Rank-Deficient Solutions for Optimal Signaling over Secure MIMO Channels

Sergey Loyka, Charalambos D. Charalambous

***Abstract*—Capacity-achieving signaling strategies for the Gaussian wiretap 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 weak eavesdropper is considered in details and the optimal covariance is established in an explicit, closed-form with no extra assumptions.**

The conditions for optimality of zero-forcing signaling are established, and the standard water-filling is shown to be optimal under those conditions. No wiretap codes are needed in this case.

The case of identical right singular vectors for the required and eavesdropper channels is studied and the optimal covariance is established in an explicit closed form. As a by-product of this analysis, we establish a generalization of celebrated Hadamard determinantal inequality using information-theoretic tools.

I. Introduction

Wide-spread use of wireless systems has initiated significant interest in their security and information-theoretic aspects of the latter [1]. In particular, the wire-tap Gaussian MIMO channel has been a subject of intensive studies and a number of results have been obtained, including the proof of optimality of Gaussian signaling [1]-[4].

The optimal transmit covariance matrix under the total power constraint has been obtained for some special cases (e.g. low/high SNR, MISO channels, full-rank or rank-1 cases) [2]-[7] but the general case remains illusive. The main difficulty lies in the fact that the underlying optimization problem is not convex in general. It was conjectured in [4] and proved in [3] using an indirect approach (via the degraded channel) that the optimal signaling is on the positive directions of the difference channel. A direct proof (based on the necessary KKT conditions) has been obtained in [5]. A weaker result (non-negative instead of positive directions) has been obtained in [7]. An exact full-rank solution for the optimal covariance has been obtained in [5] and its properties have been characterized. In particular, unlike the regular channel (no eavesdropper), the optimal power allocation does not converge to uniform one at high SNR and the latter remains sub-optimal at any finite SNR. In the case of weak eavesdropper, the optimal signaling mimics the conventional one (water-filling over the channel eigenmodes) with an adjustment for the eavesdropper channel. The case of isotropic eavesdropper is studied in details in [6], including the optimal signaling in an explicit closed form and its properties. This case is shown to be the worst-case MIMO wire-tap channel. Based on this, lower and upper (tight) capacity bounds have been obtained for the general case, which are achievable by an isotropic eavesdropper. The set of channels for which isotropic signaling is optimal has been fully characterized [6]. It turns out to be much richer than that of the conventional (no eavesdropper) MIMO channel.

The present paper extends the known results for optimal covariance in several directions. The case of weak eavesdropper is studied and the optimal covariance is obtained in an explicit closed form without any extra assumptions (e.g. full rank or rank-1). It provides a lower bound to the secrecy capacity in the general case, which is tight when the eavesdropper path loss is large and hence serves as an approximation to the true capacity. It also captures the capacity saturation effect at high SNR observed in [3][5].

The case of identical right singular vectors of the required and eavesdropper channels is investigated and the optimal covariance is established in a closed from. This case is motivated by a scenario where the legitimate receiver (Rx) and the eavesdropper (Ev) 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 (Tx) (e.g. a base station) and hence the same right SVs. This is similar to the popular Kronecker MIMO channel correlation model [8], where the overall channel correlation is a product of the independent Tx and Rx parts, which are induced by respective sets of scatterers. As a by-product of this analysis, a generalization of the celebrated Hadamard determinantal inequality is established, which applies to a ratio of two determinants, using information-theoretic tools in the spirit of [9].

Finally, the conditions for optimality of zero-forcing (ZF) signaling are established, where the Tx antenna array forms a null in the Ev direction. Under those conditions, the standard eigenmode signaling and 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 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 [10]), which insures 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

S. Loyka is with the School of Electrical Engineering and Computer Science, University of Ottawa, Ontario, Canada, K1N 6N5, e-mail: sergey.loyka@ieee.org.

C.D. Charalambous is with the ECE Department, University of Cyprus, 75 Kallipoleos Avenue, P.O. Box 20537, Nicosia, 1678, Cyprus, e-mail: chadcha@ucy.ac.cy

978-1-4799-5186-4/14/$31.00 ©2014 Crown

source-channel coding separation [11].

II. Wire-Tap Gaussian MIMO Channel Model

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

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \xi_1, \quad \mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \xi_2 \tag{1}$$

where $\mathbf{x} = [x_1, x_2, ...x_m]^T \in C^{m,1}$ is the transmitted complex-valued signal vector of dimension $m \times 1$, "T" denotes transposition, $\mathbf{y}_{1(2)} \in C^{n,1}$ are the received vectors at the receiver (eavesdropper), $\xi_{1(2)}$ is the circularly-symmetric additive white Gaussian noise at the receiver (eavesdropper) (normalized to unit variance in each dimension), $\mathbf{H}_{1(2)} \in C^{n_{1(2)},m}$ is the $n_{1(2)} \times m$ matrix of the complex channel gains between each Tx and each receive (eavesdropper) antenna, $n_{1(2)}$ and m are the numbers of Rx (eavesdropper) and Tx antennas respectively. The channels $\mathbf{H}_{1(2)}$ 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 $\mathbf{R} = E\{\mathbf{x}\mathbf{x}^+\}$, where $E\{\cdot\}$ is statistical expectation, the maximum achievable secure rate between the Tx and Rx (so that the rate between the Tx and eavesdropper is zero) is [3][4]

$$C(\mathbf{R}) = \ln\frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \mathbf{W}_2\mathbf{R}|} = C_1(\mathbf{R}) - C_2(\mathbf{R}) \tag{2}$$

where negative $C(\mathbf{R})$ is interpreted as zero rate, $\mathbf{W}_i = \mathbf{H}_i^+\mathbf{H}_i$, $()^+$ means Hermitian conjugation, and the secrecy capacity subject to the total Tx power constraint is

$$C_s = \max_{\mathbf{R}\geq 0} C(\mathbf{R}) \text{ s.t. } tr\mathbf{R} \leq P_T \tag{3}$$

where P_T is the total transmit power (also the SNR since the noise is normalized). It is well-known that the problem in (3) is not convex in general and explicit solutions for the optimal Tx covariance is 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]-[7]).

Notations: $\lambda_i(\mathbf{W})$ denotes eigenvalues of matrix $\mathbf{W}$; $(x)_+ = \max\{x, 0\}$ for a scalar x; $\mathcal{N}(\mathbf{W})$ is a null space of matrix $\mathbf{W}$; $(\mathbf{W})_+$ denotes positive eigenmodes of Hermitian matrix $\mathbf{W}$:

$$(\mathbf{W})_+ = \sum_{i:\lambda_i(\mathbf{W})>0} \lambda_i \mathbf{u}_i\mathbf{u}_i^+ \tag{4}$$

where $\mathbf{u}_i$ is i-th eigenvector of $\mathbf{W}$.

III. Weak Eavesdropper

In this section, we consider the scenario where the eavesdropper is weak. This may be due to the fact that the eavesdropper is located far away from the Tx so that its path loss is large. There is no requirement for the channel to be degraded or for the optimal covariance to be of full rank or rank 1, so that this result adds considerably to the known solutions.

Theorem 1. *Consider the problem in* (3) *when the eavesdropper is weak, i.e.* $\lambda_i(\mathbf{W}_2\mathbf{R}) \ll 1$. *The optimal covariance is given by*

$$\mathbf{R}^* = \mathbf{W}_\lambda^{-1/2}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^+\mathbf{W}_\lambda^{-1/2} \tag{5}$$

where $\mathbf{W}_\lambda = \lambda\mathbf{I} + \mathbf{W}_2$, *the columns of unitary matrix* $\mathbf{U}$ *are the eigenvectors of*[1]

$$\widetilde{\mathbf{W}}_1 = \mathbf{W}_\lambda^{-1/2}\mathbf{W}_1\mathbf{W}_\lambda^{-1/2}, \tag{6}$$

the diagonal entries of diagonal matrix $\mathbf{\Lambda}$ *are*

$$\widetilde{\lambda}_i = (1 - \lambda_i^{-1}(\widetilde{\mathbf{W}}_1))_+ \tag{7}$$

where $\lambda \geq 0$ *is found from the total power constraint:*

$$tr\mathbf{R}^* = P_T \text{ if } P_T < P_T^* \tag{8}$$

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

$$P_T^* = tr\mathbf{W}_2^{-1}(\mathbf{I} - \mathbf{W}_2^{1/2}\mathbf{W}_1^{-1}\mathbf{W}_2^{1/2})_+ \tag{9}$$

if $\mathbf{W}_2$ *is non-singular, and* $P_T^* = \infty$ *if* $\mathbf{W}_2$ *is singular and* $\mathcal{N}(\mathbf{W}_2) \notin \mathcal{N}(\mathbf{W}_1)$[2]. *The corresponding secrecy capacity is:*

$$C_s = \sum_{i:\widetilde{\lambda}_{1i}>1} \ln\widetilde{\lambda}_{1i} - w_i(1 - \widetilde{\lambda}_{1i}^{-1}) \tag{10}$$

where $\widetilde{\lambda}_{1i} = \lambda_i(\widetilde{\mathbf{W}}_1)$, w_i *is* i*-th diagonal entry of* $\mathbf{U}^+\mathbf{W}_\lambda^{-1/2}\mathbf{W}_2\mathbf{W}_\lambda^{-1/2}\mathbf{U}$.

Proof: Under the weak eavesdropper assumption, the secrecy capacity can be approximated as[3]

$$C(\mathbf{R}) \approx \ln|\mathbf{I} + \mathbf{W}_1\mathbf{R}| - tr(\mathbf{W}_2\mathbf{R}) \tag{11}$$

Using this in (3), the Lagrangian of the optimization problem becomes

$$L = \ln|\mathbf{I} + \mathbf{W}_1\mathbf{R}| - tr(\mathbf{W}_2\mathbf{R}) - \lambda(tr\mathbf{R} - P_T) + tr(\mathbf{M}\mathbf{R}) \tag{12}$$

where $\lambda \geq 0$ is a Lagrange multiplier responsible for the total power constraint and $\mathbf{M} \geq \mathbf{0}$ is a matrix Lagrange multiplier responsible for the constraint $\mathbf{R} \geq \mathbf{0}$. The corresponding KKT conditions are:

$$\partial L/\partial\mathbf{R} = (\mathbf{I} + \mathbf{W}_1\mathbf{R})^{-1}\mathbf{W}_1 - \mathbf{W}_2 - \lambda\mathbf{I} + \mathbf{M} = \mathbf{0} \tag{13}$$

$$\lambda(tr\mathbf{R} - P_T) = 0, \mathbf{M}\mathbf{R} = \mathbf{0} \tag{14}$$

$$\lambda \geq 0, \mathbf{M}, \mathbf{R} \geq \mathbf{0} \tag{15}$$

Since the objective is concave, the corresponding optimization problem is convex, and since Slater condition holds, the KKT conditions are sufficient for optimality [12]. After some manipulations, (13) can be transformed to

$$\widetilde{\mathbf{R}} - (\mathbf{I} - \widetilde{\mathbf{M}})^{-1} = -\widetilde{\mathbf{W}}_1^{-1} \tag{16}$$

[1]here we implicitly assume that $\mathbf{W}_\lambda$ is non-singular, e.i. either $\mathbf{W}_2$ is non-singular or $\lambda > 0$ if it is singular. If this is not the case, a pseudo-inverse should be used instead.

[2]this can be obtained via a limiting transition.

[3]including 2nd order terms in this approximation gives the weak eavesdropper condition $\lambda_i(\mathbf{W}_2\mathbf{R}) \ll 1$.

where

$$\widetilde{\mathbf{R}} = \mathbf{W}_\lambda^{1/2}\mathbf{R}\mathbf{W}_\lambda^{1/2},\ \widetilde{\mathbf{M}} = \mathbf{W}_\lambda^{-1/2}\mathbf{M}\mathbf{W}_\lambda^{-1/2},\quad \widetilde{\mathbf{W}}_1 = \mathbf{W}_\lambda^{-1/2}\mathbf{W}_1\mathbf{W}_\lambda^{-1/2} \tag{17}$$

Since $\widetilde{\mathbf{M}}\widetilde{\mathbf{R}} = \mathbf{0}$, these matrices commute and thus have the same eigenvectors, which, from (16), implies that these eigenvectors are the same as those of $\widetilde{\mathbf{W}}_1$. Hence, all three matrices can be simultaneously diagonalized and thus (16) can be transformed to diagonal form (where the diagonal entries are respective eigenvalues). From this and complementary slackness $\widetilde{\mathbf{M}}\widetilde{\mathbf{R}} = \mathbf{0}$, (7) follows, where $\widetilde{\lambda}_i = \lambda_i(\widetilde{\mathbf{R}})$. Combining this with (17), (5) and (6) follow. Lagrange multiplier λ is found from the total power constraint. The existence of the threshold power P_T^* follows from the fact that $tr\mathbf{R}^*$ 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 $tr\mathbf{R}^* = P_T^* < P_T$, i.e. only partial power is used (see Fig. 1 for illustration and discussion). (10) can be obtained by using (5) in (11). ■

Remark 1. It may appear that (7) requires $\widetilde{\mathbf{W}}_1$ and thus $\mathbf{W}_1$ be positive definite, i.e. singular case is not allowed. This is not so: $(\cdot)_+$ operator makes sure that $\widetilde{\lambda}_i = 0$ if $\lambda_i(\widetilde{\mathbf{W}}_1) = 0$ so that singular $\mathbf{W}_1$ is allowed. The same observation also applies to (9) and (16).

Remark 2. One way to ensure that the Ev is weak, i.e. $\lambda_i(\mathbf{W}_2\mathbf{R}) \ll 1$, is to require

$$\lambda_i(\mathbf{W}_2) \ll 1/P_T \tag{18}$$

from which it follows that this holds as long as the power (or SNR) is not too large, i.e. $P_T \ll 1/\lambda_i(\mathbf{W}_2)$; see also Fig. 1. It should be noted, however, that this approximation extends well beyond the low-SNR regime provided that the eavesdropper path loss is sufficiently large (i.e. $\lambda_i(\mathbf{W}_2)$ are small). For the scenario in Fig. 1, it works well up to about 10 dB and can extend to larger SNR for smaller α.

To illustrate Theorem 1 and also to see how accurate the approximation is, Fig. 1 shows the secrecy capacity obtained from the theorem for

$$\mathbf{W}_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix},\ \mathbf{W}_2 = \alpha \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \tag{19}$$

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 in Theorem 1, the approximate problem has also been solved by the MC-based approach. It is clear that the approximation is accurate in this case 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][5] and, in the case of $\mathbf{W}_1 > \mathbf{W}_2 > \mathbf{0}$, the saturation capacity is

$$C_s^* = \ln|\mathbf{W}_1| - \ln|\mathbf{W}_2| \tag{20}$$

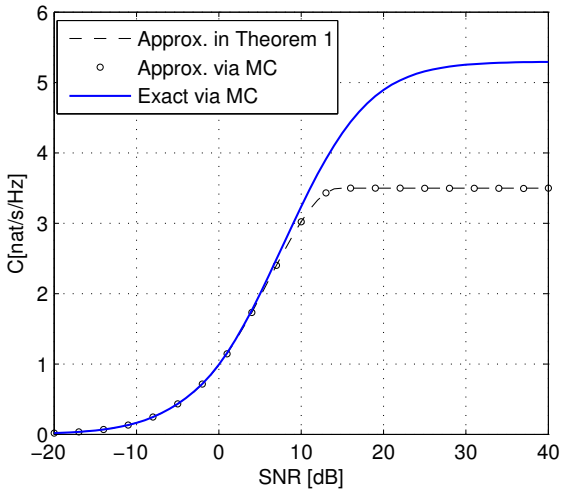

Fig. 1. Weak eavesdropper approximation and exact secrecy capacity (via MC) versus SNR. $\mathbf{W}_{1,2}$ are as in (19), $\alpha = 0.1$. The approximation is accurate if $SNR < 10$ dB. Note the capacity saturation effect at high SNR in both cases.

which follows directly from (2) by neglecting $\mathbf{I}$. In the weak eavesdropper approximation, the saturation effect is due to the fact that 2nd term in (11) is linear in P_T while 1st one is only logarithmic, so that using the full available power in 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 (10) by setting $\lambda = 0$. In the case of $\mathbf{W}_1 > \mathbf{W}_2 > \mathbf{0}$, it is given by

$$C^* = \ln|\mathbf{W}_1| - \ln|\mathbf{W}_2| - tr(\mathbf{I} - \mathbf{W}_2\mathbf{W}_1^{-1}) \tag{21}$$

By comparing (20) and (21), one concludes that the thresholds are close to each other when $tr\mathbf{W}_2\mathbf{W}_1^{-1} \approx m$.

In any case, the approximated capacity and corresponding optimal covariance in Theorem 1 provide a lower bound to the true capacity in (3) at any SNR/power and for any eavesdropper channel (weak or not):

$$C_s \geq C(\mathbf{R}^*) \tag{22}$$

which follows from $\ln(1+x) \leq x\ \forall x \geq 0$, and the bound is tight for the weak eavesdropper case.

To obtain further insight in the weak eavesdropper regime, let us consider the case when $\mathbf{W}_{1,2}$ have the same eigenvectors. This is a broader case than it may first appear as it requires $\mathbf{H}_{1,2}$ to have the same right singular vectors while leaving left ones unconstrained. In this case, the results in Theorem 1 simplify as follows.

Corollary 1. *Under the conditions of Theorem 1 and when* $\mathbf{W}_{1,2}$ *have the same eigenvectors, the optimal covariance is*

$$\mathbf{R}^* = \mathbf{U}\mathbf{\Lambda}^*\mathbf{U}^+ \tag{23}$$

where $\mathbf{U}$ *is found from the eigenvalue decompositions* $\mathbf{W}_i = \mathbf{U}\mathbf{\Lambda}_i\mathbf{U}^+$ *so that the eigenvectors of* $\mathbf{R}^*$ *are the same as those*

of $\mathbf{W}_{1,2}$*. The diagonal matrix* $\mathbf{\Lambda}^*$ *collects the eigenvalues of* $\mathbf{R}^*$*:*

$$\lambda_i(\mathbf{R}^*) = \left(\frac{1}{\lambda+\lambda_{2i}} - \frac{1}{\lambda_{1i}}\right)_+ \tag{24}$$

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

Note that the power allocation in (24) resembles that of the standard water filling, except for the λ_{2i} term. In particular, only sufficiently strong eigenmodes are active:

$$\lambda_i(\mathbf{R}^*) > 0 \text{ iff } \lambda_{1i} > \lambda + \lambda_{2i} \tag{25}$$

As P_T increases, λ decreases so that more eigenmodes become active; 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. 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:

$$\mathbf{H}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}^+ \tag{26}$$

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 Ev respectively. Therefore, when the Rx and Ev 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 [13][14]. The secrecy capacity and optimal covariance in this case can be explicitly characterized as follows.

Proposition 1. *Consider the wiretap MIMO channel as in* (1)*,* (26)*. The optimal Tx covariance for this channel takes the following form:*

$$\mathbf{R}^* = \mathbf{V}\mathbf{\Lambda}^*\mathbf{V}^+ \tag{27}$$

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

$$\lambda_i^* = \frac{\lambda_{2i}+\lambda_{1i}}{2\lambda_{2i}\lambda_{1i}}\left(\sqrt{1+\frac{4\lambda_{2i}\lambda_{1i}}{(\lambda_{2i}+\lambda_{1i})^2}\left(\frac{\lambda_{1i}-\lambda_{2i}}{\lambda}-1\right)_+}-1\right) \tag{28}$$

where $\lambda_{ki} = \sigma_{ki}^2$ *and* σ_{ki} *denotes singular values of* $\mathbf{H}_k$*;* $\lambda > 0$ *is found from the total power constraint:*

$$\sum_i \lambda_i^* = P_T \tag{29}$$

Proof: Under (26),

$$\mathbf{W}_k = \mathbf{V}\mathbf{\Lambda}_k\mathbf{V}^+ \tag{30}$$

where diagonal matrix $\mathbf{\Lambda}_k = \mathbf{\Sigma}_k^+\mathbf{\Sigma}_k$ collects eigenvalues of $\mathbf{W}_k$, so that the problem in (3) can be re-formulated as

$$C_s = \max_{tr\tilde{\mathbf{R}}\geq\mathbf{0}} \ln\frac{|\mathbf{I}+\mathbf{\Lambda}_1\tilde{\mathbf{R}}|}{|\mathbf{I}+\mathbf{\Lambda}_2\tilde{\mathbf{R}}|} \text{ s.t. } tr\tilde{\mathbf{R}} \leq P_T \tag{31}$$

where $\tilde{\mathbf{R}} = \mathbf{V}^+\mathbf{R}\mathbf{V}$. However, this is the secrecy capacity of a set of parallel Gaussian wire-tap channels as in [13][14], for which independent signaling is known to be optimal[4], so that maximizing $\tilde{\mathbf{R}}^*$ is diagonal, from which (27) follows. The optimal power allocation in (28) is essentially the same as for the equivalent parallel channels in [14]. ■

In fact, Eq. (27) says that optimal signaling is on the right SVs of $\mathbf{H}_{1,2}$ and (28) implies that only those eigenmodes are active for which

$$\sigma_{1i}^2 > \sigma_{2i}^2 + \lambda \tag{32}$$

If $\lambda_{2i} = 0$, then (28) reduces to

$$\lambda_i^* = \left(\frac{1}{\lambda} - \frac{1}{\lambda_{1i}}\right)_+ \tag{33}$$

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 out that the original channels in (26) 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 optimal covariance to be of full rank or rank-1.

Proposition 1 can be used to establish a new matrix inequality using information-theoretic tools in the spirit of [9], which we term a generalized Hadamard inequality.

Proposition 2 (generalized Hadamard inequality)**.** *Let* $\mathbf{D}_{1,2}$ *be diagonal positive semi-definite matrices, and let* $\mathbf{R}$ *be positive semi-definite. Then, the following inequality holds:*

$$\frac{|\mathbf{I}+\mathbf{D}_1\mathbf{R}|}{|\mathbf{I}+\mathbf{D}_2\mathbf{R}|} \leq \prod_{i:d_{1i}>d_{2i}} \frac{1+d_{1i}r_i}{1+d_{2i}r_i} \tag{34}$$

where d_{1i}, d_{2i}, r_i *denote diagonal entries of* $\mathbf{D}_{1,2}, \mathbf{R}$*. The equality is achieved by diagonal* $\mathbf{R}$ *with* $r_i = 0$ *if* $d_{1i} < d_{2i}$*.*

This inequality is indeed a generalization of the celebrated Hadamard determinantal inequality $|\mathbf{R}| \leq \prod_i r_i$, which (almost trivially) implies $|\mathbf{I}+\mathbf{D}_1\mathbf{R}| \leq \prod_i(1+d_{1i}r_i)$. It is far less trivial that (34) should hold as using the diagonal part of $\mathbf{R}$ maximizes the numerator but also the denominator so it's not clear what is the net result. In fact, (34) says that just retaining the diagonal part of $\mathbf{R}$ in *not* optimal: one should retain only those diagonal entries for which $d_{1i} > d_{2i}$. To the best of our knowledge, this inequality cannot be found in the

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

matrix-theoretic literature (see e.g. [15]-[17]) and furthermore, we are not aware about any matrix-theoretic way to established it.

V. When Is ZF Signaling Optimal?

In this section, we consider the case when ZF signaling is optimal, i.e. when active eigenmodes of optimal covariance $\mathbf{R}^*$ are orthogonal to those of $\mathbf{W}_2$: $\mathbf{W}_2\mathbf{R}^* = \mathbf{0}$[5]. 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 insured by well-established ZF signalling.

Proposition 3. *Consider the wire-tap MIMO channel in* (1)*. Gaussian ZF signaling is optimal, i.e.* $\mathbf{W}_2\mathbf{R}^* = \mathbf{0}$ *so that active eigenmodes of* $\mathbf{R}^*$ *are orthogonal to those of* $\mathbf{W}_2$*, iff* $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *have the same eigenvetors, so that* $\mathbf{H}_1$ *and* $\mathbf{H}_2$ *have the same right singular vectors as in* (26)*, and*

$$\lambda_{1i} \leq \lambda_{2i} + \lambda \text{ if } \lambda_{2i} > 0, \tag{35}$$

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

$$\lambda_i^* = \lambda_i(\mathbf{R}^*) = \left(\frac{1}{\lambda} - \frac{1}{\lambda_{1i}}\right)_+ \text{ if } \lambda_{2i} = 0, \tag{36}$$

and 0 otherwise. The optimal covariance is as in (27) *so that its eigenvectors are those of* $\mathbf{W}_{1,2}$.

Proof: The original problem in (3) 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 [18]. Using these conditions (see e.g. [5]), one obtains, after some manipulations, that $\mathbf{W}_1$ and $\mathbf{W}_2$ have the same eigenvectors[6] and hence the optimal covariance is as in (27). Using the optimal power allocation in (28), the condition in (35) insures $\mathbf{W}_2\mathbf{R}^* = \mathbf{0}$. Eq. (36) follows from (28) when $\lambda_{2i} = 0$. Since the necessary KKT conditions have a unique solution under the condition $\mathbf{W}_2\mathbf{R}^* = \mathbf{0}$, it is also sufficient for optimality. ∎

Remark 3. The optimal power allocation in (36) is the same as the 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. It is the smaller, the larger the set of eavesdropper positive eigenmodes is.

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 insures that the Ev 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 Ev were not there. No new wiretap codes need to be designed.

[5]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 [10].

[6]note that we do not assume here that $\mathbf{W}_1$ and $\mathbf{W}_2$ have the same eigenvectors; rather, it is a result of this proposition.

VI. Acknowledgement

The authors would like to thank M. Urlea and K. Li for running numerical experiments and generating Fig. 1.

References

[1] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.
[2] A. Khisti, G.W. Wornell, Secure Transmission With Multiple Antennas—Part I: The MISOME Wiretap Channel, IEEE Trans. Info. Theory, v. 56, No. 7, July 2010.
[3] A. Khisti, G.W. Wornell, Secure Transmission With Multiple Antennas—Part II: The MIMOME Wiretap Channel, IEEE Trans. Info. Theory, v. 56, No. 11, Nov. 2010.
[4] F. Oggier, B. Hassibi, The Secrecy Capacity of the MIMO Wiretap Channel, IEEE Trans. Info. Theory, v. 57, No. 8, Aug. 2011.
[5] S. Loyka, C.D. Charalambous, On Optimal Signaling over Secure MIMO Channels, IEEE ISIT-12, Boston, USA, July 2012.
[6] S. Loyka and C. D. Charalambous, Further Results on Optimal Signaling over Secure MIMO Channels, IEEE ISIT-13, Istanbul, Turkey, Jul. 2013.
[7] J. Li, A. Petropulu, Transmitter Optimization for Achieving Secrecy Capacity in Gaussian MIMO Wiretap Channels, arXiv:0909.2622v1, Sep 2009.
[8] J.P. Kermoal et al., A stochastic MIMO radio channel model with experimental validation, IEEE JSAC, v.20, N.6, pp. 1211-1226, Aug. 2002.
[9] T.M. Cover, J.A. Thomas, Determinantal Inequalities via Information Theory, SIAM J. Matrix Anal. Appl., v. 9, No. 3, July 1988.
[10] H.L. Van Trees, Optimum Array Processing, Wiley, New York, 2002.
[11] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, 2006.
[12] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.
[13] A. Khisti et al, Secure Broadcasting Over Fading Channels, IEEE Trans. Info. Theory, v. 54, No. 6, pp. 2453-2469, June 2008.
[14] Z. Li et al, Secrecy Capacity of Independent Parallel Channels, in R. Liu, W. Trappe (eds.), Securing Wireless Communications at the Physical Layer, Springer, 2010.
[15] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge Univ. Press, 1985.
[16] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge Univ. Press, 1991.
[17] D.S. Bernstein, Matrix Mathematics, Princeton University Press, 2009.
[18] D.P. Bertsekas, Nonlinear Programming, Athena Scientific, 2nd Ed., 2008.