On the Diversity-Multiplexing Tradeoff of Secret-Key Agreement over Multiple-Antenna Channels

Marwen Zorgui1, Zouheir Rezki1, Basel Alomair2 and Mohamed-Slim Alouini1

1 Computer, Electrical, and Mathematical Sciences and Engineering (CEMSE) Division
King Abdullah University of Science and Technology (KAUST), Saudi Arabia

2 The National Center for Cybersecurity Technology (C4C)
King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia
{marwen.zorgui,zouheir.rezki,slim.alouini}@kaust.edu.sa, alomair@kacst.edu.sa

Abstract—We consider secret-key agreement with public discussion over Rayleigh fading quasi-static channels. First, the secret-key diversity gain and the secret-key multiplexing gain are defined. Then, the secret-key diversity multiplexing tradeoff (DMT) is established. The eavesdropper is shown to “steal” only transmit antennas. We show that likewise the DMT without secrecy constraint, the secret-key DMT is the same either with or without full channel state information (CSI) at the transmitter (CSI-T). This insensitivity of secret-key DMT toward CSI-T highlights a fundamental difference between secret-key agreement and the wiretap channel whose secret DMT depends crucially on CSI-T. Several secret-key DMT-achieving schemes are presented in case of full CSI-T.

Index Terms—Secret-key agreement, outage probability, high-power, diversity multiplexing tradeoff.

I. INTRODUCTION

As the need for data transmission by mobile devices is increasing, privacy concerns and secrecy issues in wireless communications are becoming one of the key parameters and primary specifications of systems design. Traditionally, protection of the transmitted data relies on public-key cryptography, secret-key distribution and symmetric encryption, which is deemed secure based on the assumption of limited computational abilities of a wiretapper. Such techniques are implemented in the higher layers of the protocol stack with little or absent awareness of the nature of the physical medium.

As opposed to this paradigm, information-theoretic security emerged as a solution that supports physical layer security with no computational restrictions placed on the eavesdropper. Its premise is to harness the inherent and abundant randomness in the physical medium to guarantee some secrecy by means of appropriate channel coding techniques.

The notion of perfect secrecy was first introduced by Shannon [1]. Later, Wyner [2] and Csiszár and Körner [3] set the theoretical basis for information-theoretic secrecy by proving in seminal papers the existence of channel coding guaranteeing not only robustness to transmission errors but also a desired level of data confidentiality.

One of the promising applications of physical layer security is secret-key agreement in which the objective of transmitting reliably secret information to the destination can be relaxed to distilling a secret-key shared between the legitimate parties. The notion of secret sharing is formalized in [4] based on the concept of common randomness between the source and the destination. In the same context, [4] suggests two different system models, called the “source model with wiretapper” (SW) and the “channel model with wiretapper” (CW). The SW model represents a situation in which three parties, two legitimate users and an eavesdropper, observe the realizations of a discrete memoryless source, which is assumed to be outside the control of all parties. The CW model differs in that the source is controlled by one of the legitimate parties, similarly to the basic wiretap channel model [2] with an additional public feedback channel. Extension of both models to the case of secret sharing among multiple users, with the possibility of some terminals acting as helpers has been investigated in [5], [6]. The CW model for secret sharing with conditionally independent destination and eavesdropper channels and continuous channel alphabets is considered in [7]. The secret key capacity is derived. Based on this result, the secret key capacity of the ergodic fast-fading multiple-input multiple-output (MIMO) wiretap channel is deduced. It is shown that for such a channel, the source does not require the instantaneous channel state information (CSI) to the legitimate receiver and the eavesdropper. For quasi-static wireless channels, the secret-key capacity is zero in the absence of CSI [8]. In [8], with full CSI knowledge at all the terminals, closed-form expressions for the low-power and high-power regimes are established.

The use of multiple antennas in fading wireless channels not only increases robustness but also transmission rates. One of the important measures that simultaneously investigates both type of gains is the diversity-multiplexing tradeoff (DMT), introduced in [9].

The DMT is a high power performance analysis that characterizes the fundamental tradeoff between the diversity gain and the multiplexing gain. The diversity gain describes the decay rate of the probability of error, and the multiplexing gain represents the rate of increase of the transmission rate.
in the high power regime. The DMT analysis and the outage probability are closely related as it is shown that the outage event generally dominates the probability of error in the high-power regime. The MIMO wiretap channel is investigated in [10] from a DMT point of view. It is shown that in the absence of CSI at the transmitter (CSI-T), the degrees of freedom in the main channel are decreased by the degrees of freedom in the source-eavesdropper channel, and thus the secret DMT depends on the remaining degrees of freedom. Full CSI-T is proven to impact the secret DMT in which case the eavesdropper steals only the transmitter antennas. A zero-forcing scheme is proposed and shown to achieve the secret DMT in the case of full CSI-T.

Motivated by these results and by the promising applications of secret-key agreement over wireless communication channels, we consider secret-key agreement with public discussion over Rayleigh fading quasi-static MIMO channels from a DMT perspective. We define the notions of secret-key multiplexing gain and secret-key diversity gain. We study the fundamental limits in the case of no CSI-T and show that the eavesdropper is stealing antennas only from the source. We also argue that the secret-key DMT does not depend on the CSI-T. This behavior parallels the regular DMT without secrecy constraints and contrasts the behavior of the secret DMT, highlighting a fundamental difference between secret-key agreement and wiretap channel coding for secrecy. In the CSI-T case, we present several secret-key DMT-achieving schemes.

The remainder of this paper is organized as follows. Section II describes the channel model for secret-key agreement. In Section III, we investigate the secret-key DMT in the no CSI-T case and then extend the arguments in the case of available CSI-T in Section IV. We conclude in Section V.

II. SYSTEM MODEL AND PRELIMINARIES

Notations: Throughout this paper, the symbol $\dagger$ indicates the conjugate transpose of a matrix. $I_n$ and $0_{n \times m}$ denotes respectively the $n \times n$ identity matrix and the $n \times m$ matrix of zeros. For conciseness, we drop the subscripts whenever the matrix dimensions are clear from the context. Random quantities are denoted with capital letters. We use the symbol $\mathbb{H}(\cdot)$ to denote the mutual information between two random variables. We denote the entropy and differential entropy of a random quantity with $H(\cdot)$ and $h(\cdot)$, respectively. The symbols $|\cdot|$ and $\det(\cdot)$ are used to denote the determinant of a matrix and Tr denotes the trace operator. We use the notation $A \succeq 0$ to denote that $A$ is a positive semi-definite matrix. The expression $f_1(SNR) = f_2(SNR)$ is defined as

$$\lim_{SNR \to \infty} \frac{\log f_1(SNR)}{\log SNR} = \lim_{SNR \to \infty} \frac{\log f_2(SNR)}{\log SNR}$$

Inequalities are defined analogously.

We consider the 2W model for secret-key agreement between a transmitter and a legitimate receiver in a MIMO wireless network in the presence of an eavesdropper. We assume that the transmitter (Alice), the legitimate receiver (Bob) and the eavesdropper (Eve) are equipped with $m_S$, $m_P$ and $m_W$ antennas, respectively. Transmissions take place over quasi-static fading channels; that is, the MIMO channel matrices $H_D \in \mathbb{C}^{m_P \times m_S}$ and $H_W \in \mathbb{C}^{m_W \times m_S}$ are fixed for the whole duration of the communication.

For each channel use, the channel is represented as follows

$$Y_D = H_D X + N_D$$
$$Y_W = H_W X + N_W,$$

where

- $X$ is the $m_S \times 1$ complex-valued transmitted symbol vector by the source,
- $Y_D$ (resp. $Y_W$) is the $m_D \times 1$ (resp. $m_W \times 1$) complex-valued received symbol vector at the destination (resp. at the eavesdropper),
- $H_D$ (resp. $H_W$) is the $m_D \times m_S$ (resp. $m_W \times m_S$) channel matrix from the source to the destination (resp. the eavesdropper) with independent identically distributed (i.i.d) zero-mean unit variance circular-symmetric complex Gaussian entries,
- $N_D$ (resp. $N_W$) is the $m_D \times 1$ (resp. $m_W \times m_S$) noise vector with i.i.d zero-mean unit variance circular-symmetric complex Gaussian entries.

The transmitter is constrained in its total power, that is equivalent to a trace constraint on the input covariance matrix

$$K_X = E[XX^\dagger]$$

$$\text{trace}(K_X) \leq m_S \cdot SNR.$$  

To distill a secret-key from the symbols transmitted over the noisy channel, the transmitter and the receiver are allowed to communicate over an interactive, authenticated public channel with unlimited capacity, where interactive means that the channel is two-way and can be used multiple times, unlimited capacity means that it is noiseless and has infinite capacity and public and authenticated means that the eavesdropper can perfectly observe all communications over this channel but cannot tamper with the messages transmitted. For a precise description of a key-distillation strategy, we refer the reader to [4]. Concisely, it consists of transmissions over the noisy channel as well as exchanges of messages over the public channel. At the end, the source generates its secret-key $K$ and the destination generates its secret-key $L$, where $K$ and $L$ take values from the same finite set $\kappa$. We denote the messages sent over the public channel by the random variable $F$.

$R$ is an achievable secret-key rate through the channel $(X,Y_D,Y_W)$, if for every $\epsilon > 0$, there exists a permissible secret-sharing strategy of the form described above such that

1) $P\{K \neq L\} < \epsilon$,
2) $\frac{1}{n} H(K;Y_W,F) < \epsilon$,
3) $\frac{1}{n} H(K) > R - \epsilon$,
4) $\frac{1}{n} \log |\kappa| < \frac{1}{n} H(K) + \epsilon$,

for sufficiently large $n$. The key capacity is the supremum of achievable secret-key rates. Condition (1) means the legitimate parties should agree on a common key with high probability. Condition (2) requires that a negligible rate of information about the key should be leaked to the eavesdropper. Condition (3) implies that the key rate is equal to $R$ and condition (4) requires the key distribution to be
almost that of a uniform one.

In this work, we investigate the high SNR behavior of the probability of error with a target secret-key rate \( R_k^{(T)}(SNR) \) that scales with SNR. We define the secret-key multiplexing gain as
\[
\lim_{SNR \to \infty} \frac{R_k^{(T)}(SNR)}{\log SNR} \triangleq r_k.
\]

The secret-key multiplexing gain indicates how fast the target secret-key rate scales with increasing SNR. We also define the secret-key diversity gain \( d_k \) as
\[
\lim_{SNR \to \infty} \frac{\log P_e(SNR)}{\log SNR} \triangleq -d_k,
\]
where \( P_e(SNR) \) denotes the probability of error under secret-key constraints.

From the definition of the CW model for secret-key agreement, the probability of error \( P_e(SNR) \) of a system targeting a constant key rate \( R_k^{(T)}(SNR) \) that scales with SNR is, in general, due to three events:

- E1: reliability constraint violation, i.e., condition (1) of the definition of achievable key rate is not satisfied.
- E2: secret-key secrecy constraint violation, i.e., condition (2) is not satisfied.
- E3: uniformity constraint violation, i.e., condition (4) is not satisfied.

Then,
\[
P_e(SNR) = P(E1 \cup E2 \cup E3).
\]

On the other hand, for given channel realization matrices \( H_D \) and \( H_W \), the observations of the legitimate receiver and the eavesdropper are conditionally independent given the transmitted signal, i.e., \( Y_D \to X \to Y_W \) forms a Markov chain. Under the assumption that the channel matrices \( H_D \) and \( H_W \) are known to all three terminals, [7, Theorem 1] proves that the secret key rate
\[
R_k = \mathbb{I}(X;Y_D) - \mathbb{I}(Y_D;Y_W) = \mathbb{I}(X;Y_D|Y_W)
\]
(4)
is achievable for any input distribution \( p(X) \). The secret-key capacity is given by [7]
\[
C_k = \max_{\mathcal{R}(X)} \mathbb{I}(X;Y_D) - \mathbb{I}(Y_D;Y_W) \geq \mathbb{I}(X;Y_D) - \mathbb{I}(Y_D;Y_W) < \mathbb{I}(X;Y_D|Y_W).
\]
(5)

For a particular key agreement scheme with input distribution \( p(X) \), the probability of error is lower bounded by
\[
P_e(SNR) \geq P(E1 \cup E2) \geq P(\mathbb{I}(X;Y_D) - \mathbb{I}(Y_D;Y_W) < R_k^{(T)}(SNR)) \triangleq P(\text{secret-key rate outage}),
\]
show in the sequel, when a secret-key rate outage event happens, then we either have a secrecy leakage event or a key disagreement event.

In the following, we study the no CSI-T case. We evaluate the probability of the secret-key rate outage to establish a lower bound on \( P_e(SNR) \), i.e., to obtain an upper bound on the secret-key diversity gain \( d_k \), then we will proceed to achieve the upper bound. Similarly, we explore the CSI-T case in Section IV.

### III. No Channel State Information At The Transmitter

In this section, we assume that the destination and the eavesdropper have perfect CSI of their respective channels from the source. The source, on the other hand, is assumed to have no knowledge of the channel matrices realizations. The availability of a public channel with infinite capacity implies that this channel could be used to feed-back the source-destination channel matrix to the source. However, in our setting, we restrict our focus to the case where no such a CSI feedback is utilized and justifying our no CSI-T assumption. This can be regarded as a worst case scenario from a secret-key DMT perspective. Nevertheless, and as we show in Section IV, the secret-key DMT is insensitive to CSI-T so that even with full CSI-T, the secret-key DMT remains the same.

In the setting of no CSI-T, the secret-key capacity is not known. For a chosen input distribution \( p(X) \), the achievable secret-key rate in (4) becomes
\[
R_k = \mathbb{I}(X;Y_D) - \mathbb{I}(Y_D;Y_W) = h(Y_D|Y_W) - h(Y_D|X) = h(Y_D|Y_W) - m_D \log (\pi e).
\]
(7)

From [7, Lemma 1], in order to maximize \( R_k \), without any loss of optimality, the input distribution can be taken Gaussian with a covariance matrix \( K_X \) in which case,
\[
R_k = \log \det (K_{Y_D} - K_{Y_D Y_W} K_{Y_W}^{-1} K_{Y_W Y_D}),
\]
(8)
where \( K_{Y_D}, K_{Y_W}, K_{Y_D Y_W} \) and \( K_{Y_W Y_D} \) are the covariance and cross-covariance matrices defined by: \( K_{Y_D} = \mathbb{E} [Y_D Y_D] \), \( K_{Y_W} = \mathbb{E} [Y_W Y_W^\dagger], K_{Y_D Y_W} = \mathbb{E} [Y_D Y_W^\dagger] \) and \( K_{Y_W Y_D} = K_{Y_W Y_D}^\dagger \). It is easy to check the following expressions: \( K_{Y_D} = H_D K_X H_D^\dagger + I_{m_D}, K_{Y_D Y_W} = H_D K_X H_W^\dagger, K_{Y_W} = H_W K_X H_W^\dagger + I_{m_W} \) and \( K_{Y_W Y_D} = H_W K_X H_D^\dagger \) so that (8) becomes
\[
R_k = \log \det \left( I + H_D K_X H_D^\dagger - H_D K_X H_W^\dagger \left( H_W K_X H_W^\dagger + I_{m_W} \right)^{-1} H_W K_X H_D^\dagger \right) = \log \det \left( I + H_D K_X \left( I - H_W \left( H_W K_X H_W^\dagger + I_{m_W} \right)^{-1} H_W K_X \right)^\dagger \right).
\]
(9)

Optimizing over all input distributions, the secret-key rate outage probability is expressed as
\[ P_{\text{out}} \left( R_k^{(T)}(SNR) \right) = \inf_{K_X \geq 0, \text{Tr}(K_x) \leq m_S SNR} P \left( R_k < R_k^{(T)}(SNR) \right), \]

where the probability is taken over the random channel matrices \( H_D \) and \( H_W \).

**Lemma 1.** The secret-key rate function \( R_k(K_X) \) in (9) is an increasing function on the scale of covariance matrices (positive definite/semi-definite matrices).

**Proof:** Suppose \( K_X \) is non-singular. Using the matrix inversion lemma, for any non-singular covariance matrix \( K_X \), we write
\[ K_X \left[ I - H_W^\dagger (H_W K_X H_W^\dagger + I_{m_W})^{-1} H_W K_X \right] = \left( K_X^{-1} + H_W^\dagger K_X H_W^\dagger \right)^{-1}. \]

Let \( K_1 \) and \( K_2 \) be non-singular covariance matrices such that \( K_1 \geq K_2 \) (i.e., \( K_1 \) is positive semi-definite). Then, it follows that \( K_1^{-1} \geq K_2^{-1} \) and \( K_1^{-1} + H_W^\dagger K_X H_W^\dagger \geq K_2^{-1} + H_W^\dagger K_X H_W^\dagger \), which also leads to \( (K_1^{-1} + H_W^\dagger K_X H_W^\dagger)^{-1} \leq (K_2^{-1} + H_W^\dagger K_X H_W^\dagger)^{-1} \) and \( H_D(K_1^{-1} + H_W^\dagger K_X H_W^\dagger)^{-1} H_D^\dagger \geq H_D(K_2^{-1} + H_W^\dagger K_X H_W^\dagger)^{-1} H_D^\dagger \). Combining the last expression with (11) and the fact that \( \log \det(\cdot) \) is an increasing function on the cone of positive-definite Hermitian matrices, we obtain \( R_k(K_1) \leq R_k(K_2) \).

The case where \( K \) is singular can be handled by substituting \( K \) by \( K + \epsilon I \). The above derivation still holds, then it suffices to tend \( \epsilon \) to zero.

If we simply pick \( K_X = SNR I \), we get an upper bound on the key outage probability. On the other hand, \( K_X \) satisfies the power constraint \( \text{Tr}(K_X) \leq m_S SNR \), hence \( m_S SNR I - K_X \geq 0 \) and \( R_k(K_X) \leq R_k(m_S SNR I) \) by virtue of Lemma 1. Accordingly, the secret-key outage probability satisfies
\[ P \left( R_k(m_S SNR I) < R_k^{(T)}(SNR) \right) \leq P \left( R_k(SNR I) < R_k^{(T)} \right). \]

At high SNR,
\[ \lim_{SNR \to \infty} \frac{P \left( R_k(SNR I) < R_k^{(T)} \right)}{\log SNR} = \lim_{SNR \to \infty} \frac{P \left( R_k(m_S SNR I) < R_k^{(T)}(SNR) \right)}{\log(m_S SNR)} = \lim_{SNR \to \infty} \frac{P \left( R_k(m_S SNR I) < R_k^{(T)} \right)}{\log(SNR)}. \]

Therefore, the bounds are tight on the scale of interest. Hence, we have
\[ P_{\text{out}}(R_k^{(T)}) = P \left( R_k(SNR I) < R_k^{(T)} \right). \]

This shows that transmitting independent signals at equal power at each antenna is optimal at high SNR, which complies with the intuition that in the absence of CSI, the source has no preference on one direction over the other for its transmission. With this choice of input covariance matrix, the achievable secret-key rate in (9) becomes
\[ \hat{R}_k = \log \det \left( I + SNR H_D \left[ I - SNR H_W^\dagger \left( I_{m_W} + SNR H_W I_{m_W}^\dagger \right)^{-1} H_W \right] H_D^\dagger \right). \]

To establish the key DMT, we first evaluate the secret-key rate outage probability for a given target key rate \( R_k^{(T)}(SNR) = r_k \log SNR \).

**Lemma 2.** The secret-key outage probability at high SNR is given by:
\[ P_{\text{out}}(r_k \log SNR) \leq \begin{cases} SNR^{-d_{m_S-m_W,m_D}(r_k)} & \text{if } m_W < m_S \\ 1 & \text{else,} \end{cases} \]

where \( d_{n,m} \) is the DMT of a MIMO channel with \( n \) transmit and \( m \) receive antennas.

**Proof:** For the case \( m_W \geq m_S \), as \( H_W^\dagger H_W \) is invertible with probability 1 (w.p.1), it is not hard to check from (15c) that
\[ \lim_{SNR \to \infty} \hat{R}_k(SNR) = \log \frac{\det(H_D^\dagger H_D + H_W^\dagger H_W)}{\det(H_W^\dagger H_W)}. \]

In this case, the secret-key rate levels off and does not increase with SNR, hence the key outage probability does not decrease with SNR and the key DMT reduces to the single point \((0,0)\).

For the case \( m_W < m_S \), \( H_W^\dagger H_W \) is invertible w.p.1. From (15a), we can write
\[ \hat{R}_k(SNR) = \log \det \left( I + SNR H_D \left[ I - H_W^\dagger \left( \frac{1}{SNR} I_{m_W} + H_W I_{m_W}^\dagger \right)^{-1} H_W \right] H_D^\dagger \right). \]

The notation \( \hat{R}_k \) is used to indicate the key rate evaluated for \( K_X = SNR I \).
Defining
\[ \hat{R}_\infty(SNR) = \log \det \left( I + SNR H_D \left[ I - H_W H_W^\dagger (H_W H_W^\dagger)^{-1} H_W \right] H_D^\dagger \right), \]
and expressing the singular value decomposition (SVD) of \( H_W \) as \( H_W = U_W \left[ S_W 0_{m_S-m_W} \right] \left[ \hat{V}_W \hat{V}_W^\dagger \right], \) [7] shows that \( \hat{R}_\infty(SNR) \) can be expressed as
\[ \hat{R}_\infty(SNR) = \log \det \left( I + SNR H_D \hat{V}_W \hat{V}_W^\dagger H_D^\dagger \right). \]

It is also shown in [7] that
\[ 0 \leq \lim_{SNR \to \infty} \min \left( R_k(SNR) - \hat{R}_\infty(SNR) \right) \leq \lim_{SNR \to \infty} \sup \left( R_k(SNR) - \hat{R}_\infty(SNR) \right) \leq (m_D - j) \log \left( 1 + \max \left( \left[ H_W H_D^\dagger \right]^\dagger H_W H_D^\dagger \right) \right), \]
where \( j \) is the number of positive eigenvalues of \( H_D \hat{V}_W \hat{V}_W^\dagger H_D^\dagger \). Since \( H_D \hat{V}_W \hat{V}_W^\dagger H_D^\dagger \) is positive semidefinite, \( j \) is equal to its rank. Since the elements of \( H_W \) are i.i.d. and independent of the elements of \( H_D \), \( j = \text{rank}(H_D \hat{V}_W (H_D \hat{V}_W)^\dagger) = \text{rank}(H_D \hat{V}_W) = \min(m_D, m_S - m_W) \). Hence, in case \( m_D = \min(m_D, m_S - m_W) \), we obtain immediately
\[ \lim_{SNR \to \infty} \hat{R}_\infty(SNR) = \lim_{SNR \to \infty} \hat{R}_\infty(SNR). \]

In both cases, since the limit difference is finite, we have
\[ \hat{R}_k(SNR) \leq \hat{R}_\infty(SNR). \]

Hence, (14) becomes
\[ P_{out}(R) \leq P \left( \frac{\hat{R}_k(SNR)}{P} < R \right) \leq P \left( \frac{\hat{R}_\infty(SNR)}{P} < R \right) \leq P \left( \log \det \left( I + SNR H_D \hat{V}_W \hat{V}_W^\dagger H_D^\dagger \right) < R \right). \]

Conditioning on the elements of \( H_W \), \( \hat{V}_W \) is a deterministic quantity that depends on \( H_W \). Let \( \Psi = H_D \hat{V}_W \), since the columns of \( \hat{V}_W \) are orthonormal, then, we can check that \( \Psi \) is an \((m_D, m_S - m_W) \times m_D\) MIMO system without secrecy constraints whose DMT is well known [9].

\[ P \left( \log \det \left( I + SNR H_D \hat{V}_W \hat{V}_W^\dagger H_D^\dagger \right) < R \right) \leq P \left( \log \det \left( I + SNR \Psi \Psi^\dagger \right) < R \right) \leq SNR^{-d_{m_S-m_W,m_D}(r_k)}. \]

Averaging over the realizations of \( H_W \) as in [10], we obtain the result
\[ P_{out}(r_k SNR) \leq SNR^{-d_{m_S-m_W,m_D}(r_k)}. \]

The key outage probability provides an upper bound on the key DMT. The achievability of this upper bound is proved in the following proposition.

**Proposition 1.** For the multiple-antenna CW channel model, if \( m_W < m_S \), the secret-key diversity-multiplexing tradeoff is given by the piecewise linear function joining the points \((l, d_k(l))\), where \( l = 0, 1, \ldots \min(m_S - m_W, m_D) \) and \( d_k(l) = (m_S - m_W - l)(m_D - l) \).

If \( m_W \geq m_S \), then the secret-key diversity multiplexing tradeoff reduces to the single point \((0,0)\).

**Proof:** A lower bound on the probability of error is already obtained by virtue of Lemma 2. To achieve the bound, we outline a proof based on the consideration of a conceptual wiretap channel as in [7, thereom 1].

First, the source generates randomness by sending a sequence of i.i.d. symbols \( X^n \) over the wiretap channel such that \( E[|X|^2] \leq m_S SNR \) (so that the power constraint is satisfied). Then, the destination generates a sequence \( U^n \) randomly and independent of \((X^n, Y^n_0, Y^n_1)\), observes the realization \( Y^n_0 \) and sends \( U^n + Y^n_0 \) back to the source through the public channel. This creates a conceptual wiretap channel with input symbol \( U \), for which the source observes \((U + Y, X)\) and the eavesdropper observes \((U + Y, Y)\) and \( (U + Y_d, X) \), and the eavesdropper observes \((U + Y_d, Y)\). In this scheme, the error events are the same error events in a system with secrecy constraints, which is studied in [10].

\[ P_e(SNR) = P(\text{secrecy not achieved}) \]
\[ \leq P(\text{secrecy not achieved}) \]
\[ + P(\text{main channel decoding error}), \]

where, as in [10], we have
\[ P(\text{secrecy not achieved}) = P \left( I(U; U + Y_D, Y_W) > R^d \right) \]
\[ P(\text{main channel decoding error}) \]
\[ = P \left( I(U; (U + Y_D, X) < R^t \right), \]

where \( R^t = R^T + R^d \) denotes the total rate communicated through the conceptual wiretap channel and \( R^d \) is the dummy codewords rate. Since our target key rate scales as \( R^T_k = r_k \log(SNR) \), we have \( R^T(SNR) = r_k \log(SNR) + R^d(SNR) \). In order to choose the rates \( R^T \) and \( R^d \), we evaluate the mutual information involved. Straightforward derivation shows that the mutual information of the main channel can be expressed as

\[ I(U; X, Y_D, X) \]
\[ = h(U) - h(Y_D | X) + h(U + Y_D | X) - h(U | X). \]

Let \( U \) be Gaussian with covariance \( \sigma_U^2 I \). Using [11, Theorem 8.6.5] and the independence between \( U \) and \( Y_D \), we can write
\[ 0 \leq h(U + Y_D | X) - h(U | X) \leq \log \frac{\det(K_U + K_{Y_D})}{\det K_U}. \]
Now, since the input symbol $U$ has no power constraint, we can choose $\sigma^2_U$, large enough such that for small $\epsilon > 0$, combining (26) and (27), we have
\[
\mathbb{I}(U; U + Y_D, X) = h(U) - h(Y_D|X) + o(\epsilon) \\
\leq h(U) - h(Y_D|X).
\]
\[
(28)
\]
Based on that, (25) becomes
\[
P(\text{main channel decoding error}) \\
\leq P(\log \det (K_U) < r_k \log (SNR) + R^d).
\]
\[
(29)
\]
Equation (29) represents the first term of the upper bound in (23) and it will determine later the choice of the dummy-codewords rate. We evaluate the probability of secrecy not achieved. Following the same procedure as above (or simply replacing $X$ by $Y_W$), and with the choice of $K_U$ as explained above, we obtain similarly
\[
\mathbb{I}(U; U + Y_D, Y_W) \leq h(U) - h(Y_D|Y_W).
\]
\[
(30)
\]
Let $X$ Gaussian with $K_X = SNR I$. Applying [12, Lemmas 3 and 4], $[Y^T_D Y^T_W]_r$ is circular-symmetric complex Gaussian random vector. Hence, using [7, Lemma 1], we obtain
\[
h(Y_D|Y_W) = \log \det (K_{Y_D} - K_{Y_D Y_W} K_{Y_W}^{-1} K_{Y_Y} Y_D) \\
+ m_D \log(\pi e) \\
= \bar{R}_k(SNR) + m_D \log(\pi e).
\]
\[
(31)
\]
Therefore, the event of secrecy not achieved is reduced to
\[
P(\text{secrecy not achieved}) \\
\leq P(\bar{R}_k(SNR) < \log \det (K_U) - R^d).
\]
\[
(32)
\]
Equations (29) and (32) determine the upper bound on the total probability of error (23). Let
\[
R^d(SNR) = \log \det (K_U) - r_k \log (SNR),
\]
\[
(33)
\]
for $r_k = 0, \ldots, \min(m_S - m_W, m_D)$. Based on Lemma 2, we obtain
\[
P(\text{secrecy not achieved}) \leq SNR^{-d_{m_S - m_W, m_D}(r_k)},
\]
\[
P(\text{main channel decoding error}) = 0.
\]
\[
(34)
\]
Overall, the upper bound on the probability of error (23) becomes
\[
P_e(SNR) \leq SNR^{-d_{m_S - m_W, m_D}(r_k)}.
\]
\[
(35)
\]
Therefore, we conclude that the secret-key DMT is equal to $d_{m_S - m_W, m_D}(r_k)$ if $m_S > m_W$. If $m_S \leq m_W$, the key outage probability does not scale with SNR and the key DMT consequently reduces to the single point $(0,0)$.

Proposition 1 states that the eavesdropper steals $m_W$ antennas from the source only but does not affect the destination. The key DMT is clearly higher than the secret DMT for which it is shown in [10] that the eavesdropper steals $m_W$ from both legitimate parties. This advantageous behavior of the key DMT is explained by the availability of the public channel with infinite capacity that compensates for the absence of CSI-T.

When the degrees of freedom in the source-eavesdropper channel, $\min(m_S, m_W)$ is equal to $m_S$, then no degrees of freedom are left for the legitimate users and the secret-key DMT, as the key DMT, reduces to the point $(0,0)$.

On the other hand, when $m_W < m_S$, the CW MIMO system becomes equivalent to an $(m_S - m_W) \times m_D$ system, from a DMT point of view.

![Figure 1](image-url)  
**Figure 1.** The DMT without secrecy constraints, secret DMT with no CSI-T, key DMT. The source, the destination and the eavesdropper have 5, 3 and 2 antennas, respectively.

In Figure 1, the key DMT is shown for $m_S = 5, m_D = 3$ and $m_W = 2$. In this configuration, the DMT without secrecy constraints, the secret DMT with no CSI-T and the key DMT are shown to be respectively equal to $d_{5,3}(r), d_{3,1}(r_S)$ and $d_{3,3}(r_k)$. We clearly see that the secrecy constraint imposes not only a diversity gain loss but also a multiplexing gain loss for the secret DMT with no CSI-T in comparison with the DMT without secrecy constraint. On the other hand, the key DMT experiences only a diversity gain loss but still achieves all multiplexing gains. In fact, as highlighted in [10], the secret DMT with no CSI-T always experiences multiplexing gain loss ($m_W$ degrees of freedom) while the key DMT can achieve all multiplexing gains in the case $m_S - m_W \geq m_D$ but it does experience a secret-key diversity gain loss.

**Remark 1.** If the destination has access to the realization of the eavesdropper channel $H_W$, then no secret-key leakage can be guaranteed in the achievability proof. Indeed, the destination can adapt its dummy codewords rate to its channel to the eavesdropper and set $R_d = \log \det K_U - \bar{R}_K(SNR)$. Then, (29) and (32) become
\[
P(\text{main channel decoding error}) \\
\leq P(\bar{R}_k(SNR) < r_k \log (SNR)) \\
\leq P(\text{secret-key rate outage}) \\
P(\text{secrecy not achieved}) = 0.
\]
\[
(36)
\]
IV. Channel State Information At The Transmitter

In the previous section, secret-key DMT was investigated with no available CSI-T. In this section, we make the assumption that the source has full knowledge about its channel to the destination as well as its channel to the eavesdropper. Though it is arguable that such assumptions are more of a theoretical interest than of a practical regard, investigating the system under this setup will help us understand the fundamental limitations of secret-key DMT.

In the next subsections, we establish the secret-key DMT and we revisit some schemes that achieve the key DMT with CSI-T.

A. secret-key DMT with CSI-T

We recall the expression of the secret-key capacity (5), known from [8] and given by:

\[ C_K = \max_{\text{tr}(K_X) \leq m_S SNR} \mathbb{I}[X; Y_D | Y_W]. \]

Having full CSI-T enables the source to determine the optimal covariance matrix achieving the key capacity, or equivalently minimizing the secret-key rate outage probability. Hence, for \( K_X \) attaining the maximum in (5), the secret-key rate outage probability is given by

\[
P_{out}(r_k \log SNR) = P\left( K_X \geq 0, \text{tr}(K_X) \leq m_S SNR \right) \mathbb{I}(X; Y_D | Y_W) < r_k \log(SNR))
\]

\[
= \min_{\text{tr}(K_X) \leq m_S SNR} \mathbb{I}(X; Y_D | Y_W) < r_k \log(SNR))
\]

\[
\leq SNR^{-d_{m_S-m_W,m_D}(r_k)},
\]

(37)

where the last equality follows from (14) and Lemma 2. This states that even though \( K_X = SNR I \) would not necessarily be the optimal covariance matrix maximizing the secret-key capacity, from a key DMT perspective, splitting the available power equally among the different source antennas is optimal. Thus, the key outage probability is the same in the case of CSI-T. Therefore, the same upper bound on the secret-key diversity still holds in the CSI-T case. Achieving this upper bound is straightforward. The scheme based on the conceptual wiretap channel, as described in the proof of Proposition 1, can be used since they do not require CSI-T to achieve the key DMT.

Recall that for the MIMO wiretap channel, a uniform power allocation is not optimal from the secret DMT perspective [10]. This observation highlights again another difference between the key agreement setup and coding for secrecy problem. In fact, splitting the power equally does not achieve the secret outage probability, but rather results in the eavesdropper stealing both transmitter and receiver antennas. To achieve the maximum secret DMT in case of CSI-T, a more sophisticated input covariance matrix must be used as shown in [10].

B. Other Schemes achieving the key DMT

In the CSI-T case, the secret DMT coincides with the key DMT. This is relevant since a wiretap code over the wiretap channel \((X, Y_D, Y_W)\) is a key-agreement strategy for the CW model. Any scheme achieving the secret DMT could hence be employed to achieve the same key DMT without relying on the existence of the public channel.

A wiretap code over the main channel with the optimum covariance matrix maximizing the secrecy capacity was used to prove the achievability of the secret DMT in [10].

A Zero-forcing scheme is also suggested and proved to achieve the secret DMT. It consists in transmitting information in the nullspace of the eavesdropper, which guarantees not only secrecy but also a good outage performance. Artificial noise is presented in [13] in the case of the destination having more antennas than the source. It consists in transmitting information in the range space of the destination channel matrix \(H_D\) along with transmitting extra noise in the nullspace of \(H_D\). It is a semi-blind scheme since it requires only the knowledge of \(H_D\) to construct the signaling codebook. It is suggested in [10] through simulations, but without a formal proof, that the artificial noise scheme achieves the secret DMT. We present here the artificial noise for the sake of comparison.

In Figure 2, the outage performance of the described schemes achieving the key DMT is represented in the case of full CSI. The schemes are the conceptual wiretap scheme with equally-split power, the zero-forcing and the artificial noise schemes. The figure confirms that the schemes achieve a secret-key diversity 0.25 for a key multiplexing gain 1.75. Another observation is that transmitting the secret-key in the nullspace of the eavesdropper is the most attractive scheme for two reasons. First, it performs better in terms of outage probability compared to the artificial noise and a uniform power distribution schemes. Second, it does not require the public channel to achieve its performance, hence the cost of using the public channel can be saved.

V. Conclusion

In this paper, we have studied secret-key agreement diversity multiplexing tradeoff of Rayleigh fading quasi-static
MIMO channels. We have characterized the secret-key DMT for arbitrary number of antennas at the transmitter, the destination and the eavesdropper. First, we have studied the case of no-CSI-T. We have showed that transmitting signals with the same power over the antennas is optimum at the scale of interest. We have also showed that secret-key diversity is possible only if the source has strictly more antennas than the eavesdropper. In this scenario, the eavesdropper steals $m_W$ antennas from the source and the secret-key DMT is equivalent to that of a $(m_S - m_W) \times m_D$ MIMO system without secrecy constraints. We have argued that the secret-key DMT does not change in the case of CSI-T. Nonetheless, coding over the wiretap wiretap channel, without the need of the public channel, is sufficient to achieve the full secret-key DMT, since in this case the secret DMT coincides with the secret-key DMT.

REFERENCES