Linear Transceiver Design for Downlink Multiuser MIMO Systems: Downlink-Interference Duality Approach

Tadilo Endeshaw Bogale, Member, IEEE, and Luc Vandendorpe, Fellow, IEEE

Abstract—This paper considers linear transceiver design for downlink multiuser multiple-input multiple-output (MIMO) systems. We examine different transceiver design problems. We focus on two groups of design problems. The first group is the weighted sum mean-square-error (WSMSE) (i.e., symbol-wise or user-wise WSMSE) minimization problems and the second group is the minimization of the maximum weighted mean-square-error (WMSE) (symbol-wise or user-wise WMSE) problems. The problems are examined for the practically relevant scenario where the power constraint is a combination of per base station (BS) antenna and per symbol (user), and the noise vector of each mobile station is a zero-mean circularly symmetric complex Gaussian random variable with arbitrary covariance matrix. For each of these problems, we propose a novel downlink-interference duality based iterative solution. Each of these problems is solved as follows. First, we establish a new mean-square-error (MSE) downlink-interference duality. Second, we formulate the power allocation part of the problem in the downlink channel as a Geometric Program (GP). Third, using the duality result and the solution of GP, we utilize alternating optimization technique to solve the original downlink problem. For the first group of problems, we have established symbol-wise and user-wise WSMSE downlink-interference duality. These duality are established by formulating the noise covariance matrices of the interference channels as fixed point functions. On the other hand, for the second group of problems, we have established symbol-wise and user-wise WMSE downlink-interference duality. These duality are established by formulating the noise covariance matrices of the interference channels as marginally stable (convergent) discrete-time-switched systems. The proposed duality based iterative solutions can be extended straightforwardly to solve many other linear transceiver design problems. We also show that our MSE downlink-interference duality unify all existing MSE duality. In our simulation results, we have observed that the proposed duality based iterative algorithms utilize less total BS power than that of the existing algorithms.

Index Terms—Multiuser MIMO, MSE, downlink-uplink duality, downlink-interference duality, fixed point function, discrete-time-switched system, convex optimization.

I. INTRODUCTION

MULTIPLE-INPUT MULTIPLE-OUTPUT (MIMO) is a promising technique to exploit the spectral efficiency of wireless channels. This spectral efficiency can be exploited by applying signal processing at the transmitter (precoder) and receiver (decoder). Signal processing is performed to meet a certain design criterion. It is well known that most practically relevant design problems such as weighted sum rate maximization, rate or signal-to-interference-plus-noise-ratio (SINR) balancing and rate or SINR constrained power minimization can be equivalently expressed as mean-square-error (MSE) based problems (see for example [1]). Because of this, the current paper examines MSE-based problems. In general, the uplink channel MSE-based problems are better understood than those of the downlink channel. Due to this fact, most literatures focus on solving the downlink MSE-based problems. The downlink MSE-based problems can be solved by direct approach as in [2], [3] or by uplink-downlink duality approach as in [4]–[6].

For a given downlink channel system model and its MSE-based problem, the idea behind uplink-downlink duality is first to create the virtual uplink channel by exchanging the roles of the transmitter and receiver, and then to enable the precoder/decoder transformation from uplink to downlink channel and vice versa by ensuring the same MSE in both channels. Once these two tasks are performed, the downlink MSE-based problems are examined as follows: When the global optimality of the dual uplink channel MSE-based problem is guaranteed, the duality approach simply transfers the optimal uplink channel precoder/decoder pairs from uplink to downlink channel (see for example the sum MSE minimization problem in [5]). When the global optimal solution of the dual uplink channel MSE-based problem can not ensured, the duality approach examines the downlink MSE-based problems by iteratively switching between the uplink and downlink channel problems (see for example the problems in [7]).

Several MSE-based problems have been examined by duality approach [4]–[8]. However, the duality of these papers are able to solve total BS power constrained MSE-based problems only. In a practical multi-antenna base station (BS) system, the maximum power of each BS antenna is limited [9]. In some scenario allocating different powers to different users (symbols) according to their priority or protection level has some interest. This motivates [10] to solve (robust) sum MSE-based problems
with per antenna, user and symbol power constraints by duality approach. However, since the problems in [10] allocate the same MSE weight to all symbols (users), [10] ignores priority and fairness issues in terms of MSE.

In a multimedia communication, different types of information (for example, audio and video information) can be sent to a user (all users) simultaneously [11], [12]. In such a scenario, for successful transmission, more priority (power) could be given to symbols (users) corresponding to the video information. Thus, for this scenario, the design criteria may incorporate fairness/priority and power constraints for each symbol (user). On the other hand, examining combined (per antenna and symbol (user)) power constraints may have practical interest (for example in network MIMO). For these reasons, the current paper generalizes the work of [10] by incorporating both symbol-wise and user-wise MSE fairness/priority, and combined (i.e., per antenna and symbol (user)) power constraints. We examine the following problems: Minimization of symbol-wise weighted sum mean-square-error (WSMSE) constrained with per BS antenna and symbol powers ($\mathcal{P}_1$), minimization of user-wise WSMSE constrained with per BS antenna and user powers ($\mathcal{P}_2$), minimization of the maximum symbol-wise weighted mean-square-error (WMSE) constrained with per BS antenna and symbol powers ($\mathcal{P}_3$) and minimization of the maximum user-wise WMSE constrained with per BS antenna and user powers ($\mathcal{P}_4$). Each of these problems is examined for the scenario where the noise vector of each mobile station (MS) is a zero-mean circularly symmetric complex Gaussian (ZMCSCG) random variable with arbitrary covariance matrix.

To the best of our knowledge, the problems $\mathcal{P}_1 - \mathcal{P}_4$ are non-convex. Furthermore, duality based solutions for these problems with our noise covariance matrix assumptions are not known. In the current paper, we propose duality based iterative solutions to solve the problems. Each of these problems is solved as follows. First, we establish a new MSE downlink-interference duality. Second, we formulate the power allocation part of the problem in the downlink channel as a Geometric Program (GP). Third, using the duality result and the solution of GP, we utilize alternating optimization technique to solve the original downlink problem. For the problems $\mathcal{P}_1$ and $\mathcal{P}_2$, the duality are established by formulating the noise covariance matrices of the interference channels as fixed point functions. For these two problems, the noise covariance matrices of the dual interference channels are computed by modifying the approach of [10] to $\mathcal{P}_1$ and $\mathcal{P}_2$ of the current paper. On the other hand, for the problems $\mathcal{P}_3$ and $\mathcal{P}_4$, the duality are established by formulating the noise covariance matrices of the interference channels as new marginally stable (convergent) discrete-time-switched systems. The proposed duality based iterative solutions can be extended straightforwardly to solve many other linear transceiver design problems. We also show that our MSE downlink-interference duality unifies all existing MSE duality.\(^1\)

3) By employing the system model of [1] and [8], we formulate the power allocation parts of $\mathcal{P}_1 - \mathcal{P}_4$ as GPs. The GPs are formulated by applying the GP formulation approach of [1]. Consequently, we are able to solve our problems by alternating optimization technique [4], [7], [8], [10] (i.e., duality based iterative algorithm).

4) In our simulation results, we have observed that the proposed duality based iterative algorithms utilize less total BS power than that of the existing algorithms. This paper is organized as follows. In Section II, multiuser MIMO downlink and virtual interference channel system models are presented. In Section III, we formulate our problems $\mathcal{P}_1 - \mathcal{P}_4$ and discuss the general framework of our duality based iterative solutions. Sections IV–VIII present the proposed duality based iterative solutions for solving these problems. The extension of our duality based iterative algorithms to other problems is discussed in Section IX. In Section X, computer simulations are used to compare the performance of the proposed duality algorithms with that of existing algorithms. Finally, conclusions are drawn in Section XI.

**Notations:** Upper/lower case boldface letters denote matrices/column vectors. The $X_{(n,n)}$, $X_{(n,:)t}$, $tr(X)$, $X^T$, $X^H$ and $E(X)$ denote the $(n, n)$ element, $n$th row, trace, transpose, conjugate transpose and expected value of $X$, respectively. $I_n$ is an identity matrix of size $n \times n$ and $C^{M \times M} \{R^{M \times M}\}$ represent spaces of $M \times M$ matrices with complex (real) entries. The diagonal and block-diagonal matrices are represented by diag(.) and blkdiag(.) respectively. Subject to is denoted by s.t and $(.)^*$ denotes optimal solution. The superscripts $(.)^{DL}$ and $(.)^I$ denote downlink and interference, respectively.

### II. System Model

In this section, multiuser MIMO downlink and virtual interference channel system models are discussed which are shown in Fig. 1. In the downlink channel, the BS and $k$th

\(^1\)Note that the existing MSE duality are established for a total BS power based linear transceiver design problems (see [1], [4], [5]).
MS are equipped with $N$ and $M_k$ antennas, respectively. The total number of MS antennas are thus $M = \sum_{k=1}^{K} M_k$. By denoting the symbol intended for the $k$th user as $d_k \in \mathbb{C}^{S_x \times 1}$ and $S = \sum_{k=1}^{K} S_k$, the entire symbol can be written in a data vector $d \in \mathbb{C}^{S\times 1}$ as $d = \begin{bmatrix} d_1^T & \cdots & d_K^T \end{bmatrix}^T$. The BS precodes $d$ into an $N$ length vector by using its overall precoder matrix $B = [b_{K_1}, \ldots, b_{K_S}]$, where $b_{k_s} \in \mathbb{C}^{N \times 1}$ is the precoder vector of the BS for the $k$th MS $s$th symbol. The $k$th MS employs a receiver $w_k$ to estimate the symbol $d_k$. We follow the same channel matrix notations as in [8]. The estimates of the $k$th MS $s$th symbol $(\hat{d}_{k,s})$ and $k$th user $(\hat{d}_k)$ are given by

$$\hat{d}_{k,s} = w_k^H \left( H_k^H \sum_{i=1}^{K} B_i d_i + n_k \right)$$

$$\hat{d}_k = w_k^H \left( H_k^H B_k d_k + n_k \right)$$

where $H_k^H \in \mathbb{C}^{M_k \times N}$ is the channel matrix between the BS and $k$th MS, $w_k = [w_{k_1}, \ldots, w_{K_S}^H], B_k = [b_{k_1}, \ldots, b_{K_S}]$, and $n_k$ is the $k$th MS additive noise. Without loss of generality, we can assume that the entries of $d_k$ are independent and identically distributed (i.i.d) ZMCCSG random variables all with unit variance, i.e., $E\{d_k d_k^H\} = I_{S_k}, E\{d_k d_k^H\} = 0, \forall i \neq k$, and $E\{d_k n_i\} = 0, \forall i, k$. The $k$th MS noise vector is a ZMCCSG random variable with covariance matrix $R_{n_k} \in \mathbb{C}^{M_k \times M_k}$.

To establish our MSE downlink-interference duality, we model the virtual interference channel (Fig. 1(b)) by introducing precoders $[v_{k_1}, \ldots, v_{K_S}]_k$ and decoders $[t_{k_1}, \ldots, t_{K_S}]_k$, where $v_{k_s} \in \mathbb{C}^{N \times 1}$ and $t_{k_s} \in \mathbb{C}^{N \times 1}, \forall k, s$. In this channel, it is assumed that the $k$th user’s $s$th symbol $(d_{k,s})$ is an i.i.d ZMCCSG random variable with variance $\zeta_{k,s}$ and estimated independently by $t_{k_s} \in \mathbb{C}^{N \times 1}$, i.e., $E\{d_k t_k^H\} = \zeta_k, E\{d_k t_k^H\} = 0, \forall i \neq k$, and $E\{d_k n_k\} = 0, \forall i, k$. Moreover, $\{n_{k_1}, n_{K_S}\}_k$ are also ZMCCSG random variables with covariance matrices $\{\Delta_k \in \mathbb{R}^{N \times N} \} = \text{diag}(\zeta_{k_1}, \ldots, \zeta_{k_S}), \forall s$, and the channels between the $k$th transmitter and all receivers are the same (i.e., $[H_{k_1}, \ldots, H_{K_S}]_{k} \in \mathbb{C}^{N \times K}$). Note that from the system model aspect, the current paper and [10] share the same idea.

As can be seen from Fig. 1, the outputs of Fig. 1(a) and Fig. 1(b) are not the same. However, since Fig. 1(b) is a “virtual” interference channel which is introduced just to solve the downlink MSE-based problems by duality approach, the output of Fig. 1(b) is not required in practice. For this reason, the difference in the outputs of the downlink and interference channels of Fig. 1 will not affect the downlink MSE-based problem formulations and the duality based solutions.

For the downlink system model of Fig. 1(a), the symbol-wise and user-wise MSEs can be expressed as

$$\xi_{k,s}^{DL} = E_k \left\{ (d_{k,s} - \hat{d}_{k,s}) (d_{k,s} - \hat{d}_{k,s})^H \right\}$$

$$\xi_{k}^{DL} = E_k \left\{ (d_k - \hat{d}_k) (d_k - \hat{d}_k)^H \right\}$$

Using these two equations, the symbol-wise and user-wise WSMSEs can be expressed as

$$\xi_{k,s}^{WDL} = \sum_{k=1}^{K} \sum_{s=1}^{S_k} \eta_{k,s} \xi_{k,s}^{DL}$$

$$\xi_{k}^{WDL} = \sum_{k=1}^{K} \hat{\eta}_k \xi_{k,s}^{DL}$$

where $R_n = \text{blkdiag}(R_{n_1}, \ldots, R_{n_K}), \eta = \text{diag}(\eta_{k_1}, \ldots, \eta_{K_S}), \hat{\eta}_k = \text{diag}(\hat{\eta}_{k_1}, \ldots, \hat{\eta}_{k_S})$, $\zeta_k$ and $\hat{\eta}_k$ are the MSE weights of the $k$th user $s$th symbol and $k$th user, respectively. Like in the downlink channel, the interference channel symbol vector, user MSE and WSMSEs are expressed as

$$\xi_{k,s}^{I} = t_{k,s}^H \Gamma t_{k,s} + \zeta_k \Delta_k t_{k,s} + H_k v_k \zeta_k$$

$$\xi_{k}^{I} = \text{tr} \left\{ T_k^H \Gamma T_k - H_k v_k \zeta_k - \zeta_k H_k v_k \zeta_k + \zeta_k \Delta_k T_k \right\}$$

where $\Gamma = \text{blkdiag}(\Gamma_{n_1}, \ldots, \Gamma_{n_K}) = \text{diag}(\Gamma_{k_1}, \ldots, \Gamma_{k_S})$ with $\Gamma_{k_s}$ and $\Gamma_{k_s}$. In the duality based solutions, the symbol-wise and user-wise WSMSEs are expressed as

$$\xi_{k,s}^{WDL} = \text{tr} \left\{ T_k^H \Gamma T_k - H_k v_k \zeta_k - \zeta_k H_k v_k \zeta_k + \zeta_k \Delta_k T_k \right\}$$

$$\xi_{k}^{WDL} = \text{tr} \left\{ T_k^H \Gamma T_k - H_k v_k \zeta_k - \zeta_k H_k v_k \zeta_k + \zeta_k \Delta_k T_k \right\}$$

$$\text{tr} \left\{ T_k^H \Gamma T_k - H_k v_k \zeta_k - \zeta_k H_k v_k \zeta_k + \zeta_k \Delta_k T_k \right\}$$

$$\text{tr} \left\{ T_k^H \Gamma T_k - H_k v_k \zeta_k - \zeta_k H_k v_k \zeta_k + \zeta_k \Delta_k T_k \right\}$$

$$\text{tr} \left\{ T_k^H \Gamma T_k - H_k v_k \zeta_k - \zeta_k H_k v_k \zeta_k + \zeta_k \Delta_k T_k \right\}$$

$$\text{tr} \left\{ T_k^H \Gamma T_k - H_k v_k \zeta_k - \zeta_k H_k v_k \zeta_k + \zeta_k \Delta_k T_k \right\}$$

$$\text{tr} \left\{ T_k^H \Gamma T_k - H_k v_k \zeta_k - \zeta_k H_k v_k \zeta_k + \zeta_k \Delta_k T_k \right\}$$
\[ \ell_{\text{wua}} = \sum_{k=1}^{K} \lambda_k \zeta_k \]
\[ = \text{tr} \left\{ \mathbf{A}^H \Gamma c \mathbf{T} - \lambda \mathbf{A}^H \mathbf{H} \mathbf{V} \zeta - \lambda \mathbf{C} \mathbf{V}^H \mathbf{H}^H \mathbf{T} + \lambda \zeta \right\} \]
\[ + \sum_{k=1}^{K} \sum_{s=1}^{S_k} \lambda_k \mathbf{t}_k^H \Delta_k \mathbf{t}_k \]
(10)

where \( \zeta_k = \text{diag}(\zeta_{1}, \ldots, \zeta_{S_k}) \), \( \zeta = \text{blkdiag}(\zeta_{1}, \ldots, \zeta_{K}) \), \( \lambda = \text{blkdiag}(\lambda_{1}, \ldots, \lambda_{S_1}, \ldots, \lambda_{K}, \lambda_{S_K}) \), and \( \Delta_k = \sum_{s=1}^{S_k} \zeta_{1,s} \mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H \) with \( \lambda_k \) and \( \bar{\lambda}_k \) are the MSE weights of the \( k \)th user \( s \)th symbol and \( k \)th user, respectively.

### III. Problem Formulation

The aforementioned MSE-based optimization problems can be formulated as:

\[ \text{P1:} \quad \min_{\{\mathbf{B}_k\}, \{\mathbf{W}_k\}_{k=1}^{K}} \frac{1}{2} \sum_{k=1}^{K} \sum_{s=1}^{S_k} \eta_{k,s} \ell^{\text{DL}}_{k,s}, \]
\[ \text{s.t.} \quad \mathbf{B}_k^H \mathbf{B}_k \leq \bar{\rho}_k, \quad \forall k, \]
(11)

\[ \text{P2:} \quad \min_{\{\mathbf{B}_k\}, \{\mathbf{W}_k\}_{k=1}^{K}} \frac{1}{2} \sum_{k=1}^{K} \sum_{s=1}^{S_k} \eta_{k,s} \ell_{k,s}^{\text{DL}}, \]
\[ \text{s.t.} \quad \mathbf{B}_k^H \mathbf{B}_k \leq \bar{\rho}_k, \quad \forall k, \]
(12)

\[ \text{P3:} \quad \max_{\mathbf{B}_k, \mathbf{W}_k} \rho_k \ell_{k,s}^{\text{DL}}, \]
\[ \text{s.t.} \quad \mathbf{B}_k^H \mathbf{B}_k \leq \bar{\rho}_k, \quad \forall k, s, \]
(13)

\[ \text{P4:} \quad \min_{\{\mathbf{B}_k\}, \{\mathbf{W}_k\}_{k=1}^{K}} \max_{\rho_k} \ell_{k,s}^{\text{DL}}, \]
\[ \text{s.t.} \quad \mathbf{B}_k^H \mathbf{B}_k \leq \bar{\rho}_k, \quad \forall k, s, \]
(14)

where \( \bar{\rho}_k \) and \( \rho_k \) are the MSE balancing weights (maximum available power) of the \( k \)th user and \( k \)th user \( s \)th symbol, respectively, and \( \rho_k \) denotes the maximum transmitted power by the \( k \)th antenna.

For both the WSMSE minimization and min max WMSE problems, different weights are given to different symbols (users). However, at optimality the solutions of these two problems are not necessarily the same. This is due to the fact that the aim of the WSMSE minimization problem is just to minimize the WSMSE of all symbols (users) (i.e., in such a problem the minimized WSMSE of each symbol (user) depends on its corresponding channel gain), whereas the aim of min max WMSE problem is to minimize and balance the WSMSE of each symbol (user) simultaneously (i.e., in such a problem all symbols (users) achieve the same minimized WSMSE [13]). Moreover, as will be clear later, the solution approach of WSMSE minimization problem can not be extended straightforwardly to solve the min max WMSE problem. Due to these facts, we examine the WSMSE minimization and min max WMSE problems separately.

Since, the problems \( \mathcal{P}1 - \mathcal{P}4 \) are not convex, convex optimization framework can not be applied to solve them. To the best of our knowledge, duality based solutions for these problems are not known. In the following, we present an MSE downlink-interference duality based approach for solving each of these problems which is shown in Algorithm 1.2

#### Algorithm 1

**Initialization:**

For each problem, initialize \( \{\mathbf{B}_k \neq 0\}_{k=1}^{K} \) such that the power constraint functions are satisfied.3 Then, update \( \{\mathbf{W}_k\}_{k=1}^{K} \) by using minimum mean-square-error (MMSE) receiver approach, i.e.,

\[ \mathbf{W}_k = \left( \mathbf{H}_k^H \mathbf{B}_k^H \mathbf{B}_k + \mathbf{R}_{n,k} \right)^{-1} \mathbf{H}_k^H \mathbf{B}_k, \quad \forall k. \]
(15)

**Repeat:** Interference channel

1) Transfer the symbol-wise (user-wise) WSMSE or WMSE from downlink to interference channel.

2) Update the receivers of the interference channel \( \{\mathbf{t}_{k,s} \}, \forall s \) using MMSE receiver technique.

**Downlink channel**

3) Transfer the symbol-wise (user-wise) WSMSE or WMSE from interference to downlink channel.

4) Update the receivers of the downlink channel \( \{\mathbf{W}_k\}_{k=1}^{K} \) by MMSE receiver approach (15).

**Until** convergence.

The above iterative algorithm is already known in [5], [8] and [10]. However, the approaches of these papers can not ensure the power constraints of \( \mathcal{P}1 - \mathcal{P}4 \) at step 3 of Algorithm 1. Hence, one can not apply the approaches of these papers to solve \( \mathcal{P}1 - \mathcal{P}4 \). In the following sections, we establish our MSE downlink-interference duality.

#### IV. Symbol-Wise WSMSE Downlink-Interference Duality

This duality is established to solve symbol-wise WSMSE-based problems (for example \( \mathcal{P}1 \)).

**A. Symbol-Wise WSMSE Transfer (From Downlink to Interference Channel)**

In order to use this WSMSE transfer for solving \( \mathcal{P}1 \), we set the interference channel precoder, decoder, noise covariance, input covariance and MSE weight matrices as

\[ \mathbf{V} = \tilde{\beta} \mathbf{W}, \quad \mathbf{T} = \frac{\mathbf{B}}{\beta} \mathbf{I} - \mathbf{\Psi} \mathbf{I}, \quad \mathbf{\Lambda} = \mathbf{I}, \quad \mathbf{\Delta}_k = \mathbf{I}, \quad \mathbf{\Psi} = \mu_k \mathbf{I} \]
(16)

\[2\] As will be clear later in Section VIII, to solve \( \mathcal{P}3 \) and \( \mathcal{P}4 \) (and more general MSE-based problems), an additional power allocation step is required. In Algorithm 1, this step is omitted for clarity of presentation.

\[3\] For the simulation, we use \( \{\mathbf{B}_k = [\mathbf{h}_{k,1}; \ldots; \mathbf{h}_{k,S_k}]\}_{k=1}^{K} \) followed by the appropriate normalization of \( \{\mathbf{B}_k\}_{k=1}^{K} \) to ensure the power constraints.
where $\beta$, $\{\psi_n\}_{n=1}^N$ and $\{\mu_{ks}, \forall s\}_{k=1}^K$ are positive real scalars that will be determined in the sequel and $\Psi = \text{diag}(\psi_1, \ldots, \psi_N)$. Substituting (16) into (9) and equating $\xi_{ws} = \xi_{w, s}^D$ yields
\[
\text{tr}\{B^H \mathbf{W} \eta \mathbf{W}^H \mathbf{H} B - B^H \mathbf{H} \mathbf{W} \eta - \eta \mathbf{W} \mathbf{H}^H \mathbf{B} + \eta\} + \frac{1}{\beta^2} \sum_{k=1}^K \sum_{s=1}^S b_{ks}^H (\Psi + \mu_{ks} \mathbf{I}_N) b_{ks}
= \text{tr}\{\eta \mathbf{W} \mathbf{H}^H \mathbf{B} \mathbf{B}^H \mathbf{W} \mathbf{H} + \eta \}\nonumber
\]
Substituting (16) into (9) and equating $\eta$ yields
\[
\beta^2 \tau = \sum_{n=1}^N \psi_n \bar{p}_n + \sum_{k=1}^K \sum_{s=1}^S \mu_{ks} \bar{b}_{ks} = \bar{p}^T \psi + \bar{p}^T \mu \tag{17}
\]
where $\tau = \text{tr}\{\eta \mathbf{W} \mathbf{R}_n \mathbf{W}\}$, $\psi = [\psi_1, \ldots, \psi_N]^T$, $\mu = [\mu_{11}, \ldots, \mu_{1S}, \ldots, \mu_{K1}, \ldots, \mu_{KS}]^T$, $\bar{p} = [\bar{p}_{11}, \ldots, \bar{p}_{1S}, \ldots, \bar{p}_{K1}, \ldots, \bar{p}_{KS}]^T$ with $\bar{p}_{ks} = b_{ks}^H b_{ks}$, $\bar{p}_n = b_{n}^H b_{n}$ and $b_{n}^H$ is the $n$th row of $\mathbf{B}$.

The above equation shows that by choosing any $\{\psi_n\}_{n=1}^N$ and $\{\mu_{ks}, \forall s\}_{k=1}^K$ that satisfy (17), one can transfer the downlink channel precoder/decoder to the interference channel decoder/precoder ensuring $\xi_{ws} = \xi_{w, s}^D$. As will be clear later, to solve (11) with $\text{Algorithm I}$, $\beta$, $\bar{p}$ and $\mu$ should be selected as in (18). This shows that step 1 of $\text{Algorithm I}$ can be carried out with (16). To perform step 2 of $\text{Algorithm I}$, we update $t_{ks}$ of (16) by using the interference channel MMSE receiver approach which is expressed as
\[
t_{ks} = (\Gamma_c + \Delta_{ks})^{-1} \mathbf{H}_k \mathbf{V}_k \xi_{ks}
= \beta \left(\mathbf{H} \mathbf{W} \eta \mathbf{W}^H \mathbf{H} + \Psi + \mu_{ks} \mathbf{I}_N\right)^{\frac{1}{2}} \mathbf{H}_k \mathbf{w}_{ks} \xi_{ks} \tag{19}
\]
where $H_k$ is the $k$th row of $\mathbf{H}$. By doing so, the interference channel symbol-wise WSMSE is upper bounded by that of the downlink channel (i.e., $\xi_{ws}^D \leq \xi_{w, s}^D$). As will be clear later, to solve (11) with $\text{Algorithm I}$, $\beta$, $\bar{p}$ and $\mu$ should be selected as in (18). This shows that step 1 of $\text{Algorithm I}$ can be carried out with (16). To perform step 2 of $\text{Algorithm I}$, we update $t_{ks}$ of (16) by using the interference channel MMSE receiver approach which is expressed as
\[
t_{ks} = (\Gamma_c + \Delta_{ks})^{-1} \mathbf{H}_k \mathbf{V}_k \xi_{ks}
= \beta \left(\mathbf{H} \mathbf{W} \eta \mathbf{W}^H \mathbf{H} + \Psi + \mu_{ks} \mathbf{I}_N\right)^{\frac{1}{2}} \mathbf{H}_k \mathbf{w}_{ks} \xi_{ks} \tag{19}
\]
where the equality is obtained from (16). The above expression shows that by choosing $\{\mu_{ks}, \forall s\}_{k=1}^K$, $\{\psi_n\}_{n=1}^N$ and $\{\mu_{ks}, \forall s\}_{k=1}^K$, there exists $\left(\mathbf{H} \mathbf{W} \eta \mathbf{W}^H \mathbf{H} + \Psi + \mu_{ks} \mathbf{I}_N\right)^{\frac{1}{2}}$ such that $\psi_n \gg f_n$ and $\mu_{ks} \gg f_{ks}$ for all $n, k$. Now, for any given $\beta$, $\{t_{ks}^H\}_{k=1}^K$ and $\{\mu_{ks}\}_{k=1}^K$, suppose that there exist $\psi_n > 0$ and $\mu_{ks} > 0$, $\forall s$, $\forall k$ such that satisfy
\[
\psi_n \geq f_n \quad \text{and} \quad \mu_{ks} \geq f_{ks} \quad \forall n, k, s \tag{24}
\]
From the above equation one can also achieve $\psi_n \bar{p}_n = f_n \bar{p}_n$, $\mu_{ks} \bar{b}_{ks} = f_{ks} \bar{b}_{ks}$ for all $n, k, s$. Summing up these expressions for all $n, k$ and $s$ results
\[
\sum_{n=1}^N \psi_n \bar{p}_n + \sum_{k=1}^K \sum_{s=1}^S \mu_{ks} \bar{b}_{ks} = \sum_{n=1}^N f_n \bar{p}_n + \sum_{k=1}^K \sum_{s=1}^S f_{ks} \bar{b}_{ks} \tag{26}
\]
This equation shows that the solution of (25) satisfies (26). Moreover, as $\{\psi_n \gg f_n\}_{n=1}^N$ and $\{\mu_{ks} \gg f_{ks}\}_{k=1}^K$, the latter solution also ensures (18). Therefore, by choosing $\{\psi_n\}_{n=1}^N$ and $\{\mu_{ks}, \forall s\}_{k=1}^K$ such that (25) is satisfied, step 3
Algorithm I can be performed. Furthermore, one can notice from (26) that $\beta^2$ can be any positive value.

Next, we show that there exists at least a set of feasible $\{\psi_n > 0\}_{n=1}^N$ and $\{\mu_{k_n} > 0, \forall s\}_{k=1}^K$, that satisfy (25). To this end, we consider the following Theorem [14].

**Theorem 1:** Let $(\frac{X}{||x||}_2)$ be a complete metric space. We say that $F : X \to X$ is an almost contraction, if there exist $\kappa(\tilde{\chi}) \in [0,1)$ and $\chi(\tilde{\chi}) \geq 0$ such that

$$
||F(x) - F(y)||_2 \leq \kappa ||x - y||_2 + \chi ||x - F(y)||_2, \forall x, y \in X.
$$

If $F$ satisfies (27), then the following holds true:

1. $\exists x \in X : x = F(x)$.
2. For any initial $x_0 \in X$, the iteration $x_{n+1} = F(x_n)$ for $n = 0, 1, 2, \cdots$ converges to some $x^* \in X$.
3. The solution $x^*$ is not necessarily unique.

**Proof:** See Theorem 1.1 of [14]. Note that according to [15] (see (1.1) and (1.2) of [15]), the two inequalities of (27) are dual to each other.

Define $x$ and $F$ as $x \xrightarrow{\Delta} [x_1, \ldots, x_{S+N}]^T = [\psi_1, \ldots, 

\psi_N, \mu_{11}, \ldots, \mu_{1S}, \ldots, \mu_{K1}, \ldots, \mu_{KS}]^T$, $F(x) \xrightarrow{\Delta} [f_1, \ldots, f_{N+1}, f_{S+1}, \ldots, f_{S+N}]$ with

$$
\{x_n = \psi_n \in [0, (\beta^2 - \epsilon \sum_{i=1}^{N} \sum_{j=1}^{S} \mu_{i,j} p_{i,j}) / \mu_{k} p_{i,k}]\}_{n=1}^N$$

and

$$
\{x_{r,S+N+1} = \mu_{k} \in [0, (\beta^2 - \epsilon \sum_{i=1}^{K} \sum_{j=1}^{S} \mu_{i,j} p_{i,k}) / \mu_{k} p_{i,k}]\}_{k=1}^K.$$  \(4\)

As we can see from (27), when $||F(x_1) - F(x_2)||_2 = 0$ with $x_1 = x_2$ or $x_1 \neq x_2$, one can set $\kappa(\tilde{\chi}) = 0$ and $\chi(\tilde{\chi}) = 0$ to satisfy this inequality. And when $||F(x_1) - F(x_2)||_2 > 0$ (i.e., $x_1 \neq x_2$), one can select appropriate $\kappa(\tilde{\chi}) \in [0,1)$ and $\chi(\tilde{\chi}) \geq 0$ such that (27) is satisfied. This is due to the fact that in the latter case, $||x_2 - F(x_1)||_2 > 0$ and/or $||x_1 - F(x_2)||_2 > 0$ and $||x_1 - x_2||_2 > 0$ are positive and bounded. This explanation shows the existence of $\kappa(\tilde{\chi}) = 0$ and $\chi(\tilde{\chi}) \geq 0$ ensuring (27) for any $F(x_1) - F(x_2)$.

Consequently, $F(x)$ is an almost contraction which implies $x_{n+1} = F(x_n), x_0 = [x_{11}, x_{12}, \cdots, x_{(S+N)}]^T \geq \epsilon 1_{N+S}$ for $n = 0, 1, 2, \cdots$ converges (28).

where $1_{N+S}$ is an $N + S$ length vector with each element equal to unity. Thus, there exist $\{\psi_n \geq \epsilon\}_{n=1}^N$ and $\{\mu_{k} \geq \epsilon, \forall s\}_{k=1}^K$ that satisfy (25) and can be computed using (28). For numerical simulation we initialize $x_0$ as $x_{01} = x_{02} = \cdots = x_{0(S+N)}$. However, finding the optimal initialization strategy is still an open research topic.

Once the appropriate $\{\psi_n\}_{n=1}^N$ and $\{\mu_{k} \}_{k=1}^K$ are obtained, step 4 of Algorithm I is immediate and hence $P_1$ can be solved iteratively using this algorithm.

**C. Extension of the Current Duality for $P_1$ With a Total BS Power Constraint**

If the constraints of $P_1$ are modified to a total BS power, the power constraint at step 3 of Algorithm I can be ensured by applying the precoder/decoder transformation expression of [5]. The precoder/decoder transformation of [5] is performed by computing $S$ scaling factors. These scaling factors are obtained by solving $S$ systems of equations which require matrix inversion with complexity $O(S^3)$ (see (23) of [5]).

In the current paper, if the constraints of $P_1$ are modified to a total BS power, one can ensure the power constraint at step 3 of Algorithm I just by assigning $\Delta_{k_s}$ of (16) as $\Delta_{k_s} = \mathbf{1}$. By doing so, $\beta^2$ of (17) and $\beta^2$ of (21) can be expressed as $\beta^2 = \sum_{n=1}^{N} \sum_{s=1}^{N} b_n b_s^T p_{n,m}$ and $\beta^2 = \frac{p_{max}}{\tau}$ and $\beta^2 = \frac{p_{max}}{\tau}$, where $P_{max}$ is the total BS power. Now by employing (20), the total BS power at step 3 of Algorithm I can thus be given as $tr_{\{B B^H\}} - \beta^2 tr_{\{T T^H\}} - \beta^2 \tau = P_{max}$ (i.e., the total BS power constraint is satisfied). Thus, for $P_1$ (with a total BS power constraint), we do not need to use Theorem I. Moreover, our duality requires only one scaling factor to perform the precoder/decoder transformation (i.e., $\beta^2$). This shows that for this problem, the proposed duality based algorithm requires less computation compared to that of [5]. Note that the duality algorithm of [5] requires the same computation as that of [8] and less computation than that of [1] and [4]. Thus, it is sufficient to compare the current duality algorithm with the duality algorithm of [5].

For other WSMSE-based problems with a total BS power constraint function, the computational advantage of the current duality based algorithm over that of [5] can be analyzed like in this subsection.

**V. USER-WISE WSMSE DOWNLINK-INTERFERENCE DUALITY**

This duality is established to solve user-wise WSMSE-based problems (for example $P_2$).

**A. User-Wise WSMSE Transfer (From Downlink to Interference Channel)**

To apply this WSMSE transfer for solving $P_2$, we set the precoder, decoder and noise covariance matrices as

$$
V = \beta W, T = \frac{B}{\beta}, \check{\zeta} = \check{\eta}, \tilde{\lambda} = \mathbf{1}, \Delta_{k_s} = \Psi + \mu_k \mathbf{1}
$$

where $\beta, \{\psi_n\}_{n=1}^N$ and $\{\mu_{k} \}_{k=1}^K$ are real positive scalars. Substituting (29) into (10) and equating $\xi_{\text{w}}^T = \xi_{\text{w}}^D$ yields

$$
\bar{\beta}^2 \tilde{\tau} = \sum_{n=1}^{N} \psi_n p_n + \sum_{k=1}^{K} \mu_k p_k = \tilde{p}^T \psi + \tilde{p}^T \mu
$$

where $\tilde{\tau} = tr_{\{\check{\eta} W^H R_{\mathbf{w}} W\}}$, $\check{\mu} = [\mu_{1}, \ldots, \mu_{K}]^T$, $\tilde{p} = [p_{1}, \ldots, p_{K}]^T$ with $\check{\mu} = tr\{B \check{B}^H\}$. Like in Section IV-A, we perform step 1 of Algorithm I by choosing $\bar{\beta}^2$, $\psi$ and $\mu$ as

$$
\bar{\beta}^2 \tilde{\tau} \geq \tilde{p}^T \psi + \tilde{p}^T \mu.
$$

(31)
To perform step 2 of \textbf{Algorithm I}, we update $t_{k,s}$ of (29) using the interference channel MMSE receiver as

$$t_{k,s} = \hat{\beta} \left( \mathbf{H} \mathbf{W} \hat{\eta} \mathbf{W}^H \mathbf{H}^H + \Psi + \mu_k I \right)^{-1} \mathbf{H}_k \mathbf{w}_{k,s} \hat{\eta}_k.$$ (32)

This expression shows that by choosing $\{\mu_k > 0\}_{k=1}^K,\{\psi_n > 0\}_{n=1}^N$, we ensure that $(\mathbf{H} \mathbf{W} \hat{\eta} \mathbf{W}^H \mathbf{H}^H + \Psi + \mu_k I)^{-1}$ exists.

\subsection*{B. User-Wise WSMSE Transfer (From Interference to Downlink Channel)}

For a given user-wise WSMSE in the interference channel with $\zeta = \hat{\eta}$ and $\hat{\lambda} = \mathbf{I}$, we can achieve the same WSMSE in the downlink channel (with the weighting matrix $\mathbf{W}$) by using a nonzero scaling factor $\{\tilde{\beta}\}$ which satisfies

$$\mathbf{B} = \frac{\tilde{\beta}}{\beta} \mathbf{T}, \quad \mathbf{W} = \frac{\mathbf{V}}{\tilde{\beta}}.$$ (33)

In this precoder/decoder transformation, we use the notations $\mathbf{B}$ and $\mathbf{W}$ to differentiate from the precoder and decoder matrices used in Section V-A. By substituting (33) into $\mathcal{E}_{w,b}$ with $\mathbf{B} = \mathbf{B}, \mathbf{W} = \mathbf{W}$, then equating the resulting user-wise WSMSE with that of the interference channel ($\mathcal{E}_{w,b}$) and after simple manipulations, we get

$$\tilde{\beta} = \frac{\beta}{\sum_{n=1}^N \psi_n \tilde{t}_n^H \tilde{t}_n + \sum_{k=1}^K \mu_k \text{tr}\{T_k^H T_k\}}$$ (34)

where $\tilde{t}_n$ is the $n$th row of the MMSE matrix $\mathbf{T}$ (32). The power constraints of each BS antenna and user (i.e., step 3 of \textbf{Algorithm I}) in the downlink channel can be expressed as

$$\psi_n \geq \tilde{f}_n \text{ and } \mu_k \geq \tilde{f}_k, \forall k$$ (35)

where

$$\tilde{f}_n = \frac{\tilde{\beta}^2}{\sum_{n=1}^N \psi_n \tilde{t}_n^H \tilde{t}_n + \sum_{k=1}^K \mu_k \text{tr}\{T_k^H T_k\}}$$ (36)

$$\tilde{f}_k = \frac{\tilde{\beta}^2}{\sum_{n=1}^N \psi_n \tilde{t}_n^H \tilde{t}_n + \sum_{k=1}^K \mu_k \text{tr}\{T_k^H T_k\}}$$ (37)

For given $\tilde{\beta}$, $\{\tilde{t}_n^H \tilde{t}_n\}_{n=1}^N$ and $\{\text{tr}\{T_k^H T_k\}\}_{k=1}^K$, one can show that there exist $\{\psi_n\}_{n=1}^N$ and $\{\mu_k\}_{k=1}^K$ which satisfy

$$\psi_n \geq \tilde{f}_n \text{ and } \mu_k \geq \tilde{f}_k, \forall n, k.$$ (38)

The solution of (38) can be obtained exactly like that of (25). As $\{\tilde{\psi}_n \geq \tilde{f}_n\}_{n=1}^N$ and $\{	ilde{\mu}_k \geq \tilde{f}_k\}_{k=1}^K$, the latter solution also satisfies (31). Thus, $\mathcal{P}2$ can be solved using \textbf{Algorithm I}.

\section*{VI. SYMBOL-WISE MSE DOWNLINK-INTERFERENCE DUALITY}

In this section, we establish the symbol-wise MSE duality between downlink and interference channels. If all symbols are active, this duality can be applied to solve MSE based problems. However, as will be clear later, this duality requires more computation compared to the duality of Sections IV and V. Thus, we propose this duality to be employed for problems like in $\mathcal{P}3$ since this problem maintains all symbols active and can not be solved by the duality in Sections IV and V.

\subsection*{A. Symbol-Wise MSE Transfer (From Downlink to Interference Channel)}

To apply this duality for $\mathcal{P}3$, we set the interference channel precoder, decoder, noise covariance, input covariance and MSE weight matrices as

$$\mathbf{v}_{k,s} = \hat{\beta}_{k,s} \mathbf{w}_{k,s}, \quad t_{k,s} = \frac{b_{k,s}}{\hat{\beta}_{k,s}} \zeta = \mathbf{I}, \quad \Delta_{k,s} = \Psi + \mu_{k,s} \mathbf{I}, \forall k, s.$$ (39)

Substituting (39) into (7) and $\{\ell^D_{w,b} = \ell^L_{w,b}\}_{k=1}^K$ yields

$$\mathbf{w}_{k,s}^H \left( \mathbf{H}_k^H \sum_{i=1}^K \sum_{j=1}^S \gamma_{i,j} b_{i,j}^H \mathbf{H}_k + \mathbf{R}_{n,k} \right) \mathbf{w}_{k,s}$$

$$= \mathbf{w}_{k,s}^H \mathbf{H}_k^H \mathbf{b}_{k,s} - \mathbf{b}_{k,s}^H \mathbf{H}_k \mathbf{w}_{k,s} + 1$$

$$= \frac{1}{\beta_{k,s}^2} \mathbf{b}_{k,s}^H \left( \sum_{i=1}^K \sum_{j=1}^S \beta_{i,j}^2 \gamma_{i,j} w_{i,j}^H \mathbf{H}_k^H + \Psi + \mu_{k,s} \mathbf{I} \right) \mathbf{b}_{k,s}$$

$$= \frac{1}{\beta_{k,s}^2} \mathbf{b}_{k,s}^H \left( \sum_{i=1}^K \sum_{j=1}^S \beta_{i,j}^2 \gamma_{i,j} w_{i,j}^H \mathbf{H}_k^H \right) \mathbf{b}_{k,s}, \forall k, s.$$ (40)

It implies

$$\mathbf{w}_{k,s}^H \left( \mathbf{H}_k^H \sum_{i=1}^K \sum_{j=1}^S \gamma_{i,j} b_{i,j}^H \mathbf{H}_k + \mathbf{R}_{n,k} \right) \mathbf{w}_{k,s}$$

$$= \frac{1}{\beta_{k,s}^2} \mathbf{b}_{k,s}^H \left( \sum_{i=1}^K \sum_{j=1}^S \beta_{i,j}^2 \gamma_{i,j} w_{i,j}^H \mathbf{H}_k^H \right) \mathbf{b}_{k,s}, \forall k, s.$$ (41)

Collecting the above expression for all $k$ and $s$ gives

$$\left( \mathbf{Y} + \Theta \tilde{\beta} \right)^2 = [a_{11}, \cdots, a_{1S}, \cdots, a_{K1}, \cdots, a_{KS}]^T = \mathbf{P} x$$

$$\Rightarrow \mathbf{Y} = \frac{1}{\beta_{k,s}^2} \mathbf{b}_{k,s}^H \left( \sum_{i=1}^K \sum_{j=1}^S \beta_{i,j}^2 \gamma_{i,j} w_{i,j}^H \mathbf{H}_k^H \right) \mathbf{b}_{k,s}, \forall k, s.$$ (42)

$$\mathbf{P} = \frac{1}{\beta_{k,s}^2} \mathbf{b}_{k,s}^H \left( \sum_{i=1}^K \sum_{j=1}^S \beta_{i,j}^2 \gamma_{i,j} w_{i,j}^H \mathbf{H}_k^H \right) \mathbf{b}_{k,s}.$$ (43)

Theorem 2: Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $A_{(i,j)(i',j')} \leq 0, 0 \leq i(i') \leq n$. If the diagonal elements of $\mathbf{A}$ are $\mathbf{A}_{(i,i)} = 1 - \sum_{j=1,j\neq i}^n A_{(j,i)}$, then

Property 1: $\mathbf{A}^{-1} \geq 0$ (42)

Property 2: $||\mathbf{A}^{-1}||_1 = 1$ (43)

where $(.) \geq 0$ and $||.||_1$ denote matrix non-negativity and one norm, respectively.

\textit{Proof:} See Appendix A.
According to the first property of Theorem 2, if \( \{ \eta_{ks} > 0, \forall s \} \) and \( \{ \mu_{ks} > 0, \forall s \} \) of (41) are strictly positive. Now, by selecting \( \{ \eta_{ks} > 0, \forall s \} \) and \( \{ \mu_{ks} > 0, \forall s \} \) of (41) is fulfilled, we can transfer the MSE of each symbol from downlink to interference channel ensuring that the power constraint of \( \mathcal{P}_3 \) at step 3 of Algorithm 1 is satisfied. To this end, we examine the steps (2) and (3) of this algorithm.

Like in Section IV, we perform step 2 of Algorithm 1 by updating \( t_{ks} \) using MMSE receiver as

\[
\begin{align*}
\mathbf{b}_{ks} &= \frac{1}{\beta_{ks}} \mathbf{H}_k \mathbf{v}_{ks} \\
\mathbf{w}_{ks} &= \mathbf{v}_{ks} / \beta_{ks}.
\end{align*}
\]

where the second equality is obtained from (39).

The above expression shows that by choosing \( \{ \mu_{ks} > 0, \forall s \} \) and \( \{ \eta_{ks} > 0 \} \) of (41), we ensure that \( \{ \sum_{j=1}^K \beta_{kj} \mathbf{H}_i \mathbf{w}_{kj} \mathbf{v}_{kji}^H + \mathbf{v}_{ks} / \beta_{ks} \} \) exists. Next, we transfer the symbol-wise MSE from interference to downlink channel by satisfying the power constraint of \( \mathcal{P}_3 \) (i.e., we perform step 3).

**B. Symbol-Wise MSE Transfer (From Interference to Downlink Channel)**

For a given symbol MSE in the interference channel with \( \zeta = 1 \), we can achieve the same symbol MSE in the downlink channel by using a nonzero scaling factor \( \beta_{ks} \) which satisfies

\[
\mathbf{b}_{ks} = \beta_{ks} \mathbf{t}_{ks}, \quad \mathbf{w}_{ks} = \mathbf{v}_{ks} / \beta_{ks}.
\]

Here we use the notations \( \mathbf{B} \) and \( \mathbf{W} \) to differentiate with the precoder and decoder matrices used in Section IV-A. By substituting (45) into \( \xi_{ks}^{DL} \) (with \( \mathbf{B} = \mathbf{B}, \mathbf{W} = \mathbf{W} \)), then equating the resulting symbol MSE with that of the interference channel (7) and after some straightforward steps, we get

\[
\begin{align*}
\mathbf{v}_{ks}^H \mathbf{H}_k \mathbf{v}_{ks} &= \left( \sum_{i=1}^K \sum_{j=1}^S \beta_{ij} \mathbf{H}_i \mathbf{w}_{ij} \mathbf{v}_{kji}^H + \mathbf{v}_{ks} / \beta_{ks} \right)^2 \\
&= \mathbf{t}_{ks}^H \left( \sum_{i=1}^K \sum_{j=1}^S \mathbf{H}_i \mathbf{v}_{ij} \mathbf{v}_{kji}^H + \mathbf{v}_{ks} / \beta_{ks} \right) \mathbf{t}_{ks}, \quad \forall k, s.
\end{align*}
\]

By collecting the above equalities for all \( k \) and \( s, \{ \beta_{ks} > 0, \forall s \} \) can be determined by

\[
\begin{align*}
(\mathbf{Y} + \hat{\Omega}) \beta^2 &= \left[ \mathbf{v}_{11}^H \mathbf{R}_{11} \mathbf{v}_{11}, \ldots, \mathbf{v}_{11}^H \mathbf{R}_{11} \mathbf{v}_{11}, \ldots, \mathbf{v}_{KK}^H \mathbf{R}_{KK} \mathbf{v}_{KK} \right]^T \\
&= \Theta \hat{\Theta} - \Theta \hat{\Theta}^{-1} (1 + \mathbf{Y} \hat{\Theta}^{-1})^{-1} \mathbf{p}_x \\
&\Rightarrow \beta^2 = (\mathbf{Y} + \hat{\Omega})^{-1} (1 + \mathbf{Y} \hat{\Theta}^{-1})^{-1} \mathbf{p}_x \\
&= \hat{\Omega}^{-1} (1 + \mathbf{Y} \hat{\Omega})^{-1} (1 + \mathbf{Y} \hat{\Theta}^{-1})^{-1} \mathbf{p}_x
\end{align*}
\]

where the third equality follows from (41),

\[
\beta^2 = (\beta_{11}, \ldots, \beta_{1S}, \ldots, \beta_{K1}, \ldots, \beta_{KS})^T.
\]

\[
\hat{\Omega} = \mathbf{diag}(t_{1}^H \mathbf{v}_{11}, \ldots, t_{11}^H \mathbf{v}_{11}, \mathbf{v}_{k1}^H \mathbf{v}_{k1}, \ldots, t_{KK}^H \mathbf{v}_{KK}, \mathbf{v}_{kK}^H \mathbf{v}_{kK}, \ldots).\n\]

Next we show that there exists \( \beta_{ks} \) such that (51) is satisfied. Towards this end, we consider the following discrete-time switched system [16].

\[
\mathbf{x}_n + 1 = F_{\sigma} \mathbf{x}_n \quad \text{for } n = 0, 1, 2, \ldots
\]

\[5\]For \( \mathcal{P}_3 \), \( \{ \mathbf{w}_{ns}^H \mathbf{R}_{ks} \mathbf{w}_{ns}^s > 0, \forall s \} \) is always true.

\[6\]Note that the application of (43) will be clear in the sequel (see (55)).
where $\mathbf{x} \in \mathbb{R}^{n \times 1}$ is a state, $\mathbf{F}_{s_n} \in \mathbb{R}^{m \times n}$ is a switching matrix and $\sigma_n \in \{0, 1, 2, \ldots\}$. According to [16] (Remark 2 of [16]), the above system is marginally stable (convergent) if

$$\max_{\sigma_n} \| \mathbf{F}_{s_n} \|_* = 1 \text{ for } n = 0, 1, \cdots$$

(53)

where $\| . \|_*$ denotes an induced matrix norm.

Let us consider the following iteration

$$\mathbf{x}'_{n+1} = \mathbf{Z}(\mathbf{x}'_n) \mathbf{x}'_n, \text{ for } n = 0, 1, 2, \cdots$$

(54)

Now if we assume $\mathbf{Z}(\mathbf{x}'_n) = \mathbf{F}_{s_n}, \forall n,^7$ we can interpret (54) as a discrete time switched system. Consequently, the above iteration is guaranteed to converge if $\max_{\sigma_n} \| \mathbf{Z}(\mathbf{x}'_n) \|_* = 1$. It is known that $\| . \|_*$ is an induced matrix norm [17]. For any $\mathbf{x}'$, the matrix one norm of $\mathbf{Z}(\mathbf{x}')$ is given by

$$\| \mathbf{Z}(\mathbf{x}') \|_1 = \left\| \mathbf{H} \mathbf{P}^{-1} (\mathbf{I} + \mathbf{Y} \mathbf{P}^{-1})^{-1} \mathbf{P}^{-1} \right\|_1$$

$$\leq \| \mathbf{H} \|_1 \left\| (\mathbf{I} + \mathbf{Y} \mathbf{P}^{-1})^{-1} \right\|_1 \| \mathbf{P}^{-1} \|_1 = \| \mathbf{H} \|_1 \| \mathbf{P} \|_1 \leq 1$$

(55)

Thus, $\max_{\sigma_n} \| \mathbf{Z}(\mathbf{x}'_n) \|_* = 1$ holds true and (54) is guaranteed to converge. As we can see (54) is derived by using (41) and (46). Thus, the solution of (54) also satisfies (41) and (46).

C. Extension of the Current Duality for $\mathcal{P}3$ With a Total BS Power Constraint

In this subsection, we show the extension of the current duality for $\mathcal{P}3$ with a total BS power constraint. For this problem, we set $\Delta_{bs} = \mathbf{I}$, $\forall k, s$ (i.e., like in Section IV-B). Upon doing so, $\mathbf{B}_k^2$ is computed directly from the first equality of (46) (i.e., the bound (55) is not needed). By summing the left and right hand sides of this equality, one can get $\| \mathbf{B}_k^2 \|_F = F_{max}$. This shows that for $\mathcal{P}3$ with a total BS power constraint problem, the total BS power at step 3 of Algorithm I is satisfied.

7Since $\mathbf{Z}(\mathbf{x}'_n)$ is the products of stochastic matrices (see the proof of Theorem 2), $\mathbf{Z}(\mathbf{x}'_n)$ is a bounded matrix for any $\mathbf{x}' > 0$. Thus, the assumption $\mathbf{Z}(\mathbf{x}'_n) = \mathbf{F}_{s_n}, \forall s$ holds true.

Thus, one can apply Algorithm I (with the additional power allocation step) to solve the latter problem by setting $\Delta_{bs} = \mathbf{I}$, $\forall k, s$.

For other total BS power constrained WMSE-based problems, the current duality based algorithm can be applied like in this subsection. The details are omitted for conciseness. Note that for such problem types, the duality algorithm of the current paper has the same complexity as that of [5].

VII. USER-WISE MSE DOWNLINK-INTERFERENCE DUALITY

This section establishes user-wise MSE duality between downlink and interference channels. This duality is established to solve the problems of type $\mathcal{P}4$.

A. User-Wise MSE Transfer (From Downlink to Interference Channel)

To apply this MSE transfer for $\mathcal{P}4$, we set the interference channel precoder, decoder, noise covariance, input covariance and MSE weight matrices as

$$\mathbf{V}_k = \mathbf{P} \mathbf{W}_k, \mathbf{T}_k = \frac{\mathbf{B}_k}{\mathbf{B}_k} \mathbf{H}_k = \mathbf{I}, \Delta_{bs} = \mathbf{B} + \mu_k \mathbf{I} = \mathbf{W}_k$$

(56)

Like in Section VI, substituting (56) into (8), equating $\{\mathbf{R}_k = \mathbf{R}_k \}_{k=1}^K$ and after some straightforward steps, we get the following system of equations

$$\{ \mathbf{Y} + \mathbf{H} \mathbf{B}_k \mathbf{P}_k \} \mathbf{B}_k = \mathbf{P}_k, \Rightarrow \mathbf{B}_k = \mathbf{H} + \mathbf{Y}, \mathbf{B}_k = \mathbf{P}_k$$

(57)

where

$$\mathbf{Y}(k, i) = \left\{ \sum_{l=1, l \neq k}^K \| \mathbf{W}_k^H \mathbf{H}_l^H \mathbf{B}_l \|_F^2 \right\} \text{ for } k = i$$

$$\mathbf{H}(k, i) = \left\{ \sum_{l=1, l \neq k}^K \| \mathbf{W}_k^H \mathbf{H}_l^H \mathbf{B}_l \|_F^2 \right\} \text{ for } k \neq i$$

$$\hat{\mathbf{B}}_k = \text{diag}(\hat{\theta}_1, \cdots, \hat{\theta}_K), \hat{\mathbf{B}}_k = \left[ \hat{\psi}_1, \cdots, \hat{\psi}_K \right]^T, \hat{\mathbf{P}}_k = \left[ \hat{\mathbf{P}}_1, \cdots, \hat{\mathbf{P}}_K \right]$$

(58)

$$\mathbf{H}_k = \text{diag}(\hat{\psi}_1, \cdots, \hat{\psi}_K), \mathbf{W}_k = \mathbf{H}_k$$

(59)

B. User-Wise MSE Transfer (From Interference to Downlink Channel)

For a given user MSE in the interference channel with $\mathbf{V}_k = \mathbf{H}_k \mathbf{W}_k$, we can achieve the same MSE in the downlink channel by using nonzero scaling factors $\{\mathbf{H}_k \}_{k=1}^K$ that satisfy

$$\mathbf{B}_k = \hat{\mathbf{B}}_k \mathbf{T}_k, \mathbf{W}_k = \frac{\mathbf{B}_k}{\mathbf{B}_k} \mathbf{H}_k$$

(60)

Here we also use the notations $\mathbf{B}$ and $\mathbf{W}$ to differentiate with the precoder and decoder matrices used in Section VII-A. By substituting (60) into $\xi_k^{DL}$ (with $\mathbf{B} = \mathbf{B}, \mathbf{W} = \mathbf{W}$) and then
equating the resulting user-wise MSE with that of the interference channel (7) and after some steps, \( \{ \tilde{\theta}_k \}_{k=1}^K \) are determined as

\[
(1 + \tilde{Y} \Omega) \tilde{\beta} R_z^{-1} \tilde{\beta} = \left[ \text{tr}\{V_1^H R_{n_1} V_1\}, \ldots, \text{tr}\{V_K^H R_{n_K} V_K\}\right]^T \Rightarrow \tilde{\beta} = \Omega^{-1}(1 + \tilde{Y} \Omega)^{-1}(1 + \tilde{Y} \tilde{\theta})^{-1} \tilde{\beta} \tilde{\beta}',
\]

where the second equality follows from (57), \( \tilde{\beta} = [\tilde{\beta}_1, \ldots, \tilde{\beta}_K]^T \), \( \Omega' = \text{diag}\{\text{tr}\{T_1^H \Psi T_1\}, \ldots, \text{tr}\{T_K^H \Psi T_K\}\} \), \( \Omega = \text{diag}\{\mu_1 \text{tr}\{T_1^H T_1\}, \ldots, \mu_K \text{tr}\{T_K^H T_K\}\} \), and \( \Omega' = \Omega + \Delta \).

By applying Theorem 2, it can be shown that \( \{ \tilde{\beta}_k \}_{k=1}^K \) are strictly positive for \( \psi_n > 0 \) and \( \{ \tilde{\beta}_k \}_{k=1}^K \). In Section IV-B, the power constraint of the \( n \)th BS antenna and \( k \)th user can thus be expressed as

\[
\tilde{x}_n^\text{opt} = \tilde{\Omega}^{-1}(1 + \tilde{Y} \tilde{\Omega})^{-1}(1 + \tilde{Y} \tilde{\theta})^{-1} \tilde{\beta} \tilde{\beta}',
\]

where \( \tilde{P} = \text{bikdiag}\{\tilde{p}_1, \ldots, \tilde{p}_K\} \) and \( \tilde{\Omega} = [\tilde{\Omega}_1, \ldots, \tilde{\Omega}_K]^T \). Hence, we have

\[
\tilde{x}_n^\text{opt} = \tilde{\Omega}^{-1}(1 + \tilde{Y} \tilde{\Omega})^{-1}(1 + \tilde{Y} \tilde{\theta})^{-1} \tilde{\beta} \tilde{\beta}',
\]

and

\[
\tilde{x}_n^\text{opt} = \tilde{\Omega}^{-1}(1 + \tilde{Y} \tilde{\Omega})^{-1}(1 + \tilde{Y} \tilde{\theta})^{-1} \tilde{\beta} \tilde{\beta}'.
\]

In [8], for fixed transmit and receive filters, the power allocation parts of total BS power constrained MSE-based problems have been formulated as GPs by employing the approach and system model of [1] under the assumption that all symbols are strictly active. For this assumption, in [8], we show that the system model of [1] is appropriate to solve any kind of total BS power constrained MSE-based problems using duality approach (alternating optimization). This motivates us to utilize the system model of [1] in the downlink channel only and then include the power allocation step (i.e., GP) into Algorithm I. Towards this end, we decompose the precoders and decoders of the downlink channel as

\[
B_k = G_k \hat{P}_k^H, W_k = U_k \alpha_k P_k^{-1/2}, \forall k
\]

where \( P_k = \text{diag}\{p_{k1}, \ldots, p_{ks}\} \in \mathbb{R}^{S_k \times S_k}, G_k = [g_{k1}, \ldots, g_{ks}] \in \mathbb{C}^{N \times S_k}, U_k = [u_{k1}, \ldots, u_{ks}] \in \mathbb{C}^{M_k \times S_k} \) and \( \alpha_k = \text{diag}\{\alpha_{k1}, \ldots, \alpha_{ks}\} \in \mathbb{R}^{S_k \times S_k} \) are the transmit power, unity norm transmit filter, unity norm receive filter and receiver scaling factor matrices of the \( k \)th user, respectively, i.e., \( \{ g_{ks} \}_{k=1}^K, u_{ks} \in 1 \}, \forall S_k \). By employing (66) and stacking \( \xi = [\xi_{DL1}^1, \ldots, \xi_{DL2}^N]^T = [\xi_{DL1}^1, \ldots, \xi_{DL2}^{N-1}]^T \), the th downlink symbol MSE can be expressed as (see [1] and [10] for more details about (64) and the above descriptions)

\[
\xi_{DL}^l = p_l^{-1} \left[ (D + \alpha^2 \Phi^T) p_l \right]_{l} + p_l^{-1} \alpha^2 u_{l1}^H R_{u1} u_{l1}
\]

where

\[
\Phi_{l, l'} = \begin{cases} \|g_{l1}^H H_{l1}\|^2 & \text{for } l \neq l' \\ 0 & \text{for } l = l' \end{cases}
\]

and

\[
D_{l, l'} = \alpha^2 R_{l1} H_{l1} H_{l1}^H R_{l1}^H u_{l1}^H H_{l1}^H \|g_{l1}\|^2 + 1.
\]

VIII. GENERALIZED AND IMPROVED VERSION OF Algorithm I

From the discussions of Sections IV–VII, one can understand that each iteration of Algorithm I gives a non increasing sequence of symbol (user) WMSE/WSMSE. As can be seen from Section III, the objective of \( P_1 \) (\( P_2 \)) is just to minimize the total WSME of all symbols (users), whereas the objective of \( \mathcal{P}_3 \) (\( \mathcal{P}_4 \)) is to simultaneously minimize and balance the WMSE of all symbols (users). Thus, Algorithm I is appropriate to solve \( \mathcal{P}_1 \) (\( \mathcal{P}_2 \)) of the current paper. For \( \mathcal{P}_3 \) (\( \mathcal{P}_4 \)), although each iteration of Algorithm I is able to provide a non increasing sequence of symbol (user) WMSE (i.e., minimizes the maximum WMSE of all symbols (users)), each iteration of this algorithm is not able to guarantee balanced WMSEs of all symbols (users). On the other hand, for an MSE constrained total BS power minimization problem (for example \( \mathcal{P}_7 \) in Section IX), an iterative algorithm that can provide a non increasing sequence of total BS power is required. This shows that Algorithm I also can not solve the latter problem. In the following we address the drawbacks of Algorithm I just by including a power allocation step into Algorithm I as explained below.

\[\text{Note that this assumption is not always true for all MSE-based problems. However, as mentioned in [1], in practice replacing zero powers by a small value will not affect the overall optimization. Due to this reason, we replace zero powers by 10^{-6} in the simulation section.}\]
Algorithm II

Initialization: Like in Algorithm I.

Repeat: Interference channel

1) For $\mathcal{P}1$ and $\mathcal{P}2$, set $V = W$, $T = B$ (i.e., $\beta = \tilde{\beta} = 1$), then compute $\{\psi_n, \mu_k, \forall k, s, n\}$ and $\{\psi_n, \mu_k, \forall k, n\}$ using (25) and (38), respectively. For $\mathcal{P}3$ and $\mathcal{P}4$, first compute $\{\psi_n, \mu_k, \forall k, s, n\}$ and $\{\psi_n, \mu_k, \forall k, n\}$ using (54) and (63), respectively, then transfer each symbol and user MSE from downlink to interference channels by (39) and (56), respectively.

2) Update the MMSE receivers of the interference channel for $\mathcal{P}1$, $\mathcal{P}2$, $\mathcal{P}3$, and $\mathcal{P}4$ using (19), (32), (44), and (59), respectively.

Downlink channel

3) Transfer the MSE (weighted sum, user or symbol MSE) from interference to downlink channel using (20), (33), (45), and (60) for $\mathcal{P}1$, $\mathcal{P}2$, $\mathcal{P}3$, and $\mathcal{P}4$, respectively.

4) For each of the problems $\mathcal{P}1 - \mathcal{P}4$, decompose the precoder and decoder matrices of each user as in (64). Then, formulate and solve the GP power allocation part. For example, the power allocation part of $\mathcal{P}1$ can be expressed in GP form as (68).

5) For each of the problems $\mathcal{P}1 - \mathcal{P}4$, by keeping $\{P_k\}_{k=1}^K$ constant, update the receive filters $\{U_k\}_{k=1}^K$ and scaling factors $\{\alpha_k\}_{k=1}^K$ by applying downlink MMSE receiver approach i.e., $\{U_k\} \alpha_k = (H_k^H G P G H_k + R_{nk})^{-1} H_k \mathbf{G}_k P_k K_{k=1}$.

Note that in these expressions, $\{\alpha_k\}_{k=1}^K$ are chosen such that each column of $\{U_k\}_{k=1}^K$ has unity norm. Then, compute $\{B_k, W_k\}_{k=1}^K$ by (64).

Until convergence.

Convergence: It can be shown that at each iteration of this algorithm, the objective function of each of the problems $\mathcal{P}1 - \mathcal{P}4$ is non-increasing [4], [7], [19]. Thus, the above iterative algorithm is convergent. However, since $\mathcal{P}1 - \mathcal{P}4$ are non-convex, this iterative algorithm is not guaranteed to converge to the global optimum.

In this algorithm, we stop iteration (i.e., our convergence condition) when the difference between the objective functions in two consecutive iterations is smaller than some small value $\varepsilon$ (we use $\varepsilon = 10^{-6}$ for the simulation).

Computational complexity: As can be seen from this algorithm, when we increase the number of users and/or (BS and/or MS antennas), the number and size of optimization variables increase. Because of this, the computational complexity of Algorithm II increases as $K$ and/or $N$ and/or $M$ increases. However, studying the complexity of this algorithm as a function of $K$, $N$, and $M$ needs effort and time. And such a task is beyond the scope of this work and is an open research topic.

The power allocation step of Algorithm II has thus the following benefits: (1) For BS power constrained WSMSE minimization problems, this step improves the convergence speed of Algorithm II compared to that of Algorithm I (for example in $\mathcal{P}1 - \mathcal{P}2$). The degree of improvement depends on different parameters (for example $H_k$, $\Delta_k$, $\forall k$, $s$ etc). Thus, the theoretical comparison of these two algorithms in terms of convergence speed requires time and effort. And this task is beyond the scope of this work and it is an open research topic. (2) For symbol-wise (user-wise) WMSE balancing problems, this step helps to balance the WMSE of all symbols (users) (for example in $\mathcal{P}3 - \mathcal{P}4$). (3) For MSE constrained total BS power minimization problems, this step ensures a non increasing total BS power at each iteration of Algorithm II.

IX. APPLICATION OF THE PROPOSED DUALITY BASED ALGORITHM FOR OTHER PROBLEMS

A. MSE Based Problem With Entry-Wise Power Constraint

The symbol-wise WSMSE minimization constrained with entry wise power i.e., $b_{k=1}^K b_{k=1}^N \leq \tilde{p}_{k=1}^N, \forall k, s, n$ problem is formulated as

$$\mathcal{P}5: \min_{\{B_k, W_k\}_{k=1}^K} \sum_{k=1}^K \sum_{s=1}^S \eta_k \xi_k \xi_k^D, \text{s.t.} \sum_{k=1}^K \sum_{s=1}^S b_{k=1}^N b_{k=1}^N \leq \tilde{p}_{k=1}^N, (69)$$

It can be shown that this problem can be solved by Algorithm II with $\{\Delta_k = \text{diag}[^{\varepsilon}_{k=1}, \cdots, \varepsilon_{k=N}] \}, \forall s \}^K_{k=1}$.

B. Weighted Sum Rate Optimization Constrained With Per Antenna and Symbol Power Problem

By employing the approach of [11] (see (16) of [11]), one can equivalently express the weighted sum rate maximization constrained with per antenna and symbol power problem as

$$\mathcal{P}6: \min_{\{\tau_k, \tilde{v}_{k}, b_{k}, b_{k}, \forall s \}^K_{k=1}} \sum_{k=1}^K \sum_{s=1}^S \tau_k \tilde{v}_{k} \xi_k \xi_k^D, \text{s.t.} b_{k=1}^H (H_{k=1}^H \tau_k, \tilde{v}_{k}) \leq \tilde{p}_{k=1}^N, \forall s \}^K_{k=1} \leq \tilde{p}_{k=1}^N$$

where $0 < \tilde{w}_{k=1} < 1, \forall s \}^K_{k=1}$ are the rate weighting factors for all symbols, $\tilde{p}_{k=1}^N = \sum_{k=1}^K \sum_{s=1}^S \hat{p}_{k=1}^N 1/\tilde{w}_{k=1}$, and $\hat{p}_{k=1} = \bar{p}_{k=1}^N \mu_{k=1}^N$ for $\mu_{k=1}^N = (1 - \bar{w}_{k=1}^N)$ for fixed $\{\tau_k, \tilde{v}_{k}, \forall s \}^K_{k=1}$ constant, $\{b_{k=1}^N, \tilde{v}_{k=1}^N, \forall s \}^K_{k=1}$ can be optimized by applying the MSE duality discussed in Section IV. Moreover, $\{\tau_k, \tilde{v}_{k}, \forall s \}^K_{k=1}$

This is at the expense of additional computation. However, as mentioned in [11] (see Appendix A of [1]), a small desktop computer can solve a GP of 100 variables and 10000 constraints by standard interior point method under a minute. Thus, we believe that the complexity of Algorithm I and Algorithm II are almost the same.
where $\text{SINR}_{ks}(R_{ks})$ is the SINR (rate) of the $k$th user $s$th symbol, and we use the fact that $R_{ks} = \log_2[1 + \text{SINR}_{ks}]$ and $\xi_{ks} = (1 + \text{SINR}_{ks})^{-1}$. It is clearly seen that the application of Algorithm II is not limited to the problems of this paper.

Note that under imperfect channel state information (CSI) condition, the stochastic robust design versions of $\mathcal{P}1$ and $\mathcal{P}3$ can be solved like in [10]. However, to the best of our knowledge, the relationship between rate (SINR) and MSE is not known when the CSI is imperfect [7]. Hence, solving the rate (SINR)-based robust design problems (for example, robust versions of $\mathcal{P}6 - \mathcal{P}8$) by our duality approach is an open problem.

X. SIMULATION RESULTS

In this section, we present simulation results for $\mathcal{P}1 - \mathcal{P}4$. All of our simulation results are averaged over 100 randomly chosen channel realizations. We set $K = 2$, $N = 4$ and $\{M_k = S_k = 2, \eta_{ks} = \rho_{ks} = \eta_k = \rho_k = 1, \forall s\}_{k=1}^K$. It is assumed that $\mathbf{R}_{n1} = \sigma_1^2 \mathbf{I}_{M_1}$, $\mathbf{R}_{n2} = \sigma_2^2 \mathbf{I}_{M_2}$, and $\sigma_2^2 = 2\sigma_1^2$. The maximum power of each BS antenna is set to $\{\bar{p}_k = 2.5 \text{ mW}\}_{k=1}^N$. And the maximum power allocated to each symbol and user are set to $\{\bar{p}_{ks} = 2.5 \text{ mW}, \forall s\}_{k=1}^K$ and $\{\bar{p}_k = 5 \text{ mW}\}_{k=1}^K$, respectively. For better exposition, we define the Signal-to-noise ratio (SNR) as $\frac{P_{\text{max}}}{K\sigma^2_{\text{w}}}$ and it is controlled by varying $\sigma^2_{\text{w}}$, where $P_{\text{max}} = 10$ mW is the total maximum BS power and $\sigma^2_{\text{w}} = (\sigma_1^2 + \sigma_2^2)/2$. We also compare Algorithm II and the algorithm in [2].

Note that the algorithm in [2] is designed for coordinated BS systems scenario. And, the iterative algorithm of [2] is based on the per BS power constraint. However, according to [2] and [21], $B$ coordinated BS systems each with $Z$ antennas can be treated as one multiuser MIMO system with $BZ$ antennas. Thus, when $Z = 1$, the considered problem has exactly the same structure as that of [2]. To the best of our knowledge, there is no other general linear algorithm that can solve the problem types $\mathcal{P}1 - \mathcal{P}4$. On the other hand, in all problems, since there are more than one power constraints (i.e., per antenna and symbol (user) powers), all power constraints may not be active at the optimal solution. Due to these reasons, we compare Algorithm II and the algorithm in [2] both in terms of the achieved MSE (i.e., minimized MSE) and total utilized BS power at the achieved MSE.

A. Simulation Results for Problems $\mathcal{P}1$ and $\mathcal{P}2$

In this subsection, we compare the performance of our proposed algorithm with that of [2]. As can be seen from Fig. 2, the proposed algorithm and the algorithm in [2] achieve the same symbol-wise and user-wise WSMSEs. Next, we plot the total utilized powers at the BS to achieve these WSMSEs which is shown in Fig. 3. From these two figures, one can see that to achieve the same WSMSE, the proposed duality based iterative algorithm requires less total BS power than that of [2]. This scenario fits to that of [10] and [19] where the sum MSE minimization constrained with a per BS antenna power problem has been examined by duality approach.

$10$As will be clear in the sequel, the proposed algorithm and the algorithm of [2] may not utilize the entire 10 mW for all noise levels. Thus, the aforementioned SNR can be considered as the maximum SNR of the transmitted signal.
Fig. 3. Comparison of the proposed algorithm (Algorithm II) and the algorithm of [2] in terms of total BS power for: [top] \( \mathcal{P}_1 \), [bottom] \( \mathcal{P}_2 \). For this figure we compute \( \sigma^2_{\text{av}} \) (dB) as \( \sigma^2_{\text{av}} \) (dB) = 10 \log_{10} \frac{1}{W}.

B. Simulation Results for Problems \( \mathcal{P}_3 - \mathcal{P}_4 \)

Like in the above subsection, here we compare the performance of our proposed algorithm with that of [2]. For these problems, we also observe from Fig. 4 that the proposed algorithm and the algorithm in [2] achieve the same maximum symbol and user MSEs. And from Fig. 5 the proposed duality based algorithms utilize less total BS power compared to that of [2]. For all of our problems, we observe that to achieve the same MSE, the proposed duality based iterative algorithm utilizes less total BS power compared to the algorithm of [2]. This scenario has also been observed for other MSE and rate based problems in [10], [19], [20].

Note that the problems of [2] are examined directly in the downlink channel. According to [5], [13], in general, duality approach of solving downlink transceiver design problems has easier to handle mathematical structure (lower complexity) compared to that of the direct approach in [2]. For this reason, we believe that the computational complexity of the proposed duality algorithm is not higher than that of [2].

C. Convergence Speed of Algorithm II

As can be seen from Section VIII, the overall computational complexity of Algorithm II depends on the number of iterations to achieve convergence. In general, the number of iterations to achieve convergence may not be the same for all problems. On the other hand, for each problem, getting the exact number of iterations to achieve convergence analytically is very difficult. Due to these reasons, we provide numerical simulations to demonstrate the convergence speed of Algorithm II for \( \mathcal{P}_1 \). As can be seen from Fig. 6, the proposed algorithm converges within few iterations in low, medium and high SNR regions.

XI. CONCLUSIONS

In this paper, we examine different transceiver design problems for multiuser MIMO systems under generalized linear power constraints. The problems are solved for the practically relevant scenario where the noise vector of each MS is a ZMCSGC random variable with arbitrary covariance matrix. For all of our problems, we propose new downlink-interference duality based iterative solutions. The current duality are established by formulating the noise covariance matrices of the dual interference channels as fixed point functions and marginally stable (convergent) discrete-time-switched systems. We show that the proposed duality based iterative algorithms can be extended straightforwardly to solve several practically relevant linear transceiver design problems. We also show that our new MSE downlink-interference duality unify all existing MSE duality. Our simulation results demonstrate that the proposed duality based algorithms utilize less total BS power than that of existing algorithms.
where \( \mathbf{A} = \mathbf{D} \mathbf{A} \mathbf{D}^{-1} \). Since \( (\mathbf{I} - \mathbf{A}) \) is strictly diagonally dominant matrix, \( (\mathbf{I} - \mathbf{A})^{-1} \) exists [17] (page 349). Furthermore, if \( \rho(\mathbf{A}) < 1 \), \( (\mathbf{I} - \mathbf{A})^{-1} \) can be expressed as

\[
(\mathbf{I} - \mathbf{A})^{-1} = \sum_{k=0}^{\infty} \mathbf{A}^k.
\] (71)

It follows

\[
\mathbf{A}^{-1} = \mathbf{D}^{-1} \mathbf{I} - \mathbf{A}^{-1} = \mathbf{D}^{-1} \sum_{k=0}^{\infty} \mathbf{A}^k \geq 0
\] (72)

From this equation we can see that if \( \rho(\mathbf{A}) < 1 \), the nonnegativity of \( \mathbf{A}^{-1} \) can be ensured. Next we show that \( \rho(\mathbf{A}) \) is indeed less than 1. For any \( n \times n \) matrix \( \mathbf{X} \), we have [17] (pages 294 and 297 of [17])

\[
\rho(\mathbf{X}) \leq \| \mathbf{X} \|_1 \| \mathbf{X} \|_1^\Delta \max_{1 \leq j \leq n} \sum_{i=1}^{n} |x_{ij}|
\] (73)

where \( \| \cdot \| \) is any matrix norm and \( \| \cdot \|_1 \) is a matrix one norm. By using (73), we get the following bound [17]

\[
\rho(\mathbf{A}) \leq \| \mathbf{A} \|_1 < 1.
\] (74)

Since \( \mathbf{A}^{-1} \) has nonnegative elements, \( \mathbf{A} \) is also an M-matrix [22]. By defining \( \mathbf{S} \triangleq \mathbf{A}^{-1} \) and \( \mathbf{e} \triangleq 1^{n \times 1} \), we get

\[
e^T \mathbf{A} - e^T \Rightarrow e^T - e^T \mathbf{S} = \left[ \sum_{j=1}^{n} \mathbf{S}_{j1}, \ldots, \sum_{j=1}^{n} \mathbf{S}_{jn} \right]
\] (75)

where the third equality follows from the fact that \( \mathbf{S} \) is a nonnegative matrix.

References

[1] S. Shi, M. Schubert, and H. Boche, “Rate optimization for multiuser MIMO systems with linear processing,” IEEE Trans. Signal Process., vol. 56, no. 8, pp. 4020–4030, Aug. 2008.
[2] S. Shi, M. Schubert, N. Vucic, and H. Boche, “MMSE optimization with per-base-station power constraints for network MIMO systems,” in Proc. IEEE Int. Conf. Commun. (ICC), Beijing, China, May 19–23, 2008, pp. 4106–4110.
[3] P. Ubaidulla and A. Chockalingam, “Robust transceiver design for multiuser MIMO downlink,” in Proc. IEEE Global Telecommun. Conf. (GLOBECOM), Nov. 30–Dec. 4, 2008, pp. 1–5.
[4] 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, no. 11, pp. 5436–5446, Nov. 2007.
[5] R. Hunger, M. Joham, and W. Utschick, “On the MSE-duality of the broadcast channel and the multiple access channel,” IEEE Trans. Signal Process., vol. 57, no. 2, pp. 698–713, Feb. 2009.
[6] M. Schubert, S. Shi, E. A. Jorswieck, and H. Boche, “Downlink sum-MSE transceiver optimization for linear multi-user MIMO systems,” in Conf. Rec. 39th Asilomar Conf. Signals, Syst., Comput., Oct. 28–Nov. 1 2005, pp. 1424–1428.
[7] T. E. Bogale, B. K. Chalise, and L. Vandendorpe, “Robust transceiver optimization for downlink multiuser MIMO systems,” IEEE Trans. Signal Process., vol. 59, no. 1, pp. 446–453, Jan. 2011.
[8] T. E. Bogale and L. Vandendorpe, “MSE uplink-downlink duality of MIMO systems with arbitrary noise covariance matrices,” presented at the 45th Annu. Conf. Inf. Sci. Syst. (CISS), Baltimore, MD, USA, Mar. 23–25, 2011.
V. Berinde, “General constructive fixed point theorems for Ciric-type almost contractions in metric spaces,” *Carpathian, J. Math.*, vol. 24, no. 2, pp. 10–19, 2008.

V. Berinde, “Approximation fixed points of weak contractions using the Picard iteration,” *Nonlinear Anal. Forum*, vol. 9, no. 1, pp. 43–53, 2004.

Z. Sun, “A note on marginal stability of switched systems,” *IEEE Trans. Autom. Control*, vol. 53, no. 2, pp. 625–631, 2008.

R. H. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

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

T. E. Bogale and L. Vandendorpe, “Sum MSE optimization for downlink multiuser MIMO systems with per antenna power constraint: Downlink-uplink duality approach,” presented at the 22nd IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC), Toronto, ON, Canada, Sep. 11–14, 2011.

T. E. Bogale and L. Vandendorpe, “Weighted sum rate optimization for downlink multiuser MIMO systems with arbitrary power constraint: Generalized duality approach,” *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1862–1875, 2012.

T. E. Bogale and L. Vandendorpe, “Weighted sum rate optimization for downlink multiuser MIMO coordinated base station systems: Centralized and distributed algorithms,” *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1876–1889, 2012.

H. Sampath, P. Stoica, and A. Paulraj, “Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion,” *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

S. Shi, M. Schubert, and H. Boche, “Downlink MMSE transceiver optimization for multiuser MIMO systems: MMSE balancing,” *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3702–3712, Aug. 2008.

T. E. Bogale and L. Vandendorpe, “Weighted sum MSE optimization for multiuser MIMO systems: MMSE balancing,” *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 3702–3712, Aug. 2012.

T. E. Bogale and L. Vandendorpe, “Sum MSE optimization for downlink multiuser MIMO systems with per antenna power constraint: Downlink-uplink duality approach,” presented at the 22nd IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC), Toronto, ON, Canada, Sep. 11–14, 2011.

T. E. Bogale and L. Vandendorpe, “Weighted sum rate optimization for downlink multiuser MIMO systems with arbitrary power constraint: Generalized duality approach,” presented at the IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Kyoto, Japan, Mar. 25–30, 2012.

T. Tamaki, K. Seong, and J. M. Cioffi, “Downlink MIMO systems using cooperation among base stations in a slow fading channel,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, Glasgow, U.K., Jun. 24–28, 2007, pp. 4728–4733.

G. Poole and T. Bourillon, “A survey on M-matrices,” *SIAM Rev.*, vol. 16, no. 4, pp. 419–427, 1974.

**Tadilo Endeshaw Bogale** (S’09–M’13) was born in Gondar, Ethiopia. He received his B.Sc and M.Sc degree in Electrical Engineering from Jimma University, Jimma, Ethiopia and Karlstad University, Karlstad, Sweden in 2004 and 2008, respectively. From 2004–2007, he was working in Ethiopian Telecommunications Corporation (ETC) in mobile project department. Since 2009 he has been working towards his Ph.D degree and as an assistant researcher at the ICTTEAM institute, University Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium. His research interests include robust (non-robust) transceiver design for multiuser MIMO systems, centralized and distributed algorithms, and convex optimization techniques for multiuser systems.

**Luc Vandendorpe** (M’93–SM’99–F’06) was born in Mouscron, Belgium, in 1962. He received the Electrical Engineering degree (summa cum laude) and the Ph.D degree from the Universit Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 1985 and 1991, respectively. Since 1985, he has been with the Communications and Remote Sensing Laboratory of UCL, where he first worked in the field of bit rate reduction techniques for video coding. In 1992, he was a Visiting Scientist and Research Fellow at the Telecommunications and Traffic Control Systems Group of the Delft Technical University, The Netherlands, where he worked on spread spectrum techniques for personal communications systems. From October 1992 to August 1997, he was Senior Research Associate of the Belgian NSF at UCL, and invited Assistant Professor. He is currently a Professor and head of the Institute for Information and Communication Technologies, Electronics and Applied Mathematics. His current interest is in digital communication systems and more precisely resource allocation for OFDM(A)-based multicell systems, MIMO and distributed MIMO, sensor networks, turbo-based communications systems, physical layer security and UWB based positioning.

Dr. Vandendorpe was corecipient of the 1990 Biennal Alcatel-Bell Award from the Belgian NSF for a contribution in the field of image coding. In 2000, he was corecipient (with J. Louveaux and F. Deryck) of the Biennal Siemens Award from the Belgian NSF for a contribution about filter-bank-based multicarrier transmission. In 2004, he was co-winner (with J. Czyz) of the Face Authentication Competition, FAC 2004. He is or has been TPC member for numerous IEEE conferences (VTc, Globecom, SPAWC, ICC, PIMRC, WCNC) and for the Turbo Symposium. He was Co-Technical Chair (with P. Duhamel) for the IEEE ICASSP 2006. He was an Editor for Synchronization and Equalization of the IEEE TRANSACTIONS ON COMMUNICATIONS between 2000 and 2002, Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS between 2003 and 2005, and Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING between 2004 and 2006. He was Chair of the IEEE Benelux joint chapter on Communications and Vehicular Technology between 1999 and 2003. He was an elected member of the Signal Processing for Communications committee between 2000 and 2005, and an elected member of the Sensor Array and Multichannel Signal Processing committee of the Signal Processing Society between 2006 and 2008. Currently, he is an elected member of the Signal Processing for Communications committee. He is the Editor-in-Chief for the EURASIP Journal on Wireless Communications and Networking. L. Vandendorpe is a Fellow of the IEEE.