On the MSE-Duality of the Broadcast Channel and the Multiple Access Channel

Raphael Hunger, *Student Member, IEEE*, Michael Joham, *Member, IEEE*, and Wolfgang Utschick, *Senior Member, IEEE*

***Abstract*—We present a mean-square-error (MSE) duality between the broadcast channel and the multiple access channel for multiantenna users communicating with a single base station. We introduce three levels of the duality which allow for a problem specific customization with different computational complexities and resolutions. The first level preserves the *sum*-MSE during the conversion from the uplink to the downlink and vice versa, whereas the second level not only keeps the sum-MSE constant but also ensures the preservation of the individual users' MSEs. The third level involves the finest resolution and preserves the individual streams' MSEs and the individual streams' signal-to-interference-and-noise ratio (SINR) simultaneously. In contrast with hitherto existing MSE-dualities, the proposed sort of duality features a lower complexity since no MSE computation detouring is necessary during the conversion to the dual domain and is capable of handling all combinations of active and passive transmitters and receivers. Moreover, we show how two of these three dualities can be exploited to solve the unweighted total sum-MSE minimization problem and the weighted sum-MSE minimization in the broadcast channel in an efficient way by revealing the hidden convexity in the first case and drastically reducing the computational complexity in the latter case.**

***Index Terms*—Broadcast channel (BC), decentralized receivers, duality, joint minimum mean-square error (MMSE) filtering, linear precoding, projected gradient.**

I. Introduction

During the last few years, dualities have gained in importance in signal processing and information theory. The basic idea which is behind a duality consists in the conversion of an existing system setup to a dual system that has some fundamental properties in common with the original system but features additional attributes that can be exploited. Among those attributes are for example a better mathematical structure, the revealing of a hidden convexity [1], and the reduction of the computational complexity.

A. Literature Review of Existing Dualities

Three types of dualities can be found in the literature: The first one implies the equality of the signal-to-interference-and-noise ratio (SINR) region representing all feasible tuples of SINRs in a nondegraded downlink broadcast channel (BC) under a sum-power constraint and the SINR region of the dual multiple-access channel (MAC) in the uplink. Given a downlink system setup, the duality allows to construct a virtual dual uplink with reversed signal flow where all streams feature the same individual SINRs and vice versa. Early work in this field focused on the signal-to-interference ratio (SIR) neglecting the noise amount, see, for example, [2], [3], or [4]. Rashid-Farrokhi *et al.* [5] were the first introducing the SINR duality concept for nonvanishing noise in the context of power minimization under minimum SINR requirements. In [6], Visotsky *et al.* constructed a virtual uplink channel by normalizing the users' channels. However, an explicit duality framework was not presented yet, since duality aspects were proven only for simultaneously optimum beamformers and power allocation. Boche *et al.* proved the uplink-downlink SINR duality for general unit-norm beamforming vectors [7], [8], whose power allocations have to fulfill a balancing of certain SINR ratios, i.e., the SINR targets have to be set to the currently achieved SINRs. At the same time, Viswanath *et al.* derived the SINR duality between the MAC and the BC from the duality of the MIMO point-to-point system [9]. Up to that time, all dualities were applicable only for single antenna receivers. This restriction was eliminated by Tse *et al.* in [10], who extended the SINR duality to single-stream transmission for multiantenna receivers/transmitters. An interesting property that was observed already in the first contributions on SINR duality is, that when applying a set of unit-norm beamforming vectors in the downlink with an arbitrary power allocation, the achieved SINR tuple in the downlink can also be achieved in the dual uplink by means of *the same set* of beamformers as *receive filters*, but *different* power allocation at the single antenna transmitters in the dual uplink. This property is exploited by the authors in [11] to find the first "stand-alone" SINR duality with the lowest complexity available. Note that the SINR duality in [11] can be seen as a byproduct of the third level of the proposed mean-square-error (MSE) duality. Based on the SINR duality concept, the quality-of-service (QoS) sum-power minimization problem subject to minimum SINR-requirements and the *balancing* problem with given relative SINR ratios could be solved, see, e.g., [7], [12]–[16] and [12], [13], [16], respectively. For the case of nonlinear filtering, duality was shown in [9] and [17] making use of Costa's dirty paper coding [18]. The power minimization problem with nonlinear filtering was then tackled in [19] and [20].

The second kind of duality is the MSE duality where the MSEs remain constant during the conversion from uplink to downlink and vice versa. Many existing MSE dualities for linear filtering (e.g., [14], [15], [21], and [22]) are deduced from the SINR duality. Thus, such MSE-dualities fail to work

Manuscript received November 16, 2007; revised August 22, 2008. First published October 31, 2008; current version published January 30, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Daniel P. Palomar.

The authors are with the Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany (e-mail: hunger@tum.de).

Digital Object Identifier 10.1109/TSP.2008.2008253

for switched off data-streams, i.e., when there exist SINRs, that are zero. In SINR optimizations, this case does not occur, but for many MSE based optimization criteria such as joint sum-MSE minimization for example, it can be observed that individual data streams or even users are switched off if the SNR and channel quality are too low, cf. [23] and [24]. Another effect resulting from the SINR deduction is that the MSE of every user specific *stream* is maintained. The very first *direct* MSE duality based on a layer-wise MSE preservation was derived in [25]. Clearly, not only the individual streams' MSEs remain constant during the conversion, but also the user-wise MSEs and the sum-MSE do not change when the MSEs of all streams do not change. However, QoS requirements often imply requirements that are formulated user-wise, as several streams may be associated to a single user who can jointly decode them. A layer-wise MSE-duality therefore is not always necessary and even leads to a computationally more complex description of the system. The first *user-wise* MSE-duality was introduced by the authors in [16] and is proven without SINR detouring. Later, in [26] we generalized the framework from [16], treating precoders, receivers, and channels as operators in Hilbert space. For the MSE-duality of nonlinear systems, see [21], [27], [28], and [29]. But up to now, even the direct MSE dualities fail, if streams at the transmitter are switched off or receivers completely ignore streams while the other end of the link has an active receive filter or an active transmit filter, respectively.

Finally, a rate-region duality can be established in case of nonlinear filtering and Gaussian signaling. The duality between the MAC and the BC for single antenna terminals or single stream transmission with multiantenna terminals with fixed precoder/receive filters already follows from the SINR duality. The first direct duality between the Gaussian MAC and BC for single-antenna users and a single-antenna base station was given in [30] and earlier versions of it, and the first direct *sum* capacity duality proof for a multiantenna base station was given in [9] by showing that the achievable sum-rate with Costa precoding merges with the maximum sum-rate in the MAC. Vishwanath *et al.* came up with an extension to multiantenna terminals supporting an arbitrary number of data streams per user in [31]. Weingarten *et al.* afterwards proved that the dirty paper rate region exactly corresponds to the capacity region of the MIMO BC [32]. Duality in terms of outage capacity, minimum-rate capacity, and ergodic capacity for fading channels was derived in [33] by Jindal *et al.* Thanks to the duality, broadcast problems can be transferred to the MAC, where the maximization of the sum-rate boils down to a convex problem, see [34] and [35].

B. Rationale for MSE-Based Optimizations

Taking the MSE as the figure of merit offers some decisive advantages compared to SINR-based problem formulations for example. In combination with a minimum mean-square-error (MMSE) receiver, the streamwise MSE formulation inherits all properties of an SINR-based description by the bijective mapping

$$\mathrm{MMSE} = \frac{1}{1+\mathrm{SINR}}. \tag{1}$$

The standard SINR-based rate expression $\log_2(1+\mathrm{SINR})$ can conveniently be expressed by the MMSE via $-\log_2(\mathrm{MMSE})$, so rate-based optimizations can also be transformed into MSE-based optimizations. A user-wise description of the SINR does not exist as in the MSE case, there is no single scalar describing the link quality of a multistream transmission between a user and a base station. Indeed, the SINRs of the individual streams of a user could be computed. However, the significance of the resulting values is limited as different streams of a single user are treated as interference, which becomes impractical if the receivers decode their streams jointly. Unlike the SINR approach, the user-wise MSE description allows to lower bound/approximate the rate of the individual users when they jointly decode their streams: With MMSE receivers, the capacity of a user's link can be rewritten as the negative logarithm of the determinant of the MMSE error covariance matrix, which itself can be lower bounded/approximated by a function which is monotonically increasing in the MSE achieved by this user, see [16]. Hence, minimizing the MSE of a user maximizes a (tight) lower bound of his capacity. Finally, a system-wide sum-MSE description is also attractive as the minimization of this metric yields excellent uncoded bit-error ratios.

C. Contributions

We present an application specific MSE duality between the broadcast channel and the multiple access channel consisting of three different kinds corresponding to three different resolutions of the MSE conservation. In the first and likewise simplest level, the overall users *sum*-MSE is conserved. The application of this kind of duality to the sum-MSE minimization problem in the broadcast channel allows for a solution with minimum complexity. Hitherto existing dualities were targeted to the preservation of the individual streams' MSEs involving a drastically higher complexity to solve this problem. A derivation of this kind of MSE duality from an SINR duality is not possible, since any SINR duality naturally exhibits a streamwise formulation, whereas only the total sum-MSE is of interest here. The second kind features a finer resolution and preserves not only the sum-MSE, but also the MSEs of the *individual* users. To this end, K scalars are deduced from a linear system of equations, where K represents the number of users in the system. Due to the application specific nature of our duality, these scalars are a byproduct of the weighted sum-MSE minimization problem in the downlink and thus need not be computed explicitly. Again, the duality requires almost no computational effort. And note that there is no way to derive this kind of duality from any existing SINR duality due to the same reasoning as above. The finest resolution is obtained by the third kind where not only the users' MSEs are preserved, but also the MSEs of the individual *streams* of the multiantenna receivers. In principle, this third kind is similar to the existing (layer-wise) MSE dualities in [22], [25], and [29], and, as has been shown in [22], a layer-wise MSE duality results from the SINR duality *if no streams are switched off*. This level of duality becomes attractive if one focuses on rate optimization problems where the joint decoding of all streams together is not desired and separate decoding per stream has to be preferred. The reason for this follows from the fact that the third kind of the proposed duality leaves the SINRs of the individual streams unchanged and hence, the rates do not change when switching from one domain to its dual under the assumption of Gaussian signaling. The third kind, therefore, has a twofold functionality.

To the best of our knowledge, our duality is not only characterized by the smallest complexity and a direct conversion of the transmit and receive beamformers from one domain to its dual, but also the first one being capable of handling not supported data streams and/or not supported users correctly. Hitherto existing dualities have to spend extra computation time on the calculation of the currently achieved MSE values during an intermediate step although only the transmit and receive beamformers in the dual domain are of interest. Moreover, they fail if a single stream or user is actively transmitting and the respective receive filter is switched off or if a stream or user is actively receiving although the respective transmitter did not send anything at all. This generalizes previous dualities where such a case had to be intercepted. Finally, we present some applications where the proposed duality can be implemented and highlight the resulting benefits.

D. Organization

The remainder of this contribution is organized as follows: In Section II, we describe the system model of the uplink and downlink underlying our multiuser scenario. A brief overview of the three levels offered by the proposed duality is presented in Section III by turning our attention to the principles behind them and their inherent properties. Having presented the three kinds of uplink-to-downlink conversion for the strictly active users in the first part of Section IV, we extend it to the generalized case with passive receivers or transmitters in the Appendix. The downlink-to-uplink conversion counterpart is shown in Section V in a slightly shortened version due to its similarity. In Section VI, we mention two possible applications for two different kinds of the duality. The first one is the *total* sum-MSE minimization in the broadcast channel, where the problem can efficiently be solved by revealing its hidden convexity. The second application is concerned with the *weighted* sum-MSE minimization, where a weighted sum of the users' MSEs is minimized and the duality allows to reduce the computational complexity.

Notation

Throughout this paper, we use the following operations and abbreviations: Matrices and vectors are upper and lower case bold, respectively. $\mathbb{C}$ denotes the set of complex-valued numbers, $\mathrm{E}[\cdot]$ means expectation with respect to symbols and noise. The operators $\|\cdot\|_2$, $\|\cdot\|_\mathrm{F}$, $(\cdot)^\mathrm{T}$, $(\cdot)^\mathrm{H}$, and $\mathrm{tr}(\cdot)$ stand for Euclidean norm, Frobenius norm, transposition, Hermitian transposition, and trace of a matrix, respectively. $\mathbf{I}_N$ represents the $N \times N$ identity matrix.

II. System Model

A. Downlink Description

The downlink of the K-user MIMO broadcast channel with different messages is shown in Fig. 1. There, K decentralized users are served by a centralized base station, which assigns a symbol vector $\boldsymbol{s}_k \in \mathbb{S}^{B_k}$ taken from the symbol-alphabet $\mathbb{S}$ to each user $k \in \{1,\ldots,K\}$. Here, B_k denotes the number of streams allocated to user k. Spatial filtering is applied by means of precoding matrices $\boldsymbol{P}_k \in \mathbb{C}^{N\times B_k}$ to form the broadcast transmit vector

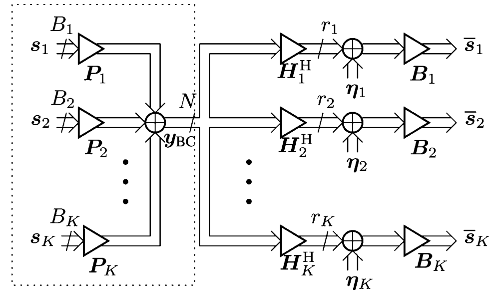

Fig. 1. Downlink system model.

$$\boldsymbol{y}_\mathrm{BC} = \sum_{i=1}^{K} \boldsymbol{P}_i \boldsymbol{s}_i \in \mathbb{C}^N \tag{2}$$

dissipating a symbol-averaged power

$$\mathrm{E}\left[\|\boldsymbol{y}_\mathrm{BC}\|_2^2\right] = \sum_{i=1}^{K} \|\boldsymbol{P}_i\|_\mathrm{F}^2 = P_\mathrm{Tx} \tag{3}$$

where we made use of the common assumption that the symbol vectors $\boldsymbol{s}_1,\ldots,\boldsymbol{s}_K$ are mutually uncorrelated with identity covariance matrix, i.e., $\mathrm{E}\left[\boldsymbol{s}_k\boldsymbol{s}_k^\mathrm{H}\right] = \mathbf{I}_{B_k}\ \forall k$. The propagation over the frequency flat channel to user k is described by the matrix $\boldsymbol{H}_k^\mathrm{H} \in \mathbb{C}^{r_k\times N}$, where r_k denotes the number of receiving antennas at mobile k. In the downlink, every precoded symbol $\boldsymbol{P}_i\boldsymbol{s}_i, i \in \{1,\ldots,K\}$, propagates over the *same* channel $\boldsymbol{H}_k^\mathrm{H}$ to user k. Zero-mean white noise $\boldsymbol{\eta}_k \in \mathbb{C}^{r_k}$ with covariance matrix $\mathrm{E}\left[\boldsymbol{\eta}_k\boldsymbol{\eta}_k^\mathrm{H}\right] = \sigma_\eta^2\mathbf{I}_{r_k}$ is added before the receiving filter $\boldsymbol{B}_k \in \mathbb{C}^{B_k\times r_k}$ generates the continuous symbol estimate $\bar{\boldsymbol{s}}_k \in \mathbb{C}^{B_k}$. Finally, this estimate is passed to the quantizer which maps the continuous $\bar{\boldsymbol{s}}_k$ to the finite alphabet $\mathbb{S}^{B_k}$. In case of an OFDM system, each frequency chunk has its own channel matrix describing the propagation in the respective frequency band. We avoid the notation of an additional superscript describing the frequency chunk for the sake of readability and because of the circumstance that the individual frequency chunks do not interfere. All dualities presented in this paper remain valid in an OFDM system, the conversions simply have to be applied for every chunk separately. Introducing $\bar{\boldsymbol{s}} = [\bar{\boldsymbol{s}}_1^\mathrm{T},\ldots,\bar{\boldsymbol{s}}_K^\mathrm{T}]^\mathrm{T} \in \mathbb{C}^{\sum_{k=1}^K B_k}$ as the stacked symbol estimates, and defining a system-wide precoding matrix $\boldsymbol{P} = [\boldsymbol{P}_1,\ldots,\boldsymbol{P}_K] \in \mathbb{C}^{N\times\sum_{k=1}^K B_k}$, a system-wide channel matrix $\boldsymbol{H} = [\boldsymbol{H}_1,\ldots,\boldsymbol{H}_K] \in \mathbb{C}^{N\times\sum_{l=1}^K r_k}$, and a system-wide blockdiagonal receive matrix $\boldsymbol{B} = \mathrm{blockdiag}\{\boldsymbol{B}_k\}_{k=1}^K \in \mathbb{C}^{\sum_{k=1}^K B_k\times\sum_{k=1}^K r_k}$, the complete receive signal $\bar{\boldsymbol{s}}$ can be expressed as (cf. Fig. 1)

$$\bar{\boldsymbol{s}} = \boldsymbol{B}\boldsymbol{H}^\mathrm{H}\boldsymbol{P}\boldsymbol{s} + \boldsymbol{B}\boldsymbol{\eta} \in \mathbb{C}^{\sum_{k=1}^K B_k},$$

where $\boldsymbol{s}$ and $\boldsymbol{\eta}$ are defined similar to $\bar{\boldsymbol{s}}$. Thus, the system-wide *sum*-MSE $\varepsilon^{\mathrm{DL}} = \sum_{k=1}^{K} \varepsilon_k^{\mathrm{DL}} = \mathrm{E}\left[\|\boldsymbol{s}-\bar{\boldsymbol{s}}\|_2^2\right]$ reads as

$$\varepsilon^{\mathrm{DL}} = \sum_{k=1}^{K} B_k - \mathrm{tr}\left[\boldsymbol{B}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P} + \boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{B}^{\mathrm{H}}\right] + \mathrm{tr}\left[\boldsymbol{B}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{B}^{\mathrm{H}} + \sigma_\eta^2 \boldsymbol{B}\boldsymbol{B}^{\mathrm{H}}\right]. \quad (4)$$

For a higher resolution, the individual user-MSEs need to be resolved and the symbol estimates $\bar{\boldsymbol{s}}_k$ have to be computed individually for every user. For user k, we find

$$\bar{\boldsymbol{s}}_k = \boldsymbol{B}_k \boldsymbol{H}_k^{\mathrm{H}} \sum_{i=1}^{K} \boldsymbol{P}_i \boldsymbol{s}_i + \boldsymbol{B}_k \boldsymbol{\eta}_k$$

and derive its downlink MSE $\varepsilon_k^{\mathrm{DL}} := \mathrm{E}\left[\|\boldsymbol{s}_k - \bar{\boldsymbol{s}}_k\|_2^2\right]$ to

$$\varepsilon_k^{\mathrm{DL}} = \mathrm{tr}\left[\mathbf{I}_{B_k} - \boldsymbol{B}_k\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{P}_k - \boldsymbol{P}_k^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{B}_k^{\mathrm{H}}\right] + \mathrm{tr}\left[\boldsymbol{B}_k\boldsymbol{H}_k^{\mathrm{H}}\sum_{i=1}^{K}\boldsymbol{P}_i\boldsymbol{P}_i^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{B}_k^{\mathrm{H}} + \sigma_\eta^2\boldsymbol{B}_k\boldsymbol{B}_k^{\mathrm{H}}\right]. \quad (5)$$

Because of the invariance property of the MSE $\varepsilon_k^{\mathrm{DL}}$ with respect to a unitary similarity transformation, the transmit and receive matrices sets $\{\boldsymbol{P}_1\boldsymbol{U}_1,\ldots,\boldsymbol{P}_K\boldsymbol{U}_K\}$ and $\{\boldsymbol{U}_1^{\mathrm{H}}\boldsymbol{B}_1,\ldots,\boldsymbol{U}_K^{\mathrm{H}}\boldsymbol{B}_K\}$ achieve the same individual user MSEs as the sets $\{\boldsymbol{P}_1,\ldots,\boldsymbol{P}_K\}$ and $\{\boldsymbol{B}_1,\ldots,\boldsymbol{B}_K\}$ for unitary $\boldsymbol{U}_1,\ldots,\boldsymbol{U}_K$ without increasing the power. For example, this can be exploited to obtain stream-balanced MSEs, see [36] and [37].

The finest granularity is obtained from resolving the MSEs per data stream. With $\boldsymbol{b}_{k,\ell}^{\mathrm{H}}$ denoting the ℓth row of the matrix $\boldsymbol{B}_k$ corresponding to the ℓth data stream, and $\boldsymbol{p}_{k,\ell}$ denoting the ℓth column of $\boldsymbol{P}_k$, the MSE of the ℓth stream of user k reads as

$$\varepsilon_{k,\ell}^{\mathrm{DL}} = 1 - \boldsymbol{b}_{k,\ell}^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{p}_{k,\ell} - \boldsymbol{p}_{k,\ell}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{b}_{k,\ell} + \boldsymbol{b}_{k,\ell}^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\sum_{i=1}^{K}\sum_{\ell=1}^{B_i}\boldsymbol{p}_{i,\ell}\boldsymbol{p}_{i,\ell}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{b}_{k,\ell} + \sigma_\eta^2\|\boldsymbol{b}_{k,\ell}\|_2^2. \quad (6)$$

B. Uplink Description

The dual uplink model is obtained by switching the roles of transmitters and receivers. Thus, we end up with the *multiple access channel* (MAC), where K decentralized users send their different messages $\boldsymbol{s}_1,\ldots,\boldsymbol{s}_K$ to one centralized receiver. Now, the precoders are denoted by $\boldsymbol{T}_k \in \mathbb{C}^{r_k\times B_k}$, and the channel from user k needs to be (Hermitian[1]) transposed for dimension matching. Consequently, the MAC receive signal $\boldsymbol{y}_{\mathrm{MAC}}$ reads as

$$\boldsymbol{y}_{\mathrm{MAC}} = \sum_{i=1}^{K}\boldsymbol{H}_i\boldsymbol{T}_i\boldsymbol{s}_i + \boldsymbol{n} \in \mathbb{C}^N \quad (7)$$

where N now represents the number of receiving antennas deployed at the base station. In the dual model the same symbol-averaged power

$$\mathrm{E}\left[\sum_{i=1}^{K}\|\boldsymbol{T}_i\boldsymbol{s}_i\|_2^2\right] = \sum_{i=1}^{K}\|\boldsymbol{T}_i\|_{\mathrm{F}}^2 = P_{\mathrm{Tx}} \quad (8)$$

[1]This is just for a more convenient notation—conventional transposition would suffice as well.

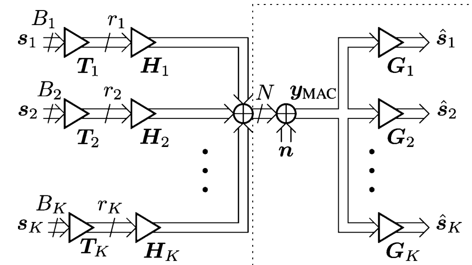

Fig. 2. Uplink system model.

shall be transmitted. Since a sum-power constraint for decentralized users is of little practical use, the dual uplink is only virtual and not a model for the link with transmission in the opposite direction of the BC. It features, however, some nice properties that can be exploited, like the revealing of a hidden convexity [1] or a reduced complexity during the filter computation. Instead of K noise vectors $\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_K$, a single vector $\boldsymbol{n}\in\mathbb{C}^N$ with covariance $\mathrm{E}\left[\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}\right] = \sigma_\eta^2\mathbf{I}_N$ is added. Linear filtering by means of the matrix $\boldsymbol{G}_k \in \mathbb{C}^{B_k\times N}$ then delivers the symbol estimate $\hat{\boldsymbol{s}}_k \in \mathbb{C}^{B_k}$ of user k.

Stacking all receive vectors yields the total receive vector $\hat{\boldsymbol{s}} = [\hat{\boldsymbol{s}}_1^{\mathrm{T}},\ldots,\hat{\boldsymbol{s}}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{\sum_{k=1}^K B_k}$. Composing the system-wide matrices $\boldsymbol{G} = [\boldsymbol{G}_1^{\mathrm{T}},\ldots,\boldsymbol{G}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{\sum_{k=1}^K B_k\times N}$ and $\boldsymbol{T} = \mathrm{blockdiag}\{\boldsymbol{T}_k\}_{k=1}^K \in \mathbb{C}^{\sum_{k=1}^K r_k\times\sum_{k=1}^K B_k}$, the total receive vector reads as (cf. Fig. 2)

$$\hat{\boldsymbol{s}} = \boldsymbol{G}\boldsymbol{H}\boldsymbol{T}\boldsymbol{s} + \boldsymbol{G}\boldsymbol{n} \in \mathbb{C}^{\sum_{k=1}^K B_k}.$$

Similar to the downlink, the system-wide uplink *sum*-MSE $\varepsilon^{\mathrm{UL}} = \sum_{k=1}^K \varepsilon_k^{\mathrm{UL}} = \mathrm{E}[\|\boldsymbol{s}-\hat{\boldsymbol{s}}\|_2^2]$ reads as

$$\varepsilon^{\mathrm{UL}} = \sum_{k=1}^{K} B_k - \mathrm{tr}\left[\boldsymbol{G}\boldsymbol{H}\boldsymbol{T} + \boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right] + \mathrm{tr}\left[\boldsymbol{G}\boldsymbol{H}\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}} + \sigma_\eta^2\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\right]. \quad (9)$$

A finer resolution is again obtained by resolving the individual users' MSEs. The symbol estimate $\hat{\boldsymbol{s}}_k$ of user k in the dual MAC can be written as

$$\hat{\boldsymbol{s}}_k = \boldsymbol{G}_k\sum_{i=1}^{K}\boldsymbol{H}_i\boldsymbol{T}_i\boldsymbol{s}_i + \boldsymbol{G}_k\boldsymbol{n} \in \mathbb{C}^{B_k}. \quad (10)$$

Thus, we obtain for the uplink MSE $\varepsilon_k^{\mathrm{UL}} := \mathrm{E}\left[\|\boldsymbol{s}_k - \hat{\boldsymbol{s}}_k\|_2^2\right]$ of user k [cf. (5)]

$$\varepsilon_k^{\mathrm{UL}} = \mathrm{tr}\left[\mathbf{I}_{B_k} - \boldsymbol{G}_k\boldsymbol{H}_k\boldsymbol{T}_k - \boldsymbol{T}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}}\right] + \mathrm{tr}\left[\boldsymbol{G}_k\sum_{i=1}^{K}\boldsymbol{H}_i\boldsymbol{T}_i\boldsymbol{T}_i^{\mathrm{H}}\boldsymbol{H}_i^{\mathrm{H}}\boldsymbol{G}_k^{\mathrm{H}} + \sigma_\eta^2\boldsymbol{G}_k\boldsymbol{G}_k^{\mathrm{H}}\right]. \quad (11)$$

Again, the invariance property of the user-wise MSEs with respect to unitary matrices $\boldsymbol{U}_k$ applied from the RHS to $\boldsymbol{T}_k$ and $\boldsymbol{U}_k^{\mathrm{H}}$ applied from the LHS to $\boldsymbol{G}_k$ is valid and allows for a distribution of the MSE of every user onto its individual streams according to the Schur-Horn theorem (e.g., [37]).

For the third level of our duality, we need to resolve the MSEs in a streamwise fashion, so we define the ℓth row of $\boldsymbol{G}_k$ as $\boldsymbol{g}_{k,\ell}^{\mathrm{H}}$ and the ℓth column of $\boldsymbol{T}_k$ as $\boldsymbol{t}_{k,\ell}$. This results in an MSE expression for the ℓth stream of user k which reads as

$$\begin{aligned}\varepsilon_{k,\ell}^{\mathrm{UL}} &= 1 - \boldsymbol{g}_{k,\ell}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{k,\ell} - \boldsymbol{t}_{k,\ell}^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{g}_{k,\ell} \\ &\quad + \boldsymbol{g}_{k,\ell}^{\mathrm{H}}\sum_{i=1}^{K}\boldsymbol{H}_i\sum_{\ell=1}^{B_i}\boldsymbol{t}_{i,\ell}\boldsymbol{t}_{i,\ell}^{\mathrm{H}}\boldsymbol{H}_i^{\mathrm{H}}\boldsymbol{g}_{k,\ell} + \sigma_\eta^2\|\boldsymbol{g}_{k,\ell}\|_2^2. \quad (12)\end{aligned}$$

III. Principle and Key Properties of the Three-Level Duality

The switching of the roles between uplink and downlink is reflected by interchanging every transmit filter in the uplink with the respective receive filter in the downlink and by choosing the transmit filters from the downlink as the respective receive filters in the uplink. As the dual domain has to consume the same amount of transmit power, we need to weight every stream's precoder with a scaling factor and the respective receive filter with the reciprocal scalar such that (3) and (8) hold simultaneously. Depending on whether we set all scaling factors to the same value, whether we allow for K different factors, one for each user, or whether we spend different factors for every stream, we arrive at different levels of the proposed duality. These different levels allow to customize the duality to a specific application by achieving different resolutions. They have the following features:

1) Level 1: System-Wide Sum-MSE Preservation: In its simplest form, the duality preserves only the total sum-MSE of the complete system when switching from uplink to downlink and vice versa. In general, the distribution of the sum-MSE onto the individual users' MSEs changes when switching from one domain to the other. Only a single scalar needs to be computed leading to an extremely low computational complexity. This scalar and its inverse are associated to the precoder and the receive filter of every stream, respectively. As a possible application of the first level of our duality, we mention the joint sum-MSE minimization in the broadcast channel. There, applying the first level of the duality, i.e., the switching from the broadcast to the multiple-access channel with one common scalar, directly reveals the hidden convexity of the problem and therefore allows for a convenient solution, see Section VI-A for further details.

2) Level 2: User-Wise MSE Preservation: Preserving the MSE of every *user* can be obtained by assigning different scalars to the individual users, but each user applies the same scalar for every single stream belonging to him. These K scalars are computed from a linear system of equations leading to a higher computational complexity than the one in the first level. However, not only the sum-MSE remains the same during the uplink-downlink conversion, but also the MSEs per user do not change. A possible area of application to which the second level of the proposed duality is tailored to is the *weighted* sum-MSE minimization, where the weighted sum of the users' MSEs is optimized. A direct minimization in the downlink is of course possible as well. However, the optimum receive filters $\boldsymbol{B}_1,\ldots,\boldsymbol{B}_K$ involve different inverses leading to a high computational complexity. Instead, the receive matrices $\boldsymbol{G}_1,\ldots,\boldsymbol{G}_K$ in the dual uplink all feature the same inverse resulting in a reduced complexity. Moreover, the K different scaling factors need not be computed explicitly in case of the weighted sum-MSE minimization because the ratio of these weights already determines the ratio of the squared scaling factors. In a nutshell, the uplink downlink conversion is for free in this case, so solving the problem in the dual uplink is clearly advantageous, see Section VI-B.

3) Level 3: Stream-Wise MSE Preservation: The full-featured version maintains the MSE of every *single* stream during the conversion. It can be interpreted as an extended form of the second kind, where we associate a virtual single-stream-user to every data stream in the system. Streams belonging to a specific real user all have the same channel matrix in common, and different scaling factors are allocated to every *virtual* user, i.e., to every data stream in the system. Consequently, $\sum_{k=1}^{K} B_k$ scaling factors need to be determined which again follow from a linear system of equations. Obviously, the third level has the highest complexity. Needless to say, the level-3 duality includes the functionality of the former two levels which means that both the sum-MSE and the user-wise MSEs are preserved as well. However, these two properties are achieved at a much higher computational effort in the third level. So if less restrictive MSE conservations are required, for example if only the users' MSEs or the sum-MSE shall be maintained, one should go for the first or the second level of our duality and save computation time. Interestingly, the proposed kind of streamwise MSE duality not only preserves the individual MSEs per stream, but also the SINRs per stream remain unchanged, see Section IV-C. Therefore, the third level of the proposed duality keeps the individual SINRs and consequently the data rate of every stream constant during the conversion. So if the focus lies on the data rate in linear transceiver design where the receivers do not decode their streams jointly, the third kind becomes attractive.

IV. Uplink-Downlink Conversion

In this section, we construct an equivalent downlink channel given a fixed uplink setup. Strictly speaking, given arbitrary $\boldsymbol{t}_{k,\ell}$ and $\boldsymbol{g}_{k,\ell}$ for all users $k=1,\ldots,K$ and all respective streams $\ell=1,\ldots,B_k$ we derive vectors $\boldsymbol{p}_{k,\ell}$ and $\boldsymbol{b}_{k,\ell}$ such that according to the level of duality, either the sum-MSE remains constant, the users' MSEs do not change, or the individual streams' MSEs are the same in the downlink and in the uplink.

A. Uplink-Downlink Conversion of the First Kind

In the first kind, the system-wide sum-MSE is preserved. Hence, we only need a single degree of freedom which turns out to leave the sum-MSE unchanged during the conversion and simultaneously preserves the sum-power constraint (3) in the dual domain. With the aforementioned positive scaling factor $\alpha \in \mathbb{R}_+$ common to all streams, we set

$$\boldsymbol{p}_{k,\ell} = \alpha\boldsymbol{g}_{k,\ell} \quad \text{and} \quad \boldsymbol{b}_{k,\ell} = \frac{1}{\alpha}\boldsymbol{t}_{k,\ell}\ \forall k,\ell$$

or, equivalently

$$\boldsymbol{P}_k = \alpha\boldsymbol{G}_k^{\mathrm{H}} \quad \text{and} \quad \boldsymbol{B}_k = \frac{1}{\alpha}\boldsymbol{T}_k^{\mathrm{H}}\ \forall k$$

leading in conjunction with the composite matrices to

$$\boldsymbol{P} = \alpha \boldsymbol{G}^{\mathrm{H}} \quad \text{and} \quad \boldsymbol{B} = \frac{1}{\alpha}\boldsymbol{T}^{\mathrm{H}}. \tag{13}$$

Inserting (13) into the downlink MSE expression (4), we obtain

$$\varepsilon^{\mathrm{DL}} = \sum_{k=1}^{K} B_k - \mathrm{tr}\left[\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}} + \boldsymbol{GHT}\right] + \mathrm{tr}\left[\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{GHT}\right] + \frac{1}{\alpha^2}\sigma_\eta^2 \mathrm{tr}\left[\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}\right].$$

Equating this to the uplink MSE expression (9), we find by means of (8) that

$$\alpha^2 = \frac{\mathrm{tr}\left[\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}\right]}{\mathrm{tr}\left[\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\right]} = \frac{P_{\mathrm{Tx}}}{\|\boldsymbol{G}\|_{\mathrm{F}}^2} = \frac{P_{\mathrm{Tx}}}{\sum_{k=1}^{K}\|\boldsymbol{G}_k\|_{\mathrm{F}}^2} \tag{14}$$

needs to hold in order to satisfy the retention of the sum-MSE. The resulting precoders read as $\boldsymbol{P}_k = \alpha \boldsymbol{G}_k^{\mathrm{H}}$ and dissipate the power $\sum_{k=1}^{K}\|\boldsymbol{P}_k\|_{\mathrm{F}}^2 = P_{\mathrm{Tx}}$, which means that the dual downlink has the same sum-power consumption as the uplink.

B. Uplink-Downlink Conversion of the Second Kind

In order to leave the individual users' MSEs unchanged during the conversion, different scalars $\alpha_k \in \mathbb{R}_{+,0}, k \in \{1,\ldots,K\}$ have to be assigned to the individual users. To this end, we set

$$\boldsymbol{p}_{k,\ell} = \alpha_k \boldsymbol{g}_{k,\ell} \quad \text{and} \quad \boldsymbol{b}_{k,\ell} = \frac{1}{\alpha_k}\boldsymbol{t}_{k,\ell}\ \forall k,\ \ell$$

corresponding to

$$\boldsymbol{P}_k = \alpha_k \boldsymbol{G}_k^{\mathrm{H}} \quad \text{and} \quad \boldsymbol{B}_k = \frac{1}{\alpha_k}\boldsymbol{T}_k^{\mathrm{H}}\ \forall k. \tag{15}$$

It will be shown later that the cases $\alpha_i = 0$ and $\alpha_i \longrightarrow \infty$ may occur. The computation of the matrices $\boldsymbol{B}_i$ and $\boldsymbol{P}_i$ has to be handled separately then. Given the uplink MSE $\varepsilon_k^{\mathrm{UL}}$ of user k in (11) and the downlink MSE $\varepsilon_k^{\mathrm{DL}}$ in (5), we observe that the first three summands in the trace expressions in (11) and (5) are again identical and cancel out when equating $\varepsilon_k^{\mathrm{DL}}$ and $\varepsilon_k^{\mathrm{UL}}$, and we get

$$\varepsilon_k^{\mathrm{UL}} = \varepsilon_k^{\mathrm{DL}} \quad \Leftrightarrow \quad \sum_{i=1}^{K}\|\boldsymbol{G}_k\boldsymbol{H}_i\boldsymbol{T}_i\|_{\mathrm{F}}^2 + \sigma_\eta^2\|\boldsymbol{G}_k\|_{\mathrm{F}}^2 = \sum_{i=1}^{K}\frac{\alpha_i^2}{\alpha_k^2}\|\boldsymbol{G}_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2 + \frac{\sigma_\eta^2}{\alpha_k^2}\|\boldsymbol{T}_k\|_{\mathrm{F}}^2\ \forall k. \tag{16}$$

It is obvious that we can neglect *nonactive* users having *simultaneously* passive transmitters ($\boldsymbol{T}_k = \boldsymbol{0}_{r_k \times B_k}$) *and* passive receivers ($\boldsymbol{G}_k = \boldsymbol{0}_{B_k \times N}$). For them, the MSE-equality in (16) is obtained for $\boldsymbol{P}_k = \boldsymbol{0}_{N \times B_k}$ and $\boldsymbol{B}_k = \boldsymbol{0}_{r_k \times B_k}$, i.e., their respective filters in the downlink also vanish, and the respective scalars $0 < \alpha_k < \infty$ are arbitrary, see (15). We observe that the duality simplifies, when a passive transmitter ($\boldsymbol{T}_k = \boldsymbol{0}_{r_k \times B_k}$) also has a passive receiver ($\boldsymbol{G}_k = \boldsymbol{0}_{N \times B_k}$), and vice versa. In the following we will see that the proof for the duality is not so simple for the case where either the transmitter $\boldsymbol{T}_k$ or the receiver $\boldsymbol{G}_k$ is inactive but the filter at the other side of the link is active. However, any *reasonable* optimization should switch off the receive filter, if the respective transmitter does not send anything at all. Conversely, if the receive filter for a certain user vanishes, the respective transmit filter should vanish as well. As mentioned above, the duality for those *nonactive* users is obvious. Thus, we assume in the following that no users are present that neither transmit nor receive. We first treat the simplified version of the uplink/downlink duality with active users ($\boldsymbol{T}_k \neq \boldsymbol{0}_{r_k \times B_k}$ and $\boldsymbol{G}_k \neq \boldsymbol{0}_{B_k \times N}$) and then extend it to the full functionality.

1) Uplink to Downlink Transformation for Strictly Active Users: For the strictly active users with $\boldsymbol{T}_k \neq \boldsymbol{0}_{r_k \times B_k}\ \forall k$ and $\boldsymbol{G}_k \neq \boldsymbol{0}_{N \times B_k}\ \forall k$ we find $0 < \alpha_k < \infty\ \forall k$, since all users need to transmit and receive in the corresponding downlink as well. We can therefore, rewrite (16)

$$\alpha_k^2\Big[\sum_{i=1, i\neq k}^{K}\|\boldsymbol{G}_k\boldsymbol{H}_i\boldsymbol{T}_i\|_{\mathrm{F}}^2 + \sigma_\eta^2\|\boldsymbol{G}_k\|_{\mathrm{F}}^2\Big] - \sum_{i=1, i\neq k}^{K}\alpha_i^2\|\boldsymbol{G}_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2 = \sigma_\eta^2\|\boldsymbol{T}_k\|_{\mathrm{F}}^2\ \forall k. \tag{17}$$

In matrix vector notation, we obtain

$$\boldsymbol{Z}\cdot\left[\alpha_1^2,\ldots,\alpha_K^2\right]^{\mathrm{T}} = \sigma_\eta^2\left[\|\boldsymbol{T}_1\|_{\mathrm{F}}^2,\ldots,\|\boldsymbol{T}_K\|_{\mathrm{F}}^2\right]^{\mathrm{T}} \tag{18}$$

with the column diagonally dominant matrix

$$[Z]_{k,j} = \begin{cases} \sum_{i=1, i\neq k}^{K}\|\boldsymbol{G}_k\boldsymbol{H}_i\boldsymbol{T}_i\|_{\mathrm{F}}^2 + \sigma_\eta^2\|\boldsymbol{G}_k\|_{\mathrm{F}}^2 & \text{for } k = j \\ -\|\boldsymbol{G}_j\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2 & \text{for } k \neq j. \end{cases} \tag{19}$$

Since $\boldsymbol{Z}$ is real-valued and has only nonpositive off-diagonal entries, it is a *Z-matrix* [38]. Furthermore, it is strictly column diagonally dominant for $\sigma_\eta^2 > 0$, so its inverse exists and $\boldsymbol{Z}$ is also an *M-matrix* [38]. As a consequence, its inverse $\boldsymbol{Z}^{-1}$ has only nonnegative entries [38]–[40]. The necessity of positive scalars α_i^2 is therefore automatically fulfilled, and one direction of the proof of the duality can always be guaranteed, namely that every MSE tuple in the uplink can also be achieved in the downlink.

Summing up all rows of (18) yields

$$\sigma_\eta^2\sum_{i=1}^{K}\alpha_i^2\|\boldsymbol{G}_i\|_{\mathrm{F}}^2 = \sigma_\eta^2\sum_{i=1}^{K}\|\boldsymbol{T}_i\|_{\mathrm{F}}^2 \tag{20}$$

and from (15), we identify the left-hand side (LHS) of (20) to be $\sigma_\eta^2\sum_{i=1}^{K}\|\boldsymbol{P}_i\|_{\mathrm{F}}^2$, i.e., the transmitted power in the downlink scaled by the noise variance σ_η^2. According to (8), the right-hand side (RHS) of (20) represents the transmit power P_{Tx} in the uplink scaled by the noise variance σ_η^2. From this, we can conclude that the sum power constraint is inherently fulfilled when we compute the scalars $\alpha_1^2,\ldots,\alpha_K^2$ via

$$\left[\alpha_1^2,\ldots,\alpha_K^2\right]^{\mathrm{T}} = \sigma_\eta^2 \boldsymbol{Z}^{-1}\cdot\left[\|\boldsymbol{T}_1\|_{\mathrm{F}}^2,\ldots,\|\boldsymbol{T}_K\|_{\mathrm{F}}^2\right]^{\mathrm{T}}. \quad (21)$$

Clearly, it is advantageous to solve for the unknowns via the LU-decomposition of $\boldsymbol{Z}$ [41] and a subsequent forward-backward substitution. Doing so, finding $\alpha_1^2,\ldots,\alpha_K^2$ given $\boldsymbol{Z}$ and $\|\boldsymbol{T}_k\|_{\mathrm{F}}^2\ \forall k$ requires $2/3K^3+3/2K^2-7/6K$ FLOPs, cf. [41] and [42]. Setting $\boldsymbol{P}_k=\alpha_k\boldsymbol{G}_k^{\mathrm{H}}$ and $\boldsymbol{B}_k=1/\alpha_k\boldsymbol{T}_k^{\mathrm{H}}\ \forall k$, we have found the transmit and receive filters in the uplink that achieve the same user-MSEs as the matrices $\boldsymbol{T}_k$ and $\boldsymbol{G}_k\ \forall k$ in the uplink with the same power consumption P_{Tx}.

2) Generalized Uplink to Downlink Transformation: The conversion for strictly active users presented in the previous section cannot handle the case when a user features a nonzero receive filter but the respective transmit filter matrix is zero or vice versa. Although such a configuration of transmit and receive matrices probably won't be applied during transmission, it might arise in an intermediate step of an iterative algorithm that applies the uplink-downlink conversion in each step. A generalized uplink to downlink transformation being capable of such semiactive users is presented in the Appendix I.

C. Uplink-Downlink Conversion of the Third Kind

Maintaining the MSE of every single stream is achieved by choosing different scaling factors for every stream's precoding and equalization vector

$$\boldsymbol{p}_{k,\ell}=\alpha_{k,\ell}\boldsymbol{g}_{k,\ell} \quad \text{and} \quad \boldsymbol{b}_{k,\ell}=\frac{1}{\alpha_{k,\ell}}\boldsymbol{t}_{k,\ell}\ \forall k,\ell. \quad (22)$$

In the following, we assume that all streams have active receivers and active transmitters. Otherwise, the concept of the generalized duality from Appendix I can be extended to the streamwise MSE duality in a straightforward fashion. Inserting (22) into the streamwise downlink MSE expression for $\varepsilon_{k,\ell}^{\mathrm{DL}}$ in (6) and equating the result with the uplink MSE $\varepsilon_{k,\ell}^{\mathrm{UL}}$ in (12), we obtain for $k\in\{1,\ldots,K\}$ and $\ell\in\{1,\ldots,B_k\}$ similar to (16)

$$\sum_{i=1}^{K}\sum_{\substack{j=1\\(i,j)\neq(k,\ell)}}^{B_i}|\boldsymbol{g}_{k,\ell}^{\mathrm{H}}\boldsymbol{H}_i\boldsymbol{t}_{i,j}|^2+\sigma_\eta^2\|\boldsymbol{g}_{k,\ell}\|_2^2 = \sum_{i=1}^{K}\sum_{\substack{j=1\\(i,j)\neq(k,\ell)}}^{B_i}\frac{\alpha_{i,j}^2}{\alpha_{k,\ell}^2}|\boldsymbol{g}_{i,j}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{k,\ell}|^2+\frac{1}{\alpha_{k,\ell}^2}\sigma_\eta^2\|\boldsymbol{t}_{k,\ell}\|_2^2. \quad (23)$$

Since both expressions in above formula also denote the interference plus noise power seen by stream ℓ of user k and since the desired signal part $|\boldsymbol{g}_{k,\ell}^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{t}_{k,\ell}|^2$ remains constant during the uplink-downlink conversion, the SINR is not affected by the conversion. Hence, the third kind of our MSE duality can also be regarded as an SINR duality. It is easy to see that (23) leads to the system of equations $\boldsymbol{Z}\cdot[\alpha_{1,1}^2,\ldots,\alpha_{K,B_K}^2]^{\mathrm{T}}=\sigma_\eta^2[\|\boldsymbol{t}_{1,1}\|_2^2,\ldots,\|\boldsymbol{t}_{K,B_K}\|_2^2]^{\mathrm{T}}$ similar to (18), which now has $\sum_{k=1}^{K}B_k$ unknowns but still features the *M-matrix* property guaranteeing feasible solutions for the unknowns $\alpha_{1,1}^2,\ldots,\alpha_{K,B_K}^2$.

V. Downlink-Uplink Conversion

The downlink to uplink transformation is important to complete the duality by showing that every MSE tuple in the downlink can also be achieved in the uplink. This means that the MSE region in the downlink is a subset of the MSE region in the uplink. In conjunction with the fact proven in the previous sections that the MSE region in the uplink is also a subset of the MSE region in the downlink, we can infer that both regions are identical and the duality is established. For our purpose, the downlink to uplink conversion is only important to show that optimal solutions of uplink and downlink are identical. However, the explicit use of the downlink-uplink conversion is necessary when repeatedly switching between uplink and downlink, cf. [29], [43], [44].

A. Downlink-Uplink Conversion of the First Kind

The first stage of the duality allows for a conservation of the system wide sum-MSE during the conversion by setting the composite matrices to

$$\boldsymbol{T}=\bar{\alpha}\boldsymbol{B}^{\mathrm{H}} \quad \text{and} \quad \boldsymbol{G}=\frac{1}{\bar{\alpha}}\boldsymbol{P}^{\mathrm{H}}. \quad (24)$$

Plugging this into the uplink sum-MSE expression (9) and equating the result with the downlink sum-MSE expression in (4), the solution for $\bar{\alpha}^2$ reads as

$$\bar{\alpha}^2=\frac{P_{\mathrm{Tx}}}{\|\boldsymbol{B}\|_{\mathrm{F}}^2}=\frac{P_{\mathrm{Tx}}}{\sum_{k=1}^{K}\|\boldsymbol{B}_k\|_{\mathrm{F}}^2}. \quad (25)$$

Again, the amount of transmitted power is invariant under the conversion.

B. Downlink-Uplink Conversion of the Second Kind

Given the precoders $\boldsymbol{P}_k$ and receive matrices $\boldsymbol{B}_k$ in the downlink, the respective filters in the dual uplink are set to

$$\boldsymbol{T}_k=\bar{\alpha}_k\boldsymbol{B}_k^{\mathrm{H}} \quad \text{and} \quad \boldsymbol{G}_k=\frac{1}{\bar{\alpha}_k}\boldsymbol{P}_k^{\mathrm{H}}. \quad (26)$$

In turn, the cases $\bar{\alpha}_k=0$ and $\bar{\alpha}_k\rightarrow\infty$ may occur and will be handled separately as for the transformation from the uplink to the downlink. By means of (26), we equate the downlink MSE $\varepsilon_k^{\mathrm{DL}}$ from (5) and the uplink MSE $\varepsilon_k^{\mathrm{UL}}$ from (11) of all users k to obtain

$$\sum_{i=1}^{K}\|\boldsymbol{P}_i^{\mathrm{H}}\boldsymbol{H}_k\boldsymbol{B}_k^{\mathrm{H}}\|_{\mathrm{F}}^2+\sigma_\eta^2\|\boldsymbol{B}_k\|_{\mathrm{F}}^2 = \sum_{i=1}^{K}\frac{\bar{\alpha}_i^2}{\bar{\alpha}_k^2}\|\boldsymbol{P}_k^{\mathrm{H}}\boldsymbol{H}_i\boldsymbol{B}_i^{\mathrm{H}}\|_{\mathrm{F}}^2+\sigma_\eta^2\frac{1}{\bar{\alpha}_k^2}\|\boldsymbol{P}_k\|_{\mathrm{F}}^2\ \forall k. \quad (27)$$

Again, we start with the simplified version of the transformation where all transmitters and receivers are active, i.e., their matrices are different from zero.

1) Downlink to Uplink Transformation for Strictly Active Users: For the strictly active users, we have $\boldsymbol{P}_k\neq\boldsymbol{0}_{N\times B_k}$ and $\boldsymbol{B}_k\neq\boldsymbol{0}_{B_k\times r_k}\ \forall k$. Consequently, $0<\bar{\alpha}_k<\infty$ holds for all k. Rearranging (27) and making use of $\|(\cdot)\|_{\mathrm{F}}=\|(\cdot)^{\mathrm{H}}\|_{\mathrm{F}}$ yields

$$\boldsymbol{Y}\cdot\left[\bar{\alpha}_1^2,\ldots,\bar{\alpha}_K^2\right]^{\mathrm{T}}=\sigma_\eta^2\left[\|\boldsymbol{P}_1\|_{\mathrm{F}}^2,\ldots,\|\boldsymbol{P}_K\|_{\mathrm{F}}^2\right]^{\mathrm{T}} \quad (28)$$

with the column diagonally dominant *M-matrix*

$$[\boldsymbol{Y}]_{k,j} = \begin{cases} \sum_{i=1,i\neq k}^{K} \|\boldsymbol{B}_k\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{P}_i\|_{\mathrm{F}}^2 + \sigma_\eta^2\|\boldsymbol{B}_k\|_{\mathrm{F}}^2 & \text{for } k=j \\ -\|\boldsymbol{B}_j\boldsymbol{H}_j^{\mathrm{H}}\boldsymbol{P}_k\|_{\mathrm{F}}^2 & \text{for } k\neq j. \end{cases} \tag{29}$$

Summing up the rows of (28), we observe that the power conservation is inherently fulfilled, and the desired weight vector reads as

$$\left[\bar{\alpha}_1^2,\ldots,\bar{\alpha}_K^2\right]^{\mathrm{T}} = \sigma_\eta^2\boldsymbol{Y}^{-1}\cdot\left[\|\boldsymbol{P}_1\|_{\mathrm{F}}^2,\ldots,\|\boldsymbol{P}_K\|_{\mathrm{F}}^2\right]^{\mathrm{T}} \tag{30}$$

whose elements are nonnegative and finite due to the properties of $\boldsymbol{Y}$.

2) Generalized Downlink to Uplink Transformation: Semi-active users where either the receive matrix or the transmit matrix is zero can be handled by the generalized version of above conversion. Similar to its uplink-downlink counterpart in Appendix I, this generalized downlink uplink transformation is presented in Appendix II.

C. Downlink-Uplink Conversion of the Third Kind

Finally, the streamwise MSE conservation is again obtained by means of individual scaling factors for every stream's precoding and receive vector:

$$\boldsymbol{t}_{k,\ell} = \bar{\alpha}_{k,\ell}\boldsymbol{b}_{k,\ell} \quad \text{and} \quad \boldsymbol{g}_{k,\ell} = \frac{1}{\bar{\alpha}_{k,\ell}}\boldsymbol{p}_{k,\ell}\ \forall k,\ell.$$

Inserting this into the uplink MSE expression from (12) and equating the result with the downlink MSE expression from (6), we arrive at a linear system of equations which can be solved for the squared scaling factors $\bar{\alpha}_{1,1}^2,\ldots,\bar{\alpha}_{K,B_K}^2$, cf. Section IV-C. In addition to the MSEs, the SINRs are not altered during the conversion.

VI. Applications for the Three-Level Duality

In the following subsections, we investigate two of the aforementioned areas of application for the proposed three-level duality in more detail. First, we apply the first kind to the sum-MSE minimization in the broadcast channel and reveal the hidden convexity of the problem. Afterwards, we show that the second kind of the proposed duality allows us to reduce the complexity of the weighted sum-MSE minimization in the broadcast channel drastically.

A. Total Sum-MSE Minimization in the Broadcast Channel

Finding the jointly optimum composite matrices $\boldsymbol{B}$ and $\boldsymbol{P}$ that minimize the downlink sum-MSE from (4) subject to a total sum-power constraint $\|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tx}}$ cannot be done in closed form in general. However, given $\boldsymbol{P}$, we can solve for the optimum $\boldsymbol{B}$ depending on the precoders $\boldsymbol{P}$. Due to the block-diagonal structure of $\boldsymbol{B}$, the optimum $\boldsymbol{B}_k$ have to be computed separately and read as

$$\check{\boldsymbol{B}}_k = \boldsymbol{P}_k^{\mathrm{H}}\boldsymbol{H}_k\left(\sigma_\eta^2\mathbf{I}_{r_k} + \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}_k\right)^{-1}$$

achieving an MSE of user k from (5) reading as

$$\varepsilon_k^{\mathrm{DL}} = B_k - \mathrm{tr}\left[\boldsymbol{P}_k^{\mathrm{H}}\boldsymbol{H}_k\left(\sigma_\eta^2\mathbf{I}_{r_k} + \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}_k\right)^{-1}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{P}_k\right]. \tag{31}$$

Unfortunately, it is not possible to merge the individual inverses $(\sigma_\eta^2\mathbf{I}_{r_k} + \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\boldsymbol{H}_k)^{-1}$ when summing up the users' MSEs to the sum-MSE $\varepsilon^{\mathrm{DL}} = \sum_{k=1}^{K}\varepsilon_k^{\mathrm{DL}}$. Thus, statements on the convexity of $\varepsilon^{\mathrm{DL}}$ are not obvious at first glance. The second approach one might think of when directly operating in the downlink is to optimize the precoders depending on the receive matrices in $\boldsymbol{B}$. Unfortunately, a closed form expression is again infeasible because of the sum-power constraint $\|\boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tx}}$ the composite precoding matrix $\boldsymbol{P}$ underlies. The only chance of revealing the hidden convexity of the problem is the splitting of $\boldsymbol{B}$ into a common scalar $b \in \mathbb{R}_+$ and the remaining matrix $\boldsymbol{B}'$ [37], i.e., $\boldsymbol{B} = b\boldsymbol{B}'$ and optimizing the precoder $\boldsymbol{P}$ and b for fixed $\boldsymbol{B}'$ simultaneously. Afterwards, a tricky modification of the cost function has to be applied to end up at a convex problem. In detail, the optimum $\boldsymbol{P}$ reads as

$$\check{\boldsymbol{P}} = \check{b}^{-1}\left(\boldsymbol{H}\boldsymbol{B}'^{\mathrm{H}}\boldsymbol{B}'\boldsymbol{H}^{\mathrm{H}} + \frac{\sigma_\eta^2}{P_{\mathrm{Tx}}}\|\boldsymbol{B}'\|_{\mathrm{F}}^2\mathbf{I}_N\right)^{-1}\boldsymbol{H}\boldsymbol{B}'^{\mathrm{H}},$$

yielding an MSE

$$\varepsilon^{\mathrm{DL}} = \sum_{k=1}^{K} B_k - N + \sigma_\eta^2\mathrm{tr}\left[\left(\sigma_\eta^2\mathbf{I}_N + P_{\mathrm{Tx}}\frac{\boldsymbol{H}\boldsymbol{B}'^{\mathrm{H}}\boldsymbol{B}'\boldsymbol{H}^{\mathrm{H}}}{\|\boldsymbol{B}'\|_{\mathrm{F}}^2}\right)^{-1}\right],$$

which still does not feature any convex structure. However, $\varepsilon^{\mathrm{DL}}$ is invariant under a scaling of $\boldsymbol{B}'$, since it will be revoked by the common scalar b. Hence, we are free to impose a constraint on the squared Frobenius norm $\|\boldsymbol{B}'\|_{\mathrm{F}}^2$ and set it arbitrarily to P_{Tx}. The problem then reads as

$$\begin{aligned} \underset{\boldsymbol{B}'_1,\ldots,\boldsymbol{B}'_K}{\text{minimize}} & \sum_{k=1}^{K} B_k - N \\ & + \sigma_\eta^2\mathrm{tr}\left[\left(\sigma_\eta^2\mathbf{I}_N + \boldsymbol{H}\boldsymbol{B}'^{\mathrm{H}}\boldsymbol{B}'\boldsymbol{H}^{\mathrm{H}}\right)^{-1}\right] \\ \text{s.t.}: & \ \mathrm{tr}\left[\boldsymbol{B}'^{\mathrm{H}}\boldsymbol{B}'\right] = P_{\mathrm{Tx}}, \end{aligned} \tag{32}$$

which now is convex in every product $\boldsymbol{B}_k'^{\mathrm{H}}\boldsymbol{B}_k'$ since $\boldsymbol{H}\boldsymbol{B}'^{\mathrm{H}}\boldsymbol{B}'\boldsymbol{H}^{\mathrm{H}} = \sum_{k=1}^{K}\boldsymbol{H}_k\boldsymbol{B}_k'^{\mathrm{H}}\boldsymbol{B}_k'\boldsymbol{H}_k^{\mathrm{H}}$ and $\mathrm{tr}\left[(\cdot)^{-1}\right]$ is a matrix convex function for positive semidefinite arguments [45].

A much more convenient way to reveal the hidden convexity and gain the resultant benefits is the application of the first kind of the proposed duality and handle the problem in the dual uplink. To this end, we solve the dual multiple-access problem and convert the resulting optimum matrices $\check{\boldsymbol{T}}$ and $\check{\boldsymbol{G}}$ into the downlink via (13) and (14). In turn, the receive filters $\boldsymbol{G}_1,\ldots,\boldsymbol{G}_K$ can be computed independently and read as

$$\check{\boldsymbol{G}}_k = \boldsymbol{T}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{X}^{-1} \tag{33}$$

with the substitution

$$\boldsymbol{X} := \sigma_\eta^2 \mathbf{I}_N + \boldsymbol{H}\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} = \sigma_\eta^2 \mathbf{I}_N + \sum_{i=1}^{K} \boldsymbol{H}_i \boldsymbol{T}_i \boldsymbol{T}_i^{\mathrm{H}} \boldsymbol{H}_i^{\mathrm{H}} \quad (34)$$

which is in contrast to the downlink solution for $\check{\boldsymbol{B}}_k$ common to all users since the channel $\boldsymbol{H}_k$ is always bound to the precoder $\boldsymbol{T}_k$ in the uplink. The resulting MSE of user k is

$$\begin{aligned}\varepsilon_k^{\mathrm{UL}} &= B_k - \mathrm{tr}\left[\boldsymbol{T}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{X}^{-1}\boldsymbol{H}_k\boldsymbol{T}_k\right] \\ &= B_k - \mathrm{tr}\left[\boldsymbol{H}_k\boldsymbol{T}_k\boldsymbol{T}_k^{\mathrm{H}}\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{X}^{-1}\right]\end{aligned}$$

involving a sum-MSE

$$\varepsilon^{\mathrm{UL}} = \sum_{k=1}^{K} B_k - N + \sigma_\eta^2 \mathrm{tr}\left[\boldsymbol{X}^{-1}\right] \quad (35)$$

being jointly convex in all transmit covariance matrices $\boldsymbol{Q}_k = \boldsymbol{T}_k\boldsymbol{T}_k^{\mathrm{H}}$. Note the similarity of the expressions for the uplink MSE $\varepsilon^{\mathrm{UL}}$ in (35) and for the cost in (32).

We solved the total sum-MSE minimization problem in the downlink in [37] by means of the concept of *alternating optimization*, where the transmitter and receivers are updated in an alternating fashion. Unfortunately, the speed of convergence turned out to be slow at high SNR values. The authors in [21], [22], [29], and [43] also make use of an *alternating optimization*. The alternation between the power allocation and the computation of the beamforming vectors or the transmit covariance matrices and repeated switching between uplink and downlink makes the proposed approaches computationally expensive. Another approach to sum-MSE minimization is due to Tenenbaum *et al.* in [46], where the interference is treated as additional noise via its covariance matrix. Then, several iterations based on the single-user MIMO sum-MSE minimization [47] are performed. In contrast to the sum-rate maximization problem, treating other users as colored noise does not yield the global optimum for the MSE minimization problem.

Here, we tackle the problem with the aid of the *projected gradient algorithm* [48], which is an extension of the conventional steepest descent algorithm for constrained optimizations. It features good convergence properties especially at large SNR values. The iteration step reads as

$$\boldsymbol{T}^{(\ell+1)} = \left[\boldsymbol{T}^{(\ell)} - \frac{\lambda^{(\ell)}}{s^{(\ell)}}\nabla^*\varepsilon^{\mathrm{UL}}\left(\boldsymbol{T}^{(\ell)}\right)\right]_\perp \quad (36)$$

where $\boldsymbol{T}^{(\ell)}$ is blockdiagonal with all precoders $\boldsymbol{T}_1^{(\ell)},\ldots,\boldsymbol{T}_K^{(\ell)}$ in the iteration step ℓ, $\varepsilon^{\mathrm{UL}}(\cdot)$ is the total sum-MSE from (35), $\lambda^{(\ell)}$ is a preconditioning scalar, and $1/s^{(\ell)}$ is the step-size. The conjugate nabla operator $\nabla^*(\cdot)$ generates all conjugate Jacobi matrices

$$\begin{aligned}\nabla_k^*\varepsilon^{\mathrm{UL}}\left(\boldsymbol{T}^{(\ell)}\right) &= \frac{\partial\varepsilon^{\mathrm{UL}}\left(\boldsymbol{T}^{(\ell)}\right)}{\partial\boldsymbol{T}_k^*} \\ &= -\sigma_\eta^2\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{X}^{-2}\boldsymbol{H}_k\boldsymbol{T}_k^{(\ell)}. \end{aligned} \quad (37)$$

The update rule (36) first performs a standard steepest descent step and then applies an *orthogonal* projection onto the constraint set defined by $\|\boldsymbol{T}^{(\ell+1)}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tx}}$. It can be shown that the unprojected gradient algorithm would always require more transmit power than available, hence the projection is always necessary. Moreover, since the constraint set is a hyperball with radius $\sqrt{P_{\mathrm{Tx}}}$, the *orthogonal* projection is very simple, it is nothing else than a scaling of all precoders by a common factor such that $\|\boldsymbol{T}^{(\ell+1)}\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tx}}$ is fulfilled *with equality*, i.e., the projection is onto the hypersphere. The preconditioning scalar $\lambda^{(\ell)} = \sqrt{P_{\mathrm{Tx}}}/\|\nabla^*\varepsilon^{\mathrm{UL}}(\boldsymbol{T}^{(\ell)})\|_{\mathrm{F}}$ is intended to speed up the convergence at high SNRs, since $\varepsilon^{\mathrm{UL}}(\cdot)$ is almost flat there, and the Jacobian has a small Frobenius norm, which makes this scaling important. Alternating optimization algorithms suffer from this flatness at high SNRs and require a large number of iterations. The inverse step-size $s^{(\ell)}$ is increased, as soon as the objective tends to increase during the iterations. The convergence of this algorithm is proven in [48] by means of a descent argument:

Theorem 1: Given $\varepsilon^{\mathrm{UL}}$ is bounded below and Lipschitzian with the Lipschitz constant C, and $0 < \lambda^{(\ell)}/s^{(\ell)} < 2/C$, then the sequence of precoders generated by the gradient projection algorithm converges. Furthermore, the limit point of this sequence satisfies the first-order KKT optimality conditions. In particular, if $\varepsilon^{\mathrm{UL}}$ is convex on the constraint set, then the global minimum is obtained.

For the proof, see [48].

From Theorem 1 we know that the projected gradient approach would yield the global optimum of the sum-MSE minimization if we were operating on the covariance matrices $\boldsymbol{Q}_k = \boldsymbol{T}_k\boldsymbol{T}_k^{\mathrm{H}}$, since the cost function is convex in the covariances. However, our projected gradient approach works on the precoders $\boldsymbol{T}_k$ instead of their outer products. Therefore, we introduce a unitary invariance of the cost function with respect to RHS multiplications of unitary matrices onto the precoding matrices. Whether the global minimum of the sum-MSE is obtained although the cost is not convex in the precoders depends on the number of transmitted streams per user:

1) As Many Data Streams as Transmit Antennas: For $B_k = r_k\ \forall k$, the covariance matrices $\boldsymbol{Q}_k$ can achieve the full rank r_k. Consequently, there is no constraint on the rank of the allowed covariance matrices $\boldsymbol{Q}_k$, the traces of all covariances only have to sum up to the maximum available transmit power. In this case, the cost in (35) is convex and we optimize over a convex set yielding a convex optimization without local nonglobal minima. The update rule from (36) can be rewritten by means of (37)

$$\boldsymbol{T}^{(\ell+1)} = \kappa^{(\ell)}\Big(\mathbf{I}_{r_k} + \frac{\lambda^{(\ell)}}{s^{(\ell)}}\sigma_\eta^2\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{X}^{-2}\boldsymbol{H}_k\Big)\boldsymbol{T}_k^{(\ell)}\ \forall k \quad (38)$$

where $\kappa^{(\ell)}$ is chosen such that $\sum_{k=1}^{K}\|\boldsymbol{T}_k^{(\ell+1)}\|_{\mathrm{F}}^2 = P_{\mathrm{Tx}}$. From (38), we see that the rank of $\boldsymbol{T}^{(\ell+1)}$ does not change in a finite number of iterations[2] since $\mathbf{I}_{r_k} + \lambda^{(\ell)}/s^{(\ell)}\sigma_\eta^2\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{X}^{-2}\boldsymbol{H}_k$ is positive definite and the LHS multiplication of $\boldsymbol{T}_k^{(\ell)}$ by a matrix of full rank does not change its algebraic rank. Consequently, if the initial $\boldsymbol{T}_k^{(0)}$ does not have full rank, but the global optimum $\check{\boldsymbol{T}}_k$ does have full rank, the projected gradient algorithm never achieves this optimum. The solution obtained when starting from a rank deficient precoder $\boldsymbol{T}_k$ corresponds

[2]In particular, the rank cannot increase, even for an infinite number of iterations.

to the minimization of the system under the assumption that less than r_k streams are transmitted, since any rank deficient $\boldsymbol{T}_k$ can be transformed via unitary rotations (that do not change the user-MSE) into a precoder with some zero-columns that correspond to switched off data streams. In addition, if we start with $\boldsymbol{T}_k^{(0)} = \mathbf{0}_{r_k \times B_k}$, the system is optimized under the assumption that this user k is not present. If we start with a full rank initialization but an optimum $\check{\boldsymbol{T}}_k$ turns out to be rank deficient, the projected gradient algorithm reduces the rank of $\boldsymbol{T}^{(\ell)}$ in the asymptotic limit. Although the algebraic rank cannot decrease during a finite number of iterations, the LHS multiplication of $\boldsymbol{T}_k^{(\ell)}$ in (38) can increase the condition number of $\boldsymbol{T}_k^{(\ell+1)}$, and after several iterations, the numerical rank may have decreased as well. In the limit $\ell \longrightarrow \infty$ after an infinite number of iterations, the distance between $\boldsymbol{T}_k^{(\ell)}$ and $\check{\boldsymbol{T}}_k$ goes down to zero, the algebraic rank of $\boldsymbol{T}_k^{(\ell)}$ has decreased and matches the one of $\check{\boldsymbol{T}}_k$. We can conclude that we need to initialize all precoders $\boldsymbol{T}_k^{(0)}$ with full rank.

Changing from the covariances $\boldsymbol{Q}_k = \boldsymbol{T}_k \boldsymbol{T}_k^{\mathrm{H}}$ to the precoders $\boldsymbol{T}_k$ does not create additional minima except the ones possible with the unitary rotations. However, additional rank-deficient matrices fulfilling the KKT conditions emerge, which can be seen from (38). These stationary points are saddle points corresponding to *rank-deficient* precoders. Moreover, a single maximum arises which corresponds to the special case of transmitting only a single stream to a single user on the eigenmode belonging the smallest eigenvalue $\min_{k\in\{1,\ldots,K\}} \min \mathrm{eigenvalues}(\boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{H}_k)$. If one starts on the stable eigendirection of these saddle points with a rank deficient precoder, the global optimum cannot be obtained. If we initialize a precoder in a rank deficient fashion and the precoder achieving the global minimum is also rank deficient, we may achieve this global optimum, if we do not start exactly on the stable eigendirection of a saddle point. Summing up, if we initialize all precoders $\boldsymbol{T}_k^{(0)}$ with full rank matrices and transmit as many data streams as transmit antennas (full multiplexing), the projected gradient algorithm converges to the global optimum.

2) More Streams Than Transmit Antennas: For $B_k > r_k, \boldsymbol{T}_k$ is a wide matrix and the covariance matrix $\boldsymbol{Q}_k$ can again achieve the full rank r_k. Hence all propositions made for $r_k = B_k$ are also valid here and the global optimum is achieved. However, for $B_k > r_k$, the MSE of this specific user k is always larger than $B_k - r_k$ for *linear*[3] filtering even if the SNR goes to infinity. Via unitary rotations, we can always transform any precoder to a precoder with $B_k - r_k$ zero-columns without increasing the respective MSE. Hence, $B_k - r_k$ streams are ignored, which obviously does not make sense.

3) Less Streams Than Transmit Antennas: If we want to exploit some (additional) diversity gain for user k with $B_k < r_k$, then $\boldsymbol{T}_k$ is a tall matrix. Therefore, the rank of the covariance matrix $\boldsymbol{Q}_k = \boldsymbol{T}_k \boldsymbol{T}_k^{\mathrm{H}} \in \mathbb{C}^{r_k \times r_k}$ of user k is upper bounded by its maximum number of active data streams B_k and hence, $\boldsymbol{Q}_k$ cannot be chosen arbitrarily. If all covariance matrices $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_K$ were allowed to achieve their full ranks $r_1, \ldots, r_K$, minimizing (35) under a sum-power constraint would be convex in the covariances. However, if $B_k < r_k$, we need to add a rank constraint $\mathrm{rank}(\boldsymbol{Q}_k) = B_k$ which obviously does not define a convex set[4], yielding a nonconvex optimization potentially featuring local minima that are not globally optimal. Limiting the maximum number of active data streams B_k to a value smaller than the number of antennas r_k inevitably leads to the fact that the resulting optimization is no longer convex in the covariance matrices.

In any case, the solution of the uplink problem is followed by the conversion to the downlink. As only the sum-MSE needs to be preserved, we can apply the first kind of our duality described in Section IV-A since it features the smallest complexity of all three kinds that preserve the sum-MSE. Regarding this kind of minimization, the application of the first kind of duality first reveals the hidden convexity of the problem, second, allows for a gradient-based algorithm with low computational complexity due to a common user-independent inverse, and third, transforms the problem back to the downlink with negligible complexity.

B. Weighted Sum-MSE Minimization in the Broadcast Channel

We intend to minimize a weighted sum-MSE in the downlink. Instead of solving the problem in the downlink, we first solve the equivalent uplink problem and then transfer it back to the downlink by means of the second kind of our duality. The squared scaling factors $\alpha_1^2, \ldots, \alpha_K^2$ from (15) need not be computed via the duality presented in the previous section, since they turn out to be a byproduct of the solution of the optimization problem, what clearly saves computational complexity.

In the direct approach in the downlink domain a huge computational complexity is required to minimize the sum of weighted user-wise MSEs:

$$\underset{\{\boldsymbol{P}_1,\ldots,\boldsymbol{P}_K\}}{\mathrm{minimize}} \sum_{k=1}^{K} w_k \varepsilon_k^{\mathrm{DL}} \quad \mathrm{s.t.:} \sum_{k=1}^{K} \|\boldsymbol{P}_k\|_{\mathrm{F}}^2 \leq P_{\mathrm{Tx}} \tag{39}$$

where the weights $w_1, \ldots, w_K$ are positive. The Lagrangian function associated to (39) reads by means of (5) in conjunction with the MMSE receiver as

$$L = \sum_{k=1}^{K} w_k B_k - \sum_{k=1}^{K} w_k \mathrm{tr}\left[\boldsymbol{P}_k^{\mathrm{H}} \boldsymbol{H}_k \boldsymbol{Z}_k^{-1} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{P}_k\right] + \mu\Big(\sum_{k=1}^{K} \|\boldsymbol{P}_k\|_{\mathrm{F}}^2 - P_{\mathrm{Tx}}\Big) \tag{40}$$

with the substitution

$$\boldsymbol{Z}_k = \sigma_\eta^2 \mathbf{I}_{r_k} + \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{P} \boldsymbol{P}^{\mathrm{H}} \boldsymbol{H}_k. \tag{41}$$

Any gradient-based approach minimizing the sum of weighted MSEs inevitably requires the evaluation and computation of the expression

$$\frac{\partial L}{\partial \boldsymbol{P}_k^*} = \mu \boldsymbol{P}_k - \boldsymbol{H}_k \boldsymbol{Z}_k^{-1} (w_k \boldsymbol{Z}_k - \boldsymbol{S}) \boldsymbol{Z}_k^{-1} \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{P}_k \tag{42}$$

where the matrix $\boldsymbol{S}$ is defined via

$$\boldsymbol{S} = \sum_{j=1}^{K} w_j \boldsymbol{H}_j^{\mathrm{H}} \boldsymbol{P}_j \boldsymbol{P}_j^{\mathrm{H}} \boldsymbol{H}_j. \tag{43}$$

[3]In case of nonlinear filtering, it may be desirable to transmit more streams than transmit antennas available, see [49].

[4]The sum of two rank-one matrices can be of rank two for example.

From (42), it is obvious to see that K different matrices $\boldsymbol{Z}_1, \dots, \boldsymbol{Z}_K$ need to be inverted, or at least the product $\boldsymbol{H}_k \boldsymbol{Z}_k^{-1}$ has to be generated which brings about an enormous computational complexity.

This large amount can be circumvented by switching to the dual MAC. There, the optimization reads as

$$\underset{\{\boldsymbol{T}_1,\dots,\boldsymbol{T}_K\}}{\text{minimize}} \sum_{k=1}^{K} w_k \varepsilon_k^{\text{UL}} \quad \text{s.t.}: \sum_{k=1}^{K} \|\boldsymbol{T}_k\|_{\text{F}}^2 \leq P_{\text{Tx}} \tag{44}$$

and the associated Lagrangian function is

$$L = \sum_{k=1}^{K} w_k B_k - \sum_{k=1}^{K} w_k \text{tr}\left[\boldsymbol{T}_k^{\text{H}} \boldsymbol{H}_k^{\text{H}} \boldsymbol{X}^{-1} \boldsymbol{H}_k \boldsymbol{T}_k\right] + \mu \Big(\sum_{k=1}^{K} \|\boldsymbol{T}_k\|_{\text{F}}^2 - P_{\text{Tx}} \Big). \tag{45}$$

Differentiating (45) with respect to $\boldsymbol{T}_k^*$ yields

$$\frac{\partial L}{\partial \boldsymbol{T}_k^*} = \mu \boldsymbol{T}_k - \boldsymbol{H}_k^{\text{H}} \boldsymbol{X}^{-1} (w_k \boldsymbol{X} - \boldsymbol{S}) \boldsymbol{X}^{-1} \boldsymbol{H}_k \boldsymbol{T}_k \tag{46}$$

where the matrix $\boldsymbol{S}$ now reads as

$$\boldsymbol{S} = \sum_{i=1}^{K} w_i \boldsymbol{H}_i \boldsymbol{T}_i \boldsymbol{T}_i^{\text{H}} \boldsymbol{H}_i^{\text{H}} \tag{47}$$

and $0 < \mu \in \mathbb{R}_+$ is the Lagrangian multiplier associated with the power constraint. Now in the dual uplink, there is only one common matrix $\boldsymbol{X}$ that needs to be inverted for the computation of *all* gradients in contrast to the downlink gradients in (42), where K different inverses $\boldsymbol{Z}_1^{-1}, \dots, \boldsymbol{Z}_K^{-1}$ have to be evaluated. This clearly reduces the computational complexity compared to the direct downlink approach.

Having solved the minimization in (44) for example by means of the projected gradient approach, it is obvious that the precoding filters satisfy the first-order KKT conditions of (44):[5]

$$\mu \check{\boldsymbol{T}}_k = \boldsymbol{H}_k^{\text{H}} \check{\boldsymbol{X}}^{-1} (w_k \check{\boldsymbol{X}} - \check{\boldsymbol{S}}) \check{\boldsymbol{X}}^{-1} \boldsymbol{H}_k \check{\boldsymbol{T}}_k \; \forall k. \tag{48}$$

LHS multiplying (48) by $\check{\boldsymbol{T}}_k^{\text{H}}$ and applying the trace operator, we obtain at any stationary point by means of $\check{\boldsymbol{G}}_k = \check{\boldsymbol{T}}_k^{\text{H}} \boldsymbol{H}_k^{\text{H}} \check{\boldsymbol{X}}^{-1}$ from (33)

$$\mu \|\check{\boldsymbol{T}}_k\|_{\text{F}}^2 = \text{tr}\big(\check{\boldsymbol{G}}_k (w_k \check{\boldsymbol{X}} - \check{\boldsymbol{S}}) \check{\boldsymbol{G}}_k^{\text{H}}\big) = w_k \sigma_\eta^2 \|\check{\boldsymbol{G}}_k\|_{\text{F}}^2 + w_k \sum_{\substack{i=1 \\ i \neq k}}^{K} \|\check{\boldsymbol{G}}_k \boldsymbol{H}_i \check{\boldsymbol{T}}_i\|_{\text{F}}^2 - \sum_{\substack{i=1 \\ i \neq k}}^{K} w_i \|\check{\boldsymbol{G}}_k \boldsymbol{H}_i \check{\boldsymbol{T}}_i\|_{\text{F}}^2 \; \forall k. \tag{49}$$

The above equation is similar to the system of equations for the uplink to downlink transformation (17), except for the constant factor $\mu \sigma_\eta^{-2}$. Therefore, we can interpret the weights w_k as the scaled and *squared* transformation coefficients α_k

$$\frac{\alpha_1^2}{w_1} = \cdots = \frac{\alpha_K^2}{w_K}. \tag{50}$$

From (50) we can observe, that we do not need to solve the system of (18) for the final uplink to downlink transformation as the scalars $\alpha_1^2, \dots, \alpha_K^2$ are already determined by the weights $w_1, \dots, w_K$ up to a constant factor α_0 which is chosen such that the resulting downlink filters $\boldsymbol{P}_1, \dots, \boldsymbol{P}_K$ fulfill the transmit power constraint (3). Summing up, we first solve the dual uplink problem for example by means of the projected gradient approach. Afterwards, we exploit the relationship (50) between the weights $w_1, \dots, w_K$ and the squared conversion coefficients $\alpha_1^2, \dots, \alpha_K^2$ and set $\alpha_k = \alpha_0 \sqrt{w_k}$ and make use of (15) to find the respective filters in the downlink. Finally, the coefficient α_0 is found from the transmit power constraint in the downlink.

[5]The check-sign $(\check{\cdot})$ denotes matrices where precoders are involved that fulfill the KKT conditions.

C. Simulation Results

Although the main contribution of this paper are the various kinds of dualities, we present some simulation results of the described algorithm for the sum-MSE minimization to illustrate the usefulness of the proposed MSE duality framework. We choose a system setup, where $K = 3$ users are served by a base station with $N = 6$ antennas. Every terminal is equipped with 2 antennas and two data streams are allocated to every user, i.e., $r_k = B_k = 2 \; \forall k$. For this configuration we compare the sum-MSE minimization algorithm from Section VI-A with the one taken from Table III in [29], which is said to be the fastest one of the three variants presented in [29], and which will act as the reference algorithm. As in [16], the inverse step-size $s^{(\ell)}$ of the gradient projection update of our algorithm in (36) is initialized with $s^{(0)} = 2$. Different to our projected gradient based approach which solves the sum-MSE minimization in the dual MAC and converts only the *final* solution back to the BC, the authors in [29] repeatedly switch between the BC and the dual MAC. Their duality therefore has to be applied in every single iteration instead of only once at the end. Moreover, the iteration itself contains another convex minimization problem which has to be solved by interior point methods as proposed in [29]. This inner optimization has to be solved with high accuracy such that the outer optimization yields accurate results as well. Otherwise, the MSE saturates above its minimum value and increasing the number of outer iterations does not bring any reduction of the sum-MSE. As mentioned in Section I-C, the currently achieved tuple of MSEs per stream has to be computed as a vehicle to convert the precoders, receive beamformers, and the power allocation of the dual uplink back to the BC in [29]. However, only the respective filters in the BC are of interest and not the specific MSEs. Finally, K eigenvalue decompositions must be performed in [29]. Our duality clearly saves complexity during the conversion by directly transforming the filters from one domain to the other. Summing up, a single iteration of the algorithm in [29] has a much higher computational complexity than an iteration of the proposed one in this paper. Nonetheless, we compare the convergence of the sum-MSE over the iteration ignoring the difference in the computational complexity per iteration.

As a figure of merit we choose the number of iterations necessary to reduce the sum-MSE until its relative error is below a certain threshold. To this end, we averaged over 10000 different channel realizations with zero-mean i.i.d. Gaussian entries and plotted the histogram of the number of iterations for three different transmit signal-to-noise ratio regimes: a low SNR regime with $10 \log_{10}(P_{\text{Tx}}/\sigma_\eta^2) = 0$ dB, a moderate SNR regime with $10 \log_{10}(P_{\text{Tx}}/\sigma_\eta^2) = 10$ dB, and a high SNR regime with $10 \log_{10}(P_{\text{Tx}}/\sigma_\eta^2) = 20$ dB. The precoders of

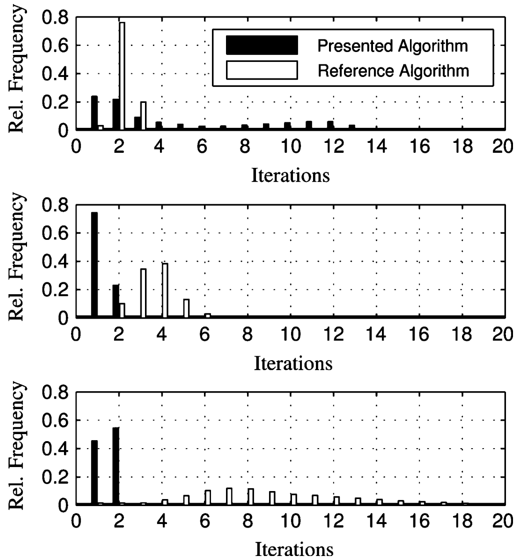

Fig. 3. Histogram of the number of iterations needed to reduce the relative error of the minimum sum-MSE below a certain threshold. The transmit SNR $10\log_{10}(P_{\rm Tx}/\sigma_\eta^2)$ is 0 dB in the upper plot, 10 dB in the middle plot, and 20 dB in the lower plot.

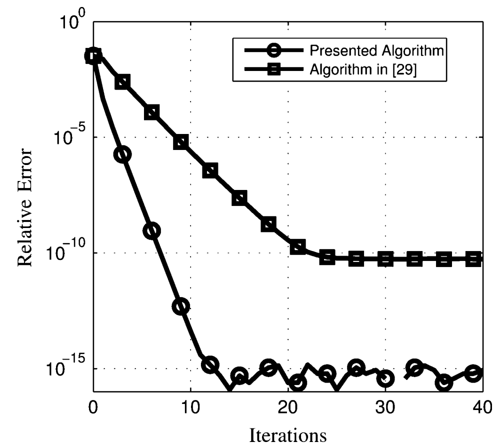

Fig. 4. Relative error versus iteration index showing the speed of convergence for a typical channel with $10\log_{10}(P_{\rm Tx}/\sigma_\eta^2) = 10$ dB.

both algorithms were initialized with the same scaled identity matrices such that the transmit power constraint is fulfilled.

In the upper plot of Fig. 3, the relative frequency of the number of iterations to reduce the relative error below 10^{-3} is plotted for the low SNR regime with $10\log_{10}(P_{\rm Tx}/\sigma_\eta^2) = 0\,{\rm dB}$. In the low power regime, the presented algorithm is outperformed by the algorithm taken from Table III in [29]. The latter one reaches the desired accuracy of 10^{-3} within 3 iterations in about 99 percent of the channel realizations, whereas the presented algorithm takes more than 3 iterations in 45 percent of the cases. Although the algorithm in [29] needs less iterations in the low power regime, its computational complexity can be reduced by applying the presented duality instead of the one in [29]. The performance of the two algorithm changes when the transmit SNR is increased.

The middle plot in Fig. 3 shows the relative frequency of the number of iterations for the moderate SNR regime and the same accuracy of 10^{-3}. In about 74 percent of the cases, only a single iteration is necessary for the proposed algorithm to reach the relative error target, in 23 percent of the cases, 2 iterations are necessary. This means that only in 3 percent of the cases, more than 2 iterations have to be run. In contrast, the algorithm in [29] needs three or more iterations in about 90 percent of the cases.

Raising the transmit SNR to 20 dB, the algorithm in [29] needs drastically more iterations. Hence, we reduced the accuracy to 10^{-2}. For this threshold, the presented algorithm reached the desired relative error target within 2 iterations for all channel realizations, whereas the histogram for the iterations of the algorithm in [29] is much broader, see the lower plot in Fig. 3.

The speed of convergence is depicted in Fig. 4, where the relative error of the two algorithms is shown over the number of iterations for a representative channel realization with $P_{\rm Tx} = 10\,{\rm dB}$. Both algorithms converge linearly, which means that the error is multiplied by a factor $0 < \beta < 1$ per iteration. It can be observed that this factor β is smaller for the presented algorithm than for the one in [29]. Because of the nonzero tolerance of the inner optimization during the iterations in the algorithm from [29], the relative error saturates with an error floor of about $5 \cdot 10^{-11}$. We used a tolerance of 10^{-10} on the cost function for the inner optimization which is solved by a nonlinear problem solver. The reason for the saturation of the presented algorithm at about 10^{-15} is the finite wordlength representation.

VII. Conclusion

We presented a complete framework for the MSE duality between the MAC and the BC, which is applicable for single and multiantenna users. Due to the three different kinds of duality, an application specific version with tailored complexity can be chosen for the solution of common optimization problems. The first kind is the simplest one which leaves the total sum-MSE constant and has lowest complexity. A conservation of the individual users' MSEs is obtained by the second kind. An even finer resolution can be achieved by the third kind which preserves every single stream's MSE when switching from uplink to downlink or vice versa. Two areas of application have been presented: The first one was the minimization of the total sum-MSE in the BC which turned out to be convex in the dual uplink MAC. The second one was the weighted sum-MSE minimization for which the computational complexity can be reduced by means of our duality.

Appendix I
Generalized Uplink to Downlink Transformation

In this section, we extend the second kind of our duality for the strictly active users in Section IV-B-I to the general case, where we also allow for active receive filters, when the respective transmit filters are zero matrices and vice versa. To this end, it is advantageous to define the following four sets of users:

$$\begin{aligned} \mathbb{P}_{\rm Tx} &= \{k | \boldsymbol{T}_k = \boldsymbol{0}_{r_k \times B_k}\} \\ \mathbb{A}_{\rm Tx} &= \{k | \boldsymbol{T}_k \neq \boldsymbol{0}_{r_k \times B_k}\} \\ \mathbb{P}_{\rm Rx} &= \{k | \boldsymbol{G}_k = \boldsymbol{0}_{B_k \times N}\} \\ \mathbb{A}_{\rm Rx} &= \{k | \boldsymbol{G}_k \neq \boldsymbol{0}_{B_k \times N}\}. \end{aligned} \quad (51)$$

The set $\mathbb{P}_{\rm Tx}$ comprises all *passive* transmitters, whereas $\mathbb{A}_{\rm Tx}$ consists of all *active* transmitters. For the receivers, the equivalent definition holds for $\mathbb{P}_{\rm Rx}$ and $\mathbb{A}_{\rm Rx}$. For the completely passive users that neither transmit nor receive, one direction of the duality is evident as aforementioned, so we can assume here that

$$\begin{aligned}&\mathbb{P}_{\mathrm{Rx}}\cap\mathbb{P}_{\mathrm{Tx}}=\emptyset,\\&\mathbb{P}_{\mathrm{Rx}}\cup\mathbb{A}_{\mathrm{Rx}}=\mathbb{P}_{\mathrm{Tx}}\cup\mathbb{A}_{\mathrm{Tx}}=\{1,\ldots,K\},\\&\mathbb{P}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}}=\mathbb{P}_{\mathrm{Rx}},\quad\mathbb{P}_{\mathrm{Tx}}\cap\mathbb{A}_{\mathrm{Rx}}=\mathbb{P}_{\mathrm{Tx}}.\end{aligned}\tag{52}$$

For the *normal users* that actively transmit and receive, i.e., $k\in\mathbb{A}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}}$, the unknowns α_k underlie the strict inequalities $0<\alpha_k<\infty$, see (15). *Passive receivers* actively transmitting ($k\in\mathbb{P}_{\mathrm{Rx}}$) have $\boldsymbol{G}_k=\boldsymbol{0}_{B_k\times N}$. From (16), we can infer that $\alpha_k\rightarrow\infty$ must hold in order to let the RHS go to zero as the left hand side is. This leads to a vanishing receive filter in the dual downlink $\boldsymbol{B}_k=\alpha_k^{-1}\boldsymbol{T}_k^{\mathrm{H}}=\boldsymbol{0}_{B_k\times r_k}, k\in\mathbb{P}_{\mathrm{Rx}}$. However, the transmit filter $\boldsymbol{P}_k=\alpha_k\boldsymbol{G}_k^{\mathrm{H}}$ in the downlink turns out to be finite! To prove this assertion, we model passive receivers and transmitters as a limit process in the sequel. We set

$$\boldsymbol{G}_k=g\boldsymbol{G}'_k\ \forall k\in\mathbb{P}_{\mathrm{Rx}}\tag{53}$$

with $g\in\mathbb{R}_+$ and an arbitrary[6], but nonzero, matrix $\boldsymbol{G}'_k\neq\boldsymbol{0}_{B_k\times N}$. We will then let g go down to zero in the limit. The respective transmitter $\boldsymbol{P}_k$ in the downlink then reads as

$$\boldsymbol{P}_k=\alpha_k\boldsymbol{G}_k^{\mathrm{H}}=\alpha_k g\boldsymbol{G}_k'^{\mathrm{H}}=\beta_k\boldsymbol{G}_k'^{\mathrm{H}}\tag{54}$$

where $\beta_k:=\alpha_k g$ and $0<\beta_k<\infty$ even in the limit $g\rightarrow 0$. For the passive receivers, the unknowns are β_k. Finally, we have the *passive transmitters* that actively receive ($k\in\mathbb{P}_{\mathrm{Tx}}\cap\mathbb{A}_{\mathrm{Rx}}$). Since $\boldsymbol{T}_k=\boldsymbol{0}_{r_k\times B_k}$, $\alpha_k\rightarrow 0$ must hold to let the RHS of (16) be different from zero as the LHS is. As a consequence

$$\boldsymbol{P}_k=\alpha_k\boldsymbol{G}_k^{\mathrm{H}}=\boldsymbol{0}_{N\times B_k}, k\in\mathbb{P}_{\mathrm{Tx}}\tag{55}$$

follows from (15), but $\boldsymbol{B}_k=\alpha_k^{-1}\boldsymbol{T}_k^{\mathrm{H}}$ is finite. Similar to the *passive receivers* in (53), we introduce

$$\boldsymbol{T}_k=g\boldsymbol{T}'_k\ \forall k\in\mathbb{P}_{\mathrm{Tx}}\tag{56}$$

with arbitrary, but nonzero $\boldsymbol{T}'_k$. In the limit process, we will let g go to zero to model the passive transmitters. The receive filter in the downlink is

$$\boldsymbol{B}_k=\alpha_k^{-1}\boldsymbol{T}_k^{\mathrm{H}}=\alpha_k^{-1}g\boldsymbol{T}_k'^{\mathrm{H}}:=\chi_k^{-1}\boldsymbol{T}_k'^{\mathrm{H}}\tag{57}$$

with the finite unknowns for the passive transmitters $\chi_k:=\alpha_k g^{-1}$ and $0<\chi_k<\infty$ even in the limit when $g\rightarrow 0$.

Starting with the *nonpassive* users ($k\in\mathbb{A}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}}$), (16) can be rewritten by means of (52)–(57) to (58):

$$\begin{aligned}&\sum_{\substack{i\in\mathbb{A}_{\mathrm{Tx}}\\i\neq k}}\|\boldsymbol{G}_k\boldsymbol{H}_i\boldsymbol{T}_i\|_{\mathrm{F}}^2+\sum_{i\in\mathbb{P}_{\mathrm{Tx}}}g^2\|\boldsymbol{G}_k\boldsymbol{H}_i\boldsymbol{T}'_i\|_{\mathrm{F}}^2+\sigma_\eta^2\|\boldsymbol{G}_k\|_{\mathrm{F}}^2\\&=\frac{1}{\alpha_k^2}\Bigg[\sum_{\substack{i\in\mathbb{A}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}}\\i\neq k}}\alpha_i^2\|\boldsymbol{G}_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2+\sum_{i\in\mathbb{P}_{\mathrm{Rx}}}\underbrace{\alpha_i^2g^2}_{\beta_i^2}\|\boldsymbol{G}'_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2\\&\qquad+\sum_{i\in\mathbb{P}_{\mathrm{Tx}}}\underbrace{\alpha_i^2}_{\chi_i^2g^2}\|\boldsymbol{G}_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2+\sigma_\eta^2\|\boldsymbol{T}_k\|_{\mathrm{F}}^2\Bigg].\end{aligned}\tag{58}$$

[6]Only the amount of generated interference is of interest.

The two sums in (58) with $i\in\mathbb{P}_{\mathrm{Tx}}$ vanish in the limit $g\rightarrow 0\Leftrightarrow\alpha_i\rightarrow 0$, hence we can drop them. Multiplying (58) by α_k^2 with $0<\alpha_k<\infty$, we find for $k\in\mathbb{A}_{\mathrm{Tx}}\cap\mathbb{A}_{\mathrm{Rx}}$

$$\begin{aligned}&\alpha_k^2\Big[\sum_{\substack{i\in\mathbb{A}_{\mathrm{Tx}}\\i\neq k}}\|\boldsymbol{G}_k\boldsymbol{H}_i\boldsymbol{T}_i\|_{\mathrm{F}}^2+\sigma_\eta^2\|\boldsymbol{G}_k\|_{\mathrm{F}}^2\Big]\\&\quad-\sum_{\substack{i\in\mathbb{A}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}}\\i\neq k}}\alpha_i^2\|\boldsymbol{G}_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2\\&\quad-\sum_{i\in\mathbb{P}_{\mathrm{Rx}}}\beta_i^2\|\boldsymbol{G}'_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2=\sigma_\eta^2\|\boldsymbol{T}_k\|_{\mathrm{F}}^2.\end{aligned}\tag{59}$$

Second, rewriting (16) for the *passive receivers* ($k\in\mathbb{P}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}},\ \alpha_k\rightarrow\infty$) and applying (52)–(57), we get

$$\begin{aligned}&\sum_{\substack{i\in\mathbb{A}_{\mathrm{Tx}}\\i\neq k}}g^2\|\boldsymbol{G}'_k\boldsymbol{H}_i\boldsymbol{T}_i\|_{\mathrm{F}}^2+\sum_{i\in\mathbb{P}_{\mathrm{Tx}}}g^2g^2\|\boldsymbol{G}'_k\boldsymbol{H}_i\boldsymbol{T}'_i\|_{\mathrm{F}}^2+\sigma_\eta^2g^2\|\boldsymbol{G}'_k\|_{\mathrm{F}}^2\\&\quad=\frac{1}{\alpha_k^2}\Big[\sum_{i\in\mathbb{A}_{\mathrm{Tx}}\cap\mathbb{A}_{\mathrm{Rx}}}\alpha_i^2\|\boldsymbol{G}_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2+\sum_{\substack{i\in\mathbb{P}_{\mathrm{Rx}}\\i\neq k}}\underbrace{g^2\alpha_i^2}_{\beta_i^2}\|\boldsymbol{G}'_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2\\&\qquad+\sum_{i\in\mathbb{P}_{\mathrm{Tx}}}\underbrace{\alpha_i^2}_{\chi_i^2g^2}\|\boldsymbol{G}_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2+\sigma_\eta^2\|\boldsymbol{T}_k\|_{\mathrm{F}}^2\Big].\end{aligned}\tag{60}$$

This time, (60) is already fulfilled in the limit when $g\rightarrow 0\Leftrightarrow\alpha_k\rightarrow\infty$. Hence, we are free to impose arbitrary constraints with (60), which may be advantageous for the duality. The key idea now is to multiply (60) by $\alpha_k^2<\infty$ *before* $g\rightarrow 0$ is applied. The resulting product $\beta_k^2=g^2\alpha_k^2$ is finite also in the limit $g\rightarrow 0\Leftrightarrow\alpha_k\rightarrow\infty$, see (54). The reason for doing so is the nice property of the resulting system of equations which again generates an *M-matrix* and hence ensures the functionality of our duality by obtaining positive values for α_i^2 and β_i^2. Multiplying (60) by α_k^2 and applying the limit $g\rightarrow 0$ *afterwards*, we obtain for $k\in\mathbb{P}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}}$

$$\begin{aligned}&\beta_k^2\Big[\sum_{\substack{i\in\mathbb{A}_{\mathrm{Tx}}\\i\neq k}}\|\boldsymbol{G}'_k\boldsymbol{H}_i\boldsymbol{T}_i\|_{\mathrm{F}}^2+\sigma_\eta^2\|\boldsymbol{G}'_k\|_{\mathrm{F}}^2\Big]\\&\quad-\sum_{i\in\mathbb{A}_{\mathrm{Tx}}\cap\mathbb{A}_{\mathrm{Rx}}}\alpha_i^2\|\boldsymbol{G}_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2\\&\quad-\sum_{\substack{i\in\mathbb{P}_{\mathrm{Rx}}\\i\neq k}}\beta_i^2\|\boldsymbol{G}'_i\boldsymbol{H}_k\boldsymbol{T}_k\|_{\mathrm{F}}^2=\sigma_\eta^2\|\boldsymbol{T}_k\|_{\mathrm{F}}^2.\end{aligned}\tag{61}$$

Interestingly, (59) and (61) suffice to determine the unknowns $\alpha_k^2, k\in\mathbb{A}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}}$ and $\beta_k^2, k\in\mathbb{P}_{\mathrm{Rx}}\cap\mathbb{A}_{\mathrm{Tx}}$, which are overall $|\mathbb{A}_{\mathrm{Tx}}|$ unknowns. Equations (59) and (61) represent a system of $|\mathbb{A}_{\mathrm{Tx}}|$ equations with $|\mathbb{A}_{\mathrm{Tx}}|$ unknowns, which is again column diagonally dominant with positive main diagonal entries and negative off-diagonal entries. Hence, a positive solution for α_k^2 and β_k^2 exists, and it is guaranteed that every user-MSE tuple in the uplink can also be obtained in the downlink. Moreover, summing up the $|\mathbb{A}_{\mathrm{Tx}}|$ equations, we find that

$$\begin{aligned}&\sigma_\eta^2\Big[\sum_{i\in\mathbb{A}_{\mathrm{Tx}}\cap\mathbb{A}_{\mathrm{Rx}}}\alpha_i^2\|\boldsymbol{G}_i\|_{\mathrm{F}}^2+\sum_{i\in\mathbb{A}_{\mathrm{Tx}}\cap\mathbb{P}_{\mathrm{Rx}}}\beta_i^2\|\boldsymbol{G}'_i\|_{\mathrm{F}}^2\Big]\\&\quad=\sigma_\eta^2\sum_{i\in\mathbb{A}_{\mathrm{Tx}}}\|\boldsymbol{T}_i\|_{\mathrm{F}}^2\quad\Leftrightarrow\quad\sum_{i\in\mathbb{A}_{\mathrm{Tx}}}\|\boldsymbol{P}_i\|_{\mathrm{F}}^2\\&\quad=\sum_{i\in\mathbb{A}_{\mathrm{Tx}}}\|\boldsymbol{T}_i\|_{\mathrm{F}}^2=P_{\mathrm{Tx}}.\end{aligned}\tag{62}$$

We see that the power constraint is inherently fulfilled, i.e., the LHS of (62), which corresponds to the consumed power in the downlink, is identical to the RHS, which represents the consumed power P_{Tx} in the uplink.

Now, the MSE equality is assured for all *active* transmitters in the uplink, and we have determined all transmit matrices $\boldsymbol{P}_k$ for the downlink. Remember that passive transmitters in the uplink are passive transmitters in the downlink as well [see (55)]. The only task remaining is to ensure the MSE equality for the passive transmitters by choosing the receive matrices $\boldsymbol{B}_k$ for $k \in \mathbb{P}_{\text{Tx}} \cap \mathbb{A}_{\text{Rx}}$. For those k, $\alpha_k = 0$ and we want to determine the values χ_k, see (57). As we will see, these χ_k can be computed separately, which becomes obvious when we reflect the system scenario: In the downlink, all transmit filters $\boldsymbol{P}_k \forall k$ have been determined. Hence, the MSE obtained by user $k \in \mathbb{P}_{\text{Tx}} \cap \mathbb{A}_{\text{Rx}}$ only depends on the choice of the receive filter $\boldsymbol{B}_k = \chi_k^{-1} \boldsymbol{T}_k'^{\text{H}}$. The downlink MSE is identical to the uplink MSE for users k with $k \in \mathbb{P}_{\text{Tx}} \cap \mathbb{A}_{\text{Rx}}$ $(\alpha_k \rightarrow 0)$, if the following equality holds, where we again apply (52)–(57):

$$
\begin{aligned}
&\sum_{i \in \mathbb{A}_{\text{Tx}}} \|\boldsymbol{G}_k \boldsymbol{H}_i \boldsymbol{T}_i\|_{\text{F}}^2 + \sum_{\substack{i \in \mathbb{P}_{\text{Tx}} \\ i \neq k}} g^2 \|\boldsymbol{G}_k \boldsymbol{H}_i \boldsymbol{T}_i'\|_{\text{F}}^2 + \sigma_\eta^2 \|\boldsymbol{G}_k\|_{\text{F}}^2 \\
&= \sum_{i \in \mathbb{A}_{\text{Rx}} \cap \mathbb{A}_{\text{Tx}}} \underbrace{\frac{g^2}{\alpha_k^2}}_{\chi_k^{-2}} \alpha_i^2 \|\boldsymbol{G}_i \boldsymbol{H}_k \boldsymbol{T}_k'\|_{\text{F}}^2 \\
&\quad + \sum_{i \in \mathbb{P}_{\text{Rx}}} \underbrace{\frac{g^2}{\alpha_k^2}}_{\chi_k^{-2}} \underbrace{\alpha_i^2 g^2}_{\beta_i^2} \|\boldsymbol{G}_i' \boldsymbol{H}_k \boldsymbol{T}_k'\|_{\text{F}}^2 \\
&\quad + \sum_{\substack{i \in \mathbb{P}_{\text{Tx}} \\ i \neq k}} \underbrace{\frac{g^2}{\alpha_k^2}}_{\chi_k^{-2}} \underbrace{\alpha_i^2}_{g^2 \chi_i^2} \|\boldsymbol{G}_i \boldsymbol{H}_k \boldsymbol{T}_k'\|_{\text{F}}^2 + \sigma_\eta^2 \underbrace{\frac{g^2}{\alpha_k^2}}_{\chi_k^{-2}} \|\boldsymbol{T}_k'\|_{\text{F}}^2.
\end{aligned} \tag{63}
$$

Multiplying (63) by χ_k^2 and applying the limit $g \rightarrow 0$, we obtain

$$
\begin{aligned}
\chi_k^2 = &\frac{\sum\limits_{i \in \mathbb{A}_{\text{Rx}} \cap \mathbb{A}_{\text{Tx}}} \alpha_i^2 \|\boldsymbol{G}_i \boldsymbol{H}_k \boldsymbol{T}_k'\|_{\text{F}}^2}{\sum\limits_{i \in \mathbb{A}_{\text{Tx}}} \|\boldsymbol{G}_k \boldsymbol{H}_i \boldsymbol{T}_i\|_{\text{F}}^2 + \sigma_\eta^2 \|\boldsymbol{G}_k\|_{\text{F}}^2} \\
&+ \frac{\sum\limits_{i \in \mathbb{P}_{\text{Rx}}} \beta_i^2 \|\boldsymbol{G}_i' \boldsymbol{H}_k \boldsymbol{T}_k'\|_{\text{F}}^2 + \sigma_\eta^2 \|\boldsymbol{T}_k'\|_{\text{F}}^2}{\sum\limits_{i \in \mathbb{A}_{\text{Tx}}} \|\boldsymbol{G}_k \boldsymbol{H}_i \boldsymbol{T}_i\|_{\text{F}}^2 + \sigma_\eta^2 \|\boldsymbol{G}_k\|_{\text{F}}^2}.
\end{aligned} \tag{64}
$$

By means of $\boldsymbol{B}_k = \chi_k^{-1} \boldsymbol{T}_k'^{\text{H}}$, we have determined all transmit and receive matrices now. Summing up, one direction of the generalized duality for the users' MSEs is obtained by first solving the system of $|\mathbb{A}_{\text{Tx}}|$ (59) and (61), which, in conjunction with the fact that passive transmitters in the uplink are passive transmitters in the downlink as well, provides all precoders and the receive matrices of all active transmitters. Then, the receive filters of the passive transmitters are computed such that the MSE equality between uplink and downlink holds. This can be done independently, see (64).

Appendix II
Generalized Downlink to Uplink Transformation

In this section, we describe the duality even for passive receivers that actively transmit and passive transmitters, that actively receive. To this end, we assume again that no *virtual users* with $k \in \mathbb{P}_{\text{Tx}} \cap \mathbb{P}_{\text{Rx}}$ are present in the system, since for them, the duality is evident, i.e., $\boldsymbol{T}_k = \boldsymbol{0}_{r_k \times B_k}$ and $\boldsymbol{G}_k = \boldsymbol{0}_{B_k \times N}$. Hence, they can be dropped, and we assume that the set properties (52) are still valid.

Passive receivers $(k \in \mathbb{P}_{\text{Rx}})$ have $\boldsymbol{B}_k = \boldsymbol{0}_{B_k \times r_k}$. From (27), we can infer that $\bar{\alpha}_k \rightarrow \infty$ must hold in order to let the RHS be zero as the LHS is. Thus, $\boldsymbol{G}_k = \bar{\alpha}_k^{-1} \boldsymbol{P}_k^{\text{H}} = \boldsymbol{0}_{N \times B_k}$, see (26). The respective receiver in the dual uplink is herewith passive as well. To show that the precoder $\boldsymbol{T}_k = \bar{\alpha}_k \boldsymbol{B}_k^{\text{H}}$ in the uplink is finite, we introduce [cf. (53)]

$$
\boldsymbol{B}_k = g \boldsymbol{B}_k' \ \forall k \in \mathbb{P}_{\text{Rx}} \cap \mathbb{A}_{\text{Tx}} \tag{65}
$$

with arbitrary nonzero $\boldsymbol{B}_k' \in \mathbb{C}^{B_k \times r_k}$ and $g \rightarrow 0$ in the limit to model the passive receivers. The precoder $\boldsymbol{T}_k$ in the uplink thus reads as

$$
\boldsymbol{T}_k = \bar{\alpha}_k \boldsymbol{B}_k^{\text{H}} = \bar{\alpha}_k g \boldsymbol{B}_k'^{\text{H}} = \bar{\beta}_k \boldsymbol{B}_k'^{\text{H}} \ \forall k \in \mathbb{P}_{\text{Rx}} \tag{66}
$$

with $\bar{\beta}_k := \bar{\alpha}_k g$ and $0 < \bar{\beta}_k < \infty$.

Passive transmitters with $k \in \mathbb{P}_{\text{Tx}}$ have $\boldsymbol{P}_k = \boldsymbol{0}_{N \times B_k}$ leading to $\bar{\alpha}_k = 0$ in order to let the RHS of (27) be different from zero, as the LHS is not zero. The respective transmit filter $\boldsymbol{T}_k$ in the dual uplink reads as $\boldsymbol{T}_k = \bar{\alpha}_k \boldsymbol{B}_k^{\text{H}} = \boldsymbol{0}_{r_k \times B_k}$, whereas the receive filter $\boldsymbol{G}_k$ can be expressed as

$$
\boldsymbol{G}_k = \bar{\alpha}_k^{-1} \boldsymbol{P}_k^{\text{H}} = \bar{\alpha}_k^{-1} g \boldsymbol{P}_k'^{\text{H}} = \bar{\chi}_k^{-1} \boldsymbol{P}_k'^{\text{H}} \tag{67}
$$

where $\bar{\chi}_k^{-1} := \bar{\alpha}_k^{-1} g$ with $0 < \bar{\chi}_k < \infty$, and

$$
\boldsymbol{P}_k = g \boldsymbol{P}_k' \ \forall k \in \mathbb{P}_{\text{Tx}} \tag{68}
$$

with arbitrary, but nonzero $\boldsymbol{P}_k' \in \mathbb{C}^{N \times B_k}$.

For the *nonpassive users*, i.e., $k \in \mathbb{A}_{\text{Tx}} \cap \mathbb{A}_{\text{Tx}}$, (27) can be rewritten by means of (52) and (65)–(68) and after letting $g \rightarrow 0$, we obtain [cf. (59)]

$$
\begin{aligned}
&\bar{\alpha}_k^2 \Big[\sum_{\substack{i \in \mathbb{A}_{\text{Tx}} \\ i \neq k}} \|\boldsymbol{B}_k \boldsymbol{H}_k^{\text{H}} \boldsymbol{P}_i\|_{\text{F}}^2 + \sigma_\eta^2 \|\boldsymbol{B}_k\|_{\text{F}}^2 \Big] \\
&\quad - \sum_{\substack{i \in \mathbb{A}_{\text{Rx}} \cap \mathbb{A}_{\text{Tx}} \\ i \neq k}} \bar{\alpha}_i^2 \|\boldsymbol{B}_i \boldsymbol{H}_i^{\text{H}} \boldsymbol{P}_k\|_{\text{F}}^2 \\
&\quad - \sum_{i \in \mathbb{P}_{\text{Rx}}} \bar{\beta}_i^2 \|\boldsymbol{B}_i' \boldsymbol{H}_i^{\text{H}} \boldsymbol{P}_k\|_{\text{F}}^2 = \sigma_\eta^2 \|\boldsymbol{P}_k\|_{\text{F}}^2.
\end{aligned} \tag{69}
$$

After multiplying with $\bar{\alpha}_k^2$, including (65)–(68), and applying $g \to 0$, the condition (27) for *passive* receivers $(k \in \mathbb{P}_{\mathrm{Rx}}, \bar{\alpha}_k \to \infty)$ transforms to [cf. (61)]

$$\begin{aligned}\bar{\beta}_k^2 \Big[\sum_{\substack{i \in \mathbb{A}_{\mathrm{Tx}} \\ i \neq k}} \|\boldsymbol{B}'_k \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{P}_i\|_{\mathrm{F}}^2 + \sigma_\eta^2 \|\boldsymbol{B}'_k\|_{\mathrm{F}}^2 \Big] \\ - \sum_{i \in \mathbb{A}_{\mathrm{Rx}} \cap \mathbb{A}_{\mathrm{Tx}}} \bar{\alpha}_i^2 \|\boldsymbol{B}_i \boldsymbol{H}_i^{\mathrm{H}} \boldsymbol{P}_k\|_{\mathrm{F}}^2 \\ - \sum_{\substack{i \in \mathbb{P}_{\mathrm{Rx}} \\ i \neq k}} \bar{\beta}_i^2 \|\boldsymbol{B}'_i \boldsymbol{H}_i^{\mathrm{H}} \boldsymbol{P}_k\|_{\mathrm{F}}^2 = \sigma_\eta^2 \|\boldsymbol{P}_k\|_{\mathrm{F}}^2 \end{aligned} \quad (70)$$

where we used $g^2 \bar{\alpha}_k^2 = \bar{\beta}_k^2$. Equations (69) and (70) in addition to (26) and (66) determine the transmit filters $\boldsymbol{T}_k$ and receive filters $\boldsymbol{G}_k$ for $k \in \mathbb{A}_{\mathrm{Tx}}$ with the same MSE for those users as in the downlink. In combination with $\boldsymbol{T}_k = \boldsymbol{0}_{r_k \times B_k}$ for $k \in \mathbb{P}_{\mathrm{Tx}}$, all precoding matrices have been determined now. The only remaining task is to find the receive matrices $\boldsymbol{G}_k$ for $k \in \mathbb{P}_{\mathrm{Tx}}$. In turn, these equations are decoupled, since after including (65)–(68), multiplying with $\bar{\chi}_k^2$, and $g \to 0$, (27) for $k \in \mathbb{P}_{\mathrm{Tx}}$ $(\bar{\alpha}_k \to 0)$ leads to

$$\bar{\chi}_k^2 = \frac{\sum_{i \in \mathbb{A}_{\mathrm{Rx}} \cap \mathbb{A}_{\mathrm{Tx}}} \bar{\alpha}_i^2 \|\boldsymbol{B}_i \boldsymbol{H}_i^{\mathrm{H}} \boldsymbol{P}'_k\|_{\mathrm{F}}^2}{\sum_{i \in \mathbb{A}_{\mathrm{Tx}}} \|\boldsymbol{B}_k \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{P}_i\|_{\mathrm{F}}^2 + \sigma_\eta^2 \|\boldsymbol{B}_k\|_{\mathrm{F}}^2} + \frac{\sum_{i \in \mathbb{P}_{\mathrm{Rx}}} \bar{\beta}_i^2 \|\boldsymbol{B}'_i \boldsymbol{H}_i^{\mathrm{H}} \boldsymbol{P}'_k\|_{\mathrm{F}}^2 + \sigma_\eta^2 \|\boldsymbol{P}'_k\|_{\mathrm{F}}^2}{\sum_{i \in \mathbb{A}_{\mathrm{Tx}}} \|\boldsymbol{B}_k \boldsymbol{H}_k^{\mathrm{H}} \boldsymbol{P}_i\|_{\mathrm{F}}^2 + \sigma_\eta^2 \|\boldsymbol{B}_k\|_{\mathrm{F}}^2}. \quad (71)$$

Summing up, the downlink to uplink transformation is achieved by first computing the $|\mathbb{A}_{\mathrm{Tx}}|$ unknowns $\bar{\alpha}_k, k \in \mathbb{A}_{\mathrm{Rx}} \cap \mathbb{A}_{\mathrm{Tx}}$ and $\bar{\beta}_k, k \in \mathbb{P}_{\mathrm{Rx}}$ from (69) and (70), respectively. This system of equations has again the nice property that the unknowns are positive and there is a unique solution as long as $\sigma_\eta^2 \neq 0$. Afterwards, we compute $\bar{\chi}_k^2$ from (71) for all $k \in \mathbb{P}_{\mathrm{Tx}}$ and determine all transmit and receive filters via (26), (66), and (67).

References

[1] D. Li, Z.-Y. Wu, H.-W. J. Lee, X.-M. Yang, and L.-S. Zhang, "Hidden convex minimization," *J. Global Optimiz.*, vol. 31, no. 2, pp. 211–233, Feb. 2005.

[2] J. Zander and M. Frodigh, "Comment on performance of optimum transmitter power control in cellular radio," *IEEE Tran. Veh. Technol.*, vol. 43, p. 636, Aug. 1994.

[3] D. Gerlach and A. Paulraj, "Base station transmitting antenna arrays for multipath environments," *Signal Process.*, vol. 54, no. 1, pp. 59–73, Oct. 1996.

[4] H. Boche and M. Schubert, "Optimum SIR balancing using extended 1-Norm beamforming optimization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Orlando, FL, May 2002, vol. 3, pp. 2945–2948.

[5] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1437–1450, Oct. 1998.

[6] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE Veh. Technol. Conf.*, May 1999, pp. 851–856.

[7] H. Boche and M. Schubert, "Solution of the SINR downlink beamforming problem," in *Proc. Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, Mar. 2002.

[8] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," in *Proc. IEEE Veh. Technol. Conf. (VTC) Fall*, Vancouver, Canada, Sep. 2002, vol. 1, pp. 87–91.

[9] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1912–1921, Aug. 2003.

[10] D. Tse and P. Viswanath, "On the capacity of the multiple antenna broadcast channel," in *Proc. Multiantenna Channels: Capacity, Coding and Signal Process., DIMACS Workshop Amer. Math. Soc. 2003*, Oct. 2002, vol. 62, pp. 87–105.

[11] D. A. Schmidt, M. Joham, R. Hunger, and W. Utschick, "Near maximum sum-rate nonzero-forcing linear precoding with successive user selection," in *Proc. 40th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2006.

[12] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 18–28, Jan. 2004.

[13] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[14] A. Khachan, A. Tenenbaum, and R. S. Adve, "Linear processing for the downlink in multiuser MIMO systems with multiple data streams," in *Proc. IEEE Int. Conf. Commun.*, 2006.

[15] A. Khachan, A. J. Tenenbaum, and R. Adve, "Joint transmitter-receiver optimization for downlink multiuser MIMO communications," in *IEEE Trans. Wireless Commun.*, 2006, submitted for publication.

[16] A. Mezghani, M. Joham, R. Hunger, and W. Utschick, "Transceiver design for multiuser MIMO systems," in *Proc. ITG/IEEE WSA 2006*, Mar. 2006.

[17] M. Schubert and H. Boche, "Joint 'Dirty paper' pre-coding and downlink beamforming," in *Proc. 7th IEEE Int. Symp. Spread Spectrum Tech. Appl. (ISSSTA 2002)*, Prague, Czech Republic, Sep. 2002, vol. 2, pp. 536–540.

[18] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[19] M. Schubert and H. Boche, "User ordering and power allocation for optimal multiantenna precoding/decoding," in *Proc. ITG/IEEE Workshop on Smart Antennas*, Munich, Germany, Mar. 2004.

[20] S. Shi and M. Schubert, "Precoding and power loading for multiantenna broadcast channels," in *Conf. Inf. Sci. Syst. (CISS)*, Princeton, NJ, Mar. 2004.

[21] M. Schubert and S. Shi, "MMSE transmit optimization with interference pre-compensation," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Stockholm, Sweden, May 2005.

[22] M. Schubert, S. Shi, E. A. Jorswieck, and H. Boche, "Downlink sum-MSE transceiver optimization for linear multiuser MIMO systems," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Monterey, CA, Sep. 2005.

[23] R. Hunger, M. Joham, and W. Utschick, "Extension of linear and nonlinear transmit filters for decentralized receivers," in *Proc. Eur. Wireless 2005*, Nicosia, Apr. 2005, vol. 1, pp. 40–46.

[24] E. Jorswieck and H. Boche, "Transmission strategies for the MIMO MAC with MMSE receiver: Average MSE optimization and achievable individual MSE region," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2872–2881, Nov. 2003.

[25] S. Shi and M. Schubert, "MMSE transmit optimization for multiuser multiantenna systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Philadelphia, PA, Mar. 2005.

[26] W. Utschick and M. Joham, "On the duality of MIMO transmission techniques for multiuser communications," *Proc. EUSIPCO 2006*, Sep. 2006.

[27] A. Mezghani, R. Hunger, M. Joham, and W. Utschick, "Iterative THP transceiver optimization for multiuser MIMO systems based on weighted sum-MSE minimization," *Proc. SPAWC 2006*, Jul. 2006.

[28] R. Doostnejad, T. J. Lim, and E. Sousa, "Precoding for the MIMO broadcast channels with multiple antennas at each receiver," in *Proc. Johns Hopkins Univ. Conf. on Inf. Sci. Syst.*, Baltimore, MD, Mar. 2005.

[29] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization," *IEEE Trans. Signal Process.*, vol. 55, pp. 5436–5446, Nov. 2007.

[30] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT 2002)*, Jul. 2002, p. 500.

[31] S. Vishwanath, N. Jindal, and A. J. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2658–2668, Oct. 2003.

[32] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," *Proc. ISIT 2004*, p. 174, Jun. 2004.

[33] N. Jindal, S. Vishwanath, and A. Goldsmith, "On the duality of Gaussian multiple-access and broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, pp. 768–783, May 2004.

[34] S. Vishwanath, W. Rhee, N. Jindal, S. A. Jafar, and A. Goldsmith, "Sum power iterative waterfilling for Gaussian vector broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2003.

[35] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. J. Goldsmith, "Sum power ierative water-filling for multiantenna Gaussian broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 1570–1580, Apr. 2005.

[36] D. Palomar, J. Cioffi, and M. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A convex optimization approach," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[37] R. Hunger, W. Utschick, M. Joham, and D. A. Schmidt, "Alternating optimization for MMSE broadcasting," *Proc. ICASSP 2006*, vol. IV, pp. IV-757–IV-760, May 2006.

[38] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York: Academic [Harcourt Brace Jovanovich], 1979.

[39] R. S. Varga, *Matrix Iterative Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1962.

[40] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Philadelphia, PA: Soc. Ind. Appl. Math., 2003.

[41] G. H. Golub and C. V. Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1991.

[42] R. Hunger, Floating Point Operations in Matrix-Vector Calculus Munich Univ. Technol., Tech. Rep. TUM-LNS-TR-05-05, 2005.

[43] S. Shi and M. Schubert, "Iterative MMSE optimization for multiuser MIMO systems," in *Proc. ITG/IEEE Workshop Smart Antennas*, Ulm, Germany, Mar. 2006.

[44] M. Schubert, S. Shi, and H. Boche, "Iterative transceiver optimization for linear multiuser MIMO channels with per-user MMSE requirements," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2006.

[45] J. Brinkhuis, Z. Luo, and S. Zhang, Matrix Convex Functions With Applications to Weighted Centers for Semidefinite Programming Dep. Syst. Eng. Eng. Manage., The Chinese Univ. Hong Kong, 2006, Technical Report SEEM2005-06.

[46] A. Tenenbaum and R. Adve, "Joint multiuser transmit-receive optimization using linear processing," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, Jun. 2004, vol. 1, pp. 588–592.

[47] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Commun. Lett.*, vol. 49, pp. 2198–2206, Dec. 2001.

[48] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computations*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[49] R. Hunger, M. Riemensberger, and W. Utschick, "Near capacity Tomlinson-Harashima precoding with uniform stream properties," in *Proc. ITG/IEEE WSA 2007*, Vienna, Austria, Feb. 2007.

Raphael Hunger (S'06) was born in Illertissen, Germany, 1979. He studied electrical engineering at the Technische Universität München (TUM), Germany, from 1999 until 2005 and, in 2004, at the RWTH Aachen, Germany. He received the Dipl.-Ing. and Master of Science degrees in electrical engineering from the TUM in 2004 and 2005, respectively.

Since 2005, he has been working toward the doctorate degree at the Associate Institute for Signal Processing, TUM. His research interests focus on the joint optimization of transmitters and receivers in multiuser MIMO communications.

Michael Joham (S'99-M'05) was born in Kufstein, Austria, 1974. He received the Dipl.-Ing. and Dr.-Ing. degrees (both *summa cum laude*) in electrical engineering from the Technische Universität München (TUM), Germany, in 1999 and 2004, respectively.

He was with the Institute of Circuit Theory and Signal Processing, TUM, from 1999 to 2004. Since 2004, he has been with the Associate Institute for Signal Processing, TUM, where he is currently a senior researcher. In summer 1998 and 2000, he visited Purdue University, West Lafayette, IN. In spring 2008, he was a guest professor at the University of A Coruña, Spain. His current research interests are precoding in mobile communications, limited rate feedback, MIMO communications, and robust signal processing.

Dr. Joham received the VDE Preis for his diploma thesis in 1999 and the Texas-Instruments-Preis for his dissertation in 2004. In 2007, he was a corecipient of the Best Paper Award of the International ITG/IEEE Workshop on Smart Antennas in Vienna.

Wolfgang Utschick (SM'06) was born on May 6, 1964. He completed several industrial education programs before he received the diploma and doctoral degrees, both with honors, in electrical engineering from the Technische Universität München, Germany (TUM), in 1993 and 1998, respectively. In this period he held a scholarship of the Bavarian Ministry of Education for exceptional students.

From 1998 to 2002, he codirected the SP Group of the Institute of Circuit Theory and Signal Processing, TUM. Since 2000, he has been consulting in 3 GPP standardization in the field of multielement antenna systems. In 2002, he was appointed Professor at the TUM where he is Head of the Associate Institute for Signal Processing. He teaches courses on signal processing, stochastic processes and optimization theory in the field of digital communications.

Dr. Utschick was awarded in 2006 for his excellent teaching records and in 2007, received the ITG Award 2007 of the German Society for Information Technology (ITG). He is a senior member of the VDE/ITG. He is also an Associate Editor the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART I: FUNDAMENTAL THEORY AND APPLICATIONS.