Artificial Noise Assisted Secure Transmission Under Training and Feedback

Hui-Ming Wang, *Member, IEEE*, Chao Wang, and Derrick Wing Kwan Ng, *Member, IEEE*

***Abstract*—This paper proposes a framework for the artificial noise assisted secure transmission in multiple-input, multiple-output, multiple antenna eavesdropper (MIMOME) wiretap channels in frequency-division duplexed (FDD) systems. We focus on a practical scenario that only the eavesdroppers' channel distribution information (CDI) is available and the imperfect channel state information (CSI) of the legitimate receiver is acquired through training and analog feedback. By taking explicitly into account the signaling overhead and training power overhead incurred by channel estimation and feedback, we define the achievable effective ergodic secrecy rate (ESR), and investigate a joint power allocation and training overhead optimization problem for the maximization of effective ESR. We first derive a deterministic approximation for the achievable effective ESR which facilitates the joint optimization. Then, efficient iterative algorithms are proposed to solve the considered nonconvex optimization problem. In particular, in the high-SNR regime, a block coordinate descent method (BCDM) is proposed to handle the joint optimization. In the low-SNR regime, we transform the problem into a sequence of geometric programmings (GPs) and locate its Karush–Kuhn–Tucker (KKT) solution using the successive convex approximation (SCA) method. For the general case of SNR, we maximize the lower bound of the achievable effective ESR. Simulation results corroborate the theoretical analysis and illustrate the secrecy performance of the proposed secure transmission scheme.**

***Index Terms*—Physical layer security, MIMOME wiretap channels, artificial noise, power allocation, training, feedback, ergodic secrecy rate.**

Manuscript received May 29, 2015; accepted July 25, 2015. Date of publication August 05, 2015; date of current version October 27, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yue Rong. Their work was partially supported by the NSFCunder Grant No. 61221063, the Foundation for the Author of National Excellent Doctoral Dissertation of Chinaunder Grant 201340, the National High-Tech Research and Development Program of Chinaunder Grant No. 2015AA011306, the New Century Excellent Talents Support Fund of Chinaunder Grant NCET-13-0458, the Fok Ying Tong Education Foundationunder Grant 141063, and the Fundamental Research Funds for the Central Universityunder Grant No. 2013jdgz11. The work of D. W. K. Ng was partially supported by the AvH Professorship Program of the Alexander von Humboldt Foundation.

H.-M. Wang and C. Wang are with the School of Electronic and Information Engineering and also the MOE Key Lab for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an, 710049, Shaanxi, China (e-mail: xjbswhm@gmail.com; wangchaoxuzhou@stu.xjtu.edu.cn).

D. W. K. Ng is with the School of Electrical Engineering and Telecommunications, The University of New South Wales, NSW 2052, Australia, and also with the Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany (e-mail: kwan@lnt.de).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2015.2465301

I. Introduction

FOLLOWING the pioneering work in [1], more and more attention has been dedicated to security issues in the physical layer from an information-theoretic perspective [2]–[19]. Recently, the multiple-input multiple-output (MIMO) processing has shown a great potential for enhancing the security of wireless transmissions [2]. Assuming the availability of full channel state information (CSI) of a multiple antenna eavesdropper, the secrecy capacity of a multiple-input, multiple-output, multiple antenna eavesdropper (MIMOME) wiretap channel is investigated in [3], and the optimal transmit signal for MIMOME wiretap channels is studied in [4]. The ergodic secrecy capacity of fast fading MIMOME wiretap channels with only statistical CSI at the transmitter has been investigated in [5]. However, in practice, eavesdroppers are usually passive and silent to hide their existences. Thus, the assumption of perfect eavesdropper CSI is too optimistic. To improve communication secrecy in practical systems, artificial noise assisted secure beamforming has been proposed in the literature when the eavesdroppers CSI is completely unkown [6], where carefully designed artificial noise is transmitted simultaneously with the confidential message to degrade the channel quality of the potential eavesdroppers.

A considerable body of literature has investigated the design of artificial noise schemes for physical-layer security under various scenarios [6]–[10]. Assuming only the channel distribution information (CDI) of the eavesdropper is available, the authors of [7] have studied the achievable ergodic secrecy rate (ESR) as well as the power allocation between data transmission and artificial noise in fast fading multiple-input single-output multiple-eavesdropper (MISOME) channels. The authors in [8] have proposed a joint design of the transmit power and the rate parameters of the wiretap code for secrecy throughput maximization in slow fading multiple-input single-output single-eavesdropper (MISOSE) channels. In [9], an optimal generalized artificial noise scheme has been investigated for ESR maximization in MISOSE channels. Assuming that the number of transmit antennas is sufficiently large, the authors in [10] have investigated the asymptotically optimal power allocation for both MISOSE and MIMOME channels. The idea of artificial noise has also been generalized to relay systems, and cooperative jamming has been extensively studied in [11]–[14].

The aforementioned works all assume that the CSI of *legitimate* receivers is known perfectly at the transmitter a prior. However, in practice, the CSI can only be collected through training and channel estimation. More importantly, the acquired

CSI is inevitably imperfect due to estimation errors. Therefore, it is critical to understand the sensitivity of the achievable secrecy rate of artificial noise assisted secure transmission schemes due to CSI imperfection, and to characterize the achievable secrecy rate under realistic CSI assumptions. Recently, several works have considered the impact of imperfect CSI of the legitimate receiver due to training and feedback errors, e.g., [15]–[19]. In [15], the optimal tradeoff between the energy used for training and data signals is studied for the MISOSE channel. Under the assumption that the perfect CSI is available at the legitimate receiver but the transmitter can obtain only quantized CSI through a feedback channel, the authors in [16] derive an analytical result of the achievable ESR for MISOSE channel. Furthermore, ESR and secrecy throughput are maximized by optimizing the power allocation between the artificial noise and confidential signal as well as the wiretap coding parameters in [17] and [18], respectively. In [19], the quantized legitimate CSI assumption has been generalized to cooperative jamming systems, and the authors optimize the feedback bit allocation for ESR maximization.

Acquiring CSI is a challenging and resource-consuming task in time-varying channels. The accuracy of the CSI acquisition through training depends heavily on the amount of training/feedback overhead and power consumption. In block-fading channels with a fixed coherence time, there is a non-trivial trade-off between CSI acquisition and data transmission: an exceedingly large amount of training/feedback overhead reduces the effective time spent on the data transmission, whereas less time spent on CSI acquisition results in a poor CSI quality, which also decreases the achievable secrecy rate. Therefore, considering the overhead for CSI acquisition, the *effective ESR* should be adopted as performance metric[1]. Similarly, a power consumption trade-off between training/feedback and secrecy date transmission also exists. However, a comprehensive investigation to the achievable effective ESR and the optimal training/feedback overhead and power consumption for artificial noise assisted secure transmission is still absent so far.

In this paper, we propose a framework for artificial noise assisted secure transmission in a frequency-division duplexed (FDD) MIMOME wiretap channel, where the legitimate CSI at the transmitter (CSIT) and receiver (CSIR) is obtained via forward training and analog feedback[2]. We provide a rigorous characterization of the achievable effective ESR by taking both the overhead and power consumption of training and feedback into consideration. Compared to the existing literatures [10]–[17], our contributions can be summarized as follows:

1) We establish a practical framework for artificial noise assisted secure transmission in the MIMOME wiretap channel by taking the CSI acquisition into consideration. Efficient algorithms are proposed for the *joint optimization of training/feedback overhead and power allocations* to maximize the achievable effective ESR,
2) In [10], [17], only the power allocation problem between the artificial noise and the information signal was considered for secrecy rate maximization. In this paper, we investigate a joint training/feedback overhead and power allocation optimization problem for effective ESR maximization. This leads to a comprehensive characterization of the tradeoff between CSI acquisition and secrecy data transmission.
3) In [10], perfect CSI of the legitimate channel is available at both the transmitter and receiver. In [15], the channel estimation error at the legitimate receiver is taken into account while the feedback channel is assumed to be ideal and error-free. Alternatively, the authors in [16], [17] assume perfect CSI of the legitimate channel at the receiver, while the transmitter obtains the quantized CSI through a feedback channel. In contrast, in this paper, we take into account both the CSI estimation error at the receiver and the feedback error at the transmitter, which constitutes a more practical and accurate model.

The remainder of the paper is organized as follows. In Section II, the considered system model is presented. The optimization of the power allocation and training overhead is investigated in Section III. Numerical results are provided in Section IV, and conclusions are drawn in Section V.

Notation: $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose, respectively. $\mathrm{tr}(\cdot)$ and $\det(\cdot)$ denote trace and determinant, respectively. $\mathrm{diag}(\mathbf{a})$ is a diagonal matrix with vector $\mathbf{a}$ on its main diagonal. $\Gamma(a,x) \triangleq \int_x^\infty t^{a-1}e^{-t}dt$ is the upper incomplete Gamma function([44], (8.35)). $\psi(\cdot)$ is Euler's digamma function ([45], (2.14)). $\mathcal{X}^2(v)$ denotes a central Chi-squared random variable with v degrees of freedom. $\mathbf{A} \sim \mathcal{W}_m(n,\mathbf{\Sigma})$ denotes a Wishart matrix with n degrees of freedom and covariance matrix $\mathbf{\Sigma}$. The factorial of a non-negative integer n is denoted by $n!$. $\mathbf{I}_N$ denotes an $N \times N$ identity matrix. $\mathbf{x} \sim \mathcal{CN}(\mathbf{\Lambda},\mathbf{\Delta})$ denotes a circular symmetric complex Gaussian vector with mean vector $\mathbf{\Lambda}$ and covariance matrix $\mathbf{\Delta}$. $||\cdot||_F$ denotes the Frobenius norm. $\mathbb{E}(\cdot)$ denotes statistical expectation, and $[\mathbf{b}]_i$ is the ith element of vector $\mathbf{b}$. $[a]^+ \triangleq \max(0,a)$. $x_n \overset{a.s.}{\to} l$ means that the series $x_1, x_2, \ldots$ converges to l almost surely.

II. System Model

We consider a standard MIMOME channel, where a transmitter (Alice), a legitimate receiver (Bob), and a passive eavesdropper (Eve) are equipped with N_T, N_B, and N_E antennas, respectively. In this paper, we consider a frequency division duplex (FDD) system and model all channels as uncorrelated Rayleigh fading. Let $\mathbf{s} \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_{N_B})$ be the unit-variance $N_B \times 1$ transmit symbol vector from Alice. Then, $\mathbf{s}$ is precoded by a precoding matrix $\mathbf{F} \in \mathbb{C}^{N_T \times N_B}$, which is transmitted

[1]Effective ESR is the achievable ESR excluding the training and feedback overhead. The mathematical definition is given by (12). In fact, the effective ergodic rate has been widely adopted as performance metric in single/multiuser downlink, interference alignment systems, and wireless power transfer system without secrecy considerations, e.g., [20]–[24], [27]–[31].

[2]Analog CSI feedback proposed in [22], [24] is an alternative CSI feedback scheme instead of the quantized channel feedback. Different from the quantized channel feedback, digitization and coding can be avoided by using the analog feedback, and therefore, analog CSI feedback is a fast and low-complexity CSI feedback scheme. It has been widely adopted in broadcast system [29], interference alignment [20], and wireless power transfer system [31].

to Bob via the legitimate MIMO channel $\mathbf{H}_B \in \mathbb{C}^{N_B \times N_T}$. Eve can eavesdrop the confidential information via a wiretap channel $\mathbf{H}_E \in \mathbb{C}^{N_E \times N_T}$. Similar to [6]–[8], [10], we assume that $N_T \geq \max(N_B, N_E)$, which is made to guarantee communication security, and the elements in $\mathbf{H}_B$ and $\mathbf{H}_E$ are modeled as independent and identically distributed (*i.i.d.*) complex Gaussian random variables with distribution $\mathcal{CN}(0,1)$.

In case that the legitimate channel $\mathbf{H}_B$ is perfectly known at Alice while the instantaneous wiretap channel $\mathbf{H}_E$ is completely unknown, artificial noise assisted transmission scheme is a promising scheme, where Alice transmits the secrecy signal using eigen-mode beamforming, and transmits artificial noise concurrently in the null-space of the signal space of Bob [6]–[8], [10]. Therefore, the precoding matrix $\mathbf{F}$ consists of the rightmost singular vectors corresponding to the N_B largest singular values of $\mathbf{H}_B$, and the artificial noise is in the form of $\mathbf{U}_N\mathbf{v}$, where $\mathbf{U}_N \in \mathbb{C}^{N_T \times (N_T - N_B)}$ should satisfy $\mathbf{H}_B\mathbf{U}_N = \mathbf{0}$, and $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_T - N_B})$. Here, we assume that Alice transmits multiple secrecy data streams with equal power[3]. When perfect $\mathbf{H}_B$ is available at Alice and Bob, the received signals at Bob and Eve are:

$$\mathbf{y}_B = \sqrt{\frac{P_S}{N_B}}\mathbf{H}_B\mathbf{F}\mathbf{s} + \mathbf{n}_B, \tag{1}$$

$$\mathbf{y}_E = \sqrt{\frac{P_S}{N_B}}\mathbf{H}_E\mathbf{F}\mathbf{s} + \sqrt{\frac{P_J}{N_T - N_B}}\mathbf{H}_E\mathbf{U}_N\mathbf{v} + \mathbf{n}_E, \tag{2}$$

respectively, where P_S is the transmit power of the secrecy signal, P_J is the transmit power of the artificial noise, $\mathbf{n}_B \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_B})$, and $\mathbf{n}_E \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_E})$ is the received noise vector at Bob and Eve, respectively.

In this paper, we consider a practical system, where the imperfect $\mathbf{H}_B$ at Bob and Alice is estimated via training and feedback. Although the eavesdropper has to estimate $\mathbf{H}_E$ as well in practice, we assume that $\mathbf{H}_E$ is perfectly known at the eavesdropper, which is the worst case for guaranteeing security. To evaluate the secrecy performance, the CDI of the eavesdropper is assumed to be available at Alice. This assumption has been adopted in the literatures for performance analysis and resource allocation [6]–[19].

In the following, we first characterize the estimation error at Alice by examining the training and feedback process in Section II.A. Then, we derive the achievable ESR with the estimated CSI in Section II.B. In Section II.C, with the channel estimation and analog feedback model, we characterize the effective ESR achieved with training and feedback.

A. Training and Analog Feedback Model

The three-phase training/feedback and transmission protocol adopted in this paper is illustrated in Fig. 1. At the beginning of a time slot, the imperfect estimate of $\mathbf{H}_B$ is acquired at Bob via the forward channel training. Then, Bob calculates the precoding matrix $\hat{\mathbf{F}}^b$ and artificial noise shaping matrix $\hat{\mathbf{U}}_N^b$ based on the imperfect estimate. Subsequently, Bob uses analog linear

[3]Although the water-filling power allocation can achieves better performance, as shown in [10], the performance gains achieved by the water-filling power allocation becomes negligible with the increasing SNR. For simplifying the theoretical analysis, equal power allocation is adopted in this paper.

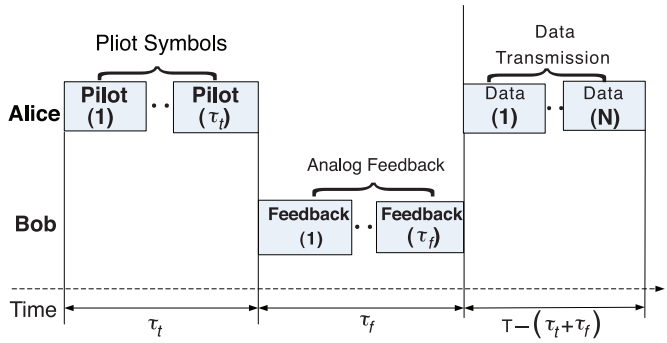

Fig. 1. A three-phase protocol for MIMO secrecy communication.

modulation (analog feedback) to transmit $\hat{\mathbf{F}}^b$ and $\hat{\mathbf{U}}_N^b$ to Alice over a noisy feedback channel, which result in imperfect estimations $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$ at Alice. Finally, Alice transmits the combination of the confidential signal and artificial noise using $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$ to Bob. This form of CSI acquisition, referred to as "closed-loop" CSI estimation [29], is relevant for FDD systems, where the forward and feedback channels are different. As an alternative, in time-division duplexing (TDD) systems, CSI can be obtained by reverse training from Bob [31]. In this paper, we focus on FDD systems, since FDD has dominated deployment in practice [29].

Remark 1: Analog CSI feedback was first proposed in [22], [24]. In contrast to the quantized channel feedback [19], digitization and coding can be avoided by using the analog feedback, and therefore, $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$ can be feedback very rapidly. Furthermore, if the feedback channel is an additive white Gaussian noise (AWGN) channel, and the CSI of the forward channel is perfectly known at the receiver, it has been shown in [22] that analog CSI feedback is optimal in that it achieves the minimum mean-squared error distortion. The studies on the application of analog CSI feedback in practical systems can be found in [22], [29], [31]. For example, in [22], an analog feedback framework for CDMA has been proposed, and in [29], an analog feedback framework for the MIMO fading broadcast channel has been proposed. Recently, the analog CSI feedback in the wireless power transfer systems has been studied in [31].

1) Forward Channel Training: In the first training phase, Alice sends an orthogonal pilot sequence matrix $\mathbf{\Phi}_f \in \mathbb{C}^{N_T \times \tau_t}$, i.e., $\mathbf{\Phi}_f\mathbf{\Phi}_f^H = \mathbf{I}_{N_T}$, over a training period of duration τ_t [27]. To obtain a reliable estimate of $\mathbf{H}_B$, the number of measurements should be no less than the number of unknowns, i.e., $\tau_t \geq N_T$. Bob observes

$$\mathbf{Y}_T = \sqrt{\frac{\tau_t\rho_t}{N_T}}\mathbf{H}_B\mathbf{\Phi}_f + \mathbf{V}_T, \tag{3}$$

where $\mathbf{V}_T \in \mathbb{C}^{N_B \times \tau_t}$ is the noise at Bob during forward channel training period with *i.i.d.* elements obeying the distribution $\mathcal{CN}(0,1)$, and ρ_t is the SNR. Based on $\mathbf{Y}_T$, Bob calculates the minimum mean square error (MMSE) estimate of the forward channel $\mathbf{H}_B$ as

$$\hat{\mathbf{H}}_B = \frac{\sqrt{\frac{\tau_t\rho_t}{N_T}}}{\frac{\tau_t\rho_t}{N_T} + 1}\mathbf{Y}_T\mathbf{\Phi}_f^H. \tag{4}$$

Considering the properties of the MMSE estimate [24], we know that $\hat{\mathbf{H}}_B$ has *i.i.d.* entries obeying $\mathcal{CN}(\mathbf{0}, 1-\sigma^2_{\mathbf{H}_B})$, and the corresponding estimation error $\tilde{\mathbf{H}}_B$ has *i.i.d.* entries obeying $\mathcal{CN}(\mathbf{0}, \sigma^2_{\mathbf{H}_B})$, where $\sigma^2_{\mathbf{H}_B} \triangleq \frac{1}{1+\tau_t\rho_t/N_T}$ and $\mathbf{H}_B = \hat{\mathbf{H}}_B + \tilde{\mathbf{H}}_B$.

Based on the imperfect estimate $\hat{\mathbf{H}}_B$, Bob calculates the precoding matrix $\hat{\mathbf{F}}^b$ and artificial noise matrix $\hat{\mathbf{U}}^b_N$. In particular, $\hat{\mathbf{F}}^b$ consists of the rightmost singular vectors corresponding to the N_B largest singular values of $\hat{\mathbf{H}}_B$, and $\hat{\mathbf{U}}^b_N$ satisfies $\hat{\mathbf{H}}_B\hat{\mathbf{U}}^b_N = \mathbf{0}$. After obtaining $\hat{\mathbf{F}}^b$ and $\hat{\mathbf{U}}^b_N$, Bob feeds back both $\hat{\mathbf{F}}^b$ and $\hat{\mathbf{U}}^b_N$ to Alice via analog feedback.

2) Analog Feedback: Instead of idealizing the feedback channel as a fixed-rate and error-free bit pipe, we explicitly consider the feedback from Bob to Alice over an additive white Gaussian noise (AWGN) channel, and the extension to the fading feedback channel is given in Appendix A[4]. For notational conciseness, we define $\mathbf{Z}^b \triangleq [\hat{\mathbf{F}}^b, \hat{\mathbf{U}}^b]$. As described in [24], we "spread" $\mathbf{Z}^b$ by post-multiplying $\mathbf{Z}^b$ by a prearranged $N_T \times \tau_f$ unitary spreading matrix $\mathbf{\Psi}$ where $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{I}_{N_T}$ (We require $\tau_f \geq N_T$ to obtain a reliable estimate). The feedback matrix $\mathbf{X}_f \in \mathbb{C}^{N_T\times\tau_f}$ transmitted by Bob can be modeled as [20], [24]

$$\mathbf{X}_f = \sqrt{\frac{\tau_f\rho_f}{N_T}}\mathbf{Z}^b\mathbf{\Psi}, \tag{5}$$

where ρ_f is the SNR during the analog feedback. We assume that the least squares (LS) estimator is adopted at Alice for estimating $\mathbf{Z}^b$ [26]. Based on the analog feedback in (5), the LS estimate of $\mathbf{Z}^b$ at Alice is obtained as

$$\hat{\mathbf{Z}} = \mathbf{Z}^b + \sqrt{\frac{N_T}{\tau_f\rho_f}}\mathbf{V}_f, \tag{6}$$

where $\mathbf{V}_f$ is the noise at Alice during the analog feedback. The corresponding LS channel estimation error matrix $\tilde{\mathbf{Z}} \triangleq \sqrt{\frac{N_T}{\tau_f\rho_f}}\mathbf{V}_f$ has *i.i.d.* entries obeying $\mathcal{CN}(\mathbf{0}, \sigma^2_{\mathbf{Z}})$, where $\sigma^2_{\mathbf{Z}} \triangleq \frac{N_T}{\tau_f\rho_f}$. We now have $\hat{\mathbf{Z}} = \mathbf{Z}^b + \tilde{\mathbf{Z}}$, and the estimates $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$ at Alice are given by $\hat{\mathbf{F}} = \hat{\mathbf{F}}^b + \tilde{\mathbf{F}}$, and $\hat{\mathbf{U}}_N = \hat{\mathbf{U}}^b_N + \tilde{\mathbf{U}}_N$, where $\tilde{\mathbf{F}}$ and $\tilde{\mathbf{U}}_N$ are the corresponding estimation error matrix, which have *i.i.d.* entries obeying $\mathcal{CN}(0, \sigma^2_{\mathbf{Z}})$. In practice, τ_f and ρ_f should be large enough such that $\sigma^2_{\mathbf{Z}} \ll 1$ for getting sufficiently accurate estimates $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$. Therefore, in the following, we approximate $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$ as unitary matrices to get a mathematical tractable result. Simulation results in Fig. 4 validate the adopted approximation.

[4]In this paper, to simplify the analysis, just as in [22], [23], [29], [30], [31], we consider a simplified setting where the feedback channel is an AWGN channel. The non-fading feedback channel can be regarded as an approximation of the feedback channel when the number of antennas equipped at Alice is large, since the channel becomes deterministic with a large number of antennas, due to the channel hardening. In Appendix A, we have shown that the fading feedback channel can be incorporated in our considered model. But the model obtained is too complicated, which does not lead to a tractable problem.

B. Artificial Noise Assisted Secure Transmission Scheme

Using the estimates $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$, Alice transmits the combination of confidential signal and artificial noise to Bob. The received signal vector at Bob is given by (7) at the bottom of the page. Due to CSI imperfection, the artificial noise cannot be eliminated completely at Bob, which deteriorates the performance of achievable secrecy rate. Furthermore, imperfect knowledge of $\mathbf{H}_B$ at Bob implies that signal misdetection may occur which reduces the effective received signal-to-interference-and-noise-ratio (SINR).

At Eve, the received signal vector is given by

$$\mathbf{y}_E = \sqrt{\frac{P_S}{N_B}}\mathbf{H}_E\hat{\mathbf{F}}\mathbf{s} + \sqrt{\frac{P_J}{N_T-N_B}}\mathbf{H}_E\hat{\mathbf{U}}_N\mathbf{v} + \mathbf{n}_E. \tag{8}$$

From [27], ([28], Lemma 1), [29], we know that the exact capacity expression under imperfect CSIR is still unavailable. Since only $\hat{\mathbf{H}}_B\hat{\mathbf{F}}^b$ is available at Bob from the forward channel training, the exact capacity of Bob is difficult to get from (7). To obtain a tractable problem, we resort to deriving a lower bound of Bob's capacity. Firstly, $\tilde{\mathbf{H}}_B$ is independent of $\hat{\mathbf{H}}_B$ due to the property of MMSE estimator. Secondly, $\tilde{\mathbf{F}}$ is independent of $\hat{\mathbf{F}}^b$, since $\tilde{\mathbf{F}}$ reflects the distortion due to noise. Thus both the estimate error terms and the artificial noise leakage terms in (7) are uncorrelated with the signal term. Then, exploiting the widely adopted result in [27], ([28], Lemma 1), a tight lower bound[5] of Bob's capacity can be obtained by considering the worst case scenario, where

$$\begin{aligned}&\sqrt{\frac{P_S}{N_B}}\tilde{\mathbf{H}}_B\hat{\mathbf{F}}^b\mathbf{s}\\ &+\sqrt{\frac{P_S}{N_B}}\hat{\mathbf{H}}_B\tilde{\mathbf{F}}\mathbf{s} + \sqrt{\frac{P_S}{N_B}}\tilde{\mathbf{H}}_B\tilde{\mathbf{F}}\mathbf{s} + \sqrt{\frac{P_J}{N_T-N_B}}\tilde{\mathbf{H}}_B\hat{\mathbf{U}}^b_N\mathbf{v}\\ &+\sqrt{\frac{P_J}{N_T-N_B}}\hat{\mathbf{H}}_B\tilde{\mathbf{U}}_N\mathbf{v} + \sqrt{\frac{P_J}{N_T-N_B}}\tilde{\mathbf{H}}_B\tilde{\mathbf{U}}_N\mathbf{v}\end{aligned}$$

[5]The lower bound has been shown to be tight in [27], [28].

$$\begin{aligned}\mathbf{y}_B &= \sqrt{\frac{P_S}{N_B}}\left(\hat{\mathbf{H}}_B+\tilde{\mathbf{H}}_B\right)\left(\hat{\mathbf{F}}^b+\tilde{\mathbf{F}}\right)\mathbf{s} + \sqrt{\frac{P_J}{N_T-N_B}}\left(\hat{\mathbf{H}}_B+\tilde{\mathbf{H}}_B\right)\left(\hat{\mathbf{U}}^b_N+\tilde{\mathbf{U}}_N\right)\mathbf{v} + \mathbf{n}_B\\ &= \sqrt{\frac{P_S}{N_B}}\hat{\mathbf{H}}_B\hat{\mathbf{F}}^b\mathbf{s} + \underbrace{\sqrt{\frac{P_S}{N_B}}\tilde{\mathbf{H}}_B\hat{\mathbf{F}}^b\mathbf{s} + \sqrt{\frac{P_S}{N_B}}\hat{\mathbf{H}}_B\tilde{\mathbf{F}}\mathbf{s} + \sqrt{\frac{P_S}{N_B}}\tilde{\mathbf{H}}_B\tilde{\mathbf{F}}\mathbf{s}}_{\text{CSI estimation error}}\\ &\quad + \underbrace{\sqrt{\frac{P_J}{N_T-N_B}}\tilde{\mathbf{H}}_B\hat{\mathbf{U}}^b_N\mathbf{v} + \sqrt{\frac{P_J}{N_T-N_B}}\hat{\mathbf{H}}_B\tilde{\mathbf{U}}_N\mathbf{v} + \sqrt{\frac{P_J}{N_T-N_B}}\tilde{\mathbf{H}}_B\tilde{\mathbf{U}}_N\mathbf{v}}_{\text{Artificial noise leakage}} + \mathbf{n}_B.\end{aligned} \tag{7}$$

are taken as uncorrelated Gaussian additive noise with the variance $(P_S+P_J)(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}})$. Hence, a capacity lower bound R_B for Bob can be calculated as

$$R_B = \log_2\left(\det\left(\mathbf{I}_{N_B} + \frac{P_S\left(1-\sigma^2_{\mathbf{H}_B}\right)}{1+\left(P_S+P_J\right)\left(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}}\right)}\frac{\bar{\mathbf{H}}\bar{\mathbf{H}}^H}{N_B}\right)\right), \quad (9)$$

where $\bar{\mathbf{H}} \triangleq \frac{\hat{\mathbf{H}}_B}{\sqrt{1-\sigma^2_{\mathbf{H}_B}}}$ is the normalized channel estimate and each entry of $\bar{\mathbf{H}}$ is *i.i.d.* complex Gaussian with distribution $\mathcal{CN}(0,1)$. Define $\rho_{\text{eff}} = \frac{P_S(1-\sigma^2_{\mathbf{H}_B})}{1+(P_S+P_J)(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}})}$, which can be regarded as the effective average SINR at Bob.

Considering the worst case of secure communication where the perfect CSI, i.e., both $\mathbf{H}_E\hat{\mathbf{F}}$ and $\mathbf{H}_E\hat{\mathbf{U}}_N$, are available at the eavesdropper, the instantaneous capacity of Eve can be calculated from (8), which is given by

$$R_E \triangleq \log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_S}{N_B}\mathbf{H}_E\hat{\mathbf{F}}\hat{\mathbf{F}}^H\mathbf{H}_E^H \times\left(\mathbf{I}_{N_E}+\frac{P_J}{N_T-N_B}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)^{-1}\right)\right). \quad (10)$$

In [6], [10], [19], it was shown that by using Gaussian inputs and stochastic encoders, the achievable ESR C_S can be calculated as

$$C_S \triangleq [\mathbb{E}(R_B)-\mathbb{E}(R_E)]^+. \quad (11)$$

C. Effective ESR

In this paper, we adopt a block fading channel model to perform the training and feedback optimization, where the channels in the considered model remain constant for a coherence time of T, but vary independently from block to block. The coherence time can be calculated as $T=\frac{1}{2f_D}$, where f_D is the block fading channels' *effective Doppler spread* [20]. In such a channel, both the channel estimation and payload data transmission should be finished within T, otherwise, the estimated CSI would be outdated. From the training and feedback model, only $T-(\tau_t+\tau_f)$ amount of time is spent on the data transmission. For such a signaling overhead model, the achievable effective ESR, C_{eff} can be calculated by

$$C_{\text{eff}} \triangleq \left(\frac{T-(\tau_t+\tau_f)}{T}\right)C_S. \quad (12)$$

Obviously, increasing the CSI acquisition overheads and power consumptions, i.e., τ_t, ρ_t, and τ_f, ρ_f, would improve the quality of the estimated CSI, but reduces the time and power for the secrecy transmission. On the contrary, less CSI acquisition overhead allows more time and power for the secrecy information transmission, but the poor quality of the estimated CSI would increase the noise variances $\sigma^2_{\mathbf{H}_B}$ and $\sigma^2_{\mathbf{Z}}$ and thus reduces C_S. As a result, there is a non-trivial trade-off in the resource allocation between the channel estimation and the data transmission for maximizing the achievable effective ESR. Besides, the power allocation between the artificial noise and the secrecy signal plays an important role in the system performance and should also be addressed.

In the following, we optimize $\tau_t, \rho_t, \tau_f, \rho_f$, P_S, and P_J jointly for maximizing the effective ESR.

III. Overhead Optimization and Power Allocation in MIMOME Channels

Before proceeding, for convenience, we denote $\tau_t = \theta\varphi T$, $\tau_f = \theta\psi T$ for $0\le\theta\le 1$, $0\le\varphi\le 1$, $0\le\psi\le 1$, and $\varphi+\psi = 1$. θT denotes the total time for CSI training and feedback. Then, the effective ESR maximization can be formulated as

$$\max_{P_S,P_J,\rho_f,\rho_t,\theta,\varphi,\psi}(1-\theta)C_S \quad (13a)$$

$$\text{s.t.} \quad (P_S+P_J)(1-\theta)T+\rho_f\tau_f+\rho_t\tau_t \le PT, \quad (13b)$$

$$N_T \le \theta\varphi T, \quad N_T \le \theta\psi T, \quad 0\le\theta\le 1, \quad \varphi+\psi=1, \quad (13c)$$

where PT denotes the maximal transmit power during the coherence time T. Note that θ, φ, ψ, and ρ_f, ρ_t characterize both the time overhead and power consumption for training and feedback, and P_S, P_J show the power allocation between confidential signal and artificial noise.

To solve problem (13), we should derive the analytical expression of C_S. Although the analytical expression of C_S can be derived by using ([32], Theorem 1), ([33], Theorem 1), the obtained result is too cumbersome and does not facilitate the algorithm design for optimizing the overhead and power allocation. As an alternative, we resort to a deterministic approximation of the achievable C_S by considering $N_T, N_B \to +\infty$. The main result is summarized in following theorem based on the random matrix theory.

Theorem 1: With ratios $\beta_1 \triangleq \frac{N_B}{N_E}$, $\beta_2 \triangleq \frac{N_T-N_B}{N_E}$, and $\beta_3 \triangleq \frac{N_T}{N_B}$, we have

$$\lim_{N_T,N_B\to+\infty}(C_S-C_S^o) \overset{a.s.}{\to} 0, \quad (14)$$

where C_S^o is given by (15),

$$\begin{aligned} C_S^o \triangleq \Bigg[& N_B\log_2\left(1+\omega-F\left(\beta_3,\frac{\omega}{\beta_3}\right)\right) \\ & + N_T\log_2\left(1+\frac{\omega}{\beta_3}-F\left(\beta_3,\frac{\omega}{\beta_3}\right)\right) \\ & + N_E\log_2\left(1+P_J-F\left(\beta_2,\frac{P_J}{\beta_2}\right)\right) \\ & - N_T\frac{\log_2(e)}{\omega}F\left(\beta_3,\frac{\omega}{\beta_3}\right) \\ & + (N_T-N_B)\log_2\left(1+\frac{P_J}{\beta_2}-F\left(\beta_2,\frac{P_J}{\beta_2}\right)\right) \\ & - (N_T-N_B)\frac{\log_2(e)}{P_J}F\left(\beta_2,\frac{P_J}{\beta_2}\right) \\ & - N_B\log_2\left(1+P_S\frac{\eta}{\beta_1}\right) \\ & - (N_T-N_B)\log_2\left(1+\frac{\eta P_J N_B}{\beta_1(N_T-N_B)}\right) \\ & + N_E\log_2(\eta) - N_E(\eta-1)\log_2(e)\Bigg]^+, \end{aligned} \quad (15)$$

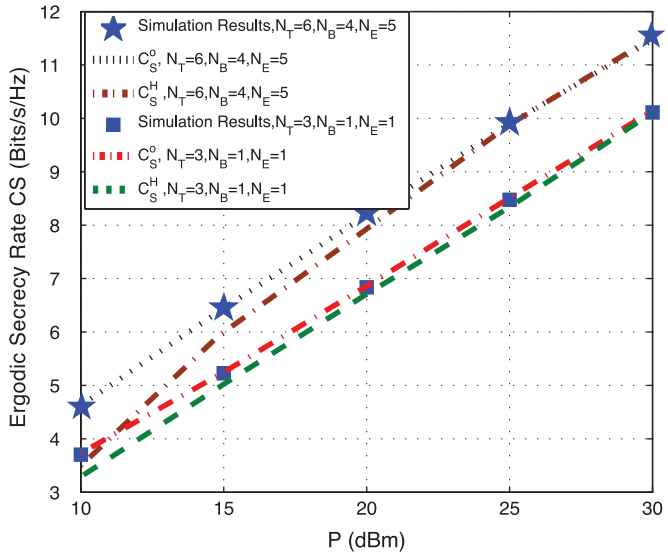

Fig. 2. Comparison between the simulations and the deterministic approximation C_S^o in (15) and the high-SNR approximation C_S^H in (19) with $\sigma^2_{\mathbf{H}_B} = \sigma^2_{\mathbf{Z}} = 10^{-3}$.

$$\omega \triangleq \frac{N_T P_S \left(1-\sigma^2_{\mathbf{H}_B}\right)}{N_B\left(1+\left(P_S+P_J\right)\left(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}}\right)\right)}, \tag{16}$$

$$F(x,y) \triangleq \frac{1}{4}\left(\sqrt{1+y\left(1+\sqrt{x}\right)^2}-\sqrt{1+y\left(1-\sqrt{x}\right)^2}\right)^2, \tag{17}$$

and η^* is the unique positive solution satisfying

$$\eta^* = \Gamma(\eta^*) \triangleq \left(1+\frac{N_B}{N_E}\frac{P_S}{P_S\eta^*+\beta_1} + \frac{N_T-N_B}{N_E}\frac{P_S P_J N_B}{P_S P_J \eta^* N_B+\beta_1 P_S\left(N_T-N_B\right)}\right)^{-1}. \tag{18}$$

Proof: The proof is given in Appendix B. ∎

Remark 2: We should highlight that η^* satisfying (18) is unique, since it can be proved that $\Gamma(\eta^*)$ satisfies the following properties of the standard interference function, i.e., a) Positivity $\Gamma(\eta^*) > 0$; b) Monotonicity. If $\eta^* > \hat{\eta}^*$, then $\Gamma(\eta^*) > \Gamma(\hat{\eta}^*)$; c) Scalability. For all $\alpha > 1$, $\alpha\Gamma(\eta^*) > \Gamma(\alpha\eta^*)$. Therefore, we can conclude that $\Gamma(\eta^*)$ is the standard interference function in [35], and η^* satisfying (18) is unique according to ([35], Theorem 1).

As shown in ([34], Table I), the asymptotic expression of C_S^o for large N_T and N_B yields an extremely accurate approximation even if the number of antennas is small. The accuracy of the adopted approximation is validated by the simulation results in Fig. 2. We can find from Fig. 2 that even though the adopted N_T, N_B, N_E are small, e.g., $N_T = 3, N_B = 1$, and $N_E = 1$, C_S^o still accurately approximates C_S. Thus with the deterministic approximation in (15), the objective function in (13) can be approximated by $C^o_{\text{eff}} = (1-\theta)C_S^o$.

Note that C_S^o is a function of the powers P_S, P_J, and the CSI estimation errors $\sigma^2_{\mathbf{H}_B}, \sigma^2_{\mathbf{Z}}$, which is determined by $\rho_f.\rho_t, \theta, \varphi$, and ψ. Therefore, C^o_{eff} is very complicated and nonconvex. In general, a brute force optimization approach may be required and the computational complexity is very high. In the following subsections, we first investigate the solution structures in the high- and low-SNR regimes, respectively, and provide efficient iterative algorithms for maximizing C^o_{eff}. Finally, we will discuss the case of general SNR, and maximize the lower bound of C^o_{eff}.

A. The High-SNR Regime

The high-SNR approximation of C_S^o is summarized in the following lemma.

Lemma 1: In the high-SNR regime, C_S^o can be approximated by

$$C_S^o \approx C_S^H \triangleq \left[g_1(P_S,P_J)+g_2\left(P_S,P_J,\sigma^2_{\mathbf{H}_B},\sigma^2_{\mathbf{Z}}\right)\right]^+, \tag{19}$$

where $g_1(P_S,P_J)$ and $g_2(P_S,P_J,\sigma^2_{\mathbf{H}_B},\sigma^2_{\mathbf{Z}})$ are given by (20) and (21).

$$\begin{aligned} g_1(P_S,P_J) &\triangleq N_B\log_2\left(\frac{N_T P_S}{N_B e}\right) - N_B\left(\beta_3-1\right)\log_2\left(1-\frac{1}{\beta_3}\right) \\ &\quad + \Omega(\beta_2) - N_B\log_2\left(1+P_S\frac{\eta^*}{\beta_1}\right) \\ &\quad - (N_T-N_B)\log_2\left(1+\frac{\eta^* P_J N_B}{\beta_1(N_T-N_B)}\right) \\ &\quad + N_E\log_2(\eta^*) - N_E\left(\eta^*-1\right)\log_2(e), \end{aligned} \tag{20}$$

$$g_2\left(P_S,P_J,\sigma^2_{\mathbf{H}_B},\sigma^2_{\mathbf{Z}}\right) \triangleq N_B\log_2\left(\frac{1-\sigma^2_{\mathbf{H}_B}}{1+\left(P_S+P_J\right)\left(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}}\right)}\right), \tag{21}$$

η^* is the solution of (18), and

$$\Omega(\beta_2) \triangleq \begin{cases} N_E\log_2\left(\frac{P_J}{e}\right) - N_E\left(\beta_2-1\right)\log_2\left(1-\frac{1}{\beta_2}\right), \\ \qquad \text{if } \beta_2 \geq 1 \\ \left(N_T-N_B\right)\log_2\left(\frac{P_J}{e}\right) - N_E\left(1-\beta_2\right)\log_2(1-\beta_2), \\ \qquad \text{if } \beta_2 \leq 1 \end{cases} \tag{22}$$

Proof: The proof is given in Appendix C. ∎

In Fig. 2, we validate the deterministic approximation C_S^o in (15) and the high-SNR approximation C_S^H in (19). From Fig. 2, we can find that C_S^H is very accurate at moderate to high P, which is the operating regions for most high speed wireless communication systems. Therefore, in the following, we approximate C_S^o by C_S^H and optimize $(1-\theta)C_S^H$ for maximizing the achievable effective ESR.

Although compared with C_S^o, C_S^H has been much simplified, the optimization problem obtained is still intractable due to the interaction between the optimization variables and the implicit parameter η^*, which is coupled through a complicated fixed-point equality constraint (18). For obtaining a tractable problem, we first rewrite C_S^H in an equivalent form via removing the fixed-point equality constraints. Then, we propose a block coordinate descent method (BCDM) ([36], Section 2.7) for handling the joint optimization problem. For brevity, we assume that $\beta_2 \geq 1$ in the following, and the obtained analytical results can be directly generalized to the case $\beta_2 < 1$.

For notational simplicity, we define

$$\begin{aligned} f(P_S,P_J,\eta) \triangleq\ & N_B\log_2\left(\frac{N_TP_S}{N_Be}\right)+N_E\log_2\left(\frac{P_J}{e}\right)\\ & -N_B\log_2\left(1+P_S\frac{\eta}{\beta_1}\right)-N_E\left(\eta-1\right)\log_2(e)\\ & -N_B\left(\beta_3-1\right)\log_2\left(1-\frac{1}{\beta_3}\right)\\ & -N_E\left(\beta_2-1\right)\log_2\left(1-\frac{1}{\beta_2}\right)\\ & -(N_T-N_B)\log_2\left(1+\frac{\eta P_JN_B}{\beta_1(N_T-N_B)}\right)\\ & +N_E\log_2(\eta), \end{aligned} \tag{23}$$

and then we have the following theorem.

Theorem 2: $g_1(P_S,P_J)$ can be expressed as

$$g_1(P_S,P_J)=\max_{\eta} f(P_S,P_J,\eta). \tag{24}$$

Proof: The proof is given in Appendix D. ∎

With Theorem 2, stacking all the optimization variables into a vector $\mathbf{y}\triangleq\{P_S,P_J,\rho_f,\rho_t,\varphi,\psi,\eta\}$, the joint optimization problem (13) in the high-SNR regime can be reformulated as

$$\begin{aligned} &\max_{\mathbf{y},\theta}\Upsilon(\mathbf{y},\theta)\triangleq(1-\theta)\left(f(P_S,P_J,\eta)+g_2\left(P_S,P_J,\sigma^2_{\mathbf{H}_B},\sigma^2_{\mathbf{Z}}\right)\right)\\ &\text{s.t.}\quad (13b)\text{ and }(13c). \end{aligned} \tag{25}$$

We can see that using Theorem 2, the fixed-point equation constraint on η has been eliminated, and we obtain an equivalent problem (25), which is still nonconvex. In the following, we will show that a stationary local optimum $P_S^*,P_J^*,\rho_f^*.\rho_t^*,\varphi^*,\psi^*,\eta^*$ and θ^* can be obtained by BCDM.

Firstly, we show that $\Upsilon(\mathbf{y},\theta)$ is a concave function of θ by the following theorem.

Theorem 3: $\Upsilon(\mathbf{y},\theta)$ is a concave function of θ.

Proof: $\Upsilon(\mathbf{y},\theta)$ can be rewritten as

$$\begin{aligned} \Upsilon(\mathbf{y},\theta)=&(1-\theta)\\ &\times\left(f(P_S,P_J,\eta)+N_B\log_2\left(\frac{a\theta^2}{a\theta^2+b\theta+c}\right)\right), \end{aligned} \tag{26}$$

where $a\triangleq\frac{\varphi\psi T^2\rho_t\rho_f}{N_T}$, $b\triangleq\psi T\rho_f+(P_S+P_J)(\psi T\rho_f+\varphi T\rho_tN_T)$, and $c\triangleq(P_S+P_J)N_T^2$.

The first derivative of $\Upsilon(\mathbf{y},\theta)$ with respect to θ is

$$\begin{aligned} \frac{d\Upsilon\left(\mathbf{y},\theta\right)}{d\theta}=&-f(P_S,P_J,\eta)-N_B\log_2\left(\frac{a\theta^2}{a\theta^2+b\theta+c}\right)\\ &+(1-\theta)\frac{N_B}{\ln 2}\frac{b\theta+2c}{\theta\left(a\theta^2+b\theta+c\right)}. \end{aligned} \tag{27}$$

It is not hard to see that the first term is irrelative to θ, and the second and the third terms are decreasing functions of θ. Therefore, we can conclude that $\frac{d^2\Upsilon(\mathbf{y},\theta)}{d^2\theta}<0$ and $\Upsilon(\mathbf{y},\theta)$ is a concave function of θ. ∎

Therefore, at each iteration in BCDM, for a fixed $\mathbf{y}$, the optimal θ for maximizing $\Upsilon(\mathbf{y},\theta)$ is the solution of $\frac{d\Upsilon(\mathbf{y},\theta)}{d\theta}=0$ which can be located by a bisection search. Then, we will show that with a given θ, the optimal $\mathbf{y}$ for maximizing $\Upsilon(\mathbf{y},\theta)$ can be located by the following procedures. Introducing the following variables $x_{P_S}=\ln(P_S)$, $x_{P_J}=\ln(P_J)$, $x_{\rho_t}=\ln(\rho_t)$, $x_{\rho_f}=\ln(\rho_f)$, $x_{\varphi}=\ln(\varphi)$, $x_{\eta}=\ln(\eta)$, and $x_{\psi}=\ln(\psi)$, and defining $\mathbf{w}_x\triangleq\{x_{P_S},x_{P_J},x_{\rho_t},x_{\rho_f},x_{\varphi},x_{\eta},x_{\psi}\}$, the objective of the problem (25) can be reformulated as $\Upsilon(\mathbf{w}_x,\theta)$ given by (28),

$$\begin{aligned} \Upsilon(\mathbf{w}_x,\theta)=&(1-\theta)\left(N_Ex_{P_J}\log_2(e)-N_E(e^{x_\eta}-1)\log_2(e)\right.\\ &+N_Bx_{P_S}\log_2(e)-N_B\log_2\left(1+\frac{e^{x_{P_S}+x_\eta}}{\beta_1}\right)\\ &+N_Ex_\eta\log_2(e)-(N_T-N_B)\log_2\left(1+\frac{e^{x_{P_J}+x_\eta}N_B}{\beta_1(N_T-N_B)}\right)\\ &+N_B(x_\varphi+x_\psi+x_{\rho_t}+x_{\rho_f})\log_2(e)\\ &-N_B\log_2\left(\frac{e^{x_\varphi+x_\psi+x_{\rho_t}+x_{\rho_f}}\theta^2T^2}{N_T}\right.\\ &+\left(e^{x_{\rho_f}+x_\psi}T+e^{x_{\rho_f}+x_\psi}T\left(e^{x_{P_S}}+e^{x_{P_J}}\right)\right.\\ &\left.\left.+N_TTe^{x_\varphi+x_{\rho_t}}\left(e^{x_{P_S}}+e^{x_{P_J}}\right)\right)\theta\right.\\ &\left.+N_T^2\left(e^{x_{P_S}}+e^{x_{P_J}}\right)\right)+\Delta\Big), \end{aligned} \tag{28}$$

where Δ denotes the collection of those constant terms independent of $\mathbf{w}_x$.

Since $\log_2(1+\sum_{i=1}^{N}e^{x_i})$ is a convex function of $\{x_1,\ldots,x_N\}$ [47], $\Upsilon(\mathbf{w}_x,\theta)$ is a concave function of $\mathbf{w}_x$. Then, given θ, considering the following problem

$$\max_{\mathbf{w}_x}\Upsilon(\mathbf{w}_x,\theta) \tag{29a}$$

$$\text{s.t.}\quad (e^{x_{P_S}}+e^{x_{P_J}})(1-\theta)T+e^{x_{\rho_f}+x_{\tau_f}}+e^{x_{\rho_t}+x_{\tau_t}}\le PT, \tag{29b}$$

$$e^{x_\varphi}+e^{x_\psi}\le 1, \tag{29c}$$

$$\ln\left(\frac{N_T}{\theta T}\right)\le x_\varphi,\ln\left(\frac{N_T}{\theta T}\right)\le x_\psi, \tag{29d}$$

in case that the constraint (29c) is active, the optimal solution of the problem (29) is also the optimal solution of the problem (25). In fact, constraint (29c) should be active, otherwise, x_φ and x_ψ can increase further to increase the objective function of the problem (29), which would lead to a contradiction. With a fixed θ, the problem (29) is a convex optimization problem, whose optimum can be located by the standard interior-point algorithm. Accordingly, the optimal $\mathbf{y}$ can be obtained.

The proposed BCDM algorithm is summarized in Algorithm 1.

Algorithm 1: Proposed BCDM Algorithm for Solving the problem (25)

1. **Initialize** $n=1$, $0<\theta(0)<1$, and accuracy $\epsilon>0$,
2. **while** $\Upsilon(\mathbf{y}(n),\theta(n))-\Upsilon(\mathbf{y}(n-1),\theta(n-1))\ge\epsilon$,
3. Solving the convex optimization problem (29), and the optimal $\mathbf{y}(n)$ can be obtained as $e^{x_{P_S}}=P_S,e^{x_{P_J}}=P_J,e^{x_{\rho_t}}=\rho_t,e^{x_{\rho_f}}=\rho_f,e^{x_\varphi}=\varphi,e^{x_\eta}=\eta$, and $e^{x_\psi}=\psi$,
4. Using the bisection search, the optimal $\theta(n)$ can be calculated by solving the following equation
$$\frac{d\Upsilon\left(\mathbf{y}(n),\theta\right)}{d\theta}=0 \tag{30}$$
5. **end while**.

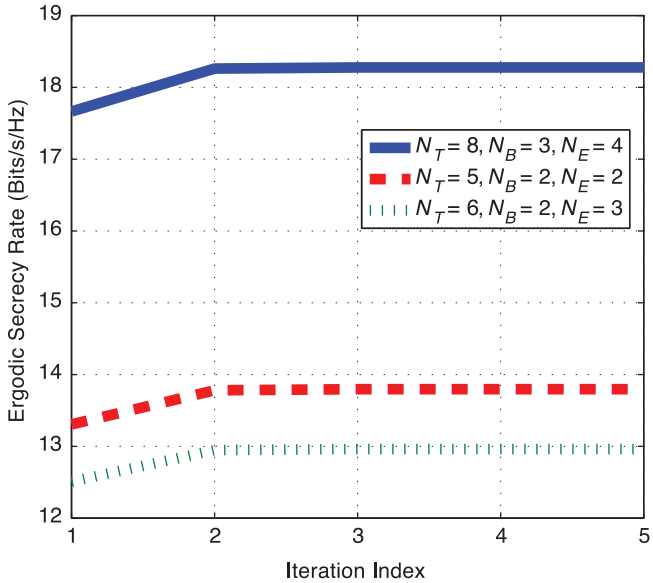

Fig. 3. The convergence rate of Algorithm 1 with $P = 30$ dBm and $\epsilon = 10^{-5}$.

For the convergence of the algorithm, we have the following corollary.

Corollary 1: The limiting solution generated by Algorithm 1 is a stationary point of the joint optimization problem (25).

Proof: It can be verified that the objective function $\Upsilon(\mathbf{y}, \theta)$ is continuously differentiable, and the feasible set is closed, nonempty, and convex. Moreover, since $\Upsilon(\mathbf{y}, \theta)$ is bounded, by Bolzano-Weierstrass theorem, we know that $\mathbf{y}(n)$ and $\theta(n)$ must have limit points. Therefore, invoking ([37], Corollary 2), we conclude that every limit point $\mathbf{y}(n), \theta(n)$ is a stationary point of the joint optimization problem (25). ∎

The convergence rate of Algorithm 1 is shown in Fig. 3. We can find that Algorithm 1 converges to a stationary point within 5 iterations for $\epsilon = 10^{-5}$ and the convergence rates are almost the same for different N_T, N_B, and N_E. Therefore, simulation results show that Algorithm 1 is an efficient optimization algorithm for handling the nonconvex problem (25).

B. The Low-SNR Regime

In this subsection, we design $P_S, P_J, \rho_f, \rho_t, \varphi, \psi, \eta$, and θ jointly to maximize C_{eff} in the low-SNR regime. In this case, we can use the first-order expansion of the mutual information around $P = 0$ to get an approximation of the achievable ESR. Using ([38], (40)), we have

$$\begin{aligned}\mathbb{E}(R_B) &= \frac{P_S\left(1-\sigma^2_{\mathbf{H}_B}\right)\mathbb{E}\left(\mathrm{tr}\left(\bar{\mathbf{H}}\bar{\mathbf{H}}^H\right)\right)}{N_B\left(1+\left(P_S+P_J\right)\left(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}}\right)\right)}+o(P)\\ &= \frac{N_T P_S\left(1-\sigma^2_{\mathbf{H}_B}\right)}{1+\left(P_S+P_J\right)\left(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}}\right)}+o(P), \qquad (31)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}(R_E) &= \frac{(1-\theta)P}{N_B}\mathbb{E}\left(\mathrm{tr}\left(\mathbf{H}_E\mathbf{F}\mathbf{F}^H\mathbf{H}_E^H\right)\right)+o(P)\\ &= N_E(1-\theta)P+o(P), \qquad (32)\end{aligned}$$

where $o(P)$ denotes the higher order infinitesimal terms with respect to P. Neglecting the higher order infinitesimal terms, the approximation of the achievable ESR in the low-SNR regime, i.e., C_S^L, is given by

$$C_S \approx C_S^L \triangleq \left[\frac{N_T P_S\left(1-\sigma^2_{\mathbf{H}_B}\right)}{1+\left(P_S+P_J\right)\left(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}}\right)}-N_E P_S\right]^+. \qquad (33)$$

We immediately have the following corollary.

Corollary 2: In the low-SNR regime, for $\frac{N_T(1-\sigma^2_{\mathbf{H}_B})}{1+(P_S+P_J)(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}})} > N_E$, the positive ESR can be achieved and the optimal $P_J^* = 0$, i.e., no artificial noise is transmitted. For $\frac{N_T(1-\sigma^2_{\mathbf{H}_B})}{1+(P_S+P_J)(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}})} \le N_E$, the achievable ESR is zero.

Proof: When $\frac{N_T(1-\sigma^2_{\mathbf{H}_B})}{1+(P_S+P_J)(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}})} > N_E$, C_S^L is a decreasing function of P_J. The corollary can be obtained from (33) directly. ∎

With $P_J^* = 0$, we optimize the remainder system parameters jointly for maximizing C_{eff} in the low-SNR regime in the following. From (33), the effective ESR can be approximated by $(1-\theta)C_S^L$ and the optimization problem can be formulated as follows

$$\begin{aligned}\max_{\substack{P_S,\theta,\varphi,\psi,\\ \rho_t,\rho_f}} \ & (1-\theta)\left(\frac{N_T P_S\left(1-\sigma^2_{\mathbf{H}_B}\right)}{1+P_S\left(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}}\right)}-N_E P_S\right)\\ \text{s.t.} \ & P_S(1-\theta)T+\rho_f\tau_f+\rho_t\tau_t \le PT \text{ and } (13c). \qquad (34)\end{aligned}$$

Introducing the slack variables t_1 and t_2 and using (26), we formulate the following problem:

$$\begin{aligned}&\min_{P_S,\theta,\varphi,\rho_t,\rho_f,\psi,t_1,t_2} \ \frac{1}{t_1 t_2} && (35a)\\ &\text{s.t.} \quad t_1+\theta \le 1, && (35b)\\ &\frac{(t_2+N_E P_S)(a\theta^2+b\theta+c)}{N_T P_S a\theta^2} \le 1, && (35c)\\ &N_T \le \theta\varphi T, N_T \le \theta\psi T, 0 \le \theta \le 1, && (35d)\\ &\varphi+\psi \le 1, && (35e)\\ &\frac{P_S T+\rho_f\tau_f+\rho_t\tau_t}{P_S\theta T+PT} \le 1. && (35f)\end{aligned}$$

In case that constraints (35b), (35c), and (35e) are all active, the optimal solution of the problem (35) is also the optimal solution of the problem (34). We use the contradiction method to show that constraints (35b), (35c), and (35e) should be active. Suppose that (35b), (35c), and (35e) are not all active at the optimal solution $\{P_S^*, \theta^*, \varphi^*, \rho_t^*, \rho_f^*, \psi^*, t_1^*, t_2^*\}$, then we can construct a feasible point $\{P_S^*, \theta^*, \nu_3\varphi^*, \rho_t^*, \rho_f^*, \nu_4\psi^*, \nu_1 t_1^*, \nu_2 t_2^*\}$ for $\nu_1, \nu_2, \nu_3, \nu_4 \ge 1$ such that constraints (35b), (35c), and (35e) are active, which is still feasible to the problem (35). It can be seen that the contructed solution $\{P_S^*, \theta^*, \nu_3\varphi^*, \rho_t^*, \rho_f^*, \nu_4\psi^*, \nu_1 t_1^*, \nu_2 t_2^*\}$ can achieve a higher objective value than that offered by the optimal point, which leads to a contradiction. Therefore, we can conclude that constraints (35b), (35c), and (35e) are active at the optimal solution, and the problem (35) is equivalent to the problem (34).

The problem (35) is a nonconvex geometric program due to constraint (35f) which is a ratio of posynomials ([41], Section IV). Although the optimal solution of the problem (35) is difficult to get, a successive convex approximation method (SCA)

can be used to transforms it into a sequence of geometric programs to locate its KKT solution [41]. To this end, we condense the denominator posynomial of (35f) into a monomial using ([41], Lemma 1) at each iteration. In particular, with the obtained $P_S(l), \theta(l)$ at the lth iteration, we condense the denominator posynomial of constraint (35f) into a monomial as

$$\left(\frac{P_S\theta T}{\alpha_1}\right)^{\alpha_1}\left(\frac{PT}{\alpha_2}\right)^{\alpha_2}, \tag{36}$$

where

$$\alpha_1 \triangleq \frac{P_S\theta T}{P_S\theta T + PT} \text{ and } \alpha_2 \triangleq \frac{PT}{P_S\theta T + PT}.$$

As indicated in [41], SCA converges very fast and the solution obtained by SCA is very close to the global optimum. The proposed SCA algorithm is summarized in Algorithm 2.

Algorithm 2: Proposed SCA for Solving the Problem (35)

1) **Initialize** feasible P_S, θ,
2) Condense the denominator posynomial of the constraint (35f) into a monomial by (36),
3) Solve the resulting GP,

$$\min_{P_S,\theta,\varphi,\rho_t,\rho_f,\psi,t_1,t_2} \frac{1}{t_1 t_2} \tag{37}$$

$$\text{s.t. } \frac{P_S T + \rho_f\tau_f + \rho_t\tau_t}{\left(\frac{P_S\theta T}{\alpha_1}\right)^{\alpha_1}\left(\frac{PT}{\alpha_2}\right)^{\alpha_2}} \le 1, \tag{38}$$

$$\text{(35b)-(35e)} \tag{39}$$

4) Go to step 2 with the obtained solution.
5) Terminate the loop if the improvement of objective function is less than ϵ, where ϵ is the error tolerance for exit condition.

C. General SNR Case

For the case of general SNR, the joint design problem is more complicated. For handling this problem, we resort to optimize the lower bound of the objective function. Specially, we first derive a simple lower bound of C_S in (11), i.e., C_{LB}, and design $P_S, P_J, \rho_f, \rho_t, \varphi, \psi, \eta$ and θ jointly for maximizing the lower bound of the effective ESR, i.e., $(1-\theta)C_{LB}$. The lower bound is given by the following theorem.

Theorem 4: A lower bound of C_S in (11) is given by

$$C_S \ge C_{LB} \triangleq h_1\left(P_S, P_J, \sigma^2_{\mathbf{H}_B}, \sigma^2_{\mathbf{Z}}\right) + h_2(P_S, P_J), \tag{40}$$

where

$$h_1\left(P_S, P_J, \sigma^2_{\mathbf{H}_B}, \sigma^2_{\mathbf{Z}}\right) \triangleq N_B \log_2\left(\frac{1-\sigma^2_{\mathbf{H}_B}}{1+\left(P_S+P_J\right)\left(\sigma^2_{\mathbf{H}_B}+N_T\sigma^2_{\mathbf{Z}}\right)}\right), \tag{41}$$

$$h_2(P_S,P_J) \triangleq \max_{\eta}\left(N_B\log_2\left(\frac{P_S}{N_B}\exp\left(\frac{1}{N_B}\Lambda_{N_B,N_T}\right)\right) + \Xi - N_B\log_2\left(1+P_S\frac{\eta}{\beta_1}\right) - N_E(\eta-1)\log_2(e) + N_E\log_2(\eta) - (N_T-N_B)\log_2\left(1+\frac{\eta P_J N_B}{\beta_1(N_T-N_B)}\right)\right), \tag{42}$$

$$\Xi \triangleq \begin{cases} (N_T-N_B)\log_2\left(\frac{P_J\exp\left(\frac{1}{N_T-N_B}\Lambda_{N_T-N_B,N_E}\right)}{N_T-N_B}\right), \\ \text{if} \quad N_E > N_T - N_B \\ N_E\log_2\left(\frac{P_J}{N_T-N_B}\exp\left(\frac{1}{N_E}\Lambda_{N_E,N_T-N_B}\right)\right), \\ \text{if} \quad N_E \le N_T - N_B. \end{cases} \tag{43}$$

$\Lambda_{m,n} \triangleq \sum_{l=0}^{m-1}\psi(n-l)$.

Proof: The proof is given in Appendix E. ∎

Optimizing the lower bound of the objective function of problem (13) can be formulated as the following problem

$$\max_{\substack{P_S,P_J,\rho_f,\\ \rho_t,\varphi,\psi,\theta}} (1-\theta)C_{LB} \quad \text{s.t.} \quad (13b) \text{ and } (13c). \tag{44}$$

Compared (44) with (25), we find that they have the same structure. Therefore, problem (44) can be solved with a similar BCDM as solving problem (25). Firstly, the following proposition shows that $(1-\theta)C_{LB}$ is a concave function of θ.

Proposition 1: $(1-\theta)C_{LB}$ is a concave function of θ, and the optimal θ satisfies the equation

$$-h_2(P_S,P_J) - N_B\log_2\left(\frac{a\theta^2}{a\theta^2+b\theta+c}\right) + (1-\theta)\frac{N_B}{\ln 2}\frac{b\theta+2c}{\theta\left(a\theta^2+b\theta+c\right)} = 0, \tag{45}$$

where a, b, and c are defined in (26).

Proof: The proof is similar as the proof of Theorem 3, which is omitted for brevity. ∎

From (43), we find that although the formulas of C_{LB} for $N_E \le N_T - N_B$ and $N_E > N_T - N_B$ are different, C_{LB} has the same optimization structure for both cases. Therefore, we consider the case $N_E \le N_T - N_B$ and the result obtained can be directly generalized to the case $N_E > N_T - N_B$. Introducing the following variables $x_{P_S} = \ln(P_S)$, $x_{P_J} = \ln(P_J)$, $x_{\rho_t} = \ln(\rho_t)$, $x_{\rho_f} = \ln(\rho_f)$, $x_\varphi = \ln(\varphi)$, $x_\eta = \ln(\eta)$, and $x_\psi = \ln(\psi)$, the objective of the problem (44) can be reformulated as $\mathrm{T}(\mathbf{w}_x)$ given by (46),

$$\begin{aligned}\mathrm{T}(\mathbf{w}_x) \triangleq\ & N_B x_{P_S}\log_2(e) - N_B\log_2\left(1+\frac{e^{P_S+x_\eta}}{\beta_1}\right) \\ & - N_E\left(e^{x_\eta}-1\right)\log_2(e) \\ & + N_E x_\eta\log_2(e) + N_E x_{P_J}\log_2(e) \\ & + N_B\left(x_\varphi + x_\psi + x_{\rho_t} + x_{\rho_f}\right)\log_2(e) \\ & - N_B\log_2\Bigg(\frac{e^{x_\varphi+x_\psi+x_{\rho_t}+x_{\rho_f}}\theta^2T^2}{N_T} + N_T^2\left(e^{x_{P_S}}+e^{x_{P_J}}\right) \\ & + \left(e^{x_{\rho_f}+x_\psi}T + e^{x_{\rho_f}+x_\psi}T\left(e^{x_{P_S}}+e^{x_{P_J}}\right)\right. \\ & \left. + N_T T e^{x_\varphi+x_{\rho_t}}\left(e^{x_{P_S}}+e^{x_{P_J}}\right)\right)\theta\Bigg) \\ & - (N_T-N_B)\log_2\left(1+\frac{e^{x_\eta+P_J}N_B}{\beta_1\left(N_T-N_B\right)}\right) + \Sigma, \end{aligned} \tag{46}$$

where Σ denotes the collection of those constant terms do not involve $\mathbf{w}_x^*$. Then, with a given θ, the optimal $\mathbf{w}_x^*$ can be obtained by solving the following convex problem via the interior-point algorithm.

$$\max_{\mathbf{w}_x^*} \mathrm{T}(\mathbf{w}_x) \quad \text{s.t.} \quad (29b)-(29d). \tag{47}$$

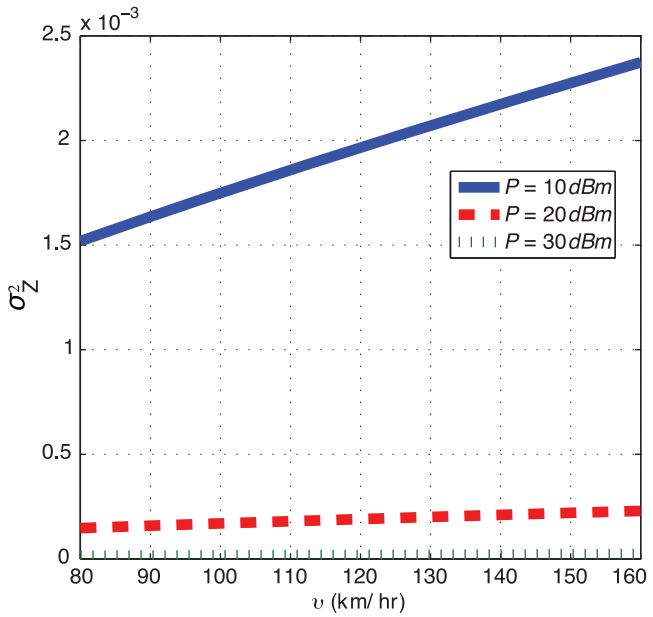

Fig. 4. Optimized $\sigma_{\mathbf{Z}}^2$ versus the velocity of Bob in the system, v, for $N_T = 5, N_B = 2$, and $N_E = 2$.

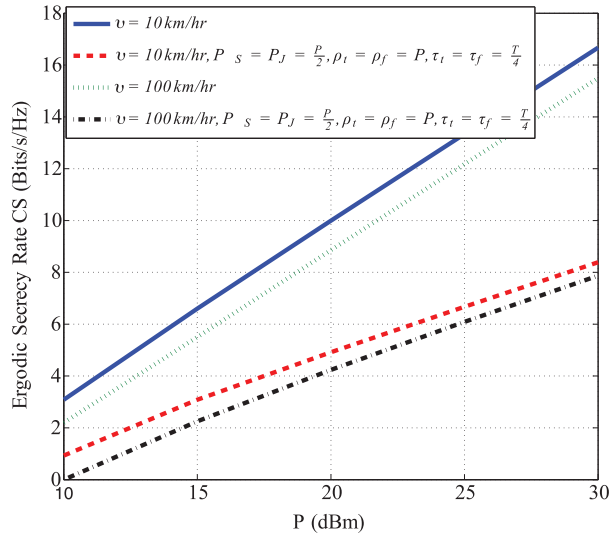

Fig. 5. Achievable effective ESR C_{eff} versus the power budget P at Alice for $N_T = 5, N_B = 2, N_E = 2$, and different velocities of nodes.

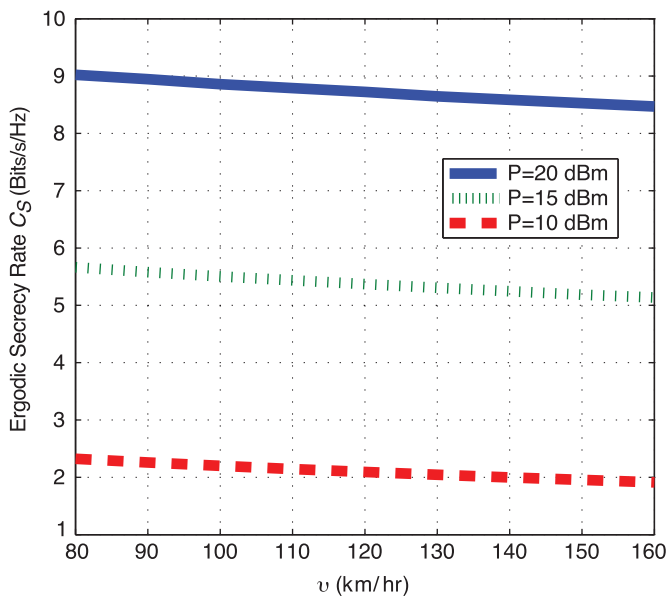

Fig. 6. Achievable effective ESR C_{eff} versus the Bob's velocity v at Alice for $N_T = 5, N_B = 2, N_E = 3$, and different velocities of nodes.

Then just as Algorithm 1, the stationary solution of the problem (44) can be located by the BCDM Algorithm which optimizes θ and $P_S, P_J, \rho_f, \rho_t, \varphi, \psi, \eta$ alternatively. The details are omitted for brevity.

IV. Simulation Results

In this section, considering typical wireless system parameters, where a wavelength of $\lambda = 0.15$ m (corresponding to a carrier frequency of 2 GHz), a coherence bandwidth of $W_C = 300$ kHz, a normalized Doppler given by $f_D = \frac{v}{\lambda W_C}$ with v Bob's velocity and $T = \frac{1}{2f_D}$, simulation results are provided to verify the performance of our proposed secure transmission scheme.

In Section II.A, we approximate $\hat{\mathbf{Z}}$ as an unitary matrix for facilitating the theoretical analysis, and we validate the approximation in Fig. 4. Specially, we plot the variance of the estimation error matrix, i.e., $\sigma_{\mathbf{Z}}^2$, versus v. From the simulation results in Fig. 4, we can find that the optimized $\sigma_{\mathbf{Z}}^2$ is very small, which can be ignored. In particular, when $P = 10$ dBm and $v = 160$ km/hr, the optimized $\sigma_{\mathbf{Z}}^2 = 2.3 \times 10^{-3}$.

Fig. 5 shows the C_{eff} achieved by the considered system for various vehicular-levels of mobilities, v. Furthermore, for showing the performance gains brought by the proposed optimization approach, we plot the achievable C_{eff} with $\tau_t = \tau_f = \frac{T}{4}, \rho_t = \rho_f = P, P_S = P_J = \frac{P}{2}$. From the simulation results in Fig. 5, we can find that with the increasing v, the achievable effective ESR with imperfect CSI decreases, due to the more severe estimation errors and shorter training overhead. Furthermore, a substantial performance gain can be achieved by the proposed algorithm over the baseline scheme with $\tau_t = \tau_f = \frac{T}{4}, \rho_t = \rho_f = P, P_S = P_J = \frac{P}{2}$, which shows the efficiency of our proposed joint optimization strategy.

Fig. 6 shows the achievable C_{eff} versus the Bob's velocity, v. With the increasing v, the coherence time T decreases, and C_{eff} decreases. However, by optimizing the power allocation and training overhead jointly, the performance degradation is very small. For example, when $P = 20$ dBm and v increases from 80 km/hr to 160 km/hr, the secrecy performance degradation is only 0.53 Bits/s/Hz, which also validates the robustness of our proposed secure transmissions scheme.

Fig. 7 shows the optimized signaling overhead fraction $\frac{\tau_t \rho_t + \tau_f \rho_f}{PT}$ versus N_T. From the simulation results in Fig. 7, we can find that with the increasing N_T, the signaling overhead fraction is increasing, since with the increasing N_T, the dimensionality of the forward channel increases and the signaling overhead should be increased for obtaining a satisfactory estimation. Furthermore, the signaling overhead increases with the increasing v. This is due to the fact that when v increases, the coherence time decreases and for protecting the system from the performance degradation due to the CSI error, more signaling overhead should be spent on the training and feedback, which leads to an increasing signaling overhead fraction.

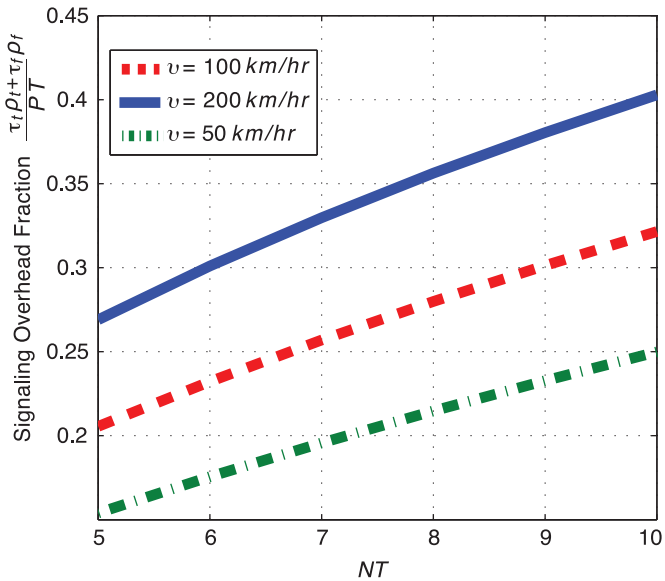

Fig. 7. Signaling overhead fraction $\frac{\tau_t\rho_t+\tau_f\rho_f}{PT}$ versus N_T for $P = 30$ dBm, and $N_B = N_E = 2$.

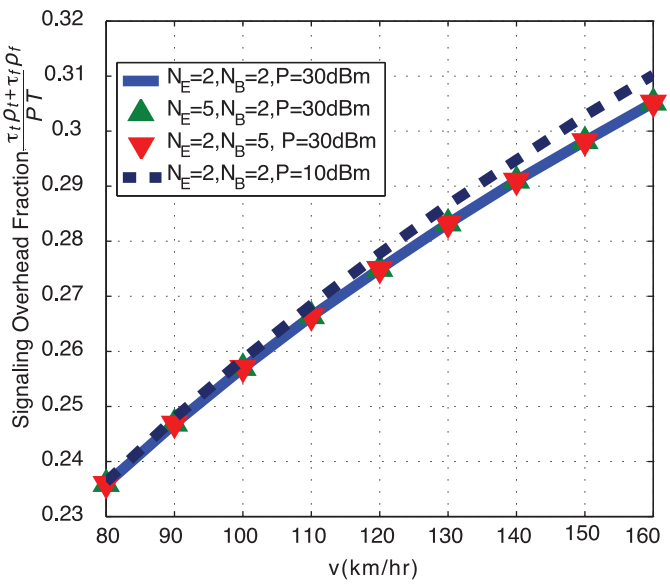

Fig. 8. Signaling overhead fraction $\frac{\tau_t\rho_t+\tau_f\rho_f}{PT}$ versus v for $P = 30$ dBm, and $N_T = 7$.

Fig. 8 shows the optimized signaling overhead fraction $\frac{\tau_t\rho_t+\tau_f\rho_f}{PT}$ versus v. Just as the simulation results in Fig. 7, the signaling overhead increases with the increasing v. However, although the increasing overhead improves the CSI estimate quality, it reduces the time duration for payload data. Therefore, it can be observed that there is a diminishing return in signaling overhead fraction as v increases. This observation again shows the trade-off between the channel estimation and data transmission. Furthermore, the simulation results show that different N_B and N_E do not affect the signaling overhead fraction and the decreasing P would increase the signaling overhead fraction.

V. Conclusion

In this paper, we proposed a practical framework for the artificial noise assisted secure transmission scheme in MIMOME channels, where the imperfect CSIs at Alice and Bob are obtained through training and analog feedback. We derive a simple closed-form analytical expression of ESR to quantify the dependence between signaling overhead/training power and the achievable effective ESR. Efficient iterative algorithms are proposed for optimizing the power allocation between data transmission and artificial noise, and signaling overhead/training power jointly to maximize the achievable effective ESR. In addition, we proved that in the low-SNR regime, exploiting artificial noise for guaranteeing secure communication is not necessary and derived a condition for achieving the positive ESR. Simulation results corroborate the theoretical analysis results and demonstrate the achievable performance of our proposed joint optimization method.

Appendix A
Extension to the Feedback Channel With Fading

Assuming that Bob sends $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}_N}$ back to Alice via a $N_T \times N_B$ feedback channel $\mathbf{G}_B$, whose elements are *i.i.d.* complex Gaussian with distribution $\mathcal{CN}(0,1)$. Here, we assume that only the imperfect estimation of $\mathbf{G}_B$, i.e., $\hat{\mathbf{G}}_B$ is obtained at Alice with MMSE estimator, which has *i.i.d.* $\mathcal{CN}(\mathbf{0},(1-\sigma^2_{\mathbf{G}_B})\mathbf{I})$, and the corresponding estimation error is $\tilde{\mathbf{G}}_B \sim \mathcal{CN}(\mathbf{0},\sigma^2_{\mathbf{G}_B}\mathbf{I})$, where $\sigma^2_{\mathbf{G}_B} = \frac{1}{1+\tau_G\rho_G/N_B}$. τ_G and ρ_G are training time and power spent on the estimation of $\mathbf{G}_B$. With the gathered observations, Alice calculates an estimate of $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$ with the LS estimator. The corresponding estimation error can be calculated with similar procedures, whose entries have zero mean and variance $(N_B\sigma^2_{\mathbf{G}_B} + \frac{N_T}{\tau_f\rho_f})\frac{1}{(1-\sigma^2_{\mathbf{G}_B})(N_T-N_B)}$. Then, the joint power allocation and training overhead optimization problem for such case can be built by setting $\sigma^2_{\mathbf{Z}} = (N_B\sigma^2_{\mathbf{G}_B} + \frac{N_T}{\tau_f\rho_f})\frac{1}{(1-\sigma^2_{\mathbf{G}_B})(N_T-N_B)}$.

Appendix B
Proof of Theorem 1

In the following, we resort to asymptotic capacity analysis results in [34] to derive the deterministic approximations of $\mathbb{E}(R_B)$ and $\mathbb{E}(R_E)$, respectively.

With ([34], (8)), the deterministic approximation of $\mathbb{E}(R_B)$ can be derived as follows

$$\begin{aligned}\mathbb{E}(R_B) \approx\ & N_B\log_2\left(1+\omega-F\left(\beta_3,\frac{\omega}{\beta_3}\right)\right)\\ &+N_T\log_2\left(1+\frac{\omega}{\beta_3}-F\left(\beta_3,\frac{\omega}{\beta_3}\right)\right)\\ &-N_T\frac{\log_2(e)}{\omega}F\left(\beta_3,\frac{\omega}{\beta_3}\right),\end{aligned} \tag{48}$$

where ω is defined in (16).

$\mathbb{E}(R_E)$ can be rewritten as (49).

$$\mathbb{E}(R_E)=\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_S}{N_T}\mathbf{H}_E\hat{\mathbf{F}}\hat{\mathbf{F}}^H\mathbf{H}_E^H+\frac{P_J}{(N_T-N_B)}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)\right)\right)-\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_S}{(N_T-N_B)}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)\right)\right). \quad (49)$$

Since $\mathbf{H}_B$, $\mathbf{G}_B$, and $\mathbf{H}_E$ have *i.i.d.* $\mathcal{CN}(0,1)$ entries, and $\hat{\mathbf{F}}$ and $\hat{\mathbf{U}}_N$ are both unitary matrices, the $N_E\times N_B$ matrix $\mathbf{H}_E\hat{\mathbf{F}}$ and $N_E\times(N_T-N_B)$ matrix $\mathbf{H}_E\hat{\mathbf{U}}_N$ both have *i.i.d.* $\mathcal{CN}(0,1)$ entries. By exploiting the asymptotic capacity analytical results in ([34], Section IV), the deterministic approximation of the minuend in (49) is given as (50),

$$\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_S}{N_T}\mathbf{H}_E\hat{\mathbf{F}}\hat{\mathbf{F}}^H\mathbf{H}_E^H+\frac{P_J}{(N_T-N_B)}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)\right)\right)\approx N_B\log_2\left(1+P_S\frac{\eta^*}{\beta_1}\right)+(N_T-N_B)\log_2\left(1+\frac{\eta^*P_JN_B}{\beta_1(N_T-N_B)}\right)-N_E\log_2(\eta^*)+N_E\left(\eta^*-1\right)\log_2(e). \quad (50)$$

where η^* is the solution to (18).

From ([34], (8)), the deterministic approximation of the subtrahend in (49) is given as follows

$$\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_J}{(N_T-N_B)}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)\right)\right)\approx N_E\log_2\left(1+P_J-F\left(\beta_2,\frac{P_S}{\beta_2}\right)\right)-\frac{(N_T-N_B)\log_2(e)}{P_S}F\left(\beta_2,\frac{P_S}{\beta_2}\right)+(N_T-N_B)\log_2\left(1+\frac{P_S}{\beta_2}-F\left(\beta_2,\frac{P_S}{\beta_2}\right)\right). \quad (51)$$

Substituting (48), (50), and (51) into (11) and the result follows immediately.

Appendix C
Proof of Lemma 1

Since $\beta_3>1$, by applying ([34], (15)), we have the following high SNR approximation

$$\mathbb{E}(R_B)\approx N_B\log_2\left(\frac{\omega}{e}\right)-N_B\left(\beta_3-1\right)\log_2\left(1-\frac{1}{\beta_3}\right). \quad (52)$$

On the other hand, $\mathbb{E}(R_E)$ is given in (49). By utilizing the result from ([34], (15)), in the high-SNR regime, the subtrahend in (49) can be approximated by $\Omega(\beta_2)$ which is defined in (22). Also, the minuend in (49) is derived in (50). Then combining the results in (50), (52), and (22), the proof is completed.

Appendix D
Proof of Theorem 2

The partial derivative of $f(\theta,\eta)$ with respect to η can be written as

$$\frac{\partial f(P_S,P_J,\eta)}{\partial\eta}=-\frac{N_E}{\eta}\left(\eta+\frac{N_B}{N_E}\frac{P_S\eta}{P_S\eta+\beta_1}+\frac{N_T-N_B}{N_E}\frac{P_JN_B\eta}{\eta P_JN_B+\beta_1(N_T-N_B)}-1\right), \quad (53)$$

and we let η^o satisfy

$$\frac{N_B}{N_E}\frac{P_S\eta^o}{P_S\eta^o+\beta_1}+\frac{N_T-N_B}{N_E}\times\frac{P_JN_B\eta^o}{\eta^oP_JN_B+\beta_1\left(N_T-N_B\right)}+\eta^o-1=0. \quad (54)$$

Defining $t(\eta)\triangleq\eta+\frac{N_B}{N_E}\frac{P_S\eta}{P_S\eta+\beta_1}+\frac{N_T-N_B}{N_E}\frac{P_JN_B\eta}{\eta P_JN_B+\beta_1(N_T-N_B)}-1$, we can find that $t(\eta)$ is a monotonically increasing with respect to η. Then, we have

$$t(\eta)=\begin{cases}<0, & 0<\eta<\eta^o,\\ >0, & \eta^o<\eta<1.\end{cases} \quad (55)$$

Therefore, we have

$$\frac{\partial f(P_S,P_J,\eta)}{\partial\eta}=\begin{cases}>0, & 0<\eta<\eta^o,\\ =0, & \eta=\eta^o,\\ >0, & \eta^o<\eta<1.\end{cases} \quad (56)$$

(56) indicates that η^o is the optimal solution for maximizing $f(P_S,P_J,\eta)$. Fortunately, from (54), we can find that η^o satisfies the fixed point (18), i.e., $\eta^*=\Gamma(\eta^*)$. Furthermore, η^* satisfying the fixed point (18) is unique, as shown by Remark 1. Therefore, we have $C_S^H=\max_\eta f(P_S,P_J,\eta)$.

Appendix E
Proof of Theorem 4

From (11) we can expressed C_S as (57).

$$C_S=\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_B}+\rho_{\text{eff}}\frac{\bar{\mathbf{H}}\bar{\mathbf{H}}^H}{N_B}\right)\right)\right)+\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_J}{(N_T-N_B)}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)\right)\right)-\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_S}{N_B}\mathbf{H}_E\hat{\mathbf{F}}\hat{\mathbf{F}}^H\mathbf{H}_E^H+\frac{P_J}{(N_T-N_B)}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)\right)\right). \quad (57)$$

In the following, we apply the Minkowshi's inequality [46] to arrive at a lower bound of C_S as follows

$$\begin{aligned}&\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_B}+\rho_{\text{eff}}\frac{\bar{\mathbf{H}}\bar{\mathbf{H}}^H}{N_B}\right)\right)\right)\\&\geq\mathbb{E}\left(\log_2\left(1+\frac{\rho_{\text{eff}}}{N_B}\det\left(\bar{\mathbf{H}}\bar{\mathbf{H}}^H\right)^{\frac{1}{N_B}}\right)^{N_B}\right)\\&=\mathbb{E}\left(N_B\log_2\left(1+\frac{\rho_{\text{eff}}}{N_B}\exp\left(\frac{1}{N_B}\ln\left(\det\left(\bar{\mathbf{H}}\bar{\mathbf{H}}^H\right)\right)\right)\right)\right)\\&\overset{(d)}{\geq}N_B\log_2\left(1+\frac{\rho_{\text{eff}}}{N_B}\exp\left(\frac{1}{N_B}\mathbb{E}\left(\ln\left(\det\left(\bar{\mathbf{H}}\bar{\mathbf{H}}^H\right)\right)\right)\right)\right)\\&>N_B\log_2\left(\frac{\rho_{\text{eff}}}{N_B}\exp\left(\frac{1}{N_B}\mathbb{E}\left(\ln\left(\det\left(\bar{\mathbf{H}}\bar{\mathbf{H}}^H\right)\right)\right)\right)\right). \quad (58)\end{aligned}$$

Since $\log_2(1+ae^x)$ is convex in x for $a>0$, applying the Jensen's inequality, step (d) can be obtained.

Besides $\bar{\mathbf{H}}\bar{\mathbf{H}}^H\sim\mathcal{W}_{N_B}(N_T,\mathbf{I}_{N_B})$, according to ([45], (2.12)), we have $\mathbb{E}(\ln(\det(\bar{\mathbf{H}}\bar{\mathbf{H}}^H)))=\Lambda_{N_B,N_T}$ and Λ_{N_B,N_T}

is defined in (40). When $N_E \leq N_T - N_B$, with a similar procedures as (58), we have

$$\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_J}{(N_T-N_B)}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)\right)\right) \geq N_E\log_2\left(\frac{P_J}{N_T-N_B}\exp\left(\frac{1}{N_E}\Lambda_{N_E,N_T-N_B}\right)\right) \quad (59)$$

However, when $N_E > N_T - N_B$, $\det(\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H)=0$. Since $\det(\mathbf{I}+\mathbf{AB})=\det(\mathbf{I}+\mathbf{BA})$ [46], when $N_E > N_T - N_B$, with a similar procedures as (58), we have

$$\mathbb{E}\left(\log_2\left(\det\left(\mathbf{I}_{N_E}+\frac{P_J}{(N_T-N_B)}\mathbf{H}_E\hat{\mathbf{U}}_N\hat{\mathbf{U}}_N^H\mathbf{H}_E^H\right)\right)\right) \geq (N_T-N_B)\log_2\left(\frac{P_J\exp\left(\frac{1}{N_T-N_B}\Lambda_{N_T-N_B,N_E}\right)}{N_T-N_B}\right) \quad (60)$$

The third term in (57) can be approximated by the deterministic approximation in (50), and according to Theorem 2, it can be written as (42). Then, combining (58), (59), and (60), the lower bound of C_S in Theorem 4 can be obtained.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] Y.-W. P. Hong, P.-C. Lan, and C.-C. J. Kuo, "Enhancing physical layer secrecy in multiantenna wireless systems: An overview of signal processing approaches," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 29–40, Sep. 2013.

[3] A. Khisti and G. Wornell, "Secure transmission with multiple antennas II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[4] Q. Li, M. Hong, H.-T. Wai, Y.-F. Liu, W.-K. Ma, and Z.-Q. Luo, "Transmit solutions for MIMO wiretap channels using alternating optimization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1714–1727, Sep. 2013.

[5] S.-C. Lin and C.-L. Lin, "On secrecy capacity of fast fading MIMOME wiretap channels with statistical CSIT," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 3293–3306, Jun. 2014.

[6] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[7] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.

[8] X. Zhang, X. Zhou, and M. R. McKay, "On the design of artificial-noise-aided secure multi-antenna transmission in slow fading channels," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2170–2181, June 2013.

[9] P.-H. Lin, S.-H. Lai, S.-C. Lin, and H.-J. Su, "On secrecy rate of the generalized artificial-noise assisted secure beamforming for wiretap channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1728–1740, Sep. 2013.

[10] S.-H. Tsai and H. V. Poor, "Power allocation for artificial-noise secure MIMO precoding systems," *IEEE Trans. Signal Process.*, vol. 62, no. 13, pp. 3479–3493, Jul. 2014.

[11] H.-M. Wang, M. Luo, and Q. Yin, "Hybrid cooperative beamforming and jamming for physical-layer security of two-way relay networks," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 12, pp. 2007–2020, Dec. 2013.

[12] C. Wang, H.-M. Wang, and X.-G. Xia, "Hybrid opportunistic relaying and jamming with power allocation for secure cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 589–605, Feb. 2015.

[13] C. Wang, H.-M. Wang, X.-G. Xia, and C. Liu, "Uncoordinated jammer selection for securing SIMOME wiretap channels: A stochastic geometry approach," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2596–2612, May 2015.

[14] H.-M. Wang and X.-G. Xia, "Enhancing wireless secrecy via cooperation: Signal design and optimization," *IEEE Commun. Mag.*, Dec. 2015, to be published.

[15] T.-Y. Liu, S.-C. Lin, T.-H. Chang, and Y.-W. P. Hong, "How much training is enough for secrecy beamforming with artificial noise," presented at the IEEE ICC, Ottawa, Canada, Jun. 2012.

[16] L. Sun and S. Jin, "On the ergodic secrecy rate of multiple-antenna wiretap channels using artificial noise and finite-rate feedback," presented at the IEEE PIMRC, Toronto, ON, Canada, Sep. 2011.

[17] S.-C. Lin, T.-H. Chang, Y.-L. Liang, Y.-W. P. Hong, and C.-Y. Chi, "On the impact of quantized channel feedback in guaranteeing secrecy with artificial noise: The noise leakage problem," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 901–915, Mar. 2013.

[18] X. Zhang, M. R. McKay, X. Zhou, and R. W. Heath, Jr, "Artificial-noise-aided secure multi-antenna transmission with limited feedback," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2742–2754, May 2015.

[19] M. Pei, A. L. Swindlehurst, D. Ma, and J. Wei, "Adaptive limited feedback for MISO wiretap channels with cooperative jamming," *IEEE Trans. Signal Process.*, vol. 62, no. 4, pp. 993–1004, Feb. 2014.

[20] O. El Ayach, A. Lozano, and R. W. Heath, Jr., "On the overhead of interference alignment: Training, feedback, and cooperation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4192–4203, Nov. 2012.

[21] R. K. Mungara, G. George, and A. Lozano, "Overhead and spectral efficiency of pilot-assisted interference alignment in time-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4884–4895, Sept. 2014.

[22] D. Samardzija and N. Mandayam, "Unquantized and uncoded channel state information feedback in multiple-antenna multiuser systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1335–1345, Jul. 2006.

[23] M. Kobayashi, N. Jindal, and G. Caire, "Training and feedback optimization for multiuser MIMO downlink," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug. 2011.

[24] T. Marzetta and B. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1268–1278, Apr. 2006.

[25] S. Sesia, M. Baker, and I. Toufik, *The UMTS Long Term Evolution: From Theory to Practice*. New York, NY, USA: Wiley-Blackwell LTE, 2009.

[26] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[27] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, pp. 951–963, Apr. 2003.

[28] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, pp. 2203–2214, May 2006.

[29] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[30] J. Zhang, J. G. Andrews, and K. B. Letaief, "Spatial intercell interference cancellation with CSI training and feedback," [Online]. Available: http://arxiv.org/pdf/1105.4206.pdf

[31] C.-F. Liu, M. Maso, S. Lakshminarayana, C.-H. Lee, and T. Q. S. Quek, "Simultaneous wireless information and power transfer under different CSI acquisition schemes," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 1911–1926, Apr. 2015.

[32] H. Shin and J. H. Lee, "Closed-form formulas for ergodic capacity of MIMO Rayleigh fading channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Anchorage, AL, USA, May 2003, pp. 2996–3000.

[33] M. Chiani, M. Z. Win, and H. Shin, "MIMO networks: The effects of interference," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 336–349, Jan. 2010.

[34] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.

[35] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE l. Set. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[36] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 1999.

[37] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Oper. Res. Lett.*, vol. 26, pp. 127–136, 2000.

[38] D. P. Palomar and S. Verd, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 52, pp. 141–154, Jan. 2006.
[39] C. Wang and H.-M. Wang, "Robust joint beamforming and jamming for secure AF networks: Low complexity design," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 2192–2198, May 2015.
[40] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming. ver. 1.21, Apr. 2011 [Online]. Available: http://cvxr.com/cvx
[41] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, July 2007.
[42] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, ser. MPS-SIAM Series on Optimization. Philadelphia, PA, USA: SIAM, 2001.
[43] C. K. A. Yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 458–462, Feb. 2007.
[44] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.
[45] A. M. Tulino and S. Verdu, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.
[46] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, MA, USA: Cambridge Univ. Press, 1990.
[47] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, MA, USA: Cambridge Univ. Press, 2004.

Hui-Ming Wang (S'07–M'10) received the B.S. and Ph.D. degrees, both in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2010, respectively. He is currently a Full Professor with the Department of Information and Communications Engineering, Xi'an Jiaotong University, and also with the Ministry of Education Key Lab for Intelligent Networks and Network Security, China. From 2007 to 2008, and 2009 to 2010, he was a Visiting Scholar at the Department of Electrical and Computer Engineering, University of Delaware, USA. His research interests include cooperative communication systems, physical-layer security of wireless communications, MIMO and space-timing coding.

Dr. Wang received the National Excellent Doctoral Dissertation Award in China in 2012, a Best Paper Award of International Conference on Wireless Communications and Signal Processing (WCSP), 2011, and a Best Paper Award of IEEE/CIC International Conference on Communications in China (ICCC), 2014.

Chao Wang received the B.S. degree in telecommunication engineering in 2008, and the M.S. degree in information and communication engineering in 2013 from Xi'an Jiaotong University, respectively. He is currently working towards the Ph.D. degree in information and communication engineering, Xi'an Jiaotong University. His research interests include cooperative communications, MIMO systems, stochastic geometry, and physical-layer security of wireless communications.

Chao Wang received a Best Paper Award of IEEE/CIC International Conference on Communications in China, 2014.

Derrick Wing Kwan Ng (S'06–M'12) received the bachelor degree with first class honors and the Master of Philosophy (M.Phil.) degree in electronic engineering from the Hong Kong University of Science and Technology (HKUST) in 2006 and 2008, respectively. He received his Ph.D. degree from the University of British Columbia (UBC) in 2012. In the summer of 2011 and spring of 2012, he was a visiting scholar at the Centre Tecnológic de Telecomunicacions de Catalunya—Hong Kong (CTTC-HK). He was a senior postdoctoral fellow in the Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany. He is now working as a Lecturer in the University of New South Wales, Sydney, Australia. His research interests include convex and non-convex optimization, physical layer security, and green (energy-efficient) wireless communications,

Dr. Ng received the Best Paper Awards at the IEEE Wireless Communications and Networking Conference (WCNC) 2012, the IEEE Global Telecommunication Conference (Globecom) 2011, and the IEEE Third International Conference on Communications and Networking in China 2008. He was awarded the IEEE Student Travel Grants for attending the IEEE WCNC 2010, the IEEE International Conference on Communications (ICC) 2011, and the IEEE Globecom 2011. He was also the recipient of the 2009 Four Year Doctoral Fellowship from the UBC, Sumida & Ichiro Yawata Foundation Scholarship in 2008, and R&D Excellence Scholarship from the Center for Wireless Information Technology in HKUST in 2006. He has served as an editorial assistant to the Editor-in-Chief of the IEEE TRANSACTIONS ON COMMUNICATIONS since Jan. 2012. He is currently an Editor of the IEEE COMMUNICATIONS LETTERS. He was a Co-Chair for the Wireless Access Track of 2014 IEEE 80th Vehicular Technology Conference. He has been a TPC member of various conferences, including the Globecom, WCNC, ICC, VTC, and PIMRC, etc. He was honoured as the top reviewer of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY in 2014 and an Exemplary Reviewer of the IEEE WIRELESS COMMUNICATIONS LETTERS for 2012, 2014.