

Article

# Minimax Duality for MIMO Interference Networks

Andreas Dotzler <sup>\*</sup>, Maximilian Riemensberger <sup>†</sup> and Wolfgang Utschick <sup>†</sup>

Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany; riemensberger@tum.de (M.R.); utschick@tum.de (W.U.)

<sup>\*</sup> Correspondence: dotzler@tum.de; Tel.: +49-89-289-28520; Fax: +49-89-289-28504

<sup>†</sup> These authors contributed equally to this work.

Academic Editors: Mikael Skoglund, Lars K. Rasmussen and Tobias Oechtering

Received: 1 November 2015; Accepted: 6 February 2016; Published: 23 March 2016

**Abstract:** A minimax duality for a Gaussian mutual information expression was introduced by Yu. An interesting observation is the relationship between cost constraints on the transmit covariances and noise covariances in the dual problem via Lagrangian multipliers. We introduce a minimax duality for general MIMO interference networks, where noise and transmit covariances are optimized subject to linear conic constraints. We observe a fully symmetric relationship between the solutions of both networks, where the roles of the optimization variables and Lagrangian multipliers are inverted. The least favorable noise covariance itself provides a Lagrangian multiplier for the linear conic constraint on the transmit covariance in the dual network, while the transmit covariance provides a Lagrangian multiplier for the constraint on the interference plus noise covariance in the dual network. The degrees of freedom available for optimization are constituted by linear subspaces, where the orthogonal subspaces induce the constraints in the dual network. For the proof of our duality we make use of the existing polite water-filling network duality and as a by-product we are able to show that maximization problems in MIMO interference networks have a zero-duality gap for a special formulation of the dual function. Our minimax duality unifies and extends several results, including the original minimax duality and other known network dualities. New results and applications are MIMO transmission strategies that manage and handle uncertainty due to unknown inter-cell interference and information theoretic proofs concerning cooperation in networks and optimality of proper signaling.

**Keywords:** MIMO; interference network; duality

## 1. Introduction

From the literature, we know network dualities of the following three types: signal-to-interference-and-noise-ratio (SINR) dualities [1–4], mean-square-error (MSE) dualities [5–7], and rate dualities [4,8–12]. The dualities show that the region of achievable SINRs, MSEs, and rates, respectively, are the same for two networks. The most prominent and important application of network dualities is the duality of the broadcast channel (BC) and the multiple access channel (MAC) with interference (pre-)subtraction, where the dualities show that SINR, MSE, and achievable rate region are identical for Gaussian signalling. The main benefit for converting the BC problem (downlink) into the MAC (uplink) domain is that for the optimal decoding order of the users, the problem formulation obtained for the uplink can be reformulated as a convex optimization problem.

The network dualities for the BC-MAC scenario with interference (pre-)subtraction and a sum-power constraint have been extended along several directions. A more general BC-MAC duality, which also holds for linear signal processing, *i.e.*, without non-linear interference cancellation, is introduced in [11]. In [12,13], dualities for general interference networks are considered.

Furthermore, the solutions are characterized by polite-waterfilling, a result we use in the proof of our duality. While most network dualities only hold for a sum-power constraint, there are some network dualities that extended to more general constraints. Dualities for scenarios with multiple linear constraints, or per antenna power constraint as a special case, can be found in [13–16]. In [10], a worst-case noise optimization is considered by a minimax duality and the connection of network duality and Lagrangian duality is revealed. Worst-case noise optimization for MIMO communications is also considered in [17,18]. Minimax expressions also appear when considering the compound capacity of a MIMO channel [19,20], where the uncertainty is in the channel knowledge instead of the noise.

SINR dualities and conic optimization, for example, semidefinite programming and second order cone programming, are frequently used in the context of power minimization problems under SINR constraints, or the related max-min SINR optimization [1,21,22]. Usually, transmit filters are reparameterized as rank-one matrices and relaxation of the rank-one constraint reveals a convex reformulation of the problem, where the SINR constraints become conic constraints. Note that this is very different to the linear conic constraints we use. Our result generalizes the original minimax duality by Yu [10] to a larger class of constraints and to interference networks, which contain broadcast and multiple access channels (with and without non-linear interference cancellation) as special cases.

Yu [10] introduced a minimax duality for the MIMO channel. For a channel matrix  $\mathbf{H}$ , noise symbol covariance  $\mathbf{R}$ , and transmit symbol covariance  $\mathbf{Q}$ , he considers a saddle point problem of the mutual information  $W(\mathbf{R}, \mathbf{Q})$  under linear constraints:

$$\min_{\mathbf{R} \succeq 0} \left\{ \max_{\mathbf{Q} \succeq 0} \{ W(\mathbf{R}, \mathbf{Q}) : \text{tr}(\mathbf{Q}\mathbf{\Omega}^*) \leq \lambda_{\mathbf{\Omega}} \} : \text{tr}(\mathbf{R}\mathbf{\Sigma}^*) \leq \lambda_{\mathbf{\Sigma}} \right\} \quad (1)$$

The dual saddle point problem with flipped channel matrix  $\mathbf{H}^H$ , noise symbol covariance  $\mathbf{\Omega}$ , and transmit symbol covariance  $\mathbf{\Sigma}$  is given by

$$\min_{\mathbf{\Omega} \succeq 0} \left\{ \max_{\mathbf{\Sigma} \succeq 0} \{ W(\mathbf{\Omega}, \mathbf{\Sigma}) : \text{tr}(\mathbf{\Sigma}\mathbf{R}^*) \leq \lambda_{\mathbf{R}} \} : \text{tr}(\mathbf{\Omega}\mathbf{Q}^*) \leq \lambda_{\mathbf{Q}} \right\} \quad (2)$$

where the constraint matrices  $\mathbf{Q}^*, \mathbf{R}^*$  are a solution of Equation (1) with  $\lambda_{\mathbf{Q}}, \lambda_{\mathbf{R}}$  as Lagrangian multipliers. The other way round, the constraint matrices  $\mathbf{\Sigma}^*, \mathbf{\Omega}^*$  of Equation (1) are solution for Equation (2) with Lagrangian multipliers  $\lambda_{\mathbf{\Sigma}} = \lambda_{\mathbf{R}}, \lambda_{\mathbf{\Omega}} = \lambda_{\mathbf{Q}}$ . By Lagrangian duality, Yu reveals the relationship between the cost constraints in one problem, the optimization variables in the dual problem, and the Lagrangian multipliers.

We consider a minimax optimization of a weighted sum of the mutual informations of the links in a network, where the minimization is over the noise covariances and the maximization is with respect to the transmit covariances. The feasible sets of the covariances are modeled by linear conic constraints, *i.e.*, the covariance matrix has to be smaller than a constraint matrix in the sense of positive-semidefiniteness. The use of the linear conic constraints is not motivated by a specific application, instead we will see that they are general enough to model a large class of constraints, including sum-power constraints, per antenna power constraints, or shaping constraints. Shaping constraints for the optimization of transmit covariances in MIMO systems are for example considered in [23–27].

The contributions of this paper are as follows. We introduce the downlink and uplink minimax problems under linear conic constraints and show that they not only exhibit an equal optimal value, but share a fully symmetric system of optimality conditions with inverted roles of the optimization variables and Lagrangian multipliers. The least favorable noise covariance itself provides a Lagrangian multiplier for the linear conic constraint on the transmit covariance in the dual network, while the transmit covariance provides a Lagrangian multiplier for the constraint on

the interference plus noise covariance in the dual network (The idea to include the interference plus noise covariances as an optimization variable, which then appears in the optimality conditions, was independently developed in [28]. Unfortunately, the current version of [28] available at arXiv contains several flaws, the most significant one being the following: The authors claim that max-min equals min-max for problem Equations (7)–(9) and the solution is a saddle point. This is wrong—here is a counter example: Consider a interference network with square channels that are full rank. For the min-max version, the outer minimization over  $\Omega$  results in  $\Omega_l = I \forall l$  which leads to  $\Sigma_l = 0$  and a shut-down of the network. This is certainly not the result of the max-min optimization). Further, we reveal the connection of noise covariance constraint in one problem and transmit covariance in the other via linear subspaces and their orthogonal complements. Unlike existing work, we directly include scenarios where the noise covariance is non-invertible by using the rather general *Lagrangian Multiplier Rule* to formulate the optimality conditions. Computing the full subdifferential of the utility allows us to parameterize the complete solution set. The direct connection of downlink and uplink solutions via the optimality conditions renders explicit transformation rules, as discussed in [9,11], obsolete. The novel minimax duality unifies and extends several existing results, for example the least favorable noise capacity [17,18] and optimization of the transmit covariances under a per antenna power constraint [14,16].

Our work extends the traditional BC-MAC duality in two dimensions: from the BC-MAC scenario to arbitrary interference networks and from simple linear power constraints to a more general type of constraints. Many optimization problems in interference networks are proven to be NP-hard, so our duality can not help to solve those efficiently. Usually, efficient solvability is tied to the BC-MAC. By modelling this scenario as a specially structured interference network we are able to state conditions for efficient solvability by the network structure and the constraints imposed. This enables future use for other scenarios, for example relay channels where interference is removed by other means or cooperative transmission of multiple transmitters (with individual constraints), single frequency broadcast networks, or network MIMO. Concerning practical applications, there are several applications and problems that are intractable without our duality for example MIMO transmission strategies that manage and handle uncertainty by inter-cell interference. Numerically computing the worst-case noise covariances was previously not possible for more general networks. The worst-case noise approach can be used to handle the otherwise complicated coupling of interference networks, for example in cognitive radio networks [29] or device-to-device networks [30]. Optimizing the BC channel for shaping constraints was not possible without the extension to the more general constraints by our duality. The shaping constraints reduce the uncertainty in interference networks. In [31] the duality is extended to multiple linear constraints and practical algorithms for numerical computation are presented in [32]. Combining the worst case noise approach and the shaping constraints facilitates a very general cognitive radio concept that also holds for receivers with multiple receive antennas and enables efficient solvability of the underlying problems, see [29] for details. Further, our duality enables information theoretic proofs concerning cooperation in networks [33] and optimality of proper signaling [26,34].

The paper is organized as follows. In Section 2, we discuss our system model and introduce our idea of linear conic constraints. In Section 3, we present our main contribution: the minimax duality with linear conic constraints is stated in Theorem 1. Applications of the minimax duality are discussed in Section 4. In Section 4.1, we relate our work to existing results and show how they can be reproduced by reformulating them for our minimax duality, and in Section 4.2 we discuss novel results enabled by our minimax duality. In Section 5, we discuss conditions for existence of a solution to the minimax problem, we state and investigate the optimality conditions, and prove the minimax duality.

The following notations are used. Vectors are denoted by lower case bold and matrices by upper case bold letters. Additionally, we use upper case bold letters for tuples of matrices, such as  $\mathbf{A} = (\mathbf{A}_1, \dots, \mathbf{A}_K)$ . The operator  $\cdot^H$  stands for Hermitian transpose,  $\text{tr}(\cdot)$  is the trace of a matrix,

and  $|\cdot|$  is the determinant of a matrix. The order relationships  $\succeq$  and  $\preceq$  are understood in the positive-semidefinite sense for matrices. For matrix tuples,  $A \succeq B$  means  $A_k \succeq B_k \forall k$  and  $A \preceq B$  means  $A_k \preceq B_k \forall k$ . We use  $I$  for the identity matrix, or a tuple of identity matrices, and  $\mathbf{0}$  for the all zero matrix, or a tuple of all zero matrices. The dimensions of  $I$  and  $\mathbf{0}$  are either clear from the context or explicitly mentioned. We denote sets by upper case calligraphic letters. The linear subspace of  $N \times N$  Hermitian matrices is denoted as  $\mathcal{S}^{N \times N}$  and the subspace of tuples of Hermitian matrices, where the  $k$ -th element is  $N_k \times N_k$ , is denoted as  $\prod_{k \in \mathcal{K}} \mathcal{S}^{N_k \times N_k}$ . For linear subspaces, the operator  $\oplus$  denotes the direct sum. Given  $S$ , let  $T, T^\perp$  such that  $[T \ T^\perp]$  unitary and full rank,  $T^H S T \succ \mathbf{0}$ , and  $T^{\perp H} S T^\perp = \mathbf{0}$ , then we say that  $T, T^\perp$  provide a full unitary basis for the fundamental subspaces of  $S$ . By this we can parameterize  $S$  as

$$S = \begin{bmatrix} T & T^\perp \end{bmatrix} \begin{bmatrix} T^H S T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} T & T^\perp \end{bmatrix}^H$$

and the pseudo inverse, denoted as  $S^+$ , is given by  $S^+ = T(T^H S T)^{-1} T^H$ .

## 2. System Model

We consider a MIMO interference network with a set of data links  $\mathcal{K}$ , with  $K = |\mathcal{K}|$ . The transmitter of the link  $k$  has  $N_k$  antennas and the receiver has  $M_k$  receive antennas. The fixed and flat-fading channel between the transmitter of link  $i$  and the receiver of link  $k$  is  $H_{ki} \in \mathbb{C}^{M_k \times N_i}$ . The zero-mean proper Gaussian input symbol of link  $k$  is  $y_k \in \mathbb{C}^{N_k}$  with covariance  $Q_k = E[y_k y_k^H]$ , and the output symbol is  $x_k \in \mathbb{C}^{M_k}$ . The input and output of a link  $k$  are related by

$$x_k = H_{kk} y_k + \sum_{i \in \mathcal{K} \setminus k} H_{ki} y_i + n_k \quad (3)$$

where  $n_k \in \mathbb{C}^{M_k}$  is additive noise that is zero mean circularly symmetric Gaussian with covariance  $R_k = E[n_k n_k^H]$ . Assuming all transmit and noise signals to be independent, the interference plus noise covariance at the receiver of link  $k$  is

$$S_k = R_k + \sum_{i \in \mathcal{K} \setminus k} H_{ki} Q_i H_{ki}^H \quad (4)$$

The dual network has the same set of links with roles of receiver and transmitter inverted and flipped channels, meaning the channel between the uplink transmitter of link  $i$  and the uplink receiver of link  $k$  is  $H_{ik}^H \in \mathbb{C}^{N_k \times M_i}$ . We refer to the original problem as the downlink and the dual network as the uplink. In the uplink, the zero-mean proper Gaussian input symbol of link  $k$  is  $\xi_k \in \mathbb{C}^{M_k}$  with covariance  $\Sigma_k = E[\xi_k \xi_k^H]$  and the output symbol is  $\rho_k \in \mathbb{C}^{N_k}$ . The input and output of a link  $k$  are related by

$$\rho_k = H_{kk}^H \xi_k + \sum_{i \in \mathcal{K} \setminus k} H_{ik}^H \xi_i + \eta_k \quad (5)$$

where  $\eta_k \in \mathbb{C}^{N_k}$  is additive noise with covariance  $\Omega_k = E[\eta_k \eta_k^H]$ . The interference plus noise covariance at the receiver of link  $k$  is

$$\Psi_k = \Omega_k + \sum_{i \in \mathcal{K} \setminus k} H_{ik}^H \Sigma_i H_{ik} \quad (6)$$

Choosing the channel matrices accordingly, this model is general enough to cover various forms of interference networks including interfering broadcast and multiple access channels as well as mixtures thereof, see [12] for details. In Section 4.1.5, we discuss in more detail how to model a broadcast channel and the corresponding multiple access channel with non-linear interference cancellation.

For the downlink, the mutual information between  $\mathbf{x}_k$  and  $\mathbf{y}_k$ , denoted as  $I_k(\mathbf{x}_k; \mathbf{y}_k)$ , can be expressed as a function of the noise covariance  $\mathbf{R}_k$  and the transmit covariances  $\mathbf{Q} = (\mathbf{Q}_1, \dots, \mathbf{Q}_K)$ . For the uplink, the mutual information between  $\rho_k$  and  $\xi_k$ , denoted as  $I_k(\rho_k; \xi_k)$ , can be expressed as a function of the noise covariance  $\Omega_k$  and the transmit covariances  $\Sigma = (\Sigma_1, \dots, \Sigma_K)$ . For the considered network optimization problems the utilities are weighted sums of the mutual informations of the links:

$$W(\mathbf{R}, \mathbf{Q}) = \sum_{k \in \mathcal{K}} w_k I_k(\mathbf{x}_k; \mathbf{y}_k) \quad (7)$$

$$W(\Omega, \Sigma) = \sum_{k \in \mathcal{K}} w_k I_k(\rho_k; \xi_k) \quad (8)$$

where  $w_1, \dots, w_K$  are non-negative weights, which can be used to prioritize links or to explore boundary points of the achievable rate region. We assume that interference is treated as additional noise, which allows us to express the utility by the interference plus noise covariances  $\mathbf{S} = (\mathbf{S}_1, \dots, \mathbf{S}_K)$  and  $\Omega = (\Omega_1, \dots, \Omega_K)$  defined by Equations (4) and (6):

$$W(\mathbf{R}, \mathbf{Q}) = \left\{ U(\mathbf{S}, \mathbf{Q}) : \mathbf{S}_k = \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ki} \mathbf{Q}_i \mathbf{H}_{ki}^H \forall k \in \mathcal{K} \right\} \quad (9)$$

$$W(\Omega, \Sigma) = \left\{ U(\Psi, \Sigma) : \Psi_k = \Omega_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ik}^H \Sigma_i \mathbf{H}_{ik} \forall k \in \mathcal{K} \right\} \quad (10)$$

where in case the mutual information is finite

$$U(\mathbf{S}, \mathbf{Q}) = \sum_{k \in \mathcal{K}} w_k \log \left| \mathbf{I} + \mathbf{S}_k^{-1} \mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \right| \quad (11)$$

$$U(\Psi, \Sigma) = \sum_{k \in \mathcal{K}} w_k \log \left| \mathbf{I} + \Psi_k^{-1} \mathbf{H}_{kk}^H \Sigma_k \mathbf{H}_{kk} \right| \quad (12)$$

The conditions for finite mutual information and a derivation of Equations (11) and (12) can be found in Appendix A.

### 2.1. Linear Conic Constraints

The transmit and noise covariances are constrained by linear conic constraints:

$$\mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z} \quad (13)$$

$$\mathbf{R} \preceq \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \quad (14)$$

$$\Sigma \preceq \mathbf{B} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y} \quad (15)$$

$$\Omega \preceq \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \quad (16)$$

where  $\mathbf{B} = (\mathbf{B}_1, \dots, \mathbf{B}_K)$  and  $\mathbf{C} = (\mathbf{C}_1, \dots, \mathbf{C}_K)$  are fixed, and the degrees of freedom available for optimization of  $\mathbf{Z} = (\mathbf{Z}_1, \dots, \mathbf{Z}_K)$ ,  $\mathbf{Z}^\perp = (\mathbf{Z}_1^\perp, \dots, \mathbf{Z}_K^\perp)$ ,  $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_K)$ , and  $\mathbf{Y}^\perp = (\mathbf{Y}_1^\perp, \dots, \mathbf{Y}_K^\perp)$  are given by  $\mathcal{Z}, \mathcal{Z}^\perp, \mathcal{Y}, \mathcal{Y}^\perp$ , which are linear subspaces of tuples of appropriately sized Hermitian matrices, in particular  $\mathcal{Z} \subseteq \prod_{k \in \mathcal{K}} \mathcal{S}^{N_k \times N_k}$ ,  $\mathcal{Z}^\perp \subseteq \prod_{k \in \mathcal{K}} \mathcal{S}^{N_k \times N_k}$ ,  $\mathcal{Y} \subseteq \prod_{k \in \mathcal{K}} \mathcal{S}^{M_k \times M_k}$ ,  $\mathcal{Y}^\perp \subseteq \prod_{k \in \mathcal{K}} \mathcal{S}^{M_k \times M_k}$ .

The prototype for the linear conic constraints Equations (13)–(16) is

$$\mathbf{W} \preceq \mathbf{A} + \mathbf{X}, \mathbf{X} \in \mathcal{X} \quad (17)$$

The choices for  $\mathbf{X} \in \mathcal{X}$  that allow for positive semidefinite  $\mathbf{W}$  are given by  $\mathbf{A} + \mathbf{X} \succeq \mathbf{0}$ . If  $\mathbf{A}$  and  $\mathbf{X}$  are matrices (instead of tuples of matrices), this is a linear matrix inequality (LMI), the intersection of an affine subspace with the cone of positive semidefinite matrices. The constant nullspace of a LMI

is the set of vectors  $\lambda$  for which  $\lambda^H(A + X)\lambda = 0 \forall X \in \mathcal{X}$  where  $A + X \succeq \mathbf{0}$ , see [35] for details. In this work, we additionally regard tuples of matrices, where the positive semidefinite constraint has to hold for every element of the tuple, while the subspace is defined for the tuple. Hence, we define a set that is similar to the constant null-space, but is defined in the matrix tuple space instead of the vector tuple space:

$$\mathcal{N}(A, \mathcal{X}) = \left\{ \left( \Lambda_1 \Lambda_1^H, \dots, \Lambda_K \Lambda_K^H \right) : \Lambda_k^H (A_k + X_k) \Lambda_k = \mathbf{0}, \forall k, \forall X \in \mathcal{X} \text{ where } A + X \succeq \mathbf{0} \right\} \quad (18)$$

Additionally, we define a set that contains all the directions where all elements of the tuple are positive semidefinite:

$$\mathcal{P}(\mathcal{X}) = \{X \in \mathcal{X} : X \succeq \mathbf{0}\} \quad (19)$$

**Lemma 1.** Let  $\mathcal{A} = \{\lambda A : \lambda \in \mathbb{R}\}$  and  $\mathcal{X}^\perp$  the orthogonal complement of  $\mathcal{A} \oplus \mathcal{X}$  with respect to the appropriately sized Hermitian matrices, then  $\mathcal{P}(\mathcal{X}^\perp) \subseteq \mathcal{N}(A, \mathcal{X})$ .

**Proof.** See Appendix B.  $\square$

Note that in general  $\mathcal{N}(A, \mathcal{X}) \not\subseteq \mathcal{P}(\mathcal{X}^\perp)$  (Take the following example:  $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ ,  $\mathcal{X} = \left\{ \lambda \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} : \lambda \in \mathbb{R} \right\}$ ,  $\mathcal{X}^\perp = \left\{ \lambda \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} : \lambda \in \mathbb{R} \right\}$  where  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathcal{N}(A, \mathcal{X})$ , but  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \notin \mathcal{P}(\mathcal{X}^\perp) = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right\}$ ). If  $A \succ \mathbf{0}$  then  $\mathcal{N}(A, \mathcal{X}) = \mathcal{P}(\mathcal{X}^\perp) = \{\mathbf{0}\}$ .

### 2.1.1. Relationship to Linear Constraints

There exist several network dualities, where the transmit covariances are constrained by one or multiple linear constraints [13–16]. So how do the linear conic constraints relate to the linear constraints? As known from semidefinite programming, a LMI can always be expressed in primal form with a finite set of linear constraints and one positive-semidefinite constraint. In principle, this can be done for Equation (17), but this does not result in linear constraints on the covariance. The following Lemma states a condition under which Equation (17) can be replaced by multiple linear constraints on  $W$ , which then can be handled by the existing dualities [13–16].

**Lemma 2.** Let  $\mathcal{A} = \{\lambda A : \lambda \in \mathbb{R}\}$  and  $\mathcal{X}^\perp$  be two subspaces such that  $\mathcal{A} \oplus \mathcal{X}^\perp$  is the orthogonal complement of  $\mathcal{X}$ . If the set  $\mathcal{E} = \{E : E \succeq \mathbf{0}, E \in \mathcal{A} \oplus \mathcal{X}^\perp\}$  is polyhedral, then the linear conic constraint  $W \preceq A + X, X \in \mathcal{X}$ , can be formulated by a finite number of linear constraints on  $W$ , i.e.,

$$\sum_{k \in \mathcal{K}} \text{tr}((W_k - A_k)E_k^e) \leq 0 \quad \forall e \in \{1, \dots, E\} \quad (20)$$

where  $E^1, \dots, E^E$  are the extreme directions of  $\mathcal{E}$ .

**Proof.** See Appendix C.  $\square$

For the inverse direction: a single linear constraint can always be replaced by a linear conic constraint, see for example Section 2.1.2. So multiple linear constraints can be transformed to multiple conic constraints, a case which is not regarded in this work but suggested as an extension a future work in Section 6. For special cases, such as in Section 2.1.2, multiple linear constraints can be represented by a single conic constraint. Linear conic constraints allow to model a variety of constraints on the covariances such as sum power constraints, per-antenna power constraints, or shaping constraints. The constraints can be joint constraints on all links, groups of links, or individual constraints per link. In the following, we present some examples how several relevant constraints can be modeled by linear conic constraint.



### 2.1.2. Examples

#### Network Sum-Power Constraint

Consider a constraint on the downlink transmit covariances that constrains the sum-power output of the network by

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{A}_k \mathbf{Q}_k) \leq P \quad (21)$$

By Lemma 2, we identify the linear conic version of the constraint as

$$\mathbf{C} = \frac{P}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{A}_k \mathbf{A}_k)} \mathbf{A}, \mathcal{Z} = \left\{ \mathbf{Z} : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{A}_k \mathbf{Z}_k) = 0 \right\}, \mathcal{Z}^\perp = \{\mathbf{0}\} \quad (22)$$

#### Per Link Sum-Power Constraint

If the power per link is constrained individually, *i.e.*,

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{Q}_k) \leq P_k \quad \forall k \in \mathcal{K} \quad (23)$$

then by Lemma 2 we identify the linear conic version of the constraint as

$$\mathbf{C} = \left( \frac{P_1}{N_1} \mathbf{I}, \dots, \frac{P_K}{N_K} \mathbf{I} \right) \quad (24)$$

$$\mathcal{Z} = \{ \mathbf{Z} : \text{tr}(\mathbf{Z}_k) = 0 \quad \forall k \in \mathcal{K} \} \quad (25)$$

$$\mathcal{Z}^\perp = \left\{ \mathbf{Z}^\perp : \mathbf{Z}^\perp = (\lambda_1 \mathbf{I}, \dots, \lambda_K \mathbf{I}), \sum_{k \in \mathcal{K}} P_k \lambda_k = 0 \right\} \quad (26)$$

#### Shaping Constraint

A shaping constraint, that is  $\mathbf{Q} \preceq \mathbf{C}$ , is already given in the form of a linear conic constraint where,  $\mathcal{Z} = \{\mathbf{0}\}$  and  $\mathcal{Z}^\perp = \{ \mathbf{Z}^\perp : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{Z}_k) = 0 \}$ . A shaping constraint cannot be formulated by a finite number of linear constraints, as the set  $\mathcal{E}$  is the cone of positive semidefinite matrices, or the tuple version thereof, which is not polyhedral and has infinitely many extreme directions.

### 3. Minimax Duality with Linear Conic Constraints

*Theorem 1.* Given is a MIMO interference network (downlink) with data links  $\mathcal{K}$  and the dual network (uplink) with the same set of data links but roles of receiver and transmitter inverted and flipped channels. Consider minimax problems with a weighted-sum of the links' mutual information as utility and linear conic constraints on the transmit and noise covariance matrices:

$$\begin{aligned} \mathbf{C}_D &= \inf_{\mathbf{R} \succeq \mathbf{0}, \mathbf{Y}^\perp} \left\{ \sup_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z}} \{ W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z} \} : \mathbf{R} \preceq \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \\ \mathbf{C}_U &= \inf_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \sup_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \{ W(\mathbf{\Sigma}, \mathbf{\Omega}) : \mathbf{\Sigma} \preceq \mathbf{B} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y} \} : \mathbf{\Omega} \preceq \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \end{aligned}$$

The subspaces that define the constraint sets are  $\mathcal{B} = \{ \lambda \mathbf{B} : \lambda \in \mathbb{R} \}$ ,  $\mathcal{Y}, \mathcal{Y}^\perp$  and  $\mathcal{C} = \{ \lambda \mathbf{C} : \lambda \in \mathbb{R} \}$ ,  $\mathcal{Z}, \mathcal{Z}^\perp$ , where  $\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k)$ . The subspace triples  $\mathcal{B}, \mathcal{Y}, \mathcal{Y}^\perp$  and  $\mathcal{C}, \mathcal{Z}, \mathcal{Z}^\perp$ , respectively, are mutually orthogonal and their direct sum is the space of all  $K$ -tuples  $\prod_{k \in \mathcal{K}} \mathcal{S}^{M_k \times M_k}$  and  $\prod_{k \in \mathcal{K}} \mathcal{S}^{N_k \times N_k}$ , respectively. If

$$(C1) \quad \mathbf{H}_{kk} \mathbf{\Lambda}_k = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{N}(\mathbf{C}, \mathcal{Z})$$

$$\begin{aligned} \text{(C2)} \quad & \mathbf{H}_{kk}\mathbf{\Lambda}_k = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{N}(\mathbf{C}, \mathcal{Z}^\perp) \\ \text{(C3)} \quad & \mathbf{\Lambda}'_k \mathbf{H}_{kk} = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda}' \in \mathcal{N}(\mathbf{B}, \mathcal{Y}) \\ \text{(C4)} \quad & \mathbf{\Lambda}'_k \mathbf{H}_{kk} = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda}' \in \mathcal{N}(\mathbf{B}, \mathcal{Y}^\perp), \end{aligned}$$

both problems have a solution and it holds  $C_D = C_U$ .

Furthermore, there exists a solution of the downlink problem

$$\left( (\check{\mathbf{R}} = \mathbf{B} + \check{\mathbf{Y}}^\perp, \check{\mathbf{Q}} = \mathbf{C} + \check{\mathbf{Z}} - \check{\mathbf{N}}, \check{\mathbf{S}}), (\check{\mathbf{\Omega}} = \check{\gamma}\mathbf{C} + \check{\mathbf{Z}}^\perp, \check{\mathbf{\Sigma}} = \check{\gamma}\mathbf{B} + \check{\mathbf{Y}} - \check{\mathbf{L}}, \check{\mathbf{\Psi}}, \check{\mathbf{K}}) \right)$$

that is connected to a solution for the uplink problem

$$\left( (\hat{\mathbf{\Omega}} = \mathbf{C} + \hat{\mathbf{Z}}^\perp, \hat{\mathbf{\Sigma}} = \mathbf{B} + \hat{\mathbf{Y}} - \hat{\mathbf{L}}, \hat{\mathbf{\Psi}}), (\hat{\mathbf{R}} = \hat{\gamma}\mathbf{B} + \hat{\mathbf{Y}}^\perp, \hat{\mathbf{Q}} = \hat{\gamma}\mathbf{C} + \hat{\mathbf{Z}} - \hat{\mathbf{N}}, \hat{\mathbf{S}}, \hat{\mathbf{M}}) \right)$$

by inverting the roles of the optimization variables and Lagrangian multipliers, that is

$$\begin{aligned} (\check{\mathbf{R}}, \check{\mathbf{S}}, \check{\mathbf{Q}}, \check{\mathbf{Z}}, \check{\mathbf{Y}}^\perp, \check{\mathbf{N}}) &= \frac{1}{\check{\gamma}} (\hat{\mathbf{R}}, \hat{\mathbf{S}}, \hat{\mathbf{Q}}, \hat{\mathbf{Z}}, \hat{\mathbf{Y}}^\perp, \hat{\mathbf{N}}) \text{ and} \\ (\check{\mathbf{\Omega}}, \check{\mathbf{\Psi}}, \check{\mathbf{\Sigma}}, \check{\mathbf{Y}}, \check{\mathbf{Z}}^\perp, \check{\mathbf{L}}) &= \gamma (\hat{\mathbf{\Omega}}, \hat{\mathbf{\Psi}}, \hat{\mathbf{\Sigma}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}}^\perp, \hat{\mathbf{L}}) \end{aligned}$$

where  $\gamma = \check{\gamma} = \hat{\gamma}$  (We introduce additional optimization variables  $\check{\mathbf{S}}$  and  $\check{\mathbf{\Psi}}$ , which are defined by Equations (4) and (6). The connection of Lagrangian multipliers to the constraints can be found in Table 1).

**Table 1.** Uplink and Downlink Optimality Conditions.

Downlink			Uplink		
Lagrangian multiplier	Constraint		Lagrangian multiplier	Constraint	
$\mathbf{\Sigma}_k$	$S_k = \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ki} \mathbf{Q}_i \mathbf{H}_{ki}^H \quad \forall k$	(27)	$\mathbf{Q}_k$	$\mathbf{\Psi}_k = \mathbf{\Omega}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ik}^H \mathbf{\Sigma}_i \mathbf{H}_{ik} \quad \forall k$	(45)
$\mathbf{\Omega}$	$\mathbf{Q} \preceq \alpha \mathbf{C} + \mathbf{Z}$	(28)	$\mathbf{R}$	$\mathbf{\Sigma} \preceq \beta \mathbf{B} + \mathbf{Y}$	(46)
$\nu \mathbf{C} + \mathbf{Z}^\perp$	$\mathbf{Z} \in \mathcal{Z}$	(29)	$\sigma \mathbf{B} + \mathbf{Y}^\perp$	$\mathbf{Y} \in \mathcal{Y}$	(47)
$\mathbf{K}$	$\mathbf{Q} \succeq \mathbf{0}$	(30)	$\mathbf{M}$	$\mathbf{\Sigma} \succeq \mathbf{0}$	(48)
$\mathbf{X}$	$\mathbf{R} = \alpha \mathbf{B} + \mathbf{Y}^\perp$	(31)	$\mathbf{W}$	$\mathbf{\Omega} = \beta \mathbf{C} + \mathbf{Z}^\perp$	(49)
$\mu \mathbf{B} + \mathbf{Y}$	$\mathbf{Y}^\perp \in \mathcal{Y}^\perp$	(32)	$\tau \mathbf{C} + \mathbf{Z}$	$\mathbf{Z}^\perp \in \mathcal{Z}^\perp$	(50)
$\mathbf{L}$	$\mathbf{R} \succeq \mathbf{0}$	(33)	$\mathbf{N}$	$\mathbf{\Omega} \succeq \mathbf{0}$	(51)
<b>Substitutions and constraints on dual variables</b>					
$\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{K} \succeq \mathbf{0}, \mathbf{L} \succeq \mathbf{0}$	(34)		$\mathbf{R} \succeq \mathbf{0}, \mathbf{M} \succeq \mathbf{0}, \mathbf{N} \succeq \mathbf{0}$	(52)	
$\mathbf{Z}^\perp \in \mathcal{Z}^\perp, \mathbf{Y} \in \mathcal{Y}$	(35)		$\mathbf{Y}^\perp \in \mathcal{Y}^\perp, \mathbf{Z} \in \mathcal{Z}$	(53)	
$\mathbf{\Psi}_k = \mathbf{\Omega}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ik}^H \mathbf{\Sigma}_i \mathbf{H}_{ik} \quad \forall k$	(36)		$S_k = \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ki} \mathbf{Q}_i \mathbf{H}_{ki}^H \quad \forall k$	(54)	
<b>Complimentary slackness</b>					
$\text{tr}(\mathbf{\Omega}_k(\mathbf{Q}_k - \alpha \mathbf{C}_k - \mathbf{Z}_k)) = 0 \quad \forall k$	(37)		$\text{tr}(\mathbf{R}_k(\mathbf{\Sigma}_k - \beta \mathbf{B}_k - \mathbf{Y}_k)) = 0 \quad \forall k$	(55)	
$\text{tr}(\mathbf{K}_k \mathbf{Q}_k) = 0 \quad \forall k$	(38)		$\text{tr}(\mathbf{M}_k \mathbf{\Sigma}_k) = 0 \quad \forall k$	(56)	
$\text{tr}(\mathbf{L}_k \mathbf{R}_k) = 0 \quad \forall k$	(39)		$\text{tr}(\mathbf{N}_k \mathbf{\Omega}_k) = 0 \quad \forall k$	(57)	
<b>Stationarity</b>					
$\mathbf{\Omega} = \nu \mathbf{C} + \mathbf{Z}^\perp$	(40)		$\mathbf{R} = \sigma \mathbf{B} + \mathbf{Y}^\perp$	(58)	
$\mathbf{X} = \mu \mathbf{B} + \mathbf{Y}$	(41)		$\mathbf{W} = \tau \mathbf{C} + \mathbf{Z}$	(59)	
$\mathbf{\Sigma} = \mathbf{X} - \mathbf{L}$	(42)		$\mathbf{Q} = \mathbf{W} - \mathbf{N}$	(60)	
$\mathbf{\Psi}_k = w_k \mathbf{H}_{kk}^H (\mathbf{S}_k + \mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H)^+ \mathbf{H}_{kk} + \mathbf{K}_k \quad \forall k$	(43)		$S_k = w_k \mathbf{H}_{kk} (\mathbf{\Psi}_k + \mathbf{H}_{kk}^H \mathbf{\Sigma}_k \mathbf{H}_{kk})^+ \mathbf{H}_{kk}^H + \mathbf{M}_k \quad \forall k$	(61)	
$\mathbf{\Sigma}_k \in \left\{ \mathbf{\Sigma}_k : \mathbf{\Sigma}_k = -w_k (\mathbf{S}_k + \mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H)^+ + \right.$ $\left. + w_k \mathbf{S}_k^+ + w_k \mathbf{\Sigma}_k^\Delta, \mathbf{\Sigma}_k^\Delta \in \mathcal{E}_k^\Delta \right\} \quad \forall k$	(44)		$\mathbf{Q}_k \in \left\{ \mathbf{Q}_k : \mathbf{Q}_k = -w_k (\mathbf{\Psi}_k + \mathbf{H}_{kk}^H \mathbf{\Sigma}_k \mathbf{H}_{kk})^+ + \right.$ $\left. + w_k \mathbf{\Psi}_k^+ + w_k \mathbf{Q}_k^\Delta, \mathbf{Q}_k^\Delta \in \mathcal{Q}_k^\Delta \right\} \quad \forall k$	(62)	



*Remark 1.* The Theorem holds for arbitrary channel matrix sizes and there are no assumptions on the rank of the channel matrices. Whether the minimax problems have a solution does not solely depend on the channel matrices, instead we need to jointly consider the constraints and channel matrices. In general, this can not be done by analyzing individual links as the tuple subspaces can couple all links of the network. The derivation of the joint conditions on channels and constraints for existence of a solution (C1)–(C4) can be found in Section 5.1.

*Remark 2.* The mutual information of a link is independent of the noise covariances of the other links and the utility is non-increasing in the noise covariance, thus we can always find a solution where the noise covariance constraint holds with equality.

*Remark 3.* If the feasible set for the downlink covariances is given by

$$\left\{ \mathbf{R} : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{R}_k) = P, \mathbf{R} \in \mathcal{B} \oplus \mathcal{Y}^\perp \right\} \quad (63)$$

this is mathematically equivalent to

$$\left\{ \mathbf{R} : \mathbf{R} = \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (64)$$

if  $\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = P$ . This formulation was used in work based on earlier versions of the duality, see for example [36].

The Theorem requires  $\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k)$ , which is not always the case for the problems considered in the following. By scaling of the channel matrices and optimization variables we can artificially enforce the condition for the downlink. By reverting this scaling of the channel in the uplink we obtain an asymmetric version of the minimax duality, which is formally stated by the following lemma:

*Lemma 3.* Let  $b = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k)$  and  $c = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k)$ , if  $p$  and  $n$  such that  $\frac{p}{n} = \frac{c}{b}$  then

$$\inf_{\mathbf{R} \succeq 0, \mathbf{Y}^\perp} \left\{ \sup_{\mathbf{Q} \succeq 0, \mathbf{Z}} \{W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z}\} : \mathbf{R} \preceq \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (65)$$

$$= \inf_{\mathbf{\Omega} \succeq 0, \mathbf{Z}^\perp} \left\{ \sup_{\mathbf{\Sigma} \succeq 0, \mathbf{Y}} \{W(\mathbf{\Sigma}, \mathbf{\Omega}) : \mathbf{\Sigma} \preceq p\mathbf{B} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y}\} : \mathbf{\Omega} \preceq n\mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (66)$$

**Proof.** See Appendix D.  $\square$

## 4. Applications

We directly proceed with applications of the minimax duality and present the rather technical proof in Section 5.

### 4.1. Relationship to Existing Work

#### 4.1.1. Original Minimax Duality

As our result is an extension and generalization of the result in [10], we are able to reestablish the original minimax duality as a special case. Specifically, we reproduce the proof of the Sato bound to be tight for the sum-rate in the broadcast with non-linear interference presubtraction. The Sato upper bound on the sum-capacity of the multi-user downlink is established by assuming full cooperation of the receivers [37], which makes the scenario equivalent to the optimization of a point-to-point link, where the minimization is with respect to the joint noise covariance, where the user noise covariances are fixed blocks on the diagonal and the off-diagonal blocks can be freely chosen.

The set of users is  $\mathcal{U}$ , with  $U = |\mathcal{U}|$ , where  $M_u$  is the number of receive antennas of user  $u$ . The transmitter has  $N$  transmit antennas and  $M = \sum_{u \in \mathcal{U}} M_u$  the number of total receive antennas. The channel of user  $u$  is  $\mathbf{H}_u \in \mathbb{C}^{N \times M_u}$  and the composite channel of the point-to-point channel where all receivers cooperate is  $\mathbf{H} = [\mathbf{H}_1^T, \dots, \mathbf{H}_U^T]^T \in \mathbb{C}^{N \times M}$ . The receive symbol covariance in the downlink, and therefore the transmit symbol covariance in the uplink, is block partitioned such that block  $i, j$  is the covariance of the symbols received by user  $i$  and  $j$  and the  $k$ -th block on the diagonal is the marginal covariance of user  $k$ . We assume the downlink noise to be the identity matrix scaled by  $\sigma^2$  for every user. The appropriate model for subspaces  $\mathcal{B}, \mathcal{Y}, \mathcal{Y}^\perp$ , which are subspaces of  $\mathcal{S}^{M \times M}$ , is  $\mathcal{B} = \sigma^2 \mathbf{I}$ .  $\mathcal{Y}^\perp$  is the subspace of Hermitian matrices that have all zero matrices as blocks on the diagonal and allow for an arbitrary choice of the other blocks, which implies  $\mathcal{Y}$  is the set of all blockdiagonal matrices with trace equal to zero. Let the structure of  $*$  and  $*^\perp$  be as follows

$$* = \begin{bmatrix} *_{1,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & *_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & *_{M,M} \end{bmatrix}, *^\perp = \begin{bmatrix} \mathbf{0} & *_{1,2} & \dots & *_{1,M} \\ *_{2,1} & \mathbf{0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & *_{M-1,M} \\ *_{M,1} & \dots & *_{M,M-1} & \mathbf{0} \end{bmatrix} \quad (67)$$

then

$$\mathcal{B} = \begin{bmatrix} \sigma^2 \mathbf{I} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \sigma^2 \mathbf{I} \end{bmatrix}, \mathcal{Y}^\perp = \{\mathbf{Y}^\perp = *^\perp\}, \mathcal{Y} = \{\mathbf{Y} : \mathbf{Y} = *, \text{tr}(\mathbf{Y}) = 0\} \quad (68)$$

According to Section 2.1.2, the linear conic version of the sum-power constraint on the downlink transmit covariances is

$$\mathcal{C} = \frac{P}{N} \mathbf{I}, \mathcal{Z} = \{\mathbf{Z} : \text{tr}(\mathbf{Z}) = 0\}, \mathcal{Z}^\perp = \{\mathbf{0}\} \quad (69)$$

where  $\mathcal{C}, \mathcal{Z}, \mathcal{Z}^\perp$  are subspaces of  $\mathcal{S}^{N \times N}$ . The downlink problem is

$$\min_{\mathbf{R} \succeq \mathbf{0}, *^\perp} \left\{ \max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z}} \left\{ W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \frac{P}{N} \mathbf{I} + \mathbf{Z}, \text{tr}(\mathbf{Z}) = 0 \right\} : \mathbf{R} \preceq \sigma^2 \mathbf{I} + *^\perp \right\} \quad (70)$$

With  $\mathbf{B} \succ \mathbf{0}$  and  $\mathbf{C} \succ \mathbf{0}$  conditions (C1)–(C4) hold true, and using the asymmetric minimax duality by Lemma 3 the uplink problem is

$$\min_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \max_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \left\{ W(\mathbf{\Omega}, \mathbf{\Sigma}) : \mathbf{\Sigma} \preceq p\sigma^2 \mathbf{I} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y} \right\} : \mathbf{\Omega} \preceq n \frac{P}{N} \mathbf{I} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (71)$$

$$= \min_{\mathbf{\Omega} \succeq \mathbf{0}} \left\{ \max_{\mathbf{\Sigma} \succeq \mathbf{0}, *} \left\{ W(\mathbf{\Omega}, \mathbf{\Sigma}) : \mathbf{\Sigma} \preceq p\sigma^2 \mathbf{I} + *, \text{tr}(*) = 0 \right\} : \mathbf{\Omega} \preceq n \frac{P}{N} \mathbf{I} \right\} \quad (72)$$

With  $b = \text{tr}(\mathbf{B}\mathbf{B}) = M\sigma^4$ ,  $c = \text{tr}(\mathbf{C}\mathbf{C}) = \frac{P^2}{N}$ , and  $\frac{c}{b} = \frac{P^2}{NM\sigma^4}$  we can select  $p = \frac{P}{\sigma^2 M}$  and  $n = \frac{\sigma^2 N}{P}$ , as  $\frac{p}{n} = \frac{P^2}{NM\sigma^2} = \frac{c}{b}$ , and we obtain

$$\min_{\mathbf{\Omega} \succeq \mathbf{0}} \left\{ \max_{\mathbf{\Sigma} \succeq \mathbf{0}, *} \left\{ W(\mathbf{\Omega}, \mathbf{\Sigma}) : \mathbf{\Sigma} \preceq \frac{P}{M} \mathbf{I} + *, \text{tr}(*) = 0 \right\} : \mathbf{\Omega} \preceq \sigma^2 \mathbf{I} \right\} \quad (73)$$

With  $\mathbf{Z}^\perp = \{\mathbf{0}\}$ ,  $\mathbf{\Omega} = \sigma^2 \mathbf{I}$  is a solution for the uplink noise, see Remark 2. For the constraint on the uplink transmit covariance, we observe that  $\mathbf{B}$  and all elements of  $\mathcal{Y}$  are blockdiagonal. Thus,  $\mathbf{\Sigma}$  will be blockdiagonal, and the users do not cooperate in the uplink. Furthermore,  $\mathbf{\Sigma} \preceq \frac{P}{M} \mathbf{I} + *, \text{tr}(*) = 0$  implies a sum-power constraint of  $P$  of the uplink transmit covariances. Let  $\mathbf{\Sigma}_1, \dots, \mathbf{\Sigma}_U$  be the blocks on the blockdiagonal  $\mathbf{\Sigma}$ , then the uplink problem is

$$\max_{\Sigma \succeq 0} \left\{ \log \left| I + \frac{1}{\sigma^2} H^H \Sigma H \right| : \text{tr}(\Sigma) \leq P, \Sigma \text{ is blockdiagonal} \right\} \quad (74)$$

$$= \max_{\Sigma \succeq 0} \left\{ \log \left| I + \frac{1}{\sigma^2} \sum_{u \in \mathcal{U}} H_u^H \Sigma_u H_u \right| : \sum_{u \in \mathcal{U}} \text{tr}(\Sigma_u) \leq P \right\} \quad (75)$$

We can see that the value of the uplink problem is equal to the sum-capacity of a multiple access channel with a sum-power constraint. It is well known that for the same sum-power constraint the same sum-rate can be achieved in the broadcast channel [4,8,9], which proves the Sato bound to be tight for the sum-rate.

#### 4.1.2. Structure of the Worst Case Noise Covariance

In [9] it is shown that Equation (80) is a solution to the worst case downlink noise covariance problem Equation (70) if  $R$  is non-singular. Further, it is conjectured that it is also a solution for the singular case, but a proof is missing. By Lemma 4 we restate this result and provide an alternative proof that also holds for the singular case. The uplink problem Equation (74) can be reformulated as (For simplicity, but without loss of generality, we set  $\sigma^2 = 1$ )

$$\max_{D, \Sigma \succeq 0} \left\{ \log |D| : D = I + H^H \Sigma H, \text{tr}(\Sigma) \leq P, \Sigma \text{ is blockdiagonal} \right\} \quad (76)$$

Let  $A, \lambda, F$ , and  $M$  be Lagrangian multipliers for the constraints  $D = I + H^H \Sigma H, \text{tr}(\Sigma) \leq P, \Sigma$  is blockdiagonal, and  $\Sigma \succeq 0$ .

*Lemma 4.* Given a solution  $D', \Sigma'$  to the uplink problem with Lagrangian multipliers  $A', \lambda'$ , the remaining Lagrangian multipliers can be selected as

$$F' = H A' H^H - \begin{bmatrix} H_1 A' H_1^H & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & H_U A' H_U^H \end{bmatrix} \quad (77)$$

$$M' = \begin{bmatrix} \lambda' I & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda' I \end{bmatrix} - \begin{bmatrix} H_1 A' H_1^H & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & H_U A' H_U^H \end{bmatrix} \quad (78)$$

and a solution to the worst case downlink noise covariance problem Equation (70) is given by

$$R = I + \frac{1}{\lambda'} F' = \frac{1}{\lambda'} H A' H^H + \frac{1}{\lambda'} M' \quad (79)$$

$$= \begin{bmatrix} I & \frac{H_1 A' H_2^H}{\lambda'} & \dots & \frac{H_1 A' H_U^H}{\lambda'} \\ \frac{H_2 A' H_1^H}{\lambda'} & I & \dots & \frac{H_2 A' H_U^H}{\lambda'} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{H_U A' H_1^H}{\lambda'} & \frac{H_U A' H_2^H}{\lambda'} & \dots & I \end{bmatrix} \quad (80)$$

**Proof.** See Appendix E.  $\square$

#### 4.1.3. Worst Case Noise Capacity

The capacity of a point-to-point link with channel  $H \in \mathbb{C}^{M \times N}$  under a worst case noise assumption is given by

$$\max_{Q \succeq 0} \left\{ \min_{R \succeq 0} \left\{ \log \left| I + R^+ H Q H^H \right| : \text{tr}(R) \leq \sigma^2 \right\} : \text{tr}(Q) \leq P \right\} \quad (81)$$

In [17], it is shown that this worst case noise capacity is equivalent to the capacity without channel knowledge and white noise.

$$\log \left| I + \frac{P}{\sigma^2} H^H H \right| \quad (82)$$

While in [17] this is proven by matrix majorization theory, it can alternatively be reproduced by the minimax duality. Computation of Equation (81) means to solve a convex-concave game with individual convex constraint sets, for which max and min can be exchanged and we can instead use

$$\min_{R \succeq 0} \left\{ \max_{Q \succeq 0} \left\{ \log \left| I + R^+ H Q H^H \right| : \text{tr}(Q) \leq P \right\} : \text{tr}(R) \leq \sigma^2 \right\} \quad (83)$$

The sum-power constraints are modelled according to Section 2.1.2:

$$B = \frac{\sigma^2}{M} I, \mathcal{Y}^\perp = \{Y^\perp : \text{tr}(Y^\perp) = 0\}, \mathcal{Y} = \{0\} \quad (84)$$

and

$$C = \frac{P}{N} I, \mathcal{Z} = \{Z : \text{tr}(Z) = 0\}, \mathcal{Z}^\perp = \{0\} \quad (85)$$

With  $B \succ 0$  and  $C \succ 0$  conditions (C1)–(C4) hold true and using the asymmetric minimax duality by Lemma 3 the uplink problem is

$$\min_{\Omega \succeq 0} \left\{ \max_{\Sigma \succeq 0} \left\{ \log \left| I + \Omega^+ H^H \Sigma H \right| : \Sigma \preceq p \frac{\sigma^2}{M} I \right\} : \Omega \preceq n \frac{P}{N} I \right\} \quad (86)$$

With  $b = \text{tr}(BB) = \frac{\sigma^4}{M}$ ,  $c = \text{tr}(CC) = \frac{P^2}{N}$ , and  $\frac{c}{b} = \frac{MP^2}{N\sigma^4}$  we can select  $p = \frac{MP}{\sigma^2}$  and  $n = \frac{N\sigma^2}{P}$ , as  $\frac{p}{n} = \frac{MP^2}{N\sigma^4} = \frac{c}{b}$ , and we obtain

$$\min_{\Omega \succeq 0} \left\{ \max_{\Sigma \succeq 0} \left\{ \log \left| I + \Omega^+ H^H \Sigma H \right| : \Sigma \preceq P I \right\} : \Omega \preceq \sigma^2 I \right\} = \log \left| I + \frac{P}{\sigma^2} H^H H \right| \quad (87)$$

#### 4.1.4. Network Duality under (Generalized) Sum-Power Constraints

Here we regard the special case, where the noise covariances are fixed and appear as a generalized sum-power constraint in the dual network. Let

$$S_D(R, \Omega) = \max_{Q \succeq 0} \left\{ W(R, Q) : \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k Q_k) = P \right\} \quad (88)$$

and

$$S_U(\Omega, R) = \max_{\Sigma \succeq 0} \left\{ W(\Omega, \Sigma) : \sum_{k \in \mathcal{K}} \text{tr}(R_k \Sigma_k) = P \right\} \quad (89)$$

where we assume  $R, \Omega$ , and the channels such that the maxima exist. This scenario is for example considered in [12], where by using a SINR duality it is proven that  $S_D(R, \Omega) = S_U(\Omega, R)$  and the polite waterfilling result is established. In the following, we show how the network maximization problems under generalized sum-power constraints, Equations (88) and (89), can be formulated in our minimax framework. It can be verified that the conditions for existence of the maxima of Equations (88) and (89) imply conditions (C1)–(C4). For the downlink, the absence of a minimization over the noise covariance can be modelled, by selecting

$$B = R, \mathcal{Y}^\perp = \{0\}, \mathcal{Y} = \left\{ Y : \sum_{k \in \mathcal{K}} \text{tr}(R_k Y_k) = 0 \right\} \quad (90)$$

which means that  $\mathbf{R}$  is a solution for the downlink noise covariances, see Remark 2. The conic version of the generalized sum-power constraint is introduced in Section 2.1.2:

$$\mathbf{C} = \frac{P}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{\Omega}_k)} \mathbf{\Omega}, \mathcal{Z} = \left\{ \mathbf{Z} : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{Z}_k) = 0 \right\}, \mathcal{Z}^\perp = \{\mathbf{0}\} \quad (91)$$

Using the asymmetric mini-max duality by Lemma 3, the uplink problem is

$$\min_{\mathbf{\Omega}' \succeq \mathbf{0}, \mathcal{Z}^\perp} \left\{ \max_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathcal{Y}} \left\{ W(\mathbf{\Sigma}, \mathbf{\Omega}') : \mathbf{\Sigma} \preceq p\mathbf{R} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y} \right\} : \mathbf{\Omega}' \preceq n \frac{P}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{\Omega}_k)} \mathbf{\Omega} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (92)$$

With  $b = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k)$ ,  $c = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k) = \frac{P^2}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{\Omega}_k)}$ , and  $\frac{c}{b} = \frac{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k) \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{\Omega}_k)}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k) \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{\Omega}_k)}$  we can select  $p = \frac{P}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k)}$  and  $n = \frac{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{\Omega}_k)}{P}$ , as  $\frac{p}{n} = \frac{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k) \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{\Omega}_k)}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k) \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{\Omega}_k)} = \frac{c}{b}$ , and we obtain

$$\min_{\mathbf{\Omega}' \succeq \mathbf{0}, \mathcal{Z}^\perp} \left\{ \max_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathcal{Y}} \left\{ W(\mathbf{\Sigma}, \mathbf{\Omega}') : \mathbf{\Sigma} \preceq \frac{P}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k)} \mathbf{R} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y} \right\} : \mathbf{\Omega}' \preceq \mathbf{\Omega} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (93)$$

which is the minimax version of Equation (89).

#### 4.1.5. Duality of Broadcast and Multiple Access Channel

As in Section 4.1.1, the set of users is  $\mathcal{U}$ , with  $U = |\mathcal{U}|$ , and the channel of user  $u$  with  $M_u$  receive antennas is  $\mathbf{H}_u \in \mathbb{C}^{N \times M_u}$ , where  $N$  is the number of antennas at the broadcast transmitter. Broadcast and multiple access channels are modelled by considering one link per user and first setting  $\mathbf{H}_{u1} = \mathbf{H}_{u2} = \dots = \mathbf{H}_{uK} = \mathbf{H}_u \forall u \in \mathcal{K}$  and  $N_k = N \forall k \in \mathcal{K}$ . Second, if interference between a user  $u$  and user  $v$  is removed by non-linear interference cancellation, we set  $\mathbf{H}_{vu} = \mathbf{0}$  instead. This leads to the well known connection: given some ordering of the users, removing interference in the downlink such that a user does not interfere with a user of higher order matches the uplink model where interference is successively canceled in the inverse order. Thus, our minimax duality provides a minimax BC-MAC duality for any type of constraints that can be modeled as linear conic constraints. It holds for linear precoding as well as non-linear interference cancellation in an arbitrary order.

A particularly interesting case is where the decoding order of the uplink is matched to the weights of the users. If the decoding order is equal to the ascending order of the weights and additionally the noise covariance is the same for every user in the uplink, then the inner maximization of the uplink problem can be reformulated as a convex optimization problem. This renders the uplink problem a convex-concave game, for which efficient algorithms to find a solution are available. The following lemma describes the structure of the linear conic constraints and the subspaces that imply an equal uplink noise covariance for all users.

**Lemma 5.** Let the subspaces  $\mathcal{C}' = \{\lambda \mathbf{C}' : \lambda \in \mathbb{R}\}$ ,  $\mathcal{Z}'$ ,  $\mathcal{Z}'^\perp$  such that they are mutually orthogonal and their direct sum is  $\mathcal{S}^{N \times N}$ . If the constraint on the downlink transmit covariances is a constraint that solely acts on the sum of the transmit covariances, i.e.,

$$\sum_{k \in \mathcal{K}} \mathbf{Q}_k \preceq \mathbf{C}' + \mathbf{Z}', \mathbf{Z}' \in \mathcal{Z}' \quad (94)$$

then the constraint on the uplink noise covariances is

$$\mathbf{\Omega}_k \preceq \mathbf{C}' + \mathbf{Z}'^\perp, \mathbf{Z}'^\perp \in \mathcal{Z}'^\perp \forall k \in \mathcal{K} \quad (95)$$

which allows for a solution where all noise covariances are equal.

**Proof.** See Appendix F.  $\square$

**Lemma 6.** Consider a broadcast problem. If the encoding order is equal to the descending order of the user weights, the constraint on the downlink transmit covariances solely acts on the sum of the transmit covariances, and conditions (C1)–(C4) hold true, then every point that fulfills the optimality conditions is a solution.

**Proof.** The uplink problem can be reformulated as convex-concave game, where every point that fulfills the optimality conditions is a solution. By Lemma 11 every point that fulfills the downlink optimality condition also fulfills the uplink optimality conditions and therefore is a solution.  $\square$

Thus, for these scenarios we have the freedom to either solve for the uplink or downlink optimality conditions, whatever is more convenient.

**Lemma 7.** Consider a broadcast problem. If the encoding order is equal to the descending order of the user weights, the constraint on the downlink transmit covariances solely acts on the sum of the transmit covariances, and conditions (C1)–(C4) hold true, then

$$\min_{\mathbf{R} \succeq \mathbf{0}, \mathbf{Y}^\perp} \left\{ \max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z}'} \left\{ W(\mathbf{R}, \mathbf{Q}) : \sum_{k \in \mathcal{K}} \mathbf{Q}_k \preceq \mathbf{C}' + \mathbf{Z}', \mathbf{Z}' \in \mathcal{Z}' \right\} : \mathbf{R} \preceq \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (96)$$

$$= \max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z}'} \left\{ \min_{\mathbf{R} \succeq \mathbf{0}, \mathbf{Y}^\perp} \left\{ W(\mathbf{R}, \mathbf{Q}) : \mathbf{R} \preceq \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} : \sum_{k \in \mathcal{K}} \mathbf{Q}_k \preceq \mathbf{C}' + \mathbf{Z}', \mathbf{Z}' \in \mathcal{Z}' \right\} \quad (97)$$

**Proof.** By Lemma 6 every point that fulfills the optimality conditions of Equation (96) is a solution. The optimality conditions are the same for Equations (96) and (97) and therefore both problems have an equal optimal value.  $\square$

## 4.2. Novel Results

### 4.2.1. Interference Robust Multi-User MIMO

Our multi-user minimax duality enables new results concerning robustness to interference in MIMO networks, where uncertainty in the spatial signature of interfering signals is a major source of performance degradation. In the companion work [36], two possible techniques to provide robustness for interfering broadcast channels are discussed. Both decouple the optimization of the network into individual problems per basestation, one by considering the interference of the other cells by a worst-case approximation and the other method enforces fixed sum-transmit covariances at each transmitter, which removes the uncertainty in the interference.

For the worst-case approximation, the minimization of the noise plus other-cell interference covariance in the downlink is constrained by a noise power constraint for every user individually, which translates to scaled identity matrices as transmit covariances in the uplink, see Section 2.1.2. The sum-power constraint on the downlink transmit covariances leads to fixed uplink noise covariances, see Section 2.1.2. Thus, the uplink problem is a maximization of the scaling coefficients of the uplink transmit covariances under a constraint on the sum of the coefficients. For details see Appendix G. This result provides a generalization of the worst-case noise capacity from point-to-point channels [17] or sum-rate in the broadcast channel [18] to the complete capacity region of the MIMO broadcast channel.

The second approach constrains the sum of the downlink transmit covariances by a shaping constraint, which can be fulfilled with equality for some solution (If the shaping constraint is fulfilled with equality depends on the number of receive antennas and the rank of the channels. Equality can be artificially enforced by adding a virtual user whose interference to the other users is canceled by non-linear interference pre-subtraction. Whether the gain by increased robustness outweighs the pain by additional interference to the other cells depends on the scenario). Assuming the channels to be

static (for some time period) the constant transmit covariances lead to a constant inter-cell interference covariance at the users, which removes the uncertainty. The fixed noise covariances in the downlink translate to a sum power constraint on the uplink transmit covariances. The shaping constraint on the sum of the transmit covariances leads to equal uplink noise covariances for all users in the uplink, where the minimization is under a noise power constraint, see Section 4.1.5.

Both approaches can be modelled by using our minimax duality, see [36] for details. As for both approaches the constraints on the downlink transmit covariance are constraints that act solely on the sum of the transmit covariance, i.e. a sum-power constraint and a sum shaping constraint, the uplink problem has a convex-concave game reformulation and can be solved efficiently for non-linear interference (pre-)subtraction.

#### 4.2.2. Information Theoretic Proofs

##### Cooperation in Wireless Networks

The connection of uplink and downlink via orthogonal subspaces allows to model cooperation in wireless networks, as for example the cooperation of the receivers to compute the Sato bound. Reconsider the result of Section 4.1.1: By letting the receivers cooperate, the correlations among the noise symbols of different users are included into the optimization and can be freely chosen by allowing all degrees of freedom on the off-diagonal blocks of the joint noise covariance, which induces a linear subspace. As a result, the joint uplink covariances are in the orthogonal subspace, which results in a block-diagonal matrix and thus no cooperation of the transmitters. Without this cooperation, the uplink is equivalent to a multiple access channel and by the BC-MAC duality translates to a broadcast channel without cooperation. Thus, whenever the additional degrees of freedom introduced by cooperation can be modeled as a linear subspace, we can use the minimax duality to transfer the problem into the uplink, which then might reveal whether cooperation is beneficial. An example can be found in [33], where carrier-cooperation is modelled in our minimax duality.

##### Optimality of Proper Signaling

Results from information theory show that improper signals are beneficial for some communication scenarios. Recent results use our minimax duality to prove the optimality of proper signaling for the MIMO broadcast channel under a shaping constraint [26] and for the MIMO relay channel with partial decode-and-forward [34]. A key ingredient for the proof is the observation that in a composite real representation a covariance matrix can be partitioned in a power shaping component and an impropriety component, such that the components are elements of linear subspaces that are orthogonal complements. Using such a formulation in our minimax duality, the problem can be transferred to an uplink problem, where the worst case noise is proper. For the considered scenarios, it can be shown that there exist solutions for the uplink with proper transmit symbols. The translation back into a downlink solution preserves this property.

## 5. Proof of the Theorem

### 5.1. Conditions for Existence of a Solution

To derive conditions (C1)–(C4), we use the generalization of the constant nullspace Equation (18), and the positive semidefinite directions of a subspace Equation (19). For the downlink, the following conditions ensure that those directions where noise or transmit covariance can become arbitrarily large have no influence on the utility



$$\mathbf{H}_{kk}\mathbf{\Lambda}_k = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{P}(\mathcal{Z}) \quad (98)$$

$$\mathbf{\Lambda}_k\mathbf{H}_{kk} = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{P}(\mathcal{Y}^\perp) \quad (99)$$

The constraints on the transmit covariance can be used to prohibit certain spatial directions of the transmit signal, resulting in a non-trivial constant nullspace  $\mathcal{N}(\mathbf{C}, \mathcal{Z}) \neq \{\mathbf{0}\}$ . By a projection of the channel matrix we can artificially enforce

$$\mathbf{H}_{kk}\mathbf{\Lambda}_k = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{N}(\mathbf{C}, \mathcal{Z}) \quad (100)$$

without changing the solutions. By imposing this condition we are able to formulate the conditions for the inner optimization to have a finite value as a condition on the channels and independent of the constraints on the transmit covariance:

$$\mathbf{\Lambda}_k\mathbf{H}_{kk} = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{N}(\mathbf{B}, \mathcal{Y}^\perp) \quad (101)$$

Assume the condition does not hold, then there exists a  $\mathbf{\Lambda} \in \mathcal{N}(\mathbf{B}, \mathcal{Y}^\perp)$  with  $\mathbf{\Lambda}_k = \lambda_k \lambda_k^H \quad \forall k$  and  $\lambda_{k'}^H \mathbf{H}_{k'k'} \neq \mathbf{0}$  for some  $k'$ . Condition Equation (100) implies that for any  $\lambda_{k'}$  where  $\lambda_{k'}^H \mathbf{H}_{kk} \neq 0$  there exists a feasible  $\mathbf{Q}'$  where  $\lambda_{k'}^H \mathbf{H}_{k'k'} \mathbf{Q}' \mathbf{H}_{k'k'}^H \lambda_{k'} > 0$ . The inequality of the transmit covariance constraint allows us to remove any interference to  $k'$  by setting all other transmit covariances to zero and we have  $\mathbf{S}_{k'} = \mathbf{R}_{k'}$ . As  $\lambda_{k'}^H \mathbf{R}_k \lambda_{k'} = 0$  for all feasible  $\mathbf{R}_{k'}$  we can use  $\lambda_{k'}^H$  as a receive filter to obtain a noise free communication channel, and the inner optimization does not have a finite value. For the uplink, the equivalent conditions are

$$\mathbf{\Lambda}_k\mathbf{H}_{kk} = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{P}(\mathcal{Y}) \quad (102)$$

$$\mathbf{H}_{kk}\mathbf{\Lambda}_k = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{P}(\mathcal{Z}^\perp) \quad (103)$$

$$\mathbf{\Lambda}_k\mathbf{H}_{kk} = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{N}(\mathbf{B}, \mathcal{Y}) \quad (104)$$

$$\mathbf{H}_{kk}\mathbf{\Lambda}_k = \mathbf{0} \quad \forall k \in \mathcal{K}, \forall \mathbf{\Lambda} \in \mathcal{N}(\mathbf{C}, \mathcal{Z}^\perp) \quad (105)$$

Combining the conditions for downlink Equations (98)–(101) with the conditions for the uplink Equations (102)–(105) we can see that by Lemma 1 the conditions Equations (98), (99), (102) and (103) are superfluous. The remaining are conditions (C1)–(C4) stated in the Theorem. For  $\mathbf{B} \succ \mathbf{0}$  and  $\mathbf{C} \succ \mathbf{0}$  we have  $\mathcal{P}(\mathcal{Y}) = \mathcal{N}(\mathbf{B}, \mathcal{Y}) = \{\mathbf{0}\}$  and  $\mathcal{P}(\mathcal{Z}^\perp) = \mathcal{N}(\mathbf{C}, \mathcal{Z}^\perp) = \{\mathbf{0}\}$  and the conditions (C1)–(C4) hold true. Throughout the paper we assume (C1)–(C4) to hold true.

## 5.2. Properties of the Solutions – Optimality Conditions

Before we present the proof of the Theorem 1 in Section 5.4, we investigate some properties of the solutions and necessary optimality conditions of the minimax problems. A good candidate for optimality conditions are the KKT conditions, but as the utility fails to be differentiable everywhere we use the more general *Lagrange Multiplier Rule*, see for example [38, Th. 6.1.1]. The conditions are necessary, but not sufficient, for optimality.

By Remark 2, there is always a solution for the noise covariances where the linear conic constraint holds with equality. This allows us to actively enforce equality, which may reduce the set of solutions but does not affect the optimal value. For the proof, we will introduce  $\mathbf{N} \succeq \mathbf{0}$  and  $\mathbf{L} \succeq \mathbf{0}$  as slack variables for the constraints  $\mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}$  and  $\mathbf{\Sigma} \preceq \mathbf{B} + \mathbf{Y}$ , such that  $\mathbf{N} = \mathbf{C} + \mathbf{Z} - \mathbf{Q}$  and  $\mathbf{L} = \mathbf{B} + \mathbf{Y} - \mathbf{\Sigma}$ . Additionally, we will use scaled versions of the optimization problems, and we explicitly include the interference plus noise covariances into the optimization:

$$\tilde{C}_D(\alpha) = \min_{\substack{R \succeq 0 \\ Y^\perp \in \mathcal{Y}^\perp}} \left\{ \max_{\substack{S \\ Q \succeq 0 \\ Z \in \mathcal{Z}}} \left\{ U(S, Q) : S_k = R_k + \sum_{i \in \mathcal{K} \setminus k} H_{ki} Q_i H_{ki}^H \forall k, Q \preceq \alpha C + Z \right\} : R = \alpha B + Y^\perp \right\} \quad (106)$$

$$\tilde{C}_U(\beta) = \min_{\substack{\Omega \succeq 0 \\ Z^\perp \in \mathcal{Z}^\perp}} \left\{ \max_{\substack{\Psi \\ \Sigma \succeq 0 \\ Y \in \mathcal{Y}}} \left\{ U(\Psi, \Sigma) : \Psi_k = \Omega_k + \sum_{i \in \mathcal{K} \setminus k} H_{ik}^H \Sigma_i H_{ik} \forall k, \Sigma \preceq \beta B + Y \right\} : \Omega = \beta C + Z^\perp \right\} \quad (107)$$

Apart from the obvious equalities  $C_D = \tilde{C}_D(\alpha) \Big|_{\alpha=1}$  and  $C_U = \tilde{C}_U(\beta) \Big|_{\beta=1}$ , the relationship between the original problem and the scaled problem is expressed by the following lemma:

**Lemma 8.**

- (a) Given  $\alpha > 0$  and a point which fulfills the optimality conditions for  $\tilde{C}_D(\alpha)$ , a point with equal utility, which fulfills the optimality conditions for  $\alpha = 1$ , can be found by scaling the optimization variables by  $\frac{1}{\alpha}$  and the Lagrangian multipliers by  $\alpha$ .
- (b) Given  $\beta > 0$  and a point which fulfills the optimality conditions for  $\tilde{C}_U(\beta)$ , a point with equal utility, which fulfills the optimality conditions for  $\beta = 1$ , can be found by scaling the optimization variables by  $\frac{1}{\beta}$  and the Lagrangian multipliers by  $\beta$ .

**Proof.** We introduce Lagrangian multipliers as shown in Table 1. The Lagrangian multiplier for a subspace constraint is element of the orthogonal complement of the subspace. Therefore, the Lagrangian multiplier for  $Z \in \mathcal{Z}$  is element of  $\mathcal{C} \oplus \mathcal{Z}^\perp$ , which we parameterize by  $\nu C + Z^\perp, Z^\perp \in \mathcal{Z}^\perp$  in order to explicitly highlight the component in the direction of  $C$ . The same explanation holds for the Lagrangian multiplier  $\mu B + Y, Y \in \mathcal{Z}^\perp$  for the constraint  $Y \in \mathcal{Y}^\perp$ . The Lagrangian function for  $\tilde{C}_D(\alpha)$  can be written as

$$\begin{aligned} L_D = & \sum_{k \in \mathcal{K}} w_k \log \left| I + S_k^+ H_{kk} Q_k H_{kk}^H \right| + \sum_{k \in \mathcal{K}} \text{tr} \left( \Sigma_k (S_k - R_k - \sum_{i \in \mathcal{K} \setminus k} H_{ki} Q_i H_{ki}^H) \right) - \\ & - \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k (Q_k - \alpha C_k - Z_k)) - \sum_{k \in \mathcal{K}} \text{tr}((\nu C_k + Z_k^\perp) Z_k) + \sum_{k \in \mathcal{K}} \text{tr}(K_k Q_k) + \\ & + \sum_{k \in \mathcal{K}} \text{tr}(X_k (R_k - \alpha B_k - Y_k^\perp)) + \sum_{k \in \mathcal{K}} \text{tr}((\mu B_k + Y_k) Y_k^\perp) - \sum_{k \in \mathcal{K}} \text{tr}(L_k R_k) \end{aligned} \quad (108)$$

where the signs are selected such that  $\Omega, K, L$  are positive semi-definite. As

$$\text{tr}(\Sigma_k H_{ki} Q_i H_{ki}^H) = \text{tr}(H_{ki}^H \Sigma_k H_{ki} Q_i) \quad (109)$$

and by introducing the auxiliary variable  $\Psi$ , where

$$\Psi_k = \Omega_k + \sum_{i \in \mathcal{K} \setminus k} H_{ik}^H \Sigma_i H_{ik}, \quad \forall k \in \mathcal{K} \quad (110)$$

the Lagrangian function can be written as

$$\begin{aligned} L_D = & \sum_{k \in \mathcal{K}} w_k \log \left| I + S_k^+ H_{kk} Q_k H_{kk}^H \right| + \\ & + \sum_{k \in \mathcal{K}} \text{tr}(\Sigma_k S_k) + \sum_{k \in \mathcal{K}} \text{tr}((K_k - \Psi_k) Q_k) + \\ & + \sum_{k \in \mathcal{K}} \text{tr}((\Omega_k - \nu C_k - Z_k^\perp) Z_k) - \sum_{k \in \mathcal{K}} \text{tr}((X_k - \mu B_k - Y_k) Y_k^\perp) - \sum_{k \in \mathcal{K}} \text{tr}((\Sigma_k + L_k - X_k) R_k) + \\ & + \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k \alpha C_k) - \sum_{k \in \mathcal{K}} \text{tr}(X_k \alpha B_k) \end{aligned} \quad (111)$$

The optimality conditions for  $\tilde{C}_D(\alpha)$  are shown in Table 1. Equations (27)–(33) are the constraints on the optimization variables, Equations (34) and (35) are the properties of the Lagrangian multipliers due to the chosen parameterization, Equation (36) is the substitution introduced by Equation (110), Equations (37)–(39) are the complimentary slackness conditions for the inequality constraints, and Equations (40)–(44) are the stationarity conditions. The Lagrangian function is differentiable with respect to  $Z, Y^\perp, R, Q$ , which in this order gives us Equations (40)–(43). The utility is not differentiable with respect to  $S$  at a point where  $S$  is not invertible and we have to compute and parameterize the complete subdifferential to obtain Equation (44). This can be done by using [38, Th. 2.5.1], see Appendix H for details and the definition of the set  $\mathcal{E}_k^\Delta$ . As a remark,  $\mathcal{E}_k^\Delta = \{0\}$  if  $S_k$  invertible.

Let

$$\left( (R = \alpha B + Y^\perp, Q = \alpha C + Z - N, S), (\Omega = \nu C + Z^\perp, \Sigma = \mu B + Y - L, \Psi, K) \right) \quad (112)$$

be a solution of the scaled downlink, and substitute  $(R, Y^\perp, Q, Z, N, S)$  by  $\alpha(\check{R}, \check{Y}^\perp, \check{Q}, \check{Z}, \check{N}, \check{S})$  and  $(\Sigma, \Omega, \nu, Z^\perp, K, X, \mu, Y, L, \Psi)$  by  $\frac{1}{\alpha}(\check{\Sigma}, \check{\Omega}, \check{\nu}, \check{Z}^\perp, \check{K}, \check{X}, \check{\mu}, \check{Y}, \check{L}, \check{\Psi})$ . Thus, the scaling factor  $\alpha$  cancels out everywhere and

$$\left( (\check{R} = B + \check{Y}^\perp, \check{Q} = C + \check{Z} - \check{N}, \check{S}), (\check{\Omega} = \check{\nu} C + \check{Z}^\perp, \check{\Sigma} = \check{\mu} B + \check{Y} - \check{L}, \check{\Psi}, \check{K}) \right) \quad (113)$$

provides a point that fulfills the optimality conditions of the unscaled problem, where  $\alpha = 1$ . The utility Equation (11) is not changed by an equal scaling of interference-plus-noise covariance and transmit covariance. This proves (a).

The conversion between scaled and unscaled problem is

$$\begin{aligned} & \left( \frac{1}{\alpha} (R = \alpha B + Y^\perp, Q = \alpha C + Z - N, S), \alpha (\Omega = \nu C + Z^\perp, \Sigma = \mu B + Y - L, \Psi, K) \right) \\ &= \left( (\check{R} = B + \check{Y}^\perp, \check{Q} = C + \check{Z} - \check{N}, \check{S}), (\check{\Omega} = \check{\nu} C + \check{Z}^\perp, \check{\Sigma} = \check{\mu} B + \check{Y} - \check{L}, \check{\Psi}, \check{K}) \right) \end{aligned} \quad (114)$$

For the uplink, we use the substitution

$$S_k = R_k + \sum_{i \in \mathcal{K} \setminus k} H_{ki} Q_k H_{ki}^H \quad \forall k \in \mathcal{K} \quad (115)$$

and the Lagrangian function can be written as

$$\begin{aligned} L_U = & \sum_{k \in \mathcal{K}} w_k \log |I + \Psi_k^+ H_{kk}^H \Sigma_k H_{kk}| + \\ & + \sum_{k \in \mathcal{K}} \text{tr}(Q_k \Psi_k) + \sum_{k \in \mathcal{K}} \text{tr}((M_k - S_k) \Sigma_k) + \\ & + \sum_{k \in \mathcal{K}} \text{tr} \left( (R_k - \sigma B_k - Y_k^\perp) Y_k \right) - \sum_{k \in \mathcal{K}} \text{tr} \left( (W_k - \tau C_k - Z_k) Z_k^\perp \right) - \sum_{k \in \mathcal{K}} \text{tr}((Q_k + N_k - W_k) \Omega_k) + \\ & + \sum_{k \in \mathcal{K}} \text{tr}(R_k \beta B_k) - \sum_{k \in \mathcal{K}} \text{tr}(W_k \beta C_k) \end{aligned} \quad (116)$$

The explanation of the optimality conditions of the uplink, the definition of  $\mathcal{Q}_k^\Delta$ , and the proof of (b) follows by symmetry. The conversion between scaled and unscaled problem is

$$\begin{aligned} & \left( \frac{1}{\beta} (\Omega = \beta C + Z^\perp, \Sigma = \beta B + Y - L, \Psi), \beta (R = \sigma B + Y^\perp, Q = \tau C + Z - N, S, M) \right) \\ &= \left( (\hat{\Omega} = C + \hat{Z}^\perp, \hat{\Sigma} = B + \hat{Y} - \hat{L}, \hat{\Psi}), (\hat{R} = \hat{\sigma} B + \hat{Y}^\perp, \hat{Q} = \hat{\tau} C + \hat{Z} - \hat{N}, \hat{S}, \hat{M}) \right) \end{aligned} \quad (117)$$

□

The characterization of the solutions to optimization problems for MIMO interference networks by polite waterfilling is introduced in [12]. The next lemma states the polite waterfilling result for the special case of an isolated point-to-point link. Although a proof can be found in the original work [12], we decided to include an alternative proof that especially highlights scenarios with non-invertible noise covariances and the conditions for which these have a solution. Further, we present the proof in a way that is directly connected to our problem, which allows us to introduce an additional result as a corollary to the proof.

**Lemma 9.** Consider a MIMO link with channel matrix  $\mathbf{H}$  and non-zero noise covariance  $\mathbf{S}$  and a link with flipped channel  $\mathbf{H}^H$  and non-zero noise covariance  $\mathbf{\Psi}$ . The matrices  $\mathbf{T}, \mathbf{T}^\perp$  and  $\mathbf{\Phi}, \mathbf{\Phi}^\perp$  provide a full unitary basis for the fundamental subspaces of  $\mathbf{S}$  and  $\mathbf{\Psi}$  respectively.

If  $\mathbf{H}\mathbf{\Phi}^\perp = \mathbf{0}$  and  $\mathbf{T}^{\perp H}\mathbf{H} = \mathbf{0}$ , then the mutual information of both links is finite, the problems

$$\sup_{\mathbf{Q} \succeq \mathbf{0}} \left\{ w \log \left| \mathbf{I} + \mathbf{S}^+ \mathbf{H} \mathbf{Q} \mathbf{H}^H \right| - \text{tr}(\mathbf{\Psi} \mathbf{Q}) \right\} \quad (118)$$

and

$$\sup_{\mathbf{\Sigma} \succeq \mathbf{0}} \left\{ w \log \left| \mathbf{I} + \mathbf{\Psi}^+ \mathbf{H}^H \mathbf{\Sigma} \mathbf{H} \right| - \text{tr}(\mathbf{S} \mathbf{\Sigma}) \right\} \quad (119)$$

have a solution, and the following equalities hold for any pair of solutions:

- (a)  $\mathbf{\Phi}^H \mathbf{H}^H \mathbf{S}^+ \mathbf{H} \mathbf{Q} \mathbf{\Phi} = \mathbf{\Phi}^H \mathbf{H}^H \mathbf{\Sigma} \mathbf{H} \mathbf{\Psi}^+ \mathbf{\Phi}$
- (b)  $\mathbf{T}^H \mathbf{S}^+ \mathbf{H} \mathbf{Q} \mathbf{H}^H \mathbf{T} = \mathbf{T}^H \mathbf{\Sigma} \mathbf{H} \mathbf{\Psi}^+ \mathbf{H}^H \mathbf{T}$
- (c)  $\mathbf{\Sigma} \in \{ \mathbf{\Sigma} : \mathbf{\Sigma} = -w \left( \mathbf{S} + \mathbf{H} \mathbf{Q} \mathbf{H}^H \right)^+ + w \mathbf{S}^+ - \mathbf{\Sigma}^\Delta, \mathbf{\Sigma}^\Delta \in \mathcal{E}^\Delta \}$
- (d)  $\mathbf{Q} \in \{ \mathbf{Q} : \mathbf{Q} = -w \left( \mathbf{\Psi} + \mathbf{H}^H \mathbf{\Sigma} \mathbf{H} \right)^+ + w \mathbf{\Psi}^+ - \mathbf{Q}^\Delta, \mathbf{Q}^\Delta \in \mathcal{Q}^\Delta \}$
- (e)  $\log \left| \mathbf{I} + \mathbf{S}^+ \mathbf{H} \mathbf{Q} \mathbf{H}^H \right| = \log \left| \mathbf{I} + \mathbf{\Psi}^+ \mathbf{H}^H \mathbf{\Sigma} \mathbf{H} \right|$
- (f)  $\text{tr}(\mathbf{\Psi} \mathbf{Q}) = \text{tr}(\mathbf{S} \mathbf{\Sigma})$

where the definition of  $\mathcal{E}^\Delta$  and  $\mathcal{Q}^\Delta$  is in Lemma 12, presented in Appendix H.

We say that  $\mathbf{Q}, \mathbf{\Sigma}, \mathbf{S}, \mathbf{\Psi}$  are connected by polite waterfilling if they fulfill the equalities of the Lemma.

**Proof.** The condition for the downlink mutual information to be finite is  $\mathbf{T}^{\perp H} \mathbf{H} = \mathbf{0}$ , see Lemma 12 in Appendix A. If  $\mathbf{H}\mathbf{\Phi}^\perp \neq \mathbf{0}$ , there exists a  $\mathbf{q}$  with  $\mathbf{H}\mathbf{q} \neq \mathbf{0}$  and  $\text{tr}(\mathbf{\Phi} \mathbf{q} \mathbf{q}^H) = 0$  and for  $\mathbf{Q} = \alpha \mathbf{q} \mathbf{q}^H$  the utility becomes arbitrarily large for increasing  $\alpha$ . So, if  $\mathbf{H}\mathbf{\Phi}^\perp \neq \mathbf{0}$ , then Equation (118) does not have a solution. Conversely,  $\mathbf{H}\mathbf{\Phi}^\perp = \mathbf{0}$  implies  $\text{tr}(\mathbf{\Psi} \mathbf{Q}) > 0$  whenever  $\mathbf{H} \mathbf{Q} \neq \mathbf{0}$ , and as the logarithm grows slower than any linear function, the supremum is finite. By symmetry, the condition for a finite uplink mutual information is  $\mathbf{H}\mathbf{\Phi}^\perp = \mathbf{0}$  and the condition for Equation (119) to have a solution is  $\mathbf{T}^{\perp H} \mathbf{H} = \mathbf{0}$ .

Assume the conditions of the lemma hold true and consider the problem

$$\max_{\mathbf{S}' = \mathbf{S}} \left\{ \max_{\mathbf{Q} \succeq \mathbf{0}} \left\{ w \log \left| \mathbf{I} + \mathbf{S}'^+ \mathbf{H} \mathbf{Q} \mathbf{H}^H \right| - \text{tr}(\mathbf{\Psi} \mathbf{Q}) \right\} \right\} \quad (120)$$

which has the same solutions as Equation (118).

As the utility fails to be differentiable everywhere, the popular KKT conditions cannot be applied, and we use the more general *Lagrange Multiplier Rule*, see for example [38, Th. 6.1.1]. Let  $\mathbf{\Sigma}$  and  $\mathbf{K}$  be the Lagrangian multipliers for  $\mathbf{S}' = \mathbf{S}$  and  $\mathbf{Q} \succeq \mathbf{0}$ , the optimality conditions are

$$\mathbf{\Sigma} \in \{ \mathbf{\Sigma} : \mathbf{\Sigma} = -w \left( \mathbf{S} + \mathbf{H} \mathbf{Q} \mathbf{H}^H \right)^+ + w \mathbf{S}^+ + w \mathbf{\Sigma}^\Delta, \mathbf{\Sigma}^\Delta \in \mathcal{E}^\Delta \} \quad (121)$$

$$\mathbf{\Psi} = w \mathbf{H}^H \left( \mathbf{S} + \mathbf{H} \mathbf{Q} \mathbf{H}^H \right)^+ \mathbf{H} + \mathbf{K} \quad (122)$$

$$\text{tr}(\mathbf{K} \mathbf{Q}) = 0, \mathbf{K} \succeq \mathbf{0}, \mathbf{Q} \succeq \mathbf{0} \quad (123)$$

where  $\mathcal{E}^\Delta$  is given by Lemma 13 in Appendix H. Furthermore, consider the following problem

$$\max_{\Psi' = \Psi} \left\{ \max_{\Sigma \succeq 0} \left\{ w \log |I + \Psi'^+ H^H \Sigma H| - \text{tr}(\Sigma \Psi) \right\} \right\} \quad (124)$$

which has the same solutions as Equation (119). Let  $Q$  and  $M$  be the Lagrangian multipliers for  $\Psi' = \Psi$  and  $\Sigma \succeq 0$ . The optimality conditions are

$$Q \in \{Q: Q = -w(\Psi + H^H \Sigma H)^+ + w\Psi^+ + wQ^\Delta, Q^\Delta \in \mathcal{Q}^\Delta\} \quad (125)$$

$$S = wH(\Psi + H^H \Sigma H)^+ H^H + M \quad (126)$$

$$\text{tr}(M\Sigma) = 0, M \succeq 0, \Sigma \succeq 0 \quad (127)$$

where  $\mathcal{Q}^\Delta$  is defined analogously to  $\mathcal{E}^\Delta$  given by Lemma 13 in Appendix H. By standard calculus, it can be shown that every solution to Equations (121)–(123) is a solution to Equations (125)–(127) and vice versa, and that these solutions fulfil the equalities of the Lemma. See Appendix I for details.  $\square$

*Corollary 1.* For both  $\tilde{C}_D(\alpha)$  and  $\tilde{C}_U(\beta)$ , if a point fulfills the optimality conditions, then for every link  $k \in \mathcal{K}$ ,  $Q_k, \Sigma_k, S_k, \Psi_k$  are connected by polite waterfilling.

**Proof.** Consider the downlink and notice that Equations (121)–(123) are fulfilled by Equations (30), (34), (38), (43), and (44), meaning that  $Q_k, \Sigma_k, S_k, \Psi_k$  are connected by polite waterfilling if the conditions for the Lemma, these are  $H_{kk}\Phi_k^\perp = 0$  and  $T_k^\perp H_{kk} = 0$ , hold. Consider the downlink. By conditions (C1)–(C4) we know that unless  $T_k^\perp H_{kk} = 0$  the mutual information is infinite, and this point does not fulfill the optimality conditions. By multiplying Equation (43) with  $\Phi_k^{\perp H}$  from the left and  $\Phi_k^\perp$  from the right we obtain

$$0 = \Phi_k^{\perp H} H_{kk}^H (S_k + H_{kk} Q_k H_{kk}^H)^+ H_{kk} \Phi_k^\perp + \Phi_k^{\perp H} K_k \Phi_k^\perp \quad (128)$$

$$= \Phi_k^{\perp H} H_{kk}^H T_k (T_k^H S_k T_k + T_k^H H_{kk} Q_k H_{kk}^H T_k)^{-1} T_k^H H_{kk} \Phi_k^\perp + \Phi_k^{\perp H} K_k \Phi_k^\perp \quad (129)$$

As both summands are positive-semidefinite matrices, they must both be equal to 0, and with  $(T_k^H S_k T_k + T_k^H H_{kk} Q_k H_{kk}^H T_k)^{-1} \succ 0$  we must have  $T_k^H H_{kk} \Phi_k^\perp = 0$ . As  $T_k^\perp H_{kk} = 0$  and  $[T_k T_k^\perp]$  full rank, we have  $[T_k T_k^\perp]^H H_{kk} \Phi_k^\perp = 0$ , and  $H_{kk} \Phi_k^\perp = 0$  follows. The proof for the uplink follows by symmetry.  $\square$

*Lemma 10.* Independent of the scaling factors  $\alpha$  and  $\beta$ , the following holds:

(a) for every point that fulfills the downlink optimality conditions, we have

$$\mu \sum_{k \in \mathcal{K}} \text{tr}(B_k B_k) = \nu \sum_{k \in \mathcal{K}} \text{tr}(C_k C_k) \quad (130)$$

(b) for every point that fulfills the uplink optimality conditions, we have

$$\sigma \sum_{k \in \mathcal{K}} \text{tr}(B_k B_k) = \tau \sum_{k \in \mathcal{K}} \text{tr}(C_k C_k) \quad (131)$$

**Proof.** The mutual orthogonality of the linear subspace triples  $\mathcal{B}, \mathcal{Y}, \mathcal{Y}^\perp$  and  $\mathcal{C}, \mathcal{Z}, \mathcal{Z}^\perp$ , respectively, implies

$$\sum_{k \in \mathcal{K}} \text{tr}(B_k Y_k) = \sum_{k \in \mathcal{K}} \text{tr}(B_k Y_k^\perp) = \sum_{k \in \mathcal{K}} \text{tr}(Y_k Y_k^\perp) = 0 \quad (132)$$

and

$$\sum_{k \in \mathcal{K}} \text{tr}(C_k Z_k) = \sum_{k \in \mathcal{K}} \text{tr}(C_k Z_k^\perp) = \sum_{k \in \mathcal{K}} \text{tr}(Z_k^\perp C_k) = 0 \quad (133)$$

respectively. Consider the downlink. By combining Equations (37) and (40), we have

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{Q}_k) = \alpha \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{C}_k) \quad (134)$$

$$\Leftrightarrow \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{Q}_k) = \alpha \sum_{k \in \mathcal{K}} \text{tr}((\nu \mathbf{C}_k + \mathbf{Z}_k^\perp) \mathbf{C}_k) \quad (135)$$

$$\Leftrightarrow \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{Q}_k) = \alpha \nu \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k) \quad (136)$$

Multiplying Equation (31) with  $\mathbf{B}$  from the right, applying the trace operator on both sides, and summing over all links, we obtain

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{B}_k) = \alpha \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) + \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{Y}_k^\perp \mathbf{B}_k) \quad (137)$$

$$= \alpha \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) \quad (138)$$

Further, combining Equations (41) and (42) we have  $\mathbf{\Sigma} = \mu \mathbf{B} + \mathbf{Y} - \mathbf{L}$  and multiplying with  $\mathbf{R}$  from the left, applying the trace operator operator on both sides, and summing over all links gives us

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{\Sigma}_k) = \mu \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{B}_k) + \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{Y}_k^\perp) - \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{L}_k) \quad (139)$$

$$= \mu \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{B}_k) \quad (140)$$

where we used  $\text{tr}(\mathbf{R}_k \mathbf{L}_k) = 0$  from Equation (39). By combining Equations (140) and (138) it holds:

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{\Sigma}_k) = \alpha \mu \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) \quad (141)$$

Multiplying Equation (27) by  $\mathbf{\Sigma}_k$  from the right and Equation (43) by  $\mathbf{Q}_k$  from the right, applying the trace operator on both sides, and summing over all links gives us

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{S}_k \mathbf{\Sigma}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{\Sigma}_k) + \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{K} \setminus k} \text{tr}(\mathbf{H}_{ki} \mathbf{Q}_i \mathbf{H}_{ki}^\text{H} \mathbf{\Sigma}_k) \quad (142)$$

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Psi}_k \mathbf{Q}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{Q}_k) + \sum_{k \in \mathcal{K}} \sum_{i \in \mathcal{K} \setminus k} \text{tr}(\mathbf{H}_{ik}^\text{H} \mathbf{\Sigma}_i \mathbf{H}_{ik} \mathbf{Q}_k) \quad (143)$$

With the trace property Equation (109) and the equivalence of the penalty terms from the polite waterfilling property Lemma 9 Equation (9), we obtain

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{\Sigma}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{Q}_k) \quad (144)$$

Combining this with Equations (136) and (141), we have

$$\alpha \mu \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \alpha \nu \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k) \quad (145)$$

The proof for the uplink follows by symmetry.  $\square$

**Lemma 11.** If  $\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k)$ , then downlink and uplink problem share the following optimality conditions:

$$\mathbf{S}_k = \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ki} \mathbf{Q}_i \mathbf{H}_{ki}^H \quad \forall k \quad (146)$$

$$\mathbf{\Psi}_k = \mathbf{\Omega}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ik}^H \mathbf{\Sigma}_i \mathbf{H}_{ik} \quad \forall k \quad (147)$$

$$\mathbf{Q} = \alpha \mathbf{C} + \mathbf{Z} - \mathbf{N} \quad (148)$$

$$\mathbf{\Sigma} = \beta \mathbf{B} + \mathbf{Y} - \mathbf{L} \quad (149)$$

$$\mathbf{R} = \alpha \mathbf{B} + \mathbf{Y}^\perp \quad (150)$$

$$\mathbf{\Omega} = \beta \mathbf{C} + \mathbf{Z}^\perp \quad (151)$$

$$\text{tr}(\mathbf{K}_k \mathbf{Q}_k) = 0 \quad \forall k \quad (152)$$

$$\text{tr}(\mathbf{L}_k \mathbf{R}_k) = 0 \quad \forall k \quad (153)$$

$$\text{tr}(\mathbf{M}_k \mathbf{\Sigma}_k) = 0 \quad \forall k \quad (154)$$

$$\text{tr}(\mathbf{N}_k \mathbf{\Omega}_k) = 0 \quad \forall k \quad (155)$$

$$\mathbf{S}_k = w_k \mathbf{H}_{kk} \left( \mathbf{\Psi}_k + \mathbf{H}_{kk}^H \mathbf{\Sigma}_k \mathbf{H}_{kk} \right)^+ \mathbf{H}_{kk}^H + \mathbf{M}_k \quad \forall k \quad (156)$$

$$\mathbf{\Psi}_k = w_k \mathbf{H}_{kk}^H \left( \mathbf{S}_k + \mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \right)^+ \mathbf{H}_{kk} + \mathbf{K}_k \quad \forall k \quad (157)$$

$$\mathbf{Q}_k \in \{ \mathbf{Q}_k : \mathbf{Q}_k = -w_k \left( \mathbf{\Psi}_k + \mathbf{H}_{kk}^H \mathbf{\Sigma}_k \mathbf{H}_{kk} \right)^+ + w_k \mathbf{\Psi}_k^+ - w_k \mathbf{Q}_k^\Delta, \mathbf{Q}_k^\Delta \in \mathcal{Q}_k^\Delta \} \quad \forall k \quad (158)$$

$$\mathbf{\Sigma}_k \in \{ \mathbf{\Sigma}_k : \mathbf{\Sigma}_k = -w_k \left( \mathbf{S}_k + \mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \right)^+ + w_k \mathbf{S}_k^+ - w_k \mathbf{\Sigma}_k^\Delta, \mathbf{\Sigma}_k^\Delta \in \mathcal{E}_k^\Delta \} \quad \forall k \quad (159)$$

$$\mathbf{R} \succeq \mathbf{0}, \mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Q} \succeq \mathbf{0}, \mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{S} \succeq \mathbf{0}, \mathbf{\Psi} \succeq \mathbf{0}, \mathbf{K} \succeq \mathbf{0}, \mathbf{M} \succeq \mathbf{0}, \mathbf{L} \succeq \mathbf{0}, \mathbf{N} \succeq \mathbf{0} \quad (160)$$

$$\mathbf{Z} \in \mathcal{Z}, \mathbf{Y} \in \mathcal{Y}, \mathbf{Y}^\perp \in \mathcal{Y}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \quad (161)$$

where  $\mathbf{T}_k, \mathbf{T}_k^\perp$  and  $\mathbf{\Phi}_k, \mathbf{\Phi}_k^\perp$  provide a full unitary basis for the fundamental subspaces of  $\mathbf{S}_k$  and  $\mathbf{\Psi}_k$ , respectively. Specifically, the following relationships hold:

- (a) given  $\alpha > 0, \beta > 0$ , and a point that fulfills the shared conditions, this point fulfills Equations (27)–(44) for  $\nu = \mu = \beta$ , and Equations (45)–(62) for  $\tau = \sigma = \alpha$ .
- (b) given a scaling factor  $\alpha > 0$  any point that fulfills the downlink optimality conditions Equations (27)–(44) also fulfills the shared conditions with  $\beta = \nu = \mu$ ,
- (c) given a scaling factor  $\beta > 0$ , any point that fulfills the uplink optimality conditions Equations (45)–(62) also fulfills the shared conditions with  $\alpha = \tau = \sigma$ .

**Proof.** First, note that if  $\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k)$ , then Lemma 10 implies  $\mu = \nu$  and  $\sigma = \tau$ . To prove (a) for the downlink, we show that all conditions in Equations (27)–(44) that are not explicitly stated in the shared conditions hold for any point that fulfills the shared conditions. For the downlink this is condition Equation (37), which holds by combining Equations (148) and (155), and conditions Equations (41) and (42), which are equivalent to Equation (149) by eliminating the superfluous variable  $\mathbf{X}$ . The proof for the uplink follows by symmetry.

To prove (b), we show that all shared conditions that are not explicitly stated in Equations (27)–(44) hold for any point that fulfills Equations (27)–(44). Conditions Equations (148) and (155) hold by Equation (37) and  $\mathbf{N} = \alpha \mathbf{C} + \mathbf{Z} - \mathbf{Q}$  as  $\mathbf{N}$  is the slack variable for Equation (148).  $\mathbf{N} \succeq \mathbf{0}$  in Equation (160) is by definition of  $\mathbf{N}$ . The conditions Equations (154), (156), and (158), hold true by the polite waterfilling connection due to Corollary 1. The proof for (c) follows from the proof for (b) by symmetry.  $\square$



### 5.3. Proof of the Network Duality for Fixed Noise and (Generalized) Sum-Power Constraint

Before we prove  $C_D = C_U$  for the general minimax problems, we discuss the special case where the noise covariances are fixed and appear as a generalized sum-power constraint in the dual network. This is the scenario from [12] introduced in Section 4.1.4, where the functions  $S_D(\mathbf{R}, \mathbf{\Omega})$  and  $S_U(\mathbf{\Omega}, \mathbf{R})$  are defined by Equations (88) and (89) respectively. As we have used the polite waterfilling result from [12] to prove Lemma 11, it is not surprising that we can in turn use Lemma 11 to prove  $S_D(\mathbf{R}, \mathbf{\Omega}) = S_U(\mathbf{\Omega}, \mathbf{R})$ . To do so, we have to reformulate  $S_D(\mathbf{R}, \mathbf{\Omega})$  and  $S_U(\mathbf{\Omega}, \mathbf{R})$  in the minimax framework, see Section 4.1.4 for details. By Lemma 11 combined with Corollary 1, every solution to Equation (88) provides a point with equal utility that fulfills the optimality conditions of Equation (89) and vice versa. Assume there is a solution to Equation (88) and the point is not a solution to Equation (89), then there exists a point with higher utility in Equation (89). As this point needs to fulfill the optimality conditions, there must be a point with equal utility in Equation (88), this contradicts with the original point being a solution, and therefore  $S_D(\mathbf{R}, \mathbf{\Omega}) = S_U(\mathbf{\Omega}, \mathbf{R})$ .

### 5.4. Proof of the Minimax Duality

Unfortunately, the argumentation of the previous section cannot be extended to minimax problems, which might have multiple points fulfilling the optimality conditions with different utilities. In addition to the solutions, there might exist points which fulfill the optimality conditions with higher or lower utility than the optimal value. To prove  $C_D = C_U$  for general minimax problems, we first show that both problems share an upper bound, and second, we show that this upper bound is tight. Consider a reformulation of the downlink problem as

$$C_D = \min_{\mathbf{R} \succeq \mathbf{0}, \mathbf{Y}^\perp} \left\{ \max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z} \in \mathcal{Z}} \{W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z}\} : \mathbf{R} \preceq \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (162)$$

$$= \min_{\mathbf{R} \succeq \mathbf{0}, \mathbf{Y}^\perp} \left\{ F_D(\mathbf{R}) : \mathbf{R} \preceq \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (163)$$

where

$$F_D(\mathbf{R}) = \max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z} \in \mathcal{Z}} \{W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}\} \quad (164)$$

We upper bound the inner, potentially non-convex, maximization from above with a Lagrangian dual problem:

$$F_D(\mathbf{R}) \leq \inf_{\mathbf{\Omega} \succeq \mathbf{0}} \sup_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z} \in \mathcal{Z}} \left\{ W(\mathbf{R}, \mathbf{Q}) - \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k(\mathbf{Q}_k - \mathbf{C}_k - \mathbf{Z}_k)) \right\} \quad (165)$$

where we assumed  $\mathbf{R}$  such that the maximum of Equation (164) exists.

The infimum does not change if we remove those  $\mathbf{\Omega}$  where the inner problem does not have a solution. As  $\mathbf{Z}$  is only constrained by  $\mathbf{Z} \in \mathcal{Z}$ , the inner problem of Equation (165) does not have a solution unless  $\mathbf{\Omega}$  is in the orthogonal complement of  $\mathcal{Z}$ , which we can parameterize by  $\mathbf{\Omega} = \nu \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp$ .

Thus, we can formulate the upper bound as

$$F_D(\mathbf{R}) \leq \inf_{\nu, \mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \sup_{\mathbf{Q} \succeq \mathbf{0}} \left\{ W(\mathbf{R}, \mathbf{Q}) - \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{Q}_k) + \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{C}_k) \right\} : \mathbf{\Omega} = \nu \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (166)$$

Furthermore, by Corollary 1, we know there exists a solution to  $C_D$  where optimization variables and Lagrangian multipliers are connected by polite waterfilling. Using such a solution for  $\mathbf{\Omega}$  to evaluate the inner problem of Equation (166),  $\mathbf{\Omega}$  and  $\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Omega}_k \mathbf{C}_k)$  are constants for the inner problem. As we assumed that  $\mathbf{R}$  is such that the mutual information of all links is finite for finite

transmit symbols, the inner problem has a solution. This implies that the infimum and supremum of Equation (166) are attained.

For every minimizing dual variable, the following conditions can be obtained from the complementary slackness conditions:

$$\sum_{k \in \mathcal{K}} \text{tr}(\Omega_k(Q_k - C_k - Z_k)) = 0 \quad (167)$$

$$\Leftrightarrow \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k Q_k) = \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k C_k) \quad (168)$$

since  $\sum_{k \in \mathcal{K}} \text{tr}(\nu C_k + Z_k^\perp) Z_k = 0$ . The upper bound can be reformulated as

$$F_D(\mathbf{R}) \leq \min_{\Omega \succeq 0, Z^\perp} \left\{ \max_{Q \succeq 0} \left\{ W(\mathbf{R}, Q) : \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k Q_k) = \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k C_k) \right\} : \Omega = \nu \mathbf{C} + Z^\perp, Z^\perp \in \mathcal{Z}^\perp \right\} \quad (169)$$

The value of the inner problem is the same if  $\Omega$  is scaled (leading to an inverse scaling of the optimization variables). This allows us to fix the scaling by picking a  $P$  with

$$P = \sum_{k \in \mathcal{K}} \text{tr}(C_k C_k) \quad (170)$$

and scale  $\Omega$  such that

$$\sum_{k \in \mathcal{K}} \text{tr}(\Omega_k Q_k) = \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k C_k) = P \quad (171)$$

Combining Equations (170), (171), and  $\Omega = \nu \mathbf{C} + Z^\perp, Z^\perp \in \mathcal{Z}^\perp$ , we obtain  $\nu = 1$ , and the upper bound is

$$F_D(\mathbf{R}) \leq \min_{\Omega \succeq 0, Z^\perp} \left\{ \max_{Q \succeq 0} \left\{ W(\mathbf{R}, Q) : \sum_{k \in \mathcal{K}} \text{tr}(\Omega_k Q_k) = P \right\} : \Omega = \mathbf{C} + Z^\perp, Z^\perp \in \mathcal{Z}^\perp \right\} \quad (172)$$

$$= \min_{\Omega \succeq 0, Z^\perp} \left\{ S_D(\mathbf{R}, \Omega) : \Omega = \mathbf{C} + Z^\perp, Z^\perp \in \mathcal{Z}^\perp \right\} \quad (173)$$

where  $S_D(\mathbf{R}, \Omega)$  is defined by Equation (88). Thus, we can formulate an upper bound on  $C_D$  by

$$C_D \leq \min_{R \succeq 0, Y^\perp} \left\{ \min_{\Omega \succeq 0, Z^\perp} \left\{ S_D(\mathbf{R}, \Omega) : \Omega = \mathbf{C} + Z^\perp, Z^\perp \in \mathcal{Z}^\perp \right\} : \mathbf{R} = \mathbf{B} + Y^\perp, Y^\perp \in \mathcal{Y}^\perp \right\} \quad (174)$$

$$= \min_{R \succeq 0, Y^\perp, \Omega \succeq 0, Z^\perp} \left\{ S_D(\mathbf{R}, \Omega) : \Omega = \mathbf{C} + Z^\perp, Z^\perp \in \mathcal{Z}^\perp, \mathbf{R} = \mathbf{B} + Y^\perp, Y^\perp \in \mathcal{Y}^\perp \right\} \quad (175)$$

Now consider the uplink problem. By using  $\mathbf{R} = \tau \mathbf{B} + Y^\perp, Y^\perp \in \mathcal{Y}^\perp$  as Lagrangian multiplier to relax the constraint on the uplink transmit covariances and selecting

$$P = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \Sigma_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{B}_k) \quad (176)$$

we can formulate an upper bound for the uplink as

$$C_U \leq \min_{R \succeq 0, Y^\perp, \Omega \succeq 0, Z^\perp} \left\{ S_U(\Omega, \mathbf{R}) : \Omega = \mathbf{C} + Z^\perp, Z^\perp \in \mathcal{Z}^\perp, \mathbf{R} = \mathbf{B} + Y^\perp, Y^\perp \in \mathcal{Y}^\perp \right\} \quad (177)$$

where  $S_U(\Omega, \mathbf{R})$  is defined by Equation (89). We have  $P = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k)$  and as  $S_D(\mathbf{R}, \Omega) = S_U(\Omega, \mathbf{R})$ , see Section 5.3, the bounds Equations (175) and (177) have the same value, and the problems  $C_D$  and  $C_U$  share an upper bound.

We can prove  $C_D = C_U$  by showing that the upper bound is tight. The bound by the Lagrangian dual problem on the inner problem of the downlink can be reformulated as an uplink minimax problem:

$$F_D(\mathbf{R}) = \max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z} \in \mathcal{Z}} \{W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}\} \quad (178)$$

$$\leq \min_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ S_D(\mathbf{R}, \mathbf{\Omega}) : \mathbf{\Omega} = \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (179)$$

$$= \min_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ S_U(\mathbf{\Omega}, \mathbf{R}) : \mathbf{\Omega} = \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (180)$$

$$= \min_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \max_{\mathbf{\Sigma} \succeq \mathbf{0}} \left\{ W(\mathbf{\Omega}, \mathbf{\Sigma}) : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{\Sigma}_k) = P \right\} : \mathbf{\Omega} = \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (181)$$

$$= \min_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \max_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \{W(\mathbf{\Omega}, \mathbf{\Sigma}) : \mathbf{\Sigma} \preceq \bar{\mathbf{B}} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y}\} : \mathbf{\Omega} \preceq \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (182)$$

$$= \min_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \max_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \{W(\mathbf{\Omega}, \mathbf{\Sigma}) : \mathbf{\Sigma} \preceq \bar{\beta} \bar{\mathbf{B}} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y}\} : \mathbf{\Omega} \preceq \bar{\beta} \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (183)$$

where we used  $S_D(\mathbf{R}, \mathbf{\Omega}) = S_U(\mathbf{\Omega}, \mathbf{R})$  and replaced the generalized sum-power constraint  $\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{\Sigma}_k)$  by a conic constraint with

$$\bar{\mathbf{B}} = \frac{P}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k)} \mathbf{R} \quad (184)$$

and

$$\mathcal{Y} = \{\mathbf{Y} : \text{tr}(\bar{\mathbf{B}}_k \mathbf{Y}_k) = 0\} \quad (185)$$

which implies  $\mathcal{Y}^\perp = \{\mathbf{0}\}$ , see Section 2.1.2 for an explanation. Equivalence of Equations (182) and (183) is given by Lemma 8.

By Lemma 11, every solution of Equation (183) fulfills the conditions Equations (45)–(62) and provides Lagrangian multipliers  $\bar{\mathbf{S}}, \bar{\mathbf{R}}, \bar{\mathbf{Q}}, \bar{\mathbf{Z}}, \bar{\mathbf{N}}, \bar{\tau}$  with

$$\bar{\mathbf{Q}} = \bar{\tau} \mathbf{C} + \bar{\mathbf{Z}} - \bar{\mathbf{N}}, \bar{\mathbf{Z}} \in \mathcal{Z}, \bar{\mathbf{Q}} \succeq \mathbf{0}, \bar{\mathbf{N}} \succeq \mathbf{0} \quad (186)$$

$$\Rightarrow \bar{\mathbf{Q}} \preceq \bar{\tau} \mathbf{C} + \bar{\mathbf{Z}}, \bar{\mathbf{Z}} \in \mathcal{Z}, \bar{\mathbf{Q}} \succeq \mathbf{0} \quad (187)$$

and as  $\mathcal{Y}^\perp = \{\mathbf{0}\}$ , we have

$$\bar{\mathbf{S}}_k = \bar{\mathbf{R}}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ki} \bar{\mathbf{Q}}_i \mathbf{H}_{ki}^H \forall k \quad (188)$$

$$\bar{\mathbf{R}} = \bar{\sigma} \bar{\mathbf{B}} \quad (189)$$

Lemma 10 holds true even if  $\sum_{k \in \mathcal{K}} \text{tr}(\bar{\mathbf{B}}_k \bar{\mathbf{B}}_k) \neq \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k)$ , and by Equation (131), we have

$$\bar{\sigma} = \bar{\tau} \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k) \frac{1}{\sum_{k \in \mathcal{K}} \text{tr}(\bar{\mathbf{B}}_k \bar{\mathbf{B}}_k)} \quad (190)$$

$$= \bar{\tau} P \frac{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k)}{P^2} \quad (191)$$

$$= \bar{\tau} \frac{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k)}{P} \quad (192)$$

where we used  $\sum_{k \in \mathcal{K}} \text{tr}(\bar{\mathbf{B}}_k \bar{\mathbf{B}}_k) = \frac{P^2}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k)}$ , which can be derived from Equation (184). Combined with Equation (189), we obtain

$$\bar{\mathbf{R}} = \bar{\sigma} \frac{P}{\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{R}_k \mathbf{R}_k)} \mathbf{R} = \bar{\tau} \mathbf{R}. \quad (193)$$

By Lemma 8(b),  $\bar{\tau}$  scales inversely with  $\bar{\beta}$  without affecting the utility, so we are free to pick  $\bar{\beta}$  such that  $\bar{\tau} = 1$ . Given such a solution to Equation (183), the Lagrangian multipliers  $\bar{\mathbf{S}}, \bar{\mathbf{Q}}, \bar{\mathbf{Z}}$  fulfill Equations (187) and (188). By using  $\bar{\tau} = 1$  and  $\bar{\mathbf{R}} = \mathbf{R}$ , Equations (187) and (188) can be written as

$$\bar{\mathbf{Q}} \preceq \mathbf{C} + \bar{\mathbf{Z}}, \bar{\mathbf{Z}} \in \mathcal{Z}, \bar{\mathbf{Q}} \succeq \mathbf{0} \quad (194)$$

$$\bar{\mathbf{S}}_k = \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ki} \bar{\mathbf{Q}}_i \mathbf{H}_{ki}^H \quad \forall k \quad (195)$$

Directly including the explicit parameterization of the interference plus noise covariance in Equation (178), we have

$$F_D(\mathbf{R}) = \max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{S}, \mathbf{Z} \in \mathcal{Z}} \left\{ U(\mathbf{S}, \mathbf{Q}) : \mathbf{S}_k = \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ki} \mathbf{Q}_i \mathbf{H}_{ki}^H \quad \forall k \in \mathcal{K}, \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z} \right\} \quad (196)$$

and we can see that  $\bar{\mathbf{Q}}, \bar{\mathbf{Z}}$  are a feasible solution to Equation (178) and have the same utility by the polite waterfilling connection, Lemma 9(e). Thus, Equation (165) holds with equality and the upper bound Equation (175) is tight for the downlink. Tightness of the upper bound for the uplink follows by symmetry and we have proven  $C_D = C_U$ .

By Lemma 10, we have  $\hat{\sigma} = \hat{\tau}$ , and by introducing  $\hat{\gamma} = \hat{\sigma} = \hat{\tau}$ , a solution to the unscaled uplink problem can be stated as

$$\left( (\hat{\mathbf{\Omega}} = \mathbf{C} + \hat{\mathbf{Z}}^\perp, \hat{\mathbf{\Sigma}} = \mathbf{B} + \hat{\mathbf{Y}} - \hat{\mathbf{L}}, \hat{\mathbf{\Psi}}), (\hat{\mathbf{R}} = \hat{\gamma} \mathbf{B} + \hat{\mathbf{Y}}^\perp, \hat{\mathbf{Q}} = \hat{\gamma} \mathbf{C} + \hat{\mathbf{Z}} - \hat{\mathbf{N}}, \hat{\mathbf{S}}) \right) \quad (197)$$

For any  $\beta$  and the conversion rule Equation (117), introduced with Lemma 8, a solution to  $C_U(\beta)$  is given by

$$\left( \beta(\hat{\mathbf{\Omega}} = \mathbf{C} + \hat{\mathbf{Z}}^\perp, \hat{\mathbf{\Sigma}} = \mathbf{B} + \hat{\mathbf{Y}} - \hat{\mathbf{L}}, \hat{\mathbf{\Psi}}), \frac{1}{\beta}(\hat{\mathbf{R}} = \hat{\gamma} \mathbf{B} + \hat{\mathbf{Y}}^\perp, \hat{\mathbf{Q}} = \hat{\gamma} \mathbf{C} + \hat{\mathbf{Z}} - \hat{\mathbf{N}}, \hat{\mathbf{S}}) \right) \quad (198)$$

Inverting the roles of Lagrangian multipliers and optimization variables we obtain a solution to  $C_D(\alpha)$  for some  $\alpha$  as

$$\left( \frac{1}{\beta}(\hat{\mathbf{R}} = \hat{\gamma} \mathbf{B} + \hat{\mathbf{Y}}^\perp, \hat{\mathbf{Q}} = \hat{\gamma} \mathbf{C} + \hat{\mathbf{Z}} - \hat{\mathbf{N}}, \hat{\mathbf{S}}), \beta(\hat{\mathbf{\Omega}} = \mathbf{C} + \hat{\mathbf{Z}}^\perp, \hat{\mathbf{\Sigma}} = \mathbf{B} + \hat{\mathbf{Y}} - \hat{\mathbf{L}}, \hat{\mathbf{\Psi}}) \right) \quad (199)$$

By the conversion rules Equation (114) a solution to the unscaled downlink problem is given by

$$\left( \frac{1}{\alpha\beta}(\hat{\mathbf{R}} = \hat{\sigma} \mathbf{B} + \hat{\mathbf{Y}}^\perp, \hat{\mathbf{Q}} = \hat{\gamma} \mathbf{C} + \hat{\mathbf{Z}} - \hat{\mathbf{N}}, \hat{\mathbf{S}}), \alpha\beta(\hat{\mathbf{\Omega}} = \mathbf{C} + \hat{\mathbf{Z}}^\perp, \hat{\mathbf{\Sigma}} = \mathbf{B} + \hat{\mathbf{Y}} - \hat{\mathbf{L}}, \hat{\mathbf{\Psi}}) \right) \quad (200)$$

Introducing  $\check{\gamma} = \check{\nu} = \check{\mu}$  and comparing the solution obtained from the uplink problem to the prototype of the downlink solution

$$\left( (\check{\mathbf{R}} = \mathbf{B} + \check{\mathbf{Y}}^\perp, \check{\mathbf{Q}} = \mathbf{C} + \check{\mathbf{Z}} - \check{\mathbf{N}}, \check{\mathbf{S}}), (\check{\mathbf{\Omega}} = \check{\gamma} \mathbf{C} + \check{\mathbf{Z}}^\perp, \check{\mathbf{\Sigma}} = \check{\gamma} \mathbf{B} + \check{\mathbf{Y}} - \check{\mathbf{L}}, \check{\mathbf{\Psi}}, \check{\mathbf{K}}) \right) \quad (201)$$

we identify  $\frac{\hat{\gamma}}{\alpha\beta} = 1$  and  $\alpha\beta = \check{\gamma}$ . Introducing  $\gamma = \hat{\gamma} = \check{\gamma} = \alpha\beta$ , we can see that uplink and downlink solution are connected by inverting the roles optimization variables and Lagrangian multipliers and a scaling by  $\gamma$ , respectively  $\frac{1}{\gamma}$ , which proves the last part of the Theorem.

### 5.5. Zero Duality Gap

The upper bound Equation (165) by a Lagrangian dual problem was shown to be tight. Therefore, maximization problems in interference networks with linear conic constraints have a zero-duality gap for this special choice of the dual function. The zero duality gap result does not hold in general. For example it does not hold if the dual function is selected as optimization of the Lagrangian function given by Equation (108). Although we have  $L_D = C_D$  for every solution, the duality gap does not disappear as

$$\sup_{S_k} \left\{ w_k \log \left| I + S_k^+ H_{kk} Q_k H_{kk}^H \right| + \text{tr}(\Sigma_k S_k) \right\} = \infty \quad (202)$$

Here, the linear penalty term is not powerful enough to prevent the exponential growth of the utility for decreasing  $S_k$ .

A Zero-Duality gap can be observed for the optimization of linear precoders for power minimization under SINR constraints [14,22], which allows to transfer the apparently non-convex problem into a convex problem that can be solved efficiently. For our scenario, the dual problem is always convex in the dual variables, but evaluating the dual function  $S_D(\mathbf{R}, \mathbf{\Omega})$  still requires to solve a non-convex problem Equation (88). So, the zero duality gap is a property worth mentioning, but it is completely useless for the design of efficient algorithms.

### 5.6. Uplink-Downlink Transformation Rules

Several network dualities include some transformation rule for the solution of one network to the dual network. In our approach, the complete characterization of the uplink-downlink duality by a shared system of optimality condition reveals such a transformation as a connection between optimization variables and Lagrangian multipliers. For example, the transformation of the transmit covariances from one problem to the other is given by Equations (44) and (62). If the noise covariance of a user  $k$  has full rank, the only valid choice for  $Q_k^\Delta$ , respectively  $\Sigma_k^\Delta$ , is the zero matrix and the transformations are identical to those in [12]. Otherwise,  $Q_k^\Delta$  and  $\Sigma_k^\Delta$  describe the additional degrees of freedom: this is where the conic constraint is not binding but the choice of has no influence on the utility. To describe the full dependencies, we start by the slack variables  $L, N$  being the Lagrangian multipliers for the positive semi-definiteness constraint of the noise covariances. So  $L$  and  $N$  indicate if and how the interference plus noise can be rank deficient, which sets the degrees of freedom for  $Q_k^\Delta$  and  $\Sigma^\Delta$  and therefore the degrees of freedom in the linear conic constraint on  $Q$  and  $\Sigma$ , implying the set of possible slack variables.

### 5.7. Possible Extensions

Another direction for future work is to investigate if other types of constraints can be included in the duality. This is straight forward for the noise covariance. Let  $\mathcal{R}$  be an arbitrary constraint set for the noise covariance  $\mathbf{R}$ . We restate the downlink problem as

$$\inf_{\mathbf{R}' \succeq 0, \mathbf{R}' \in \mathcal{R}} \left\{ \sup_{\mathbf{Q} \succeq 0, \mathbf{Z}} \{W(\mathbf{R}', \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z}\} \right\} \quad (203)$$

$$= \inf_{\mathbf{R}' \succeq 0} \inf_{\mathbf{R} \succeq 0} \left\{ \sup_{\mathbf{Q} \succeq 0, \mathbf{Z}} \{W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z}\} : \mathbf{R} = \mathbf{R}' \right\} \quad (204)$$

$$= \inf_{\mathbf{R}' \succeq 0, \mathbf{R} \in \mathcal{R}} \inf_{\mathbf{R} \succeq 0, \mathbf{Y}^\perp} \left\{ \sup_{\mathbf{Q} \succeq 0, \mathbf{Z}} \{W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z}\} : \mathbf{R} = \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (205)$$

where  $\mathbf{B} = \mathbf{R}'$  and  $\mathcal{Y}^\perp = \{\mathbf{0}\}$ . Now, we can apply the minimax duality to the inner problem and obtain

$$\inf_{\mathbf{R}' \succeq \mathbf{0}, \mathbf{R}' \in \mathcal{R}} \inf_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \sup_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \left\{ W(\mathbf{\Omega}, \mathbf{\Sigma}) : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Sigma}_k \mathbf{R}') \leq P \right\} : \mathbf{\Omega} \preceq \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (206)$$

$$= \inf_{\mathbf{R}' \succeq \mathbf{0}, \mathbf{R}' \in \mathcal{R}} \inf_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \sup_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \left\{ W(\mathbf{\Omega}, \mathbf{\Sigma}) : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{\Sigma}_k \mathbf{R}') \leq P \right\} : \mathbf{\Omega} \preceq \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (207)$$

where the formulation by the network sum-power constraint is explained in Section 2.1.2. Concerning the generalization of constraints on the transmit covariances, we suggest to first extend the duality to multiple conic constraints for each link. Furthermore, an especially interesting direction is to include more general (convex) constraints on the transmit covariance into our duality. In [16] convex constraints are handled by replacing them with a single linear constraint, where the parameter needs to be found algorithmically. If the multiple linear constraints can be combined into a single linear conic constraint this is inherently done by our model. The general case is an open issue and suggested as future work.

## 6. Conclusions and Future Work

We presented a minimax duality for MIMO interference networks, where noise and transmit covariances are optimized subject to linear conic constraints. We illustrated how they are general enough to not only model typically used constraints such as the network sum-power constraint, the per link sum-power constraint, and shaping constraints, but also constraints on cooperation within the network or the propriety of signals. The interconnection of the downlink and uplink problems is by the flipped channel matrices, the common constraint matrices, and orthogonal subspaces. We showed that every solution has a polite water-filling structure and the downlink and uplink problem share a system of optimality conditions, where the solutions are connected by inverting the roles of the optimization variables and Lagrangian multipliers. We proof our minimax duality by using an upper bound from Lagrangian duality, which can be shown to be tight, thus revealing a zero-duality gap of maximization problems in MIMO interference networks under linear conic constraints.

Especially interesting is the BC-MAC duality, which can also be modeled in our framework. We were able to extend the existing BC-MAC dualities in two directions. First, we allow for more general constraints (for example shaping constraints) on the downlink transmit covariance and investigate the conditions under which the uplink problem has a convex-concave game structure. Second, we extended the worst case noise capacity from the sum-rate to the complete capacity region of the MIMO broadcast with non-linear interference pre-subtraction. New applications that are enabled by our duality are discussed in Section 4.2. These are the design of interference robust multi-user MIMO transmission strategies for interfering broadcast channels and information theoretic proofs concerning cooperation in networks and optimality of proper signals.

Concerning the generalization of constraints on the transmit covariances, we suggest to first extend the duality to multiple conic constraints for each link. Furthermore, an especially interesting direction is to include more general (convex) constraints on the transmit covariance into our duality. In [16], convex constraints are handled by replacing them with a single linear constraint, where the parameter needs to be found algorithmically. Furthermore, we believe that our result can serve as the theoretical basis for further research in the fields of robust transmission strategies, cognitive radio networks, and secrecy capacity, where conic constraints and minimax problems appear naturally in the proofs and numerical computations.

**Author Contributions:** Andreas Dotzler discovered the Min-Max Duality; Maximilian Riemensberger assisted with the proof; Wolfgang Utschick contributed his expertise in interference networks and network dualities.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A Mutual Information For Non-Invertible Noise Covariances

**Lemma 12.** Consider a MIMO link with channel matrix  $\mathbf{H}$  and noise covariance  $\mathbf{S}$ . The transmit covariance  $\mathbf{Q}$  is constrained by  $\mathbf{Q} \in \mathcal{Q}$ , but can be freely chosen into all directions of the channel. Let  $\mathbf{T}, \mathbf{T}^\perp$  provide a full unitary basis for the fundamental subspaces of  $\mathbf{S}$ . If  $\mathbf{T}^{\perp\mathrm{H}}\mathbf{H} = \mathbf{0}$ , then the mutual information is finite and is given by

$$\max_{\mathbf{Q} \in \mathcal{Q}} \log |\mathbf{I} + \mathbf{S}^+ \mathbf{H} \mathbf{Q} \mathbf{H}^{\mathrm{H}}| \quad (\text{A1})$$

**Proof.** The mutual information is infinite unless we have  $\mathbf{T}^{\perp\mathrm{H}}\mathbf{H} = \mathbf{0}$ , as otherwise there exists a  $\mathbf{Q}$  for which we obtain a noise free communication channel by using  $\mathbf{T}^{\perp\mathrm{H}}$  as a receive filter. For finite mutual information, we can compute the mutual information for any  $\mathbf{Q}$  by taking a limit:

$$\lim_{\epsilon \rightarrow 0} \log |\mathbf{I} + (\mathbf{S} + \epsilon \mathbf{T} \mathbf{T}^{\perp\mathrm{H}})^{-1} \mathbf{H} \mathbf{Q} \mathbf{H}^{\mathrm{H}}| \quad (\text{A2})$$

$$= \lim_{\epsilon \rightarrow 0} \log \left| \mathbf{I} + \begin{bmatrix} \mathbf{T} \mathbf{T}^{\perp} \end{bmatrix} \begin{bmatrix} (\mathbf{T}^{\mathrm{H}} \mathbf{S} \mathbf{T})^{-1} & \mathbf{0} \\ \mathbf{0} & \epsilon^{-1} \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{T} \mathbf{T}^{\perp} \end{bmatrix}^{\mathrm{H}} \mathbf{H} \mathbf{Q} \mathbf{H}^{\mathrm{H}} \right| \quad (\text{A3})$$

$$= \lim_{\epsilon \rightarrow 0} \log \left| \mathbf{I} + \mathbf{H}^{\mathrm{H}} \begin{bmatrix} \mathbf{T} \mathbf{T}^{\perp} \end{bmatrix} \begin{bmatrix} (\mathbf{T}^{\mathrm{H}} \mathbf{S} \mathbf{T})^{-1} & \mathbf{0} \\ \mathbf{0} & \epsilon^{-1} \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{T} \mathbf{T}^{\perp} \end{bmatrix}^{\mathrm{H}} \mathbf{H} \mathbf{Q} \right| \quad (\text{A4})$$

$$= \lim_{\epsilon \rightarrow 0} \log \left| \mathbf{I} + \begin{bmatrix} \mathbf{H}^{\mathrm{H}} \mathbf{T} & \mathbf{0} \end{bmatrix} \begin{bmatrix} (\mathbf{T}^{\mathrm{H}} \mathbf{S} \mathbf{T})^{-1} & \mathbf{0} \\ \mathbf{0} & \epsilon^{-1} \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{T}^{\mathrm{H}} \mathbf{H} \\ \mathbf{0} \end{bmatrix} \mathbf{Q} \right| \quad (\text{A5})$$

$$= \log |\mathbf{I} + \mathbf{H}^{\mathrm{H}} \mathbf{T} (\mathbf{T}^{\mathrm{H}} \mathbf{S} \mathbf{T})^{-1} \mathbf{T}^{\mathrm{H}} \mathbf{H} \mathbf{Q}| \quad (\text{A6})$$

$$= \log |\mathbf{I} + \mathbf{S}^+ \mathbf{H} \mathbf{Q} \mathbf{H}^{\mathrm{H}}| \quad (\text{A7})$$

□

## Appendix B Proof of Lemma 1

For every  $\tilde{\mathbf{\Lambda}} \in \mathcal{P}(\mathcal{X}^\perp)$ , we have  $\tilde{\mathbf{\Lambda}} \succeq \mathbf{0}$ , which implies every element of the tuple is Hermitian. If  $\mathbf{A} + \mathbf{X} \succeq \mathbf{0}$ , then  $\text{tr}(\tilde{\mathbf{\Lambda}}_k(\mathbf{A}_k + \mathbf{X}_k)) \geq 0 \forall k \in \mathcal{K}$ . Further, by the orthogonal subspaces we have

$$\sum_{k \in \mathcal{K}} \text{tr}(\tilde{\mathbf{\Lambda}}_k(\mathbf{A}_k + \mathbf{X}_k)) = 0 \quad \forall \mathbf{X} \in \mathcal{X} \quad (\text{B1})$$

and therefore

$$\text{tr}(\tilde{\mathbf{\Lambda}}_k(\mathbf{A}_k + \mathbf{X}_k)) = 0 \quad \forall k \in \mathcal{K}, \forall \mathbf{X} \in \mathcal{X}, \mathbf{A} + \mathbf{X} \succeq \mathbf{0} \quad (\text{B2})$$

For  $\tilde{\mathbf{\Lambda}}_k \in \mathbb{C}^{N_k \times N_k}$  let  $\tilde{\mathbf{\Lambda}}_k = \sum_{i=1}^{N_k} \gamma_{ki} \boldsymbol{\lambda}_{ki} \boldsymbol{\lambda}_{ki}^{\mathrm{H}}$  such that  $\gamma_{ki} \geq 0, \forall k, i$ . Then it follows  $\forall k \in \mathcal{K}, \forall \mathbf{X} \in \mathcal{X}, \mathbf{A} + \mathbf{X} \succeq \mathbf{0}$ .

$$\text{tr}(\tilde{\mathbf{\Lambda}}_k(\mathbf{A}_k + \mathbf{X}_k)) = 0 \quad (\text{B3})$$

$$\Leftrightarrow \sum_{i=1}^{N_k} \text{tr}(\sqrt{\gamma_{ki}} \boldsymbol{\lambda}_{ki}^{\mathrm{H}} (\mathbf{A}_k + \mathbf{X}_k) \boldsymbol{\lambda}_{ki} \sqrt{\gamma_{ki}}) = 0 \quad (\text{B4})$$

$$\Leftrightarrow \sum_{i=1}^{N_k} \sqrt{\gamma_{ki}} \boldsymbol{\lambda}_{ki}^{\mathrm{H}} (\mathbf{A}_k + \mathbf{X}_k) \boldsymbol{\lambda}_{ki} \sqrt{\gamma_{ki}} = 0 \quad (\text{B5})$$

$$\Leftrightarrow \boldsymbol{\Lambda}_k^{\mathrm{H}} (\mathbf{A}_k + \mathbf{X}_k) \boldsymbol{\Lambda}_k = \mathbf{0} \quad (\text{B6})$$

and it follows that for  $\tilde{\mathbf{\Lambda}}_k = \boldsymbol{\Lambda}_k \boldsymbol{\Lambda}_k^{\mathrm{H}} \forall k \in \mathcal{K}$  we have  $\tilde{\mathbf{\Lambda}} \in \mathcal{N}(\mathbf{A}, \mathcal{X})$ .



### Appendix C Proof of Lemma 2

$W \succeq 0$  fulfills the linear conic constraint if

$$\exists X \in \mathcal{X} : W \preceq A + X \quad (C1)$$

$$\Leftrightarrow \exists X \in \mathcal{X} : \text{tr}(E_k(W_k - A_k - X_k)) \leq 0, \forall E \succeq 0, \forall k \in \mathcal{K} \quad (C2)$$

$$\Leftrightarrow \exists X \in \mathcal{X} : \sum_{k \in \mathcal{K}} \text{tr}(E_k(W_k - A_k - X_k)) \leq 0, \forall E \succeq 0 \quad (C3)$$

$$\Leftrightarrow \exists X \in \mathcal{X} : \sum_{k \in \mathcal{K}} \text{tr}(E_k(W_k - A_k - X_k)) \leq 0, \forall E \in \mathcal{E} \quad (C4)$$

$$\Leftrightarrow \sum_{k \in \mathcal{K}} \text{tr}(E_k(W_k - A_k)) \leq 0, \forall E \in \mathcal{E} \quad (C5)$$

Equivalence of Equations (C1) and (C2) is clear by Equations (B3)–(B5). By summing over all links, Equation (C3) follows from Equation (C2). For the inverse direction, as Equation (C3) holds for any  $E$  we can select an arbitrary  $E_k$  and set all other elements of the tuple to the all zero matrix and Equation (C2) follows. As Equation (C3) holds for all  $E$  it obviously holds for those where  $E \in \mathcal{E}$ , this is Equation (C4). For the inverse direction, if  $E \notin \mathcal{E}$  then  $\sum_{k \in \mathcal{K}} \text{tr}(E_k X_k) \neq 0$  and as  $X$  can be scaled arbitrarily and there exists a  $X$ , that fulfills the inequality. Equivalence of Equations (C4) and (C5) holds as  $\sum_{k \in \mathcal{K}} \text{tr}(E_k X_k) = 0$  for any  $X \in \mathcal{X}, E \in \mathcal{E}$  by definition. If  $\mathcal{E}$  is polyhedral, every point can be expressed by a positive combination of the finite extreme points and Equation (C5) is equivalent to Equation (20).

### Appendix D Proof of Lemma 3

Consider the downlink. By a scaling of the constraint matrices and channels by scaling factors  $\theta$  and  $\eta$  such that

$$\bar{H}_{ki} = \frac{\sqrt{\eta}}{\sqrt{\theta}} H_{ki}, \bar{B} = \frac{1}{\theta} B, \bar{C} = \frac{1}{\eta} C \quad (D1)$$

and substituting

$$(\bar{R}, \bar{S}, \bar{Q}, \bar{Z}, \bar{Y}^\perp) = \left( \frac{1}{\theta} R, \frac{1}{\theta} S, \frac{1}{\eta} Q, \frac{1}{\eta} Z, \frac{1}{\theta} Y^\perp \right) \quad (D2)$$

we obtain the following equivalences:

$$\inf_{\substack{\bar{R} \succeq 0 \\ \bar{Y}^\perp \in \mathcal{Y}^\perp}} \left\{ \sup_{\substack{\bar{S} \\ \bar{Q} \succeq 0 \\ \bar{Z} \in \mathcal{Z}}} \left\{ \sum_{k \in \mathcal{K}} w_k \log |I + \bar{S}^+ \bar{H}_{kk} \bar{Q}_k \bar{H}_{kk}^H| : \right. \right. \\ \left. \left. \bar{S}_k = \bar{R}_k + \sum_{i \in \mathcal{K} \setminus k} \bar{H}_{ki} \bar{Q}_i \bar{H}_{ki}^H \forall k, \bar{Q} \preceq \bar{C} + \bar{Z} \right\} : \bar{R} = \bar{B} + \bar{Y}^\perp \right\} \quad (D3)$$

$$= \inf_{\substack{\bar{\mathbf{R}} \succeq \mathbf{0} \\ \mathbf{Y}^\perp \in \mathcal{Y}^\perp}} \left\{ \sup_{\substack{\bar{\mathbf{S}} \\ \bar{\mathbf{Q}} \succeq \mathbf{0} \\ \bar{\mathbf{Z}} \in \mathcal{Z}}} \left\{ \sum_{k \in \mathcal{K}} w_k \log \left| \mathbf{I} + \bar{\mathbf{S}}^+ \bar{\mathbf{H}}_{kk} \bar{\mathbf{Q}}_k \bar{\mathbf{H}}_{kk}^H \right| : \right. \right. \\ \left. \left. \bar{\mathbf{S}}_k = \bar{\mathbf{R}}_k + \sum_{i \in \mathcal{K} \setminus k} \bar{\mathbf{H}}_{ki} \bar{\mathbf{Q}}_i \bar{\mathbf{H}}_{ki}^H \forall k, \eta \bar{\mathbf{Q}} \preceq \eta \bar{\mathbf{C}} + \eta \bar{\mathbf{Z}} \right\} : \theta \bar{\mathbf{R}} = \theta \bar{\mathbf{B}} + \theta \bar{\mathbf{Y}}^\perp \right\} \quad (\text{D4})$$

$$= \inf_{\substack{\mathbf{R} \succeq \mathbf{0} \\ \mathbf{Y}^\perp \in \mathcal{Y}^\perp}} \left\{ \sup_{\substack{\mathbf{S} \\ \mathbf{Q} \succeq \mathbf{0} \\ \mathbf{Z} \in \mathcal{Z}}} \left\{ \sum_{k \in \mathcal{K}} w_k \log \left| \mathbf{I} + \left( \frac{1}{\theta} \mathbf{S} \right)^+ \bar{\mathbf{H}}_{kk} \frac{1}{\eta} \mathbf{Q}_k \bar{\mathbf{H}}_{kk}^H \right| : \right. \right. \\ \left. \left. \frac{1}{\theta} \mathbf{S}_k = \frac{1}{\theta} \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \bar{\mathbf{H}}_{ki} \frac{1}{\eta} \mathbf{Q}_i \bar{\mathbf{H}}_{ki}^H \forall k, \mathbf{Q} \preceq \eta \bar{\mathbf{C}} + \mathbf{Z} \right\} : \mathbf{R} = \theta \bar{\mathbf{B}} + \mathbf{Y}^\perp \right\} \quad (\text{D5})$$

$$= \inf_{\substack{\mathbf{R} \succeq \mathbf{0} \\ \mathbf{Y}^\perp \in \mathcal{Y}^\perp}} \left\{ \sup_{\substack{\mathbf{S} \\ \mathbf{Q} \succeq \mathbf{0} \\ \mathbf{Z} \in \mathcal{Z}}} \left\{ \sum_{k \in \mathcal{K}} w_k \log \left| \mathbf{I} + \mathbf{S}^+ \sqrt{\theta} \bar{\mathbf{H}}_{kk} \frac{1}{\sqrt{\eta}} \mathbf{Q}_k \frac{1}{\sqrt{\eta}} \bar{\mathbf{H}}_{kk}^H \sqrt{\theta} \right| : \right. \right. \\ \left. \left. \mathbf{S}_k = \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \sqrt{\theta} \bar{\mathbf{H}}_{ki} \frac{1}{\sqrt{\eta}} \mathbf{Q}_i \frac{1}{\sqrt{\eta}} \bar{\mathbf{H}}_{ki}^H \sqrt{\theta} \forall k, \mathbf{Q} \preceq \eta \bar{\mathbf{C}} + \mathbf{Z} \right\} : \mathbf{R} = \theta \bar{\mathbf{B}} + \mathbf{Y}^\perp \right\} \quad (\text{D6})$$

$$= \inf_{\substack{\mathbf{R} \succeq \mathbf{0} \\ \mathbf{Y}^\perp \in \mathcal{Y}^\perp}} \left\{ \sup_{\substack{\mathbf{S} \\ \mathbf{Q} \succeq \mathbf{0} \\ \mathbf{Z} \in \mathcal{Z}}} \left\{ \sum_{k \in \mathcal{K}} w_k \log \left| \mathbf{I} + \mathbf{S}^+ \mathbf{H}_{kk} \mathbf{Q}_k \mathbf{H}_{kk}^H \right| : \right. \right. \\ \left. \left. \mathbf{S}_k = \mathbf{R}_k + \sum_{i \in \mathcal{K} \setminus k} \mathbf{H}_{ki} \mathbf{Q}_i \mathbf{H}_{ki}^H \forall k, \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z} \right\} : \mathbf{R} = \mathbf{B} + \mathbf{Y}^\perp \right\} \quad (\text{D7})$$

Let the functions  $W(\mathbf{R}, \mathbf{Q}; s)$  and  $W(\mathbf{\Omega}, \mathbf{\Sigma}; s)$  follow the definition of  $W(\mathbf{R}, \mathbf{Q})$  and  $W(\mathbf{\Omega}, \mathbf{\Sigma})$  respectively, these are Equations (9) and (10), but all channels are scaled by  $s$ . Let  $\mathbf{B}_n = \frac{1}{\sqrt{b}} \mathbf{B}$  and  $\mathbf{C}_n = \frac{1}{\sqrt{c}} \mathbf{C}$ , where  $b = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k)$  and  $c = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k)$ , then

$$\inf_{\mathbf{R} \succeq \mathbf{0}, \mathbf{Y}^\perp} \left\{ \sup_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z}} \{W(\mathbf{R}, \mathbf{Q}) : \mathbf{Q} \preceq \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z}\} : \mathbf{R} \preceq \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (\text{D8})$$

$$= \inf_{\mathbf{R} \succeq \mathbf{0}, \mathbf{Y}^\perp} \left\{ \sup_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z}} \left\{ W\left(\mathbf{R}, \mathbf{Q}; \frac{\sqrt{\sqrt{c}}}}{\sqrt{\sqrt{b}}}\right) : \mathbf{Q} \preceq \frac{1}{\sqrt{c}} \mathbf{C} + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z} \right\} : \mathbf{R} \preceq \frac{1}{\sqrt{b}} \mathbf{B} + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (\text{D9})$$

$$= \inf_{\mathbf{R} \succeq \mathbf{0}, \mathbf{Y}^\perp} \left\{ \sup_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Z}} \left\{ W\left(\mathbf{R}, \mathbf{Q}; \frac{\sqrt{\sqrt{c}}}}{\sqrt{\sqrt{b}}}\right) : \mathbf{Q} \preceq \mathbf{C}_n + \mathbf{Z}, \mathbf{Z} \in \mathcal{Z} \right\} : \mathbf{R} \preceq \mathbf{B}_n + \mathbf{Y}^\perp, \mathbf{Y}^\perp \in \mathcal{Y}^\perp \right\} \quad (\text{D10})$$

For the uplink, we have

$$\inf_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \sup_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \{W(\mathbf{\Sigma}, \mathbf{\Omega}) : \mathbf{\Sigma} \preceq \mathbf{B} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y}\} : \mathbf{\Omega} \preceq \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (\text{D11})$$

$$= \inf_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \sup_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \left\{ W\left(\mathbf{\Sigma}, \mathbf{\Omega}; \frac{\sqrt{\sqrt{b}}}}{\sqrt{\sqrt{c}}}\right) : \mathbf{\Sigma} \preceq \frac{1}{\sqrt{b}} \mathbf{B} + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y} \right\} : \mathbf{\Omega} \preceq \frac{1}{\sqrt{c}} \mathbf{C} + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (\text{D12})$$

$$= \inf_{\mathbf{\Omega} \succeq \mathbf{0}, \mathbf{Z}^\perp} \left\{ \sup_{\mathbf{\Sigma} \succeq \mathbf{0}, \mathbf{Y}} \left\{ W\left(\mathbf{\Sigma}, \mathbf{\Omega}; \frac{\sqrt{\sqrt{b}}}}{\sqrt{\sqrt{c}}}\right) : \mathbf{\Sigma} \preceq \mathbf{B}_n + \mathbf{Y}, \mathbf{Y} \in \mathcal{Y} \right\} : \mathbf{\Omega} \preceq \mathbf{C}_n + \mathbf{Z}^\perp, \mathbf{Z}^\perp \in \mathcal{Z}^\perp \right\} \quad (\text{D13})$$

Combining the reformulations, we can prove the asymmetric minimax duality:

$$\inf_{R \succeq 0, Y^\perp} \left\{ \sup_{Q \succeq 0, Z} \{W(R, Q) : Q \preceq C + Z, Z \in \mathcal{Z}\} : R \preceq B + Y^\perp, Y^\perp \in \mathcal{Y}^\perp \right\} \quad (D14)$$

$$= \inf_{R \succeq 0, Y^\perp} \left\{ \sup_{Q \succeq 0, Z} \left\{ W\left(R, Q; \frac{\sqrt{\sqrt{c}}}{\sqrt{\sqrt{b}}}\right) : Q \preceq C_n + Z, Z \in \mathcal{Z} \right\} : R \preceq B_n + Y^\perp, Y^\perp \in \mathcal{Y}^\perp \right\} \quad (D15)$$

$$= \inf_{\Omega \succeq 0, Z^\perp} \left\{ \sup_{\Sigma \succeq 0, Y} \left\{ W\left(\Sigma, \Omega; \frac{\sqrt{\sqrt{c}}}{\sqrt{\sqrt{b}}}\right) : \Sigma \preceq B_n + Y, Y \in \mathcal{Y} \right\} : \Omega \preceq C_n + Z^\perp, Z^\perp \in \mathcal{Z}^\perp \right\} \quad (D16)$$

$$= \inf_{\Omega \succeq 0, Z^\perp} \left\{ \sup_{\Sigma \succeq 0, Y} \{W(\Sigma, \Omega) : \Sigma \preceq \sqrt{c}B_n + Y, Y \in \mathcal{Y}\} : \Omega \preceq \sqrt{b}C_n + Z^\perp, Z^\perp \in \mathcal{Z}^\perp \right\} \quad (D17)$$

$$= \inf_{\Omega \succeq 0, Z^\perp} \left\{ \sup_{\Sigma \succeq 0, Y} \left\{ W(\Sigma, \Omega) : \Sigma \preceq \frac{\sqrt{c}}{\sqrt{b}}B + Y, Y \in \mathcal{Y} \right\} : \Omega \preceq \frac{\sqrt{b}}{\sqrt{c}}C + Z^\perp, Z^\perp \in \mathcal{Z}^\perp \right\} \quad (D18)$$

$$= \inf_{\Omega \succeq 0, Z^\perp} \left\{ \sup_{\Sigma \succeq 0, Y} \left\{ W(\Sigma, \Omega) : \Sigma \preceq \beta \frac{\sqrt{c}}{\sqrt{b}}B + Y, Y \in \mathcal{Y} \right\} : \Omega \preceq \beta \frac{\sqrt{b}}{\sqrt{c}}C + Z^\perp, Z^\perp \in \mathcal{Z}^\perp \right\} \quad (D19)$$

From Equations (D14) to Equation (D15) is by the equivalence of Equations (D8) and (D10). With  $\sum_{k \in \mathcal{K}} \text{tr}(B_{nk}B_{nk}) = \sum_{k \in \mathcal{K}} \text{tr}(C_{nk}C_{nk}) = 1$ , we can apply the minimax duality by Theorem 1 to show the equivalence of Equations (D15) and (D16). We obtain Equation (D17) by inverting Equation (D11) to Equation (D12). For Equation (D18) we applied the definition of  $B_n$  and  $C_n$  and equivalence of Equations (D18) and (D19) is by Lemma 8. Let  $p = \beta \frac{\sqrt{c}}{\sqrt{b}}$  and  $n = \beta \frac{\sqrt{b}}{\sqrt{c}}$ . By choosing  $\beta$ , we can select any  $p$  and  $n$  as long as  $\frac{p}{n} = \beta \frac{\sqrt{c}}{\sqrt{b}} \frac{1}{\beta} \frac{\sqrt{c}}{\sqrt{b}} = \frac{c}{b}$ .

#### Appendix E Proof of Lemma 4

The optimality conditions of Equation (76) are

$$D^{-1} = A \quad (E1)$$

$$\lambda' I + F' = HA'H^H + M' \quad (E2)$$

$$F' = \begin{bmatrix} 0 & \dots & * \\ \vdots & \ddots & \vdots \\ * & \dots & 0 \end{bmatrix} \quad (E3)$$

$$\Sigma' = \begin{bmatrix} \Sigma'_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \Sigma'_U \end{bmatrix} \quad (E4)$$

$$M' \succeq 0 \quad (E5)$$

$$\text{tr}(M'\Sigma') = 0 \quad (E6)$$

$$\text{tr}(\Sigma') = P \quad (E7)$$

$$\Sigma' \succeq 0 \quad (E8)$$

First, we show that Equations (77) and (78) provide valid Lagrangian multipliers, by showing they fulfill Equations (E2), (E3), (E5), and (E6). It is straight-forward to verify condition Equations (E2) and (E3). Condition (E2) combined with Equation (E3) implies

$$\lambda' I \succeq H_u A' H_u^H \quad \forall u \quad (E9)$$

as otherwise there would be no  $M'$  such that  $M' \succeq 0$ . Comparing Equation (E9) to definition of  $M'$  implies  $M' \succeq 0$ , this is Equation (E5). From Equations (E3) and (E4) follows  $\text{tr}(F'\Sigma') = 0$  and from Equation (E2) we obtain

$$\lambda' \text{tr}(\Sigma') + \text{tr}(F'\Sigma') = \text{tr}(HA'H^H\Sigma') + \text{tr}(M'\Sigma') \quad (\text{E10})$$

$$\Rightarrow \lambda' P = \text{tr}(HA'H^H\Sigma') \quad (\text{E11})$$

where we used Equations (E6) and (E7). Finally, we combine Equations (78) and (E6). As the offdiagonal blocks do not contribute to the trace we have

$$\text{tr}(M'\Sigma') = \lambda' \text{tr}(\Sigma') - \text{tr}(HA'H^H\Sigma') \quad (\text{E12})$$

and by Equations (E7) and (E11) we have  $\text{tr}(M'\Sigma') = 0$ , which is Equation (E6). Second, we show that Equation (80) fulfills the optimality conditions of Equation (70). Introducing  $\Sigma$  as a Lagrangian multiplier for the structure of  $R$  and  $R \succeq 0$ , which implies  $\Sigma$  is blockdiagonal and  $\Sigma \succeq 0$ , and  $\mu$  for  $\text{tr}(Q) \leq P$  and  $K$  for  $Q \succeq 0$ , these optimality conditions are

$$\Sigma = -\left(R + HQH^H\right)^+ + R^+ + \Sigma^\Delta \quad (\text{E13})$$

$$\mu I = H^H\left(R + HQH^H\right)^+ H + K \quad (\text{E14})$$

$$\text{tr}(KQ) = 0, K \succeq 0, Q \succeq 0 \quad (\text{E15})$$

We now introduce  $\Psi' = I$  and

$$Q' = (\Phi' + H^H\Sigma'H)^+ + \Psi' \quad (\text{E16})$$

$$= (I + H^H\Sigma'H)^{-1} + I \quad (\text{E17})$$

$$S' = H(\Psi' + H^H\Sigma'H)^+ H^H + M' \quad (\text{E18})$$

$$= HA'H^H + M' \quad (\text{E19})$$

By the proof of Lemma 9 in Appendix I, we know that Equations (E16), (E18), (E5), (E6), and (E8), imply there exist  $\Sigma^\Delta, K'$  that fulfill the following set of equations:

$$\Sigma' = -\left(S' + HQ'H^H\right)^+ + S'^+ + \Sigma^{\Delta'} \quad (\text{E20})$$

$$I = H^H\left(S' + HQ'H^H\right)^+ H + K' \quad (\text{E21})$$

$$\text{tr}(K'Q') = 0, K' \succeq 0, Q' \succeq 0 \quad (\text{E22})$$

Comparing Equations (E20)–(E22) to Equations (E13)–(E15), we can see that  $\mu = \lambda'$ ,  $R = \frac{1}{\lambda'} S'$ ,  $Q = \frac{1}{\lambda'} Q'$ ,  $\Sigma = \lambda' \Sigma'$ ,  $K = \lambda' K'$  fulfill the optimality conditions for the downlink problem. As  $R = \frac{1}{\lambda'} S'$  matches Equation (80) the proof is complete.

## Appendix F BC-MAC Duality – Conditions for a Convex Reformulation of the MAC Problem

Let  $C = (C', \dots, C') \in \prod_{k \in \mathcal{K}} \mathcal{S}^{N \times N}$  and  $\mathcal{Z}^\perp \subseteq \prod_{k \in \mathcal{K}} \mathcal{S}^{N \times N}$  with  $C' \in \mathcal{S}^{N \times N}$  and

$$\mathcal{Z}^\perp = \left\{ Z^\perp : Z_k^\perp = Z'^\perp \forall k \in \mathcal{K}, Z'^\perp \in \mathcal{Z}'^\perp \right\} \quad (\text{F1})$$

where  $\mathcal{Z}'^\perp$  is a subspace of  $\mathcal{S}^{N \times N}$  that is orthogonal to  $C' = \{\lambda C' : \lambda \in \mathbb{R}\}$ . Then the uplink noise constraint is

$$\Omega_k \preceq \frac{1}{K} C' + Z'^\perp, Z'^\perp \in \mathcal{Z}'^\perp \forall k \in \mathcal{K} \quad (\text{F2})$$

and there exists an uplink solution where all noise covariances are equal, see Remark 2 in Section 3. For this special choice, the subspace  $\mathcal{Z}$  is given by

$$\mathcal{Z} = \left\{ \mathbf{Z} : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}' \mathbf{Z}_k) = 0, \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{Z}'^\perp \mathbf{Z}_k) = 0 \forall \mathbf{Z}'^\perp \in \mathcal{Z}'^\perp \right\} \quad (\text{F3})$$

In the following, we show that Equation (94) can be expressed by this model. The constraint on the downlink transmit covariance is

$$\mathbf{Q}_k \preceq \frac{1}{K} \mathbf{C}' + \mathbf{Z}_k, \mathbf{Z} \in \mathcal{Z} \forall k \in \mathcal{K} \quad (\text{F4})$$

which implies

$$\sum_{k \in \mathcal{K}} \mathbf{Q}_k \preceq \frac{1}{K} \sum_{k \in \mathcal{K}} \mathbf{C}' + \sum_{k \in \mathcal{K}} \mathbf{Z}_k \quad (\text{F5})$$

$$\Leftrightarrow \mathbf{Q}' \preceq \mathbf{C}' + \mathbf{Z}' \quad (\text{F6})$$

where  $\mathbf{Q}' = \sum_{k \in \mathcal{K}} \mathbf{Q}_k$  and  $\mathbf{Z}' = \sum_{k \in \mathcal{K}} \mathbf{Z}_k$ . In the following, we show that not only Equation (F4) implies Equation (F6), but both are equivalent. We do so by showing that if given  $\mathbf{Q} = (\mathbf{Q}_1, \dots, \mathbf{Q}_K)$ ,  $\mathbf{Q}' = \sum_{k \in \mathcal{K}} \mathbf{Q}_k$  and  $\mathbf{Z}'$  that fulfill Equation (F6) we can construct a  $\mathbf{Z} = (\mathbf{Z}_1, \dots, \mathbf{Z}_K)$  with  $\sum_{k \in \mathcal{K}} \mathbf{Z}_k = \mathbf{Z}'$  such that Equation (F4) holds. For the first  $K-1$  users, we select  $\mathbf{Z}_k$  as  $\mathbf{Z}_k = \mathbf{Q}_k - \frac{1}{K} \mathbf{C}'$ , so that Equation (F4) is fulfilled for  $k = 1, \dots, K-1$ , and we have

$$\sum_{k \in \mathcal{K} \setminus \{K\}} \mathbf{Q}_k = (K-1) \frac{1}{K} \mathbf{C}' + \sum_{k \in \mathcal{K} \setminus \{K\}} \mathbf{Z}_k \quad (\text{F7})$$

By subtracting Equation (F7) from Equation (F6), we obtain

$$\mathbf{Q}' - \sum_{k \in \mathcal{K} \setminus K} \mathbf{Q}_k \preceq \frac{1}{K} \mathbf{C}' + \mathbf{Z}' - \sum_{k \in \mathcal{K} \setminus K} \mathbf{Z}_k \quad (\text{F8})$$

$$\Rightarrow \mathbf{Q}_K \preceq \frac{1}{K} \mathbf{C}' + \mathbf{Z}_K \quad (\text{F9})$$

By Equation (F3) we have

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}' \mathbf{Z}_k) = 0 \Rightarrow \text{tr}(\mathbf{C}' \mathbf{Z}') = 0 \quad (\text{F10})$$

and

$$\sum_{k \in \mathcal{K}} \text{tr}(\mathbf{Z}'^\perp \mathbf{Z}_k) = 0 \Rightarrow \text{tr}(\mathbf{Z}'^\perp \mathbf{Z}') = 0 \quad (\text{F11})$$

and therefore  $\mathbf{Z}' \in \mathcal{Z}'$ , where  $\mathcal{Z}'$  is the orthogonal complement to  $\mathcal{C}' \oplus \mathcal{Z}'^\perp$ . Thus, if the constraint on the downlink transmit covariance is a constraint that solely acts on the sum of the transmit covariances, i.e.,

$$\sum_{k \in \mathcal{K}} \mathbf{Q}_k \preceq \mathbf{C}' + \mathbf{Z}', \mathbf{Z}' \in \mathcal{Z}' \quad (\text{F12})$$

then the uplink has a solution where all noise covariances are equal.

## Appendix G Details of the Worst Case Noise Approximation in the Broadcast

The optimization under a worst-case noise approximation is given by

$$\sup_{\mathbf{Q} \succeq \mathbf{0}} \left\{ \inf_{\mathbf{R} \succeq \mathbf{0}} \left\{ W(\mathbf{R}, \mathbf{Q}) : \text{tr}(\mathbf{R}_k) \leq \sigma_k^2 \forall k \right\} : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{Q}_k) \leq P \right\} \quad (\text{G1})$$

According to Section 2.1.2, the conic versions of the downlink constraint sets are

$$\mathbf{B} = \left( \frac{\sigma_1^2}{M_1} \mathbf{I}, \dots, \frac{\sigma_K^2}{M_K} \mathbf{I} \right) \quad (\text{G2})$$

$$\mathcal{Y}^\perp = \{ \mathbf{Y}^\perp : \text{tr}(\mathbf{Y}_k^\perp) = 0 \ \forall k \} \quad (\text{G3})$$

$$\mathcal{Y} = \left\{ \mathbf{Y} = (\lambda_1 \mathbf{I}, \dots, \lambda_K \mathbf{I}), \sum_{k \in \mathcal{K}} \lambda_k \sigma_k^2 = 0 \right\} \quad (\text{G4})$$

$$\mathbf{C} = \frac{P}{KN} \mathbf{I} \quad (\text{G5})$$

$$\mathcal{Z} = \left\{ \mathbf{Z} : \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{Z}_k) = 0 \right\} \quad (\text{G6})$$

$$\mathcal{Z}^\perp = \{ \mathbf{0} \} \quad (\text{G7})$$

Assuming the optimal encoding order and considering that the constraint on the downlink transmit covariance is a constraint on the sum covariance Lemma 7, allows us to exchange infimum and supremum. The asymmetric mini-max duality by Lemma 3 translates the downlink problem to the following uplink constraints

$$\mathbf{\Sigma} \preceq p \left( \frac{\sigma_1^2}{M_1} \mathbf{I}, \dots, \frac{\sigma_K^2}{M_K} \mathbf{I} \right) + (\lambda_1 \mathbf{I}, \dots, \lambda_K \mathbf{I}), \sum_{k \in \mathcal{K}} \lambda_k \sigma_k^2 = 0 \quad (\text{G8})$$

$$\mathbf{\Omega} \preceq n \frac{P}{KN} \mathbf{I} \quad (\text{G9})$$

By substituting

$$\frac{P_k}{\sigma_k^2} = p \frac{\sigma_k^2}{M_k} + \lambda_k \quad (\text{G10})$$

$$\Leftrightarrow \lambda_k = \frac{P_k}{\sigma_k^2} - p \frac{\sigma_k^2}{M_k} \quad (\text{G11})$$

We have

$$\sum_{k \in \mathcal{K}} \lambda_k \sigma_k^2 = 0 \quad (\text{G12})$$

$$\Leftrightarrow \sum_{k \in \mathcal{K}} \left( \frac{P_k}{\sigma_k^2} - p \frac{\sigma_k^2}{M_k} \right) \sigma_k^2 = 0 \quad (\text{G13})$$

$$\Leftrightarrow \sum_{k \in \mathcal{K}} P_k = p \sum_{k \in \mathcal{K}} \frac{\sigma_k^4}{M_k} \quad (\text{G14})$$

With  $b = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{B}_k \mathbf{B}_k) = \sum_{k \in \mathcal{K}} \frac{\sigma_k^4}{M_k}$ ,  $c = \sum_{k \in \mathcal{K}} \text{tr}(\mathbf{C}_k \mathbf{C}_k) = \frac{P^2}{KN}$  and  $\frac{c}{b} = \frac{p}{n}$ , selecting  $p = \frac{P}{\sum_{k \in \mathcal{K}} \frac{\sigma_k^4}{M_k}}$

implies  $n = \frac{KN}{P}$ . The uplink problem becomes

$$\inf_{\mathbf{\Omega} \succeq \mathbf{0}} \left\{ \sup_{\substack{\mathbf{\Sigma} \succeq \mathbf{0} \\ P_1, \dots, P_K}} \left\{ W(\mathbf{\Sigma}, \mathbf{\Omega}) : \mathbf{\Sigma}_k \preceq \frac{P_k}{\sigma_k^2} \mathbf{I} \ \forall k, \sum_{k \in \mathcal{K}} P_k = P \right\} : \mathbf{\Omega} \preceq \mathbf{I} \right\} \quad (\text{G15})$$

$$= \sup_{\substack{\mathbf{\Sigma} \succeq \mathbf{0} \\ P_1, \dots, P_K}} \left\{ W(\mathbf{\Sigma}, \mathbf{I}) : \mathbf{\Sigma}_k \preceq \frac{P_k}{\sigma_k^2} \mathbf{I} \ \forall k, \sum_{k \in \mathcal{K}} P_k = P \right\} \quad (\text{G16})$$

Assuming users sorted matched to the encoding order we introduce  $\alpha_k = w_k - w_{k+1}$ ,  $\forall k = 1, \dots, K-1$  and  $\alpha_K = w_K$ , and the uplink problem can be rewritten as

$$\sup_{\substack{\Sigma \succeq \mathbf{0} \\ P_1, \dots, P_K}} \left\{ \sum_{k \in \mathcal{K}} \alpha_k \log \left| \mathbf{I} + \sum_{j \leq k} \mathbf{H}_{jj}^H \Sigma_j \mathbf{H}_{jj} \right| : \Sigma_k \preceq \frac{P_k}{\sigma_k^2} \mathbf{I} \forall k, \sum_{k \in \mathcal{K}} P_k = P \right\} \quad (\text{G17})$$

$$= \sup_{P_1, \dots, P_K} \left\{ \sum_{k \in \mathcal{K}} \alpha_k \log \left| \mathbf{I} + \sum_{j \leq k} \frac{P_j}{\sigma_j^2} \mathbf{H}_{jj}^H \mathbf{H}_{jj} \right| : \sum_{k \in \mathcal{K}} P_k = P \right\} \quad (\text{G18})$$

where we have used that the reformulated utility is non-decreasing in the transmit covariances and therefore  $\Sigma_k = \frac{P_k}{\sigma_k^2} \mathbf{I}$ .

#### Appendix H Subdifferential of the Mutual Information at Points where the Noise Covariances are not Invertible

*Lemma 13.* Consider a MIMO channel  $\mathbf{H}$  with given noise covariance  $\mathbf{S}$  and transmit covariance  $\mathbf{Q}$ , such that the mutual information is finite and given by

$$f(\mathbf{S}) = \log \left| \mathbf{I} + \mathbf{S}^+ \mathbf{H} \mathbf{Q} \mathbf{H}^H \right| \quad (\text{H1})$$

Let  $\mathbf{T}, \mathbf{T}^\perp$  provide a full unitary basis for the fundamental subspaces of  $\mathbf{S}$ . The subdifferential of  $f$  at  $\mathbf{S}$  is given by

$$\frac{\partial f(\mathbf{S})}{\partial \mathbf{S}} = \left\{ \Sigma : \Sigma = - \left( \mathbf{S} + \mathbf{H} \mathbf{Q} \mathbf{H}^H \right)^+ + \mathbf{S}^+ + \Sigma^\Delta, \Sigma^\Delta \in \mathcal{E}^\Delta \right\} \quad (\text{H2})$$

where

$$\mathcal{E}^\Delta = \text{conv}_j \left( \left\{ \Sigma^\Delta : \Sigma^\Delta = \mathbf{L} \mathbf{T} \mathbf{J} \mathbf{T}^{\perp H} - \mathbf{T}^\perp \mathbf{J}^H \mathbf{T}^H \mathbf{L} + \mathbf{T}^\perp \mathbf{J}^H \mathbf{T}^H \mathbf{L} \mathbf{T} \mathbf{J} \mathbf{T}^{\perp H}, \mathbf{L} = \left( \mathbf{S} + \mathbf{H} \mathbf{Q} \mathbf{H}^H \right)^+ + \mathbf{S}^+ \right\} \right) \quad (\text{H3})$$

and  $\mathbf{T}^H \Sigma^\Delta \mathbf{T} = \mathbf{0} \forall \Sigma^\Delta \in \mathcal{E}^\Delta$ .

**Proof.** According to [38, Th. 2.5.1], the subdifferential of a function  $f$  at  $\mathbf{x}$  is the convex hull of the limit points of the sequence  $\frac{\partial f(\mathbf{x}_i)}{\partial \mathbf{x}_i}$  where the limit  $\mathbf{x}_i \rightarrow \mathbf{x}$  exists and  $\mathbf{x}_i$  avoids points where  $f$  is not differentiable. We introduce

$$\tilde{\mathbf{S}} = \begin{bmatrix} \mathbf{T} & \mathbf{T}^\perp \end{bmatrix} \begin{bmatrix} \mathbf{T}^H \mathbf{S} \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{T} & \mathbf{T}^\perp \end{bmatrix}^H + \epsilon \begin{bmatrix} \mathbf{T} & \mathbf{T}^\perp \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{T} & \mathbf{T}^\perp \end{bmatrix}^H \quad (\text{H4})$$

where  $\mathbf{C} \succ \mathbf{0}$  and

$$\mathbf{\Delta} = \mathbf{A} - \mathbf{B} \mathbf{C}^{-1} \mathbf{B}^H \succeq \mathbf{0} \quad (\text{H5})$$

We use  $\tilde{\mathbf{S}}$  to parameterize  $\mathbf{S}_i$  by the choice of  $\epsilon$  and  $\mathbf{S}_i \rightarrow \mathbf{S}$  for  $\epsilon \rightarrow 0$ . As  $\tilde{\mathbf{S}} \succ \mathbf{0}$  for every  $\epsilon > 0$ , we have

$$f(\tilde{\mathbf{S}}) = \log \left| \mathbf{I} + \tilde{\mathbf{S}}^{-1} \mathbf{H} \mathbf{Q} \mathbf{H}^H \right| \quad (\text{H6})$$

$$= \log \left| \tilde{\mathbf{S}} + \mathbf{H} \mathbf{Q} \mathbf{H}^H \right| - \log |\tilde{\mathbf{S}}| \quad (\text{H7})$$



and

$$\frac{\partial f(\tilde{S})}{\tilde{S}} = (\tilde{S} + HQH^H)^{-1} - \tilde{S}^{-1} \quad (H8)$$

Using  $T^{\perp H}H = 0$  and substituting

$$X = T^HST + T^H HQH^H T \quad (H9)$$

We obtain

$$[T \ T^{\perp}]^H \frac{\partial f(\tilde{S})}{\tilde{S}} [T \ T^{\perp}] \quad (H10)$$

$$= \begin{bmatrix} X + \epsilon A & \epsilon B \\ \epsilon B^H & \epsilon C \end{bmatrix}^{-1} - \begin{bmatrix} T^HST + \epsilon A & \epsilon B \\ \epsilon B^H & \epsilon C \end{bmatrix}^{-1} \quad (H11)$$

$$= \begin{bmatrix} (X + \epsilon \Delta)^{-1} & -(X + \epsilon \Delta)^{-1} BC^{-1} \\ -C^{-1} B^H (X + \epsilon \Delta)^{-1} & \frac{1}{\epsilon} C^{-1} + C^{-1} B^H (X + \epsilon \Delta)^{-1} BC^{-1} \end{bmatrix} - \quad (H12)$$

$$- \begin{bmatrix} (T^HST + \epsilon \Delta)^{-1} & -(T^HST + \epsilon \Delta)^{-1} BC^{-1} \\ -C^{-1} B^H (T^HST + \epsilon \Delta)^{-1} & \frac{1}{\epsilon} C^{-1} - C^{-1} B^H (T^HST + \epsilon \Delta)^{-1} BC^{-1} \end{bmatrix} \quad (H13)$$

$$= \begin{bmatrix} (X + \epsilon \Delta)^{-1} - (T^HST + \epsilon \Delta)^{-1} & -[(X + \epsilon \Delta)^{-1} - (T^HST + \epsilon \Delta)^{-1}] BC^{-1} \\ -C^{-1} B^H [(X + \epsilon \Delta)^{-1} - (T^HST + \epsilon \Delta)^{-1}] & +C^{-1} B^H [(X + \epsilon \Delta)^{-1} - (T^HST + \epsilon \Delta)^{-1}] BC^{-1} \end{bmatrix} \quad (H14)$$

and

$$\lim_{\epsilon \rightarrow 0} \frac{\partial f(\tilde{S})}{\tilde{S}} = [T \ T^{\perp}] \begin{bmatrix} X^{-1} - (T^HST)^{-1} & -[X^{-1} - (T^HST)^{-1}] BC^{-1} \\ -C^{-1} B^H [X^{-1} - (T^HST)^{-1}] & C^{-1} B^H [X^{-1} - (T^HST)^{-1}] BC^{-1} \end{bmatrix} [T \ T^{\perp}]^H \quad (H15)$$

$$\begin{aligned} &= T [X^{-1} - (T^HST)^{-1}] T^H - \\ &\quad - T [X^{-1} - (T^HST)^{-1}] BC^{-1} T^{\perp H} - \\ &\quad - T^{\perp} C^{-1} B^H [X^{-1} - (T^HST)^{-1}] T^H + \\ &\quad + T^{\perp} C^{-1} B^H [X^{-1} - (T^HST)^{-1}] BC^{-1} T^{\perp H} \end{aligned} \quad (H16)$$

$$\begin{aligned} &= T [(T^HST + T^H HQH^H T)^{-1} - (T^HST)^{-1}] T^H - \\ &\quad - T [(T^HST + T^H HQH^H T)^{-1} - (T^HST)^{-1}] T^H TBC^{-1} T^{\perp H} - \\ &\quad - T^{\perp} C^{-1} B^H T^H T [(T^HST + T^H HQH^H T)^{-1} - (T^HST)^{-1}] T^H + \\ &\quad + T^{\perp} C^{-1} B^H T^H T [(T^HST + T^H HQH^H T)^{-1} - (T^HST)^{-1}] T^H TBC^{-1} T^{\perp H} \end{aligned} \quad (H17)$$

$$\begin{aligned} &= (S + HQH^H)^+ - S^+ - \\ &\quad - [(S + HQH^H)^+ - S^+] TBC^{-1} T^{\perp H} - \\ &\quad - T^{\perp} C^{-1} B^H T^H [(S + HQH^H)^+ - S^+] + \\ &\quad + T^{\perp} C^{-1} B^H T^H [(S + HQH^H)^+ - S^+] TBC^{-1} T^{\perp H} \end{aligned} \quad (H18)$$

We observe that the limit is independent of  $A$ , and when considering all possible choices for limit points, we can freely select  $B, C \succ 0$  as we can always select  $A$  such that  $\tilde{S} \succ 0$ . If  $C$  is not strictly positive,  $\tilde{S}$  is not strictly positive and  $f$  is not differentiable. Therefore, the limit point of this sequence is not relevant. We conclude that Equation (H4) allows to parameterize all relevant limit points.

We introduce the following substitutions

$$\Sigma^\Delta = L - LTJT^{\perp H} - T^\perp J^H T^H L + T^\perp J^H T^H LTJT^{\perp H} \quad (H19)$$

$$L = \left[ (S + HQH^H)^+ - S^+ \right] \quad (H20)$$

$$J = BC^{-1} \quad (H21)$$

As  $C \succ 0, B$  can have arbitrary values,  $J$  can take any value, and we can parameterize the subdifferential, given by the convex hull of the limit points, as

$$\frac{\partial f(S)}{\partial S} = \text{conv}_J \left( \left\{ L - LTJT^{\perp H} - T^\perp J^H T^H L + T^\perp J^H T^H LTJT^{\perp H}, L = (S + HQH^H)^+ - S^+ \right\} \right) \quad (H22)$$

$$= \left\{ \Sigma : \Sigma = - (S + HQH^H)^+ + S^+ - \Sigma^\Delta, \Sigma^\Delta \in \mathcal{E}^\Delta \right\} \quad (H23)$$

where

$$\mathcal{E}^\Delta = \text{conv}_J \left( \left\{ \Sigma^\Delta : \Sigma^\Delta = LTJT^{\perp H} - T^\perp J^H T^H L + T^\perp J^H T^H LTJT^{\perp H}, L = (S + HQH^H)^+ - S^+ \right\} \right) \quad (H24)$$

With  $T^H T^\perp = 0$  we have  $T^H \Sigma^\Delta T = 0 \forall \Sigma^\Delta \in \mathcal{E}^\Delta$ .  $\square$

## Appendix I Details of the Proof for the Polite Waterfilling Lemma 9

In the following, we show that every solution to Equations (121)–(123) is a solution to Equation (125)–(127) and vice versa and that these solutions fulfill the equalities of Lemma 9. There are several implications of the conditions of the lemma that we use in the following:

$$[TT^\perp] \text{ unitary and full rank} \Rightarrow T^H T = I \quad (I1)$$

$$[\Phi\Phi^\perp] \text{ unitary and full rank} \Rightarrow \Phi^H \Phi = I \quad (I2)$$

$$T^H S T \succ 0, T^{\perp H} S T^\perp = 0 \Rightarrow S^+ = T(T^H S T)^{-1} T^H \text{ and } S = T T^H S T T^H \quad (I3)$$

$$\Phi^H \Psi \Phi \succ 0, \Phi^{\perp H} \Psi \Phi^\perp = 0 \Rightarrow \Psi^+ = \Phi(\Phi^H \Psi \Phi)^{-1} \Phi \text{ and } \Psi = \Phi \Phi^H \Psi \Phi \Phi^H \quad (I4)$$

$$H\Phi^\perp = 0 \Rightarrow H = H\Phi\Phi^H \quad (I5)$$

$$T^\perp H = 0 \Rightarrow H = T T^H H \quad (I6)$$

$$H\Phi^\perp = 0, \Phi^H Q^\Delta \Phi = 0 \Rightarrow H Q^\Delta H^H = 0 \quad (I7)$$

$$T^\perp H = 0, T^H \Sigma^\Delta T = 0 \Rightarrow H^H \Sigma^\Delta H = 0 \quad (I8)$$

With  $H^H \Sigma^\Delta H = 0$  and Equation (121), we have

$$H^H \Sigma H = -w H^H (S + HQH^H)^+ H + w H^H S^+ H \quad (I9)$$

and by adding it to Equation (122), we obtain

$$\Psi + H^H \Sigma H = w H^H S^+ H + K \quad (I10)$$

Furthermore,  $\text{tr}(KQ) = 0, K \succeq 0, Q \succeq 0$  implies  $KQ = 0$ . Multiplying Equation (122) with  $\Phi^{\perp H}$  from the left and  $\Phi^\perp$  from the right and using  $H\Phi^\perp = 0$  and  $\Phi^{\perp H}\Psi\Phi^\perp = 0$  yields  $\Phi^{\perp H}K\Phi^\perp = 0$ , which implies  $K = \Phi\Phi^H K \Phi\Phi^H$ . Combined with  $KQ = 0$  and  $\Phi^H\Phi = I$  we have  $\Phi^H K \Phi\Phi^H Q = 0$ .

Now we are prepared to show that any solution of Equations (121)–(123) also fulfills Equation (125). As  $Q^\Delta$  is only constrained by  $\Phi^H Q^\Delta \Phi = 0$  and can be freely chosen otherwise, as long as  $Q \succeq 0$ , it is sufficient to verify

$$\Phi^H Q \Phi = -w \Phi^H (\Psi + H^H \Sigma H)^+ \Phi + w (\Phi^H \Psi + \Phi) \quad (I11)$$

$$\Leftrightarrow \Phi^H Q \Phi = -w (\Phi^H (\Psi + H^H \Sigma H) \Phi)^{-1} + w (\Phi^H \Psi \Phi)^{-1} \quad (I12)$$

$$\Leftrightarrow \Phi^H (\Psi + H^H \Sigma H) \Phi \Phi^H Q \Phi \Phi^H \Psi \Phi = -w \Phi^H \Psi \Phi + w \Phi^H (\Psi + H^H \Sigma H) \Phi \quad (I13)$$

$$\Leftrightarrow \Phi^H (w H^H S^+ H + K) \Phi \Phi^H Q \Phi \Phi^H \Psi \Phi = w \Phi^H H^H \Sigma H \Phi \quad (I14)$$

$$\Leftrightarrow \Phi^H (H^H S^+ H) \Phi \Phi^H Q \Phi \Phi^H \Psi \Phi = \Phi^H H^H \Sigma H \Phi \quad (I15)$$

where we used Equation (122) and  $\Phi^H K \Phi \Phi^H Q = 0$ . The left side of Equation (I15) is

$$\begin{aligned} & \Phi^H (H^H S^+ H) \Phi \Phi^H Q \Phi \Phi^H \Psi \Phi \\ &= \Phi^H (H^H S^+ H) \Phi \Phi^H Q \Phi \Phi^H (w H^H (S + H Q H^H)^+ H + K) \Phi \end{aligned} \quad (I16)$$

$$= w \Phi^H (H^H S^+ H) \Phi \Phi^H Q \Phi \Phi^H H^H (S + H Q H^H)^+ H \Phi \quad (I17)$$

$$= w \Phi^H H^H T (T^H S T)^{-1} T^H H \Phi \Phi^H Q \Phi \Phi^H H^H T (T^H S T + T^H H \Phi \Phi^H Q \Phi \Phi^H H^H T)^{-1} T^H H \Phi \quad (I18)$$

$$= w \bar{H}^H \bar{S}^{-1} \bar{H} \bar{Q} \bar{H}^H (\bar{S} + \bar{H} \bar{Q} \bar{H}^H)^{-1} \bar{H} \quad (I19)$$

where we substituted  $\bar{H} = T^H H \Phi$ ,  $\bar{S} = T^H S T$ , and  $\bar{Q} = \Phi^H Q \Phi$ . By Equation (I9), the right side of Equation (I15) can be reformulated as

$$\begin{aligned} & \Phi^H H^H \Sigma H \Phi \\ &= -w \Phi^H H^H (S + H Q H^H)^+ H \Phi + w \Phi^H H^H S^+ H \Phi \end{aligned} \quad (I20)$$

$$= -w \Phi^H H^H T (T^H S T + T^H H \Phi \Phi^H Q \Phi \Phi^H H^H T)^{-1} T^H H \Phi + w \Phi^H H^H T (T^H S T)^{-1} T^H H \Phi \quad (I21)$$

$$= -w \bar{H}^H (\bar{S} + \bar{H} \bar{Q} \bar{H}^H)^{-1} \bar{H} + w \bar{H}^H \bar{S}^{-1} \bar{H} \quad (I22)$$

Now, combining Equations (I16) and (I20) we can rewrite Equation (I15) as

$$\bar{H}^H \bar{S}^{-1} \bar{H} \bar{Q} \bar{H}^H (\bar{S} + \bar{H} \bar{Q} \bar{H}^H)^{-1} \bar{H} = -\bar{H}^H (\bar{S} + \bar{H} \bar{Q} \bar{H}^H)^{-1} \bar{H} + \bar{H}^H \bar{S}^{-1} \bar{H} \quad (I23)$$

$$\Leftrightarrow (\bar{H}^H \bar{S}^{-1} \bar{H} \bar{Q} + I) \bar{H}^H (\bar{S} + \bar{H} \bar{Q} \bar{H}^H)^{-1} \bar{H} = \bar{H}^H \bar{S}^{-1} \bar{H} \quad (I24)$$

$$\Leftrightarrow \bar{H}^H (\bar{S} + \bar{H} \bar{Q} \bar{H}^H)^{-1} \bar{H} = (\bar{H}^H \bar{S}^{-1} \bar{H} \bar{Q} + I)^{-1} \bar{H}^H \bar{S}^{-1} \bar{H} \quad (I25)$$

which holds true by the matrix inversion lemma.

Any solution of Equations (121)–(123) will satisfy Equation (126) by selecting  $M$  as

$$M = S - w H (\Psi + H^H \Sigma H)^+ H^H \quad (I26)$$

which leaves us to show that this choice is consistent with Equation (127).

By  $HQ^\Delta H^H = \mathbf{0}$  and multiplying Equation (125) with  $H$  from the left and with  $H^H$  from the right, we have

$$-wH(\Psi + H^H \Sigma H)^+ H^H = HQH^H - wH\Psi^+ H^H \quad (I27)$$

and combined with Equation (I26), we have

$$M = S + HQH^H - wH\Psi^+ H^H \quad (I28)$$

$$\Rightarrow \Sigma M = \Sigma S + \Sigma HQH^H - w\Sigma H\Psi^+ H^H \quad (I29)$$

$$\Rightarrow \text{tr}(\Sigma M) = \text{tr}(\Sigma S) + \text{tr}(\Sigma HQH^H) - w \text{tr}(\Sigma H\Psi^+ H^H) \quad (I30)$$

From Equation (121) we have

$$T^H \Sigma T = -w(T^H S T + T^H HQH^H T)^{-1} + w(T^H S T)^{-1} \quad (I31)$$

$$\Leftrightarrow T^H S T T^H \Sigma T + T^H HQH^H T T^H \Sigma T = w T^H HQH^H T (T^H S T)^{-1} \quad (I32)$$

$$\Rightarrow \text{tr}(T^H S T T^H \Sigma T) + \text{tr}(T^H HQH^H T T^H \Sigma T) = w \text{tr}(T^H HQH^H T (T^H S T)^{-1}) \quad (I33)$$

$$\Leftrightarrow \text{tr}(T T^H S T T^H \Sigma) + \text{tr}(T T^H HQH^H T T^H \Sigma) = w \text{tr}(HQH^H T (T^H S T)^{-1} T^H) \quad (I34)$$

$$\Leftrightarrow \text{tr}(\Sigma S) + \text{tr}(HQH^H \Sigma) = w \text{tr}(HQH^H S^+) \quad (I35)$$

Combining Equations (I30) and (I35) we have

$$\text{tr}(\Sigma M) = w \text{tr}(S^+ HQH^H) - w \text{tr}(\Sigma H\Psi^+ H^H) \quad (I36)$$

Thus,  $\text{tr}(\Sigma M) = 0$  holds if

$$\Phi^H H^H S^+ HQ\Phi = \Phi^H H^H \Sigma H\Psi^+ \Phi \quad (I37)$$

and

$$\Phi^{\perp H} H^H S^+ HQ\Phi^{\perp} = \Phi^{\perp H} H^H \Sigma H\Psi^+ \Phi^{\perp} \quad (I38)$$

As  $\Phi^{\perp H} H = \mathbf{0}$  and  $\Psi^+ \Phi^{\perp} = \mathbf{0}$ , it is sufficient to verify

$$\Phi^H H^H S^+ HQ\Phi = \Phi^H H^H \Sigma H\Psi^+ \Phi \quad (I39)$$

$$\Leftrightarrow \Phi^H H^H S^+ HQ\Phi = \Phi^H H^H \Sigma H\Phi (\Phi^H \Psi \Phi)^{-1} \quad (I40)$$

$$\Leftrightarrow \Phi^H H^H S^+ H\Phi \Phi^H Q\Phi \Phi^H \Psi \Phi = \Phi^H H^H \Sigma H\Phi \quad (I41)$$

which is Equation (I15) and was shown to hold true.

We have shown that every solution to Equations (121)–(123) is a solution to Equations (125)–(127) and the inverse direction follows by symmetry. It remains to show that Equations (121)–(123) and (125)–(127) imply the equalities of the Lemma. Equality (a) is Equation (I37) and a proof for (b) follows by symmetry. Equalities (c) and (d) are Equations (121) and (125). Equality (e) follows from (a) or (b). By symmetry, the same way we derived Equation (I35) we derive

$$\text{tr}(\Psi Q) + \text{tr}(H^H \Sigma H Q) = w \text{tr}(H^H \Sigma H\Psi^+) \quad (I42)$$

and combining this with Equation (I35) and (a) implies (f).

## References

1. Rashid-Farrokhi, F.; Liu, K.; Tassiulas, L. Transmit beamforming and power control for cellular wireless systems. *IEEE J. Sel. Areas Commun.* **1998**, *16*, 1437–1450.
2. Visotsky, E.; Madhow, U. Optimum beamforming using transmit antenna arrays. In Proceedings of the IEEE 49th Vehicular Technology Conference, Houston, TX, USA, 16–20 May 1999; pp. 851–856.

3. Boche, H.; Schubert, M. A general duality theory for uplink and downlink beamforming. In Proceedings of the IEEE 56th IEEE Vehicular Technology Conference (VTC 2002-Fall), Vancouver, BC, Canada, 24–28 September 2002; pp. 87–91.
4. Viswanath, P.; Tse, D. Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. *IEEE Trans. Inf. Theory* **2003**, *49*, 1912–1921.
5. Jorswieck, E.; Boche, H. Transmission strategies for the MIMO MAC with MMSE receiver: Average MSE optimization and achievable individual MSE region. *IEEE Trans. Signal Process.* **2003**, *51*, 2872–2881.
6. Shi, S.; Schubert, M.; Boche, H. Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization. *IEEE Trans. Signal Process.* **2007**, *55*, 5436–5446.
7. Hunger, R.; Joham, M.; Utschick, W. On the MSE-duality of the broadcast channel and the multiple access channel. *IEEE Trans. Signal Process.* **2009**, *57*, 698–713.
8. Jindal, N.; Vishwanath, S.; Goldsmith, A. On the duality of Gaussian multiple-access and broadcast channels. *IEEE Trans. Inf. Theory* **2004**, *50*, 768–783.
9. Vishwanath, S.; Jindal, N.; Goldsmith, A. Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels. *IEEE Trans. Inf. Theory* **2003**, *49*, 2658–2668.
10. Yu, W. Uplink-downlink duality via minimax duality. *IEEE Trans. Inf. Theory* **2006**, *52*, 361–374.
11. Hunger, R.; Joham, M. A general rate duality of the MIMO multiple access channel and the MIMO broadcast channel. In Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM), New Orleans, LA, USA, 30 November–4 December 2008.
12. Liu, A.; Liu, Y.; Xiang, H.; Luo, W. Duality, Polite Water-Filling, and Optimization for MIMO B-MAC Interference Networks and Itree Networks. **2010**, arXiv:1004.2484.
13. Liu, A.; Liu, Y.; Xiang, H.; Luo, W. Polite water-filling for weighted sum-rate maximization in MIMO B-MAC networks under multiple linear constraints. *IEEE Trans. Signal Process.* **2012**, *60*, 834–847.
14. Yu, W.; Lan, T. Transmitter optimization for the multi-antenna downlink with per-antenna power constraints. *IEEE Trans. Signal Process.* **2007**, *55*, 2646–2660.
15. Huh, H.; Papadopoulos, H.C.; Caire, G. Multiuser MISO transmitter optimization for intercell interference mitigation. *IEEE Trans. Signal Process.* **2010**, *58*, 4272–4285.
16. Zhang, L.; Zhang, R.; Liang, Y.C.; Xin, Y.; Poor, H. On Gaussian MIMO BC-MAC duality with multiple transmit covariance constraints. *IEEE Trans. Inf. Theory* **2012**, *58*, 2064–2078.
17. Jorswieck, E.A.; Boche, H. Performance analysis of capacity of MIMO systems under multiuser interference based on worst-case noise behavior. *EURASIP J. Wirel. Commun. Netw.* **2004**, *2004*, 273–285.
18. Vishwanath, S.; Boyd, S.; Goldsmith, A. Worst-case capacity of Gaussian vector channels. In Proceedings of the 2003 Canadian Workshop on Information Theory, Waterloo, ON, Canada, 18–21 May 2003.
19. Loyka, S.; Charalambous, C. On the compound capacity of a class of MIMO channels subject to normed uncertainty. *IEEE Trans. Inf. Theory* **2012**, *58*, 2048–2063.
20. Wiesel, A.; Eldar, Y.; Shamai, S. Optimization of the MIMO compound capacity. *IEEE Trans. Wirel. Commun.* **2007**, *6*, 1094–1101.
21. Bengtsson, M.; Ottersten, B. Optimal downlink beamforming using semidefinite optimization. In Proceedings of the 37th Allerton Conference on Communication, Control, and Computing, Urbana-Champaign, IL, USA, 22–24 September 1999; pp. 987–996.
22. Wiesel, A.; Eldar, Y.; Shamai, S. Linear precoding via conic optimization for fixed MIMO receivers. *IEEE Trans. Signal Process.* **2006**, *54*, 161–176.
23. Palomar, D. Unified framework for linear MIMO transceivers with shaping constraints. *IEEE Commun. Lett.* **2004**, *8*, 697–699.
24. Scutari, G.; Palomar, D.; Barbarossa, S. MIMO cognitive radio: A game theoretical approach. In Proceedings of the 9th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Recife, Brazil, 6–9 July 2008; pp. 426–430.
25. Hammarwall, D.; Bengtsson, M.; Ottersten, B. On downlink beamforming with indefinite shaping constraints. *IEEE Trans. Signal Process.* **2006**, *54*, 3566–3580.
26. Hellings, C.; Weiland, L.; Utschick, W. Optimality of proper signaling in Gaussian MIMO broadcast channels with shaping constraints. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 3474–3478.

27. Lameiro, C.; Santamaria, I.; Utschick, W. Interference shaping constraints for underlay MIMO interference channels. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 7313–7317.
28. Chen, L.; You, S. The weighted sum rate maximization in MIMO interference networks: The minimax Lagrangian duality and algorithm. **2013**, arXiv: 1309.4034.
29. Dotzler, A.; Newinger, M.; Utschick, W. Covariance shapes as a cognitive radio concept for receivers with multiple antennas. In Proceedings of the 19th International ITG Workshop on Smart Antennas (WSA), Ilmenau, Germany, 3–5 March 2015.
30. Newinger, M.; Dotzler, A.; Utschick, W. Interference shaping for device-to-device communication in cellular networks. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 4120–4125.
31. Brunner, H.H.; Dotzler, A.; Utschick, W.; Nossek, J.A. Weighted sum rate maximization with multiple linear conic constraints. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 4635–4640.
32. Brunner, H.H.; Dotzler, A.; Utschick, W.; Nossek, J.A. Intercell interference robustness tradeoff with loosened covariance shaping. In Proceedings of the 18th International ITG Workshop on Smart Antennas (WSA), Erlangen, Germany, 12–13 March 2014.
33. Hellings, C.; Utschick, W. On carrier-cooperation in parallel gaussian mimo relay channels with partial decode-and-forward. In Proceedings of the 48th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 2–5 November 2014; pp. 220–224.
34. Hellings, C.; Gerdes, L.; Weiland, L.; Utschick, W. On optimal Gaussian signaling in MIMO relay channels with partial decode-and-forward. *IEEE Trans. Signal Process.* **2014**, *62*, 3153–3164.
35. Ramana, M.; Goldman, A.J. Some geometric results in semidefinite programming. *J. Glob. Optim.* **1995**, *7*, 33–50.
36. Dotzler, A.; Riemensberger, M.; Utschick, W.; Dietl, G. Interference robustness for cellular MIMO networks. In Proceedings of the 13th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cesme, Turkey, 17–20 June 2012; pp. 229–233.
37. Sato, H. An outer bound to the capacity region of broadcast channels. *IEEE Trans. Inf. Theory* **1978**, *24*, 374–377.
38. Clarke, F.H. *Optimization and Nonsmooth Analysis*; Classics in Applied Mathematics; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1990.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).