5 Secrecy Rate Maximization in Gaussian MIMO Wiretap Channels

Sergey Loyka and Charalambos D. Charalambous

Secrecy rate maximization in Gaussian MIMO wiretap channels is considered. While the optimality of Gaussian signaling and a general expression for the secrecy capacity have been well established, closed-form solutions for the optimal transmit covariance matrix are known for some special cases only, while the general case remains an open problem. This chapter reviews known closed-form solutions and presents a numerical algorithm for the general case with guaranteed convergence to the global optimum. The known solutions include full-rank and rank-1 cases (which, when combined, provide a complete solution for the case of two transmit antennas), the case of identical right singular vectors for the eavesdropper and legitimate channels, and the cases of weak, isotropic, and omnidirectional eavesdroppers, which also provide lower and upper bounds to the general case. Necessary optimality conditions and a tight upper bound for the rank of the optimal covariance matrix in the general case are discussed. Sufficient and necessary conditions for the optimality of three popular signaling strategies over MIMO channels, namely, isotropic and zero-forcing signaling as well as water-filling over the legitimate channel eigenmodes, are presented. The chapter closes with a detailed description of a numerical globally convergent algorithm to solve the general case, and gives some illustrative examples.

5.1 Introduction

Due to their high spectral efficiency, wireless MIMO (multiple input, multiple output) systems are widely adopted by academia and industry. The broadcast nature of wireless channels stimulated significant interest in their security aspects and the Gaussian MIMO wiretap channel (WTC) has emerged as a popular model to study information theoretic secrecy aspects of wireless systems [1]. A number of results have been obtained for this model, including the proof of optimality of Gaussian signaling [1–4], which is far from trivial and significantly more involved than that of the regular (no wiretap) MIMO channel. Once the functional form of the optimal input is established, the only unknown is its covariance matrix since the mean is always zero. This latter part has not been solved yet in the general case; only a number of special cases have been settled.

In this chapter, we review the well-known as well as recent results on an optimal transmit covariance matrix for the MIMO WTC. Several new results will be reported as well.

The optimal transmit covariance matrix under the total power constraint has been obtained for some special cases (low/high signal to noise ratio (SNR), multiple input, single output (MISO) channels, full-rank or rank-1 solutions) [2–12], but the general case is still open. The main difficulty lies in the fact that, unlike the regular MIMO channel, the underlying optimization problem for the MIMO WTC is not convex in general, in addition to the fact that the respective Karush–Kuhn–Tucker (KKT) optimality conditions are a system of non-linear matrix equalities and inequalities. 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 [10], 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 [2, 7, 10]. The low-SNR regime has been studied in detail in [9]. An exact full-rank solution for the optimal covariance matrix has been obtained in [10] and its properties have been characterized. In particular, unlike the regular channel (no eavesdropper), the optimal power allocation does not converge to a uniform one at high SNR and the latter remains sub-optimal at any finite SNR.

Finally, while no analytical solution is known in the general case, a globally convergent numerical algorithm was proposed in [13] to find an optimal covariance for any Gaussian MIMO wiretap channel (degraded or not), and its convergence to a global optimum, which takes only a moderate or small number of steps in practice, was proved.

The rest of this chapter is organized as follows. Section 5.2 introduces the MIMO WTC model. Rank-1 and full-rank solutions are discussed in Sections 5.3 and 5.4. The weak eavesdropper case is considered in Section 5.5, which is motivated by a scenario where the Tx–Rx distance is much smaller than the Tx–Ev one. Section 5.6 discusses an isotropic eavesdropper model, whereby the Tx does not know the directional properties of the Ev and hence assumes it is isotropic. Section 5.7 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 sub-space. The case of identical right singular vectors of the Rx and Ev channels is investigated in Section 5.8. In Sections 5.9–5.11, we consider three popular signaling techniques: zero-forcing (ZF), standard water-filling (WF) over the eigenmodes of the legitimate channel, and isotropic signaling (whereby the covariance matrix is a scaled identity), and discuss sufficient and necessary conditions under which they are optimal for the MIMO WTC. These techniques are appealing for a number of reasons, including their lower complexity and existing solutions. Finally, Section 5.12 presents an algorithm for numerical evaluation of the Tx covariance matrix with guaranteed convergence to a global optimum in the general case.

Notation

$\lambda_i(\mathbf{W})$ denotes eigenvalues of a matrix $\mathbf{W}$; $(x)_+ = \max\{x, 0\}$ for a real 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 positive

eigenmodes of a Hermitian matrix $\mathbf{W}$:

$$(\mathbf{W})_+ = \sum_{i:\lambda_i(\mathbf{W})>0} \lambda_i \mathbf{u}_i \mathbf{u}_i^+, \tag{5.1}$$

where λ_i is the ith largest eigenvalue of $\mathbf{W}$ and $\mathbf{u}_i$ is its corresponding eigenvector. $\mathbf{A} > \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive definite; $|\mathbf{A}|$ is the determinant of $\mathbf{A}$, while $\mathbf{A}'$ and $\mathbf{A}^+$ are its transposition and Hermitian conjugation.

5.2 MIMO Wiretap Channel

Let us consider the standard wiretap Gaussian MIMO channel model,

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \boldsymbol{\xi}_1, \quad \mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \boldsymbol{\xi}_2, \tag{5.2}$$

where $\mathbf{x} = [x_1, x_2, \ldots x_m]' \in C^{m,1}$ is the transmitted complex-valued signal vector of dimension $m \times 1$, $\mathbf{y}_{1(2)} \in C^{n_{1(2)},1}$ are the received vectors at the receiver (eavesdropper), $\boldsymbol{\xi}_{1(2)}$ 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, and $n_{1(2)}$ and m are the numbers of Rx (eavesdropper) and Tx antennas respectively; see Fig. 5.1. The channels $\mathbf{H}_{1(2)}$ are assumed to be quasi-static (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. With slight modifications, this model can also include the case of spatially correlated noise.

The main performance indicator for a wiretap channel is its secrecy capacity, defined as follows. A secrecy rate is achievable if it satisfies the *secrecy* criterion (the information leakage to the eavesdropper approaches zero as the block length increases) in addition to the traditional *reliability* criterion (the error probability of the legitimate receiver approaches zero); see [1] for more details. The secrecy capacity is the supremum of all achievable secrecy rates, subject to the total power constraint.

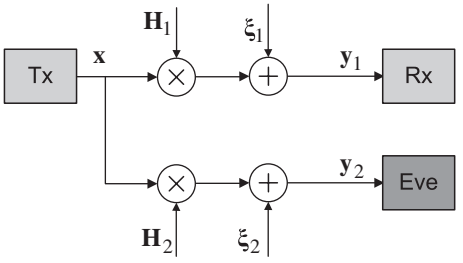

Figure 5.1 A block diagram of the Gaussian MIMO wiretap channel. Full channel state information is available at the transmitter. $\mathbf{H}_{1(2)}$ is the channel matrix to the legitimate receiver (eavesdropper); $\mathbf{x}$ is the transmitted signal and $\mathbf{y}_{1(2)}$ is the received (eavesdropper) signal; $\boldsymbol{\xi}_{1(2)}$ is the additive white Gaussian noise at the receiver (eavesdropper). The information leakage to the eavesdropper is required to approach zero asymptotically.

Gaussian signaling is known to be optimal for the Gaussian MIMO WTC (the proof of this is significantly more complicated than for the regular (no eavesdropper) Gaussian MIMO channel) [2–4], so that the only unknown part is its covariance (since the mean is always zero). For a given transmit covariance matrix $\mathbf{R} = E\{\mathbf{x}\mathbf{x}^+\}$, where $E\{\cdot\}$ is statistical expectation, the maximum achievable secrecy rate between the Tx and Rx (so that the rate between the Tx and eavesdropper is zero) is [3, 4]

$$C(\mathbf{R}) = \ln \frac{|\mathbf{I}+\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+\mathbf{W}_2\mathbf{R}|} = C_1(\mathbf{R}) - C_2(\mathbf{R}), \tag{5.3}$$

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

$$C_s = \max_{\mathbf{R}\geq 0} C(\mathbf{R}) \text{ s.t. } \mathrm{tr}\,\mathbf{R} \leq P_T, \tag{5.4}$$

where P_T is the total transmit power (also the SNR since the noise is normalized). It is well known that the problem in (5.4) is not convex in general and an explicit solution 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 cases [2–13]). In fact, the problem in (5.4) is still open even for the degraded (and hence convex) but otherwise general case of $\mathbf{W}_1 \geq \mathbf{W}_2$.

The optimization problem in (5.4) is the main subject of this chapter. The following theorem gives the necessary optimality conditions in the general case, which are instrumental for further development and allow one to established closed-form solutions for the optimal covariance in many cases.

THEOREM 5.1 *Let* $\mathbf{R}^*$ *be an optimal covariance in* (5.4),

$$\mathbf{R}^* = \arg\max_{\mathbf{R}\geq 0} C(\mathbf{R}) \;\; s.t.\; \mathrm{tr}\,\mathbf{R} \leq P_T,$$

and let $\mathbf{u}_{i+}$ *be its active eigenvector (i.e., corresponding to a positive eigenvalue). Then,*

$$\mathbf{U}_{r+}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_{r+} > \mathbf{0}, \tag{5.5}$$

where the columns of semi-unitary matrix $\mathbf{U}_{r+}$ *are the active eigenvectors* $\{\mathbf{u}_{i+}\}$*, so that* $\mathbf{x}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{x} > \mathbf{0} \;\; \forall \mathbf{x} \in \mathrm{span}\{\mathbf{u}_{i+}\}$*, i.e., a necessary condition for an optimal signaling strategy in* (5.4) *is to transmit into the positive directions of* $\mathbf{W}_1 - \mathbf{W}_2$ *(where the legitimate channel is stronger than the eavesdropper).*

Proof Based on the necessary KKT optimality conditions ; see [10] for details.

It was demonstrated in [4] that $\mathrm{rank}(\mathbf{R}^*) < m$ unless $\mathbf{W}_1 > \mathbf{W}_2$, i.e., an optimal transmission is of low rank over a non-degraded channel. The corollary below gives a more precise characterization based on the necessary optimality condition above.

COROLLARY 5.1 *Let* $\mathbf{W}_1 - \mathbf{W}_2 = \mathbf{W}_+ + \mathbf{W}_-$*, where* $\mathbf{W}_{+(-)}$ *collects positive (negative and zero) eigenmodes of* $\mathbf{W}_1 - \mathbf{W}_2$ *(found from its eigenvalue decomposition). Then,*

$$\mathrm{rank}(\mathbf{R}^*) \leq \mathrm{rank}(\mathbf{W}_+) \leq m, \tag{5.6}$$

i.e., the rank of an optimal covariance $\mathbf{R}^*$ *does not exceed the number of strictly positive eigenvalues of* $\mathbf{W}_1 - \mathbf{W}_2$ *(the rank of* $\mathbf{W}_+$*).*

5.3 Rank-1 Solution

Using Corollary 5.1, one immediately obtains an optimal covariance when $\text{rank}(\mathbf{W}_+) = 1$, e.g., for MISO or MIMO rank-1 channels.

COROLLARY 5.2 *Let* $\text{rank}(\mathbf{W}_+) = 1$. *The secrecy capacity and optimal covariance are*

$$C_\text{s} = \ln \lambda_1,\ \mathbf{R}^* = P_\text{T} \mathbf{u}_1 \mathbf{u}_1^+, \tag{5.7}$$

where λ_1, $\mathbf{u}_1$ *are the largest eigenvalue and corresponding eigenvector of* $(\mathbf{I} + P_\text{T}\mathbf{W}_2)^{-1}(\mathbf{I} + P_\text{T}\mathbf{W}_1)$ *or, equivalently, the largest generalized eigenvalue and corresponding eigenvector of* $(\mathbf{I} + P_\text{T}\mathbf{W}_1, \mathbf{I} + P_\text{T}\mathbf{W}_2)$*, so that transmit beamforming on* $\mathbf{u}_1$ *is the optimal strategy.*

Proof Corollary 5.1 ensures that $\text{rank}(\mathbf{R}^*) = 1$; the optimal covariance $\mathbf{R}^*$ in (5.7) follows in the same way as in [2].

Note that MISO channels (single-antenna channel at the receiver or eavesdropper) considered in [2, 6, 8] are special cases of this corollary with, e.g., $\mathbf{W}_1 = \mathbf{h}_1 \mathbf{h}_1^+$. The corollary allows not only MIMO channels with $\text{rank}(\mathbf{W}_1) = 1$ but also any higher-rank $\mathbf{W}_1$ and $\mathbf{W}_2$ provided that rank $(\mathbf{W}_+) = 1$.

Furthermore, the signaling in (5.7) is also optimal for any $\text{rank}(\mathbf{W}_+) \geq 1$ at sufficiently small SNR, where λ_1, $\mathbf{u}_1$ become the largest eigenvalue and corresponding eigenvector of the difference channel $\mathbf{W}_1 - \mathbf{W}_2$. The appeal of this signaling is due to its low complexity.

It should be emphasized that the solution in (5.7) is not zero-forcing (i.e., $\mathbf{W}_2 \mathbf{u}_1 \neq 0$) in general, i.e., the Tx does not form null in the Ev direction. Intuitively, doing so results in loss of power at the Rx and hence is not optimal in general. Such a solution may be optimal in some special cases; see Section 5.9.

5.4 Full-Rank Solution

The full-rank solution of the optimization problem in (5.4) is given by the following theorem.

THEOREM 5.2 *Let* $\mathbf{W}_1 > \mathbf{W}_2$ *and* $P_\text{T} > P_{\text{T}0}$*, where* $P_{\text{T}0}$ *is a threshold power given by* (5.12)*. Then,* $\mathbf{R}^*$ *is of full rank and is given by*

$$\mathbf{R}^* = \mathbf{U} \Lambda_1 \mathbf{U}^+ - \mathbf{W}_1^{-1}, \tag{5.8}$$

where the columns of the unitary matrix $\mathbf{U}$ *are the eigenvectors of* $\mathbf{Z}=\mathbf{W}_2+\mathbf{W}_2(\mathbf{W}_1-\mathbf{W}_2)^{-1}\mathbf{W}_2$, $\Lambda_1=\mathrm{diag}\{\lambda_{1i}\}>\mathbf{0}$ *is a diagonal positive-definite matrix,*

$$\lambda_{1i}=\frac{2}{\lambda}\left(\sqrt{1+\frac{4\mu_i}{\lambda}}+1\right)^{-1}, \tag{5.9}$$

and $\mu_i\geq 0$ *are the eigenvalues of* $\mathbf{Z}$; $\lambda>0$ *is found from the total power constraint* $\mathrm{tr}\,\mathbf{R}^*=P_\mathrm{T}$ *as a unique solution of the equation*

$$\frac{2}{\lambda}\sum_i\left(\sqrt{1+\frac{4\mu_i}{\lambda}}+1\right)^{-1}=P_\mathrm{T}+\mathrm{tr}\,\mathbf{W}_1^{-1}. \tag{5.10}$$

The corresponding secrecy capacity is

$$C_\mathrm{s}=\ln\frac{|\mathbf{W}_1||\Lambda_1|}{|\mathbf{I}-\mathbf{W}_2(\mathbf{W}_1^{-1}-\mathbf{U}\boldsymbol{\Lambda}_1\mathbf{U}^+)|}=\ln\frac{|\mathbf{W}_1|}{|\mathbf{W}_2|}+\ln\frac{|\Lambda_1|}{|\Lambda_2|}, \tag{5.11}$$

where $\Lambda_2=\Lambda_1+\mathrm{diag}\{\mu_i^{-1}\}$ *and the second equality holds when* $\mathbf{W}_2$ *is positive definite,* $\mathbf{W}_2>0$. $P_{\mathrm{T}0}$ *can be expressed as follows:*

$$P_{\mathrm{T}0}=\frac{2(\mu_1+\lambda_\mathrm{min})}{\lambda_\mathrm{min}^2}\sum_i\left(\sqrt{1+\frac{4\mu_i(\mu_1+\lambda_\mathrm{min})}{\lambda_\mathrm{min}^2}}+1\right)^{-1}-\mathrm{tr}\,\mathbf{W}_1^{-1}, \tag{5.12}$$

where λ_min *is the minimum eigenvalue of* $\mathbf{W}_1$ *and* μ_1 *is the maximum eigenvalue of* $\mathbf{Z}$.

Proof Based on the KKT conditions , which are sufficient for optimality in this case (since the channel is degraded and hence the problem is convex); see [10] for details.

It should be pointed out that Theorem 5.2 gives an exact (not approximate) optimal covariance at finite SNR ($P_\mathrm{T}\to\infty$ is not required) since $P_{\mathrm{T}0}$ is a finite constant that depends only on $\mathbf{W}_1$ and $\mathbf{W}_2$ and this constant is small in some cases: it follows from (5.12) that $P_{\mathrm{T}0}\to 0$ if $\lambda_\mathrm{min}\to\infty$, i.e., $P_{\mathrm{T}0}$ is small if λ_min is large. In particular, $P_{\mathrm{T}0}$ can be bounded above as

$$P_{\mathrm{T}0}\leq\frac{m\mu_1}{\lambda_\mathrm{min}^2}+\frac{m-1}{\lambda_\mathrm{min}}, \tag{5.13}$$

and if $\lambda_\mathrm{min}\gg\mu_1$, then

$$P_{\mathrm{T}0}\approx\frac{m}{\lambda_\mathrm{min}}-\mathrm{tr}\,\mathbf{W}_1^{-1}\leq\frac{m-1}{\lambda_\mathrm{min}}\leq 1, \tag{5.14}$$

where the last inequality holds if $\lambda_\mathrm{min}\geq m-1$. Figure 5.2 illustrates this case. On the other hand, when $\mathbf{W}_1-\mathbf{W}_2$ approaches a singular matrix, it follows that $P_{\mathrm{T}0}\to\infty$, so that $P_{\mathrm{T}0}$ is large iff $\mathbf{W}_1-\mathbf{W}_2$ is close to singular.

Theorem 5.2, in combination with the rank-1 solution, provides the complete solution for the optimal covariance in the $m=2$ case: if the channel is not strictly degraded or

if the SNR is not above the threshold, the rank-1 solution in (5.7) applies; otherwise, Theorem 5.2 applies. Figure 5.2 illustrates this for the following channel:

$$\mathbf{W}_1 = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}, \quad \mathbf{W}_2 = \begin{bmatrix} 0.35 & 0.15 \\ 0.15 & 0.35 \end{bmatrix}. \tag{5.15}$$

Note that the transition to full-rank covariance takes place at low SNR of about -6 dB, i.e., $P_{\mathrm{T}0}$ is not high at all in this case.

We further observe that the first term in (5.11), $C_\infty = \ln \frac{|\mathbf{W}_1|}{|\mathbf{W}_2|}$, is independent of SNR, and the second one, $\Delta C = \ln \frac{|\Lambda_1|}{|\Lambda_2|} < 0$, monotonically increases with the SNR. Furthermore, $C_{\mathrm{s}} \to C_\infty$, $\Delta C \to 0$ as $P_{\mathrm{T}} \to \infty$, in agreement with Theorem 2 in [3]. This is also clear from Fig. 5.2.

Note also that the second term in (5.8) de-emphasizes weak eigenmodes of $\mathbf{W}_1$. Since λ is monotonically decreasing as P_{T} increases [this follows from (5.10)], λ_{1i} monotonically increases with P_{T}, and approaches $\lambda_{1i} \approx 1/\sqrt{\mu_i \lambda}$, $i = 1, \ldots, m$, at sufficiently high SNR, which is in contrast with the conventional WF solution, where the uniform power allocation is optimal at high SNR. Furthermore, it follows from (5.9) that λ_{1i} decreases with μ_i, i.e., stronger eigenmodes of $\mathbf{W}_2^{-1} - \mathbf{W}_1^{-1} = \mathbf{Z}^{-1}$ (which correspond to larger eigenmodes of $\mathbf{W}_1$ and weaker ones of $\mathbf{W}_2$) get allocated more power, which follows the same tendency as the conventional WF. It further follows from (5.8) that when $\mathbf{W}_1$ and $\mathbf{W}_2$ have the same eigenvectors, $\mathbf{R}^*$ also has the same eigenvectors, i.e., the optimal signaling is on the eigenvectors of $\mathbf{W}_{1(2)}$.

The case of singular $\mathbf{W}_1$ can also be included by observing that, under certain conditions, $\mathbf{R}^*$ puts no power on the null space of $\mathbf{W}_1$ so that all matrices can be projected, without loss of generality, on the positive eigenspace of $\mathbf{W}_1$ and Theorem 5.2 will apply to the projected channel.

It is instructive to consider the case when the required channel is much stronger than the eavesdropper one, $\mathbf{W}_1 \gg \mathbf{W}_2$, meaning that all eigenvalues of $\mathbf{W}_1$ are much larger than those of $\mathbf{W}_2$.

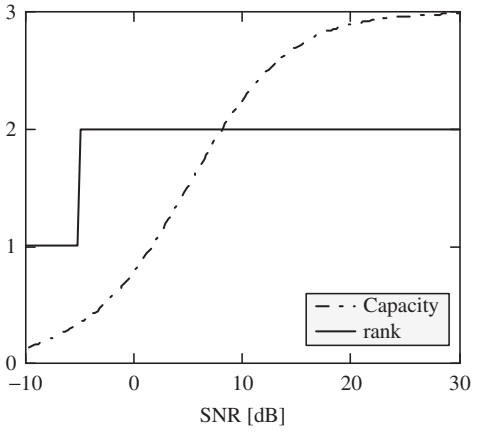

Figure 5.2 Secrecy capacity and the rank of $\mathbf{R}^*$ vs. SNR [dB] for the channel in (5.15). The transition to full-rank covariance takes place at about -6 dB.

COROLLARY 5.3 *Consider the MIMO WTC in (5.2) under the conditions of Theorem 5.2 and when the eavesdropper channel is much weaker than the required one,*

$$\lambda_1(\mathbf{W}_2) \ll m(P_\mathrm{T} + \mathrm{tr}\,\mathbf{W}_1^{-1})^{-1}/4 \le m/(4P_\mathrm{T}), \tag{5.16}$$

where $\lambda_i(\mathbf{W}_2)$ is ith largest eigenvalue of $\mathbf{W}_2$, e.g., when $\mathbf{W}_2 \to \mathbf{0}$ with fixed $\mathbf{W}_1$. Then the optimal covariance in (5.8) becomes

$$\mathbf{R}^* \approx (\lambda^{-1}\mathbf{I} - \mathbf{W}_1^{-1}) - \lambda^{-2}\mathbf{W}_2. \tag{5.17}$$

Proof See [10].

The approximation in (5.17) should be understood in Frobenius or any other norm (since all norms are equivalent). An interpretation of (5.17) is immediate: the first term

$$\mathbf{R}_\mathrm{WF} = \lambda^{-1}\mathbf{I} - \mathbf{W}_1^{-1} \tag{5.18}$$

is the standard water-filling on the eigenmodes of $\mathbf{W}_1$ (which is the capacity-achieving strategy for the regular MIMO channel) and the second term is a correction due to the secrecy requirement: those modes that spill over into the eavesdropper channel get less power to accommodate the secrecy constraint.

Let us now consider the high-SNR regime.

COROLLARY 5.4 *When $\mathbf{W}_2 > 0$, the optimal covariance $\mathbf{R}^*$ in (5.8) in the high-SNR regime*

$$P_\mathrm{T} \gg \mu_m^{-1/2} \sum\nolimits_i \mu_i^{-1/2} \tag{5.19}$$

(e.g., when $P_\mathrm{T} \to \infty$), where $\mu_m = \min_i \mu_i$, simplifies to

$$\mathbf{R}^* \approx \mathbf{U}\mathrm{diag}\{d_i\}\mathbf{U}^+, \quad d_i = \frac{P_\mathrm{T}\mu_i^{-1/2}}{\sum_i \mu_i^{-1/2}}. \tag{5.20}$$

The corresponding secrecy capacity is

$$C_\mathrm{s} \approx \ln\frac{|\mathbf{W}_1|}{|\mathbf{W}_2|} - \frac{1}{P_\mathrm{T}}\left(\sum\nolimits_i \frac{1}{\sqrt{\mu_i}}\right)^2, \tag{5.21}$$

where we have neglected the second- and higher-order effects in $1/P_\mathrm{T}$.

Proof Follows from Theorem 5.2 along the same lines as that of Corollary 5.3.

Note that the optimal signaling is on the eigenmodes of $\mathbf{W}_2^{-1} - \mathbf{W}_1^{-1}$ with the optimal power allocation given by $\{d_i\}$. This somewhat resembles the conventional water-filling, but also has a remarkable difference: unlike the conventional WF, the secure WF in (5.20) does not converge to the uniform allocation in the high-SNR regime.[1] However, strong eigenmodes of $\mathbf{W}_2^{-1} - \mathbf{W}_1^{-1}$ (which correspond to weak modes of $\mathbf{W}_2$ and strong ones of $\mathbf{W}_1$) do get more power, albeit in a form different from that of the conventional WF.

[1] The sub-optimality of the isotropic signaling suggested in Theorem 2 of [3] is hidden in the $o(1)$ term. The second term of Eq. (5.21) above refines that $o(1)$ term.

5.5 Weak Eavesdropper

Motivated by a scenario where the legitimate receiver is closer to the transmitter than the eavesdropper so that its path loss is large, see, e.g., Fig. 5.4, the case of a weak eavesdropper is considered in this section. There are no additional assumptions here (e.g., for the channel to be degraded, etc.). The weak Ev case 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, 10].

THEOREM 5.3 *Consider the problem in* (5.4) *when the eavesdropper is weak,* $\lambda_i(\mathbf{W}_2\mathbf{R}) \ll 1$*, e.g., when* $\lambda_1(\mathbf{W}_2) \ll 1/P_{\mathrm{T}}$*. The optimal covariance is given by*

$$\mathbf{R}^* \approx \mathbf{W}_\lambda^{-1/2}(\lambda^{-1}\mathbf{I} - \widehat{\mathbf{W}}_1^{-1})_+\mathbf{W}_\lambda^{-1/2}, \tag{5.22}$$

where $\mathbf{W}_\lambda = \mathbf{I} + \lambda^{-1}\mathbf{W}_2$, $\widehat{\mathbf{W}}_1 = \mathbf{W}_\lambda^{-1/2}\mathbf{W}_1\mathbf{W}_\lambda^{-1/2}$, $\lambda \geq 0$ *is found from the total power constraint,*[2]

$$\operatorname{tr}\mathbf{R}^* = P_{\mathrm{T}} \text{ if } P_{\mathrm{T}} < P_{\mathrm{T}}^*, \tag{5.23}$$

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

$$P_{\mathrm{T}}^* = \operatorname{tr}\mathbf{W}_2^{-1}(\mathbf{I} - \mathbf{W}_2^{1/2}\mathbf{W}_1^{-1}\mathbf{W}_2^{1/2})_+ \tag{5.24}$$

if $\mathbf{W}_2$ *is non-singular;* $P_{\mathrm{T}}^* = \infty$ *if* $\mathbf{W}_2$ *is singular and* $\mathcal{N}(\mathbf{W}_2) \not\subseteq \mathcal{N}(\mathbf{W}_1)$*. The corresponding secrecy capacity is*

$$C_{\mathrm{s}} \approx \sum_{i:\widehat{\lambda}_{1i}>\lambda} \ln(\widehat{\lambda}_{1i}/\lambda) - \operatorname{tr}\widehat{\mathbf{W}}_2(\mathbf{I} - \widehat{\mathbf{W}}_1^{-1})_+ \tag{5.25}$$

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

Proof The proof is based on the weak eavesdropper approximation

$$C(\mathbf{R}) \approx \ln|\mathbf{I} + \mathbf{W}_1\mathbf{R}| - \operatorname{tr}(\mathbf{W}_2\mathbf{R}), \tag{5.26}$$

which holds if $\lambda_i(\mathbf{W}_2\mathbf{R}) \ll 1$, and on the respective KKT optimality conditions. See [12] for details.

REMARK 5.1 *It may appear that* (5.22) *requires* $\widehat{\mathbf{W}}_1$ *and thus* $\mathbf{W}_1$ *to be positive definite, i.e., the singular case is not allowed. This is not so: the* $(\cdot)_+$ *operator makes sure that zero eigenmodes of* $\widehat{\mathbf{W}}_1$ *are eliminated so that singular* $\mathbf{W}_1$ *is allowed. The same observation also applies to* (5.24) *and* (5.25)*.*

REMARK 5.2 *One way to ensure that the Ev is weak, i.e.,* $\lambda_i(\mathbf{W}_2\mathbf{R}) \ll 1$*, is to require*

$$\lambda_1(\mathbf{W}_2) \ll 1/P_{\mathrm{T}} \tag{5.27}$$

[2] Here we implicitly assume that $\mathbf{W}_\lambda$ is non-singular, i.e., 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.

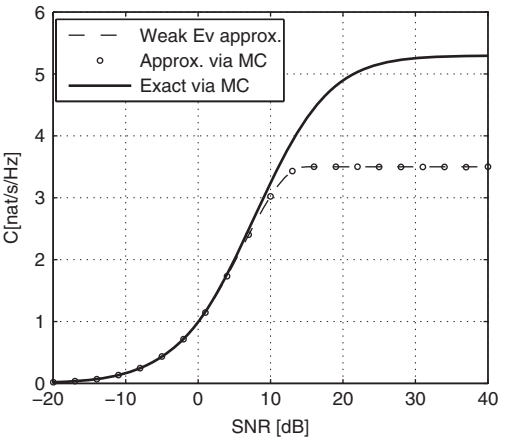

Figure 5.3 Weak eavesdropper approximation in (5.25) and exact secrecy capacity (via MC) versus SNR. $\mathbf{W}_{1,2}$ are as in (5.28), $\alpha = 0.1$. The approximation is accurate if SNR < 10 dB. Note the capacity saturation effect at high SNR in both cases.

[since $\lambda_i(\mathbf{W}_2\mathbf{R}) \le \lambda_i(\mathbf{W}_2)\lambda_1(\mathbf{R}) \le P_T\lambda_1(\mathbf{W}_2)$], 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. 5.3. 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., $\lambda_1(\mathbf{W}_2)$ is small). For the scenario in Fig. 5.3, it works well up to about 10 dB and can extend to larger SNR for smaller α.

REMARK 5.3 *When the optimal covariance in* (5.22) *is full rank , it takes on the same form as in* (5.17)*, thus revealing similarity with the standard water-filling over the channel eigenmodes in* (5.18)*.*

To illustrate Theorem 5.3, and also to see how accurate the approximation is, Fig. 5.3 shows the secrecy capacity obtained from the theorem for

$$\mathbf{W}_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{W}_2 = \alpha \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}. \tag{5.28}$$

In addition, 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 5.3, 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 already been observed in [3, 10], and, in the case of $\mathbf{W}_1 > \mathbf{W}_2 > \mathbf{0}$,

the saturation capacity is

$$C_{\mathrm{s}}^{*} = \ln|\mathbf{W}_1| - \ln|\mathbf{W}_2|, \tag{5.29}$$

which follows directly from (5.3) by neglecting $\mathbf{I}$. In the weak eavesdropper approximation, the saturation effect is due to the fact that the second term in (5.26) is linear in P_{T} while the first is only logarithmic. So using the full available power is not optimal when that power is sufficiently high. Roughly speaking, the approximation is accurate before it reaches the saturation point, i.e., for $P_{\mathrm{T}} < P_{\mathrm{T}}^{*}$. The respective saturation capacity is obtained from (5.25) by setting $\lambda = 0$. In the case of $\mathbf{W}_1 > \mathbf{W}_2 > \mathbf{0}$, it is given by

$$C^{*} = \ln|\mathbf{W}_1| - \ln|\mathbf{W}_2| - \mathrm{tr}(\mathbf{I} - \mathbf{W}_2\mathbf{W}_1^{-1}). \tag{5.30}$$

By comparing (5.29) and (5.30), one concludes that the thresholds are close to each other when $\mathrm{tr}\,\mathbf{W}_2\mathbf{W}_1^{-1} \approx m$.

REMARK 5.4 *In the general case, the approximated capacity and corresponding optimal covariance in Theorem 5.3 provide a lower bound to the true secrecy capacity in* (5.4) *at any SNR/power and for any eavesdropper channel (weak or not):*

$$C_{\mathrm{s}} \geq C(\mathbf{R}^{*}), \tag{5.31}$$

where $C(\mathbf{R})$ *is as in* (5.26)*, which follows from* $\ln(1+x) \leq x\ \forall x \geq 0$*. The sub-optimality gap can be bounded above as follows:*

$$0 \leq C_{\mathrm{s}} - C(\mathbf{R}^{*}) \leq \frac{P_{\mathrm{T}}^{2}}{2}\lambda_1^{2}(\mathbf{W}_2), \tag{5.32}$$

so that the bound is tight for a weak eavesdropper or/and low SNR .

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 the left ones unconstrained (see Section 5.8 for more details on this scenario). In this case, the results of Theorem 5.3 simplify as follows.

COROLLARY 5.5 *Let* $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *have the same eigenvectors. Then, under the conditions of Theorem 5.3, the optimal covariance is*

$$\mathbf{R}^{*} = \mathbf{U}\boldsymbol{\Lambda}^{*}\mathbf{U}^{+}, \tag{5.33}$$

where $\mathbf{U}$ *is found from the eigenvalue decompositions* $\mathbf{W}_i = \mathbf{U}\boldsymbol{\Lambda}_i\mathbf{U}^{+}$ *so that the eigenvectors of* $\mathbf{R}^{*}$ *are the same as those of* $\mathbf{W}_{1,2}$*. The diagonal matrix* $\boldsymbol{\Lambda}^{*}$ *collects the eigenvalues of* $\mathbf{R}^{*}$*:*

$$\lambda_i(\mathbf{R}^{*}) = \left(\frac{1}{\lambda + \lambda_{2i}} - \frac{1}{\lambda_{1i}}\right)_{+}, \tag{5.34}$$

where λ_{ki} *is ith eigenvalue of* $\mathbf{W}_k$*.*

Note that the power allocation in (5.34) resembles the standard water-filling solution, except for the λ_{2i} term. In particular, only sufficiently strong eigenmodes are active:

$$\lambda_i(\mathbf{R}^*) > 0 \text{ iff } \lambda_{1i} > \lambda + \lambda_{2i}. \tag{5.35}$$

As P_T increases, λ decreases so that more eigenmodes become active; legitimate channel eigenmodes are active provided that they are stronger than 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.

5.6 Isotropic Eavesdropper and Capacity Bounds

The model above requires 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 [14–16]. In this section, an alternative approach is considered in which the eavesdropper is characterized by a channel again assumed to be identical in all directions. We term this an "isotropic eavesdropper." This minimizes the amount of CSI to be available at the transmitter (one scalar parameter and no directional properties). Based on this, lower and upper (tight) capacity bounds are given for the general case, which are achievable by an isotropic eavesdropper.

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. 5.4. 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 the transmit direction $\mathbf{x}$ is $\mathbf{x}^+\mathbf{W}_2\mathbf{x} = |\mathbf{H}_2\mathbf{x}|^2$ (assuming $|\mathbf{x}| = 1$), and since $\max_{|\mathbf{x}|=1} \mathbf{x}^+\mathbf{W}_2\mathbf{x} = \epsilon_1$ (from the variational characterization of eigenvalues [17]), 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.

In combination with the matrix monotonicity of the log-det function, the latter inequality ensures that $\epsilon_1\mathbf{I}$ is the worst possible $\mathbf{W}_2$ that attains the capacity lower bound in (5.39), 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. 5.4, the eavesdropper channel matrix $\mathbf{H}_2$ can be presented in the form

$$\mathbf{H}_2 = \sqrt{\frac{\alpha}{R_2^\nu}}\widetilde{\mathbf{H}}_2, \tag{5.36}$$

where α/R_2^ν 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.) [18], and $\widetilde{\mathbf{H}}_2$ is a properly normalized channel matrix (includes local scattering/multipath effects but excludes the average path loss)

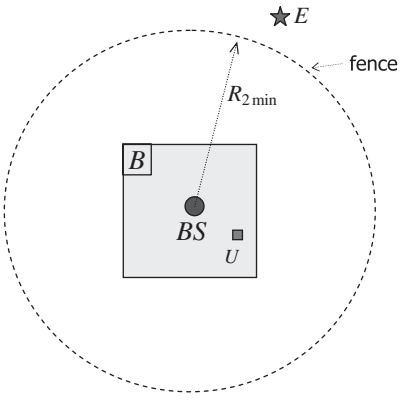

Figure 5.4 Physical scenario for a secret communication system: base station BS (the transmitter) is located on the rooftop of a secure building B, legitimate user U (the receiver) is inside the building B, and eavesdropper E is beyond the fence so that $R_2 \geq R_{2\min}$.

so that $\operatorname{tr} \widetilde{\mathbf{H}}_2^+ \widetilde{\mathbf{H}}_2 \leq n_2 m$ [19]. With this in mind, one obtains

$$\mathbf{W}_2 = \mathbf{H}_2^+ \mathbf{H}_2 = \frac{\alpha}{R_2^\nu} \widetilde{\mathbf{H}}_2^+ \widetilde{\mathbf{H}}_2 \leq \frac{\alpha}{R_{2\min}^\nu} \widetilde{\mathbf{H}}_2^+ \widetilde{\mathbf{H}}_2 \leq \frac{\alpha n_2 m}{R_{2\min}^\nu} \mathbf{I}, \tag{5.37}$$

so that one can choose $\epsilon_1 = \alpha n_2 m R_{2\min}^{-\nu}$ 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 practically relevant than the full Tx CSI model.

The isotropic eavesdropper model is closely related to the parallel channel setting in [20,21]: even though the original channel is not parallel, it can be transformed (via an information-preserving transformation) into a parallel channel, for which independent signaling is known to be optimal [20,21]. 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 [21]. These properties in combination with the bounds in (5.38) 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 given above to construct lower and upper capacity bounds for the general case using the standard matrix inequalities

$$\epsilon_m \mathbf{I} \leq \mathbf{W}_2 \leq \epsilon_1 \mathbf{I}, \tag{5.38}$$

where $\epsilon_i = \lambda_i(\mathbf{W}_2)$ denotes the ith 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 5.1 *The MIMO WTC secrecy capacity in* (5.4) *is bounded as follows:*

$$C^*(\epsilon_1) \leq C_s \leq C^*(\epsilon_m), \tag{5.39}$$

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

$$C^*(\epsilon) = \max_{\mathbf{R}\geq 0,\ \mathrm{tr}\mathbf{R}\leq P_T} \ln \frac{|\mathbf{I}+\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+\epsilon\mathbf{R}|} = \sum_i \ln \frac{1+g_i\lambda_i^*}{1+\epsilon\lambda_i^*}, \tag{5.40}$$

$g_i = \lambda_i(\mathbf{W}_1)$*, and* $\lambda_i^* = \lambda_i(\mathbf{R}^*)$ *are the eigenvalues of the optimal transmit covariance* $\mathbf{R}^* = \mathbf{U}_1\Lambda^*\mathbf{U}_1^+$*,*

$$\lambda_i^* = \frac{\epsilon+g_i}{2\epsilon g_i}\left(\sqrt{1+\frac{4\epsilon g_i}{(\epsilon+g_i)^2}\left(\frac{g_i-\epsilon}{\lambda}-1\right)_+}-1\right), \tag{5.41}$$

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

The bounds gap in (5.39) *can be bounded above as*

$$\Delta C = C^*(\epsilon_m) - C^*(\epsilon_1) \leq m_+ \ln \frac{1+\epsilon_1 P_T/m_+}{1+\epsilon_m P_T/m_+} \leq m_+ \ln \frac{\epsilon_1}{\epsilon_m}, \tag{5.42}$$

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

Proof Use the matrix monotonicity of the log-det function and a unitary transformation to put this model into the parallel channel setting; see [11] for details.

Thus, the optimal signaling 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 conventional water-filling. The latter is further elaborated for the high and low SNR regimes below. 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) but only on its eigenvalues, so that the optimal signaling problem here separates into two independent parts: (1) optimal signaling directions are selected as the eigenvectors of $\mathbf{W}_1$, and (2) 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 (5.39) coincide when $\epsilon_1 = \epsilon_m$, thus giving the secrecy capacity of the isotropic eavesdropper. Furthermore, as follows from (5.42), 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 secrecy capacity, see Fig. 5.5. Referring to Fig. 5.4, one can also set $\epsilon_1 = \alpha n_2 m R_{2\min}^{-\nu}$ and proceed with a conservative system design to achieve secrecy rate $C^*(\epsilon_1)$. Note that this design requires only the knowledge of n_2 and $R_{2\min}$ at the transmitter instead of the 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 too large, as follows from (5.42). It can be further shown that $C^*(\epsilon_1)$ is the compound secrecy capacity for the class of eavesdroppers with bounded spectral norm (maximum channel gain), $\mathbf{W}_2 \leq \epsilon_1\mathbf{I}$, and that signaling on the worst-case channel ($\mathbf{W}_2 = \epsilon_1\mathbf{I}$) achieves the capacity [16].

We note that the power allocation in (5.41) has properties similar to those of the conventional water-filling, which are as follows.

PROPOSITION 5.2 *Properties of the optimum power allocation:*

1. λ_i^* *is an increasing function of* g_i *(strictly increasing unless* $\lambda_i^* = 0$ *or* $P_{\mathrm{T}} = 0$*) , i.e., stronger eigenmodes get allocated 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* $\lambda_1^* = P_{\mathrm{T}}$ *as* $P_{\mathrm{T}} \to 0$ *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* $P_{\mathrm{T}} \to \infty$*, 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* $P_{\mathrm{T}} \to \infty$ *and* $\lambda \to g_1 - \epsilon$ *as* $P_{\mathrm{T}} \to 0$*.*
5. *There are* m_+ *active eigenmodes if the following inequalities hold:*

$$P_{m_+} < P_{\mathrm{T}} \le P_{m_+ + 1}, \tag{5.43}$$

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

$$P_{m_+} = \sum_{i=1}^{m_+ - 1} \frac{\epsilon + g_i}{2\epsilon g_i} \left(\sqrt{1 + \frac{4\epsilon g_i}{(\epsilon + g_i)^2} \frac{g_i - g_{m_+}}{(g_{m_+} - \epsilon)_+}} - 1 \right), m_+ = 2, \ldots, m, \tag{5.44}$$

and $P_1 = 0$*, so that* m_+ *is an increasing function of* P_{T}*.*

It follows from Proposition 5.2 that there is only one active eigenmode, i.e., beamforming is optimal, if $g_2 > \epsilon$ and

$$P_{\mathrm{T}} \le \frac{\epsilon + g_1}{2\epsilon g_1} \left(\sqrt{1 + \frac{4\epsilon g_1}{(\epsilon + g_1)^2} \frac{g_1 - g_2}{g_2 - \epsilon}} - 1 \right). \tag{5.45}$$

For example, this holds 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., the Newton–Raphson method) will do so efficiently.

As a side benefit of Proposition 5.2, one can use (5.43) as a condition for having m_+ active eigenmodes under the regular eigenmode transmission (no eavesdropper) with standard water-filling by taking $\epsilon \to 0$ in (5.44),

$$P_{m_+} = \sum_{i=1}^{m_+ - 1} \left(\frac{1}{g_{m_+}} - \frac{1}{g_i} \right), \tag{5.46}$$

and (5.46) approximates (5.44) when the eavesdropper is weak, $\epsilon \ll g_{m+}$. To the best of our knowledge, the expression (5.46) for the threshold powers of the standard water-filling has not previously appeared in the literature.

5.6.1 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, (5.40) simplifies to

$$C_\infty^* = \sum_{i_+} \ln \frac{g_i}{\epsilon}, \tag{5.47}$$

where the summation is over all active eigenmodes, $i_+ = \{i : g_i > \epsilon\}$, so that the secrecy 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 $P_\mathrm{T} \to \infty$, so that (5.41) becomes

$$\lambda_i^* = \frac{P_\mathrm{T}\sqrt{\epsilon^{-1} - g_i^{-1}}}{\sum_{i_+}\sqrt{\epsilon^{-1} - g_i^{-1}}}(1 + o(1)), \tag{5.48}$$

where

$$\sqrt{\lambda} = \frac{1}{P_\mathrm{T}} \sum_{i_+} \sqrt{\epsilon^{-1} - g_i^{-1}}(1 + o(1)) \tag{5.49}$$

from the total power constraint. Using (5.48), the secrecy capacity becomes

$$C^*(\epsilon) = \sum_{i_+} \ln \frac{g_i}{\epsilon} - \frac{1}{P_\mathrm{T}} \left(\sum_{i_+} \sqrt{\frac{1}{\epsilon} - \frac{1}{g_i}} \right)^2 + o\left(\frac{1}{P_\mathrm{T}}\right), \tag{5.50}$$

which is a refinement of (5.47). The saturation takes place when the second term is much smaller than the first one, so that

$$P_\mathrm{T} \gg \frac{\sum_{i_+}\sqrt{\epsilon^{-1} - g_i^{-1}}}{\sum_{i_+} \ln \frac{g_i}{\epsilon}} \tag{5.51}$$

and $C^*(\epsilon) \approx C_\infty^*$ under this condition. This effect is illustrated in Fig. 5.5.

Note that, from (5.48), the optimal power allocation behaves almost like water-filling in this case, due to the $\sqrt{\epsilon^{-1} - g_i^{-1}}$ term.

Using (5.47), the gap ΔC_∞^* between the lower and upper bounds in (5.39) becomes

$$\Delta C_\infty^* = C_\infty^*(\epsilon_m) - C_\infty^*(\epsilon_1) = m_1 \ln \frac{\epsilon_1}{\epsilon_m} + \sum_{i=m_1+1}^{m_2} \ln \frac{g_i}{\epsilon_m}, \tag{5.52}$$

where $m_{1(2)}$ is the number of active eigenmodes when $\epsilon = \epsilon_{1(m)}$. Note that this gap is SNR independent, and if $m_1 = m_2 = m_+$, which is the case if $g_{m+} > \epsilon_1$, then

$$\Delta C_\infty^* = m_+ \ln \frac{\epsilon_1}{\epsilon_m}, \tag{5.53}$$

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 (5.42) are tight.

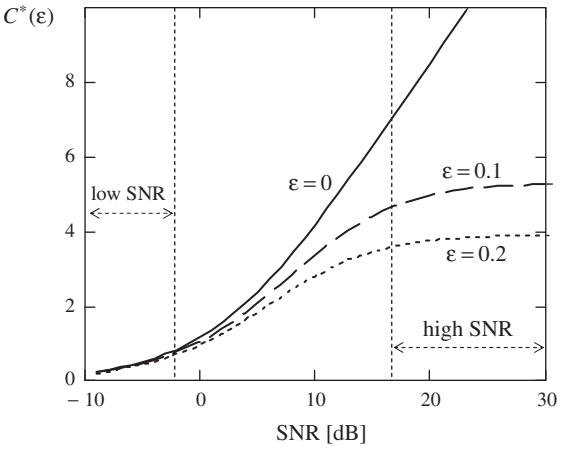

Figure 5.5 Secrecy capacity for the isotropic eavesdropper and the capacity of the regular MIMO channel (no eavesdropper, $\epsilon = 0$) vs. the SNR ($= P_{\mathrm{T}}$ since the noise variance is unity); $g_1 = 2$, $g_2 = 1$. Note the saturation effect at high SNR, where the capacity strongly depends on ϵ but not on the SNR, and the negligible impact of the eavesdropper at low SNR.

5.6.2 When Is the Eavesdropper Negligible?

It is clear from (5.40) that under fixed $\{g_i\}$, 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 is negligible since one can always select sufficiently high P_{T} to make the saturation effect important (see Fig. 5.5). To answer the question in the section's title, we use (5.40) to obtain:

$$\begin{aligned} C^*(\epsilon) &= \max_{\{\lambda_i\}} \sum_i \ln\left(1 + \frac{1 + (g_i - \epsilon)\lambda_i}{1 + \epsilon\lambda_i}\right) \text{ s.t. } \lambda_i \geq 0, \sum_i \lambda_i = P_{\mathrm{T}} \\ &\overset{(a)}{\approx} \max_{\{\lambda_i\}} \sum_i \ln(1 + (g_i - \epsilon)\lambda_i) \\ &\overset{(b)}{\approx} \max_{\{\lambda_i\}} \sum_i \ln(1 + g_i\lambda_i) = C^*(0), \end{aligned} \tag{5.54}$$

where (a) holds if

$$P_{\mathrm{T}} \ll 1/\epsilon \tag{5.55}$$

(since $\lambda_i \leq P_{\mathrm{T}}$), i.e., if the SNR is not too large, and (b) holds if

$$\epsilon \ll g_{i+}, \tag{5.56}$$

where i_+ is the set of active eigenmodes, i.e., if the eavesdropper eigenmodes are much weaker than the active eigenmodes of the legitimate channel. It is the combination of (5.55) and (5.56) that ensures that the eavesdropper is negligible. Neither condition alone is able to do so. Figure 5.5 illustrates this point. Equation (5.54) also indicates that the impact of the eavesdropper is the per-eigenmode gain loss of ϵ. Unlike the high

SNR regime in (5.47), where the loss is multiplicative (i.e., very significant and never negligible), here it is additive (mild or negligible in many cases).

5.6.3 Low SNR Regime

Let us now consider the low SNR regime, which is characteristic for CDMA-type systems [22]. Traditionally, this regime is defined via $P_T \to 0$. We, however, use a more relaxed definition requiring that $m_+ = 1$, which holds under (5.45). In this regime, assuming $g_1 > \epsilon$,

$$C^*(\epsilon) = \ln \frac{1+g_1 P_T}{1+\epsilon P_T} = \ln\left(1 + \frac{(g_1-\epsilon)P_T}{1+\epsilon P_T}\right) \overset{(a)}{\approx} \ln(1+(g_1-\epsilon)P_T), \qquad (5.57)$$

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 (5.47), where this impact is never negligible. Figure 5.5 illustrates this difference.

Note further from (5.57) that the difference between the lower and upper bounds in (5.39) is the SNR gap of $(\epsilon_1 - \epsilon_m)P_T$, which is negligible if $g_1 \gg \epsilon_1 - \epsilon_m$. This may be the case even if the condition number ϵ_1/ϵ_m is large. 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.

When $g_1 - \epsilon \ll P_T$, (a) in (5.57) gives $C^*(\epsilon) \approx (g_1 - \epsilon)P_T$, which is linear in P_T. A similar capacity scaling at low SNR has been obtained in [23] for an 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 first two equalities in (5.57) do not require $P_T \to 0$ but only to satisfy (5.45).

5.7 Omnidirectional Eavesdropper

In this section, we consider a scenario where the eavesdropper has equal gain in all directions of a certain sub-space. This model accounts for two points: (1) 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; (2) 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.[3]

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

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

$$|\mathbf{H}_2\mathbf{x}|^2 = \mathbf{x}^+\mathbf{W}_2\mathbf{x} = \text{const.}\ \forall \mathbf{x} \in \mathcal{N}(\mathbf{W}_2)^\perp, \tag{5.58}$$

where $\mathcal{N}(\mathbf{W}_2)^\perp$ is the sub-space orthogonal to the nullspace $\mathcal{N}(\mathbf{W}_2)$ of $\mathbf{W}_2$, i.e., its active sub-space, whose dimensionality is $r_2 = \text{rank}(\mathbf{W}_2)$. In particular, when the eavesdropper is isotropic, $\mathcal{N}(\mathbf{W}_2)$ is an empty set so that $\mathcal{N}(\mathbf{W}_2)^\perp$ is the entire space and $r_2 = m$. The condition in (5.58) implies that

$$\mathbf{W}_2 = \varepsilon \mathbf{U}_{2+}\mathbf{U}_{2+}^+, \tag{5.59}$$

where $\mathbf{U}_{2+}$ is a semi-unitary matrix that collects active eigenvectors of $\mathbf{W}_2$, and $\mathcal{N}(\mathbf{W}_2)^\perp = \text{span}\{\mathbf{U}_{2+}\}$. Note that the model in (5.59) allows $\mathbf{W}_2$ to be rank-deficient: $r_2 < m$ is allowed. ε can be evaluated from, e.g., (5.37): $\varepsilon = \alpha n_2 m R_{2\min}^{-\nu}$.

THEOREM 5.4 *Consider the omnidirectional eavesdropper in* (5.58)*,* (5.59) *and let* $\mathcal{R}(\mathbf{W}_1) \subseteq \mathcal{R}(\mathbf{W}_2)$. *Then the MIMO WTC secrecy capacity can be expressed as*

$$C_\text{s} = \max_{\text{tr}\mathbf{R} \le P_\text{T}} \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \mathbf{W}_2\mathbf{R}|} = \max_{\text{tr}\mathbf{R} \le P_\text{T}} \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \epsilon\mathbf{R}|} = C^*(\epsilon). \tag{5.60}$$

Proof See Appendix.

Note that the secrecy capacity as well as the optimal signaling for an omnidirectional eavesdropper in Theorem 5.4 is the same as for the isotropic one in Proposition 5.1, 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 (which is not the case in general, as can be shown via examples).

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

$$|\mathbf{H}\mathbf{x}|^2 = \mathbf{x}^+\mathbf{W}\mathbf{x} \neq 0\ \forall \mathbf{x} \in \mathcal{R}(\mathbf{W}); \tag{5.61}$$

the condition $\mathcal{R}(\mathbf{W}_1) \in \mathcal{R}(\mathbf{W}_2)$ 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 $|\mathbf{H}_1\mathbf{x}| \neq 0$ implies $|\mathbf{H}_2\mathbf{x}| \neq 0$, 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 (5.58) does not require $\mathbf{U}_2 = \mathbf{U}_1$, i.e., the eigenvectors of the legitimate channel and of the eavesdropper can be different.

5.8 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 form

$$\mathbf{H}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}^+, \tag{5.62}$$

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 are determined 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 the different left SVs) while the same scattering environment around the Tx induces the same right SVs. This is similar to the popular Kronecker MIMO channel correlation model [24], where the overall channel correlation is a product of the independent Tx and Rx parts, which are induced by respective sets of scatterers. In this section, we make no weak eavesdropper or any other assumptions.

After unitary (and thus information-preserving) transformations, this scenario can be put into the parallel channel setting of [20, 21]. In this case, the secrecy capacity and optimal covariance can be explicitly characterized.

PROPOSITION 5.3 *Consider the wiretap MIMO channel in* (5.2), (5.62). *The optimal Tx covariance for this channel takes the form*

$$\mathbf{R}^* = \mathbf{V}\boldsymbol{\Lambda}^*\mathbf{V}^+, \tag{5.63}$$

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

$$\lambda_i^* = \frac{\lambda_{2i}+\lambda_{1i}}{2\lambda_{2i}\lambda_{1i}}\left(\sqrt{1+\frac{4\lambda_{2i}\lambda_{1i}}{(\lambda_{2i}+\lambda_{1i})^2}\left(\frac{\lambda_{1i}-\lambda_{2i}}{\lambda}-1\right)_+}-1\right) \tag{5.64}$$

where $\lambda_{ki}=\sigma_{ki}^2$ *and* σ_{ki} *denotes a singular value of* $\mathbf{H}_k$*;* $\lambda>0$ *is found from the total power constraint* $\sum_i \lambda_i^* = P_{\mathrm{T}}$.

Proof After a unitary transformation, the problem can be put into the parallel channel setting; see [12] for details.

In fact, Eq. (5.63) says that optimal signaling is on the right SVs of $\mathbf{H}_{1,2}$, and (5.64) implies that only those eigenmodes are active for which

$$\sigma_{1i}^2 > \sigma_{2i}^2 + \lambda. \tag{5.65}$$

If $\lambda_{2i}=0$, then (5.64) reduces to

$$\lambda_i^* = \left(\frac{1}{\lambda}-\frac{1}{\lambda_{1i}}\right)_+, \tag{5.66}$$

i.e., as for 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 (5.62) 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.

5.9 When Is ZF Signaling Optimal?

In this section, we consider the case when popular ZF signaling is optimal for the MIMO WTC, i.e., when active eigenmodes of optimal covariance $\mathbf{R}^*$ are orthogonal to those of $\mathbf{W}_2$: $\mathbf{W}_2\mathbf{R}^* = \mathbf{0}$.[4] It is clear that this does not hold in general. However, the importance of this scenario comes 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 the secrecy requirement imposes no extra complexity penalty but is rather ensured by well-established ZF signaling.

PROPOSITION 5.4 *Consider the wiretap MIMO channel in* (5.2) *and let* $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *have the same eigenvectors [so that* $\mathbf{H}_1$ *and* $\mathbf{H}_2$ *have the same right singular vectors as in* (5.62)*] and*

$$\lambda_{1i} \le \lambda_{2i} + \lambda \text{ if } \lambda_{2i} > 0, \tag{5.67}$$

where λ *is found from the total power constraint* $\sum_i \lambda_i^* = P_{\mathrm{T}}$,

$$\lambda_i^* = \lambda_i(\mathbf{R}^*) = \left(\frac{1}{\lambda} - \frac{1}{\lambda_{1i}}\right)_+ \text{ if } \lambda_{2i} = 0, \tag{5.68}$$

and 0 otherwise. Then, the 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* $\mathbf{W}_2$ *and the optimal covariance is as in* (5.63)*, so that its eigenvectors are those of* $\mathbf{W}_{1,2}$.

The necessary condition of ZF optimality is that active eigenvectors of $\mathbf{R}^*$ *are also the active eigenvectors of* $\mathbf{W}_1$ *and inactive eigenvectors of* $\mathbf{W}_2$*, and that the power allocation is given by* (5.68).

Proof Based on the necessary KKT conditions, which, under the ZF condition $\mathbf{W}_2\mathbf{R} = \mathbf{0}$, have a unique solution; see [12] for details.

REMARK 5.5 *The optimal power allocation in* (5.68) *is the same as standard WF. However, a subtle difference here is the condition for an eigenmode to be active,* $\lambda_i^* > 0$*: while standard WF requires* $\lambda_{1i} > \lambda$*, the solution above additionally requires* $\lambda_{2i} = 0$*, so that the set of active eigenmodes is generally smaller; the smaller the set of active eigenmodes, the larger the set of eavesdropper positive 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 Ev gets no signal (via the ZF), and part 2 is the standard signaling and WF on the active eigenmodes of the legitimate channel as if the Ev were not there. No new wiretap codes need to be designed as regular coding on the required channel suffices, so that the secrecy requirement does

[4] This simply means that the Tx antenna array puts null in the direction of the eavesdropper, which is known as null forming in antenna array literature [25]. This can also be considered as a special case of interference alignment, so that Proposition 5.4 establishes its optimality.

not impose an extra complexity penalty (beyond the standard ZF). This is reminiscent of the classical source–channel coding separation [26].

5.10 When Is Standard Water-Filling Optimal?

Motivated by the fact that the transmitter may be unaware of the presence of an eavesdropper and hence uses the standard transmission on the eigenmodes of $\mathbf{W}_1$ with power allocated via the 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 $\mathbf{R}_{\mathrm{WF}}$ be the optimal Tx covariance matrix for transmission on $\mathbf{W}_1$ only, which is given by standard WF over the eigenmodes of $\mathbf{W}_1$:

$$\mathbf{R}_{\mathrm{WF}} = \mathbf{U}_1 \mathbf{\Lambda}^* \mathbf{U}_1^+, \ \lambda_i^* = \left\{ \frac{1}{\lambda} - \frac{1}{\lambda_{1i}} \right\}_+, \tag{5.69}$$

where $\mathbf{\Lambda}^* = \mathrm{diag}\{\lambda_i^*\}$ is a diagonal matrix of the eigenvalues of $\mathbf{R}_{\mathrm{WF}}$, and λ is found from the total power constraint $\sum_i \lambda_i^* = P_{\mathrm{T}}$. Alternatively, one can consider P_{T} as parameterized by λ, where $P_{\mathrm{T}}(\lambda)$ is monotonically decreasing in λ, with $P_{\mathrm{T}} \to 0$ as $\lambda \to \lambda_{11}$ and $P_{\mathrm{T}} \to \infty$ as $\lambda \to 0$.

THEOREM 5.5 *The standard WF transmit covariance matrix in* (5.69) *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*

$$\lambda_{2i} = \frac{\lambda_{1i}}{1 + \alpha \lambda_{1i}} < \lambda_{1i}, \text{ for some } \alpha > 0, \tag{5.70}$$

   or, equivalently, $\lambda_{2i}^{-1} = \lambda_{1i}^{-1} + \alpha$*;*
3. *for inactive eigenmodes* $\lambda_i^* = 0$*, the eigenvalues* λ_{1i} *and* λ_{2i} *are related either as in* (5.70) *or* $\lambda_{1i} \le \lambda_{2i}$.

Proof See Appendix.

Note that the conditions of Theorem 5.5 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 (5.70) implies that larger λ_{1i} corresponds to larger λ_{2i}, so that, over the active signaling subspace, the channel is degraded.

The first condition in Theorem 5.5 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 base station and the legitimate channel as well as the eavesdropper experience the same scattering with their own individual scatterers around their own receivers (which determine the left singular vectors), as in Section 5.8.

5.11 When Is Isotropic Signaling Optimal?

In the regular MIMO channel ($\mathbf{W_2} = \mathbf{0}$), isotropic signaling 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, identically distributed codewords, so that no precoding, no Tx CSI, and thus no feedback is required), we consider isotropic signaling over the wiretap 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 5.5 *Consider the MIMO wiretap channel in* (5.2). *The isotropic signaling is optimal, i.e.,* $\mathbf{R}^* = a\mathbf{I}$ *in* (5.4), *for the set of channels* $\{\mathbf{W}_1, \mathbf{W}_2\}$ *that can be characterized as follows:*

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 the ordered eigenvalues of* $\mathbf{W}$.
3. *Take any* $b_1 > 0$ *and* $a_1 < b_1$, *and set*

$$\lambda = (a_1 + a)^{-1} - (b_1 + a)^{-1} > 0. \tag{5.71}$$

4. *For* $i = 2, \ldots, m$, *take any* b_i *such that*

$$b_i > \lambda a^2 (1 - \lambda a)^{-1} > 0, \tag{5.72}$$

and set

$$a_i = -a + (\lambda + (b_i + a)^{-1})^{-1} > 0. \tag{5.73}$$

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

Proof See [11].

Note that a 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. Unlike the regular channel, however, there is also a large set of channels with distinct eigenvalues which dictate 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 first channel is compensated for by a carefully adjusted non-isotropy of the second one.

5.12 An Algorithm for Global Maximization of Secrecy Rates

Although a number of analytical solutions are available for an optimal Tx covariance matrix in the MIMO WTC (as discussed above), the general case remains an open problem. In this section, we introduce an algorithm for the global maximization of secrecy rates with guaranteed convergence [13].

Due to the non-convex nature of the optimization problem in (5.4) in the general case, constructing a numerical algorithm faces an immediate difficulty as global convergence cannot be guaranteed (see, e.g., [27, 28] for such algorithms).

Instead, we adopt an equivalent minimax reformulation of (5.4) [3]:

$$C_{\rm s} = \max_{\mathbf{R}} \min_{\mathbf{K}} f(\mathbf{R},\mathbf{K}) = \min_{\mathbf{K}} \max_{\mathbf{R}} f(\mathbf{R},\mathbf{K}), \tag{5.74}$$

where[5]

$$f(\mathbf{R},\mathbf{K}) = \frac{1}{2}\ln\frac{|\mathbf{I}+\mathbf{K}^{-1}\mathbf{H}\mathbf{R}\mathbf{H}'|}{|\mathbf{I}+\mathbf{W}_2\mathbf{R}|} \geq C(\mathbf{R}), \tag{5.75}$$

$$\mathbf{K} = \begin{pmatrix} \mathbf{I} & \mathbf{K}_{21}' \\ \mathbf{K}_{21} & \mathbf{I} \end{pmatrix} \geq \mathbf{0},\ \mathbf{H} = \begin{pmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{pmatrix}, \tag{5.76}$$

and the optimization is over the set $\mathcal{S}$ of all feasible $\mathbf{R},\mathbf{K}$:

$$\mathcal{S} = \{(\mathbf{R},\mathbf{K}) : \mathrm{tr}\,\mathbf{R} \leq P,\ \mathbf{R},\mathbf{K} \geq \mathbf{0},\ \mathbf{K} \text{ as in (5.76)}\}. \tag{5.77}$$

Even if this reformulation may appear to be more difficult to solve (since it involves both min and max), this is not the case: the problem in (5.74) is convex (since $f(\mathbf{R},\mathbf{K})$ is convex–concave in the right way) so that the KKT conditions are sufficient for global optimality and the global convergence of a numerical algorithm for this reformulated problem is within reach. The KKT conditions (of which there are two sets: one for min and one for max; see [13] for details) are solved below via a numerical algorithm, which is based on the barrier method and the primal/dual version of the Newton method.

To account for the positive semi-definite constraints $\mathbf{R},\mathbf{K} \geq \mathbf{0}$, we use the barrier function $\psi_t(\mathbf{R}) = t^{-1}\ln|\mathbf{R}|$ so that the modified objective f_t is

$$f_t(\mathbf{R},\mathbf{K}) = f(\mathbf{R},\mathbf{K}) + \psi_t(\mathbf{R}) - \psi_t(\mathbf{K}), \tag{5.78}$$

where $t > 0$ is the barrier parameter (see [13, 29]). Let us introduce the following variables and gradients/Hessians [13]:

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix},\ \nabla f_t = \begin{bmatrix} \nabla_x f_t \\ \nabla_y f_t \end{bmatrix},\ \nabla^2 f_t = \begin{bmatrix} \nabla^2_{xx} f_t & \nabla^2_{xy} f_t \\ \nabla^2_{yx} f_t & \nabla^2_{yy} f_t \end{bmatrix}, \tag{5.79}$$

where $\mathbf{x} = \mathrm{vech}(\mathbf{R})$ and vech stacks column-wise all lower-triangular entries into a single column vector, and we use only $\mathbf{K}_{21}$ as independent variables: $\mathbf{y} = \mathrm{vec}(\mathbf{K}_{21})$. The expressions for gradients $\nabla_{x(y)} f_t$ and Hessians $\nabla^2_{xx(yy)} f_t$ can be found in [13] (alternatively, they can be evaluated numerically). The total power constraint $\mathrm{tr}\,\mathbf{R} = P_{\rm T}$ can be expressed in the following form: $\mathbf{a}'\mathbf{z} = P_{\rm T}$, where $\mathbf{a} = [\mathrm{vech}(\mathbf{I}_m)', \mathbf{0}']'$, $\mathbf{0}$ is the $n_1 n_2 \times 1$ all-zero vector, and $\mathbf{I}_m$ is the $m \times m$ identity matrix.

With this choice of variables and new objective in (5.78), the barrier method transforms the inequality-constrained problem in (5.74)–(5.77) into the following

[5] In this section, we use the real-valued channel model, so that all entries of $\mathbf{H},\mathbf{R},\mathbf{K}$ are real. Equivalently, one can consider real and imaginary parts as independent variables.

problem without inequality constraints:

$$\max_{\mathbf{x}} \min_{\mathbf{y}} f_t(\mathbf{x},\mathbf{y}),\ \text{s.t. } \mathbf{a}'\mathbf{z} = P_{\mathrm{T}}. \tag{5.80}$$

To solve this problem via the primal/dual Newton method, let $\mathbf{w} = [\mathbf{z}', \lambda]'$ be the vector of aggregated (primal/dual) variables, where $\lambda \geq 0$ is the dual variable (Lagrange multiplier responsible for the power constraint), and $\Delta\mathbf{w} = (\Delta\mathbf{z}', \Delta\lambda)'$ is its update, which is found as the solution of the following system of linear equations at each step:

$$\mathbf{T}\Delta\mathbf{w} = -\mathbf{r}(\mathbf{w}), \tag{5.81}$$

where $\mathbf{T}$ is the KKT matrix,

$$\mathbf{T} = \begin{bmatrix} \nabla^2 f_t & \mathbf{a} \\ \mathbf{a}' & \mathbf{0} \end{bmatrix}, \tag{5.82}$$

and $\mathbf{r}(\mathbf{w})$ is the residual

$$\mathbf{r}(\mathbf{w}) = [(\nabla f_t + \mathbf{a}\lambda)', (\mathbf{a}'\mathbf{z} - P_{\mathrm{T}})]'. \tag{5.83}$$

The solution of the KKT conditions corresponds to $\mathbf{r}(\mathbf{w}) = \mathbf{0}$. At each step, the variables are updated as $\mathbf{w} \to \mathbf{w} + \Delta\mathbf{w}$.

Since the algorithm requires an initial point to begin with, we use the following point:

$$\mathbf{R}_0 = m^{-1}P\mathbf{I} \to \mathbf{x}_0 = \mathrm{vech}(\mathbf{R}_0),\ \mathbf{K}_0 = \mathbf{I} \to \mathbf{y}_0 = \mathbf{0},\ \lambda_0 = 0, \tag{5.84}$$

which is clearly feasible ($\mathbf{R}_0$ corresponds to isotropic signaling).

With this choice of variables and initial points, Algorithm 5.3, in combination with Algorithms 5.1 and 5.2, can now be used to solve numerically the minimax problem in (5.80); α, β in Algorithms 5.1–5.3 are backtracking line search parameters, s is the step size, ϵ is the desired accuracy, and μ is the barrier increase parameter [13, 29]. Global convergence of this algorithm has been proved in [13].

Algorithm 5.1 Backtracking line search

Require: $\mathbf{w}$, $0 < \alpha < 1/2$, $0 < \beta < 1$, $s = 1$.
  while $|\mathbf{r}(\mathbf{w} + s\Delta\mathbf{w})| > (1 - \alpha s)|\mathbf{r}(\mathbf{w})|$ **do** $s := \beta s$
  end while

Algorithm 5.2 Newton method for minimax optimization

Require: $\mathbf{z}_0, \lambda_0, \alpha, \beta, \epsilon$
  repeat
    1. Find $\Delta\mathbf{z}, \Delta\lambda$ using Newton step in (5.81).
    2. Find s using the backtracking line search (Algorithm 5.1).
    3. Update variables: $\mathbf{z}_{k+1} = \mathbf{z}_k + s\Delta\mathbf{z}, \lambda_{k+1} = \lambda_k + s\Delta\lambda$.
  until $|\mathbf{r}(\mathbf{z}_{k+1}, \lambda_{k+1})| \leq \epsilon$.

Algorithm 5.3 Barrier method

Require: $\mathbf{z}, \lambda, \epsilon > 0, t > 0, \mu > 1$

repeat

1. Solve the problem in (5.80) using the Newton method (Algorithm 5.2) starting at $\mathbf{z}, \lambda$.

2. Update variables: $\mathbf{z} := \mathbf{z}^*(t), \lambda := \lambda^*(t),\ t := \mu t$.

until $1/t < \epsilon$.

To demonstrate the algorithm's performance, we consider the following example:

$$\mathbf{H}_1 = \begin{bmatrix} 0.77 & -0.30 \\ -0.32 & -0.64 \end{bmatrix}, \mathbf{H}_2 = \begin{bmatrix} 0.54 & -0.11 \\ -0.93 & -1.71 \end{bmatrix}. \tag{5.85}$$

Convergence of the Newton method for different values of the barrier parameter t is demonstrated in Fig. 5.6, which shows the Euclidian norm of the residual $\mathbf{r}$ versus Newton steps. Even though this channel is not degraded, since the eigenvalues of $\mathbf{W}_1 - \mathbf{W}_2$ are $\{0.395, -3.293\}$, the algorithm does find the global optimum (this particular channel was selected because it is "difficult" for optimization; note also that the channel is not degraded so that the problem in (5.4) is not convex). Note the presence of two convergence phases: linear and quadratic, which is typical for Newton methods in general. After the quadratic phase is reached, the convergence is very fast (waterfall region). It takes 10–20 Newton steps to reach a very low residual (at the level of machine precision). This is in agreement with the observations in [29] (although obtained for different problems).

Figure 5.7 shows the corresponding secrecy rate evaluated via the upper bound in (5.75) and the actual achievable rate via $C(\mathbf{R}(t))$ in (5.3), where $\mathbf{R}(t)$ is an optimal

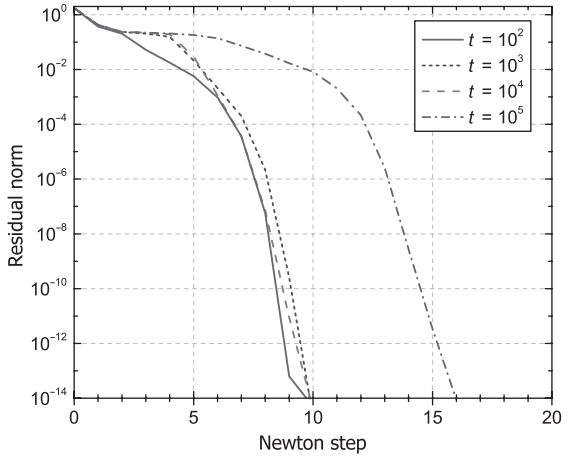

Figure 5.6 Convergence of the Newton method for different values of t; $m = 2$, $P = 10$, $\alpha = 0.3$, $\beta = 0.5$, $\mathbf{H}_1, \mathbf{H}_2$ as in (5.85). Note the presence of two convergence phases: linear and quadratic. It takes only 10 to 20 Newton steps to reach the machine precision level.

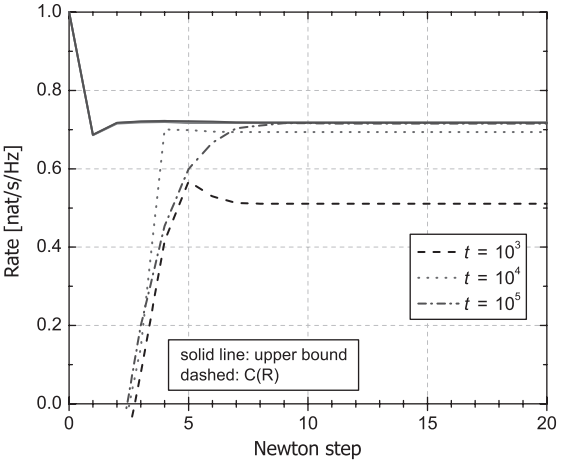

Figure 5.7 Secrecy rates for the same settings as in Fig. 5.6. Solid line: via the upper bound in (5.75) (the lines coincide for different t); dashed line: via $C(\mathbf{R})$ in (5.3). While $t = 10^3$ is sufficient for accurate computation of the upper bound, $t = 10^4$–10^5 is needed for the accurate computation of $C(\mathbf{R})$.

covariance at a particular step of the Newton method and for a given t. As the algorithm converges, they become almost equal if t is sufficiently large (in this case, about 10^4–10^5). While t has negligible impact on the upper bound, it significantly affects the corresponding $C(\mathbf{R}(t))$ so that the choice of t is not critical if the secrecy capacity is the only quantity of interest (since the upper bound is quite tight even for moderate t). However, if a transmitter is implemented with the optimal covariance $\mathbf{R}(t)$ returned by the algorithm, it is $C(\mathbf{R}(t))$ that determines the achievable rate and this choice is important. We attribute this fact to higher sensitivity of $C(\mathbf{R})$ to $\mathbf{R}$ compared to that of $f(\mathbf{R}, \mathbf{K})$. Similar observations apply to the number of Newton steps required to achieve a certain performance: if C_s is the quantity of interest, the upper bound converges to it in 3–5 steps. However, when using $\mathbf{R}$ in a system design, $C(\mathbf{R})$ should be used as a performance metric and, in addition to proper choice of t, it takes 5–10 steps to achieve the convergence for $C(\mathbf{R})$. Note that in both cases the number of steps is not large and the execution time is small (a few seconds). In general, larger t and m, n_1, n_2 require more steps to achieve the same accuracy. As expected, the behavior of the upper bound is not monotonic, while the residual norm decreases monotonically in each step.

5.13 Appendix

5.13.1 Proof of Theorem 5.4

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

$$C_s = \max_{\mathrm{tr}\mathbf{R} \leq P_T} \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \mathbf{W_2}\mathbf{R}|} \geq \max_{\mathrm{tr}\mathbf{R} \leq P_T} \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \epsilon\mathbf{R}|} = C^*(\epsilon). \tag{5.86}$$

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

$$C(\mathbf{R}) = \ln\frac{|\mathbf{I}+\mathbf{P}_2\mathbf{W}_1\mathbf{P}_2\mathbf{R}|}{|\mathbf{I}+\mathbf{P}_2\mathbf{W}_2\mathbf{P}_2\mathbf{R}|} = \ln\frac{|\mathbf{I}+\tilde{\mathbf{W}}_1\tilde{\mathbf{R}}|}{|\mathbf{I}+\varepsilon\tilde{\mathbf{R}}|} = \tilde{C}(\tilde{\mathbf{R}}), \tag{5.87}$$

where $\tilde{\mathbf{R}} = \mathbf{U}_{2+}^{+}\mathbf{R}\mathbf{U}_{2+}$ 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

$$\mathrm{tr}\,\tilde{\mathbf{R}} = \mathrm{tr}\,\mathbf{U}_{2+}^{+}\mathbf{R}\mathbf{U}_{2+} = \sum_i \lambda_i(\mathbf{R})|\mathbf{u}_{2+i}^{+}\mathbf{u}_{Ri}|^2 \le \sum_i \lambda_i(\mathbf{R}) = \mathrm{tr}\mathbf{R} \le P_{\mathrm{T}}, \tag{5.88}$$

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

$$C_{\mathrm{s}} = \max_{\mathrm{tr}\mathbf{R}\le P_{\mathrm{T}}} C(\mathbf{R}) \le \max_{\mathrm{tr}\tilde{\mathbf{R}}\le P_{\mathrm{T}}} \tilde{C}(\tilde{\mathbf{R}}) = \max_{\lambda_i\ge 0,\ \sum_i\lambda_i\le P_{\mathrm{T}}} \sum_i \ln\frac{1+\tilde{g}_i\lambda_i}{1+\epsilon\lambda_i} = \tilde{C}^*(\varepsilon), \tag{5.89}$$

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

$$\lambda_i(\tilde{\mathbf{W}}_1) = \lambda_i(\mathbf{U}_{2+}^{+}\mathbf{W}_1\mathbf{U}_{2+}) = \lambda_i([\mathbf{U}_2^{+}\mathbf{W}_1\mathbf{U}_2]_{r_2\times r_2}) \le \lambda_i(\mathbf{U}_2^{+}\mathbf{W}_1\mathbf{U}_2) = \lambda_i(\mathbf{W}_1), \tag{5.90}$$

where $[\mathbf{A}]_{k\times k}$ denotes the $k\times k$ principal sub-matrix of $\mathbf{A}$, $r_2 = \mathrm{rank}(\mathbf{W}_2)$. The inequality is due to the Cauchy eigenvalue interlacing theorem [17], and the last equality is due to the fact that $\mathbf{U}_2\mathbf{W}_1\mathbf{U}_2^{+}$ and $\mathbf{W}_1$ have the same eigenvalues. Based on this, one obtains

$$C_{\mathrm{s}} \le \tilde{C}^*(\varepsilon) \le \max_{\lambda_i\ge 0,\ \sum_i\lambda_i\le P_{\mathrm{T}}} \sum_i \ln\frac{1+g_i\lambda_i}{1+\epsilon\lambda_i} = C^*(\varepsilon), \tag{5.91}$$

thus establishing $C_{\mathrm{s}} = C^*(\varepsilon)$ under an omnidirectional eavesdropper with $\mathcal{R}(\mathbf{W}_1) \in \mathcal{R}(\mathbf{W}_2)$.

5.13.2 Proof of Theorem 5.5

We assume that $\mathbf{W}_1$ and $\mathbf{W}_2$ are non-singular; the singular case will follow from the standard continuity argument. The KKT conditions for the optimal covariance $\mathbf{R} = \mathbf{R}_{\mathrm{WF}}$, which are necessary for optimality in (5.4), can be expressed as

$$(\mathbf{W}_1^{-1}+\mathbf{R})^{-1} - (\mathbf{W}_2^{-1}+\mathbf{R})^{-1} = \lambda'\mathbf{I} - \mathbf{M} \tag{5.92}$$

$$\lambda'(\mathrm{tr}\,\mathbf{R} - P_{\mathrm{T}}) = 0,\ \mathbf{M}\mathbf{R} = 0 \tag{5.93}$$

$$\lambda' \ge 0,\ \mathbf{M},\mathbf{R} \ge 0,\ \mathrm{tr}\,\mathbf{R} \le P_{\mathrm{T}}, \tag{5.94}$$

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 $\mathrm{tr}\,\mathbf{R} \le P_{\mathrm{T}}$. Multiplying both sides of (5.92) by $\mathbf{U}_1^{+}$ on the left and by $\mathbf{U}_1$ on the right, one obtains

$$(\mathbf{\Lambda}_1^{-1}+\mathbf{\Lambda}^*)^{-1} - (\mathbf{\Lambda}_2^{-1}+\mathbf{\Lambda}^*)^{-1} = \lambda'\mathbf{I} - \mathbf{U}_1^{+}\mathbf{M}\mathbf{U}_1 = \lambda'\mathbf{I} - \mathbf{\Lambda}_M, \tag{5.95}$$

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 $\mathbf{U}_1^+\mathbf{M}\mathbf{U}_1$ are diagonal so that the last term has to be diagonal too: $\mathbf{U}_1^+\mathbf{M}\mathbf{U}_1 = \mathbf{\Lambda}_M$, i.e., $\mathbf{M}$ has the same eigenvectors as $\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}$. The complementary slackness in (5.93) implies that $\lambda_i^* \lambda_{Mi} = 0$, where λ_{Mi} is the ith eigenvalue of $\mathbf{M}$, i.e., if $\lambda_i^* > 0$ (active eigenmode) then $\lambda_{Mi} = 0$ so that, after some manipulations, (5.95) can be expressed as

$$\begin{aligned} \lambda_i^* &= \frac{1}{(\lambda_{2i}^{-1} + \lambda_i^*)^{-1} + \lambda'} - \frac{1}{\lambda_{1i}} \\ &= \lambda^{-1} - \lambda_{1i}^{-1} \end{aligned} \tag{5.96}$$

for each $\lambda_i^* > 0$, where the second equality follows from (5.69). Therefore,

$$\lambda = (\lambda_{2i}^{-1} + \lambda_i^*)^{-1} + \lambda' \tag{5.97}$$

and hence

$$\lambda_i^* = (\lambda - \lambda')^{-1} - \lambda_{2i}^{-1} = \lambda^{-1} - \lambda_{1i}^{-1}, \tag{5.98}$$

so that $\lambda_{2i}^{-1} = \lambda_{1i}^{-1} + \alpha$ with $\alpha = (\lambda - \lambda')^{-1} - \lambda^{-1} > 0$ satisfies both equalities in (5.96).

For inactive eigenmodes $\lambda_i^* = 0$, it follows from (5.95) that

$$\lambda_{1i} - \lambda_{2i} = \lambda' - \lambda_{Mi} \leq \lambda'. \tag{5.99}$$

Observe that this inequality is satisfied when $\lambda_{1i} \leq \lambda_{2i}$ (since $\lambda' > 0$). To see that it also holds under (5.70), observe that

$$\lambda_{1i} - \lambda_{2i} = \frac{\alpha \lambda_{1i}^2}{1 + \alpha \lambda_{1i}} \leq \frac{\alpha \lambda^2}{1 + \alpha \lambda} = \lambda', \tag{5.100}$$

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 (5.99) and hence the KKT conditions in (5.92)–(5.94) have a unique solution which also satisfies (5.69). This proves the optimality of $\mathbf{R}_{\mathrm{WF}}$.

References

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

[2] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas 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] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[6] Z. Li, W. Trappe, and R. Yates, "Secret communication via multi-antenna transmission," in *Proc. Conf. Inf. Sciences Systems*, Baltimore, MD, USA, Mar. 2007, pp. 905–910.
[7] J. Li and A. Petropulu, "Transmitter optimization for achieving secrecy capacity in Gaussian MIMO wiretap channels," Sep. 2009. [Online]. Available: http://arxiv.org/abs/0909.2622
[8] 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.
[9] M. C. Gursoy, "Secure communication in the low-SNR regime," *IEEE Commun.*, vol. 60, no. 4, pp. 1114–1123, Apr. 2012.
[10] S. Loyka and C. D. Charalambous, "On optimal signaling over secure MIMO channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, Jul. 2012, pp. 443–447.
[11] S. Loyka and C. D. Charalambous, "Further results on optimal signaling over secure MIMO channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2019–2023.
[12] S. Loyka and C. D. Charalambous, "Rank-deficient solutions for optimal signaling over secure MIMO channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun. 2014, pp. 201–205.
[13] 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.
[14] A. Khisti, "Interference alignment for the multiantenna compound wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2976–2993, May 2011.
[15] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy results for compound wiretap channels," *Probl. Inf. Transmission*, vol. 49, no. 1, pp. 73–98, Mar. 2013.
[16] R. F. Schaefer and S. Loyka, "The secrecy capacity of compound MIMO Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5535–5552, Dec. 2015.
[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1999.
[18] T. S. Rappaport, *Wireless Communications*, 2nd edn. Upper Saddle River, NJ: Prentice Hall, 2002.
[19] 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.
[20] 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.
[21] Z. Li, R. Yates, and W. Trappe, *Securing Wireless Communications at the Physical Layer*. Boston, MA: Springer US, 2010, pp. 1–18.
[22] D. N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge: Cambridge University Press, 2005.
[23] 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.
[24] 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, Jun. 2002.
[25] H. L. van Trees, *Optimum Array Processing*. Chichester: Wiley, 2002.
[26] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd edn. Chichester: Wiley & Sons, 2006.
[27] 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.

[28] 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 Int. Conf. Acoustics, Speech, Signal Process.*, Florence, Italy, May 2014, pp. 5686–5690.

[29] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2004.