_1 \in Y_1(j)} \sum_{y_0 \in Y_0(j)} \sum_{x \in \mathcal{X}} \Bigg\{ \mathcal{H}_{\mathrm{GDC}} \Bigg(\Bigg[-\infty \times \mathbf{1}_{\{y_0=0\}} + 2y_0 - M_2 + \kappa_{cs}(M_1 - 1 - 2x) + \kappa_c(M_2 - 1 - 2y_1) \Bigg] \sqrt{\overline{\gamma}_{x,y_1}}, L, K \Bigg) - \mathcal{H}_{\mathrm{GDC}} \Bigg(\Bigg[\infty \times \mathbf{1}_{\{y_0=M_2-1\}} + 2y_0 + 2 - M_2 + \kappa_{cs}(M_1 - 1 - 2x) + \kappa_c(M_2 - 1 - 2y_1) \Bigg] \sqrt{\overline{\gamma}_{x,y_1}}, L, K \Bigg) \Bigg\}. \tag{77}$$

as shown in (76) at the top of the page, and in (77) also at the top of the page. Finally, the average BEP can be obtained as

$$\overline{P}_{\mathrm{bit},M-QAM} = \frac{\sum_{j=0}^{k_1-1} \overline{P}_b(a_j) + \sum_{j=0}^{k_2-1} \overline{P}_b(b_j)}{\log_2(M)}. \tag{78}$$

2) Remarks and Discussion: Recently, [20] presented an analysis of BEP for 16-QAM with MRC diversity and estimation errors. Unfortunately, the results are not in closed form, and a 2-D numerical integration is needed to evaluate the average BEP [20, eqs. (35) and (37)]. A simple closed-form solution for [20], involving no numerical integration, was reported in [44]. Note that the results of [20] are valid only for 16-QAM, whereas using (76)–(78) derived here, one can obtain a simple closed-form expression valid for arbitrary rectangular QAM constellations with GDC and estimation errors. The average BEP expressions for M-QAM in [22], which are based on

Gaussian quadratic forms [21, Appendix B], are valid only for MRC, and the methodology in [22] does not appear to be extendable to GDC, whereas (76)–(78) are valid for arbitrary GDC(L, K).

In [45], the authors present expressions for the exact BEP of hierarchical QAM constellations (i.e., embedded QAM constellations) on fading channels with perfect CSI. We are currently investigating the impact of GDC and channel estimation errors on the average BEP and SEP performances of the embedded constellations of [45].

D. Arbitrary Two-Dimensional Constellations

When s_m belongs to an arbitrary 2-D constellation, we rewrite (12) as

$$r = s_m + \underbrace{s_m\left(\frac{R_c + jR_{cs}}{\sigma_p^2} - 1\right) + \eta_1 + j\eta_2}_{\triangleq \mathcal{W} = W_1 + jW_2} \tag{79}$$

where, conditioned on β^2 and s_m, $\mathcal{W}$ is a complex Gaussian r.v. with the conditional mean

$$\mu_I(s_m) + j\mu_Q(s_m) = s_m\left(\frac{R_c + jR_{cs}}{\sigma_p^2} - 1\right)$$

and its conditional variance-per-dimension given by (13). The joint pdf of (W_1, W_2), conditioned on s_m, in polar coordinates is

$$f_{W_1,W_2|s_m}(R\cos\theta, R\sin\theta|s_m) = \frac{e^{-\frac{|R\cos\theta + jR\sin\theta - \mu_I(s_m) - j\mu_Q(s_m)|^2}{2\sigma^2_{|s_m|}}}}{2\pi\sigma^2_{|s_m|}}. \tag{80}$$

We now express $\sigma^2_{|s_m|}$ of (13) as $\sigma^2_{|s_m|} = \Delta_{|s_m|}/\beta^2$, where

$$\Delta_{|s_m|} = \frac{|s_m|^2\sigma_e^2 + \sigma^2}{2\sigma_p^2}. \tag{81}$$

For 2-D constellations having polygonal decision regions, the probability of error for the jth decision boundary when s_m is the transmitted signal is given by the joint pdf of the equivalent noise that is superimposed on s_m evaluated for that decision region [36]. The error probability over the pth subregion $P_{e,s_m,p}(\beta^2)$, can be expressed as [36]

$$\begin{aligned} P_{e,s_m,p}(\beta^2) &= \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} \int_{R=A_p(\theta)}^{\infty} RdRd\theta \\ &\quad \times f_{W_1,W_2|s_m}(R\cos\theta, R\sin\theta|s_m) \\ &= \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} \int_{R=A_p(\theta)}^{\infty} \frac{R\beta^2}{2\pi\Delta_{|s_m|}} dRd\theta \\ &\quad \times e^{-\frac{\beta^2|R\cos\theta + jR\sin\theta - \mu_I(s_m) - j\mu_Q(s_m)|^2}{2\Delta_{|s_m|}}}. \end{aligned} \tag{82}$$

In (82), the amplitude parameter $A_p(\theta)$ is defined as $A_p(\theta) = \frac{A_j\sin(\psi_j)}{\sin(|\theta - \theta_{L,p}| + \psi_j)}$ [36], and the variables $\theta_{L,p}$, $\theta_{U,p}$, A_j, and ψ_j are the constellation parameters for the pth subregion [36]. As an example, for 16-star-QAM, these parameters can be found in [46].

In order to average (82) over β^2, we note that

$$E[\beta^2\exp(-a\beta^2)] = -d/ds\mathcal{L}_{\beta^2}(s)|_{s=a}.$$

With this, the expectation of (82) over β^2 yields

$$\begin{aligned} \overline{P}_{e,s_m,p} &= \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} \int_{R=A_p(\theta)}^{\infty} \frac{R}{2\pi\Delta_{|s_m|}} dRd\theta \\ &\quad \times -\frac{d}{ds}\mathcal{L}_{\beta^2}(s)\Big|_{s=\frac{|R\cos\theta + jR\sin\theta - \mu_I(s_m) - j\mu_Q(s_m)|^2}{2\Delta_{|s_m|}}}. \end{aligned} \tag{83}$$

Using (31), the derivative of $\mathcal{L}_{\beta^2}(s)$ can be obtained as

$$\begin{aligned} -\frac{d}{ds}\mathcal{L}_{\beta^2}(s) &= \frac{K}{(1+s)^{K+1}} \prod_{l=K+1}^{L} \frac{1}{1+sK/l} \\ &\quad + \frac{1}{(1+s)^K} \sum_{l_1=K+1}^{L} \frac{K}{l_1} \frac{1}{(1+sK/l_1)^2} \\ &\quad \times \prod_{l_2=K+1, l_2\neq l_1}^{L} \frac{1}{1+sK/l_2}. \end{aligned} \tag{84}$$

Invoking the partial fractions method, (84) can be simplified as

$$\begin{aligned} -\frac{d}{ds}\mathcal{L}_{\beta^2}(s) &= \frac{K}{(s+1)^{K+1}} \sum_{l=K+1}^{L} \frac{\lambda(l)}{s+\frac{l}{K}} \\ &\quad + \frac{1}{(s+1)^K} \sum_{l_1=K+1}^{L} \frac{l_1}{K} \times \frac{1}{(s+l_1/K)^2} \\ &\quad \times \sum_{l_2=K+1, l_2\neq l_1}^{L} \frac{\zeta(l_1, l_2)}{s+\frac{l_2}{K}} \end{aligned} \tag{85}$$

where

$$\lambda(l) = \frac{K}{l} \prod_{m=K+1, m\neq l}^{L} \frac{m}{m-l} \tag{86}$$

$$\text{and} \quad \zeta(l_1, l_2) = \frac{K}{l_2} \prod_{m=K+1, m\neq l_1, m\neq l_2}^{L} \frac{m}{m-l_2}. \tag{87}$$

To proceed further, let us define the following functions:

$$\begin{aligned} \mathcal{I}(x, a, b, c, n) &\triangleq \frac{1}{(x^2 - ax + b)^n} \\ &\quad \times \frac{1}{x^2 - ax + b + c} \end{aligned} \tag{88}$$

$$\text{and} \quad \mathcal{J}(x, a, b, n) \triangleq \frac{1}{(x^2 - ax + b)^n}. \tag{89}$$

We note that, throughout this section, n is assumed to be a positive integer. Using the identity

$$\begin{aligned}\mathcal{I}(x,a,b,c,n) &= \frac{1}{c}\left\{\frac{1}{(x^2-ax+b)^n} - \frac{1}{(x^2-ax+b)^{n-1}(x^2-ax+b+c)}\right\} \\ &= \frac{1}{c}\left[\mathcal{J}(x,a,b,n) - \mathcal{I}(x,a,b,c,n-1)\right] \end{aligned} \quad (90)$$

recursively, (88) can be simplified to

$$\begin{aligned}\mathcal{I}(x,a,b,c,n) &= \sum_{l=0}^{n-1}\frac{(-1)^l}{c^{l+1}}\mathcal{J}(x,a,b,n-l) + \frac{(-1)^n}{c^n}\mathcal{J}(x,a,b+c,1) \\ &= \sum_{l=0}^{n-2}\frac{(-1)^l}{c^{l+1}}\mathcal{J}(x,a,b,n-l) + \frac{(-1)^{n+1}}{c^{n-1}}\mathcal{J}_1(x,a,b,c), \end{aligned} \quad (91)$$

where

$$\mathcal{J}_1(x,a,b,c) \triangleq \frac{1}{(x^2-ax+b)(x^2-ax+b+c)}. \quad (92)$$

Assuming $4b-a^2>0$, $n>1$ and $c>0$, we now define the following integrals:

$$\begin{aligned}\mathcal{K}(a,b,n,R) &\triangleq \int_R^\infty x\mathcal{J}(x,a,b,n)dx \\ &= \int_R^\infty \frac{xdx}{(x^2-ax+b)^n}, \end{aligned} \quad (93)$$

$$\begin{aligned}\mathcal{K}_1(a,b,c,R) &\triangleq \int_R^\infty x\mathcal{J}_1(x,a,b,c)dx \\ &= \int_R^\infty \frac{xdx}{(x^2-ax+b)(x^2-ax+b+c)}. \end{aligned} \quad (94)$$

In Appendix IV, we derive the following expressions for $\mathcal{K}(a,b,n,R)$ and $\mathcal{K}_1(a,b,c,R)$:

$$\begin{aligned}&\mathcal{K}(a,b,n,R) \\ &= \frac{a(\sqrt{4b-a^2})^{1-2n}}{2^{3-2n}} \times \left\{\mathsf{Beta}\left(\frac{1}{2},n-\frac{1}{2}\right) - \mathsf{Beta}_{\text{inc}}\left(\frac{(2R-a)^2}{(2R-a)^2+4b-a^2},\frac{1}{2},n-\frac{1}{2}\right)\right\} \\ &\quad + \frac{1}{2n-2}\times\frac{4^{n-1}}{[(2R-a)^2+4b-a^2]^{n-1}} \end{aligned} \quad (95)$$

and

$$\begin{aligned}&\mathcal{K}_1(a,b,c,R) \\ &= \frac{a\pi}{c\sqrt{4b-a^2}} \times I_0\left(\frac{\pi}{2}-\tan^{-1}\left(\frac{2R-a}{\sqrt{4b-a^2}}\right);0,\frac{4b-a^2}{4c}\right) + \frac{1}{2c}\log\left(1+\frac{4c}{4b-a^2+(2R-a)^2}\right). \end{aligned} \quad (96)$$

In (95)

$$\mathsf{Beta}_{\text{inc}}(x,m,n) = \int_{u=0}^{x} u^{m-1}(1-u)^{n-1}du$$

is the incomplete beta integral [47], and $\mathsf{Beta}(m,n) = \mathsf{Beta}_{\text{inc}}(\infty,m,n)$ is the complete beta function [47]. In (96), $I_0(\cdot;\cdot,\cdot)$ is defined in (34). Similar to (88), let us define the following functions:

$$\begin{aligned}\mathcal{Q}_1(x,a,b,c,d,n) &\triangleq \frac{1}{(x^2-ax+b)^n} \times \frac{1}{(x^2-ax+b+c)} \times \frac{1}{(x^2-ax+b+d)^2} \end{aligned} \quad (97)$$

$$\begin{aligned}\text{and } \mathcal{Q}_2(x,a,b,d,n) &\triangleq \frac{1}{(x^2-ax+b)^n} \times \frac{1}{(x^2-ax+b+d)^2}. \end{aligned} \quad (98)$$

Using (88) and (98), we can write (97) as

$$\mathcal{Q}_1(x,a,b,c,d,n) = \frac{[\mathcal{I}(x,a,b,c,n)-\mathcal{I}(x,a,b,d,n)]}{(d-c)^2} - \frac{\mathcal{Q}_2(x,a,b,d,n)}{d-c}. \quad (99)$$

Similar to (93), let us define the three integrals in (100)–(102) at the top of the following page. The last equality in (102) is due to the relationship between $\mathcal{I}(\cdot,\cdot,\cdot,\cdot,\cdot)$, $\mathcal{J}(\cdot,\cdot,\cdot,\cdot)$, and $\mathcal{J}_1(\cdot,\cdot,\cdot,\cdot)$, as given in (91), and then using (93) and (94). Since $\mathcal{K}(\cdot,\cdot,\cdot,\cdot)$ has a closed-form solution, as given in (95), and $\mathcal{K}_1(\cdot,\cdot,\cdot,\cdot)$ has a closed-form solution, as given in (96), we can evaluate (102) in closed form.

Observe that $\mathcal{C}_2(a,b,c,n,R) = -\frac{\partial}{\partial c}\mathcal{C}_3(a,b,c,n,R)$. As a result, we can express (100) only in terms of $\mathcal{C}_3(a,b,c,n,R)$ as

$$\mathcal{C}_1(a,b,c,d,n,R) = \frac{\mathcal{C}_3(a,b,c,n,R)-\mathcal{C}_3(a,b,d,n,R)}{(d-c)^2} + \frac{1}{d-c}\times\frac{\partial}{\partial d}\mathcal{C}_3(a,b,d,n,R). \quad (103)$$

Since $\mathcal{C}_1(\cdot,\cdot,\cdot,\cdot,\cdot,\cdot)$ is only a function of $\mathcal{K}(\cdot,\cdot,\cdot,\cdot)$ and $\mathcal{K}_1(\cdot,\cdot,\cdot,\cdot)$, via $\mathcal{C}_3(\cdot,\cdot,\cdot,\cdot,\cdot)$, (103) can also be evaluated in closed form.

To proceed for the final derivation of symbol error rates, let us now define the integral in (104) at the top of the following page, where

$$s_0 = (|R\cos\theta + jR\sin\theta - \mu_I(s_m) - j\mu_Q(s_m)|^2)/(2\Delta_{|s_m|})$$

$$|\mu_{s_m}| = \sqrt{\mu_I^2(s_m)+\mu_Q^2(s_m)}$$

$$\mathcal{C}_1(a,b,c,d,n,R) \triangleq \int_R^\infty x\mathcal{Q}_1(x,a,b,c,d,n)dx$$
$$= \int_R^\infty \frac{xdx}{(x^2-ax+b)^n(x^2-ax+b+c)(x^2-ax+b+d)^2} \tag{100}$$

$$\mathcal{C}_2(a,b,d,n,R) \triangleq \int_R^\infty x\mathcal{Q}_2(x,a,b,d,n)dx = \int_R^\infty \frac{xdx}{(x^2-ax+b)^n(x^2-ax+b+d)^2} \tag{101}$$

$$\mathcal{C}_3(a,b,c,n,R) \triangleq \int_R^\infty x\mathcal{I}(x,a,b,c,n)dx = \int_R^\infty \frac{xdx}{(x^2-ax+b)^n(x^2-ax+b+c)}$$
$$= \sum_{l=0}^{n-2} \frac{(-1)^l}{c^{l+1}}\mathcal{K}(a,b,n-l,R) + \frac{(-1)^{n+1}}{c^{n-1}}\mathcal{K}_1(a,b,c,R), \tag{102}$$

$$\mathcal{N}_1(a,n,\theta_{L,p},\theta_{U,p},A_p(\theta),\mu_I(s_m),\mu_Q(s_m),\Delta_{|s_m|})$$
$$\triangleq \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} \frac{d\theta}{2\pi\Delta_{|s_m|}} \int_{R=A_p(\theta)}^{\infty} \frac{R\,dR}{(s+1)^n(s+a)}\bigg|_{s=s_0}$$
$$= \frac{(2\Delta_{|s_m|})^n}{\pi} \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} d\theta \int_{R=A_p(\theta)}^{\infty} RdR$$
$$\times \frac{1}{\left(R^2-2R|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right)+|\mu_{s_m}|^2+2\Delta_{|s_m|}\right)^n}$$
$$\times \frac{1}{\left(R^2-2R|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right)+|\mu_{s_m}|^2+2a\Delta_{|s_m|}\right)} \tag{104}$$

and

$$\phi_{\mu_{s_m}} = \tan^{-1}\left(\mu_Q(s_m)/\mu_I(s_m)\right).$$

Using (88), (89), and (93), (104) can be written as

$$\mathcal{N}_1(a,n,\theta_{L,p},\theta_{U,p},A_p(\theta),\mu_I(s_m),\mu_Q(s_m),\Delta_{|s_m|})$$
$$= \frac{(2\Delta_{|s_m|})^n}{\pi} \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} d\theta \int_{R=A_p(\theta)}^{\infty} RdR$$
$$\times \mathcal{I}\left(R, 2|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right), |\mu_{s_m}|^2 + 2\Delta_{|s_m|}, 2(a-1)\Delta_{|s_m|}, n\right)$$
$$= \frac{(2\Delta_{|s_m|})^n}{\pi} \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} d\theta$$
$$\times \left\{ \sum_{l=0}^{n-2} \frac{(-1)^l}{c^{l+1}} \mathcal{K}\left(2|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right), |\mu_{s_m}|^2 + 2\Delta_{|s_m|}, n-l, A_p(\theta)\right) \right.$$
$$\left. + \frac{(-1)^{n+1}}{c^{n-1}} \mathcal{K}_1\left(2|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right), |\mu_{s_m}|^2 + 2\Delta_{|s_m|}, 2(a-1)\Delta_{|s_m|}, A_p(\theta)\right) \right\} \tag{105}$$

where $c = 2(a-1)\Delta_{|s_m|}$.

Similar to (104), for $a, b > 1$, consider the integral shown in (106) at the top of the following page. Using (100), we can express (106) as

$$\mathcal{N}_2(a,b,n,\theta_{L,p},\theta_{U,p},A_p(\theta),\mu_I(s_m),\mu_Q(s_m),\Delta_{|s_m|})$$
$$= \frac{(2\Delta_{|s_m|})^{n+2}}{\pi} \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} d\theta$$
$$\times \mathcal{C}_1\left(2|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right), |\mu_{s_m}|^2 + 2\Delta_{|s_m|}, 2(a-1)\Delta_{|s_m|}, 2(b-1)\Delta_{|s_m|}, n, A_p(\theta)\right). \tag{107}$$

That is, evaluation of (105) and (107) requires a single integration over θ, similar to what is needed for the AWGN channel [36, eq. (3.125)].

Upon using (105), (107), and (85) in (83), the average probability of symbol error of (83) can be expressed as (108) also at the top of the following page. The average probability of symbol error can then be obtained by summing (108) over all possible decision regions, and averaging the resulting expression for every symbol in the constellation. This leads to

$$\overline{P}_e = \frac{1}{M}\sum_{m=1}^{M}\sum_{p=1}^{N_m} \overline{P}_{e,s_m,p} \tag{109}$$

$$\begin{aligned}
&\mathcal{N}_2(a,b,n,\theta_{L,p},\theta_{U,p},A_p(\theta),\mu_I(s_m),\mu_Q(s_m),\Delta_{|s_m|}) \\
&\quad\triangleq \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} \frac{d\theta}{2\pi\Delta_{|s_m|}} \int_{R=A_p(\theta)}^{\infty} \frac{RdR}{(s+1)^n(s+a)(s+b)^2}\bigg|_{s=s_0} \\
&\quad= \frac{(2\Delta_{|s_m|})^{n+2}}{\pi} \int_{\theta=\theta_{L,p}}^{\theta_{U,p}} d\theta \int_{R=A_p(\theta)}^{\infty} RdR \\
&\qquad\times \frac{1}{\left(R^2-2R|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right)+|\mu_{s_m}|^2+2\Delta_{|s_m|}\right)^n} \\
&\qquad\times \frac{1}{\left(R^2-2R|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right)+|\mu_{s_m}|^2+2a\Delta_{|s_m|}\right)} \\
&\qquad\times \frac{1}{\left(R^2-2R|\mu_{s_m}|\cos\left(\theta-\phi_{\mu_{s_m}}\right)+|\mu_{s_m}|^2+2b\Delta_{|s_m|}\right)^2}.
\end{aligned} \tag{106}$$

$$\begin{aligned}
\overline{P}_{e,s_m,p} = K \sum_{l=K+1}^{L} \lambda(l)\mathcal{N}_1\left(\frac{l}{K},K+1,\theta_{L,p},\theta_{U,p},A_p(\theta),\mu_I(s_m),\mu_Q(s_m),\Delta_{|s_m|}\right) \\
+ \sum_{l_1=K+1}^{L} \frac{l_1}{K} \sum_{l_2=K+1,l_2\neq l_1}^{L} \zeta(l_1,l_2)\mathcal{N}_2\left(\frac{l_2}{K},\frac{l_1}{K},K,\theta_{L,p},\theta_{U,p},A_p(\theta),\mu_I(s_m),\mu_Q(s_m),\Delta_{|s_m|}\right).
\end{aligned} \tag{108}$$

where N_m is the number of nonintersecting decision regions for signal s_m.

Note again that by including the effects of signal-dependent noise, (108) and (109) improve upon the prior work. The average SEP expressions for 2-D signal constellations on Rayleigh fading without diversity but with channel estimation error are given in [25]. It is also easy to show that the final expressions in [25] are a special case of the results presented here in (108) and (109) when $L = K = 1$.

IV. Noncoherent and Differentially Coherent Receivers

We now extend the results of Section III to noncoherent and differentially coherent receivers. The receiver structure for M-ary orthogonal signaling and noncoherent detection is shown in Fig. 2, whereas the structure for an M-DPSK receiver with the conventional two-symbol detection is shown in Fig. 3. For M-DPSK, similar to [11]–[13], we also assume that both the channel and its estimate remain constant over the detection interval. Based on the relative strengths of the channel estimates, $|p_1|,\ldots,|p_L|$, the demodulator outputs from the K out of the L available channels, for each of the M possible hypotheses, are simply combined algebraically. One key observation to make regarding Figs. 2 and 3 is that the channel estimates play no role in the detection stage. First we start with binary FSK (i.e., $M = 2$) signaling.

A. Binary FSK

Assume that the branches corresponding to the estimates $|p_{(1)}|,\ldots,|p_{(K)}|$ are chosen for square-law combining. Then, conditioned on $|p_{(1)}|,\ldots,|p_{(K)}|$, the average probability of error is given by [21]

$$\begin{aligned}
&P_b(|p_{(1)}|,\ldots,|p_{(K)}|) \\
&\quad= \frac{1}{2}E_{|g_{(1)}|,\ldots,|g_{(K)}|\,\big||p_{(1)}|,\ldots,|p_{(K)}|} \\
&\qquad\times\left[\exp\left(-\frac{E_s}{2N_0}\sum_{k=1}^{K}|g_{(k)}|^2\right)\right]
\end{aligned} \tag{110}$$

where, from (2)

$$g_{(k)} = (x_k + jy_k) + \left(\frac{R_c + jR_{cs}}{\sigma_p^2}\right)p_{(k)}. \tag{111}$$

Clearly, conditioned on $|p_{(k)}|$, $|g_{(k)}|^2$ is noncentral χ^2 distributed with two degrees of freedom. The Laplace transform of the conditional density function $f_{|g_{(k)}|^2||p_{(k)}|}(x)$ of $|g_{(k)}|^2 \mid |p_{(k)}|$ is given by [21]

$$\mathcal{L}_{|g_{(k)}|^2||p_{(k)}|}(s) = \frac{\exp\left(-\frac{s(\rho^2/\kappa)|p_{(k)}|^2}{1+s2\sigma_e^2}\right)}{1+s2\sigma_e^2}. \tag{112}$$

Using (112) in (110), and the fact that, for $k = 1,\ldots,K$, the pairs $(|p_{(k)}|,|g_{(k)}|)$ are independent, we obtain

$$\begin{aligned}
P_b(|p_{(1)}|,\ldots,|p_{(K)}|) &= \frac{1}{2}\prod_{k=1}^{K} E_{|g_{(k)}||\,|p_{(k)}|}\left[e^{-\frac{E_s}{2N_0}|g_{(k)}|^2}\right] \\
&= \frac{1}{2}\prod_{k=1}^{K}\mathcal{L}_{|g_{(k)}|^2||p_{(k)}|}\left(\frac{E_s}{2N_0}\right) \\
&= \frac{1}{2}\left(\frac{2}{2+\overline{\gamma}(1-\rho^2)}\right)^K \\
&\quad\times\exp\left(-\frac{\overline{\gamma}\rho^2\beta^2}{2+\overline{\gamma}(1-\rho^2)}\right).
\end{aligned} \tag{113}$$

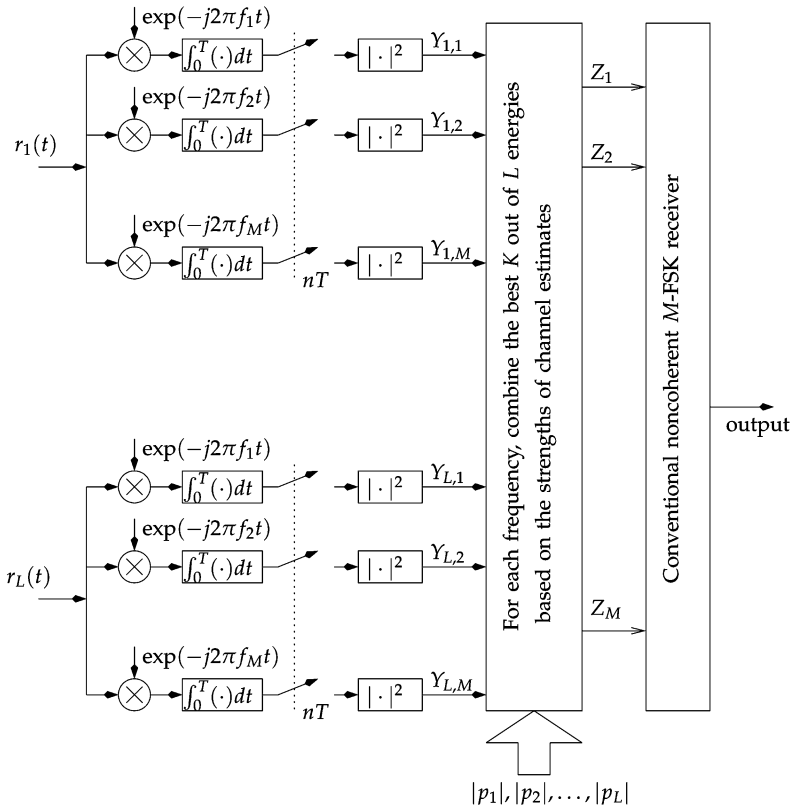

Fig. 2. GDC receiver for M-ary orthogonal FSK signaling with noncoherent detection. Note that the complex channel estimates $\{p_1, \ldots, p_L\}$ play no role in the demodulation process, whereas their magnitudes $\{|p_1|, \ldots, |p_L|\}$ are used to combine only a subset of the demodulator outputs.

The average probability of error is obtained by averaging over the statistics of β^2

$$\overline{P}_b(\text{Binary FSK}) = \frac{1}{2}\left(\frac{2}{2+\overline{\gamma}(1-\rho^2)}\right)^K \times \mathcal{L}_{\beta^2}\left(\frac{\overline{\gamma}\rho^2}{2+\overline{\gamma}(1-\rho^2)}\right). \quad (114)$$

Note from (114) that, when $\rho = 0$, we have

$$\overline{P}_b(\rho=0) = \frac{1}{2}\left(\frac{2}{2+\overline{\gamma}}\right)^K \quad (115)$$

which is the same as the performance of binary FSK signaling with Kth-order diversity. This is expected, and is explained as follows: When the channel estimate is completely decorrelated from the actual fade (as is evidenced by $\rho^2 \to 0$), picking the best K branches based on $|p_{(1)}|, \ldots, |p_{(K)}|$ is equivalent to picking K branches *randomly*. Consequently, we obtain Kth-order diversity performance. Note that, in contrast to (115), [13, eq. (16)] concludes that $\overline{P}_b = 1/(2+\overline{\gamma})$ with $\rho = 0$.

When $K = L$, using (32) for $\mathcal{L}_{\beta^2}(s)$, we obtain

$$\overline{P}_b(K=L) = \frac{1}{2}\left(\frac{2}{2+\overline{\gamma}}\right)^L \quad (116)$$

which is the same as the performance of binary FSK signaling with Lth-order square-law combining. This is also to be expected, since when all the branches are chosen, the channel estimates play no role in deciding the receiver performance, as the latter is employed with estimate-independent square-law detection (also see Fig. 2). In contrast, the authors in [5, eq. (26)] conclude that the L-branch binary NCFSK receiver is affected by channel estimation errors.

B. Binary DPSK

The main thing to notice for BDPSK signaling is that, conditioned on $|p_{(1)}|, \ldots, |p_{(K)}|$, (110) changes to [21]

$$P_b(|p_{(1)}|, \ldots, |p_{(K)}|) = \frac{1}{2}E_{|g_{(1)}|,\ldots,|g_{(K)}|||p_{(1)}|,\ldots,|p_{(K)}|} \times \left[\exp\left(-\frac{E_s}{N_0}\sum_{k=1}^{K}|g_{(k)}|^2\right)\right]. \quad (117)$$

Now, upon following the steps of (111)–(114), we arrive at the final expression for the average probability of error as

$$\overline{P}_b(\text{Binary DPSK}) = \frac{1}{2}\left(\frac{1}{1+\overline{\gamma}(1-\rho^2)}\right)^K \times \mathcal{L}_{\beta^2}\left(\frac{\overline{\gamma}\rho^2}{1+\overline{\gamma}(1-\rho^2)}\right). \quad (118)$$

When $K = L$, and for a given value of ρ, similar to (116), we obtain the average BEP as

$$\overline{P}_b(K=L) = \frac{1}{2}\left(\frac{1}{1+\overline{\gamma}}\right)^L. \quad (119)$$

When $\rho = 0$, (118) reduces to

$$\overline{P}_b(\rho=0) = \frac{1}{2}\left(\frac{1}{1+\overline{\gamma}}\right)^K \quad (120)$$

which is exactly the same as the performance of ideal GDC(L, K). The intuitive explanations for (119) and (120) are the same as given for binary FSK. By averaging the conditional BEP with the pdf of the instantaneous SNR r.v., [11]–[13] showed

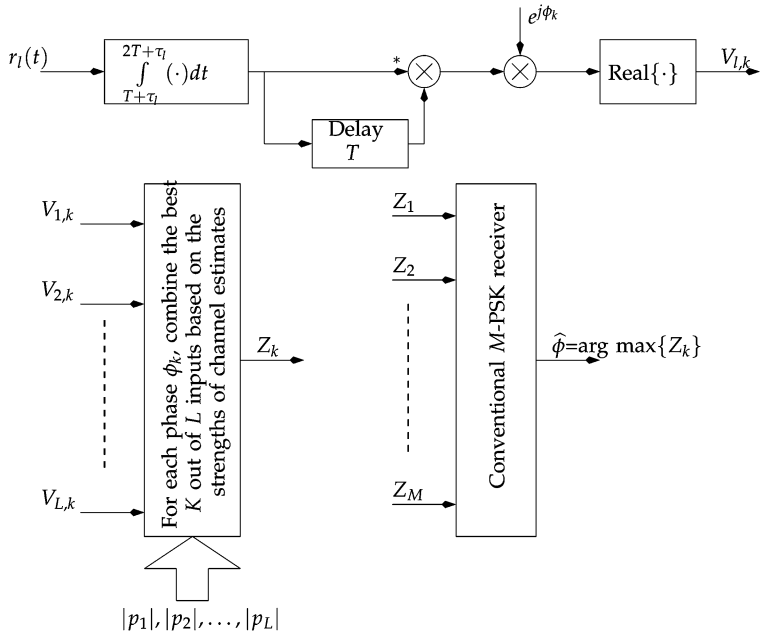

Fig. 3. GDC receiver for two-symbol M-ary DPSK signaling. Here, τ_l is the propagation delay on the lth channel and $\phi_k = 2\pi k/M$, $k = 1, \ldots, M$, is the phase of the information symbol. Note that the complex channel estimates $\{p_1, \ldots, p_L\}$ play no role in the demodulation process, whereas their magnitudes $\{|p_1|, \ldots, |p_L|\}$ are used to combine only a subset of the demodulator outputs.

that, when $\rho = 0$, the average BEP of BDPSK reduces to $(1/2)(1+\overline{\gamma})^{-1}$ (i.e., single-channel performance), whereas the actual performance is given by (119).[5]

C. M-Ary FSK

Conditioned on $|p_{(1)}|, \ldots, |p_{(K)}|$, the average symbol error probability for M-FSK signaling with noncoherent reception is given by [21] (121) at the bottom of the page. Using (112) in (121), we obtain the following simplification:

$$P_s(|p_{(1)}|, \ldots, |p_{(K)}|) = \sum_{n=1}^{M-1} \frac{(-1)^{n+1}}{n+1} \binom{M-1}{n} \times \left(\frac{n+1}{n+1+n\overline{\gamma}(1-\rho^2)}\right)^K e^{-\frac{n\rho^2\overline{\gamma}\beta^2}{n+1+n\overline{\gamma}(1-\rho^2)}}. \quad (122)$$

Invoking the Laplace transform of β^2, the average SEP with noncoherent M-FSK is

$$\overline{P}_s = \sum_{n=1}^{M-1} \frac{(-1)^{n+1}\binom{M-1}{n}}{n+1} \left(\frac{n+1}{n+1+n\overline{\gamma}(1-\rho^2)}\right)^K \times \mathcal{L}_{\beta^2}\left(\frac{n\rho^2\overline{\gamma}}{n+1+n\overline{\gamma}(1-\rho^2)}\right) \quad (123)$$

[5]It is to be noted that [31, eq. (22)] concludes that when $\varrho = 0$ the average BEP of BDPSK approaches 0.5.

using which the average BEP can be obtained as $\overline{P}_b = ((M/2)/(M-1))\,\overline{P}_s$[21].

The following two special cases are worth mentioning: *a)* $K = L$, and *b)* $\rho = 0$. When $K = L$, with the help of (32), (123) reduces to

$$\overline{P}_s(K = L) = \sum_{n=1}^{M-1} \frac{(-1)^{n+1}\binom{M-1}{n}}{n+1} \left(\frac{n+1}{n+1+n\overline{\gamma}}\right)^L. \quad (124)$$

Notice that (124) is exactly the same as the performance of an L-branch square-law receiver [39]. This shows that imperfect CSI does not have any effect on the performance of the M-FSK receiver.

When $\rho = 0$, using $\mathcal{L}_{\beta^2}(0) = 1$, (123) can be simplified to

$$\overline{P}_s(\rho = 0) = \sum_{n=1}^{M-1} \frac{(-1)^{n+1}\binom{M-1}{n}}{n+1} \left(\frac{n+1}{n+1+n\overline{\gamma}}\right)^K. \quad (125)$$

Comparing (125) with (124) we conclude that with $\rho = 0$, GDC(L, K) has the same performance as that of GDC(K, K). In contrast, [13] concludes that, with $\rho = 0$, GDC(L, K) has the performance of GDC(1,1) (i.e., no diversity). The reason for this is the same as given for (115).

$$P_s(|p_{(1)}|, \ldots, |p_{(K)}|) = E_{|g_{(1)}|,\ldots,|g_{(K)}|\,\big||p_{(1)}|,\ldots,|p_{(K)}|}\left\{\sum_{n=1}^{M-1} \frac{(-1)^{n+1}}{n+1}\binom{M-1}{n} e^{-\frac{nE_s}{(n+1)N_0}\sum_{k=1}^{K}|g_{(k)}|^2}\right\}. \quad (121)$$

D. M-Ary DPSK

To derive the average SEP with MDPSK reception and GDC, we use the following simple expression, due to [48, eq. (3)] and [49, eq. (11b)], for the average SEP of MDPSK on an AWGN channel:

$$P_s = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} e^{-\frac{E_s}{N_0}\times\frac{\sin^2\frac{\pi}{M}}{1+\cos\frac{\pi}{M}\cos\theta}}\,d\theta = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} e^{-\frac{E_s}{N_0}\times\frac{\sin^2\frac{\pi}{M}}{\sin^2\theta+\sin^2\left(\theta+\frac{\pi}{M}\right)}}\,d\theta. \quad (126)$$

Now observe that, conditioned on $\{|p_{(1)}|,\ldots,|p_{(K)}|\}$, the average SEP of MDPSK is

$$P_s(|p_{(1)}|,\ldots,|p_{(K)}|) = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} E_{|g_{(1)}|,\ldots,|g_{(K)}|\left||p_{(1)}|,\ldots,|p_{(K)}|\right.}\left\{e^{-\frac{E_s}{N_0}\sum_{l=1}^{K}|g_{(l)}|^2\times\frac{\sin^2\frac{\pi}{M}}{\sin^2\theta+\sin^2\left(\theta+\frac{\pi}{M}\right)}}\right\}d\theta. \quad (127)$$

Using (112), we have

$$\mathcal{L}_{\sum_{l=1}^{K}|g_{(l)}|^2||p_{(1)}|,\ldots,|p_{(K)}|}(s) = \prod_{l=1}^{K}\mathcal{L}_{|g_{(l)}|^2||p_{(l)}|}(s) = \left(\frac{1}{1+s2\sigma_e^2}\right)^K e^{-\frac{s(\rho^2/\kappa^2)\sum_{l=1}^{K}|p_{(l)}|^2}{1+s2\sigma_e^2}}. \quad (128)$$

Upon using (128) in (127), we obtain (129), shown at the bottom of the page. Upon letting $\sigma_e^2 = \sigma_g^2(1-\rho^2)$ and $2\sigma_g^2 E_s/N_0 = \overline{\gamma}$, and averaging (129) over β^2, we obtain the following simplification for the average SEP of MDPSK shown in (130), also at the bottom of the page. For the special cases of $\rho = 0$ and $K = L$, we obtain

$$\overline{P}_s(\rho=0) = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} d\theta \times\left(\frac{\sin^2\theta+\sin^2(\theta+\pi/M)}{\sin^2\theta+\sin^2(\theta+\pi/M)+\overline{\gamma}\sin^2(\pi/M)}\right)^K \quad (131)$$

and

$$\overline{P}_s(K=L) = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} d\theta \times\left(\frac{\sin^2\theta+\sin^2(\theta+\pi/M)}{\sin^2\theta+\sin^2(\theta+\frac{\pi}{M})+\overline{\gamma}\sin^2\pi/M}\right)^L. \quad (132)$$

That is, (131) shows that by randomly picking K branches, we obtain the performance of GDC(K, K) (see Fig. 3), whereas (132) shows that channel estimates play no role in SEP when combining all the branches.

V. Results and Discussion

In this section, we compare and contrast some of the results published in the literature against the ones presented in this paper. Fig. 4 plots the average output SNR of an MRC receiver with combiner weights derived from pilot-based MMSE channel estimation (see Section II-A). In Fig. 4, we assume $L = 4$ branches, and set the average received pilot SNR per branch, $\overline{\gamma}_{\text{pilot}}$, to 20 dB. From Section II-A, we have

$$\rho^2 = \frac{\overline{\gamma}_{\text{pilot}}}{1+\overline{\gamma}_{\text{pilot}}} \quad \text{and} \quad \sigma_p^2 = \sigma_g^2\left(1+\frac{1}{\overline{\gamma}_{\text{pilot}}}\right).$$

The average output SNR, derived in [10], is compared against the results presented in this paper in Appendix I. From Fig. 4, we observe that [10] predicts a linear increase in the average output SNR as a function of the average input SNR, whereas, in reality, signal-dependent noise due to imperfect channel estimation leads to a saturation of the output SNR. For the same set of parameters as that of Fig. 4, in Fig. 5 we compare the outage probability reported in [10, eq. (48)] against (137) derived in this paper. From Fig. 5, we observe that, due to imperfect channel

$$P_s(|p_{(1)}|,\ldots,|p_{(K)}|) = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} d\theta \times\left(\frac{\sin^2\theta+\sin^2(\theta+\pi/M)}{\sin^2\theta+\sin^2(\theta+\pi/M)+2\sigma_e^2\frac{E_s}{N_0}\sin^2(\pi/M)}\right)^K e^{-\frac{E_s/N_0\times 2\sigma_g^2\rho^2\beta^2\sin^2\pi/M}{\sin^2\theta+\sin^2(\theta+\frac{\pi}{M})+2\sigma_e^2\frac{E_s}{N_0}\sin^2\pi/M}}. \quad (129)$$

$$\overline{P}_s = \frac{1}{\pi}\int_0^{\frac{(M-1)\pi}{M}} d\theta \times\left(\frac{\sin^2\theta+\sin^2(\theta+\pi/M)}{\sin^2\theta+\sin^2(\theta+\pi/M)+(1-\rho^2)\overline{\gamma}\sin^2(\pi/M)}\right)^K \times\mathcal{L}_{\beta^2}\left(\frac{\overline{\gamma}\rho^2\sin^2\pi/M}{\sin^2\theta+\sin^2(\theta+\frac{\pi}{M})+(1-\rho^2)\overline{\gamma}\sin^2\pi/M}\right). \quad (130)$$

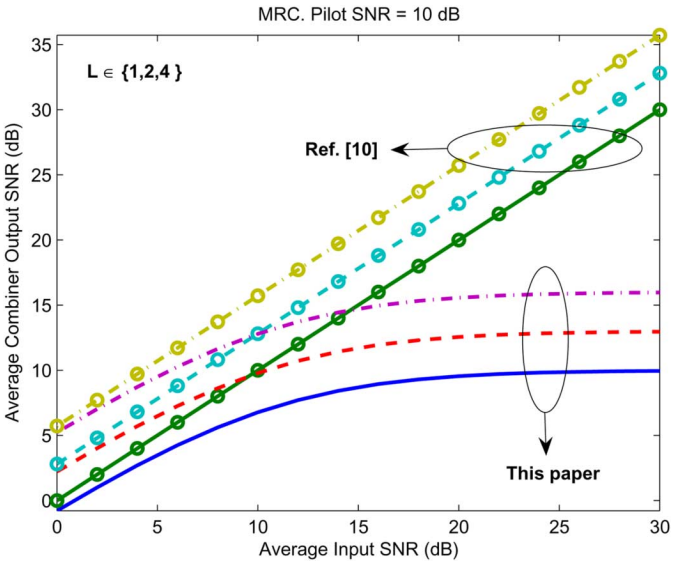

Fig. 4. Average output SNR as a function of the average input SNR for MRC receiver with the combiner weights based on MMSE channel estimation.

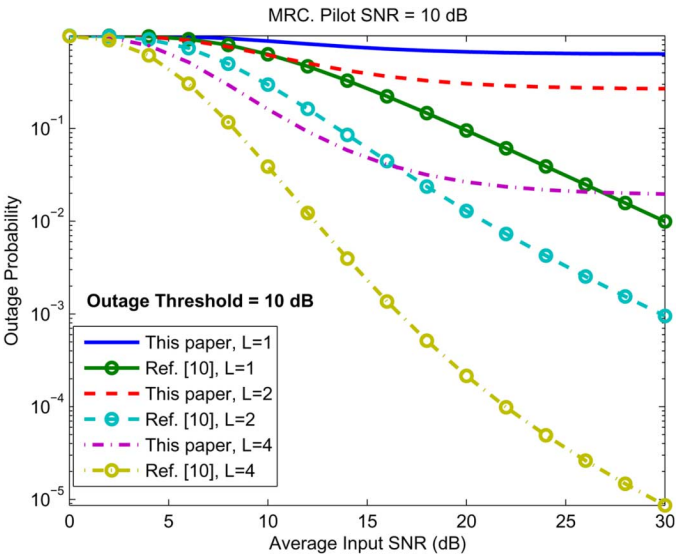

Fig. 5. Outage probability of MRC receiver with the combiner weights based on MMSE channel estimation. The legend labeled "This paper" corresponds to the outage probability expression derived in Appendix I, whereas the legend labeled "[10]" corresponds to the outage probability derived in [10, eq. (48)].

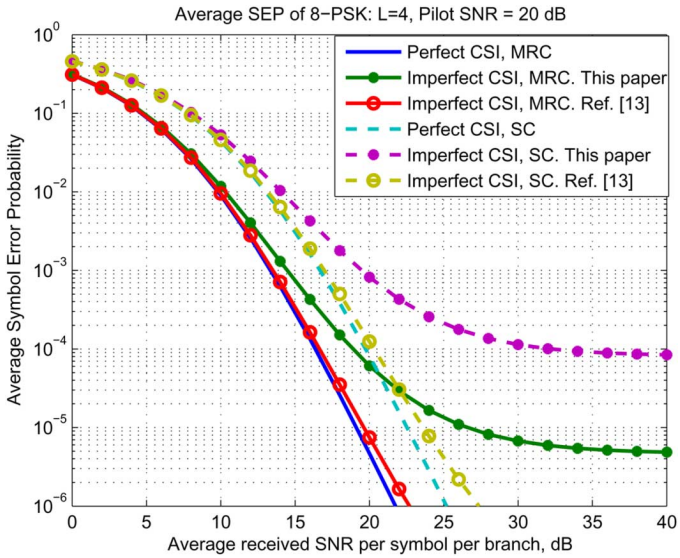

Fig. 6. Average SEP of 8-PSK with MRC and SC receivers, assuming $L = 4$ branches. The combiner weights are based on MMSE channel estimation. The legend containing "[13]" corresponds to the expression derived in [13, eq.(17)].

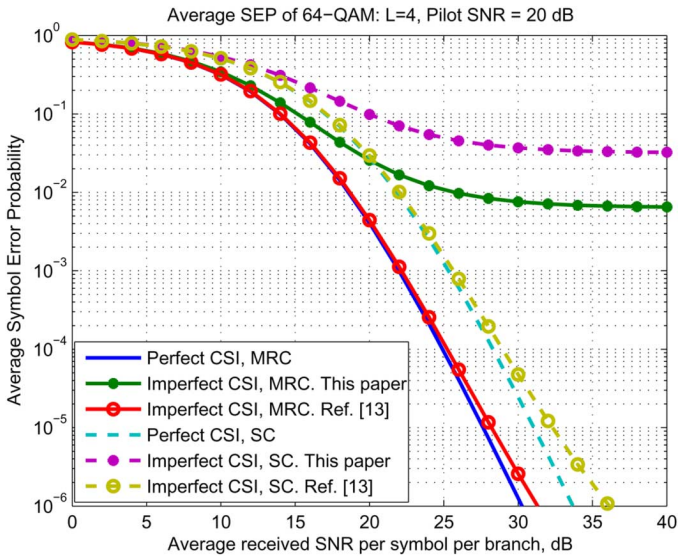

Fig. 7. Average SEP of 64-QAM with MRC and SC receivers, assuming $L = 4$ branches. The combiner weights are based on MMSE channel estimation. The legend containing "[13]" corresponds to the expression derived in [13, eq. (19)].

estimation, the actual outage probability suffers from an error floor.

The average SEP performance of 8-PSK modulation with MRC and SC receivers, and with $L = 4$ channels, is presented in Fig. 6. Similar to Figs. 4 and 5, MMSE channel estimation is assumed with $\overline{\gamma}_{\text{pilot}} = 20$ dB. The ideal performance (i.e., without estimation errors), and the performance based on the analysis in [13, eq. (17)] are also compared against the results derived in this paper. From Fig. 6, our analysis shows that the receiver incurs a severe degradation in performance due to an error floor. For the same set of system and channel parameters Fig. 7 shows the SEP performance of 64-QAM constellation, with a conclusion similar to Fig. 6.

We now plot the average SEP performance of M-ary DPSK and NCFSK modulations in Figs. 8 and 9, respectively. We set $M = 8$, $L = 4$, and choose $K \in \{2, L\}$. We also assume that $\rho = 0$ (i.e., a completely noisy channel estimate is provided to the *conventional* noncoherent/differentially coherent receivers). We note, from Figs. 2 and 3, that the channel estimates are used only for selecting the diversity channels but not for the signal detection process. From Fig. 8, we notice that, with $\rho = 0$, our result reveals that Kth-order diversity performance can be achieved with a completely noisy channel estimate. Similar results can be seen in Fig. 9 for the 8-ary NCFSK receiver. In short, our results establish that the effective diversity order of the receiver is equal to the number of branches the receiver combines. As reasoned in Section IV, with $\rho = 0$ and i.i.d. channel estimates, randomly choosing K channels from L channels is tantamount to having only K branches to start with. It follows

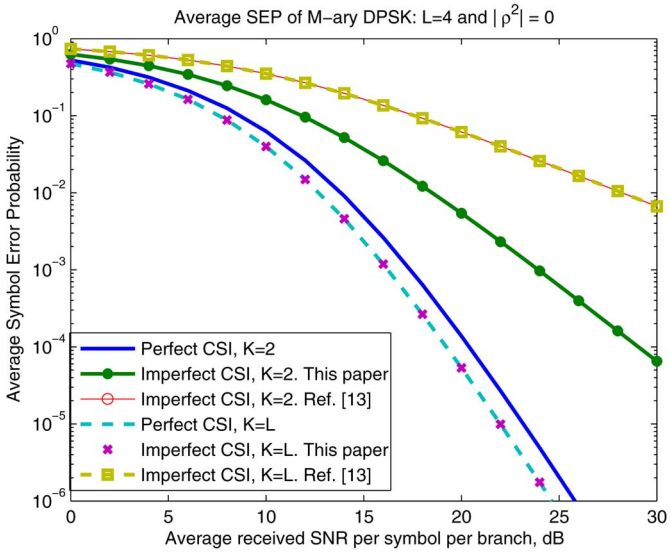

Fig. 8. Average SEP of 8-ary DPSK with GDC(L, K) reception. We assume $L = 4$ branches and $K \in \{2, L\}$. We consider the case with $\rho = 0$ (i.e., a completely noisy channel estimate is supplied to the *conventional* differential detector). The legend containing "[13]" corresponds to the expression derived in [13, eq.(20)].

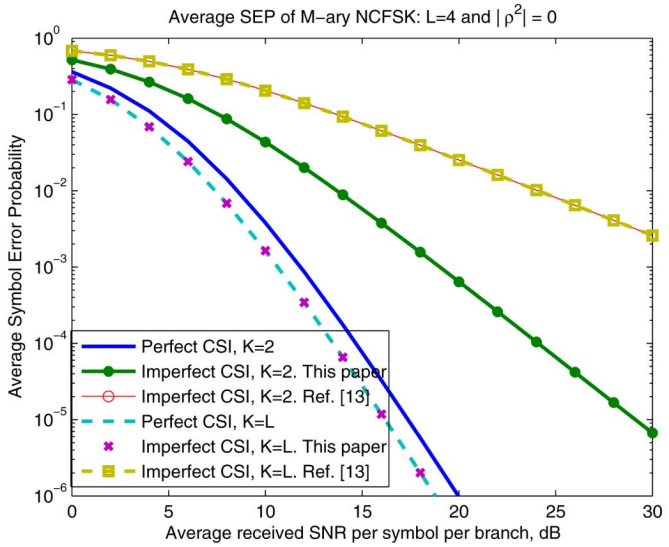

Fig. 9. Average SEP of 8-ary NCFSK with GDC(L, K) reception. We assume $L = 4$ branches and $K \in \{2, L\}$. We consider the case with $\rho = 0$ (i.e., a completely noisy channel estimate is supplied to the *conventional* noncoherent detector). The legend containing "[13]" corresponds to the expression derived in [13, eq.(18)].

that the latter system, with a conventional noncoherent/differentially coherent detection, yields a diversity of K[21]. Until now, we have assumed that the pilot SNR is fixed, irrespective of the operating data SNR. In this regime, the performance is limited by the quality of the channel estimates. However, in some practical wireless standards, the pilot SNR is continuously boosted relative to the data SNR.[6] In this case, asymptotically as the data SNR goes to infinity the pilot SNR also goes to infinity, and

[6]For example, in the emerging IEEE 802.16e WiMax standard [50], the pilot SNR is boosted by a variable factor relative to the data SNR.

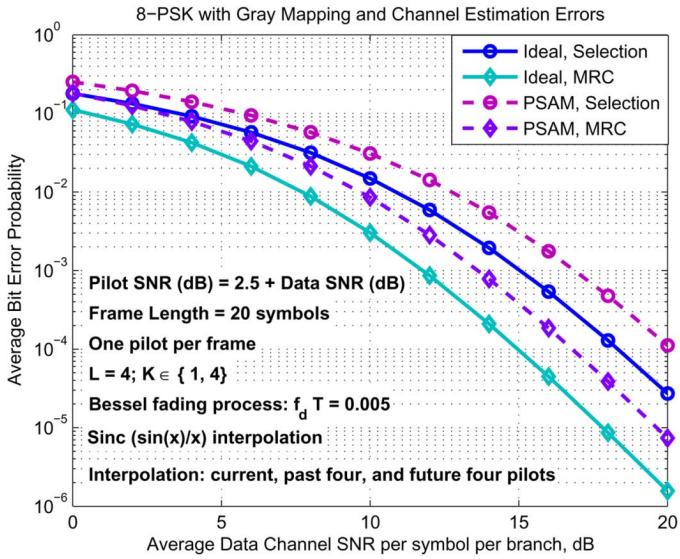

Fig. 10. Average probability of bit error for Gray coded 8-PSK with PSAM. The pilot SNR is continuously boosted relative to the data SNR by a factor of 2.5 dB.

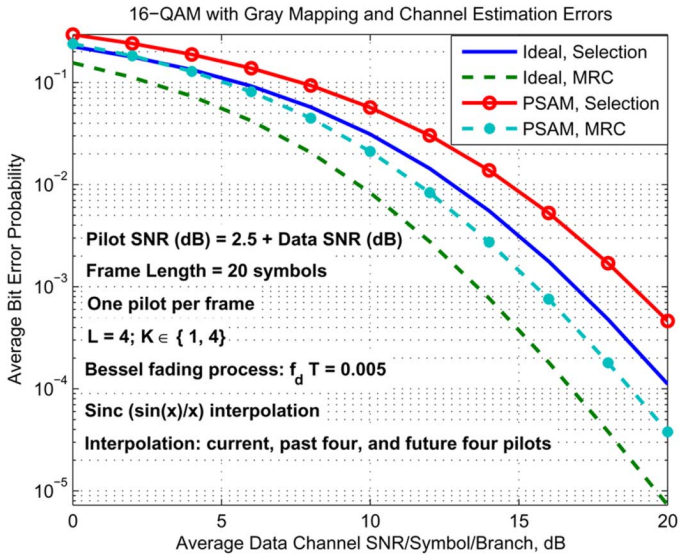

Fig. 11. Average probability of bit error for Gray coded 16-QAM with PSAM. The pilot SNR is continuously boosted relative to the data SNR by a factor of 2.5 dB.

hence the estimation errors vanish. As a result, there will not be any error floor.[7] Figs. 10 and 11 numerically verify this observation for 8-PSK and 16-QAM constellations with Gray code mapping. Here, the pilot SNR is assumed to be boosted by a factor of 2.5 dB. We let $L = 4$, and focus on the BEP performance with MRC and SC receivers. For channel estimation, we use the PSAM technique of [15] with the following parameters: Bessel fading correlation with a normalized fading bandwidth of $f_d T_s = 0.005$, a frame length of 20 symbols, one pilot symbol per frame, fading interpolation using the pilots of the current, past four, and future four pilots, and a $\sin(x)/x$ interpolation filter. Figs. 10 and 11 show that, except for a penalty in output SNR, there is no noticeable loss in diversity performance.

[7]This observation can in fact be proven analytically. However, for brevity we skip the proof.

VI. Conclusion

In this paper, using a fundamental decision-variable approach, we presented a rigorous analysis of the performance of GDC receivers on time-correlated Rayleigh fading channels with noisy channel estimates. We derived several new results on the error probabilities of coherent, noncoherent, and differentially coherent receivers with GDC and imperfect CSI. Our expressions for coherent receivers are also shown to be generalizations of some of the published results that were valid either for small constellation sizes, or for specific combining schemes (i.e., MRC or no-diversity systems), or both. We showed that the final analytical expressions were simple, requiring at most a single numerical integration with finite integration range, and have the complexity of evaluation that is comparable to that for the AWGN channel.

With completely decorrelated channel estimates, our results differ from prior literature in the following manner: i) While [10], [12], and [13] show that the outage probability, and the average SEP of M-ary PSK signaling, with a GDC(L, K) receiver vary inverse linearly with the average SNR, we prove that outage occurs with probability one, whereas the average SEP reduces to $(M-1)/M$. ii) With noncoherent and differentially coherent signaling, [5], [9], [11]–[13] conclude that the average probability of error varies inverse linearly with the average SNR, whereas our results show that the receiver performance is identical to that of an ideal GDC(K, K) receiver.

Appendix I
Outage Analysis for Coherent Reception

In this appendix, we present a simple analysis for the outage probability of a coherent diversity receiver whose output signal is given by (12). Let us denote by $\gamma^{(S)}(\beta^2)$ the instantaneous SNR r.v. at the output of the combiner, conditioned on S and β^2. Using (12), $\gamma^{(S)}(\beta^2)$ can be written as

$$\begin{aligned}\gamma^{(S)}(\beta^2) &= |S|^2 \frac{\left(R_c^2+R_{cs}^2\right)}{\sigma_p^4} \times \frac{1}{2\sigma_{|S|}^2} = \frac{|S|^2\rho^2\sigma_g^2\beta^2}{\sigma^2+|S|^2\sigma_e^2} \\ &= \frac{|S|^2\rho^2\sigma_g^2\beta^2}{\sigma^2+|S|^2\sigma_g^2(1-\rho^2)} \\ &= \underbrace{\frac{(|S|^2\sigma_g^2/\sigma^2)\rho^2}{1+(|S|^2\sigma_g^2/\sigma^2)(1-\rho^2)}}_{\triangleq \overline{\gamma}_{\mathrm{eff},S}} \beta^2. \end{aligned} \tag{133}$$

Let us denote by $\overline{\gamma}^{(S)} = |S|^2 2\sigma_g^2/(2\sigma^2) = 2\sigma_g^2|S|^2/N_0$ the ideal received SNR when S is transmitted, so that

$$\overline{\gamma}_{\mathrm{eff},S} = \frac{\overline{\gamma}^{(S)}\rho^2}{1+\overline{\gamma}^{(S)}(1-\rho^2)}, \tag{134}$$

is the effective SNR due to noisy CSI. Note that, for an M-PSK constellation, all the signal points have identical effective SNR.

Let us denote by $\gamma(\beta^2)$ the instantaneous SNR averaged over the constellation. For simplicity, we assume each $S \in \mathcal{S}$ is equiprobable, so that $\gamma(\beta^2)$ can be calculated as

$$\begin{aligned}\gamma(\beta^2) &= \frac{1}{|\mathcal{S}|}\sum_{j=1}^{|\mathcal{S}|}\gamma^{(S)}(\beta^2) \\ &= \beta^2\frac{1}{|\mathcal{S}|}\sum_{j=1}^{|\mathcal{S}|}\overline{\gamma}_{\mathrm{eff},S_j} = \beta^2\widehat{\overline{\gamma}_{\mathrm{eff}}} \end{aligned} \tag{135}$$

where $|\mathcal{S}|$ in (135) denotes the size of the set $\mathcal{S}$ and

$$\widehat{\overline{\gamma}_{\mathrm{eff}}} \triangleq \frac{1}{|\mathcal{S}|}\sum_{j=1}^{|\mathcal{S}|}\overline{\gamma}_{\mathrm{eff},S_j} \tag{136}$$

denotes the constellation-averaged effective SNR. For an M-PSK constellation, $\widehat{\overline{\gamma}_{\mathrm{eff}}}$ is identical to $\overline{\gamma}_{\mathrm{eff}}$ of (23). Using (135), the probability that $\gamma(\beta^2)$ falls below a predetermined threshold γ_{th} is

$$\begin{aligned}\mathrm{P}_{\mathrm{out}}(\gamma_{\mathrm{th}}) &\triangleq \mathrm{Prob}\left(\gamma(\beta^2) \le \gamma_{\mathrm{th}}\right) \\ &= \mathrm{Prob}\left(\beta^2 \le \gamma_{\mathrm{th}}/\widehat{\overline{\gamma}_{\mathrm{eff}}}\right), \qquad \gamma_{\mathrm{th}} > 0 \\ &= F_{\beta^2}\left(\frac{\gamma_{\mathrm{th}}}{\widehat{\overline{\gamma}_{\mathrm{eff}}}}\right) \end{aligned} \tag{137}$$

where $F_{\beta^2}(x)$ is the cdf of β^2 which is computed in [38], in closed form, as shown in (138) at the bottom of the page. We now perform some sanity checks on (137). First, let $\rho = 0$. Using (134) and (136), we have $\widehat{\overline{\gamma}_{\mathrm{eff}}} = 0$. Substituting $\widehat{\overline{\gamma}_{\mathrm{eff}}} = 0$ in (137), we see that $\mathrm{P}_{\mathrm{out}}(\gamma_{\mathrm{th}}) = 1$, which is in agreement with our intuition, as argued below (14). Next, when $\rho = 1$, (136) with (134) gives us

$$\widehat{\overline{\gamma}_{\mathrm{eff}}} = \left(2\sigma_g^2/N_0\right)(1/|\mathcal{S}|)\sum_{j=1}^{|\mathcal{S}|}|S_j|^2 = (E_s/N_0)2\sigma_g^2 = \overline{\gamma}.$$

Upon substituting $\widehat{\overline{\gamma}_{\mathrm{eff}}} = \overline{\gamma}$ in (137), we see that $\mathrm{P}_{\mathrm{out}}(\gamma_{\mathrm{th}}) = F_{\beta^2}(\gamma_{\mathrm{th}}/\overline{\gamma})$, which is the same as the outage probability for an ideal GDC [39]. For values of $\rho \notin \{0,1\}$, from (137), the outage probability is given by $F_{\beta^2}(\gamma_{\mathrm{th}}/\widehat{\overline{\gamma}_{\mathrm{eff}}})$, which indicates that, except for replacing the ideal SNR $\overline{\gamma}$ by the effective SNR $\widehat{\overline{\gamma}_{\mathrm{eff}}}$, noisy CSI does not reduce the diversity order of a coherent GDC.

Note that, when $\rho = 0$, the expressions for $\mathrm{P}_{\mathrm{out}}(\gamma_{\mathrm{th}})$ from [10] and [13] are given by

$$\mathrm{P}_{\mathrm{out}}(\gamma_{\mathrm{th}})(\rho=0) = 1-\exp\left(-\frac{\gamma_{\mathrm{th}}}{\overline{\gamma}}\right) \quad \text{([10, eq. (50)], [13, eq. (34)])} \tag{139}$$

$$\begin{aligned}F_{\beta^2}(x) = \binom{L}{K}\Bigg[1-e^{-x}\sum_{l=0}^{K-1}\frac{x^l}{l!} + \sum_{l=1}^{L-K}(-1)^{K+l-1}\binom{L-K}{l}\left(\frac{K}{l}\right)^{K-1} \\ \cdot\left(\frac{1-e^{-(1+\frac{l}{K})x}}{1+\frac{l}{K}} - \sum_{m=0}^{K-2}\left(\frac{-l}{K}\right)^m\left(1-e^{-x}\sum_{k=0}^{m}\frac{x^k}{k!}\right)\right)\Bigg]. \end{aligned} \tag{138}$$

which leads to the incorrect conclusion that single-channel performance results when $\rho = 0$.

Notice that when $\rho = 0$, using (134) in (136), the average SNR at the output of the GDC receiver becomes zero, irrespective of the GDC parameters L and K (This also follows from (14), since there is no signal component). On the other hand, the results in [10, p. 497] and [13, p. 505] show that the average SNR at the output of the combiner is equal to the per-branch average SNR $\overline{\gamma}$.

Appendix II
Derivation of (22)

To use Lemma 1, we need to compute s^2, Σ^2, and ψ. The parameter s^2 is given by $\mu_X^2 + \mu_Y^2$, where $\mu_X = m_{Z_I}$, as given by (18), and $\mu_Y = m_{Z_Q}$, as given by (19). This yields

$$\begin{aligned} s^2 &= m_{Z_I}^2 + m_{Z_Q}^2 \\ &= \left(\sqrt{E_s} \frac{R_c \cos\theta_m - R_{cs} \sin\theta_m}{\sigma_p^2} \right)^2 \\ &\quad + \left(\sqrt{E_s} \frac{R_c \sin\theta_m + R_{cs} \cos\theta_m}{\sigma_p^2} \right)^2 \\ &= \frac{E_s}{\sigma_p^4} \left(R_c^2 + R_{cs}^2 \right). \end{aligned} \tag{140}$$

The phase angle $\psi = \tan^{-1}(\mu_Y/\mu_X)$ is a function of the transmitted information phase θ_m and is denoted by ϕ_m. This is given by

$$\begin{aligned} \phi_m &= \tan^{-1} \left(\frac{R_c \sin\theta_m + R_{cs} \cos\theta_m}{R_c \cos\theta_m - R_{cs} \sin\theta_m} \right) \\ &= \tan^{-1} \left(\frac{\sqrt{R_c^2 + R_{cs}^2} \sin(\theta_m + \phi_\rho)}{\sqrt{R_c^2 + R_{cs}^2} \cos(\theta_m + \phi_\rho)} \right) \\ &= \theta_m + \phi_\rho. \end{aligned} \tag{141}$$

The term Σ^2 is simply given by $\sigma_{Z_I}^2$ (or $\sigma_{Z_Q}^2$), as in (20). The ratio $s^2/(2\Sigma^2)$ is then

$$\begin{aligned} \frac{s^2}{2\Sigma^2} &= \frac{E_s}{\sigma_p^4} \left(R_c^2 + R_{cs}^2 \right) \times \frac{2\sigma_p^2 \beta^2}{2(\sigma^2 + \sigma_e^2 E_s)} \\ &= \frac{R_c^2 + R_{cs}^2}{\sigma_p^2 \sigma_g^2} \times \frac{\frac{E_s 2\sigma_g^2}{2\sigma^2}}{1 + \frac{E_s 2\sigma_g^2}{2\sigma^2}(1-\rho^2)} \\ &= \frac{\rho^2 \overline{\gamma}}{1 + \overline{\gamma}(1-\rho^2)} \beta^2 = \overline{\gamma}_{\text{eff}} \beta^2 \end{aligned} \tag{142}$$

where $\overline{\gamma}_{\text{eff}}$ is given in (23). In the second step of (142) we have used $\rho = \sqrt{R_c^2 + R_{cs}^2}/(\sigma_p \sigma_g)$, $N_0 = 2\sigma^2$ and $\sigma_e^2 = \sigma_g^2(1-\rho^2)$. In the third step, we have used $\overline{\gamma} = 2\sigma_g^2 E_s/N_0$. Upon substituting (141) and (142) in Lemma 1, we arrive at (22).

Appendix III
Derivation of $\mathcal{G}(|a|, |b|, L, K)$ of (68)

Let a and b denote two real-valued constants. Let us consider the following integral:

$$\begin{aligned} \mathcal{G}(|a|, |b|, L, K) &\triangleq E\left[Q(|a|\sqrt{\beta^2}) Q(|b|\sqrt{\beta^2}) \right] \\ &= \int_{x=0}^{\infty} Q(|a|\sqrt{x}) Q(|b|\sqrt{x}) f_{\beta^2}(x) dx \\ &= F_{\beta^2}(x) Q(|a|\sqrt{x}) Q(|b|\sqrt{x}) \Big|_0^\infty \\ &\quad + \frac{|b|}{2\sqrt{2\pi}} \int_{x=0}^{\infty} F_{\beta^2}(x) Q(|a|\sqrt{x}) e^{-\frac{b^2 x}{2}} x^{-\frac{1}{2}} dx \\ &\quad + \frac{|a|}{2\sqrt{2\pi}} \int_{x=0}^{\infty} F_{\beta^2}(x) Q(|b|\sqrt{x}) e^{-\frac{a^2 x}{2}} x^{-\frac{1}{2}} dx \\ &= \frac{|b|}{2\sqrt{2\pi}} \int_{x=0}^{\infty} F_{\beta^2}(x) Q(|a|\sqrt{x}) e^{-\frac{b^2 x}{2}} x^{-\frac{1}{2}} dx \\ &\quad + \frac{|a|}{2\sqrt{2\pi}} \int_{x=0}^{\infty} F_{\beta^2}(x) Q(|b|\sqrt{x}) e^{-\frac{a^2 x}{2}} x^{-\frac{1}{2}} dx \end{aligned} \tag{143}$$

where the simplification is due to integration-by-parts.

Recognizing that every term in $F_{\beta^2}(x)$ of (138) can be expressed as a linear combination of $e^{-n_1 x} x^{n_2}$, we can find a solution for (143) as a linear combination of the solution for the following integral:

$$\begin{aligned} \mathcal{Z}(|a|, |b|, n_1, n_2) \triangleq & \frac{|b|}{2\sqrt{2\pi}} \int_{x=0}^{\infty} e^{-n_1 x} x^{n_2} \\ & \times Q(|a|\sqrt{x}) e^{-\frac{b^2 x}{2}} x^{-\frac{1}{2}} dx \\ & + \frac{|a|}{2\sqrt{2\pi}} \int_{x=0}^{\infty} e^{-n_1 x} x^{n_2} \\ & \times Q(|b|\sqrt{x}) e^{-\frac{a^2 x}{2}} x^{-\frac{1}{2}} dx. \end{aligned} \tag{144}$$

Fortunately, each of the above integrals can be simplified using the following result: First, using the definition

$$Q(x) = \frac{1}{\pi} \int_{\theta=0}^{\pi/2} \exp\left(-\frac{x^2}{2\sin^2\theta} \right) d\theta$$

[51], we simplify the following integral as:

$$\begin{aligned} & \int_{x=0}^{\infty} Q\left(|a|\sqrt{x}\right) e^{-b^2 x} x^{n-1} dx \\ &= \frac{1}{\pi} \int_{\theta=0}^{\frac{\pi}{2}} d\theta \int_{x=0}^{\infty} e^{-x\left[b^2 + \frac{a^2}{2\sin^2\theta}\right]} x^{n-1} dx \\ &= \frac{\Gamma(n)}{\pi b^{2n}} \int_{\theta=0}^{\frac{\pi}{2}} \left(\frac{\sin^2\theta}{\sin^2\theta + \frac{a^2}{2b^2}} \right)^n d\theta \end{aligned} \tag{145}$$

where $n \geq 1$ and $b > 0$. Upon using [39, eq. (5.17b)], we arrive at the following simplification for (145):

$$\int_{x=0}^{\infty} Q\left(|a|\sqrt{x}\right) e^{-b^2 x} x^{n-1} dx = \frac{\Gamma\left(n+\frac{1}{2}\right)|a|2^{n-1}}{n\sqrt{\pi}(a^2+2b^2)^{n+\frac{1}{2}}} \times {}_2F_1\left(1, n+\frac{1}{2}; n+1; \frac{2b^2}{a^2+2b^2}\right) \quad (146)$$

where ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ is the Gauss hypergeometric function [47], and $\Gamma(n)$ is the standard Gamma function [47].

Using (146), (144) can be expressed as

$$\begin{aligned}\mathcal{Z}(|a|,|b|,n_1,n_2) &= \frac{|b|\Gamma(n_2+1)}{(2n_2+1)2\sqrt{2}\pi} \times \sqrt{\frac{a^2}{a^2+b^2+2n_1}} \left(\frac{2}{a^2+b^2+2n_1}\right)^{n_2+\frac{1}{2}} \times {}_2F_1\left(1, n_2+1; n_2+\frac{3}{2}; \frac{b^2+2n_1}{a^2+b^2+2n_1}\right) \\ &+ \frac{|a|\Gamma(n_2+1)}{(2n_2+1)2\sqrt{2}\pi} \sqrt{\frac{b^2}{a^2+b^2+2n_1}} \times \left(\frac{2}{a^2+b^2+2n_1}\right)^{n_2+\frac{1}{2}} \times {}_2F_1\left(1, n_2+1; n_2+\frac{3}{2}; \frac{a^2+2n_1}{a^2+b^2+2n_1}\right). \quad (147)\end{aligned}$$

Using (147) along with (138), an expression for $\mathcal{G}(|a|,|b|,L,K)$ can be obtained as shown in (148) at the bottom of the page. For the case of SC, we have

$$F_{\beta^2}(x) = L\sum_{k=0}^{L}(-1)^k \binom{L}{k} e^{-kx}$$

[39], whereas with MRC

$$F_{\beta^2}(x) = 1 - e^{-x}\sum_{k=0}^{L-1} x^k/k!$$

[39]. Accordingly, (148) reduces to

$$\mathcal{G}(|a|,|b|,L,1) = L\sum_{k=0}^{L}(-1)^k \binom{L}{k} \times \mathcal{Z}(|a|,|b|,k,0) \text{ (SC)} \quad (149)$$

and

$$\mathcal{G}(|a|,|b|,L,L) = \mathcal{Z}(|a|,|b|,0,0) - \sum_{k=0}^{L-1} \frac{\mathcal{Z}(|a|,|b|,1,k)}{k!} \text{ (MRC)}. \quad (150)$$

Appendix IV
Derivation of (95) and (96)

Consider the following integral:

$$\mathcal{T} \triangleq \int_{x=R}^{\infty} \frac{x dx}{[(x-x_1)^2+y_1^2]^n}, \qquad n > 1. \quad (151)$$

By setting $x_1 = a/2$ and $y_1 = \sqrt{4b-a^2}/2$ in (151), we obtain (93).

By changing the integration variables from x to θ through the transformation $x = x_1 + y_1\tan\theta$ in (151), we obtain

$$\mathcal{T} = x_1 y_1^{1-2n} \int_{\theta=\theta_0}^{\pi/2} \cos^{2n-2}\theta d\theta + y_1^{2-2n} \int_{\theta=\theta_0}^{\pi/2} \cos^{2n-3}\theta \sin\theta d\theta \quad (152)$$

where $\theta_0 = \tan^{-1}\left(\frac{R-x_1}{y_1}\right)$. The second integral in (152) can easily be evaluated to $\frac{\cos^{2n-2}\theta_0}{2n-2}$, whereas the first integral can be rewritten as

$$\begin{aligned}\int_{\theta=\theta_0}^{\pi/2} \cos^{2n-2}\theta d\theta &= \int_{\theta=0}^{\pi/2} \cos^{2n-2}\theta d\theta - \int_{\theta=0}^{\theta_0} \cos^{2n-2}\theta d\theta \\ &= \frac{1}{2}\mathsf{Beta}\left(\frac{1}{2}, n-\frac{1}{2}\right) - \frac{1}{2}\mathsf{Beta}_{\text{inc}}\left(\sin^2\theta_0, \frac{1}{2}, n-\frac{1}{2}\right) \quad (153)\end{aligned}$$

$$\begin{aligned}\mathcal{G}(|a|,|b|,L,K) &= E\left[Q(|a|\sqrt{\beta^2})Q(|b|\sqrt{\beta^2})\right] \\ &= \binom{L}{K}\left[\mathcal{Z}(|a|,|b|,0,0) - \sum_{l=0}^{K-1}\frac{\mathcal{Z}(|a|,|b|,1,l)}{l!} + \sum_{l=1}^{L-K}(-1)^{K+l-1}\binom{L-K}{l}\left(\frac{K}{l}\right)^{K-1}\left(\frac{\mathcal{Z}(|a|,|b|,0,0) - \mathcal{Z}(|a|,|b|,1+\frac{l}{K},0)}{1+\frac{l}{K}}\right.\right. \\ &\left.\left. - \sum_{m=0}^{K-2}\left(-\frac{l}{K}\right)^m \times \left\{\mathcal{Z}(|a|,|b|,0,0) - \sum_{k=0}^{m}\frac{\mathcal{Z}(|a|,|b|,1,k)}{k!}\right\}\right)\right]. \quad (148)\end{aligned}$$

where we have used the definitions of $\mathsf{Beta}(m,n)$ and $\mathsf{Beta}_{\rm inc}(x,m,n)$ in the second step of (153).

Finally, after combining terms, (151) can be simplified to

$$\mathcal{T} = \frac{x_1 y_1^{1-2n}}{2}\left\{\mathsf{Beta}\left(\frac{1}{2}, n-\frac{1}{2}\right) - \mathsf{Beta}_{\rm inc}\left(\frac{(R-x_1)^2}{(R-x_1)^2+y_1^2}, \frac{1}{2}, n-\frac{1}{2}\right)\right\} + \frac{1}{2n-2}\times\frac{1}{[(R-x_1)^2+y_1^2]^{n-1}} \tag{154}$$

which, with $x_1 = a/2$ and $y_1 = \sqrt{4b-a^2}/2$, reduces to (95).

We now derive (96). From (94), we have

$$\mathcal{K}_1(a,b,c,R) = \int_R^\infty \frac{x\,dx}{(x^2-ax+b)(x^2-ax+b+c)}, \quad \text{where } 4b-a^2>0 \text{ and } c>0. \tag{155}$$

We write $x^2-ax+b = (x-x_1)^2+y_1^2$, where $x_1 = a/2$ and $y_1 = \sqrt{4b-a^2}/2$, and use the change of integration variable, from x to θ, via $x - x_1 = y_1\tan\theta$. With this, (155) becomes

$$\begin{aligned}\mathcal{K}_1(a,b,c,R) &= \int_{\theta_0}^{\frac{\pi}{2}} \frac{(x_1+y_1\tan\theta)y_1\sec^2\theta}{y_1^2\sec^2\theta\,(c+y_1^2\sec^2\theta)}d\theta \\ &= \frac{x_1}{y_1}\int_{\theta_0}^{\frac{\pi}{2}} \frac{\cos^2\theta d\theta}{c\times\cos^2\theta+y_1^2} + \int_{\theta_0}^{\frac{\pi}{2}} \frac{\sin\theta\cos\theta d\theta}{c\times\cos^2\theta+y_1^2} \\ &= \frac{x_1}{cy_1}\int_0^{\frac{\pi}{2}-\theta_0} \frac{\sin^2\theta d\theta}{\sin^2\theta+\frac{y_1^2}{c}} + \int_0^{\cos\theta_0}\frac{tdt}{ct^2+y_1^2} \\ &= \frac{x_1\pi}{cy_1} I_0\left(\frac{\pi}{2}-\theta_0; 0, \frac{y_1^2}{c}\right) + \frac{1}{2c}\log\left(1+\frac{c}{y_1^2}\cos^2\theta_0\right)\end{aligned} \tag{156}$$

where, in (156), $\theta_0 = \tan^{-1}\left(\frac{2R-a}{\sqrt{4b-a^2}}\right)$, and we have used (34) in the last step of (156). Equation (96) follows upon setting $x_1 = a/2$ and $y_1 = \sqrt{4b-a^2}/2$.

ACKNOWLEDGMENT

The authors wish to thank one of the reviewers for his/her meticulous reading of the manuscript, and for pointing out an error in an earlier version of this manuscript.

REFERENCES

[1] A. F. Molisch, Ed., *Wideband Wireless Digital Communications*. Upper Saddle River, NJ: Prentice-Hall, 2000.

[2] N. Kong, T. Eng, and L. B. Milstein, "A selection combining scheme for rake receivers," in *Proc. IEEE ICUPC'95*, Tokyo, Japan, Nov. 1995, pp. 426–430.

[3] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[4] A. Lapidoth and S. Shamai (Shitz), "Fading channels: How perfect need perfect side information be?," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[5] P. Bello and B. D. Nelin, "Predetection diversity combining with selectively fading channels," *IEEE Trans. Commun. Syst.*, vol. COM-10, no. 3, pp. 32–42, Mar. 1962.

[6] R. Annavajjala and L. B. Milstein, "Performance analysis of linear diversity combining schemes on Rayleigh fading channels with binary signaling and Gaussian weighting errors," *IEEE Trans. Wireless. Commun.*, vol. 4, no. 5, pp. 2267–2278, Sep. 2005.

[7] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.

[8] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[9] S. Roy and P. Fortier, "Maximal-ratio combining architectures and performance with channel estimation based on a training sequence," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1154–1164, Jul. 2004.

[10] M. J. Gans, "The effect of Gaussian error in maximal ratio combiners," *IEEE Trans. Commun. Technol.*, vol. COM-19, no. 8, pp. 492–500, Aug. 1971.

[11] B. R. Tomiuk, N. C. Beaulieu, and A. A. Abu-Dayya, "General forms of maximal ratio diversity with weighting errors," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 488–492, Apr. 1999.

[12] A. Annamalai, "The effect of Gaussian error on the selection diversity combiners," *Wiley J. Wireless Commun. Mobile Comput.*, vol. 1, no. 3, pp. 419–435, Jul. 2001.

[13] A. Annamalai and C. Tellambura, "Analysis of hybrid selection/maximal-ratio diversity with Gaussian errors," *IEEE Trans. Wireless Commun.*, vol. 1, no. 3, pp. 498–512, Jul. 2002.

[14] E. G. Larsson, "Diversity and channel estimation errors," *IEEE. Trans. Commun.*, vol. 52, no. 2, pp. 205–208, Feb. 2004.

[15] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 6, pp. 686–693, Nov. 1991.

[16] K. Yu, J. Evans, and I. Collings, "Performance analysis of pilot symbol aided QAM for Rayleigh fading channels," in *Proc. IEEE Int. Conf. Communications*, New York, Apr./May 2002, pp. 1731–1735.

[17] A. Aghamohammadi and H. Meyr, "On the error probability of linearly modulated signals on Rayleigh frequency-flat fading channels," *IEEE Trans. Commun.*, vol. 38, no. 11, pp. 1966–1970, Nov. 1990.

[18] M. G. Shayesteh and A. Aghamohammadi, "On the error probability of linearly modulated signals on frequency-flat Ricean, Rayleigh, and AWGN channels," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1454–1466, Feb./Mar./Apr. 1995.

[19] X. Tang, M.-S. Alouini, and A. J. Goldsmith, "Effect of channel estimation error on M-QAM BER performance on Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, no. 12, pp. 1856–1864, Dec. 1999.

[20] L. Cao and N. C. Beaulieu, "Exact error-rate analysis of diversity 16-QAM with channel estimation error," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 1019–1029, Jun. 2004.

[21] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[22] B. Xia and J. Wang, "Effect of channel-estimation error on QAM systems with antenna diversity," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 481–488, Mar. 2005.

[23] L. Cao and N. C. Beaulieu, "Closed-form BER results for MRC diversity with channel estimation errors in Rician fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 7, pp. 1440–1447, Jul. 2005.

[24] Y. Ma, D. Zhang, and R. Schober, "Effect of channel estimation errors on M-QAM with GSC diversity in fading channels," in *Proc. IEEE Int. Conf. Communications*, Seoul, Korea, May 2005, pp. 2190–2194.

[25] X. Dong and N. C. Beaulieu, "SER of two-dimensional signalings in Rayleigh fading with channel estimation errors," in *Proc. IEEE Int. Conf. Communications*, Anchorage, AK, May 2003, pp. 2763–2767.

[26] S. K. Wilson and J. M. Cioffi, "Probability density functions for analyzing multi-amplitude constellations in Rayleigh and Rician channels," *IEEE Trans. Commun.*, vol. 47, no. 3, pp. 380–386, Mar. 1999.

[27] J. G. Proakis, "Probabilities of error for adaptive reception of M-phase signals," *IEEE Trans. Comm. Technol.*, vol. COM-16, no. 1, pp. 71–81, Feb. 1968.

[28] J. Lassing, E. Strom, E. Agrell, and T. Ottosson, "Computation of the exact bit error rate of coherent M-ary PSK with Gray code bit mapping," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1758–1760, Nov. 2003.

[29] Y. Ma, R. Schober, and S. Pasupathy, "Performance of M-PSK with GSC and EGC with Gaussian weighting errors," *IEEE Trans. Veh. Technol.*, vol. 54, no. 1, pp. 149–162, Jan. 2005.

[30] X. Dong, N. C. Beaulieu, and P. H. Wittke, "Signal constellations for fading channels," *IEEE Trans. Commun.*, vol. 47, no. 5, pp. 703–714, May 1999.

[31] Y. Ma, R. Schober, and S. Pasupathy, "Effect of imperfect channel estimation on MRC diversity in fading channels," in *Proc. IEEE Int. Conf. Communications*, Paris, France, Jun. 2004, pp. 3163–3167.

[32] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, ser. Probability and Statistics. New York: Wiley, 2003.

[33] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Clifffs, NJ: Prentice-Hall, 1993.

[34] N. Kong and L. B. Milstein, "Average SNR of a generalized diversity selection combining scheme," *IEEE Commun. Lett.*, vol. 3, no. 3, pp. 57–59, Mar. 1999.

[35] M. Win and J. Winters, "Analysis of hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, no. 12, pp. 1773–1776, Dec. 1999.

[36] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques: Signal Design and Detection*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[37] R. F. Pawula, S. O. Rice, and J. H. Roberts, "Distribution of the phase angle between two vectors perturbed by Gaussian noise," *IEEE Trans. Commun.*, vol. COM-30, no. 8, pp. 1828–1841, Aug. 1982.

[38] M. S. Alouini and M. K. Simon, "An MGF-based performance analysis of generalized selection combining over Rayleigh fading," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 401–415, Mar. 2000.

[39] M. K. Simon and M.-S. Alouini, *Digital Communications over Generalized Fading Channels: A Unified Approach to Performance Analysis*. New York: Wiley, 2000.

[40] P. J. Lee, "Computation of the bit error rate of coherent M-ary PSK with Gray code bit mapping," *IEEE Trans. Commun.*, vol. COM-34, no. 5, pp. 488–491, May 1986.

[41] P. K. Vitthaladevuni and M.-S. Alouini, "Exact BER computation of generalized hierarchical PSK constellations," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2030–2037, Dec. 2003.

[42] M. K. Simon and R. Annavajjala, "On the optimality of bit detection of certain digital modulations," *IEEE Trans. Commun.*, vol. 53, no. 2, pp. 299–307, Feb. 2005.

[43] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, Jul. 2002.

[44] R. Annavajjala, "Comments on exact error-rate analysis of diversity 16-QAM with channel estimation error," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 393–396, Mar. 2006.

[45] P. K. Vitthaladevuni and M.-S. Alouini, "A recursive algorithm for the exact BER computation of generalized hierarchical QAM constellations," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 297–307, Jan. 2003.

[46] X. Dong, N. C. Beaulieu, and P. H. Wittke, "Error probabilities of two-dimensional M-ary signaling in fading," *IEEE Trans. Commun.*, vol. 47, no. 3, pp. 352–355, Mar. 1999.

[47] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, ser. Applied Mathematics Series 55. New York: National Bureau of Standard, 1964.

[48] R. F. Pawula, "A new formula for MDPSK symbol error probability," *IEEE Commun. Lett.*, vol. 2, no. 10, pp. 271–272, Oct. 1998.

[49] R. F. Pawula, "Generic error probabilities," *IEEE Trans. Commun.*, vol. 47, no. 5, pp. 697–702, May 1999.

[50] *IEEE Std. for Local and Metropolitan Area Networks, Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, IEEE Std. 802 16-2005.

[51] J. W. Craig, "A new simple and exact result for calculating the probability of error for two-dimensional signal constellation," in *Proc. IEEE MILCOM'91*, McLean, VA, Nov. 1991, pp. 571–575, .