Received 23 May 2020; revised 27 June 2020; accepted 12 July 2020. Date of publication 20 July 2020; date of current version 5 August 2020.

Digital Object Identifier 10.1109/OJCOMS.2020.3010514

Low-Resolution Quantization in Phase Modulated Systems: Optimum Detectors and Error Rate Analysis

SAMIRU GAYAN [1] (Graduate Student Member, IEEE), RAJITHA SENANAYAKE [1] (Member, IEEE), HAZER INALTEKIN [2] (Member, IEEE), AND JAMIE EVANS [1] (Senior Member, IEEE)

[1]Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC 3100, Australia

[2]School of Engineering, Macquarie University, North Ryde, NSW 2109, Australia

CORRESPONDING AUTHOR: S. GAYAN (e-mail: hewas@student.unimelb.edu.au)

This work was supported in part by the Australian Research Council through the Discovery Early Career Researcher Award under Grant DE180100501.

ABSTRACT This paper studies optimum detectors and error rate analysis for wireless systems with low-resolution quantizers in the presence of fading and noise. A *universal* lower bound on the average symbol error probability (SEP), correct for all M-ary modulation schemes, is obtained when the number of quantization bits is not enough to resolve M signal points. In the special case of M-ary phase shift keying (M-PSK), the maximum likelihood detector is derived. Utilizing the structure of the derived detector, a general average SEP expression for M-PSK modulation with n-bit quantization is obtained when the wireless channel is subject to fading with a circularly-symmetric distribution. For the Nakagami-m fading, it is shown that a transceiver architecture with n-bit quantization is *asymptotically* optimum in terms of communication reliability if $n \geq \log_2 M + 1$. That is, the decay exponent for the average SEP is the same and equal to m with infinite-bit and n-bit quantizers for $n \geq \log_2 M + 1$. On the other hand, it is only equal to $\frac{1}{2}$ and 0 for $n = \log_2 M$ and $n < \log_2 M$, respectively. An extensive simulation study is performed to illustrate the accuracy of the derived results, energy efficiency gains obtained by means of low-resolution quantizers, performance comparison of phase modulated systems with independent in-phase and quadrature channel quantization and robustness of the derived results under channel estimation errors.

INDEX TERMS Low-resolution ADCs, maximum likelihood detectors, symbol error probability, diversity order.

I. INTRODUCTION

A. BACKGROUND AND MOTIVATION

ANALOG-TO-DIGITAL converters (ADCs) are known to consume most of the power dissipated at a base station [1]. It is shown that the power consumed by ADCs grows exponentially with their resolution level and linearly with their sampling rate [2], [3]. Thus, using high-resolution quantization with high sampling rates can significantly degrade the energy efficiency of a communication system. With the introduction of massive multiple-input-multiple-output (MIMO) and millimeter wave (mmWave) technology, this is even more prominent in next generation wireless systems. Because, massive MIMO systems use hundreds of antennas where each antenna is connected to a dedicated radio frequency (RF) chain equipped with high-resolution ADCs. MmWave systems, on the other hand, use much larger bandwidths that require higher sampling rates. In fact, the typical power consumption of a high speed ($\geq$ 20 GSamples/s) and high-resolution (8-12 bits) ADC is around 500 [mWatts]. As an RF chain contains one ADC for the in-phase (I) component and another ADC for the quadrature (Q) component of the received signal, the total power consumption of the ADCs in a single RF chain is around 1 [Watts]. Therefore, a future mmWave massive MIMO system with 256 RF chains and 512 ADCs will require around 256 [Watts] of power [4], which is potentially unaffordable.

To overcome this challenge, one can consider two possible solutions. The first solution is to use high-resolution but

low-speed sub-ADCs, whereas the second one is to use high-speed but low-resolution ADCs. Since high-resolution but low-speed sub-ADC architecture introduces error floors due to mismatches among sub-ADCs [4], the research in the field has been more focused on the second solution. Consequently, the idea of replacing power hungry high-resolution ADCs with low-resolution ADCs could provide a viable solution to the power consumption concerns in future wireless systems.

Indeed, low-resolution ADCs have long been known to provide significant energy savings in digital transceiver implementations [5]–[7]. Their other benefits include simplification in design (especially with 1-bit ADCs) and reduction in transceiver form-factor [4], [8], [9]. On the other hand, using low-resolution ADCs will limit the amount of data that has to be transferred over the link that connects the RF components and the baseband-processing unit [10]. For an example, let us consider a base station which is equipped with 256 antenna arrays. At each antenna element, the in-phase and the quadrature components are sampled with 10-bit ADCs operating at 20 GSamples/s. Such a system produces over 10 Tbit/s of data. However, this data rate exceeds the rate supported by the common public radio interface (CPRI) used over today's fiber-optical fronthaul links [11]. This is also a severe problem in cloud radio access network (C-RAN) architecture where the baseband processing is moved away to a central unit from the base stations [12]. Furthermore, the future long-term evolution (LTE) networks are also expected to support a wide range of Internet-of-Things (IoT) applications through protocols such as LTE-M, NB-IoT and EC-GSM, where devices are usually battery power-limited [13]. In these future application scenarios, low-resolution ADC based digital transceivers have the ability to prolong the battery lifetime of remote IoT devices as well, and thereby lessening the operating costs and the need for frequent human intervention.

Motivated by the potential of low resolution ADCs, in this paper, we aim to understand the fundamental trade-off between using low resolution quantization and achieving a higher communication reliability. More specifically, we analyse the error probability performance of a communication system, when the receiver is equipped with a phase quantizer that only outputs the knowledge of the phase quantization region on which the received signal lands. The phase quantization is a generalization of the one-bit ADCs, i.e., one for each in-phase and quadrature components of the received signal, considered in several previous works [5], [9], [14]–[26], in which the quantizer only outputs the quadrant on which the received signal lands. Using the derived structure of the maximum likelihood (ML) detector, we fully characterize the symbol error probability performance of a single-input single-output (SISO) system when the receiver is equipped with an n-bit phase quatizer.

We note that the SISO system provides the basic building block of a more complex MIMO system. We adopt the SISO system to draw fundamental insights. It allows us to conduct a rigorous mathematical investigation to provide exact analytical expressions for the error probability performance. The performance results obtained in this paper hold in the general case of n-bit quantization, M-ary phase shift keying (M-PSK) modulation and Nakagami-m fading channels. We note that some work on low-resolution ADCs have previously focused on MIMO systems [5], [9], [14]–[25], but under the case of one-bit quantization. In this paper, our focus is on identifying the trade-off between the number of quantization bits and the error probability performance, which has not been considered previously. Therefore, we take a different approach in which we allow the number of bits in the quantizer to vary until the transceiver architecture becomes asymptotically optimum in terms of communication reliability.

B. RELATED WORK

Over the years, using low-resolution ADCs in wireless communication systems has been investigated under various aspects. The performance of communication systems with low-resolution ADCs is lower than that of the idealized systems without quantization or traditional systems with high-resolution ADCs. It was shown in [21] that the capacity of a point-to-point MIMO channel with 1-bit ADCs is lower bounded by the rank of the channel in the high signal-to-noise ratio (SNR) regime. Results in [23] and [24] show that the channel capacity reduces by a factor of $2/\pi$ (1.96 dB) in the low-SNR regime for a MIMO system with 1-bit ADCs, when compared to a conventional high-resolution system. Further, the results in [25] establish the fact that the performance loss due to employing 1-bit ADCs can be overcome by having approximately 2.5 times more antennas at the base station.

In [27], authors focus on the information rate of a quantized block non-coherent channel with 1-bit ADCs. The results in this paper show that around $80-85\%$ of the mutual information attained with unquantized observations can also be attained with 3-bit quantization for QPSK modulation and SNR greater than 2-3 dB. In [28], authors presented a mixed-ADC architecture for MIMO systems in which some of the high-resolution ADCs were replaced with 1-bit ADCs. Their results show that the proposed architecture can achieve a near-similar performance as conventional architecture while reducing the energy consumption considerably.

Signal detection rules developed for receivers with high-resolution ADCs often become sub-optimal for receivers with low-resolution ADCs [4]. In [29], the authors propose a linear minimum mean square error (LMMSE) receiver when in-phase and quadrature components of the received signal are independently quantized by using a low-resolution ADC. They provide an approximation for the mean squared error between the transmitted symbol and the received one, and derive an optimized linear receiver which performs better than the conventional Weiner filter. Results in [29] were further extended to an iterative decision feedback receiver with quantized observations in [30].

For the same quantizer structure of independent quantization of in-phase and quadrature signal components, an ML detector was obtained in [14] by using *only* 1-bit ADCs. The complexity of the ML detector proposed in [14] grows exponentially with high signal constellations, number of transmit antennas and network size, which is not practical for real-world deployments. To overcome this difficulty, a near-optimum ML detector was proposed in [15] by using the convex optimization techniques. Although the SEP performance of the proposed near-optimum ML detector is better than the performance of linear detectors, it has been numerically observed that the proposed near-optimum ML detector still suffers from an error floor as SNR increases [4], [15]. Complementing this critical observation, in our work, we show the existence of a universal error floor below which the average SEP cannot be pushed down for any M-ary modulation scheme and quantizer structure if the number of quantization bits is less than $\log_2 M$.

For a massive MIMO system with one-bit ADCs, a sphere decoding based low-complexity near-ML detection method is presented in [31]. In the proposed algorithm, which is called one-bit-sphere decoding (OSD), the complexity of the construction of the sphere is reduced by dividing the received signal vector into multiple subvectors. The authors discuss the trade-off between performance and complexity with the dimension of the sphere and the length of the subvectors. In [32], the authors focus on the development of soft-output detection methods for low-precision ADCs and propose another near-optimal detection method for a coded mmWave MIMO system. More recently, there has been progress in machine learning based approaches as well [33]–[35]. In [33], [34] a reinforcement learning approach is used to design a robust likelihood function learning method for MIMO systems with one-bit ADCs. For a similar system, a semi-supervised learning detector is proposed in [35] which is further improved to an online-learning detector.

C. MAIN CONTRIBUTIONS

In this paper, we consider a point-to-point wireless communication system, in which the receiver is equipped with a low-resolution ADC that quantizes the phase of the received signal. Our main contributions are summarized as follows.

- For any M-ary modulation scheme and quantizer structure, we show the existence of an error floor below which the average symbol error probability (SEP) cannot be pushed if the number of quantization bits n is less than $\log_2 M$. This provides a fundamental performance limit for the SEP of a communication system that does not have enough quantization bits to resolve the M possible transmitted signals. This result is presented in Theorem 1.
- Motivated by the capacity achieving property of circularly symmetric input distributions for low-resolution ADCs [36], we adopt M-PSK modulation and derive the optimum ML detection rule for signal detection with low-resolution ADCs. For the general n-bit quantization, we obtain exact analytical expressions for the average SEP attained by the derived ML rule, when the wireless channel is subjected to Nakagami-m fading. We use extensive numerical examples to compare the analytical expressions against the Monte-Carlo simulations. These results are presented in Theorems 2 and 3 respectively.
- We establish a fundamental ternary behavior in the average SEP performance of a wireless communication system with low-resolution ADCs and M-PSK modulation. Based on the Nakagami-m fading model, we show that the decay exponent of the average SEP is the same with that of an infinite-bit quantization, which is equal to m, when n is larger than or equal to $\log_2 M + 1$. We also show that it is equal to $\frac{1}{2}$ and 0 for $n = \log_2 M$ and $n < \log_2 M$, respectively. This result is presented in Theorem 4.
- By analysing the error probability performance in the high-SNR regime, we characterize the diversity order for low-resolution ADC based communication systems. From a system design point of view, our results show that using one additional bit on top of $\log_2 M$ of them can achieve optimum communication robustness in the high-SNR regime. In particular, for fading environments with a large value of m, using an extra quantization bit improves communication reliability significantly. On the other hand, for fading environments with a low value of m, it may be more beneficial to use $\log_2 M$ bits, without sacrificing from communication robustness too much but doubling system energy efficiency.
- We perform a numerical analysis to compare the performance of phase quantization to the more common situation where I and Q channels are independently quantized. For this analysis, we focus on Rayleigh fading ($m = 1$) with QPSK ($M = 4$) modulation to obtain some initial key insights into the potential bit savings that can be provided by the phase quantization approach employed in the paper. While clearly 2-bit phase quantization is identical to using 1-bit quantization on each of the I and Q arms, we show that 3-bit phase quantization performs very similarly to the optimized independent I and Q quantization using 4-bits. This indicates that phase quantization provides a 1-bit saving in this particular case. Further, our analysis also demonstrates that the optimized independent quantization with 6-bits in total performs the same as 4-bit phase quantization, i.e., a saving of 2 bits. An important conclusion from this analysis is that not only does phase quantization have the potential to save bits, but the independent case needs to be optimized and adjusted as a function of the channel statistics, something that is not required for phase quantization.
- We provide a numerical analysis establishing an important robustness property for our analytical results under channel state estimation errors. A salient feature arising from this numerical analysis is that the derived

fundamental ternary behaviour of the decay exponent of $p(\mathsf{SNR})$ stays the same if $n+1$ bits are used during the channel estimation stage, while n being the number of bits used for data decoding. Moreover, the performance difference between using n and $n+1$ bits for channel estimation becomes negligibly small as n increases. We note that the use of different number of bits for data decoding and channel estimation is possible by employing mixed-ADC receiver architectures commonly used in the literature. Furthermore, the same effect to improve channel estimation accuracy can be achieved by using carefully designed training sequences to perform a bisection search during the channel estimate phase.

The material in this paper was presented in part at [37] and [38]. In [37] we focus only on QPSK modulation for Rayleigh fading environments and in [38] we generalize the error probability results to M-PSK modulation under any circularly symmetric fading distributions. The present paper expands upon the results presented in [37] and [38] and provides a rigorous error probability performance and diversity order analysis under Nakagami-m fading. Our results in Theorem 4 (together with Theorem 1) establishes a fundamental *ternary* behaviour for the symbol error probability in the high SNR regime for low-resolution ADC based receiver architectures. These results do not appear in our previous work [37] and [38], or exist in any other previous paper in the literature. Further, we present a numerical performance comparison of phase quantization to the situation where the in-phase and quadrature channels are independently quantized. We further extend our numerical analysis to evaluate the robustness of our findings to imperfect channel state information. This analysis did not exist in our previous work [37] and [38].

D. NOTATION

We use uppercase letters to represent random variables and calligraphic letters to represent sets. We use $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{N}$ to denote the real line, 2-dimensional Euclidean space and natural numbers, respectively. For a pair of integers $i \leq j$, we use $\left[i : j\right]$ to denote the discrete interval $\{i, i+1, \ldots, j\}$. For two functions f and g, we will say $f(x) = O(g(x))$ as $x \to x_0$ if $|f(x)| \leq c|g(x)|$ for some $c > 0$ when x is sufficiently close to x_0. Similarly, we will say $f(x) = \Omega(g(x))$ as $x \to x_0$ if $|f(x)| \geq c|g(x)|$ for some $c > 0$ when x is sufficiently close to x_0. We write $f(x) = \Theta(g(x))$ as $x \to x_0$ if $f(x) = O(g(x))$ and $f(x) = \Omega(g(x))$ as $x \to x_0$. Finally, we will say $f(x) = o(g(x))$ as $x \to x_0$ if $\lim_{x \to x_0} \left|\frac{f(x)}{g(x)}\right| = 0$.

The set of complex numbers $\mathbb{C}$ is $\mathbb{R}^2$ equipped with the usual complex addition and complex multiplication. We write $z = z_{\mathrm{re}} + \jmath z_{\mathrm{im}}$ to represent a complex number $z \in \mathbb{C}$, where $\jmath = \sqrt{-1}$ is the imaginary unit of $\mathbb{C}$, and z_{re} and z_{im} are called, respectively, *real* and *imaginary* parts of z [39]. Every $z \in \mathbb{C}$ has also a *polar* representation $z = |z|\mathrm{e}^{\jmath\theta} = |z|(\cos(\theta) + \jmath \sin(\theta))$, where $|z| \triangleq \sqrt{z_{\mathrm{re}}^2 + z_{\mathrm{im}}^2}$ is the magnitude of z and $\theta = \mathsf{Arg}(z) \in [-\pi, \pi)$ is called the (principle) argument of z.[1] As is common in the communications and signal processing literature, $\mathsf{Arg}(z)$ will also be called the phase of z (modulo 2π). For a complex random variable $Z = Z_{\mathrm{re}} + \jmath Z_{\mathrm{im}}$, we define its mean and variance as $\mathsf{E}[Z] \triangleq \mathsf{E}[Z_{\mathrm{re}}] + \jmath \mathsf{E}[Z_{\mathrm{im}}]$ and $\mathsf{Var}(Z) \triangleq \mathsf{E}\left[|Z - \mathsf{E}[Z]|^2\right]$, respectively. We say that Z is *circularly-symmetric* if Z and $\mathrm{e}^{\jmath\theta}Z$ induce the same probability distribution over $\mathbb{C}$ for all $\theta \in \mathbb{R}$ [40], [41]. For $x > 0$, $\log x$ and $\log_2 x$ will denote natural logarithm of x and logarithm of x in base 2, respectively.

II. SYSTEM SETUP

A. CHANNEL MODEL AND SIGNAL MODULATION

We consider the classical point-to-point wireless channel model with flat-fading. For this channel, the received discrete-time baseband equivalent signal Y can be expressed by

$$Y = \sqrt{\mathsf{SNR}}HX + W, \tag{1}$$

where $X \in \mathcal{C} \subset \mathbb{C}$ is the transmitted signal, $\mathcal{C}$ is the constellation set of information signals in $\mathbb{C}$, SNR is the ratio of the transmitted signal energy to the additive white Gaussian noise (AWGN) spectral density, $H \in \mathbb{C}$ is the unit power channel gain between the transmitter and the receiver, and W is the circularly-symmetric zero-mean unit-variance AWGN, i.e., $W \sim \mathcal{CN}(0, 1)$. In order to formalize the receiver architecture and the optimum signal detection problem below, we will assume that $\mathcal{C} = \left\{\mathrm{e}^{\jmath\pi\left(\frac{2k+1}{M}-1\right)}\right\}_{k=0}^{M-1}$ in the remainder of the paper, which is the classical M-ary phase shift keying (M-PSK) signal constellation[2] and for ease of exposition, we only consider the case in which M is an integer power of 2.[3]

B. RECEIVER ARCHITECTURE

The receiver architecture is based on a low-resolution ADC. As illustrated in Fig. 1, the received signal Y is first sent through a low-resolution quantizer, and then the resulting quantized signal information is used to determine the transmitted symbol X. More specifically, if n bits are used to quantize Y, the quantizer Q divides the complex domain $\mathbb{C}$ into 2^n quantization regions and outputs the index of the region in which Y lies as an input to the detector. As such, we declare $Q(Y) = k$ if $Y \in \mathcal{R}_k$ for $k \in [0 : 2^n - 1]$, where $\mathcal{R}_k \subseteq \mathbb{C}$ is the kth quantization region. Since information is encoded in the phase of X with the above choice of constellation points, we choose $\mathcal{R}_k$ as the convex cone given by

$$\mathcal{R}_k = \left\{z \in \mathbb{C} : \frac{2\pi}{2^n}k \leq \mathsf{Arg}(z) + \pi < \frac{2\pi}{2^n}(k+1)\right\}. \tag{2}$$

1. The range of $\mathsf{Arg}(z)$ can be taken to be any interval of length 2π. For our purposes, taking its range to be $[-\pi, \pi)$ will help to simplify the notation for some integral expressions.

2. This choice of $\mathcal{C}$ ensures that the phase of X always lies in $[-\pi, \pi)$.

3. Extensions of our results to the more general case of M being any positive integer is straightforward, albeit with more complicated notation and separate analyses in some special cases.

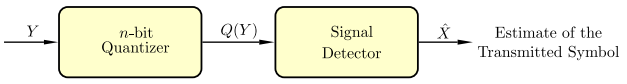

FIGURE 1. The receiver architecture with low-resolution quantization. The signal detector observes only the *n*-bit quantized versions of *Y* to estimate the transmitted signal.

The use of phase quantization in our receiver architecture is further motivated by the following two factors. First, considering channel impairments as phase rotations in transmitted signals, quantization and decision regions for M-PSK modulation are conveniently modelled as convex cones in the complex plane [42] as described above, and without requiring the use of automatic gain control. Second, phase quantizers can be implemented using one-bit ADCs that consist of simple comparators, and they consume negligible power (in the order of mWatts).

We also assume that full channel state information (CSI) is available at the receiver. The motivation behind our model with full-precision CSI in the paper is two-fold. First, in previous work [8], it was shown that it is possible to attain a high channel estimation precision (for $\mathsf{SNR} = 10$ dB, the estimator based on 3-bit ADC gives a mean square error of -20 dB with 11 training symbols) with the use of low-resolution ADCs by increasing the number of training symbols in the closed-loop estimation process.

Second, the mixed-ADC architectures are also commonly investigated in the literature, and they can be employed to achieve high channel estimation accuracy [28]. In a mixed-ADC architecture, high-resolution ADCs (structured through either serial or parallel connections) are used during the channel estimation stage [5], [8], [28] and during the data transmission phase, the receiver switches to low-resolution operation by using less number of quantization bits. Although the energy consumption is increased in this approach, this is not a restrictive degradation for our purposes. Each fading state will span a large group of information bits at the target multiple Gbits per second data rates in next generation wireless systems. Hence, the energy saving during data transmission is more significant than the increased energy consumption during channel estimation. Based on the existing results for channel estimation accuracy with low-resolution quantizers, the assumption on the availability of full channel state information at the receiver is also commonly used by the pioneering papers in the field [8], [43]–[45]. The effect of channel estimation errors on the SEP performance are further investigated in Section VII-C.

III. OPTIMUM SIGNAL DETECTION

The aim of the detector is to minimize the SEP by using the knowledge of $Q(Y)$ and channel state information, which can be represented as selecting a signal point $\hat{x}(k,h)$ satisfying

$$\hat{x}(k,h) \in \arg\max_{x\in\mathcal{C}} \mathsf{Pr}\big\{X = x \big| Q(Y) = k, H = h\big\}, \tag{3}$$

for $h \in \mathbb{C}$ and $k \in [0 : 2^n - 1]$. The main performance figure of merit for the optimum detector is the average SEP given by

$$p(\mathsf{SNR}) = \mathsf{Pr}\big\{X \neq \hat{x}(Q(Y), H)\big\}. \tag{4}$$

It is important to note that $p(\mathsf{SNR})$ depends on SNR as well as the number of quantization bits. Our first result indicates that there is an SNR-independent error floor such that the average SEP values below which cannot be attained for $n < \log_2 M$. The following theorem establishes this result formally.

Theorem 1: Let $p_{\min}$ be the probability of the least probable transmitted symbol. If $n < \log_2 M$, then for any choice of modulation scheme and quantizer structure

$$p(\mathsf{SNR}) \geq \frac{M - 2^n}{2^n} p_{\min} \tag{5}$$

for all $\mathsf{SNR} \geq 0$.

Proof: See Appendix A. ∎

Firstly, we note that the error floor in (5) is always a valid lower bound since $p_{\min} \leq \frac{1}{M}$. Secondly, it does not depend on the fading model. The average SEP values below $\frac{M-2^n}{2^n} p_{\min}$ cannot be achieved due to the inherent inability of low-resolution ADC receivers to resolve different signal points when $n < \log_2 M$. We also note that the Fano's inequality can also be used to obtain similar, perhaps tighter, lower bounds on $p(\mathsf{SNR})$ [46]. However, this will require the calculation of equivocation between X and $Q(Y)$ for each choice of modulation scheme and quantizer structure. Hence, it is not clear how the minimization can be carried out over the modulation and quantizer selections in this approach.

Next, we will assume that all signal points in $\mathcal{C}$ are equiprobable, with probability $\frac{1}{M}$, and hence the optimum detector in (3) is equivalent to the ML detector given by

$$\hat{x}(k,h) \in \arg\max_{x\in\mathcal{C}} \mathsf{Pr}\big\{Q(Y) = k \big| X = x, H = h\big\} \tag{6}$$

for $h \in \mathbb{C}$ and $k \in [0 : 2^n - 1]$. Since Y is a proper complex Gaussian random variable with mean $\mathsf{E}[Y] = \sqrt{\mathsf{SNR}}hx$ and variance $\mathsf{Var}(Y) = 1$, we can write the probability in (6) as

$$\begin{aligned}&\mathsf{Pr}\big\{Q(Y) = k \big| X = x, H = h\big\} \\ &= \int_{\mathcal{R}_k} \frac{1}{\pi} \exp\Big(-\big|y - \sqrt{\mathsf{SNR}}hx\big|^2\Big) dy,\end{aligned} \tag{7}$$

where the integral in (7) is with respect to the standard Borel measure in $\mathbb{C}$ [47]. The next theorem describes the operation of the ML detector for the above signal detection problem.

Theorem 2: Assume H has a continuous probability density function (pdf). Then, $\hat{x}(k,h)$ is unique with probability one, i.e., the set of h values for which $\arg\max_{x\in\mathcal{C}} \mathsf{Pr}\{Q(Y) = k | X = x, H = h\}$ is singleton has probability one, and the ML detection rule for the low-resolution ADC based receiver architecture can be given as

$$\hat{x}(k,h) = \arg\min_{x\in\mathcal{C}} \mathsf{dist}\Big(\sqrt{\mathsf{SNR}}hx, \mathcal{H}_k\Big), \tag{8}$$

where $h \in \mathbb{C}$, $k \in [0{:}2^n - 1]$, $\mathsf{dist}(z, \mathcal{A})$ is the distance between a point $z \in \mathbb{C}$ and a set $\mathcal{A} \subseteq \mathbb{C}$,

which is defined as $\mathsf{dist}(z, \mathcal{A}) \triangleq \inf_{s\in\mathcal{A}}|z-s|$, and $\mathcal{H}_k = \left\{z \in \mathbb{C} : \mathsf{Arg}(z) + \pi = \frac{\pi}{2^n}(2k+1)\right\}$.

Proof: See Appendix B. ■

We first note that although the ML detection rule given in Theorem 2 is written as dependent on the full channel knowledge, we only need to know the phase of H in order to determine $\hat{x}(k, h)$. This can be seen by means of simple scaling arguments as the set $\mathcal{H}_k$ is invariant under any amplitude scaling. The half-hyperplane $\mathcal{H}_k$ in Theorem 2 bisects the kth quantization region $\mathcal{R}_k$ into two symmetric regions. This is an important observation since it indicates that we are only required to estimate the channel phase to implement the ML detector. For phase estimation with circularly symmetric fading distributions, the estimation errors due to low-resolution quantization can be modelled as bounded uniform random variables, which lead to a key robustness property for our analytical results under channel estimation errors as explained in Section VII-C.

Secondly, given $X = x$ and $H = h$, the conditional probability of the event $Q(Y) = k$ corresponds to the integral of unit-variance and mean $\sqrt{\mathsf{SNR}}hx$ proper complex Gaussian distribution over $\mathcal{R}_k$. As we vary x in $\mathcal{C}$, we change the mean value $\sqrt{\mathsf{SNR}}hx$. Hence, Theorem 2 indicates that the most probability mass is accumulated in the region $\mathcal{R}_k$ when the unit-variance proper complex Gaussian distribution with mean $\sqrt{\mathsf{SNR}}hx$, $x \in \mathcal{C}$, closest to $\mathcal{H}_k$ is integrated over $\mathcal{R}_k$, which coincides with the intuition. We will use the structure of the ML detection rule to derive integral expressions for $p(\mathsf{SNR})$ for $M \geq 2$ in Section IV. Further, in order to characterize the communication robustness with low-resolution ADCs in the high SNR regime, we will also provide a detailed analysis on the asymptotic decay exponent of $p(\mathsf{SNR})$ in Section V.

IV. AVERAGE SYMBOL ERROR PROBABILITY

A. SYMBOL ERROR PROBABILITY FOR $n \geq \log_2 M$

We first obtain a key lemma that simplifies the calculations for deriving $p(\mathsf{SNR})$ when the number of quantization bits is at least $\log_2 M$. We note that this lemma holds for general circularly-symmetric fading processes without assuming any specific functional form.

Lemma 1: Let $H = Re^{j\Lambda}$ be a circularly-symmetric fading coefficient with R and Λ denoting the magnitude and the phase of H, respectively. Let the joint pdf of R and Λ be given by $f_{R,\Lambda}(r, \lambda) = \frac{1}{2\pi} f_R(r)$ for $\lambda \in [-\pi, \pi)$ and $r \geq 0$. Then, $p(\mathsf{SNR})$ is equal to

$$p(\mathsf{SNR}) = \frac{2^{n-1}}{\pi} \int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}} \int_0^\infty \Pr\left\{\sqrt{\mathsf{SNR}}re^{j\theta} + W \notin \mathcal{E}\right\} f_R(r)\, dr\, d\theta, \tag{9}$$

where $\mathcal{E} = \left\{z \in \mathbb{C} : 0 \leq \mathsf{Arg}(z) < \frac{2\pi}{M}\right\}$.

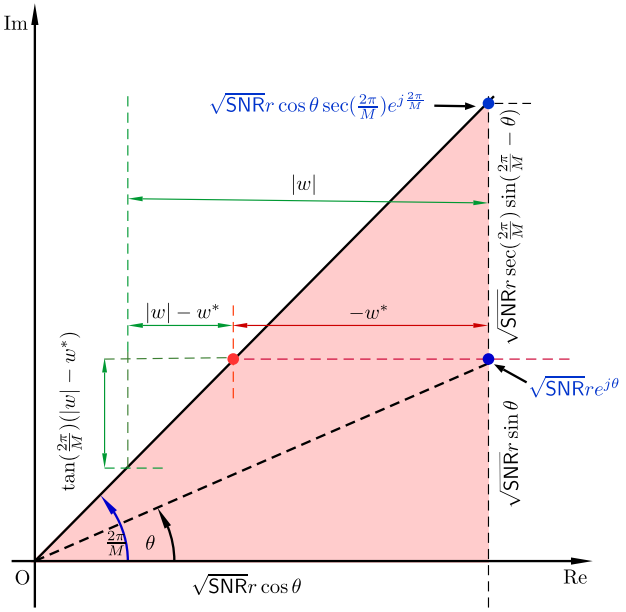

FIGURE 2. An illustration of average SEP calculations. If the noise does not drag the original M-PSK constellation point rotated by the channel h beyond the region $\mathcal{E}$ (shaded area), there will not be any errors in decoding.

Proof: See Appendix C. ■

Using Lemma 1, next we obtain integral expressions for $p(\mathsf{SNR})$ when H is circularly-symmetric with the generalized Nakagami-m fading magnitude. We note that the Nakagami-m fading model characterizes a broad range of fading phenomena ranging from severe to moderate and no fading conditions as m varies over $[0.5, \infty)$ [48], [49] and it reduces to Rayleigh fading for $m = 1$.

Considering these advantages, we will focus on the Nakagami-m fading model for H to derive integral expressions for $p(\mathsf{SNR})$ in the remainder of the paper. This will be done so for all parameter combinations of $M \geq 2$ (as an integer power of 2), $n \geq \log_2 M$ and $m \geq 0.5$. It will be seen that the derived integral expressions are easy to calculate numerically and they reduce to simple closed-form expressions in some special cases. Further, we will also show that using $\log_2 M + 1$ bits is enough to achieve the maximum communication robustness achieved by using infinite number of quantization bits.

Theorem 3: Assume H is a unit-power fading coefficient distributed according to a circularly-symmetric distribution with Nakagami-m fading magnitude. Let $\mathcal{Q}(\cdot)$ be the complementary distribution function of the standard normal random variable and $\Gamma(\cdot)$ be the gamma function [50]. Then, for $n \geq \log_2 M$ and $M \geq 2$, $p(\mathsf{SNR})$ is given according to (10), as shown at the bottom of the page.

Proof: In the following we provide the proof for $M \geq 4$. Please note that the proof for $M = 2$ is similar and simpler.

$$p(\mathsf{SNR}) = \begin{cases} p_1(\mathsf{SNR}) + p_2(\mathsf{SNR}) - p_3(\mathsf{SNR}) + p_4(\mathsf{SNR}) & M \geq 4 \\ p_2(\mathsf{SNR}) & M = 2 \end{cases}, \text{ where} \tag{10}$$

With a slight abuse of notation, we define

$$p(\mathsf{SNR}, h) = \mathsf{Pr}\Big\{\sqrt{\mathsf{SNR}}r\,\mathrm{e}^{J\theta} + W \notin \mathcal{E}\Big\}, \tag{15}$$

where the set $\mathcal{E}$ is defined as in Lemma 1. The probability in (15) can be calculated by conditioning on the real part of W, which is denoted by W_{re}. By using Fig. 2 as a visual guide, we can write $p(\mathsf{SNR}, h)$ after conditioning on W_{re} as (16) shown at the bottom of the page, for $w \geq -\sqrt{\mathsf{SNR}}r\cos\theta$.

Similarly, for $w < -\sqrt{\mathsf{SNR}}r\cos\theta$, we get

$$\mathsf{Pr}\Big\{\sqrt{\mathsf{SNR}}re^{j\theta} + W \notin \mathcal{E} \,\Big|\, W_{\mathrm{re}} = w\Big\} = 1. \tag{17}$$

Integrating (16) and (17) with respect to the pdf of W_{re}, which is given by $f_{W_{\mathrm{re}}}(w) = \frac{1}{\sqrt{\pi}}\mathrm{e}^{-w^2}$, we obtain $p(\mathsf{SNR}, h)$ as (18) shown at the bottom of the page. For Nakagami-m fading distribution with shape parameter $m \geq 0.5$ and spread parameter $\Omega > 0$ [51], we can write the pdf of the fading magnitude as $f_R(r) = \frac{2m^m}{\Gamma(m)\,\Omega^m} r^{2m-1}\mathrm{e}^{-\frac{m}{\Omega}r^2}$ for $r \geq 0$. We set $\Omega = 1$ in our calculations to make sure that H has unit-power. We average $p(\mathsf{SNR}, h)$ over the fading distribution and solve the resulting integral based on Lemma 1, and the fact that θ lies between 0 and $\frac{2\pi}{M}$, to obtain $p(\mathsf{SNR})$ in Theorem 3. ■

B. CENTERING PROPERTY: IMPACT OF QUANTIZATION BITS ON THE AVERAGE SEP

In this subsection, we will present an intuitive explanation as to why $p(\mathsf{SNR})$ improves with increasing number of quantization bits. In particular, we will observe that one extra bit, on top of $\log_2 M$ of them, provides a desirable centering property that steers the received signal away from the error-prone decision boundaries. This intuition will help to understand the underlying dynamics leading to the ternary behaviour for the decay exponent of $p(\mathsf{SNR})$ that we establish in the high SNR regime in Section V.

For $i \in [0 : M-1]$, let $x_i = \mathrm{e}^{J\pi\left(\frac{2i+1}{M}-1\right)}$ be the ith signal point in the constellation set $\mathcal{C}$ and $\mathcal{E}_i = \big\{z \in \mathbb{C} : \mathsf{Arg}(x_i) - \frac{\pi}{M} \leq \mathsf{Arg}(z) < \mathsf{Arg}(x_i) + \frac{\pi}{M}\big\}$. It can be shown (i.e., see Appendix C) that the regions defined by $\mathcal{E}_{i,k} \triangleq \exp\Big(J\big(k - 2^{n-1}\big)\frac{2\pi}{2^n}\Big)\mathcal{E}_i$ for $i \in [0 : M-1]$ and $k \in [0 : 2^n - 1]$ contains all $\mathcal{H}_k$'s to which $\sqrt{\mathsf{SNR}}hx_i$ is the closest for $h \in \mathcal{D}_k$, where

$$\mathcal{D}_0 = \Big\{z \in \mathbb{C} : \pi - \frac{\pi}{2^n} \leq \mathsf{Arg}(z) < \pi\Big\} \bigcup \Big\{z \in \mathbb{C} : -\pi \leq \mathsf{Arg}(z) < \frac{\pi}{2^n} - \pi\Big\} \tag{19}$$

and

$$\mathcal{D}_k = \Big\{z \in \mathbb{C} : (2k-1)\frac{\pi}{2^n} \leq \mathsf{Arg}(z) + \pi < (2k+1)\frac{\pi}{2^n}\Big\}. \tag{20}$$

This means that all the received signal points in $\mathcal{E}_{i,k}$ will be detected as x_i, and hence $\mathcal{E}_{i,k}$ can be considered as the *region of attraction* of x_i. This also means that if the received signal lands in $\mathcal{E}_{i,k}$ when x_i is transmitted, then there will not be any detection errors.

Let us consider an example for QPSK modulation with 2-bit and 3-bit quantization. Without loss of generality, we will assume that $x_3 = \mathrm{e}^{J\frac{\pi}{M}}$ is the transmitted signal. Our analysis will be for two cases of $\lambda = \frac{\pi}{18}$ and $\lambda = \frac{4\pi}{18}$, where $\lambda = \mathsf{Arg}(h)$. Table 1 summarizes these two cases, and Fig. 3 illustrates them. In this figure, we show both the

$$p_1(\mathsf{SNR}) = \frac{2^{n-1}m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\left(\frac{\mathsf{SNR}}{\sin^2\beta}\cos^2\theta + m\right)^{-m} d\theta d\beta \tag{11}$$

$$p_2(\mathsf{SNR}) = \frac{2^{n-1}m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\left(\frac{\mathsf{SNR}}{\sin^2\beta}\sin^2\theta + m\right)^{-m} d\theta d\beta \tag{12}$$

$$p_3(\mathsf{SNR}) = \frac{2^{n-1}m^m}{\pi^3}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\left(\frac{\mathsf{SNR}\cos^2\theta}{\sin^2\beta} + \frac{\mathsf{SNR}\sin^2\theta}{\sin^2\gamma} + m\right)^{-m} d\theta d\beta d\gamma \tag{13}$$

$$p_4(\mathsf{SNR}) = \frac{2^n m^m}{\pi\sqrt{\pi}\,\Gamma(m)}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\int_0^{\infty}\int_{-\sqrt{\mathsf{SNR}}r\cos\lambda}^{\infty}\mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\sec\left(\frac{2\pi}{M}\right)\sin\left(\frac{2\pi}{M}-\theta\right) + \sqrt{2}w\tan\left(\frac{2\pi}{M}\right)\right) \cdot \exp\Big(-\Big(w^2 + mr^2\Big)\Big) dw dr d\theta \tag{14}$$

$$\mathsf{Pr}\Big\{\sqrt{\mathsf{SNR}}re^{j\theta} + W \notin \mathcal{E} \,\Big|\, W_{\mathrm{re}} = w\Big\} = \mathcal{Q}\Big(\sqrt{2\mathsf{SNR}}r\sin\theta\Big) + \mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\sec\left(\frac{2\pi}{M}\right)\sin\left(\frac{2\pi}{M}-\theta\right) + \sqrt{2}w\tan\left(\frac{2\pi}{M}\right)\right) \tag{16}$$

$$p(\mathsf{SNR}, h) = \mathcal{Q}\Big(\sqrt{2\mathsf{SNR}}r\cos\theta\Big) + \mathcal{Q}\Big(\sqrt{2\mathsf{SNR}}r\sin\theta\Big) - \mathcal{Q}\Big(\sqrt{2\mathsf{SNR}}r\cos\theta\Big)\mathcal{Q}\Big(\sqrt{2\mathsf{SNR}}r\sin\theta\Big) + \frac{1}{\sqrt{\pi}}\int_{-\sqrt{\mathsf{SNR}}r\cos\theta}^{\infty}\mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\sec\left(\frac{2\pi}{M}\right)\sin\left(\frac{2\pi}{M}-\theta\right) + \sqrt{2}w\tan\left(\frac{2\pi}{M}\right)\right)\mathrm{e}^{-w^2}\,dw \tag{18}$$

TABLE 1. Centering property for QPSK modulation with 2-bit and 3-bit quantization. $\mathcal{E}_{i,k}$ is the region of attraction of the symbol x_i when the quantizer output $Q(Y) = k$.

	$\lambda = \frac{\pi}{18}$	$\lambda = \frac{4\pi}{18}$
$n = 2$	$h \in \mathcal{D}_2$ $i = 3,\ k = 2$ $\mathcal{E}_{i,k} = \mathcal{E}_3$	$h \in \mathcal{D}_4$ $i = 3,\ k = 4$ $\mathcal{E}_{i,k} = \mathcal{E}_3$
$n = 3$	$h \in \mathcal{D}_2$ $i = 3,\ k = 2$ $\mathcal{E}_{i,k} = \mathcal{E}_3$	$h \in \mathcal{D}_4$ $i = 3,\ k = 5$ $\mathcal{E}_{i,k} = e^{j\frac{\pi}{4}}\mathcal{E}_3$

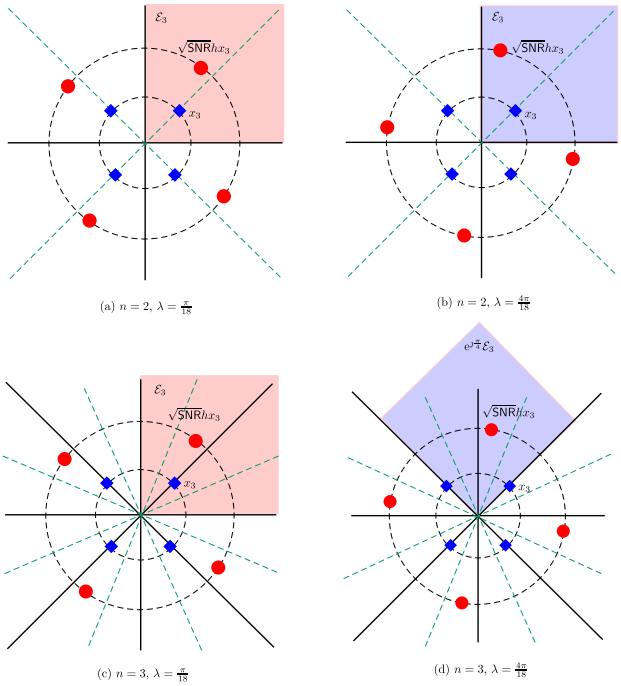

FIGURE 3. An illustration of the centering property for QPSK modulation with 2-bit and 3-bit quantization. Original signal points are indicated by '$\diamond$', whereas the rotated ones after multiplication with $\sqrt{\mathsf{SNR}}$ and h are indicated by '$\bullet$'. Quantization region boundaries and the corresponding bisectors are indicated in solid black lines and green dash lines, respectively. The shaded area represents the region of attraction of the transmitted symbol x_3.

original signal points (indicated by '$\diamond$') and the rotated ones (indicated by '$\bullet$') after multiplying with $\sqrt{\mathsf{SNR}}$ and h.

For both 2-bit and 3-bit quantization, we observe that $h \in \mathcal{D}_2$ and $h \in \mathcal{D}_4$ for $\lambda = \frac{\pi}{18}$ and $\lambda = \frac{4\pi}{18}$, respectively. Therefore, for 2-bit quantization, the region of attraction for x_3 will be $\mathcal{E}_3$ for both cases. Here, we can see that the rotated constellation point $\sqrt{\mathsf{SNR}}hx_3$ is very close to the decision boundary when $\lambda = \frac{4\pi}{18}$. Hence, there is a high probability that the received signal $\sqrt{\mathsf{SNR}}hx_3 + w$ lands in the adjacent quantization region for $\lambda = \frac{4\pi}{18}$. In this instance, we will have a detection error. However, with the addition of one bit to the quantizer (i.e., with 3-bit quantization), the region of attraction of x_3 will be $e^{j\frac{\pi}{4}}\mathcal{E}_3$, and hence the ML detector can correctly decode the transmitted signal even if the received one lands in the adjacent quantization region. This is illustrated in Fig. 3(d). Therefore, the addition of one extra bit to the quantizer, steers the received signal away from the error-prone decision boundaries to improve $p(\mathsf{SNR})$. Similarly, when the number of bits in the quantizer continues to increase, the quantization regions will become thinner, and hence the regions of attraction will be better centered around the received signal points. This is the fundamental phenomenon that explains why the average SEP improves with a larger number of quantization bits.

V. THE DECAY EXPONENT FOR THE AVERAGE SYMBOL ERROR PROBABILITY

In this section, we will analyze the communication robustness that can be achieved with low-resolution ADCs by focusing on the decay exponent for $p(\mathsf{SNR})$, which is given by[4]

$$\mathsf{DVO} = -\lim_{\mathsf{SNR}\to\infty} \frac{\log p(\mathsf{SNR})}{\log \mathsf{SNR}}. \tag{21}$$

Following the convention in the field, we will call DVO the *diversity order*, although there is only a single diversity branch in our system. It should be noted that Nakagami-m amplitude distribution can be obtained as the envelope distribution of m independent Rayleigh faded signals for integer values of m [48]. Hence, visualizing a Nakagami-m wireless channel as a pre-detection analog square-law diversity combiner will put the results of this section into context. We devote the rest of the current section to the proof of this important finding. We will first start with a definition that will simplify the notation below.

Definition 1: We say a function f is *exponentially equal* to SNR^d if $\lim_{\mathsf{SNR}\to\infty} \frac{\log f(\mathsf{SNR})}{\log \mathsf{SNR}} = d$ for some $d \in \mathbb{R}$. We write $f(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^d$ to indicate exponential equality whenever this limit exists. Similarly, we also write $f(\mathsf{SNR}) \stackrel{\mathrm{e}}{\leq} \mathsf{SNR}^d$ and $f(\mathsf{SNR}) \stackrel{\mathrm{e}}{\geq} \mathsf{SNR}^d$ if $\lim_{\mathsf{SNR}\to\infty} \frac{\log f(\mathsf{SNR})}{\log \mathsf{SNR}} \leq d$ and $\lim_{\mathsf{SNR}\to\infty} \frac{\log f(\mathsf{SNR})}{\log \mathsf{SNR}} \geq d$, respectively.

The following lemma establishes two important properties for exponential equality.

Lemma 2: Let $f(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^d$ and $f_i(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^{d_i}$ for $i \in [1:N]$. Then,

(i) For any $\alpha > 0$, $\alpha f(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^d$ (i.e., invariance with scaling property).

(ii) $\sum_{i=1}^{N} f_i(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^{d_{\max}}$, where $d_{\max} = \max_{i\in[1:N]} d_i$ (i.e., summation property).

Proof: See Appendix D. ∎

The next two lemmas establish the decay rates for $p_1(\mathsf{SNR})$ and $p_2(\mathsf{SNR})$ in Theorem 3 in terms of exponential equalities.

Lemma 3: For $M \geq 4$, $p_1(\mathsf{SNR})$ is exponentially equal to

$$p_1(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \begin{cases} \mathsf{SNR}^{-\frac{1}{2}} & \text{if } M = 4 \text{ and } n = 2, \\ \mathsf{SNR}^{-m} & \text{if } M = 4 \text{ and } n > 2, \\ \mathsf{SNR}^{-m} & \text{if } M > 4 \text{ and } n \geq \log_2 M. \end{cases} \tag{22}$$

Proof: See Appendix E. ∎

4. We will show that the limit in (21) exists, and hence there is no ambiguity in the definition of DVO.

Lemma 4: For $M \geq 4$, $p_2(\mathsf{SNR})$ is exponentially equal to

$$p_2(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \begin{cases} \mathsf{SNR}^{-\frac{1}{2}} & \text{if } n = \log_2 M, \\ \mathsf{SNR}^{-m} & \text{if } n > \log_2 M. \end{cases} \tag{23}$$

Proof: See Appendix F. ∎

The following lemma establishes lower and upper bounds on SEP in (10). We note that the bounds in Lemma 5 hold for all circularly-symmetric fading processes, including Nakagami-m magnitude pdf as a special case.

Lemma 5: For $M \geq 4$ and $n \geq \log_2 M$, let $L(\mathsf{SNR}) = p_1(\mathsf{SNR}) + \frac{1}{2}p_2(\mathsf{SNR})$ and $U(\mathsf{SNR}) = p_1(\mathsf{SNR}) + 2p_2(\mathsf{SNR})$. Then,

$$L(\mathsf{SNR}) \leq p(\mathsf{SNR}) \leq U(\mathsf{SNR}). \tag{24}$$

Proof: See Appendix G. ∎

The upper and lower bounds obtained in Lemma 5 are instrumental for us in providing a proof for the Theorem 4. In particular, they decay at the same rate as SNR grows to infinity. Technically speaking, they are *exponentially equal* to $p(\mathsf{SNR})$ (i.e., see Definition 1), and this exponential equality property enables the derivation of ternary behaviour for DVO as established in Theorem 4. Further, our numerical analysis in Section VII shows that the upper bound $U(\mathsf{SNR})$ becomes also a very tight bound for $p(\mathsf{SNR})$ for large SNR values in some certain circumstances.

Theorem 4: The DVO of a low-resolution ADC based receiver architecture with M-PSK modulation and Nakagami-m fading is given by

$$\mathsf{DVO} = \begin{cases} \frac{1}{2} & n = \log_2 M, \\ m & n \geq \log_2 M + 1. \end{cases} \tag{25}$$

Proof: The proof for $M \geq 4$ directly follows from Lemmas 2, 3, 4 and 5. For BPSK modulation (i.e., $M = 2$) and $n = 1$, we have

$$\begin{aligned} p(\mathsf{SNR}) &= \frac{2^{n-1}m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_0^{\pi}\left(\frac{\mathsf{SNR}}{\sin^2\beta}\sin^2\theta + m\right)^{-m} d\theta d\beta \\ &= \frac{2^{n-1}m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\left(\frac{\mathsf{SNR}}{\sin^2\beta}\sin^2\theta + m\right)^{-m} d\theta d\beta \\ &\quad + \frac{2^{n-1}m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_{\frac{\pi}{2}}^{\pi}\left(\frac{\mathsf{SNR}}{\sin^2\beta}\sin^2\theta + m\right)^{-m} d\theta d\beta \end{aligned} \tag{26}$$

By using the change of variables $\hat{\theta} = \theta - \frac{\pi}{2}$ in the second integral term of (26), we have

$$\begin{aligned} p(\mathsf{SNR}) &= \frac{2^{n-1}m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\left(\frac{\mathsf{SNR}}{\sin^2\beta}\sin^2\theta + m\right)^{-m} d\theta d\beta \\ &\quad + \frac{2^{n-1}m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\left(\frac{\mathsf{SNR}}{\sin^2\beta}\cos^2\hat{\theta} + m\right)^{-m} d\hat{\theta} d\beta \end{aligned} \tag{27}$$

This expression is equivalent to $p_1(\mathsf{SNR}) + p_2(\mathsf{SNR})$ for $M = 4$ and $n = 2$. Hence, by using Lemma 2, we can conclude that

$$\lim_{\mathsf{SNR}\to\infty} -\frac{\log(p(\mathsf{SNR}))}{\log(\mathsf{SNR})} = \frac{1}{2} \tag{28}$$

for BPSK modulation with 1-bit quantization and $m \geq \frac{1}{2}$.

For BPSK modulation with $n > \log_2(M)$, we have

$$\begin{aligned} &p(\mathsf{SNR}) \\ &= \frac{2^{n-1}m^m}{\pi^2}(\mathsf{SNR})^{-m}\int_0^{\frac{\pi}{2}}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\left(\frac{\sin^2\theta}{\sin^2\beta} + \frac{m}{\mathsf{SNR}}\right)^{-m} d\theta d\beta \end{aligned} \tag{29}$$

Therefore

$$\begin{aligned} &\log(p(\mathsf{SNR})) \\ &= c - m\log(\mathsf{SNR}) \\ &\quad + \log\left(\int_0^{\frac{\pi}{2}}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\left(\frac{\sin^2\theta}{\sin^2\beta} + \frac{m}{\mathsf{SNR}}\right)^{-m} d\theta d\beta\right), \end{aligned} \tag{30}$$

where $c = \log\left(\frac{2^{n-1}m^m}{\pi^2}\right)$. Define the function $g_{\mathsf{SNR}}(\theta,\beta) \triangleq \left(\frac{\sin^2\theta}{\sin^2\beta} + \frac{m}{\mathsf{SNR}}\right)^{-m}$, indexed by SNR. Since it is positive and increases to the limiting function $g_\infty(\theta,\beta) = \left(\frac{\sin^2\theta}{\sin^2\beta}\right)^{-m}$ as SNR increases, we can use the monotone convergence theorem [52] to write

$$\begin{aligned} &\lim_{\mathsf{SNR}\to\infty}\log\left(\int_0^{\frac{\pi}{2}}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\left(\frac{\sin^2\theta}{\sin^2\beta} + \frac{m}{\mathsf{SNR}}\right)^{-m} d\theta d\beta\right) \\ &= \log\left(\int_0^{\frac{\pi}{2}}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\left(\frac{\sin^2\theta}{\sin^2\beta}\right)^{-m} d\theta d\beta\right). \end{aligned}$$

We note that the last integral is finite since $g_\infty(\theta,\beta)$ is continuous and finite over the range of integration. Therefore, after observing that the last term in (30) converges to a finite value for each fixed $m \geq \frac{1}{2}$, we get

$$\lim_{\mathsf{SNR}\to\infty} -\frac{\log(p(\mathsf{SNR}))}{\log(\mathsf{SNR})} = m \tag{31}$$

for BPSK modulation with $n > \log_2(M)$. ∎

The DVO analysis above helps to discover the first-order effects of the low-resolution ADC based receivers on the SEP system performance. In particular, we observe that it is enough to use $\log_2 M + 1$ bits for quantizing the received signal to extract full diversity, which is equal to m for Nakagami-m faded wireless channels. Considering the fact that energy consumption increases exponentially with the number of quantization bits [53], this finding indicates that a significant energy saving is possible by means of low-resolution ADC based receivers without any (first order) loss in communication robustness.

We also observed that the DVO is only equal to $\frac{1}{2}$ when $n = \log_2 M$. Together with the universal bound obtained in Theorem 1, the discovered *ternary* behaviour has significant implications in terms of how to choose the number of quantization bits for low-resolution ADC based receivers. In particular, for fading environments with m close to $\frac{1}{2}$,

a system designer may decide to trade off reliability for energy consumption, without having too much degradation in average SEP by using $\log_2 M$ bits. On the other hand, for fading environments with large m, it is more beneficial to use one extra bit to have a major improvement in average SEP.

VI. PERFORMANCE ANALYSIS FOR QPSK MODULATION

In this section, we conduct a performance analysis for QPSK modulation with low-resolution ADCs by using our results in previous sections. We first present a simplified version of the average SEP expression in (11) for QPSK modulation, and then we analyze the effect of low-resolution quantization under Rayleigh fading.

A. SYMBOL ERROR PROBABILITY FOR QPSK MODULATION

Nakagami-m Fading: In the special case of QPSK modulation (i.e., $M = 4$), the average SEP expression in (11) can be further simplified to produce

$$p(\mathsf{SNR}) = p_1(\mathsf{SNR}) + p_2(\mathsf{SNR}) - p_3(\mathsf{SNR}), \quad (32)$$

because $\tan\left(\frac{2\pi}{M}\right) = \infty$ for $M = 4$. By using hypergeometric function ${}_2F_1[\,\cdot\,]$ [50], we can simplify (32) for 2-bit quantization (i.e., $M = 4$ and $n = 2$) as

$$p(\mathsf{SNR}) = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} {}_2F_1\left[\frac{1}{2}, m, 1, \frac{-\mathsf{SNR}}{m\sin^2\beta}\right] d\beta - \frac{m^m}{\pi}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} \left(\frac{\mathsf{SNR}}{\sin^2\gamma} + m\right)^{-m} {}_2F_1\left[\frac{1}{2}, m, 1, z\right] d\beta\, d\gamma,$$

where $z = \frac{\mathsf{SNR}(\sin^2\beta - \sin^2\gamma)}{\mathsf{SNR} + m\sin^2\gamma}$.

Rayleigh Fading: For special case of Rayleigh fading, which is obtained by setting $m = 1$, the expression in (32) can be re-expressed as (33) shown at the bottom of the page.

Furthermore, for 2-bit quantization with Rayleigh fading (i.e., $M = 4$, $n = 2$ and $m = 1$), we can obtain $p(\mathsf{SNR})$ in closed form as

$$p(\mathsf{SNR}) = \frac{2}{\pi}\arctan\left(\frac{1}{\sqrt{\mathsf{SNR}}}\right) - \left(\frac{1}{\pi}\arctan\left(\frac{1}{\sqrt{\mathsf{SNR}}}\right)\right)^2.$$

This closed-form analytical expression is very easy to compute without resorting to any numerical integration.

B. ANALYSIS OF QUANTIZATION PENALTY FOR QPSK MODULATION

By using the Taylor series expansion for high SNR values, we can re-express the average SEP expressions for QPSK modulation under Rayleigh fading given in (33) as (34) shown at the bottom of the page.

While phase quantization with less number of quantization bits is desirable, due to less processing complexity at the receiver, it deteriorates the average SEP performance of the system. In the following, we quantify the increase in the average SEP as a quantization penalty defined as

$$\Psi(\mathsf{SNR}, n) = 10\log\left(\frac{p_A(\mathsf{SNR}, n)}{p_A(\mathsf{SNR}, \infty)}\right), \quad (35)$$

where $p_A(\mathsf{SNR}, \infty)$ is the average SEP with infinite number of quantization bits. Based on (34), we can derive $p_A(\mathsf{SNR}, \infty)$ as

$$p_A(\mathsf{SNR}, \infty) = \left(\frac{4\pi - 1}{\pi^2}\right)\mathsf{SNR}^{-1} + o\left(\mathsf{SNR}^{-1}\right), \quad (36)$$

where we have used the small-angle approximation $\tan(x) = x$ as $x \to 0$. Substituting (34) and (36) into (35) and doing some mathematical manipulations, we can derive the quantization penalty in terms of average SEP with n-bit quantization as (37) shown at the bottom of the page.

$$p(\mathsf{SNR}) = \frac{2^n}{\pi^2}\int_0^{\frac{\pi}{2}} \frac{\sin\beta}{\sqrt{\mathsf{SNR} + \sin^2\beta}} \arctan\left(\frac{2\sin\beta\sqrt{\mathsf{SNR} + \sin^2\beta}}{\mathsf{SNR} + 2\sin^2\beta}\tan\left(\frac{\pi}{2^{n-1}}\right)\right) d\beta - \frac{2^{n-1}}{\pi^3}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} \sqrt{\frac{\sin^2\beta\sin^2\gamma}{(\mathsf{SNR} + \sin^2\beta)(\mathsf{SNR} + \sin^2\gamma)}} \cdot \arctan(\vartheta)\, d\beta\, d\gamma \quad (33)$$

where $\vartheta = \dfrac{2\sin\beta\sin\gamma\sqrt{(\mathsf{SNR} + \sin^2\beta)(\mathsf{SNR} + \sin^2\gamma)}}{\mathsf{SNR}(\sin^2\beta + \sin^2\gamma) + 2\sin^2\beta\sin^2\gamma}\tan\left(\frac{\pi}{2^{n-1}}\right)$

$$p_A(\mathsf{SNR}, n) = \begin{cases} \frac{2}{\pi}\mathsf{SNR}^{-\frac{1}{2}} + o\left(\mathsf{SNR}^{-\frac{1}{2}}\right) & n = 2 \\ \frac{2^{n-1}(4\pi - 1)}{\pi^3}\tan\left(\frac{\pi}{2^{n-1}}\right)\mathsf{SNR}^{-1} + o\left(\mathsf{SNR}^{-1}\right) & n \geq 3 \end{cases} \quad (34)$$

$$\Psi(\mathsf{SNR}, n) = \begin{cases} 10\log\left(\left(\frac{2\pi}{4\pi - 1}\right)\mathsf{SNR}^{\frac{1}{2}}\right) + o\left(\mathsf{SNR}^{\frac{1}{2}}\right) & n = 2 \\ 10\log\left(\frac{2^{n-1}}{\pi}\tan\left(\frac{\pi}{2^{n-1}}\right)\right) + o(1) & n \geq 3 \end{cases} \quad (37)$$

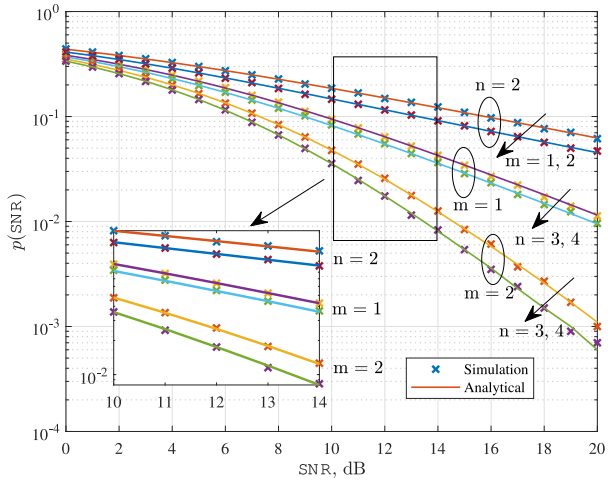

FIGURE 4. Average SEP curves as a function of SNR for QPSK modulation. $n = 2, 3, 4$ and $m = 1, 2$.

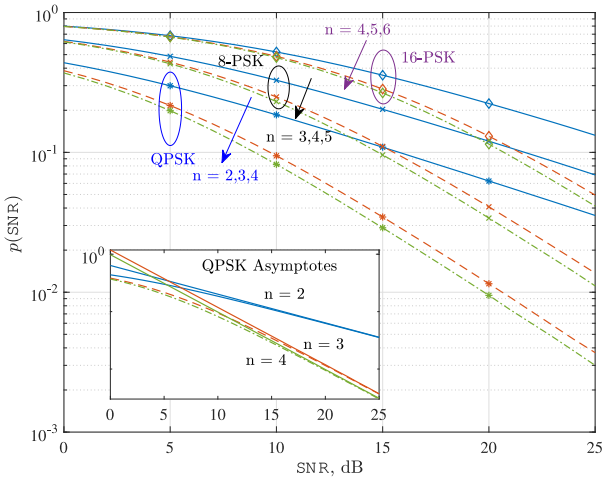

FIGURE 5. Average SEP curves as a function of SNR for different modulation schemes. $n = \log_2 M, \log_2 M + 1, \log_2 M + 2$ and $m = 1$.

In Section VII, we use $\Psi(\mathsf{SNR}, n)$ to quantify the increase in average SEP as we change from infinite-bit to n-bit quantization.

We further notice that, in order to achieve the same average SEP as with n-bit quantization, we need to transmit the signal using a higher power if we use only $(n-1)$-bit quantization. In the following, we quantify the increase in the transmit power as another quantization penalty defined by

$$\Phi(\mathsf{SEP}, n) = 10\log\left(\frac{\mathsf{SNR}_{n-1}}{\mathsf{SNR}_n}\right), \tag{38}$$

where SNR_n and SNR_{n-1} are the SNR values required to achieve a certain average SEP with n and $n-1$ quantization bits, respectively. Substituting (34) into (38) and doing some mathematical manipulations, we can derive the quantization penalty with n-bit quantization as

$$\Phi(\mathsf{SEP}, n) = \begin{cases} 10\log\left(\frac{\pi^2}{2(4\pi-1)}(\mathsf{SNR}_2)^{\frac{1}{2}}\right) & n = 3 \\ 10\log\left(\frac{1}{2}\tan\left(\frac{\pi}{2^{n-2}}\right)\cot\left(\frac{\pi}{2^{n-1}}\right)\right) & n \geq 4. \end{cases} \tag{39}$$

In Section VII, we use $\Phi(\mathsf{SEP}, n)$ to quantify the required transmit power increase as we change from n-bit to $(n-1)$-bit quantization.

VII. NUMERICAL RESULTS

In this section, we present analytical and simulated SEP results for M-PSK modulation with n-bit quantization. Channel fading is unit-power and circularly-symmetric with Nakagami-m distributed magnitude, and additive noise is complex Gaussian with zero mean and unit variance.

A. EFFECT OF QUANTIZATION BITS ON SYSTEM PERFORMANCE

First, we will investigate the effect of the number of quantization bits on the system performance and illustrate the accuracy of our analytical results obtained in previous sections. Fig. 4 plots the average SEP as a function of SNR for QPSK modulation with $n = 2, 3, 4$-bit quantization under Nakagami-m fading with shape parameter $m = 1$ and 2. The simulated results are generated using Monte Carlo simulation, while the analytical results are generated using our expression in (11). As the plot illustrates, the analytical results accurately follow the simulated results for all cases. We observe a noteworthy improvement in the average SEP when n changes from 2 to 3-bit quantization for QPSK modulation in both $m = 1$ and 2. This jump in the average SEP performance is expected in the light of Theorem 4, which states that using one extra bit, on top of $\log_2 M$ bits, improves the DVO from $\frac{1}{2}$ to m. We also observe that the average SEP reduces as we increase n, but the amount by which it reduces also gets smaller as we increase n. This can be clearly observed from the zoomed-in section in Fig. 4. As expected, $\mathsf{DVO} = m$ for all $n \geq 3$. Furthermore, $\mathsf{DVO} = \frac{1}{2}$ for any m, when $n = 2$.

Fig. 5 plots the average SEP as a function of SNR for QPSK, 8-PSK and 16-PSK modulations schemes while keeping the Nakagami-m shape parameter fixed at $m = 1$, which is the classical Rayleigh fading scenario. We plot the average SEP for each modulation scheme by using $n = \log_2 M$, $\log_2 M + 1$ and $\log_2 M + 2$ bits. From the plots, we can clearly observe that QPSK with 2-bit, 8-PSK with 3-bit and 16-PSK with 4-bit quantization have a DVO of $\frac{1}{2}$. Further, we can observe that QPSK with 3 or more bits, 8-PSK with 4 or more bits, 16-PSK with 5 or more bits quantizations have a DVO of 1, which is equal to m in this case. To further emphasize this point, the zoomed-in section in Fig. 5 illustrates the asymptotic average SEP versus SNR for QPSK modulation. As stated in Theorem 4, these numerical observations clearly indicate the ternary behaviour in the decay exponent for $p(\mathsf{SNR})$ depending on whether $n \geq \log_2 M + 1$, $n = \log_2 M$, or $n < \log_2 M$.

In order to illustrate the accuracy of upper and lower bounds on $p(\mathsf{SNR})$, derived in Lemma 5, in Fig. 6 we

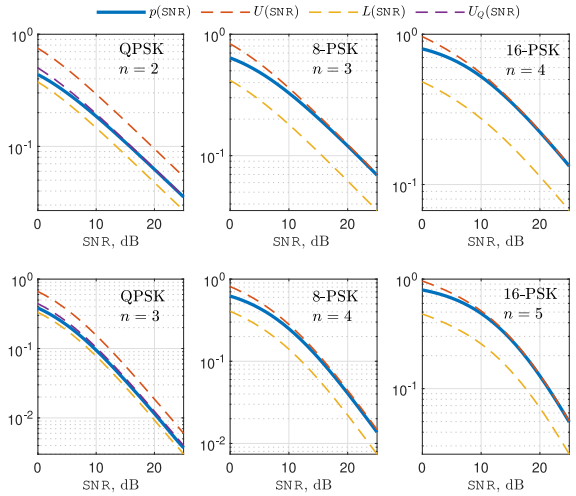

FIGURE 6. Upper and lower bounds on $p(\mathsf{SNR})$ as a function of SNR for QPSK, 8-PSK and 16-PSK modulations. $n = 2, 3, 4, 5$ and $m = 1$.

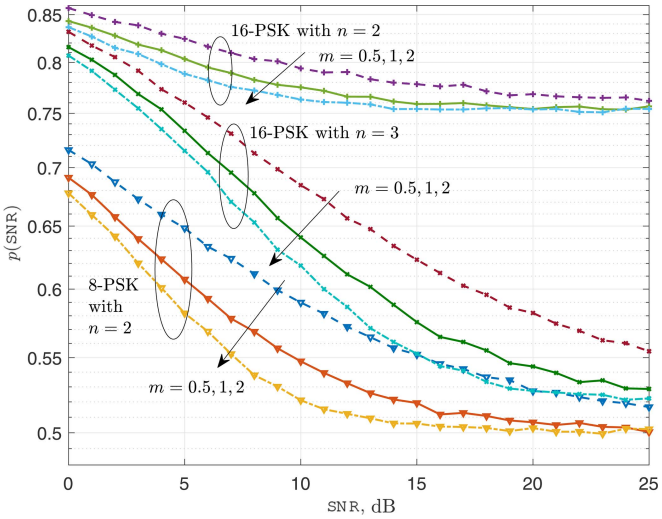

FIGURE 7. Average SEP as a function of SNR for 8-PSK and 16-PSK modulations. $n = 2 < \log_2 M$ and $m = 0.5, 1$ and 2.

plot the expressions in (24), alongside the exact $p(\mathsf{SNR})$ curve, as a function of SNR for QPSK, 8-PSK and 16-PSK modulations under Rayleigh fading (i.e., $m = 1$). This figure clearly shows that $U(\mathsf{SNR})$ becomes a very tight upper bound for 8-PSK and 16-PSK in the high SNR regime. The figure also confirms that the decay exponents of both $L(\mathsf{SNR})$ and $U(\mathsf{SNR})$ are the same as that of the $p(\mathsf{SNR})$.

Next, in Fig. 7, we plot the simulated average SEP curves as a function of SNR for 8-PSK modulation with 2-bit quantization and 16-PSK modulation with 2 and 3-bit quantization. We consider equi-probable transmitted symbols and Nakagami-m fading channel model with $m = 0.5, 1$ and 2. The simulated results are again generated by using Monte Carlo simulations. We can clearly observe an error floor for high SNR values when $n < \log_2 M$, as established by Theorem 1. In particular, the average SEP for 8-PSK has

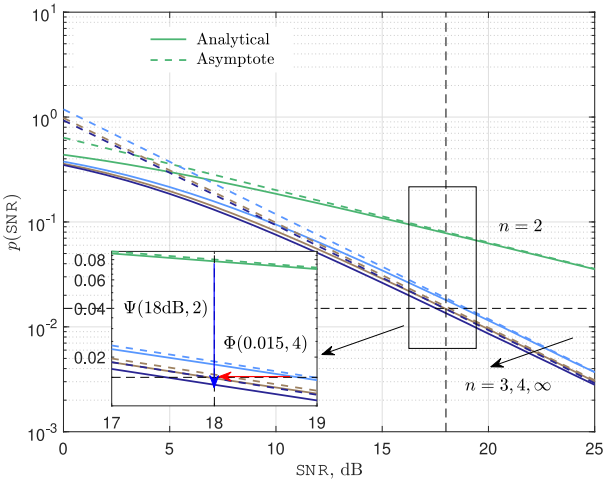

FIGURE 8. Quantization penalty for QPSK modulation and the asymptotic average SEP curves. $n = 2, 3, 4, \infty$ and at $m = 1$.

a lower bound of 0.5 with 2-bit quantization. Similarly, the average SEP for 16-PSK has a lower bound of 0.75 with 2-bit quantization and a lower bound of 0.5 with 3-bit quantization. It should be noted that the error floor given in Theorem 1 is more conservative than those observed in Fig. 7. This is because it is a universal lower bound that holds for all modulation schemes, quantizer types and fading environments, not only for very specific ones used to plot average SEP curves in Fig. 7.

Fig. 8 illustrates the quantization penalty and plots the asymptotic average SEP curves as a function of SNR for QPSK modulation with $n = 2$, 3, 4 and ∞ under Rayleigh fading. The asymptotic plots are generated by using the expressions in (34). We observe that a DVO of half is achieved with 2-bit quantization, and the full DVO of one is achieved with $n > 2$. When the SNR is fixed at 18 dB, we observe a quantization penalty of $\Psi(18\text{ dB}, 2) \approx 6.35$ dB as we change from $n = 2$ to ∞, i.e., we get a 5-fold increase in the average SEP as we change from $n = 2$ to ∞. When the average SEP is fixed at 0.015, we observe a quantization penalty of $\Phi(0.015, 4) \approx 0.8$ dB as we change from $n = 3$ to 4, i.e., to achieve an average SEP of 0.015 with 4-bit quantization, we need 0.8 dB more transmit power than what is required with 3-bit quantization.

B. COMPARISON WITH INDEPENDENT I AND Q QUANTIZATION

In this part, we will compare and contrast the system performance results obtained by using M-PSK modulation with the commonly adopted approach using independent I and Q channel quantization. To this end, Fig. 9 provides performance comparison curves for Rayleigh fading and QPSK modulation. We set the number of quantization bits to $n = 2, 3, 4$ and 6 by numerically optimizing the quantization regions for independent I and Q quantization. In order to cover all possibilities in the latter case, we also consider,

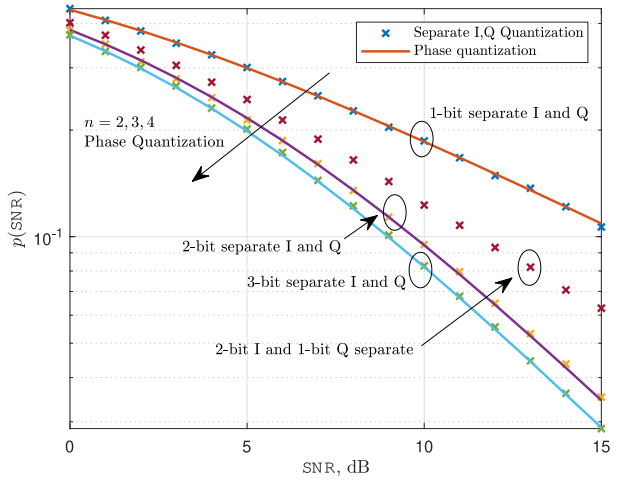

FIGURE 9. Average SEP as a function of SNR for QPSK modulation with phase quantization and separate I and Q quantization, $m = 1$.

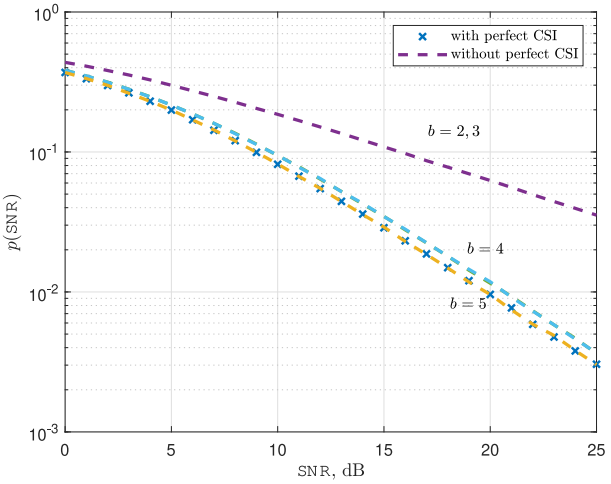

FIGURE 10. Average SEP as a function of SNR for QPSK modulation with 4-bit quantization and imperfect CSI, $m = 1$, b is the number of quantization bits used for channel estimation.

in addition to the identical bit allocation strategy, unbalanced allocation of quantization bits to the I and Q arms.

For $n = 2$, 2-bit phase quantization SEP performance coincides with that of using 1-bit quantization on each of the I and Q arms. The reason for this behaviour is that the optimized quantization regions coincide with the phase quantization regions, as introduced in Section II in the form of convex cones in $\mathbb{C}$. When $n = 3$, the 3-bit phase quantization clearly outperforms independent I and Q quantization with 2 bits on the I arm and 1 bit on the Q arm (and with optimized quantization regions). When $n = 4$, perhaps more interestingly, we observe that independent I and Q quantization with optimized quantization regions for 2 bits on both arms performs very similarly to the 3-bit phase quantization. Hence, for this particular allocation of quantization bits to the I and Q arms, the phase quantization approach introduced in this paper has the advantange of saving one quantization bit when compared to the commonly used independent I and Q quantization.

To test the robustness of our observations above, we have also obtained the SEP curves for the 6-bit independent I and Q quantization, with 3 bits on each of the I and Q arms. Our results indicate that the SEP performance for 6-bit independent I and Q quantization is similar to 4-bit phase quantization. Surprisingly, we need two more additional bits in the case of independent I and Q quantization in order to match the SEP performance attained by 4-bit phase quantization. Similar results with an increased performance gap continue to hold with unbalanced allocation of bits on the I and Q arms. While these results provides some initial insights into the bit-advantage of using phase quantization over independent I and Q quantization, a more thorough investigation is required to obtain general proofs establishing this phenomenon rigorously.

C. EFFECT OF CHANNEL ESTIMATION ERRORS ON DVO

In this part, we provide a robustness analysis for our analytical results under channel estimation errors. As we discussed above, we have assumed the availability of full-precision CSI at the receiver to derive the DVO results given in Theorem 4. This assumption was motivated by the previous work establishing the feasibility of accurate channel estimation by means of carefully designed training sequences for low-resolution ADC receivers as well as by the availability of mixed-ADC architectures. Below, we will focus on Rayleigh fading ($m = 1$) and QPSK ($M = 4$) modulation but our conclusions continue to hold more generally.

In order to model channel estimation errors in the numerical study below, we first recall that our ML detector in Theorem 2 only requires the knowledge of channel phase to determine the most likely transmitted symbol. Hence, we can model the channel estimation problem as that of estimating the phase information of H.[5] For the purpose of analyzing the effect of channel estimation errors on the SEP performance, this shows that the channel estimation errors due to low-resolution quantization can be modelled as bounded uniform random variables with the MMSE estimator being the mid-point of the quantization regions (i.e., H has a circularly symmetric distribution). Hence, the Figs. 10 and 11 are generated based on the assumption of bounded uniform channel estimation errors, with span of $\frac{2\pi}{2^b}$ when b bits are used for channel estimation.[6]

A salient feature arising from the SEP curves presented in Figs. 10 and 11 is that the derived fundamental ternary behaviour for the decay exponent of $p(\mathrm{SNR})$ in Theorem 4 stays the same if $n + 1$ bits are used during the channel estimation stage. We recall that n is the number of bits used for data decoding. Moreover, the performance difference between using n and $n+1$ bits for channel estimation becomes negligibly small as n increases. We note that

5. In addition to assisting in modeling channel estimation errors, this is also practically an important observation since it signifies that the same phase quantizer in a mixed-ADC architecture can be used for both channel estimation and data decoding by only dynamically changing the number of quantization bits.

6. Here, we ignore the channel noise to focus on channel estimation errors due to quantization noise only.

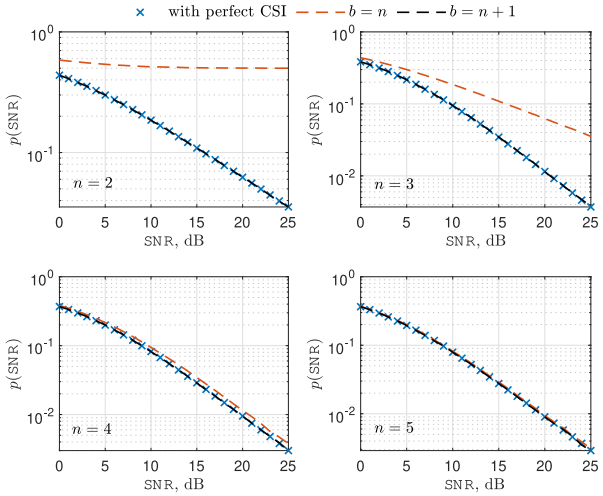

FIGURE 11. Average SEP as a function of SNR for QPSK modulation with n-bit quantization with and without perfect CSI, $m = 1$, $n = 2, 3, 4, 5$, $b = n, n+1$, where b is the number of quantization bits used for channel estimation.

the use of different number of bits for data decoding and channel estimation is possible by employing mixed-ADC receiver architectures commonly used in the literature [28]. Furthermore, the same effect to improve channel estimation accuracy can be achieved by using carefully designed training sequences to perform a bisection search in $[-\pi, \pi)$ during the channel estimation stage.

The robustness of our results against channel estimation errors in Figs. 10 and 11 arises from the following two key properties of the approach employed in the paper. The first property leading to the observed robustness behaviour is the fact that the optimum ML detector we obtained in Theorem 2 is only dependent on the channel phase information. The second one is the centering property achieved by the number of quantization bits steering the received symbols away from the decision boundaries and leading to the observed ternary behaviour for the decay exponent of $p(\mathsf{SNR})$ with full-precision CSI, as explained in Section IV-B in detail. The same effect continues to hold for the channel estimation, with one more additional bit providing the required accuracy in channel phase estimation and thereby preventing the bounded channel phase estimation errors from causing the received symbols to flip over the adjacent quantization regions.

VIII. CONCLUSION AND FUTURE GENERALIZATIONS

In this paper, we performed a theoretical analysis of a low-resolution based ADC communication system and obtained fundamental performance limits, optimum ML detectors and a general analytical expression for the average SEP for M-PSK modulation with n-bit quantization. These results were further investigated for Nakagami-m fading model in detail. We conducted an asymptotic analysis to show that the decay exponent for the average SEP is the same and equal to m with infinite-bit and n-bit quantizers for $n \geq \log_2 M + 1$. We also performed an extensive numerical study to illustrate the accuracy of the derived analytical expressions, comparison with independent I and Q channel quantization and robustness under channel estimation errors.

In most parts of the paper, we have focused on phase modulated communications. Phase modulation has an important and practical layering feature enabling the quantizer and detector design separation in low-resolution ADC communications. For a given number of bits, the quantizer needs to be designed only once, and can be kept constant for all channel realizations. The detector can be implemented digitally as a table look-up procedure using channel knowledge and quantizer output. On the other hand, this feature is lost in joint phase and amplitude modulation schemes such as QAM. The quantizer needs to be dynamically updated for each channel realization in low-resolution ADC based QAM systems. This is because the fading channel amplitude may vary over a wide range, but the phase always varies over $[-\pi, \pi)$. However, phase modulation is historically known to be optimum only up to modulation order 16 under peak power limitations [54]. Hence, it is a notable future research direction to extend the results of this paper to higher order phase and amplitude modulations by taking practical design considerations into account.

A major result of this paper is the discovery of a ternary SEP behaviour, indicating the sufficiency of $\log_2 M + 1$ bits for achieving asymptotically optimum M-ary communication reliability. Hence, without modifying the conventional RF chain, we can use one extra bit and still achieve the asymptotically optimum communication performance. Another important future research direction is to compare and contrast the backward-compatible receiver design approach of using one extra bit proposed in this paper with other approaches that can potentially modify the conventional RF chain and manipulate the received signals in the waveform domain by introducing extra analog components. This study needs to be done in detail by considering accuracy and agility of analog domain operations, energy consumption of analog and digital circuit components, different modulation schemes and the average SEP performance curves resulting from different low-resolution ADC based receiver architectures. Similarly, utilizing the results of this paper, a further detailed study on the receiver architecture design to determine where to place the diversity combiner (before or after quantizer or detector) and its type is needed when multiple diversity branches are available for data reception.

APPENDIX A
PROOF OF THEOREM 1

Let us consider a class of hypothetical genie-aided detectors $g : \mathbb{C}^2 \times \left[0 : 2^{n-1}\right] \to [0 : M-1]$ that has the knowledge of channel noise $W \in \mathbb{C}$, fading coefficient $H \in \mathbb{C}$ and quantizer output $Q(Y) \in \left[0 : 2^{n-1}\right]$. We also let $\mathcal{S}_{w,h,k} = \left\{x \in \mathcal{C} : \sqrt{\mathsf{SNR}}hx + w \in \mathcal{R}_k\right\}$ be the set of received signal points resulting in $Q(Y) = k$ for particular realizations of $H = h$ and $W = w$. We first observe that since $n < \log_2 M$, there exists at least one quantization region $\mathcal{R}_{\tilde{k}}$ (depending

on w and h) such that $\mathcal{S}_{w,h,\tilde{k}}$ contains at least $\frac{M}{2^n}$ signal points. We note that $\frac{M}{2^n}$ is always an integer greater than 2 since M is assumed to be an integer power of 2. Then, the conditional SEP of any detector g given $W = w$ and $H = h$, which we will denote by $p_g(\mathsf{SNR}, h, w)$, can be lower-bounded as (40) shown at the bottom of the page.

By averaging with respect to w and h, we have $p_g(\mathsf{SNR}) \geq \frac{M-2^n}{2^n} p_{\min}$, where $p_g(\mathsf{SNR})$ is the average SEP corresponding to detector g. This concludes the proof since the obtained lower bound does not depend on the choice of modulation scheme, quantizer structure and detector rule, and hence holds for detectors not utilizing the knowledge of W for any choice of modulation scheme and quantizer structure.

APPENDIX B
PROOF OF THEOREM 2

To prove Theorem 2, we will first obtain the following result.

Lemma 6: Let $\mathcal{R}$ be a convex cone given by $\mathcal{R} = \{z \in \mathbb{C}: \alpha_1 \leq \mathsf{Arg}(z) \leq \alpha_2\}$ for $\alpha_1, \alpha_2 \in [-\pi, \pi)$, and $W_1 \sim \mathcal{CN}(\mu_1, 1)$ and $W_2 \sim \mathcal{CN}(\mu_2, 1)$ be proper complex Gaussian random variables with means satisfying $|\mu_1| = |\mu_2| = r$ for some $r > 0$. Then, $\mathsf{Pr}\{W_1 \in \mathcal{R}\} \geq \mathsf{Pr}\{W_2 \in \mathcal{R}\}$ if $|\mu_1 - z_{\text{mid}}| \leq |\mu_2 - z_{\text{mid}}|$, where $z_{\text{mid}} = re^{J\frac{\alpha_1+\alpha_2}{2}}$.

Proof: It is enough to show this result only for $\alpha_2 = -\alpha_1 = \alpha$. Otherwise, we can first rotate W_1, W_2 and $\mathcal{R}$ with $e^{-J\frac{\alpha_1+\alpha_2}{2}}$ and repeat the same calculations below. Let $g(\mu_i) = \mathsf{Pr}\{W_i \in \mathcal{R}\}$ for $i = 1, 2$, and assume $|\mu_1 - z_{\text{mid}}| \leq |\mu_2 - z_{\text{mid}}|$. There are multiple cases in which the inequality $|\mu_1 - z_{\text{mid}}| \leq |\mu_2 - z_{\text{mid}}|$ holds, which we will analyze one-by-one below.

First, we will consider the case in which both μ_1 and μ_2 lie outside $\mathcal{R}^\circ$, where $\mathcal{R}^\circ$ is the set of interior points of $\mathcal{R}$. This is the case shown in Fig. 12. To start with, we will assume $0 \leq \mathsf{Arg}(\mu_1) \leq \mathsf{Arg}(\mu_2) < \pi$. Then, for any $y \in \mathcal{R}$, the angle between the line segments $\mathcal{L}_{Oy}$ and $\mathcal{L}_{O\mu_1}$ is smaller than the one between the line segments $\mathcal{L}_{Oy}$ and $\mathcal{L}_{O\mu_2}$.[7] Hence, applying the cosine rule for the triangle formed by O, y and μ_1, and for the triangle formed by O, y and μ_2, it can be seen that $|y - \mu_1| \leq |y - \mu_2|$ for all $y \in \mathcal{R}$.[8] Therefore, $g(\mu_1) = \frac{1}{\pi}\int_{\mathcal{R}} \exp(-|y-\mu_1|^2)dy \geq \frac{1}{\pi}\int_{\mathcal{R}} \exp(-|y-\mu_2|^2)dy = g(\mu_2)$. Next, we assume $\mathsf{Arg}(\mu_2) \in [-\pi, 0)$ and $0 \leq \mathsf{Arg}(\mu_1) \leq |\mathsf{Arg}(\mu_2)| \leq \pi$. Let $\widetilde{W}$ be the auxiliary random variable distributed according to $\mathcal{CN}(\tilde{\mu}, 1)$ with $\tilde{\mu} = re^{J|\mathsf{Arg}(\mu_2)|}$, i.e., $\tilde{\mu}$ is the reflection of μ_2 around the real line. Symmetry around the real line implies that $g(\mu_2)$ is equal to $g(\tilde{\mu}) = \mathsf{Pr}\{\widetilde{W} \in \mathcal{R}\}$, which is less than $g(\mu_1)$ due to our arguments above. For $\mathsf{Arg}(\mu_1) \in [-\pi, 0)$, the same analysis still holds after reflecting μ_1 around the real line, leading to $g(\mu_1) \geq g(\mu_2)$ for all $\mu_1, \mu_2 \notin \mathcal{R}^\circ$ satisfying $|\mu_1 - z_{\text{mid}}| \leq |\mu_2 - z_{\text{mid}}|$.

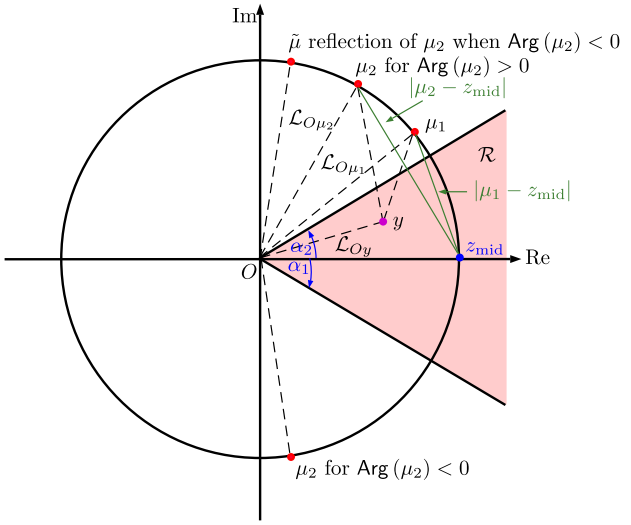

FIGURE 12. An illustration for the proof of Lemma 6 when μ_1 and μ_2 lie outside $\mathcal{R}^\circ$. $|\mu_1| = |\mu_2| = r$, $|\mu_1 - z_{\text{mid}}| \leq |\mu_2 - z_{\text{mid}}|$ and $\alpha_2 = -\alpha_1 = \alpha$.

Second, we consider the case where $\mu_1 \in \mathcal{R}^\circ$ but $\mu_2 \notin \mathcal{R}^\circ$. This is the case shown in Fig. 13. It is enough to establish the desired result only for $0 \leq \mathsf{Arg}(\mu_1) \leq \mathsf{Arg}(\mu_2) < \pi$. When μ_1 or μ_2 has a negative phase angle, the same analysis below still holds after reflecting the mean with negative phase around the real line. Let $\widetilde{W}$ be the auxiliary random variable distributed according to $\mathcal{CN}(\tilde{\mu}, 1)$ with $\tilde{\mu} = re^{J\alpha}$, i.e., $\tilde{\mu}$ is located at the upper boundary of $\mathcal{R}$. Our analysis in the first case shows that $g(\tilde{\mu}) = \mathsf{Pr}\{\widetilde{W} \in \mathcal{R}\} \geq g(\mu_2)$ since both $\tilde{\mu}$ and μ_2 are outside $\mathcal{R}^\circ$ and $0 \leq \mathsf{Arg}(\tilde{\mu}) \leq \mathsf{Arg}(\mu_2) < \pi$. We next divide $\mathcal{R}$ into two disjoint regions: $\mathcal{R}_1 = \{z \in \mathbb{C} : \mathsf{Arg}(\mu_1) \leq \mathsf{Arg}(z) \leq \alpha\}$ and $\mathcal{R}_2 = \{z \in \mathbb{C} : -\alpha \leq \mathsf{Arg}(z) < \mathsf{Arg}(\mu_1)\}$. Then, we have $\mathsf{Pr}\{\widetilde{W} \in \mathcal{R}_1\} = \mathsf{Pr}\{W_1 \in \mathcal{R}_1\}$ due to symmetry around the line bisecting $\mathcal{R}_1$ and $\mathsf{Pr}\{\widetilde{W} \in \mathcal{R}_2\} \leq \mathsf{Pr}\{W_1 \in \mathcal{R}_2\}$ since $|y - \mu_1| \leq |y - \tilde{\mu}|$ for all $y \in \mathcal{R}_2$. Hence, $g(\mu_1) \geq$

7. The line segment $\mathcal{L}_{z_1 z_2}$ between the points $z_1 \in \mathbb{C}$ and $z_2 \in \mathbb{C}$ is defined as $\mathcal{L}_{z_1 z_2} = \{(1-t)z_1 + tz_2 : t \in [0, 1]\}$.

8. This statement is correct even when both y and μ_1 lies on the boundary of $\mathcal{R}$ and the triangle formed by O, y and μ_1 reduces to a line segment.

$$
\begin{aligned}
p_g(\mathsf{SNR}, h, w) &\geq p_{\min} \sum_{x_i \in \mathcal{S}_{w,h,\tilde{k}}} \mathsf{Pr}\Big\{g\Big(h, w, \tilde{k}\Big) \neq x_i \,\Big|\, W = w, H = h, X = x_i\Big\} \\
&\geq p_{\min}\left(\frac{M}{2^n} - 1\right) \\
&= \frac{M - 2^n}{2^n} p_{\min}
\end{aligned} \tag{40}
$$

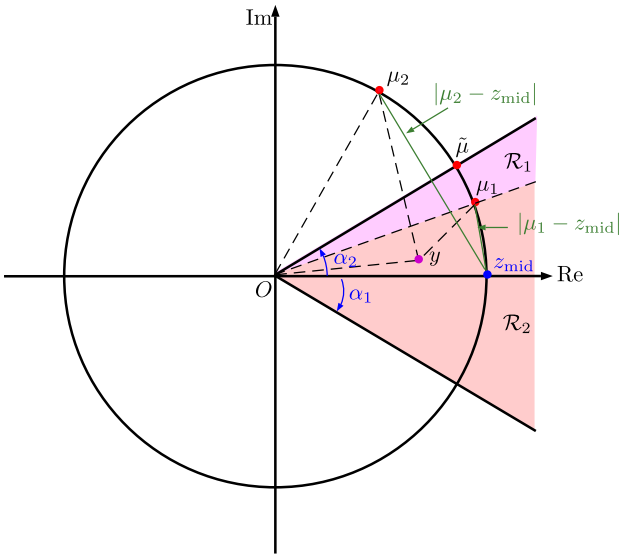

FIGURE 13. An illustration for the proof of Lemma 6 when $\mu_1 \in \mathcal{R}^\circ$ and $\mu_2 \notin \mathcal{R}^\circ$. $|\mu_1| = |\mu_2| = r$, $|\mu_1 - z_{\text{mid}}| \leq |\mu_2 - z_{\text{mid}}|$ and $\alpha_2 = -\alpha_1 = \alpha$.

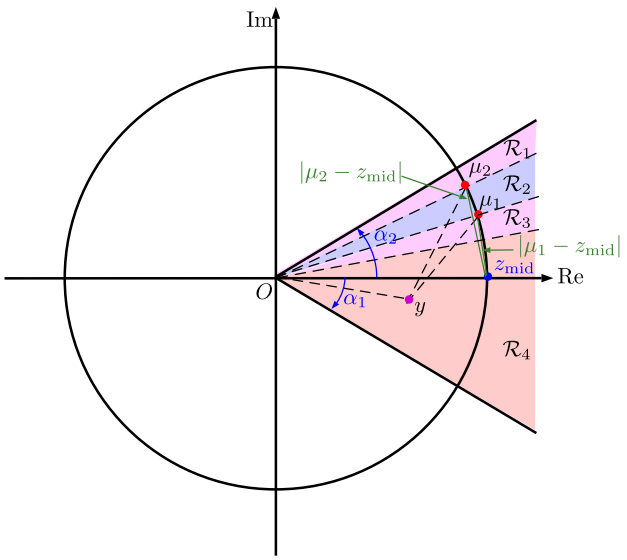

FIGURE 14. An illustration for the proof of Lemma 6 when μ_1 and μ_2 lie inside $\mathcal{R}^\circ$. $|\mu_1| = |\mu_2| = r$, $|\mu_1 - z_{\text{mid}}| \leq |\mu_2 - z_{\text{mid}}|$ and $\alpha_2 = -\alpha_1 = \alpha$.

$g(\tilde{\mu}) \geq g(\mu_2)$. This establishes the desired results for all $\mu_1 \in \mathcal{R}^\circ$, $\mu_2 \notin \mathcal{R}^\circ$ satisfying $|\mu_1 - z_{\text{mid}}| \leq |\mu_2 - z_{\text{mid}}|$.

Finally, we will consider the third case where both μ_1 and μ_2 lie inside $\mathcal{R}^\circ$. This is the case shown in Fig. 14. Similar to the first two cases, it is enough to focus only on $0 \leq \mathsf{Arg}(\mu_1) \leq \mathsf{Arg}(\mu_2) \leq \alpha$. We divide $\mathcal{R}$ into four disjoint regions given by

$$\mathcal{R}_1 = \{z \in \mathbb{C} : \mathsf{Arg}(\mu_2) \leq \mathsf{Arg}(z) \leq \alpha\}$$
$$\mathcal{R}_2 = \{z \in \mathbb{C} : \mathsf{Arg}(\mu_1) \leq \mathsf{Arg}(z) < \mathsf{Arg}(\mu_2)\}$$
$$\mathcal{R}_3 = \{z \in \mathbb{C} : \mathsf{Arg}(\mu_1) + \mathsf{Arg}(\mu_2) - \alpha \leq \mathsf{Arg}(z) < \mathsf{Arg}(\mu_1)\}$$
$$\mathcal{R}_4 = \{z \in \mathbb{C} : -\alpha \leq \mathsf{Arg}(z) < \mathsf{Arg}(\mu_1) + \mathsf{Arg}(\mu_2) - \alpha\}$$

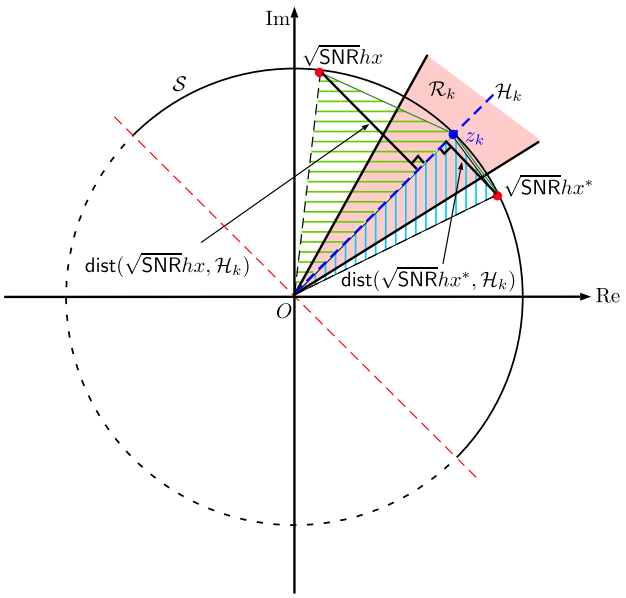

FIGURE 15. An illustration for the proof of Theorem 2 where $\left|\sqrt{\mathsf{SNR}}hx^\star - z_k\right| \leq \left|\sqrt{\mathsf{SNR}}hx - z_k\right|$.

Using the symmetry in the problem, we have $\mathsf{Pr}\{W_1 \in \mathcal{R}_1\} = \mathsf{Pr}\{W_2 \in \mathcal{R}_3\}$, $\mathsf{Pr}\{W_1 \in \mathcal{R}_2\} = \mathsf{Pr}\{W_2 \in \mathcal{R}_2\}$ and $\mathsf{Pr}\{W_1 \in \mathcal{R}_3\} = \mathsf{Pr}\{W_2 \in \mathcal{R}_1\}$. On the other hand, $\mathsf{Pr}\{W_1 \in \mathcal{R}_4\} \geq \mathsf{Pr}\{W_2 \in \mathcal{R}_4\}$ since $|y - \mu_1| \leq |y - \mu_2|$ for all $y \in \mathcal{R}_4$. Hence, $g(\mu_1) \geq g(\mu_2)$ when both μ_1 and μ_2 lie inside $\mathcal{R}^\circ$, which completes the proof. ■

Now, we will utilize Lemma 6 to prove Theorem 2. For $Q(Y) = k$, the ML detector given in (6) reduces to finding a signal point in $\mathcal{C}$ maximizing the probability $\mathsf{Pr}\left\{\sqrt{\mathsf{SNR}}hx + W \in \mathcal{R}_k\right\}$, i.e.,

$$\hat{x}(k, h) \in \arg\max_{x \in \mathcal{C}} \mathsf{Pr}\left\{\sqrt{\mathsf{SNR}}hx + W \in \mathcal{R}_k\right\}.$$

By Lemma 6, $\hat{x}(k, h)$ is the signal point in $\mathcal{C}$ such that $\sqrt{\mathsf{SNR}}h\hat{x}(k, h)$ is closest to $z_k = \sqrt{\mathsf{SNR}}re^{\jmath\left(\frac{2\pi}{2^n}k + \frac{\pi}{2^n}\right)}$, where $r = |h|$. Further, $\hat{x}(k, h)$ is unique with probability one due to the continuity assumption of the fading distribution. Consider now the semi-circle $\mathcal{S} = \{z \in \mathbb{C} : |z|\}$ such that $|z| = \sqrt{\mathsf{SNR}}r$ and $\left(\frac{2\pi}{2^n}k + \frac{\pi}{2^n} - \frac{\pi}{2}\right) \leq \mathsf{Arg}(z) \leq \left(\frac{2\pi}{2^n}k + \frac{\pi}{2^n} + \frac{\pi}{2}\right)$ centered around z_k and having $\mathcal{H}_k$ as its bisector. The semi-circle $\mathcal{S}$ is illustrated in Fig 15. Let $x^\star \in \arg\min_{x \in \mathcal{C}} \mathsf{dist}\left(\sqrt{\mathsf{SNR}}hx, \mathcal{H}_k\right)$. For the M-PSK modulation scheme ($M \geq 2$) with regularly spaced signal points on the unit circle, we always have $\sqrt{\mathsf{SNR}}hx^\star \in \mathcal{S}$ and $\sqrt{\mathsf{SNR}}h\hat{x}(k, h) \in \mathcal{S}$. Take now another signal point $x \in \mathcal{C}$ different than $x^\star$ and satisfying $\sqrt{\mathsf{SNR}}hx \in \mathcal{S}$. Consider the triangle formed by 0, z_k and $\sqrt{\mathsf{SNR}}hx^\star$, and the one formed by 0, z_k and $\sqrt{\mathsf{SNR}}hx$. We first observe that the area of the first triangle is smaller than the area of the second one since they share the line segment $\mathcal{L}_{Oz_k}$ as their common base but the height of the first one $\mathsf{dist}\left(\sqrt{\mathsf{SNR}}hx^\star, \mathcal{H}_k\right)$

corresponding to this base is smaller than the height of the second one $\mathsf{dist}\left(\sqrt{\mathsf{SNR}}hx, \mathcal{H}_k\right)$ corresponding to the same base. This is also illustrated in Fig. 15. This observation, in turn, implies $\left|\sqrt{\mathsf{SNR}}hx^\star - z_k\right| \leq \left|\sqrt{\mathsf{SNR}}hx - z_k\right|$ because the remaining side lengths of both triangles are equal to $\sqrt{\mathsf{SNR}}r$. Since this is correct for any $x \in \mathcal{C}$ satisfying $\sqrt{\mathsf{SNR}}hx \in \mathcal{S}$, we conclude that $x^\star$ is unique and equal to $x^\star = \hat{x}(k, h)$.

APPENDIX C
PROOF OF LEMMA 1

The proof of Lemma 1 is based on an application of the law of total probability [55]. To this end, we consider a partition $\{\mathcal{D}_k\}_{k=0}^{2^n-1}$ of $\mathbb{C}$, where each element of this partition is given by $\mathcal{D}_k = \left\{z \in \mathbb{C} : (2k-1)\frac{\pi}{2^n} \leq \mathsf{Arg}(z) + \pi < (2k+1)\frac{\pi}{2^n}\right\}$ for $k \in [1 : 2^n - 1]$, and $\mathcal{D}_0 = \left\{z \in \mathbb{C} : \pi - \frac{\pi}{2^n} \leq \mathsf{Arg}(z) < \pi\right\} \bigcup \left\{z \in \mathbb{C} : -\pi \leq \mathsf{Arg}(z) < \frac{\pi}{2^n} - \pi\right\}$. Let $x_i = \mathrm{e}^{\jmath\pi\left(\frac{2i+1}{M}-1\right)}$ be the ith signal point in the constellation set $\mathcal{C}$ for $i \in [0 : M-1]$. Then, we can express $p\big(\mathsf{SNR}\big)$ according to (41), as shown at the bottom of the page.

We will show that all the terms in (41), as shown at the bottom of this page, are equal to each other. Next, we define $\mathcal{E}_i = \left\{z \in \mathbb{C} : \mathsf{Arg}(x_i) - \frac{\pi}{M} \leq \mathsf{Arg}(z) < \mathsf{Arg}(x_i) + \frac{\pi}{M}\right\}$ for $i \in [0 : M-1]$. Note that $\mathcal{E}_i$ contains all $\mathcal{H}_k$'s (i.e., bisectors of quantization regions) to which x_i is the closest signal point since x_i's are uniformly spaced on the unit circle in $\mathbb{C}$. Furthermore, this statement continues to be true for $\sqrt{\mathsf{SNR}}hx_i$ as long as $\mathsf{Arg}(h) \in \left[-\frac{\pi}{2^n}, \frac{\pi}{2^n}\right)$ since the angular spacing between $\mathcal{H}_k$'s is uniform and equal to $\frac{2\pi}{2^n}$. Notice that $\mathsf{Arg}(h) \in \left[-\frac{\pi}{2^n}, \frac{\pi}{2^n}\right)$ if and only if $h \in \mathcal{D}_{2^{n-1}}$. On the other hand, if $\mathsf{Arg}(h) \in \left[\frac{\pi}{2^n}, \frac{3\pi}{2^n}\right)$, the region $\mathrm{e}^{\jmath\frac{2\pi}{2^n}}\mathcal{E}_i = \left\{\mathrm{e}^{\jmath\frac{2\pi}{2^n}}z \in \mathbb{C} : \mathsf{Arg}(x_i) - \frac{\pi}{M} \leq \mathsf{Arg}(z) < \mathsf{Arg}(x_i) + \frac{\pi}{M}\right\}$ contains all $\mathcal{H}_k$'s to which $\sqrt{\mathsf{SNR}}hx_i$ is the closest. Notice also that $\mathsf{Arg}(h) \in \left[\frac{\pi}{2^n}, \frac{3\pi}{2^n}\right)$ if and only if $h \in \mathcal{D}_{2^{n-1}+1}$. Similarly, $\mathrm{e}^{-\jmath\frac{2\pi}{2^n}}\mathcal{E}_i$ contains all $\mathcal{H}_k$'s to which $\sqrt{\mathsf{SNR}}hx_i$ is closest if $\mathsf{Arg}(h) \in \left[-\frac{3\pi}{2^n}, -\frac{\pi}{2^n}\right)$, and $\mathsf{Arg}(h) \in \left[-\frac{3\pi}{2^n}, -\frac{\pi}{2^n}\right)$ if and only if $h \in \mathcal{D}_{2^{n-1}-1}$. The same idea extends to any $\mathcal{D}_k$, and we define

$$\mathcal{E}_{i,k} \triangleq \exp\left(\jmath\left(k - 2^{n-1}\right)\frac{2\pi}{2^n}\right)\mathcal{E}_i, \tag{42}$$

for $i \in [0 : M-1]$ and $k \in [0{:}2^n - 1]$. We will use the sets defined in (42) to show that all the terms in (41) are equal.

To complete the proof, we let $p_{i,k} = \int_{\mathcal{D}_k} \mathsf{Pr}\{x_i \neq \hat{x}(Q(Y), h)|H = h, X = x_i\} f_H(h)dh$ for $i \in [0 : M-1]$ and $k \in [0 : 2^n - 1]$. We also define $\theta_i' = -\pi\left(\frac{2i}{M} - 1\right)$, $\theta_k'' = -\left(k - 2^{n-1}\right)\frac{2\pi}{2^n}$, and $\theta_{i,k} = \theta_i' + \theta_k''$ for $i \in [0 : M-1]$ and $k \in [0 : 2^n - 1]$. We first observe that $\mathrm{e}^{\jmath\theta_{i,k}}\mathcal{E}_{i,k} = \mathcal{E}_{\frac{M}{2}}$ since multiplication with $\mathrm{e}^{\jmath\theta_i'}$ rotates the ith signal point to $x_{\frac{M}{2}}$ and multiplication with $\mathrm{e}^{\jmath\theta_k''}$ removes the effect of partition selection for h. Secondly, we observe that when $h \in \mathcal{D}_k$, the event $\left\{x_i \neq \hat{x}(Q(Y), h)\right\}$ is equivalent to $\left\{Y \notin \mathcal{E}_{i,k}\right\}$ since $\mathcal{E}_{i,k}$ contains all bisectors to which $\sqrt{\mathsf{SNR}}hx_i$ is closest for this range of h values. Hence, the following chain of equalities hold:

$$\begin{aligned}
p_{i,k} &\overset{(a)}{=} \int_{\mathcal{D}_k} \mathsf{Pr}\left\{\sqrt{\mathsf{SNR}}hx_i + W \notin \mathcal{E}_{i,k}\right\} f_H(h)dh \\
&= \int_{\mathcal{D}_k} \mathsf{Pr}\left\{W \notin \mathcal{E}_{i,k} - \sqrt{\mathsf{SNR}}hx_i\right\} f_H(h)dh \\
&= \int_{\mathcal{D}_k} \mathsf{Pr}\left\{\mathrm{e}^{\jmath\theta_{i,k}}W \notin \mathrm{e}^{\jmath\theta_{i,k}}\mathcal{E}_{i,k} - \mathrm{e}^{\jmath\theta_{i,k}}\sqrt{\mathsf{SNR}}hx_i\right\} \\
&\quad f_H(h)dh \\
&\overset{(b)}{=} \int_{\mathcal{D}_k} \mathsf{Pr}\left\{W \notin \mathcal{E}_{\frac{M}{2}} - \sqrt{\mathsf{SNR}}\mathrm{e}^{\jmath\theta_k''}hx_{\frac{M}{2}}\right\} f_H(h)dh,
\end{aligned} \tag{43}$$

where (a) follows from the independence of W, H and X, and (b) follows from above observations and the circular symmetry property of W. Let us now define $z = \mathrm{e}^{\jmath\theta_k''}h$ in (43). Since multiplication with a unit magnitude complex number is a unitary transformation (i.e., rotation) over the complex plane, we have

$$\begin{aligned}
p_{i,k} &= \int_{\mathcal{D}_k} \mathsf{Pr}\left\{W \notin \mathcal{E}_{\frac{M}{2}} - \sqrt{\mathsf{SNR}}\mathrm{e}^{\jmath\theta_k''}hx_{\frac{M}{2}}\right\} f_H(h)dh \\
&= \int_{\mathrm{e}^{\jmath\theta_k''}\mathcal{D}_k} \mathsf{Pr}\left\{W \notin \mathcal{E}_{\frac{M}{2}} - \sqrt{\mathsf{SNR}}zx_{\frac{M}{2}}\right\} f_H\left(\mathrm{e}^{-\jmath\theta_k''}z\right)dz \\
&\overset{(a)}{=} \int_{\mathcal{D}_{2^{n-1}}} \mathsf{Pr}\left\{W \notin \mathcal{E}_{\frac{M}{2}} - \sqrt{\mathsf{SNR}}zx_{\frac{M}{2}}\right\} f_H(z)dz \\
&\overset{(b)}{=} \int_{\mathcal{D}_{2^{n-1}}} \mathsf{Pr}\left\{\sqrt{\mathsf{SNR}}zx_{\frac{M}{2}} + W \notin \mathcal{E}_{\frac{M}{2},2^{n-1}}\right\} f_H(z)dz \\
&\overset{(c)}{=} p_{\frac{M}{2},2^{n-1}},
\end{aligned} \tag{44}$$

where (a), (b) and (c) follow from the circular symmetry of H [40], [41] and the corresponding definitions for $\mathcal{D}_k$, $\mathcal{E}_{i,k}$ and $p_{i,k}$ for $i \in [0 : M-1]$ and $k \in [0 : 2^n - 1]$. This shows $p\big(\mathsf{SNR}\big) = 2^n p_{\frac{M}{2},2^{n-1}}$. For a circularly-symmetric pdf $f_H(h)$, it is well-known that $rf_H(r\cos\lambda, r\sin\lambda) = \frac{1}{2\pi}f_R(r)$ [56]. Switching to polar coordinates, and using the identities $rf_H(r\cos\lambda, r\sin\lambda) = \frac{1}{2\pi}f_R(r)$, $x_{\frac{M}{2}} = \mathrm{e}^{\jmath\frac{\pi}{M}}$ and $\mathcal{E}_{\frac{M}{2},2^{n-1}} = \mathcal{E}$, we have

$$\begin{aligned}
p(\mathsf{SNR}) = \frac{2^{n-1}}{\pi} \int_{-\frac{\pi}{2^n}}^{\frac{\pi}{2^n}} \int_0^\infty \mathsf{Pr}\left\{\sqrt{\mathsf{SNR}}r\mathrm{e}^{\jmath\left(\frac{\pi}{M}+\lambda\right)} + W \notin \mathcal{E}\right\} \\
f_R(r)\, dr\, d\lambda.
\end{aligned}$$

$$p\big(\mathsf{SNR}\big) = \frac{1}{M}\sum_{i=0}^{M-1}\sum_{k=0}^{2^n-1}\int_{\mathcal{D}_k} \mathsf{Pr}\{x_i \neq \hat{x}(Q(Y), h)|H = h, X = x_i\} f_H(h)dh \tag{41}$$

By using the change of variables $\theta = \frac{\pi}{M} + \lambda$, we conclude the proof.

APPENDIX D
PROOF OF LEMMA 2

The proof of part (i) follows immediately from Definition 1:

$$\lim_{\mathsf{SNR}\to\infty} \frac{\log\big(\alpha f(\mathsf{SNR})\big)}{\log \mathsf{SNR}} = \lim_{\mathsf{SNR}\to\infty} \frac{\log \alpha}{\log \mathsf{SNR}} + \lim_{\mathsf{SNR}\to\infty} \frac{\log f(\mathsf{SNR})}{\log \mathsf{SNR}} = d_i.$$

For the proof of part (ii), given any $\epsilon > 0$, let $c > 0$ be such that

$$\mathsf{SNR}^{d_i-\epsilon} \le f_i(\mathsf{SNR}) \le \mathsf{SNR}^{d_i+\epsilon} \tag{45}$$

for all $\mathsf{SNR} \ge c$ and $i \in [1{:}N]$. Let $i' = \arg\max_{i\in[1:N]} d_i$. Then, as $\mathsf{SNR} \to \infty$, we can write

$$\begin{aligned}
&\log\left(\sum_{i=1}^{N} f_i(\mathsf{SNR})\right) \\
&\le \log\left(\sum_{i=1}^{N} \mathsf{SNR}^{d_i+\epsilon}\right) \\
&= \log \mathsf{SNR}^{d_{\max}+\epsilon} + \log\left(1 + \sum_{\substack{i=1\\ i\neq i'}}^{N} \mathsf{SNR}^{d_i-d_{\max}}\right) \\
&= (d_{\max}+\epsilon)\log \mathsf{SNR} + o(1),
\end{aligned}$$

which implies $\sum_{i=1}^{N} f_i(\mathsf{SNR}) \overset{\mathrm{e}}{\le} d_{\max} + \epsilon$. Since $\epsilon > 0$ is arbitrary, we conclude that $\sum_{i=1}^{N} f_i(\mathsf{SNR}) \overset{\mathrm{e}}{\le} d_{\max}$. The other direction $\sum_{i=1}^{N} f_i(\mathsf{SNR}) \overset{\mathrm{e}}{\ge} d_{\max}$ follows from the same arguments, which completes the proof.

APPENDIX E
PROOF OF LEMMA 3

We start with the case $M = 4$ and $n = 2$, and obtain upper and lower bounds on $p_1(\mathsf{SNR})$ that will lead to the same exponential equality. For the upper bound, we write $p_1(\mathsf{SNR})$ as

$$\begin{aligned}
p_1(\mathsf{SNR}) &= \frac{2m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} \left(\frac{\mathsf{SNR}}{\sin^2\beta}\cos^2\theta + m\right)^{-m} d\theta d\beta \\
&\le \frac{m^m}{\pi}\int_0^{\frac{\pi}{2}} \Big(\mathsf{SNR}\cos^2\theta + m\Big)^{-m} d\theta.
\end{aligned}$$

Let $\theta^*(\mathsf{SNR})$ be such that $\cos\big(\theta^*(\mathsf{SNR})\big) = \mathsf{SNR}^{-\frac{1}{2}}$. Then, $\theta^*(\mathsf{SNR}) = \arccos\Big(\mathsf{SNR}^{-\frac{1}{2}}\Big) = \frac{\pi}{2} - \mathsf{SNR}^{-\frac{1}{2}} - o\Big(\mathsf{SNR}^{-\frac{1}{2}}\Big)$ as $\mathsf{SNR} \to \infty$. Using the fact that the line $1 - \frac{2}{\pi}\theta$ is a lower bound for $\cos\theta$ for $\theta \in \big[0, \frac{\pi}{2}\big]$, we have (46), as shown at the bottom of the page, for large values of SNR. Equation (46) shows that $p_1(\mathsf{SNR}) \overset{\mathrm{e}}{\le} \mathsf{SNR}^{-\frac{1}{2}}$ for all $m \ge \frac{1}{2}$ when $M = 4$ and $n = 2$. For the other direction, we obtain a lower bound on $p_1(\mathsf{SNR})$ as (47) shown at the bottom of the page.

We observe that $\int_0^{\frac{\pi}{2}} (\sin\beta)^{2m} d\beta = \frac{\sqrt{\pi}\,\Gamma\left(m+\frac{1}{2}\right)}{2\Gamma(m+1)}$ for $m > -\frac{1}{2}$ and let $c = \frac{m^m}{\pi^{1.5}}\frac{\Gamma\left(m+\frac{1}{2}\right)}{\Gamma(m+1)}$. We first consider the case $m = \frac{1}{2}$. Then, as $\mathsf{SNR} \to \infty$, we have

$$\begin{aligned}
p_1(\mathsf{SNR}) &\ge c\mathsf{SNR}^{-\frac{1}{2}} \int_0^{\frac{\pi}{2}} \Big(\cos^2\theta + 1\Big)^{-\frac{1}{2}} d\theta \\
&= \mathsf{SNR}^{-\frac{1}{2}}\Omega(1).
\end{aligned} \tag{48}$$

For $m > \frac{1}{2}$, we define $\theta^*(\mathsf{SNR})$ as above and lower bound $p_1(\mathsf{SNR})$ for large values of SNR as

$$\begin{aligned}
p_1(\mathsf{SNR}) &\ge c\mathsf{SNR}^{-m} \int_{\theta^*(\mathsf{SNR})}^{\frac{\pi}{2}} \Big(\cos^2\theta + \frac{m}{\mathsf{SNR}}\Big)^{-m} d\theta \\
&\ge c\mathsf{SNR}^{-m} \int_{\theta^*(\mathsf{SNR})}^{\frac{\pi}{2}} \left(\frac{1}{\mathsf{SNR}} + \frac{m}{\mathsf{SNR}}\right)^{-m} d\theta \\
&= \mathsf{SNR}^{-\frac{1}{2}} c(1+m)^{-m}(1+o(1)) \\
&= \mathsf{SNR}^{-\frac{1}{2}}\Omega(1).
\end{aligned} \tag{49}$$

$$\begin{aligned}
p_1(\mathsf{SNR}) &\le \frac{m^m}{\pi}\int_0^{\theta^*(\mathsf{SNR})} \Big(\mathsf{SNR}\cos^2\theta + m\Big)^{-m} d\theta + \frac{m^m}{\pi}\int_{\theta^*(\mathsf{SNR})}^{\frac{\pi}{2}} \Big(\mathsf{SNR}\cos^2\theta + m\Big)^{-m} d\theta \\
&\le \frac{m^m}{\pi}\mathsf{SNR}^{-m}\int_0^{\theta^*(\mathsf{SNR})} \Big(1 - \frac{2}{\pi}\theta\Big)^{-2m} d\theta + \frac{1}{\pi}\mathsf{SNR}^{-\frac{1}{2}}(1+o(1)) \\
&= -\frac{m^m}{2}\mathsf{SNR}^{-m}\int_1^{1-\frac{2}{\pi}\theta^*(\mathsf{SNR})} u^{-2m} du + \frac{1}{\pi}\mathsf{SNR}^{-\frac{1}{2}}(1+o(1)) \\
&= \begin{cases} \mathsf{SNR}^{-\frac{1}{2}} O(\log \mathsf{SNR}) & \text{if } m = \frac{1}{2} \\ \mathsf{SNR}^{-\frac{1}{2}} O(1) & \text{if } m > \frac{1}{2} \end{cases}
\end{aligned} \tag{46}$$

$$\begin{aligned}
p_1(\mathsf{SNR}) &= \frac{2m^m}{\pi^2}\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}} (\sin\beta)^{2m}\Big(\mathsf{SNR}\cos^2\theta + m\sin^2\beta\Big)^{-m} d\theta d\beta \\
&\ge \frac{2m^m}{\pi^2}\int_0^{\frac{\pi}{2}} (\sin\beta)^{2m} d\beta \int_0^{\frac{\pi}{2}} \Big(\mathsf{SNR}\cos^2\theta + m\Big)^{-m} d\theta
\end{aligned} \tag{47}$$

Using (48) and (49), we conclude that $p_1(\mathsf{SNR}) \stackrel{\mathrm{e}}{\geq} \mathsf{SNR}^{-\frac{1}{2}}$ for all $m \geq \frac{1}{2}$ when $M = 4$ and $n = 2$. Since $p_1(\mathsf{SNR})$ also satisfies $p_1(\mathsf{SNR}) \stackrel{\mathrm{e}}{\leq} \mathsf{SNR}^{-\frac{1}{2}}$ in this case, we have $p_1(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^{-\frac{1}{2}}$ for $m \geq \frac{1}{2}$, $M = 4$ and $n = 2$.

Next, we consider $M = 4$ and $n > 2$. In this case, we write $p_1(\mathsf{SNR})$ as

$$p_1(\mathsf{SNR}) = \frac{2^{n-1}m^m}{\pi^2}\mathsf{SNR}^{-m} \int_0^{\frac{\pi}{2}} \int_{\frac{\pi}{4}-\frac{\pi}{2^n}}^{\frac{\pi}{4}+\frac{\pi}{2^n}} \left(\frac{\cos^2\theta}{\sin^2\beta} + \frac{m}{\mathsf{SNR}}\right)^{-m} d\theta d\beta. \quad (50)$$

Let $g_{\mathsf{SNR}}(\theta,\beta) = \left(\frac{\cos^2\theta}{\sin^2\beta} + \frac{m}{\mathsf{SNR}}\right)^{-m}$ be a collection of functions indexed by SNR. These functions increase to $g_\infty(\theta,\beta) = \left(\frac{\cos\theta}{\sin\beta}\right)^{-2m}$ as $\mathsf{SNR} \to \infty$. Further, $\int_0^{\frac{\pi}{2}} \int_{\frac{\pi}{4}-\frac{\pi}{2^n}}^{\frac{\pi}{4}+\frac{\pi}{2^n}} g_\infty(\theta,\beta) d\theta d\beta < \infty$. Hence, as $\mathsf{SNR} \to \infty$, we conclude that

$$p_1(\mathsf{SNR}) = \mathsf{SNR}^{-m}\Theta(1) \quad (51)$$

by using the monotone convergence theorem [52], which implies $p_1(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^{-m}$. The proof for $M > 4$ and $n \geq \log_2 M$ is similar, and we omit it to avoid repetition.

APPENDIX F
PROOF OF LEMMA 4

For $M = 4$ and $n = 2$, $p_2(\mathsf{SNR}) = p_1(\mathsf{SNR})$, and hence the proof of $p_2(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^{-\frac{1}{2}}$ in this case directly follows from Lemma 3. For $M > 4$ and $n > \log_2 M$, a similar argument using the monotone convergence theorem as in the proof of Lemma 3 readily shows that $p_2(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^{-m}$ for all $m \geq \frac{1}{2}$. Therefore, we will only focus on the case $M > 4$ and $n = \log_2 M$ to complete the proof of Lemma 4.

For $M > 4$ and $n = \log_2 M$, we will obtain upper and lower bounds on $p_2(\mathsf{SNR})$ leading to the same exponential equality. For the upper bound, we have

$$p_2(\mathsf{SNR}) = \frac{2^{n-1}m^m}{\pi^2} \int_0^{\frac{\pi}{2}} \int_0^{\frac{2\pi}{M}} \left(\frac{\mathsf{SNR}}{\sin^2\beta}\sin^2\theta + m\right)^{-m} d\theta d\beta \leq \frac{2^{n-1}m^m}{\pi^2} \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \left(\frac{\mathsf{SNR}}{\sin^2\beta}\sin^2\theta + m\right)^{-m} d\theta d\beta$$

$$\stackrel{\mathrm{e}}{=} \mathsf{SNR}^{-\frac{1}{2}}, \quad (52)$$

where the last equality follows from the fact that $p_2(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^{-\frac{1}{2}}$ for $M = 4$ and $n = 2$ and Lemma 2.

For the lower bound, we have (53), as shown at the bottom of the page, where $c = \frac{2^{n-2}m^m}{\pi^{1.5}} \frac{\Gamma\left(m+\frac{1}{2}\right)}{\Gamma(m+1)}$.

Let $\theta^*(\mathsf{SNR})$ be such that $\sin(\theta^*(\mathsf{SNR})) = \mathsf{SNR}^{-\frac{1}{2}}$. Then, $\theta^*(\mathsf{SNR}) = \arcsin\left(\mathsf{SNR}^{-\frac{1}{2}}\right) = \mathsf{SNR}^{-\frac{1}{2}} + o\left(\mathsf{SNR}^{-\frac{1}{2}}\right)$ as $\mathsf{SNR} \to \infty$. Hence, as $\mathsf{SNR} \to \infty$, we have

$$p_2(\mathsf{SNR}) \geq c\mathsf{SNR}^{-m} \int_0^{\theta^*(\mathsf{SNR})} \left(\sin^2\theta + \frac{m}{\mathsf{SNR}}\right)^{-m} d\theta \geq c\mathsf{SNR}^{-m} \int_0^{\theta^*(\mathsf{SNR})} \left(\frac{1}{\mathsf{SNR}} + \frac{m}{\mathsf{SNR}}\right)^{-m} d\theta = \mathsf{SNR}^{-\frac{1}{2}}\Omega(1), \quad (54)$$

which implies $p_2(\mathsf{SNR}) \stackrel{\mathrm{e}}{\geq} \mathsf{SNR}^{-\frac{1}{2}}$. Since $p_2(\mathsf{SNR})$ also satisfies $p_2(\mathsf{SNR}) \stackrel{\mathrm{e}}{\leq} \mathsf{SNR}^{-\frac{1}{2}}$ in this case, we have $p_2(\mathsf{SNR}) \stackrel{\mathrm{e}}{=} \mathsf{SNR}^{-\frac{1}{2}}$ for $m \geq \frac{1}{2}$, $M > 4$ and $n = \log_2 M$.

APPENDIX G
PROOF OF LEMMA 5

We will prove this lemma for general circularly-symmetric fading processes. To this end, let $H = Re^{\jmath\Theta}$ be the circularly-symmetric fading coefficient with the joint phase and magnitude pdf $f_{R,\Theta}(r,\theta) = \frac{1}{2\pi} f_R(r)$ for $\theta \in [-\pi,\pi)$ and $r \geq 0$. In the proof of Theorem 3, we obtained the expression $p(\mathsf{SNR},h)$ with $h = re^{\jmath\theta}$ given by (55) as shown at the bottom of the page.

Below, we will always use $re^{\jmath\theta}$ as the polar coordinate representation of h, i.e., $r = |h|$ and $\theta = \mathsf{Arg}(h)$. Integrating $2^n p(\mathsf{SNR},h)$ with respect to $f_{R,\Theta}(r,\theta)$ for $\theta \in \left[\frac{\pi}{M} - \frac{\pi}{2^n}, \frac{\pi}{M} + \frac{\pi}{2^n}\right)$ and $r \geq 0$, and using the Nakagami-m pdf for $f_R(r)$ together with Craig's formula, we obtained the resulting $p(\mathsf{SNR})$ expression in Theorem 3. Here, we will not assume any specific functional form for $f_R(r)$.

We start with obtaining the lower bound $L(\mathsf{SNR})$ on $p(\mathsf{SNR})$. Let

$$p_1(\mathsf{SNR},h) = \mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\cos\theta\right)$$

$$p_2(\mathsf{SNR}) = \frac{2^{n-1}m^m}{\pi^2} \int_0^{\frac{\pi}{2}} \int_0^{\frac{2\pi}{M}} (\sin\beta)^{2m}\left(\mathsf{SNR}\sin^2\theta + m\sin^2\beta\right)^{-m} d\theta d\beta \geq c\mathsf{SNR}^{-m} \int_0^{\frac{2\pi}{M}} \left(\sin^2\theta + \frac{m}{\mathsf{SNR}}\right)^{-m} d\theta \quad (53)$$

$$p(\mathsf{SNR},h) = \mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\cos\theta\right) + \mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\sin\theta\right) - \mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\cos\theta\right)\mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\sin\theta\right) + \frac{1}{\sqrt{\pi}} \int_{-\sqrt{\mathsf{SNR}}r\cos\theta}^{\infty} \mathcal{Q}\left(\sqrt{2\mathsf{SNR}}r\sec\left(\frac{2\pi}{M}\right)\sin\left(\frac{2\pi}{M} - \theta\right) + \sqrt{2}\tan\left(\frac{2\pi}{M}\right)w\right)e^{-w^2} dw \quad (55)$$

$$p_2(\mathsf{SNR},h)=\mathcal{Q}\Big(\sqrt{2\mathsf{SNR}}r\sin\theta\Big)$$
$$p_3(\mathsf{SNR},h)=\mathcal{Q}\Big(\sqrt{2\mathsf{SNR}}r\cos\theta\Big)\mathcal{Q}\Big(\sqrt{2\mathsf{SNR}}r\sin\theta\Big)$$

and $p_4(\mathsf{SNR},h)$ be the last integral term in (55). For $i\in[1{:}4]$, $p_i(\mathsf{SNR})$ is defined to be the integral of $2^n p_i(\mathsf{SNR},h)$ with respect to $f_{R,\Theta}(r,\theta)$ for $\theta\in\left[\frac{\pi}{M}-\frac{\pi}{2^n},\frac{\pi}{M}+\frac{\pi}{2^n}\right)$ and $r\geq 0$. For the given integration range, $p_3(\mathsf{SNR},h)\leq\frac{1}{2}p_2(\mathsf{SNR},h)$ since the argument of the $\mathcal{Q}$-function is always positive. Hence, we have

$$\begin{aligned}p(\mathsf{SNR},h)&\geq p_1(\mathsf{SNR},h)+p_2(\mathsf{SNR},h)-p_3(\mathsf{SNR},h)\\&\geq p_1(\mathsf{SNR},h)+\frac{1}{2}p_2(\mathsf{SNR},h).\end{aligned}\tag{56}$$

After scaling with 2^n and integrating (56) with respect to $f_{R,\Theta}(r,\theta)$ over the above integration range, we have

$$\begin{aligned}p(\mathsf{SNR})&\geq p_1(\mathsf{SNR})+\frac{1}{2}p_2(\mathsf{SNR})\\&=L(\mathsf{SNR}).\end{aligned}\tag{57}$$

Next, we establish that $U(\mathsf{SNR})=p_1(\mathsf{SNR})+2p_2(\mathsf{SNR})$ is an upper bound on $p(\mathsf{SNR})$. To this end, we will show that $p_4(\mathsf{SNR})\leq p_2(\mathsf{SNR})$ for all $M\geq 4$. For $M=4$, this is trivial since $p_4(\mathsf{SNR},h)=0\leq p_2(\mathsf{SNR},h)$. For $M>4$, we define

$$p_5(\mathsf{SNR},h)=\tfrac{1}{\sqrt{\pi}}\textstyle\int_{-\infty}^{\infty}\mathcal{Q}(\varphi)\mathrm{e}^{-w^2}dw,$$

where $\varphi=\sqrt{2\mathsf{SNR}}r\sec\left(\frac{2\pi}{M}\right)\sin\left(\frac{2\pi}{M}-\theta\right)+\sqrt{2}w\tan\left(\frac{2\pi}{M}\right)$. We also define $p_5(\mathsf{SNR})$ to be the integral of $2^n p_5(\mathsf{SNR},h)$ with respect to $f_{R,\Theta}(r,\theta)$ for $\theta\in\left[\frac{\pi}{M}-\frac{\pi}{2^n},\frac{\pi}{M}+\frac{\pi}{2^n}\right)$ and $r\geq 0$.

We observe that $p_4(\mathsf{SNR})\leq p_5(\mathsf{SNR})$ since the integrands are always positive and the integral with respect to w is over the whole real line for $p_5(\mathsf{SNR},h)$. Thus, it will be enough to show $p_2(\mathsf{SNR})=p_5(\mathsf{SNR})$ to conclude the proof. For $\mathsf{SNR}=0$, this can be verified by using the identity $\mathcal{Q}(x)=1-\mathcal{Q}(-x)$. To prove the equality for all $\mathsf{SNR}\geq 0$, we define the function $f(\mathsf{SNR})=p_2(\mathsf{SNR})-p_5(\mathsf{SNR})$. It is enough to show that the derivative of $f(\mathsf{SNR})$, which we represent by $f'(\mathsf{SNR})$, is equal to zero everywhere in order to show $p_2(\mathsf{SNR})=p_5(\mathsf{SNR})$. This is because if $f'(\mathsf{SNR})$ is equal to zero for all $\mathsf{SNR}\geq 0$, then $f(\mathsf{SNR})$ must be a constant function. Since $f(0)=0$, we have $f(\mathsf{SNR})=p_2(\mathsf{SNR})-p_5(\mathsf{SNR})=0$ for all $\mathsf{SNR}\geq 0$. We devote the rest of the proof to showing that $f'(\mathsf{SNR})=0$ for all $\mathsf{SNR}\geq 0$.

Using the definition of the $\mathcal{Q}$-function, the derivative of $p_2(\mathsf{SNR})$ with respect to SNR, which we represent by $p_2'(\mathsf{SNR})$, is given by

$$\begin{aligned}&p_2'(\mathsf{SNR})\\&=\frac{2^{n-1}}{\pi}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\int_0^{\infty}\frac{dp_2(\mathsf{SNR},re^{\jmath\theta})}{d\mathsf{SNR}}f_R(r)drd\theta\\&=\frac{2^{n-1}}{\pi}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\int_0^{\infty}\frac{-r\sin\theta}{2\sqrt{\pi\mathsf{SNR}}}\mathrm{e}^{-\mathsf{SNR}r^2\sin^2\theta}f_R(r)drd\theta.\end{aligned}\tag{58}$$

Similarly, $p_5'(\mathsf{SNR})$ can be written as

$$\begin{aligned}&p_5'(\mathsf{SNR})\\&=\frac{2^{n-1}}{\pi}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\int_0^{\infty}\frac{dp_5(\mathsf{SNR},re^{\jmath\theta})}{d\mathsf{SNR}}f_R(r)drd\theta\\&=\frac{2^{n-1}}{\pi\sqrt{\pi}}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\int_0^{\infty}\frac{-A(r,\theta)}{2\sqrt{\pi\mathsf{SNR}}}I(r,\theta)f_R(r)drd\theta,\end{aligned}\tag{59}$$

where $A(r,\theta)=r\sec\left(\frac{2\pi}{M}\right)\sin\left(\frac{2\pi}{M}-\theta\right)$, $I(r,\theta)=\int_{-\infty}^{\infty}\mathrm{e}^{-\left(w^2+\left(A(r,\theta)\sqrt{\mathsf{SNR}}+Bw\right)^2\right)}dw$ and $B=\tan\left(\frac{2\pi}{M}\right)$. After completing the term in the exponent in $I(r,\theta)$ to square and using the affinity of the resulting expression to a Gaussian pdf, $I(r,\theta)$ can be shown to be equal to

$$I(r,\theta)=\sqrt{\pi}\cos\left(\tfrac{2\pi}{M}\right)\mathrm{e}^{-\mathsf{SNR}r^2\sin^2\left(\frac{2\pi}{M}-\theta\right)}.\tag{60}$$

Using (60) in (59), we have

$$\begin{aligned}p_5'(\mathsf{SNR})=\frac{2^{n-1}}{\pi}\int_{\frac{\pi}{M}-\frac{\pi}{2^n}}^{\frac{\pi}{M}+\frac{\pi}{2^n}}\int_0^{\infty}\frac{-r\sin\left(\frac{2\pi}{M}-\theta\right)}{2\sqrt{\pi\mathsf{SNR}}}\\\mathrm{e}^{-\mathsf{SNR}r^2\sin^2\left(\frac{2\pi}{M}-\theta\right)}f_R(r)drd\theta.\end{aligned}\tag{61}$$

Change of variables $u=\frac{2\pi}{M}-\theta$ in (61) shows that $p_2'(\mathsf{SNR})=p_5'(\mathsf{SNR})$, and hence $f'(\mathsf{SNR})=0$ as desired.

REFERENCES

[1] Q. Bai and J. A. Nossek, "Energy efficiency maximization for 5G multi-antenna receivers," *Trans. Emerg. Telecommun. Technol.*, vol. 26, pp. 3–14, Jan. 2015.

[2] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[3] B. Murmann, *ADC Performance Survey 1997–2020*. Accessed: Jul. 28, 2020. [Online]. Available: http://web.stanford.edu/~murmann/adcsurvey.html

[4] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 205–211, Jul. 2018.

[5] M. T. Ivrlac and J. A. Nossekh, "On MIMO channel estimation with single-bit signal-quantization," in *Proc. Int. ITG/IEEE Workshop Smart Antennas (WSA)*, Feb. 2007, pp. 1–7.

[6] T. M. Lok and V. K.-W. Wei, "Channel estimation with quantized observations," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, Aug. 1998, p. 333.

[7] J. Singh, S. Ponnuru, and U. Madhow, "Multi-gigabit communication: The ADC bottleneck[1]," in *Proc. IEEE Int. Conf. Ultra-Wideband*, Vancouver, BC, Canada, Sep. 2009, pp. 22–27.

[8] O. Dabeer and U. Madhow, "Channel estimation with low-precision analog-to-digital conversion," in *Proc. IEEE Int. Conf. Commun.*, Cape Town, South Africa, May 2010, pp. 1–6.

[9] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, "Channel estimation in millimeter wave MIMO systems with one-bit quantization," in *Proc. 48th Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2014, pp. 957–961.

[10] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.

[11] *Common Public Radio Interface (CPRI); Interface Specification, CPRI*. Accessed: Jul. 9, 2019. [Online]. Available: http://www.cpri.info/spec.html

[12] S. Park, O. Simeone, O. Sahin, and S. S. Shitz, "Fronthaul compression for cloud radio access networks: Signal processing advances inspired by network information theory," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 69–79, Nov. 2014.

[13] *Mobile IoT in the 5G Future*. Accessed: Nov. 6, 2018. [Online]. Available: https://www.ericsson.com/assets/local/networks/documents/gsma-5g-mobile-iot.pdf

[14] J. Choi, D. J. Love, D. R. Brown, and M. Boutin, "Quantized distributed reception for MIMO wireless systems using spatial multiplexing," *IEEE Trans. Signal Process.*, vol. 63, no. 13, pp. 3537–3548, Jul. 2015.

[15] J. Choi, J. Mo, and R. W. Heath, "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2005–2018, May 2016.

[16] J. Choi, D. J. Love, and D. R. Brown, "Channel estimation techniques for quantized distributed reception in MIMO systems," in *Proc. 48th Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2014, pp. 1066–1070.

[17] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[18] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "On one-bit quantized ZF precoding for the multiuser massive MIMO downlink," in *Proc. IEEE Sens. Array Multichannel Signal Process. Workshop (SAM)*, Rio de Janeiro, Brazil, Jul. 2016, pp. 1–5.

[19] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4624–4634, Sep. 2017.

[20] A. Swindlehurst, A. Saxena, A. Mezghani, and I. Fijalkow, "Minimum probability-of-error perturbation precoding for the one-bit massive MIMO downlink," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, New Orleans, LA, USA, Mar. 2017, pp. 6483–6487.

[21] J. Mo and R. W. Heath, "High SNR capacity of millimeter wave MIMO systems with one-bit quantization," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2014, pp. 1–5.

[22] J. Mo and R. W. Heath, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, Oct. 2015.

[23] A. Mezghani and J. A. Nossek, "Analysis of 1-bit output noncoherent fading channels in the low SNR regime," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, Jun. 2009, pp. 1080–1084.

[24] A. Mezghani and J. A. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1286–1289.

[25] Y. Li, C. Tao, A. L. Swindlehurst, A. Mezghani, and L. Liu, "Downlink achievable rate analysis in massive MIMO systems with one-bit DACs," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1669–1672, Jul. 2017.

[26] S. Krone and G. Fettweis, "Capacity of communications channels with 1-bit quantization and oversampling at the receiver," in *Proc. 35th IEEE Sarnoff Symp.*, Newark, NJ, USA, 2012, pp. 1–7.

[27] J. Singh and U. Madhow, "Phase-quantized block noncoherent communication," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2828–2839, Jul. 2013.

[28] N. Liang and W. Zhang, "Mixed-ADC massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 983–997, Apr. 2016.

[29] A. Mezghani, M. S. Khoufi, and J. A. Nossek, "A modified MMSE receiver for quantized MIMO systems," in *Proc. Int. ITG/IEEE Workshop Smart Antennas (WSA)*, Vienna, Austria, 2007, pp. 1–5.

[30] A. Mezghani, M. Rouatbi, and J. A. Nossek, "An iterative receiver for quantized MIMO systems," in *Proc. 16th IEEE Mediterr. Electrotechn. Conf.*, Yasmine Hammamet, Tunisia, Mar. 2012, pp. 1049–1052.

[31] Y. Jeon, N. Lee, S. Hong, and R. Heath, "One-bit sphere decoding for uplink massive MIMO systems with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4509–4521, Jul. 2018.

[32] Y. Jeon, H. Do, S. Hong, and N. Lee, "Soft-output detection methods for sparse millimeter-wave MIMO systems with low-precision ADCs," *IEEE Trans. Commun.*, vol. 67, no. 4, pp. 2822–2836, Apr. 2019.

[33] Y. Jeon, N. Lee, and V. Poor, "Robust data detection for MIMO systems with one-bit ADCs: A reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1663–1676, Mar. 2020.

[34] Y. Jeon, H. Lee, and N. Lee, "Robust MLSD for wideband SIMO systems with one-bit ADCs: Reinforcement-learning approach," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Kansas City, MO, USA, May 2018, pp. 1–6.

[35] S. K. J. Chae and S. Hong, "Machine learning detectors for MU-MIMO systems with one-bit ADCs," *IEEE Access*, vol. 8, pp. 86608–86616, 2020.

[36] S. Krone and G. Fettweis, "Fading channels with 1-bit output quantization: Optimal modulation, ergodic capacity and outage probability," in *Proc. IEEE Inf. Theory Workshop*, Dublin, Ireland, Aug. 2010, pp. 1–5.

[37] S. Gayan, R. Senanayake, and J. Evans, "On the symbol error probability for QPSK with quantized observations," in *Proc. 27th Int. Telecommun. Netw. Appl. Conf. (ITNAC)*, Melbourne, VIC, Australia, Nov. 2017, pp. 1–6.

[38] S. Gayan, H. Inaltekin, R. Senanayake, and J. Evans, "Phase modulated communication with low-resolution ADCs," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, May 2019, pp. 1–6. [Online]. Available: http://arxiv.org/abs/1907.00789

[39] R. Remmert, *Theory of Complex Functions*. New York, NY, USA: Springer, 1991.

[40] B. Picinbono, "On circularity," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[41] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5597–5625, Nov. 2012.

[42] R. G. Gallager, *Principles of Digital Communication*. New York, NY, USA: Cambridge Univ. Press, 2008.

[43] A. Mezghani, F. Antreich, and J. A. Nossek, "Multiple parameter estimation with quantized channel output," in *Proc. Int. ITG Workshop Smart Antennas (WSA)*, Bremen, Germany, Feb. 2010, pp. 143–150.

[44] J. Mo, P. Schniter, and R. W. Heath, "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141–1154, Mar. 2018.

[45] C. Wen, C. Wang, S. Jin, K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2541–2556, May 2016.

[46] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[47] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[48] M. Nakagami, "The *m*-distribution—A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*. London, U.K.: Pergamon Press, 1960, pp. 7–36.

[49] G. L. Stüber, *Principles of Mobile Communication (2nd Ed.)*, 2nd ed. Norwell, MA, USA: Kluwer Acad. Publ., 2001.

[50] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. Amsterdam, The Netherlands: Academic, 2007.

[51] N. C. Beaulieu and C. Cheng, "Efficient Nakagami-m fading channel simulation," *IEEE Trans. Veh. Technol.*, vol. 54, no. 2, pp. 413–424, Mar. 2005.

[52] W. Rudin, *Real and Complex Analysis*, 3rd ed. New York, NY, USA: McGraw-Hill, 1987.

[53] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[54] R. W. Lucky and J. C. Hancock, "On the optimum performance of *n*-ary systems having two-degrees of freedom," *IRE Trans. Commun. Syst.*, vol. 10, no. 2, pp. 185–192, Jun. 1962.

[55] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*, 2nd ed. Nashua, NH, USA: Athena Sci., 2008.

[56] K. T. Fang, S. Kotz, and K. W. Ng, *Symmetric Multivariate and Related Distributions*, 1st ed. New York, NY, USA: CRC Press, 1990.

SAMIRU GAYAN (Graduate Student Member, IEEE) received the Bachelor of Science degree in engineering and the Master of Philosophy degree in telecommunication from the University of Moratuwa, Sri Lanka, in 2011 and 2015, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Electronic Engineering, University of Melbourne, where he received an international research scholarship and focusing on wireless communications with low-resolution quantization. His research interests are in communications theory and signal processing for wireless communications.

RAJITHA SENANAYAKE (Member, IEEE) received the B.E. degree in electrical and electronics engineering from the University of Peradeniya, Sri Lanka, in 2009, the B.I.T. degree in information technology from the University of Colombo, Sri Lanka, in 2010, and the Ph.D. degree in electrical and electronics engineering from the University of Melbourne, Australia, in 2015. From 2009 to 2011, she was with the research and development team with Excel Technology, Sri Lanka. From 2015 to 2016, she was with the Department of Electrical and Computer Systems Engineering, Monash University, Australia. She is currently a Research Fellow with the Department of Electrical and Electronics Engineering, University of Melbourne. Her research interests are in cooperative communications, distributed antenna systems and vehicular communications. She is a recipient of the Australian Research Council Discovery Early Career Researcher Award.

HAZER INALTEKIN (Member, IEEE) received the B.S. degree (High Hons.) in electrical and electronics engineering from Bogazici University, Istanbul, Turkey, in 2001, and the M.S./Ph.D. degree in electrical and computer engineering from Cornell University, Ithaca, NY, USA, in 2006. He is a Senior Lecturer with Macquarie University. He held various researcher and faculty positions in Australia, Europe, and USA. His research interests include fog/edge computing, satellite and airborne networks, IoT technology, wireless communications, and information theory.

JAMIE EVANS (Senior Member, IEEE) was born in Newcastle, Australia, in 1970. He received the B.S. degree in physics and the B.E. degree in computer engineering from the University of Newcastle, in 1992 and 1993, respectively, where he received the University Medal upon graduation. He received the M.S. and the Ph.D. degrees in electrical engineering from the University of Melbourne, Australia, in 1996 and 1998, respectively, and was awarded the Chancellor's Prize for excellence for his Ph.D. thesis. From March 1998 to June 1999, he was a Visiting Researcher with the Department of Electrical Engineering and Computer Science, University of California, Berkeley. Since returning to Australia in July 1999, he has held academic positions with the University of Sydney, the University of Melbourne, and Monash University. He is currently a Professor and the Deputy Dean with the Melbourne School of Engineering, University of Melbourne. His research interests are in communications theory, information theory, and statistical signal processing with a focus on wireless communications networks.