Phase-Quantized Block Noncoherent Communication

Jaspreet Singh and Upamanyu Madhow, *Fellow, IEEE*

***Abstract*—Analog-to-digital conversion (ADC) is a key bottleneck in scaling DSP-centric receiver architectures to multiGigabit/s speeds. Recent information-theoretic results, obtained under ideal channel conditions (perfect synchronization, no dispersion), indicate that low-precision ADC (1-4 bits) could be a suitable choice for designing such high speed systems. In this work, we study the impact of employing low-precision ADC in a *carrier asynchronous* system. Specifically, we consider transmission over the block noncoherent additive white Gaussian noise channel, and investigate the achievable performance under low-precision output quantization. We focus attention on an architecture in which the receiver quantizes *only the phase* of the received signal: this has the advantage of being implementable without automatic gain control, using multiple 1-bit ADCs preceded by analog multipliers. For standard uniform Phase Shift Keying (PSK) modulation, we study the structure of the transition density of the phase-quantized block noncoherent channel. Several results, based on the symmetry inherent in the channel model, are provided to characterize this transition density. Low-complexity procedures for computing the channel information rate, and for block demodulation, are obtained using these results. Numerical computations are performed to compare the performance of quantized and unquantized systems, for different quantization precisions, and different block lengths. With QPSK modulation, it is observed, for example, that for SNR larger than 2-3 dB, 8-bin phase quantization of the received signal recovers about 80-85% of the mutual information attained with unquantized observations, while 12-bin phase quantization recovers more than 90% of the unquantized mutual information. Dithering the constellation is shown to improve the performance in the face of drastic quantization.**

***Index Terms*—Quantization, analog-to-digital conversion, noncoherent communication, channel capacity, phase noise.**

I. INTRODUCTION

THE economies of scale provided by integrated circuit implementation of sophisticated digital signal processing (DSP) algorithms have propelled mass market deployment of cellular and wireless local area network systems over the last two decades. As we now look to scale up the speeds of such DSP-centric architectures by orders of magnitude (e.g., to build multigigabit/sec systems by exploiting the wide swath of spectrum in the 60 GHz band [1]), the analog-to-digital converter (ADC), which converts the received analog waveform into the digital domain, becomes a bottleneck: high-speed high-precision ADC is costly and power-hungry [2], [3]. It is of interest, therefore, to explore the feasibility of system design with low-precision ADC. Recent information-theoretic results [4] show that for an ideal channel model (perfect synchronization, no dispersion), the loss in the Shannon capacity incurred by using 1-3 bits ADC is acceptable, even at moderately high signal-to-noise ratio (SNR). In this paper, we consider a system without *a priori* carrier synchronization, i.e., there is a small frequency offset between the receiver's local oscillator and the incoming carrier wave, and investigate the impact of low-precision ADC on the achievable performance.

In order to model the effect of carrier asynchronism, we consider a discrete-time complex baseband *block noncoherent* additive white Gaussian noise (AWGN) channel. If the receiver's local oscillator is not synchronized with that of the transmitter, the phase after downconversion is *a priori* unknown. However, for practical values of carrier offset, it is well approximated as constant over a block of symbols. For example, consider a frequency uncertainty of 100 parts per million (ppm) and a bandwidth equal to 10% of the carrier frequency f_c. Then $\Delta f \in [-10^{-4} f_c, 10^{-4} f_c]$ and $\frac{1}{Ts} \approx 0.1 f_c$, so that the phase change ϕ per symbol lies in [-0.0062, 0.0062] radians, or [-0.36, 0.36] degrees, which is small enough to assume that the phase remains constant over a block of, say, 4-6 consecutive symbols. The classical approach to noncoherent communication is to approximate the phase as constant over two symbols, and to apply differential modulation and demodulation. Divsalar and Simon [5] were the first to point out the gains that may be achieved by performing multiple symbol differential demodulation over a block of $L > 2$ symbols, under the assumption that the phase remains constant over the entire block. The block length L is termed the channel coherence length. More recent work [6], [7], [8] has shown that block demodulation, even for large L, can be implemented efficiently, and exhibits excellent performance for both coded and uncoded systems.[1]

We study the effect of low-precision receiver quantization

Manuscript received November 8, 2011; revised July 25 and December 31, 2012. The editor coordinating the review of this paper and approving it for publication was L. Yang.

J. Singh is with Samsung Research America, 1301, E Lookout Dr, Richardson, TX 75082, USA (e-mail: jsingh@sta.samsung.com).

U. Madhow is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA (e-mail: madhow@ece.ucsb.edu).

This research was supported in part by the National Science Foundation under grants CCF-0729222 and CNS-0832154, and in part by the Systems on Nanoscale Information fabriCs (SONIC), one of six centers supported by the STARnet phase of the Focus Center Research Program (FCRP), a Semiconductor Research Corporation program sponsored by MARCO and DARPA.

This work was presented in part at the IEEE International Symposium on Information Theory (ISIT 2009, Seoul, Korea) and at the IEEE International Conference on Ultra-WideBand (ICUWB 2009, Vancouver, Canada).

Digital Object Identifier 10.1109/TCOMM.2013.052013.110755

[1]Note that the block length L here refers to the size of the block over which noncoherent modulation/demodulation would be performed, and is distinct from the notion of *packet size* used in any modern packet-switched communication system.

for the block noncoherent channel under M-ary Phase Shift Keying (MPSK) modulation. Since PSK encodes information in the phase of the transmitted symbols, and since the channel impairment is also being modeled as a phase rotation, we investigate an architecture in which the receiver quantizes *only the phase* of the received samples, disregarding the amplitude information. Such phase-only quantization is attractive because of its ease of implementation, since it eliminates the need for automatic gain control at baseband. Further, it can be performed using 1-bit ADCs preceded by analog pre-multipliers that provide linear combinations of the I and Q channel samples (obtained after the low noise amplification and downconversion operations). Specifically, every linear combination, $\alpha I + \beta Q$ (for analog multiplicative constants α and β), followed by a 1-bit ADC, partitions the complex plane into two halves (separated by a straight line with slope $-\frac{\alpha}{\beta}$). Thus, any desired phase sectorization of the plane may be attained using appropriately chosen linear combinations. For example, 1-bit ADC on both I and Q channels implements uniform 4-sector phase quantization, and uniform 8-sector quantization is achieved by adding two new linear combinations, I+Q and I-Q (using trivial analog pre-multipliers $+1$ and -1), as depicted in Fig. 1.

Our focus here is on computing information-theoretic benchmarks (applicable for heavily coded systems) as well as uncoded symbol error rates for communication over the phase-quantized block noncoherent channel [2]. While brute force computation of these quantities has complexity that scales exponentially in the channel coherence length, we find that significant complexity reduction can be attained by understanding the nature of the channel block transition probability. A summary of our analytical and numerical results for M-PSK modulation and uniform K-sector phase quantization, is as follows:

1) We begin by studying the structure of the input-output relationship of the phase-quantized block noncoherent AWGN channel. For the special case when M divides K, we exploit the symmetry inherent in the channel model to derive several results characterizing the output probability distribution over a block of symbols, with and without conditioning on the input. These results are used to obtain a low-complexity procedure for computing the mutual information between the channel input and output (brute force computation has complexity exponential in the block length L).
2) We obtain low-complexity optimal block noncoherent demodulation rules. These rules are obtained by specializing existing low-complexity procedures for block demodulation with unquantized observations, to our setting with quantized observations.
3) A close analysis of the block demodulator reveals that, depending on the number of quantization sectors, the symmetries inherent in the channel model, while helping us reduce the computational complexities, can also have dire consequences on performance: they can make it impossible to distinguish between the effect of the unknown phase offset and the phase modulation. As a result, we may have two equally likely inputs for certain outputs, irrespective of the block length and the SNR, leading to severe performance degradation. In order to break such undesirable symmetries, we investigate the performance with a *dithered*-PSK input scheme, in which we rotate the PSK constellation across the different symbols in a block.
4) Numerical results are obtained for QPSK input with 8 sector and 12 sector phase quantization, for different choices of the block length L, and compared with the unquantized performance (studied earlier in [10]). We find that 8-sector quantization with dithered-QPSK achieves more than 80-85 % of the mutual information achieved with unquantized observations (with identical block length, and for SNR larger than 2-3 dB), while with 12-sector quantization we can get as much as 90-95 % of the unquantized mutual information without requiring dithering. The corresponding loss in terms of SNR, for fixed mutual information, varies between 2-4 dB for 8-sector quantization, and between 0.5-2 dB with 12 sectors. In terms of uncoded symbol error rates (SER), the performance degradation is of the same order. For instance, at $\text{SER} = 10^{-3}$, the loss for 8 and 12 sector quantization, compared to unquantized observations, is about 4 dB and 2 dB respectively.

Related work: Our preliminary results on phase-quantized noncoherent communication appeared in [11], [12]. Two papers that address problems similar to our work are [13], [14]. Both consider a phase offset between the transmitter and the receiver, and investigate the impact of output quantization on the achievable performance. In [13], the authors study a mixed-signal receiver architecture, wherein the unknown phase offset is first estimated in post-ADC DSP, and then compensated for in the analog domain prior to the ADC. In [14], the authors do not resort to analog domain compensation (e.g., due to hardware complexity constraints), but use knowledge of the phase offset (assumed to be somehow available) in post-ADC processing. They observe that, in such a system, output (amplitude) quantization can cause significant degradation in the achievable transmission rate. Note that while the methods in [13], [14] focus on *explicit* estimation/knowledge and use of the phase offset, the block noncoherent receiver studied here amounts to *implicit* joint estimation of the unknown phase offset and the unknown data symbols.

In this work, we consider only a phase offset (induced by the asynchronous local oscillators) between the transmitter and the receiver. However, the block noncoherent model extends to the setting of communication over narrowband slow *fading* channels as well, where the channel state, while unknown, can be assumed to be constant over a block of symbols. Indeed, this block fading model has been investigated extensively in the recent literature, ranging from capacity analysis [15], to efficient architectures for block demodulation and decoding ([6], [7], [8], [16], and references therein). While our work draws upon this extensive literature, it addresses, for the first time, the unique channel characteristics that arise due to (drastic) output quantization of the block noncoherent channel.

[2]In addition to being of fundamental interest for the communication theorist, uncoded error rates provide an important benchmark for scenarios where the receiver power consumption due to decoding of capacity-achieving codes may be excessive (e.g., for short-range communication) [9].

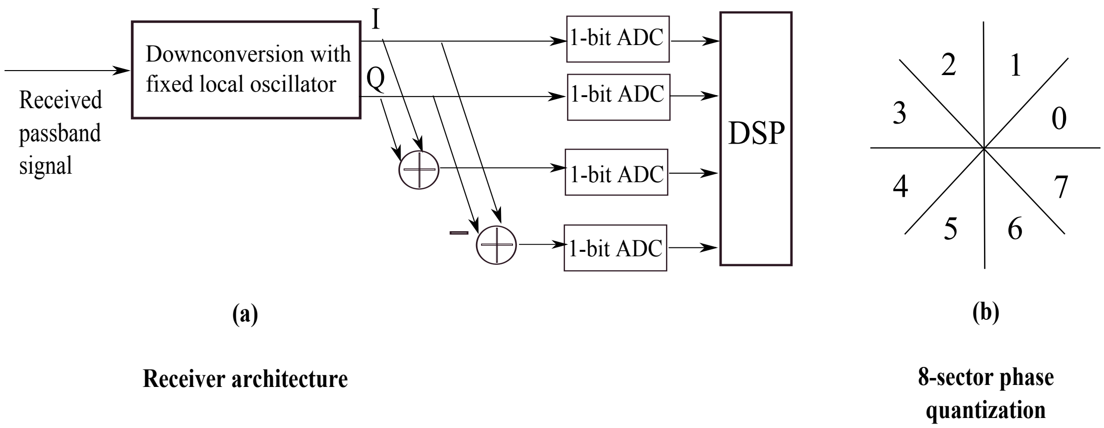

Fig. 1. Receiver architecture for 8-sector quantization.

Prior information-theoretic studies to understand the impact of output quantization on the capacity of the ideal AWGN channel include [17], [18], [19], [4]. More recent work that explores the impact of quantization on other aspects of communication system design includes, amongst others, the study of oversampling [20], [21], asymmetric quantization [22], automatic gain control [23], channel estimation performance [24], performance over fading channels [25], [26], broadcast and MAC channels [27], and channels with memory [28], constellation shaping methods [29], and study of system performance under the generalized mutual information criterion [30].

Organization of the Paper: The rest of the paper is organized as follows. In Section II, we describe the channel model and the receiver architecture for phase quantization. In Section III, we study the properties of the channel's transition probability distribution. Efficient mechanisms for computing the channel information rate and for block noncoherent demodulation are described in Sections IV and V, respectively. Numerical results are presented in Section VI, followed by the conclusions and a list of open issues in Section VII.

Notation for random variables/vectors: Throughout the paper, a random variable is denoted by a capital letter, while the specific value it takes is denoted by a small letter. Bold face notation is used for vectors of random variables. For example, X is a random variable and x is the value it takes, while $\mathbf{X}$ is a random vector and $\mathbf{x}$ is the value it takes. We use the notation $\mathsf{P}(X)$ to denote the probability distribution function of X, while $\mathsf{P}(x)$ is used to denote the value of the PDF at $X = x$. $\mathbb{E}$ is the expectation operator.

II. Channel Model and Receiver Architecture

The received signal over a block of length L, after quantization, is represented as

$$Z_l = \mathsf{Q}(S_l e^{j\Phi} + N_l)\ ,\ l = 0, 1, \cdots, L-1, \tag{1}$$

where,

- $\mathbf{S} := [S_0\ S_1\ \cdots\ S_{L-1}]$ is the vector of transmitted symbols,
- Φ is an unknown phase offset with uniform distribution on $[0, 2\pi)$,
- $\mathbf{N} := [N_0\ \cdots\ N_{L-1}]$ is a vector of i.i.d. complex Gaussian noise with variance $\sigma^2 = N_0/2$ in each dimension,
- $\mathsf{Q} : \mathbb{C} \to \mathcal{K} = \{0, 1, \cdots, K-1\}$ denotes a quantization function that maps each point in the complex plane to one of the K quantization indices, and
- $\mathbf{Z} := [Z_0\ Z_1\ \cdots\ Z_{L-1}]$ is the vector of quantized received symbols, so that each $Z_l \in \mathcal{K}$.

Each S_l is picked in an i.i.d. manner from a uniform M-PSK constellation denoted by the set of points $\mathcal{A} = \{e^{j\theta_0}, e^{j\theta_1}, \cdots, e^{j\theta_{M-1}}\}$, where $\theta_m = (\theta_{m-1} + \frac{2\pi}{M})$ [3], for $m = 1, 2, \cdots, M-1$.

In order to represent the channel input as a vector of indices (rather than as a vector of complex phasors), we introduce the random vector $\mathbf{X} = [X_0\quad X_1\quad \cdots\quad X_{L-1}]$, with each X_i picked in an i.i.d. manner from a uniform distribution on the set $\{0, 1, \cdots, M-1\}$. Our channel model (1) can now equivalently be written as

$$Z_l = \mathsf{Q}(e^{j\theta_{X_l}} e^{j\Phi} + N_l)\ ,\ l = 0, 1, \cdots, L-1\ , \tag{2}$$

with every output symbol $Z_l \in \{0, 1, \cdots, K-1\}$ as before, and every input symbol $X_l \in \{0, 1, \cdots, M-1\}$. The set of all possible input vectors is denoted by $\mathcal{X}$, while $\mathcal{Z}$ denotes the set of all possible output vectors.

We consider K-bin (or K-sector) phase quantization. The quantizer divides the interval $[0, 2\pi)$ into K equal parts (or sectors), and the quantization indices go from 0 to $K-1$ in the counter-clockwise direction. Fig. 1(b) depicts the scenario for K=8. Thus, our quantization function is $\mathsf{Q}(c) = \lfloor \frac{\arg(c)}{(\frac{2\pi}{K})} \rfloor$, where $c \in \mathbb{C}$ (the set of all complex numbers), and $\lfloor p \rfloor$ denotes the greatest integer less than or equal to p. As discussed in Section I, such phase quantization can be implemented using 1-bit ADCs preceded by analog multipliers which provide linear combinations of the I and Q channel samples. For instance, employing 1-bit ADC on I and Q channels results in uniform 4-sector phase quantization, while uniform 8-sector quantization can be achieved simply by adding two new linear combinations, I+Q and I-Q (requiring trivial

[3]Unless stated otherwise, arithmetic operations for phase angles are assumed to be performed modulo 2π. For the output symbols Z_l, the arithmetic is modulo K, while for the input symbols X_l (introduced immediately after in the text), it is modulo M.

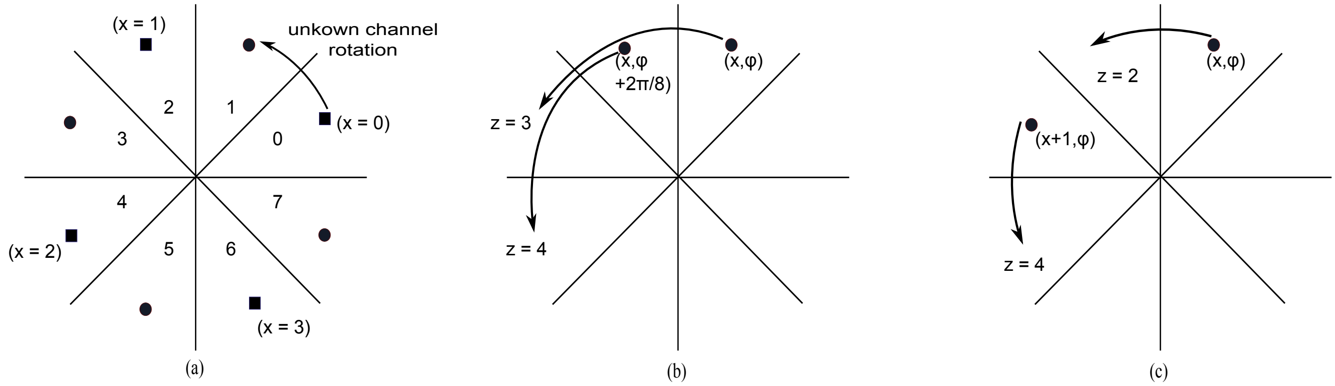

Fig. 2. QPSK with 8-sector quantization (i.e., M=4, K=8). (a) depicts how the unknown channel phase ϕ results in a rotation of the transmitted symbol (square : original constellation , circle : rotated constellation). (b) and (c) depict the circular symmetry induced in the conditional probability $\mathsf{P}(z|x,\phi)$ due to the circular symmetry of the complex Gaussian noise. (b) shows that increasing ϕ by $2\pi/K = (\pi/4)$ and z by 1 will keep the conditional probability unchanged, i.e., $\mathsf{P}(z=3|x,\phi) = \mathsf{P}(z=4|x,\phi+2\pi/K)$. (c) shows that increasing x by 1 and z by $K/M = 2$ will keep the conditional probability unchanged, i.e., $\mathsf{P}(z=2|x,\phi) = \mathsf{P}(z=4|x+1,\phi)$.

analog multipliers $+1$ and -1), corresponding to a $\pi/4$ rotation of I/Q axes, as shown in Fig. 1(a).

We begin our investigation by studying the inherent symmetry in the relationship between the channel input and output. This yields several results that govern the structure of the output probability distribution, both conditioned on the input (i.e., $\mathsf{P}(\mathbf{Z}|\mathbf{X})$), and without conditioning (i.e., $\mathsf{P}(\mathbf{Z})$). These distributions are integral to our later computations of the input-output mutual information, as well as for soft decision decoding (not considered here). While brute force computation (computing $\mathsf{P}(\mathbf{z}|\mathbf{x})$ for every $\mathbf{z} \in \mathcal{Z}$ and every $\mathbf{x} \in \mathcal{X}$) of these distributions has exponential complexity in block length, our results allow their computation with significant reduction in complexity.

Note: Throughout the paper, we assume that the PSK constellation size M, and the number of quantization bins K, are such that $K = aM$ for some positive integer a. We illustrate our results with the running example of QPSK with 8-sector quantization (so that $a = 2$), depicted in Fig. 2(a).

III. Input-Output Relationship

Conditioned on the channel phase Φ, $\mathsf{P}(\mathbf{Z}|\mathbf{X},\Phi)$ is a product of individual symbol probabilities $\mathsf{P}(Z_l|X_l,\Phi)$. We therefore begin by analyzing the symmetries in the latter.

A. *Properties of* $\mathsf{P}(Z_l|X_l,\Phi)$

We have that $\mathsf{P}(z_l|x_l,\phi)$ is the probability that $\arg(e^{j(\theta_{x_l}+\phi)}+N_l)$ belongs to the interval $[\frac{2\pi}{K}z_l,\ \frac{2\pi}{K}(z_l+1))$. In other words, it is the probability that, on adding complex Gaussian noise N_l to the point $e^{j(\theta_{x_l}+\phi)}$, we get a point whose phase belongs to $[\frac{2\pi}{K}z_l,\ \frac{2\pi}{K}(z_l+1))$. Due to the circular symmetry of the complex Gaussian noise, this is the same as the probability that, on adding N_l to the point $e^{j(\theta_{x_l}+\phi+\frac{2\pi}{K}i)}$, we get a point whose phase belongs to $[\frac{2\pi}{K}(z_l+i),\ \frac{2\pi}{K}(z_l+1+i))$, where i is an integer. We thus get our first two results.

Property A-1: $\mathsf{P}(z_l|x_l,\phi) = \mathsf{P}(z_l+i|x_l,\phi+i\frac{2\pi}{K})$.

Property A-2: $\mathsf{P}(z_l|x_l,\phi) = \mathsf{P}(z_l+ia|x_l+i,\phi)$.

Note that $\theta_{x_l+i} = \theta_{x_l} + \frac{2\pi}{M}i = \theta_{x_l} + \frac{2\pi}{K}(ia)$, which gives Property A-2.

Property A-2 simply states that if we jump from one point in the M-PSK constellation to the next, then we must jump $a = \frac{K}{M}$ quantization sectors in order to keep the conditional probability invariant. This is because the separation between consecutive points in the input constellation is $2\pi/M$, while each quantization sector covers an angle of $2\pi/K$. For QPSK with $K = 8$, Fig. 2(b) and 2(c) depict example scenarios for the two properties.

If we put $i = -x_l$ in Property A-2, we get the following special case, which relates the conditioning on a general x_l to the conditioning on 0.

Property A-3: $\mathsf{P}(z_l|x_l,\phi) = \mathsf{P}(z_l - ax_l|0,\phi)$.

To motivate the final property, we consider our example of QPSK with $K = 8$. While we have 8 distinct quantization sectors, we see from Fig. 2(a) that the orientation of these 8 sectors relative to the 4 constellation points (shown as squares) can be described by dividing the sectors into 2 groups : $\{0,2,4,6\}$, and $\{1,3,5,7\}$. For instance, the positioning of the first sector ($z = 0$) w.r.t. $x = 0$ is identical to the positioning of the third sector ($z = 2$) w.r.t. $x = 1$ (and similarly $z = 4$ w.r.t $x = 2$, and $z = 6$ w.r.t $x = 3$). On the other hand, the positioning of the second sector ($z = 1$) w.r.t. $x = 0$ is identical to the positioning of the fourth sector ($z = 3$) w.r.t. $x = 1$ (and similarly $z = 5$ w.r.t $x = 2$, and $z = 7$ w.r.t $x = 3$). In terms of the conditional probabilities, this implies, for example, that $\mathsf{P}(z_l = 7|x_l = 3,\phi) = \mathsf{P}(z_l = 1|x_l = 0,\phi)$, and similarly, $\mathsf{P}(z_l = 6|x_l = 3,\phi) = \mathsf{P}(z_l = 0|x_l = 0,\phi)$. In general, we can relate the conditional probability of every odd z_l with that of $z_l = 1$, and similarly of every even z_l with that of $z_l = 0$, with corresponding rotations of the symbol x_l. For general values of K and M, the number of groups equals $a = \frac{K}{M}$, and we can relate the probability of any z_l with that of $z_l \bmod a$.

Property A-4: Let $z_l = q_l a + r_l$, where q_l is the quotient on dividing z_l by a, and r_l is the remainder, i.e, $r_l = z_l \bmod a$. Then, $\mathsf{P}(z_l|x_l,\phi) = \mathsf{P}(z_l \bmod a|x_l - q_l,\phi)$.

While this result follows directly from Property A-2 by setting $i = -q_l$, it is an important special case, enabling us to restrict attention to only the first a sectors ($Z_l \in \{0,1,\cdots,a-1\}$) rather than having to work with all K sectors. As detailed later, this leads to significant complexity reduction in computation of the mutual information.

We now use these properties to present results for $\mathsf{P}(\mathbf{Z}|\mathbf{X})$.

B. Properties of $\mathsf{P}(\mathbf{Z}|\mathbf{X})$

Property B-1: Let $\mathbf{1}$ denote the row vector with all entries as 1. Then $\mathsf{P}(\mathbf{z}|\mathbf{x}) = \mathsf{P}(\mathbf{z} + i\mathbf{1}|\mathbf{x})$.

Proof: For a fixed $\mathbf{x}$, increasing each z_l by the same number i leaves the conditional probability unchanged, because the phase Φ in the channel model (1) is uniformly distributed in $[0, 2\pi)$. A detailed proof follows. We have

$$\begin{aligned}\mathsf{P}(\mathbf{z}|\mathbf{x}) &= \mathbb{E}_\Phi\left(\mathsf{P}(\mathbf{z}|\mathbf{x},\Phi)\right) = \mathbb{E}_\Phi\left(\prod_{l=0}^{L-1}\mathsf{P}(z_l|x_l,\Phi)\right)\\ &= \mathbb{E}_\Phi\left(\prod_{l=0}^{L-1}\mathsf{P}(z_l+i|x_l,\Phi+i\frac{2\pi}{K})\right)\\ &= \mathbb{E}_{\hat\Phi}\left(\prod_{l=0}^{L-1}\mathsf{P}(z_l+i|x_l,\hat\Phi)\right)\\ &= \mathbb{E}_{\hat\Phi}\left(\mathsf{P}(\mathbf{z}+i\mathbf{1}|\mathbf{x},\hat\Phi))\right) = \mathsf{P}(\mathbf{z}+i\mathbf{1}|\mathbf{x}).\end{aligned}$$

The second equality follows because the components of $\mathbf{Z}$ are conditionally independent, conditioned on $\mathbf{X}$ and Φ. Property A-1 gives the third equality. A change of variables, $\hat\Phi = \Phi + i\frac{2\pi}{K}$ gives the fourth equality (since Φ is uniformly distributed on $[0, 2\pi)$, so is $\hat\Phi$), thereby completing the proof. ∎

Remark 1: For the rest of the paper, we refer to the operation $\mathbf{z} \to \mathbf{z} + i\mathbf{1}$ as *constant addition*.

Next, we observe that the conditional probability remains invariant under an *identical* permutation of the components of the vectors $\mathbf{z}$ and $\mathbf{x}$.

Property B-2: Let Π denote a permutation operation, and $\Pi\mathbf{x}$ ($\Pi\mathbf{z}$) the vector obtained on permuting $\mathbf{x}$ ($\mathbf{z}$) under this operation. Then, $\mathsf{P}(\mathbf{z}|\mathbf{x}) = \mathsf{P}(\Pi\mathbf{z}|\Pi\mathbf{x})$.

Proof: As in the proof of Property 1, the idea is to condition on Φ and work with the symbol probabilities $\mathsf{P}(z_l|x_l,\Phi)$. Consider $\mathsf{P}(\mathbf{z}|\mathbf{x},\Phi) = \prod_{l=0}^{L-1}\mathsf{P}(z_l|x_l,\Phi)$, and $\mathsf{P}(\Pi\mathbf{z}|\Pi\mathbf{x},\Phi) = \prod_{l=0}^{L-1}\mathsf{P}((\Pi\mathbf{z})_l|(\Pi\mathbf{x})_l,\Phi)$. Since multiplication is a commutative operation, we have $\mathsf{P}(\mathbf{z}|\mathbf{x},\Phi) = \mathsf{P}(\Pi\mathbf{z}|\Pi\mathbf{x},\Phi)$. Taking expectation w.r.t. Φ completes the proof. ∎

The next two results extend properties A-3 and A-4.

Property B-3: Define the input vector $\mathbf{x}_0 = [0\cdots 0]$. Then, $\mathsf{P}(\mathbf{z}|\mathbf{x}) = \mathsf{P}(\mathbf{z} - a\mathbf{x}|\mathbf{x}_0)$, where $a = \frac{K}{M}$, and the subtraction is performed modulo K.

Property B-4: Let $z_l = q_l a + r_l$, where q_l is the quotient on dividing z_l by a, and r_l is the remainder, i.e, $r_l = z_l \bmod a$. Define $\mathbf{q} = [q_0, \cdots, q_{L-1}]$, and, $\mathbf{z} \bmod a = [z_0 \bmod a \ \cdots \ z_{L-1} \bmod a]$. Then $\mathsf{P}(\mathbf{z}|\mathbf{x}) = \mathsf{P}(\mathbf{z} \bmod a \mid \mathbf{x} - \mathbf{q})$.

Proofs: The properties follow from A-3 and A-4 respectively, by first noting that the vector probability $\mathsf{P}(\mathbf{z}|\mathbf{x},\Phi)$ is the product of the scalar ones, and then integrating over Φ . ∎

C. Properties of $\mathsf{P}(\mathbf{Z})$

We now consider the unconditional distribution $\mathsf{P}(\mathbf{z})$. The first result states that $\mathsf{P}(\mathbf{z})$ is invariant under constant addition.

Property C-1: $\mathsf{P}(\mathbf{z}) = \mathsf{P}(\mathbf{z} + i\mathbf{1})$.

Proof: Using Property B-1, this follows directly by taking expectation over $\mathbf{X}$ on both sides. ∎

We now extend Property B-2 along similar lines to show that $\mathsf{P}(\mathbf{z})$ is invariant under any permutation of $\mathbf{z}$.

Property C-2: $\mathsf{P}(\mathbf{z}) = \mathsf{P}(\Pi\mathbf{z})$.

Proof: We have $\mathsf{P}(\mathbf{z}) = \frac{1}{M^L}\sum_{\mathbf{x}\in\mathcal{X}}\mathsf{P}(\mathbf{z}|\mathbf{x})$. Using Property B-2, we get $\mathsf{P}(\mathbf{z}) = \frac{1}{M^L}\sum_{\mathbf{x}\in\mathcal{X}}\mathsf{P}\left(\Pi\mathbf{z}|\Pi\mathbf{x}\right)$. Since the permutation operation results in a one-to-one mapping (every unique choice of $\mathbf{x}\in\mathcal{X}$ results in a unique $\Pi\mathbf{x}\in\mathcal{X}$), we can rewrite the last equation as $\mathsf{P}(\mathbf{z}) = \frac{1}{M^L}\sum_{\mathbf{x}\in\mathcal{X}}\mathsf{P}(\Pi\mathbf{z}|\mathbf{x}) = \mathsf{P}(\Pi\mathbf{z})$. ∎

Our final result extends Property B-4.

Property C-3: Let $a = \frac{K}{M}$. Then $\mathsf{P}(\mathbf{z}) = \mathsf{P}(\mathbf{z} \bmod a)$.

Proof: Using the same notation as in Property B-4, we have $\mathsf{P}(\mathbf{z}|\mathbf{x}) = \mathsf{P}(\mathbf{z} \bmod a \mid \mathbf{x} - \mathbf{q})$. Noting that the transformation $\mathbf{x} \to \mathbf{x} - \mathbf{q}$ is a one-to-one mapping, the proof follows on the same lines as the proof of Property C-2. ∎

Example: For QPSK with $K = 8$ and $L = 4$, $\mathsf{P}(z = [5\ 7\ 2\ 4]) = \mathsf{P}(z = [1\ 1\ 0\ 0])$.

We now apply these results for low-complexity computation of the input-output mutual information.

IV. Mutual Information Computation

We wish to compute the mutual information

$$I(\mathbf{X};\mathbf{Z}) = H(\mathbf{Z}) - H(\mathbf{Z}|\mathbf{X}).$$

We first discuss computation of the conditional entropy.

A. Conditional Entropy

We have $H(\mathbf{Z}|\mathbf{X}) = \sum_{\mathcal{X}} H(\mathbf{Z}|\mathbf{x})\mathsf{P}(\mathbf{x})$, where $H(\mathbf{Z}|\mathbf{x}) = -\sum_{\mathcal{Z}}\mathsf{P}(\mathbf{z}|\mathbf{x})\log\mathsf{P}(\mathbf{z}|\mathbf{x})$ is the entropy of the output when the input vector $\mathbf{X}$ takes on the specific value $\mathbf{x}$. Our main result in this section is that $H(\mathbf{Z}|\mathbf{x})$ is constant $\forall\mathbf{x}$.

Property D-1: $H(\mathbf{Z}|\mathbf{x})$ is a constant.

Proof: We show that for any input vector $\mathbf{x}, H(\mathbf{Z}|\mathbf{x}) = H(\mathbf{Z}|\mathbf{x}_0)$, where $\mathbf{x}_0 = [0\cdots 0]$ as defined before. We have

$$\begin{aligned}H(\mathbf{Z}|\mathbf{x}) = & - \sum_{\mathcal{Z}}\mathsf{P}(\mathbf{z}|\mathbf{x})\log\mathsf{P}(\mathbf{z}|\mathbf{x})\\ = & - \sum_{\mathcal{Z}}\mathsf{P}(\mathbf{z}-a\mathbf{x}|\mathbf{x}_0)\log\mathsf{P}(\mathbf{z}-a\mathbf{x}|\mathbf{x}_0)\ ,\end{aligned}\tag{3}$$

where the second equality follows from Property B-3. Noting that the transformation $\mathbf{z} \to \mathbf{z} - a\mathbf{x}$ is a one-to-one mapping, we can rewrite (3) as

$$H(\mathbf{Z}|\mathbf{x}) = -\sum_{\mathcal{Z}}\mathsf{P}(\mathbf{z}|\mathbf{x}_0)\log\mathsf{P}(\mathbf{z}|\mathbf{x}_0) = H(\mathbf{Z}|\mathbf{x}_0)\tag{4}$$

∎

Thus, $H(\mathbf{Z}|\mathbf{X}) = H(\mathbf{Z}|\mathbf{x}_0)$, but brute force computation of $H(\mathbf{Z}|\mathbf{x}_0)$ still has exponential complexity: $\mathsf{P}(\mathbf{Z}|\mathbf{x}_0)$ must be computed for each of the K^L possible output vectors $\mathbf{Z}$. However, we find that it suffices to compute $\mathsf{P}(\mathbf{Z}|\mathbf{x}_0)$ for a much smaller set of $\mathbf{Z}$ vectors.

Using Property B-2, we have $\mathsf{P}(\mathbf{z}|\mathbf{x}_0) = \mathsf{P}(\Pi\mathbf{z}|\Pi\mathbf{x}_0)$. Since $\mathbf{x}_0 = [0..0]$, any permutation of $\mathbf{x}_0$ gives back $\mathbf{x}_0$. Hence, $\mathsf{P}(\mathbf{z}|\mathbf{x}_0) = \mathsf{P}(\Pi\mathbf{z}|\mathbf{x}_0)$. Combined with Property B-1, we thus get that it suffices to compute $\mathsf{P}(\mathbf{z}|\mathbf{x}_0)$ for a set of vectors $S_\mathbf{Z}$ in which no vector can be obtained from another by performing the joint operations of constant addition and permutation. While we do not have a method to obtain the set $S_\mathbf{Z}$ exactly, we exploit the preceding properties to obtain, and, restrict attention to, a sub-optimal set $S_{\mathbf{Z}_2}$ (constructed by considering the operations of constant addition and permutation in a sequential manner, rather than jointly) having cardinality which is still significantly lower than the exponential figure of K^L. Specifically, the cardinality of $S_{\mathbf{Z}_2}$ is $\binom{K+L-2}{L-1}$, where $\binom{A}{B}$ denotes the number of possible combinations of A objects, taking B at a time. As an example, with $K = 8$, and $L = \{3, 4, 5, 6, 7\}$, the cardinality of $S_{\mathbf{Z}_2}$ is $\{36, 120, 330, 792, 1716\}$, whereas the exponential figure K^L is $\{512, 4096, 32768, 2.6 \times 10^5, 2.1 \times 10^6\}$. Further details regarding the construction of the set $S_{\mathbf{Z}_2}$ are provided in Appendix A.

Once we have constructed the set $S_{\mathbf{Z}_2}$, the probability $\mathsf{P}(\mathbf{z}|\mathbf{x}_0)$ can be numerically computed for every vector $\mathbf{z}$ in this set. Computation of the entropy $H(\mathbf{Z}|\mathbf{x}_0)$ is then straightforward. See Appendix A for the details.

B. Output Entropy

The output entropy is $H(\mathbf{Z}) = -\sum_{\mathcal{Z}} \mathsf{P}(\mathbf{z}) \log \mathsf{P}(\mathbf{z})$. Brute force computation requires knowledge of $\mathsf{P}(\mathbf{z})\ \forall \mathbf{z} \in \mathcal{Z}$, which has exponential complexity. However, using Properties C-1, C-2 and C-3, it suffices to compute $\mathsf{P}(\mathbf{z})$ for a set of vectors $\tilde{S}_{\mathbf{Z}}$ in which no vector can be obtained from another by performing the operations of constant addition and permutation, and also, the vector components $\in \{0, 1, \cdots, a-1\}$. This is similar to the situation encountered earlier in the last subsection, except that the vector components there were allowed to be in $\{0, 1, \cdots, K-1\}$. To exploit this for further complexity reduction, we can begin by defining the set $\tilde{\mathcal{Z}}$ to be the set of vectors in which the vector components take values in $\{0, 1, \cdots, a-1\}$ only. Since $\mathsf{P}(\mathbf{z}) = \mathsf{P}(\mathbf{z} \bmod a)$, a moment's thought reveals that each vector in $\tilde{\mathcal{Z}}$ has the same probability as a set of $(\frac{K}{a})^L = M^L$ distinct vectors in $\mathcal{Z}$, and the sets corresponding to different vectors are disjoint. Thus $H(Z) = -\sum_{\mathcal{Z}} \mathsf{P}(\mathbf{z}) \log \mathsf{P}(\mathbf{z}) = -M^L \sum_{\tilde{\mathcal{Z}}} \mathsf{P}(\mathbf{z}) \log \mathsf{P}(\mathbf{z})$. To obtain $\{\mathsf{P}(\mathbf{z})\}$ for $\mathbf{z} \in \tilde{\mathcal{Z}}$, we can follow exactly the same procedure as described in the last subsection, with K being replaced by a. In particular, we need to compute $\mathsf{P}(\mathbf{z})$ only for $\binom{a+L-2}{L-1}$ vectors.

Example: For QPSK with 8 sectors (so $a = 2$), the relevant vectors for block length 2 are $[0\ \ 0]$ and $[0\ \ 1]$.

Computation of $\mathsf{P}(\mathbf{Z})$*:* We now need to compute $\mathsf{P}(\mathbf{z}) = \frac{1}{M^L} \sum_{\mathbf{x} \in \mathcal{X}} \mathsf{P}(\mathbf{z}|\mathbf{x})$ for each of the $\binom{a+L-2}{L-1}$ vectors. A brute force approach is to compute $\mathsf{P}(\mathbf{z}|\mathbf{x})$ for each $\mathbf{x}$, but again, has exponential complexity. We exploit the structure in $\mathbf{z}$ to reduce the number of vectors $\mathbf{x}$ for which we need $\mathsf{P}(\mathbf{z}|\mathbf{x})$. Specifically, we have that each $z_i \in \{0, 1, \cdots, a-1\}$. Since there are only a different types of components in $\mathbf{z}$, for block length $L > a$, some of the components in $\mathbf{z}$ will be repeated. For any $\mathbf{x}$, we can then use Property B-2 to rearrange the components at those locations for which the components in $\mathbf{z}$ are identical, without changing the conditional probability. For instance, let $z_m = z_n$ for some m, n. Then, $\mathsf{P}(\mathbf{z}|\mathbf{x}) = \mathsf{P}(\mathbf{z}|\Pi\mathbf{x})$, where $\Pi\mathbf{x}$ is obtained from $\mathbf{x}$ by rearranging the components at locations m and n. To sum up, we can restrict attention to a set of vectors $S_{\mathbf{X}}$ in which no vector can be obtained from another one by permutations between those locations for which the elements in $\mathbf{z}$ are identical. Construction of this set $S_{\mathbf{X}}$, and the subsequent computation of $\mathsf{P}(\mathbf{Z})$, is discussed in Appendix B. For the example scenario of QPSK with 8-sector quantization (so that $a = 2$), and block length $L = 8$, the worst case cardinality of the set $S_{\mathbf{X}}$ is 1225, compared to the exponential figure of $M^L = 65536$.

In the next section, we consider efficient block noncoherent demodulation for the phase-quantized channel. This allows us to evaluate the uncoded symbol error rates. Numerical results for uncoded performance and input-output mutual information are subsequently provided in Section VI.

V. Block Noncoherent Demodulation

We consider the generalized likelihood ratio test (GLRT) for block noncoherent demodulation. This entails joint maximum likelihood estimation of the unknown block of input symbols and the unknown channel phase. Specifically, given the received vector $\mathbf{z}$, the GLRT estimate for $\mathbf{x}$ is given by

$$\hat{\mathbf{x}}(\mathbf{z}) = \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmax}} \max_{\phi \in [0, 2\pi)} \mathsf{P}(\mathbf{z}|\mathbf{x}, \phi) \ . \tag{5}$$

Brute force computation of the solution to (5) has prohibitive complexity, since the cardinality of the input space $\mathcal{X}$ grows exponentially with the block length. For unquantized observations, it is known that the solution can be obtained with linear-logarithmic complexity [7]. The key idea used to obtain this complexity reduction works for quantized observations as well, as we soon demonstrate.

First, we make some observations resulting due to the symmetry of our channel model. As before, we let $a = \frac{K}{M}$, and $\mathbf{q} = [q_0 \cdots q_{L-1}]$, where q_l is the quotient obtained on dividing z_l by a. Using Property A-4, we get

$$\hat{\mathbf{x}}(\mathbf{z}) = \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmax}} \max_{\phi \in [0, 2\pi)} \mathsf{P}(\mathbf{z} \bmod a | \mathbf{x} - \mathbf{q}, \phi) \ , \tag{6}$$

which in turn gives

$$\hat{\mathbf{x}}(\mathbf{z}) = \hat{\mathbf{x}}(\mathbf{z} \mod a) + \mathbf{q}(\mathbf{z}) \ , \tag{7}$$

where we have explicitly noted that $\mathbf{q}$ is a function of $\mathbf{z}$. This result implies that the solution for a received vector $\mathbf{z}$ can be easily obtained if the solution for $\mathbf{z} \bmod a$ is known, since computing $\mathbf{q}(\mathbf{z})$ is trivial. Hence, we restrict attention to computing the GLRT solution only for vectors $\mathbf{z}$ with components in $\{0, 1, \cdots, a-1\}$. Further, observe that $\mathsf{P}(\mathbf{z}|\mathbf{x}, \phi) = \mathsf{P}(\mathbf{z}|\mathbf{x} + i, \phi - i\frac{2\pi}{M})$. This implies that the demodulator cannot distinguish between two input vectors that are related by the operation of constant addition. This is well known (for unquantized observations), and is the basis for using techniques such as differential modulation.

The key to low-complexity block demodulation is to interchange the order of maximization in (5). Consider

$$\max_{\phi} \max_{\mathbf{x} \in \mathcal{X}} \mathsf{P}(\mathbf{z}|\mathbf{x}, \phi) \ . \tag{8}$$

For a fixed ϕ, the inner maximization over $\mathbf{x}$ is straightforward, since it is accomplished using coherent symbol by symbol demodulation. For $\phi = 0$, denote the coherent solution by $\mathbf{c}(0) = [c_0(0) \cdots c_{L-1}(0)]$, dropping the dependence on $\mathbf{z}$ to simplify notation. Note that $c_l(0) = \underset{x_l \in \{0, \cdots, M-1\}}{\operatorname{argmax}} \mathsf{P}(z_l|x_l, \phi = 0)$. As ϕ increases, the coherent solution $\mathbf{c}$ changes at certain crossover points. This happens

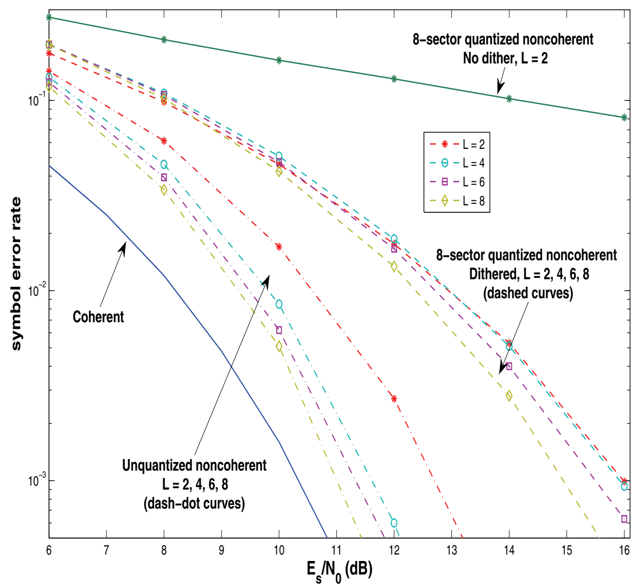

Fig. 3. Symbol error rate performance for QPSK with 8-sector phase quantization, for block lengths varying from 2 to 8. Also shown for comparison are the curves for coherent QPSK, and unquantized block noncoherent QPSK.

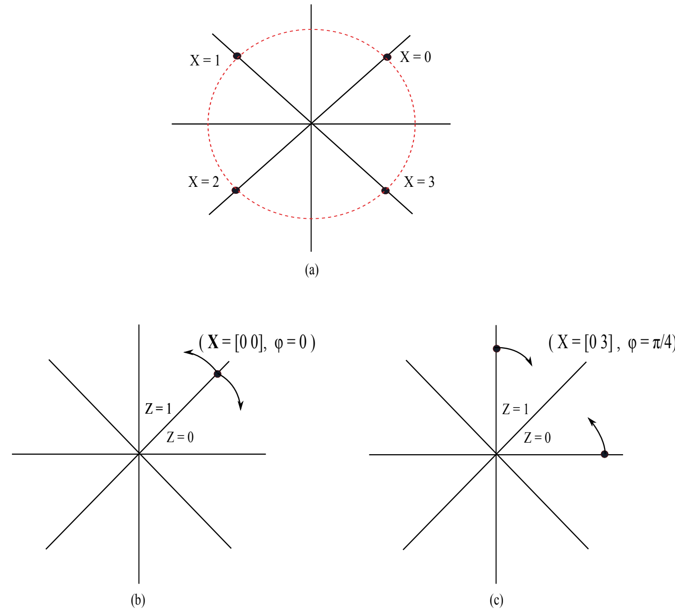

Fig. 4. Ambiguity in the block noncoherent demodulator. If the received vector is $[1\ 0]$, then $(\mathbf{X} = [0\ 0], \Phi = 0)$, and, $(\mathbf{X} = [0\ 3], \Phi = \frac{\pi}{4})$ are both equally likely solutions. In other words, both of these preceding combinations maximize the probability $\mathsf{P}(\mathbf{Z} = [1\ 0] \mid (\mathbf{X}, \Phi))$. [(a) depicts the QPSK constellation, while (b) and (c) are intended to illustrate the equality of the conditional probabilities $\mathsf{P}(\mathbf{Z} = [1\ 0] \mid (\mathbf{X} = [0\ 0], \Phi = 0))$ and $\mathsf{P}(\mathbf{Z} = [1\ 0] \mid (\mathbf{X} = [0\ 3], \Phi = \pi/4))$.]

only when any of the individual solutions c_l changes. The crucial observation now is that as ϕ is varied from 0 to $\frac{2\pi}{M}$, each of the individual solutions $c_l(\phi)$ changes only once. In other words, for each l, there is a *crossover angle* [7] α_l, such that

$$\begin{aligned} c_l(\phi) &= c_l(0)\ , \quad \text{if} \quad 0 \leq \phi \leq \alpha_l \\ &= c_l(0) - 1\ , \quad \text{if} \quad \alpha_l < \phi < \frac{2\pi}{M}\ . \end{aligned} \tag{9}$$

The exact crossover angles are easy to obtain as functions of z_l, K, M and the locations of the input constellation points. Now, since we only consider those $\mathbf{z}$ vectors for which every component $\in \{0, \cdots, a-1\}$, there can be at most a distinct crossover angles. Hence, when ϕ is varied between $[0, \frac{2\pi}{M})$, the number of distinct coherent solutions to the inner maximization in (8) is at most a, and these solutions can be obtained simply by sorting the crossover angles in an ascending order. For each of these (at most) a input vectors, we can now numerically compute the metric $\max\limits_{\phi \in [0,2\pi)} \mathsf{P}(\mathbf{z}|\mathbf{x}, \phi)$, and pick the one with the largest metric as the GLRT solution. This numerical computation can be done, for example, by fine discretization of the interval $[0, 2\pi)$, and computing $\mathsf{P}(\mathbf{z}|\mathbf{x}, \phi)$ for every ϕ in this discrete set. The number of computations (multiplications) required to obtain $\max\limits_{\phi \in [0,2\pi)} \mathsf{P}(\mathbf{z}|\mathbf{x}, \phi)$ then scales linearly in the block length L.

Note that it suffices to restrict attention to $\phi \in [0, \frac{2\pi}{M})$ while performing the inner maximization in (8). This is because, for $\phi > \frac{2\pi}{M}$, any new solution that we obtain, say $\mathbf{c}_1$, is related to one of the existing solutions, say $\mathbf{c}_2$, by the operation of constant addition, so that the noncoherent demodulator cannot distinguish between $\mathbf{c_1}$ and $\mathbf{c_2}$.

VI. Numerical Results

We present results for QPSK with 8-sector and 12-sector phase quantization, for different block lengths L. We begin with the symbol error rate (SER) plots for block demodulation, obtained by averaging over 10^5 blocks of data. For block length L, the number of information carrying symbols transmitted, therefore, is, $10^5 \times (L-1)$. Fig. 3 shows the results for 8-sector quantization. Looking at the topmost curve, which corresponds to $L = 2$, we find that the performance is extremely poor. As SNR is increased, the SER falls off extremely slowly. A close analysis of the block demodulator reveals that the reason behind this is an ambiguity in the demodulator decision rule: for certain outputs $\mathbf{z}$, irrespective of the SNR, the demodulator always returns two equally likely solutions for the input $\mathbf{x}$. While we do not provide a complete analysis of this ambiguous behavior, an example scenario is shown in Fig. 4 to give insight. If the quantized output vector is $\mathbf{Z} = [1\ 0]$, then we find that $\mathsf{P}(\mathbf{Z}|\mathbf{X}, \Phi)$ is maximized by two equally likely pairs, $(\mathbf{X}_1, \Phi_1) = ([0\ 0], 0)$, and $(\mathbf{X}_2, \Phi_2) = ([0\ 3], \pi/4)$. Thus, the block demodulator, which does joint maximum likelihood estimation over the input and the unknown phase, becomes ambiguous. In other words, the symmetry inherent in the channel model, which on the one hand helped us reduce the complexity of the required computations, is also making it impossible to distinguish between the effect of the unknown phase offset and the phase modulation on the received signal, resulting in poor performance. While we have shown results only for block length $L = 2$, we find that the ambiguity persists for larger block lengths, and that SER performance does not improve with increasing block length.

One possible approach to break undesirable symmetries is the use of *dithering*, where our usage of the term denotes a deterministic *phase shift* in the PSK constellation, from one symbol to the next. Specifically, consider a dither scheme in which we shift the PSK constellation by an angle of $\frac{1}{L}(\frac{2\pi}{K})$

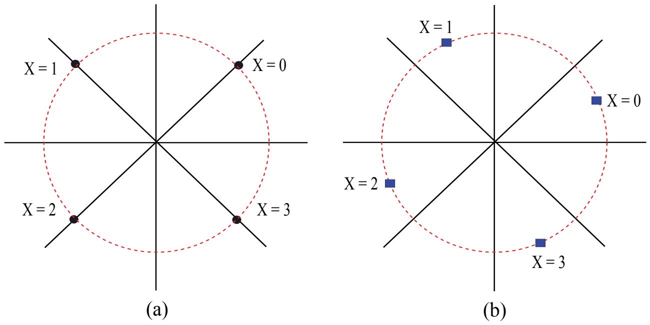

Fig. 5. Constellation dithering : The constellations used for the two symbols are not identical, but the one used on the second symbol (depicted in (b)) is a dithered version of the one used on the first symbol (depicted in (a)).

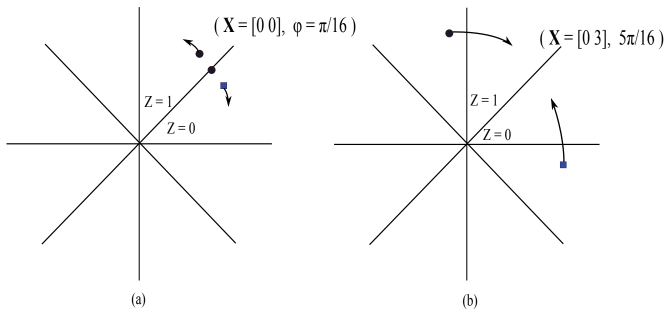

Fig. 6. Use of the dithered constellation on the second symbol alleviates the ambiguity in the block noncoherent demodulator. The figure illustrates that the conditional probability $\mathsf{P}(\mathbf{Z} = [1\ 0] \mid (\mathbf{X} = [0\ 0], \Phi = \pi/16)) > \mathsf{P}(\mathbf{Z} = [1\ 0] \mid (\mathbf{X} = [0\ 3], \Phi = 5\pi/16))$.]

from one symbol to the next (the shift may be in clockwise direction, or counterclockwise direction). Fig. 5 shows this scheme for QPSK with block length $L = 2$ and $K = 8$. The constellation used for the second symbol (shown by the squares) is dithered from the constellation used for the first symbol by an angle of $\pi/8$. Fig. 6 indicates how this choice of transmit constellations helps remove the ambiguity in the demodulator output. Considering again the exemplary scenario of the quantized output vector $\mathbf{Z} = [1\ 0]$, we find that the conditional probability $\mathsf{P}(\mathbf{Z}|\mathbf{X}, \Phi)$ is now maximized uniquely, using the combination $(\mathbf{X}, \Phi) = ([0\ 0], \pi/16)$. As far as the input symbol pair $\mathbf{X} = [0\ 3]$ is concerned, the optimal Φ is $5\pi/16$, however the conditional probability $\mathsf{P}(\mathbf{Z} \mid (\mathbf{X} = [0\ 3], \Phi = 5\pi/16)) < \mathsf{P}(\mathbf{Z} \mid (\mathbf{X} = [0\ 0], \Phi = \pi/16))$. Since the ambiguity in the demodulator has been eliminated, the performance is expected to improve. The results in Fig. 3 indeed show a significant performance improvement compared to the no-dithering case. With dithering, at SER of 10^{-3}, 8-sector quantization with $L = 8$ results in a loss of about 4 dB compared to unquantized observations.

Another approach for breaking undesirable symmetries is through choice of quantizer. For example, if we consider the performance with 12-sector quantization, it is observed that the block demodulator performs well, and that dithering is not required. This suggests that 12-sector quantization does not result in undesirable symmetries (although we do not prove this). Fig. 7 shows the performance for different block lengths. At SER of 10^{-3}, and $L = 8$, the loss compared to the unquantized observations is reduced to about 2 dB.

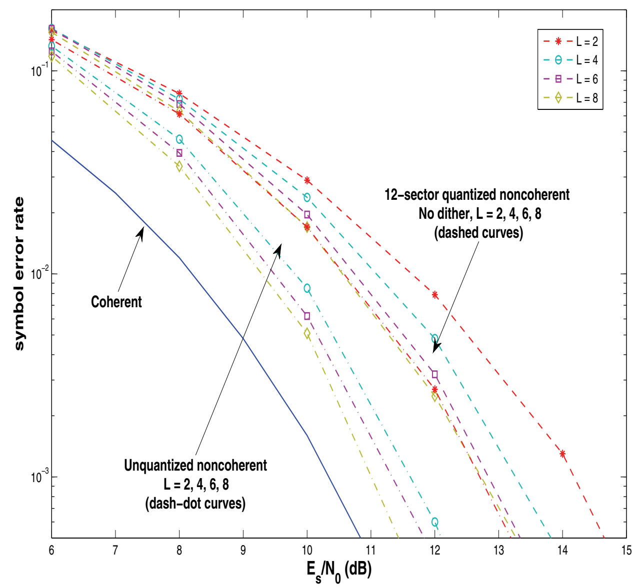

Fig. 7. Symbol error rate performance for QPSK with 12-sector phase quantization, for block lengths varying from 2 to 8. Also shown for comparison are the curves for coherent QPSK, and unquantized block noncoherent QPSK.

Note that, in both of the preceding sets of quantized block noncoherent results (8-sector quantization with dithering, and 12-sector quantization without dithering), increasing the block length provides performance gains at higher SNRs. However, at low SNRs, we observed that increasing the block length can result in slight performance degradation (visible to some extent in Fig. 3). This is contrary to the performance with unquantized observations, where increasing the block length provides gains across all SNRs. We do not have a clear understanding of this behavior with quantized observations. One possible explanation might be that, while increasing the block length increases the minimum separation between quantized output sequences, thus improving performance at higher SNR, it might at the same time increase the number of nearest neighbors, leading to poorer performance at lower SNR. We leave further investigation of this behavior as an open issue.

Next we show plots for input-output mutual information. For all our results, we normalize the mutual information $I(\mathbf{X}; \mathbf{Z})$ by L-1 to obtain the per symbol mutual information, since in practice successive blocks can be overlapped by one symbol due to slow phase variation from one block to the next [4]. Fig. 8 shows the results for 8-sector quantization (without dithering), for block lengths $L = \{2, 4, 6, 8\}$. Also shown for reference are the mutual information values for coherent communication, and for unquantized block nonco-

[4]If the phase varies slowly across blocks, the last symbol from the previous transmitted block can act as a “reference” for the current block, and hence does not need to be transmitted again. For example, consider differential modulation with block length $L = 2$. The information is conveyed by the difference in the phase of the two symbols in the block, so that the first symbol acts as a reference. If the channel varies slowly across the blocks, the second symbol sent in the previous block may be used as the reference for the current block, and hence does not need to be sent again. See, for example, [6], [10] for detailed discussion.

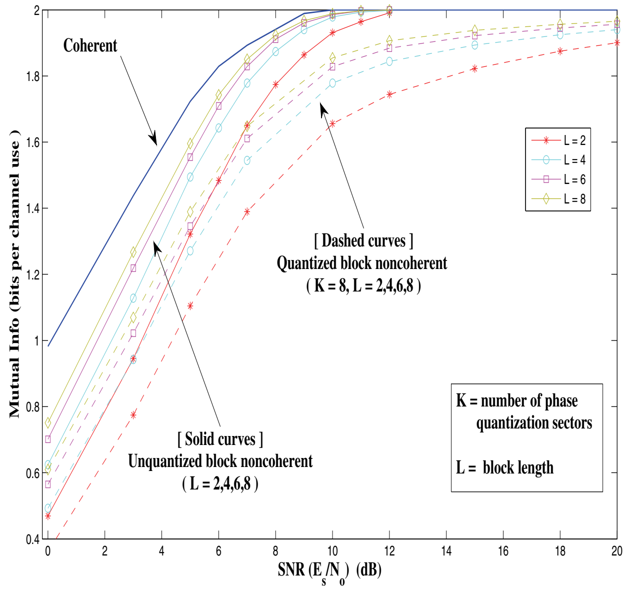

Fig. 8. Mutual information for QPSK with 8-sector phase quantization, for block lengths varying from 2 to 8. Also shown for comparison are the curves for coherent QPSK, and unquantized block noncoherent QPSK.

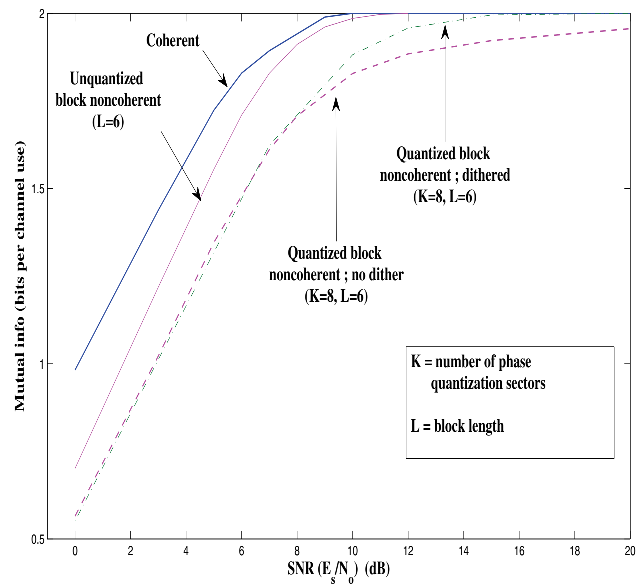

Fig. 9. Mutual information with dithering : For block length $L = 6$, plots depict the mutual information for QPSK with 8-sector quantization (with and without dithering). Also shown are the curves for coherent QPSK, and unquantized block noncoherent QPSK.

herent communication [5]. Despite the poor performance of 8-sector quantization for uncoded SER (in the absence of any dithering) witnessed earlier, we see that, in terms of the input-output mutual information, 8-sector quantization scheme recovers more than 80-85% of the mutual information obtained with unquantized observations (for identical block length), for SNR $>$ 2-3 dB. Further, the mutual information increases with increase in the block length (the SER performance did not show any improvement with increasing block length). However, note that, as the SNR is increased, the mutual information with 8-sector quantization approaches 2 bits/channel use extremely slowly (evident for all block lengths). Since $H(\mathbf{X})$ is constant, this implies that $H(\mathbf{X}|\mathbf{Z})$ falls off very slowly as SNR $\to \infty$, which is consistent with the earlier observation that there is significant ambiguity in $\mathbf{X}$, given $\mathbf{Z}$, even at high SNR. The mutual information performance with dithered-QPSK is shown in Fig. 9 [6]. (To avoid clutter, we show results only for $L = 6$.) It is seen that the slow increase of the mutual information towards 2 bits/channel use is eliminated by dithering.

While the simple constellation dither scheme considered here has provided performance gains (in terms of both SER and mutual information), we hasten to add that there is no claim of optimality associated with it. A more detailed investigation of different dithering schemes and their potential gains is therefore an important topic for future research.

In Fig. 10, we plot the mutual information curves for QPSK with 12-sector quantization (without constellation dithering) for block length $L = \{2, 4, 6, 8\}$. Also shown for reference are the coherent and unquantized block noncoherent performance curves. For identical block lengths, the loss in mutual information (at a fixed SNR $>$ 2-3 dB) compared to the unquantized scenario is less than 5-10 %, while the loss in power efficiency (for fixed mutual information) varies between 0.5-2 dB.

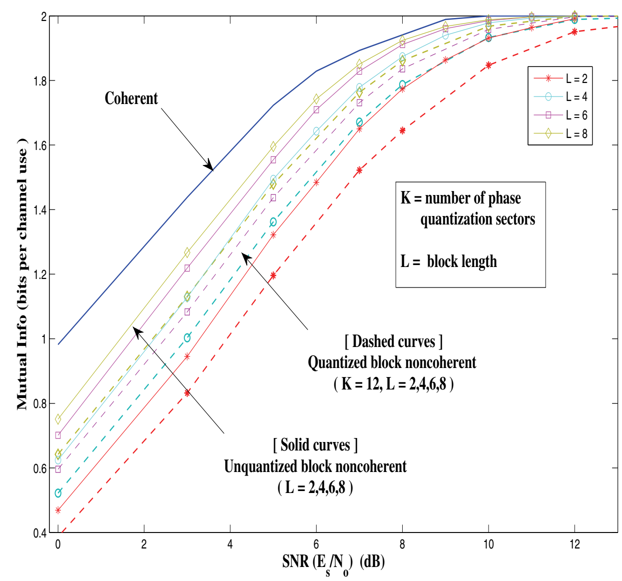

Fig. 10. Mutual information for QPSK with 12-sector phase quantization, for block lengths varying from 2 to 8. Also shown for comparison are the curves for coherent QPSK, and unquantized block noncoherent QPSK.

VII. Conclusions

We have investigated information-theoretic benchmarks and uncoded error rates for block noncoherent communication with low-precision quantization. Building up on our earlier

[5]Unquantized mutual information was computed using Monte Carlo simulation, as done previously in [10] as well. See Appendix C for details.

[6]The low-complexity mutual information computation procedure in Section IV was developed without considering any constellation dithering. Since we have not studied the applicability of these complexity reduction techniques for dithered constellations, we use Monte Carlo simulations to compute mutual information with dithering. See Appendix C for details.

results for ADC-constrained communication under ideal channel conditions [4], the results here indicate that it is possible to handle carrier phase/frequency asynchronism despite drastic quantization at the receiver. Our results show that such quantization might lead to ambiguities causing error floors, but that the symmetries leading to such ambiguities can be broken by dithering at the transmitter, or by appropriate choice of quantizer at the receiver. It is interesting to note that dithering has also been found to be useful for other aspects of ADC-constrained receiver design, such as channel estimation [24] and automatic gain control [23]. Thus, an important topic for further investigation is to develop a deeper understanding of symmetry breaking strategies such as dithering and asymmetric receiver quantization for ADC-constrained communication.

Other topics for future work include tackling the problem of timing synchronization as well as carrier synchronization, together with extension to QAM constellations using amplitude as well as phase quantization. Study of the phase quantization architecture, when the PSK constellation size does not divide the number of quantization sectors (an assumption we made throughout this work) might also be worth pursuing.

Our block noncoherent model is based on fixed analog preprocessing prior to quantization (e.g., the analog linear combinations followed by one bit quantization used to implement a given phase quantizer), which allows us to compute analytical performance benchmarks. However, from the point of view of transceiver design, it is also of interest to explore feedback-based adaptive analog preprocessing followed by low-precision quantization. For example, one could accomplish carrier synchronization using derotation of the received samples prior to quantization, with the derotator block adapted based on the output of the quantizer (while this architecture is similar to conventional DSP-based carrier synchronization, the number of bits at the output of the quantizer would be much smaller).

Our focus in this paper was on fundamental communication-theoretic questions. An interesting topic for future research is to perform a complete design of an ADC-constrained communication link building on the insights developed here and in related papers such as [4], [24], [23]. For example, the non-dispersive AWGN channel model considered here is a good approximation for a short-range, near-line-of-sight, 60 GHz link exploiting asymmetry (e.g., a powerful laptop transmitting to a low-power handheld receiver), where the transmitter can employ precoding to significantly reduce the channel dispersion seen by the receiver.

Appendix A
Computation of $H(\mathbf{Z}|\mathbf{x}_0)$

Instead of jointly accounting for constant addition and permutation, we first account for constant addition, and then for permutation. Specifically, we first note that using Property B-1, it suffices to compute $\mathsf{P}(\mathbf{z}|\mathbf{x}_0)$ only for a set of vectors $S_{\mathbf{Z}_1}$ for which the first symbol is 0. Next, using the fact that $\mathsf{P}(\mathbf{z}|\mathbf{x}_0) = \mathsf{P}(\Pi\mathbf{z}|\mathbf{x}_0)$, within the set $S_{\mathbf{Z}_1}$, we can further restrict attention to a subset $S_{\mathbf{Z}_2}$ in which no vector can be obtained from another one by a permutation operation. Specifically, since permutations don't matter, all we are interested in is how many symbols of each type are picked, so that obtaining the set $S_{\mathbf{Z}_2}$ is equivalent to the well-known problem of distributing L–1 identical balls into K distinct boxes, with empty boxes allowed. The number of ways to do this is $\binom{K+L-2}{L-1}$, and each of these combinations can be obtained easily using standard known procedures.

Once we have the set $S_{\mathbf{Z}_2}$, we can numerically compute the probability $\mathsf{P}(\mathbf{z}|\mathbf{x}_0)$ for every vector in $S_{\mathbf{Z}_2}$. The entropy $H(\mathbf{Z}|\mathbf{x}_0)$ can then be obtained as follows. For $\mathbf{z} \in S_{\mathbf{Z}_2}$, let $n(\mathbf{z})$ denote the number of distinct permutations that can be generated from it, while keeping the first symbol fixed. This is equal to $\frac{(L-1)!}{\prod_{i=0}^{K-1} r_i}$, where r_i is the number of times the symbol i occurs in $\mathbf{z}$. The conditional entropy then is $H(\mathbf{Z}|\mathbf{x}_0) = -\sum_{\mathcal{Z}} \mathsf{P}(\mathbf{z}|\mathbf{x}_0)\log \mathsf{P}(\mathbf{z}|\mathbf{x}_0) = -\sum_{S_{\mathbf{Z}_1}} K\mathsf{P}(\mathbf{z}|\mathbf{x}_0)\log \mathsf{P}(\mathbf{z}|\mathbf{x}_0) = -\sum_{S_{\mathbf{Z}_2}} K\ n(\mathbf{z})\ \mathsf{P}(\mathbf{z}|\mathbf{x}_0)\log \mathsf{P}(\mathbf{z}|\mathbf{x}_0)$.

Appendix B
Computation of $P(\mathbf{Z})$

To obtain the set $S_{\mathbf{X}}$, we divide the L locations into a groups, and permutations are allowed only between locations belonging to the same group. The problem then breaks down into a sub-problems. Specifically, let the number of locations in the groups be $n_0, n_1, \cdots, n_{a-1}$, then we need to distribute n_i identical balls into M distinct boxes, for each i. The required number of combinations is the product of the individual solutions. While for large a, the reduction in complexity may not be huge, for small values of a (which is the paradigm of interest in this work), the savings will be significant. For instance, for QPSK with $L = 8$, and $a = 2$, the worst case (which happens when $n_0 = n_1$) number of combinations is 1225, compared to the exponential figure of $M^L = 65536$. Once the set S_X has been obtained, we can get $\mathsf{P}(\mathbf{z}) = \frac{1}{M^L}\sum_{\mathbf{x}\in S_X} q(x)\mathsf{P}(\mathbf{z}|\mathbf{x})$. Here, $q(x) = \prod_{i=0}^{a-1} \frac{(n_i)!}{\prod_{j=0}^{M-1}(r_{i,j}(x))!}$, where $r_{i,j}(x)$ is the number of times the input symbol j occurs in the locations belonging to group i, for the vector x.

Appendix C
Monte Carlo Simulation for Mutual Information Computation

Unquantized Block Noncoherent: Referring to the notation in Section II, the unquantized block noncoherent channel is

$$Y_l = e^{j\theta X_l} e^{j\Phi} + N_l\ ,\ l = 0, 1, \cdots, L-1\ , \tag{10}$$

where $\mathbf{X} = [X_0 \quad X_1 \quad \cdots \quad X_{L-1}]$ is the block of input symbols, and $\mathbf{Y} = [Y_0 \ \ Y_1 \ \ \cdots \ \ Y_{L-1}]$ is the block of output symbols. The mutual information $I(\mathbf{X};\mathbf{Y})$ can be written as

$$I(\mathbf{X};\mathbf{Y}) = \mathbb{E}_{\mathbf{X},\mathbf{Y}}\left[\log_2 \frac{\mathsf{P}(\mathbf{Y}|\mathbf{X})}{\mathsf{P}(\mathbf{Y})}\right]\ , \tag{11}$$

where the conditional probability $\mathsf{P}(\mathbf{Y}|\mathbf{X})$ can be analytically expressed in a closed form, in terms of the modified Bessel function of the first kind (see [10, eq. (10)], [5, eq. (9)]), and the unconditional probability $\mathsf{P}(\mathbf{Y}) = \sum_{\mathbf{X}} \mathsf{P}(\mathbf{Y}|\mathbf{X})\mathsf{P}(\mathbf{X})$. Numerical computation of the mutual information is performed by generating several random samples of $(\mathbf{X}, \mathbf{Y})$ (according to the channel model 10), and computing the average of

$\log_2 \frac{\mathsf{P}(\mathbf{Y}|\mathbf{X})}{\mathsf{P}(\mathbf{Y})}$ over this sample set. In our simulations (carried out for block lenghts $L = \{2, 4, 6, 8\}$), we averaged over a sample set of size 20,000 (the results did not show significant variation beyond 5,000-10,000 samples.)

Quantized Block Noncoherent: The quantized block noncoherent channel is given in (2). We use Monte Carlo simulation to compute the mutual information of this channel, under input constellation dithering (explained in Section VI). To compute the mutual information $I(\mathbf{X}; \mathbf{Z}) = H(\mathbf{Z}) - H(\mathbf{Z}|\mathbf{X})$, we need the conditional probabilities $\mathsf{P}(\mathbf{Z}|\mathbf{X})$ for all values of $\mathbf{Z}$ and $\mathbf{X}$. We compute these probabilities numerically. Specifically, for each $\mathbf{X}$, we generate N_0 samples of $\mathbf{Z}$ (based on the channel model (2)). For $\mathbf{X} = \mathbf{x}$, the number of times $\mathbf{Z} = \mathbf{z}$ occurs (normalized by N_0) gives an estimate of $\mathsf{P}(\mathbf{z}|\mathbf{x})$.

The simulation is sped up to some extent by exploiting the simple symmetries in the channel model, which hold even when constellation dithering is employed. Specifically, due to the circular symmetry of the Gaussian noise, it is easy to verify that, even under constellation dithering, the conditional probability $\mathsf{P}(\mathbf{Z}|\mathbf{X})$ remains invariant under a constant addition to the input vector, or, to the output vector. Hence, we need to compute $\mathsf{P}(\mathbf{Z}|\mathbf{X})$ only for a subset of input vectors (resp. a subset of output vectors) in which no input vector (resp. no output vector) can be obtained from another by constant addition. Such input and output subsets are obtained trivially by considering only those vectors for which the first element is 0, cutting down the number of input vectors from M^L to M^{L-1} (for M-PSK with block length L), and the number of output vectors from K^L to K^{L-1} (for K-sector quantization.) Note that, in the simulation set up described in the preceding paragraph, for each input vector, we generate N_0 samples of $\mathbf{Z}$ and count the number of occurrences of different outputs to get the conditional probabilities. The analysis here implies that, (a) we need to do this for M^{L-1} input vectors, and, (b) the counts need to be stored for K^{L-1} output vectors. (An occurrence of an output vector outside of this subset can be trivially mapped to one of the vectors in this subset, by performing a constant addition operation that results in the first element being 0. In this case, the final counts corresponding to each of the output vectors in the subset are further normalized by K to get the conditional probabilities.) For our simulation results depicted in Fig. 9, we took N_0, i.e., the number of output vector samples generated for each input vector to be $K^{L-1} \times 100$. (With $L = 6$, we tried $N_0 = K^{L-1} \times R$, for $R = 1, 10, 30, 60, 100$, and did not observe significant variation in the results beyond $R = 10$.)

ACKNOWLEDGMENTS

The authors are grateful to the anonymous reviewers for their valuable comments.

REFERENCES

[1] P. Smulders, “Exploiting the 60 GHz band for local wireless multimedia access: prospects and future directions,” *IEEE Commun. Mag.*, pp. 140–147, Jan. 2002.

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

[3] L. Bin, T. W. Rondeau, J. H. Reed, and C. W. Bostian, “Analog-to-digital converters,” *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 69–77, Nov 2005.

[4] J. Singh, O. Dabeer, and U. Madhow, “On the limits of communication with low-precision analog-to-digital conversion at the receiver,” *IEEE Trans. Commun.*, vol. 57, no. 12, Dec. 2009.

[5] D. Divsalar and M. Simon, “Multiple-symbol differential detection of MPSK,” *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, Mar. 1990.

[6] D. Warrier and U. Madhow, “Spectrally efficient noncoherent communication,” *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 651–668, Mar. 2002.

[7] W. Sweldens, “Fast block noncoherent decoding,” *IEEE Commun. Lett.*, vol. 5, pp. 132–134, 2001.

[8] R.-R. Chen, R. Koetter, U. Madhow, and D. Agrawal, “Joint noncoherent demodulation and decoding for the block fading channel: a practical framework for approaching Shannon capacity,” *IEEE Trans. Commun.*, Oct. 2003.

[9] P. Grover, K. A. Woyach, and A. Sahai, “Towards a communication-theoretic understanding of system-level power consumption,” *IEEE J. Sel. Areas Commun.*, Sep. 2011.

[10] M. Peleg and S. Shamai, “On the capacity of the blockwise incoherent MPSK channel,” *IEEE Trans. Commun.*, vol. 46, pp. 603–609, Mar. 1998.

[11] J. Singh and U. Madhow, “On block noncoherent communication with low-precision phase quantization at the receiver,” in *2009 IEEE International Symposium on Information Theory*.

[12] J. Singh, P. Sandeep, and U. Madhow, “Multi-gigabit communication: the ADC bottleneck,” in *2009 IEEE International Conference on Ultra-Wideband*, invited paper, 2009.

[13] D. A. Sobel and R. W. Brodersen, “A 1 Gb/s mixed-signal baseband analog front-end for a 60 GHz wireless receiver,” *IEEE J. Solid-State Circuits,*, vol. 44, no. 4, pp. 1281–1289, Apr. 2009.

[14] S. Krone and G. Fettweis, “Achievable rate of single-carrier systems with optimal uniform quantization at the receiver,” in *2010 IEEE Information Theory Workshop*.

[15] T. Marzetta and B. Hochwald, “Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading,” *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[16] N. Jacobsen and U. Madhow, “Coded noncoherent communication with amplitude/phase modulation: from Shannon theory to practical turbo architectures,” *IEEE Trans. Commun.*, vol. 56, no. 12, pp. 2040-2049, Dec. 2008.

[17] O. Dabeer, J. Singh, and U. Madhow, “On the limits of communication performance with one-bit analog-to-digital conversion,” in *2006 IEEE Workshop on Signal Processing Advances in Wireless Communication*.

[18] J. Singh, O. Dabeer, and U. Madhow, “Communication limits with low-precision analog-to-digital conversion at the receiver,” in *2007 IEEE International Conference on Communications*.

[19] ——, “Capacity of the discrete-time AWGN channel under output quantization,” in *2008 IEEE International Symposium on Information Theory*.

[20] T. Koch and A. Lapidoth, “Increased capacity per unit-cost by oversampling,” in *2010 IEEE Convention of Electrical and Electronics Engineers in Israel*.

[21] S. Krone and G. Fettweis, “Achievable rate with 1-bit quantization and oversampling at the receiver,” in *2010 IEEE Communication Theory Workshop*.

[22] T. Koch and A. Lapidoth, “Asymmetric quantizers are better at low SNR,” in *2011 IEEE International Symposium on Information Theory*.

[23] F. Sun, J. Singh, and U. Madhow, “Automatic gain control for ADC-limited communication,” in *2010 IEEE Global Telecommunications Conference*.

[24] O. Dabeer and U. Madhow, “Channel estimation with low-precision analog-to-digital conversion,” in *2010 IEEE International Conference on Communicaion*.

[25] S. Krone and G. Fettweis, “Fading channels with 1-bit output quantization: optimal modulation, ergodic capacity and outage probability,” in *2010 IEEE Information Theory Workshop*.

[26] A. Mezghani and J. Nossek, “Analysis of Rayleigh-fading channels with 1-bit quantized output,” in *2008 IEEE International Symposium on Information Theory*.

[27] S. Chandrasekaran, S. K. Mohammed, and A. Chockalingam, “On the capacity of quantized Gaussian MAC channels with finite input alphabet,” in *2011 IEEE International Conference on Communicaion*.

[28] G. Zeitler, A. C. Singer, and G. Kramer, “Low-precision A/D conversion for maximum information rate in channels with memory,” *IEEE Trans. Commun.*, vol. 60, no. 9, Sep. 2012.

[29] C. Nannapaneni, M. C. Valenti, and X. Xingyu, “Constellation shaping for communication channels with quantized outputs,” in *2011 Annual Conference on Information Sciences and Systems*.

[30] W. Zhang, "A general framework for transmission with transceiver distortion and some applications," *IEEE Trans. Commun.*, vol. 60, no. 2, Feb. 2012.

Jaspreet Singh received his bachelor's degree in electrical engineering from the Indian Institute of Technology, Delhi, in 2004. He received the M. A. degree in statistics in 2009, and the M. S. and Ph.D. degrees in electrical and computer engineering, in 2005 and 2009, respectively, from the University of California, Santa Barbara. He is currently a Senior Researcher at Samsung Research America-Dallas, Texas. He was an intern at the Swiss Federal Institute of Technology, Lausanne, and at Qualcomm Flarion Technologies, New Jersey, during the summers of 2003 and 2008, respectively. In summer 2007, he was a visiting graduate student at the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge. Dr. Singh's main research interests are in wireless communication, information theory, and, signal processing, with current focus on LTE/LTE-A terminal modem algorithms, and system design for millimeter wave/Massive MIMO communication.

Upamanyu Madhow is Professor of Electrical and Computer Engineering at the University of California, Santa Barbara. His research interests broadly span communications, signal processing and networking, with current emphasis on millimeter wave communication and bio-inspired approaches to networking and inference. He received his bachelor's degree in electrical engineering from the Indian Institute of Technology, Kanpur, in 1985, and his Ph. D. degree in electrical engineering from the University of Illinois, Urbana-Champaign in 1990. He has worked as a research scientist at Bell Communications Research, Morristown, NJ, and as a faculty at the University of Illinois, Urbana-Champaign. Dr. Madhow is a recipient of the 1996 NSF CAREER award, and co-recipient of the 2012 IEEE Marconi prize paper award in wireless communications. He has served as Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON INFORMATION THEORY, and the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. He is the author of the textbook *Fundamentals of Digital Communication* (Cambridge University Press, 2008).