On the Ergodic Secret-Key Agreement over Spatially Correlated Multiple-Antenna Channels with Public Discussion

Marwen Zorgui, *Student Member, IEEE*, Zouheir Rezki, *Member, IEEE*, Basel Alomair, *Member, IEEE*, Eduard A. Jorswieck, *Senior Member, IEEE* and Mohamed-Slim Alouini, *Fellow, IEEE*

Abstract

We consider secret-key agreement with public discussion over multiple-input multiple-output (MIMO) Rayleigh fast-fading channels under correlated environment. We assume that transmit, legitimate receiver and eavesdropper antennas are correlated. The legitimate receiver and the eavesdropper are assumed to have perfect channel knowledge while the transmitter has only knowledge of the correlation matrices. First, we derive the expression of the secret-key capacity under the considered setup. We prove that the optimal transmit strategy achieving the secret-key capacity consists in transmitting independent Gaussian signals along the eigenvectors of the transmit correlation matrix. The powers allocated to each channel mode are determined as the solution to a numerical optimization problem. A necessary and sufficient condition for beamforming (i.e., transmitting along the strongest channel mode) to be capacity-achieving is derived. Moreover, we analyze the impact of correlation matrices on the system performance. Finally, we study the system's performance in the two extreme power regimes. In the high-power regime, we provide closed-form expressions of the gain/loss due to correlation. In the low signal-to-noise ratio (SNR) regime, we investigate the energy efficiency of the system by determining the minimum energy required for sharing a secret-key bit and the wideband slope while highlighting the impact of correlation matrices.

Index Terms

Secret-key agreement, MIMO systems, optimal signaling, correlation, beamforming, high-power regime, energy efficiency.

I. INTRODUCTION

Information-theoretic security has recently spurred a lot of interest in the research community. Its premise is to exploit the inherent randomness of the physical medium and the difference between a legitimate receiver channel and an eavesdropper channel in order to securely transmit confidential messages. The wiretap channel investigated in the seminal paper [1] demonstrates the possibility of such a promising idea in case the source-to-eavesdropper channel is a degraded version of the source-to-destination channel. Subsequent works extend this result to different channels, such as the Gaussian wiretap

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

This work has been supported by a grant from King Abdulaziz City of Science and Technology (KACST), Riyadh, Saudi Arabia. The work of E. Jorswieck is partly supported by the German Research Foundation (DFG) within the Cluster of Excellence "Center for Advancing Electronics Dresden (cfaed)".

M. Zorgui, Z. Rezki and M.-S. Alouini are with the Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia 23955-6900. (e-mail: marwen.zorgui@kaust.edu.sa; zouheir.rezki@kaust.edu.sa; slim.alouini@kaust.edu.sa).

B. Alomair is with the Computer Research Institute (CRI), King Abdulaziz City for Science and Technology (KACST), PO Box 6086, Riyadh 11442, Saudi Arabia. (e-mail: alomair@uw.edu).

E. A. Jorswieck is with the Communications Laboratory, Technische Universität Dresden, Dresden 01062, Germany (e-mail: eduard.jorswieck@tu-dresden.de).

channel [2], the general discrete memoryless broadcast channel [3], the fading channel [4] and also to the multiple input multiple output (MIMO) channel [5]–[7]. With the availability of a feedback channel, [8] introduced the idea of distilling a secret-key shared between legitimate parties through the concept of common randomness. Sharing a secret-key between legitimate parties can enhance significantly the security of a communication system. Indeed, the shared secret-key can be used, for example, as a one-time pad to secure a confidential message. A common system model for the key agreement problem is the “channel model with wiretapper” (CW), introduced in [8]. For the CW model, the source of randomness is controlled by one of the legitimate parties, similar to the basic wiretap channel model [1] with an additional public feedback channel available to the legitimate parties for exchanging public messages.

A. Related work

Several works have built upon the concept of common randomness. Extensions to the case of secret sharing among multiple users, with the possibility of some terminals acting as helpers have been investigated in [9], [10]. We shall focus on a particular extension of the CW model. Indeed, in case of fast-fading MIMO wiretap channels, [11] shows that the transmitter does not require the instantaneous channel state information to achieve a positive secret-key rate. The secret-key capacity is provided and it is also shown that distributing power uniformly across transmit antennas is secret-key capacity-achieving. In [11], it is assumed that the channel matrices have independent and identically distributed (i.i.d.) rows and also i.i.d. columns. Such an assumption may be hard to prove in practice. Practical systems demonstrate strong correlation due to several impairments, to mention a lack of sufficient local reflectors or insufficient spacing across transmit and/or receive antennas. Taking correlation into consideration, a wide body of work has focused on the effect of spatial correlation on the performance of wireless communication systems, but without secrecy constraint. For instance, in [12], the multiple input single output (MISO) is studied in the case of transmit correlation. Reference [12] highlighted the benefit that partial knowledge of the channel provides. In [13], the MIMO system with transmit correlation and imperfect feedback is studied. The optimal input covariance matrix is characterized and involves transmitting independent Gaussian symbols along the eigenvectors of the channel covariance matrix, which parallels the result obtained in [12] for the MISO case. Reference [14] studies the MIMO system under the additional assumption of receive correlation. Under this setup, [14] determines the optimal transmit strategy and analyzes the impact of both correlation matrices.

To the best of our knowledge, the effect of spatial correlation on the secret key rate from the channel model has not been addressed before. Analyzing the impact of correlation and identifying optimal signaling strategies under such assumptions is interesting from both information-theoretic and system design perspectives.

In addition to this, energy efficiency (formally defined later) is a key parameter for wireless communication system design. The presence of spatial correlation impacts also the performance of the system from an energy efficiency perspective. It is worth noting that for many channel models, energy efficiency improves as one operates at lower signal-to-noise ratio (SNR) regime, and moreover, the minimum energy per bit is achieved as SNR vanishes [15]. Closely related to that, [15] introduces the wideband slope measure, that not only describes the rate of convergence of the energy required for transmitting a bit reliably to its minimum value, but also allows a good approximation of the capacity in the low but non vanishing SNRs. In

[16], the secrecy capacity per unit-cost is defined in order to study cost-efficient wide-band secrecy communication, extending the concept of channel capacity per unit cost, introduced in [17]. In [18], the secrecy capacity of a multiple-antenna wiretap channel in the low-SNR regime has been characterized in both quasi-static and ergodic fading. In [19], the secret-key capacity is studied in the low power regime in the quasi-static fading. The first and second derivatives of the secret-key capacity are determined along with optimal transmission schemes achieving them.

Motivated by the importance of wireless physical layer security and the relevance of spatial correlation in practical wireless systems, we analyze the performance of secret-key agreement in a correlated environment. For that, we consider the problem of secret-sharing over a fast-fading MIMO wiretap channel. Using the Kronecker model, we assume that transmit, legitimate receiver and eavesdropper antennas are all correlated, and we refer to them as transmit, receive and eavesdropper correlation, respectively. In [20], a shorter version of the present paper, we studied the secret-key capacity expression and determined the general optimal transmission strategy. We derived a necessary and sufficient condition for which beamforming is optimal and analyzed briefly the impact of correlation matrices on the secret-key capacity.

B. Contributions of the paper

In this paper, we consider the same setup as in [20]. First, we revisit the secret-key capacity expression and provide detailed proofs of the optimal transmission strategy. Interestingly, we show that secrecy constraint does not impact the optimal beamforming directions which consist of transmitting signals along the eigenvectors of the transmit correlation matrix. Moreover, optimal power allocation along these beamforming directions is shown to depend on the eigenvalues of all correlation matrices. Next, we provide a necessary and sufficient condition under which transmitting along the eigenvector corresponding to the maximum eigenvalue of the transmit correlation matrix is secret-key capacity achieving. In addition, we analyze the impact of each correlation matrix on the system performance and show in particular that eavesdropper correlation increases the capacity. Furthermore, we investigate the system performance in the high and low-power regimes. In the high-power regime, we show that transmit correlation decreases the capacity and we provide closed-form expressions of the asymptotic difference between secret-key capacities with and without spatial correlation. In the low-power regime, we provide a second-order Taylor series approximation of the secret-key capacity and identify the optimal transmission strategies. Based on that, we evaluate the minimum energy required for sharing reliably a secret-key bit as well as the wideband slope. Overall, the contributions lead to system design guidelines on how to place and operate the multiple antenna transmitter and receiver in a secret key generation scheme.

C. Outline of the paper

The organization of this paper is as follows. In Section II, the system model and the corresponding assumptions are described and the secret-key capacity expression is derived. Section III contains the solution to the transmitter optimization problem. In Section IV, we develop a necessary and sufficient condition under which beamforming is optimal. In Section V, we study the impact of correlation matrices on the secret-key capacity. Section VI analyzes the system in the high-power regime. In Section VII, we study low-SNR behavior of the secret-key capacity. Conclusions are provided in Section VIII. Appendices

Fig. 1. System model. The transmitter, the receiver and the eavesdropper have m_S , m_D and m_E antennas, respectively. A public channel is available for both legitimate parties.

contain detailed proofs.

Notations: Throughout this paper, the symbol † indicates the conjugate transpose operator. $\mathbf{I}_n$ denotes the $n \times n$ identity matrix and $\text{diag}[\lambda_1, \dots, \lambda_n]$ denotes the $n \times n$ diagonal matrix whose entries are $\lambda_1, \dots, \lambda_n$. Vectors are denoted with small bold letters while matrices are denoted with capital bold letters. The capital letter $\mathbb{E}$ denotes the expectation of a random quantity. We use the symbol $\mathbb{I}(\cdot; \cdot)$ to denote the mutual information between two random variables. $\mathbf{H}_{i\bullet}$ and $\mathbf{H}_{\bullet j}$ represent, respectively, the i th row and the j th column of a matrix $\mathbf{H}$. The symbol $|\cdot|$ is used to denote the determinant of a matrix and $\text{Tr}(\cdot)$ denotes the trace operator. We use the notation $\mathbf{A} \succeq 0$ to denote that $\mathbf{A}$ is a positive semi-definite matrix. $\mathbf{A}^{\frac{1}{2}}$ denotes a square-root of a matrix. We use the notation $\mathbf{x} \succ \mathbf{y}$ to denote that the vector $\mathbf{x}$ majorizes the vector $\mathbf{y}$, in the sense of [21]. The notation $\mathbf{x} \sim \mathcal{CN}(0, \Sigma)$ denotes that the random vector $\mathbf{x}$ is complex Gaussian with zero mean and covariance matrix Σ .

II. SYSTEM MODEL AND PRELIMINARIES

As illustrated in Fig. 1, we consider the problem of secret-key agreement between a transmitter and a legitimate receiver in the presence of an eavesdropper who overhears transmissions broadcasted over the wireless medium [8], [11]. The objective of the communication is that both legitimate users generate a key at the end of the communication process that should be kept secure from the eavesdropper. To this end, both users have access to a wireless channel and public channel over which they can exchange messages [8], [11]. The transmitter, destination and eavesdropper have m_S , m_D and m_E antennas, respectively. For each channel use, the channel is represented as follows

$$\begin{aligned} \mathbf{y}_D(i) &= \mathbf{H}_D(i)\mathbf{x}(i) + \mathbf{n}_D(i) \\ \mathbf{y}_E(i) &= \mathbf{H}_E(i)\mathbf{x}(i) + \mathbf{n}_E(i), \end{aligned} \tag{1}$$

where index $i, i = 1, \dots, n$, designates time instant i , and

- $\mathbf{x}(i)$ is the $m_S \times 1$ complex-valued transmitted symbol vector,
- $\mathbf{y}_D(i)$ (resp. $\mathbf{y}_E(i)$) is the $m_D \times 1$ (resp. $m_E \times 1$) complex-valued received symbol vector at the destination (resp. at the eavesdropper),
- $\mathbf{n}_D(i)$ (resp. $\mathbf{n}_E(i)$) is the $m_D \times 1$ (resp. $m_E \times 1$) noise vector with i.i.d. circular-symmetric complex Gaussian entries $\sim \mathcal{CN}(0, \sigma_D^2)$ (resp. $\sim \mathcal{CN}(0, \sigma_E^2)$),
- $\mathbf{H}_D(i)$ (resp. $\mathbf{H}_E(i)$) is the $m_D \times m_S$ (resp. $m_E \times m_S$) channel matrix from the source to the destination (resp. the eavesdropper).

For ease of notation we drop the index hereafter. The transmitter is constrained in its total power, that is equivalent to a trace constraint on the input covariance matrix $\mathbf{Q} = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$

$$\text{Tr}(\mathbf{Q}) \leq P. \quad (2)$$

To capture the effect of correlation, many channel models exist in the literature [22]. One of the commonly used models is the Kronecker model that is characterized by its tractability and its relevance in many wireless environments [23] [24]. Therefore, in the presence of spatial correlation at all terminals, the channel matrices are modeled as [25]

$$\begin{aligned} \mathbf{H}_D &= \mathbf{R}_D^{1/2} \mathbf{W}_D \mathbf{R}_T^{1/2} \\ \mathbf{H}_E &= \mathbf{R}_E^{1/2} \mathbf{W}_E \mathbf{R}_T^{1/2}, \end{aligned} \quad (3)$$

where $\mathbf{W}_D$ (resp. $\mathbf{W}_E$) is $m_D \times m_S$ (resp. $m_E \times m_S$) matrix with i.i.d. entries $\sim \mathcal{CN}(0, 1)$. $\mathbf{R}_T$ is the $m_S \times m_S$ transmit correlation matrix, $\mathbf{R}_D$ is the $m_D \times m_D$ destination correlation matrix and $\mathbf{R}_E$ is the $m_E \times m_E$ eavesdropper correlation matrix. All correlation matrices are normalized such that all diagonal entries are unity. This model characterizes a situation in which the channel distribution changes much more slowly than the channel itself. In such scenario, we are interested in finding the secret-key capacity for each given and fixed channel distribution. The destination and the eavesdropper have perfect CSI about their incoming channels while the transmitter has imperfect feedback about the legitimate link and statistical channel state information about the eavesdropper link. Precisely, the transmitter is assumed to know the correlation matrices. We note that the assumption of correlation matrices knowledge can be justified in scenarios in which a base station, having knowledge of each user channel, tries to share a secret-key with one particular user, treating the other users as eavesdroppers. Moreover, we note that for passive eavesdroppers the channel statistics can be computed from the scenario geometry, e.g. if some directions are impossible because of obstacles. Let $\mathbf{y} = \begin{bmatrix} \mathbf{y}_D & \mathbf{H}_D \end{bmatrix}$ be the observation of the legitimate receiver for each channel use, and $\mathbf{z} = \begin{bmatrix} \mathbf{y}_E & \mathbf{H}_E \end{bmatrix}$ be the observation of the eavesdropper for each channel use. The wireless channel modeled in (1) is used n times. The n uses of the channel are assumed independent and identically distributed, resulting in the memoryless property of the channel $(\mathbf{x}, \mathbf{y}, \mathbf{z})$. Since the transmitter does not have knowledge of the current channels realizations, we have that $\mathbf{x}$ is independent of $\mathbf{H}_D$, $\mathbf{H}_E$, $\mathbf{n}_D$ and $\mathbf{n}_E$. The signals satisfy the Markov chain $\mathbf{y} - \mathbf{x} - \mathbf{z}$. Henceforth, using [11, Theorem 1],

the secret-key capacity is obtained as

$$C = \max_{\substack{\text{Tr}(\mathbf{Q}) \leq P \\ \mathbf{Q} \succeq 0}} [\mathbb{I}(\mathbf{x}; \mathbf{y}) - \mathbb{I}(\mathbf{y}; \mathbf{z})]. \quad (4)$$

Following along similar lines as [11], it can be seen that the optimal input distribution is a circular-symmetric complex Gaussian $\sim \mathcal{CN}(0, \mathbf{Q})$. Taking into account the additional correlation matrices, the secret-key capacity can be expressed as

$$C = \max_{\substack{\text{Tr}(\mathbf{Q}) \leq P \\ \mathbf{Q} \succeq 0}} \left\{ \mathbb{E} [\log |\mathbf{K}_{\mathbf{y}_D} - \mathbf{K}_{\mathbf{y}_D \mathbf{y}_E} \mathbf{K}_{\mathbf{y}_E}^{-1} \mathbf{K}_{\mathbf{y}_E \mathbf{y}_D}|] - m_D \log(\sigma_D^2) \right\}, \quad (5)$$

where the expectation is taken over the realizations of $(\mathbf{H}_D, \mathbf{H}_E)$, $\mathbf{K}_{\mathbf{y}_D}$, $\mathbf{K}_{\mathbf{y}_E}$, $\mathbf{K}_{\mathbf{y}_D \mathbf{y}_E}$ and $\mathbf{K}_{\mathbf{y}_E \mathbf{y}_D}$ are the conditional covariances and cross-covariances. After some algebraic manipulations as in [19, Lemma 1], the secret-key capacity can be expressed as

$$C = \max_{\substack{\text{Tr}(\mathbf{Q}) \leq P \\ \mathbf{Q} \succeq 0}} \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \mathbf{H} \mathbf{Q} \mathbf{H}^\dagger|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{H}_E \mathbf{Q} \mathbf{H}_E^\dagger|} \right], \text{ with } \mathbf{H} = \begin{bmatrix} \frac{1}{\sigma_D} \mathbf{H}_D \\ \frac{1}{\sigma_E} \mathbf{H}_E \end{bmatrix}. \quad (6)$$

We recall that achieving the ergodic secret-key capacity is performed over many channel realizations. For ease of notation, the term secret-key capacity throughout this paper should be understood in the sense of an ergodic secret-key capacity. Note that secret-key agreement schemes can be used in conjunction with wiretap codes to enhance the security of confidential messages. For instance, the transmitter can exploit the secret-key generated at the end of the process as a one-time pad encryption as [26, Equation 4.4]. In the following section, we identify the optimal transmission scheme solving (6) and achieving the secret-key capacity.

III. OPTIMAL SIGNALING STRATEGY

Let the eigenvalue decomposition of the correlation matrix $\mathbf{R}_l$ be as $\mathbf{R}_l = \mathbf{U}_l \mathbf{\Lambda}_l \mathbf{U}_l^\dagger$, for $l \in \{T, D, E\}$, where $\mathbf{U}_l$ is unitary and $\mathbf{\Lambda}_l$ is a diagonal matrix containing the corresponding eigenvalues. We further assume, without loss in generality, that the eigenvalues of the transmit correlation matrix $\{\lambda_i^T, i = 1, \dots, m_S\}$ are arranged in decreasing order, i.e., $\lambda_1^T \geq \lambda_2^T \geq \dots \geq \lambda_{m_S}^T$. The following theorem characterizes the optimal transmission scheme.

Theorem 1. Let $\mathbf{H}_D = \mathbf{R}_D^{1/2} \mathbf{W}_D \mathbf{R}_T^{1/2}$, $\mathbf{H}_E = \mathbf{R}_E^{1/2} \mathbf{W}_E \mathbf{R}_T^{1/2}$ with $\mathbf{W}_D \sim \mathcal{CN}(0, \mathbf{I}_{m_D})$, $\mathbf{W}_E \sim \mathcal{CN}(0, \mathbf{I}_{m_E})$ and assume that $\mathbf{R}_T$ is of full rank, then the optimal input covariance matrix can be expressed as $\mathbf{Q} = \mathbf{U}_T \mathbf{\Lambda}_Q \mathbf{U}_T^\dagger$, where $\mathbf{\Lambda}_Q = \text{diag}[\lambda_1^Q, \dots, \lambda_{m_S}^Q]$.

That is, the optimal signaling strategy is to transmit independent complex circular Gaussian inputs along the eigenvectors of $\mathbf{R}_T$. The optimal power allocation is obtained by solving a numerical optimization problem subject to a trace constraint.

Proof: We note that the proof builds upon the techniques used in [13], [14]. Starting from the expressions of $\mathbf{H}_D$ and $\mathbf{H}_E$, $\mathbf{H}$ can be expressed as

$$\begin{aligned}\mathbf{H} &= \begin{bmatrix} \frac{1}{\sigma_D} \mathbf{R}_D^{1/2} \mathbf{W}_D \mathbf{R}_T^{1/2} \\ \frac{1}{\sigma_E} \mathbf{R}_E^{1/2} \mathbf{W}_E \mathbf{R}_T^{1/2} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{R}_D^{1/2} & 0 \\ 0 & \mathbf{R}_E^{1/2} \end{bmatrix}}_{\mathbf{R}_{DE}^{1/2}} \underbrace{\begin{bmatrix} \frac{1}{\sigma_D} \mathbf{W}_D \\ \frac{1}{\sigma_E} \mathbf{W}_E \end{bmatrix}}_{\mathbf{W}} \mathbf{R}_T^{1/2} \\ &= \mathbf{R}_{DE}^{1/2} \mathbf{W} \mathbf{R}_T^{1/2}.\end{aligned}\quad (7)$$

Given the eigenvalue decomposition of $\mathbf{R}_D$ and $\mathbf{R}_E$, $\mathbf{R}_{DE}$ can be diagonalized as

$$\begin{aligned}\mathbf{R}_{DE} &= \underbrace{\begin{bmatrix} \mathbf{U}_D & 0 \\ 0 & \mathbf{U}_E \end{bmatrix}}_{\mathbf{U}_{DE}} \underbrace{\begin{bmatrix} \mathbf{\Lambda}_D & 0 \\ 0 & \mathbf{\Lambda}_E \end{bmatrix}}_{\mathbf{\Lambda}_{DE}} \begin{bmatrix} \mathbf{U}_D^\dagger & 0 \\ 0 & \mathbf{U}_E^\dagger \end{bmatrix} \\ &= \mathbf{U}_{DE} \mathbf{\Lambda}_{DE} \mathbf{U}_{DE}^\dagger.\end{aligned}\quad (8)$$

Since $\mathbf{W} \stackrel{d}{=} \mathbf{W} \mathbf{V} \stackrel{d}{=} \mathbf{V} \mathbf{W}$ for any unitary matrix $\mathbf{V}$ (the symbol $\stackrel{d}{=}$ means equality in distribution), using $|\mathbf{I} + \mathbf{A} \mathbf{B}| = |\mathbf{I} + \mathbf{B} \mathbf{A}|$, we obtain

$$R(\mathbf{Q}) = \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda}_T^{1/2} \mathbf{U}_T^\dagger \mathbf{Q} \mathbf{U}_T \mathbf{\Lambda}_T^{1/2} \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda}_T^{1/2} \mathbf{U}_T^\dagger \mathbf{Q} \mathbf{U}_T \mathbf{\Lambda}_T^{1/2} \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2}|} \right].$$

We define unitary $\mathbf{U}$ and diagonal $\mathbf{\Lambda}$ by the spectral decomposition

$$\mathbf{\Lambda}_T^{1/2} \mathbf{U}_T^\dagger \mathbf{Q} \mathbf{U}_T \mathbf{\Lambda}_T^{1/2} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\dagger, \quad (9)$$

where $\mathbf{\Lambda} = \text{diag}[\lambda_1, \dots, \lambda_{m_S}]$, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{m_S}$. It follows that

$$R(\mathbf{Q}) = \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda} \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda} \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2}|} \right]. \quad (10)$$

Also, since the transmit correlation is assumed to be invertible, the trace constraint can be expressed as

$$\text{Tr}(\mathbf{Q}) = \text{Tr}(\mathbf{U}_T \mathbf{\Lambda}_T^{-1/2} \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\dagger \mathbf{\Lambda}_T^{-1/2} \mathbf{U}_T^\dagger) = \text{Tr}(\mathbf{U}^\dagger \mathbf{\Lambda}_T^{-1} \mathbf{U} \mathbf{\Lambda}), \quad (11)$$

where the last equality is obtained using the property $\text{Tr}(\mathbf{A} \mathbf{B}) = \text{Tr}(\mathbf{B} \mathbf{A})$. Henceforth, the optimization problem becomes

$$C = \max_{\text{Tr}(\mathbf{U}^\dagger \mathbf{\Lambda}_T^{-1} \mathbf{U} \mathbf{\Lambda}) \leq P} \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda} \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda} \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2}|} \right]. \quad (12)$$

To complete the proof, we need to show that $\mathbf{U} = \mathbf{I}_{m_S}$ is optimal. Proceeding as in [27], and using [21, Th.H.1.h], it can be shown that $\text{Tr}(\mathbf{U}^\dagger \mathbf{\Lambda}_T^{-1} \mathbf{U} \mathbf{\Lambda})$ is simultaneously minimized for all $\mathbf{\Lambda}$ by $\mathbf{U} = \mathbf{I}_{m_S}$. Therefore, for $\mathbf{U} = \mathbf{I}_{m_S}$, (9) leads to

$\mathbf{Q} = \mathbf{U}_T \mathbf{\Lambda}_T^{-1} \mathbf{\Lambda} \mathbf{U}_T^\dagger$ with $\mathbf{\Lambda}_Q = \mathbf{\Lambda}_T^{-1} \mathbf{\Lambda}$, which is the result we seek to prove. Therefore, the optimization problem becomes

$$C = \max_{\text{Tr}(\mathbf{\Lambda}_T^{-1} \mathbf{\Lambda}) \leq P} \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda} \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda} \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2}|} \right] \\ = \max_{\text{Tr}(\mathbf{\Lambda}_Q) \leq P} \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda}_T \mathbf{\Lambda}_Q \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda}_T \mathbf{\Lambda}_Q \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2}|} \right] \quad (13)$$

$$= \max_{\text{Tr}(\mathbf{\Lambda}_Q) \leq P} R_1(\mathbf{\Lambda}_Q) - R_2(\mathbf{\Lambda}_Q), \quad (14)$$

where

$$R_1(\mathbf{\Lambda}_Q) = \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda}_T \mathbf{\Lambda}_Q \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda}_T \mathbf{\Lambda}_Q \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2}|} \right] \\ R_2(\mathbf{\Lambda}_Q) = \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda}_T \mathbf{\Lambda}_Q \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda}_T \mathbf{\Lambda}_Q \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2}|} \right]. \quad (15)$$

■

Remark 1. We note that the power constraint should be satisfied with equality. Indeed, since the function $C(\mathbf{\Lambda}_Q)$, defined in (14), is strictly increasing on the set of positive definite matrices [28, Lemma 1], we can simply allocate the rest of the available power to the strongest mode to achieve a higher secret-key rate. The optimal eigenvalues have to be computed by solving the optimization problem

$$\mathbf{\Lambda}_Q^{opt} = \arg \max_{\mathbf{\Lambda}_Q: \sum_{i=1:m_S} \mathbf{\Lambda}_Q(i) = P} \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \sum_{i=1}^{m_S} \lambda_i^T \lambda_i^Q \tilde{\mathbf{W}}_{\bullet i} \tilde{\mathbf{W}}_{\bullet i}^\dagger|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \sum_{i=1}^{m_S} \lambda_i^T \lambda_i^Q \tilde{\mathbf{W}}_{E\bullet i} \tilde{\mathbf{W}}_{E\bullet i}^\dagger|} \right], \quad (16)$$

where

$$\tilde{\mathbf{W}}_{\bullet i} = \mathbf{\Lambda}_{RE}^{1/2} \mathbf{W}_{\bullet i} = \begin{pmatrix} \frac{1}{\sigma_D} \mathbf{\Lambda}_D^{1/2} \mathbf{W}_{D\bullet i} \\ \frac{1}{\sigma_E} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_{E\bullet i} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_D} \tilde{\mathbf{W}}_{D\bullet i} \\ \frac{1}{\sigma_E} \tilde{\mathbf{W}}_{E\bullet i} \end{pmatrix} \\ \tilde{\mathbf{W}}_{E\bullet i} = \mathbf{\Lambda}_E^{1/2} \mathbf{W}_{E\bullet i}, \quad \tilde{\mathbf{W}}_{D\bullet i} = \mathbf{\Lambda}_D^{1/2} \mathbf{W}_{D\bullet i}. \quad (17)$$

Remark 2. The optimal choice of eigenvectors depends only on the transmit antenna correlation matrix $\mathbf{R}_T$ and is independent of the receive correlation matrix $\mathbf{R}_D$ and of the eavesdropper correlation matrix $\mathbf{R}_E$. This follows because the optimal beamforming directions are essentially dictated by the trace constraint which itself depends on the transmit correlation irrespectively of the secrecy constraint. Nonetheless, the optimal power allocation scheme depends on all the three correlation matrices. This will be highlighted by the necessary and sufficient condition of optimality of beamforming, derived in the sequel.

In the next section, we study the optimality of beamforming. Beamforming corresponds to transmitting along the strongest eigenvector. In this case, the input covariance matrix has rank one. We provide a necessary and sufficient condition under which beamforming is secret-key capacity-achieving.

IV. OPTIMALITY OF BEAMFORMING

The form of the optimization problem (16) does not lend itself to a simple closed-form solution. Knowing the eigenvalues of the correlation matrices $\mathbf{R}_T$, $\mathbf{R}_D$ and $\mathbf{R}_E$, the optimal power allocation is difficult to compute. For that, we study the condition under which allocating all the power to the strongest eigenvector is optimal. The following theorem formalizes our result.

Theorem 2. For the problem of secret-key agreement over fast-fading MIMO wiretap channel with known transmit correlation matrix $\mathbf{R}_T$, known receive correlation matrix $\mathbf{R}_D$, and known eavesdropper correlation matrix $\mathbf{R}_E$, with $\mathcal{A}$ defined as in (19), we have the following properties:

- if $\mathcal{A} \leq 0$, then beamforming is always optimal,
- else: beamforming is optimal if and only if

$$\lambda_2^T \mathcal{A} + \frac{1}{P} \left\{ \mathbb{E} \left[\frac{1}{1 + P\lambda_1^T \left(\frac{\|\tilde{\mathbf{W}}_{D\bullet 1}\|^2}{\sigma_D^2} + \frac{\|\tilde{\mathbf{W}}_{E\bullet 1}\|^2}{\sigma_E^2} \right)} \right] - \mathbb{E} \left[\frac{1}{1 + \frac{P\lambda_1^T}{\sigma_E^2} \|\tilde{\mathbf{W}}_{E\bullet 1}\|^2} \right] \right\} \leq 0. \quad (18)$$

$$\mathcal{A} = \frac{m_D}{\sigma_D^2} - \mathbb{E} \left[\frac{P\lambda_1^T \left[\frac{\|\Lambda_D^{1/2} \tilde{\mathbf{W}}_{D\bullet 1}\|^2}{\sigma_D^4} + \frac{\|\Lambda_E^{1/2} \tilde{\mathbf{W}}_{E\bullet 1}\|^2}{\sigma_E^4} \right]}{1 + P\lambda_1^T \left(\frac{\|\tilde{\mathbf{W}}_{D\bullet 1}\|^2}{\sigma_D^2} + \frac{\|\tilde{\mathbf{W}}_{E\bullet 1}\|^2}{\sigma_E^2} \right)} \right] + \frac{1}{\sigma_E^2} \mathbb{E} \left[\frac{\frac{P\lambda_1^T}{\sigma_E^2} \|\Lambda_E^{1/2} \tilde{\mathbf{W}}_{E\bullet 1}\|^2}{1 + \frac{P\lambda_1^T}{\sigma_E^2} \|\tilde{\mathbf{W}}_{E\bullet 1}\|^2} \right]. \quad (19)$$

Proof: See Appendix A. ■

We note that $\|\Lambda_D^{1/2} \tilde{\mathbf{W}}_{D\bullet 1}\|^2 = \sum_{i=1}^{m_D} \lambda_i^D w_i$, where the w_i 's are independent standard exponential random variables. Hence, the distribution of $\|\Lambda_D^{1/2} \tilde{\mathbf{W}}_{D\bullet 1}\|^2$ can be evaluated in closed-form [29]. The terms $\|\tilde{\mathbf{W}}_{D\bullet 1}\|^2$, $\|\Lambda_E^{1/2} \tilde{\mathbf{W}}_{E\bullet 1}\|^2$ and $\|\tilde{\mathbf{W}}_{E\bullet 1}\|^2$ can be expressed in a similar way.

Remark 3. We note that in the absence of an eavesdropper, i.e., $\mathbf{R}_E = 0$, one obtains the same optimality condition derived for the MIMO channel with no secrecy constraint [14]. Indeed, in this case, $\mathcal{A}$ reduces to

$$\mathcal{A} = \frac{1}{\sigma_D^2} (m_D - \frac{P\lambda_1^T}{\sigma_D^2} \sum_{k=1}^{m_D} \lambda_k^D \tau_k), \quad (20)$$

where $\tau_k = \mathbb{E} \left[\frac{\lambda_k^D w_k}{1 + \frac{P\lambda_1^T}{\sigma_D^2} \sum_{j=1}^{m_D} \lambda_j^D w_j} \right]$ and λ_k^D 's are the eigenvalues of $\mathbf{R}_D$. $\mathcal{A} > 0$, therefore, the optimality condition of beamforming reduces to evaluating the second condition in Theorem 2, which coincides with [14, Equation (14)].

We illustrate the result of Theorem 2 through the example below.

Example 1. We consider uncorrelated receive and eavesdropper antennas, and assume $\sigma_D^2 = \sigma_E^2 = \sigma^2$. Simplification of the expression of $\mathcal{A}$ yields

$$\mathcal{A} = \frac{1}{\sigma^2} (m_D + \alpha - \beta), \quad (21)$$

where $\alpha = \mathbb{E} \left[\frac{1}{1 + \frac{P\lambda_1^T}{\sigma^2} (\|\mathbf{W}_{D\bullet 1}\|^2 + \|\mathbf{W}_{E\bullet 1}\|^2)} \right]$, and $\beta = \mathbb{E} \left[\frac{1}{1 + \frac{P\lambda_2^T}{\sigma^2} \|\mathbf{W}_{E\bullet 1}\|^2} \right]$. As $0 < \beta - \alpha < 1$, then $\mathcal{A} > 0$. Thus, the necessary and sufficient condition for optimality of beamforming becomes

$$\gamma \triangleq \frac{\lambda_2^T}{\sigma^2} (m_D + \alpha - \beta) - \frac{1}{P} (\beta - \alpha) \leq 0 \quad (22)$$

$$\Leftrightarrow \frac{P\lambda_2^T}{\sigma^2} \leq \frac{(\beta - \alpha)}{(m_D + \alpha - \beta)}. \quad (23)$$

Since α involves a chi-squared random variable with $2(m_D + m_E)$ degrees of freedom, we can evaluate α in closed form as

$$\alpha = \left(\frac{\sigma^2}{P\lambda_1^T} \right)^{m_D + m_E} e^{\frac{\sigma^2}{P\lambda_1^T}} \Gamma \left(1 - m_D - m_E, \frac{\sigma^2}{P\lambda_1^T} \right), \quad (24)$$

where $\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$ is the upper incomplete Gamma function. Similarly, β can be expressed as

$$\beta = \left(\frac{\sigma^2}{P\lambda_1^T} \right)^{m_E} e^{\frac{\sigma^2}{P\lambda_1^T}} \Gamma \left(1 - m_E, \frac{\sigma^2}{P\lambda_1^T} \right). \quad (25)$$

We note that if $m_E = 0$, we end up having exactly the same condition developed for the peaceful MIMO case in [27, Equation 37]. Starting from (23), we plot in Fig. 2, the optimality condition of beamforming in the space $(P\lambda_1^T/\sigma^2, P\lambda_2^T/\sigma^2)$ for different receive and eavesdropper antenna configurations. The beamforming optimality region (BOR) for a given configuration corresponds to all the points that are below the corresponding curve in the space $(P\lambda_1^T/\sigma^2, P\lambda_2^T/\sigma^2)$. It is clear that as the number of receive and/or eavesdropper antennas increase, the BOR gets narrower. Similarly, one can show that receive correlation and eavesdropper correlation reduces the BOR. Moreover, Fig. 2 suggests a different impact of transmit correlation on the BOR than in the peaceful setup. For instance, as can be seen from Fig. 2, one may start from a non-optimal beamforming scenario, and by increasing λ_1^T , one can move to an optimal beamforming situation and then again to a non-optimal beamforming situation.

Theorem 2 provides insight into the optimality of beamforming in the low-power regime. For instance, we have $\mathcal{A} \simeq \frac{m_D}{\sigma_D^2}$ as $P \rightarrow 0$. Let γ be the LHS of (19). Thus, we have

$$\gamma \underset{P \rightarrow 0}{\simeq} \frac{\lambda_2^T m_D}{\sigma_D^2} - \frac{\lambda_1^T \mathbb{E}(\|\tilde{\mathbf{W}}_{D\bullet 1}\|^2)}{\sigma_D^2} = \frac{m_D(\lambda_2^T - \lambda_1^T)}{\sigma_D^2} \leq 0. \quad (26)$$

Thus beamforming is optimal in the low power regime. On the other hand, one can see that beamforming is suboptimal in the high-power regime. For instance, assuming for simplicity as in Example 1 that $\sigma_D^2 = \sigma_E^2 = \sigma^2$, then we have $\alpha - \beta \rightarrow 0$ as $P \rightarrow \infty$. Thus, from (23), we have that $\gamma \underset{P \rightarrow \infty}{\simeq} \frac{\lambda_2^T m_D}{\sigma_D^2} > 0$, which implies that beamforming is not optimal in the high-power regime.

In the sequel, these observations are confirmed through our numerical results and the analysis of low-power and high-power

Fig. 2. Optimality region of beamforming.

regimes.

V. IMPACT OF CORRELATION ON SECRET-KEY CAPACITY

In this section, we analyze the impact spatial correlation on the secret-key capacity.

A. Impact of spatial transmit correlation

The impact of transmit correlation depends on the operating regime. Indeed, simulations suggest that transmit correlation helps increase the secret-key capacity in the low power regime. However, as the power increases, the beneficial impact of transmit correlation reduces and eventually transmit correlation becomes detrimental. Indeed, in the high-power regime, we show later in Proposition 4 that transmit correlation induces secret-key capacity loss for all configurations of antennas. Moreover, we show in Section VII that, in the low-power regime, transmit correlation does increase the secret-key capacity. In [30], it has been shown that for a MISO system without secrecy constraint and with covariance feedback, spatial correlation improves the capacity for any power. As we show in this paper, when secrecy is taken into account, this behavior of capacity does not hold even in the presence of a single-antenna eavesdropper.

B. Impact of receive correlation

To characterize the impact of receive correlation, we will use the measure for correlation as in [30]. We say that a correlation matrix $\mathbf{R}_1$ is larger than $\mathbf{R}_2$ if the vector of eigenvalues of $\mathbf{R}_{D,1}$ majorizes the vector of eigenvalues of $\mathbf{R}_{D,2}$. Using this measure, we show that receive correlation decreases the secret-key capacity. This is formalized in the following proposition.

Proposition 1. If the receive correlation $\mathbf{R}_{D,1}$ is larger than $\mathbf{R}_{D,2}$, then the secret-key capacity is lower. The maximum is achieved for uncorrelated antennas, the minimum for completely correlated.

Proof: For fixed $\mathbf{\Lambda}_T$, $\mathbf{\Lambda}_E$ and $\mathbf{\Lambda}_Q$, the secret-key rate function, defined in (14), is Schur-concave in λ_D , where λ_D is the vector of eigenvalues of $\mathbf{R}_D$. Indeed, starting from (14), we write

$$R(\mathbf{\Lambda}_D) = R_1(\mathbf{\Lambda}_D) - R_2 \quad (27)$$

$$= \mathbb{E} \left\{ \log \left| \mathbf{I}_{m_S} + \frac{1}{\sigma_D^2} \mathbf{\Lambda}_Q^{1/2} \mathbf{\Lambda}_T^{1/2} \mathbf{W}_D^\dagger \mathbf{\Lambda}_D \mathbf{W}_D \mathbf{\Lambda}_T^{1/2} \mathbf{\Lambda}_Q^{1/2} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_Q^{1/2} \mathbf{\Lambda}_T^{1/2} \mathbf{W}_E^\dagger \mathbf{\Lambda}_E \mathbf{W}_E \mathbf{\Lambda}_T^{1/2} \mathbf{\Lambda}_Q^{1/2} \right| \right\} - R_2, \quad (28)$$

where R_2 is independent of $\mathbf{\Lambda}_D$. The rate function in (28) is concave in $\mathbf{\Lambda}_D$. Also, it is a symmetric function. Therefore, it is Schur-concave [21, Chapter 3 section C.2]. Given that, consider two receive correlation matrices $\mathbf{R}_{D,1}$ and $\mathbf{R}_{D,2}$ such that $\mathbf{R}_{D,2}$ is larger than $\mathbf{R}_{D,1}$, i.e., $\lambda_{D,2} \succ \lambda_{D,1}$. The secret-key capacity given $\mathbf{R}_{D,i}$ is denoted as $C_i = R(\mathbf{R}_{D,i}, \mathbf{R}_E, \mathbf{Q}_i^{opt})$, $i \in 1, 2$. We write

$$\begin{aligned} C_2 &= R(\mathbf{R}_{D,2}, \mathbf{R}_E, \mathbf{Q}_2^{opt}) \\ &\leq R(\mathbf{R}_{D,1}, \mathbf{R}_E, \mathbf{Q}_2^{opt}) \leq R(\mathbf{R}_{D,1}, \mathbf{R}_E, \mathbf{Q}_1^{opt}) = C_1, \end{aligned} \quad (29)$$

where the first inequality is due to the Schur-concavity property of the secret-key rate as a function of $\mathbf{R}_D$ and the second inequality is obtained by maximizing over $\mathbf{Q}$ to obtain the secret-key capacity with respect to $\mathbf{R}_{D,1}$. Therefore, the secret-key capacity is Schur-concave with respect to the receive correlation. Moreover, since the uniform vector, which is a Schur-minimal element (an element majorized by any other element) in $\Omega_{m_D} = \{\lambda_i \geq 0 \text{ and } \sum_{i=1}^{m_D} \lambda_i = m_D\}$, it follows that the secret-key capacity is largest with uncorrelated receive antennas. Likewise, the secret-key capacity is lowest for fully correlated receive antennas. ■

It is well-known that receive correlation reduces capacity in MIMO systems with no secrecy constraint [14]. Thus, one would expect the same result to hold true when secrecy constraint is added to the system.

C. Impact of the eavesdropper correlation matrix

To characterize the impact of eavesdropper correlation, we use the same measure for correlation as in the previous section. Unlike receive correlation that is shown to decrease secret-key capacity, we show that eavesdropper correlation increases secret-key capacity. Intuitively, correlation creates dependency between the observations, henceforth it is plausible that eavesdropper correlation would increase the secret-key capacity. The following proposition establishes the relevance of such intuition.

Proposition 2. If the eavesdropper correlation $\mathbf{R}_{E,1}$ is larger than $\mathbf{R}_{E,2}$, then the secret-key capacity is higher. The maximum is achieved for fully correlated antennas, the minimum for uncorrelated antennas.

Proof: The proof follows along similar steps as the proof of Proposition 1. We need to prove that for fixed $\mathbf{\Lambda}_T$, $\mathbf{\Lambda}_D$ and $\mathbf{\Lambda}_Q$, the secret-key rate, as a function of λ_E (vector of eigenvalues of $\mathbf{R}_E$) is Schur-convexe. Starting from (14), we write

$$\begin{aligned} R(\mathbf{\Lambda}_E) &= R_1(\mathbf{\Lambda}_E) - R_2(\mathbf{\Lambda}_E) \\ &= \mathbb{E} \log \left| \mathbf{I}_{m_S} + \frac{1}{\sigma_D^2} \mathbf{\Lambda}_Q^{1/2} \mathbf{\Lambda}_T^{1/2} \mathbf{W}_D^\dagger \mathbf{\Lambda}_D \mathbf{\Lambda}_T^{1/2} \mathbf{\Lambda}_Q^{1/2} \mathbf{W}_D \right. \end{aligned} \quad (30)$$

$$\begin{aligned}
& + \frac{1}{\sigma_E^2} \Lambda_Q^{1/2} \Lambda_T^{1/2} \mathbf{W}_E^\dagger \Lambda_E \mathbf{W}_E \Lambda_T^{1/2} \Lambda_Q^{1/2} \Big| \\
& - \mathbb{E} \log \left| \mathbf{I}_{m_S} + \frac{1}{\sigma_E^2} \Lambda_Q^{1/2} \Lambda_T^{1/2} \mathbf{W}_E^\dagger \Lambda_E \mathbf{W}_E \Lambda_T^{1/2} \Lambda_Q^{1/2} \right|,
\end{aligned} \tag{31}$$

where (31) results from applying the identity $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$. Let $\mathbf{B} = \frac{1}{\sigma_D^2} \Lambda_Q^{1/2} \Lambda_T^{1/2} \mathbf{W}_D^\dagger \Lambda_D \mathbf{W}_D \Lambda_T^{1/2} \Lambda_Q^{1/2}$ and $\mathbf{F} = \mathbf{W}_E \Lambda_T^{1/2} \Lambda_Q^{1/2}$. Then, the secret-key rate becomes

$$R(\Lambda_E) = \mathbb{E} \log \frac{|\mathbf{I}_{m_S} + \mathbf{B} + \mathbf{F}^\dagger \Lambda_E \mathbf{F}|}{|\mathbf{I}_{m_S} + \mathbf{F}^\dagger \Lambda_E \mathbf{F}|} \tag{32}$$

$$= \mathbb{E} \log \left| \mathbf{I} + \mathbf{B} \left(\mathbf{I} + \mathbf{F}^\dagger \Lambda_E \mathbf{F} \right)^{-1} \right|, \tag{33}$$

where (33) is obtained using the identity $\frac{|\mathbf{A}+\mathbf{B}|}{|\mathbf{A}|} = |\mathbf{I} + \mathbf{BA}^{-1}|$. Since $\mathbf{B} \succeq 0$, we know from [31, Lemma II.3] that the mapping $\mathbf{X} \mapsto \log |\mathbf{I} + \mathbf{BX}^{-1}|$ is convex in $\mathbf{X} \succ 0$. Moreover, the mapping $\Lambda_E \mapsto \mathbf{I} + \mathbf{F}^\dagger \Lambda_E \mathbf{F}$ is convex in $\Lambda_E \succeq 0$. Thus, applying the composition theorem implies that the function $R(\Lambda_E)$ is convex. Moreover, $R(\Lambda_E)$ is a symmetric function. Hence, it is Schur-convex. Similarly, consider two eavesdropper correlation matrices $\mathbf{R}_{E,1}$ and $\mathbf{R}_{E,2}$ such that $\mathbf{R}_{E,2}$ is larger than $\mathbf{R}_{E,1}$, i.e., $\lambda_{E,2} \succ \lambda_{E,1}$. The secret-key capacity given $\mathbf{R}_{E,i}$ is denoted as $C_i = R(\mathbf{R}_{E,i}, \mathbf{R}_D, \mathbf{Q}_i^{opt})$, $i \in 1, 2$. We write

$$\begin{aligned}
C_1 &= R(\mathbf{R}_{E,1}, \mathbf{R}_D, \mathbf{Q}_1^{opt}) \\
&\leq R(\mathbf{R}_{E,2}, \mathbf{R}_D, \mathbf{Q}_1^{opt}) \leq R(\mathbf{R}_{E,2}, \mathbf{R}_D, \mathbf{Q}_2^{opt}) = C_2,
\end{aligned} \tag{34}$$

where the first inequality is due to the Schur-convexity property of the secret-key rate as a function of $\mathbf{R}_E$ and the second inequality is obtained by maximizing over $\mathbf{Q}$ to obtain the secret-key capacity with respect to $\mathbf{R}_{E,2}$. The results of the proposition follows immediately. ■

Remark 4. Unlike transmit correlation which is beneficial in the low power regime, eavesdropper correlation is helpful regardless of the operating power regime. The more correlated eavesdropper antennas, the higher the secret-key capacity.

Remark 5. We note that the vector $\lambda_E^{max} = \{m_E, 0, \dots, 0\}$ majorizes every vector in Ω_{m_E} . Hence, it follows from the Schur-convexity property proved in Proposition 2, that the secret-key capacity is highest for fully correlated eavesdropper antennas, i.e., for $\lambda_E = \lambda_E^{max}$. Any permutation of λ_E^{max} achieves likewise the same secret-key rate as the secret-key rate function is symmetric in λ_E .

Example 2. We now provide numerical simulations illustrating our results. In Fig. 3, we study the combined impact of all correlation matrices. In this configuration, the transmitter, the legitimate receiver and the eavesdropper are equipped with two antennas, three antennas and two antennas, respectively. The correlation matrices are $\Lambda_T = \text{diag}[1.7, 0.3]$, $\Lambda_D = \text{diag}[1.8, 0.9, 0.3]$ and $\Lambda_E = \text{diag}[1.2, 0.8]$. Note that the secret-key capacity is obtained by solving (17) through an exhaustive search. In the low power regime, we experience a gain in secret-key capacity with respect to the case of uncorrelated nodes and we observe that beamforming is optimal. Nonetheless, in the high-power, we experience a loss in capacity, which is asymptotically equal to 0.08 bits per channel use. Moreover, a uniform power allocation is optimal in the high-power regime.

Fig. 3. Impact of combined correlation on secret-key rates. The transmitter, the destination and the eavesdropper are equipped with 2, 3 and 2 antennas, respectively: weak correlation at the eavesdropper.

In Fig. 4, we keep the same configuration. The only change is that the eavesdropper experiences a larger correlation with $\Lambda_E = \text{diag}[1.9, 0.1]$. In the low power regime, the same observations holds true. Beamforming is secret-key capacity-achieving with a gain in capacity. The new effect lies in the high-power regime. We experience also a substantial gain in capacity, that is asymptotically equal to 2 bits per channel use. This gain is explained by the heavily correlated eavesdropper matrix that balances the negative impact of transmit and receive correlation matrices. Therefore, depending on the eigenvalues' distribution of correlation matrices, one can experience either a gain, or loss or no impact on secret-key capacity in the high-power regime. In the next section, we evaluate analytically this impact on secret-key capacity in the high-power regime in terms of system parameters.

VI. HIGH-POWER ANALYSIS

In this section, we study the impact of correlation matrices on secret-key capacity in the high-power regime. For this purpose, the secret-key capacity will be evaluated using a uniform power allocation in the high-power regime. The justification of this choice is twofold. First, in case $m_E \geq m_S$, assuming the optimal covariance matrix is of full rank in the high-power regime (otherwise one cannot achieve the maximum secret-key multiplexing gain), we show in Proposition 3 that any non-singular covariance matrix achieves the secret-key capacity. In particular, a uniform power allocation strategy is asymptotically optimal. Second, in case $m_E < m_S$, the secret-key capacity scales with P . Determining the optimal input covariance matrix is not straightforward. Based on our simulation results, we observe that uniform power allocation performs comparatively very close to the optimal solution obtained through an exhaustive search in most configurations. Henceforth, we assume that uniform power allocation is secret-key capacity-achieving in the high-power regime.

Fig. 4. Impact of combined correlation on secret-key rates. The transmitter, the destination and the eavesdropper are equipped with 2, 3 and 2 antennas, respectively: strong correlation at the eavesdropper.

Proposition 3. If $m_E \geq m_S$, assuming $\mathbf{R}_T$ and $\mathbf{R}_E$ are invertible, the secret-key capacity in the high-power regime is given by

$$\lim_{P \rightarrow \infty} C(P) = \mathbb{E} \left\{ \frac{\left| \mathbf{W}_E^\dagger \mathbf{\Lambda}_E \mathbf{W}_E + \frac{\sigma_E^2}{\sigma_D^2} \mathbf{W}_D^\dagger \mathbf{\Lambda}_D \mathbf{W}_D \right|}{\left| \mathbf{W}_E^\dagger \mathbf{\Lambda}_E \mathbf{W}_E \right|} \right\}. \quad (35)$$

Moreover, it is achieved by any non-singular input covariance matrix.

Proof: See Appendix B. ■

We note that (35) generalizes [11, Equation 30]. Based on this result, we have that uniform power allocation is secret-key capacity-achieving in the high-power regime in case when the eavesdropper possess as many antennas as the transmitter. We now study the impact of correlation matrices. In the next proposition, we show that transmit correlation is harmful in the high-power regime.

Proposition 4. In case $m_E > m_S$, in the high-power regime, if the transmit correlation $\mathbf{R}_{T,1}$ is larger than $\mathbf{R}_{T,2}$, then the secret-key capacity is lower. The maximum is achieved for uncorrelated antennas, the minimum for completely correlated.

Proof: In case $m_E > m_S$, Proposition 4 shows that uniform power allocation is optimal in the high-power regime. Given that, the secret-key capacity is written as

$$C = \mathbb{E} \left[\log \frac{\left| \mathbf{I}_{m_D+m_E} + \frac{P}{m_S} \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda}_T \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2} \right|}{\left| \mathbf{I}_{m_E} + \frac{P}{m_S} \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda}_T \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2} \right|} \right]. \quad (36)$$

As a function of $\mathbf{\Lambda}_T$, the right-hand side of (36), denoted $f(\lambda_T)$ with λ_T containing the elements of the diagonal matrix

Λ_T , can be shown to be Schur-concave in λ_T . Indeed, let $\mathbf{Z}_D = \Lambda_D^{1/2} \mathbf{W}_D$ and $\mathbf{Z}_E = \Lambda_E^{1/2} \mathbf{W}_E$. One can show (the same way one obtains (6) from (5)) the following equality

$$f(\lambda_T) = \mathbb{E} \{ \log |\mathbf{A} - \mathbf{B} \mathbf{C}^{-1} \mathbf{D}| \} - m_D \log(\sigma_D^2), \quad (37)$$

where $\mathbf{A} = \sigma_D^2 \mathbf{I} + \mathbf{Z}_D \Lambda_T \mathbf{Z}_D^\dagger$, $\mathbf{B} = \mathbf{Z}_D \Lambda_T \mathbf{Z}_E^\dagger$, $\mathbf{C} = \sigma_E^2 \mathbf{I} + \mathbf{Z}_E \Lambda_T \mathbf{Z}_E^\dagger$, $\mathbf{D} = \mathbf{B}^\dagger$. Using similar arguments as in [11, Lemma 2], the mapping f is concave in $\Omega = \{\lambda_T : \lambda_i \geq 0 \text{ for } i = 1, 2, \dots, m_S\}$. Moreover, f is symmetric. Therefore, f is Schur-concave. Proceeding similarly as in the proof of Proposition 2, it is clear that the larger the transmit correlation is (in the sense of [30]), the lower the secret-key capacity is. The maximum is achieved for uncorrelated transmit antennas and the minimum for completely correlated transmit antennas. Finally, we note that, based on the near-optimal behavior of uniform power allocation in the high-power regime, the previous analysis can be justifiably be applied to the case $m_S \geq m_E$. ■

Remark 6. Proposition 4 characterizes the impact of transmit correlation in the high-power regime. In particular, we note that in the high-power regime, in case of no eavesdropper correlation, i.e., $\mathbf{R}_E = \mathbf{I}_{m_E}$, we can not experience a gain in the secret-key capacity as both transmit correlation and receive correlation are detrimental. This observation is similar to the correlated MIMO channel without secrecy constraint [14], in which the high-power correlated capacity can not exceed the uncorrelated capacity.

In the general case, when eavesdropper correlation comes into the picture, its beneficial impact may balance out the detrimental effect of transmit and receive correlations, and we may experience considerable secret-key rate gain as shown in Fig. 4. Assuming non-singular correlation matrices, let $\Delta(P) = C_{uncor}(P) - C(P)$ denote the difference between secret-key capacities without and with spatial correlation. Using again a uniform power allocation strategy to approximate the secret-key capacity, Theorem 3 characterizes the behavior of $\Delta(P)$ in the high-power regime.

Theorem 3. For non-singular correlation matrices, $\mathbf{R}_T$, $\mathbf{R}_D$ and $\mathbf{R}_E$, we have

$$\lim_{P \rightarrow \infty} \Delta(P) = \begin{cases} \mathbb{E} \log \frac{|\mathbf{W}_E^\dagger \Lambda_E \mathbf{W}_E|}{|\mathbf{W}_E^\dagger \mathbf{W}_E|} - \mathbb{E} \log \frac{|\mathbf{W}^\dagger \Lambda_{DE} \mathbf{W}|}{|\mathbf{W}^\dagger \mathbf{W}|} & \text{if } m_E \geq m_S \\ -\log |\Lambda_T| + \log |\Lambda_E| + \mathbb{E} \log \frac{|\mathbf{W}_E \Lambda_T \mathbf{W}_E^\dagger|}{|\mathbf{W}_E \mathbf{W}_E^\dagger|} \\ \quad - \mathbb{E} \log \frac{|\mathbf{W}^\dagger \Lambda_{DE} \mathbf{W}|}{|\mathbf{W}^\dagger \mathbf{W}|} & \text{if } m_E < m_S \leq m_E + m_D \\ -\log |\Lambda_D| - \mathbb{E} \log \frac{|\mathbf{W} \Lambda_T \mathbf{W}^\dagger|}{|\mathbf{W} \mathbf{W}^\dagger|} + \mathbb{E} \log \frac{|\mathbf{W}_E \Lambda_T \mathbf{W}_E^\dagger|}{|\mathbf{W}_E \mathbf{W}_E^\dagger|} & \text{else.} \end{cases} \quad (38)$$

Proof: See Appendix C. ■

Theorem 3 quantifies the combined effect of correlation matrices in the high-power regime. Evaluating the gap expression in Theorem 3 determines whether the detrimental impact of both transmit and receive correlation or the positive impact of eavesdropper correlation dominates in the high-power regime.

Remark 7. In the high-power regime, likewise a MIMO channel with no secrecy constraint, while transmit correlation does not affect the degrees of freedom under a trace constraint, the transmit processing gain is undermined due to spatial correlation. Hence, transmit correlation is harmful in this regime. This impact is not true in the low-power regime, as we explain in the

following section.

VII. LOW-POWER ANALYSIS

In this section, we present a summary of some results of [32]. We identify the first and second derivatives of the fast-fading secret-key capacity at SNR=0 in terms of the correlation matrices. Based on that, we provide a low-SNR second-order approximation to the secret-key capacity. Moreover, we evaluate the minimum energy required for sharing reliably a secret-key bit as well as the wideband slope and their dependence on each of the correlation matrices. We define the signal-to-noise ration (SNR), as $\text{SNR} = \frac{P}{\sigma_D^2}$. Hereafter, we express the secret-key capacity as a function of SNR.

A. First and second derivatives

Determining the first and second derivatives with respect to SNR at SNR= 0 allows us to accurately approximate the secret-key capacity function as follows [17]

$$C(\text{SNR}) = \dot{C}(0)\text{SNR} + \frac{\ddot{C}(0)}{2}\text{SNR}^2 + o(\text{SNR}^2). \quad (39)$$

For a given covariance matrix $\mathbf{Q}$ satisfying the power constraint, we introduce the standardized covariance matrix

$$\hat{\mathbf{Q}} = P^{-1} \mathbf{Q} \text{ such that } \text{Tr}(\hat{\mathbf{Q}}) \leq 1. \quad (40)$$

Thus, the secret-key capacity, in terms of $\hat{\mathbf{Q}}$ and SNR, becomes

$$C(\text{SNR}) = \max_{\text{Tr}(\hat{\mathbf{Q}}) \leq 1} \mathbb{E}[\mathbb{I}(\text{SNR})], \quad (41)$$

where

$$\mathbb{I}(\text{SNR}) = \log \frac{|\mathbf{I}_{m_D+m_E} + \text{SNR} \hat{\mathbf{H}} \hat{\mathbf{Q}} \hat{\mathbf{H}}^\dagger|}{|\mathbf{I}_{m_E} + \text{SNR} \hat{\mathbf{H}}_E \hat{\mathbf{Q}} \hat{\mathbf{H}}_E^\dagger|}, \quad (42)$$

where $\hat{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_D \\ \mathbf{H}_E \end{bmatrix}$ and $\hat{\mathbf{H}}_E = \frac{\sigma_D}{\sigma_E} \mathbf{H}_E$. For a given input covariance matrix $\hat{\mathbf{Q}}$, following the steps of [19], the first and second derivatives of the expression (42) at SNR= 0 are

$$\dot{\mathbb{I}}(0) = \text{Tr}(\hat{\mathbf{H}} \hat{\mathbf{Q}} \hat{\mathbf{H}}^\dagger) - \text{Tr}(\hat{\mathbf{H}}_E \hat{\mathbf{Q}} \hat{\mathbf{H}}_E^\dagger) \quad (43)$$

$$\ddot{\mathbb{I}}(0) = -\text{Tr}((\hat{\mathbf{H}} \hat{\mathbf{Q}} \hat{\mathbf{H}}^\dagger)^2) + \text{Tr}((\hat{\mathbf{H}}_E \hat{\mathbf{Q}} \hat{\mathbf{H}}_E^\dagger)^2). \quad (44)$$

It follows that

$$C = \max_{\text{Tr}(\hat{\mathbf{Q}}) \leq 1} \mathbb{E} \left[\dot{\mathbb{I}}(0)\text{SNR} + \ddot{\mathbb{I}}(0) \frac{\text{SNR}^2}{2} \right] + o(\text{SNR}^2). \quad (45)$$

From (45), the first derivative of the secret-key capacity is expressed as

$$\dot{C}(0) = \max_{\substack{\text{Tr}(\mathbf{Q}) \leq 1 \\ \mathbf{Q} \succeq 0}} \mathbb{E} [\dot{\mathbb{I}}(0)] \quad (46)$$

Theorem 4. The first derivative of the secret-key capacity in (6) with respect to SNR at SNR= 0 is

$$\dot{C}(0) = m_D \lambda_1^T, \quad (47)$$

where λ_1^T denotes the maximum eigenvalue of $\mathbf{R}_T$. Moreover, $\dot{C}(0)$ can be achieved by any input covariance matrix of the form

$$\mathbf{Q} = P \sum_{i=1}^L \alpha_i \mathbf{u}_i \mathbf{u}_i^\dagger, \quad (48)$$

where $\mathbf{u}_i$ are the eigenvectors that span the maximal-eigenvalue eigenspace of $\mathbf{R}_T$, α_i are non-negative coefficients that sum to unity and L denotes the multiplicity of λ_1^T .

Proof: See Appendix D. ■

Thus, transmitting in the maximal-eigenvalue eigenspace is necessary to achieve $\dot{C}(0)$, regardless of how power is distributed between the different eigenvectors associated with λ_1^T .

Remark 8. Unlike its negative impact in the high-power regime shown in Proposition 4, Theorem 4 shows that the larger transmit correlation is, the higher the secret-key capacity is. This can be explained by the observation that beamforming is optimal in the low-power regime. The channel gain associated with beamforming is given by λ_1^T . Moreover, λ_1^T is higher for larger transmit correlation matrix, hence the result.

We now consider the second derivative. For that, we assume the covariance matrix is chosen as in (48) and our result is summarized in the following Theorem 5.

Theorem 5. The second derivative of the secret-key capacity in (6) with respect to SNR at SNR= 0 is

$$\ddot{C}(0) = -\frac{\lambda_1^{T^2}}{L} \left[L \text{Tr}(\mathbf{R}_D^2) + 2 \frac{\sigma_D^2}{\sigma_E^2} m_D m_E + m_D^2 \right]. \quad (49)$$

Moreover, the optimal input covariance matrix is the form

$$\mathbf{Q} = P \sum_{i=1}^L \frac{1}{L} \mathbf{u}_i \mathbf{u}_i^\dagger \quad (50)$$

Proof: See Appendix E. ■

This says that it is necessary to distribute power equally in each orthogonal direction in the maximal-eigenvalue eigenspace of $\mathbf{R}_T$ for the second derivative to be maximized. If λ_1^T has multiplicity one, $\dot{C}(0)$ is achieved with $\mathbf{Q} = P \mathbf{u}_1 \mathbf{u}_1^\dagger$ where $\mathbf{u}_1$ is the eigenvector that corresponds to λ_1^T . Thus, beamforming in the direction in which $\mathbf{R}_T$ is maximized is optimal in the sense of achieving the first and the second derivatives of the secret-key capacity in the low-SNR regime.

Remark 9. We note that one can derive explicit achievable schemes in the low-power regime such as an on-off signaling [17]. Due to space limitation, we do not follow this direction in this paper.

B. Minimum energy per secret-key bit and wideband slope

The first and second derivatives of the secret-key capacity at vanishing SNR are closely related to the energy performance measures, namely the minimum secret-key bit energy (energy required for sharing a secret-key bit) and the wideband slope. We start by stating the following lemma.

Lemma 1. The secret-key capacity achieved under the average power constraint $\mathbb{E}\|\mathbf{x}\|^2 \leq P$ is a concave function of SNR.

Proof: A time-sharing argument as in [18, Appendix C] can be used to prove the result. ■

The energy per secret-key bit normalized by the noise variance at the legitimate receiver is defined

$$\frac{E_b}{N_{0\ sk}} \triangleq \frac{\text{SNR}}{C(\text{SNR})} \log 2. \quad (51)$$

By virtue of Lemma 1, the secret-key capacity is a concave function of SNR. Therefore, the minimum secret-key bit energy is achieved in the limit of low SNR; i.e., $\text{SNR} \rightarrow 0$. Thus, the minimum secret-key bit energy and the wideband slope (in bit/s/Hz/(3 dB)) at the point $\frac{E_b}{N_{0\ sk, \min}}$ are given by [15, Theorem 5]

$$\frac{E_b}{N_{0\ sk, \min}} = \frac{\log 2}{\dot{C}(0)}, \quad \mathcal{S}_0 = \frac{2 [\dot{C}(0)]^2}{-\ddot{C}(0)}. \quad (52)$$

Using the expressions of (47) and (49), we obtain:

$$\frac{E_b}{N_{0\ sk, \min}} = \frac{\log 2}{m_D \lambda_1^T}. \quad (53)$$

$$\mathcal{S}_0 = \frac{2m_D^2 L}{L \text{Tr}(\mathbf{R}_D^2) + 2 \frac{\sigma_D^2}{\sigma_E^2} m_D m_E + m_D^2}. \quad (54)$$

Obviously, the receive and eavesdropper correlations have no impact of the minimum $\frac{E_b}{N_{0\ sk}}$. Moreover, $\frac{E_b}{N_{0\ sk}}$ is Schur-concave with respect to transmit correlation. This follows from the Schur-convexity of the application: $\lambda \rightarrow \max(\lambda)$. Therefore, the more correlated the transmit antennas, the lower $\frac{E_b}{N_{0\ sk}}$. Given that, we note that the lowest possible value for $\frac{E_b}{N_{0\ sk}}$ is achieved under fully correlated transmit antennas, which yields $\left(\frac{E_b}{N_{0\ sk, \min}}\right)_{\min} = \frac{\log 2}{m_D m_S}$ ¹. On the other hand, the wideband slope is independent of the eavesdropper correlation. It is Schur-concave with respect to the receive correlation. This follows from the Schur-convexity of the dispersion measure of Hermitian matrices $\zeta(\mathbf{A}) = \text{Tr}(\mathbf{A}^2)$ [33]. Thus, $\mathcal{S}_0$ is highest for uncorrelated receive antennas. Moreover, $\mathcal{S}_0$ depends on the transmit correlation through the multiplicity of the maximum eigenvalue. The higher the multiplicity is, the higher the wideband slope is. Therefore, transmit correlation (i.e., smaller L) reduces the wideband slope.

¹While in Theorem 1 we require the transmit correlation matrix to be full rank, we note that the low SNR analysis is valid for any given positive semi-definite transmit correlation matrix.

Fig. 5. Secret key rates in bit/s/Hz vs. SNR. The transmitter, the destination and the eavesdropper have 2, 2 and 2 antennas, respectively.

Now, we study the effect of each correlation matrix on the energy performance of the system. In Fig. 5, we visualize the effect of receive correlation through the spectral efficiency versus $\frac{E_b}{N_{0,sk}}$ curve. For that, we considered a system with uncorrelated transmit and uncorrelated eavesdropper antennas, while the legitimate receiver experiences a correlation of the form

$$\mathbf{R}_D = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}, \text{ such that } \lambda_D = \text{diag}[1 + \alpha, 1 - \alpha]. \quad (55)$$

We increase α from 0 to 1 by a step of 0.1, i.e., the receive antennas are getting more correlated as α increases. Note that the secret-key capacity is obtained simply by a uniform power allocation, as the transmit antennas are uncorrelated. Also, the minimum secret-key energy is independent of the receive correlation, which explains the observation that all curves start at the same point $\frac{E_b}{N_{0,sk,min}} = \frac{\log 2}{m_D} = -4.6020$ dB. The dependence of the wideband slope on the receive correlation can also be visualized for values close to $\frac{E_b}{N_{0,sk,min}}$. Indeed, the different curves start diverging due to the impact of receive correlation. Moreover, as the secret-key capacity value increases and as the correlation factor α increases, the energy loss due to receive correlation also increases.

In Fig. 6, we reverse the situation. We keep the same number of antennas and we consider the case of uncorrelated transmit and uncorrelated receive antennas while the eavesdropper antennas are correlated as in (55). As expected, the minimum energy per bit remains the same for all curves as it is independent of the eavesdropper correlation. Also, the non-dependence of the wideband slope on the eavesdropper correlation is clear as all curves start diverging at a value above $\frac{E_b}{N_{0,sk,min}}$. Besides, as the secret-key capacity value increases, the energy required per secret-key bit increases as the eavesdropper correlation decreases. Hence, eavesdropper correlation improves the performance of the system, which confirms the result of Proposition 2.

Finally, we visualize the effect of transmit correlation of the energy performance. Fig. 7 confirms that transmit correlation

Fig. 6. Secret key rates in bit/s/Hz vs. SNR. The transmitter, the destination and the eavesdropper have 2, 2 and 2 antennas, respectively.

Fig. 7. Secret key rates in bit/s/Hz vs. SNR. The transmitter, the destination and the eavesdropper has 2, 2 and 2 antennas, respectively.

helps decrease the minimum secret-key bit energy, as discussed earlier. The minimum of the minimum energy per secret-key bit is obtained under fully correlated transmit antennas, and equals $\left(\frac{E_b}{N_0 \text{ sk,min}}\right)_{\min} = \frac{\log 2}{m_D m_S} = -7.6123$ dB. As the secret-key capacity value increases, the benefit of transmit correlation decreases gradually and induces a loss in the energy performance compared to the uncorrelated case.

VIII. CONCLUSIONS

In this paper, we have considered secret-key agreement over Rayleigh fast-fading MIMO channels in a correlated environment. We have determined the expression of the secret-key capacity. Then, we have determined the optimal transmit strategy achieving the secret-key capacity. In particular, we have shown that transmitting independent Gaussian signals along the eigenvectors of the transmit correlation matrix is optimal. The corresponding power allocated to each channel direction can be determined as the solution to a numerical optimization problem. Next, we have derived a necessary and sufficient condition for beamforming (i.e., transmitting along the strongest channel mode) to be capacity-achieving. In addition, we have analyzed the impact of correlation matrices on secret-key capacity. For instance, we have proved that eavesdropper correlation increases secret-key capacity, receive correlation decreases capacity, while the effect of transmit correlation depends on the operating power regime. Moreover, we have investigated the system in both extreme power regimes. In the high-power regime, we have argued that uniform power allocation is optimal, and quantified in closed-forms the gain/loss induced by spatial correlation matrices on the secret-key capacity. In the low power regime, we have derived a second order approximation of the secret-key capacity and determined explicitly the related optimal transmission schemes. In particular, we have proved that achieving the first and the second derivatives requires a uniform power distribution between the orthogonal eigenvectors spanning the maximal-eigenvalue eigenspace of the transmit correlation matrix. Moreover, we have shown that transmit correlation improves the secret-key capacity in this regime. From an energy efficiency point of view, we have determined the minimum energy required for secret-key agreement and proved that is Schur-concave with respect to transmit correlation and it is independent of receive and eavesdropper correlation. Beyond the minimum energy point, eavesdropper correlation improves the spectral efficiency of the system, receive correlation decreases it while transmit correlation impact becomes detrimental as the transmit power increases.

APPENDICES

A. Necessary and sufficient condition for optimality of beamforming

Proof: The proof builds upon the techniques used in [27]. Suppose we allocate power $P - p$ to the dominant eigenvalue of $\mathbf{R}_T$ and distribute the rest of the power budget among the remaining eigenvalues so that $\lambda_i^Q = \alpha_i p, i \in \{2, 3, \dots, m_S\}$. The convex combination coefficients α_i are all nonnegative and sum to one. The achievable secret-key rate with such power allocation is given by

$$R(p) = R_1(p) - R_2(p), \quad (56)$$

where

$$R_1(p) = \mathbb{E} \left[\log \left| \mathbf{I}_{m_D+m_E} + \mathbf{\Lambda}_{DE}^{1/2} \mathbf{W} \mathbf{\Lambda}_Q \mathbf{\Lambda}_T \mathbf{W}^\dagger \mathbf{\Lambda}_{DE}^{1/2} \right| \right] \quad (57)$$

$$= \mathbb{E} \left[\log \left| \mathbf{I}_{m_D+m_E} + (P-p) \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger + \sum_{i=2}^{m_S} \lambda_i^T p \alpha_i \tilde{\mathbf{W}}_{\bullet i} \tilde{\mathbf{W}}_{\bullet i}^\dagger \right| \right] \quad (58)$$

$$R_2(p) = \mathbb{E} \left[\log \left| \mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E^{1/2} \mathbf{W}_E \mathbf{\Lambda}_Q \mathbf{\Lambda}_T \mathbf{W}_E^\dagger \mathbf{\Lambda}_E^{1/2} \right| \right] \\ = \mathbb{E} \left[\log \left| \mathbf{I}_{m_E} + \frac{(P-p) \lambda_1^T}{\sigma_E^2} \tilde{\mathbf{W}}_{E\bullet 1} \tilde{\mathbf{W}}_{E\bullet 1}^\dagger + \sum_{i=2}^{m_S} \frac{p \alpha_i \lambda_i^T}{\sigma_E^2} \tilde{\mathbf{W}}_{E\bullet i} \tilde{\mathbf{W}}_{E\bullet i}^\dagger \right| \right]. \quad (59)$$

To derive a necessary condition for optimality, we need to compute $\frac{\partial R(p)}{\partial p}$. Indeed, if beamforming is optimal, then we should have $\frac{\partial R(p)}{\partial p} \Big|_{p=0} \leq 0$. Now, we have

$$\frac{\partial R(p)}{\partial p} \Big|_{p=0} = \frac{\partial R_1(p)}{\partial p} \Big|_{p=0} - \frac{\partial R_2(p)}{\partial p} \Big|_{p=0}. \quad (60)$$

The derivative of R_1 can be obtained following the steps of [27]. However, as we are dealing with different noise variances in general, some care needs to be taken in the derivations.

$$\frac{\partial R_1(p)}{\partial p} \Big|_{p=0} \\ = \text{Tr} \left\{ \mathbb{E} \left[\left(\mathbf{I}_{m_D+m_E} + P \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right)^{-1} \right] \mathbb{E} \left[\sum_{i=2}^{m_S} \alpha_i \lambda_i^T \tilde{\mathbf{W}}_{\bullet i} \tilde{\mathbf{W}}_{\bullet i}^\dagger \right] \right. \\ \left. - \mathbb{E} \left[\left(\mathbf{I}_{m_D+m_E} + P \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right)^{-1} \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right] \right\} \quad (61)$$

$$= \text{LHS1} - \text{RHS2}. \quad (62)$$

$$= \text{LHS1} - \text{RHS2}. \quad (63)$$

where

LHS1

$$= \text{Tr} \left\{ \mathbb{E} \left[\left(\mathbf{I} + P \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right)^{-1} \right] \mathbb{E} \left[\sum_{i=2}^{m_S} \alpha_i \lambda_i^T \tilde{\mathbf{W}}_{\bullet i} \tilde{\mathbf{W}}_{\bullet i}^\dagger \right] \right\} \quad (64)$$

$$= \sum_{i=2}^{m_S} \alpha_i \lambda_i^T \mathbb{E} \left[\text{Tr} \left\{ \left(\mathbf{I} + P \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right)^{-1} \begin{pmatrix} \frac{1}{\sigma_D^2} \mathbf{\Lambda}_D & 0 \\ 0 & \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E \end{pmatrix} \right\} \right] \quad (65)$$

$$= \sum_{i=2}^{m_S} \alpha_i \lambda_i^T \mathbb{E} \left[\frac{m_D}{\sigma_D^2} + \frac{m_E}{\sigma_E^2} - \frac{P \lambda_1^T \left[\frac{\|\mathbf{\Lambda}_D^{1/2} \tilde{\mathbf{W}}_{D\bullet 1}\|^2}{\sigma_D^4} + \frac{\|\mathbf{\Lambda}_E^{1/2} \tilde{\mathbf{W}}_{E\bullet 1}\|^2}{\sigma_E^4} \right]}{1 + P \lambda_1^T \left(\frac{\|\tilde{\mathbf{W}}_{D\bullet 1}\|^2}{\sigma_D^2} + \frac{\|\tilde{\mathbf{W}}_{E\bullet 1}\|^2}{\sigma_E^2} \right)} \right], \quad (66)$$

where (66) is obtained using the Sherman-Morrison formula again and then evaluating the trace of the involved product of matrices. The derivation of RHS1 below follows along similar lines as RHS2.

$$\text{RHS1} = \text{Tr} \left\{ \mathbb{E} \left[\left(\mathbf{I} + P \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right)^{-1} \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right] \right\} \quad (67)$$

$$= \text{Tr} \left\{ \mathbb{E} \left[\left(\mathbf{I} + P \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right)^{-1} \left(\mathbf{I} + P \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger - \mathbf{I} \right) \frac{1}{P} \right] \right\} \quad (68)$$

$$= \frac{1}{P} \mathbb{E} \left[\text{Tr} \left\{ \mathbf{I} - \left(\mathbf{I} + P \lambda_1^T \tilde{\mathbf{W}}_{\bullet 1} \tilde{\mathbf{W}}_{\bullet 1}^\dagger \right)^{-1} \right\} \right] \quad (69)$$

$$= \frac{1}{P} \left\{ 1 - \mathbb{E} \left[\frac{1}{1 + P \lambda_1^T \|\tilde{\mathbf{W}}_{\bullet 1}\|^2} \right] \right\} \quad (70)$$

$$= \frac{1}{P} \left\{ 1 - \mathbb{E} \left[\frac{1}{1 + P\lambda_1^T \left(\frac{\|\tilde{\mathbf{W}}_{D\bullet 1}\|^2}{\sigma_D^2} + \frac{\|\tilde{\mathbf{W}}_{E\bullet 1}\|^2}{\sigma_E^2} \right)} \right] \right\}. \quad (71)$$

In a similar way, we write

$$\left. \frac{\partial R_2(p)}{\partial p} \right|_{p=0} = \text{LHS2} - \text{RHS2}, \quad (72)$$

where

$$\text{LHS2} = \sum_{i=2}^{m_S} \alpha_i \frac{\lambda_i^T}{\sigma_E^2} \mathbb{E} \left[m_E - \frac{\frac{P\lambda_1^T}{\sigma_E^2} \|\Lambda_E^{1/2} \tilde{\mathbf{W}}_{E\bullet 1}\|^2}{1 + \frac{P\lambda_1^T}{\sigma_E^2} \|\tilde{\mathbf{W}}_{E\bullet 1}\|^2} \right] \quad (73)$$

$$\text{RHS2} = \frac{1}{P} \left\{ 1 - \mathbb{E} \left[\frac{1}{1 + \frac{P\lambda_1^T}{\sigma_E^2} \|\tilde{\mathbf{W}}_{E\bullet 1}\|^2} \right] \right\}. \quad (74)$$

Finally, we obtain

$$\left. \frac{\partial R(p)}{\partial p} \right|_{p=0} = \text{LHS1} - \text{RHS1} - \text{LHS2} + \text{RHS2}. \quad (75)$$

Factoring (75) yields

$$\begin{aligned} \left. \frac{\partial R(p)}{\partial p} \right|_{p=0} = & \mathcal{A} \sum_{i=2}^{m_S} \alpha_i \lambda_i^T - \frac{1}{P} \left\{ \mathbb{E} \left[\frac{1}{1 + \frac{P\lambda_1^T}{\sigma_E^2} \|\tilde{\mathbf{W}}_{E\bullet 1}\|^2} \right] \right. \\ & \left. - \mathbb{E} \left[\frac{1}{1 + P\lambda_1^T \left(\frac{\|\tilde{\mathbf{W}}_{D\bullet 1}\|^2}{\sigma_D^2} + \frac{\|\tilde{\mathbf{W}}_{E\bullet 1}\|^2}{\sigma_E^2} \right)} \right] \right\}, \end{aligned} \quad (76)$$

where $\mathcal{A}$ is defined in (19), for convenience. Recall that if beamforming is optimal, then we necessarily have $\left. \frac{\partial R(p)}{\partial p} \right|_{p=0} \leq 0$ for all positive combination α_i that sum to unity. In particular, we can check easily from the expression of $\left. \frac{\partial R(p)}{\partial p} \right|_{p=0}$ that in the case $\mathcal{A} \leq 0$, the condition $\left. \frac{\partial R(p)}{\partial p} \right|_{p=0} \leq 0$ is always satisfied. Thus, the necessary condition is always satisfied in such scenarios. In the other case where $\mathcal{A} > 0$, maximizing over all possible α_i is equivalent to considering $\alpha_2 = 1$ and the rest of the α_i are zero. That means the next strongest eigenvector is allocated the power p . Thus, the necessary condition for optimality of beamforming in such scenarios can be expressed as (18). Next we show that the necessary condition is also sufficient. For that, we evaluate the second derivative $\frac{\partial^2 R(p)}{\partial p^2}$ and prove that the second derivative is nonpositive for all p and all α_i . This implies the sufficiency of the necessary condition for optimality of beamforming obtained from $\left. \frac{\partial R(p)}{\partial p} \right|_{p=0} \leq 0$. Indeed, such condition eliminates the existence of a maximum of $R(p)$ outside $p = 0$ for all α_i that sum to unity. To show that, we could compute the second derivative starting from (56) and then using the results of [27] with proper modifications. However, this would result in a cumbersome long expression. Instead, we prove the non-positivity of $\frac{\partial^2 R(p)}{\partial p^2}$ using a concavity argument given by the following lemma.

Lemma 2. The function $R(p)$, defined in (56), is concave in p over the interval $[0, P]$.

Proof: We define the function $f(\Lambda_Q)$ given by

$$f(\Lambda_Q) = \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \Lambda_{DE}^{1/2} \mathbf{W} \Lambda_T \Lambda_Q \mathbf{W}^\dagger \Lambda_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \Lambda_E^{1/2} \mathbf{W}_E \Lambda_T \Lambda_Q \mathbf{W}_E^\dagger \Lambda_E^{1/2}|} \right], \quad (77)$$

We also define the function $g(p)$ over the interval $[0, P]$ as $g(p) = \text{diag}[P - p, \alpha_2 p, \dots, \alpha_{m_S} p]$. Thus, we have $R(p) = f \circ g(p)$. Proceeding as in the proof of Proposition 4, we obtain that f is concave (we let $\mathbf{Z}_D = \Lambda_D^{1/2} \mathbf{W}_D \Lambda_T^{1/2}$ and $\mathbf{Z}_E = \Lambda_E^{1/2} \mathbf{W}_E \Lambda_T^{1/2}$). Moreover, the function g is affine in p . Therefore, applying the composition theorem [34] implies that the mapping $R(p)$ is concave in $[0, P]$. ■

Based on Lemma 2, we conclude that $\frac{\partial^2 R(p)}{\partial p^2} \leq 0$, which proves that the necessary condition is also sufficient. ■

B. Asymptotic expression of the secret-key capacity

Proof: From (13), the secret-key capacity can be written as $C = \max_{\text{Tr}(\Lambda_Q) \leq P} R(\Lambda_Q)$

$$\begin{aligned} R(\Lambda_Q) &= \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \Lambda_{DE}^{1/2} \mathbf{W} \Lambda_T \Lambda_Q \mathbf{W}^\dagger \Lambda_{DE}^{1/2}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \Lambda_E^{1/2} \mathbf{W}_E \Lambda_T \Lambda_Q \mathbf{W}_E^\dagger \Lambda_E^{1/2}|} \right] \\ &= \mathbb{E} \left[\log \frac{|\mathbf{I}_{m_D+m_E} + \Lambda_T \Lambda_Q \mathbf{W}^\dagger \Lambda_{DE} \mathbf{W}|}{|\mathbf{I}_{m_E} + \frac{1}{\sigma_E^2} \Lambda_T \Lambda_Q \mathbf{W}_E^\dagger \Lambda_E \mathbf{W}_E|} \right] \end{aligned} \quad (78)$$

$$\begin{aligned} &= \mathbb{E} \left[\log \left| \mathbf{I}_{m_S} + \left(\mathbf{I}_{m_S} + \frac{1}{\sigma_E^2} \Lambda_T \Lambda_Q \mathbf{W}_E^\dagger \Lambda_E \mathbf{W}_E \right)^{-1} \right. \right. \\ &\quad \left. \left. \times \frac{1}{\sigma_D^2} \Lambda_T \Lambda_Q \mathbf{W}_D^\dagger \Lambda_D \mathbf{W}_D \right| \right] \end{aligned} \quad (79)$$

$$\begin{aligned} &= \mathbb{E} \left[\log \left| \mathbf{I}_{m_S} + \left(\frac{1}{P} \mathbf{I}_{m_S} + \frac{1}{\sigma_E^2} \tilde{\Lambda}_T \Lambda_Q \mathbf{W}_E^\dagger \Lambda_E \mathbf{W}_E \right)^{-1} \right. \right. \\ &\quad \left. \left. \times \frac{1}{\sigma_D^2} \tilde{\Lambda}_T \Lambda_Q \mathbf{W}_D^\dagger \Lambda_D \mathbf{W}_D \right| \right], \end{aligned} \quad (80)$$

where (78) is obtained using $|\mathbf{I} + \mathbf{A}\mathbf{B}| = |\mathbf{I} + \mathbf{B}\mathbf{A}|$. Equation (79) is obtained by replacing $\mathbf{W}$ and Λ_{DE} by their expressions and factorizing the term in parenthesis. Equation (80) is obtained by writing, without loss of generality, $\Lambda_T = P \tilde{\Lambda}$ such that $\text{Tr}(\tilde{\Lambda}) = P$. Therefore, as $m_E \geq m_S$, we have $\mathbf{W}_E^\dagger \Lambda_E \mathbf{W}_E$ is invertible. Since $\tilde{\Lambda}$ and Λ_T are invertible, we obtain

$$\lim_{P \rightarrow \infty} R(\Lambda_Q) = \mathbb{E} \left[\log \left| \mathbf{I}_{m_S} + \left(\mathbf{W}_E^\dagger \Lambda_E \mathbf{W}_E \right)^{-1} \frac{\sigma_E^2}{\sigma_D^2} \mathbf{W}_D^\dagger \Lambda_D \mathbf{W}_D \right| \right]. \quad (81)$$

The result in (81) does not depend on the structure of Λ_Q . It only requires Λ_Q to be invertible. Moreover, (81) is equivalent to (35), which is the result we seek to prove. ■

C. Asymptotic difference between secret-key capacities with and without correlation

We assume in this section that all correlation matrices are assumed invertible. We prove the result in two steps. First, we consider a Rayleigh-fading MIMO channel and we establish the asymptotic difference between the correlated and non-correlated secret-key rates achieved with uniform power allocation. Second, observing that the secret-key capacity can be expressed as

a difference between two different Rayleigh-fading MIMO channels, we derive the asymptotic difference between secret-key capacities in case of correlation and in case no correlation.

1) *Asymptotic rate difference of a MIMO channel:* let

$$\delta = \log |\mathbf{I} + P\mathbf{H}\mathbf{H}^\dagger| - \log |\mathbf{I} + P\mathbf{R}_D\mathbf{H}\mathbf{R}_T\mathbf{H}^\dagger|. \quad (82)$$

We are interested in evaluating $\lim_{P \rightarrow \infty} \mathbb{E}(\delta)$. We start by proving the following lemma.

Lemma 3. Let $\mathbf{H}$ be an $m_D \times m_S$ matrix, $\mathbf{R}_T$ (resp. $\mathbf{R}_D$) is an $m_S \times m_S$ (resp. $m_D \times m_D$) positive definite matrix. Then we have

$$\lim_{P \rightarrow \infty} \delta = -\log |\mathbf{I} + \mathbf{H}^+ \mathbf{R}_D \mathbf{H} \mathbf{R}_T - \mathbf{H}^+ \mathbf{H}|, \quad (83)$$

where $\mathbf{H}^+$ denotes the pseudo-inverse of $\mathbf{H}$.

Proof:

$$\begin{aligned} & |\mathbf{I} + P\mathbf{R}_D\mathbf{H}\mathbf{R}_T\mathbf{H}^\dagger| \\ &= |\mathbf{I} + P(\mathbf{R}_D - \mathbf{I} + \mathbf{I})\mathbf{H}(\mathbf{R}_T - \mathbf{I} + \mathbf{I})\mathbf{H}^\dagger| \\ &= |\mathbf{I} + P\mathbf{H}\mathbf{H}^\dagger + P(\mathbf{R}_D - \mathbf{I})\mathbf{H}\mathbf{R}_T\mathbf{H}^\dagger + P\mathbf{H}(\mathbf{R}_T - \mathbf{I})\mathbf{H}^\dagger| \\ &= |\mathbf{I} + P\mathbf{H}\mathbf{H}^\dagger| |\mathbf{I} + (\mathbf{I} + P\mathbf{H}\mathbf{H}^\dagger)^{-1} \\ &\quad \times \{P(\mathbf{R}_D - \mathbf{I})\mathbf{H}\mathbf{R}_T\mathbf{H}^\dagger + P\mathbf{H}(\mathbf{R}_T - \mathbf{I})\mathbf{H}^\dagger\}|. \end{aligned} \quad (84)$$

Therefore, we obtain

$$\begin{aligned} \delta &= -\log |\mathbf{I} + (\mathbf{I} + P\mathbf{H}\mathbf{H}^\dagger)^{-1} \\ &\quad \times \{P(\mathbf{R}_D - \mathbf{I})\mathbf{H}\mathbf{R}_T\mathbf{H}^\dagger + P\mathbf{H}(\mathbf{R}_T - \mathbf{I})\mathbf{H}^\dagger\}| \\ &= -\log |\mathbf{I} + \mathbf{H}^\dagger (\frac{1}{P}\mathbf{I} + \mathbf{H}\mathbf{H}^\dagger)^{-1} (\mathbf{R}_D - \mathbf{I})\mathbf{H}\mathbf{R}_T \\ &\quad + \mathbf{H}^\dagger (\frac{1}{P}\mathbf{I} + \mathbf{H}\mathbf{H}^\dagger)^{-1} \mathbf{H}(\mathbf{R}_T - \mathbf{I})|. \end{aligned} \quad (85)$$

Now, using the limit property of a pseudo-inverse [35], we have $\lim_{P \rightarrow \infty} \mathbf{H}^\dagger (\frac{1}{P}\mathbf{I} + \mathbf{H}\mathbf{H}^\dagger)^{-1} = \mathbf{H}^+$. The continuity of the function $\mathbf{\Gamma} \mapsto \log |\mathbf{\Gamma}|$ implies that

$$\begin{aligned} \lim_{P \rightarrow \infty} \delta &= -\log |\mathbf{I} + \mathbf{H}^+ (\mathbf{R}_D - \mathbf{I})\mathbf{H}\mathbf{R}_T + \mathbf{H}^+ \mathbf{H}(\mathbf{R}_T - \mathbf{I})| \\ &= -\log |\mathbf{I} + \mathbf{H}^+ \mathbf{R}_D \mathbf{H} \mathbf{R}_T - \mathbf{H}^+ \mathbf{H}|. \end{aligned} \quad (86)$$

■

For full rank matrix $\mathbf{H}$, the last result could be simplified as follows.

Corollary 1. 1) If $\mathbf{H}$ is full column rank, then

$$\lim_{P \rightarrow \infty} \delta = -\log \frac{|\mathbf{R}_T| |\mathbf{H}^\dagger \mathbf{R}_D \mathbf{H}|}{|\mathbf{H}^\dagger \mathbf{H}|}. \quad (87)$$

2) If $\mathbf{H}$ is full row rank, then

$$\lim_{P \rightarrow \infty} \delta = -\log \frac{|\mathbf{R}_D| |\mathbf{H} \mathbf{R}_T \mathbf{H}^\dagger|}{|\mathbf{H} \mathbf{H}^\dagger|}. \quad (88)$$

Proof: If $\mathbf{H}$ is full column rank, then $\mathbf{H}^\dagger \mathbf{H}$ is invertible, $\mathbf{H}^+ = (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger$ and $\mathbf{H}^+ \mathbf{H} = \mathbf{I}$. Substituting that in (83), we obtain the result.

In case $\mathbf{H}$ is full row rank, $\mathbf{H} \mathbf{H}^\dagger$ is invertible, $\mathbf{H}^+ = \mathbf{H}^\dagger (\mathbf{H} \mathbf{H}^\dagger)^{-1}$ and $\mathbf{H} \mathbf{H}^+ = \mathbf{I}$. Substituting that in (83) yields the desired relationship. ■

Based on Lemma 3, we now establish an analytical expression of the asymptotic loss.

Corollary 2. Consider an ergodic Rayleigh-fading MIMO channel in which the transmitter has m_S antennas and the destination has m_D antennas. With δ defined as in (82), we have

$$\lim_{P \rightarrow \infty} \mathbb{E}_{\mathbf{H}}(\delta) = \begin{cases} -\log |\mathbf{R}_T| - \mathbb{E} \left\{ \log \frac{|\mathbf{H}^\dagger \mathbf{R}_D \mathbf{H}|}{|\mathbf{H}^\dagger \mathbf{H}|} \right\} & \text{if } m_D \geq m_S \\ -\log |\mathbf{R}_D| - \mathbb{E} \left\{ \log \frac{|\mathbf{H} \mathbf{R}_T \mathbf{H}^\dagger|}{|\mathbf{H} \mathbf{H}^\dagger|} \right\} & \text{else.} \end{cases} \quad (89)$$

In particular, if $m_D = m_S$, then we obtain immediately

$$\lim_{P \rightarrow \infty} \mathbb{E}_{\mathbf{H}}(\delta) = -\log |\mathbf{R}_T| - \log |\mathbf{R}_D|. \quad (90)$$

The last result characterizes the ergodic capacity loss due to correlation in the high-power regime in MIMO systems with no secrecy constraint, which has been derived in [36] by lower-bounding the capacity with an expression that is tight in the high-power regime.

Proof:

$$\begin{aligned} \log |\mathbf{I} + P \mathbf{R}_D \mathbf{H} \mathbf{R}_T \mathbf{H}^\dagger| &\leq \log |\mathbf{I} + m_S P \mathbf{H}^\dagger \mathbf{R}_D \mathbf{H}| \\ &\leq \log |\mathbf{I} + m_S m_D P \mathbf{H}^\dagger \mathbf{H}| \\ &\leq m_S \log(m_S m_D) + \log |\mathbf{I} + P \mathbf{H}^\dagger \mathbf{H}|, \end{aligned} \quad (91)$$

where the first inequality is obtained using the fact that $\mathbf{R}_T \preceq m_S \mathbf{I}$ and the identity $|\mathbf{I} + \mathbf{A} \mathbf{B}| = |\mathbf{I} + \mathbf{B} \mathbf{A}|$. The second inequality follows from $\mathbf{R}_D \preceq m_D \mathbf{I}$. Based on that, we obtain

$$\delta \geq -m_S \log(m_S m_D). \quad (92)$$

Similarly, with λ_{min}^T and λ_{min}^D denoting the minimum eigenvalue of $\mathbf{R}_T$ and $\mathbf{R}_D$, respectively, we have

$$\log |\mathbf{I} + P \mathbf{R}_D \mathbf{H} \mathbf{R}_T \mathbf{H}^\dagger| \geq \log |\mathbf{I} + \lambda_{min}^T P \mathbf{H}^\dagger \mathbf{R}_D \mathbf{H}|$$

$$\begin{aligned} &\geq \log |\mathbf{I} + \lambda_{min}^T \lambda_{min}^D P \mathbf{H}^\dagger \mathbf{H}| \\ &\geq m_S \log(\lambda_{min}^T \lambda_{min}^D) + \log |\mathbf{I} + P \mathbf{H}^\dagger \mathbf{H}|, \end{aligned} \quad (93)$$

where the first and second inequalities follows from the $\mathbf{R}_T \succeq \lambda_{min}^T \mathbf{I}$ and $\mathbf{R}_D \succeq \lambda_{min}^D \mathbf{I}$. The last inequality follows from the fact that $\lambda_{min}^T \leq 1$ and $\lambda_{min}^D \leq 1$. Thus, we obtain

$$\delta \leq -m_S \log(\lambda_{min}^T \lambda_{min}^D). \quad (94)$$

Let $c = \max(m_S \log(m_S m_D), -m_S \log(\lambda_{min}^T \lambda_{min}^D))$. It follows immediately from (92) and (94) that for any $\mathbf{H}$ and $P \geq 0$, we have

$$|\delta| \leq c. \quad (95)$$

Now, since the elements of $\mathbf{H}$ are generated independently according to a Gaussian distribution, $\mathbf{H}$ is full rank almost sure (a.s.). Henceforth, based on Corollary 1, the Bounded Convergence Theorem (BCT) and (95), we obtain

$$\lim_{P \rightarrow \infty} \mathbb{E}_{\mathbf{H}}(\delta) = \mathbb{E}_{\mathbf{H}}(\lim_{P \rightarrow \infty} \delta), \quad (96)$$

which implies the desired result. ■

2) *Application to secret-key agreement*: based on the previous derivations, we establish the limit in the high-power regime of our quantity of interest $\Delta(P)$. From (13), with $Q = \frac{P}{m_S} \mathbf{I}_{m_S}$, the secret-key rate is expressed as $R(p) = R_1(P) - R_2(P)$, where

$$R_1(P) = \mathbb{E} \log |\mathbf{I} + \frac{P}{m_S} \mathbf{\Lambda}_{DE} \mathbf{W} \mathbf{\Lambda}_T \mathbf{W}^\dagger| \quad (97)$$

$$R_2(P) = \mathbb{E} \log |\mathbf{I} + \frac{P}{m_S} \frac{1}{\sigma_E^2} \mathbf{\Lambda}_E \mathbf{W}_E \mathbf{\Lambda}_T \mathbf{W}_E^\dagger|. \quad (98)$$

In case of no correlation, we write similarly

$$C_{\text{uncor}}(P) = \mathbb{E} \log \frac{|\mathbf{I} + \frac{P}{m_S} \mathbf{W} \mathbf{W}^\dagger|}{|\mathbf{I} + \frac{P}{m_S \sigma_E^2} \mathbf{W}_E \mathbf{W}_E^\dagger|} = C_{1,\text{uncor}}(P) - C_{2,\text{uncor}}(P). \quad (99)$$

Analogously, we write $\Delta(P) = \Delta_1(P) - \Delta_2(P)$, where

$$\Delta_1(P) = C_1(P) - C_{1,\text{uncor}}(P) \quad (100)$$

$$\Delta_2(P) = C_2(P) - C_{2,\text{uncor}}(P). \quad (101)$$

It follows that $\lim_{P \rightarrow \infty} \Delta(P) = \lim_{P \rightarrow \infty} [\Delta_1(P) - \Delta_2(P)]$.

With the result of Corollary 2, taking into consideration all possible configurations of transmit, receive and eavesdropper antennas, one obtains the result in Theorem 3.

D. First derivative of the secret-key capacity

Consider the spectral decomposition of the input covariance matrix $\hat{\mathbf{Q}} = \sum_{i=1}^{m_S} \alpha_i \mathbf{u}_i \mathbf{u}_i^\dagger$, where the vectors $\mathbf{u}_i$ are columns of a unitary matrix, and the non-negative coefficients α_i sum to unity. Then, we write

$$\dot{C}(0) = \max_{\text{Tr}(\hat{\mathbf{Q}}) \leq 1} \mathbb{E} [\dot{\mathbb{I}}(0)] \quad (102)$$

$$= \max_{\text{Tr}(\hat{\mathbf{Q}}) \leq 1} \mathbb{E} [\text{Tr}(\hat{\mathbf{H}} \hat{\mathbf{Q}} \hat{\mathbf{H}}^\dagger) - \text{Tr}(\hat{\mathbf{H}}_E \hat{\mathbf{Q}} \hat{\mathbf{H}}_E^\dagger)] \quad (103)$$

$$= \max_{\{\alpha_i\}, \{\mathbf{u}_i\}} \mathbb{E} \left[\sum_{i=1}^{m_S} \alpha_i \left(\text{Tr}(\hat{\mathbf{H}} \mathbf{u}_i \mathbf{u}_i^\dagger \hat{\mathbf{H}}^\dagger) - \text{Tr}(\hat{\mathbf{H}}_E \mathbf{u}_i \mathbf{u}_i^\dagger \hat{\mathbf{H}}_E^\dagger) \right) \right] \quad (104)$$

$$= \max_{\{\alpha_i\}, \{\mathbf{u}_i\}} \mathbb{E} \left[\sum_{i=1}^{m_S} \alpha_i \left(\mathbf{u}_i^\dagger \left(\mathbf{H}_D^\dagger \mathbf{H}_D + \hat{\mathbf{H}}_E^\dagger \hat{\mathbf{H}}_E \right) \mathbf{u}_i \right. \right. \quad (105)$$

$$\left. - \mathbf{u}_i^\dagger \hat{\mathbf{H}}_E^\dagger \hat{\mathbf{H}}_E \mathbf{u}_i \right) \mathbf{u}_i \right] \quad (106)$$

$$= \max_{\{\alpha_i\}, \{\mathbf{u}_i\}} \left[\sum_{i=1}^{m_S} \alpha_i \mathbf{u}_i^\dagger \mathbb{E} \left(\mathbf{H}_D^\dagger \mathbf{H}_D \right) \mathbf{u}_i \right] \quad (107)$$

$$= m_D \max_{\{\alpha_i\}, \{\mathbf{u}_i\}} \left[\sum_{i=1}^{m_S} \alpha_i \mathbf{u}_i^\dagger \mathbf{R}_T \mathbf{u}_i \right].$$

Note that (107) is maximized by choosing one or more $\mathbf{u}_i$ within the eigenspace corresponding to the maximal eigenvalue of the transmit correlation matrix $\lambda_{\max}(\mathbf{R}_T) = \lambda_1^T$, and by setting the remaining $\alpha_i = 0$. In general, if λ_1^T has multiplicity $L > 1$, then any covariance matrix in the form $\mathbf{Q} = P \sum_{i=1}^L \alpha_i \mathbf{u}_i \mathbf{u}_i^\dagger$ achieves the first derivative, where $\mathbf{u}_i$ are the eigenvectors that span the maximal-eigenvalue eigenspace of $\mathbf{R}_T$, and α_i are non-negative coefficients that sum to unity.

E. Second derivative of the secret-key capacity

$$\ddot{C}(0) = \max_{\{\alpha_i\}} \mathbb{E} \left[-\text{Tr}((\hat{\mathbf{H}} \hat{\mathbf{Q}} \hat{\mathbf{H}}^\dagger)^2) + \text{Tr}((\hat{\mathbf{H}}_E \hat{\mathbf{Q}} \hat{\mathbf{H}}_E^\dagger)^2) \right] \quad (108)$$

$$= -\min_{\{\alpha_i\}} \mathbb{E} \left[\text{Tr} \left(\left(\sum_{i=1}^L \alpha_i \hat{\mathbf{H}} \mathbf{u}_i \mathbf{u}_i^\dagger \hat{\mathbf{H}}^\dagger \right)^2 \right) - \text{Tr} \left(\left(\sum_{i=1}^L \alpha_i \hat{\mathbf{H}}_E \mathbf{u}_i \mathbf{u}_i^\dagger \hat{\mathbf{H}}_E^\dagger \right)^2 \right) \right] \quad (109)$$

$$= -\min_{\{\alpha_i\}} \mathbb{E} \left[\sum_{i,j} \alpha_i \alpha_j \left(|\mathbf{u}_j^\dagger (\mathbf{H}_D^\dagger \mathbf{H}_D + \hat{\mathbf{H}}_E^\dagger \hat{\mathbf{H}}_E) \mathbf{u}_i|^2 - |\mathbf{u}_j^\dagger \hat{\mathbf{H}}_E^\dagger \hat{\mathbf{H}}_E \mathbf{u}_i|^2 \right) \right] \quad (110)$$

$$= -\min_{\{\alpha_i\}} \sum_{i,j} \alpha_i \alpha_j \left(\mathbb{E} \left\{ |\mathbf{u}_j^\dagger \mathbf{H}_D^\dagger \mathbf{H}_D \mathbf{u}_i|^2 \right\} + 2\Re \left\{ \mathbb{E} \left\{ \left(\mathbf{u}_j^\dagger \mathbf{H}_D^\dagger \mathbf{H}_D \mathbf{u}_i \right) \left(\mathbf{u}_j^\dagger \hat{\mathbf{H}}_E^\dagger \hat{\mathbf{H}}_E \mathbf{u}_i \right)^\dagger \right\} \right\} \right) \quad (111)$$

$$= -\lambda_1^{T^2} \min_{\{\alpha_i\}} \sum_{i,j} \alpha_i \alpha_j \left(\mathbb{E} \left\{ |\mathbf{u}_j^\dagger W_D^\dagger \mathbf{R}_D W_D \mathbf{u}_i|^2 \right\} \right. \quad (112)$$

$$\left. + 2 \frac{\sigma_D^2}{\sigma_E^2} \Re \left\{ \mathbb{E} \left(\mathbf{u}_j^\dagger W_D^\dagger \mathbf{R}_D W_D \mathbf{u}_i \right) \mathbb{E} \left(\mathbf{u}_j^\dagger W_E^\dagger \mathbf{R}_E W_E \mathbf{u}_i \right)^\dagger \right\} \right) \quad (113)$$

Equation (113) is obtained using the following identities.

$$\mathbb{E} \left\{ |\mathbf{u}_j^\dagger W_D^\dagger \mathbf{R}_D W_D \mathbf{u}_i|^2 \right\} = \text{Tr}(\mathbf{R}_D^2) + m_D^2 \delta_{ij} \quad (114)$$

$$\mathbb{E} \left(\mathbf{u}_j^\dagger W_D^\dagger \mathbf{R}_D W_D \mathbf{u}_i \right) = m_D \delta_{ij} \quad (115)$$

$$\mathbb{E} \left(\mathbf{u}_j^\dagger W_E^\dagger \mathbf{R}_E W_E \mathbf{u}_i \right) = m_E \delta_{ij}, \quad (116)$$

where δ_{ij} is the Kronecker Delta symbol. The minimization problem in (113) is a convex quadratic programming. It can be expressed as

$$\min_{\alpha_i \in [0,1], \sum_{i=1}^L \alpha_i = 1} \alpha^T \mathbf{G} \alpha, \quad (117)$$

where α^T denotes the transpose of α , and $\mathbf{G}$ is an $L \times L$ matrix whose components are $\mathbf{G}_{ii} = \text{Tr}(\mathbf{R}_D^2) + m_D^2 + m_D m_E$ and $\mathbf{G}_{ij} = 2 \frac{\sigma_D^2}{\sigma_E^2} \text{Tr}(\mathbf{R}_D^2)$ for $i, j \in \{1, \dots, L\}, i \neq j$. The optimal solution can be obtained by noticing that the objective function in (113) is convex and symmetric. Hence, it is Schur-convex. Therefore, the uniform vector minimizes the objective function with a sum constraint. Substituting the optimal solution into (113) yields the second derivative of the key capacity in (49).

REFERENCES

- [1] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.
- [2] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451–456, Jul 1978.
- [3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, 1978.
- [4] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.
- [5] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas-Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, 2010.
- [6] F. Oggier and B. Hassibi, "The secrecy capacity of the mimo wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, 2011.
- [7] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, 2009.
- [8] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography—part I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, 1993.
- [9] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 344–366, 2000.
- [10] —, "Secrecy capacities for multiterminal channel models," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2437–2452, 2008.
- [11] T. F. Wong, M. Bloch, and J. M. Shea, "Secret sharing over fast-fading MIMO wiretap channels," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, p. 8, 2009.
- [12] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep 2001.
- [13] S. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *IEEE International Conference on Communications, 2001 (ICC' 01), Helsinki, Finland*, pp. 2266–2270 vol.7.
- [14] E. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in mimo wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1543–1553, Sept 2004.
- [15] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun 2002.
- [16] M. El-Halabi, T. Liu, and C. Georgiades, "Secrecy capacity per unit cost," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1909–1920, September 2013.

- [17] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1030, Sep 1990.
- [18] M. Gursoy, "Secure Communication in the Low-SNR Regime," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1114–1123, April 2012.
- [19] F. Renna, M. Bloch, and N. Laurenti, "Semi-blind key-agreement over MIMO fading channels," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 620–627, February 2013.
- [20] M. Zorghi, Z. Rezki, B. Alomair, and M.-S. Alouini, "Secret-key agreement over spatially correlated fast-fading multiple-antenna channels with public discussion," in *IEEE International Symposium on Information Theory Proceedings (ISIT'2015)*, Hong Kong, Jun 2015.
- [21] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications: Theory of Majorization and Its Applications*. Springer, 2010.
- [22] W. Weichselberger, M. Herdin, H. Ozelik, and E. Bonek, "A stochastic mimo channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan 2006.
- [23] C.-N. Chuah, D. Tse, J. Kahn, and R. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar 2002.
- [24] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge university press, 2003.
- [25] S. Tomasin and E. Jorswieck, "Pilot-based secret key agreement for reciprocal correlated mimome block fading channels," in *Globecom Workshops (GC Wkshps)*, 2014, Dec 2014, pp. 1343–1348.
- [26] M. Bloch and J. Barros, *Physical-layer security: from information theory to security engineering*. Cambridge University Press, 2011.
- [27] S. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1165–1175, July 2004.
- [28] M. Zorghi, Z. Rezki, B. Alomair, and M.-S. Alouini, "On the diversity-multiplexing tradeoff of secret-key agreement over multiple-antenna channels," in *52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2014, Sept 2014, pp. 175–182.
- [29] S. Amari and R. Misra, "Closed-form expressions for distribution of sum of exponential random variables," *IEEE Trans. on Reliability*, vol. 46, no. 4, pp. 519–522, Dec 1997.
- [30] E. Jorswieck and H. Boche, "Optimal transmission strategies and impact of correlation in multiantenna systems with different types of channel state information," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3440–3453, Dec 2004.
- [31] S. Diggavi and T. Cover, "The worst additive noise under a covariance constraint," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3072–3081, Nov 2001.
- [32] M. Zorghi, Z. Rezki, B. Alomair, E. Jorswieck, and M.-S. Alouini, "Ergodic secret-key agreement over spatially correlated multiple-antenna channels with public discussion." [Online]. Available: sri-uq.kaust.edu.sa/Pages/Paper_Zorgui_Rezki_Alomair_Jorswieck_Alouini_1.aspx
- [33] E. Jorswieck and H. Boche, "Multiple-antenna capacity in the low-power regime: channel knowledge and correlation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05)*, vol. 3, March 2005, pp. iii/385–iii/388 Vol. 3.
- [34] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [35] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU Press, 2012, vol. 3.
- [36] O. Oyman, R. Nabar, H. Bolcskei, and A. Paulraj, "Characterizing the statistical properties of mutual information in mimo channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2784–2795, Nov 2003.

Marwen Zorgui (S'13) was born in Tunis, Tunisia. He received the Eng. degree in electronic and communication engineering from the Ecole Polytechnique de Tunisie, Tunisia, in 2013, the M.S. degree in electrical engineering from King Abdullah University of Science and Technology, Thuwal, Makkah Province, Saudi Arabia, in 2015. He is now pursuing his Ph.D. at the University of California at Irvine, CA, USA. His research interests include interference networks, performance limits of communication systems, physical layer security and information theory.

Zouheir Rezki (S'01-M'08-SM'13) was born in Casablanca, Morocco. He received the Diplôme d'Ingénieur degree from the École Nationale de l'Industrie Minérale (ENIM), Rabat, Morocco, in 1994, the M.Eng. degree from École de Technologie Supérieure, Montreal, Québec, Canada, in 2003, and the Ph.D. degree from École Polytechnique, Montreal, Québec, in 2008, all in electrical engineering. From October 2008 to September 2009, he was a postdoctoral research fellow with Data Communications Group, Department of Electrical and Computer Engineering, University of British Columbia. He is now a Senior Research Scientist at King Abdullah University of Science and Technology (KAUST), Thuwal, Mekkah Province, Saudi Arabia. His research interests include:

performance limits of communication systems, physical-layer security, cognitive and sensor networks and low-complexity detection algorithms.

Basel Alomair is an Assistant Professor and Founding Director of the National Center for Cybersecurity Technology (C4C) in King Abdulaziz City for Science and Technology (KACST), an Affiliate Professor and co-director of the Network Security Lab (NSL) at the University of Washington-Seattle, an Affiliate Professor at King Saud University (KSU), and a cryptographic consultant at various companies. He was recognized by the IEEE Technical Committee on Fault-Tolerant Computing (TC-FTC) and the IFIP Working Group on Dependable Computing and Fault Tolerance (WG 10.4) with the 2010 IEEE/IFIP William Carter Award for his significant contributions in the area of dependable computing. His research in information security was recognized with the 2011 Outstanding Research Award from the University of Washington. He was also the recipient of the 2012 Distinguished Dissertation Award from

the Center for Information Assurance and Cybersecurity at the University of Washington (UW CIAC). He was awarded the 2015 Early Career Award in Cybersecurity by the NSA/DHS Center of Academic Excellence in Information Assurance Research for his contributions to Modern Cryptographic Systems and Visionary Leadership. He authored/co-authored multiple best paper awards and he is a member of the IEEE.

Eduard Jorswieck (S'01 – M'03 – SM'08) received the Diplom-Ingenieur (M.S.) degree and Doktor-Ingenieur (Ph.D.) degree, both in electrical engineering and computer science, from the Technische Universitt Berlin, Germany, in 2000 and 2004, respectively. He was with the Broadband Mobile Communication Networks Department, Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut, Berlin, from 2000 to 2008. From 2005 to 2008, he was a Lecturer with the Technische Universitt Berlin. From 2006 to 2008, he was with the Department of Signals, Sensors and Systems, Royal Institute of Technology, as a Post-Doctoral Researcher and an Assistant Professor. Since 2008, he has been the Head of the Chair of Communications Theory and a Full Professor with the Technische Universitt Dresden, Germany. He is principal investigator in the excellence cluster center for Advancing Electronics

Dresden (cfAED) and founding member of the 5G lab Germany (5Glab.de).

His main research interests are in the area of signal processing for communications and networks, applied information theory, and communications theory. He has co-authored more than 80 journal papers, 7 book chapters, some 220 conference papers and 3 monographs on these research topics. Eduard was a co-recipient of the IEEE Signal Processing Society Best Paper Award in 2006 and co-authored papers that won the Best Paper or Best Student Paper Awards at IEEE WPMC 2002, Chinacom 2010, IEEE CAMSAP 2011, IEEE SPAWC 2012, and IEEE WCSP 2012.

Dr. Jorswieck was a member of the IEEE SPCOM Technical Committee (2008-2013), and has been a member of the IEEE SAM Technical Committee since 2015. Since 2011, he has been an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. Since 2008, continuing until 2011, he has served as an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS, and until 2013, as a Senior Associate Editor. Since 2013, he has served as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Mohamed-Slim Alouini (S'94, M'98, SM'03, F09) was born in Tunis, Tunisia. He received the Ph.D. degree in Electrical Engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1998. He served as a faculty member in the University of Minnesota, Minneapolis, MN, USA, then in the Texas A&M University at Qatar, Education City, Doha, Qatar before joining King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia as a Professor of Electrical Engineering in 2009. His current research interests include the modeling, design, and performance analysis of wireless communication systems.