PHY Layer Anonymous Precoding: Sender Detection Performance and Diversity-Multiplexing Tradeoff

Zhongxiang Wei, *Member, IEEE,* Christos Masouros, *Senior Member, IEEE,* Xu Zhu, *Senior Member, IEEE,*
Ping Wang, *Member, IEEE,* and Athina P. Petropulu, *Fellow, IEEE*

Abstract

Departing from traditional data security-oriented designs, the aim of anonymizing techniques is to conceal the transmitters' identities during communications to all possible receivers. In this work, joint anonymous transceiver design at the physical (PHY) layer is investigated. We first present sender detection error rate (DER) performance analysis, where closed-form expression of DER is derived for a generic precoding scheme applied at the transmitter side. Based on the tight DER expression, a fully DER-tunable anonymous transceiver design is demonstrated. An alias channel-based combiner is first proposed, which helps the receiver find a Euclidean space that is close to the propagation channel of the received signal for high quality reception, but does not rely on the recognition of the real sender's channel. Then, two novel anonymous precoders are proposed under a given DER requirement, one being able to provide full multiplexing performance, and the other flexibly adjusting the number of multiplexing streams with further consideration of the receive-reliability. Simulation demonstrates that the proposed joint transceiver design can always guarantee the subscribed DER performance, while well striking the trading-off among the multiplexing, diversity and anonymity performance.

Zhongxiang Wei and Ping Wang are with the College of Electronic and Information Engineering, at Tongji University, Shanghai, China. Email: {z_wei, pwang}@tongji.edu.cn

Christos Masouros is with the department of Electronic and Electrical Engineering at the University College London, London, UK. Email: c.masouros@ucl.ac.uk

Xu Zhu is with the School of Electronic and Information Engineering, Harbin Institute of Technology, Shenzhen, China. Email: xuzhu@ ieee.org

Athina P. Petropulu is with the Department of Electrical and Computer Engineering, Rutgers University, New Brunswick, NJ 08901 USA (e-mail: athinap@rutgers.edu)

Index Terms

Physical Layer Anonymity, Tunable DER, Anonymous Transceiver Design, Multiplexing and Diversity Trade-off

I. INTRODUCTION

In last decades, wireless communications security has been extensively investigated at all network layers, from the upper layers to the physical layer (PHY) [1]. Related topics range from cryptographic primitives to information-theoretic designs, including but not limited to encryption, authentication [2], secure precoding plus artificial noise [3], cooperative jamming [4] [5] [6], PHY authentication [7], covert communications [8], among others. In general, the aim of data security is to prevent confidential data from being exploited by external eavesdroppers. With 5G and looking towards 6G, new applications have emerged, requiring new types of security and privacy. For example, users may need to offload their data to a legitimate edge receiver for obtaining utility, such as e-voting, remote-health, computing and recording [9]. During that process, a curious receiver may infer the user's identity (ID) or other non-shared data, such as the individual's lifestyle, habits, and whereabouts. This constitutes privacy leakage towards a legitimate but curious communication party. Different from ensuring data security, the aim of privacy protection is to guarantee accuracy of the released data for utility, while minimizing the receiver's capability to infer the non-shared information [10]. For example, the well-known "differential privacy" was first proposed in querying databases, aiming at answering queries while ensuring privacy of individual records in the datasets [11]. The design principle is to suppress the receiver's gain in terms of the probability of correctly guessing the non-shared sensitive information after observing the disclosed data, by perturbing the released data. The concept of differential privacy recently has been extended to maximal leakage logarithmic gain [12], α-leakage and maximal α-leakage [13], and other divergence-based metrics. Nevertheless, this mechanism reduces the fidelity of the released data, and thus is mainly used for data statistics, such as average and variance of income [13] [14].

To countermeasure privacy leakage while guaranteeing data accuracy, the concept of anonymous communication has attracted attention in recent years. It is also termed as user anonymity design, referring to the absence of identifying information of an individual in the transmitted signal [15]. The design principle is to mask the user's ID and other associated characteristics

towards a legitimate receiver, while ensuring reliable detection of the shared data for communication by the same receiver. For example, the anonymous authentication and encryption designs at higher layers let the sender apply pseudo accounts, instead of its real ID, during the authentication and encryption process [16]. However, a curious receiver may analyze the data traffic at the network layer, and associate the traffic pattern with a specific user' ID. For stronger anonymity, a user can complicate the routing path via a number of proxy servers, such as the onion networks [17], where the traffic characteristics are hidden by the extended routing length. However, merely removing users' IDs and higher layer network information (routes) may still not provide sufficient protection. Indeed, the released information, when coupled with a user's unique channel characteristic, can reveal the identity of an individual at the PHY layer. As a result, a receiver can analyze the signalling patterns of the received signal to unmask the data sender, referred to as PHY sender detection [15]. To counteract the PHY sender detection, the concept of anonymous precoding was proposed in [18]. Different from the classic throughput maximization [19], power minimization [20], minimization of weighted-sum of mean square error [21], or other anonymity-agnostic precoders [22], anonymous precoding incorporates a so-called anonymous constraint. Its purpose is to eliminate the user-dependent channel characteristics from the received signal, so that aliases can be intentionally created [18]. As per [18], aliases are a subset of the multiple access channel users, that the precoder mimics, to prohibit sender identification at the receiver side. As a result, when the receiver tries to associate certain channel characteristic to a specific user for sender detection, the detection error rate (DER) performance is significantly degraded. As a further step, the work in [23] investigated the anonymous precoding design from the perspective of anonymity entropy, which aims at scrambling the receiver's detection as much as possible by an iterative algorithm.

There are still open challenges in the area of anonymous precoder design. 1) The DER performance of the anonymity-agnostic or anonymous precoders are only numerically evaluated so far. As there is no DER performance analysis for generic precoders, the anonymity performance gain of the anonymous precoder has not been quantified yet. Hence, the precoder of [18] relies on an empirical anonymous constraint, which provides qualitative DER only. Also, the target of the anonymous precoder in [23] is to scramble the DER performance as much as possible. Its anonymity comes at the cost of significant degree-of-freedom (DoF) reduction of the precoder design. 2) The existing anonymous precoding cannot provide a fully tunable DER performance. In practice, heterogeneous anonymity performance may be required. For example,

reporting physiological signal in e-Health has high anonymity requirement, but offloading non-sensitive data has low anonymity requirement. 3) The existing anonymous precoders cannot strike a good tradeoff among the anonymity, multiplexing, and diversity performance. With joint precoder and combiner design, the classic anonymity-agnostic precoders are able to multiplex up to $\min\{N_r, N_t\}$ streams, with N_r and N_t denoting the number of receive- and transmit-antennas. However, in anonymous communications, as the receiver is unaware who the real sender is, it is challenging to design a channel-dependent combiner at the receiver-side. The existing anonymous precoders either use an equal-gain combiner, where only one data stream is conveyed and have poor multiplexing performance, or the existing anonymous precoders treat each receive-antenna as an individual receiver for multiplexing (thus no combiner is performed), where per stream receive-reliability is not guaranteed and have low diversity performance.

Motivated by the above challenges, in this work we present a DER-tunable anonymous transceiver design, and strike the balance among the anonymity, multiplexing, and diversity performance. Our contributions can be summarized as follows.

- We first present analytic anonymity DER performance, where the closed-form of the DER is derived for generic precoders. It is demonstrated that the DER acts as a function of PHY parameters, i.e., block length, applied precoder, and noise statistics. The derived closed-form expression is shown to be tight to the true DER result, regardless of the system antenna configuration.
- Aided by the quantitative DER analysis, we then propose a framework for DER-tunable joint transceiver design. Explicitly, with a threshold DER requirement, we first calculate the minimum number of user aliases and formulate a corresponding anonymous constraint towards the dissipated signalling pattern. This constraint creates a set of artificial alias channels that mask the true channel of the sender. Then, an alias channel based combiner is proposed for high quality reception. This combiner finds a Euclidean space that is close to the propagation channel of the received signal, but does not rely on the recognition of the real sender's channel. Hence, the receiver only needs to build a combiner for an approximate channel based on the set of alias channels, to enable reliable shared-data detection, while it does not need to infer the sender's identity.
- A so called lower-bound anonymity (LBA) precoder is designed to multiplex $\min\{N_r, N_t\}$ spatial streams, while ensuring that the obtained DER is strictly higher than the minimum required for anonymity. As a further step, we demonstrate that the upper bound of the shared-

data error probability directly depends on that of each spatial stream, which is then used to build a per-stream receive-SNR constraint for the purpose of diversity (reliability). Then, a diversity-multiplexing-tradeoff lower-bound anonymity (DM-LBA) precoder is further proposed, which adaptively finds the reasonable number of multiplexing streams with system anonymity as well as diversity requirements. Hence, the DM-LBA precoder well trades-off the diversity, multiplexing and anonymity performance.

Notations: Matrices and vectors are represented by boldface capital and lower case letters, respectively. $|\cdot|$ calculates the absolute value of a complex number or denotes cardinality of a set. $||\cdot||_p$ calculates the p-norm. $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian transpose of a matrix. $\boldsymbol{I}_n$ denotes an n-by-n identity matrix. $\mathbb{E}(\cdot)$ and $\mathbb{V}(\cdot)$ represent expectation and variance of a random variable. $\mathcal{N}\{\cdot\}$ denotes Gaussian distribution.

II. System Model and Sender Detection

In this section, system model and sender detection are demonstrated in subsection II-A and B, respectively.

A. System Model

We consider an anonymous multiuser multiple-input and multiple-output (MIMO) scenario, where K ($|\mathbb{K}| = K$, $\mathbb{K}$ denotes the user set) users anonymously transmit shared-data to a base station (BS) in a time-division manner, without leaking their identities. Assume that the BS is equipped with N_r receive-antennas, while each user is equipped with N_t transmit-antennas. Typically, the number of the receive-antennas is larger than that of the transmit-antennas ($N_r > N_t$) at uplink transmission. Define $\boldsymbol{H}_k \in \mathbb{C}^{N_r \times N_t}$ as the channel between the k-th user and the BS. Define $\boldsymbol{W}_k$ and $\boldsymbol{S}$ as the precoding matrix and transmitted symbol matrix at the k-th user, i.e., $\boldsymbol{W}_k\boldsymbol{S} \in \mathbb{C}^{N_t \times L}$ with L denoting block-length. The received signal at the BS is then calculated as

$$\boldsymbol{Y} = \boldsymbol{H}_k\boldsymbol{W}_k\boldsymbol{S} + \boldsymbol{Z}, \tag{1}$$

where $\boldsymbol{Z} \in \mathbb{C}^{N_r \times L}$ denotes the circularly symmetric complex Gaussian (CSCG) noise with variance σ^2.

B. Sender Detection

For completeness, let us briefly describe the least-Euclidean distance based detector [18]. The BS has knowledge of all users propagation channels $\boldsymbol{H}_k, k \in \mathbb{K}$, and only analyzes PHY information, i.e., the inherent characteristics of the received signal to reveal the identity of the sender k. A multiple hypotheses testing (MHT) problem is formulated as

$$\mathcal{Y} = \begin{cases} \mathcal{H}_0: & \boldsymbol{Z}, \\ \mathcal{H}_1: & \boldsymbol{H}_1\boldsymbol{W}_1\boldsymbol{S} + \boldsymbol{Z}, \\ \quad\vdots & \\ \mathcal{H}_K: & \boldsymbol{H}_K\boldsymbol{W}_K\boldsymbol{S} + \boldsymbol{Z}, \end{cases} \tag{2}$$

Explicitly, the hypothesis $\mathcal{H}_0$ denotes that there is no transmission and only noise appears at the BS, while hypothesis $\mathcal{H}_k$ means there is a signal coming from the k-th user. The distinction between hypothesis $\mathcal{H}_0$ and the rest can be performed through classic energy detection [24], where the test statistic is compared against a threshold β, i.e.,

$$\Lambda(\boldsymbol{Y}) = \frac{||\boldsymbol{Y}||_F^2}{N_r L} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1 \sim \mathcal{H}_K}{\gtrless}} \beta, \tag{3}$$

where $||\cdot||_F$ denotes the Frobenius norm. The value of threshold β is set based on the Neyman-Pearson criterion. Once $\mathcal{H}_0$ is decided as a false hypothesis, the BS turns to detect the origin of the signal. As shown in (1), the characteristic of the received signal is coupled to the channel of the sender. Suppose that the BS utilizes the correct propagation channel for testing, the maximum likelihood estimation (MLE) of the transmitted signal is given as $\hat{\boldsymbol{X}}_k = \boldsymbol{H}_k^{\dagger}\boldsymbol{Y}$, where $\boldsymbol{H}_k^{\dagger} = (\boldsymbol{H}_k^H\boldsymbol{H}_k)^{-1}\boldsymbol{H}_k^H$. Then, a re-constructed signal is given as $\hat{\boldsymbol{Y}}_k = \boldsymbol{H}_k\hat{\boldsymbol{X}}_k = \boldsymbol{H}_k\boldsymbol{W}_k\boldsymbol{S} + \boldsymbol{H}_k\boldsymbol{H}_k^{\dagger}\boldsymbol{Z}$. The Euclidean distance between the re-constructed signal $\hat{\boldsymbol{Y}}_k$ and the actual signal $\boldsymbol{Y}$ is calculated as $d_k = ||\boldsymbol{Y} - \hat{\boldsymbol{Y}}_k||^2 = ||(\boldsymbol{H}_k\boldsymbol{H}_k^{\dagger} - \boldsymbol{I}_{N_r})\boldsymbol{Z}||^2$. On the other hand, if the BS uses the i-th user's channel for testing, $i \in \mathbb{K}, i \neq k$, the Euclidean distance between the actual signal $\boldsymbol{Y}$ and re-constructed signal $\hat{\boldsymbol{Y}}_i$ is calculated as $d_i = ||(\boldsymbol{H}_i\boldsymbol{H}_i^{\dagger} - \boldsymbol{I}_{N_r})\boldsymbol{H}_k\boldsymbol{W}_k\boldsymbol{S} + (\boldsymbol{H}_i\boldsymbol{H}_i^{\dagger} - \boldsymbol{I}_{N_r})\boldsymbol{Z}||_F^2$. As d_k only contains a colored-noise term, there is high probability that the value of d_i is larger than that of d_k. Hence, the sender detector in [18] points out that the BS can use different possible channels for testing, and classifies the one having the smallest Euclidean distance to the received signal, i.e., $\min_{k\in\mathbb{K}}\{||\boldsymbol{Y} - \boldsymbol{H}_1\boldsymbol{H}_1^{\dagger}\boldsymbol{Y}||_F^2, ..., ||\boldsymbol{Y} - \boldsymbol{H}_K\boldsymbol{H}_K^{\dagger}\boldsymbol{Y}||_F^2\}$ as the sender.

III. Analytical Closed-form Derivation of the Sender Detection Error Rate

In this section, we analyze the PHY DER of the BS, where the result is used to aid the joint anonymous transceiver design to be presented in Section IV. Define type-k error probability as the probability that, given event $\mathcal{H}_k$, the receiver falsely declares either that no one sends, or that a user other than user k sends. For the considered MHT problem, the type-k error probability measures the DER performance, given as

$$\tau = 1 - \underbrace{\Pr(\Lambda(\boldsymbol{Y}) \geq \beta | \mathcal{H}_k)}_{a} \underbrace{\prod_{i=1, i\neq k}^{K} \Pr(d_i \geq d_k | \mathcal{H}_k)}_{b}, \tag{4}$$

where the term "*a*" represents the probability that, under event $\mathcal{H}_k$, the BS correctly declares the presence of an incoming signal. The term "*b*" represents the probability that, given event $\mathcal{H}_k$, the BS correctly identifies the signal coming from user k. In practice, though the sender detection is always performed at the block-level, the precoder may change at block- or symbol-level. Hence, in the following, we analyze the DER of generic block- and symbol-level precoders respectively.

A. DER of Generic Block-Level Precoders

A generic block-level (BL) precoder $\boldsymbol{W}_k$ is a function of the sender's channel [21], i.e., $\boldsymbol{W}_k = f(\boldsymbol{H}_k)$, which thus remains constant in each block[1]. Hence, for a block duration consisting of L symbol vectors, the term "*a*" is the complement of the probability of miss detection, calculated as

$$\Pr(\Lambda(\boldsymbol{Y}) \geq \beta | \mathcal{H}_k) = 1 - \Pr(\Lambda(\boldsymbol{Y}) < \beta | \mathcal{H}_k) = 1 - \mathcal{F}(\frac{2\beta L N_r}{\sigma^2}), \tag{5}$$

where the proof can be similarly found in [25] [26] and thus is omitted to avoid repetition. $\mathcal{F}(\cdot)$ denotes the cdf of a non-central Chi-Square random variable with $2LN_r$ DoF and non-centrality parameter $\frac{2||\boldsymbol{H}_k \boldsymbol{W}_k \boldsymbol{S}||_F^2}{\sigma^2}$. Now, we analyze the probability that, under event $\mathcal{H}_k$, the BS correctly declares event $\mathcal{H}_i$ being false, i.e., $\Pr(d_i \geq d_k | \mathcal{H}_k)$ in term "*b*". Evidently, we first need to investigate the statistical distributions of the variables d_k and d_i, respectively.

For the simplicity of notation, let $\boldsymbol{\Xi}_k = \boldsymbol{H}_k \boldsymbol{H}_k^{\dagger} - \boldsymbol{I}_{N_r}$, $\forall k \in \mathbb{K}$. Recall that d_k is in a quadratic form with respect to (w.r.t.) the noise term. Assuming independent and identically distributed (i.i.d.) channel and noise statistics, the expectation and variance of d_k are calculated as

[1][21] formulated a series of linear precoders for the class of Schur-concave and Schur-convex cost functions, which encompass most of the existing precoders. We refer readers to [21] for details.

$$\mathbb{E}\{||\boldsymbol{\Xi}_k\boldsymbol{Z}||_F^2\} = L\sigma^2\mathrm{tr}\Big(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\Big), \tag{6}$$

and

$$\mathbb{V}\{||\boldsymbol{\Xi}_k\boldsymbol{Z}||_F^2\} = L\sigma^4\mathrm{tr}\Big(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\Big), \tag{7}$$

where proof is shown in Appendix A. On the other hand, the value of d_i is related to the precoding matrix $\boldsymbol{W}_k$ and noise. Let $\boldsymbol{V}_i = \boldsymbol{\Xi}_i\boldsymbol{H}_k\boldsymbol{W}_k\boldsymbol{S}$. d_i can be calculated as $d_i = ||\boldsymbol{V}_i + \boldsymbol{\Xi}_i\boldsymbol{Z}||_F^2$. Define an operator $\mathrm{vec}(\cdot)$ which stacks columns of a matrix into a vector, and thus we have $d_i = ||\boldsymbol{V}_i + \boldsymbol{\Xi}_i\boldsymbol{Z}||_F^2 = ||\mathrm{vec}(\boldsymbol{V}_i + \boldsymbol{\Xi}_i\boldsymbol{Z})||_2^2$. The expectation of d_i is given as

$$\begin{aligned}\mathbb{E}\{d_i\} &= \mathbb{E}\{\mathrm{tr}(\mathrm{vec}(\boldsymbol{V}_i + \boldsymbol{\Xi}_i\boldsymbol{Z})\mathrm{vec}(\boldsymbol{V}_i + \boldsymbol{\Xi}_i\boldsymbol{Z})^H)\}\\ &= \mathrm{tr}(\mathbb{E}\{\mathrm{vec}(\boldsymbol{V}_i + \boldsymbol{\Xi}_i\boldsymbol{Z})\mathrm{vec}(\boldsymbol{V}_i + \boldsymbol{\Xi}_i\boldsymbol{Z})^H\})\\ &= L\sigma^2\mathrm{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i) + \mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i),\end{aligned} \tag{8}$$

and its variance is given as

$$\mathbb{V}\{d_i\} = L\sigma^4\mathrm{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i) + 2\sigma^2\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i\boldsymbol{V}_i), \tag{9}$$

where the derivation of the variance is similar to that in Appendix, and thus is omitted due to page limitation.

Now, we have obtained the expectation and variance of d_k and d_i, but their exact statistic distribution may still be difficult to know. In fact, the values of d_k and d_i are contributed by N_rL samples. Leveraging the central limit theorem by allowing L to grow large, we thus approximate d_k and d_i by a Gaussian distribution. On defining a variable $\zeta_i = d_k - d_i$, we have that $\Pr(d_i \geq d_k|\mathcal{H}_k) = \Pr(\zeta_i \leq 0|\mathcal{H}_k)$. Since the difference of d_i and d_k still follows a Gaussian distribution, the expectation of ζ_i is given as

$$\mathbb{E}\{\zeta_i\} = \mathbb{E}\{d_k\} - \mathbb{E}\{d_i\} = L\sigma^2\mathrm{tr}(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k - \boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i) - \mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i), \tag{10}$$

where the term $\mathrm{tr}(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k - \boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i)$ can be reduced to

$$\begin{aligned}
&\mathrm{tr}(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k-\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i)\\
&=\mathrm{tr}\Big((\boldsymbol{H}_k\boldsymbol{H}_k^\dagger-\boldsymbol{I}_{N_r})^H(\boldsymbol{H}_k\boldsymbol{H}_k^\dagger-\boldsymbol{I}_{N_r})\Big)-\mathrm{tr}\Big((\boldsymbol{H}_i\boldsymbol{H}_i^\dagger-\boldsymbol{I}_{N_r})^H(\boldsymbol{H}_i\boldsymbol{H}_i^\dagger-\boldsymbol{I}_{N_r})\Big)\\
&=\mathrm{tr}\Big(\boldsymbol{I}_{N_r}-\boldsymbol{H}_k(\boldsymbol{H}_k^H\boldsymbol{H}_k)^{-1}\boldsymbol{H}_k^H)\Big)-\mathrm{tr}\Big(\boldsymbol{I}_{N_r}-\boldsymbol{H}_i(\boldsymbol{H}_i^H\boldsymbol{H}_i)^{-1}\boldsymbol{H}_i^H)\Big)\\
&=\mathrm{tr}(\boldsymbol{I}_{N_r})-\mathrm{tr}(\boldsymbol{I}_{N_t})-\Big(\mathrm{tr}(\boldsymbol{I}_{N_r})-\mathrm{tr}(\boldsymbol{I}_{N_t})\Big)=0.
\end{aligned}\tag{11}$$

Hence, (10) can be simplified into

$$\mathbb{E}\{\zeta_i\}=-\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i).\tag{12}$$

Also, the variance of ζ_i is given as

$$\begin{aligned}
\mathbb{V}\{\zeta_i\}=\mathbb{V}\{d_k\}+\mathbb{V}\{d_i\}-2\mathrm{cov}\{d_k,d_i\}\overset{\mathrm{c}}{\simeq}\\
\sigma^4L\mathrm{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i+\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k)+2\sigma^2\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i\boldsymbol{V}_i),
\end{aligned}\tag{13}$$

where step "c" is due to ignoring the covariance term of two weakly correlated variables. Similar to the derivation in (11), we find that $\mathrm{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i)=\mathrm{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i)$ and $\mathrm{tr}(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k)=\mathrm{tr}(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k)$. Thus, (13) can be simplified into

$$\mathbb{V}\{\zeta_i\}=\sigma^4L\mathrm{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i+\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k)+2\sigma^2\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i).\tag{14}$$

For the Gaussian distributed variable ζ_i, the value of $\Pr(\zeta_i\leq 0|\mathcal{H}_k)$ is determined by its cumulative density function (cdf), calculated as

$$\Pr(\zeta_i\leq 0|\mathcal{H}_k)=\int_{-\infty}^{0}f_{\zeta_i}(t)\mathrm{d}t=\frac{1}{2}\Big(1+\mathrm{erf}(\frac{0-\mathbb{E}(\zeta_i)}{\sqrt{2\mathbb{V}(\zeta_i)}})\Big),\tag{15}$$

where $f_{\zeta_i}(\cdot)$ denotes the probability distribution function (pdf) of the variable ζ_i, and $\mathrm{erf}(\cdot)$ denotes the erf function, i.e., $\mathrm{erf}(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\mathrm{d}t$. Substituting (12) and (14) into (15) yields

$$\Pr(\zeta_i\leq 0|\mathcal{H}_k)=\frac{1}{2}\Big(1+\mathrm{erf}(\frac{\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i)}{\sqrt{2\sigma^4L\mathrm{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i+\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k)+4\sigma^2\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i)}})\Big).\tag{16}$$

Substituting (5) and (16) into (4), the DER with a generic BL precoder is given in the closed-form of

$$\tau_{\mathrm{BL}} = 1 - (1 - \mathcal{F}(\frac{2\beta L N_r}{\sigma^2})) \prod_{i,i\neq k}^{K} \frac{1 + \mathrm{erf}(\frac{\mathrm{tr}(\boldsymbol{V}_i^H \boldsymbol{V}_i)}{\sqrt{2\sigma^4 L \mathrm{tr}(\boldsymbol{\Xi}_i^H \boldsymbol{\Xi}_i + \boldsymbol{\Xi}_k^H \boldsymbol{\Xi}_k) + 4\sigma^2 \mathrm{tr}(\boldsymbol{V}_i^H \boldsymbol{V}_i)}})}{2}, \tag{17}$$

With a small valued β, the term $\mathcal{F}(\frac{2\beta L N_r}{\sigma^2})$ approaches 0, which denotes that the miss detection rate can be ignored. Though Neyman-Pearson criterion indicates that a small value of β may raise the probability of false alarm, its effect can be significantly mitigated due to the multiple antennas at the BS [26] [24]. Ignoring the effect of miss detection, a tight bound of the DER is given as

$$\tau_{\mathrm{BL}} = 1 - \prod_{i,i\neq k}^{K} \frac{1 + \mathrm{erf}(\frac{\mathrm{tr}(\boldsymbol{V}_i^H \boldsymbol{V}_i)}{\sqrt{2\sigma^4 L \mathrm{tr}(\boldsymbol{\Xi}_i^H \boldsymbol{\Xi}_i + \boldsymbol{\Xi}_k^H \boldsymbol{\Xi}_k) + 4\sigma^2 \mathrm{tr}(\boldsymbol{V}_i^H \boldsymbol{V}_i)}})}{2}. \tag{18}$$

B. DER of Generic Symbol-Level Precoders

The symbol-level (SL) precoder is able to exploit the correlation among the channels and the transmitted symbols for its precoder design [28], which is written as function of the channel and the transmitted symbol vector. Hence, we now introduce a superscript l as the index of symbol slot, i.e., $\boldsymbol{S} = [\boldsymbol{s}^{(1)}, ..., \boldsymbol{s}^{(L)}]$ and $\boldsymbol{s}^{(l)} \in \mathbb{C}^{N\times 1}$. The SL precoder is given as $\boldsymbol{W}_k^{(l)} = f(\boldsymbol{H}_k, \boldsymbol{s}_k^{(l)})$, $l = 1, ..., L$. Evidently, the term "a" still follows Chi-square distribution, but with a non-central parameter $\frac{2\sum_{l=1}^{L} ||\boldsymbol{H}_k \boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)}||^2}{\sigma^2}$. As shown by (6)-(7), the statistics of d_k is dependent from the precoder design, and thus we only need to re-calculate d_i. Define $\boldsymbol{v}_i^{(l)} = \boldsymbol{\Xi}_i \boldsymbol{H}_k \boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)}$, $\forall i$. The value of d_i is calculated as $d_i = \sum_{l=1}^{L} ||\boldsymbol{v}_i^{(l)} + \boldsymbol{\Xi}_i \boldsymbol{z}||^2$. Note that as noise is independent of the symbol slot, the index l is omitted from the noise term. We have

$$\begin{aligned} \mathbb{E}\{d_i\} &= \mathbb{E}\{\sum_{l=1}^{L} \mathrm{tr}((\boldsymbol{v}_i^{(l)} + \boldsymbol{\Xi}_i \boldsymbol{z})(\boldsymbol{v}_i^{(l)} + \boldsymbol{\Xi}_i \boldsymbol{z})^H)\} \\ &= L\sigma^2 \mathrm{tr}(\boldsymbol{\Xi}_i^H \boldsymbol{\Xi}_i) + \sum_{l=1}^{L} (\boldsymbol{v}_i^{(l)})^H \boldsymbol{v}_i^{(l)}, \end{aligned} \tag{19}$$

On the other hand, the variance of d_i is written as

$$\mathbb{V}\{d_i\} = L\sigma^4 \mathrm{tr}(\boldsymbol{\Xi}_i^H \boldsymbol{\Xi}_i) + 2\sigma^2 \sum_{l=1}^{L} (\boldsymbol{v}_i^{(l)})^H \boldsymbol{\Xi}_i^H \boldsymbol{\Xi}_i \boldsymbol{v}_i^{(l)}, \tag{20}$$

Similarly, let $\zeta_i = d_k - d_i$ for the considered block. Its expectation is given as

$$\mathbb{E}\{\zeta_i\} = -\sum_{l=1}^{L}(\boldsymbol{v}_i^{(l)})^H \boldsymbol{v}_i^{(l)}, \tag{21}$$

and its variance is given as

$$\mathbb{V}\{\zeta_i\} = \sigma^4 L \mathrm{tr}(\boldsymbol{\Xi}_i^H \boldsymbol{\Xi}_i + \boldsymbol{\Xi}_k^H \boldsymbol{\Xi}_k) + 2\sigma^2 \sum_{l=1}^{L}(\boldsymbol{v}_i^{(l)})^H \boldsymbol{v}_i^{(l)}. \tag{22}$$

Substituting (21) and (22) into (16), the DER with a generic SL precoder is given as

$$\tau_{\mathrm{SL}} = 1 - \prod_{i,i\neq k}^{K} \frac{1 + \mathrm{erf}(\frac{\sum_{l=1}^{L}(\boldsymbol{v}_i^{(l)})^H \boldsymbol{v}_i^{(l)},}{\sqrt{2\sigma^4 L \mathrm{tr}(\boldsymbol{\Xi}_i^H \boldsymbol{\Xi}_i + \boldsymbol{\Xi}_k^H \boldsymbol{\Xi}_k) + 4\sigma^2 \sum_{l=1}^{L}(\boldsymbol{v}_i^{(l)})^H \boldsymbol{v}_i^{(l)}}})}{2}. \tag{23}$$

To verify the above analysis, Fig. 1 shows the simulation and theoretic DER performance for some classic precoders applied at the user side. Explicitly, we use minimum mean square error (MMSE) and singular value decomposition (SVD) precoders as the representatives of BL precoders, and use constructive interference (CI) precoder as the representative of SL precoder. It is observed that regardless of BLP and SLP, the derived analytic DER is tight to the simulation result, where the deviation between the simulation and theoretic results is below the level of 10^{-2}. Also, the DER approaches 0 at SNR regions above 5 dB, where the reasons are summarized in below Remarks.

Remark 1: For anonymity-agnostic precoders, the received signal $\boldsymbol{H}_k \boldsymbol{W}_k \boldsymbol{S}$ excluding noise generally does not lie in the null-space of $\boldsymbol{\Xi}_i$, $\forall i \neq k$. Hence, $\mathrm{tr}(\boldsymbol{V}_i^H \boldsymbol{V}_i)$ is a non-zero finite valued number. At moderate and high transmit-SNR regions, a small value of noise variance makes the value of the erf function in (18) approach 1. As a result, by generic anonymity-agnostic precoders, the value of $\tau = 1 - \prod_{i,i\neq k}^{K} \frac{1+1}{2}$ becomes 0, meaning that the BS can perfectly reveal the real sender. The some observation also applies to SL precoders. ■

Remark 2: A large value of block length L helps reduce the value of DER. It is because the erf function is a non-decreasing function with w.r.t L. A extreme case would be $L \to \infty$. It equivalently means that there are infinite numbers of samples for testing, and thus the DER by generic anonymity-agnostic precoders approaches 0 at all SNR regions. Also, the noise status has impact on the value of DER. ■

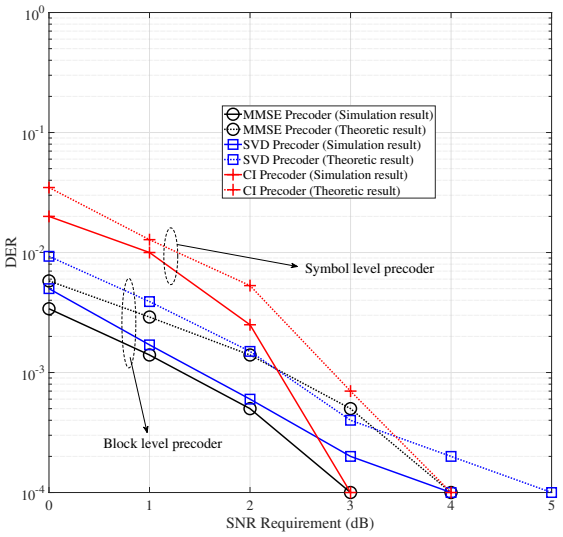

Fig. 1. Simulations vs. theoretic result of the DER performance, where $K = 5$ users, $N_r = 6$ and $N_t = 5$.

Following Remarks 1 and 2, the only manageable variable at the transmitter side for scrambling the DER performance is the precoder [2]. Hence, the design principle of the anonymous precoders is manipulating the transmitted signaling pattern, so that a user (which is termed to as alias sender) other than user k becomes an equally likely sender from the perspective of the BS.

IV. Anonymous Joint Transceiver Design

Under a threshold DER performance, we aim at multiplexing $N = \min\{N_t, N_r\}$ streams as that of anonymity-agnostic precoders, while providing reasonable per-stream SINR performance for communication. Aided by a combiner $\boldsymbol{C} \in \mathbb{C}^{N \times N_r}$, the combined signal is given as

$$\boldsymbol{R} = \boldsymbol{C}\boldsymbol{Y} = \tilde{\boldsymbol{H}}_k \boldsymbol{W}_k \boldsymbol{S} + \tilde{\boldsymbol{Z}}, \tag{24}$$

where $\tilde{\boldsymbol{H}}_k = \boldsymbol{C}\boldsymbol{H}_k = [\boldsymbol{h}_{k1}^H, ..., \boldsymbol{h}_{kN_t}^H]^H$ denotes the equivalent propagation channel of user k, and the vector $\boldsymbol{h}_{kn} \in \mathbb{C}^{1 \times N_t}$ denotes the channel of the n stream. Decompose $\boldsymbol{W}_k = [\boldsymbol{w}_{k1}, ..., \boldsymbol{w}_{kN}]$, where $\boldsymbol{w}_{kn} \in \mathbb{C}^{N_t \times 1}$ denotes the precoding vector of the n-th stream of user k. $\tilde{\boldsymbol{Z}} = \boldsymbol{C}\boldsymbol{Z}$ denotes the equivalent noise. Hence, the SINR of the n-th stream is calculated as

[2]Remark 2 states that the DER is also related to block-length L. Though the block-length optimization is popular in the topic of delay-sensitive networks, in this paper we consider fixed block-length.

$$\Gamma_n = \frac{|\boldsymbol{h}_{kn}\boldsymbol{w}_{kn}|^2}{\sum_{n'\neq n}^{N}|\boldsymbol{h}_{kn}\boldsymbol{w}_{kn'}|^2+\tilde{\sigma}^2}, \forall n, \tag{25}$$

where $\tilde{\sigma}^2$ denotes variance of the equivalent noise with the combiner. Now, the anonymous joint transceiver design is formulated as

$$\begin{aligned} P1 \ &: \underset{\boldsymbol{W}_k,\boldsymbol{C}}{\text{argmax}} \ \min_{\forall n\in N} \ \frac{|\boldsymbol{h}_{kn}\boldsymbol{w}_{kn}|^2}{\sum_{n'\neq n}^{N}|\boldsymbol{h}_{kn}\boldsymbol{w}_{kn'}|^2+\tilde{\sigma}^2}, \\ \text{s.t. } &(\text{C1}) : \tau(\boldsymbol{W}_k) \geq \bar{\tau}, \\ &(\text{C2}) : N = \min\{N_r, N_t\}, \\ &(\text{C3}) : ||\boldsymbol{W}_k\boldsymbol{S}||_F^2 \leq p_t, \end{aligned} \tag{26}$$

where constraint (C1) guarantees that the lower-bound DER is higher than a threshold $\bar{\tau}$ for the purpose of user anonymity. (C2) denotes that we need to multiplex N streams, as that of anonymity-agnostic MIMO designs. (C3) confines the power budget p_t. Evidently, the difficulty of solving P1 lies in the anonymity requirement in (C1), i.e., making the obtained DER with the precoder $\boldsymbol{W}_k$ higher than the threshold. Also for per-stream SINR to be optimized in the objective function, since the BS may not know the exact channel that the received signal propagates, it is difficult to design a combiner $\boldsymbol{C}$ to equalize the received streams in (C2). In the following, we first construct a link between the precoder and the subscribed DER for handling (C1). Then, we propose an alias-channel based combiner for multiplexing $\min\{N_t, N_r\}$ for handling (C2). Finally, a DER-tunable anonymous precoder is designed.

A. Anonymous Constraint with DER Threshold

Revisiting (16) of a BL Precoder, we write $\text{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i)$ as a quadratic function w.r.t. $\Pr(\zeta_i \leq 0|\mathcal{H}_k)$, given as

$$\begin{aligned} &\text{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i)^2 - [\text{erf}^{-1}(2\Pr(\zeta_i \leq 0|\mathcal{H}_k)-1)]^2 4\sigma^2\text{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i) - \\ &[\text{erf}^{-1}(2\Pr(\zeta_i \leq 0|\mathcal{H}_k)-1)]^2 2\sigma^4 L\text{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i + \boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k) = 0, \end{aligned} \tag{27}$$

Finding the root of the quadratic function of (27) yields

$$\begin{aligned} &\text{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i) = \sigma^2\text{erf}^{-1}(2\Pr(\zeta_i \leq 0|\mathcal{H}_k)-1)\cdot \\ &\left(2\text{erf}^{-1}(2\Pr(\zeta_i \leq 0|\mathcal{H}_k)-1) + \sqrt{4\Big(\text{erf}^{-1}(2\Pr(\zeta_i \leq 0|\mathcal{H}_k)-1)\Big)^2 + 2L\text{tr}(\boldsymbol{\Xi}_i^H\boldsymbol{\Xi}_i + \boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k)}\right), \end{aligned} \tag{28}$$

where the negative root is ignored due to the value of $\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i) \geq 0$. The above result also applies to SL precoder by replacing $\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i)$ by $\sum_{l=1}^{L}(\boldsymbol{v}_i^{(l)})^H\boldsymbol{v}_i^{(l)}$. (28) leads to the following statements in Lemmas 1-4.

Lemma 1: By manipulating the value of $\boldsymbol{V}_i^H\boldsymbol{V}_i$, the probability that, under event $\mathcal{H}_k$, the receiver falsely declares that user i other than user k sends, i.e., $\Pr(\zeta_i \leq 0|\mathcal{H}_k)$, is constrained in-between [0.5, 1]. ■

Proof of Lemma 1 is straightforward. Based on (10) and (14), the expectation of ζ_i becomes 0 if and only if (iif) $\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i) = 0$, which physically denotes that $\Pr(\zeta_i < 0|\mathcal{H}_k) = 0.5$. In other words, the BS finds user i and real sender k as equally probable senders, where user i is thus termed as an alias sender. This can also be explained by our analysis in subsection II-B. When $\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i) = 0$, d_i is reduced to $d_i = ||(\boldsymbol{H}_i\boldsymbol{H}_i^{\dagger} - \boldsymbol{I}_{N_r})\boldsymbol{Z}||_F^2$, which becomes only related to a colored-noise. It is easy to prove that in this case d_i has the same expectation and variance with d_k, and thus the BS is unable to distinguish sender k from user i. Also, for any other value $\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i) > 0$, the value of $\Pr(\zeta_i \leq 0|\mathcal{H}_k)$ locates in-between (0.5, 1]. □

Lemma 1 in fact discusses the achievable DER when one alias sender is constructed (user i in the above case), and now we extend the conclusion into a multi-alias case.

Lemma 2: By introducing M alias senders, the achievable DER is upper-bounder by $\tau = 1 - (1/2)^M$. ■

The proof of Lemma 2 is given as follows. Introduce M ($M = |\mathbb{M}|$, $\mathbb{M} \subseteq \mathbb{K}/k$) aliases and let $\mathrm{tr}(\boldsymbol{V}_i^H\boldsymbol{V}_i) = 0$, $\forall i \in \mathbb{M}$. Then, all the M aliases become equally likely senders, while other users not belonging $\mathbb{M}$ can be detected as false events by the detector. Based on (17), the achievable DER is upper bounded by $\tau = 1 - (1/2)^M$. □

Lemma 3: Given a DER requirement $\bar{\tau}$, the minimum required number of alias senders as

$$M = \lceil \log_{\frac{1}{2}}(1 - \bar{\tau}) \rceil, \forall \tau \in [0, 1), \tag{29}$$

where the operator $\lceil \cdot \rceil$ denotes the roundup function. ■

The proof of Lemma 3 follows the Lemma 2, and is omitted due to the limit of page. An illustration of the required number of alias users is plotted in Fig. 2, demonstrating a stairstep graph. In particular, no alias is needed when $\bar{\tau} = 0$ (no anonymity requirement). This reduces to conventional anonymity-agnostic transceiver design. When $\bar{\tau} = 1$, the required number M

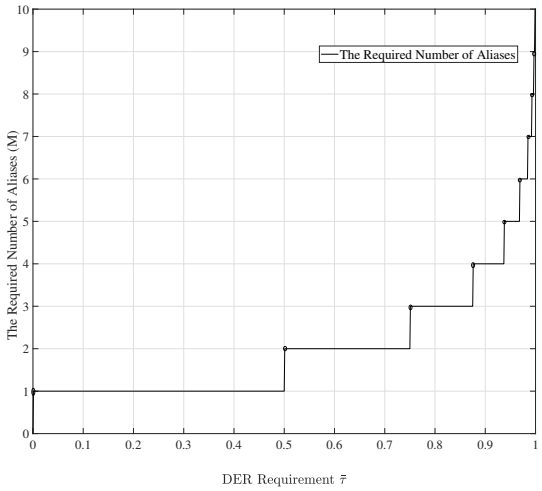

Fig. 2. DER requirement vs. the required number of aliases. Note that the required number of aliases is independent to antenna configuration.

approaches infinity. In other words, $\bar{\tau} = 1$ serves as an upper bound of the achievable DER in practice.

Lemma 4: The required number of aliases only depends on the DER requirement $\bar{\tau}$ but is independent to the precoder or other PHY parameters. Hence, given a subscribed DER, one can first calculate the required number of aliases, and then design its precoder accordingly to satisfy the anonymous constraints. ■

The proof of Lemma 4 follows the Lemma 3 and thus is omitted. Now, we are ready to devise how (C1) is handled for achieving the subscribed DER requirement. With a DER requirement $\bar{\tau}$, we select M users as aliases based on (29), which bounds the achievable DER τ in-between $[0, 1-(1/2)^M]$. Hence, we set constraint $\mathrm{tr}(\boldsymbol{V}_i^H \boldsymbol{V}_i) = 0$ (which is equivalently given as $\boldsymbol{V}_i = \boldsymbol{0}_{N_r \times L}$), for the first $i = 1, ..., M-1$ aliases. While for the M-th alias, it should provide $\Pr(d_M \geq d_k | \mathcal{H}_k) = \frac{1-\bar{\tau}}{(\frac{1}{2})^{M-1}}$, so that the composite DER equals to that of the subscribed DER. In particular, the anonymous constraint for the M-th alias is obtained from (28), by substituting $\Pr(\zeta_M \leq 0 | \mathcal{H}_k) = \Pr(d_M \geq d_k | \mathcal{H}_k) = \frac{1-\bar{\tau}}{(\frac{1}{2})^{M-1}}$ into (28). Finally, recalling $\boldsymbol{V}_i = \boldsymbol{\Xi}_i \boldsymbol{H}_k \boldsymbol{W}_k \boldsymbol{S}$, $\forall i \in \mathbb{M}$, anonymous constraint (C1) can be equivalently transformed into

$$
\begin{aligned}
&(\text{C1a}): M = \lceil \log_{\frac{1}{2}}(1-\bar{\tau}) \rceil, \\
&(\text{C1b}): \boldsymbol{\Xi}_i \boldsymbol{H}_k \boldsymbol{W}_k \boldsymbol{S} = \mathbf{0}_{N_r \times L}, \text{for } i = 1, ..., M-1, \\
&(\text{C1c}): ||\boldsymbol{\Xi}_M \boldsymbol{H}_k \boldsymbol{W}_k \boldsymbol{S}||_F^2 = \sigma^2 p_M \left(2p_M + \sqrt{4p_M^2 + 2L\text{tr}(\boldsymbol{\Xi}_M^H \boldsymbol{\Xi}_M + \boldsymbol{\Xi}_k^H \boldsymbol{\Xi}_k)} \right),
\end{aligned} \tag{30}
$$

where $p_M = \text{erf}^{-1}(2\frac{1-\bar{\tau}}{(\frac{1}{2})^{M-1}} - 1)$. Note that the discussion above directly applies to SL precoder, and thus is not discussed for brevity.

The whole procedure of handling (C1) is summarized in Algorithm 1. For the purpose of illustration, we show a toy example. Assuming a DER requirement $\bar{\tau} = 0.6$ and user k is the real sender, (29) indicates two aliases are needed. Hence, we set constraint $\boldsymbol{\Xi}_1 \boldsymbol{H}_k \boldsymbol{W}_k \boldsymbol{S} = \mathbf{0}_{N_r \times L}$ for the first alias, while the constraint of the second alias is calculated by (C1c) with $p_2 = \text{erf}^{-1}(2\frac{1-\bar{\tau}}{(\frac{1}{2})^{M-1}} - 1)$. As a result, we have $\Pr(d_1 \geq d_k | \mathcal{H}_k) = 0.5$ and $\Pr(d_2 \geq d_k | \mathcal{H}_k) = 0.8$. Finally, the obtained DER is strictly lower-bounded by $1 - 0.5 \times 0.8 = 0.6$, thereby guaranteeing the subscribed anonymity requirement.

Algorithm 1 Alias Senders Generation

Input: DER requirement $\bar{\tau}$.

1: Randomly select M users as alias senders according to (29).

2: Let $\boldsymbol{\Xi}_i \boldsymbol{H}_k \boldsymbol{W}_k \boldsymbol{S}_k = \mathbf{0}_{N_r \times L}$ for the first $M-1$ alias senders, while the anonymous constraint of the last alias is calculated by (30).

Output: The tackable form of anonymous constraint (C1).

B. Alias Channel based Combiner Design

In anonymous communications, the precoder mimics a set of alias channels, and thus the BS may not correctly know the exact channel that the signal comes from, which in particular inhibits the design of combiner at the receiver side. A approach is to apply channel-independent equal-gain post-coder at the receiver, which however makes the equivalent channel of rank-1 [23]. As a result, only single data-stream can be conveyed. Alternatively, [18] proposed a transmit-equalizer, which treats each receive-antenna as an individual receiver, thus always transmitting N_r streams regardless of the value of N_t. Nevertheless, the per-stream receive-performance degrades significantly, as the channel characteristic is not exploited at the receiver side.

As suggested in subsection IV-A, imposing M aliases makes these aliases and the sender k equally likely senders, from the perspective of the BS. Hence, the combiner can be designed based on an "average channel", that has a minimum Euclidean distance to the channels of all the probable senders. The construction of the average channel $\boldsymbol{H}_a$ can be formulated as a least-squares problem

$$P2 \ : \underset{\boldsymbol{H}_a}{\operatorname{argmin}} \ ||\boldsymbol{H}_a - \boldsymbol{H}_k||_F + \sum\nolimits_{i=1}^{M} ||\boldsymbol{H}_a - \boldsymbol{H}_i||_F. \tag{31}$$

As P2 is an unconstrained quadratic programming, it can be directly solved by a standard solver, such as CVX.

Remark 3: P2 in fact finds the barycentre of a high-dimensional space confined by all the probable senders' channels, where the optimal result of P2 can also be directly obtained as $\boldsymbol{H}_a = \frac{\boldsymbol{H}_k + \sum_{i=1}^{M} \boldsymbol{H}_i}{M+1}$. Hence, for typical i.i.d. Rayleigh MIMO channels, we have $\mathrm{rank}(\boldsymbol{H}_a) = \min\{N_r, N_t\}$. ■

Remark 3 essentially means that the alias channel $\boldsymbol{H}_a$ is still of full-rank, without sacrificing the capability of multiplexing. Hence, a combiner obtained from $\boldsymbol{H}_a$ on one hand multiplexes $\min\{N_r, N_t\}$ streams as that of classic anonymity-agnostic precoders, and on the other hand does not reveal the real sender's channel.

Remark 4: It is easy to prove that the obtained average channel $\boldsymbol{H}_a$ has an equivalent distance to all the possible channels. In other words, though the BS may not know which is the correct channel that the received signal propagates, it can construct a channel that has similar spatial characteristics to all the possible channels. Hence, the combiner devised by the average channel provides a near-optimal performance, compared to that devised by the real channel $\boldsymbol{H}_k$. More importantly, the combiner based on the average channel does rely on the recognition of the real channel, thus maintaining the anonymity. ■

Applying singular-value-decomposition (SVD) onto $\boldsymbol{H}_a$ yields

$$\boldsymbol{H}_a = \boldsymbol{U}_a \boldsymbol{\Lambda}_a \boldsymbol{V}_a^H, \tag{32}$$

where $\boldsymbol{U}_a \in \mathbb{C}^{N_r \times N_r}$ and $\boldsymbol{V}_a \in \mathbb{C}^{N_t \times N_t}$ are unitary-matrices. $\boldsymbol{\Lambda}_a \in \mathbb{C}^{N_r \times N_t}$ contains singular values in an descending order on its diagonal. Write $\boldsymbol{U}_a$ in the form of $\boldsymbol{U}_a = [\boldsymbol{U}_a^{(1)}, \boldsymbol{U}_a^{(2)}]$, where $\boldsymbol{U}_a^{(1)} \in \mathbb{C}^{N_r \times \mathrm{rank}(\boldsymbol{H}_a)}$ corresponds to the first $\mathrm{rank}(\boldsymbol{H}_a)$ left singular vectors and provides

an ortho-normal basis for the column space of $\boldsymbol{H}_a$. Hence, the combiner can be accordingly designed as

$$\boldsymbol{C} = (\boldsymbol{U}_a^{(1)})^H, \tag{33}$$

where the row of $\boldsymbol{C}$ contains the $\mathrm{rank}(\boldsymbol{H}_a)$ dominant left singular vectors of $\boldsymbol{H}_a$, and thus demonstrates high gain towards receive-direction.

C. Lower-Bound Anonymity (LBA) Based Precoder

In this subsection, we turn to design anonymous precoder to handle the objective function. In general, the objective function in P1 can be handled by classic semi-definite programming (SDP) with a procedure of semi-definite relaxation, and it requires eigen-decomposition for the optimal result [27]. Instead, we leverage the concept of CI to transform the SINR into a linear form, where handling SINR becomes much easier than that with SDP expression. Briefly speaking, CI based precoders let interference act as a constructive element to push the received signal into constructive regions. Due to an increased distance to the detection threshold of demodulation, CI based precoders [28] [29] provide significant SINR enhancement over the interference mitigation based precoders [19] [21]. Without loss of generality, we use quadrature phase shift keying (QPSK) modulation as an example. Then, the received signal falls into a constructive region if and only if the trigonometry holds

$$|\mathrm{Im}\{\boldsymbol{h}_{kn}\boldsymbol{W}_k^{(l)}\boldsymbol{s}^{(l)}\}| \leq (\mathrm{Re}\{\boldsymbol{h}_{kn}\boldsymbol{W}_k^{(l)}\boldsymbol{s}^{(l)}\} - \tilde{\sigma}\sqrt{\Gamma_n})\tan(\frac{\pi}{X}), \forall n \in N, \forall l \in L, \tag{34}$$

Note that since the CI-based design belongs to the family of SL precoder, the superscript l is introduced for both precoder and transmitted symbol vector. X represents constellation size. The operators $\mathrm{Im}(\cdot)$ and $\mathrm{Re}(\cdot)$ take the real and imaginary parts of a complex variable. We have noise variance $\tilde{\sigma}^2 = \frac{\sigma^2}{\mathrm{rank}(\boldsymbol{H}_b)}$ due to the effect of combiner.

Exactly, $\gamma_n = \tilde{\sigma}\sqrt{\Gamma_n}$ measures the Euclidean distance between the originate and the detection thresholds of the signal constellation of the n-th data stream. Define $\gamma = \min\{\gamma_1, ..., \gamma_N\}$, serving as the lower bound of the SINRs of the N streams. Hence, maximizing the lower bound of SINRs in P1 is equivalent to maximizing γ. Now, constraints (C1) and (C2) have been transformed into tractable forms, and thus P1 is re-formulated as

$$
\begin{aligned}
P3 \ &: \underset{\boldsymbol{W}_k^{(l)}, \boldsymbol{C}}{\text{argmax}} \ \gamma, \\
\text{s.t. } &(\text{C1a}) : M = \lceil \log_{\frac{1}{2}}(1-\bar{\tau}) \rceil, \\
&(\text{C1b}) : \boldsymbol{\Xi}_i \boldsymbol{H}_k \boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)} = \boldsymbol{0}_{N_r \times 1}, \text{for } i = 1, ..., M-1, \\
&(\text{C1c}) : ||\boldsymbol{\Xi}_M \boldsymbol{H}_k \boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)}||_F^2 = \frac{\sigma^2 p_M}{L} \left(2p_M + \sqrt{4p_M^2 + 2L\text{tr}(\boldsymbol{\Xi}_M^H \boldsymbol{\Xi}_M + \boldsymbol{\Xi}_k^H \boldsymbol{\Xi}_k)} \right), \\
&(\text{C2}) : |\text{Im}\{\boldsymbol{h}_{kn} \boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)}\}| \leq (\text{Re}\{\boldsymbol{h}_{kn} \boldsymbol{W}_k^{(l)} \boldsymbol{s}_k^{(l)}\} - \gamma) \tan(\frac{\pi}{X}), \forall n, \\
&(\text{C3}) : ||\boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)}||_F^2 \leq p_t / L,
\end{aligned}
\tag{35}
$$

where (C1a)-(C1c) denote anonymity constraints based on (30), while (C2) relates the per-stream receive-SINR with the objective function. Now, the last difficulty of solving P3 lies in the non-convex constraint (C1c). Hence, we relax constraint (C1c) into a second order cone (SOC) constraint

$$
(\tilde{\text{C1c}}) : ||\boldsymbol{\Xi}_M \boldsymbol{H}_k \boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)}||_F \leq \left(\frac{\sigma^2 p_M}{L} \left(2p_M + \sqrt{4p_M^2 + 2L\text{tr}(\boldsymbol{\Xi}_M^H \boldsymbol{\Xi}_M + \boldsymbol{\Xi}_k^H \boldsymbol{\Xi}_k)} \right) \right)^{\frac{1}{2}}. \tag{36}
$$

Remark 5: A smaller value of $||\boldsymbol{\Xi}_M \boldsymbol{H}_k \boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)}||_F$ makes the value of $\Pr(\zeta_M \leq 0 | \mathcal{H}_k)$ decrease, thereby increasing the value of DER. In other words, by the relaxed constraint $(\tilde{\text{C1c}})$ in (36), the obtained DER is in fact lower bounded by the original result solved with (C1c), leading to better anonymity performance. ■

Replacing (C1c) by $(\tilde{\text{C1c}})$, now P3 maximizes a linear objective function, subject to linear constraints as well as SOC constraints. Hence, P3 can be readily solved by CVX, and the whole algorithm of the anonymous LBA transceiver design is summarized in Algorithm 2.

Algorithm 2 Anonymous LBA Transceiver

Input: Power budget p_t, CSI, and DER requirement $\bar{\tau}$.

1: Call Algorithm 1 to calculate the number of aliases M.

2: Formulate anonymous constraints (C1a), (C1b), and $(\tilde{\text{C1c}})$.

3: Solve optimization P2 to obtain the alias channel $\boldsymbol{H}_a$.

4: Do SVD of the average channel $\boldsymbol{H}_a$, and calculate the anonymous combiner $\boldsymbol{C}$ by (33).

5: Solve optimization P3 to obtain the optimal anonymous precoder.

Output: Optimal anonymous combiner and precoder results.

The proposed LBA transceiver design is able to multiplex $N = \min\{N_r, N_t\}$ streams, while providing subscribed DER performance. However, the achievable receive-reliability may not be well guaranteed when the number of receive-antenna becomes larger than that of the transmit-antenna, as discussed in the following Remark 6.

Remark 6: Assume that there are M aliases. $(\mathrm{C1b})$ means that the received signal, i.e., $\boldsymbol{H}_k\boldsymbol{W}_k\boldsymbol{S}$ should lie in the orthogonal space of $\boldsymbol{\Xi}_i$, $\forall i = 1, ..., M-1$, and $(\tilde{\mathrm{C1c}})$ denotes that the received signal should lie in the space that is close to the orthogonal space of $\boldsymbol{\Xi}_M$ (as the right hand of $(\tilde{\mathrm{C1c}})$ is a small valued variable). Hence, with the increase of N_r, the length of the orthogonal basis of $\boldsymbol{\Xi}_i$ increases, $\forall i \in M$. It further reduces the DoF of precoder design and leads to degraded receive-reliability performance. ■

V. Diversity-Multiplexing Trade-off in Anonymous Precoding

MIMO can boost the reliability of reception for a given data rate (providing diversity gain) or boost the data rate for a given reliability of reception (providing multiplexing gain). Maximizing one type of gain may not necessarily maximize the other [30]. In Section-IV, we have demonstrated a LBA design to multiplex $\min\{N_r, N_t\}$ streams under a flexible anonymity constraint. This high multiplexing gain comes at the price of sacrificing diversity. By contrast, the work in [23] implements a high diversity oriented anonymous precoder, where only one stream is conveyed through N_rN_t channels at the cost of low multiplexing performance. In a nutshell, the existing anonymous work focuses on designing schemes to extract either maximal diversity gain [18] or maximal spatial multiplexing gain [23] [3].

In this section we target at better trading-off the diversity and multiplexing for anonymous communications. Defining SNR as the average SNR per receive-antenna, a scheme is said to have an asymptotic diversity gain g_d if the average error probability PE decays like SNR^{-g_d}, mathematically given as $g_d = -\lim_{SNR\to\infty} \frac{\log PE}{\log SNR}$ [30] [31]. Considering arbitrary N ($N \leq \min\{N_r, N_t\}$) multiplexing streams, we have

$$\max_{1\leq n\leq N} \mathrm{PE}_n \leq \mathrm{PE} \leq \sum_{n=1}^{N} \mathrm{PE}_n, \tag{37}$$

[3]The design in [23] multiplexes N_r streams when $N_r > N_t$, where combiner is not considered at the receiver side. Hence, this comes at low reliability performance, especially when N_r is large.

where PE_n denotes the error probability of the n-th steam, and it is calculated as $\mathrm{PE}_n = \mathrm{PO}_n \mathrm{Pr}_n(\mathrm{error}|\mathrm{outage}) + \mathrm{Pr}_n(\mathrm{error}, \mathrm{no\ outage})$ [32]. Outage probability PO_n represents the probability that the mutual information between the input and the output of the channel is smaller than the data rate, while $\mathrm{Pr}_n(\mathrm{error}, \mathrm{no\ outage})$ denotes the error probability averaged over the no-outage channel on the n-th channel [32] [33]. Hence, the per-stream error probability PE_n is bounded by

$$\mathrm{PO}_n \leq \mathrm{PE}_n \leq \mathrm{PO}_n + \mathrm{Pr}_n(\mathrm{error}, \mathrm{no\ outage}). \tag{38}$$

For the considered scenario, as CSI is available at the transmitter side, there is no outage because the user can compute the instantaneous channel capacity and adapt the data rate accordingly. On the other hand, the term $\mathrm{Pr}_n(\mathrm{error}, \mathrm{no\ outage})$ is upper bounded by the pairwise error probability (PEP) averaged over the no-outage channel [33], i.e., PEP_n. Recalling (24), the n-th stream in the l-th symbol vector is given as

$$r_n^{(l)} = \boldsymbol{h}_{kn} \sum_{n=1}^{N} (\boldsymbol{w}_{kn} s_n^{(l)}) + \tilde{z}_n, \tag{39}$$

where we have $\boldsymbol{s}^{(l)} = [s_1^{(l)}, ..., s_N^{(l)}]^T$ with underscript denoting the index of the stream. Thus, PEP of the n-th stream is bounded as

$$\begin{aligned} \mathrm{PEP}_n &\overset{\mathrm{d}}{\leq} \mathrm{Pr}\left(|\frac{\tilde{z}_n}{\boldsymbol{h}_{kn}\boldsymbol{w}_{kn}}| > \frac{d_{\min}}{2} \right) \\ &= \mathrm{Pr}\left(|\tilde{z}_n| > \frac{|\boldsymbol{h}_{kn}\boldsymbol{w}_{kn}| d_{\min}}{2} \right), \end{aligned} \tag{40}$$

where we have $\tilde{\sigma}^2 = \frac{\sigma^2}{N}$. The step "d" is due to under the provision of (34), the intra-stream interference, i.e., $\boldsymbol{h}_{kn} \sum_{n'=1, n' \neq n}^{N} (\boldsymbol{w}_{kn'} s_{n'}^{(l)})$, contributes constructively. Hence, the PEP is upper-bound by that achieved without the constructive interference, i.e., $\boldsymbol{h}_{kn}\boldsymbol{w}_{kn} s_n^{(l)} + \tilde{z}_n, \forall n$. As the amplitude term $|\tilde{z}_n|$ follows Rayleigh distribution, (40) is further given as

$$\begin{aligned} \mathrm{PEP}_n &\leq \int_{|\boldsymbol{h}_{kn}\boldsymbol{w}_{kn}| d_{\min}/2}^{\infty} 2x \exp(-x^2) \mathrm{d}x \\ &= \exp(\frac{|\boldsymbol{h}_{kn}\boldsymbol{w}_{kn}|^2 d_{\min}^2}{4\tilde{\sigma}^2}), \end{aligned} \tag{41}$$

where $d_{\min}$ is related to the signal demodulation procedure. For example, $d_{\min} = \sqrt{2}$ for QPSK and $d_{\min} = 1/\sqrt{\frac{2^X-1}{6}}$ for 2^X-order QAM.

As suggested by (37)-(41), MIMO diversity performance can be guaranteed by suppressing the upper-bound of communication error probability. This is equivalent to maintaining the value of $|\boldsymbol{h}_{kn}\boldsymbol{w}_{kn}|^2$ for each stream, which is directly equivalent to guaranteeing the minimum SINRs value of all the multiplexing streams above a threshold. Hence, in the following, our target is to find a reasonable number of multiplexing streams for optimizing MIMO multiplexing gain, under anonymity and diversity constraints. With arbitrary number of multiplexing streams N ($N = |\mathbb{N}|$), the optimization is formulated in the form of

$$\begin{aligned} P4 \ &: \underset{\boldsymbol{W}_k^{(l)}, \boldsymbol{C}}{\text{argmax}} \ N, \\ \text{s.t. } &(\text{C4}) : \tau(\boldsymbol{W}_k^{(l)}) \geq \bar{\tau}, \ (\text{C5}) : \Gamma_n \geq \bar{\Gamma}, \forall n \in \mathbb{N}, \\ &(\text{C6}) : ||\boldsymbol{W}_k^{(l)} \boldsymbol{s}^{(l)}||^2 \leq p_t/L, \ (\text{C7}) : N \leq \min\{N_r, N_t\}, \end{aligned} \tag{42}$$

where (C5) denotes that the per stream SINR should be higher than a target $\bar{\Gamma}$, with the consideration of diversity gain performance. Revisiting (32), split $\boldsymbol{U}_a$ in the form of left singular vector, i.e., $\boldsymbol{U}_a = [\boldsymbol{u}_b^{(1)}, \boldsymbol{u}_b^{(2)}, ..., \boldsymbol{u}_b^{(N_r)}]$. With N streams multiplexed by the system, the average-channel based combiner can be re-calculated as

$$\boldsymbol{C} = [\boldsymbol{u}_b^{(1)}, ..., \boldsymbol{u}_b^{(N)}]^H, \tag{43}$$

which abstracts N streams from the N_r-dimension received signal. Also, the anonymity constraint (C1) can be simplified into constraints (C1a)-($\tilde{\text{C1c}}$) as we presented in Section III, while (C5) can be handled by the CI constraint in a different form of

$$|\text{Im}\{\boldsymbol{h}_{k,n}\boldsymbol{W}_k^{(l)}\boldsymbol{s}^{(l)}\}| \leq (\text{Re}\{\boldsymbol{h}_{k,n}\boldsymbol{W}_k^{(l)}\boldsymbol{s}^{(l)}\} - \tilde{\sigma}\sqrt{\bar{\Gamma}_n})\tan(\frac{\pi}{X}), \forall n \in \mathbb{N}, \tag{44}$$

Note that the key difference to (34) is that per-stream SINR requirement $\bar{\Gamma}$ is embedded for guaranteeing diversity performance, instead of being a variable to be optimized. Evidently, maximizing N is equivalent to maximizing the number of constraints (the cardinality of $\mathbb{N}$) in (C2) while checking the feasibility of the optimization problem, given as

$$
\begin{aligned}
P5 \ : & \underset{\boldsymbol{W}_k^{(l)}}{\operatorname{argmax}} \ |\mathbb{N}|, \\
\text{s.t. } & (\mathrm{C4}) : (30) \text{ and } (36), \ (\mathrm{C5}) : (44), \forall n \in \mathbb{N}, \text{and } (\mathrm{C7}),
\end{aligned} \tag{45}
$$

Now, we are able to devise the diversity-multiplexing-tradeoff lower-bound anonymity (DM-LBA) transceiver in Algorithm 3.

Algorithm 3 The DM-LBA Transceiver

Input: CSI, power budget p_{t}, SINR threshold requirement $\bar{\Gamma}$.

1: Call Algorithm 1 to calculate the number of aliases M.

2: Call Algorithm 2 to obtain the average channel.

3: Initialize the number of multiplexing streams N.

4: **repeat**

5: Calculate the combiner by (43), and check the feasibility of P5.

6: Enlarge the cardinality of $\mathbb{N}$ (multiplex more streams) if P5 is feasible and vice versa, i.e., by bisection or Dinkelbach search algorithm.

7: **until** Converge to the maximum number of multiplexing streams.

Output: Optimal anonymous transceiver design.

Note that given a stringent DER performance, more users are needed to act as aliases. It reduces the DoF of the precoder design, thereby yielding a small value of N. In an extreme case, there might be no feasible solution, even only one steam is conveyed from the user. Hence, one can properly reduce the anonymity or per stream receive-SINR quality requirement, so that the DoF can be relaxed to find a feasible solution.

VI. Simulation Results

We present the Monte-Carlo simulation results in this section. The power budget is normalized to 1 and QPSK is employed for modulation. Assume that each block has 50 symbols. There are $K = 5$ senders, and the communication user at each block is randomly generated. Rayleigh block fading channel is considered. The anonymity threshold is set to $\bar{\tau} = 0.5$ and 0.3 for the LBA design. For the DM-LBA transceiver, its anonymity threshold is set to $\bar{\tau} = 0.3$, and the SINR requirement to SNR ($\bar{\Gamma} = \frac{p_t}{\sigma^2}$), or to 5 dB higher than SNR ($\bar{\Gamma} = \frac{p_t}{\sigma^2} + 5$ (in dB)). The antenna configuration is set as $N_t = 10$ and $N_r = 11$, everywhere except in Figs. 5 and 6. The following anonymous and anonymity-agnostic precoders are selected as benchmarks: 1) The

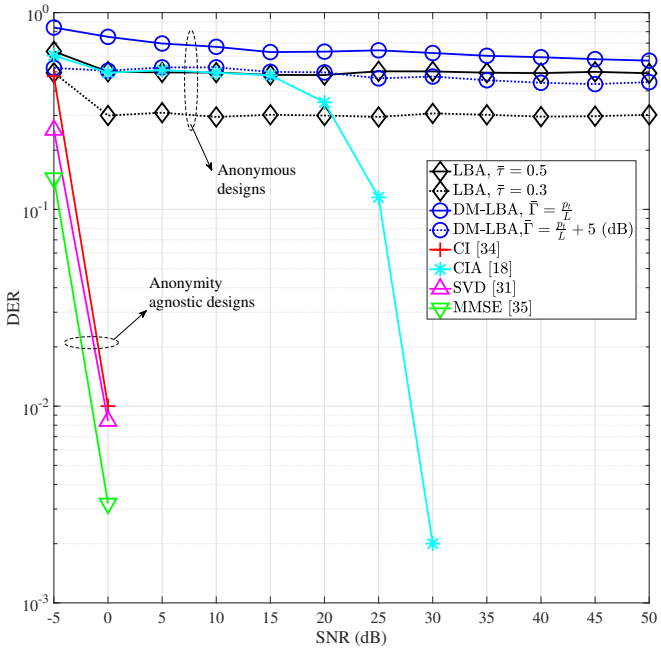

Fig. 3. The impact of transmit-SNR on the DER performance.

CIA precoder [18], where the anonymous precoder always multiplexes N_r streams, under an empirical anonymous constraint. 2) The CI precoder [34], where the precoder is designed based on the signal constellation of modulation. 4) The SVD precoder [31], where the precoder and combiner are designed based on the SVD of the sender's channel. 5) The MMSE precoder [35].

In Fig. 3, the impact of transmit-SNR on the DER performance is demonstrated. It is observed that the proposed LBA and DM-LBA transceivers always guarantee the subscribed anonymity threshold. With a higher threshold, such as $\bar{\tau} = 0.5$, the obtained DER is strictly higher than that with $\bar{\tau} = 0.3$. Also, the achieved DER of the LBA and DM-LBA is slightly higher than the anonymity thresholds. This is because they set the anonymity threshold as a lower bound, and the resulted DER may not necessarily be equal to the threshold. In particular, as the DM-LBA transceiver aims at finding a reasonable number of multiplexing streams, it may not use full transmission power. Hence, this equivalently reduces the transmit-SNR, and lets the achieved DER always higher than that of the LBA transceiver. For the the benchmarks, the anonymous CIA precoder only sets an empirical anonymous constraint to scramble the BS's detection, and fails to provide anonymity with 20 dB or higher SNRs. In particular, the BS can perfectly reveal

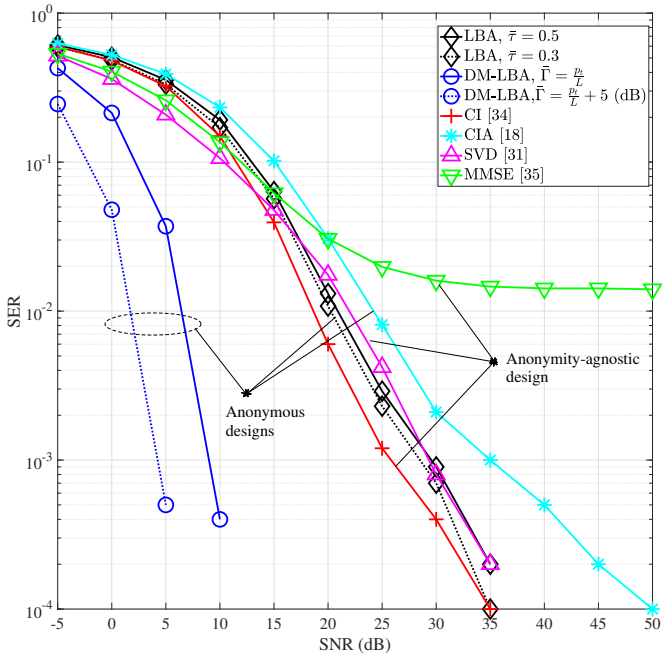

Fig. 4. The impact of transmit-SNR on the SER performance.

the real sender with 5 dB or higher SNRs if anonymity-agnostic precoders are applied at the users, which verifies our analysis in Remark 2.

In Fig. 4, the impact of transmit-SNR on the SER performance is demonstrated. As the proposed LBA and DM-LBA transceivers can provide fully-tunable DER performance and receive signals aided by the alias channel based combiner, the achieved SER performance is much enhanced over the anonymous CIA design. In particular, since the DM-LBA transceiver adaptively finds a reasonable number of multiplexing streams, it achieves the best SER performance among all designs, and even outperforms the anonymity-agnostic designs. For the LBA transceiver, it always multiplexes $\min\{N_r, N_t\}$ streams, and thus the obtained SER is inferior to the DM-LBA transceiver. Nevertheless, it still obtain 2-5 dB SNR gain over the anonymous CIA precoder, which tries to multiplex N_r streams without the aid of a combiner and thus the DoF of its precoder design is overly constrained. Finally, it shows that with a stricter anonymous threshold (such as $\bar{\tau} = 0.5$ for LBA) or lower receive-quality (such as threshold $\bar{\Gamma} = \frac{p_t}{L\sigma^2}$ for DM-LBA), the obtained SER performance reduces in order to satisfy the anonymity or receive-quality requirement. The trade-off of these metrics is further demonstrated in Fig. 7.

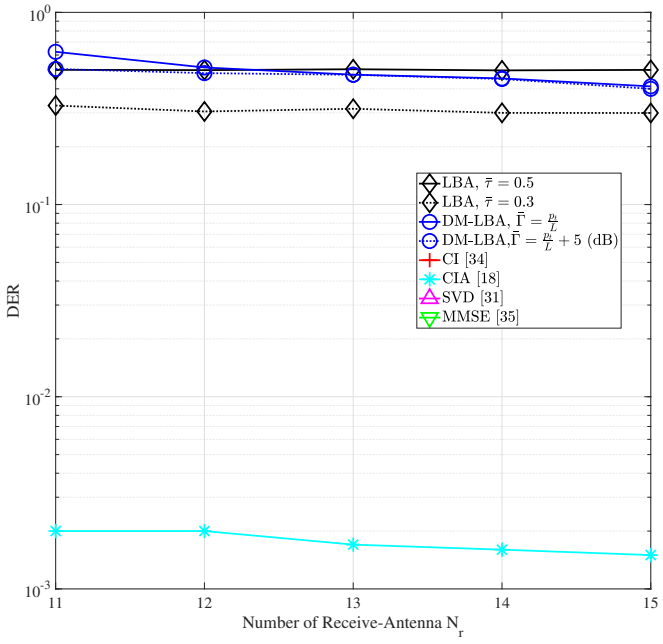

Fig. 5. The impact of number of receive-antennas N_r on the DER performance. $N_t = 10$ and transmit-SNR is set to be 30 dB. Note that the DER of the anonymity-agnostic designs equals to 0, which are not visible in logarithmic coordinates.

In Fig. 5, the impact of the number of receive-antennas is demonstrated. While a larger number of receive-antennas enhances the BS's detection capability, the proposed LBA and DM-LBA transceivers still guarantee the required anonymity level. As a comparison, the DER of the anonymous CIA precoders is slightly reduced with more receive-antennas. In particular, the anonymity-agnostic designs can not provide anonymity for users, and their associated DER equals to 0, which are not visible in logarithmic coordinates.

Fig. 6 verifies Remark 6 that, with a larger number of receive-antennas, it becomes difficult to satisfy the anonymous constraint while providing a high SER performance. In order to satisfy the anonymous constraints, the DoF of the anonymous precoder design is further constrained. As a result, the SER performance of the anonymous LBA and the benchmark CIA is reduced with increase of number of receive-antennas. However, the DM-LBA can adaptively adjust the number of multiplexing streams taking system setup into consideration. It is observed that the DM-LBA still provides high SER performance, and its SER equals to 0, which is not visible in logarithmic coordinates. In other words, when the number of receive-antenna is much higher than that of transmit-antennas, the DM-LBA can well trade-off the anonymity and communication

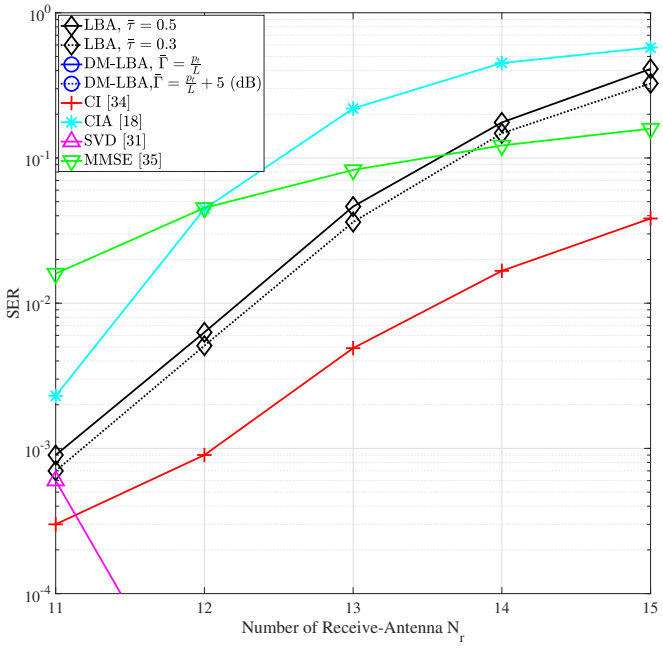

Fig. 6. The impact of number of receive-antennas N_r on the SER performance. $N_t = 10$ and transmit-SNR is set to be 30 dB. Note that the SER of the DM-LBA, SVD designs equals to 0, which are not visible in logarithmic coordinates.

diversity, and multiplexing performance. Also, the SER of the CI and MMSE designs increase with the increase of N_r, as they try to multiplex N_r streams, where a solution is to multiplex less streams and receive signal with a combiner, in the style of SVD transceiver.

Figs. 5-6 have verified the DER (anonymity) and SER (diversity) performance of the proposed designs, and now we present their multiplexing performance with different numbers of antennas. For guaranteeing the subscribed anonymity and receive-quality requirement, the DM-LBA adaptively reduces its number of multiplexing streams in Fig. 7(b), and thus maintains a high SER (diversity) performance in 7(a). In a different manner, the LBA transceiver always multiplex $\min\{N_t, N_t\}$ streams, but its diversity performance is in fact inferior to that of DM-LBA transceiver. It is because with more receive-antennas, it becomes difficult to satisfy the anonymity constraint, and thus always multiplexing $\min\{N_t, N_t\}$ streams limits the DoFs of precoder and leads to degraded SER performance. Also, as the CI, CIA, and MMSE always multiplex N_r streams, their throughput performance degrade significantly when N_r increases.

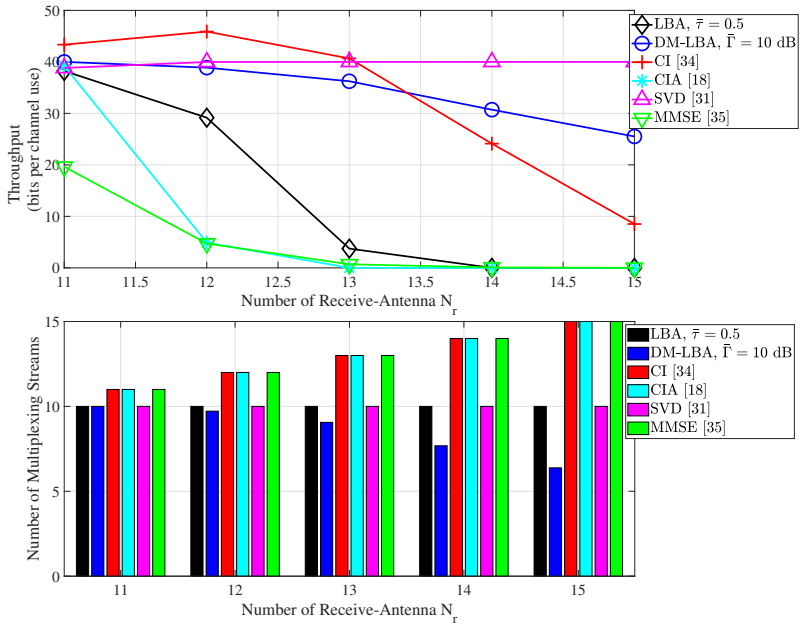

Fig. 7. The impact of number of receive-antennas N_r on the multiplexing performance. $N_t = 10$ and transmit-SNR is set to be 30 dB.

VII. Conclusion

In this work, we have investigated the anonymous joint transceiver design with fully tunable DER performance. We first quantify the DER performance of generic BL and SL precoders. By providing the closed-form of DER as a function of precoder, blocklength, and noise status, we were able to set exact anonymity constraint for guaranteeing a certain DER performance. Aided by an alias channel based combiner, an anonymous LBA precoder has been introduced to multiplex $\min\{N_r, N_t\}$ streams without loss of the sender's anonymity. Then, to well tradeoff the anonymity, diversity and multiplexing performance, a so-called DM-LBA anonymous transceiver has been further proposed, which flexibly adjusts the number of multiplexing streams with the consideration of receive-reliability. Simulation demonstrated that the proposed anonymous transceiver designs can provide superior anonymity performance over the existing anonymous and anonymity-agnostic precoders, while at the same time achieves close multiplexing and diversity performance to the classic anonymity-agnostic designs.

APPENDIX

Since $\mathbb{E}\{||\boldsymbol{\Xi}_k\boldsymbol{Z}||^2\} = \mathbb{E}\{\mathrm{tr}(\boldsymbol{\Xi}_k\boldsymbol{Z}\boldsymbol{Z}^H\boldsymbol{\Xi}_k^H)\} = \mathrm{tr}(\mathbb{E}\{\boldsymbol{\Xi}_k\boldsymbol{Z}\boldsymbol{Z}^H\boldsymbol{\Xi}_k^H\}) = \mathrm{tr}(\mathbb{E}\{\boldsymbol{Z}\boldsymbol{Z}^H\}\boldsymbol{\Xi}_k\boldsymbol{\Xi}_k^H) = \sigma^2 L\{\mathrm{tr}(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k)\}$. On the other hand, we use the moment generating function (MGF) to calculate the variance. Let $\boldsymbol{C}(x) = \boldsymbol{I}_{N_r} - 2x\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma} = \sigma^2 L\boldsymbol{I}_{N_r}$. Since $\mathbb{E}\{\boldsymbol{Z}\} = \boldsymbol{0}$, the MGF of $\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{\Xi}^H\boldsymbol{\Xi}\boldsymbol{Z})$ is written as $M_{\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{\Xi}^H\boldsymbol{\Xi}\boldsymbol{Z})}(x) = |\boldsymbol{C}|^{-\frac{1}{2}}$. We further let $k(x) = \ln(M_{\mathrm{tr}(\boldsymbol{Z}^H\boldsymbol{\Xi}^H\boldsymbol{\Xi}\boldsymbol{Z})}(x)) = -\frac{1}{2}\ln|C|$, where the second-order derivative of $k(x)$ is calculated as $k^{''}(x) = \frac{1}{2}\frac{1}{|\boldsymbol{C}|^2}[\frac{\boldsymbol{d}|\boldsymbol{C}|}{\boldsymbol{d}x}]^2 - \frac{1}{2}\frac{1}{|\boldsymbol{C}|}\frac{\boldsymbol{d}^2|\boldsymbol{C}|}{\boldsymbol{d}x^2}$. Substituting the value of $|\boldsymbol{C}||_{x=0}$, $\frac{\mathrm{d}|\boldsymbol{C}|}{\mathrm{d}x}|_{x=0}$ and $\frac{\mathrm{d}^2|\boldsymbol{C}|}{\mathrm{d}x^2}|_{x=0}$ into $k^{''}(x)$, we have $k^{''}(0) = \mathrm{tr}\Big(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\Big) = L\sigma^4\mathrm{tr}\Big(\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\boldsymbol{\Xi}_k^H\boldsymbol{\Xi}_k\Big)$.

REFERENCES

[1] X. Chen, D. Ng, and H. Chen, "A survey on multiple-antenna techniques for physical layer security, *IEEE Commun. Survey & Tut.*, vol. 19, no. 2, pp. 1027–1053, Nov. 2016.

[2] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: technical challenges, recent advances, and future trends," *Proc. of IEEE,* vol. 104, no. 9, pp. 1727-1765, Sept. 2016.

[3] Z. Wei, C. Masouros, F. Liu, S. Chatzinotas, and B. Ottersten, "Energy and cost-efficient physical layer security in the era of IoT: the role of interference," *IEEE Commun. Mag.*, vol. 58, issue. 4, pp. 81-87, Apr. 2020.

[4] L. Dong, Z. Han, A. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. on Signal Process.*, vol. 58 , no. 3, pp. 1875-1888, Mar. 2010.

[5] S. Luo, J. Li, and A. Petropulu, "Uncoordinated cooperative jamming for secrecy communications," *IEEE Trans. Inf. Forensics Secur.,* vol. 8, no. 7, pp. 1081-1090, Jul. 2013.

[6] Y. Liu and A. Petropulu, "Destination assisted cooperative jamming for wireless physical layer security," *IEEE Trans. Inf. Forensics Secur.*, vol. 8, no. 4, pp. 682-694, Apr. 2013.

[7] V. Kumar, J. J. Park, and K. Bian, "PHY-layer authentication using duobinary signaling for spectrum enforcement, *IEEE Trans. Inf. Foren. Sec.,* vol. 11, no. 5, pp. 1027-1038, May 2016.

[8] M. V. Jamali and H. Mahdavifar, "Covert millimeter-wave communication: design strategies and performance analysis, *IEEE Trans. Wireless Commun.,* vol. 21, no. 6, pp. 3691-3704, Jun. 2022.

[9] T. Taleb, K. Samdanis, B Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: a survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 3, pp. 1657-1681, 2017.

[10] M. Bloch et al., "An overview of information-theoretic security and privacy: metrics, limits and applications," *IEEE J. Sel. Inf. Theory,* vol. 2, no. 1, pp. 1-22, Mar. 2021.

[11] C. Dwork, "Differential privacy," in Proc. *Int. Colloquium Automata Lang. Program.*, Venice, Italy, Jul. 2006.

[12] I. Issa, A. B. Wagner, and S. Kamath, "An operational approach to information leakage ," *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1625-1657, Dec. 2019.

[13] J. Liao, O. Kosut, L. Sankar, and F. P. Calmon, "Tunable measures for information leakage and applications to privacy-utility tradeoffs," *IEEE Trans. Inf. Theory*, vol. 65, no. 12, pp. 8043-8066, Dec. 2019.

[14] M. Diaz, H. Wang, F. P. Calmon, and L. Sankar, "On the robustness of information -theoretic privacy measures and mechanisms," *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 1949-1978, Apr. 2020.

[15] Z. Wei, C. Masouros, H. V. Poor, A. P. Petropulu, and L. Hanzo, "Physical layer anonymous precoding: the path to privacy-preserving communications," *IEEE Wireless Commun.,* vol. 29, no. 2, pp. 154-160, Apr. 2022.

[16] K. Emura et al., "Secure and anonymous communication technique: formal model and its prototype implementation," *IEEE Trans. Emerging Topics Comput.*, vol. 4, no. 1, pp. 88–101, Mar. 2016.

[17] K. Sakai et al., "On anonymous routing in delay tolerant networks," *IEEE Trans. Mobile Comput.*, vol. 18, no. 12, pp. 2926–2940, Dec. 2019.

[18] Z. Wei, F. Liu, and C. Masouros, "Fundamentals of physical layer anonymous communications: sender detection and anonymous precoding," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 64-79, Jan. 2022.

[19] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Sig. Proc.*, vol. 10, no. 3, pp. 501-513, Apr. 2016.

[20] Z. Wei, C. Masouros, K. Wong, and X. Kang, "Multi-cell interference exploitation: a new dimension in cell coordination", *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 1303-1312, Oct. 2019.

[21] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint RX-TX beamforming design for multi carrier MIMO channel: a unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[22] Q. Xu, P. Ren, A. L. Swindlehurst, "Rethinking secure precoding via interference exploitation: a smart eavesdropper perspective," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 585-600, Aug. 2020.

[23] Z. Wei, C. Masouros, P. Wang, X. Zhu, J. Wang, and A. P. Petropulu, "Physical layer anonymous precoding design: from the perspective of anonymity entropy," *IEEE J. Sel. Area Commun.*, vol. 40, no. 4, Nov. 2022.

[24] R. Zhang and Y. C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networksm," *IEEE J. Sel. Top. Signal Process.,* vol. 2, no. 1, pp. 88-102, Feb. 2008.

[25] Y. C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing throughput trade-off for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[26] E. Axell, G. Leus, E. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio," *IEEE Sig. Proc. Mag.*, vol. 57, no. 6, pp. 101-116, May 2012.

[27] Y. Huang and D. P. Palomar, "Rank-constrained separable semi-definite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.

[28] C. Masouros, "Correlation rotation linear precoding for MIMO broadcast communications," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 252–262, Jan. 2011.

[29] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization" *IEEE Trans. Sig. Proc.*, vol. 63, no. 14, pp. 3668-3680, Jul 2015.

[30] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073-1096, May 2003.

[31] D. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1859-1874, Sept. 2004.

[32] A. Lim and K. N. Lau, "On the fundamental tradeoff of spatial diversity and spatial multiplexing of MISO/SIMO links with imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 110-117, Jan. 2008.

[33] L. G. Ordonez, A. P-Zamora, and J. R. Fononllosa, "Diversity and multiplexing tradeoff of spatial multiplexing MIMO systems with CSI," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2959-2975, Jul. 2008.

[34] A. Li and C. Masouros, "Interference exploitation precoding made practical: optimal closed-form solutions for PSK modulations," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7661-7676, Nov. 2018.

[35] A. Wiesel, Y. C. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses, " *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409-4418, Sep. 2008.