IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 15, 2016, accepted August 30, 2016, date of publication September 13, 2016, date of current version September 28, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2608907

Toward Efficient 5G Transmission: SER Performance Analysis for Asynchronous Physical-Layer Network Coding

LEI GUO[1], ZHAOLONG NING[2,3], QINGYANG SONG[1], YAYUN CUI[1], AND ZHIKUI CHEN[2]

[1]School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China
[2]School of Software, Dalian University of Technology, Dalian 116620, China
[3]State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China

Corresponding author: Z. Ning (zhaolongning@dlut.edu.cn) and Q. Song (songqingyang@cse.neu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61471109, Grant 91438110, Grant 61302072, Grant 61502075, and Grant 61672123, in part by the Fundamental Research Funds for the Central Universities under Grant N150401002, Grant N140405005, Grant N140405007, Grant DUT15RC(3)009, in part by the China Post-Doctoral Science Foundation Project under Grant 2015M580224, in part by the Liaoning Province Doctor Startup Fund under Grant 201501166, and in part by the State Key Laboratory for Novel Software Technology, within the Nanjing University under Grant KFKT2015B12.

ABSTRACT People's pursuit of higher quality communication has never ceased, which challenges the layout of 5G networks. Physical-layer network coding (PNC), as a key technology for 5G, supplies a powerful platform through leveraging the broadcast nature of wireless media. However, the symbol error rate (SER) of PNC is not well investigated, which would seriously influence user's quality of experience due to packet loss in wireless environment. In this paper, considering both phase and symbol misalignments, we perform analysis on SER of asynchronous PNC. By assuming part of information is known to the relay, and we derive the lower bound for SER. Afterward, through applying the concept of error vector and eliminating redundant terms, we derive the upper bound for SER. Both the lower and upper bounds are applicable to either multiuser detection-based network coding or belief propagation-based maximum a posteriori decoding. Finally, Monte Carlo simulation verifies our results and demonstrates that the bounds are relatively tight. The analytical results derived in this paper can facilitate future studies of practical and theoretical issues on PNC.

INDEX TERMS Asynchronous PNC, belief propagation, performance bounds, SER analysis, 5G.

I. INTRODUCTION

The great requirements for dramatic delay reduction and throughput enhancement call a new area of communication systems. It is acknowledged that 5G will bring an army of performance improvement in wireless coverage, spectrum utilization, transmission delay, user experience, and so on [1]. Recently, research on improving spectrum efficiency is now being motivated by novel physical layer techniques, among which physical-layer network coding (PNC) has attracted much interest [2], [3]. PNC's specific strength in improving throughput comes from exploiting the superposition of electromagnetic waves and the broadcast nature of wireless channels during its multiple access (MA) phase and broadcast (BC) phase, respectively, as shown in Fig. 1.

Physical layer is fundamental for wireless communications [4], [5]. Although a lot of literatures focused on throughput improvement in PNC, current network design falls in the fairly narrow set of objections, limiting the effectiveness and feasibility of the resource demanding applications of

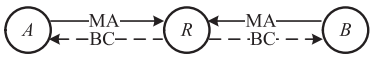


FIGURE 1. PNC over a bidirectional relay network.

5G networks [6], [7], and the symbol error rate (SER) of PNC is not well investigated, which would seriously influence user's quality of experience due to packet loss in wireless environment [8]. Under the context of error performance analysis for PNC, current works focus on PNC schemes with symbol alignment [9]–[12]. Assuming phase synchronization and symbol alignment, SER for binary phase-shift keying (BPSK) and quadrature phase-shift keying (QPSK) was analyzed in [9]. Considering the effect of phase errors (which is unknown to the receiver) during the synchronization process, the SER for arbitrary M-ary quadrature amplitude modulation (M-QAM) was analyzed in [10]. Assuming that individual phases of the two overlapped signals can be trackable, references [11] and [12] analyzed the SER of PNC when performing minimum distance decoding

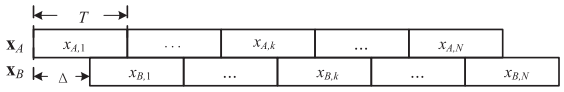


FIGURE 2. Superposed complex baseband signal with *N*-symbol length.

(as in [13]) and maximum likelihood decoding, respectively, over fading channels.

Other than the PNC schemes requiring symbol alignment, there are also designs based on trackable symbol misalignment, which is also known as asynchronous PNC, such as [14], which takes advantage of symbol misalignment to improve the decoding performance under arbitrary phase differences. However, to the best of our knowledge, existing work have not focused on SER performance analysis for asynchronous PNC, whereas analytical results on error performance are expected to facilitate future practical and theoretical studies on higher-layer designs. To this end, this paper focuses on SER performance analysis for asynchronous PNC. We focus on BPSK modulation in this paper as an initial but significant step towards this direction, and higher level modulations can be considered in the future. Our main contributions can be summarized as follows.

- Considering trackable relative symbol and phase offsets, we derive an approximate lower bound and an exact upper bound of SER for asynchronous PNC with BPSK, where we consider two candidate decoding and mapping methods: multiuser detection (MUD) based exclusive-or (XOR) network coding (MUD-XOR) [15], [16] and belief propagation (BP) algorithm based maximum a posteriori (MAP) decoding (BP-MAP) [17]. The lower and upper bounds are respectively derived based on BP-MAP and MUD-XOR, and are suitable for either decoding method, as will be discussed in Section II-B.
- Specifically, the lower bound is obtained by assuming that part of the overlapped messages (received by the relay) is known to the relay, and the upper bound is obtained by expressing the error probability as a sum of upper bounds and eliminating redundant terms.
- Our simulation results indicate that the lower and upper bounds are relatively tight.

The remainder of this paper is organized as follows. In Section II, we introduce the signal model and the two different decoding methods, i.e. MUD-XOR and BP-MAP, and discuss their relationship. Subsequently, Sections III and IV respectively derive the lower and upper bounds of SER for asynchronous PNC. Afterwards, the analytical results are verified by simulations in Section V. Conclusions are drawn in Section VI.

II. PRELIMINARIES

A. SIGNAL MODEL FORMULATION

We consider a typical bidirectional relaying network with slow flat fading channel, and the channel power gain is assumed to be one. As shown in Fig. 2, the superposed complex baseband signal (which carries a total of N symbols) received by relay R can be expressed as

$$y_R(t) = \sum_{k=1}^{N} x_{A,k} p(t-kT) + x_{B,k} e^{j\phi} p(t-\Delta-kT) + n(t), \tag{1}$$

where $x_{A,k}$ and $x_{B,k}$ respectively denote the signals from source nodes A and B, and $x_{A,k}, x_{B,k} \in \{-1, 1\}$ for BPSK; ϕ and Δ respectively denote relative carrier phase and symbol offsets, and $0 < \Delta < T$ (where T denotes the symbol interval); the function $p(t)$ is the rectangular pulse shape; and $n(t)$ is complex additive white Gaussian noise.

B. DECODING METHODS FOR SUPERPOSED SYMBOLS

This subsection first reviews the MUD-XOR and BP-MAP methods, and then illustrates the relationship between these two methods. In this paper, the lower bound for SER is derived under the BP-MAP decoding method and the upper bound for SER is derived under the MUD-XOR decoding method. As we will discuss later, the lower and upper bounds are correct for cases when performing either MUD-XOR or BP-MAP.

1) THE MUD-XOR METHOD

MUD-XOR exploits MUD technique to estimate individual packets from the superposed signal. The outputs of the correlators when respectively synchronized with symbols from A and B are given by,

$$\begin{aligned} y_{A,k} &= \int_{(k-1)T}^{kT} y_R(t)\, p(t-kT)\, dt \\ &= x_{A,k} + \frac{\Delta}{T} x_{B,k-1} e^{j\phi} + \frac{T-\Delta}{T} x_{B,k} e^{j\phi} \\ &\quad + \int_{(k-1)T}^{kT} n(t)\, p(t-kT)\, dt \end{aligned} \tag{2}$$

$$\begin{aligned} y_{B,k} &= \int_{(k-1)T+\Delta}^{kT+\Delta} y_R(t)\, p(t-\Delta-kT)\, dt \\ &= x_{B,k} e^{j\phi} + \frac{\Delta}{T} x_{A,k+1} + \frac{T-\Delta}{T} x_{A,k} \\ &\quad + \int_{(k-1)T+\Delta}^{kT+\Delta} n(t)\, p(t-kT-\Delta)\, dt \end{aligned} \tag{3}$$

where $k \in \{1, 2, \ldots, N\}$ and $\int_{(k-1)T}^{kT} p^2(t-kT)\, dt = 1$. The received signals $y_{A,k}$ and $y_{B,k}$ are complex Gaussian random variables with variance σ^2 for both real and imaginary components. Let vector $\mathbf{x} = (x_{A,1}, x_{B,1}, \ldots, x_{A,N}, x_{B,N})$ denote the transmitted sequences from end nodes and vector $\mathbf{y}_1 = (y_{A,1}, y_{B,1}, \ldots, y_{A,N}, y_{B,N})$ denote the outputs from the correlator, the estimated $\hat{\mathbf{x}}$ can be achieved by

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} P(\mathbf{y}_1 \mid \mathbf{x}), \tag{4}$$

which is essentially the maximum-likelihood sequence estimation. For notational simplicity, let

$$s(t, \mathbf{x}) = \sum_{k=1}^{N} x_{A,k} p(t-kT) + x_{B,k} e^{j\phi} p(t-\Delta-kT), \tag{5}$$

then we have

$$P(\mathbf{y}_1 \mid \mathbf{x}) \propto \exp\left(-\frac{1}{2\sigma^2}\int_0^{NT+\Delta} |y_R(t) - s(t,\mathbf{x})|^2 dt\right), \quad (6)$$

where the proportionality notation "$\propto$" indicates that the two sides of (6) are proportional to each other and the proportional factor is independent from the symbol sequence $\mathbf{x}$. Thus, selecting $\hat{\mathbf{x}}$ from (4) is equivalent to choosing $\mathbf{x}$ which maximizes the right-hand side of (6). After decoding the individual packets, bit-wise XOR operation is applied to achieve the network-coded symbol $x_{R,k} = x_{A,k} \oplus x_{B,k}$.

2) THE BP-MAP METHOD

BP-MAP as introduced in [17] exploits the BP algorithm, one of the sum-product algorithms which is efficient for obtaining the marginal probability [18]. By using BP-MAP, we can find the exact marginal a posteriori probability of the k th pair of overlapped signals (i.e. $x_{A,k}$ and $x_{B,k}$), and then obtain the network-coded symbol $x_{R,k}$ directly, without decoding the individual symbols from nodes A and B.

Different from the MUD-XOR method, when BP-MAP is performed, the correlator integrates over the intervals Δ and $T-\Delta$ respectively, rather than over the whole symbol interval T. The outputs of the correlators are demonstrated by:

$$\begin{aligned} y_{R,2k-1} &= \frac{T}{\Delta}\int_{(k-1)T}^{(k-1)T+\Delta} y_R(t)\, p(t-kT)\, dt \\ &= x_{A,k} + x_{B,k-1}e^{j\phi} + \frac{T}{\Delta} \\ &\quad \times \int_{(k-1)T}^{(k-1)T+\Delta} n(t)\, p(t-kT) dt \end{aligned} \quad (7)$$

$$\begin{aligned} y_{R,2k} &= \frac{T}{T-\Delta}\int_{(k-1)T+\Delta}^{kT} y_R(t)\, p(t-\Delta-kT)\, dt \\ &= x_{A,k} + x_{B,k}e^{j\phi} + \frac{T}{T-\Delta} \\ &\quad \times \int_{(k-1)T+\Delta}^{kT} n(t)\, p(t-kT-\Delta) dt \end{aligned} \quad (8)$$

where $k \in \{1, 2, \ldots, N+1\}$, and $T/\Delta \cdot \int_{(k-1)T}^{(k-1)T+\Delta} p^2(t-kT)\, dt = 1$. Thus, $y_{R,2k-1}$ and $y_{R,2k}$ are complex Gaussian random variables respectively with variance $\sigma^2 \cdot T/\Delta$ and $\sigma^2 \cdot T/(T-\Delta)$, for both real and imaginary components. Obviously, $y_{R,2k-1}$ relates to the pair of $x_{A,k}$ and $x_{B,k-1}$, and $y_{R,2k}$ relates to the pair of $x_{A,k}$ and $x_{B,k}$, where $x_{B,0} = x_{A,N+1} = x_{B,N+1} = y_{R,2N+2} = 0$, as shown in Fig. 2.

Let $\mathbf{y}_2 = \left(y_{R,1}, y_{R,2}, \ldots, y_{R,2N+1}\right)$, the marginal a posteriori probability of the k th pair of overlapped signals (i.e. $x_{A,k}$ and $x_{B,k}$) is given by $P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right)$, which can be obtained by marginalizing all the other estimators in $P(\mathbf{x} \mid \mathbf{y}_2)$, namely,

$$P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right) = \sum_{\sim\{x_{A,k}, x_{B,k}\}} P(\mathbf{x} \mid \mathbf{y}_2), \quad (9)$$

where the notation $\sim \{x_{A,k}, x_{B,k}\}$ denotes the traversal of all $x_{A,i}, x_{B,i} \in \{-1, 1\}$ $(i = 1, 2, \ldots, N)$ except for $x_{A,k}$ and $x_{B,k}$. The exact expression of $P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right)$ can also be factorized as (9).

Thus, according to (9) and the iterative processes of BP algorithm, $P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right)$ can be rewritten as

$$P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right) = \mu_{2k-1} \cdot P\left(x_{A,k}, x_{B,k} \mid y_{R,2k}\right) \cdot \nu_{2k+1}, \quad (11)$$

where μ_{2k-1} and ν_{2k+1} are iterative processes (which can be seen in (11), given by Equations (12) and (13), as shown at the top of the next page.

The BP algorithm provides exact solution to $P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right)$ as expressed in (14). After computing $P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right)$, we can perform MAP estimation by

$$\begin{aligned} &P\left(x_{A,k}=1, x_{B,k}=1 \mid \mathbf{y}_2\right) \\ &\quad + P\left(x_{A,k}=-1, x_{B,k}=-1 \mid \mathbf{y}_2\right) \underset{\hat{x}_{R,k}=-1}{\overset{\hat{x}_{R,k}=1}{\gtrless}} \\ &P\left(x_{A,k}=-1, x_{B,k}=1 \mid \mathbf{y}_2\right) \\ &\quad + P\left(x_{A,k}=1, x_{B,k}=-1 \mid \mathbf{y}_2\right), \end{aligned} \quad (14)$$

which yields the estimated network-coded symbol $\hat{x}_{R,k}$.

$$\begin{aligned} &P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right) \\ &= \sum_{x_{A,1}}\sum_{x_{B,1}} \cdots \sum_{x_{A,k-1}}\sum_{x_{B,k-1}}\sum_{x_{A,k+1}}\sum_{x_{B,k+1}} \cdots \sum_{x_{A,N}}\sum_{x_{B,N}} \frac{P\left(x_{A,1}, x_{B,1}, \ldots, x_{A,k}, x_{B,k}, \ldots, x_{A,N}, x_{B,N}, \mathbf{y}_2\right)}{P(\mathbf{y}_2)} \\ &= \sum_{x_{B,k-1}} \left[P\left(x_{A,k}, x_{B,k-1} \mid y_{R,2k-1}\right) \cdots \sum_{x_{B,1}} \left(P\left(x_{A,2}, x_{B,1} \mid y_{R,3}\right) \sum_{x_{A,1}} \left(P\left(x_{A,1}, x_{B,1} \mid y_{R,2}\right) P\left(x_{A,1} \mid y_{R,1}\right) \right) \right) \right] \\ &\quad \cdot P\left(x_{A,k}, x_{B,k} \mid y_{R,2k}\right) \cdot \sum_{x_{A,k+1}} \left[P\left(x_{A,k+1}, x_{B,k} \mid y_{R,2k+1}\right) \right. \\ &\quad \left. \cdots \sum_{x_{A,N}} \left(P\left(x_{A,N}, x_{B,N-1} \mid y_{R,2N-1}\right) \sum_{x_{B,N}} \left(P\left(x_{A,N}, x_{B,N} \mid y_{R,2N}\right) P\left(x_{B,N} \mid y_{R,2N+1}\right) \right) \right) \right] \end{aligned} \quad (10)$$

$$\begin{aligned}\mu_{2k-1} &= P\left(x_{A,k} \mid y_{R,1}, y_{R,2}, \ldots, y_{R,2k-1}\right)\\ &= \sum_{x_{B,k-1}} \left[P\left(x_{A,k}, x_{B,k-1} \mid y_{R,2k-1}\right) \cdot P\left(x_{B,k-1} \mid y_{R,1}, y_{R,2}, \ldots, y_{R,2k-1}\right)\right] \quad (12)\\ \nu_{2k+1} &= P\left(x_{B,k} \mid y_{R,2k+1}, y_{R,2k+2}, \ldots, y_{R,2N+1}\right)\\ &= \sum_{x_{A,k+1}} \left[P\left(x_{A,k+1}, x_{B,k} \mid y_{R,2k+1}\right) \cdot P\left(x_{A,k+1} \mid y_{R,2k+1}, y_{R,2k+2}, \ldots, y_{R,2N+1}\right)\right] \quad (13)\\ P\left(\mathbf{y}_2 \mid \mathbf{x}\right) &= P\left(y_{R,1} \mid x_{A,1}\right) P\left(y_{R,2} \mid x_{A,1}, x_{B,1}\right) P\left(y_{R,3} \mid x_{A,2}, x_{B,1}\right) \cdots P\left(y_{R,2N+1} \mid x_{B,N}\right)\\ &\propto \exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{N+1}\left(\left|y_{R,2k} - x_{A,k} - x_{B,k}e^{j\phi}\right|^2 \frac{T-\Delta}{T} + \left|y_{R,2k-1} - x_{A,k} - x_{B,k-1}e^{j\phi}\right|^2 \frac{\Delta}{T}\right)\right) \quad (16)\end{aligned}$$

$$\begin{aligned}\int_0^{NT+\Delta} |s\left(t, \mathbf{x}\right)|^2 \, dt &= \int_0^{NT+\Delta} \left(\sum_{k=1}^{N} x_{A,k} p\left(t - kT\right) + x_{B,k} e^{j\phi} p\left(t - \Delta - kT\right)\right)\\ &\quad \cdot \left(\sum_{k=1}^{N} x_{A,k} p\left(t - kT\right) + x_{B,k} e^{-j\phi} p\left(t - \Delta - kT\right)\right) dt\\ &= \sum_{k=1}^{N}\left(x_{A,k}^2 + x_{B,k}^2 + 2 \cdot \frac{T-\Delta}{T} x_{A,k} x_{B,k} \cos\phi + 2 \cdot \frac{\Delta}{T} x_{A,k} x_{B,k-1} \cos\phi\right). \quad (17)\end{aligned}$$

$$\begin{aligned}&\left(\sum_{\sim\{x_{A,k}, x_{B,k}\}} P\left(\mathbf{y}_1 \mid \mathbf{x}\right)\right)\Big|_{x_{A,k}=1, x_{B,k}=1} + \left(\sum_{\sim\{x_{A,k}, x_{B,k}\}} P\left(\mathbf{y}_1 \mid \mathbf{x}\right)\right)\Big|_{x_{A,k}=-1, x_{B,k}=-1}\\ &\underset{\hat{x}_{R,k}=-1}{\overset{\hat{x}_{R,k}=1}{\gtrless}} \left(\sum_{\sim\{x_{A,k}, x_{B,k}\}} P\left(\mathbf{y}_1 \mid \mathbf{x}\right)\right)\Big|_{x_{A,k}=-1, x_{B,k}=1} + \left(\sum_{\sim\{x_{A,k}, x_{B,k}\}} P\left(\mathbf{y}_1 \mid \mathbf{x}\right)\right)\Big|_{x_{A,k}=1, x_{B,k}=-1}. \quad (19)\end{aligned}$$

3) RELATIONSHIP BETWEEN MUD-XOR AND BP-MAP

MUD-XOR cannot provide lower SER than BP-MAP, because MUD-XOR first decodes the individual symbols and then encodes them, whereas BP-MAP estimates the network-coded symbol directly. The relay finally wants to obtain the coded symbol and BP-MAP is directly optimized for the coded symbol, hence BP-MAP would not perform worse than MUD-XOR.

However, the performances of both approaches remain similar, which is shown as follows.

Lemma 1: When the received sequences $\mathbf{y}_1$ *and* $\mathbf{y}_2$ *are given, we have*

$$P\left(\mathbf{y}_2 \mid \mathbf{x}\right) \propto P\left(\mathbf{y}_1 \mid \mathbf{x}\right). \quad (15)$$

Proof: In the expansion of (16), as shown at the top of this page, the terms $|y_{R,2k-1}|^2\Delta/T$ and $\left|y_{R,2k}\right|^2 (T-\Delta)/T$ are independent from (14). We can omit these terms because when we estimate the values of $x_{R,k}$ from (14), these terms appear as constants because our received signals are given. According to (2), (3), (7) and (8), we have $y_{A,k} = y_{R,2k-1}\Delta/T + y_{R,2k}(T-\Delta)/T$ and $y_{B,k} = y_{R,2k}(T-\Delta)/T + y_{R,2k+1}\Delta/T$. Accordingly, their corresponding conjugate terms are also equal. We can also obtain Equation (17), as shown at the top of this page.

Hence, the right-hand side of (16) can be written as $\exp\left(\frac{1}{2\sigma^2}\int_0^{NT+\Delta} y_R^*(t)s(t,\mathbf{x}) + y_R(t)s^*(t,\mathbf{x}) - |s(t,\mathbf{x})|^2 dt\right)$ which is proportional to $P\left(\mathbf{y}_1 \mid \mathbf{x}\right)$, because we can omit the term $\frac{1}{2\sigma^2}\int_0^{NT+\Delta} |y_R(t)|^2 \, dt$ at the right-hand side of (6) when deciding the value of $\mathbf{x}$. Therefore, Lemma 1 is proved. ■

Claim 1: At high SNR, MUD-XOR and BP-MAP methods are approximately equivalent for estimating the network-coded symbol $x_{R,k}$.

Proof: For equiprobable symbols and a given received sequence $\mathbf{y}_2$, (9) can be rewritten as

$$P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right) \propto \sum_{\sim\{x_{A,k}, x_{B,k}\}} P\left(\mathbf{y}_2 \mid \mathbf{x}\right). \quad (18)$$

According to (15) and (18), the probability terms within (14), which is used for BP-MAP, can be replaced with $\sum_{\sim\{x_{A,k}, x_{B,k}\}} P\left(\mathbf{y}_1 \mid \mathbf{x}\right)$. Namely, (14) can be rewritten as (19), shown at the top of this page.

From (4) and (19), we can see that the MUD-XOR and BP-MAP methods relate to the common likelihood probability $P\left(\mathbf{y}_1 \mid \mathbf{x}\right)$.

We note that, for real numbers α and β, when $\alpha > \beta$, we have $\ln(\exp(\alpha) + \exp(\beta)) = \ln(\exp(\alpha) \cdot (1 + \exp(\beta - \alpha))) = \alpha + \ln(1 + \exp(-(\alpha - \beta)))$. The case is similar when $\beta > \alpha$. Hence, we have

$$\begin{aligned}&\ln\left(\exp\left(\alpha\right) + \exp\left(\beta\right)\right)\\ &\quad = \max\left\{\alpha, \beta\right\} + \ln\left(1 + \exp\left(-\left|\alpha - \beta\right|\right)\right). \quad (20)\end{aligned}$$

Therefore, $\ln(\exp(\alpha)+\exp(\beta)) \approx \max\{\alpha,\beta\}$ when $|\alpha-\beta|$ is large, which is also known as max-log approximation [19].

According to (6), the left- and right-hand sides of (19) are both proportional to the sum of exponential functions. It is also obvious that the exponent in these exponential functions have a common coefficient $1/2\sigma^2$. When the SNR is large, σ is small, and the difference between the exponents becomes large. In this case, we can use the max-log approximation to approximate the logarithm of both sides of (19).

Let vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{2^{2N-1}}$ denote possible transmitted sequences where the k th pair of overlapped signals $x_{A,k}$ and $x_{B,k}$ satisfies $x_{A,k}=1, x_{B,k}=1$ or $x_{A,k}=-1$, $x_{B,k}=-1$, and vectors $\mathbf{x}_{2^{2N-1}+1}, \mathbf{x}_{2^{2N-1}+2}, \ldots, \mathbf{x}_{2^{2N}}$ denote possible transmitted sequences where $x_{A,k}=-1, x_{B,k}=1$ or $x_{A,k}=1, x_{B,k}=-1$. After the max-log approximation, (19) can be written as

$$\begin{aligned}&\max\left\{\ln P\left(\mathbf{y}_1 \mid \mathbf{x}_1\right), \ln P\left(\mathbf{y}_1 \mid \mathbf{x}_2\right), \ldots, \ln P\left(\mathbf{y}_1 \mid \mathbf{x}_{2^{2N-1}}\right)\right\} \underset{\hat{x}_{R,k}=-1}{\overset{\hat{x}_{R,k}=1}{\gtrless}} \\ &\max\left\{\ln P\left(\mathbf{y}_1 \mid \mathbf{x}_{2^{2N-1}+1}\right), \ln P\left(\mathbf{y}_1 \mid \mathbf{x}_{2^{2N-1}+2}\right), \ldots, \ln P\left(\mathbf{y}_1 \mid \mathbf{x}_{2^{2N}}\right)\right\} \end{aligned} \tag{21}$$

Recall that MUD-XOR first performs (4) and then the XOR operation for the detected symbols, which is equivalent to the operation in (21). Because (21) approximates the BP-MAP method (particularly under high SNR), MUD-XOR and BP-MAP methods are approximately equivalent. ■

Also because MUD-XOR does not provide lower SER than BP-MAP (as discussed earlier), in the remaining part of this paper, we derive the lower bound from BP-MAP and the upper bound from MUD-XOR. The resulting bounds hold for asynchronous PNC that may perform either BP-MAP or MUD-XOR.

III. LOWER BOUND OF SER

This section derives an approximate lower SER bound for asynchronous PNC, by considering the BP-MAP method and assuming that part of the message is known to the relay. Such an assumption reduces the uncertainty of the symbols to be estimated. The approximate lower bound is confirmed to be correct in the simulations is Section V.

It can be observed in (11)-(13) that estimation regarding $x_{A,k}$ and $x_{B,k}$ relates to the certainty of adjacent symbols $x_{B,k-1}$ and $x_{A,k+1}$. To obtain the upper bound, when we focus on the error probability of $x_{R,k}=x_{A,k}\oplus x_{B,k}$, we assume that $x_{B,k-1}$ and $x_{A,k+1}$ are known, which reduces the uncertainty of $x_{A,k}$ and $x_{B,k}$. With this assumption, the marginal a posteriori probabilities[1] $P(x_{B,k-1} \mid y_{R,1}, y_{R,2}, \ldots, y_{R,2k-1}) = 1$ and $P\left(x_{A,k+1} \mid y_{R,2k+1}, y_{R,2k+2}, \ldots, y_{R,2N+1}\right) = 1$. Then, we have $\mu_{2k-1} = \sum_{x_{B,k-1}}[P\left(x_{A,k}, x_{B,k-1} \mid y_{R,2k-1}\right) \cdot P\left(x_{B,k-1} \mid y_{R,1}, y_{R,2}, \ldots, y_{R,2k-1}\right)] = P\left(x_{A,k}, x_{B,k-1} \mid y_{R,2k-1}\right)$ and $\nu_{2k+1} = P\left(x_{A,k+1}, x_{B,k} \mid y_{R,2k+1}\right)$. Thus, (11) can be written as

$$\begin{aligned}P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right) &= P\left(x_{A,k}, x_{B,k-1} \mid y_{R,2k-1}\right) \\ &\times P\left(x_{A,k}, x_{B,k} \mid y_{R,2k}\right) P\left(x_{A,k+1}, x_{B,k} \mid y_{R,2k+1}\right).\end{aligned} \tag{22}$$

We use the right-hand side of (22) to substitute the corresponding terms in the MAP decision rule given by (14), which is used for estimating $x_{R,k}$. However, it is not easy to directly obtain the analytical expression for the error probability of $x_{R,k}$. Hence, we approximate (14) with max-log approximation [19].

Lemma 2: Assuming that $x_{B,k-1}$ and $x_{A,k+1}$ are known, the MAP estimation can be written as follows after utilizing max-log approximation:

$$\begin{aligned}&\max\{r_1+r_2-\upsilon, -r_1-r_2-\upsilon\} \\ &\underset{\hat{x}_{R,k}=-1}{\overset{\hat{x}_{R,k}=1}{\gtrless}} \max\{-r_1+r_2+\upsilon, r_1-r_2+\upsilon\},\end{aligned} \tag{23}$$

where $r_1 = |y_{R,2k}| \cdot \cos\left(\angle y_{R,2k}\right) \cdot (T-\Delta)/T + |y_{R,2k-1}| \cos\left(\angle y_{R,2k-1}\right) \cdot \Delta/T - x_{B,k-1}\cos\phi \cdot \Delta/T$, $r_2 = |y_{R,2k}| \cdot \cos\left(\angle y_{R,2k}-\phi\right) \cdot (T-\Delta)/T + |y_{R,2k+1}| \cos\left(\angle y_{R,2k+1}-\phi\right) \cdot \Delta/T - x_{A,k+1}\cos\phi \cdot \Delta/T$, *and* $\upsilon = \cos\phi \cdot (T-\Delta)/T$. *The notation "$\angle$" indicates the argument of a complex number.*

Proof: We expand the right-hand side of (22) and omit the constant terms such as $\left|y_{R,2k}\right|^2$, $\left|y_{R,2k-1}-x_{B,k-1}e^{j\phi}\right|^2$, $\left|y_{R,2k+1}-x_{A,k+1}\right|^2$, as well as $\left|x_{A,k}\right|^2$ and $\left|x_{B,k}e^{j\phi}\right|^2$ because $\left|x_{A,k}\right|^2 = \left|x_{B,k}e^{j\phi}\right|^2 = 1$ for BPSK. Then we can simplify the right-hand side of (22) into (24), as shown at the top the next page.

Substituting the corresponding terms in (14) with (24), we have

$$\begin{aligned}&\exp\left(\frac{2r_1+2r_2-2\upsilon}{2\sigma^2}\right)+\exp\left(\frac{-2r_1-2r_2-2\upsilon}{2\sigma^2}\right) \underset{\hat{x}_{R,k}=-1}{\overset{\hat{x}_{R,k}=1}{\gtrless}} \\ &\exp\left(\frac{-2r_1+2r_2+2\upsilon}{2\sigma^2}\right)+\exp\left(\frac{2r_1-2r_2+2\upsilon}{2\sigma^2}\right).\end{aligned} \tag{25}$$

Similar to [12] which adopts max-log approximation and decision regions in a two-dimensional space to analyze the SER of synchronous PNC, we also adopt max-log approximation to approximate (25) and obtain (23). ■

Then we express (23) as decision regions in a two-dimensional space as shown in Fig. 3. Assuming $x_{A,k+1}$ and $x_{B,k-1}$ known reduces the uncertainty of $x_{A,k}$ and $x_{B,k}$ to be estimated, thus the analytical solution of the error probability of $x_{R,k}$ based on Fig. 3 is the lower bound of SER. It is obviously that the lower bound of SER can be obtained by integrating the joint probability density function (PDF) of r_1 and r_2 over the corresponding decision regions of Fig. 3.

[1]Note that the marginal a posteriori probabilities in iterative process such as μ_{2k-1} and ν_{2k+1} as shown in (12) and (13) are normalized, so that they can conform to the probability axioms, e.g., $\sum_{x_{A,k}}\mu_{2k-1}=1$ and $\sum_{x_{B,k}}\nu_{2k+1}=1$. Additionally, the likelihood functions for the a posteriori probabilities such as $P\left(x_{A,1}, x_{B,1} \mid y_{R,2}\right)$ as shown in (10), at the bottom of page 3, are normalized.

$$
\begin{aligned}
&P\left(x_{A,k}, x_{B,k} \mid \mathbf{y}_2\right) \\
&\propto \exp\left(-\frac{1}{2\sigma^2}\left(\left|y_{R,2k}-\left(x_{A,k}+x_{B,k}e^{j\phi}\right)\right|^2\cdot\frac{T-\Delta}{T}+\left|\left(y_{R,2k-1}-x_{B,k-1}e^{j\phi}\right)-x_{A,k}\right|^2\cdot\frac{\Delta}{T}\right.\right. \\
&\quad\left.\left.+\left|\left(y_{R,2k+1}-x_{A,k+1}\right)-x_{B,k}e^{j\phi}\right|^2\cdot\frac{\Delta}{T}\right)\right) \\
&\propto \exp\left[\frac{1}{2\sigma^2}\left(\left(y_{R,2k}\left(x_{A,k}+x_{B,k}e^{j\phi}\right)^*+y_{R,2k}^*\left(x_{A,k}+x_{B,k}e^{j\phi}\right)-\left|x_{A,k}+x_{B,k}e^{j\phi}\right|^2\right)\right.\right. \\
&\quad\cdot\frac{T-\Delta}{T}+\left(x_{A,k}\left(y_{R,2k-1}-x_{B,k-1}e^{j\phi}\right)^*+x_{A,k}^*\cdot\left(y_{R,2k-1}-x_{B,k-1}e^{j\phi}\right)-\left|x_{A,k}\right|^2\right) \\
&\quad\left.\left.\cdot\frac{\Delta}{T}+\left(x_{B,k}e^{j\phi}\left(y_{R,2k+1}-x_{A,k+1}\right)^*+\left(x_{B,k}e^{j\phi}\right)^*\cdot\left(y_{R,2k+1}-x_{A,k+1}\right)-\left|x_{B,k}e^{j\phi}\right|^2\right)\cdot\frac{\Delta}{T}\right)\right] \\
&\propto \exp\left[\frac{1}{2\sigma^2}\left(\left(x_{A,k}\cdot\left(y_{R,2k}^*+y_{R,2k}\right)+x_{B,k}\cdot\left(y_{R,2k}^*e^{j\phi}+y_{R,2k}e^{-j\phi}\right)-\left(x_{A,k}x_{B,k}e^{j\phi}+x_{A,k}x_{B,k}e^{-j\phi}\right)\right)\right.\right. \\
&\quad\cdot\frac{T-\Delta}{T}+\left(x_{A,k}\cdot\left(y_{R,2k-1}^*+y_{R,2k-1}\right)-x_{A,k}\cdot\left(x_{B,k-1}e^{j\phi}+x_{B,k-1}e^{-j\phi}\right)\right) \\
&\quad\left.\left.\cdot\frac{\Delta}{T}+\left(x_{B,k}\cdot\left(y_{R,2k+1}^*e^{j\phi}+y_{R,2k+1}e^{-j\phi}\right)-x_{B,k}\cdot\left(x_{A,k+1}e^{j\phi}+x_{A,k+1}e^{-j\phi}\right)\right)\cdot\frac{\Delta}{T}\right)\right] \\
&= \exp\left[\frac{1}{2\sigma^2}\left(x_{A,k}\cdot\left(2\cdot\frac{T-\Delta}{T}\left|y_{R,2k}\right|\cos\left(\angle y_{R,2k}\right)+2\cdot\frac{\Delta}{T}\left|y_{R,2k-1}\right|\cos\left(\angle y_{R,2k-1}\right)-2\cdot\frac{\Delta}{T}x_{B,k-1}\cos\phi\right)+x_{B,k}\right.\right. \\
&\quad\cdot\left(2\cdot\frac{T-\Delta}{T}\left|y_{R,2k}\right|\cos\left(\angle y_{R,2k}-\phi\right)+2\cdot\frac{\Delta}{T}\left|y_{R,2k+1}\right|\right. \\
&\quad\left.\left.\left.\cdot\cos\left(\angle y_{R,2k+1}-\phi\right)-2\cdot\frac{\Delta}{T}x_{A,k+1}\cos\phi\right)-2\cdot\frac{T-\Delta}{T}x_{A,k}x_{B,k}\cos\phi\right)\right] \\
&= \exp\left(\frac{2x_{A,k}r_1+2x_{B,k}r_2-2x_{A,k}x_{B,k}\upsilon}{2\sigma^2}\right). \qquad (24)
\end{aligned}
$$

$$
\begin{aligned}
\rho &= \frac{E\left[\left(r_1-E_{r_1}\right)\left(r_2-E_{r_2}\right)\right]}{\sigma^2} \\
&= \frac{1}{\sigma^2}E\left[\int_{(k-1)T}^{kT}p\left(t_1-kT\right)\left|n\left(t_1\right)\right|\cos\angle n\left(t_1\right)\,dt_1\cdot\int_{(k-1)T+\Delta}^{kT+\Delta}p\left(t_2-kT-\Delta\right)\left|n\left(t_2\right)\right|\cos\left(\angle n\left(t_2\right)-\phi\right)\,dt_2\right] \\
&= \frac{1}{\sigma^2}E\left[\frac{1}{2}\int_{(k-1)T}^{kT}p\left(t_1-kT\right)\left(n\left(t_1\right)+n^*\left(t_1\right)\right)\,dt_1\cdot\frac{1}{2}\int_{(k-1)T+\Delta}^{kT+\Delta}p\left(t_2-kT-\Delta\right)\left(n\left(t_2\right)e^{-j\phi}+n^*\left(t_2\right)e^{j\phi}\right)\,dt_2\right] \\
&= \frac{1}{4\sigma^2}\int_{(k-1)T}^{kT}\int_{(k-1)T+\Delta}^{kT+\Delta}p\left(t_1-kT\right)p\left(t_2-kT-\Delta\right)\left(E\left[n\left(t_1\right)n\left(t_2\right)e^{-j\phi}\right]+E\left[n^*\left(t_1\right)n\left(t_2\right)e^{-j\phi}\right]\right. \\
&\quad\left.+E\left[n\left(t_1\right)n^*\left(t_2\right)e^{j\phi}\right]+E\left[n^*\left(t_1\right)n^*\left(t_2\right)e^{j\phi}\right]\right)dt_1dt_2 \\
&= \frac{1}{4\sigma^2}\int_{(k-1)T}^{kT}\int_{(k-1)T+\Delta}^{kT+\Delta}p\left(t_1-kT\right)p\left(t_2-kT-\Delta\right)\left(0+2\sigma^2\delta\left(t_1-t_2\right)e^{-j\phi}+2\sigma^2\delta\left(t_1-t_2\right)e^{j\phi}+0\right)dt_1dt_2 \\
&= \frac{1}{2}\int_{(k-1)T+\Delta}^{kT}p\left(t-kT\right)p\left(t-kT-\Delta\right)\left(e^{-j\phi}+e^{j\phi}\right)dt=\frac{T-\Delta}{T}\cos\phi. \qquad (26)
\end{aligned}
$$

Lemma 3: *The joint PDF of r_1 and r_2 can be written as*

$$
P_r\left(r_1,r_2;\rho\right)=\frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\exp\left(-\frac{(r_1-E_{r_1})^2+(r_2-E_{r_2})^2-2\rho(r_1-E_{r_1})(r_2-E_{r_2})}{2\sigma^2(1-\rho^2)}\right). \qquad (27)
$$

where $E_{r_1}=x_{B,k}\cos\phi\cdot(T-\Delta)/T+x_{A,k}$ and $E_{r_2}=x_{A,k}\cos\phi\cdot(T-\Delta)/T+x_{B,k}$, are the mean values of r_1 and r_2, respectively; the σ^2 is the variance of r_1 and r_2; and $\rho=(T-\Delta)/T\cdot\cos\phi$ is the correlation between r_1 and r_2.

Proof: We can simplify r_1 and r_2 further. It is obviously that r_1 is related to the real components of $y_{R,2k-1}$ and $y_{R,2k}$ which are respectively given by (7) and (8).

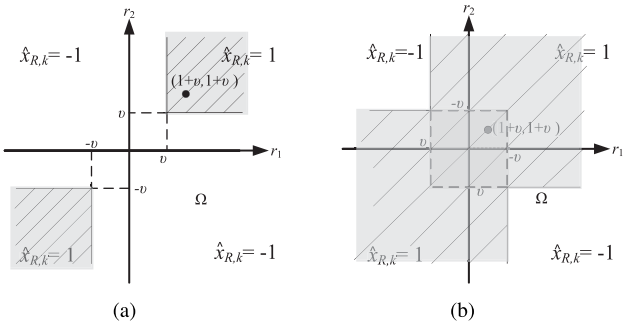


FIGURE 3. Geometric presentation of (23). The shaded areas denote the decision regions for the case that $x_{R,k}$ is estimated as 1 (i.e. $\hat{x}_{R,k} = 1$). (a) $\upsilon \geq 0$. (b) $\upsilon \leq 0$.

Therefore, according to the real components of (7) and (8), r_1 can be written as in (28), shown at the bottom of this page.

Likewise, we can obtain that $r_2 = x_{B,k} + x_{A,k} \cdot (T-\Delta)/T \cdot \cos\phi + \int_{(k-1)T+\Delta}^{kT+\Delta} p\,(t-kT-\Delta)\,|n\,(t)|\cos\left(\angle n\,(t)-\phi\right)\,dt$.

According to (28) where one of terms is Gaussian noise, r_1 and r_2 can be treated as real-valued Gaussian random variables. The mean values of r_1 and r_2 are $E_{r_1} = x_{B,k}\cos\phi \cdot (T-\Delta)/T + x_{A,k}$ and $E_{r_2} = x_{A,k}\cos\phi \cdot (T-\Delta)/T + x_{B,k}$, respectively; the variance of r_1 and r_2 is σ^2; the correlation ρ between r_1 and r_2 can be obtained by (26), as shown at the top of the previous page, where $\delta(t_1 - t_2)$ is Dirac delta function. Note that, since real and imaginary components of Gaussian noise are independent, in (26), $E[n(t_1)n(t_2)] = E[|n(t_1)|\cos(\angle n(t_1)) \cdot |n(t_2)|\cos(\angle n(t_2)) - |n(t_1)|\sin(\angle n(t_1)) \cdot |n(t_2)|\sin(\angle n(t_2))] = \sigma^2 \cdot \delta(t_1 - t_2) - \sigma^2 \cdot \delta(t_1 - t_2) = 0$. Likewise, $E[n^*(t_1)n^*(t_2)] = 0$, and $E[n^*(t_1)n(t_2)] = E[n(t_1)n^*(t_2)] = 2\sigma^2\delta(t_1 - t_2)$. ■

Subsequently, we integrate the joint probability density function of r_1 and r_2 over the corresponding decision regions of Fig. 3 to obtain the lower bound of SER. We need to consider the cases of $(x_{A,k}, x_{B,k})$, namely, $(1,1)$, $(-1,-1)$, $(-1,1)$ and $(1,-1)$, as well as the cases of $\upsilon \geq 0$ and $\upsilon \leq 0$.

Proposition 1: The approximate lower bound of SER can be written as

$$\begin{aligned} P_L &= \sum_{x_{A,k},x_{B,k}\in\{1,-1\}} P_L\left(x_{R,k} \neq \hat{x}_{R,k}, x_{A,k}, x_{B,k}\right) \\ &= \frac{1}{2} - \frac{1}{2}Q\left(-1,-1;\rho\right) - \frac{1}{2}Q\left(2\upsilon+1, 2\upsilon+1;\rho\right) \\ &\quad + Q\left(2\upsilon-1, 1;\rho\right), \end{aligned} \tag{29}$$

where $\rho = \upsilon = (T-\Delta)/T \cdot \cos\phi$ *and* $Q(x_1, y_1; \rho) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \int_{x_1}^{\infty}\int_{y_1}^{\infty} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2\sigma^2(1-\rho^2)}\right) dxdy$ *denotes the two-dimensional Gaussian probability integral.*

Proof: We first study the case that $\upsilon \geq 0$ and $(x_{A,k}, x_{B,k}) = (1,1)$ is transmitted. Thus the mean values of r_1 and r_2 are $E_{r_1} = E_{r_2} = 1 + (T-\Delta)/T \cdot \cos\phi$. The notation $\Omega_1 = \left\{\Omega : \hat{x}_{R,k} = 1\right\}$ is taken to represent the areas (the shaded areas in Fig. 3) to be mapped into network-coded symbol 1. Considering that symbols from end nodes are transmitted equiprobably, the error probability of $x_{R,k}$ is obtained by

$$\begin{aligned} &P_L\left(x_{R,k} \neq \hat{x}_{R,k}, x_{A,k} = 1, x_{B,k} = 1\right) \\ &= P\left(x_{A,k} = 1, x_{B,k} = 1\right)\left(1 - \iint_{\Omega_1} P_r\left(r_1, r_2; \rho\right) dr_1 dr_2\right) \\ &= \frac{1}{4}\left(1 - \left(\int_{-\infty}^{-\upsilon}\int_{-\infty}^{-\upsilon} + \int_{\upsilon}^{\infty}\int_{\upsilon}^{\infty}\right) P_r\left(r_1, r_2; \rho\right) dr_1 dr_2\right) \\ &= \frac{1}{4}\left(1 - Q\left(-1,-1;\rho\right) - Q\left(2\upsilon+1, 2\upsilon+1;\rho\right)\right). \end{aligned} \tag{30}$$

In a similar way, we obtain $P_L\left(x_{R,k} \neq \hat{x}_{R,k}, x_{A,k} = -1, x_{B,k} = -1\right)$, $P_L\left(x_{R,k} \neq \hat{x}_{R,k}, x_{A,k} = -1, x_{B,k} = 1\right)$, $P_L\left(x_{R,k} \neq \hat{x}_{R,k}, x_{A,k} = 1, x_{B,k} = -1\right)$. Then, we achieve (29).

Considering the case that $\upsilon \leq 0$, as shown in the Fig. 9, the result achieved is identical to (29). ■

Moreover, the two-dimensional Gaussian probability integral in (29) can be approximated by the method in [20].

IV. UPPER BOUND OF SER

In this section, we derive the analytical solution of upper bound P_U for asynchronous PNC, considering that the MUD-XOR method is used to decode the superposed symbols.

In [21] which aims to obtain the upper bound of SER for asynchronous Gaussian multiple-access channels like code division multiple access (CDMA), the author introduces the concepts of error vector and decomposition, then omits the redundant probabilities regarding the decomposable error vectors, and obtain the upper bound. In this paper, we also take use of the concepts of error vector and decomposition, but we derive the upper bound of SER for asynchronous PNC with our own way which is suitable for PNC and obtain its analytical result.

$$\begin{aligned} r_1 &= |y_{R,2k}|\cos\left(\angle y_{R,2k}\right) \cdot \frac{T-\Delta}{T} + |y_{R,2k-1}|\cos\left(\angle y_{R,2k-1}\right) \cdot \frac{\Delta}{T} - x_{B,k-1}\cos\phi \cdot \frac{\Delta}{T} \\ &= \left(x_{A,k} + x_{B,k}\cos\phi\right) \cdot \frac{T-\Delta}{T} + \int_{(k-1)T+\Delta}^{kT} p(t-kT-\Delta)\,|n\,(t)|\cos\angle n\,(t)\;dt + \left(x_{A,k} + x_{B,k-1}\cos\phi\right) \cdot \frac{\Delta}{T} \\ &\quad + \int_{(k-1)T}^{(k-1)T+\Delta} p(t-kT)\,|n\,(t)|\cos\angle n\,(t)\;dt - x_{B,k-1}\cos\phi \cdot \frac{\Delta}{T} \\ &= x_{A,k} + x_{B,k}\cos\phi \cdot \frac{T-\Delta}{T} + \int_{(k-1)T}^{kT} p\,(t-kT)\,|n\,(t)|\cos\angle n\,(t)\,dt. \end{aligned} \tag{28}$$

$$
\begin{aligned}
&E\Bigg[\frac{1}{2\sigma^2}\int_0^{NT+\Delta}\left(n^*(t_1)\,s\left(t_1,\hat{\mathbf{x}}-\mathbf{x}\right)+n(t_1)\,s^*\left(t_1,\hat{\mathbf{x}}-\mathbf{x}\right)\right)dt_1\\
&\quad\cdot\frac{1}{2\sigma^2}\int_0^{NT+\Delta}\left(n^*(t_2)\,s\left(t_2,\hat{\mathbf{x}}-\mathbf{x}\right)+n(t_2)\,s^*\left(t_2,\hat{\mathbf{x}}-\mathbf{x}\right)\right)dt_2\Bigg]\\
&=\frac{1}{4\sigma^4}\int_0^{NT+\Delta}\int_0^{NT+\Delta}\Big(E[n^*(t_1)\,n^*(t_2)]s\left(t_1,\hat{\mathbf{x}}-\mathbf{x}\right)s\left(t_2,\hat{\mathbf{x}}-\mathbf{x}\right)+E[n^*(t_1)\,n(t_2)]s\left(t_1,\hat{\mathbf{x}}-\mathbf{x}\right)s^*\left(t_2,\hat{\mathbf{x}}-\mathbf{x}\right)\\
&\quad+E[n(t_1)\,n^*(t_2)]s^*\left(t_1,\hat{\mathbf{x}}-\mathbf{x}\right)\cdot s\left(t_2,\hat{\mathbf{x}}-\mathbf{x}\right)+E[n(t_1)\,n(t_2)]s^*\left(t_1,\hat{\mathbf{x}}-\mathbf{x}\right)s^*\left(t_2,\hat{\mathbf{x}}-\mathbf{x}\right)\Big)\,dt_1dt_2\\
&=\frac{1}{4\sigma^4}\int_0^{NT+\Delta}\int_0^{NT+\Delta}\left(2\sigma^2\cdot\delta(t_1-t_2)\,s\left(t_1,\hat{\mathbf{x}}-\mathbf{x}\right)s^*\left(t_2,\hat{\mathbf{x}}-\mathbf{x}\right)+2\sigma^2\cdot\delta(t_1-t_2)\,s^*\left(t_1,\hat{\mathbf{x}}-\mathbf{x}\right)s\left(t_2,\hat{\mathbf{x}}-\mathbf{x}\right)\right)dt_1dt_2\\
&=\frac{1}{\sigma^2}\int_0^{NT+\Delta}\left|s\left(t,\hat{\mathbf{x}}-\mathbf{x}\right)\right|^2dt. \qquad (31)
\end{aligned}
$$

The k th network-coded symbol is detected erroneous (i.e. $x_{R,k}\neq\hat{x}_{R,k}$), only when either $\hat{x}_{A,k}$ or $\hat{x}_{B,k}$ in the detected sequence $\hat{\mathbf{x}}$ is wrong due to the property of XOR operation: $x_{A,k}\oplus\overline{x_{B,k}}=\overline{x_{A,k}\oplus x_{B,k}}$, while, the other elements in $\hat{\mathbf{x}}$ can be right or wrong. In the following part, we use $\mathcal{J}_{\mathbf{x},k}$ to denote the set of all possible estimations of $\mathbf{x}$, in which the k th pair of $\hat{x}_{A,k}$ and $\hat{x}_{B,k}$ satisfies $\hat{x}_{A,k}\neq x_{A,k},\hat{x}_{B,k}=x_{B,k}$ or $\hat{x}_{A,k}=x_{A,k},\hat{x}_{B,k}\neq x_{B,k}$. Thus, for certain transmitted sequence $\mathbf{x}$, the probability of the k th estimation $\hat{x}_{R,k}$ being erroneous can be expressed as

$$
P\left(\hat{x}_{R,k}\neq x_{R,k}\mid\mathbf{x}\right)=\sum_{\hat{\mathbf{x}}\in\mathcal{J}_{\mathbf{x},k}}P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right). \qquad (32)
$$

Let the set $\left\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\right\}$ denote all possible transmitted sequences. Considering all kinds of transmitted sequence, the SER $P_{e,k}$ for $x_{R,k}$ can be achieved by

$$
\begin{aligned}
P_{e,k}&=\sum_{\mathbf{x}\in\left\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\right\}}P(\mathbf{x})\,P\left(\hat{x}_{R,k}\neq x_{R,k}\mid\mathbf{x}\right)\\
&=\sum_{\mathbf{x}\in\left\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\right\}}P(\mathbf{x})\sum_{\hat{\mathbf{x}}\in\mathcal{J}_{\mathbf{x},k}}P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right), \qquad (33)
\end{aligned}
$$

where $P(\mathbf{x})=1/2^{2N}$, because all the transmitted symbols are independent and equiprobable.

According to (33), it is observed that $P_{e,k}$ is related to the term $P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right)$. Therefore, in this paper, the idea for deriving P_U is based on the upper bound of $P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right)$. In detail, firstly, we obtain the upper bound of $P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right)$, then, traversing $\mathbf{x}\in\left\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\right\}$ and $\hat{\mathbf{x}}\in\mathcal{J}_{\mathbf{x},k}$, we accumulate upper bounds of all kinds of $P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right)$. However, some of those upper bounds are redundant, which is analyzed in IV-A. Finally, we omit the redundancy, and obtain the analytical result of P_U.

The upper bound of $P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right)$ can be derived as follows. Let $\mathcal{L}(\mathbf{x})=\ln P(\mathbf{y}_1\mid\mathbf{x})$, according to (4), the upper bound of $P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right)$ is achieved by

$$
\begin{aligned}
P\left(\hat{\mathbf{x}}\mid\mathbf{x}\right)&=P\left(\hat{\mathbf{x}}=\arg\max_{\mathbf{x}'\in\left\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\right\}}\mathcal{L}\left(\mathbf{x}'\right)\,\middle|\,\mathbf{x}\right)\\
&=P\left(\mathcal{L}\left(\hat{\mathbf{x}}\right)\geq\mathcal{L}(\mathbf{x}_1),\mathcal{L}\left(\hat{\mathbf{x}}\right)\right.\\
&\qquad\left.\geq\mathcal{L}(\mathbf{x}_2),\ldots,\mathcal{L}\left(\hat{\mathbf{x}}\right)\geq\mathcal{L}\left(\mathbf{x}_{2^{2N}}\right)\mid\mathbf{x}\right)\\
&\leq P\left(\mathcal{L}\left(\hat{\mathbf{x}}\right)\geq\mathcal{L}(\mathbf{x})\mid\mathbf{x}\right). \qquad (34)
\end{aligned}
$$

Lemma 4: The value of $P\left(\mathcal{L}\left(\hat{\mathbf{x}}\right)\geq\mathcal{L}(\mathbf{x})\mid\mathbf{x}\right)$ is obtained by

$$
P\left(\mathcal{L}\left(\hat{\mathbf{x}}\right)\geq\mathcal{L}(\mathbf{x})\mid\mathbf{x}\right)=Q\left(\frac{S\left(\hat{\mathbf{x}}-\mathbf{x}\right)}{2\sigma}\right), \qquad (35)
$$

where $Q(\cdot)$ denotes one-dimensional Gaussian probability integral and $S\left(\hat{\mathbf{x}}-\mathbf{x}\right)=\sqrt{\int_0^{NT+\Delta}\left|s\left(t,\hat{\mathbf{x}}-\mathbf{x}\right)\right|^2dt}$.

Proof: According to (6), we can obtain

$$
\begin{aligned}
&\mathcal{L}\left(\hat{\mathbf{x}}\right)-\mathcal{L}(\mathbf{x})=\ln P\left(\mathbf{y}_1\mid\hat{\mathbf{x}}\right)-\ln P(\mathbf{y}_1\mid\mathbf{x})\\
&=\frac{1}{2\sigma^2}\int_0^{NT+\Delta}\left|y_R(t)-s(t,\mathbf{x})\right|^2dt\\
&\quad-\frac{1}{2\sigma^2}\int_0^{NT+\Delta}\left|y_R(t)-s\left(t,\hat{\mathbf{x}}\right)\right|^2dt\\
&=\frac{1}{2\sigma^2}\int_0^{NT+\Delta}\Big(s(t,\mathbf{x})\,s^*(t,\mathbf{x})-y_R(t)\,s^*(t,\mathbf{x})\\
&\qquad-y_R^*(t)\,s(t,\mathbf{x})+y_R(t)\,s^*\left(t,\hat{\mathbf{x}}\right)\\
&\qquad+y_R^*(t)\,s\left(t,\hat{\mathbf{x}}\right)-s\left(t,\hat{\mathbf{x}}\right)s^*\left(t,\hat{\mathbf{x}}\right)\Big)dt.\\
&=\frac{1}{2\sigma^2}\int_0^{NT+\Delta}\Big(n^*(t)\,s\left(t,\hat{\mathbf{x}}-\mathbf{x}\right)+n(t)\,s^*\left(t,\hat{\mathbf{x}}-\mathbf{x}\right)\\
&\qquad-\left|s\left(t,\hat{\mathbf{x}}-\mathbf{x}\right)\right|^2\Big)\,dt, \qquad (36)
\end{aligned}
$$

where $s\left(t,\hat{\mathbf{x}}-\mathbf{x}\right)=s\left(t,\hat{\mathbf{x}}\right)-s(t,\mathbf{x})$. Note that in (36), we substitute $y_R(t)$ with $s(t,\mathbf{x})+n(t)$ based on (1) and (5). Since $\mathcal{L}\left(\hat{\mathbf{x}}\right)-\mathcal{L}(\mathbf{x})$ includes Gaussian noise, it is a real-valued Gaussian random variable and the mean value is

$-1/2\sigma^2 \cdot \int_0^{NT+\Delta} \left|s\left(t, \hat{\mathbf{x}}-\mathbf{x}\right)\right|^2 dt$. Its variance is computed by (31), as shown at the top of the previous page, where $E[n(t_1)\, n(t_2)] = E[n^*(t_1)\, n^*(t_2)] = 0$ and $E[n^*(t_1)\, n(t_2)] = E[n(t_1)\, n^*(t_2)] = 2\sigma^2\delta(t_1 - t_2)$ referring to the analysis for (26).

Therefore, we can obtain the probability density function (PDF) of $\mathcal{L}\left(\hat{\mathbf{x}}\right) - \mathcal{L}(\mathbf{x})$, then integrate the PDF over the integral $[0, \infty)$ to obtain (35). ■

A. UPPER BOUND FOR SER

After obtaining the upper bound of $P\left(\hat{\mathbf{x}} \mid \mathbf{x}\right)$ by (35), according to (33), we can obtain

$$P_{e,k} \leq \frac{1}{2^{2N}} \sum_{\mathbf{x}\in\left\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\right\}} \sum_{\hat{\mathbf{x}}\in\mathcal{J}_{\mathbf{x},k}} Q\left(\frac{S\left(\hat{\mathbf{x}}-\mathbf{x}\right)}{2\sigma}\right). \tag{37}$$

However, in this subsection, we prove that some terms in (37) are redundant.

Prior to further discussions, we first introduce the concept of error vector, which records the difference between the detected sequence and transmitted sequence, i.e. the value of $\hat{\mathbf{x}} - \mathbf{x}$. In the following parts, we use $\boldsymbol{\lambda} = \left(\lambda_{A,1}, \lambda_{B,1}, \ldots, \lambda_{A,N}, \lambda_{B,N}\right)$ to denote a certain error vector, where $\lambda_{A,i}, \lambda_{B,i} \in \{-2, 0, 2\}$ and $i \in \{1, 2, \ldots, N\}$. With the concept of error vectors, according to (33) and (37), the SER can be written as:

$$\begin{aligned} P_{e,k} &= \frac{1}{2^{2N}} \sum_{\mathbf{x}\in\left\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\right\}} \sum_{\boldsymbol{\lambda}:\mathbf{x}+\boldsymbol{\lambda}\in\mathcal{J}_{\mathbf{x},k}} P(\mathbf{x}+\boldsymbol{\lambda} \mid \mathbf{x}) \\ &\leq \frac{1}{2^{2N}} \sum_{\mathbf{x}\in\left\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\right\}} \sum_{\boldsymbol{\lambda}:\mathbf{x}+\boldsymbol{\lambda}\in\mathcal{J}_{\mathbf{x},k}} Q\left(\frac{S(\boldsymbol{\lambda})}{2\sigma}\right) \end{aligned} \tag{38}$$

We then define a certain $\boldsymbol{\lambda}$ as a decomposable error vector (which can be decomposed into the other two error vectors $\boldsymbol{\lambda}'$ and $\boldsymbol{\lambda}''$), provided that $\boldsymbol{\lambda}$ satisfies

$$\begin{cases} \boldsymbol{\lambda} = \boldsymbol{\lambda}' + \boldsymbol{\lambda}'' & \text{(39a)} \\ \text{if } \lambda_{A,i} = 0, \text{ then } \lambda'_{A,i} = \lambda''_{A,i} = 0 & \text{(39b)} \\ \text{if } \lambda_{B,i} = 0, \text{ then } \lambda'_{B,i} = \lambda''_{B,i} = 0 & \text{(39c)} \\ f(\boldsymbol{\lambda}', \boldsymbol{\lambda}'') \triangleq \int_0^{NT+\Delta} \left(s^*\left(t, \boldsymbol{\lambda}'\right) s\left(t, \boldsymbol{\lambda}''\right)\right. & \\ \left. + s\left(t, \boldsymbol{\lambda}'\right) s^*\left(t, \boldsymbol{\lambda}''\right)\right) dt \geq 0 & \text{(39d)} \end{cases}$$

where $s(t, \boldsymbol{\lambda}) = \sum_{k=1}^{N} \lambda_{A,k} p(t - kT) + \lambda_{B,k} e^{j\phi} p(t - \Delta - kT)$.

According to (38), we can obtain that the SER of k th network-coded symbol $x_{R,k}$ only relates to the error vectors in the set $\left\{\boldsymbol{\lambda} : \mathbf{x} + \boldsymbol{\lambda} \in \mathcal{J}_{\mathbf{x},k}\right\}$ given certain transmitted vector $\mathbf{x}$, thus, the k th pair of elements in $\boldsymbol{\lambda}$ (i.e., $\lambda_{A,k}$ and $\lambda_{B,k}$) satisfies $\lambda_{A,k} \neq 0, \lambda_{B,k} = 0$ or $\lambda_{A,k} = 0, \lambda_{B,k} \neq 0$ and other elements can be any values in $\{-2, 0, 2\}$. It is observed in (38) that all the error vectors $\boldsymbol{\lambda}$, which satisfies $\mathbf{x}+\boldsymbol{\lambda} \in \mathcal{J}_{\mathbf{x},k}$, are traversed, including decomposable and indecomposable vectors when we aim to obtain the upper bound of $P_{e,k}$. However, the following Lemma 5 and 6 reveal that the upper bound of $P_{e,k}$ only relates to the terms regarding indecomposable error vectors at the right-hand side of (38), and the terms regarding decomposable error vectors are redundant.

Lemma 5: A certain decomposable error vector $\boldsymbol{\lambda}$ *satisfying* $\mathbf{x}+\boldsymbol{\lambda} \in \mathcal{J}_{\mathbf{x},k}$ *can be decomposed into an indecomposable error vector* $\boldsymbol{\lambda}^{in}$ *and another error vector* $(\boldsymbol{\lambda} - \boldsymbol{\lambda}^{in})$*, where* $\boldsymbol{\lambda}^{in}$ *satisfies* $\mathbf{x} + \boldsymbol{\lambda}^{in} \in \mathcal{J}_{\mathbf{x},k}$.

Proof: We first discuss that if we decompose $\boldsymbol{\lambda}$ step by step, we can obtain an indecomposable error vector at last. Then, we will prove that $\boldsymbol{\lambda}$ can be directly decomposed into the indecomposable error vector and another error vector.

If $\boldsymbol{\lambda}$ is decomposable, there are at least one kind of decomposition, for example, $\boldsymbol{\lambda}$ can be decomposed into $\boldsymbol{\lambda}^a + \boldsymbol{\lambda}^b$ or $\boldsymbol{\lambda}' + \boldsymbol{\lambda}''$, etc. Thus, we can always find one of aforementioned decompositions to maximize the value of $f(\cdot)$ in (39a). In the following discussion, we assume that the decomposition (not necessarily unique) $\boldsymbol{\lambda} = \boldsymbol{\lambda}^a + \boldsymbol{\lambda}^b$ can maximize the value of $f(\cdot)$, that is

$$f(\boldsymbol{\lambda}^a, \boldsymbol{\lambda}^b) \geq f(\boldsymbol{\lambda}', \boldsymbol{\lambda}'') \geq 0 \tag{40}$$

for any other decomposition $\boldsymbol{\lambda} = \boldsymbol{\lambda}' + \boldsymbol{\lambda}''$. Note that only one of $\mathbf{x} + \boldsymbol{\lambda}^a$ and $\mathbf{x} + \boldsymbol{\lambda}^b$ will belong to $\mathcal{J}_{\mathbf{x},k}$, since if $\lambda_{A,k} \neq 0, \lambda_{B,k} = 0$, it can be only decomposed into $\lambda^a_{A,k} \neq 0, \lambda^a_{B,k} = 0$ and $\lambda^b_{A,k} = 0, \lambda^b_{B,k} = 0$ $(\mathbf{x}+\boldsymbol{\lambda}^a \in \mathcal{J}_k)$, or $\lambda^a_{A,k} = 0, \lambda^a_{B,k} = 0$ and $\lambda^b_{A,k} \neq 0, \lambda^b_{B,k} = 0$ $(\mathbf{x}+\boldsymbol{\lambda}^b \in \mathcal{J}_k)$. We assume $\mathbf{x} + \boldsymbol{\lambda}^a \in \mathcal{J}_k$. Then, if $\boldsymbol{\lambda}^a$ is indecomposable (i.e., $f\left(\boldsymbol{\lambda}^a_1, \boldsymbol{\lambda}^a_2\right) < 0$ for any decomposition $\boldsymbol{\lambda}^a = \boldsymbol{\lambda}^a_1 + \boldsymbol{\lambda}^a_2$), it is obtained that $\boldsymbol{\lambda}^{\text{in}} = \boldsymbol{\lambda}^a$. Otherwise, $\boldsymbol{\lambda}^a$ can be decomposed into $\boldsymbol{\lambda}^{a_1}$ and $\boldsymbol{\lambda}^{a_2}$ which maximize the value of corresponding $f(\cdot)$, and $\mathbf{x} + \boldsymbol{\lambda}^{a_1} \in \mathcal{J}_{\mathbf{x},k}$. If $\boldsymbol{\lambda}^{a_1}$ is indecomposable, we stop decomposing because we have found an indecomposable error vector. Otherwise, $\boldsymbol{\lambda}^{a_1}$ can be decomposed into $\boldsymbol{\lambda}^{a_{11}}$ and $\boldsymbol{\lambda}^{a_{12}}$ which maximize the value of corresponding $f(\cdot)$, and $\mathbf{x} + \boldsymbol{\lambda}^{a_{11}} \in \mathcal{J}_{\mathbf{x},k}$. If $\boldsymbol{\lambda}^{a_{11}}$ is indecomposable, we stop decomposing. Otherwise, we can continue to decompose $\boldsymbol{\lambda}^{a_{11}}$. At last, we will obtain an indecomposable error vector, otherwise, the aforementioned process will continue until forever.

After obtaining an indecomposable error vector, we prove that $\boldsymbol{\lambda}$ can be directly decomposed into the indecomposable error vector and another error vector. Assuming $\boldsymbol{\lambda}^a$ and $\boldsymbol{\lambda}^{a_1}$ decomposable, firstly, we prove that $\boldsymbol{\lambda}$ can be directly decomposed into $\boldsymbol{\lambda}^{a_1}$ and $\left(\boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^b\right)$. According to the additivity of $f(\cdot)$, and $\boldsymbol{\lambda}^a = \boldsymbol{\lambda}^{a_1} + \boldsymbol{\lambda}^{a_2}$, we achieve the following equation,

$$\begin{aligned} &f(\boldsymbol{\lambda}^b + \boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2}) - f(\boldsymbol{\lambda}^a, \boldsymbol{\lambda}^b) \\ &= f(\boldsymbol{\lambda}^b + \boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2}) - f(\boldsymbol{\lambda}^{a_1} + \boldsymbol{\lambda}^{a_2}, \boldsymbol{\lambda}^b) \\ &= f(\boldsymbol{\lambda}^b, \boldsymbol{\lambda}^{a_2}) + f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2}) - \left(f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^b) + f(\boldsymbol{\lambda}^{a_2}, \boldsymbol{\lambda}^b)\right) \\ &= f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2}) - f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^b) \\ &= 2f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2}) - f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^b). \end{aligned} \tag{41}$$

Since $\boldsymbol{\lambda}^a$ can be directly decomposed into $\boldsymbol{\lambda}^{a_1}$ and $\boldsymbol{\lambda}^{a_2}$, thus $f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2}) \geq 0$. According to (41), if $\boldsymbol{\lambda}$ cannot be directly decomposed into $\boldsymbol{\lambda}^{a_1}$ and $\left(\boldsymbol{\lambda}^b + \boldsymbol{\lambda}^{a_2}\right)$, namely, $f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^b) < 0$, thus $f(\boldsymbol{\lambda}^b + \boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2}) - f(\boldsymbol{\lambda}^a, \boldsymbol{\lambda}^b) = 2f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2}) - f(\boldsymbol{\lambda}^{a_1}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^b) > 0$ which contradicts with choosing

$\boldsymbol{\lambda} = \boldsymbol{\lambda}^a + \boldsymbol{\lambda}^b$ as the decomposition maximizing the value of corresponding $f(\cdot)$. Therefore, $\boldsymbol{\lambda}$ can be directly decomposed into $\boldsymbol{\lambda}^{a_1}$ and $\left(\boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^b\right)$, likewise, $\boldsymbol{\lambda}^a$ can be directly decomposed into $\boldsymbol{\lambda}^{a_{11}}$ and $(\boldsymbol{\lambda}^{a_{12}} + \boldsymbol{\lambda}^{a_2})$. Secondly, we prove that $\boldsymbol{\lambda}$ can be directly decomposed into $\boldsymbol{\lambda}^{a_{11}}$ and $\left(\boldsymbol{\lambda}^b + \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^{a_{12}}\right)$. According to $\boldsymbol{\lambda} = \boldsymbol{\lambda}^a + \boldsymbol{\lambda}^b$, $\boldsymbol{\lambda}^a = \boldsymbol{\lambda}^{a_1} + \boldsymbol{\lambda}^{a_2}$, and $\boldsymbol{\lambda}^{a_1} = \boldsymbol{\lambda}^{a_{11}} + \boldsymbol{\lambda}^{a_{12}}$, we achieve that

$$\begin{aligned} &f(\boldsymbol{\lambda}^b + \boldsymbol{\lambda}^{a_{11}}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^{a_{12}}) - f(\boldsymbol{\lambda}^a, \boldsymbol{\lambda}^b) \\ &= f(\boldsymbol{\lambda}^b + \boldsymbol{\lambda}^{a_{11}}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^{a_{12}}) - f(\boldsymbol{\lambda}^{a_{11}} + \boldsymbol{\lambda}^{a_{12}} + \boldsymbol{\lambda}^{a_2}, \boldsymbol{\lambda}^b) \\ &= f(\boldsymbol{\lambda}^{a_{11}}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^{a_{12}}) - f(\boldsymbol{\lambda}^{a_{11}}, \boldsymbol{\lambda}^b) \\ &= 2f(\boldsymbol{\lambda}^{a_{11}}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^{a_{12}}) - f(\boldsymbol{\lambda}^{a_{11}}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^{a_{12}} + \boldsymbol{\lambda}^b). \quad (42) \end{aligned}$$

Because we have proved that $\boldsymbol{\lambda}^a$ can be directly decomposed into $\boldsymbol{\lambda}^{a_{11}}$ and $(\boldsymbol{\lambda}^{a_{12}} + \boldsymbol{\lambda}^{a_2})$, thus $f(\boldsymbol{\lambda}^{a_{11}}, \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^{a_{12}}) \geq 0$. Similar to the analysis for (41), we can prove that $\boldsymbol{\lambda}$ indeed can be directly decomposed into $\boldsymbol{\lambda}^{a_{11}}$ and $\left(\boldsymbol{\lambda}^b + \boldsymbol{\lambda}^{a_2} + \boldsymbol{\lambda}^{a_{12}}\right)$. If $\boldsymbol{\lambda}^{a_{11}}$ is indecomposable, we can obtain that $\boldsymbol{\lambda}^{\text{in}} = \boldsymbol{\lambda}^{a_{11}}$. According to induction hypothesis and above analysis, we can prove that $\boldsymbol{\lambda}$ can be directly decomposed into an indecomposable error vector $\boldsymbol{\lambda}^{\text{in}}$ and another error vector. ■

Lemma 6: If decomposable error vectors $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_m \in \left\{\boldsymbol{\lambda} : \mathbf{x} + \boldsymbol{\lambda} \in \mathcal{J}_{\mathbf{x},k}\right\}$ can be respectively decomposed into a common indecomposable error vector $\boldsymbol{\lambda}^{in}$ (where $\mathbf{x} + \boldsymbol{\lambda}^{in} \in \mathcal{J}_{\mathbf{x},k}$) and corresponding vectors $(\boldsymbol{\lambda}_i - \boldsymbol{\lambda}^{in})$ (where $i = 1, 2, \ldots, m$), it is obtained that

$$P\left(\mathbf{x} + \boldsymbol{\lambda}^{in} \mid \mathbf{x}\right) + \sum_{i=1}^{m} P\left(\mathbf{x} + \boldsymbol{\lambda}_i \mid \mathbf{x}\right) \leq Q\left(\frac{S\left(\boldsymbol{\lambda}^{in}\right)}{2\sigma}\right). \quad (43)$$

Proof: We first prove that when $\mathbf{x} + \boldsymbol{\lambda}^i$ is detected, we can also obtain $\mathcal{L}\left(\mathbf{x} + \boldsymbol{\lambda}^{\text{in}}\right) \geq \mathcal{L}(\mathbf{x})$. Thus, on a constellation diagram, the decision region where $\mathbf{x} + \boldsymbol{\lambda}^{\text{in}}$ or $\mathbf{x} + \boldsymbol{\lambda}_i$ is detected, is within the decision region where $\mathcal{L}\left(\mathbf{x} + \boldsymbol{\lambda}^{\text{in}}\right) \geq \mathcal{L}(\mathbf{x})$. Then, based on the above conclusion, we derive (43).

For the i th decomposable error vector $\boldsymbol{\lambda}_i$, let $\boldsymbol{\lambda}_i' = \boldsymbol{\lambda}_i - \boldsymbol{\lambda}^{\text{in}}$. According to (36), we can obtain that

$$\begin{aligned} &\mathcal{L}\left(\mathbf{x} + \boldsymbol{\lambda}^{\text{in}}\right) - \mathcal{L}(\mathbf{x}) \\ &= \frac{1}{2\sigma^2}\int_0^{NT+\Delta} \left(n^*(t)\, s\left(t, \boldsymbol{\lambda}^{\text{in}}\right) + n(t)\, s^*\left(t, \boldsymbol{\lambda}^{\text{in}}\right) \right. \\ &\quad \left. - \left|s\left(t, \boldsymbol{\lambda}^{\text{in}}\right)\right|^2\right) dt \\ &= \frac{1}{2\sigma^2}\int_0^{NT+\Delta} \left(n^*(t)\, s\left(t, \boldsymbol{\lambda}_i - \boldsymbol{\lambda}_i'\right) + n(t)\, s^*\left(t, \boldsymbol{\lambda}_i - \boldsymbol{\lambda}_i'\right) \right. \\ &\quad \left. - \left|s\left(t, \boldsymbol{\lambda}_i - \boldsymbol{\lambda}_i'\right)\right|^2\right) dt \\ &= \frac{1}{2\sigma^2}\int_0^{NT+\Delta} \left(\left(n^*(t)s(t, \boldsymbol{\lambda}_i) + n(t)s^*(t, \boldsymbol{\lambda}_i)\right| - s(t, \boldsymbol{\lambda}_i)|^2\right) \\ &\quad - \left(n^*(t)\, s\left(t, \boldsymbol{\lambda}_i'\right) + n(t)\, s^*\left(t, \boldsymbol{\lambda}_i'\right) + \left|s\left(t, \boldsymbol{\lambda}_i'\right)\right|^2\right) \\ &\quad + s^*\left(t, \boldsymbol{\lambda}^{\text{in}} + \boldsymbol{\lambda}_i'\right) s\left(t, \boldsymbol{\lambda}_i'\right) + s\left(t, \boldsymbol{\lambda}^{\text{in}} + \boldsymbol{\lambda}_i'\right) s^*\left(t, \boldsymbol{\lambda}_i'\right)\Big) dt \\ &= \frac{1}{2\sigma^2}\int_0^{NT+\Delta} \left(n^*(t)s(t, \boldsymbol{\lambda}_i) + n(t)s^*(t, \boldsymbol{\lambda}_i) - |s(t, \boldsymbol{\lambda}_i)|^2\right) \\ &\quad - \left(n^*(t)\, s\left(t, \boldsymbol{\lambda}_i'\right) + n(t)\, s^*\left(t, \boldsymbol{\lambda}_i'\right) - \left|s\left(t, \boldsymbol{\lambda}_i'\right)\right|^2\right) \\ &\quad + s^*\left(t, \boldsymbol{\lambda}^{\text{in}}\right) s\left(t, \boldsymbol{\lambda}_i'\right) + s\left(t, \boldsymbol{\lambda}^{\text{in}}\right) s^*\left(t, \boldsymbol{\lambda}_i'\right)\Big) dt \\ &= \left[\mathcal{L}(\mathbf{x} + \boldsymbol{\lambda}_i) - \mathcal{L}(\mathbf{x})\right] - \left[\mathcal{L}\left(\mathbf{x} + \boldsymbol{\lambda}_i'\right) - \mathcal{L}(\mathbf{x})\right] \\ &\quad + \frac{1}{2\sigma^2}\int_0^{NT+\Delta} \left(s^*\left(t, \boldsymbol{\lambda}^{\text{in}}\right) s\left(t, \boldsymbol{\lambda}_i'\right) \right. \\ &\quad \left. + s\left(t, \boldsymbol{\lambda}^{\text{in}}\right) s^*\left(t, \boldsymbol{\lambda}_i'\right)\right) dt \geq 0 \quad (44) \end{aligned}$$

because $\mathcal{L}(\mathbf{x} + \boldsymbol{\lambda}_i) \geq \mathcal{L}\left(\mathbf{x} + \boldsymbol{\lambda}_i'\right)$ when $\mathbf{x} + \boldsymbol{\lambda}_i$ is detected (i.e. the most likely vector), and $\int_0^{NT+\Delta} \left(s^*\left(t, \boldsymbol{\lambda}^{\text{in}}\right) s\left(t, \boldsymbol{\lambda}_i'\right) + s\left(t, \boldsymbol{\lambda}^{\text{in}}\right) s^*\left(t, \boldsymbol{\lambda}_i'\right)\right) dt \geq 0$ according to (39a). Equation (44) indicates that $\mathcal{L}\left(\mathbf{x} + \boldsymbol{\lambda}^{\text{in}}\right) \geq \mathcal{L}(\mathbf{x})$ when $\mathbf{x} + \boldsymbol{\lambda}_i$ is detected.

Then we prove (43) based on (44). In the subsequent discussion, for notational simplicity, we define the events $D = \{\mathbf{x} \text{ is transmitted}\}$, $H = \left\{\mathcal{L}\left(\mathbf{x} + \boldsymbol{\lambda}^{\text{in}}\right) \geq \mathcal{L}(\mathbf{x})\right\}$, $F_i = \{\mathbf{x} + \boldsymbol{\lambda}_i \text{ is detected}\}$, and $G = \left\{\mathbf{x} + \boldsymbol{\lambda}^{\text{in}} \text{ is detected}\right\}$. It is obvious that $\Pr(H \mid G, D) = 1$. According to (44), we obtain that $\Pr(H \mid F_i, D) = 1$. Since

$$\begin{aligned} &\Pr(G \mid H, D)\Pr(H \mid D) = \Pr(H \mid G, D)\Pr(G \mid D) \\ &\Pr(F_i \mid H, D)\Pr(H \mid D) = \Pr(H \mid F_i, D)\Pr(F_i \mid D), \quad (45) \end{aligned}$$

we obtain

$$\begin{aligned} &\Pr(H \mid D)\left(\Pr(G \mid H, D) + \sum_{i=1}^{m} \Pr(F_i \mid H, D)\right) \\ &\quad = \Pr(G \mid D) + \sum_{i=1}^{m} \Pr(F_i \mid D). \quad (46) \end{aligned}$$

Since if traverse all possible detected vectors in $\mathcal{J}_{\mathbf{x},k}$, we obtain

$$\sum_{\boldsymbol{\lambda}:\mathbf{x}+\boldsymbol{\lambda}\in\mathcal{J}_{\mathbf{x},k}} \Pr\left(\mathbf{x}+\boldsymbol{\lambda} = \arg\max_{\mathbf{x}'\in\mathcal{J}_{\mathbf{x},k}} \mathcal{L}\left(\mathbf{x}'\right) \mid H, D\right) = 1, \quad (47)$$

thus $\Pr(G \mid H, D) + \sum_{i=1}^{m} \Pr(F_i \mid H, D) \leq 1$. Therefore, (46) can be rewritten as

$$\Pr(G \mid D) + \sum_{i=1}^{m} \Pr(F_i \mid D) \leq \Pr(H \mid D). \quad (48)$$

Replacing $\Pr(H \mid D)$ with (35), (43) can be proven. ■

According to Lemma 6, we know that $Q\left(S\left(\boldsymbol{\lambda}^{\text{in}}\right)/2\sigma\right)$ is a tighter upper bound than $Q\left(S\left(\boldsymbol{\lambda}^{\text{in}}\right)/2\sigma\right) + \sum_{i=1}^{m} Q\left(S\left(\boldsymbol{\lambda}_i\right)/2\sigma\right)$ for $P\left(\mathbf{x} + \boldsymbol{\lambda}^{\text{in}} \mid \mathbf{x}\right) + \sum_{i=1}^{m} P\left(\mathbf{x} + \boldsymbol{\lambda}_i \mid \mathbf{x}\right)$. Therefore, we can treat $\sum_{i=1}^{m} Q\left(S\left(\boldsymbol{\lambda}_i\right)/2\sigma\right)$ as the redundant term and omit it. Moreover, according to Lemma 5, every decomposable error vector belonging to $\left\{\boldsymbol{\lambda} : \mathbf{x} + \boldsymbol{\lambda} \in \mathcal{J}_{\mathbf{x},k}\right\}$ can be decomposed into an indecomposable error vector also belonging to $\left\{\boldsymbol{\lambda} : \mathbf{x} + \boldsymbol{\lambda} \in \mathcal{J}_{\mathbf{x},k}\right\}$ and another error vector. Thus, according to Lemma 6, we can omit all of the terms regarding decomposable error vector at the right-hand side of (38). Therefore, the upper bound of SER can be achieved by

$$P_{e,k} \leq \frac{1}{2^{2N}} \sum_{\mathbf{x}\in\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{2^{2N}}\}} \sum_{\boldsymbol{\lambda}\in\mathcal{F}\cap\{\boldsymbol{\lambda}:\mathbf{x}+\boldsymbol{\lambda}\in\mathcal{J}_{\mathbf{x},k}\}} Q\left(\frac{S(\boldsymbol{\lambda})}{2\sigma}\right). \quad (49)$$

The set $\mathcal{F}$ denotes the set of indecomposable error vectors where $\lambda_{A,k} \neq 0, \lambda_{B,k} = 0$ or $\lambda_{A,k} = 0, \lambda_{B,k} \neq 0$.

It is observed in (49) that we need to traverse all kinds of transmitted vectors to obtain the upper bound of $P_{e,k}$, while, we prefer to simplify (49) so that upper bound of $P_{e,k}$ is independent on transmitted vectors.

Lemma 7: The upper bound of SER can be rewritten as

$$P_{e,k} \leq \sum_{\boldsymbol{\lambda} \in \mathcal{F}} 2^{-\omega(\boldsymbol{\lambda})} Q\left(\frac{S(\boldsymbol{\lambda})}{2\sigma}\right), \tag{50}$$

where $\omega(\boldsymbol{\lambda})$ denotes the number of non-zero elements in $\boldsymbol{\lambda}$.

Proof: We focus on $\sum_{\mathbf{x} \in \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{2^{2N}}\}} \sum_{\hat{\mathbf{x}} \in \mathcal{J}_{\mathbf{x},k}} Q\left(S\left(\hat{\mathbf{x}} - \mathbf{x}\right)/2\sigma\right)$. When traversing $\mathbf{x} \in \left\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{2^{2N}}\right\}$, we can combine like terms. In detail, for some elements (not include $x_{A,k}$ and $x_{B,k}$ simultaneously) in $\mathbf{x}$, if the corresponding elements in $\hat{\mathbf{x}}$ are detected right, it is observed that the value of $\hat{\mathbf{x}} - \mathbf{x}$ will not change no matter what are the values of those elements in $\mathbf{x}$ because the corresponding elements in $\hat{\mathbf{x}} - \mathbf{x}$ are always zero. Therefore, the number of like terms relates to the number of zero element in the common $\hat{\mathbf{x}} - \mathbf{x}$ of the like terms. Let $\omega(\cdot)$ denote the number of non-zero element in a vector, for a certain $\boldsymbol{\lambda} = \hat{\mathbf{x}} - \mathbf{x}$, according to the above analysis, it is observed that there are $2^{2N-\omega(\boldsymbol{\lambda})}$ terms of $Q(S(\boldsymbol{\lambda})/2\sigma)$ in $\sum_{\mathbf{x} \in \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{2^{2N}}\}} \sum_{\hat{\mathbf{x}} \in \mathcal{J}_{\mathbf{x},k}} Q\left(S\left(\hat{\mathbf{x}} - \mathbf{x}\right)/2\sigma\right)$. Likewise, (49) can be written as

$$\begin{aligned} P_{e,k} &\leq \frac{1}{2^{2N}} \sum_{\boldsymbol{\lambda} \in \mathcal{F}} 2^{2N-\omega(\boldsymbol{\lambda})} Q\left(\frac{S(\boldsymbol{\lambda})}{2\sigma}\right) \\ &= \sum_{\boldsymbol{\lambda} \in \mathcal{F}} 2^{-\omega(\boldsymbol{\lambda})} Q\left(\frac{S(\boldsymbol{\lambda})}{2\sigma}\right). \end{aligned} \tag{51}$$

Therefore, (50) is proved. ■

B. ANALYTICAL EXPRESSION OF UPPER BOUND FOR SER

In this subsection, we expand the left-hand side of (50) and carry out the analytical expression of upper bound of $P_{e,k}$.

Proposition 2: The upper bound for SER is obtained by

$$\begin{aligned} P_U = &Q\left(\frac{\sqrt{2 - 2\cos\phi \cdot \Delta/T}}{\sigma}\right) + \frac{2}{3} Q\left(\frac{\sqrt{3 - 2\cos\phi}}{\sigma}\right) \\ &+ 2Q\left(\frac{1}{\sigma}\right) + \frac{2}{7} Q\left(\frac{\sqrt{4 - 2\cos\phi - 2\cos\phi \cdot \Delta/T}}{\sigma}\right). \end{aligned} \tag{52}$$

Proof: Firstly, we present all kinds of the indecomposable error vector. Judging whether a certain error vector $\boldsymbol{\lambda}$ can be decomposed into $\boldsymbol{\lambda}'$ and $\boldsymbol{\lambda}''$ or not lies in judging whether (39a) is true, subject to (39a)-(39a). Similar to (17), we obtain

$$\begin{aligned} &f(\boldsymbol{\lambda}', \boldsymbol{\lambda}'') \\ &= \int_0^{NT+\Delta} \left(s^*\left(t, \boldsymbol{\lambda}'\right) s\left(t, \boldsymbol{\lambda}''\right) + s\left(t, \boldsymbol{\lambda}'\right) s^*\left(t, \boldsymbol{\lambda}''\right)\right) dt \\ &= \int_0^{NT+\Delta} \left|s\left(t, \boldsymbol{\lambda}' + \boldsymbol{\lambda}''\right)\right|^2 - \left|s\left(t, \boldsymbol{\lambda}'\right)\right|^2 - \left|s\left(t, \boldsymbol{\lambda}''\right)\right|^2 dt \\ &= \sum_{k=1}^{N} \left(2 \cdot \frac{T - \Delta}{T} \left(\lambda'_{A,k} + \lambda''_{A,k}\right)\left(\lambda'_{B,k} + \lambda''_{B,k}\right) \cos\phi \right. \\ &\quad + \left(\lambda'_{A,k} + \lambda''_{A,k}\right)^2 + \left(\lambda'_{B,k} + \lambda''_{B,k}\right)^2 \\ &\quad \left. + 2 \cdot \frac{\Delta}{T} \left(\lambda'_{A,k} + \lambda''_{A,k}\right)\left(\lambda'_{B,k-1} + \lambda''_{B,k-1}\right) \cos\phi \right) \\ &\quad - \sum_{k=1}^{N} \left(2 \cdot \frac{T - \Delta}{T} \lambda'_{A,k}\lambda'_{B,k} \cos\phi + \left(\lambda'_{A,k}\right)^2 \right. \\ &\quad \left. + \left(\lambda'_{B,k}\right)^2 + 2 \cdot \frac{\Delta}{T} \lambda'_{A,k}\lambda'_{B,k-1} \cos\phi \right) \\ &\quad - \sum_{k=1}^{N} \left(2 \cdot \frac{T - \Delta}{T} \lambda''_{A,k}\lambda''_{B,k} \cos\phi + \left(\lambda''_{A,k}\right)^2 \right. \\ &\quad \left. + \left(\lambda''_{B,k}\right)^2 + 2 \cdot \frac{\Delta}{T} \lambda''_{A,k}\lambda''_{B,k-1} \cos\phi \right) \\ &= \sum_{k=1}^{N} \left(\frac{T - \Delta}{T} \cdot \left(2\lambda'_{A,k}\lambda''_{B,k} + 2\lambda'_{B,k}\lambda''_{A,k}\right) \cdot \cos\phi \right. \\ &\quad + 2\lambda'_{A,k}\lambda''_{A,k} + 2\lambda'_{B,k}\lambda''_{B,k} + \frac{\Delta}{T} \\ &\quad \left. \cdot \left(2\lambda'_{A,k}\lambda''_{B,k-1} + 2\lambda''_{A,k}\lambda'_{B,k-1}\right) \cos\phi \right). \end{aligned} \tag{53}$$

Note that a certain element in $\boldsymbol{\lambda}$ can be only decomposed into itself and 0 no matter whether the element is equal to 0 or not. Therefore, it is obvious that the terms $\lambda'_{A,k}\lambda''_{A,k}$ and $\lambda'_{B,k}\lambda''_{B,k}$ in (53) equal 0 since $\lambda'_{A,k}$ or $\lambda''_{A,k}$ is 0 and $\lambda'_{B,k}$ or $\lambda''_{B,k}$ is 0.

Considering the k th network-coding symbol $x_{R,k}$ is erroneously detected and the case that $\lambda_{A,k} \neq 0, \lambda_{B,k} = 0$, then, $(\lambda_{A,1}, \lambda_{B,1}, \ldots, \lambda_{B,k-1}, \lambda_{A,k}, \lambda_{B,k}, \lambda_{A,k+1}, \ldots, \lambda_{B,N})$ can be decomposed into $(\lambda'_{A,1}, \lambda'_{B,1}, \ldots, \lambda'_{B,k-1}, \lambda'_{A,k}, 0, 0, \ldots, 0)$ and $(0, 0, \ldots, 0, 0, \lambda''_{B,k}, \lambda''_{A,k+1}, \ldots, \lambda''_{B,N})$ where $\lambda'_{A,k} \neq 0$ and $\lambda''_{B,k} = 0$, since $f(\boldsymbol{\lambda}', \boldsymbol{\lambda}'') = 0$. Therefore, indecomposable error vectors have the same style as $(\lambda'_{A,1}, \lambda'_{B,1}, \ldots, \lambda'_{B,k-1}, \lambda'_{A,k}, 0, 0, \ldots, 0)$ and a certain element between any two non-zero elements can not be 0, otherwise, the error vector can be decomposed just as above analysis. In conclusion, the elements in an indecomposable error vector satisfy that all of the non-zero element must be adjacent.

Focusing on other terms in (53) and considering the carrier phase offset $\phi \in [0, \pi/2]$ (i.e. $\cos\phi \geq 0$), we can obtain that $f(\boldsymbol{\lambda}', \boldsymbol{\lambda}'') < 0$ if any adjacent two non-zero elements (e.g., $\lambda_{B,k-1}$ and $\lambda_{A,k}$ are adjacent two elements, as well as $\lambda_{A,k}$ and $\lambda_{B,k}$) in $\boldsymbol{\lambda}$ are numbers with opposite signs, since $\lambda'_{A,k}\lambda''_{B,k}$, $\lambda''_{A,k}\lambda'_{B,k}$, $\lambda'_{A,k}\lambda''_{B,k-1}$ and $\lambda''_{A,k}\lambda'_{B,k-1}$ are not more than 0. For example, the values of $\lambda'_{A,k}\lambda''_{B,k}$ and $\lambda''_{A,k}\lambda'_{B,k}$ can only be in the set $\left\{0, \lambda_{A,k}\lambda_{B,k}\right\}$, if $\lambda_{A,k}$ and $\lambda_{B,k}$ are numbers with opposite signs, thus, $\lambda_{A,k}\lambda_{B,k} = -4$. While, $\lambda'_{A,k}\lambda''_{B,k}$, $\lambda''_{A,k}\lambda'_{B,k}$, $\lambda'_{A,k}\lambda''_{B,k-1}$ and $\lambda''_{A,k}\lambda'_{B,k-1}$ can not be zero simultaneously, otherwise, $\boldsymbol{\lambda}$ will be decomposed into $\boldsymbol{\lambda}$ and corresponding null vector, which is an insignificant decomposition.

	$\lambda_{A,1}$	$\cdots$	$\lambda_{A,k-1}$	$\lambda_{B,k-1}$	$\lambda_{A,k}$	$\lambda_{B,k}$	$\cdots$	$\lambda_{B,N}$
$2(2k-1)+1$ combinations	0	$\cdots$	0	0	2	0	$\cdots$	0
	0	$\cdots$	0	0	-2	0	$\cdots$	0
	0	$\cdots$	0	-2	2	0	$\cdots$	0
	0	$\cdots$	0	2	-2	0	$\cdots$	0
	0	$\cdots$	2	-2	2	0	$\cdots$	0
	0	$\cdots$	-2	2	-2	0	$\cdots$	0
	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$
	2	$\cdots$	2	-2	2	0	$\cdots$	0
	-2	$\cdots$	-2	2	-2	0	$\cdots$	0
				$2N$				

FIGURE 4. Indecomposable error vectors when $\lambda_{A,k} \neq 0, \lambda_{B,k} = 0$, and $\cos\phi \geq 0$ (i.e. phase offset $\phi \in [0, \pi/2]$). Each row represents elements of an indecomposable error vector.

Thus, considering different number of non-zero elements, the set $\mathcal{F}\,|_{\lambda_{A,k}\neq 0,\lambda_{B,k}=0}$ can be shown in Fig. 4. For the error vectors in the $2i$ th and $2i-1$ th rows ($i \in \{1, 2, \ldots, 2k-1\}$) of Fig. 4, the number of non-zero elements in the error vector is equal i. Likewise, considering the case that $\lambda_{A,k} = 0, \lambda_{B,k} \neq 0$, we can obtain $\mathcal{F}\,|_{\lambda_{A,k}=0,\lambda_{B,k}\neq 0}$.

Secondly, we expand the right-hand side of (50) based on Fig. 4. According to $S(\boldsymbol{\lambda}) = \sqrt{\int_0^{NT+\Delta} |s(t,\boldsymbol{\lambda})|^2\, dt}$ and (17), we obtain (54), as shown at the bottom of this page. If a certain element in $\boldsymbol{\lambda}$ is non-zero, thus its square is equal to 4. Because the adjacent two non-zero elements in $\boldsymbol{\lambda}$ are numbers with opposite signs, thus the product of the two numbers is -4.

For the number of non-zero element $i \in \{1, 2, \ldots, 2k-1\}$, if i is even number, regarding the error vectors in $2i$ th and $2i-1$ th rows,

$$S(\boldsymbol{\lambda}) = \sqrt{4\left(i - 2\cdot\left(\frac{i}{2}-1\right)\cos\phi - 2\cdot\frac{\Delta}{T}\cdot\cos\phi\right)}. \tag{55}$$

If i is odd number, regarding the error vectors in $2i$ th and $2i-1$ th rows,

$$S(\boldsymbol{\lambda}) = \sqrt{4\,(i-(i-1)\cos\phi)}. \tag{56}$$

According to (55) and (56),

$$\begin{aligned}
&\sum_{\boldsymbol{\lambda}\in\mathcal{F}|_{\lambda_{A,k}\neq 0,\lambda_{B,k}=0}} 2^{-\omega(\boldsymbol{\lambda})} Q\left(\frac{S(\boldsymbol{\lambda})}{2\sigma}\right) \\
&= 2\sum_{i\in\{2,4,\ldots,2k-2\}} 2^{-i} Q \\
&\quad\times\left(\frac{\sqrt{i-2\cdot\left(\frac{i}{2}-1\right)\cos\phi-2\cdot\frac{\Delta}{T}\cdot\cos\phi}}{\sigma}\right) \\
&\quad+2\sum_{i\in\{1,3,\ldots,2k-1\}} 2^{-i} Q\left(\frac{\sqrt{i-(i-1)\cos\phi}}{\sigma}\right).
\end{aligned} \tag{57}$$

$$\begin{aligned}
S(\boldsymbol{\lambda}) &= \sqrt{\sum_{k=1}^{N}\left(\lambda_{A,k}^2+\lambda_{B,k}^2+2\cdot\frac{T-\Delta}{T}\lambda_{A,k}\lambda_{B,k}\cos\phi+2\cdot\frac{\Delta}{T}\lambda_{A,k}\lambda_{B,k-1}\cos\phi\right)} \\
&= \sqrt{\lambda_{A,1}^2+2\cdot\frac{T-\Delta}{T}\lambda_{A,1}\lambda_{B,1}\cos\phi+\lambda_{B,1}^2+2\cdot\frac{\Delta}{T}\lambda_{A,2}\lambda_{B,1}\cos\phi+\lambda_{A,2}^2+\ldots+\lambda_{B,N}^2}.
\end{aligned} \tag{54}$$

$$\begin{aligned}
\sum_{\boldsymbol{\lambda}\in\mathcal{F}} 2^{-\omega(\boldsymbol{\lambda})} Q\left(\frac{|S(\boldsymbol{\lambda})|}{2\sigma}\right) &= \sum_{\boldsymbol{\lambda}\in\mathcal{F}|_{\lambda_{A,k}\neq 0,\lambda_{B,k}=0}} 2^{-\omega(\boldsymbol{\lambda})} Q\left(\frac{|S(\boldsymbol{\lambda})|}{2\sigma}\right) + \sum_{\boldsymbol{\lambda}\in\mathcal{F}|_{\lambda_{A,k}=0,\lambda_{B,k}\neq 0}} 2^{-\omega(\boldsymbol{\lambda})} Q\left(\frac{|S(\boldsymbol{\lambda})|}{2\sigma}\right) \\
&< 4\sum_{i\in\{2,4,\ldots,2N-2\}} 2^{-i} Q\left(\frac{\sqrt{i-2\cdot\left(\frac{i}{2}-1\right)\cos\phi-2\cdot\frac{\Delta}{T}\cdot\cos\phi}}{\sigma}\right) + 4\sum_{i\in\{1,3,\ldots,2N-1\}} 2^{-i} Q\left(\frac{\sqrt{i-(i-1)\cos\phi}}{\sigma}\right).
\end{aligned} \tag{59}$$

$$\begin{aligned}
\sum_{\boldsymbol{\lambda}\in\mathcal{F}} 2^{-\omega(\boldsymbol{\lambda})} Q\left(\frac{|S(\boldsymbol{\lambda})|}{2\sigma}\right) &< 4\cdot 2^{-i} Q\left(\frac{\sqrt{i-2\cdot\left(\frac{i}{2}-1\right)\cos\phi-2\cdot\frac{\Delta}{T}\cdot\cos\phi}}{\sigma}\right)\Bigg|_{i=2} + 4\cdot 2^{-i} Q\left(\frac{\sqrt{i-(i-1)\cos\phi}}{\sigma}\right)\Bigg|_{i=1} \\
&\quad+\lim_{N\to\infty} 4\sum_{i\in\{4,6,\ldots,2N-2\}} 2^{-i}\left(Q\left(\frac{\sqrt{i-2\cdot\left(\frac{i}{2}-1\right)\cos\phi-2\cdot\frac{\Delta}{T}\cdot\cos\phi}}{\sigma}\right)\Bigg|_{i=4}\right) \\
&\quad+\lim_{N\to\infty} 4\sum_{i\in\{3,5,\ldots,2N-1\}} 2^{-i}\left(Q\left(\frac{\sqrt{i-(i-1)\cos\phi}}{\sigma}\right)\Bigg|_{i=3}\right) \\
&= 2Q\left(\frac{1}{\sigma}\right)+Q\left(\frac{\sqrt{2-2\cos\phi\cdot\Delta/T}}{\sigma}\right)+\frac{2Q\left(\frac{\sqrt{3-2\cos\phi}}{\sigma}\right)}{3}+\frac{2Q\left(\frac{\sqrt{4-2\cos\phi-2\cos\phi\cdot\Delta/T}}{\sigma}\right)}{7}.
\end{aligned} \tag{60}$$

Likewise, we can obtain

$$\sum_{\boldsymbol{\lambda}\in\mathcal{F}|\lambda_{A,k}=0,\lambda_{B,k}\neq 0} 2^{-\omega(\boldsymbol{\lambda})} Q\left(\frac{|S(\boldsymbol{\lambda})|}{2\sigma}\right)$$
$$= 2 \sum_{i\in\{2,4,\ldots,2N-2k\}} 2^{-i} Q$$
$$\times \left(\frac{\sqrt{i-2\cdot\left(\frac{i}{2}-1\right)\cos\phi-2\cdot\frac{\Delta}{T}\cdot\cos\phi}}{\sigma}\right)$$
$$+2 \sum_{i\in\{1,3,\ldots,2N-2k+1\}} 2^{-i} Q\left(\frac{\sqrt{i-(i-1)\cos\phi}}{\sigma}\right). \quad (58)$$

According to sum of (57) and (58), the upper bound of SER for k th network-coded symbol $x_{R,k}$ relates to the value of k. In order to derive a common upper bound suitable for the error probability of $x_{R,k}$ no matter what is the value of k, we can add additional terms in right-hand side of (57) and (58) as follows, shown at the bottom of this page.

Since $Q(\cdot)$ in (59), as shown at the bottom of the previous page, is monotonically decreasing, the terms at the right-hand side of (59) satisfy that $Q(\cdot)\mid_{i=4} > Q(\cdot)\mid_{i>4}$ and $Q(\cdot)\mid_{i=3} > Q(\cdot)\mid_{i>3}$, and if $N\to\infty$, the right-hand side of (59) will be convergent. Therefore, we make $N\to\infty$, then (59) can be written as

When the phase offset $\phi\in[\pi/2,\pi]$ (i.e. $\cos\phi\leq 0$), we can still obtain the same analytical expression as in (52). ■

V. SIMULATION RESULTS

In this section, we provide numerical results with BP-MAP method to verify our analytical results derived in Sections III and IV. We compare our analytical bounds for SER with the exact values obtained by Monte Carlo simulations. During our simulation, the size of each frame N is set to 1,024 bytes, and we simulate 300,000 frames. T is set to 1.

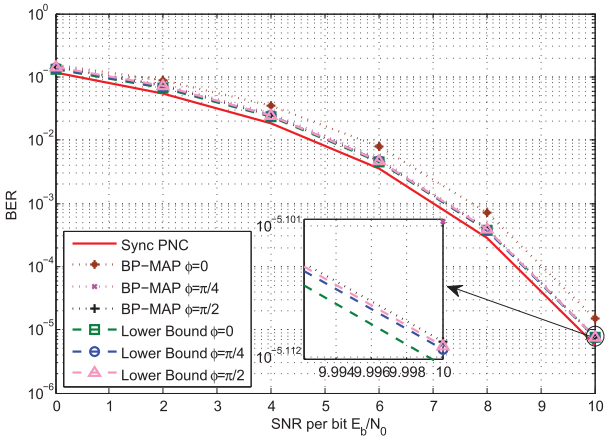


FIGURE 5. Analytical results of the lower bound versus simulated results when $\Delta = 0.5$.

Fig. 5 investigates the SERs at the relay during the MA phase in various signal-to-noise ratio (SNR) regions when the relative phase offsets ϕ equal 0, $\pi/4$, and $\pi/2$. During this study, the symbol offset is set to 0.5. It is observed in Fig. 5 that the lower bounds for SER under different phase offsets are less than the exact SER of asynchronous PNC (which is decoded through BP-MAP), which verifies our analytical results. Additionally, it is known that the SER performance of synchronous PNC outperforms that of asynchronous PNC [17], and Fig. 5 also illustrates that the analytical lower bounds for different phase offsets are always greater than the exact SER of synchronous PNC. This phenomenon indicates that our analytical results are adequate accurate and can be approximate values for SER of asynchronous PNC, which should be greater than the exact SER of synchronous PNC.

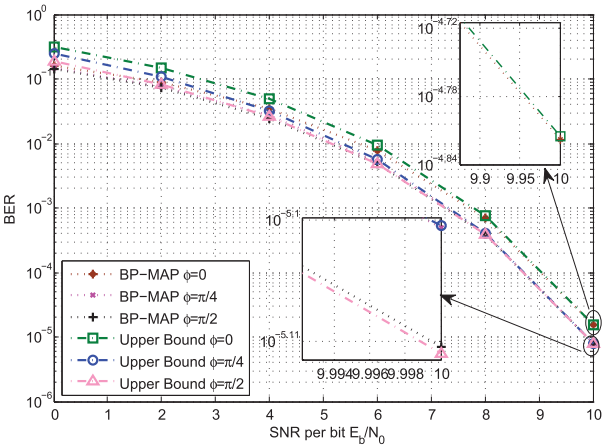


FIGURE 6. Analytical results of the upper bound versus simulated results when $\Delta = 0.5$.

Similarly, Fig. 6 verifies our analytical results on the upper bound for SER. It is depicted in Fig. 6 that all the exact SERs of asynchronous PNC under different phase offsets are upper-bounded by corresponding analytical results.[2] The small differences between the simulation results and their corresponding bounds demonstrate that both the upper and the lower bounds are tight enough to facilitate the analysis on SER performance at the relay. Since BP-MAP method approximates MUD-XOR method as the analysis in II-B.3 when SNR increases, the upper bounds approach the corresponding simulation results as the SNR increases, as shown in Fig. 6.

Fig. 7 exhibits the SERs of asynchronous PNC and their corresponding upper bounds and lower bounds at the relay for different symbol offsets when SNR per bit equals 8 dB. We do not consider the cases $\Delta > 0.5$, because if $\Delta > 0.5$, we can decode $x_{A,k+1}\oplus x_{B,k}$ rather than $x_{A,k}\oplus x_{B,k}$ so that the performance for $x_{A,k+1}\oplus x_{B,k}$ with symbol offset Δ is similar to that for $x_{A,k}\oplus x_{B,k}$ with symbol offset $1-\Delta$ [14]. In the scenarios that phase offsets are 0 and $\pi/4$, the SERs are nearly

[2]The upper bound approaches the corresponding lower bound as the SNR increases, hence, if the simulation value of SER is not precise enough, the simulation result will be slightly higher than the analytical result, just as the phase offset is $\pi/2$.

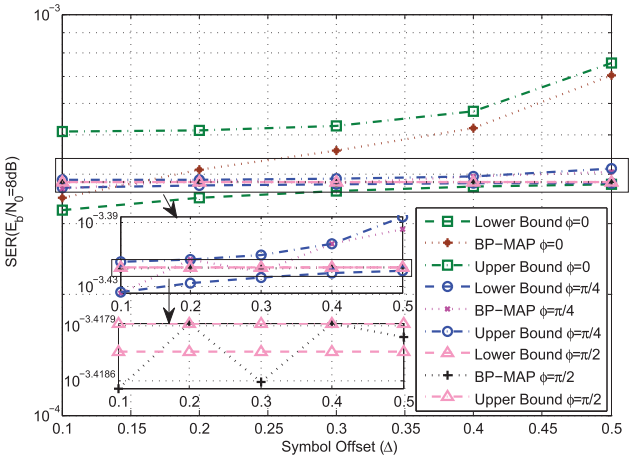


FIGURE 7. SER versus symbol offsets, when SNR per bit is 8 dB.

invariable (for QPSK, difference between the SERs will be great) when the symbol offset is small (e.g. less than 0.3), while the SER curves increase obviously when the symbol offset becomes larger. However, in the case that the phase offset is $\pi/2$, the SER value almost stays the same level with the increase of the symbol offset,[3] and the curve of corresponding upper bound follows the trend of the exact value. This is because when the relative phase offset equals $\pi/2$, the two discrete-time complex signals (with BPSK modulation from the two end nodes) are orthogonal in the 2-dimensional plane. Therefore, we can make the decoding more robust against phase offset through controlling the relative symbol offset, when performing asynchronous PNC.

VI. CONCLUSIONS

PNC is viewed as one of the key 5G key enabling technologies. In this paper, we have analyzed SER performance of asynchronous PNC during the MA phase at the relay and derived analytical results of the lower bound and upper bound for SER. The lower bound and upper bound obtained in this paper are applicable to decoding method based on either MUD-XOR or MAP. Numerical results have verified the analytical results and demonstrated that our results are adequate accurate and can be applied when investigating SER performance of asynchronous PNC with various symbol and phase offsets. Focusing on asynchronous PNC, this paper complement issues on the performance analysis of PNC, and can facilitate future practical and theoretical studies.

REFERENCES

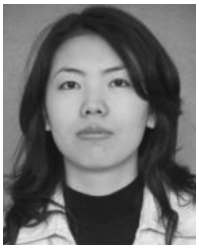
[1] X. Liu, P. Wang, Z. Lan, and B. Shao, "Biological characteristic online identification technique over 5G network," *IEEE Wireless Commun.*, vol. 22, no. 6, pp. 84–90, Dec. 2015.
[2] W. Mao, X. Wang, A. Tang, and H. Qian, "ANC-ERA: Random access for analog network coding in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 15, no. 1, pp. 45–59, Jan. 2016.
[3] J. Wang, K. Lu, J. Wang, and C. Qiao, "Optimal local data exchange in fiber-wireless access network: A joint network coding and device association design," in *Proc. IEEE INFOCOM*, Apr. 2016, pp. 1–9.

[3]The simulation value of SER fluctuates as Δ increases, however, the fluctuation is so little that we can treat it invariable.

[4] L. Wan, G. Han, L. Shu, S. Chan, and T. Zhu, "The application of doa estimation approach in patient tracking systems with high patient density," *IEEE Trans. Ind. Informat.*, vol. PP, no. 99, pp. 1–10, May 2016, doi: 10.1109/TII.2016.2569416.
[5] Z. Ning, Q. Song, L. Guo, and A. Jamalipour, "Throughput improvement by network coding and spatial reuse in wireless mesh networks," in *Proc. IEEE GLOBECOM*, Dec. 2013, pp. 5072–5077.
[6] J. Hansen, D. E. Lucani, J. Krigslund, M. Médard, and F. H. P. Fitzek, "Network coded software defined networking: Enabling 5G transmission and storage networks," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 100–107, Sep. 2015.
[7] Z. Ning, L. Liu, F. Xia, B. Jedari, I. Lee, and W. Zhang, "CAIS: A copy adjustable incentive scheme in community-based socially-aware networking," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–14. Jul. 2016, doi: 10.1109/TVT.2016.2593051.
[8] L. Wan, G. Han, L. Shu, S. Chan, and N. Feng, "PD source diagnosis and localization in industrial high-voltage insulation system via multimodal joint sparse representation," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2506–2516, Apr. 2016.
[9] K. Lu, S. Fu, Y. Qian, and H.-H. Chen, "SER performance analysis for physical layer network coding over AWGN channels," in *Proc. IEEE GLOBECOM*, Nov./Dec. 2009, pp. 1–6.
[10] Y. Huang, Q. Song, S. Wang, and A. Jamalipour, "Symbol error rate analysis for M-QAM modulated physical-layer network coding with phase errors," in *Proc. IEEE PIMRC*, Sep. 2012, pp. 2003–2008.
[11] M. Park, I. Choi, and I. Lee, "Exact BER analysis of physical layer network coding for two-way relay channels," in *Proc. IEEE VTC*, May 2011, pp. 1–5.
[12] M. Ju and I.-M. Kim, "Error performance analysis of BPSK modulation in physical-layer network-coded bidirectional relay networks," *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2770–2775, Oct. 2010.
[13] T. Koike-Akino, P. Popovski, and V. Tarokh, "Optimized constellations for two-way wireless relaying with physical network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 773–787, Jun. 2009.
[14] L. Lu and S. C. Liew, "Asynchronous physical-layer network coding," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 819–831, Feb. 2012.
[15] S. C. Liew, S. Zhang, and L. Lu, "Physical-layer network coding: Tutorial, survey, and beyond," *Phys. Commun.*, vol. 6, pp. 4–42, Mar. 2013.
[16] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near-single-user performance," *IEEE Trans. Commun.*, vol. 46, no. 12, pp. 1693–1699, Dec. 1998.
[17] L. Lu, S. C. Liew, and S. Zhang, "Optimal decoding algorithm for asynchronous physical-layer network coding," in *Proc. IEEE ICC*, Jun. 2011, pp. 1–6.
[18] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.
[19] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 2004.
[20] J. Park and S. Park, "Approximation for the two-dimensional Gaussian Q-function and its applications," *ETRI J.*, vol. 32, no. 1, pp. 145–147, 2010.
[21] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 32, no. 1, pp. 85–96, Jan. 1986.

LEI GUO (M'06) received the Ph.D. degree from the University of Electronic Science and Technology of China in 2006. He is a Professor with Northeastern University, Shenyang, China. His current research interests include communication networks, optical communications, and wireless communications. He has published over 200 technical papers in the above areas in international journals and conferences, such as the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE JOURNAL OF LIGHTWAVE TECHNOLOGY, the IEEE JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING, the IEEE GLOBECOM, and the IEEE ICC. He is a member of the OSA, and he is also a Senior Member of the CIC. He is now serving as an Editor for several international journals, such as *Photonic Network Communications* and *The Open Optics Journal*.

ZHAOLONG NING (M'14) received the Ph.D. degree from Northeastern University, China, in 2014. From 2013 to 2014, he was a Research Fellow with Kyushu University, Japan. He is currently an Assistant Professor with the Dalian University of Technology. His current research interests include social network, network optimization, and big data. He has authored over 40 papers in the above areas. He is a member of the ACM.



QINGYANG SONG (SM'12) received the Ph.D. degree in telecommunications engineering from the University of Sydney, Australia. She is a Professor with Northeastern University, China. She has authored over 50 papers in major journals and international conferences. These papers have been cited over 1000 times in the scientific literature. Her current research interests are in radio resource management, cognitive radio networks, cooperative communications, ad-hoc networks, heterogeneous cellular networks, and protocol design.

YAYUN CUI received the master's degree from Northeastern University, China, in 2015. He is now with HUAWEI company. His research focuses on physical layer design and network coding.

ZHIKUI CHEN (M'07) received the Ph.D. degree in digital signal processing and the M.S. degree in mechanics from Chongqing University in 1998 and 1993, respectively. He is currently a Full Professor with the Dalian University of Technology, China. He is leading the Institute of Ubiquitous Network and Computing, Dalian University of Technology. He was a General Chair of the IEEE ithings2011, an Advisor Chair of the IEEE ithings 2012–2016, and a Program Chair of the IEEE ICDH2014. He is a General Chair of the IEEE Smartdata2015–2016. His research interests are big data processing, mobile cloud computing, and ubiquitous network and its computing.