Rank-Deficient Solutions for Optimal Signaling Over Wiretap MIMO Channels

Sergey Loyka and Charalambos D. Charalambous

***Abstract*— Capacity-achieving signaling strategies for the Gaussian wiretap multiple-input multple-output (MIMO) channel are investigated without the degradedness assumption. In addition to known solutions, a number of new rank-deficient solutions for the optimal transmit covariance matrix are obtained. The case of a weak eavesdropper is considered in detail, and the optimal covariance is established in an explicit, closed form with no extra assumptions. This provides lower and upper bounds to the secrecy capacity in the general case with a bounded gap, which are tight for a weak eavesdropper or/and low SNR. Closed-form solutions are also obtained for isotropic and omnidirectional eavesdroppers, based on which lower and upper bounds to the secrecy capacity are established in the general case. Sufficient and necessary conditions for the optimality of three popular transmission techniques, namely, the zero-forcing (ZF), the standard water-filling over the channel eigenmodes, and the isotropic signaling (IS), are established for the MIMO wiretap channel. These solutions are appealing due to their lower complexity. In particular, no wiretap codes are needed for the ZF transmission, and no precoding or feedback is needed for the isotropic signaling.**

***Index Terms*— MIMO, wiretap channel, secrecy capacity, optimal signalling.**

I. Introduction

WIDESPREAD use of wireless systems on one hand and their broadcast nature on the other have initiated significant interest in their security. Information-theoretic studies of the secrecy aspects of wireless systems have recently attracted significant interest [1]. Due to the high spectral efficiency of wireless MIMO systems and their wide adoption by the academia and industry, the Gaussian MIMO wire-tap channel (WTC) has emerged as a popular model and a number of results have been obtained for this model, including the proof of optimality of the Gaussian signaling [1]–[4].

An optimal transmit covariance matrix under the total power constraint has been obtained for some special cases (low/high SNR, MISO channels, full-rank or rank-1 solutions) [2]–[7], but the general case remains elusive. The main difficulty lies in the fact that, unlike the regular MIMO channel, the underlying optimization problem for the MIMO-WTC is generally not convex. It was conjectured in [4] and proved in [3] using an indirect approach (via a degraded channel) that the optimal signaling is on the positive directions of the difference channel. A direct proof (based on the necessary Karush-Kuhn-Tucker (KKT) optimality conditions) has been obtained in [6], while the optimality of signaling on non-negative directions has been established in [7] via an indirect approach. Closed form solutions for MISO and rank-1 MIMO channels have been obtained in [2] and [6]–[8]. The 2-2-1 channel (2 transmit, 2 receive, 1 eavesdropper antenna) has been studied earlier in [5]. The low-SNR regime has been studied in detail in [9]. An exact full-rank solution for the optimal covariance and several of its properties have been obtained in [6]. In particular, unlike the regular channel (no eavesdropper), the optimal power allocation does not converge to the uniform one at high SNR and the latter remains sub-optimal at any finite SNR. In the case of a weak eavesdropper, the optimal signaling mimics the conventional one (water-filling over the channel eigenmodes) with an adjustment for the eavesdropper channel.

Finally, while no analytical solution for the optimal covariance is known in the general case, numerical algorithms have been developed to attack the problem in [10]–[13], which however suffer from the lack of provable global convergence due to the non-convex nature of the optimization problem in the general case. A globally-convergent numerical algorithm for the general case, which is based on an equivalent min-max reformulation of the original problem, was proposed in [14] and its convergence was proved, which takes only a moderate or small number of steps in practice.

The present paper extends the known analytical results for the optimal covariance in several directions. First, motivated by a scenario where the legitimate receiver (Rx) is closer to the transmitter (Tx) than the eavesdropper, the case of a weak eavesdropper is studied and its optimal covariance is obtained in an explicit closed form without any extra assumptions in Section III. It provides novel lower and upper bounds to the secrecy capacity in the general case with a bounded gap, which are tight when the eavesdropper is weak or/and the SNR is low and hence serve as an approximation to the true capacity. It also captures the capacity saturation effect at high SNR observed in [3] and [6]. The range of validity of this model is indicated.

The presence of the eavesdropper channel state information (CSI) at the transmitter is in question when the eavesdropper does not cooperate (e.g. to hide its presence).

Manuscript received September 2, 2015; revised December 21, 2015, March 22, 2016, and April 18, 2016; accepted April 20, 2016. Date of publication April 28, 2016; date of current version June 14, 2016. This paper was presented at the IEEE International Symposium Information Theory, Istanbul, Turkey, Jul. 2013 and the IEEE International Symposium on Information Theory, Honolulu, HI, USA, Jun. 2014. The associate editor coordinating the review of this paper and approving it for publication was V. Y. F. Tan.

S. Loyka is with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: sergey.loyka@uottawa.ca).

C. D. Charalambous is with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia 1678, Cyprus (e-mail: chadcha@ucy.ac.cy).

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

Digital Object Identifier 10.1109/TCOMM.2016.2560173

To address this issue, we consider in Section IV an isotropic eavesdropper model, whereby the Tx does not know the directional properties of the eavesdropper and hence assumes it is isotropic, i.e. the eavesdropper channel gain is the same in all directions. The secrecy capacity as well as an optimal signaling to achieve it and its properties are established in an explicit closed form. This case is shown to be the worst-case MIMO wire-tap channel. Based on this, lower and upper capacity bounds are obtained for the general case, which are achievable by the isotropic eavesdropper. The properties of the optimal power allocation are pointed out.

The case of isotropic eavesdropper above requires the number of its antennas to be not less than the number of Tx antennas (which is necessary for a full-rank eavesdropper channel), which may not be the case in practice. To address this issue, Section V studies an omnidirectional eavesdropper, which may have a smaller number of antennas (and hence rank-deficient channel) and which has the same gain in any direction of a given subspace. The secrecy capacity and the optimal signaling are established in a closed form.

The case of identical right singular vectors of the Rx and eavesdropper channels is investigated and the optimal covariance is established in a closed from in Section VI. This case is motivated by a scenario where the legitimate receiver and the eavesdropper are spatially separated so that each has its own set of local scatterers inducing its own left singular vectors (SV), while both channel are subject to the same set of scatterers around the transmitter (e.g. a base station) and hence the same right SVs. This is similar to the popular Kronecker MIMO channel correlation model, see e.g. [15], where the overall channel correlation is a product of the independent Tx and Rx parts, which are induced by the respective sets of scatterers.

In Section VII, the conditions for optimality of popular zero-forcing (ZF) signaling are established, whereby the Tx antenna array forms a null in the eavesdropper direction. Under those conditions, the standard eigenmode signaling and the water-filling (WF) power allocation on what remains of the required channel (after the ZF) are optimal. Furthermore, no wiretap codes are required as regular coding on the required channel suffices, so that the secrecy requirement imposes no extra complexity penalty (beyond the standard ZF). In this case, the optimal secure signaling is decomposed into two parts: part 1 is the ZF (null forming in the terminology of antenna array literature [16]), which ensures the secrecy requirement, and part 2 is the standard signaling (eigenmode transmission, WF power allocation and coding) on the required channel, which maximizes the rate of required transmission. This is reminiscent of the classical source-channel coding separation [17].

In Sections VIII and IX, we consider two other popular signaling techniques: the standard water-filling over the eigenmodes of the legitimate channel and the isotropic signaling (IS, whereby the covariance matrix is a scaled identity) and establish sufficient and necessary conditions under which they are optimal for the MIMO WTC. These techniques are also appealing due to a number of reasons. While the standard WF does require wiretap codes, standard solutions can be used for power allocation and eigenmode transmission (i.e. spatial modulation); the isotropic signaling is appealing due to its low complexity: no eavesdropper CSI is required at the transmitter as independent, identically distributed data streams are launched by each antenna. The set of channels for which the isotropic signaling is optimal is fully characterized in Section IX. It turns out to be much richer than that of the conventional (no eavesdropper) MIMO channel.

Notations: Lower case bold letters denote vectors while bold capitals denote matrices. $\lambda_i(\mathbf{W})$ denotes the eigenvalues of a matrix $\mathbf{W}$ in decreasing order unless indicated otherwise; $(x)_+ = \max\{x, 0\}$ for a scalar x; $\mathcal{N}(\mathbf{W})$ and $\mathcal{R}(\mathbf{W})$ are the null space and the range of a matrix $\mathbf{W}$; $(\mathbf{W})_+$ denotes the positive eigenmodes of a Hermitian matrix $\mathbf{W}$:

$$(\mathbf{W})_+ = \sum_{i:\lambda_i(\mathbf{W})>0} \lambda_i \mathbf{u}_i \mathbf{u}_i^\dagger \tag{1}$$

where $\mathbf{u}_i$ is i-th eigenvector of $\mathbf{W}$; $\mathrm{tr}\,\mathbf{W}$ and $|\mathbf{W}|$ denote the trace and the determinant of $\mathbf{W}$; $\mathbf{W}^\dagger$ is the Hermitian conjugation of $\mathbf{W}$.

II. Wire-Tap Gaussian MIMO Channel Model

Let us consider the standard wire-tap Gaussian MIMO channel model,

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \boldsymbol{\xi}_1, \quad \mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \boldsymbol{\xi}_2 \tag{2}$$

where $\mathbf{x} = [x_1, x_2, \dots x_m]^T \in \mathbb{C}^{m,1}$ is the transmitted complex-valued signal vector of dimension $m \times 1$, "T" denotes transposition, $\mathbf{y}_k \in \mathbb{C}^{n_k}$, $k = 1, 2$, are the received vectors at the receiver and eavesdropper, $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are the circularly-symmetric additive white Gaussian noise at the receiver and eavesdropper (normalized to unit variance in each dimension), $\mathbf{H}_k \in \mathbb{C}^{n_k,m}$ is the $n_k \times m$ matrix of the complex channel gains between each Tx and each receive (eavesdropper) antenna, n_1, n_2 and m are the numbers of Rx, eavesdropper and Tx antennas respectively. The channels $\mathbf{H}_k$ are assumed to be quasistatic (i.e., constant for a sufficiently long period of time so that the infinite horizon information theory assumption holds) and frequency-flat, with full channel state information (CSI) at the Rx and Tx ends.

For a given transmit covariance matrix $\mathbf{R} = E\left\{\mathbf{x}\mathbf{x}^\dagger\right\}$, where $E\{\cdot\}$ is the statistical expectation, the maximum achievable secrecy rate between the Tx and Rx (so that the rate between the Tx and the eavesdropper is zero) is [3], [4]

$$C(\mathbf{R}) = \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \mathbf{W}_2\mathbf{R}|} = C_1(\mathbf{R}) - C_2(\mathbf{R}) \tag{3}$$

where $C_k(\mathbf{R}) = \ln|\mathbf{I} + \mathbf{W}_k\mathbf{R}|$, $k = 1, 2$, negative $C(\mathbf{R})$ is interpreted as zero rate, $\mathbf{W}_k = \mathbf{H}_k^\dagger\mathbf{H}_k$, and the secrecy capacity subject to the total Tx power constraint is

$$C_s = \max_{\mathbf{R}\geq 0} C(\mathbf{R}) \quad \text{s.t. } \mathrm{tr}\,\mathbf{R} \leq P_T \tag{4}$$

where P_T is the total transmit power (also the SNR since the noise is normalized). It is well-known that the problem in (4) is not convex in general and explicit solutions for the optimal Tx covariance are not known for the general case, but only for some special cases (e.g. low/high SNR, MISO channels, full-rank or rank-1 case [2]–[6]).

III. Weak Eavesdropper and Capacity Bounds

In this section, we obtain novel lower and upper bounds to the secrecy capacity in the general case and show that the bounds are tight when the eavesdropper is weak or if the SNR is low. The weak eavesdropper case is motivated by a scenario where the eavesdropper is located far away from the Tx so that its propagation path loss is large, see e.g. Fig. 2. This is the case when the presence of the eavesdropper does not result in a large capacity loss so that the physical-layer secrecy approach is feasible (while in the case of a strong eavesdropper, the capacity loss is large and other approaches may be preferable, e.g. cryptography). There is no requirement here for the channel to be degraded or for the optimal covariance to be of full rank or rank 1, so that these results extend the known closed form solutions.

To this end, let

$$\begin{aligned} C_w(\mathbf{R}) &= \ln|\mathbf{I}+\mathbf{W}_1\mathbf{R}| - \mathrm{tr}(\mathbf{W}_2\mathbf{R}) \\ C_w &= \max_{\mathbf{R}} C_w(\mathbf{R}) \\ \mathbf{R}^* &= \arg\max_{\mathbf{R}} C(\mathbf{R}), \quad \mathbf{R}_w^* = \arg\max_{\mathbf{R}} C_w(\mathbf{R}) \end{aligned} \tag{5}$$

all subject to $\mathbf{R} \geq 0, \mathrm{tr}\,\mathbf{R} \leq P_T$, i.e. $\mathbf{R}^*$ is the optimal covariance and $\mathbf{R}_w^*$ maximizes $C_w(\mathbf{R})$. Using $\ln(1+x) \approx x$ when $0 \leq x \ll 1$, it can been seen that $C_w(\mathbf{R})$ is a weak eavesdropper approximation of $C(\mathbf{R})$:

$$C(\mathbf{R}) \approx C_w(\mathbf{R}) \quad \text{if } \lambda_1(\mathbf{W}_2\mathbf{R}) \ll 1 \tag{6}$$

so that C_w is the weak eavesdropper secrecy capacity. The following Theorem establishes novel secrecy capacity bounds based on C_w.

Theorem 1: The secrecy capacity C_s in (4) *for the general Gaussian MIMO-WTC in* (2) *is bounded as follows:*

$$C_w \leq C(\mathbf{R}_w^*) \leq C_s \leq C_w + \frac{P_T^2}{2}\lambda_1^2(\mathbf{W}_2) \tag{7}$$

where

$$\mathbf{R}_w^* = \mathbf{Q}^{1/2}(\mathbf{I} - \widehat{\mathbf{W}}_1^{-1})_+\mathbf{Q}^{1/2} \tag{8}$$

$$\widehat{\mathbf{W}}_1 = \mathbf{Q}^{1/2}\mathbf{W}_1\mathbf{Q}^{1/2} \tag{9}$$

and $\mathbf{Q}$ is the (Moore-Penrose) pseudo-inverse of $\mathbf{W}_\lambda = \lambda\mathbf{I} + \mathbf{W}_2$; $\lambda \geq 0$ is found from the total power constraint:

$$\mathrm{tr}\,\mathbf{R}_w^* = P_T \quad \text{if } P_T < P_T^* \tag{10}$$

and $\lambda = 0$ otherwise; the threshold power

$$P_T^* = \mathrm{tr}\,\mathbf{W}_2^{-1}(\mathbf{I} - \mathbf{W}_2^{1/2}\mathbf{W}_1^{-1}\mathbf{W}_2^{1/2})_+ \tag{11}$$

if $\mathbf{W}_2$ is non-singular. When $\mathbf{W}_2$ is singular, $P_T^* = \infty$ if $\mathcal{N}(\mathbf{W}_2) \nsubseteq \mathcal{N}(\mathbf{W}_1)$; otherwise, $\mathbf{W}_1$ and $\mathbf{W}_2$ are projected orthogonally to $\mathcal{N}(\mathbf{W}_2)$ and the projected matrices are used in (11)*. The weak eavesdropper secrecy capacity can be expressed as*

$$C_w = \sum_{i:\widehat{\lambda}_{1i}>1} \ln\widehat{\lambda}_{1i} - \mathrm{tr}\,\widehat{\mathbf{W}}_2(\mathbf{I} - \widehat{\mathbf{W}}_1^{-1})_+ \tag{12}$$

where $\widehat{\lambda}_{1i} = \lambda_i(\widehat{\mathbf{W}}_1)$, $\widehat{\mathbf{W}}_2 = \mathbf{Q}^{1/2}\mathbf{W}_2\mathbf{Q}^{1/2}$.

Proof: See the Appendix. □

Remark 1: It may appear that (8) requires $\widehat{\mathbf{W}}_1$ and thus $\mathbf{W}_1$ to be positive definite, i.e. singular case is not allowed. This is not so since $(\cdot)_+$ operator eliminates singular eigenmodes of $\widehat{\mathbf{W}}_1$ so that $(\mathbf{I}-\widehat{\mathbf{W}}_1^{-1})_+$ is well-defined even if $\mathbf{W}_1$ is singular: one can use $\widehat{\mathbf{W}}_{1\delta} = \widehat{\mathbf{W}}_1 + \delta\mathbf{I} > 0$ instead of $\widehat{\mathbf{W}}_1$, where $\delta > 0$, evaluate $(\mathbf{I} - \widehat{\mathbf{W}}_{1\delta}^{-1})_+$ and take the limit $\delta \to 0$ to see that the singular modes of $\widehat{\mathbf{W}}_1$ are eliminated so that

$$(\mathbf{I} - \widehat{\mathbf{W}}_1^{-1})_+ = \mathbf{U}_+\mathbf{D}\mathbf{U}_+^\dagger \tag{13}$$

where $\mathbf{U}_+$ is a semi-unitary matrix whose columns are the eigenvectors of $\widehat{\mathbf{W}}_1$ corresponding to its positive eigenvalues, $\mathbf{D}$ is a $r \times r$ diagonal matrix whose i-th diagonal entry is $(1 - \lambda_i^{-1}(\widehat{\mathbf{W}}_1))_+$, $i = 1 \ldots r$, where r is the rank of $\widehat{\mathbf{W}}_1$. The same observation also applies to (11).

Remark 2: The 1st inequality in (7) bounds the sub-optimality gap of using $\mathbf{R}_w^*$, for which an achievable rate is $C(\mathbf{R}_w^*)$, instead of the true optimal covariance $\mathbf{R}^*$:

$$|C_s - C(\mathbf{R}_w^*)| \leq \lambda_1^2(\mathbf{W}_2)P_T^2/2 \tag{14}$$

so that $C(\mathbf{R}_w^*) \to C_s$ as $\lambda_1(\mathbf{W}_2)P_T \to 0$.

Using Theorem 1, we can now approximate the secrecy capacity via its weak eavesdropper counterpart.

Corollary 1: The secrecy capacity of the general Gaussian MIMO-WTC can be expressed as follows:

$$C_s = C_w + \Delta C \tag{15}$$

where ΔC is the inaccuracy of the weak eavesdropper approximation, which is bounded as

$$0 \leq \Delta C \leq \lambda_1^2(\mathbf{W}_2)P_T^2/2 \tag{16}$$

so that $\Delta C \to 0$ and $C_s/C_w \to 1$ as $P_T \to 0$ or/and $\lambda_1(\mathbf{W}_2) \to 0$.

Proof: (15) and (16) follow from the bounds in (7), which also implies $\Delta C \to 0$ as $P_T\lambda_1(\mathbf{W}_2) \to 0$. To show that $C_s/C_w \to 1$ as $P_T \to 0$ observe that

$$C_s = P_T\lambda_1(\mathbf{W}_1 - \mathbf{W}_2) + o(P_T) = C_w + o(P_T)$$

from which the desired result follows (here, we implicitly assume that $\lambda_1(\mathbf{W}_1 - \mathbf{W}_2) > 0$; otherwise, $C_s = 0$ and there is nothing to prove). When $\lambda_1(\mathbf{W}_2) \to 0$, note that both $C(\mathbf{R})$ and $C_w(\mathbf{R})$ converge to $\ln|\mathbf{I} + \mathbf{W}_1\mathbf{R}|$ so that taking $\max_{\mathbf{R}}$ results in $C_s/C_w \to 1$ (since the objectives are continuous and the feasible set is compact). □

Using this Corollary, the secrecy capacity can be approximated as

$$C_s \approx C_w \tag{17}$$

and the approximation is accurate for a weak eavesdropper or/and low SNR: $\lambda_1(\mathbf{W}_2)P_T \ll 1$, when the bounds in (7) are also tight, see Fig. 1.

Remark 3: Since $\lambda_1(\mathbf{W}_2\mathbf{R}) \leq \lambda_1(\mathbf{W}_2)\lambda_1(\mathbf{R}) \leq P_T\lambda_1(\mathbf{W}_2)$, one way to ensure that the eavesdropper is weak, i.e. $\lambda_1(\mathbf{W}_2\mathbf{R}) \ll 1$ so that $\ln|\mathbf{I} + \mathbf{W}_2\mathbf{R}| \approx \mathrm{tr}\,\mathbf{W}_2\mathbf{R}$, is to require $\lambda_1(\mathbf{W}_2) \ll 1/P_T$ from which it follows that this holds as long as the power (or SNR) is not too large, i.e. $P_T \ll 1/\lambda_1(\mathbf{W}_2)$; see also Fig. 1. It should be noted, however, that the approximation in (17) extends well beyond

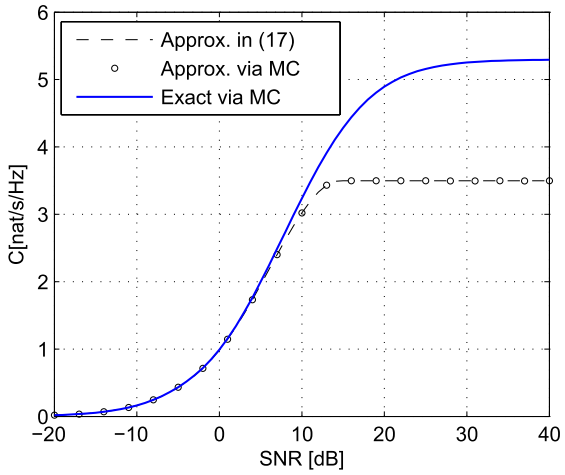

Fig. 1. Weak eavesdropper approximation in (17) and exact secrecy capacity (via MC) versus SNR. $\mathbf{W}_1$ and $\mathbf{W}_2$ are as in (18), $\alpha = 0.1$, $\lambda_1(\mathbf{W}_2) \approx 0.25$. The approximation is accurate if SNR < 10 dB. Note the capacity saturation effect at high SNR in both cases.

the low-SNR regime provided that the eavesdropper propagation path loss is sufficiently large (i.e. $\lambda_1(\mathbf{W}_2)$ is small). For the scenario in Fig. 1, it works well up to about 10 dB and this can extend to larger SNR for smaller path loss factor α.

To illustrate Theorem 1 and Corrolary 1 and also to see how accurate the approximation is, Fig. 1 shows the secrecy capacity obtained from the approximation in (17) for

$$\mathbf{W}_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{W}_2 = \alpha \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \tag{18}$$

also, its exact values (without the weak eavesdropper approximation) obtained by brute force Monte-Carlo (MC) based approach (where a large number of covariance matrices are randomly generated, subject to the total power constraint, and the best one is selected) are shown for comparison. To validate the analytical solution for C_w in Theorem 1, the weak eavesdropper case has also been solved by the MC-based approach. It is clear that the approximation $C_s \approx C_w$ is accurate for the channel in (18) provided that SNR < 10 dB. Also note the capacity saturation effect, for both the approximate and exact values. This saturation effect has been already observed in [3] and [6] and, in the case of $\mathbf{W}_1 > \mathbf{W}_2 > \mathbf{0}$, the saturation capacity is

$$C_s^* = \ln|\mathbf{W}_1| - \ln|\mathbf{W}_2| \tag{19}$$

which follows directly from (3) by neglecting $\mathbf{I}$. In the weak eavesdropper approximation, the saturation effect is due to the fact that the 2nd term in (5) is linear in P_T while the 1st one is only logarithmic, so that using the full available power is not optimal when it is sufficiently high. Roughly, the approximation is accurate before it reaches the saturation point, i.e. for $P_T < P_T^*$. The respective saturation capacity is obtained from (12) by setting $\lambda = 0$. In the case of $\mathbf{W}_1 > \mathbf{W}_2 > \mathbf{0}$, it is given by

$$C_w = \ln|\mathbf{W}_1| - \ln|\mathbf{W}_2| - \mathrm{tr}(\mathbf{I} - \mathbf{W}_2\mathbf{W}_1^{-1}) \tag{20}$$

By comparing (19) and (20), one concludes that the thresholds are close to each other when $\mathrm{tr}\,\mathbf{W}_2\mathbf{W}_1^{-1} \approx m$.

To obtain further insight in the weak eavesdropper regime, let us consider the case when $\mathbf{W}_1$ and $\mathbf{W}_2$ have the same eigenvectors. This is a broader case than it may first appear as it requires $\mathbf{H}_1$ and $\mathbf{H}_2$ to have the same right singular vectors while leaving left ones unconstrained (see Section VI for more details on this scenario). In this case, the results in Theorem 1 and Corollary 1 simplify as follows.

Corollary 2: Under the weak eavesdropper condition $\lambda_1(\mathbf{W}_2) \ll 1/P_T$ and when $\mathbf{W}_1$ and $\mathbf{W}_2$ have the same eigenvectors, the optimal covariance is

$$\mathbf{R}^* \approx \mathbf{R}_w^* = \mathbf{U}\mathbf{\Lambda}^*\mathbf{U}^\dagger \tag{21}$$

where $\mathbf{U}$ is found from the eigenvalue decompositions $\mathbf{W}_i = \mathbf{U}\mathbf{\Lambda}_i\mathbf{U}^\dagger$ so that the eigenvectors of $\mathbf{R}_w^*$ are the same as those of $\mathbf{W}_1$ and $\mathbf{W}_2$. The diagonal matrix $\mathbf{\Lambda}^*$ collects the eigenvalues of $\mathbf{R}_w^*$:

$$\lambda_i(\mathbf{R}_w^*) = \left(\frac{1}{\lambda + \lambda_{2i}} - \frac{1}{\lambda_{1i}}\right)_+ \tag{22}$$

where λ_{ki} is i-th eigenvalue of $\mathbf{W}_k$.

Proof: Using $\mathbf{W}_i = \mathbf{U}\mathbf{\Lambda}_i\mathbf{U}^\dagger$ in (8) results in (21) and (22). □

Note that the power allocation in (22) resembles that of the standard water filling, except for the λ_{2i} term. In particular, only sufficiently strong eigenmodes are active:

$$\lambda_i(\mathbf{R}_w^*) > 0 \quad \text{iff } \lambda_{1i} > \lambda + \lambda_{2i} \tag{23}$$

As P_T increases, λ decreases so that more eigenmodes become active; the legitimate channel eigenmodes are active provided that they are stronger that those of the eavesdropper: $\lambda_{1i} > \lambda_{2i}$. Only the strongest eigenmode (for which the difference $\lambda_{1i} - \lambda_{2i}$ is largest) is active at low SNR.

IV. Isotropic Eavesdropper and Capacity Bounds

The model in Section III requires the full eavesdropper CSI at the transmitter. This becomes questionable if the eavesdropper does not cooperate (e.g. when it is hidden in order not to compromise its eavesdropping ability). One approach to address this issue is via a compound channel model [23]–[25]. An alternative approach is considered here, where the eavesdropper is characterized by its channel gain identical in all directions, which we term "isotropic eavesdropper." This minimizes the amount of CSI available at the transmitter (one scalar parameter and no directional properties).

A further physical justification for this model comes from an assumption that the eavesdropper cannot approach the transmitter too closely due to e.g. some minimum protection distance, see Fig. 2. This ensures that the gain of the eavesdropper channel does not exceed a certain threshold in any transmit direction due to the minimum propagation path loss (induced by the minimum distance constraint). Since the channel power gain in transmit direction $\mathbf{u}$ is $\mathbf{u}^\dagger\mathbf{W}_2\mathbf{u} = |\mathbf{H}_2\mathbf{u}|^2$ (assuming $|\mathbf{u}| = 1$) and since $\max_{|\mathbf{u}|=1} \mathbf{u}^\dagger\mathbf{W}_2\mathbf{u} = \epsilon_1$ (from the variational characterization of eigenvalues [21]), where ϵ_1 is the largest eigenvalue of $\mathbf{W}_2$, $\mathbf{W}_2 \le \epsilon_1\mathbf{I}$ ensures that the eavesdropper channel power gain does not exceed ϵ_1 in any direction.

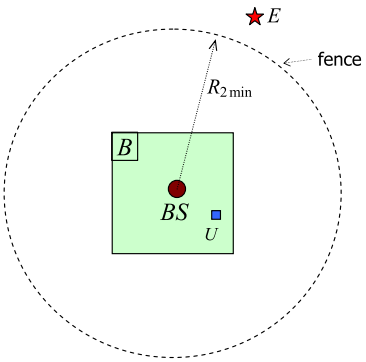

Fig. 2. Physical scenario for a secret communication system: base station BS (the transmitter) is located on the rooftop of a secure building B, legitimate user U (the receiver) is inside the building B, and eavesdropper E is beyond the fence so that $R_2 \geq R_{2\min}$.

In combination with matrix monotonicity of the log-det function, the latter inequality ensures that $\epsilon_1\mathbf{I}$ is the worst possible $\mathbf{W}_2$ that results in the smallest capacity (the lower bound in (27)), i.e. the isotropic eavesdropper with the maximum channel gain is the worst possible one among all eavesdroppers with a bounded spectral norm. Referring to Fig. 2, the eavesdropper channel matrix $\mathbf{H}_2$ can be presented in the following form:

$$\mathbf{H}_2 = \sqrt{\alpha R_2^{-\nu}}\widetilde{\mathbf{H}}_2 \tag{24}$$

where $\alpha R_2^{-\nu}$ represents the average propagation path loss, R_2 is the eavesdropper-transmitter distance, ν is the path loss exponent (which depends on the propagation environment), α is a constant independent of distance (but dependent on frequency, antenna height, etc.) [27], and $\widetilde{\mathbf{H}}_2$ is a properly normalized channel matrix (includes local scattering/multipath effects but excludes the average path loss) so that $\mathrm{tr}\,\widetilde{\mathbf{H}}_2^{\dagger}\widetilde{\mathbf{H}}_2 \leq n_2 m$ [28]. With this in mind, one obtains:

$$\mathbf{W}_2 = \mathbf{H}_2^{\dagger}\mathbf{H}_2 = \frac{\alpha}{R_2^{\nu}}\widetilde{\mathbf{H}}_2^{\dagger}\widetilde{\mathbf{H}}_2 \leq \frac{\alpha n_2 m}{R_{2\min}^{\nu}}\mathbf{I} \tag{25}$$

so that one can take $\epsilon_1 = \alpha n_2 m R_{2\min}^{-\nu}$ in this scenario, where $R_{2\min}$ is the minimum transmitter-eavesdropper distance. Note that the model captures the impact of the number of transmit and eavesdropper antennas, in addition to the minimum distance and propagation environment. In our view, the isotropic eavesdropper model is more practical than the full Tx CSI model.

The isotropic eavesdropper model is closely related to the parallel channel setting in [19] and [20]: even though the original channel is not parallel, it can be transformed into a parallel channel,[1] for which independent signaling is known to be optimal [19], [20]. This shows that signaling on the eigenvectors of $\mathbf{W}_1$ is optimal in this case while an optimal power allocation is different from the standard water filling [20].

[1]Via an information-preserving transformation: using a unitary transmit precoding with the unitary matrix whose columns are the eigenvectors of $\mathbf{W}_1$ and unitary post-codings at the receiver and eavesdropper with unitary matrices whose columns are the left singular vectors of $\mathbf{H}_1$ and $\mathbf{H}_2$ respectively.

These properties in combination with the bounds in (26) are exploited below.

While it is a challenging analytical task to evaluate the secrecy capacity in the general case, one can use the isotropic eavesdropper model above to construct lower and upper capacity bounds for the general case using the standard matrix inequalities,

$$\epsilon_m\mathbf{I} \leq \mathbf{W}_2 \leq \epsilon_1\mathbf{I} \tag{26}$$

where $\epsilon_i = \lambda_i(\mathbf{W}_2)$ denotes i-th largest eigenvalue of $\mathbf{W}_2$, and the equalities are achieved when $\epsilon_1 = \epsilon_m$, i.e. by the isotropic eavesdropper. This is formalized below.

Proposition 1: The secrecy capacity of the general MIMO-WTC in (4) *is bounded as follows:*

$$C^*(\epsilon_1) \leq C_s \leq C^*(\epsilon_m) \tag{27}$$

where $C^*(\epsilon)$ is the secrecy capacity if the eavesdropper were isotropic, i.e. under $\mathbf{W}_2 = \epsilon\mathbf{I}$,

$$C^*(\epsilon) = \max_{\mathbf{R}\geq 0,\ tr\mathbf{R}\leq P_T} \ln\frac{|\mathbf{I}+\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+\epsilon\mathbf{R}|} = \sum_i \ln\frac{1+g_i\lambda_i^*}{1+\epsilon\lambda_i^*} \tag{28}$$

$g_i = \lambda_i(\mathbf{W}_1)$, and $\lambda_i^* = \lambda_i(\mathbf{R}^*) = f(g_i,\epsilon)$ are the eigenvalues of the optimal transmit covariance $\mathbf{R}^* = \mathbf{U}_1\Lambda^*\mathbf{U}_1^{\dagger}$ under the isotropic eavesdropper, where

$$f(x,y) = \frac{x+y}{2yx}\left(\sqrt{1+\frac{4xy}{(x+y)^2}\left(\frac{x-y}{\lambda}-1\right)_+}-1\right) \tag{29}$$

and $\lambda > 0$ is found from the total power constraint $\sum_i \lambda_i^* = P_T$.

The gap in the bounds of (27) *is upper bounded as follows:*

$$\begin{aligned}\Delta C = C^*(\epsilon_m) - C^*(\epsilon_1) &\leq m_+ \ln\frac{1+\epsilon_1 P_T/m_+}{1+\epsilon_m P_T/m_+} \\ &\leq m_+ \ln\frac{\epsilon_1}{\epsilon_m}\end{aligned} \tag{30}$$

where m_+ is the number of eigenmodes such that $g_i > \epsilon_m$. Both bounds are tight at high SNR if $g_{m+} > \epsilon_1$.

Proof: See the Appendix.

Thus, the optimal signaling for the isotropic eavesdropper case is on the eigenvectors of $\mathbf{W}_1$ (or right singular vectors of $\mathbf{H}_1$), identically to the regular MIMO channel, with the optimal power allocation somewhat similar (but not identical) to the conventional water filling. The latter is further elaborated below for the high and low SNR regimes. Unlike the general case (of non-isotropic eavesdropper), the secrecy capacity of the isotropic eavesdropper case does not depend on the eigenvectors of $\mathbf{W}_1$ (but the optimal signaling does), only on its eigenvalues, so that the optimal signaling problem here separates into 2 independent parts: (i) optimal signaling directions are selected as the eigenvectors of $\mathbf{W}_1$, and (ii) optimal power allocation is done based on the eigenvalues of $\mathbf{W}_1$ and the eavesdropper channel gain ϵ. It is the lack of this separation that makes the optimal signaling problem so difficult in the general case.

The bounds in (27) coincide when $\epsilon_1 = \epsilon_m$ thus giving the secrecy capacity of the isotropic eavesdropper.

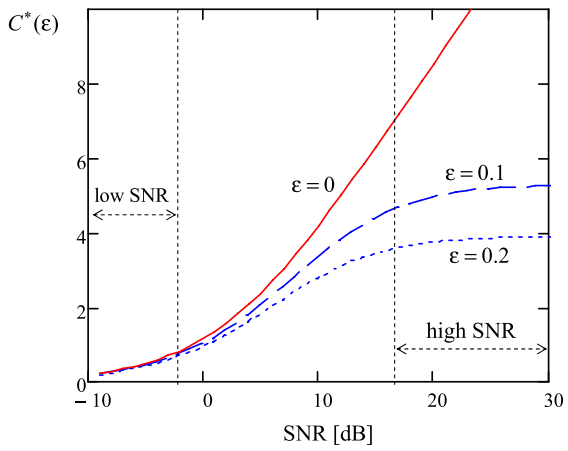

Fig. 3. Secrecy capacity for the isotropic eavesdropper and the capacity of the regular MIMO channel (no eavesdropper, $\epsilon = 0$) vs. the SNR ($= P_T$ since the noise variance is unity); $g_1 = 2, g_2 = 1$. Note the saturation effect at high SNR , where the capacity strongly depends on ϵ but not the SNR, and the negligible impact of the eavesdropper at low SNR.

Furthermore, as follows from (30), they are close to each other when the condition number ϵ_1/ϵ_m of $\mathbf{W}_2$ is not too large, thus providing a reasonable estimate of the capacity, see Fig. 3. Referring to Fig. 2, one can also set $\epsilon_1 = \alpha n_2 m R_{2\min}^{-\nu}$ and proceed with a conservative system design to achieve the secrecy rate $C^*(\epsilon_1)$. Note that this design requires only the knowledge of n_2 and $R_{2\min}$ at the transmitter, not full CSI ($\mathbf{W}_2$) and hence is more realistic. This signaling strategy does not incur significant penalty (compared to the full CSI case) provided that the condition number ϵ_1/ϵ_m is not large, as follows from (30). It can be further shown that $C^*(\epsilon_1)$ is the compound channel capacity for the class of eavesdroppers with bounded spectral norm (maximum channel gain), $\mathbf{W}_2 \le \epsilon_1 \mathbf{I}$, and that signaling on the worst-case channel ($\mathbf{W}_2 = \epsilon_1 \mathbf{I}$) achieves the capacity for the whole class of channels with $\mathbf{W}_2 \le \varepsilon \mathbf{I}$ [25].

We note that the power allocation in (29) has properties similar to those of the conventional water-filling, which follow from Proposition 1.

Proposition 2: Properties of the optimum power allocation in (29) *for the isotropic eavesdropper:*

1. λ_i^* is an increasing function of g_i (strictly increasing unless $\lambda_i^* = 0$ or P_T) , i.e. stronger eigenmodes get more power (as in the standard WF).

2. λ_i^* is an increasing function of P_T (strictly increasing unless $\lambda_i^* = 0$). $\lambda_i^* = 0$ for $i > 1$ and $\lambda_1^* = P_T$ as $P_T \to 0$ if $g_1 > g_2$, i.e. only the strongest eigenmode is active at low SNR, and $\lambda_i^* > 0$ if $g_i > \epsilon$ as $P_T \to \infty$, i.e. all sufficiently strong eigenmodes are active at high SNR.

3. $\lambda_i^* > 0$ only if $g_i > \epsilon$, i.e. only the eigenmodes stronger than the eavesdropper ones can be active.

4. λ is a strictly decreasing function of P_T and $0 < \lambda < g_1 - \epsilon$; $\lambda \to 0$ as $P_T \to \infty$ and $\lambda \to g_1 - \epsilon$ as $P_T \to 0$.

5. There are m_+ active eigenmodes if the following inequalities hold:

$$P_{m_+} < P_T \le P_{m_+ + 1} \tag{31}$$

where P_{m_+} is a threshold power (to have at least m_+ active eigenmodes):

$$P_{m_+} = \sum_{i=1}^{m_+ - 1} \frac{\epsilon + g_i}{2\epsilon g_i} \left(\sqrt{1 + \frac{4\epsilon g_i}{(\epsilon + g_i)^2} \frac{g_i - g_{m_+}}{(g_{m_+} - \epsilon)_+}} - 1 \right), \tag{32}$$

where $m_+ = 2 \ldots m$ and $P_1 = 0$, so that m_+ increases with P_T.

It follows from Proposition 2 that there is only one active eigenmode, i.e. beamforming is optimal, if $g_2 > \epsilon$ and

$$P_T \le \frac{\epsilon + g_1}{2\epsilon g_1} \left(\sqrt{1 + \frac{4\epsilon g_1}{(\epsilon + g_1)^2} \frac{g_1 - g_2}{g_2 - \epsilon}} - 1 \right) \tag{33}$$

e.g. in the low SNR regime (note however that the single-mode regime extends well beyond low SNR if $\epsilon \to g_2$ and $g_1 > g_2$), or at any SNR if $g_1 > \epsilon$ and $g_2 \le \epsilon$.

While it is difficult to evaluate λ analytically from the power constraint, Property 4 ensures that any suitable numerical algorithm (e.g. Newton-Raphson method) will do so efficiently.

As a side benefit of Proposition 2, one can use (31) as a condition for having m_+ active eigenmodes under the regular eigenmode transmission (no eavesdropper) with the standard water-filling by taking $\epsilon \to 0$ in (32):

$$P_{m_+} = \sum_{i=1}^{m_+ - 1} \left(\frac{1}{g_{m_+}} - \frac{1}{g_i} \right) \tag{34}$$

and (34) approximates (32) when the eavesdropper is weak, $\epsilon \ll g_{m+}$. To the best of our knowledge, expression (34) for the threshold powers of the standard water-filling has not appeared in the literature before.

A. High SNR Regime

Let us now consider the isotropic eavesdropper model when the SNR grows large, so that $g_i \lambda_i^* \gg 1, \epsilon \lambda_i^* \gg 1$. In this case, (28) simplifies to

$$C_\infty^* = \sum_{i: g_i > \epsilon} \ln \frac{g_i}{\epsilon} \tag{35}$$

where the summation is over active eigenmodes only, so that the capacity is independent of the SNR (saturation effect) and the impact of the eavesdropper is the multiplicative SNR loss, which is never negligible. To obtain a threshold value of P_T at which the saturation takes place, observe that $\lambda \to 0$ as $P_T \to \infty$ so that (29) becomes

$$\lambda_i^* = P_T \sqrt{\epsilon^{-1} - g_i^{-1}} / \beta (1 + o(1)) \tag{36}$$

for $i : g_i > \epsilon$, where $\beta = \sum_{i: g_i > \epsilon} \sqrt{\epsilon^{-1} - g_i^{-1}}$ and $\sqrt{\lambda} = \beta P_T^{-1} (1 + o(1))$ from the total power constraint. Using (36), the capacity becomes

$$C^*(\epsilon) = \sum_{i: g_i > \epsilon} \ln \frac{g_i}{\epsilon} - \frac{\beta^2}{P_T} + o\left(\frac{1}{P_T}\right) \tag{37}$$

which is a refinement of (35). The saturation takes place when the second term is much smaller than the first one, so that

$$P_T \gg \beta^2 / \sum_{i:g_i>\epsilon} \ln \frac{g_i}{\epsilon} \tag{38}$$

and $C^*(\epsilon) \approx C^*_\infty$ under this condition. This effect in illustrated in Fig. 3. Note that, from (36), the optimal power allocation behaves almost like water-filling in this case, due to the $\sqrt{\epsilon^{-1} - g_i^{-1}}$ term.

Using (35), the gap ΔC^*_∞ between the lower and upper bounds in (27) becomes

$$\begin{aligned} \Delta C^*_\infty &= C^*_\infty(\epsilon_m) - C^*_\infty(\epsilon_1) \\ &= m_1 \ln \frac{\epsilon_1}{\epsilon_m} + \sum_{i=m_1+1}^{m_2} \ln \frac{g_i}{\epsilon_m} \end{aligned} \tag{39}$$

where m_1 and m_2 are the numbers of active eigenmodes when $\epsilon = \epsilon_1$ and $\epsilon = \epsilon_m$. Note that this gap is SNR-independent and if $m_1 = m_2 = m_+$, which is the case if $g_{m+} > \epsilon_1$, then

$$\Delta C^*_\infty = m_+ \ln \frac{\epsilon_1}{\epsilon_m} \tag{40}$$

i.e. also independent of the eigenmode gains of the legitimate user and is determined solely by the condition number of the eavesdropper channel and the number of active eigenmodes. Note that, in this case, the upper bounds in (30) are tight.

B. When Is the Eavesdropper's Impact Negligible?

It is clear from (28) that under fixed $\{g_i\}$ and P_T, the secrecy capacity converges to the conventional one $C^*(0)$ as $\epsilon \to 0$. However, no fixed ϵ (does not matter how small) can ensure by itself that the eavesdropper's impact on the capacity is negligible since one can always select sufficiently high P_T to make the saturation effect important (see Fig. 3). To answer the question in the section's title, we use (28) to obtain:

$$\begin{aligned} C^*(\epsilon) &= \max_{\{\lambda_i\}} \sum_i \ln \left(1 + \frac{1 + (g_i - \epsilon)\lambda_i}{1 + \epsilon \lambda_i}\right) \\ &\overset{(a)}{\approx} \max_{\{\lambda_i\}} \sum_i \ln(1 + (g_i - \epsilon)\lambda_i) \\ &\overset{(b)}{\approx} \max_{\{\lambda_i\}} \sum_i \ln(1 + g_i \lambda_i) = C^*(0) \end{aligned} \tag{41}$$

where $\max_{\{\lambda_i\}}$ is subject to $\lambda_i \geq 0, \sum_i \lambda_i = P_T$, and (a) holds if

$$P_T \ll 1/\epsilon \tag{42}$$

(since $\lambda_i \leq P_T$), i.e. if the SNR is not too large, and (b) holds if

$$\epsilon \ll g_i \tag{43}$$

for all active eigenmodes, i.e. if the eavesdropper is much weaker than the legitimate active eigenmodes. It is the combination of (42) and (43) that ensures that the eavesdropper's impact is negligible. Neither condition alone is able to do so. Fig. 3 illustrates this point. Eq. (41) also indicates that the impact of the eavesdropper is the per-eigenmode gain loss of ϵ. Unlike the high-SNR regime in (35) where the loss is multiplicative (i.e. very significant and never negligible), here it is additive (mild or negligible in many cases).

C. Low SNR Regime

Let us now consider the low-SNR regime, which is characteristic for CDMA-type systems [26]. Traditionally, this regime is defined via $P_T \to 0$. We, however, use a more relaxed definition requiring that $m_+ = 1$, which holds under (33). In this regime, assuming $g_1 > \epsilon$,

$$\begin{aligned} C^*(\epsilon) &= \ln \frac{1 + g_1 P_T}{1 + \epsilon P_T} = \ln \left(1 + \frac{(g_1 - \epsilon) P_T}{1 + \epsilon P_T}\right) \\ &\overset{(a)}{\approx} \ln(1 + (g_1 - \epsilon) P_T) \end{aligned} \tag{44}$$

where (a) holds when $P_T \ll 1/\epsilon$. It is clear from the last expression that the impact of the eavesdropper is an additive SNR loss of ϵP_T, which is negligible when $\epsilon \ll g_1$. Note a significant difference to the high SNR regime in (35), where this impact is never negligible. Fig. 3 illustrates this difference.

It follows from (44)(a) that the difference between the lower and upper bounds in (27) at low SNR is the SNR gap of $(\epsilon_1 - \epsilon_m) P_T$. This difference is negligible if $g_1 \gg \epsilon_1 - \epsilon_m$, which may be the case even if the condition number ϵ_1/ϵ_m is large (in which case the difference is significant at high SNR, see (40)). Therefore, we conclude that the impact of the eavesdropper is more pronounced in the high-SNR regime and is negligible in the low-SNR one if its channel is weaker than the strongest eigenmode of the legitimate user, $g_1 \gg \varepsilon_1$.

When $g_1 - \epsilon \ll 1/P_T$, (44)(a) gives $C^*(\epsilon) \approx (g_1 - \epsilon) P_T$, i.e. linear in P_T. A similar capacity scaling at low SNR has been obtained in [29] for i.i.d. block-fading single-input single-output (SISO) WTC, without however explicitly identifying the capacity but via establishing upper/lower bounds. Also note that the 1st two equalities in (44) do not require $P_T \to 0$ but only to satisfy (33).

V. Omnidirectional Eavesdropper

In this section, we consider a scenario where the eavesdropper has equal gain in all directions of a certain subspace. This model accounts for 2 points: (i) when the transmitter has no particular knowledge about the directional properties of the eavesdropper, which is most likely from the practical perspective, it is reasonable to assume that its gain is the same in all directions; (ii) on the other hand, when the eavesdropper has a small number of antennas (less than the number of transmit antennas), its channel rank, which does not exceed the number of transmit or receive antennas, is limited by this number so that the isotropic model of the previous section does not apply.[2]

For an omnidirectional eavesdropper, its channel gain is the same in all directions of its active subspace, i.e.

$$|\mathbf{H}_2 \mathbf{x}|^2 = \mathbf{x}^\dagger \mathbf{W}_2 \mathbf{x} = \text{const} \quad \forall \mathbf{x} \in \mathcal{N}(\mathbf{W}_2)^\perp \tag{45}$$

where $\mathcal{N}(\mathbf{W}_2)^\perp$ is the subspace orthogonal to the nullspace $\mathcal{N}(\mathbf{W}_2)$ of $\mathbf{W}_2$, i.e. its active subspace, whose dimensionality

[2]This was pointed out by A. Khisti.

is $r_2 = \text{rank}(\mathbf{W}_2)$. In particular, when the eavesdropper is isotropic, $\mathcal{N}(\mathbf{W}_2)$ is empty so that $\mathcal{N}(\mathbf{W}_2)^\perp$ is the entire space and $r_2 = m$. The condition in (45) implies that

$$\mathbf{W}_2 = \varepsilon \mathbf{U}_{2+}\mathbf{U}_{2+}^\dagger \tag{46}$$

where $\mathbf{U}_{2+}$ is a semi-unitary matrix whose columns are the active eigenvectors of $\mathbf{W}_2$, and $\mathcal{N}(\mathbf{W}_2)^\perp = \text{span}\{\mathbf{U}_{2+}\}$. Note that the model in (46) allows $\mathbf{W}_2$ to be rank-deficient: $r_2 < m$ is allowed. ε can be evaluated from e.g. (25): $\varepsilon = \alpha n_2 m R_{2\min}^{-\nu}$.

Theorem 2: Under the omnidirectional eavesdropper setting in (45), (46) *and when* $\mathcal{R}(\mathbf{W}_1) \subseteq \mathcal{R}(\mathbf{W}_2)$, *the MIMO-WTC secrecy capacity can be expressed as follows:*

$$\begin{aligned} C_s &= \max_{\text{tr}\,\mathbf{R}\le P_T} \ln \frac{|\mathbf{I}+\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+\mathbf{W_2}\mathbf{R}|} = \max_{\text{tr}\,\mathbf{R}\le P_T} \ln \frac{|\mathbf{I}+\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+\epsilon\mathbf{R}|} \\ &= C^*(\epsilon) \end{aligned} \tag{47}$$

i.e. the capacity and optimal signaling to achieve it are the same as for the isotropic eavesdropper as in Proposition 1.

Proof: First note that, for the omnidirectional eavesdropper, $\mathbf{W}_2 \le \varepsilon\mathbf{I}$ so that $|\mathbf{I}+\mathbf{W}_2\mathbf{R}| \le |\mathbf{I}+\varepsilon\mathbf{R}|$ and hence

$$\begin{aligned} C_s &= \max_{\text{tr}\,\mathbf{R}\le P_T} \ln \frac{|\mathbf{I}+\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+\mathbf{W_2}\mathbf{R}|} \ge \max_{\text{tr}\,\mathbf{R}\le P_T} \ln \frac{|\mathbf{I}+\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+\epsilon\mathbf{R}|} \\ &= C^*(\epsilon) \end{aligned} \tag{48}$$

To prove the reverse inequality, let $\mathbf{P}_2$ be a projection matrix on $\mathcal{R}(\mathbf{W}_2)$, i.e. $\mathbf{P}_2 = \mathbf{U}_{2+}\mathbf{U}_{2+}^\dagger$. Then, $\mathbf{P}_2\mathbf{W}_k\mathbf{P}_2 = \mathbf{W}_k$, $k = 1, 2$, so that

$$C(\mathbf{R}) = \ln \frac{|\mathbf{I}+\mathbf{P}_2\mathbf{W}_1\mathbf{P}_2\mathbf{R}|}{|\mathbf{I}+\mathbf{P}_2\mathbf{W}_2\mathbf{P}_2\mathbf{R}|} = \ln \frac{|\mathbf{I}+\tilde{\mathbf{W}}_1\tilde{\mathbf{R}}|}{|\mathbf{I}+\varepsilon\tilde{\mathbf{R}}|} = \tilde{C}(\tilde{\mathbf{R}})$$

where $\tilde{\mathbf{R}} = \mathbf{U}_{2+}^\dagger\mathbf{R}\mathbf{U}_{2+}$ and likewise for $\tilde{\mathbf{W}}_k$, so that $\tilde{\mathbf{W}}_2 = \varepsilon\mathbf{I}$, where we used $|\mathbf{I}+\mathbf{AB}| = |\mathbf{I}+\mathbf{BA}|$. Further note that

$$\text{tr}\,\tilde{\mathbf{R}} = \text{tr}\,\mathbf{U}_{2+}^\dagger\mathbf{R}\mathbf{U}_{2+} = \sum_i \lambda_i(\mathbf{R})|\mathbf{u}_{2i}^\dagger\mathbf{u}_{Ri}|^2 \tag{49}$$

$$\le \sum_i \lambda_i(\mathbf{R}) = \text{tr}\,\mathbf{R} \le P_T \tag{50}$$

where $\mathbf{u}_{2i}$ and $\mathbf{u}_{Ri}$ are i-th eigenvectors of $\mathbf{W}_2$ and $\mathbf{R}$, and we have used $\mathbf{R} = \sum_i \lambda_i(\mathbf{R})\mathbf{u}_{Ri}\mathbf{u}_{Ri}^\dagger$ and $|\mathbf{u}_{2i}^\dagger\mathbf{u}_{Ri}|^2 \le |\mathbf{u}_{2i}|^2|\mathbf{u}_{Ri}|^2 = 1$. Hence, $\tilde{\mathbf{R}}$ satisfies power constraint if $\mathbf{R}$ does and thus

$$\begin{aligned} C_s &= \max_{\text{tr}\,\mathbf{R}\le P_T} C(\mathbf{R}) \le \max_{\text{tr}\,\tilde{\mathbf{R}}\le P_T} \tilde{C}(\tilde{\mathbf{R}}) \\ &= \max_{\lambda_i\ge 0,\ \sum_i \lambda_i \le P_T} \sum_i \ln \frac{1+\tilde{g}_i\lambda_i}{1+\epsilon\lambda_i} = \tilde{C}^*(\varepsilon) \end{aligned} \tag{51}$$

where $\tilde{g}_i = \lambda_i(\tilde{\mathbf{W}}_1)$, and $\tilde{C}^*(\varepsilon)$ is the secrecy capacity under $\tilde{\mathbf{W}}_1$ and isotropic eavesdropper $\tilde{\mathbf{W}}_2 = \varepsilon\mathbf{I}$. Note that

$$\begin{aligned} \lambda_i(\tilde{\mathbf{W}}_1) &= \lambda_i(\mathbf{U}_{2+}^\dagger\mathbf{W}_1\mathbf{U}_{2+}) = \lambda_i([\mathbf{U}_2^\dagger\mathbf{W}_1\mathbf{U}_2]_{r_2\times r_2}) \\ &\le \lambda_i(\mathbf{U}_2^\dagger\mathbf{W}_1\mathbf{U}_2) = \lambda_i(\mathbf{W}_1) \end{aligned} \tag{52}$$

where $[\mathbf{A}]_{k\times k}$ denotes $k \times k$ principal sub-matrix of $\mathbf{A}$, $r_2 = \text{rank}(\mathbf{W}_2)$, and $\mathbf{U}_2$ is a unitary matrix whose columns are the eigenvectors of $\mathbf{W}_2$. The inequality is due to Cauchy eigenvalue interlacing theorem [21] and the last equality is due to the fact that $\mathbf{U}_2\mathbf{W}_1\mathbf{U}_2^\dagger$ and $\mathbf{W}_1$ have the same eigenvalues. Based on this, one obtains:

$$\begin{aligned} C_s \le \tilde{C}^*(\varepsilon) &\le \max_{\lambda_i\ge 0,\sum_i \lambda_i \le P_T} \sum_i \ln \frac{1+g_i\lambda_i}{1+\epsilon\lambda_i} \\ &= C^*(\varepsilon) \end{aligned} \tag{53}$$

thus establishing $C_s = C^*(\varepsilon)$ under an omnidirectional eavesdropper with $\mathcal{R}(\mathbf{W}_1) \subseteq \mathcal{R}(\mathbf{W}_2)$. □

Note that the secrecy capacity as well as the optimal signaling for the omnidirectional eavesdropper in Theorem 2 is the same as those for the isotropic one (which is not the case in general, as can be shown via examples), i.e. the fact that the rank of the eavesdropper channel is low has no impact provided that $\mathcal{R}(\mathbf{W}_1) \subseteq \mathcal{R}(\mathbf{W}_2)$ holds.

Since $\mathcal{R}(\mathbf{W})$ collects directions where the channel gain is not zero:

$$|\mathbf{H}\mathbf{x}|^2 = \mathbf{x}^\dagger\mathbf{W}\mathbf{x} \ne 0 \quad \forall \mathbf{x} \in \mathcal{R}(\mathbf{W}) \tag{54}$$

the condition $\mathcal{R}(\mathbf{W}_1) \subseteq \mathcal{R}(\mathbf{W}_2)$ means that $|\mathbf{H}_2\mathbf{x}| = 0$ implies $|\mathbf{H}_1\mathbf{x}| = 0$ (but the converse is not true in general) and hence $|\mathbf{H}_1\mathbf{x}| \ne 0$ implies $|\mathbf{H}_2\mathbf{x}| \ne 0$, i.e. the eavesdropper can "see" in any direction where the receiver can "see" (but there is no requirement here for the eavesdropper to be degraded with respect to the receiver so that the channel is not necessarily degraded).

Further note that the condition in (45) does not require $\mathbf{U}_2 = \mathbf{U}_1$, i.e. the eigenvectors of the legitimate channel and of the eavesdropper can be different.

VI. Identical Right Singular Vectors

In this section, we consider the case when $\mathbf{H}_{1,2}$ have the same right singular vectors (SV), so that their singular value decomposition takes the following form:

$$\mathbf{H}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}^\dagger \tag{55}$$

where the unitary matrices $\mathbf{U}_k, \mathbf{V}$ collect left and right singular vectors respectively and diagonal matrix $\mathbf{\Sigma}_k$ collects singular values of $\mathbf{H}_k$. In this model, the left singular vectors can be arbitrary. This is motivated by the fact that right singular vectors are determined by scattering around the Tx while left ones - by scattering around the Rx and eavesdropper respectively. Therefore, when the Rx and eavesdropper are spatially separated, their scattering environments may differ significantly (and hence different left SVs) while the same scattering environment around the Tx induces the same right SVs. We make no weak eavesdropper or other assumptions here. After unitary (and thus information-preserving) transformations, this scenario can be put into the parallel channel setting of [19] and [20]. The secrecy capacity and the optimal covariance in this case can be explicitly characterized as follows.

Proposition 3: Consider the wiretap MIMO channel as in (2), (55). *The optimal Tx covariance for this channel takes the following form:*

$$\mathbf{R}^* = \mathbf{V}\mathbf{\Lambda}^*\mathbf{V}^\dagger \tag{56}$$

where the diagonal matrix $\mathbf{\Lambda}^*$ *collects its eigenvalues* λ_i^**:*

$$\lambda_i^* = f(\lambda_{1i}, \lambda_{2i}) \tag{57}$$

where $f(x, y)$ *is as in* (29)*,* $\lambda_{ki} = \sigma_{ki}^2$ *and* σ_{ki} *denotes singular values of* $\mathbf{H}_k$*;* $\lambda > 0$ *is found from the total power constraint:* $\sum_i \lambda_i^* = P_T$.

Proof: Under (55), $\mathbf{W}_k = \mathbf{V}\mathbf{\Lambda}_k\mathbf{V}^\dagger$, where diagonal matrix $\mathbf{\Lambda}_k = \mathbf{\Sigma}_k^\dagger\mathbf{\Sigma}_k$ collects eigenvalues of $\mathbf{W}_k$, so that the problem in (4) can be re-formulated as

$$C_s = \max_{\mathrm{tr}\,\tilde{\mathbf{R}}\geq\mathbf{0}} \ln \frac{|\mathbf{I} + \mathbf{\Lambda}_1\tilde{\mathbf{R}}|}{|\mathbf{I} + \mathbf{\Lambda}_2\tilde{\mathbf{R}}|} \quad \text{s.t. } \mathrm{tr}\,\tilde{\mathbf{R}} \leq P_T \tag{58}$$

where $\tilde{\mathbf{R}} = \mathbf{V}^\dagger\mathbf{R}\mathbf{V}$. However, this is the secrecy capacity of a set of parallel Gaussian wire-tap channels as in [19] and [20], for which independent signaling is known to be optimal,[3] so that maximizing $\tilde{\mathbf{R}}^*$ is diagonal, from which (56) follows. The optimal power allocation in (57) is essentially the same as for the equivalent parallel channels in [20]. □

In fact, Eq. (56) says that optimal signaling is on the right SVs of $\mathbf{H}_{1,2}$ and (57) implies that only those eigenmodes are active for which

$$\sigma_{1i}^2 > \sigma_{2i}^2 + \lambda \tag{59}$$

If $\lambda_{2i} = 0$, then (57) reduces to

$$\lambda_i^* = \left(\lambda^{-1} - \lambda_{1i}^{-1}\right)_+ \tag{60}$$

i.e. as in the standard WF. This implies that when $\lambda_{2i} = 0$ for all active eigenmodes, then the standard WF power allocation is optimal.

It should be stressed that the original channels in (55) are not parallel (diagonal). They become equivalent to a set of parallel independent channels after performing information-preserving transformations. Also, there is no assumption of degradedness here and no requirement for the optimal covariance to be of full rank or rank-1.

VII. When Is ZF Signaling Optimal?

In this section, we consider the case when ZF signaling is optimal, i.e. when active eigenmodes of the optimal covariance $\mathbf{R}^*$ are orthogonal to those of $\mathbf{W}_2$: $\mathbf{W}_2\mathbf{R}^* = \mathbf{0}$.[4] It is clear that this does not hold in general. However, the importance of this scenario is coming from the fact that such signaling does not require wiretap codes: since the eavesdropper gets no signal, regular coding on the required channel suffices. Hence, the system design follows the well-established standard framework and secrecy requirement imposes no extra complexity penalty but is rather ensured by the well-established ZF signaling.

Proposition 4: A sufficient condition for Gaussian ZF signaling being optimal for the Gaussian MIMO-WTC in (2) *is that* $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *have the same eigenvectors or, equivalently,* $\mathbf{H}_1$ *and* $\mathbf{H}_2$ *have the same right singular vectors as in* (55)*, and*

$$\lambda_{1i} \leq \lambda_{2i} + \lambda \quad \textit{if } \lambda_{2i} > 0, \tag{61}$$

where λ *is found from the total power constraint* $\sum_i \lambda_i^* = P_T$*, and*

$$\lambda_i^* = \lambda_i(\mathbf{R}^*) = \left(\lambda^{-1} - \lambda_{1i}^{-1}\right)_+ \quad \textit{if } \lambda_{2i} = 0, \tag{62}$$

and 0 otherwise. The optimal covariance is as in (56) *so that its eigenvectors are those of* $\mathbf{W}_1$ *and* $\mathbf{W}_2$*.*

A necessary condition of ZF optimality is that the active eigenvectors of $\mathbf{R}^*$ *are also the active eigenvectors of* $\mathbf{W}_1$ *and the inactive eigenvectors of* $\mathbf{W}_2$*, and that the power allocation is given by* (62)*.*

Proof: See the Appendix. □

Remark 4: The optimal power allocation in (62) is the same as standard water filling. However, a subtle difference here is the condition for an eigenmode to be active, $\lambda_i^* > 0$: while the standard WF requires $\lambda_{1i} > \lambda$, the solution above requires in addition $\lambda_{2i} = 0$, so that the set of active eigenmodes is generally smaller: the larger the set of eavesdropper positive eigenmodes, the smaller the set of active eigenmodes.

It is gratifying to see that the standard WF over the eigenmodes of the required channel is optimal if ZF is optimal. In a sense, the optimal transmission strategy in this case is separated into two independent parts: part 1 ensures that the eavesdropper gets no signal (via the ZF) and part 2 is the standard eigenmode signaling and WF on what remains of the required channel as if the eavesdropper were not there. No new wiretap codes need to be designed.

VIII. When Is the Standard Water Filling Optimal?

Motivated by the fact that the transmitter may be unaware about the presence of an eavesdropper and hence uses the standard transmission on the eigenmodes of $\mathbf{W}_1$ with power allocated via the water-filling (WF) algorithm, we ask the question: is it possible for this strategy to be optimal for the MIMO-WTC? The affirmative answer and conditions for this to happen are given below. To this end, let $\mathbf{R}_{WF}$ be the optimal Tx covariance matrix for transmission on $\mathbf{W}_1$ only, which is given by the standard water-filling over the eigenmodes of $\mathbf{W}_1$:

$$\mathbf{R}_{WF} = \mathbf{U}_1\mathbf{\Lambda}^*\mathbf{U}_1^\dagger, \quad \lambda_i^* = \left\{\lambda^{-1} - \lambda_{1i}^{-1}\right\}_+ \tag{63}$$

where $\mathbf{\Lambda}^* = diag\{\lambda_i^*\}$ is a diagonal matrix of the eigenvalues of $\mathbf{R}_{WF}$, and λ is found from the total power constrain $\sum_i \lambda_i^* = P_T$.

Theorem 3: The standard WF Tx covariance matrix in (63) *is also optimal for the Gaussian MIMO-WTC if:*

1) the eigenvectors of $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *are the same:* $\mathbf{U}_1 = \mathbf{U}_2$*;*

2) for active eigenmodes $\lambda_i^* > 0$*, their eigenvalues* λ_{1i} *and* λ_{2i} *are related as follows:*

$$\lambda_{2i} = \frac{\lambda_{1i}}{1 + \alpha\lambda_{1i}} < \lambda_{1i}, \quad \textit{for some } \alpha > 0, \tag{64}$$

or, equivalently, $\lambda_{2i}^{-1} = \lambda_{1i}^{-1} + \alpha$*;*

[3] The authors would like to thank A. Khisti for pointing out this line of argument.

[4] This simply means that the Tx antenna array puts null in the direction of eavesdropper, which is known as null forming in antenna array literature [16]. This can also be considered as a special case of interference alignment, so that Proposition 4 establishes its optimality.

3) for inactive eigenmodes $\lambda_i^* = 0$, the eigenvalues λ_{1i} and λ_{2i} are related either as in (64) *or* $\lambda_{1i} \leq \lambda_{2i}$.

Proof: We assume that $\mathbf{W}_1$ and $\mathbf{W}_2$ are non-singular; the singular case will be considered below (using a standard continuity argument). The KKT conditions for the optimal covariance $\mathbf{R} = \mathbf{R}_{WF}$, which are necessary for optimality in (4), can be expressed as:

$$(\mathbf{W}_1^{-1} + \mathbf{R})^{-1} - (\mathbf{W}_2^{-1} + \mathbf{R})^{-1} = \lambda' \mathbf{I} - \mathbf{M} \tag{65}$$

$$\lambda'(\mathrm{tr}\,\mathbf{R} - P_T) = 0, \quad \mathbf{MR} = 0 \tag{66}$$

$$\lambda' \geq 0, \quad \mathbf{M}, \mathbf{R} \geq 0, \ tr\mathbf{R} \leq P_T \tag{67}$$

where $\mathbf{M} \geq 0$ is the Lagrange multiplier matrix responsible for the constraint $\mathbf{R} \geq 0$ while $\lambda' \geq 0$ is the Lagrange multiplier responsible for the total power constraint $tr\mathbf{R} \leq P_T$. Multiplying both sides of (65) by $\mathbf{U}_1^\dagger$ on the left and by $\mathbf{U}_1$ on the right, one obtains:

$$\begin{aligned}(\mathbf{\Lambda}_1^{-1} + \mathbf{\Lambda}^*)^{-1} - (\mathbf{\Lambda}_2^{-1} + \mathbf{\Lambda}^*)^{-1} &= \lambda' \mathbf{I} - \mathbf{U}_1^\dagger \mathbf{M} \mathbf{U}_1 \\ &= \lambda' \mathbf{I} - \mathbf{\Lambda}_M \end{aligned} \tag{68}$$

where $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \mathbf{\Lambda}_M$ are diagonal matrices of eigenvalues of $\mathbf{W}_1, \mathbf{W}_2, \mathbf{M}$. The last equality follows from the fact that all terms but $\mathbf{U}_1^\dagger \mathbf{M} \mathbf{U}_1$ are diagonal so that the last term has to be diagonal too: $\mathbf{U}_1^\dagger \mathbf{M} \mathbf{U}_1 = \mathbf{\Lambda}_M$, i.e. $\mathbf{M}$ has the same eigenvectors as $\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}$. The complementary slackness in (66) implies that $\lambda_i^* \lambda_{Mi} = 0$, where λ_{Mi} is i-th eigenvalue of $\mathbf{M}$, i.e. if $\lambda_i^* > 0$ (active eigenmode) then $\lambda_{Mi} = 0$ so that, after some manipulations, (68) can be expressed as

$$\lambda_i^* = \frac{1}{(\lambda_{2i}^{-1} + \lambda_i^*)^{-1} + \lambda'} - \frac{1}{\lambda_{1i}} = \lambda^{-1} - \lambda_{1i}^{-1}$$

for each $\lambda_i^* > 0$, where the 2nd equality follows from (63). Therefore, $\lambda = (\lambda_{2i}^{-1} + \lambda_i^*)^{-1} + \lambda'$ and hence

$$\lambda_i^* = (\lambda - \lambda')^{-1} - \lambda_{2i}^{-1} = \lambda^{-1} - \lambda_{1i}^{-1} \tag{69}$$

so that $\lambda_{2i}^{-1} = \lambda_{1i}^{-1} + \alpha$ with $\alpha = (\lambda - \lambda')^{-1} - \lambda^{-1} > 0$ satisfies both equalities in (69).

For inactive eigenmodes $\lambda_i^* = 0$, it follows from (68) that

$$\lambda_{1i} - \lambda_{2i} = \lambda' - \lambda_{Mi} \leq \lambda' \tag{70}$$

Observe that this inequality is satisfied when $\lambda_{1i} \leq \lambda_{2i}$ (since $\lambda' > 0$). To see that it also holds under (64), observe that

$$\lambda_{1i} - \lambda_{2i} = \frac{\alpha \lambda_{1i}^2}{1 + \alpha \lambda_{1i}} \leq \frac{\alpha \lambda^2}{1 + \alpha \lambda} = \lambda' \tag{71}$$

where the inequality is due to $\lambda_{1i} \leq \lambda$ (which holds for inactive eigenmodes) and the fact that $\frac{\alpha \lambda_{1i}^2}{1 + \alpha \lambda_{1i}}$ is increasing in λ_{1i}. Thus, one can always select $\lambda_{Mi} \geq 0$ to satisfy (70) and hence the KKT conditions in (65)-(67) have a unique solution which also satisfies (63). This proves the optimality of $\mathbf{R}_{WF}$.

In the case of singular $\mathbf{W}_1$ or/and $\mathbf{W}_2$, one can use a standard continuity argument, see [30] for details. □

Note that the conditions of Theorem 3 do not require $\mathbf{W}_1 = a\mathbf{W}_2$ for some scalar $a > 1$; they also allow for the WTC to be non-degraded. However, the condition in (64) implies that larger λ_{1i} corresponds to larger λ_{2i}, so that, over the active signaling subspace, the channel is degraded.

The 1st condition in Theorem 3 implies that $\mathbf{H}_1$ and $\mathbf{H}_2$ have the same right singular vectors but imposes no constraints on their left singular vectors. This may represent a scenario where the transmitter is a basestation where the legitimate channel and the eavesdropper experience the same scattering while having their own individual scatterers around their own receivers (which determine the left singular vectors), as in Section VI.

IX. When Is Isotropic Signaling Optimal?

In the regular MIMO channel ($\mathbf{W_2} = \mathbf{0}$), the isotropic signaling (IS) is optimal ($\mathbf{R}^* = a\mathbf{I}$) iff $\mathbf{W}_1 = b\mathbf{I}$, i.e. $\mathbf{W}_1$ has identical eigenvalues. Since this transmission strategy is appealing due to its low complexity (all antennas send independent data streams, no precoding, no Tx CSI and thus no feedback is required), we consider the isotropic signaling over the wire-tap MIMO channel and characterize the set of channels on which it is optimal. It turns out to be much richer than that of the regular MIMO channel.

Proposition 5: Consider the MIMO wire-tap channel in (2)*. The isotropic signaling is optimal, i.e.* $\mathbf{R}^* = a\mathbf{I}$ *in* (4)*, for the set of channels* $\{\mathbf{W}_1, \mathbf{W}_2\}$ *that satisfy all of the following:*

1. $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *have the same (otherwise arbitrary) eigenvectors,* $\mathbf{U}_1 = \mathbf{U}_2$.

2. $\mathbf{W}_1 > \mathbf{W}_2$ *so that* $\lambda_i(\mathbf{W}_1) = a_i^{-1} > \lambda_i(\mathbf{W}_2) = b_i^{-1}$*, where* $\lambda_i(\mathbf{W})$ *are ordered eigenvalues of* $\mathbf{W}$.

3. Take any $b_1 > 0$ *and* $a_1 < b_1$ *and set* $\lambda = (a_1 + a)^{-1} - (b_1 + a)^{-1} > 0$*,*

4. For $i = 2 \ldots m$*, take any* b_i *such that* $b_i > \lambda a^2 (1 - \lambda a)^{-1} > 0$*, and set*

$$a_i = -a + (\lambda + (b_i + a)^{-1})^{-1} > 0 \tag{72}$$

This gives the complete characterization of the set of channels for which isotropic signaling is optimal.

Proof: It is straightforward to see that any channel in the given set satisfies the conditions of [6, Th. 2] and the corresponding optimal covariance is isotropic, which proves the sufficiency. The converse (necessity) follows from [6, Th. 1], which requires $\mathbf{W}_1 > \mathbf{W}_2$, so that the optimization problem is strictly convex and thus has a unique solution. For isotropic signaling to be optimal, the corresponding KKT conditions (see the proofs of Theorems 1 and 2 in [6]) imply the conditions stated above. □

Note that the special case of this Proposition is when $\mathbf{W}_1$ and $\mathbf{W}_2$ have identical eigenvalues, as in the case of the regular MIMO channel, but, unlike the regular channel, there is also a large set of channels with distinct eigenvalues which dictates the isotropic signaling as well. It is the interplay between the legitimate user and the eavesdropper that is responsible for this phenomenon, i.e. a non-isotropic nature of the 1st channel is compensated for by a carefully-adjusted non-isotropy of the 2nd one.

Table I summarizes the conditions for the optimality of the ZF, the WF and the IS in the Gaussian MIMO-WTC. Clearly, the requirement for $\mathbf{W}_1$ and $\mathbf{W}_2$ to have the same eigenvectors is the key condition. It is satisfied when the legitimate receiver and the eavesdropper are subject to the same scattering around the base station (the transmitter) while

TABLE I
THE CONDITIONS OF OPTIMALITY OF THE ZF, THE WF AND THE IS IN THE GAUSSIAN MIMO-WTC

Strategy	Optimality conditions
WF	$\mathbf{U}_1 = \mathbf{U}_2$; $\lambda_{1i}, \lambda_{2i}$ as in Theorem 3
ZF	$\mathbf{U}_1 = \mathbf{U}_2$; $\lambda_{1i}, \lambda_{2i}$ as in Proposition 4
IS	$\mathbf{U}_1 = \mathbf{U}_2$; $\lambda_{1i}, \lambda_{2i}$ as in Proposition 5

they may have their own sets of scatterers around their own units.

APPENDIX

A. Proof of Theorem 1

Applying the inequalities

$$x - x^2/2 \le \ln(1+x) \le x \tag{73}$$

which hold for any $x \ge 0$, to

$$\ln|\mathbf{I} + \mathbf{W}_2\mathbf{R}| = \sum_i \ln(1 + \lambda_i(\mathbf{W}_2\mathbf{R})) \tag{74}$$

one obtains:

$$C_w(\mathbf{R}) \le C(\mathbf{R}) \le C_w(\mathbf{R}) + \frac{1}{2}\sum_i \lambda_i^2(\mathbf{W}_2\mathbf{R}) \tag{75}$$

from which the 1st inequality in (7) follows by using $\mathbf{R} = \mathbf{R}_w^*$; the 2nd inequality follows from the fact that $C(\mathbf{R})$ is maximized by $\mathbf{R}^*$: $C_s = C(\mathbf{R}^*) \ge C(\mathbf{R}_w^*)$. To obtain the last inequality, we need the following lemma.

Lemma 1: Let $\lambda_i \ge 0$ and $\sum_i \lambda_i \le P_T$. Then, $\sum_i \lambda_i^2 \le P_T^2$.

Using this Lemma and observing that $\lambda_i(\mathbf{W}_2\mathbf{R}) \le \lambda_1(\mathbf{W}_2)\lambda_i(\mathbf{R})$ (see e.g. [21]), one obtains:

$$\sum_i \lambda_i^2(\mathbf{W}_2\mathbf{R}) \le \lambda_1^2(\mathbf{W}_2)\sum_i \lambda_i^2(\mathbf{R}) \le \lambda_1^2(\mathbf{W}_2)P_T^2 \tag{76}$$

since $\sum_i \lambda_i(\mathbf{R}) \le P_T$, so that

$$\begin{aligned} C_s = C(\mathbf{R}^*) &\le C_w(\mathbf{R}^*) + \lambda_1^2(\mathbf{W}_2)P_T^2/2 \\ &\le C_w + \lambda_1^2(\mathbf{W}_2)P_T^2/2 \end{aligned} \tag{77}$$

since $C_w = C_w(\mathbf{R}_w^*) \ge C_w(\mathbf{R}^*)$, which establishes the last inequality in (7).

To establish the closed form solution for C_w in (12), consider the optimization problem in (5), for which the KKT conditions are:

$$(\mathbf{I} + \mathbf{W}_1\mathbf{R})^{-1}\mathbf{W}_1 - \mathbf{W}_2 - \lambda\mathbf{I} + \mathbf{M} = \mathbf{0} \tag{78}$$

$$\lambda(\mathrm{tr}\,\mathbf{R} - P_T) = 0, \quad \mathbf{M}\mathbf{R} = \mathbf{0} \tag{79}$$

$$\lambda \ge 0, \quad \mathbf{M}, \mathbf{R} \ge \mathbf{0} \tag{80}$$

where $\lambda \ge 0$ is a Lagrange multiplier responsible for the total power constraint and $\mathbf{M} \ge \mathbf{0}$ is a matrix Lagrange multiplier responsible for the constraint $\mathbf{R} \ge \mathbf{0}$. Since the objective is concave, the corresponding optimization problem is convex, and since Slater condition holds (e.g. take $\mathbf{R} = P_T\mathbf{I}/2 > 0, tr\mathbf{R} < P_T$), the KKT conditions are sufficient for optimality [18]. After some manipulations, (78) can be transformed to

$$\widehat{\mathbf{R}} - (\mathbf{I} - \widehat{\mathbf{M}})^{-1} = -\widehat{\mathbf{W}}_1^{-1} \tag{81}$$

$$\widehat{\mathbf{R}} = \mathbf{W}_\lambda^{1/2}\mathbf{R}\mathbf{W}_\lambda^{1/2}, \quad \widehat{\mathbf{M}} = \mathbf{W}_\lambda^{-1/2}\mathbf{M}\mathbf{W}_\lambda^{-1/2},$$
$$\widehat{\mathbf{W}}_1 = \mathbf{W}_\lambda^{-1/2}\mathbf{W}_1\mathbf{W}_\lambda^{-1/2} \tag{82}$$

where we implicity assume that $\mathbf{W}_1$ and $\mathbf{W}_\lambda$ are non-singular, so that $\mathbf{Q} = \mathbf{W}_\lambda^{-1}$; the singular case will be considered below. Since $\widehat{\mathbf{M}}\widehat{\mathbf{R}} = \mathbf{0}$ (which follows from $\mathbf{M}\mathbf{R} = \mathbf{0}$), these matrices commute and thus have the same eigenvectors, which, from (81), implies that these eigenvectors are the same as those of $\widehat{\mathbf{W}}_1$. Hence, all three matrices can be simultaneously diagonalized and thus (81) can be transformed to diagonal form from which (8) follows after some manipulations [30].

The existence of the threshold power P_T^* follows from the fact that $\mathrm{tr}\,\mathbf{R}^*$ is monotonically decreasing in λ so that its largest value corresponds to $\lambda \to 0$ and equals P_T^*. When $P_T > P_T^*$, $\lambda = 0$ and $\mathrm{tr}\,\mathbf{R}^* = P_T^* < P_T$, i.e. only partial power is used (see Fig. 1 for illustration and discussion). The case of singular $\mathbf{W}_2$ and $\mathbf{W}_\lambda$ is considered in [30].

B. Proof of Proposition 1

The 1st equality in (28) follows from (4). The 2nd equality follows from the Hadamard inequality applied to $|\mathbf{I} + \mathbf{W}_1\mathbf{R}|$ in the same way as for the regular MIMO channel, and the equality is achieved when $\mathbf{R}$ has the same eigenvectors as $\mathbf{W}_1$, $\mathbf{R}^* = \mathbf{U}_1\Lambda^*\mathbf{U}_1^\dagger$, which maximizes the numerator and leaves the denominator unchanged. The remaining part is the optimal power allocation in (29), which can be formulated as

$$C^*(\epsilon) = \max_{\{\lambda_i\}} \sum_i \ln\frac{1 + g_i\lambda_i}{1 + \epsilon\lambda_i}, \quad \text{s.t. } \lambda_i \ge 0, \ \sum_i \lambda_i = P_T \tag{83}$$

This, however, represents an optimal power allocation for parallel channels which can be found in [20].

The lower/upper bounds follow from the fact that $|\mathbf{I}+\mathbf{W}\mathbf{R}|$ is a matrix-monotone function of $\mathbf{W}$ [21], so that $|\mathbf{I}+\mathbf{W}_b\mathbf{R}| \ge |\mathbf{I}+\mathbf{W}_a\mathbf{R}|$ $\forall \mathbf{W}_b \ge \mathbf{W}_a \ge \mathbf{0}$. The gap bound in (30) is proved in [30].

C. Proof of Proposition 4

The original problem in (4) is not convex in general. However, since the objective is continuous, the feasible set is compact and Slater condition holds, KKT conditions are necessary for optimality [22]. They take on the following form (see e.g. [6]):

$$\lambda\mathbf{W}_1\mathbf{R} = \mathbf{W}_1 - \mathbf{W}_2 + \mathbf{M} - \lambda\mathbf{I} \tag{84}$$

$$\lambda(\mathrm{tr}\,\mathbf{R} - P_T) = 0, \quad \mathbf{M}\mathbf{R} = 0 \tag{85}$$

$$\lambda \ge 0, \quad \mathbf{M}, \mathbf{R} \ge 0, \ \ \mathrm{tr}\,\mathbf{R} \le P_T \tag{86}$$

where $\mathbf{M} \ge 0$ is the Lagrange multiplier matrix responsible for the constraint $\mathbf{R} \ge 0$ while $\lambda \ge 0$ is the Lagrange multiplier responsible for the total power constraint $tr\mathbf{R} \le P_T$, and we used the orthogonality condition $\mathbf{W}_2\mathbf{R} = 0$.

To prove sufficiency, note from Proposition 3 that if $\mathbf{W}_1, \mathbf{W}_2$ have the same eigenvectors so is $\mathbf{R}$ and hence $\mathbf{M}$ and also the KKT conditions are sufficient for optimality (since they have a unique solution). Hence, (84) can be transformed to a diagonal form:

$$\lambda\lambda_{1i}\lambda_i = \lambda_{1i} - \lambda_{2i} + \lambda_{Mi} - \lambda \tag{87}$$

where λ_i, λ_{Mi} are the eigenvalues of $\mathbf{R}, \mathbf{M}$. Complementary slackness in (85) gives $\lambda_i \lambda_{Mi} = 0$ so that $\lambda_i > 0$ (active eigenmodes) implies $\lambda_{Mi} = 0$ and hence

$$\lambda_i = \frac{\lambda_{1i} - \lambda_{2i} - \lambda}{\lambda \lambda_{1i}} = \frac{1}{\lambda} - \frac{1}{\lambda_{1i}} \tag{88}$$

where the 2nd equality follows from the orthogonality condition $\lambda_{2i}\lambda_i = 0$. For inactive eigenmodes $\lambda_i = 0$, one obtains $\lambda_{Mi} = \lambda - \lambda_{1i} + \lambda_{2i} \geq 0$ so that $\lambda_{1i} \leq \lambda + \lambda_{2i}$. A proof of the necessary part can be found in [30].

ACKNOWLEDGEMENT

The authors would like to thank M. Urlea and K. Li for running numerical experiments and generating Fig. 1, and A. Khisti for suggesting the problem formulation in Section V.

REFERENCES

[1] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

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

[3] 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, Nov. 2010.

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

[5] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4033–4039, Sep. 2009.

[6] S. Loyka and C. D. Charalambous, "On optimal signaling over secure MIMO channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Boston, MA, USA, Jul. 2012, pp. 443–447.

[7] J. Li and A. Petropulu. (Sep. 2009). "Transmitter optimization for achieving secrecy capacity in Gaussian MIMO wiretap channels." [Online]. Available: http://arxiv.org/abs/0909.2622

[8] J. Li and A. Petropulu, "Optimal input covariance for achieving secrecy capacity in Gaussian MIMO wiretap channels," in *Proc. IEEE ICASSP*, Mar. 2010, pp. 3362–3365.

[9] M. C. Gursoy, "Secure communication in the low-SNR regime," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1114–1123, Apr. 2012.

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

[11] A. Khabbazibasmenj, M. A. Girnyk, S. A. Vorobyov, M. Vehkaperae, and L. K. Rasmussen, "On the optimal precoding for MIMO Gaussian wire-tap channels," in *Proc. 10th Int. Symp. Wireless Commun. Syst. (ISWCS)*, Ilmenau, Germany, Aug. 2013, pp. 1–4.

[12] J. Steinwandt, S. A. Vorobyov, and M. Haardt, "Secrecy rate maximization for MIMO Gaussian wiretap channels with multiple eavesdroppers via alternating matrix POTDC," in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 5686–5690.

[13] A. Alvarado, G. Scutari, and J.-S. Pang, "A new decomposition method for multiuser DC-programming and its applications," *IEEE Trans. Signal Process.*, vol. 62, no. 11, pp. 2984–2998, Jun. 2014.

[14] S. Loyka and C. D. Charalambous, "An algorithm for global maximization of secrecy rates in Gaussian MIMO wiretap channels," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2288–2299, Jun. 2015.

[15] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[16] H. L. Van Trees, *Optimum Array Processing: Detection, Estimation, and Modulation Theory*. New York, NY, USA: Wiley, 2002.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 2006.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[19] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453–2469, Jun. 2008.

[20] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," in *Securing Wireless Communications at the Physical Layer*, R. Liu and W. Trappe, Eds. New York, NY, USA: Springer, 2010.

[21] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[22] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 2008.

[23] A. Khisti, "Interference alignment for the multiantenna compound wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2976–2993, May 2011.

[24] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy results for compound wiretap channels," *Problems Inf. Transmiss.*, vol. 49, no. 1, pp. 73–98, Mar. 2013.

[25] R. F. Schaefer and S. Loyka, "The secrecy capacity of a compound MIMO Gaussian channel," in *Proc. IEEE Inf. Theory Workshop*, Seville, Spain, Sep. 2013, pp. 1–5.

[26] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[27] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[28] S. Loyka and G. Levin, "On physically-based normalization of MIMO channel matrices," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1107–1112, Mar. 2009.

[29] Z. Rezki, A. Khisti, and M.-S. Alouini, "Ergodic secret message capacity of the wiretap channel with finite-rate feedback," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3364–3379, Jun. 2014.

[30] S. Loyka and C. D. Charalambous. (Apr. 2016). "Rank-deficient solutions for optimal signaling over wiretap MIMO channels." [Online]. Available: https://arxiv.org/abs/1604.06785

Sergey Loyka was born in Minsk, Belarus. He received the Ph.D. degree in radio engineering from the Belorussian State University of Informatics and Radioelectronics (BSUIR), Minsk, in 1995, and the M.S. (Hons.) degree from the Minsk Radioengineering Institute, Minsk, in 1992. He was a Research Fellow with the Laboratory of Communications and Integrated Microelectronics, Ecole de Technologie Superieure, Montreal, Canada, a Senior Scientist with the Electromagnetic Compatibility Laboratory, BSUIR, and an Invited Scientist with the Laboratory of Electromagnetism and Acoustic, Swiss Federal Institute of Technology, Lausanne, Switzerland. Since 2001, he has been a Faculty Member with the School of Electrical Engineering and Computer Science, University of Ottawa, Canada. His research areas are wireless communications and networks, in particular, MIMO systems and security aspects of such systems, in which he has published extensively. He received a number of awards from the URSI, the IEEE, the Belarus, Switzerland, and former USSR governments, and the Soros Foundation.

Charalambos D. Charalambous received the B.S. degree in electrical engineering, the M.E. degree, and the Ph.D. degree from the Department of Electrical Engineering, Old Dominion University, Norfolk, VA, in 1987, 1988, and 1992, respectively. From 1993 to 1995, he was a Post-Doctoral Fellow with the Engineering Department, Idaho State University, Pocatello. He served with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, as a Nontenure Faculty Member, from 1995 to 1999. He was an Associate Professor with the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada, from 1999 to 2003. In 2003, he joined the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus, where he is currently a Professor. His research group is interested in theoretical and technological developments concerning large scale distributed communication and control systems and networks in science and engineering. He is currently an Associate Editor of *Systems and Control Letters* and *Mathematics of Control, Signals, and Systems*, and the Vice Chair of the IFAC Technical Committee of Stochastic Systems. He served as an Associate Editor of the IEEE COMMUNICATIONS LETTERS and the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.