# Linear Precoding of Data and Artificial Noise in Secure Massive MIMO Systems

Jun Zhu, *Student Member, IEEE*, Robert Schober, *Fellow, IEEE*, and Vijay K. Bhargava, *Life Fellow, IEEE*

***Abstract*—In this paper, we consider secure downlink transmission in a multicell massive multiple-input multiple-output (MIMO) system where the numbers of base station (BS) antennas, mobile terminals, and eavesdropper antennas are asymptotically large. The channel state information of the eavesdropper is assumed to be unavailable at the BS and hence, linear precoding of data and artificial noise (AN) are employed for secrecy enhancement. Four different data precoders (i.e., selfish zero-forcing (ZF)/regularized channel inversion (RCI) and collaborative ZF/RCI precoders) and three different AN precoders (i.e., random, selfish/collaborative null-space-based precoders) are investigated and the corresponding achievable ergodic secrecy rates are analyzed. Our analysis includes the effects of uplink channel estimation, pilot contamination, multicell interference, and path-loss. Furthermore, to strike a balance between complexity and performance, linear precoders that are based on matrix polynomials are proposed for both data and AN precoding. The polynomial coefficients of the data and AN precoders are optimized, respectively, for minimization of the sum-mean-squared-error of and the AN leakage to the mobile terminals in the cell of interest using tools from free probability and random matrix theory. Our analytical and simulation results provide interesting insights for the design of secure multicell massive MIMO systems and reveal that the proposed polynomial data and AN precoders closely approach the performance of selfish RCI data and null-space-based AN precoders, respectively.**

***Index Terms*—Physical layer security, massive MIMO, multicell systems, artificial noise, ergodic secrecy rate, pilot contamination, computational complexity, and matrix polynomial.**

## I. Introduction

MASSIVE multiple-input multiple-output (MIMO) systems employing simple linear precoding and combining schemes offer significant performance gains in terms of bandwidth, power, and energy efficiency compared to conventional multiuser MIMO systems as impairments such as fading, noise, and interference are averaged out for very large numbers of base station (BS) antennas [3]–[5]. Furthermore, in time-division duplex (TDD) systems, channel reciprocity can be exploited to estimate the downlink channels via uplink training so that the training overhead scales only linearly with the number of users and is independent of the number of BS antennas [4]. However, if the pilot sequences employed in different cells are not orthogonal, so-called pilot contamination occurs and impairs the channel estimates, which ultimately limits the achievable performance of massive MIMO systems [4], [6].

Since secrecy and privacy are critical concerns for the design of future communication systems [7], it is of interest to investigate how the large number of spatial degrees of freedom in massive MIMO systems can be exploited for secrecy enhancement [8]–[10]. If the eavesdropper (Eve) remains passive to hide its existence, neither the transmitter (Alice) nor the legitimate receiver (Bob) will be able to learn Eve's channel state information (CSI). In this situation, it is advantageous to inject artificial noise (AN) at the transmitter to degrade Eve's channel and to use linear precoding to avoid impairment to Bob's channel as was shown in [11]–[17] for single user and single-cell multiuser systems. However, in multi-cell massive MIMO systems, multi-cell interference and pilot contamination will hamper Alice's ability to degrade Eve's channel and to protect Bob's channel. This problem was studied first in [18] for simple matched-filter (MF) data precoding and null-space (NS) and random AN precoding. However, it is well known that MF data precoding suffers from a large loss in the achievable information rate compared to other linear data precoders such as zero-forcing (ZF) and regularized channel inversion (RCI) precoders as the number of mobile terminals (MTs) increases [19]. Since it is expected that this loss in information rate also translates into a loss in secrecy rate, studying the secrecy performance of ZF and RCI data precoders in massive MIMO systems is of interest. Furthermore, while NS AN precoding was shown to achieve a better performance compared to random AN precoding [18], it also entails a much higher complexity. Similarly, the improved performance of ZF and RCI data precoding compared to MF data precoding comes at the expense of a higher complexity. Hence, the design of novel data and AN precoders which allow a flexible tradeoff between complexity and secrecy performance is desirable.

Related work on physical layer security in massive MIMO systems includes [20] where the authors use the channel between Alice and Bob as secrete key and show that the complexity required by Eve to decode Alice's message is at

Manuscript received April 6, 2015; revised August 31, 2015; accepted November 8, 2015. Date of publication November 13, 2015; date of current version March 8, 2016. This work was supported in part by the National Science and Engineering Research Council of Canada (NSERC) Strategic under Grant STPGP 396545-10, in part by the Alexander von Humboldt Professorship Program, and in part by the German Science Foundation under Grant SCHO 831/5-1. This work was presented in part at the European Wireless Conference, Barcelona, Spain, 2014 [1] and the International Symposium on Communications, Control, and Signal Processing, Athens, Greece, 2014 [2]. The associate editor coordinating the review of this paper and approving it for publication was W. Zhang.

J. Zhu and V. K. Bhargava are with the Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, BC, Canada (e-mail: zhujun@ece.ubc.ca; vijayb@ece.ubc.ca).

R. Schober is with the Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, BC, Canada, and also with the Institute for Digital Communications (IDC), Friedrich-Alexander-University (FAU), Erlangen, Germany (e-mail: rschober@ece.ubc.ca).

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/TWC.2015.2500578

least of the same order as a worst-case lattice problem. Physical layer security in a downlink multi-cell MIMO system and a relay-assisted system were considered in [21]–[23] and [24], respectively. However, unlike our work, perfect knowledge of Eve's channel was assumed, AN injection was not considered, and pilot contamination was not taken into account. Furthermore, ZF and RCI data precoding were analyzed in the large system limit in [25], [26]. However, neither pilot contamination nor AN were taken into account and the secrecy rate was not analyzed. Using a concept that was originally conceived for code division multiple access (CDMA) uplink systems in [27] and later extended to MIMO systems in [28], reduced complexity linear data precoders that are based on matrix polynomials were investigated for use in massive MIMO systems in [29]–[31]. However, [29]–[31] did not take into account the effect of AN leakage for precoder design and did not study the secrecy performance. Moreover, in [32]–[34], active eavesdropping in massive MIMO systems was considered, whereas the eavesdropper in this paper is assumed to be passive in order to hide its presence. Hence, the results presented in [20]–[34] are not directly applicable to the system studied in this paper.

In this paper, we consider secure downlink transmission in a multi-cell massive MIMO system employing linear data and AN precoding in the presence of a passive multi-antenna eavesdropper. We study the achievable ergodic secrecy rate of such systems for different linear precoding schemes taking into account the effects of uplink channel estimation, pilot contamination, multi-cell interference, and path-loss. The main contributions of this paper are summarized as follows:

- We study the performance-complexity tradeoff of selfish and collaborative data and AN precoders. Selfish precoders require only the CSI of the MTs in the local cell but cause inter-cell interference and inter-cell AN leakage. In contrast, collaborative precoders require the CSI between the local BS and the MTs in all cells, but reduce inter-cell interference and inter-cell AN leakage. However, since the additional CSI required for the collaborative precoders can be estimated directly by the local BS, the additional overhead and complexity incurred compared to selfish precoders is limited.
- We derive novel closed-form expressions for the asymptotic ergodic secrecy rate which facilitate the performance comparison of different combinations of linear data precoders (i.e., MF, selfish and collaborative ZF/RCI) and AN precoders (i.e., random, selfish and collaborative NS), and provide significant insight for system design and optimization.
- In order to avoid the computational complexity and potential stability issues in fixed point implementations entailed by the large-scale matrix inversions required for ZF and RCI data precoding and NS AN precoding, we propose polynomial (POLY) data and AN precoders and optimize their coefficients. Unlike [30] and [31], which considered polynomial data precoders for massive MIMO systems without AN generation, we use free probability theory [29], [35] to obtain the POLY coefficients. This allows us to express the POLY coefficients as simple functions of the channel and system parameters. Simulation results reveal that these precoders are able to closely approach the performance of selfish RCI data and NS AN precoders, respectively.

The remainder of this paper is organized as follows. In Section II, we outline the considered system model and review some basic results from [18]. In Sections III and IV, the considered linear data and AN precoders are investigated, respectively. In Section V, the ergodic secrecy rates of different linear precoders are compared analytically for a simple path-loss model. Simulation and numerical results are presented in Section VI, and some conclusions are drawn in Section VII.

*Notation:* Superscripts $T$ and $H$ stand for the transpose and conjugate transpose, respectively. $\mathbf{I}_N$ is the $N$-dimensional identity matrix. The expectation operation and the variance of a random variable are denoted by $\mathbb{E}[\cdot]$ and $\mathrm{var}[\cdot]$, respectively. $\mathrm{diag}\{\mathbf{x}\}$ denotes a diagonal matrix with the elements of vector $\mathbf{x}$ on the main diagonal. $\mathrm{tr}\{\cdot\}$ and $\mathrm{rank}\{\cdot\}$ denote trace and rank of a matrix, respectively. $\mathbb{C}^{m\times n}$ represents the space of all $m \times n$ matrices with complex-valued elements. $\mathbf{x} \sim \mathbb{CN}(\mathbf{0}_N, \mathbf{\Sigma})$ denotes a circularly symmetric complex Gaussian vector $\mathbf{x} \in \mathbb{C}^{N\times 1}$ with zero mean and covariance matrix $\mathbf{\Sigma}$. $[\mathbf{A}]_{kl}$ denotes the element in the $k^{\text{th}}$ row and $l^{\text{th}}$ column of matrix $\mathbf{A}$, and $[x]^+ = \max\{x, 0\}$.

## II. System Model and Preliminaries

In this section, we introduce the considered system model as well as the adopted channel estimation scheme, and review some ergodic secrecy rate results.

### A. System Model

We consider the downlink of a multi-cell massive MIMO system with $M$ cells and a frequency reuse factor of one, i.e., all BSs use the same spectrum. Each cell includes one $N_T$-antenna BS, $K \leq N_T$ single-antenna MTs, and potentially an $N_E$-antenna eavesdropper. The eavesdroppers try to hide their existence and hence remain passive. As a result, the BSs cannot estimate the eavesdroppers' CSI. To overcome this limitation, each BS generates AN to mask its information-carrying signal and to prevent eavesdropping [11]. In the following, the $k^{\text{th}}$ MT, $k = 1, \ldots, K$, in the $n^{\text{th}}$ cell, $n = 1, \ldots, M$, is the MT of interest and we assume that an eavesdropper tries to decode the signal intended for this MT. We note that neither the BSs nor the MTs are assumed to know which MT is targeted by the eavesdropper. The signal vector, $\mathbf{x}_n \in \mathbb{C}^{N_T\times 1}$, transmitted by the BS in the $n^{\text{th}}$ cell (also referred to as the $n^{\text{th}}$ BS in the following) is given by

$$\mathbf{x}_n = \sqrt{p}\mathbf{F}_n\mathbf{s}_n + \sqrt{q}\mathbf{A}_n\mathbf{z}_n, \tag{1}$$

where $\mathbf{s}_n \sim \mathbb{CN}(\mathbf{0}_K, \mathbf{I}_K)$ and $\mathbf{z}_n \sim \mathbb{CN}(\mathbf{0}_{N_T}, \mathbf{I}_{N_T})$ denote the data and AN vectors for the $K$ MTs in the $n^{\text{th}}$ cell, respectively. $\mathbf{F}_n = [\mathbf{f}_{n1}, \cdots, \mathbf{f}_{nK}] \in \mathbb{C}^{N_T\times K}$ and $\mathbf{A}_n = [\mathbf{a}_{n1}, \cdots, \mathbf{a}_{nN_T}] \in \mathbb{C}^{N_T\times N_T}$ are the data and AN precoding matrices, respectively, and the efficient design of these matrices is the main scope of this paper. Thereby, the structure of both types of precoding matrices does not depend on which MT is targeted by

the eavesdropper. The AN precoding matrix $\mathbf{A}_n$ has rank $L = \text{rank}\{\mathbf{A}_n\} \leq N_T$, i.e., $L$ dimensions of the $N_T$-dimensional signal space spanned by the $N_T$ BS antennas are exploited for jamming of the eavesdropper. The data and AN precoding matrices are normalized as $\text{tr}\{\mathbf{F}_n^H \mathbf{F}_n\} = K$ and $\text{tr}\{\mathbf{A}_n^H \mathbf{A}_n\} = L$, i.e., their average power per dimension is one. The average powers $p$ and $q$ allocated to the information-carrying signal for each MT and each AN signal, respectively, can be written as $p = \frac{\phi P_T}{K}$ and $q = \frac{(1-\phi)P_T}{L}$, where $P_T$ is the total transmit power and $\phi \in (0, 1]$ is a power allocation factor which can be optimized. For the sake of clarity, in this paper, we assume that all cells utilize the same value of $\phi$.

The vectors collecting the received signals at the $K$ MTs and the $N_E$ antennas of the eavesdropper in the $n^{\text{th}}$ cell are given by

$$\mathbf{y}_n = \sum_{m=1}^{M} \mathbf{G}_{mn}\mathbf{x}_m + \mathbf{n}_n \quad \text{and} \quad \mathbf{y}_E = \sum_{m=1}^{M} \mathbf{G}_{mE}\mathbf{x}_m + \mathbf{n}_E, \tag{2}$$

respectively, with Gaussian noise vectors $\mathbf{n}_n \in \mathbb{CN}(\mathbf{0}_K, \sigma_n^2 \mathbf{I}_K)$ and $\mathbf{n}_E \in \mathbb{CN}(\mathbf{0}_{N_E}, \sigma_E^2 \mathbf{I}_{N_E})$, where $\sigma_n^2$ and $\sigma_E^2$ denote the noise variances at one MT and one eavesdropper receive antenna, respectively. Furthermore, $\mathbf{G}_{mn} = \mathbf{D}_{mn}^{1/2}\mathbf{H}_{mn} \in \mathbb{C}^{K \times N_T}$ and $\mathbf{G}_{mE} = \sqrt{\beta_{mE}}\mathbf{H}_{mE} \in \mathbb{C}^{N_E \times N_T}$ are the matrices modeling the channels from the $m^{\text{th}}$ BS to the $K$ MTs and the eavesdropper in the $n^{\text{th}}$ cell, respectively. Thereby, $\mathbf{D}_{mn} = \text{diag}\{\beta_{mn}^1, \ldots, \beta_{mn}^K\}$ and $\beta_{mE}$ represent the path-losses from the $m^{\text{th}}$ BS to the $K$ MTs and the eavesdropper in the $n^{\text{th}}$ cell, respectively. Matrix $\mathbf{H}_{mn} \in \mathbb{C}^{K \times N_T}$, with row vector $\mathbf{h}_{mn}^k \in \mathbb{C}^{1 \times N_T}$ in the $k^{\text{th}}$ row, and matrix $\mathbf{H}_{mE} \in \mathbb{C}^{N_E \times N_T}$ represent the corresponding small-scale fading components. Their elements are modeled as mutually independent and identically distributed (i.i.d.) complex Gaussian random variables (RVs) with zero mean and unit variance.

For the design of the data and noise precoders, we consider two different approaches: *Selfish* designs and *collaborative* designs. For the selfish designs, each BS designs its precoders only based on the estimate of the CSI in its own cell, $\mathbf{G}_{nn}$, and without regard for the interference and the AN it causes to other cells. In contrast, for the collaborative designs, each BS designs its precoders based on the estimates of the CSI to the MTs in all cells, $\mathbf{G}_{mn}$, $m = 1, \ldots, M$, in an effort to avoid excessive interference and AN to other cells. Although collaborative designs introduce more channel estimation overhead at the BS, they may not always outperform selfish designs because of the imperfection of the CSI and the limited number of spatial degrees of freedom available for precoder design.

### B. Channel Estimation and Pilot Contamination

As is customary for massive MIMO systems, we assume that the downlink and uplink channels are reciprocal and the CSI is estimated in an uplink training phase [3]–[6]. To this end, all MTs emit pilot sequences of length $\tau \geq K$ and with pilot symbol power $p_\tau$. We assume that the pilot sequences of the $K$ MTs in a given cell are mutually orthogonal but the same pilot sequences are used in all cells. This gives rise to so-called pilot contamination [3]–[6]. Furthermore, we assume that the path-loss information changes on a much slower time scale than the small-scale fading. Hence, the path-loss matrices $\mathbf{D}_{nm}$, $m = 1, \ldots, M$, can be estimated perfectly and are assumed to be known at the BS for minimum mean-square error (MMSE) estimation of the small-scale fading gains [6]. At the $n^{\text{th}}$ BS, the small-scale fading vector to the $k^{\text{th}}$ MT in the $m^{\text{th}}$ cell, $\mathbf{h}_{nm}^k$, can be expressed as

$$\mathbf{h}_{nm}^k = \hat{\mathbf{h}}_{nm}^k + \tilde{\mathbf{h}}_{nm}^k, \tag{3}$$

where the estimate $\hat{\mathbf{h}}_{nm}^k$ and the estimation error $\tilde{\mathbf{h}}_{nm}^k$ are mutually independent and can be statistically characterized as $\hat{\mathbf{h}}_{nm}^k \sim \mathbb{CN}(\mathbf{0}_{N_T}, \frac{p_\tau \tau \beta_{nm}^k}{1+p_\tau \tau \sum_{l=1}^{M} \beta_{nl}^k}\mathbf{I}_{N_T})$ and $\tilde{\mathbf{h}}_{nm}^k \sim \mathbb{CN}(\mathbf{0}_{N_T}, \frac{1+p_\tau \tau \sum_{l\neq m}^{M} \beta_{nl}^k}{1+p_\tau \tau \sum_{l=1}^{M} \beta_{nl}^k}\mathbf{I}_{N_T})$, respectively, cf. [18]. For future reference, we collect the estimates and the estimation errors at the $n^{\text{th}}$ BS corresponding to all $K$ MTs in the $m^{\text{th}}$ cell in matrices $\hat{\mathbf{H}}_{nm} = [(\hat{\mathbf{h}}_{nm}^1)^T, \ldots, (\hat{\mathbf{h}}_{nm}^K)^T]^T \in \mathbb{C}^{K \times N_T}$ and $\tilde{\mathbf{H}}_{nm} = [(\tilde{\mathbf{h}}_{nm}^1)^T, \ldots, (\tilde{\mathbf{h}}_{nm}^K)^T]^T \in \mathbb{C}^{K \times N_T}$, respectively.

### C. Ergodic Secrecy Rate

The performance metric adopted in this paper is the ergodic secrecy rate [10]. In this section, we review some results for the ergodic secrecy rate in multi-cell massive MIMO systems employing linear data and AN precoding from [18], as these results will be needed throughout this paper. Combining (1) and (2) we observe that the downlink channel comprising the BS, the $k^{\text{th}}$ MT, and the eavesdropper in the $n^{\text{th}}$ cell is an instance of a multiple-input, single-output, multi-eavesdropper (MISOME) wiretap channel [8]. Hence, the achievable secrecy rate of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell is bounded by the difference of the capacities of the channel between the BS and the MT and the channel between the BS and the eavesdropper, see [18, Lemma 1], [22, Lemma 2]. Thus, a lower bound on the ergodic secrecy rate of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell is given by [18]

$$R_{nk}^{\text{sec}} = [R_{nk} - C_{nk}^{\text{eve}}]^+, k = 1, \ldots, K, \tag{4}$$

where $R_{nk}$ denotes an achievable rate of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell and $C_{nk}^{\text{eve}}$ denotes the ergodic capacity of the channel between the BS and the eavesdropper. In order to obtain a tractable lower bound on the ergodic secrecy rate, we lower bound the achievable rate of the MT as $R_{nk} = \log_2(1 + \gamma_{nk})$ with signal-to-interference-and-noise ratio (SINR) [18, Eq. (10)] given in (5) shown at the bottom of the next page. Furthermore, we make the pessimistic assumption that the eavesdropper is able to cancel the received signals of all in-cell and out-of-cell MTs except the signal intended for the MT of interest. This leads to an upper bound for the eavesdropper's capacity, and consequently, to a lower bound for the ergodic secrecy rate.[1] Hence, the ergodic capacity of the eavesdropper is given by [18, Eq. (7)]

$$C_{nk}^{\text{eve}} = \mathbb{E}\left[\log_2\left(1 + p\mathbf{f}_{nk}^H \mathbf{G}_{nE}^H \mathbf{X}^{-1} \mathbf{G}_{nE}\mathbf{f}_{nk}\right)\right], \tag{6}$$

[1]This lower bound is achievable if the eavesdropper has access to the data of all interfering in-cell and out-of-cell MTs, which might be the case e.g. if the interfering MTs cooperate with the eavesdropper.

where $\mathbf{X} = q\sum_{m=1}^{M}\mathbf{G}_{mE}\mathbf{A}_m\mathbf{A}_m^H\mathbf{G}_{mE}^H \in \mathbb{C}^{N_T\times N_T}$ denotes the noise correlation matrix at the eavesdropper under the worst-case assumption that the receiver noise at the eavesdropper is negligible, i.e., $\sigma_E^2 \to 0$. Denoting the normalized number of eavesdropper antennas by $\alpha = N_E/N_T$, a necessary condition for the invertibility of matrix $\mathbf{X}$ is $\alpha \le ML/N_T$. Hence, a non-zero secrecy rate can only be achieved if this condition is met. Consequently, a larger $L$ implies that the BS is able to tolerate more eavesdropper antennas.

If $\mathbf{H}_{nE}\mathbf{f}_{nk}$ and matrix $\mathbf{X}$ are statistically independent, which in turn means for the data and AN precoders that vector $\mathbf{f}_{nk}$ and the subspace spanned by the columns of $\mathbf{A}_n$ are mutually orthogonal, a simple and tight upper bound on (6) can be obtained. Since any efficient data/AN precoder pair has to keep the AN self-interference at the desired MT small, this orthogonality condition holds at least approximately in practice. In this case, for $\alpha < a^2L/(cN_T)$ and $N_T \to \infty$, where $a = 1+\sum_{m\neq n}^{M}\beta_{mE}/\beta_{nE}$ and $c = 1+\sum_{m\neq n}^{M}(\beta_{mE}/\beta_{nE})^2$, a simple and tight upper bound for $C_{nk}^{\text{eve}}$ is given by [18, Theorem 1]

$$C_{nk}^{\text{eve}} \le \log_2\left(1+\frac{\alpha p}{aqL/N_T - c\alpha q/a}\right) = \log_2\left(1+\frac{\alpha\phi}{\beta(1-\phi)(a-c\alpha N_T/(La))}\right). \quad (7)$$

For $M = 1$, we have $a^2/c = M = 1$, i.e., the bound in (7) is applicable in the entire range of $\alpha$ where $C_{nk}^{\text{eve}}$ in (6) is finite. For $M > 1$, we have $a^2/c \le M$, i.e., the bound is not applicable for $La^2/(cN_T) \le \alpha \le ML/N_T$. However, for strong inter-cell interference, we have $\beta_{mE} \approx \beta_{nE}$ and $a^2/c \approx M$, i.e., the bound is applicable for all $\alpha$ for which $C_{nk}^{\text{eve}}$ in (6) is finite. On the other hand, for weak inter-cell interference, we have $\beta_{mE} \ll \beta_{nE}$, and matrix $\mathbf{X}$ will be ill-conditioned for $L/N_T \le \alpha \le ML/N_T$ and $C_{nk}^{\text{eve}}$ will become very large. Hence, the bound is again applicable for the values of $\alpha$ (i.e., $0 \le \alpha \le L/N_T$), for which $C_{nk}^{\text{eve}}$ in (6) assumes practically relevant values. More generally, [18, Figs. 2–4] and Section VI suggest that, for $N_T \to \infty$, (7) is applicable and tight for all values of $\alpha$ which permit a non-vanishing secrecy rate.

Combining (4), (5), and (7), we obtain a tight and tractable lower bound on the secrecy rate [18]. It is noteworthy that the upper bound on the capacity of the eavesdropper in (7) is only affected by the dimensionality of the AN precoder, $L$, but not by the exact structures of $\mathbf{A}_n$ and $\mathbf{F}_n$, as long as $\mathbf{f}_{nk}$ and the subspace spanned by the columns of $\mathbf{A}_n$ are orthogonal. On the other hand, the achievable rate of the MT in (5) is affected by both the data and the AN precoders. In the following two sections, we analyze the impact of the most important existing data and AN precoder designs on the achievable rate $R_{nk}$ as $N_T \to \infty$, respectively, and propose novel low-complexity data and AN precoders that are based on a polynomial matrix expansion.

## III. Linear Data Precoders for Secure Massive MIMO

In this section, we analyze the achievable rate of selfish and collaborative ZF/RCI data precoding, respectively, and develop a novel POLY data precoder. In contrast to existing analyses and designs of data precoders for massive MIMO, e.g. [25], [26], [29]–[31], the results presented in this section account for the effect of AN leakage, which is only present if AN is injected at the BS for secrecy enhancement. We are interested in the asymptotic regime where $K, N_T \to \infty$ but $\beta = K/N_T$ and $\alpha = N_E/N_T$ are finite.

### A. Analysis of Existing Data Precoders

For $N_T \to \infty$, analyzing the achievable rate is equivalent to analyzing the SINR in (5). Thereby, the effect of the AN precoder can be captured by the term

$$Q = \sum_{m=1}^{M}\sum_{i=1}^{N_t}\mathbb{E}\left[\left|\sqrt{\beta_{mn}^k}\mathbf{h}_{mn}^k\mathbf{a}_{mi}\right|^2\right] = \sum_{m=1}^{M}\beta_{mn}^k\mathbb{E}\left[\mathbf{h}_{mn}^k\mathbf{A}_m\mathbf{A}_m^H\left(\mathbf{h}_{mn}^k\right)^H\right] \quad (8)$$

in the denominator of (5), which represents the inter-cell and intra-cell AN leakage. This term is assumed to be given in this section and will be analyzed in detail for different AN precoders in Section IV.

*1) Selfish ZF/RCI Data Precoding:* The selfish RCI (SRCI) data precoder for the $n^{\text{th}}$ cell is given by

$$\mathbf{F}_n = \gamma_1\mathbf{L}_{nn}\hat{\mathbf{H}}_{nn}^H, \quad (9)$$

where $\mathbf{L}_{nn} = (\hat{\mathbf{H}}_{nn}^H\hat{\mathbf{H}}_{nn} + \kappa_1\mathbf{I}_{N_T})^{-1}$, $\gamma_1$ is a scalar normalization constant, and $\kappa_1$ is a regularization constant. In the following proposition, we provide the resulting SINR of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell.

*Proposition 1:* For SRCI data precoding, the received SINR at the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell is given by $\gamma_{nk}^{\text{SRCI}} =$

$$\frac{1}{\frac{\hat{\Gamma}_{\text{SRCI}}+(1+\mathcal{G}(\beta,\kappa_1))^2}{\mathcal{G}(\beta,\kappa_1)\left(\hat{\Gamma}_{\text{SRCI}}+\frac{\hat{\Gamma}_{\text{SRCI}}\kappa_1}{\beta}(1+\mathcal{G}(\beta,\kappa_1))^2\right)}+\sum_{m\neq n}\beta_{mn}^k/\beta_{nn}^k}, \quad (10)$$

where

$$\mathcal{G}(\beta,\kappa_1) = \frac{1}{2}\left[\sqrt{\frac{(1-\beta)^2}{\kappa_1^2}+\frac{2(1+\beta)}{\kappa_1}+1}+\frac{1-\beta}{\kappa_1}-1\right], \quad (11)$$

$$\gamma_{nk} = \frac{\left|\mathbb{E}\left[\sqrt{\beta_{nn}^k p}\mathbf{h}_{nn}^k\mathbf{f}_{nk}\right]\right|^2}{\text{var}\left[\sqrt{\beta_{nn}^k p}\mathbf{h}_{nn}^k\mathbf{f}_{nk}\right]+\sum_{m=1}^{M}\sum_{i=1}^{N_t}\mathbb{E}\left[\left|\sqrt{\beta_{mn}^k q}\mathbf{h}_{mn}^k\mathbf{a}_{mi}\right|^2\right]+\sum_{\{m,l\}\neq\{n,k\}}\mathbb{E}\left[\left|\sqrt{\beta_{mn}^k p}\mathbf{h}_{mn}^k\mathbf{f}_{ml}\right|^2\right]+1}. \quad (5)$$

and $\hat{\Gamma}_{\mathrm{SRCI}} = \frac{\Gamma_{\mathrm{SRCI}}\theta_{nk}}{\Gamma_{\mathrm{SRCI}}\vartheta_{nk}+1}$ with $\Gamma_{\mathrm{SRCI}} = \frac{\beta_{nn}^k K}{\sum_{m\neq n}^{M}\sum_{l\neq k}\beta_{mn}^k+\eta Q+\frac{K}{\phi P_T}}$, $\theta_{mk} = \frac{p_\tau \tau(\beta_{mn}^k)^2}{1+p_\tau\tau\sum_{l=1}^{M}\beta_{ml}^k}$, $\vartheta_{mk} = \beta_{mn}^k \times \frac{1+p_\tau\tau\sum_{l\neq m}^{M}\beta_{ml}^k}{1+p_\tau\tau\sum_{l=1}^{M}\beta_{ml}^k}$, and $\eta = q/p$.

*Proof:* Please refer to Appendix A. ■

Regularization constant $\kappa_1$ can be optimized for maximization of the lower bound on the secrecy rate in (4), which is equivalent to maximizing the SINR in (10). Setting the derivative of $\gamma_{nk}^{\mathrm{SRCI}}$ with respect to $\kappa_1$ to zero, the optimal regularization parameter is found as $\kappa_{1,\,\mathrm{opt}} = \beta/\hat{\Gamma}_{\mathrm{SRCI}}$, and the corresponding maximum SINR is given by

$$\gamma_{nk}^{\mathrm{SRCI}} = \frac{1}{1/\mathcal{G}(\beta,\kappa_{1,\,\mathrm{opt}}) + \sum_{m\neq n}\beta_{mn}^k/\beta_{nn}^k}. \tag{12}$$

On the other hand, for $\kappa_1 \to 0$, the SRCI data precoder in (9) reduces to the selfish ZF (SZF) data precoder. The corresponding received SINR is provided in the following corollary.

*Corollary 1:* Assuming $\beta \leq 1$, for SZF data precoding, the received SINR at the $k^{\mathrm{th}}$ MT in the $n^{\mathrm{th}}$ cell is given by

$$\gamma_{nk}^{\mathrm{SZF}} = \frac{1}{\frac{\beta}{(1-\beta)\hat{\Gamma}_{\mathrm{SRCI}}} + \sum_{m\neq n}\beta_{mn}^k/\beta_{nn}^k}. \tag{13}$$

*Proof:* $\gamma_{nk}^{\mathrm{SZF}}$ in (13) can be obtained from (10) as $\gamma_{nk}^{\mathrm{SZF}} = \lim_{\kappa_1\to 0}\gamma_{nk}^{\mathrm{SRCI}}$. ■

*2) Collaborative ZF/RCI Precoding:* The collaborative RCI (CRCI) precoder for the $n^{\mathrm{th}}$ cell is given by

$$\mathbf{F}_n = \gamma_2 \mathbf{L}_n \hat{\mathbf{H}}_{nn}^H, \tag{14}$$

where $\mathbf{L}_n = (\hat{\mathbf{H}}_n^H\hat{\mathbf{H}}_n + \kappa_2\mathbf{I}_{N_T})^{-1}$ with $\hat{\mathbf{H}}_n = [\hat{\mathbf{H}}_{n1}^T \ldots \hat{\mathbf{H}}_{nM}^T]^T \in \mathbb{C}^{MK\times N_T}$, $\gamma_2$ is a normalization constant, and $\kappa_2$ is a regularization constant. The corresponding SINR of the $k^{\mathrm{th}}$ MT in the $n^{\mathrm{th}}$ cell is provided in the following proposition.

*Proposition 2:* For CRCI data precoding, the received SINR at the $k^{\mathrm{th}}$ MT in the $n^{\mathrm{th}}$ cell is given by $\gamma_{nk}^{\mathrm{CRCI}} =$

$$\frac{1}{\frac{\hat{\Gamma}_{\mathrm{CRCI}}+(1+\mathcal{G}(M\beta,\kappa_2))^2}{\mathcal{G}(M\beta,\kappa_2)\left(\hat{\Gamma}_{\mathrm{CRCI}}+\frac{\hat{\Gamma}_{\mathrm{CRCI}}\kappa_2}{\beta}(1+\mathcal{G}(M\beta,\kappa_2))^2\right)} + \sum_{m\neq n}\beta_{mn}^k/\beta_{nn}^k}, \tag{15}$$

where $\hat{\Gamma}_{\mathrm{CRCI}} = \frac{\Gamma_{\mathrm{CRCI}}\theta_{nk}}{\Gamma_{\mathrm{CRCI}}\vartheta_{nk}+1}$ with $\Gamma_{\mathrm{CRCI}} = \frac{\beta_{nn}^k K}{\eta Q+\frac{K}{\phi P_T}}$.

*Proof:* The proof is similar to that for the SINR for the SRCI data precoder given in Appendix A and omitted here for brevity. ■

Furthermore, the optimal regularization constant maximizing the SINR (and thus the secrecy rate) in (15) is obtained as $\kappa_{2,\,\mathrm{opt}} = M\beta/\hat{\Gamma}_{\mathrm{CRCI}}$, and the corresponding maximum SINR is given by

$$\gamma_{nk}^{\mathrm{CRCI}} = \frac{1}{1/\mathcal{G}(M\beta,\kappa_{2,\,\mathrm{opt}}) + \sum_{m\neq n}\beta_{mn}^k/\beta_{nn}^k}. \tag{16}$$

On the other hand, for $\kappa_2 \to 0$, the CRCI precoder in (14) reduces to the collaborative ZF (CZF) precoder. The corresponding received SINR is provided in the following corollary.

*Corollary 2:* Assuming $\beta \leq 1/M$, for CZF data precoding, the received SINR at the $k^{\mathrm{th}}$ MT in the $n^{\mathrm{th}}$ cell is given by

$$\gamma_{nk}^{\mathrm{CZF}} = \frac{1}{\frac{M\beta}{(1-M\beta)\hat{\Gamma}_{\mathrm{CRCI}}} + \sum_{m\neq n}\beta_{mn}^k/\beta_{nn}^k}. \tag{17}$$

*Proof:* $\gamma_{nk}^{\mathrm{CZF}}$ in (17) is obtained by letting $\kappa_2 \to 0$ in (15). ■

*Remark 1:* Selfish data precoders require estimation of in-cell CSI, i.e., $\hat{\mathbf{H}}_{nn}$, only. In contrast, collaborative data precoders require estimation of both in-cell and inter-cell CSI at the BS, i.e., $\hat{\mathbf{H}}_n$. Furthermore, since collaborative data precoders attempt to avoid interference not only to in-cell users but also to out-of-cell users, more BS antennas are needed to achieve high performance. This is evident from Corollaries 1 and 2, which reveal that $N_T > K$ and $N_T > MK$ are necessary for SZF and CZF data precoding, respectively. On the other hand, if successful, trying to avoid out-of-cell interference is beneficial for the overall performance. Hence, whether selfish or collaborative precoders are preferable depends on the parameters of the considered system, cf. Sections V and VI.

### *B. Polynomial Data Precoder*

The RCI and ZF data precoders introduced in the previous section achieve a higher performance than simple MF data precoding [18]. However, they require a matrix inversion which entails a high computational complexity for the large values of $K$ and $N_T$ desired in massive MIMO. Hence, in this section, we propose a low-complexity POLY data precoder which avoids the matrix inversion. As the goal is a low-complexity design, we focus on selfish POLY precoders, although the extension to collaborative designs is possible.

The proposed POLY precoder, $\mathbf{F}_n$, for the $n^{\mathrm{th}}$ BS can be expressed as

$$\mathbf{F}_n = \frac{1}{\sqrt{N_T}}\hat{\bar{\mathbf{H}}}_{nn}^H \sum_{i=0}^{\mathcal{J}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^i, \tag{18}$$

where $\hat{\bar{\mathbf{H}}}_{nn} = \frac{1}{\sqrt{N_T}}\hat{\mathbf{H}}_{nn}$, and $\boldsymbol{\mu} = [\mu_0, \ldots, \mu_{\mathcal{J}}]^T$ are the real-valued coefficients of the precoder matrix polynomial, which have to be optimized. In the following, we show that, for $K, N_T \to \infty$, the optimum coefficients $\boldsymbol{\mu}$ do not depend on the instantaneous channel estimates but are constant and can be determined by exploiting results from free probability [35] and random matrix theory [38]. To this end, we define the asymptotic average mean-square error (MSE) of the users in the $n^{\mathrm{th}}$ cell as $\mathrm{mse}_n = \lim_{K\to\infty}\frac{1}{K}\mathbb{E}\left[\|\mathbf{e}_n\|^2\right]$ with error vector

$$\mathbf{e}_n = \varsigma\mathbf{y}_n - \mathbf{s}_n = \varsigma(\mathbf{G}_{nn}(\sqrt{p}\mathbf{F}_n\mathbf{s}_n + \sqrt{q}\mathbf{A}_n\mathbf{z}_n) + \tilde{\mathbf{n}}_n) - \mathbf{s}_n, \tag{19}$$

where $\tilde{\mathbf{n}}_n = \sum_{m\neq n}\mathbf{G}_{mn}\mathbf{x}_m + \mathbf{n}_n$ includes Gaussian noise, inter-cell interference, and inter-cell AN leakage. Furthermore, $\varsigma$ is a normalization constant at the receiver, which does not impact detection performance. The optimal coefficient vector $\boldsymbol{\mu}$ minimizes $\mathrm{mse}_n$ for a given power budget $\phi P_T$ for the information-carrying signal, i.e.,

$$\min_{\boldsymbol{\mu},\varsigma}\ \mathrm{mse}_n \qquad \text{s.t.: } \mathrm{Tr}\{\mathbf{F}_n^H\mathbf{F}_n\} = 1, \tag{20}$$

where we use the notation $\mathrm{Tr}\{\cdot\} = \lim_{K\to\infty} \mathrm{tr}\{\cdot\}/K$. The optimal coefficient vector, $\boldsymbol{\mu}_{\mathrm{opt}}$, is provided in the following theorem.

*Theorem 1:* For $K, N_T \to \infty$, the optimal coefficient vector minimizing the asymptotic average MSE of the users in the $n^{\mathrm{th}}$ cell for the POLY precoder in (18) is given by

$$\boldsymbol{\mu}_{\mathrm{opt}} = \gamma_3 \boldsymbol{\Pi}^{-1} \boldsymbol{\psi}, \tag{21}$$

where $\boldsymbol{\psi} = [\zeta, \zeta^2, \ldots, \zeta^{\mathcal{J}+1}]^T$, $[\boldsymbol{\Pi}]_{i,j} = \mathrm{Tr}\{\mathbf{D}_{nn}\}\zeta^{i+j} + \left(\mathrm{Tr}\{\mathbf{D}_{nn}\boldsymbol{\Delta}_n\} + \frac{\mathrm{Tr}\{\boldsymbol{\Sigma}_n\}+P_{\mathrm{AN}}}{N_T p}\right)\zeta^{i+j-1}$, $\boldsymbol{\Sigma}_n = \mathbb{E}[\tilde{\mathbf{n}}_n \tilde{\mathbf{n}}_n^H]$, $\boldsymbol{\Delta}_n = \mathrm{diag}\left\{\frac{1+p_\tau\tau\sum_{m\neq n}\beta_{nm}^1}{1+p_\tau\tau\sum_{m=1}^M \beta_{nm}^1}, \cdots, \frac{1+p_\tau\tau\sum_{m\neq n}\beta_{nm}^K}{1+p_\tau\tau\sum_{m=1}^M \beta_{nm}^K}\right\}$, and $P_{\mathrm{AN}} = q\mathbb{E}\left[\mathrm{Tr}\left\{\mathbf{G}_{nn}\mathbf{A}_n\mathbf{A}_n^H\mathbf{G}_{nn}^H\right\}\right]$. Furthermore, $\zeta^l$ denotes the $l^{\mathrm{th}}$-order moment of the sum of the eigenvalues of $\hat{\mathbf{H}}_{nn}\hat{\mathbf{H}}_{nn}^H$, i.e., $\zeta^l = \lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^K \lambda_k^l$, which converges to $\zeta^l = \sum_{i=0}^{l-1}\binom{l}{i}\binom{l}{i+1}\frac{\beta^i}{l}$ for $K\to\infty$ [29, Theorem 2]. Finally, $\gamma_3$ is chosen such that $\mathrm{Tr}\{\mathbf{F}_n^H\mathbf{F}_n\} = 1$ holds.

*Proof:* Please refer to Appendix B. ■

We note that $\boldsymbol{\mu}_{\mathrm{opt}}$ does not depend on instantaneous channel estimates, and hence, can be computed offline.

### C. Computational Complexity of Data Precoding

We compare the computational complexity of the considered data precoders in terms of the number of floating point operations (FLOPs) [37]. Each FLOP represents one scalar complex addition or multiplication. We assume that the coherence time of the channel is $T$ symbol intervals of which $\tau$ are used for training and $T-\tau$ are used for data transmission. Hence, the complexity required for precoding in one coherence interval is comprised of the complexity required for generating one precoding matrix and $T-\tau$ precoded vectors. A similar complexity analysis was conducted in [29, Section IV] for selfish data precoders without AN injection at the BS. Since the AN injection does not affect the structure of the data precoders, we can directly adapt the results from [29, Section IV] to the case at hand. In particular, the selfish MF, the SZF/SRCI, and the CZF/CRCI precoders require $(2K-1)N_T(T-\tau)$, $0.5(K^2+K)(2N_T-1)+K^3+K^2+K+N_TK(2K-1)+(2K-1)N_T(T-\tau)$, and $0.5(M^2K^2+MK)(2N_T-1)+M^3K^3+M^2K^2+MK+N_TMK(2MK-1)+(2K-1)N_T(T-\tau)$ FLOPs per coherence interval, see [29, Section IV]. In contrast, for the POLY data precoder, we obtain for the overall computational complexity $(T-\tau)((\mathcal{J}+1)(2K-1)N_T+\mathcal{J}(2N_T-1)K)$ FLOPs, which assumes implementation of the precoding operation by Horner's rule [29, Section IV].

The above complexity expressions reveal that the additional complexity introduced by collaborative data precoders compared to selfish data precoders is at most a factor of $M^3$. In addition, the complexity savings achieved with the POLY data precoder compared to the SZF/SRCI data precoders increase with increasing $K$ for a given $T$. We note however that, regardless of their complexity, POLY data precoders are attractive as they avoid the stability issues that may arise in fixed point implementation of large matrix inverses.

## IV. Linear AN Precoders for Secure Massive MIMO

In this section, we investigate the performance of selfish and collaborative NS (S/CNS) and random AN precoders. In addition, a novel POLY AN precoder is derived. To the best of the authors' knowledge, POLY AN precoding has not been considered in the literature before.

### A. Analysis of Existing AN Precoders

For a given dimensionality of the AN precoder, $L$, the secrecy rate depends on the AN precoder only via the AN leakage, $Q$, given in (8), which affects the SINR of the MT. Furthermore, the optimal POLY data precoder coefficients in (21) are affected by the AN precoder via the leakage term $P_{\mathrm{AN}}$. In this subsection, for $N_T\to\infty$, we will provide closed-form expressions for $Q$ and $P_{\mathrm{AN}}$ for the SNS, CNS, and random AN precoders.

*1) SNS AN Precoder:* The SNS AN precoder of the $n^{\mathrm{th}}$ BS is given by [11]

$$\mathbf{A}_n = \mathbf{I}_{N_T} - \hat{\mathbf{H}}_{nn}^H\left(\hat{\mathbf{H}}_{nn}\hat{\mathbf{H}}_{nn}^H\right)^{-1}\hat{\mathbf{H}}_{nn}, \tag{22}$$

which has rank $L = N_T - K$ and exists only if $\beta < 1$. We divide the corresponding AN leakage $Q_{\mathrm{SNS}}$ into an inter-cell AN leakage $Q_o^{\mathrm{SNS}}$ and an intra-cell AN leakage $Q_i^{\mathrm{SNS}}$, where $Q_{\mathrm{SNS}} = Q_o^{\mathrm{SNS}} + Q_i^{\mathrm{SNS}}$. For the SNS AN precoder, $Q_o^{\mathrm{SNS}}$ is obtained as

$$\begin{aligned} Q_o^{\mathrm{SNS}} &= \sum_{m\neq n}\beta_{mn}^k \mathbb{E}\left[\mathbf{h}_{mn}^k\mathbf{A}_m\mathbf{A}_m^H(\mathbf{h}_{mn}^k)^H\right] \\ &= \mathbb{E}\left[\mathrm{tr}\left\{\mathbf{A}_m\mathbf{A}_m^H\right\}\right]\sum_{m\neq n}^M \beta_{mn}^k \\ &= (N_T-K)\sum_{m\neq n}^M \beta_{mn}^k, \end{aligned} \tag{23}$$

where we exploited [30, Lemma 11] and the independence of $\mathbf{A}_m$ and $\mathbf{h}_{mn}^k$. In contrast, the intra-cell AN leakage power is given by

$$\begin{aligned} Q_i^{\mathrm{SNS}} &= \beta_{nn}^k \mathbb{E}\left[\mathbf{h}_{nn}^k\mathbf{A}_n\mathbf{A}_n^H(\mathbf{h}_{nn}^k)^H\right] \\ &= \beta_{nn}^k \mathbb{E}\left[\tilde{\mathbf{h}}_{nn}^k\mathbf{A}_n\mathbf{A}_n^H(\tilde{\mathbf{h}}_{nn}^k)^H\right] \\ &= (N_T-K)\beta_{nn}^k\frac{1+p_\tau\tau\sum_{m\neq n}^M\beta_{nm}^k}{1+p_\tau\tau\sum_{m=1}^M\beta_{nm}^k}, \end{aligned} \tag{24}$$

as the SNS AN precoder matrix lies in the null space of the estimated channels of all $K$ MTs in the $n^{\mathrm{th}}$ cell. Similarly, the AN leakage relevant for computation of the POLY data precoder is obtained as

$$P_{\mathrm{AN}}^{\mathrm{SNS}} = (1-\phi)P_T\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^K\beta_{nn}^k\frac{1+p_\tau\tau\sum_{m\neq n}^M\beta_{nm}^k}{1+p_\tau\tau\sum_{m=1}^M\beta_{nm}^k}. \tag{25}$$

*2) CNS AN Precoder:* For the CNS AN precoder at the $n^{\text{th}}$ BS, the AN is designed to lie in the null space of the estimated channels between all $MK$ MTs and the BS, i.e.,

$$\mathbf{A}_n = \mathbf{I}_{N_T} - \hat{\mathbf{H}}_n^H \left(\hat{\mathbf{H}}_n \hat{\mathbf{H}}_n^H\right)^{-1} \hat{\mathbf{H}}_n, \tag{26}$$

which has rank $L = N_T - MK$ and exists only if $\beta < 1/M$. The corresponding AN leakage to the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell is given by

$$\begin{aligned} Q_{\text{CNS}} &= \sum_{m=1}^{M} \beta_{mn}^k \mathbb{E}\left[\mathbf{h}_{mn}^k \mathbf{A}_m \mathbf{A}_m^H (\mathbf{h}_{mn}^k)^H\right] \\ &= (N_T - MK) \sum_{m=1}^{M} \beta_{mn}^k \frac{1 + p_\tau \tau \sum_{l \neq m}^{M} \beta_{ml}^k}{1 + p_\tau \tau \sum_{l=1}^{M} \beta_{ml}^k}. \end{aligned} \tag{27}$$

Furthermore, the CNS AN precoder results in the same $P_{\text{AN}}$ as the SNS AN precoder, cf. (25).

*3) Random AN Precoder:* For the random precoder, all elements of $\mathbf{A}_n$ are i.i.d. random variables independent of the channel [18], i.e., $\mathbf{A}_n$ has rank $L = N_T$. Hence, $\mathbf{h}_{mn}^k$ and $\mathbf{A}_m$, $\forall m$, are mutually independent, and we obtain

$$Q_{\text{random}} = \sum_{m=1}^{M} \beta_{mn}^k \mathbb{E}\left[\mathbf{h}_{mn}^k \mathbf{A}_m \mathbf{A}_m^H (\mathbf{h}_{mn}^k)^H\right] = N_T \sum_{m=1}^{M} \beta_{mn}^k. \tag{28}$$

Furthermore, we obtain

$$P_{\text{AN}}^{\text{random}} = (1 - \phi) P_T \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \beta_{nn}^k. \tag{29}$$

*Remark 2:* If the power and time allocated to channel estimation are very small, i.e., $\tau p_\tau \to 0$, the S/CNS AN precoders yield the same $qQ$ and $P_{\text{AN}}$ as the random AN precoder. This suggests that in this regime all considered AN precoders achieve a similar SINR performance for a given MT. However, for $\tau p_\tau > 0$, the S/CNS AN precoders cause less AN leakage resulting in an improved SINR performance compared to the random precoder at the expense of a higher complexity.

### B. POLY AN Precoder

To mitigate the high computational complexity imposed by the matrix inversion required for the S/CNS AN precoders, while achieving an improved performance compared to the random AN precoder, we propose a POLY AN precoder. Similar to the POLY data precoder, we concentrate on the selfish design because of the desired low complexity, and hence, set $L = N_T - K$. The proposed POLY AN precoder is given by

$$\mathbf{A}_n = \mathbf{I}_{N_T} - \hat{\bar{\mathbf{H}}}_{nn}^H \left(\sum_{i=0}^{\mathcal{J}} \nu_j \left(\hat{\bar{\mathbf{H}}}_{nn} \hat{\bar{\mathbf{H}}}_{nn}^H\right)^j\right) \hat{\bar{\mathbf{H}}}_{nn}, \tag{30}$$

where $\boldsymbol{\nu} = [\nu_0, \ldots, \nu_{\mathcal{J}}]^T$ contains the real-valued coefficients of the AN precoder polynomial, which have to be optimized. In particular, $\boldsymbol{\nu}$ is optimized for minimization of the asymptotic average AN leakage caused to all MTs in the $n^{\text{th}}$ cell $P_{\text{AN}}$. The corresponding optimization problem is formulated as

$$\begin{aligned} &\min_{\boldsymbol{\nu}} P_{\text{AN}} = q\mathbb{E}\left[\text{Tr}\{\mathbf{G}_{nn} \mathbf{A}_n \mathbf{A}_n^H \mathbf{G}_{nn}^H\}\right] \\ &\text{s.t.: } \text{Tr}\{\mathbf{A}_n^H \mathbf{A}_n\} = 1/\beta - 1. \end{aligned} \tag{31}$$

The solution of (31) is provided in the following theorem.

*Theorem 2:* For $K, N_T \to \infty$, the optimal coefficient vector minimizing the asymptotic average AN leakage caused to the users in the $n^{\text{th}}$ cell for the AN precoder structure in (30) is given by

$$\boldsymbol{\nu}_{\text{opt}} = \boldsymbol{\Sigma}^{-1} \boldsymbol{\omega}, \tag{32}$$

where $[\boldsymbol{\Sigma}]_{i,j} = \zeta^{i+j+1} + \epsilon \zeta^{i+j}$ and $\boldsymbol{\omega} = [\zeta^2 + \epsilon\zeta, \ldots, \zeta^{\mathcal{J}+2} + \epsilon\zeta^{\mathcal{J}+1}]$. Here, $\zeta^l$ denotes again the $l^{\text{th}}$ order moment of the sum of the eigenvalues of matrix $\hat{\bar{\mathbf{H}}}_{nn} \hat{\bar{\mathbf{H}}}_{nn}^H$, cf. Theorem 1. $\epsilon$ is chosen such that $\text{Tr}\{\mathbf{A}_n^H \mathbf{A}_n\} = 1/\beta - 1$.

*Proof:* Please refer to Appendix C. ■

### C. Computational Complexity of AN Precoding

Similarly to the data precoders, the complexity of the AN precoders is evaluated in terms of the number of flops required per coherence interval $T$. For the SNS AN precoder, the computation of $\mathbf{A}_n$ in (22) requires the computation and inversion of a $K \times K$ positive definite matrix, which entails $0.5(K^2 + K)(2N_T - 1) + K^3 + K^2 + K$ FLOPs [37], and the multiplication of an $N_T \times K$, an $K \times K$, and an $K \times N_T$ matrix, which entails $N_T(N_T + K)(2K - 1)$ FLOPs [37]. Furthermore, the $T - \tau$ vector-matrix multiplications required for AN precoding entail a complexity of $(2N_T - 1)N_T$ FLOPs [37], respectively. Hence, the overall complexity is $0.5(K^2 + K)(2N_T - 1) + K^3 + K^2 + K + N_T(N_T + K)(2K - 1) + (2N_T - 1)N_T(T - \tau)$ FLOPs. Similarly, for the CNS AN precoder, we obtain a complexity of $0.5((MK)^2 + MK)(2N_T - 1) + (MK)^3 + (MK)^2 + MK + N_T(N_T + MK)(2MK - 1) + (2N_T - 1)N_T(T - \tau)$ FLOPs, whereas the random AN precoder entails a complexity of $(2N_T - 1)N_T(T - \tau)$ FLOPs as only the AN vector-matrix multiplications are required.

Similar to the precoded data vector [29, Section IV], the POLY precoded AN vector can be generated using Horner's rule. Hence, based on (30), the transmitted AN vector in the $n^{\text{th}}$ cell can be obtained as

$$\mathbf{A}_n \mathbf{z}_n = \mathbf{z}_n - \left(\nu_0 \hat{\bar{\mathbf{H}}}_{nn}^H \hat{\bar{\mathbf{H}}}_{nn} \left(\mathbf{z}_n + \frac{\nu_1}{\nu_0} \hat{\bar{\mathbf{H}}}_{nn}^H \hat{\bar{\mathbf{H}}}_{nn} (\mathbf{z}_n + \ldots)\right)\right). \tag{33}$$

Hence, $\mathbf{A}_n \mathbf{z}_n$ can be computed efficiently by first multiplying $\hat{\bar{\mathbf{H}}}_{nn}$ with $\mathbf{z}_n$, which requires $(2N_T - 1)K$ FLOPs, then multiplying $\hat{\bar{\mathbf{H}}}_{nn}^H$ with the resulting vector, which requires $(2K - 1)N_T$ FLOPs, adding $\mathbf{z}_n$ to the newly resulting vector, and repeating similar operations $(\mathcal{J} + 1)$ times, see [27], [29] for details of Horner's rule. Overall, this leads to a complexity of $(\mathcal{J} + 1)\left((2K - 1)N_T + (2N_T - 1)K\right)(T - \tau)$ FLOPs.

TABLE I
SINR OF THE $k^{\text{TH}}$ MT IN THE $n^{\text{TH}}$ CELL FOR LINEAR DATA PRECODING AND THE SIMPLIFIED PATH-LOSS MODEL IN (34). FOR THIS MODEL, $\hat{\Gamma}_{\text{SRCI}}$ AND $\hat{\Gamma}_{\text{CRCI}}$ SIMPLIFY TO $\hat{\Gamma}_{\text{SRCI}} = \frac{\Gamma_{\text{SRCI}}\theta}{\Gamma_{\text{SRCI}}\vartheta+1}$ AND $\hat{\Gamma}_{\text{CRCI}} = \frac{\Gamma_{\text{CRCI}}\theta}{\Gamma_{\text{CRCI}}\vartheta+1}$ WHERE $\Gamma_{\text{SRCI}} = \frac{\beta\phi}{\beta\phi\rho(M-1)+(1-\phi)\beta\tilde{Q}+\frac{\beta}{P_T}}$, $\Gamma_{\text{CRCI}} = \frac{\beta\phi}{(1-\phi)\beta\tilde{Q}+\frac{\beta}{P_T}}$, $\theta = \frac{p_\tau\tau}{1+ap_\tau\tau}$, AND $\vartheta = \frac{1+(M-1)\rho p_\tau\tau}{1+ap_\tau\tau}$

| Data Precoder | $\gamma_{nk}$ |
|---|---|
| SZF | $\frac{\theta\phi(1-\beta)}{(1-\phi)\beta\tilde{Q}+\beta\phi(a-\theta)+(M-1)\rho^2\theta\phi(1-\beta)+\beta/P_T}$ |
| SRCI | $\frac{1}{1/\mathcal{G}(\beta,\beta/\hat{\Gamma}_{\text{SRCI}})+(M-1)\rho}$ |
| CZF | $\frac{\theta\phi(1-M\beta)}{(1-\phi)\beta\tilde{Q}+\beta\phi a(1-\theta)+(M-1)\rho^2\theta\phi(1-M\beta)+\beta/P_T}$ |
| CRCI | $\frac{1}{1/\mathcal{G}(a\beta,a\beta/\hat{\Gamma}_{\text{CRCI}})+(M-1)\rho}$ |
| MF | $\frac{\theta\phi}{(1-\phi)\beta\tilde{Q}+\beta\phi a+(M-1)\rho^2\theta\phi+\beta/P_T}$ |

TABLE II
AN LEAKAGE FOR SIMPLIFIED PATH-LOSS MODEL IN (34). $\theta$ AND $\vartheta$ ARE DEFINED IN THE CAPTION OF TABLE I

| AN Precoder | $\tilde{Q}$ | $P_{\text{AN}}$ | $L$ |
|---|---|---|---|
| SNS | $(a-\theta)$ | $(1-\phi)P_T\vartheta$ | $N_T-K$ |
| CNS | $a(1-\theta)$ | $(1-\phi)P_T\vartheta$ | $N_T-MK$ |
| Random | $a$ | $(1-\phi)P_T$ | $N_T$ |

## V. COMPARISON OF LINEAR DATA AND AN PRECODERS

In this subsection, we compare the secrecy performances of the considered data and AN precoders. Thereby, in order to get tractable results, we focus on the relative performances of SZF, CZF, and MF [18] data precoders and SNS, CNS, and random AN precoders. The performances of SRCI, CRCI, and POLY data precoders and the POLY AN precoder will be investigated via numerical and simulation results in Section VI.

In order to gain some insight for system design and analysis, we adopt a simplified path-loss model. In particular, we assume the path losses are given by

$$\beta_{mn}^k = \begin{cases} 1, & m = n \\ \rho, & \text{otherwise} \end{cases} \tag{34}$$

where $\rho \in [0, 1]$ denotes the inter-cell interference factor. For this simplified model, $a$ and $c$ in (7) simplify to $a = 1 + (M-1)\rho$ and $c = 1 + (M-1)\rho^2$. Furthermore, the SINR expressions of the linear data precoders considered in Section III-a and the MF precoder considered in [18] can be simplified considerably and are provided in Table I, where we use the normalized AN leakage $\tilde{Q} = Q/L$. The expressions for the normalized AN leakage $\tilde{Q}$, the asymptotic average AN leakage $P_{\text{AN}}$, and the dimensionality $L$ of the considered linear AN precoders are given in Table II.

### A. Comparison of SZF, CZF, and MF Data Precoders

In this subsection, we compare the performances achieved with SZF, CZF, and MF data precoders for a given AN precoder, i.e., $L$ and $\tilde{Q}$ are fixed. Since the upper bound on the capacity of the eavesdropper channel is independent of the adopted data precoder, cf. Section II-C, we compare the considered data precoders based on their SINRs. Exploiting the results in Table I, we obtain the following relations between $\gamma_{nk}^{\text{SZF}}$, $\gamma_{nk}^{\text{CZF}}$, and $\gamma_{nk}^{\text{MF}}$:

$$\frac{\gamma_{nk}^{\text{SZF}}}{\gamma_{nk}^{\text{MF}}} = 1 + \beta(c\gamma_{nk}^{\text{SZF}} - 1),$$
$$\frac{\gamma_{nk}^{\text{CZF}}}{\gamma_{nk}^{\text{SZF}}} = \frac{1-M\beta}{1-\beta} + \frac{a(a-1)\beta}{1-\beta}\gamma_{nk}^{\text{CZF}}. \tag{35}$$

Hence, for $\gamma_{nk}^{\text{SZF}} > \gamma_{nk}^{\text{MF}}$, we require $\gamma_{nk}^{\text{SZF}} > 1/c = 1/(1+\rho^2(M-1))$, and for $\gamma_{nk}^{\text{CZF}} > \gamma_{nk}^{\text{SZF}}$, we need $\gamma_{nk}^{\text{CZF}} > 1/(\rho a) = 1/[\rho(1+\rho(M-1))]$. As expected, (35) suggests that for a lightly loaded system, i.e., $\beta \to 0$, all three precoders have a similar performance, i.e., $\gamma_{nk}^{\text{CZF}} \approx \gamma_{nk}^{\text{SZF}} \approx \gamma_{nk}^{\text{MF}}$. In the following, we investigate the impact of the number of MTs and the pilot power on the relative performances of the considered data precoders.

*1) Number of MTs:* From (35), we find that for $\gamma_{nk}^{\text{SZF}} > \gamma_{nk}^{\text{MF}}$ and $\gamma_{nk}^{\text{CZF}} > \gamma_{nk}^{\text{SZF}}$ to hold, the number of MTs has to meet $K < K_{\text{SZF}>\text{MF}}$ and $K < K_{\text{CZF}>\text{SZF}}$, where

$$K_{\text{SZF}>\text{MF}} = \frac{\theta\phi N_T}{(1-\phi)\tilde{Q} + a\phi + 1/P_T}$$
$$K_{\text{CZF}>\text{SZF}} = \frac{\rho\phi\theta N_T}{(1-\phi)\tilde{Q} + [a(1-\theta)+\rho\theta M]\phi + 1/P_T}, \tag{36}$$

respectively. Interestingly, both the maximum numbers of MTs for which the SZF data precoder is advantageous compared to the MF data precoder, $K_{\text{SZF}>\text{MF}}$, and the maximum number of MTs for which the CZF data precoder is advantageous compared to the SZF data precoder, $K_{\text{CZF}>\text{SZF}}$, decrease with increasing AN leakage, $\tilde{Q}$, and increasing number of cells, $M$, but increase with the amount of resources dedicated to channel estimation, $p_\tau\tau$ (via $\theta$), and consequently with the channel estimation quality. However, while $K_{\text{SZF}>\text{MF}}$ decreases with increasing inter-cell interference factor, $\rho$ (via $a$), $K_{\text{CZF}>\text{SZF}}$ increases.

*2) Pilot Energy:* From (35), we find that for $\gamma_{nk}^{\text{SZF}} > \gamma_{nk}^{\text{MF}}$ and $\gamma_{nk}^{\text{CZF}} > \gamma_{nk}^{\text{SZF}}$ to hold, pilot energy $p_\tau\tau$ has to fulfill

$$p_\tau\tau > (p_\tau\tau)_{\text{SZF}>\text{MF}} = \frac{1}{\frac{\phi(1-\beta)/\beta+1}{a+1/P_T} - a}$$
$$p_\tau\tau > (p_\tau\tau)_{\text{CZF}>\text{SZF}} = \frac{1}{\frac{\rho\phi(1-\beta)/\beta+1}{a+1/P_T} - a}, \tag{37}$$

where we have assumed that SNS AN precoding is adopted, i.e., $\tilde{Q} = a - \theta$, to arrive at insightful expressions. Similar results can be obtained for other AN precoders. From (37), we observe that MF, SZF, and CZF data precoding are preferable if $0 < p_\tau\tau < (p_\tau\tau)_{\text{SZF}>\text{MF}}$, $(p_\tau\tau)_{\text{SZF}>\text{MF}} \le p_\tau\tau < (p_\tau\tau)_{\text{CZF}>\text{SZF}}$, and $p_\tau\tau \ge (p_\tau\tau)_{\text{CZF}>\text{SZF}}$, respectively. In general, the more MTs are in the system (i.e., the larger $\beta$), the larger the pilot energy has to be to make SZF and CZF data precoding beneficial. In fact, from (37) we observe that if $\beta$ exceeds $\beta_{\text{MF}} = \phi/[a^2 + a/P_T + \phi - 1]$, MF data precoding is always preferable regardless of the value of $p_\tau\tau$. Similarly, if $\beta$ exceeds $\beta_{\text{SZF}} = \phi\rho/[a^2 + a/P_T + \phi\rho - 1]$, SZF data precoding is always preferable compared to CZF data precoding regardless of the value of $p_\tau\tau$.

*3) Comparison of SNS, CNS, and MF AN Precoding:* In this subsection, we analyze the impact of the AN precoders on the secrecy rate. AN precoders affect the ergodic capacity of the eavesdropper via $L$ and the achievable rate of the MT via the leakage, $\tilde{Q}$. Since the upper bound on the ergodic secrecy rate of the eavesdropper in (7) is a decreasing function in $L$, we have

$$C_{nk}^{\text{eve}}|_{\text{random}} \leq C_{nk}^{\text{eve}}|_{\text{SNS}} \leq C_{nk}^{\text{eve}}|_{\text{CNS}}. \tag{38}$$

On the other hand, from Table II, we observe $\tilde{Q}_{\text{random}} \geq \tilde{Q}_{\text{SNS}} \geq \tilde{Q}_{\text{CNS}}$. Since according to Table I the SINRs for all data precoders are decreasing functions of $\tilde{Q}$, for a given data precoder, we obtain for the lower bound on the ergodic rate of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell

$$R_{nk}|_{\text{random}} \leq R_{nk}|_{\text{SNS}} \leq R_{nk}|_{\text{CNS}}. \tag{39}$$

Considering (38), (39), and the expression for the ergodic secrecy rate, $R_{nk}^{\text{sec}} = [R_{nk} - C_{nk}^{\text{eve}}]^+$, it is not a priori clear which AN precoder has the best performance. In fact, our numerical results in Section VI confirm that it depends on the system parameters (e.g. $\alpha$, $\beta$, $M$, $p_\tau \tau$, and $\rho$) which AN precoder is preferable.

*4) Ergodic Secrecy Rate Analysis:* In this subsection, we provide closed-form results for the ergodic secrecy rate for SZF, CZF, and MF data precoding for the simplified path-loss model in (34). Thereby, the simplified path-loss model is extended also to the eavesdropper, i.e., $\beta_{nE} = 1$ and $\beta_{mE} = \rho$, $m \neq n$, is assumed.

Combining (4), (7), and the results in Table I, we obtain the lower bounds for the ergodic secrecy rate of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell, given in (40) shown at the bottom of the page, where $\chi = \frac{a\beta}{\alpha} - \frac{\beta c N_T}{aL}$, and $\tilde{Q}$ and $L$ are given in Table II for the considered AN precoders. Eq. (40) is easy to evaluate and reveals how the ergodic secrecy rate of the three considered data precoders depends on the various system parameters. To gain more insight, we determine the maximum value of $\alpha$ which admits a non-zero secrecy rate. This value is denoted by $\alpha_s$ in the following, and can be shown to be a decreasing function of $\phi$ for all conidered data precoders. Hence, we find $\alpha_s$ by setting $R_{nk}^{\text{sec}} = 0$ in (40) and letting $\phi \to 0$. This leads to

$$\alpha_s = \begin{cases} \frac{a^2\theta}{\tilde{Q}a + c\theta N_T/L + a/P_T} & \text{for MF,} \\ \frac{(1-\beta)a^2\theta}{\tilde{Q}a + c\theta(1-\beta)N_T/L + a/P_T} & \text{for SZF,} \\ \frac{(1-M\beta)a^2\theta}{\tilde{Q}a + c\theta(1-M\beta)N_T/L + a/P_T} & \text{for CZF.} \end{cases} \tag{41}$$

Eq. (41) reveals that for a given AN precoder, independent of the system parameters, the MF data precoder can always tolerate a larger number of eavesdropper antennas than the SZF data precoder, which in turn can always tolerate a larger number of eavesdropper antennas than the CZF data precoder. This can be explained by the fact that the high AN transmit power required to combat a large number of eavesdropper antennas drives the receiver of the desired MT into the noise-limited regime, where the MF data precoder has a superior performance compared to the S/CZF data precoders. On the other hand, since $\alpha_s$ depends on both $\tilde{Q}$ and $L$, it is not a priori clear which AN precoder can tolerate the largest number of eavesdropper antennas. For a lightly loaded network with small $\beta$ and small $M$, according to Table II, we have $L \approx N_T$ for all three AN precoders. Hence, in this case, we expect the CNS AN precoder to outperform the SNS and random AN precoders as it achieves a smaller $\tilde{Q}$. On the other hand, for a heavily loaded network with large $\beta$ and $M$, the value of $\alpha_s$ of the CNS AN precoder is compromised by its small value of $L$ and SNS and even random AN precoders are expected to achieve a larger $\alpha_s$.

## VI. Performance Evaluation

In this section, we evaluate the performance of the considered secure multi-cell massive MIMO system. We consider cellular systems with $M = 2$ and $M = 7$ hexagonal cells, respectively, and to gain insight for system design, we adopt the simplified path-loss model introduced in Section V, i.e., the severeness of the inter-cell interference is only characterized by the parameter $\rho \in (0, 1]$. The pilot sequence length is $\tau = K$. The simulation results for the ergodic secrecy rate of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell are based on (4), (6), and the expression for the ergodic rate of the MT [18, Eq. (8)] and are averaged over 5,000 random channel realizations. Note that, in this paper, we consider the ergodic secrecy rate of a certain MT, i.e., the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell. The cell sum secrecy rate can be obtained by multiplying the secrecy rate of the $k^{\text{th}}$ MT by the number of MTs, $K$, as for the considered channel model, all MTs in the $n^{\text{th}}$ cell achieve the same secrecy rate. The values of all relevant system parameters are provided in the captions of the figures. To enable a fair comparison, throughout this section, we adopted the selfish SNS AN precoder when we compare different data precoders and the selfish ZF data precoder when we compare different AN precoders.

### A. Ergodic Capacity of the Eavesdropper for Conventional AN Precoders

In Fig. 1, we show the ergodic capacity of the eavesdropper for the considered conventional AN precoders. First, we note that the upper bound in (7) is very tight since the number of BS antennas is large ($N_T = 200$) and $\alpha < a^2 L/(cN_T)$

$$R_{nk}^{\text{sec}} \geq \begin{cases} \left[\log_2\left(\frac{(\tilde{Q}+1/P_T)\beta + (a-\tilde{Q})\beta\phi + c\theta\phi}{(\tilde{Q}+1/P_T)\beta + (a-\tilde{Q})\beta\phi + (c-1)\theta\phi} \cdot \frac{-\chi\phi+\chi}{(1-\chi)\phi+\chi}\right)\right]^+ & \text{for MF,} \\ \left[\log_2\left(\frac{(\tilde{Q}+1/P_T)\beta + (a-\theta-\tilde{Q})\beta\phi + c\theta(1-\beta)\phi}{(\tilde{Q}+1/P_T)\beta + (a-\theta-\tilde{Q})\beta\phi + (c-1)\theta(1-\beta)\phi} \cdot \frac{-\chi\phi+\chi}{(1-\chi)\phi+\chi}\right)\right]^+ & \text{for SZF,} \\ \left[\log_2\left(\frac{(\tilde{Q}+1/P_T)\beta + (a-a\theta-\tilde{Q})\beta\phi + c\theta(1-M\beta)\phi}{(\tilde{Q}+1/P_T)\beta + (a-a\theta-\tilde{Q})\beta\phi + (c-1)\theta(1-M\beta)\phi} \cdot \frac{-\chi\phi+\chi}{(1-\chi)\phi+\chi}\right)\right]^+ & \text{for CZF.} \end{cases} \tag{40}$$

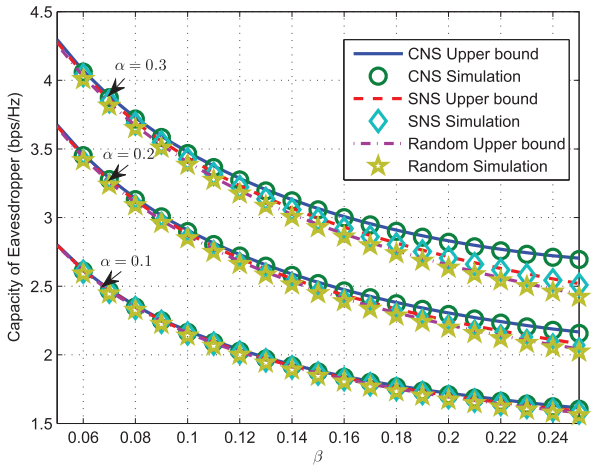



Fig. 1. Ergodic capacity of the eavesdropper vs. the normalized number of MTs in the cell, $\beta$, for a system with $N_T = 200$, $\phi = 0.75$, $P_T = 10$ dB, $\rho = 0.3$, and $M = 2$.

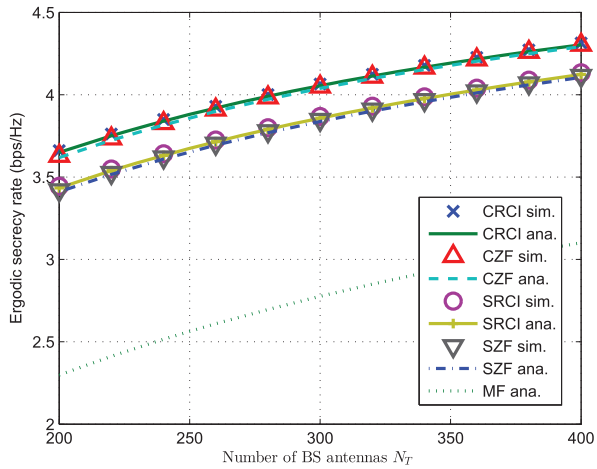



Fig. 2. Analytical and simulation results for the ergodic secrecy rate vs. the number of BS antennas, $N_T$, for a lightly loaded network with $\phi = 0.75$, $P_T = 10$ dB, $p_\tau = P_T/K$, $\alpha = 0.1$, $K = 10$, $\rho = 0.1$, and $M = 2$.

holds for all considered AN precoders and all consider values of $\alpha$ and $\beta$. Furthermore, as $\beta$ increases, the ergodic capacity of all AN precoders decreases since the power allocated to the information-carrying signal of the user that the eavesdropper tries to intercept decreases with increasing $\beta$ as the total power allocated to the information-carrying signals of all users is fixed. As expected, the eavesdropper's capacity benefits from larger values of $\alpha$. Furthermore, as predicted in (38), because of their different values of $L$, the CNS AN precoder yields the largest eavesdropper capacity, while the random AN precoder yields the lowest. The performance differences between the different AN precoders diminish for small values of $\alpha$ and $\beta$ as the dependence of the eavesdropper capacity on $L$ becomes negligible for small $\alpha$, cf. (7), and $L \approx N_T$ holds for all precoders for small $\beta$, cf. Table II.

### *B. Ergodic Secrecy Rate for Conventional Linear Data Precoders*

In Figs. 2 and 3, we show the ergodic secrecy rates of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell vs. the number of BS antennas for the MF, SZF, CZF, SRCI, and CRCI data precoders for a lightly loaded and a dense network, respectively, and a fixed power allocation factor of $\phi = 0.75$. In both figures, the analytical results were obtained from (4), (6), and (12) for the SRCI data precoder, (16) for the CRCI data precoder, and (40) for the MF, SZF, and CZF data precoders. For all considered precoders, the analytical results provide a tight lower bound for the ergodic secrecy rates obtained by simulations. Furthermore, as expected, the RCI data precoders outperform the ZF data precoders for both the selfish and the collaborative strategies, but the performance gap diminishes with increasing number of BS antennas.

For the lightly loaded network in Fig. 2, we assume $M = 2$ cells, $K = 10$ users, and a small inter-cell interference factor of $\rho = 0.1$. For this scenario, the collaborative designs outperform the selfish designs and C/SZF precoding yields a large performance gain compared to MF precoding. This is expected from our analysis in Section V-A as for the parameters valid for Fig. 2, we obtain from (36), $K_{\text{SZF}>\text{MF}} \approx 250$ and $K_{\text{CZF}>\text{SZF}} \approx 60$ for $N_T = 400$. Intuitively, as the network is only lightly loaded, the collaborative data precoder can efficiently reduce interference to the other cell despite the pilot contamination.

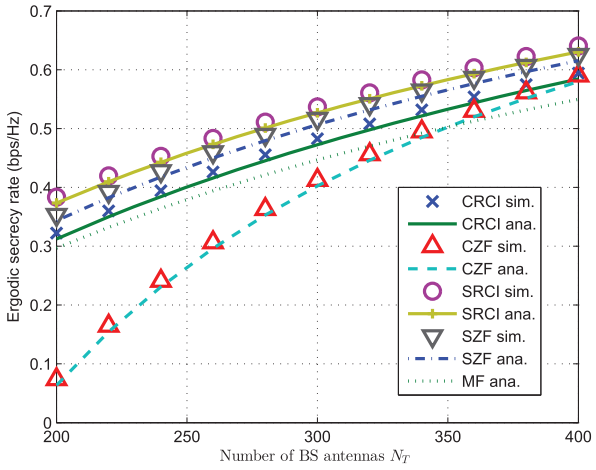



Fig. 3. Analytical and simulation results for the ergodic secrecy rate vs. the number of BS antennas, $N_T$, for a dense network with $\phi = 0.75$, $P_T = 10$ dB, $p_\tau = P_T/K$, $\alpha = 0.1$, $K = 20$, $\rho = 0.3$, and $M = 7$.

For the dense network in Fig. 3, we assume $M = 7$ cells, $K = 20$ users, and a larger inter-cell interference factor of $\rho = 0.3$. In this case, for the considered range of $N_T$, the collaborative precoder designs are not able to suppress inter-cell interference and AN leakage to other cells sufficiently well to outperform the selfish precoder designs. In fact, for $N_T = 400$, we obtain from (36) $K_{\text{CZF}>\text{SZF}} \approx 16$, i.e., our analytical results suggest that the SZF precoder outperforms the CZF precoder for $K = 20$ which is confirmed by Fig. 3. Nevertheless, for $N_T > 400$, the ergodic secrecy rate for the CZF data precoder will eventually surpass that for the SZF data precoder.

### *C. Optimal Power Allocation*

In this subsection, we investigate the dependence of the ergodic secrecy rate on the power allocation factor $\phi$ and study

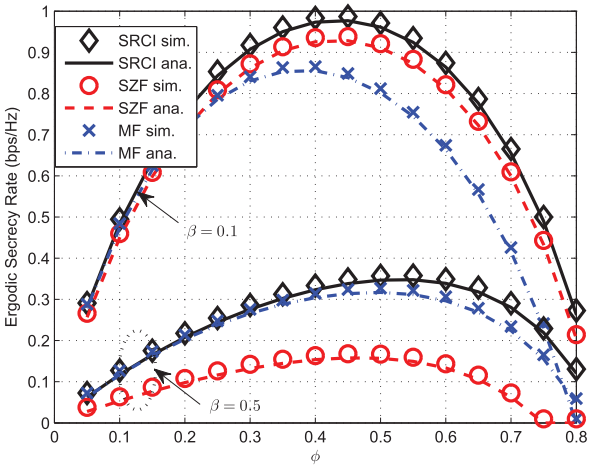



Fig. 4. Ergodic secrecy rate vs. $\phi$ for different selfish data precoders for a network with $P_T = 10$ dB, $N_T = 100$, $p_\tau = P_T/K$, $\alpha = 0.1$, $\rho = 0.1$, and $M = 7$.

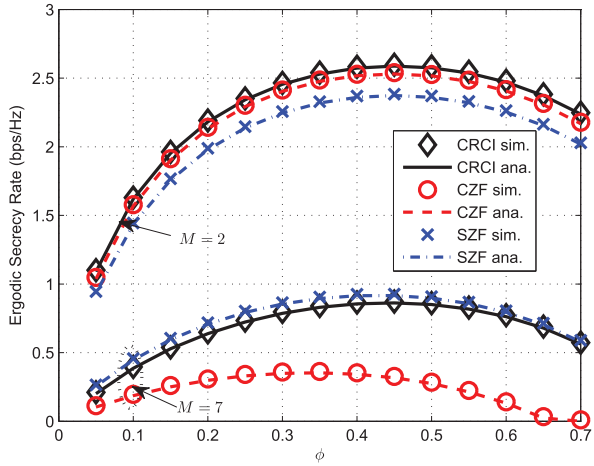



Fig. 5. Ergodic secrecy rate vs. $\phi$ for different data precoders for a network with $P_T = 10$ dB, $N_T = 100$, $p_\tau = P_T/K$, $\alpha = 0.1$, $\beta = 0.1$, and $\rho = 0.1$.

the impact of system parameters such as $\beta$, $M$, and $\rho$ on the optimal $\phi$ that maximizes the ergodic secrecy rate. The results in this subsection were generated based on the analytical expressions in (4), (6), and (12) for the SRCI data precoder, (16) for the CRCI data precoder, and (40) for the MF, SZF, and CZF data precoders.

Fig. 4 depicts the ergodic secrecy rate of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell for the selfish data precoders SRCI, SZF, and MF as a function of the power allocation factor $\phi$. All curves are concave and have a single maximum. For $\phi = 0$ only AN is transmitted, hence $R_{nk}^{\text{sec}} = 0$ results since no data can be transmitted. For $\phi = 1$, no AN is transmitted, hence $R_{nk}^{\text{sec}} = 0$ results since the capacity of the eavesdropper becomes unbounded (recall that we make the worst-case assumption that the eavesdropper can receive noise-free). For $0 < \phi < 1$, a positive secrecy rate may result depending on the system parameters and the precoding schemes. Since we keep the total transmit power fixed, the transmit power per MT decreases with increasing $\beta$. To compensate for this effect, the portion of the total transmit power allocated to data transmission should increase. This is confirmed by Fig. 4 where the optimal value of $\phi$ for $\beta = 0.5$ is larger than that for $\beta = 0.1$. Furthermore, for a given $\beta$, the optimal $\phi$ is the larger, the better the performance of the adopted data precoder is, i.e., for a more effective data precoder, transmitting the data signal with higher power is more beneficial, whereas for a less effective data precoder impairing the eavesdropper with a higher AN power is more beneficial.

In Fig. 5, we show the ergodic secrecy rate vs. $\phi$ for the CRCI, CZF, and SZF precoders. Similar to our observations in Fig. 4, for given system parameters, the optimal $\phi$ tends to be larger for more effective precoders that achieve a better performance. For the system with $M = 7$, this can be observed by comparing the optimal $\phi$ for the SZF and CZF precoders. Furthermore, while for the smaller system with $M = 2$ cells collaborative precoding is always preferable, for $M = 7$, SZF precoding outperforms CZF and CRCI precoding for all considered values of $\phi$, as the collaborative designs are not able to effectively suppress the interference and AN leakage to

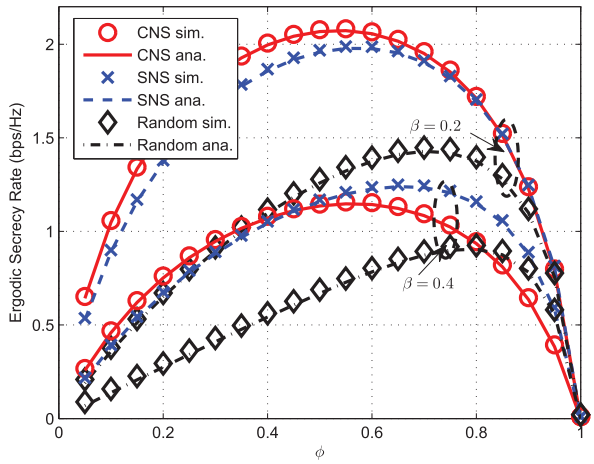



Fig. 6. Ergodic secrecy rate vs. $\phi$ for different AN precoders for a network with $P_T = 10$ dB, $N_T = 100$, $p_\tau = P_T/K$, $M = 2$, $\rho = 0.1$, and $\alpha = 0.1$.

the $(M-1)K = 60$ users in the other cells with the available $N_T = 100$ antennas. In particular, from (36), we obtain $K_{\text{CZF}>\text{SZF}} \leq 18$ for $M = 2$ and $K_{\text{CZF}>\text{SZF}} \leq 5$ for $M = 7$, which confirms the results shown in Fig. 5.

Fig. 6 depicts the ergodic secrecy rate vs. $\phi$ for the considered conventional AN precoder structures. We consider a lightly loaded network with $\beta = 0.2$ and a moderately loaded network with $\beta = 0.4$. For $\beta = 0.2$, the CNS AN precoder outperforms the SNS AN precoder since, in this case, for the CNS AN precoder, the negative impact of having (slightly) fewer dimensions available for degrading the eavesdropper's channel (smaller value of $L$) is outweighed by the positive impact of causing less AN leakage (smaller value of $\tilde{Q}$). On the other hand, for $\beta = 0.4$, the CNS AN precoder has a substantially smaller $L$ than the SNS precoder which cannot be compensated by its larger $\tilde{Q}$. Despite having the largest value of $L$, the random AN precoder has the worst performance for both considered cases because of its large AN leakage.

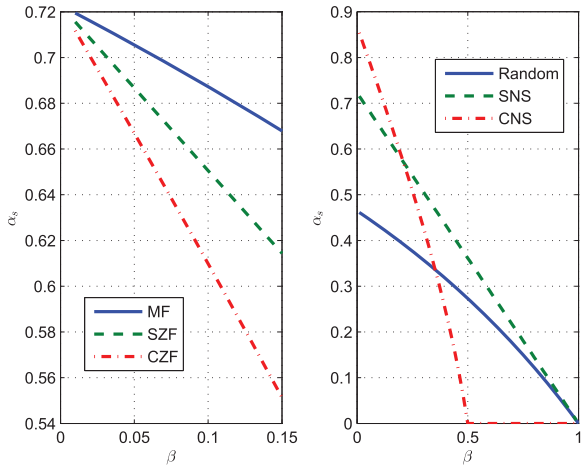



Fig. 7. $\alpha_s$ vs. $\beta$ for different data and AN precoders for a network with $P_T = 10$ dB, $N_T = 100$, $p_\tau = P_T/K$, $\rho = 0.3$, and $M = 2$.

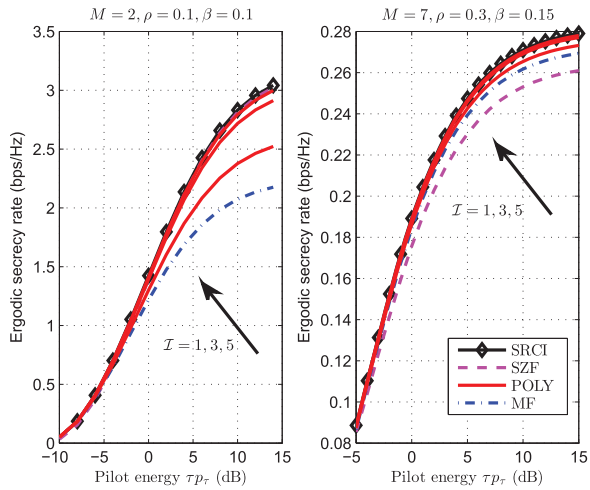



Fig. 8. Ergodic secrecy rate for POLY and conventional selfish data precoders for a network employing the optimal $\phi$, $P_T = 10$ dB, $N_T = 200$, and $\alpha = 0.1$.

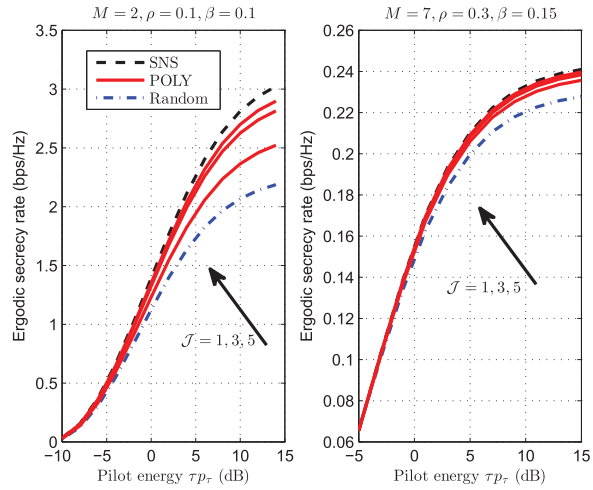



Fig. 9. Ergodic secrecy rate for POLY, SNS, and random AN precoders for a network employing the optimal $\phi$, $P_T = 10$ dB, $N_T = 200$, and $\alpha = 0.1$.

### D. Conditions for Non-Zero Secrecy Rate

In Section V-C, we showed that a positive ergodic secrecy rate is possible only if $\alpha < \alpha_s$. In Fig. 7, using (41), we plot $\alpha_s$ as a function of $\beta$. In the left hand side subfigure, we compare MF, SZF, and CZF data precoding for SNS AN precoding, and in the right hand side subfigure, we compare random, SNS, and CNS AN precoding for SZF data precoding. The comparison of the data precoders reveals that although SZF and CZF entail a much higher complexity, MF precoding achieves a larger $\alpha_s$. Therefore, if the eavesdropper has a large number of antennas and small ergodic secrecy rates are targeted, simple MF precoding is always preferable. On the other hand, whether SNS or CNS AN precoder is preferable depends on the system load. For small values of $\beta$, CNS AN precoding can tolerate more eavesdropper antennas, whereas for large values of $\beta$, SNS AN precoding is preferable. Random AN precoding is outperformed by SNS AN precoding for any value of $\beta$. A closer examination of (41) reveals that this is always true if S/CZF data precoders are employed. However, for the MF data precoder, there are parameter combination for which random AN precoding outperforms SNS and CNS AN precoding.

### E. Low-Complexity POLY Data and AN Precoders

In this subsection, we evaluate the ergodic secrecy rates of the proposed low-complexity POLY data and AN precoders. To this end, we consider again a lightly loaded network with little inter-cell interference ($M = 2$, $\beta = 0.1$, $\rho = 0.1$) and a dense network with more inter-cell interference ($M = 7$, $\beta = 0.15$, $\rho = 0.3$). All results shown in this section were obtained by simulation. For each simulation point, the optimal value of $\phi$ was found numerically and applied. In Figs. 8 and 9, we show the ergodic secrecy rate of the $k^{\text{th}}$ MT in the $n^{\text{th}}$ cell as a function of the pilot energy, $\tau p_\tau$. As expected, for all considered schemes, the ergodic secrecy rate is monotonically increasing in the pilot energy since more accurate channel estimates improve the performance.

In Fig. 8, we depict the ergodic secrecy rates for the proposed POLY data precoder for different values of $\mathcal{I}$ and compare them to those of conventional selfish data precoders. For the sake of comparison, all data precoders are combined with the SNS AN precoder. As the number of terms of the polynomial $\mathcal{I}$ increase, the performance of the POLY data precoder quickly improves and approaches that of the SRCI data precoder. The convergence is faster for the dense network considered in the right hand side subfigure, where the performance difference between all precoders is smaller in general since interference cannot be as efficiently avoided as for the lightly loaded network.

In Fig. 9, we show the ergodic secrecy rates for the proposed POLY AN precoder for different values of $\mathcal{J}$ and compare them to those of the random and SNS AN precoders. For the sake of comparison, all AN precoders are combined with SZF data precoding. The POLY AN precoder quickly approaches the performance of the SNS AN precoder as the polynomial order $\mathcal{J}$ increases. Similar to the POLY data precoders, the convergence is faster for the dense network where the performance differences between different AN precoders are also smaller. For the denser network, even the random AN precoder is a viable option and suffers only from a small loss in performance compared to the SNS AN precoder.

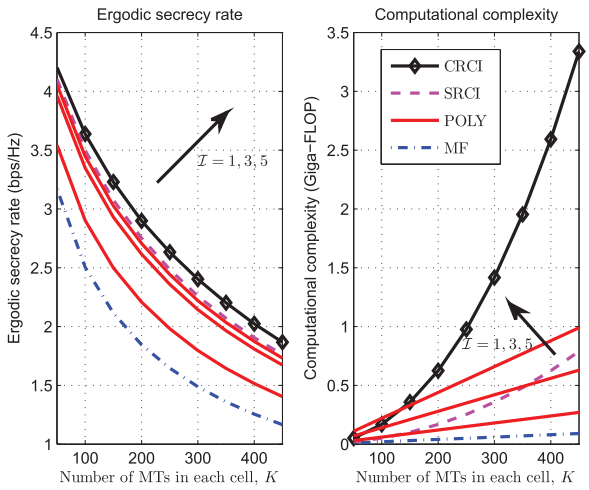



Fig. 10. Ergodic secrecy rate (left hand side) and computational complexity (right hand side) of various linear data precoders for a network employing $P_T = 10$ dB, $N_T = 1000$, $p_\tau = P_T/K$, $M = 2$, $\rho = 0.1$, $T - \tau = 100$, and an SNS AN precoder.

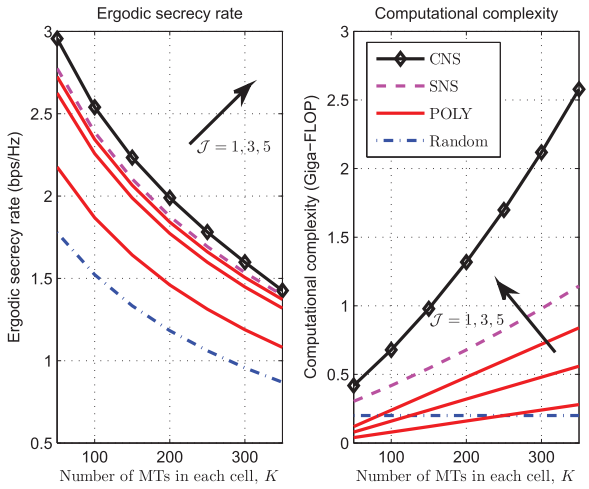



Fig. 11. Ergodic secrecy rate (left hand side) and computational complexity (right hand side) of various linear AN precoders for a network employing $P_T = 10$ dB, $N_T = 1000$, $p_\tau = P_T/K$, $M = 2$, $\rho = 0.1$, $T - \tau = 100$, and an SZF data precoder.

### *F. Complexity-Performance Tradeoff*

In this subsection, we investigate the tradeoff between the ergodic secrecy rate performance and the computational complexity of the proposed data and AN precoders in Figs. 10 and 11, respectively. In particular, Figs. 10 and 11 depict the ergodic secrecy rate on the left hand side and the computational complexity (in Giga FLOP) on the right hand side, both as a function of the numbers of users in a cell. For the considered setting, the performance gains of collaborative data and AN precoding compared to selfish strategies are moderate, but the associated increase in complexity is substantial, especially for large $K$.

Fig. 10 illustrates that for the considered setting a POLY data precoder with $\mathcal{I} = 1$ achieves a better performance than the MF precoder but has substantially lower complexity than the SRCI precoder. For large $\mathcal{I}$, the POLY data precoder has a lower complexity than the SRCI precoder for large $K$. However, even for small $K$, the POLY precoder may be preferable as it does not incur the stability issues that may arise in the implementation of the large-scale matrix inversions required for the SRCI precoder.

Fig. 11 shows that for the considered setting the proposed POLY AN precoder with $\mathcal{J} = 1$ outperforms the random AN precoder. The POLY AN precoder with $\mathcal{J} = 5$ achieves almost the same performance as the SNS AN precoder but with a substantially lower complexity. We further observe that for small $K$, because of its efficient implementation via Horner's scheme, cf. (33), the proposed POLY AN precoder requires an even lower complexity than the random AN precoder.

## VII. Conclusion

In this paper, we considered downlink multi-cell massive MIMO systems employing linear data and AN precoding for physical layer security provisioning. We analyzed and compared the achievable ergodic secrecy rate of various conventional data and AN precoders in the presence of pilot contamination. To this end, we also optimized the regularization constants of the selfish and collaborative RCI precoders in the presence of AN and multi-cell interference. In addition, we derived linear POLY data and AN precoders which offer a good compromise between complexity and performance in massive MIMO systems. Interesting findings of this paper include: 1) Collaborative data precoders outperform selfish designs only in lightly loaded systems where a sufficient number of degrees of freedom for suppressing inter-cell interference and sufficient resources for training are available. 2) Similarly, CNS AN precoding is preferable over SNS AN precoding in lightly loaded systems as it causes less AN leakage to the information-carrying signal, whereas in more heavily loaded systems, CNS AN precoding does not have sufficient degrees of freedom for effectively degrading the eavesdropper channel and SNS AN precoding is preferable. 3) For a large number of eavesdropper antennas, where only small positive secrecy rates are achievable, MF data precoding is always preferable compared to SZF and CZF data precoding. 4) The proposed POLY data and AN precoders approach the performances of the SRCI data and SNS AN precoders with only a few terms in the respective matrix polynomials and are attractive options for practical implementation.

## Appendix

### *A. Proof of Proposition 1*

Considering (3) and (9), the effective signal power, i.e., the numerator in (5), can be expressed as [26]

$$\begin{aligned}\mathbb{E}^2[\mathbf{h}_{nn}^k \mathbf{f}_{nk}] &= \gamma_1^2 \mathbb{E}^2[\mathbf{h}_{nn}^k \mathbf{L}_{nn} (\hat{\mathbf{h}}_{nn}^k)^H] \\ &= \gamma_1^2 \mathbb{E}^2 \left[ \frac{\mathbf{h}_{nn}^k \mathbf{L}_{n,k} (\hat{\mathbf{h}}_{nn}^k)^H}{1 + \hat{\mathbf{h}}_{nn}^k \mathbf{L}_{n,k} (\hat{\mathbf{h}}_{nn}^k)^H} \right] \\ &= \frac{\gamma_1^2 (X_{nk} + A_{nk})^2}{(1 + X_{nk})^2}, \end{aligned} \tag{42}$$

where $\mathbf{L}_{n,k} = (\hat{\mathbf{H}}_{nn}\hat{\mathbf{H}}_{nn}^H - (\hat{\mathbf{h}}_{nn}^k)^H\hat{\mathbf{h}}_{nn}^k + \kappa_1\mathbf{I}_{N_T})^{-1}$, $X_{nk} = \mathbb{E}[\hat{\mathbf{h}}_{nn}^k\mathbf{L}_{n,k}(\hat{\mathbf{h}}_{nn}^k)^H]$, and $A_{nk} = \mathbb{E}[\tilde{\mathbf{h}}_{nn}^k\mathbf{L}_{n,k}(\hat{\mathbf{h}}_{nn}^k)^H]$. On the other hand, the intra-cell interference term in the denominator of (5) can be expressed as

$$\mathbb{E}\left[\sum_{l\neq k}|\mathbf{h}_{nn}^k\mathbf{f}_{nl}|^2\right] = \gamma_1^2\mathbb{E}\left[\frac{\mathbf{h}_{nn}^k\mathbf{L}_{n,k}\hat{\mathbf{H}}_{n,k}^H\hat{\mathbf{H}}_{n,k}\mathbf{L}_{n,k}(\mathbf{h}_{nn}^k)^H}{\left(1+\hat{\mathbf{h}}_{nn}^k\mathbf{L}_{n,k}(\hat{\mathbf{h}}_{nn}^k)^H\right)^2}\right] = \frac{\gamma_1^2(Y_{nk}+B_{nk})}{(1+X_{nk})^2}, \tag{43}$$

where $\hat{\mathbf{H}}_{n,k}$ is equal to $\hat{\mathbf{H}}_{nn}$ with the $k^{\text{th}}$ row removed, and $Y_{nk} = \mathbb{E}[\hat{\mathbf{h}}_{nn}^k\mathbf{L}_{n,k}\hat{\mathbf{H}}_{n,k}^H\hat{\mathbf{H}}_{n,k}\mathbf{L}_{n,k}(\hat{\mathbf{h}}_{nn}^k)^H]$ and $B_{nk} = \mathbb{E}[\tilde{\mathbf{h}}_{nm}^k\mathbf{L}_{n,k}\hat{\mathbf{H}}_{n,k}^H\hat{\mathbf{H}}_{n,k}\mathbf{L}_{n,k}(\tilde{\mathbf{h}}_{nn}^k)^H]$.

Due to pilot contamination, the data precoding matrix of the $m^{\text{th}}$ BS is a function of the channel vectors between the $m^{\text{th}}$ BS and the $k^{\text{th}}$ MTs in all cells. Hence, the inter-cell interference from the BSs in adjacent cells is obtained as

$$\mathbb{E}[|\mathbf{h}_{mn}^k\mathbf{f}_{mk}|^2] = \frac{\gamma_1^2(X_{nk}+A_{nk})^2}{(1+X_{nk})^2} + \frac{1+p_\tau\tau\sum_{l\neq m}^M\beta_{ml}^k}{1+p_\tau\tau\sum_{l=1}^M\beta_{ml}^k}. \tag{44}$$

Meanwhile, by exploiting (42), (44), and the definition of the variance, i.e., $\text{var}[x] = \mathbb{E}[x^2] - \mathbb{E}^2[x]$, we obtain for the first term of the denominator of (5), $\text{var}[\mathbf{h}_{nn}^k\mathbf{f}_{nk}] = \frac{1+p_\tau\tau\sum_{m\neq n}^M\beta_{nm}^k}{1+p_\tau\tau\sum_{m=1}^M\beta_{nm}^k}$. According to [26, Eq. (16)] and [36, Theorem 7], for $N_T\to\infty$ and constant $\beta$, $X_{nk}$ converges to $\mathcal{G}(\beta,\kappa_1)$ defined in (11) and $A_{nk}\to 0$. Similarly, $Y_{nk}$ and $B_{nk}$ approach

$$Y_{nk} \overset{N_T\to\infty}{=} \mathcal{G}(\beta,\kappa_1) + \kappa_1\frac{\partial}{\partial\kappa_1}\mathcal{G}(\beta,\kappa_1) \tag{45}$$

and

$$B_{nk} \overset{N_T\to\infty}{=} \frac{\vartheta_{nk}}{\theta_{nk}}(1+\mathcal{G}(\beta,\kappa_1))^2\left(\mathcal{G}(\beta,\kappa_1) + \kappa_1\frac{\partial}{\partial\kappa_1}\mathcal{G}(\beta,\kappa_1)\right), \tag{46}$$

respectively, where $\frac{\partial}{\partial\kappa_1}\mathcal{G}(\beta,\kappa_1) = -\frac{\mathcal{G}(\beta,\kappa_1)(1+\mathcal{G}(\beta,\kappa_1))^2}{\beta+\kappa_1(1+\mathcal{G}(\beta,\kappa_1))^2}$.

Moreover, the inter-cell interference from other MTs (i.e., not the $k^{\text{th}}$ MTs) is calculated as

$$\mathbb{E}\left[\mathbf{h}_{mn}^k\mathbf{F}_{m,k}\mathbf{F}_{m,k}^H(\mathbf{h}_{mn}^k)^H\right] = \mathbb{E}\left[\text{ tr}\left\{\mathbf{F}_{m,k}\mathbf{F}_{m,k}^H\right\}\right] = K-1, \tag{47}$$

where $\mathbf{F}_{m,k}$ is equal to $\mathbf{F}_m$ with the $k^{\text{th}}$ column removed. The first equality in (47) is due to the fact that the precoding matrix for the other MTs (i.e., not the $k^{\text{th}}$ MTs) in adjacent cells are independent of $\mathbf{h}_{mn}^k$ and [30, Lemma 11], while the second equality holds for $N_T\to\infty$.

On the other hand, the constant scaling factor $\gamma_1$ for SRCI precoding is given by [26, Eq. (22)]

$$\gamma_1^2 = \frac{1}{\mathcal{G}(\beta,\kappa_1) + \kappa_1\frac{\partial}{\partial\kappa_1}\mathcal{G}(\beta,\kappa_1)}. \tag{48}$$

Hence, employing (42)–(48) in (5), the received SINR in (10) is obtained, which completes the proof of Proposition 1.

### *B. Proof of Theorem 1*

The objective function in (20) can be rewritten in (49) shown at the bottom of the page, where we exploited $\mathbb{E}[\mathbf{s}_n\mathbf{s}_n^H] = \mathbf{I}_K$, the definition of $P_{\text{AN}}$ given in Theorem 1, the definition of $\mathbf{F}_n$ in (18), the definition $\frac{1}{\sqrt{N_T}}\mathbf{H}_{nn} = \hat{\bar{\mathbf{H}}}_{nn} + \tilde{\bar{\mathbf{H}}}_{nn}$, and $\tilde{\bar{\mathbf{H}}}_{nn} = \frac{1}{\sqrt{N_T}}\tilde{\mathbf{H}}_{nn}$.

In the following, we simplify the right hand side (RHS) of (49) term by term. To this end, we denote the first three terms on the RHS of (49) by $t_1$, $t_2$, and $t_3$, respectively. Using a result from free probability theory [35], the first term converges to [29, Theorem 1]

$$t_1 = \varsigma^2 p\,\text{Tr}\{\mathbf{D}_{nn}\}\,\mathbb{E}\left[\text{ Tr}\left\{\left(\sum_{i=0}^{\mathcal{I}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^{i+1}\right)^2\right\}\right], \tag{50}$$

as matrix $\mathbf{D}_{nn}$ is free from $\sum_{i=0}^{\mathcal{I}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^{i+1}$. Similarly, the third term converges to

$$t_3 = -2\varsigma\sqrt{p}\,\text{Tr}\left\{\mathbf{D}_{nn}^{1/2}\right\}\mathbb{E}\left[\text{ Tr}\left\{\sum_{i=0}^{\mathcal{I}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^{i+1}\right\}\right]. \tag{51}$$

Furthermore, the second term can be rewritten in (52) shown at the bottom of the next page, where (a) follows again from [29, Theorem 1] and (b) results from $\mathbb{E}[\text{ Tr}\{\tilde{\bar{\mathbf{H}}}_{nn}^H\mathbf{D}_{nn}\tilde{\bar{\mathbf{H}}}_{nn}\}] = \text{Tr}\{\mathbf{D}_{nn}\mathbf{\Delta}_n\}$, where $\mathbf{\Delta}_n$ is defined in Theorem 1, (18), and the constraint in (20).

$$\begin{aligned}\text{mse}_n = {}& \varsigma^2 p\mathbb{E}\left[\text{ Tr}\left\{\sum_{i=0}^{\mathcal{I}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^{i+1}\mathbf{D}_{nn}\sum_{i=0}^{\mathcal{I}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^{i+1}\right\}\right] \\ & + \varsigma^2 p\mathbb{E}\left[\text{ Tr}\left\{\sum_{i=0}^{\mathcal{I}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^{i}\hat{\bar{\mathbf{H}}}_{nn}\tilde{\bar{\mathbf{H}}}_{nn}^H\mathbf{D}_{nn}\tilde{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\sum_{i=0}^{\mathcal{I}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^{i}\right\}\right] \\ & - 2\varsigma\sqrt{p}\mathbb{E}\left[\text{ Tr}\left\{\mathbf{D}_{nn}^{1/2}\sum_{i=0}^{\mathcal{I}}\mu_i\left(\hat{\bar{\mathbf{H}}}_{nn}\hat{\bar{\mathbf{H}}}_{nn}^H\right)^{i+1}\right\}\right] + 1 + \varsigma^2 P_{\text{AN}} + \varsigma^2\,\text{Tr}\{\mathbf{\Sigma}_n\}.\end{aligned} \tag{49}$$

Exploiting (50)–(52) and the eigen-decomposition of matrix $\hat{\tilde{\mathbf{H}}}_{nn}\hat{\tilde{\mathbf{H}}}_{nn}^H = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^H$, where diagonal matrix $\mathbf{\Lambda} = \text{diag}\,(\lambda_1, \ldots, \lambda_K)$ contains all eigenvalues and unitary matrix $\mathbf{T}$ contains the corresponding eigenvectors, the asymptotic average MSE becomes

$$\begin{aligned}\text{mse}_n = \mathbb{E}\Bigg[ & \varsigma^2 p\ \text{Tr}\{\mathbf{D}_{nn}\}\ \text{Tr}\left\{\mathbf{\Lambda}^2\left(\sum_{i=0}^{\mathcal{J}}\mu_i\mathbf{\Lambda}^i\right)^2\right\} \\ & -2\varsigma\sqrt{p}\ \text{Tr}\left\{\mathbf{D}_{nn}^{1/2}\right\}\ \text{Tr}\left\{\sum_{i=0}^{\mathcal{J}}\mu_i\mathbf{\Lambda}^{i+1}\right\}\Bigg] + \varsigma^2 P_{\text{AN}} \\ & + 1 + \varsigma^2\ \text{Tr}\{\mathbf{\Sigma}_n\} + \varsigma^2 p N_T\ \text{Tr}\{\mathbf{D}_{nn}\mathbf{\Delta}_n\}. \end{aligned} \tag{53}$$

Next, we introduce the Vandermonde matrix $\mathbf{C}_1 \in \mathbb{R}^{K\times(\mathcal{J}+1)}$, where $[\mathbf{C}_1]_{i,j} = \lambda_i^{j-1}$, and $\mathbf{\lambda} = [\lambda_1, \ldots, \lambda_K]^T$, which allows us to rewrite (53) in compact form as

$$\begin{aligned}\text{mse}_n = \lim_{K\to\infty}\frac{1}{K}\mathbb{E}\Big[ & \varsigma^2 p\ \text{Tr}\{\mathbf{D}_{nn}\}\,\boldsymbol{\mu}^T\mathbf{C}_1^T\mathbf{\Lambda}^2\mathbf{C}_1\boldsymbol{\mu} \\ & -2\varsigma\sqrt{p}\ \text{Tr}\left\{\mathbf{D}_{nn}^{1/2}\right\}\boldsymbol{\mu}^T\mathbf{C}_1^T\mathbf{\lambda}\Big] + \varsigma^2 P_{\text{AN}} \\ & + 1 + \varsigma^2\ \text{Tr}\{\mathbf{\Sigma}_n\} + \varsigma^2 p N_T\ \text{Tr}\{\mathbf{D}_{nn}\mathbf{\Delta}_n\}. \end{aligned} \tag{54}$$

Similarly, the constraint in (20) can be expressed as

$$\lim_{K\to\infty}\frac{1}{K}\mathbb{E}\left[\boldsymbol{\mu}^T\mathbf{C}_1^T\mathbf{\Lambda}\mathbf{C}_1\boldsymbol{\mu}\right] = N_T. \tag{55}$$

Thus, the Lagrangian function of primal problem (20) can be expressed as $\mathcal{L}_1(\boldsymbol{\mu}, \varsigma) = \text{mse}_n + \epsilon_1(\lim_{K\to\infty}\frac{1}{K}\mathbb{E}[\boldsymbol{\mu}^T\mathbf{C}_1^T\mathbf{\Lambda}\mathbf{C}_1\boldsymbol{\mu}] - N_T)$, where $\epsilon_1$ is the Lagrangian multiplier. Taking the gradient of the Lagrangian function with respect to $\boldsymbol{\mu}$, and setting the result to zero, we obtain for the optimal coefficient vector $\boldsymbol{\mu}_{\text{opt}}$:

$$\begin{aligned}&\lim_{K\to\infty}\frac{1}{K}\mathbb{E}\left[\mathbf{C}_1^T\mathbf{\Lambda}\left(\mathbf{\Lambda} + \frac{\epsilon_1}{\text{Tr}\{\mathbf{D}_{nn}\}\,\varsigma^2 p}\mathbf{I}_K\right)\mathbf{C}_1\right]\boldsymbol{\mu} \\ &\quad = \frac{\text{Tr}\left\{\mathbf{D}_{nn}^{1/2}\right\}}{\varsigma\sqrt{p}\ \text{Tr}\{\mathbf{D}_{nn}\}}\lim_{K\to\infty}\frac{1}{K}\mathbb{E}\left[\mathbf{C}_1^T\mathbf{\lambda}\right]. \end{aligned} \tag{56}$$

Furthermore, taking the derivative of $\mathcal{L}_1(\boldsymbol{\mu}, \varsigma)$ with respect to $\varsigma$ and equating it to zero, and multiplying both sides of (56) by $\boldsymbol{\mu}^T$ and applying (55), we obtain

$$\frac{\epsilon_1}{\varsigma^2 p} = \text{Tr}\{\mathbf{D}_{nn}\mathbf{\Delta}_n\} + \frac{P_{\text{AN}} + \text{Tr}\{\mathbf{\Sigma}_n\}}{N_T p}. \tag{57}$$

The expressions involving $\mathbf{C}_1$, $\mathbf{\Lambda}$, and $\mathbf{\lambda}$ in (56) can be further simplified. For example, we obtain

$$\lim_{K\to\infty}\mathbb{E}\left[\frac{1}{K}\left[\mathbf{C}_1^T\mathbf{\Lambda}\mathbf{C}_1\right]_{m,n}\right] = \lim_{K\to\infty}\mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K}\lambda_k^{m+n-1}\right]. \tag{58}$$

Simplifying the other terms in (56) in a similar manner and inserting (57) into (56) we obtain the result in Theorem 1.

### C. Proof of Theorem 2

Exploiting $\mathbb{E}[\mathbf{z}_n\mathbf{z}_n^H] = \mathbf{I}_{N_T}$, the constraint in (31), and a similar approach as was used to arrive at (25), the objective function in (31) can be simplified as

$$\begin{aligned}P_{\text{AN}} &= q\mathbb{E}\left[\ \text{Tr}\left\{\mathbf{G}_{nn}\mathbf{A}_n\mathbf{A}_n^H\mathbf{G}_{nn}^H\right\}\right] \\ &= q\mathbb{E}\left[\ \text{Tr}\left\{\mathbf{D}_{nn}\hat{\mathbf{H}}_{nn}\mathbf{A}_n\mathbf{A}_n^H\hat{\mathbf{H}}_{nn}^H\right\}\right] \\ &\quad + (1-\phi)P_T\ \text{Tr}\{\mathbf{D}_{nn}\mathbf{\Delta}_n\}. \end{aligned} \tag{59}$$

Using (30) and a similar approach as in Appendix B, (59) can be rewritten as

$$\begin{aligned}P_{\text{AN}} = {} & (1-\phi)P_T\ \text{Tr}\{\mathbf{D}_{nn}\mathbf{\Delta}_n\} \\ & + qN_T\ \text{Tr}\{\mathbf{D}_{nn}\}\,\mathbb{E}\Bigg[-2\ \text{Tr}\left\{\sum_{j=0}^{\mathcal{J}}\nu_j\mathbf{\Lambda}^{j+2}\right\} \\ & + \text{Tr}\{\mathbf{\Lambda}\} + \ \text{Tr}\left\{\mathbf{\Lambda}\left(\sum_{i=0}^{\mathcal{J}}\nu_j\mathbf{\Lambda}^{j+1}\right)^2\right\}\Bigg] \end{aligned} \tag{60}$$

Defining Vandermode matrix $\mathbf{C}_2 \in \mathbb{R}^{K\times(\mathcal{J}+1)}$, where $[\mathbf{C}_2]_{i,j} = \lambda_i^{j-1}$, we can rewrite (60) in compact form as

$$\begin{aligned}P_{\text{AN}} = {} & (1-\phi)P_T\ \text{Tr}\{\mathbf{D}_{nn}\mathbf{\Delta}_n\} + qN_T\ \text{Tr}\{\mathbf{D}_{nn}\}\lim_{K\to\infty} \\ & \frac{1}{K}\mathbb{E}\left[-2\boldsymbol{\nu}^T\mathbf{C}_2^T\mathbf{\Lambda}\mathbf{\lambda} + \mathbf{1}^T\mathbf{\lambda} + \boldsymbol{\nu}^T\mathbf{C}_2^T\mathbf{\Lambda}^3\mathbf{C}_2\boldsymbol{\nu}\right], \end{aligned} \tag{61}$$

where $\mathbf{1}$ denotes the all-ones column vector. Taking into account the constraint in (31), we can formulate the Lagrangian as $\mathcal{L}_2(\boldsymbol{\nu}) = P_{\text{AN}} + \epsilon_2(\lim_{K\to\infty}\frac{1}{K}\mathbb{E}[\boldsymbol{\nu}^T\mathbf{C}_2^T\mathbf{\Lambda}^2\mathbf{C}_2\boldsymbol{\nu} - 2\boldsymbol{\nu}^T\mathbf{C}_2^T\mathbf{\lambda}] + 1)$ with Lagrangian multiplier $\epsilon_2$. The optimal coefficient vector $\boldsymbol{\nu}_{\text{opt}}$ is then obtained by taking the gradient of the Lagrangian function with respect to $\boldsymbol{\nu}$ and setting it to zero:

$$\lim_{K\to\infty}\mathbb{E}\left[\mathbf{C}_2^T\mathbf{\Lambda}^2\left(\mathbf{\Lambda} + \epsilon\mathbf{I}_K\right)\mathbf{C}_2\right]\boldsymbol{\nu} = \lim_{K\to\infty}\mathbb{E}\left[\mathbf{C}_2^T\left(\mathbf{\Lambda} + \epsilon\mathbf{I}_K\right)\mathbf{\lambda}\right], \tag{62}$$

where we used $\epsilon = \frac{\epsilon_2}{qN_T\ \text{Tr}\{\mathbf{D}_{nn}\}}$. Simplifying the terms in (62) by exploiting a similar approach as in Appendix B, we obtain the result in Theorem 2.

$$\begin{aligned}t_2 &\overset{(a)}{=} \varsigma^2 p\mathbb{E}\left[\ \text{Tr}\left\{\tilde{\hat{\mathbf{H}}}_{nn}^H\mathbf{D}_{nn}\tilde{\hat{\mathbf{H}}}_{nn}\right\}\ \text{Tr}\left\{\sum_{i=0}^{\mathcal{J}}\mu_i\left(\hat{\tilde{\mathbf{H}}}_{nn}\hat{\tilde{\mathbf{H}}}_{nn}^H\right)^i\hat{\tilde{\mathbf{H}}}_{nn}\hat{\tilde{\mathbf{H}}}_{nn}^H\sum_{i=0}^{\mathcal{J}}\mu_i\left(\hat{\tilde{\mathbf{H}}}_{nn}\hat{\tilde{\mathbf{H}}}_{nn}^H\right)^i\right\}\right] \\ &\overset{(b)}{=} \varsigma^2 p N_T\ \text{Tr}\{\mathbf{D}_{nn}\mathbf{\Delta}_n\}. \end{aligned} \tag{52}$$

## REFERENCES


[1] J. Zhu, R. Schober, and V. K. Bhargava, "Secure downlink transmission in massive MIMO system with zero-forcing precoding," *in Proc. Eur. Wireless Conf.*, Barcelona, Spain, May 2014, pp. 1–6.

[2] J. Zhu, R. Schober, and V. K. Bhargava, "Secrecy analysis of multi-cell massive MIMO systems with RCI precoding and artificial noise transmission," *in Proc. Int. Symp. Commun. Control Signal Process.*, Athens, Greece, May 2014, pp. 101–104.

[3] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–46, Jan. 2013.

[4] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of BS antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[6] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[7] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1550–1573, Feb. 2014.

[8] A. Khisti and G. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[9] W. Xu, Z. Peng, and S. Jin, "On secrecy of a multi-antenna system with eavesdropper in close proximity," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1525–1529, Oct. 2015.

[10] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

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

[12] 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.

[13] A. Mukherjee and A. L. Swindlehurst, "Robust beamforming for security in MIMO wiretap channels with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 351–361, Jan. 2011.

[14] H. Wang, C. Wang, and D. W. K. Ng, "Artificial noise assisted secure transmission under training and feedback," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6285–6298, Dec. 2015.

[15] T. Zheng, H. Wang, J. Yuan, D. Towsley, and M. H. Lee, "Multi-antenna transmission with artificial noise against randomly distributed eavesdroppers," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 4347–4362, Nov. 2015.

[16] W. Liao, T. Chang, W. Ma, and C. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1202–1216, Mar. 2011.

[17] M. Pei, J. Wei, K.-K. Wong, and X. Wang, "Masked beamforming for multiuser MIMO wiretap channels with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 544–549, Feb. 2012.

[18] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multi-cell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4766–4781, Sep. 2014.

[19] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in UL/DL cellular systems: How many antennas do we need," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[20] T. Dean and A. Goldsmith, "Physical layer cryptography through massive MIMO," *in Proc. IEEE Inf. Theory Workshop*, Sevilla, OH, USA, Sep. 2013, pp. 1–5.

[21] G. Geraci, H. S. Dhillon, J. G. Andrews, J. Yuan, and I. B. Collings, "Physical layer security in downlink multi-antenna cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 2006–2021, Jun. 2014.

[22] G. Geraci, M. Egan, J. Yuan, A. Razi, and I. B. Collings, "Secrecy sum-rates for multi-user MIMO regularized channel inversion precoding," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3472–3482, Nov. 2012.

[23] G. Geraci, J. Yuan, and I. B. Collings, "Large system analysis of linear precoding in MISO broadcast channels with confidential messages," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1660–1671, Sep. 2013.

[24] X. Chen, L. Lei, H. Zhang, and C. Yuen, "Large-scale MIMO relaying techniques for physical layer security: AF or DF?," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5135–5146, Sep. 2015.

[25] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems, "*IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[26] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: A large system analysis," in *Proc. IEEE Global Commun. Conf.*, New Orleans, LO, USA, Dec. 2008, pp. 1–4.

[27] R. R. Müller and S. Verdu, "Design and analysis of low-complexity interference mitigation on vector channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1429–1441, Aug. 2001.

[28] R. R. Müller, "Polynomial expansion equalizers for communication via large antenna arrays," in *Eur. Pers. Mobile Commun. Conf. (EPMCC)*, Feb. 2001, pp. 1–7.

[29] S. Zarei, W. Gerstacker, R. R. Muller, and R. Schober, "Low-complexity linear precoding for downlink large-scale MIMO systems," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 1119–1124.

[30] A. Müller, A. Kammoun, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Reducing complexity in massive MIMO," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, 2014, pp. 273–276.

[31] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO Systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 861–875, Oct. 2014.

[32] D. Kapetanovic, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 21–27, Jun. 2015.

[33] Y. Basciftci, C. Koksal, and A. Ashikhmin, "Securing massive MIMO at the physical layer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4766–4781, Sep. 2014.

[34] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission in the presence of an active eavesdropper," in *Proc. IEEE Int. Commun. Conf. (ICC'15)*, London, U.K., Jun. 2015, pp. 1434–1440.

[35] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*. Providence, RI, USA: AMS, 2006.

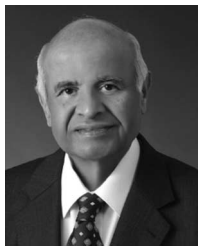
[36] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2059–2078, Sep. 2000.

[37] R. Hunger, "*Floating point operations in matrix-vector calculus*," Technische Universität München, Associate Institute for Signal Processing, Tech. Rep. TUM-LNS-TR-05-05, Oct. 2005.

[38] A. M. Tulino and S. Verdu, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.


**Jun Zhu** (S'10) was born in Nanjing, China. He received the B.Sc. degree (with high distinction) in information engineering from Southeast University, Nanjing, China, and the M.A.Sc. degree (with high distinction) in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 2008 and 2011, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at the University of British Columbia, Vancouver, BC, Canada. He was a Visiting Student at the Institute for Digital Communications (IDC), Friedrich Alexander University (FAU), Erlangen, Germany, between 2014 and 2015. His research interests include MIMO–OFDM wireless systems, massive MIMO, energy-efficient (green) communications, and physical layer security. He was a Finalist of the Lieutenant Governors Silver Medal for Outstanding Master Thesis at the University of Victoria and was the recipient of the Dr. Esme Foord Scholarship in 2011, and also the recipient of Four-Year-Fellowship (FYF) at the University of British Columbia since 2011. He was also the holder of the Pei-Huang Tung and Tan-Wen Tung Fellowship in 2012, Graduate Support Initiative Award in 2013, Chinese Government Award for Outstanding Self-Financed Students Abroad in 2014, German Academic Exchange Service (DAAD) Grant, and UBC International Research Award in 2015.

**Robert Schober** (S'98–M'01–SM'08–F'10) was born in Neuendettelsau, Germany, in 1971. He received the Diplom (Univ.) and Ph.D. degrees in electrical engineering from the University of Erlangen-Nuermberg, Erlangen, Germany, in 1997 and 2000, respectively. From May 2001 to April 2002, he was a Postdoctoral Fellow with the University of Toronto, ON, Canada, sponsored by the German Academic Exchange Service (DAAD). Since May 2002, he has been with the University of British Columbia (UBC), Vancouver, BC, Canada, where he is now a Full Professor. Since January 2012, he is an Alexander von Humboldt Professor and the Chair for Digital Communication with the Friedrich Alexander University (FAU), Erlangen, Germany. His research interests include communication theory, wireless communications, and statistical signal processing. He is a Fellow of the Canadian Academy of Engineering and the Engineering Institute of Canada. He is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON COMMUNICATIONS. He was the recipient of several awards for his work including the 2002 Heinz Maier Award of the German Science Foundation (DFG), the 2004 Innovations Award of the Vodafone Foundation for Research in Mobile Communications, the 2006 UBC Killam Research Prize, the 2007 Wilhelm Friedrich Bessel Research Award of the Alexander von Humboldt Foundation, the 2008 Charles McDowell Award for Excellence in Research from UBC, a 2011 Alexander von Humboldt Professorship, and a 2012 NSERC E.W.R. Steacie Fellowship. He was also the recipient of best paper awards from the German Information Technology Society (ITG), the European Association for Signal, Speech, and Image Processing (EURASIP), the IEEE WCNC 2012, the IEEE Globecom 2011, the IEEE ICUWB 2006, the International Zurich Seminar on Broadband Communications, and European Wireless 2000.



**Vijay K. Bhargava** (S'70–M'74–SM'82–F'92–LF'13) was born in Beawar, India, in 1948. He received the B.A.Sc., M.A.Sc., and Ph.D. degrees from Queens University at Kingston, Kingston, ON, Canada, in 1970, 1972, and 1974, respectively. He is a Professor with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada, where he served as the Department Head from 2003 to 2008. Previously, he was with the University of Victoria, Victoria, BC, Canada (1984–2003), Concordia University, Montreal, QC, Canada (1976–1984), the University of Waterloo, Waterloo, ON, Canada (1976), and the Indian Institute of Science, Bengaluru, India (1974–1975). He has held visiting appointments at Ecole Polytechnique de Montral, NTT Research Laboratory, Tokyo Institute of Technology, Tokyo, Japan, the University of Indonesia, Jawa Barat, Indonesia, the Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, and Tohoku University, Sendai, Japan. He is an Honourary Professor with UESTC, Chengdu, China, and a Gandhi Distinguished Professor at IIT Bombay, Mumbai, India. For the academic year 2015–2016, he is on sabbatical leave at the Friedrich Alexander University (FAU), Erlangen, Germany. He is in the Institute for Scientific Information (ISI) Highly Cited list. He served as the Founder and the President of Binary Communications Inc. (1983–2000). He is a coauthor of *Digital Communications by Satellite* (Wiley, 1981), which was translated in Chinese and Japanese. He is a coeditor of *Reed Solomon Codes and their Applications* (IEEE Press, 1994), *Communications, Information and Network Security* (Kluwer, 2003), *Cognitive Wireless Communication Networks* (Springer, 2007), *Cooperative Wireless Communications Networks* (Cambridge University Press, 2011), and *Green Radio Communications Networks* (Cambridge University Press, 2012). He is a Fellow of the Royal Society of Canada, the Canadian Academy of Engineering and the Engineering Institute of Canada. He is a Foreign Fellow of the National Academy of Engineering (India) and has served as a Distinguished Visiting Fellow of the Royal Academy of Engineering (U.K.). He has served on the Board of Governors of the IEEE Information Theory Society and the IEEE Communications Society. He has held important positions in these societies. He has served as an Associate Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS. He played a major role in the creation of the IEEE Communications and Networking Conference (WCNC) and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, for which he served as the Editor-in-Chief (2007–2009). He is a past President of the IEEE Information Theory Society and a Past President of the IEEE Communications Society. He was the recipient of awards for his teaching, research, and service to the IEEE. He was also the recipient of the Killam Prize in Engineering and the Humboldt Research Prize.