A Novel Receiver Design With Joint Coherent and Non-Coherent Processing

Wanchun Liu, *Student Member, IEEE*, Xiangyun Zhou, *Member, IEEE*,
Salman Durrani, *Senior Member, IEEE*, and Petar Popovski, *Fellow, IEEE*

Abstract—In this paper, we propose a novel splitting receiver, which involves a joint processing of coherently and non-coherently received signals. Using a passive RF power splitter, the received signal at each receiver antenna is split into two streams, which are then processed by a conventional coherent detection (CD) circuit and a power-detection (PD) circuit, respectively. The streams of the signals from all the receiver antennas are then jointly used for information detection. We show that the splitting receiver creates a 3-D received signal space due to the joint coherent and non-coherent processing. We analyze the achievable rate of a splitting receiver, which shows that the splitting receiver provides a rate gain of 3/2 compared with either the conventional (CD-based) coherent receiver or the PD-based non-coherent receiver in the high SNR regime. We also analyze the symbol error rate (SER) for practical modulation schemes, which shows that the splitting receiver achieves asymptotic SER reduction by a factor of at least $\sqrt{M}-1$ for M-QAM compared with either the conventional (CD-based) coherent receiver or the PD-based non-coherent receiver.

Index Terms—Receiver design, energy/power-detection based communications, joint coherent and non-coherent processing, RF signal splitting.

I. INTRODUCTION

A. Background and Motivation

WIRELESS communications has witnessed several major theoretical advancements in the last few decades, which have been quickly incorporated into communication standards, e.g., multiple input multiple output (MIMO) systems and orthogonal frequency division multiple access (OFDMA) [1]. The baseband receiver design, underlying these technologies, has been the receiver based on a coherent detection (CD), i.e., coherent receiver for short, which has been adopted exclusively in nearly all the popular wireless communication standards. The coherent receiver design has stayed virtually unchanged throughout the evolution of wireless communication systems to date.

The recent trend towards a large number of antennas at the transmitter or receiver (such as in massive MIMO and millimeter wave systems [2], [3]) provides incentive to rethink the coherent receiver design for future communication systems. This is because provision of accurate channel state information (CSI) becomes challenging in scenarios with a massive number of antennas. In this regard, *power/intensity-detection* (PD) based *non-coherent* receivers have been proposed [4] and [5]. Although, in general, non-coherent receivers suffer from a performance loss compared to coherent receivers, their low cost and low power consumption makes them attractive for systems with a large number of antennas. It is proved in [4] that PD-based non-coherent modulation in a massive single input multiple output (SIMO) system can achieve a scaling law which is the same as the coherent modulation scheme with an increasing number of antennas. It is also shown in [5] that the performance of the PD-based non-coherent modulation asymptotically approaches the performance of coherent detection in high SNR regimes. These studies show that PD-based non-coherent receivers can offer a performance comparable to coherent receivers in future wireless systems.

Another motivation for considering PD-based non-coherent receivers comes from the recent interest in simultaneous wireless information and power transfer (SWIPT) [6]–[9]. In such systems, a user is assumed to be equipped with an energy receiver which is based on non-coherent RF (radio frequency)-to-DC (direct current) conversion, and a conventional coherent (information) receiver. In SWIPT systems, the user employs the energy receiver and the coherent receiver *separately*: (i) In time-switching, the user switches between the two receivers, depending on whether it is in RF energy harvesting (EH) mode or information detection mode, or (ii) in power-splitting, the user splits the received signal into two streams, and then sends one stream to the energy receiver and the other to the coherent receiver.

Although the PD-based non-coherent receiver has received more attention in wireless communication systems recently, it has played an important role in optical communications for a long time. For the low-cost wireless infrared communication system [10], *intensity modulation* (IM) is the most commonly adopted modulation scheme. For IM-based wireless infrared communication, information is carried by the instantaneous power of the carrier, and the receiver uses a photodetector to produce a current proportional to the received instantaneous power directly, i.e., PD-based non-coherent demodulation. The wireless infrared communication channel is usually referred

to as the intensity channel or the *non-coherent additive white Gaussian noise (AWGN) channel*, echoing the coherent AWGN channel. The modulation schemes and the capacity of the non-coherent AWGN channel were studied in [11] and [12], respectively.

B. Novel Contributions

Motivated by the recent interest in the PD-based non-coherent receiver, we consider a basic point-to-point communication system and revisit the design of the communication receiver. Rather than focusing on an improved design for either coherent receiver or non-coherent receiver alone, we consider a receiver with joint coherent and PD-based non-coherent processing. To the best of our knowledge, this is an open problem in the literature and it is not immediately clear whether joint processing will be better than either coherent or PD-based non-coherent processing alone. In this work, we show that it can in fact significantly improve the achievable rate and also reduce the symbol error rate (SER).

The main contributions of the paper are summarized as follows:

1) We propose a novel information receiver architecture for a K-antenna receiver called *splitting receiver*. The received signal at each antenna is split into two streams by a passive power splitter with a certain *splitting ratio*. One stream is processed by a conventional (coherent) CD circuit, and the other is processed by a (non-coherent) PD circuit, and then the $2K$ streams of processed signal are jointly used for information detection.

2) As a variant of the splitting receiver, we also propose a simplified receiver where no power splitters are required and a fixed number of antennas are connected to CD circuits and the remaining antennas are connected to PD circuits. Analytically, the simplified receiver can be treated as a special case of a splitting receiver, where the splitting ratio at each antenna can only take 1 or 0.

3) We show that the splitting receiver (and also the simplified receiver), increases the dimension of the received signal space, since the noise adds linearly to the signal in the coherent receiver part and the noise adds to the squared amplitude of the signal in the PD-based non-coherent receiver part. This results in improved communication performance.

4) From an information-theoretic perspective, we model the channel introduced by the splitting receiver as a splitting channel. Assuming a Gaussian input to the splitting channel, in the high signal-to-noise-ratio (SNR) regime, we show analytically that: (i) The asymptotic maximum mutual information of the splitting channel is 3/2 times that of either the coherent AWGN channel or the non-coherent AWGN channel, under the same average received signal power constraint. (ii) For a splitting receiver with a single receiver antenna, the asymptotic optimal power splitting ratio is 1/3. (iii) For the simplified receiver with a large number of receiver antennas, connecting half the antennas to the CD circuits and the other half to the PD circuits is the optimal strategy.

5) For transmissions based on practical modulations, we analyze the symbol decision region and the symbol error rate (SER) at the splitting receiver. Considering high SNR regime, we derive the SER expression for a general modulation scheme. The analytical results show that, compared with the conventional coherent receiver, the splitting receiver achieves asymptotic SER reduction by a factor of $M - 1$ for M-PAM (pulse amplitude modulation) and $\sqrt{M} - 1$ for M-QAM (quadrature amplitude modulation).

C. Paper Organization and Notation

This paper is organized as follows. Section II presents the system model, the proposed receiver architectures and the splitting channel. Section III analyzes the mutual information of the splitting channel with a Gaussian input. Section IV presents the received signal constellation at the splitting receiver for practical modulation schemes. Section V shows the SER results of practical modulation schemes. Finally, Section VI concludes the paper.

Notation: $\tilde{\cdot}$ denotes a complex number. $(\cdot)^*$ and $|\cdot|$ denote the conjugate and the absolute-value norm of a complex number, respectively. Real$\{\cdot\}$ and Imag$\{\cdot\}$ denotes the real part and the imaginary part of a complex number, respectively. Pr$(\cdot)$ denotes the probability of an event. $h(\cdot)$, $h(\cdot, \cdot)$, $h(\cdot|\cdot)$ denote the differential entropy, joint and conditional differential entropy, respectively. $I(\cdot; \cdot)$ denotes the mutual information. Random variables and their realizations are denoted by upper and lower case letters, respectively. erfc$(\cdot)$ is the complementary error function, and $Q(x) \triangleq \frac{1}{2}\text{erfc}(\frac{x}{\sqrt{2}})$ is the Q-function.

II. SYSTEM MODEL

Consider the communication between a single-antenna transmitter and a K-antenna receiver. The average received signal power at each antenna is denoted by $\mathcal{P}$. The channel coefficient at the kth receiver antenna is denoted by $\tilde{h}_k$.

A. Proposed Receiver Architecture

1) Splitting Receiver: The proposed splitting receiver architecture is illustrated in Figs. 1(a) and (b). In the first stage, the received signal at each antenna is split into two streams by an ideal *passive RF power splitter*. We assume there is no power loss or noise introduced during the splitting process [13]–[15]. One stream is sent to the (conventional) CD circuit and the other to the PD circuit. The signals in the CD and PD circuits are first converted to the baseband signals and then sampled and digitized by the analog-to-digital converters (ADCs) accordingly, for further processing. Specifically, the rectifier-based PD circuit converts the RF signal into a DC signal with a conversion efficiency η. In the second stage, all the $2K$ streams of signal of the K antennas are jointly used for information detection.[1] Note that although we focus on the wireless communication application in this paper, the proposed splitting receiver with single-antenna ($K = 1$) is also applicable to cable and fibre-optical communication systems.

---

[1] Although the CD and PD circuits may have different detection sensitivity level in practice [16], we assume both the circuits are able to detect arbitrarily small power signal for tractability.
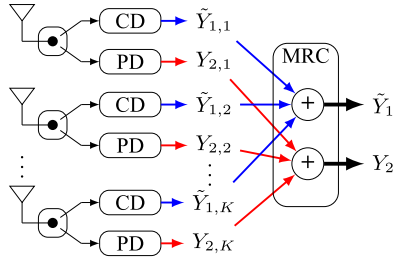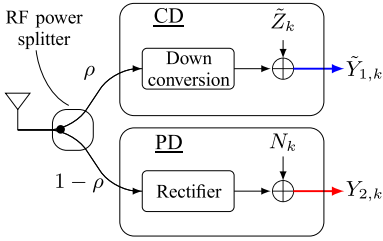
(a) K-antenna splitting receiver architecture.

(b) The kth antenna receiver architecture.

Fig. 1. The proposed splitting receiver architecture.

(a) K-antenna simplified receiver architecture.

(b) The K_1th and the $(K_1 + 1)$th antenna receiver architecture.

Fig. 2. The proposed simplified receiver architecture.

2) Simplified Receiver: We also propose a simplified receiver, as a variant of the splitting receiver, where no power splitters are required. This is illustrated in Fig. 2. In the simplified receiver, K_1 antennas $(1 \leq K_1 < K)$ are connected to the CD circuit and the remaining antennas are connected to the PD circuits. This is illustrated in Fig. 2. We assume that the connections are determined offline, hence do not depend on the instantaneous channel coefficients at each antenna.

B. Signal Model

In this section, we present the signal model for the splitting receiver. Note that the simplified receiver can be analytically treated as a special case of the splitting receiver with power splitting ratios taking binary values only, i.e., $\rho_k \in \{0, 1\}$, for all $k = 1, 2, \ldots, K$.

Based on [12], [15], the output signals from the CD and PD circuits at the kth antenna are given by, respectively,

$$\tilde{Y}_{1,k} = \sqrt{\rho_k \mathcal{P}} \tilde{h}_k \tilde{X} + \tilde{Z}_k, \tag{1}$$

$$Y_{2,k} = \eta(1 - \rho_k)|\tilde{h}_k|^2 \mathcal{P}|\tilde{X}|^2 + N_k', \tag{2}$$

where $\rho_k \in [0, 1]$ is the power splitting ratio. $\tilde{X}$ is the transmitted signal with normalized variance and $\tilde{X} \in \mathcal{X}$, where $\mathcal{X}$ denotes the set of all possible transmitted signals. $\tilde{Z}_k$ is the post-processing complex AWGN of the CD circuit with the mean of zero and the variance of σ_{cov}^2, which includes both the RF band to baseband conversion noise and the ADC noise. N_k' is the post-processing noise of the PD circuit which is also assumed to be real Gaussian noise [12], which includes both the rectifier noise and the ADC noise. Note that we only consider the post-processing noise $\tilde{Z}_k$ and N_k', i.e., we ignore the pre-processing noise, such as the antenna noise which is almost at the thermal noise level and is much smaller than the post-processing noise [15].
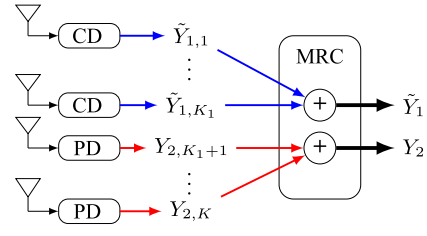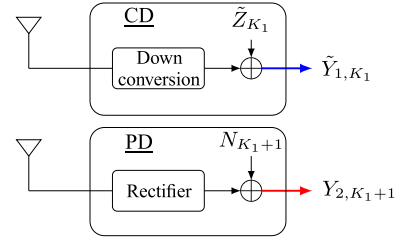
Without loss of generality, scaling (2) by η, the received signal $Y_{2,k}$ can be rewritten as

$$Y_{2,k} = (1 - \rho_k)|\tilde{h}_k|^2 \mathcal{P}|\tilde{X}|^2 + N_k, \tag{3}$$

where $N_k \triangleq N_k'/(\eta)$ is the equivalent rectifier conversion AWGN with the mean of zero and the variance σ_{rec}^2.

C. Maximal Ratio Combining of Splitting Receiver

To detect the transmitted signal $\tilde{X}$, similar with a conventional SIMO receiver, the optimal method is the maximal ratio combining (MRC). We assume that the receiver has perfect channel state information (CSI), i.e., knowledge of $\tilde{h}_k$. Since the K-antenna received signals $\tilde{Y}_{1,k}$ and $Y_{2,k}$, $k = 1, 2, \ldots, K$, lie in different signal spaces, we use MRC for coherently processed signals (i.e., $\tilde{Y}_{1,k}$) and non-coherent signals (i.e., $Y_{2,k}$) separately. Based on (1) and (3), the combined coherently and non-coherently processed signals are given by, respectively,

$$\tilde{Y}_1 = \left(\sum_{k=1}^{K} \rho_k |\tilde{h}_k|^2 \right) \sqrt{\mathcal{P}} \tilde{X} + \sum_{k=1}^{K} \sqrt{\rho_k} \tilde{h}_k^* \tilde{Z}_k,$$

$$\tag{4}$$

$$Y_2 = \left(\sum_{k=1}^{K} (1-\rho_k)^2 |\tilde{h}_k|^4 \right) \mathcal{P}|\tilde{X}|^2 + \sum_{k=1}^{K} (1-\rho_k)|\tilde{h}_k|^2 N_k.$$

For convenience of analysis, after linear scaling, (4) can be rewritten as

$$\tilde{Y}_1 = \sqrt{\Theta_1} \sqrt{\mathcal{P}} \tilde{X} + \tilde{Z}, \quad Y_2 = \sqrt{\Theta_2} \mathcal{P}|\tilde{X}|^2 + N, \tag{5}$$

where

$$\Theta_1 = \sum_{k=1}^{K} \rho_k |\tilde{h}_k|^2, \quad \Theta_2 = \sum_{k=1}^{K} (1-\rho_k)^2 |\tilde{h}_k|^4, \tag{6}$$

and $\tilde{Z}$ and N follow the same distributions as $\tilde{Z}_k$ and N_k, respectively. The two-dimensional signal $\tilde{Y}_1$ and the
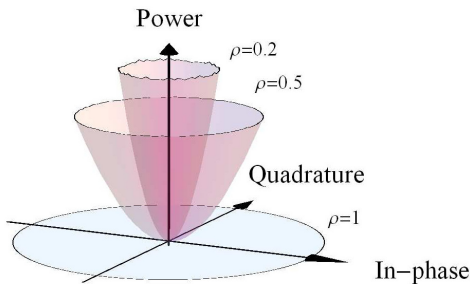
Fig. 3.   Illustration of the signal space of the splitting receiver, $\rho = 0.2, 0.5, 1$.

one-dimensional signal Y_2 form a triple $(\tilde{Y}_1, Y_2)$, which is the equivalent received signal of the K-antenna splitting receiver.

It is interesting to see that since the two-dimensional signal $\tilde{Y}_1$ lies on the *in-phase-quadrature* (I-Q) plane and the one-dimensional signal Y_2 lines on the *power* (P)-axis, the equivalent received signal $(\tilde{Y}_1, Y_2)$ lies in the three-dimensional I-Q-P space. This is different from the conventional coherent (two-dimensional) and non-coherent (one-dimensional) receiver signal spaces. Thus, the splitting receiver expands the received signal space and fundamentally changes the way in which the signal is processed compared with the conventional receivers.

Considering the noiseless signal, i.e., letting $\tilde{Z}$ and $N = 0$ in (5), we have $Y_2 = \frac{\sqrt{\Theta_2}}{\Theta_1}|\tilde{Y}_1|^2$ from (5), which is a paraboloid equation. From a geometric point of view, defining $\vec{\rho} \triangleq [\rho_1, \rho_2, \dots \rho_K]$, $\vec{1} \triangleq \underbrace{[1, 1, \cdots, 1]}_{K}$, and $\vec{0} \triangleq \underbrace{[0, 0, \cdots, 0]}_{K}$, the splitting receiver is actually bending the noiseless received signal space into a paraboloid with $\vec{\rho}$ as illustrated in Fig. 3. When $\vec{\rho} = \vec{1}$, i.e., a non-splitting case, the splitting receiver degrades to the coherent receiver. As the parameter $\sqrt{\Theta_2}/\Theta_1$ increases, e.g., each element of $\vec{\rho}$ decreases, the splitting receiver bends the signal plane to a paraboloid, which is taller and thinner with a larger $\sqrt{\Theta_2}/\Theta_1$, e.g., a smaller $\vec{\rho}$. When $\vec{\rho} = \vec{0}$, the splitting receiver degrades to the PD-based non-coherent receiver.

In this paper, PD-based non-coherent receiver is named as *non-coherent receiver* for short, and we refer to both the K-antenna coherent receiver (i.e., $\vec{\rho} = \vec{1}$) and the K-antenna non-coherent receiver (i.e., $\vec{\rho} = \vec{0}$) as the *conventional receivers*.

D. Splitting Channel

From an information theory perspective, (5) can be rewritten as

$$\begin{bmatrix} \tilde{Y}_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & |\cdot|^2 \end{bmatrix} \begin{bmatrix} \sqrt{\Theta_1} \\ \sqrt[4]{\Theta_2} \end{bmatrix} \sqrt{\mathcal{P}}\tilde{X} + \begin{bmatrix} \tilde{Z} \\ N \end{bmatrix}, \tag{7}$$

where $|\cdot|^2$ is the squared magnitude operator. We name (7) as the *splitting channel*, and the input and output of the splitting channel regarded as random variables, are $\sqrt{\mathcal{P}}\tilde{X}$ and $(\tilde{Y}_1, Y_2)$, respectively.

The splitting channel can be treated as a SIMO channel, since the channel has one input $\sqrt{\mathcal{P}}\tilde{X}$ and two outputs

$\tilde{Y}_1$ and Y_2. It can also be treated as a degraded (due to the power splitting) SISO channel with the output $\tilde{Y}_1$ and a side information Y_2.

E. Performance Metrics and SNR Definition

We study the mutual information between the input and output of the splitting channel for an ideal Gaussian input, and study the SER performance for practical modulation schemes.

For convenience of analysis, we define the operating SNR as

$$\mathsf{SNR} \triangleq \min \{\mathsf{SNR}_{\mathrm{coh}}, \mathsf{SNR}_{\mathrm{non-coh}}\}, \tag{8}$$

where

$$\mathsf{SNR}_{\mathrm{coh}} \triangleq H_2\frac{\mathcal{P}}{\sigma_{\mathrm{cov}}^2}, \quad \mathsf{SNR}_{\mathrm{non-coh}} \triangleq \sqrt{H_4}\frac{\mathcal{P}}{\sigma_{\mathrm{rec}}}, \tag{9}$$

$$H_2 \triangleq \sum_{k=1}^{K} |\tilde{h}_k|^2, \quad H_4 \triangleq \sum_{k=1}^{K} |\tilde{h}_k|^4.$$

$\mathsf{SNR}_{\mathrm{coh}}$ and $\mathsf{SNR}_{\mathrm{non-coh}}$ are the SNRs of the conventional receivers, i.e., $\vec{\rho} = \vec{1}$ for the coherent receiver and $\vec{\rho} = \vec{0}$ for the non-coherent receiver, respectively. Specifically, the definition of $\mathsf{SNR}_{\mathrm{non-coh}}$ is consistent with [12]. Note that although $\sqrt{H_4}\mathcal{P}$ and σ_{rec} correspond to the standard deviation (not variance) of the signal $\sqrt{H_4}\mathcal{P}|\tilde{X}|^2$ and the noise N at the PD receiver, respectively, $\sqrt{H_4}\mathcal{P}$ still has the physical meaning of "power". Thus, the signal-to-noise ratio is defined as $\sqrt{H_4}\frac{\mathcal{P}}{\sigma_{\mathrm{rec}}}$ not $H_4\frac{\mathcal{P}^2}{\sigma_{\mathrm{rec}}^2}$.

In the following, we refer to the high SNR regime as $\mathsf{SNR} \to \infty$ which is obtained by letting $\mathcal{P} \to \infty$. Our analysis will focus on the splitting receiver which includes the simplified receiver as a special case.

III. SPLITTING CHANNEL: MUTUAL INFORMATION

In this section, we study the mutual information of the splitting channel to determine the gain due to the joint coherent and non-coherent processing. We also provide a discussion to intuitively explain this processing gain.

Based on (7), the mutual information between the input and outputs of the splitting channel with the splitting ratio $\vec{\rho}$ is

$$\begin{aligned}
I&\left(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2\right) \\
&= h(\tilde{Y}_1, Y_2) - h(\tilde{Y}_1, Y_2|\sqrt{\mathcal{P}}\tilde{X}) \\
&= h(\tilde{Y}_1, Y_2) - h(\tilde{Z}, N|\sqrt{\mathcal{P}}\tilde{X}) \\
&= h(\tilde{Y}_1, Y_2) - h(\tilde{Z}, N) \\
&= h(\tilde{Y}_1, Y_2) - h(\tilde{Z}) - h(N) \\
&= -\int_{Y_2}\int_{\tilde{Y}_1} f_{\tilde{Y}_1, Y_2}(\tilde{y}_1, y_2) \log_2\left(f_{\tilde{Y}_1, Y_2}(\tilde{y}_1, y_2)\right) \mathrm{d}\tilde{y}_1 \mathrm{d}y_2 \\
&\quad - \log_2(\pi e\sigma_{\mathrm{cov}}^2) - \frac{1}{2}\log_2(2\pi e\sigma_{\mathrm{rec}}^2). \tag{10}
\end{aligned}$$

The joint probability density function (pdf) of $(\tilde{Y}_1, Y_2)$ is

$$\begin{aligned}
f_{\tilde{Y}_1, Y_2}&(\tilde{y}_1, y_2) \\
&= \int_{\tilde{X}} f_1\left(\sqrt{\Theta_1 \mathcal{P}}\tilde{x}, \tilde{y}_1\right) f_2\left(\sqrt{\Theta_2}\mathcal{P}|\tilde{x}|^2, y_2\right) f_{\tilde{X}}(\tilde{x})\mathrm{d}\tilde{x}, \tag{11}
\end{aligned}$$

where $f_{\tilde{X}}(\tilde{x})$ is the pdf of $\tilde{X}$, and $f_1\left(\sqrt{\Theta_1 \mathcal{P}}\tilde{x}, \cdot\right)$ and $f_2\left(\sqrt{\Theta_2}\mathcal{P}|\tilde{x}|^2, \cdot\right)$ are the pdfs of the distributions $\mathcal{CN}\left(\sqrt{\Theta_1 \mathcal{P}}\tilde{x}, \sigma_{\text{cov}}^2\right)$ and $\mathcal{N}\left(\sqrt{\Theta_2}\mathcal{P}|\tilde{x}|^2, \sigma_{\text{rec}}^2\right)$, respectively.

The mutual information expression in (10) needs five integrals to evaluate, which is cumbersome and thus the maximal mutual information theoretically achieved by the optimal distribution of $\tilde{X}$, cannot be obtained.

A. Mutual Information and Joint Processing Gain

For tractability, in the following analysis, we consider the mutual information with a Gaussian input, i.e., $\tilde{X} \sim \mathcal{CN}(0, 1)$, and we have:

1) Letting $\vec{\rho} = \vec{1}$, the splitting channel is degraded to the coherent AWGN channel, and the mutual information is well-known as [17]

$$I\left(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2\right) = h(\tilde{Y}_1) - h(\tilde{Z}) = \log_2\left(1 + H_2 \frac{\mathcal{P}}{\sigma_{\text{cov}}^2}\right),$$
(12)

which is exactly the capacity of the coherent AWGN channel, i.e., $\mathcal{C}(\vec{\rho} = \vec{1})$.

2) Letting $\vec{\rho} = 0$, the splitting channel is degraded to the conventional intensity channel in free-space optical communications [12]. Recall that we refer to the intensity channel as the non-coherent AWGN channel, echoing the coherent AWGN channel in this paper.[2] The mutual information of the non-coherent AWGN channel is [12]

$$I\left(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2\right)$$
(13)

$$= h(Y_2) - h(N)$$
(14)

$$= -\int_{-\infty}^{\infty} f_{Y_2}(y_2)\log_2(f_{Y_2}(y_2))\mathrm{d}y_2 - \frac{1}{2}\log_2(2\pi e \sigma_{\text{rec}}^2)$$

$$\overset{(a)}{\geq} \frac{1}{2}\log_2\left(1 + H_4\frac{\mathcal{P}^2 e}{2\pi \sigma_{\text{rec}}^2}\right),$$
(15)

where $Y_2 = \sqrt{H_4}\mathcal{P}|\tilde{X}|^2 + N$ follows an exponential modified Gaussian distribution [18]:

$$f_{Y_2}(y_2) = \frac{1}{2\sqrt{H_4}\mathcal{P}}\exp\left(\frac{1}{2\sqrt{H_4}\mathcal{P}}\left(\frac{\sigma_{\text{rec}}^2}{\sqrt{H_4}\mathcal{P}} - 2y_2\right)\right)$$

$$\cdot \operatorname{erfc}\left(\frac{\frac{\sigma_{\text{rec}}^2}{\sqrt{H_4}\mathcal{P}} - y_2}{\sqrt{2}\sigma_{\text{rec}}}\right).$$
(16)

The inequality (a) is given in [12], and (15) is the asymptotic mutual information in the high SNR regime, which is also the asymptotic capacity (with gap less than $\left|\frac{1}{2}\log_2\left(\frac{e}{2\pi}\right)\right|$ bits) of the non-coherent AWGN channel, i.e., $\mathcal{C}(\vec{\rho} = 0)$.

Comparing (12) and (15), it is easy to see that as SNR $\rightarrow \infty$, *the coherent and non-coherent AWGN channels have the same asymptotic capacity*, i.e., $\lim_{\text{SNR}\to\infty} \mathcal{C}(\vec{\rho} = \vec{1})/\mathcal{C}(\vec{\rho} = 0) = 1$. In the following, we will show that

[2]Note that in this paper, the non-coherent channel refers to the intensity channel, and it does not refer to the kind of channel without CSI at the transmitter or the receiver.
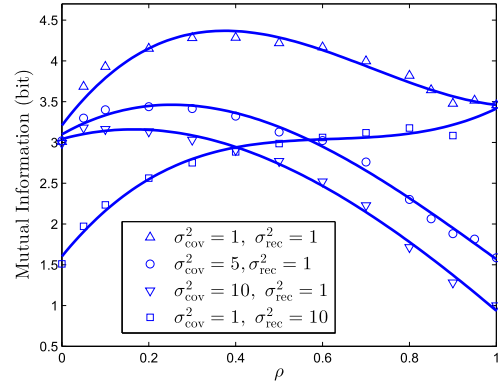

Fig. 4. $I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2)$ versus ρ, $\mathcal{P} = 10$, $K = 1$, $|\tilde{h}_1| = 1$. The simulation results are marked with 'o's, and are curve fitted by polynomials of degree of 3.

the splitting receiver with $\vec{\rho} \neq 0$ nor $\vec{1}$ provides a gain in the mutual information compared with the conventional receivers. Firstly, we need the following definition.

Definition 1: The joint processing gain of the splitting receiver is

$$G \triangleq \frac{\sup\{I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2) : \vec{\rho} \in [0, 1]^K\}}{\max\{I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2)|_{\vec{\rho}=0}, I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2)|_{\vec{\rho}=\vec{1}}\}}, (17)$$

where $\sup\{\cdot\}$ denotes for the supremum, and $[0, 1]^K$ is the K-product space generated by the interval $[0, 1]$.

If the joint processing gain $G > 1$, the splitting receiver achieves higher mutual information compared with the conventional receivers. If the joint processing gain $G = 1$ which means the joint coherent and non-coherent processing is unnecessary, the splitting receiver should be degraded to either one of the conventional receivers.

Due to the complicated form of (10), it is not possible to accurately evaluate the mutual information[3] for $\vec{\rho} \in [0, 1]^K$ and prove whether G is greater than 1 or not. Hence, we first use Monte Carlo based histogram method to simulate the results. In Fig. 4, considering the $K = 1$ case, it is observed that when SNR is reasonably high, e.g., $\mathcal{P} = 10$, $\sigma_{\text{cov}}^2 = 1$ and $\sigma_{\text{rec}} = 1$, the joint processing gain G is greater than 1. Inspired by this, we will focus on the analysis on the mutual information in (10) and the joint processing gain in Definition 1 in the high SNR regime in the following subsection.

B. High SNR Analysis

Lemma 1: In the high SNR regime, $I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2)$ with $\vec{\rho} \in [0, 1]^K \setminus \{\vec{0}, \vec{1}\}$ is given by

$$I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2)$$

$$\approx \log_2\left(\frac{\Theta_1 \mathcal{P}}{\sigma_{\text{cov}}^2}\right) + \frac{1}{2\log(2)}\exp\left(\frac{\Theta_1 \sigma_{\text{rec}}^2}{\Theta_2 2\sigma_{\text{cov}}^2 \mathcal{P}}\right)\operatorname{Ei}\left(\frac{\Theta_1 \sigma_{\text{rec}}^2}{\Theta_2 2\sigma_{\text{cov}}^2 \mathcal{P}}\right)$$
(18a)

$$\approx \log_2\left(\frac{\sqrt{2}\mathcal{P}^{\frac{3}{2}}\sqrt{\Theta_1 \Theta_2}}{\sigma_{\text{cov}}\sigma_{\text{rec}}}\right) - \frac{\gamma}{2\ln 2},$$
(18b)

[3]A lower bound and an upper bound of $I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2)$ with explicit expressions can be found based on the basic inequalities $I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2) > I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1)$, $I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2) > I(\sqrt{\mathcal{P}}\tilde{X}; Y_2)$ and $I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2) < I(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1) + I(\sqrt{\mathcal{P}}\tilde{X}; Y_2)$ [17]. Since the bounds are loose, we do not pursue them here.

where $\mathrm{Ei}(x) \triangleq \int_x^\infty \frac{e^{-t}}{t} dt$ is the exponential integral function, and γ is Euler's constant.

Proof: See Appendix. ∎

From Lemma 1, it is clear that the mutual information of the splitting channel increases linearly with $\log_2(\mathcal{P})$ and decreases linearly with $\log_2(\sigma_{\mathrm{cov}})$ and $\log_2(\sigma_{\mathrm{rec}})$ in the high SNR regime. Moreover, since the mutual information depends on the power splitting ratio $\vec{\rho}$, which is contained in the term $\Theta_1\Theta_2$, it is interesting to find the optimal $\vec{\rho}$ that maximizes the mutual information.

Based on Lemma 1, the following optimization problem is proposed to obtain the optimal splitting ratio $\vec{\rho}$ in the high SNR regime:

$$(\text{P1}) \quad \max_{\vec{\rho}\in[0,1]^K\setminus\{\vec{0},\vec{1}\}} \Theta_1\Theta_2$$

$$\Leftrightarrow \max_{\vec{\rho}\in[0,1]^K\setminus\{\vec{0},\vec{1}\}} \sum_{k=1}^K \rho_k|\tilde{h}_k|^2 \sum_{k=1}^K (1-\rho_k)^2 |\tilde{h}_k|^4. \quad (19)$$

It can be shown that (P1) is not a convex optimization problem. Thus, the optimal splitting ratio can be obtained by numerical methods. In what follows, we first focus on two special scenarios and then discuss the joint processing gain for a general splitting receiver.

1) Splitting Receiver With Single Receiver Antenna: When $K = 1$, $\Theta_1 = \rho_1|\tilde{h}_1|^2$ and $\Theta_2 = (1 - \rho_1)^2|\tilde{h}_1|^4$, and thus, the optimal power-splitting ratio is obtained by solving the equation $\frac{\partial \rho^{\frac{1}{3}}(1-\rho)^{\frac{2}{3}}}{\partial \rho} = 0$. It is straightforward to obtain the following results.

Proposition 1: For the splitting receiver with single receiver antenna, the optimal splitting ratio in the high SNR regime is

$$\rho^\star = \frac{1}{3}, \quad (20)$$

and the maximal mutual information is given by

$$I\left(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2\right)|_{\rho^\star} \approx \log_2\left(\frac{2\sqrt{2}}{3\sqrt{3}} \frac{|\tilde{h}_1|^3 \mathcal{P}^{\frac{3}{2}}}{\sigma_{\mathrm{cov}}\sigma_{\mathrm{rec}}}\right) - \frac{\gamma}{2\ln 2}. \quad (21)$$

2) Simplified Receiver With a Large Number of Antennas: For the simplified receiver with a large number of antennas, (19) can be rewritten as

$$\max_{1\le K_1 < K} \sum_{k=1}^{K_1} |\tilde{h}_k|^2 \sum_{k=K_1+1}^K |\tilde{h}_k|^4. \quad (22)$$

Assuming that $\tilde{h}_k$, $k = 1, 2, \ldots K$, are independent and identically distributed (i.i.d.) random variables, i.e., *the uncorrelated scenario*, due to the law of large numbers when K is sufficiently large, we have

$$\lim_{K\to\infty} \sum_{k=1}^{K_1} |\tilde{h}_k|^2 \sum_{k=K_1+1}^K |\tilde{h}_k|^4$$

$$= \lim_{K\to\infty} K_1 K_2 \sum_{k=1}^{K_1} |\tilde{h}_k|^2 / K_1 \sum_{k=K_1+1}^K |\tilde{h}_k|^4 / K_2$$

$$\overset{(a)}{=} K_1 K_2 \mathbb{E}\left(|\tilde{h}_k|^2\right) \mathbb{E}\left(|\tilde{h}_k|^4\right), \quad (23)$$

where $K_2 \triangleq K - K_1$, and (a) is because both K_1 and K_2 are sufficiently large. Assuming that $|\tilde{h}_k|$, $k = 1, 2, \ldots K$, are identical with each other, i.e., *the free-space scenario which is also a fully-spatially-correlated scenario*, we have the same expression with (23). Thus, $K_1 K_2$ is maximized when $K_1 = K_2 = K/2$, and we have the following proposition.

Proposition 2: For the simplified receiver with a large number of antennas, the optimal strategy for the spatially-uncorrelated channel or the fully-spatially-correlated channel (i.e., the free-space scenario) in the high SNR regime is to connect half of the antennas to the CD circuits and the other half to the PD circuits, and the maximum mutual information is given by

$$I\left(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2\right)|_{\vec{\rho}^\star}$$

$$\approx \log_2\left(\frac{K\mathcal{P}^{\frac{3}{2}}\sqrt{\mathbb{E}\left(|\tilde{h}_k|^2\right)\mathbb{E}\left(|\tilde{h}_k|^4\right)}}{\sqrt{2}\sigma_{\mathrm{cov}}\sigma_{\mathrm{rec}}}\right) - \frac{\gamma}{2\ln 2}. \quad (24)$$

Note that for the general spatially-correlated scenario, the optimal strategy in the high SNR regime is not immediately clear. We will investigate this scenario for future study.

3) Joint Processing Gain of Splitting Receiver With K Receiver Antennas: We assume that $\vec{\rho} \neq \vec{0}$ nor $\vec{1}$, thus, $\Theta_1 \neq 0$ and $\Theta_2 \neq 0$. Then, based on (18b) of Lemma 1, (12) and (15), we can show that

$$\lim_{\mathrm{SNR}\to\infty} \frac{I\left(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2\right)|_{\vec{\rho}\in[0,1]^K\setminus\{\vec{0},\vec{1}\}}}{\max\{I\left(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2\right)|_{\vec{\rho}=\vec{0}}, I\left(\sqrt{\mathcal{P}}\tilde{X}; \tilde{Y}_1, Y_2\right)|_{\vec{\rho}=\vec{1}}\}}$$

$$= \frac{3}{2}. \quad (25)$$

In other words, the asymptotic gain is the same no matter what value $\vec{\rho}$ takes, as long as $\vec{\rho} \neq \vec{0}$ nor $\vec{1}$. Therefore, the asymptotic optimal splitting ratio $\vec{\rho}^\star \in [0, 1]^K\setminus\{\vec{0}, \vec{1}\}$, and we have the following result based on Definition 1.

Proposition 3: In the high SNR regime, the asymptotic joint processing gain for a splitting receiver with K receiver antennas is

$$\lim_{\mathrm{SNR}\to\infty} G = \frac{3}{2}. \quad (26)$$

C. Explanation of the Joint Processing Gain

The result of Proposition 3 shows that in the high SNR regime, since $G = 3/2 > 1$, the splitting receiver provides a processing gain. Note that although the joint processing gain at any given SNR depends on the specific value of the received signal power $\mathcal{P}$, the noise variances σ_{cov}^2 and σ_{rec}^2, the asymptotic joint processing gain is independent of the specific noise variances at the CD and PD circuits in the high SNR regime. *This implies that the reason for the performance improvement lies in the joint coherent and non-coherent processing.* This is explained in detail using intuitive and geometric arguments as follows.

1) Intuitive Explanation of the Rate Improvement: Since the degrees of freedom of a channel is commonly defined as the dimension of the received signal space [1], the coherent AWGN channel has two degrees of freedom (I-Q plane) while the non-coherent AWGN channel has one degree of freedom (P-axis). For the splitting channel created by jointly utilizing both the coherent and non-coherent AWGN channels, the received signals are spread into a three-dimensional space, i.e., the I-Q-P space. Thus, the splitting channel can be treated as a channel with three degrees of freedom. Therefore, the splitting channel with a properly designed splitting ratio can take better advantage of the I-Q-P space, and achieve a better channel rate performance compared with either the coherent or non-coherent AWGN channel.

We would like to highlight that a 'splitting receiver', which splits the received signal at each antenna into two streams and sends both streams to CD circuits (i.e., two coherent AWGN channels), does not provide any rate improvement. After MRC, it is straightforward to see that the received signal space still lies on the I-Q plane. Thus, the received signal space is the same as for the conventional coherent receiver. For instance, consider a single-antenna receiver for ease of illustration. It can be proved that the best 'splitting' strategy is to send the entire signal to the CD circuit with a smaller noise variance, instead of splitting and sending signals to both CD circuits [1]. Therefore, there is no joint processing gain by using two coherent AWGN channels, i.e., $G = 1$. The same argument holds for a 'splitting receiver' which splits the received signal at each antenna into two streams and sends them to two PD circuits (i.e., two non-coherent AWGN channels).

Therefore, the key to the rate improvement is the increased dimension of the received signal space achieved by joint coherent and non-coherent processing, where the coherent channel adds noise linearly to the signal, and the noncoherent channel adds noise to the square amplitude of the signal.

2) A Geometric Explanation of the Asymptotic Gain: As discussed in Sec. II.C, a splitting receiver with the splitting ratio $\vec{\rho}$ maps the noiseless received signal space, i.e., the I-Q plane, to a paraboloid in the I-Q-P space with parameter $\sqrt{\Theta_2}/\Theta_1$ which depends on $\vec{\rho}$. Considering a disk with radius R and center $(0, 0)$ in the I-Q plane, the area of the disk is πR^2, where R is proportional to $\sqrt{P}$ in this paper. After the mapping, the disk is converted into a paraboloid with parameter $\sqrt{\Theta_2}/\Theta_1$ which is restricted by the condition that the projection of the paraboloid in the I-Q plane should lie within the disk with radius $\sqrt{\Theta_1}R$. When R is sufficiently large, the area of the paraboloid can be shown to be approximated by $3\pi\sqrt{\Theta_1\Theta_2}R^3$ for $\vec{\rho} \neq \vec{0}$ nor $\vec{1}$. It is well known that the optimal constellation design for the I-Q space is equivalent to a sphere-packing problem, i.e., packing two-dimensional spheres (disks) with a certain radius, which is related to the detection error rate, on the surface of the disk (i.e., the disk on the I-Q plane). The number of spheres that can be packed is proportional to the area of the disk. Thus, the communication rate can be written as $O\left(\log(\pi R^2)\right) \sim 2O(\log R)$. Similarly, for the paraboloid,

the number of three-dimensional spheres[4] (balls) that can be packed on the surface is proportional to the paraboloid area, and the rate can be written as $O\left(\log(3\pi\sqrt{\Theta_1\Theta_2}R^3)\right) \sim 3O(\log R)$. Therefore, it is straightforward to see that there is a $3/2$ fold rate gain provided by the splitting receiver when R is sufficiently large. To sum up, bending the signal space from a two-dimensional plane to a three-dimensional paraboloid increases the effective area of the signal space which boosts the communication rate.

3) The Complexity of Splitting Receiver: Although the splitting receiver is able to provide a performance gain, it is clear that for the information detection in the digital domain, the splitting receiver requires a three-dimensional detection, while the conventional CD/PD receiver only needs a two/one-dimensional detection, respectively. Specifically, when applying the minimum distance detection for practical modulation, the splitting receiver needs to calculate the distance between two signal points in the three-dimensional space, while the conventional CD/PD receiver only needs to calculate the distance in the two/one-dimensional space. Thus, the splitting receiver requires a higher computation complexity to achieve the performance gain. Regarding the circuit complexity, for each antenna branch, the splitting receiver requires two detection circuits, while the conventional CD/PD receiver only needs one detection circuit. On the other hand, the proposed simplified receiver has a lower complexity than the CD receiver and a higher complexity than the PD receiver. Therefore, we should consider both the performance gain and the complexity (and the cost) when adopting a splitting receiver in practical systems.

D. Numerical Results

In the last two subsections, we have shown and explained that the splitting receiver achieves a $3/2$ fold rate gain compared with the non-splitting channels in the high SNR regime. This suggests that a notable performance improvement can be found within a moderate SNR range, which is verified as follows. Also, we verify the tightness of the asymptotic analytical results presented in Sec. III.B.

1) Single-Antenna Scenario: We set the channel power gain $|\tilde{h}_1|^2 = 1$ for simplicity.

Fig. 5 depicts the mutual information approximation given in (18a) and also the simulated mutual information with different received signal power. We see that the approximation and simulation results have the same general trend, and the percentage difference between the approximation and simulation results decreases as P increase (e.g., from $P = 10$ to 100). Also we see that the optimal splitting ratios are (almost) the same for both the approximation and simulation results. When SNR is sufficiently large, e.g., 20 dB, $\rho = 0.33$ makes the mutual information at least 20% larger than that of the conventional cases (i.e., $\rho = 0$ or 1), and the joint processing gain is shown to be $G \approx 1.3$. When SNR = 30 dB (i.e., $P = 1000$), the approximation is tight, and the joint processing

---

[4]Note that sphere-packing is considered only if $\sigma_{\text{cov}}^2 = 2\sigma_{\text{rec}}^2$, i.e., a uniform three-dimensional noise sphere, otherwise, it is ellipsoid-packing. Here we use sphere-packing for ease of illustration.
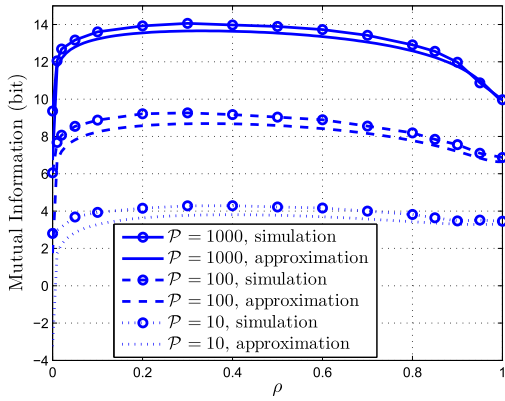
Fig. 5. Mutual information versus ρ, $\sigma_{\text{cov}}^2 = \sigma_{\text{rec}}^2 = 1$.
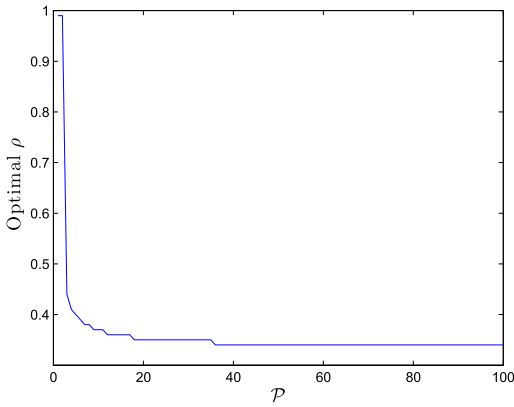

Fig. 6. Optimal ρ versus $\mathcal{P}$, $\sigma_{\text{cov}}^2 = \sigma_{\text{rec}}^2 = 1$.
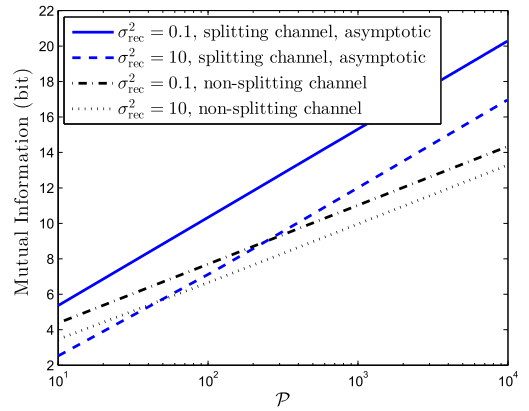

Fig. 7. Mutual information of the splitting channel and the non-splitting channel versus $\mathcal{P}$, $\sigma_{\text{cov}}^2 = 1$.
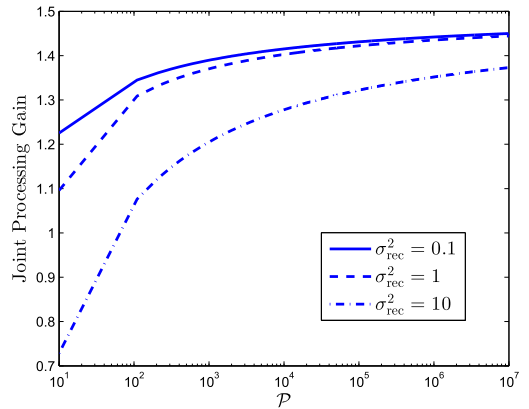

Fig. 8. Joint processing gain versus $\mathcal{P}$, $\sigma_{\text{cov}}^2 = 1$.

gain with $\rho = 0.33$ is close to 1.5. Thus, the tightness of the mutual information expressions in Lemma 1 and Proposition 1 (which is obtained by taking $\rho = 1/3$ into Lemma 1) and also the asymptotic joint processing gain given by Proposition 3 are verified.

Fig. 6 depicts the optimal splitting ratio ρ (obtained by simulation) versus the received signal power $\mathcal{P}$. It is observed that the optimal splitting ratio approaches $1/3$ quickly as $\mathcal{P}$ increases, i.e., $\rho = 1/3$ when $\mathcal{P} > 35$. *Thus, $\rho = 1/3$ is a near-optimal choice even at moderate SNRs, and the tightness of the asymptotic optimal splitting ratio in Proposition 1 is verified.*

Fig. 7 depicts the approximation of the splitting channel mutual information given in (18a) with $\rho = 1/3$, and the optimal non-splitting channel mutual information, i.e., $\max\{(12), (15)\}$. It is observed that the splitting channel mutual information increases much faster w.r.t. $\mathcal{P}$ as compared with the coherent and non-coherent AWGN channels. When SNR > 20 dB (e.g., $\mathcal{P} > 100$, $\sigma_{\text{cov}}^2 = 1$ and $\sigma_{\text{rec}}^2 = 0.1$, or $\mathcal{P} > 1000$, $\sigma_{\text{cov}}^2 = 1$ and $\sigma_{\text{rec}}^2 = 10$), one can clearly see the mutual information improvement due to splitting.

Fig. 8 depicts the joint processing gain obtained by taking (18a) into Definition 1. It is observed that the joint processing gain increases with $\mathcal{P}$ and slowly approaches the constant $3/2$. The gain at a practically high SNR, e.g., 30 dB, is notable, e.g., $1.2 \sim 1.4$.

2) Multi-Antenna Scenario: Fig. 9 depicts the average mutual information over 10^3 channel realizations using (18b), where the channel power gain $|\tilde{h}_k|^2$ is assumed to follow an exponential distribution with the mean of 1, and $\mathcal{P} = 100$. Three splitting strategies are considered: (i) the numerically searched optimal splitting ratios by solving (P1) for every channel realization (i.e., optimal splitting), (ii) the simplified receiver with the strategy in Proposition 2, and (iii) $\rho_k = 1/3$ for all $k = 1, 2, \ldots, K$. It is observed that the splitting receiver with the optimal splitting strategy is better than the simplified receiver. On the other hand, the splitting receiver can perform worse than the simplified receiver if some sub-optimal splitting strategy is used, e.g., $\rho = 1/3$, which is the optimal for a single-antenna receiver, but generally not necessarily optimal for a multi-antenna receiver.

Fig. 10 depicts the optimal ratio of antennas allocated for coherent processing for a simplified receiver obtained by simulation using 10^4 random channel realizations. *It shows that the optimal ratio is within the range of $(0.45, 0.55)$ when $K > 40$, and the optimal ratio converges to $1/2$ as K increases further, which verifies Proposition 2.*

IV. SPLITTING RECEIVER: PRACTICAL MODULATION

In this section, we consider commonly used modulation schemes and assume that each signal of the constellation
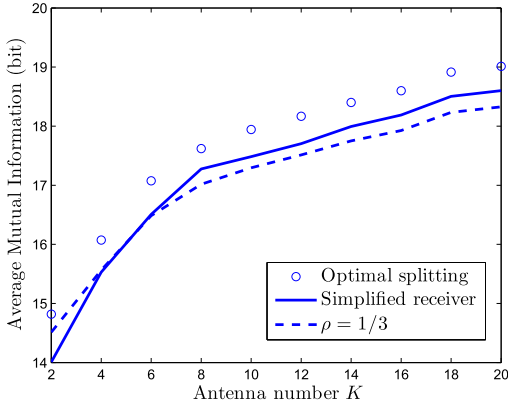
Fig. 9. Mutual information of the splitting channel with different splitting strategies, $\sigma_{\text{cov}}^2 = \sigma_{\text{rec}}^2 = 1$.
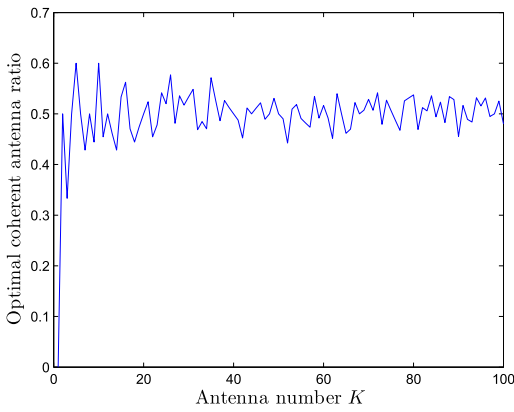

Fig. 10. Average optimal ratio of antennas allocated for coherent processing versus K.

is transmitted with the same probability. Note that, in this section, x and y denote the in-phase and quadrature signals in the CD circuit, respectively, and z denotes the signal in the PD circuit, which are different from the notation in Sec. III. This change of notation is adopted for ease of presentation of results.

A. Transmitted Signal Constellation

We consider the transmitted signal constellation of a general M-ary modulation scheme is Ω_{gen}, which is a two-dimensional constellation placed on the I-Q plane. The ith symbol is denoted by the tuple (x_i, y_i) on the I-Q plane, $i = 1, 2, \ldots, M$. Specifically:

(i) For the M-PAM scheme [1], which is a one-dimensional modulation scheme on the I-axis, we have $x_i = 2i-1$ for $i = 1, 2, \ldots, M/2$, and $x_i = -x_{i-M/2}$ for $i = M/2 + 1, \ldots M$, and $y_i = 0$ for all i.

(ii) For the M-QAM scheme, which is a two-dimensional modulation scheme on the I-Q plane, we have $x_i = 2\left(i \mod \frac{\sqrt{M}}{2}\right) - 1$, $y_i = 2\left\lfloor \frac{i-1}{\sqrt{M}/2} \right\rfloor - 1$, $i = 1, 2, \ldots, M/4$, which are the first quadrant symbols on the I-Q plane. Due to the symmetry property of M-QAM, the other quadrant symbol expressions are omitted for brevity.

(iii) For the M-IM scheme, which is a one-dimensional modulation scheme on the positive I-axis where the information is carried by the signal power but not phase, we have $x_i = \sqrt{2(i-1)}$ and $y_i = 0$, $i = 1, 2, \ldots, M$.

B. Noiseless Received Signal Constellation

Based on the received signal expression after MRC in (5), the average signal power of the coherently and non-coherently processed signals are $\Theta_1 \mathcal{P}$ and $\sqrt{\Theta_2} \mathcal{P}$, respectively. Thus, with such average power constraints, we define the noiseless received signal constellation $\check{\Omega}_{\text{gen}}$, and the ith symbol in $\check{\Omega}_{\text{gen}}$ is denoted by the tuple $(\check{x}_i, \check{y}_i, \check{z}_i)$, and $\check{x}_i = k_1 \sqrt{\Theta_1 \mathcal{P}} x_i$, $\check{y}_i = k_1 \sqrt{\Theta_1 \mathcal{P}} y_i$, $\check{z}_i = k_2 \sqrt{\Theta_2} \mathcal{P} (x_i^2 + y_i^2)$. Here k_1 and $k_2 \triangleq k_1^2$ are the power normalization parameters determined only by the geometric property of a certain modulation scheme. Specifically, we have

$$
k_1 = \begin{cases}
\sqrt{\dfrac{3}{M^2-1}}, & M\text{-PAM}, \\
\sqrt{\dfrac{3}{2(M-1)}}, & M\text{-QAM}, \\
\sqrt{\dfrac{1}{M-1}}, & M\text{-IM}.
\end{cases} \tag{27}
$$

For the single-antenna scenario, Figs. 11(a) and 11(c) show that the splitting ratio $\rho \in (0, 1)$ bends the received signal constellation from the I-axis to a paraboloid in the I-P plane, for 4-PAM and 4-IM, respectively. Fig. 11(b) shows that the splitting ratio $\rho \in (0, 1)$ bends the received signal constellation from the I-Q plane to a paraboloid in the I-Q-P space.

C. Decision Region

Since all the transmitted symbols are of equal probability, the optimal signal detection method is the maximum likelihood (ML) method [1]. The decision region for the ith symbol is defined as

$$
\mathcal{V}_i \triangleq \left\{ \boldsymbol{v} \mid f(\boldsymbol{v}|i) \geq f(\boldsymbol{v}|j), \forall j \neq i, \boldsymbol{v} \in \mathbb{R}^3 \right\}, \tag{28}
$$

where $\boldsymbol{v} \triangleq (x, y, z)$ is the three-dimensional post-processing (noise added) signal in the splitting receiver, and $f(\cdot|\cdot)$ is the conditional pdf.

From Sec. III, since both the CD and PD circuits introduce additive Gaussian noise, the received signal is surrounded by the noise sphere (ellipsoid) in the I-Q-P space, and $f(\boldsymbol{v}|i)$ is thus given by

$$
f(\boldsymbol{v}|i) = \frac{1}{\sigma_{\text{cov}}^2 \pi \sqrt{2\sigma_{\text{rec}}^2 \pi}}
$$
$$
\cdot \exp\left(-\frac{(x-x_i)^2}{\sigma_{\text{cov}}^2} - \frac{(y-y_i)^2}{\sigma_{\text{cov}}^2} - \frac{(z-z_i)^2}{2\sigma_{\text{rec}}^2} \right). \tag{29}
$$

Therefore, (28) is rewritten as

$$
\mathcal{V}_i \triangleq \left\{ (x, y, z) : d_i(x, y, z) \leq d_j(x, y, z), \forall j \neq i \right\}, \tag{30}
$$

where

$$
d_j(x, y, z) \triangleq \frac{(x-x_j)^2}{\sigma_{\text{cov}}^2/2} + \frac{(y-y_j)^2}{\sigma_{\text{cov}}^2/2} + \frac{(z-z_j)^2}{\sigma_{\text{rec}}^2}. \tag{31}
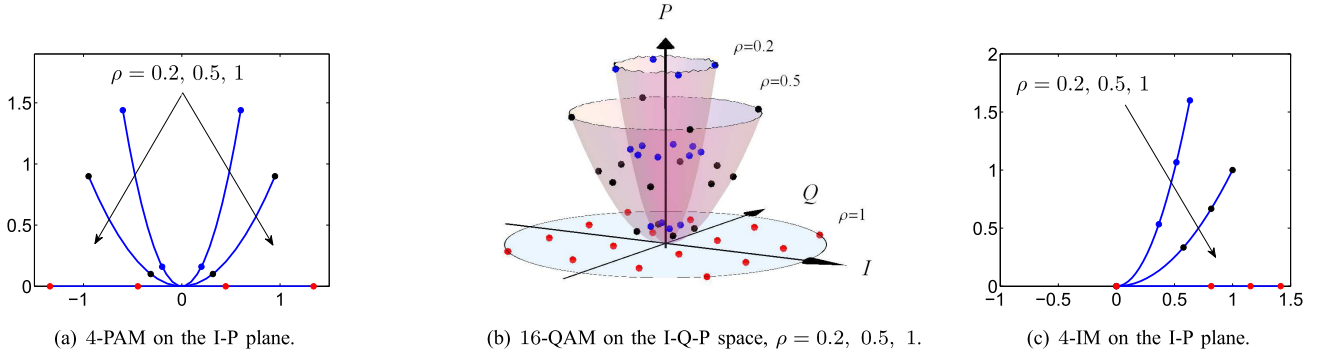$$

(a) 4-PAM on the I-P plane.

(b) 16-QAM on the I-Q-P space, $\rho = 0.2,\ 0.5,\ 1$.

(c) 4-IM on the I-P plane.

Fig. 11. Noiseless received signal constellation with different $\vec{\rho}$, $K = 1$, $|\tilde{h}_1| = 1$, $\mathscr{P} = 10$.

(a) 8-PAM decision regions in the I-P plane.

(b) Illustration of the decision regions for 36-QAM in the first quadrant, i.e., $x > 0$ and $y > 0$.
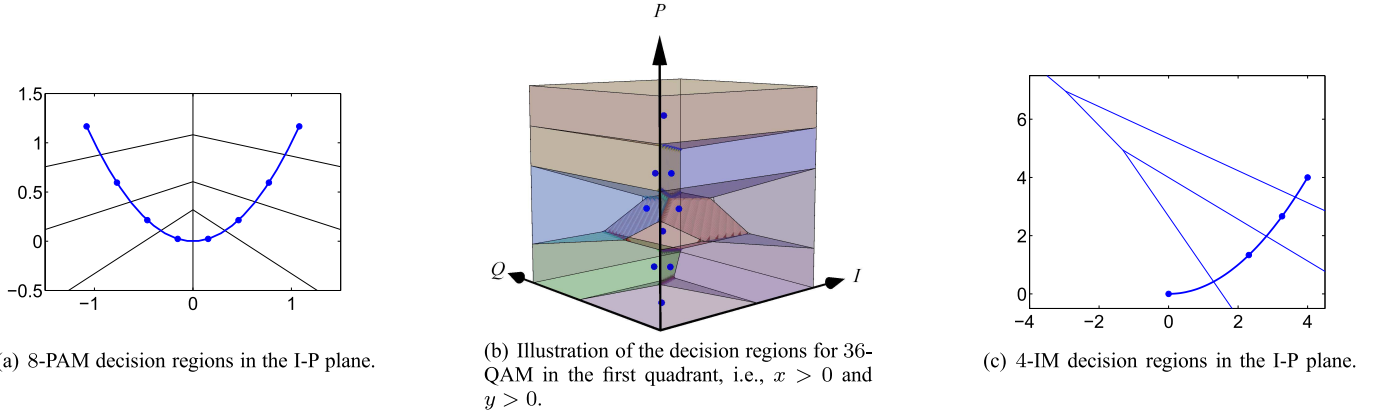
(c) 4-IM decision regions in the I-P plane.

Fig. 12. Decision regions for 8-PAM, 36-QAM and 4-IM, $\mathscr{P} = 10$, $\sigma_{\text{cov}}^2 = 2$, $\sigma_{\text{rec}}^2 = 1$.

From (30) and (31), after simplification, the decision region of the ith symbol, $\mathcal{V}_i$ is given by

$$\mathcal{V}_i = \left\{ (x, y, z) : \ \frac{\check{x}_j - \check{x}_i}{\sigma_{\text{cov}}^2} x + \frac{\check{y}_j - \check{y}_i}{\sigma_{\text{cov}}^2} y + \frac{\check{z}_j - \check{z}_i}{2\sigma_{\text{rec}}^2} z \right.$$
$$\left. \leq \frac{\check{x}_j^2 + \check{y}_j^2 - \check{x}_i^2 - \check{y}_i^2}{2\sigma_{\text{cov}}^2} + \frac{\check{z}_j^2 - \check{z}_i^2}{4\sigma_{\text{rec}}^2}, \forall j \neq i \right\}, \quad (32)$$

where $\check{x}_i$, $\check{y}_i$ and $\check{z}_i$ are defined in Sec. IV.B above (27). It is easy to see that $\mathcal{V}_i$ is bounded by planes. The plane implied in (32), which divides the decision region between the ith and jth receive symbols, is given by

$$\mathcal{A}_{i-j} \triangleq \left\{ (x, y, z) : \ \frac{\check{x}_j - \check{x}_i}{\sigma_{\text{cov}}^2} x + \frac{\check{y}_j - \check{y}_i}{\sigma_{\text{cov}}^2} y + \frac{\check{z}_j - \check{z}_i}{2\sigma_{\text{rec}}^2} z \right.$$
$$\left. = \frac{\check{x}_j^2 + \check{y}_j^2 - \check{x}_i^2 - \check{y}_i^2}{2\sigma_{\text{cov}}^2} + \frac{\check{z}_j^2 - \check{z}_i^2}{4\sigma_{\text{rec}}^2} \right\}. \quad (33)$$

The decision regions for 8-PAM, 36-QAM (only for the symbols within the first quadrant of the I-Q-P space) and 4-IM are illustrated in Figs. 12(a), 12(b), and 12(c), respectively.

D. Joint Processing Gain in SER

To quantify the reduction in SER by the splitting receiver, we define the joint processing gain in terms of SER as:

Definition 2 (Joint processing gain in SER): Given a certain modulation scheme, the joint processing gain of the splitting receiver is

$$G \triangleq \frac{\min_{\vec{\rho}=\vec{0},\vec{1}} P_e}{\inf\{P_e : \vec{\rho} \in [0, 1]^K\}}, \quad (34)$$

where $\inf\{\cdot\}$ denotes for the infimum, and P_e is the SER for a given $\vec{\rho}$.

The joint processing gain represents the maximum SER reduction provided by the splitting receiver, compared with the best of the conventional receivers.

V. SPLITTING RECEIVER: SER ANALYSIS

In this section, we derive the SER at a splitting receiver for practical modulation schemes with the transmitted signal constellation Ω_{gen} in the I-Q plane and the received signal constellation $\check{\Omega}_{\text{gen}}$ in the I-Q-P space.

The SER can be written as

$$P_e = \frac{1}{M} \sum_{i=1}^{M} (1 - P_i), \quad (35)$$

where P_i is the symbol success probability of the ith symbol, which is given by

$$P_i = \iiint_{\mathcal{V}_i} \exp\left(-\frac{(x - \check{x}_j)^2}{\sigma_{\text{cov}}^2/2} - \frac{(y - \check{y}_j)^2}{\sigma_{\text{cov}}^2/2} - \frac{(z - \check{z}_j)^2}{\sigma_{\text{rec}}^2} \right) dx\,dy\,dz. \quad (36)$$
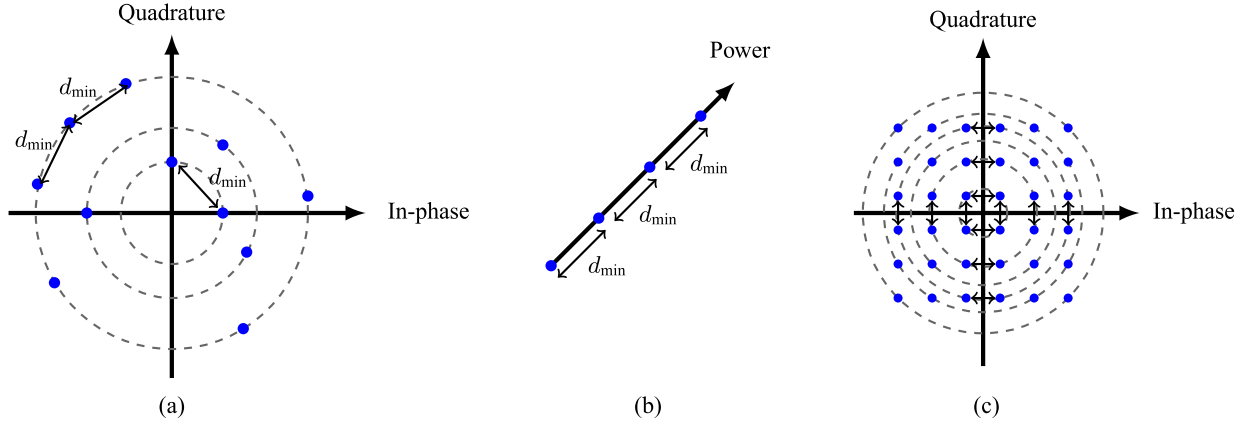
Fig. 13. (a) and (b) Two transmitted signal constellation maps with 3 pairs of symbols that are dominant on detection error probability, plotted on the I-Q plane and the P-axis, respectively. (c) The transmitted signal constellation for 36-QAM, where 12 pairs of symbol that are dominant on detection error probability are illustrated.

Based on Sec. IV.B, when $\mathsf{SNR} \to \infty$ and $\vec{\rho} \in [0, 1]^K \setminus \{\vec{0}, \vec{1}\}$, the received symbols $(\check{x}_i, \check{y}_i, \check{z}_i)$ and $(\check{x}_j, \check{y}_j, \check{z}_j)$ belonging to different power tiers, i.e., $\check{z}_i \neq \check{z}_j$, are easily distinguished because they are separated by a distance proportional to $\mathcal{P}$ in the power domain. In contrast, the symbols belonging to the same power tier are only separated with a distance proportional to $\sqrt{\mathcal{P}}$ on the I-Q plane. Thus, the intra-tier detection error probability dominates the overall SER in the high SNR regime.

Therefore, there are basically two cases for the SER analysis in the high SNR regime:

1) For Ω_{gen} having symbols that belong to the same tier as illustrated in Fig. 13(a), such as M-PAM, M-QAM and M-PSK (phase-shift keying), the intra-tier detection error probability is dominant. Moreover, the detection error caused by the pair symbols with the minimum distance on the I-Q plane is dominant (see Fig. 13(a)).

2) For Ω_{gen} in which every symbol belongs to a different tier as illustrated in Fig. 13(b), such as a M-IM, the inter-tier detection error probability is dominant. Moreover, the detection error caused by the pair of symbols with the minimum distance in the P-axis (power domain) is dominant.

Consider a transmitted signal constellation Ω_{gen} with W pairs of dominant symbols as mentioned above and the minimum distance being $d_{\min}$. The approximated SER is calculated as [1]

$$P_e \approx \frac{1}{M} \sum_{i=1}^{W} 2Q\left(\frac{d_{\min}}{2\sigma}\right), \tag{37}$$

where $\sigma = \sqrt{\sigma_{\text{cov}}^2/2}$ or σ_{rec} for cases 1) and 2), respectively.

It is straightforward to see that: For M-PAM, we have $W = 1$, i.e., the pair of symbols with the lowest power having the minimum distance given by $d_{\min} = 2\check{x}_1$. For M-QAM, we have $W = 2\sqrt{M}$, as illustrated in Fig. 13(c), and $d_{\min} = 2\check{x}_1$. For M-IM, we have $W = M - 1$, as illustrated in Fig. 13(b) and $d_{\min} = \check{x}_2 - \check{x}_1$. Then, based on (37), we can obtain the following results.

A. M-PAM

Proposition 4: For the M-PAM scheme and $\vec{\rho} \in [0, 1]^K \setminus \{\vec{0}, \vec{1}\}$, the SER in the high SNR regime is given by

$$P_e \approx \frac{2}{M} Q\left(\frac{\sqrt{2}\check{x}_1}{\sigma_{\text{cov}}}\right). \tag{38}$$

Based on Proposition 4, we can see that when $\vec{\rho} \neq \vec{1}$ and $\vec{\rho} \to \vec{1}$, $\check{x}_1 \approx \sqrt{3H_2\mathcal{P}/(M^2 - 1)}$, and $P_e \approx \frac{2}{M} Q\left(\sqrt{\frac{6H_2\mathcal{P}}{\sigma_{\text{cov}}^2(M^2-1)}}\right)$, which is smaller than the SER of the $\vec{\rho} = \vec{1}$ case, i.e., $\frac{2(M-1)}{M} Q\left(\sqrt{\frac{6H_2\mathcal{P}}{\sigma_{\text{cov}}^2(M^2-1)}}\right)$, and is also smaller than the SER of the $\vec{\rho} = \vec{0}$ case in which the SER can be as large as 0.5. Thus, we have the following proposition:

Proposition 5: For M-PAM, the asymptotic joint processing gain in the high SNR regime is

$$G_{\text{PAM}} = \frac{\min\left\{0.5, \frac{2(M-1)}{M} Q\left(\sqrt{\frac{6H_2\mathcal{P}}{\sigma_{\text{cov}}^2(M^2-1)}}\right)\right\}}{\frac{2}{M} Q\left(\sqrt{\frac{6H_2\mathcal{P}}{\sigma_{\text{cov}}^2(M^2-1)}}\right)}$$
$$= M - 1. \tag{39}$$

Note that although the joint processing gain depends on $\mathcal{P}$, σ_{cov}^2 and σ_{rec}^2, the asymptotic joint processing gain is independent of the specific noise variance at the CD and PD circuits in the high SNR regime.

B. M-QAM

Proposition 6: For M-QAM and $\vec{\rho} \in [0, 1]^K \setminus \{\vec{0}, \vec{1}\}$, the SER in the high SNR regime is given by

$$P_e \approx \frac{4}{\sqrt{M}} Q\left(\frac{\sqrt{2}\check{x}_1}{\sigma_{\text{cov}}}\right). \tag{40}$$

Letting $\vec{\rho} \to \vec{1}$, i.e., $\check{x}_1 \to \sqrt{\frac{3H_2\mathcal{P}}{2(M-1)}}$, the approximated SER in Proposition 6 is minimized as $P_e \approx \frac{4}{\sqrt{M}} Q\left(\sqrt{\frac{3H_2\mathcal{P}}{(M-1)\sigma_{\text{cov}}^2}}\right)$, which is smaller than the SER obtained by setting $\vec{\rho} = \vec{0}$ or $\vec{1}$. Thus, we have the following result:

Proposition 7: For M-QAM, the asymptotic joint processing gain in the high SNR regime is

$$G_{\text{QAM}} = \lim_{\text{SNR}\to\infty} \frac{1 - \left(1 - 2\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3H_2\mathcal{P}}{(M-1)\sigma_{\text{cov}}^2}}\right)\right)^2}{\frac{4}{\sqrt{M}} Q\left(\sqrt{\frac{3H_2\mathcal{P}}{(M-1)\sigma_{\text{cov}}^2}}\right)}$$

$$= \lim_{\text{SNR}\to\infty} \frac{4\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3H_2\mathcal{P}}{(M-1)\sigma_{\text{cov}}^2}}\right)}{\frac{4}{\sqrt{M}} Q\left(\sqrt{\frac{3H_2\mathcal{P}}{(M-1)\sigma_{\text{cov}}^2}}\right)}$$

$$= \sqrt{M} - 1. \tag{41}$$

Therefore, in the high SNR regime, for M-PAM and M-QAM, there always exists a non-trivial $\vec{\rho}$ that $\vec{\rho} \in [0, 1]^K \setminus \{\vec{0}, \vec{1}\}$ and achieves a lower SER than the conventional receivers, i.e., $\vec{\rho} = \vec{0}$ or $\vec{1}$, no matter what values σ_{cov}^2 and σ_{rec} take.

C. M-IM

Proposition 8: For M-IM, the SER in the high SNR regime is given by

$$P_e \approx \frac{2(M-1)}{M} Q\left(\frac{\sqrt{\Theta_2}\mathcal{P}}{(M-1)\sigma_{\text{rec}}}\right). \tag{42}$$

From Proposition 8, as $\vec{\rho} \to \vec{0}$, the minimum approximated SER is obtained as $P_e = \frac{2(M-1)}{M} Q\left(\frac{\sqrt{H_4}\mathcal{P}}{(M-1)\sigma_{\text{rec}}}\right)$, which equal to the SER when $\vec{\rho} = \vec{0}$. Thus, the splitting receiver cannot improve the SER performance compared with the conventional receivers, and we have the result:

Proposition 9: For M-IM, the asymptotic joint processing gain in the high SNR regime is equal to one.

D. Numerical Results

We present the numerical results using M-QAM for (i) the splitting receiver with single receiver antenna assuming $|\tilde{h}_1|^2 = 1$, and (ii) the simplified receiver with multiple receiver antennas. The SER results for M-QAM are plotted based on Monte Carlo simulation with 10^9 points using the detection rule (32). The results for M-PAM and M-IM are omitted due to space limitation.

1) Splitting Receiver With Single Receiver Antenna: Fig. 14 plots the SER versus the splitting ratio ρ for different M and different $\mathcal{P}$, where the approximation results are plotted using Proposition 6. It shows that the SER first decreases and then increases as ρ increases. We can see that the optimal ρ that minimizes the SER, increases with $\mathcal{P}$ and approaches 1 but decreases with the increasing of the order of constellation M. Also we can see that when SNR is sufficiently large, e.g., 23 dB (i.e., $\mathcal{P} = 200$, and $\sigma_{\text{cov}}^2 = \sigma_{\text{rec}}^2 = 1$), the approximation of the SER is very close with the accurate SER for the value of ρ in the range $(0, \rho^\star)$, where $\rho^\star$ is the optimal splitting ratio. Note that $\rho^\star$ approaches 1 as SNR increases. This means the mismatch around $\rho = 1$ is minimized as SNR increases. *Therefore, the approximation in Proposition 6 is accurate for the values of $\rho \in (0, 1)$ when SNR is sufficiently large.*
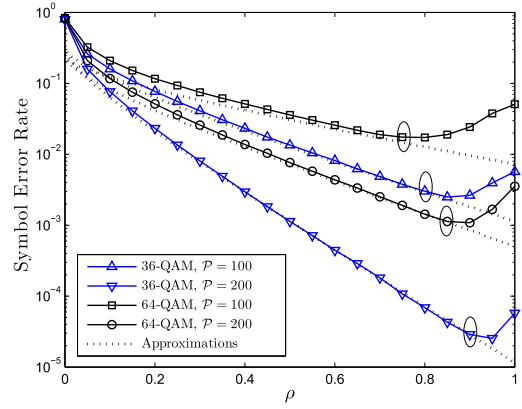

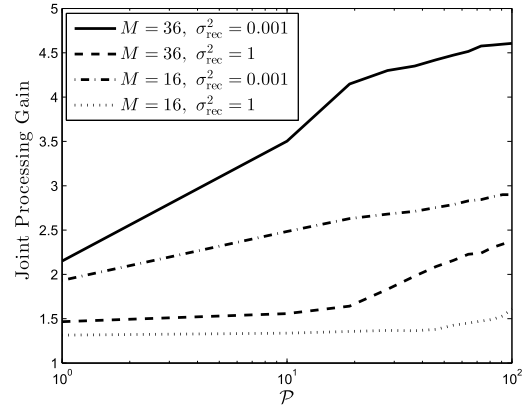Fig. 14.   SER versus ρ, $\sigma_{\text{cov}}^2 = 1$ and $\sigma_{\text{rec}}^2 = 1$.

Fig. 15. G_{QAM} versus $\mathcal{P}$, $\sigma_{\text{cov}}^2 = 1$.

Fig. 15 shows the joint processing gain versus $\mathcal{P}$ using Definition 2. We can see that the joint processing gain increase with $\mathcal{P}$ and approaches 3 and 5 for 16-QAM and 36-QAM, respectively, when $\mathcal{P} = 100$, $\sigma_{\text{cov}}^2 = 1$ and σ_{rec}^2 is sufficiently small, e.g., 10^{-3}. These results approach the asymptotic joint processing gain in Proposition 7. Also we see that only half the joint processing gain is achieved when $\mathcal{P} = 100$ and σ_{rec}^2 is large, e.g., $\sigma_{\text{rec}}^2 = 1$. However, the increasing trend of the joint processing gain in Fig. 15 suggest that the asymptotic joint processing gain can be eventually achieved, when $\mathcal{P}$ is much larger than 100. *Therefore, the asymptotic joint processing gain in Proposition 7 may not be approached in a normal range of the received signal power and the noise variance, but half of the joint processing gain is achievable.*

2) Simplified Receiver With Multiple Receiver Antennas: For the simplified receiver, we assume that the channel power gain at each antenna is independent and follows an exponential distribution with the mean of 1. Fig. 16 plots the optimal number of antennas allocated for coherent processing, i.e., $K_1^\star$, versus the total number of antennas for the 36-QAM scheme obtained by using 10^3 random channel realizations. It shows that $K_1^\star$ increases with K, and approaches to $K - 1$ in the high SNR regime, e.g, $K_1^\star \approx K - 5$, $K - 2$ and $K - 1$ when $\mathcal{P} = 2$, 10 and 200, respectively. This is because that the optimal ratio $\vec{\rho} \to \vec{1}$ but never reaches $\vec{1}$ based on Proposition 4 in the high SNR regime. In other words, for the simplified receiver (where $\rho_k \in \{0, 1\}$), most of the antennas
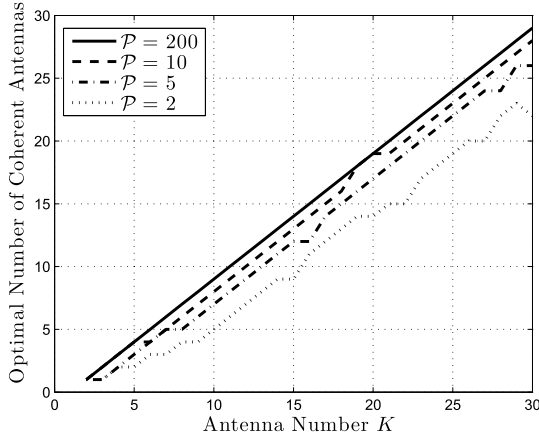
Fig. 16. Optimal number of antennas allocated for coherent processing versus K for 16-QAM, $\sigma_{\text{cov}}^2 = \sigma_{\text{rec}}^2 = 1$.

should be connected to the CD circuits and at least one antenna should be connected to a PD circuit to achieve the highest joint processing gain.

Note that in practice, the degradation of ADC noise is usually modeled by the signal-to-quantization-noise ratio (SQNR), approximately given by $6K$ dB, where K is the number of quantization bits [15]. Here, by assuming $\mathcal{P} = 2$ and the noise variance of the ADC equals to 0.1 (i.e., less than σ_{cov}^2 and σ_{rec}^2), the SQNR equals to 13 dB, which implies $K \approx 2$ bits. Similarly, by assuming $\mathcal{P} = 200$, the SQNR equals to 33 dB, which implies $K \approx 5$ bits. Therefore, the parameter settings are practical.

VI. CONCLUSIONS

In this paper, we have proposed a splitting receiver, which fundamentally changes the way in which the signal is processed. With the same received signal power, the analytical results show that the splitting receiver provides excellent performance gain in the sense of both the mutual information (Gaussian input) and the SER (practical modulation), compared with the conventional coherent and non-coherent receivers. Future research may focus on the topics such as the MIMO system with a multi-antenna splitting receiver, and the design of constellations and coding schemes for the communication systems with splitting receiver. Moreover, some practical issues can also be taken into account, such as the effects of the antenna noise, the power splitter losses and the different receive sensitivity level at the CD and PD circuits of the splitting receiver.

APPENDIX
PROOF OF LEMMA 1

We assume that $\vec{\rho} \in [0, 1]^K \setminus \{\vec{0}, \vec{1}\}$. Thus, based on (6), $\Theta_1 > 0$ and $\Theta_2 > 0$.

A. Proof of (18a)

Based on the property of mutual information invariance under scaling of random variables [19], a scaled received

signal expression based on (7) is given by

$$\tilde{Y}_1 = \sqrt{\Theta_1}\tilde{X} + \frac{\tilde{Z}}{\sqrt{\mathcal{P}}},$$

$$Y_2 = \sqrt{\Theta_2}k\sqrt{\mathcal{P}}|\tilde{X}|^2 + k\frac{N}{\sqrt{\mathcal{P}}}, \quad \text{(A.1)}$$

where $k \triangleq \frac{\sigma_{\text{cov}}}{\sqrt{2}\sigma_{\text{rec}}}$. Thus, it is easy to verify that the real and imaginary parts of $\frac{\tilde{Z}}{\sqrt{\mathcal{P}}}$ and $k\frac{N}{\sqrt{\mathcal{P}}}$ are independent with each other and follow the same distribution $\mathcal{N}(0, \frac{\sigma_{\text{cov}}^2}{2\mathcal{P}})$.

We define two random variables as

$$\tilde{X}_1 \triangleq \sqrt{\Theta_1}\tilde{X}, \quad \text{and} \quad X_2 \triangleq \sqrt{\Theta_2}k\sqrt{\mathcal{P}}|\tilde{X}|^2. \quad \text{(A.2)}$$

Because of the Markov chain $\sqrt{\mathcal{P}}\tilde{X} \to (\tilde{X}_1, X_2) \to (\tilde{Y}_1, Y_2)$ and the smooth and uniquely invertible map from $\sqrt{\mathcal{P}}\tilde{X}$ to $(\tilde{X}_1, X_2)$, we have

$$I(\sqrt{\mathcal{P}}X; \tilde{Y}_1, Y_2) = I(\tilde{X}_1, X_2; \tilde{Y}_1, Y_2). \quad \text{(A.3)}$$

Before the analysis of $I(\tilde{X}_1, X_2; \tilde{Y}_1, Y_2)$, we first define a new coordinate system named as paraboloid-normal (PN) coordinate system which is based on a paraboloid $\mathcal{U}$. The paraboloid $\mathcal{U}$ is defined by the equation

$$c_P = k\sqrt{\mathcal{P}}\frac{\sqrt{\Theta_2}}{\Theta_1}(c_I^2 + c_Q^2), \quad \text{(A.4)}$$

where c_I, c_Q and c_P are the three axes of Cartesian coordinate system of the I-Q-P space. By changing coordinate system, the point (c_1, c_2, c_3) is represented as $(\tilde{a}, l)$ in the PN coordinate system, where $\tilde{a}$ is the nearest point on the paraboloid $\mathcal{U}$ to the point (c_1, c_2, c_3), and $|l|$ is the distance. In other words, the point (c_1, c_2, c_3) is on the normal line at the point $\tilde{a}$ on the paraboloid. Specifically, the sign of l is positive when the point (c_1, c_2, c_3) is above the parabolic, otherwise, it is negative.

Based on the property of mutual information invariance under a change of coordinates [19], representing Cartesian coordinate based random variables $(\tilde{X}_1, X_2)$ and $(\tilde{Y}_1, Y_2)$ under the PN coordinate system as $(\tilde{A}_X, L_X)$ and $(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}, L_{\tilde{Z},N})$, respectively, gives

$$I(\tilde{X}_1, X_2; \tilde{Y}_1, Y_2) = I(\tilde{A}_X, L_X; \tilde{A}_X + \tilde{A}_{\tilde{Z},N}, L_{\tilde{Z},N}), \text{(A.5)}$$

where the noise-related random variables $\tilde{A}_{\tilde{Z},N}$ and $L_{\tilde{Z},N}$, which are generated by $\tilde{Z}$ and N, are correlated with the random variable $\tilde{A}_X$. Since $X_2 = k\sqrt{\mathcal{P}}\frac{\sqrt{\Theta_2}}{\Theta_1}|\tilde{X}_1|^2$, $(\tilde{X}_1, X_2)$ lies on the paraboloid $\mathcal{U}$, i.e., L_X is a constant which is equal to zero, $\tilde{A}_X$ can be represented by $(\tilde{X}_1, X_2)$ for brevity. Thus, we have

$$I(\tilde{X}_1, X_2; \tilde{Y}_1, Y_2)$$

$$= I(\tilde{A}_X; \tilde{A}_X + \tilde{A}_{\tilde{Z},N}, L_{\tilde{Z},N})$$

$$= h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}, L_{\tilde{Z},N}) - h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}, L_{\tilde{Z},N}|\tilde{A}_X)$$

$$= h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}) + h(L_{\tilde{Z},N}|\tilde{A}_X + \tilde{A}_{\tilde{Z},N})$$

$$\quad - \Big(h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}|\tilde{A}_X) + h(L_{\tilde{Z},N}|\tilde{A}_X, \tilde{A}_X + \tilde{A}_{\tilde{Z},N})\Big). \text{(A.6)}$$

Since the expectations $\mathbb{E}\left(\tilde{Z}\right) = (0,0)$ and $\mathbb{E}(N) = 0$, and the variances $\mathrm{Var}(\frac{\tilde{Z}}{\sqrt{\mathcal{P}}}) \to 0$ and $\mathrm{Var}(k\frac{N}{\sqrt{\mathcal{P}}}) \to 0$ as $\mathcal{P} \to \infty$, it is easy to see that the noise variable $\tilde{A}_{\tilde{Z},N}$ converges in probability towards $(0,0)$. Thus, $\tilde{A}_X + \tilde{A}_{\tilde{Z},N}$ converges in probability towards $\tilde{A}_X$.

Furthermore, since convergence in probability implies convergence in distribution and the entropy function $h(\cdot)$ is continuous and defined based on the probability distribution of the input random variable [17], we have $h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}) \to h(\tilde{A}_X)$ as $\mathcal{P} \to \infty$. Similarly, we have the convergence of the random variable, i.e., $\left(L_{\tilde{Z},N}, \tilde{A}_X + \tilde{A}_{\tilde{Z},N}\right) \to \left(L_{\tilde{Z},N}, \tilde{A}_X\right)$, hence the convergence of entropy, i.e., $h(L_{\tilde{Z},N}, \tilde{A}_X + \tilde{A}_{\tilde{Z},N}) \to h(L_{\tilde{Z},N}, \tilde{A}_X)$.

Therefore, the conditional entropy $h(L_{\tilde{Z},N}|\tilde{A}_X + \tilde{A}_{\tilde{Z},N}) \triangleq h(L_{\tilde{Z},N}, \tilde{A}_X + \tilde{A}_{\tilde{Z},N}) - h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N})$ converges to $h(L_{\tilde{Z},N}|\tilde{A}_X) \triangleq h(L_{\tilde{Z},N}, \tilde{A}_X) - h(\tilde{A}_X)$, i.e., $h(L_{\tilde{Z},N}|\tilde{A}_X + \tilde{A}_{\tilde{Z},N}) \to h(L_{\tilde{Z},N}|\tilde{A}_X)$. Similarly, we have $h(L_{\tilde{Z},N}|\tilde{A}_X, \tilde{A}_X + \tilde{A}_{\tilde{Z},N}) \to h(L_{\tilde{Z},N}|\tilde{A}_X, \tilde{A}_X)$. Together with the fact that $h(L_{\tilde{Z},N}|\tilde{A}_X) = h(L_{\tilde{Z},N}|\tilde{A}_X, \tilde{A}_X)$ [17], we have $h(L_{\tilde{Z},N}|\tilde{A}_X + \tilde{A}_{\tilde{Z},N}) - h(L_{\tilde{Z},N}|\tilde{A}_X, \tilde{A}_X + \tilde{A}_{\tilde{Z},N}) \to 0$ as $\mathcal{P} \to \infty$. Thus, the mutual information in (A.6) can asymptotically be rewritten as

$$I(\tilde{X}_1, X_2; \tilde{Y}_1, Y_2) = h(\tilde{A}_X) - h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}|\tilde{A}_X). \quad \text{(A.7)}$$

Then we calculate $h(\tilde{A}_X)$ and $h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}|\tilde{A}_X)$ as follows.

1) $h(\tilde{A}_X)$: Due to the fact that the probability contained in a differential area should not alter under a change of variables, we have

$$|f_{\tilde{X}}(\tilde{x})\mathrm{d}S| = |f_{\tilde{A}_X}(\tilde{a})\mathrm{d}\Sigma|, \quad \text{(A.8)}$$

where $\mathrm{d}S = \mathrm{d}u\mathrm{d}v$, $u = \mathrm{Real}\{\tilde{x}\}$, $v = \mathrm{Imag}\{\tilde{x}\}$, $\tilde{a}$ is the PN coordinate system representation of the point $(\tilde{x}_1, x_2)$, $\mathrm{d}\Sigma$ is the differential area on the paraboloid $\mathcal{U}$, $f_{\tilde{A}_X}(\tilde{a})$ and $f_{\tilde{X}}(\tilde{x})$ are the pdfs of $\tilde{A}_X$ and $\tilde{X}$, respectively, and

$$f_{\tilde{X}}(\tilde{x}) = \frac{1}{\pi} \exp\left(-|\tilde{x}|^2\right). \quad \text{(A.9)}$$

Assuming that $\mathbf{r} = (\mathrm{Real}\{\tilde{x}_1\}, \mathrm{Imag}\{\tilde{x}_1\}, x_2)$, which is a point on the paraboloid $\mathcal{U}$, and thus, based on (A.2), we have

$$\frac{\partial \mathbf{r}}{\partial u} = (\sqrt{\Theta_1}, 0, 2k\sqrt{\mathcal{P}}\sqrt{\Theta_2}u),$$

$$\frac{\partial \mathbf{r}}{\partial v} = (0, \sqrt{\Theta_1}, 2k\sqrt{\mathcal{P}}\sqrt{\Theta_2}v),$$

$$\mathrm{d}\Sigma = \left|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right| \mathrm{d}u\mathrm{d}v = \Theta_1\sqrt{4\frac{k^2\mathcal{P}\Theta_2}{\Theta_1^2}|\tilde{x}_1|^2 + 1} \ \mathrm{d}u\mathrm{d}v,$$

$$\text{(A.10)}$$

where $\times$ is the cross product operator. Taking (A.10), (A.9) and $\tilde{x} = \frac{\tilde{x}_1}{\sqrt{\Theta_1}}$ into (A.8), after simplification, we have

$$f_{\tilde{A}_X}(\tilde{a}) = f_{\tilde{A}_X}(\tilde{x}_1, x_2)$$

$$= \frac{1}{\pi \Theta_1\sqrt{4\frac{k^2\mathcal{P}\Theta_2}{\Theta_1^2}|\tilde{x}_1|^2 + 1}} \exp\left(-\frac{|\tilde{x}_1|^2}{\Theta_1}\right). \quad \text{(A.11)}$$

The differential entropy of $\tilde{A}_X$ is derived as

$$h(\tilde{A}_X) = \iint -f_{\tilde{A}_X}(\tilde{a}) \log_2\left(f_{\tilde{A}_X}(\tilde{a})\right)\mathrm{d}\Sigma. \quad \text{(A.12)}$$

Taking (A.11) and (A.10) into (A.12), we have

$h(\tilde{A}_X)$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} -\frac{\exp\left(-\frac{|\tilde{x}_1|^2}{\Theta_1}\right)}{\pi} \log_2\left(\frac{\exp\left(-\frac{|\tilde{x}_1|^2}{\Theta_1}\right)}{\pi\Theta_1\sqrt{4\frac{k^2\mathcal{P}\Theta_2}{\Theta_1^2}|\tilde{x}_1|^2 + 1}}\right)\mathrm{d}u\mathrm{d}v$$

$$\overset{(a)}{=} 2\pi\int_0^{\infty} -\frac{\exp\left(-\frac{r^2}{\Theta_1}\right)}{\pi\Theta_1} \log_2\left(\frac{\exp\left(-\frac{r^2}{\Theta_1}\right)}{\pi\Theta_1\sqrt{4\frac{k^2\mathcal{P}\Theta_2}{\Theta_1^2}r^2 + 1}}\right)r\mathrm{d}r$$

$$= \log_2(\pi e\Theta_1) + \frac{1}{2\log(2)}\exp\left(\frac{\Theta_1}{4k^2\mathcal{P}\Theta_2}\right)\mathrm{Ei}\left(\frac{\Theta_1}{4k^2\mathcal{P}\Theta_2}\right),$$

$$\text{(A.13)}$$

where (a) is because of the polar transformation.

2) Asymptotic $h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}|\tilde{A}_X)$: For a given value of $\tilde{A}_X$, based on the definition of the PN coordinate system, the random variable $\tilde{A}_X + \tilde{A}_{\tilde{Z},N}$ is treated as the projection of the three-dimensional circular symmetric Gaussian noise, i.e., $(\frac{\tilde{Z}}{\sqrt{\mathcal{P}}}, k\frac{N}{\sqrt{\mathcal{P}}})$ shifted by $\tilde{A}_X$, on the parabolic $\mathcal{U}$ by the normal vectors of it.

As $\mathcal{P} \to \infty$, $\tilde{A}_X + \tilde{A}_{\tilde{Z},N}$ converges in probability toward $\tilde{A}_X$, thus, for a given value of $\tilde{A}_X$, the effective range of the random variable $\tilde{A}_X + \tilde{A}_{\tilde{Z},N}$ on the paraboloid $\mathcal{U}$, is very small, which is close to the tangent plane of $\mathcal{U}$ at the point $\tilde{A}_X$. Therefore, the random variable $\tilde{A}_X + \tilde{A}_{\tilde{Z},N}$ converges in probability toward the random variable generated by the projection of the three-dimensional circular symmetric Gaussian noise on the tangent plane of $\mathcal{U}$ at the point $\tilde{A}_X$ by the normal vector of the point $\tilde{A}_X$, which is the well-known two-dimensional complex Gaussian random variable with variance $\sigma_{\mathrm{cov}}^2/\mathcal{P}$. Therefore, given $\tilde{A}_X$, the entropy $h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N})$ is approaching to $\log_2(\pi e\sigma_{\mathrm{cov}}^2/\mathcal{P})$ which does not rely on $\tilde{A}_X$. Thus, as $\mathcal{P} \to \infty$, the asymptotic conditional entropy is

$$h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}|\tilde{A}_X) = \mathbb{E}_{\tilde{A}_X}\left[h(\tilde{A}_X + \tilde{A}_{\tilde{Z},N}|\tilde{A}_X = \tilde{a}_X)\right]$$

$$\approx \log_2\left(\frac{\pi e\sigma_{\mathrm{cov}}^2}{\mathcal{P}}\right). \quad \text{(A.14)}$$

3) Asymptotic $I(\sqrt{\mathcal{P}}X; \tilde{Y}_1, Y_2)$: Taking (A.13) and (A.14) into (A.7), (18a) is obtained.

B. Proof of (18b)

Based on the power series expansion of the exponential integral function [20]

$$\mathrm{Ei}(x) = -\gamma - \ln x - \sum_{n=1}^{\infty}\frac{(-x)^n}{n\,n!}, \quad x > 0, \quad \text{(A.15)}$$

where $\gamma \approx 0.5772$ is Euler's constant, as $\mathcal{P}$ is sufficiently large, we have

$$\lim_{\mathcal{P}\to\infty} \exp\left(\frac{\Theta_1\sigma_{\text{rec}}^2}{2\sigma_{\text{cov}}^2\mathcal{P}\Theta_2}\right) \text{Ei}\left(\frac{\Theta_1\sigma_{\text{rec}}^2}{2\sigma_{\text{cov}}^2\mathcal{P}\Theta_2}\right)$$

$$= -\gamma + \ln\left(\frac{2\sigma_{\text{cov}}^2\mathcal{P}\Theta_2}{\Theta_1\sigma_{\text{rec}}^2}\right). \tag{A.16}$$

Substituting (A.16) into (18a), (18b) is obtained.

REFERENCES

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[2] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath, Jr., "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014.

[3] J. G. Andrews, T. Bai, M. Kulkarni, A. Alkhateeb, A. Gupta, and R. W. Heath, Jr. (2016). "Modeling and analyzing millimeter wave cellular systems." [Online]. Available: https://arxiv.org/pdf/1605.04283v1.pdf

[4] M. Chowdhury, A. Manolakos, and A. Goldsmith, "Scaling laws for noncoherent energy-based communications in the SIMO MAC," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1980–1992, Apr. 2016.

[5] L. Jing, E. D. Carvalho, P. Popovski, and À. O. Martínez, "Design and performance analysis of noncoherent detection systems with massive receiver arrays," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5000–5010, Oct. 2016.

[6] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 117–125, Apr. 2015.

[7] K. Huang and X. Zhou, "Cutting the last wires for mobile communications by microwave power transfer," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 86–93, Jun. 2015.

[8] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, "Simultaneous wireless information and power transfer in modern communication systems," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 104–110, Nov. 2014.

[9] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, 2nd Quart., 2015.

[10] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.

[11] K. Szczerba *et al.*, "4-PAM for high-speed short-range optical communications," *J. Opt. Commun. Netw.*, vol. 4, no. 11, pp. 885–894, Nov. 2012.

[12] A. Lapidoth, S. Moser, and M. Wigger, "On the capacity of free-space optical intensity channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.

[13] Y. Wu, Y. Liu, Q. Xue, S. Li, and C. Yu, "Analytical design method of multiway dual-band planar power dividers with arbitrary power division," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 12, pp. 3832–3841, Dec. 2010.

[14] Agilent Technologies. *DC to 18 GHz Power Splitter*, accessed on 2012. [Online]. Available: http://cp.literature.agilent.com/litweb/pdf/5990-5351EN.pdf

[15] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: Architecture design and rate-energy trade-off," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754–4767, Nov. 2013.

[16] W. Liu, X. Zhou, S. Durrani, and P. Popovski, "SWIPT with practical modulation and RF energy harvesting sensitivity," in *Proc. IEEE ICC*, May 2016, pp. 1–7.

[17] T. Cover and J. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2006.

[18] E. Grushka, "Characterization of exponentially modified Gaussian peaks in chromatography," *Anal. Chem.*, vol. 44, no. 11, pp. 1733–1738, 1972.

[19] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 6, p. 066138, Jun. 2004.

[20] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, vol. 55. North Chelmsford, MA, USA: Courier Corporation, 1964.

Wanchun Liu (S'13) received the B.S. and M.S.E. degrees in electronics and information engineering from Beihang University, Beijing, China. She is currently pursuing the Ph.D. degree with the Research School of Engineering, Australian National University, Canberra, Australia. Her research interest lies in wireless communications and includes the application of wireless communication networks powered by energy harvesting, and backscatter communications.

Xiangyun Zhou (M'11) received the Ph.D. degree from Australian National University (ANU) in 2010. He is currently a Senior Lecturer with (ANU). His research interests are in communication theory and wireless networks. He was a recipient of the Best Paper Award at ICC'11 and the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2016. He currently serves on the Editorial Board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE COMMUNICATIONS LETTERS. He also served as a Symposium/Track and Workshop Co-Chair of major IEEE conferences. He was the Chair of the ACT Chapter of the IEEE Communications Society and Signal Processing Society, from 2013 to 2014. He served as a Guest Editor of the *IEEE Communications Magazine*'s feature topic on wireless physical layer security in 2015.

Salman Durrani (S'00–M'05–SM'10) received the B.Sc. degree (Hons.) degree in electrical engineering from the University of Engineering and Technology at Lahore, Lahore, Pakistan, in 2000, and the Ph.D. degree in electrical engineering from The University of Queensland, Brisbane, Australia, in 2004. He has been with the Australian National University, Canberra, Australia, since 2005, where he is currently a Senior Lecturer with the College of Engineering and Computer Science, Research School of Engineering. His research interests are in wireless communications and signal processing, including machine-to-machine and device-to-device communication, wireless energy harvesting systems, the stochastic geometry modeling of finite area networks, and synchronization in communication systems. He is a member of Engineers Australia and a Senior Fellow of The Higher Education Academy, U.K. He was a recipient of the 2016 IEEE ComSoc Asia Pacific Outstanding Paper Award. He was the Chair of the ACT Chapter of the IEEE Signal Processing and Communications Societies, from 2015 to 2016. He currently serves as an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS.

Petar Popovski (S'97–A'98–M'04–SM'10–F'16) received the Dipl.-Ing. degree in electrical engineering and the Magister Ing. degree in communication engineering from the Saints Cyril and Methodius University of Skopje, Skopje, Macedonia, in 1997 and 2000, respectively, and the Ph.D. degree from Aalborg University, Denmark, in 2004. He is currently a Professor in wireless communications with Aalborg University. He has over 290 publications in journals, conference proceedings, and books. He has over 30 patents and patent applications. His research interests are wireless communication and networking and communication/information theory. From 2014 to 2016, he was a Steering Committee Member of the IEEE INTERNET OF THINGS. He currently serves as a Steering Committee Member of the IEEE SmartGridComm. He is a holder of a Consolidator Grant from the European Research Council. He was a recipient of the Elite Researcher Award in Denmark in 2016. He is currently an Area Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. From 2012 to 2014, he served as the Chair of the IEEE ComSoc Emerging Technology Committee on Smart Grid Communications.