Weighted Sum MSE Minimization under Per-BS Power Constraint for Network MIMO Systems

Haewook Park, Seok-Hwan Park, Han-Bae Kong, and Inkyu Lee, *Senior Member, IEEE*

***Abstract*—We study joint processing (JP) for network MIMO systems where base stations exchange the user's message and channel state information under per-BS power constraint. In this letter, we propose a weighted sum mean square error (WS-MSE) minimization algorithm for the JP systems by considering the channel gain as the weight factor in the MSE metric. To efficiently solve the formulated WS-MSE problem, an alternating optimization method which iteratively finds a local optimal solution is employed in our algorithm. The simulation results confirm that the proposed algorithm provides the sum rate performance close to the near-optimal gradient ascent approach and outperforms conventional schemes. In addition, we also propose a modified WS-MSE design which is robust to channel mismatch caused by channel estimation and feedback errors.**

***Index Terms*—Network MIMO, weighted sum mean square error (MSE) minimization.**

I. Introduction

COOPERATIVE signal processing among several base stations (BS), called network multiple input multiple output (MIMO), was proposed to overcome the inter-cell interference and to achieve high spectral efficiency [1]. Depending on the BS cooperation level, the network MIMO can be classified into two categories [2]. One is coordinated beamforming where BSs design their transmission strategies such as power control and beamforming by exchanging user's channel state information (CSI) only. The other one is joint processing (JP) where BSs share both user's message and CSI [3]–[5]. This letter focuses on the latter case which provides the improved performance.

Conventional single-cell downlink systems usually impose total power constraint [6]–[8]. On the other hand, the JP systems, where the cooperating BSs act as a single distributed super-BS, should be designed under per-BS power constraint, since each BS has its own power amplifier. Recently, several iterative precoder designs for the JP systems with the per-BS power constraint have been proposed [3]–[5]. The author in [3] introduced a block diagonalization (BD) precoder, and in [4], the transceiver optimization with the criterion of minimizing the sum mean square error (S-MSE) was presented. Since the maximization of the sum rate is not a convex problem in general, the optimization for the first order Taylor expansion of a sum rate expression is proposed in [5].

Manuscript received November 13, 2011. The associate editor coordinating the review of this letter and approving it for publication was W. Zhang.

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No.2010-0017909).

H. Park, H. Kong, and I. Lee are with the School of Electrical Eng., Korea University, Seoul, Korea (e-mail: {jetaime01, kongbae, inkyu}@korea.ac.kr).

S. Park is with the Agency for Defense Development, Daejeon, Korea (e-mail: psk1501@korea.ac.kr).

Digital Object Identifier 10.1109/LCOMM.2012.010512.112300

In this letter, we propose a weighted sum MSE (WS-MSE) minimization algorithm for the JP systems. The idea of minimizing WS-MSE was originally developed for single-cell downlink systems under total power constraint [7] with a goal of regularizing a zero-forcing (ZF) scheme. We extend this concept to the JP systems with per-BS power constraint. For an efficient regularization process on transceiver designs, the weight terms in the WS-MSE metric need to be computed from the effective channel gains of the optimal ZF scheme. However, since the optimal ZF scheme identifies the solution with an iterative process due to the per-BS power constraints [3], we instead employ the channel gain obtained from the generalized minimum mean squared error channel inversion (GMI) solution [8] as a weight matrix in the WS-MSE metric for simple computation. Then, the proposed algorithm based on the alternating optimization which iteratively finds a local optimal solution efficiently minimizes the formulated WS-MSE metric. It is shown from simulation results that the proposed scheme provides the sum rate performance close to the near-optimal gradient approach and outperforms conventional BD and S-MSE schemes. In addition, a modified WS-MSE transceiver design is also provided in the presence of channel mismatch caused by channel estimation and feedback errors.

The following notations are used throughout the letter. We employ uppercase boldface letters for matrices, lowercase boldface for vectors and normal letters for scalar quantities. The superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ and $|\cdot|$ stand for conjugate, transpose, conjugate transpose and determinant, respectively. $\mathbf{I}_d$ indicates an identity matrix of size d, $E[\cdot]$ accounts for expectation and $\|\cdot\|$ denotes the Euclidean 2-norm of a vector. In addition, $\text{tr}(\mathbf{A})$ represents trace of a matrix $\mathbf{A}$.

II. System Model

We consider a JP system where B BSs with M antennas cooperatively serve total K users with N receive antennas by sharing both data and CSI. We refer to this configuration as the $(M \times B) \times (N \times K)$ network MIMO system. Under the frequency-flat fading model, the received signal vector at user k $(k = 1, \cdots, K)$ can be expressed as

$$\mathbf{y}_k = \sum_{i=1}^{B} \mathbf{H}_k^{(i)} \mathbf{T}_k^{(i)} \mathbf{s}_k + \sum_{i=1}^{B} \mathbf{H}_k^{(i)} \sum_{m \neq k, m=1}^{K} \mathbf{T}_m^{(i)} \mathbf{s}_m + \mathbf{n}_k \qquad (1)$$

where $\mathbf{H}_k^{(i)} \in \mathbb{C}^{N \times M}$ is the channel coefficient matrix from BS i to user k, $\mathbf{T}_k^{(i)} \in \mathbb{C}^{M \times N}$ stands for the precoding matrix for user k from BS i, $\mathbf{s}_k \in \mathbb{C}^N$ indicates the data symbol vector for user k, and $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_N)$ represents the additive Gaussian noise vector at user k. It is assumed that elements of the channel matrices are sampled from independent identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, so that the probability of event that the channel is rank-deficient converges to zero.

This system can be modeled as a conventional multiuser downlink system where a BS is equipped with MB antennas supporting K users with N antennas each. The only difference is that the network MIMO system imposes per-BS power constraint as

$$\sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{T}_k^{(i)}\mathbf{T}_k^{(i)H}\right) \leq P_{\max} \quad \text{for} \quad i = 1, \cdots, B \tag{2}$$

where $P_{\max}$ is the maximum transmit power allowed for each BS. Then, the achievable sum rate R_Σ can be given as

$$R_\Sigma = \sum_{k=1}^{K} \log_2 \frac{\left|\sum_{m=1}^{K} \mathbf{\Omega}_{km}\mathbf{\Omega}_{km}^H + \sigma_n^2 \mathbf{I}_N\right|}{\left|\sum_{m\neq k} \mathbf{\Omega}_{km}\mathbf{\Omega}_{km}^H + \sigma_n^2 \mathbf{I}_N\right|} \tag{3}$$

where we define $\mathbf{\Omega}_{km} \triangleq \sum_{i=1}^{B} \mathbf{H}_k^{(i)}\mathbf{T}_m^{(i)}$.

III. Weighted Sum MSE Minimization Algorithm

In this section, we propose the WS-MSE minimization algorithm with the per-BS power constraint. At user k, the receive filter $\mathbf{R}_k \in \mathbb{C}^{N\times N}$ is multiplied to $\mathbf{y}_k$ to obtain $\hat{\mathbf{s}}_k = \mathbf{R}_k\mathbf{y}_k$ as

$$\hat{\mathbf{s}}_k = \mathbf{R}_k\sum_{i=1}^{B} \mathbf{H}_k^{(i)}\mathbf{T}_k^{(i)}\mathbf{s}_k + \mathbf{R}_k\sum_{i=1}^{B} \mathbf{H}_k^{(i)}\sum_{m\neq k, m=1}^{K} \mathbf{T}_m^{(i)}\mathbf{s}_m + \mathbf{R}_k\mathbf{n}_k. \tag{4}$$

Consider that ZF based transceiver design such as BD [3] is employed to decompose the desired block channel into parallel subchannels and cancel out the inter-user interference terms, i.e., $\mathbf{R}_{\mathrm{BD},k}\sum_{i=1}^{B} \mathbf{H}_k^{(i)}\mathbf{T}_{\mathrm{BD},k}^{(i)} = \mathbf{\Lambda}_{\mathrm{BD},k}$, and $\mathbf{R}_{\mathrm{BD},k}\sum_{i=1}^{B} \mathbf{H}_k^{(i)}\mathbf{T}_{\mathrm{BD},m}^{(i)} = \mathbf{0}$ for all $m \neq k$. Here, $\{\mathbf{R}_{\mathrm{BD},k}\}$ and $\{\mathbf{T}_{\mathrm{BD},k}\}$ are the transceiver computed from [3], $\mathbf{\Lambda}_{\mathrm{BD},k} \in \mathbb{C}^{N\times N}$ is a diagonal matrix whose entries consist of the equivalent channel gains of the BD scheme. Then, (4) can be expressed as

$$\hat{\mathbf{s}}_k = \mathbf{\Lambda}_{\mathrm{BD},k}\mathbf{s}_k + \mathbf{R}_{\mathrm{BD},k}\mathbf{n}_k.$$

To give an regularization effect on the transceiver design, we may utilize $\mathbf{\Lambda}_{\mathrm{BD},k}$ as the weight matrix in the WS-MSE metric. However, due to per-BS power constraint, the BD scheme identifies the solution in an iterative manner [3]. Thus, for a simple computation of weight matrix, we employ the GMI method which generates good sum rate performance with a one-shot precoder design [8]. Actually, this GMI method acts like BD scheme at high SNR region [8].

Since the GMI scheme was originally introduced under total power constraint, we multiply the common scaling parameter η to enforce per-BS power constraint. Denoting $\{\mathbf{T}_{\mathrm{GMI,k}}^{(i)}\}$ as the precoder computed from [8], the scaling parameter η is calculated as

$$\eta = \sqrt{\frac{P_{\max}}{\max_{i=1,\cdots,B} \mathrm{tr}(\sum_{m=1}^{K} \mathbf{T}_{\mathrm{GMI},m}^{(i)}\mathbf{T}_{\mathrm{GMI},m}^{(i)H})}}.$$

Then, the weight matrix $\mathbf{\Lambda}_k$ is simply calculated using singular value decomposition (SVD) of the effective channel as

$$\eta\sum_{i=1}^{B} \mathbf{H}_k^{(i)}\mathbf{T}_{\mathrm{GMI},k}^{(i)} = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^H \quad \text{for} \quad k = 1, \cdots, K. \tag{5}$$

where $\mathbf{U}_k \in \mathbb{C}^{N\times N}$ and $\mathbf{V}_k \in \mathbb{C}^{N\times N}$ are unitary matrices, and $\mathbf{\Lambda}_k \in \mathbb{C}^{N\times N}$ is a diagonal matrix whose entries consist of the equivalent channel gains of the scaled GMI solution.

Finally, the proposed WS-MSE can be represented as

$$\Xi_\Sigma = \sum_{k=1}^{K} E\left[\|\hat{\mathbf{s}}_k - \mathbf{\Lambda}_k\mathbf{s}_k\|^2\right] = \sum_{k=1}^{K}\mathrm{tr}\Big\{\mathbf{R}_k\Big(\sum_{l=1}^{K}\mathbf{\Omega}_{kl}\mathbf{\Omega}_{kl}^H\Big)\mathbf{R}_k^H - \mathbf{\Lambda}_k\mathbf{\Omega}_{kk}^H\mathbf{R}_k^H - \mathbf{R}_k\mathbf{\Omega}_{kk}\mathbf{\Lambda}_k^H + \sigma_n^2\mathbf{R}_k\mathbf{R}_k^H\Big\}. \tag{6}$$

With an aid of the weight matrix $\mathbf{\Lambda}_k$, we can assign more power to stronger subchannels with larger channel gains [7]. When $\mathbf{\Lambda}_k = \mathbf{I}_N$, the problem (7) reduces to the S-MSE minimization. Since the transceiver matrices $\{\mathbf{T}_k^{(i)}\}$ and $\{\mathbf{R}_k\}$ are related to each other in general, it is difficult to optimize them simultaneously with the WS-MSE metric. Thus, we adopt an alternating optimization method which iteratively finds a local optimal solution of the transceivers.

First, we design the optimal precoder when the receive filters $\{\mathbf{R}_k\}$ are given. Then, the WS-MSE minimization under the per-BS power constraint can be formulated as

$$\min_{\{\mathbf{T}_k^{(i)}\}} \ \Xi_\Sigma$$

$$\text{subject to} \ \sum_{k=1}^{K}\mathrm{tr}\left(\mathbf{T}_k^{(i)}\mathbf{T}_k^{(i)H}\right) \leq P_{\max} \ \ \text{for } i = 1, \cdots, B. \tag{7}$$

Note that (7) is a convex optimization problem [9].

To solve (7), we construct the Lagrangian function as

$$\mathcal{L}\left(\{\mathbf{T}_k^{(i)}\},\{\lambda_i\}\right) = \Xi_\Sigma + \sum_{i=1}^{B}\lambda_i\Big(\sum_{k=1}^{K}\mathrm{tr}\ \left(\mathbf{T}_k^{(i)}\mathbf{T}_k^{(i)H} - P_{\max}\right)\Big). \tag{8}$$

Then, the Karush-Kuhn-Tucker (KKT) conditions [9] for (8) for $i = 1, \cdots, B$ are given by

$$\sum_{k=1}^{K}\mathrm{tr}\big(\mathbf{T}_k^{(i)}\mathbf{T}_k^{(i)H}\big) - P_{\max} \ \leq \ 0, \quad \lambda_i \ \geq \ 0,$$

$$\lambda_i\Big(\sum_{k=1}^{K}\mathrm{tr}\big(\mathbf{T}_k^{(i)}\mathbf{T}_k^{(i)H}\big) - P_{\max}\Big) = 0,$$

$$\frac{\partial\mathcal{L}}{\partial\mathbf{T}_k^{(i)*}} = \Big(\sum_{n=1}^{K}\mathbf{H}_n^{(i)H}\mathbf{R}_n^H\mathbf{R}_n\mathbf{H}_n^{(i)}\Big)\mathbf{T}_k^{(i)} + \lambda_i\mathbf{T}_k^{(i)} - \mathbf{\Upsilon}_k^{(i)} = 0 \quad k = 1, \cdots, K \tag{9}$$

where $\mathbf{\Upsilon}_k^{(i)} = -\big(\sum_{n=1}^{K}\mathbf{H}_n^{(i)H}\mathbf{R}_n^H\mathbf{R}_n(\sum_{m\neq i}\mathbf{H}_n^{(m)}\mathbf{T}_k^{(m)})\big) + \mathbf{H}_k^{(i)H}\mathbf{R}_k^H\mathbf{\Lambda}_k$.

Finally, the optimal precoding matrix $\mathbf{T}_k^{(i)}$ is computed as

$$\mathbf{T}_k^{(i)} = \Big(\sum_{n=1}^{K}\mathbf{H}_n^{(i)H}\mathbf{R}_n^H\mathbf{R}_n\mathbf{H}_n^{(i)} + \lambda_i\mathbf{I}_M\Big)^{-1}\mathbf{\Upsilon}_k^{(i)}. \tag{10}$$

Here, we compute the Lagrange multipliers in a similar fashion described in [6] and [10]. First, we define the SVD of $\sum_{n=1}^{K}\mathbf{H}_n^{(i)H}\mathbf{R}_n^H\mathbf{R}_n\mathbf{H}_n^{(i)}$ $(i = 1,\cdots,B)$ as $\sum_{n=1}^{K}\mathbf{H}_n^{(i)H}\mathbf{R}_n^H\mathbf{R}_n\mathbf{H}_n^{(i)} = \mathbf{U}^{(i)}\mathbf{\Sigma}^{(i)}\mathbf{U}^{(i)H}$ where $\mathbf{U}^{(i)} \in \mathbb{C}^{N\times N}$ is an unitary matrix and $\mathbf{\Sigma}^{(i)} \in \mathbb{C}^{N\times N}$ is a diagonal matrix. Then, the power constraint equality for BS i in (7) is given by

$$
\begin{aligned}
\sum_{k=1}^{K}\mathrm{tr}\big(\mathbf{T}_k^{(i)}\mathbf{T}_k^{(i)H}\big) &= \sum_{k=1}^{K}\mathrm{tr}\Big(\big(\mathbf{U}^{(i)}\mathbf{\Sigma}^{(i)}\mathbf{U}^{(i)H}+\lambda_i\mathbf{I}\big)^{-2}\mathbf{\Upsilon}_k^{(i)}\mathbf{\Upsilon}_k^{(i)H}\Big)\\
&= \sum_{k=1}^{K}\mathrm{tr}\Big(\big(\mathbf{\Sigma}^{(i)}+\lambda_i\mathbf{I}\big)^{-2}\mathbf{D}_k^{(i)}\Big)\\
&= \sum_{k=1}^{K}\sum_{n=1}^{N}\frac{[\mathbf{D}_k^{(i)}]_{n,n}}{\big([\mathbf{\Sigma}^{(i)}]_{n,n}+\lambda_i\big)^2} = P_{\max} \qquad (11)
\end{aligned}
$$

where $\mathbf{D}_k^{(i)} = \mathbf{U}^{(i)H}\mathbf{\Upsilon}_k^{(i)}\mathbf{\Upsilon}_k^{(i)H}\mathbf{U}^{(i)}$ and $[\mathbf{A}]_{n,n}$ indicates the n-th diagonal element of a matrix $\mathbf{A}$.

Since $\sum_{k=1}^{K}\sum_{n=1}^{N}\frac{[\mathbf{D}_k^{(i)}]_{n,n}}{([\mathbf{\Sigma}^{(i)}]_{n,n}+\lambda_i)^2}$ is a monotonically decreasing function with respect to λ_i, the Lagrange multiplier λ_i can be efficiently solved by a bisection method. Due to the second KKT condition in (9), the Lagrange multipliers $\{\lambda_i\}$ should be non-negative real values. Therefore, the optimal $\{\lambda_i\}$ can be determined as a solution of equation (11). If such solutions do not exist, λ_i is set to zero.

Next, we illustrate the optimization of the receive filter $\mathbf{R}_k$ with the given precoders $\{\mathbf{T}_k^{(i)}\}$. In this case, the WS-MSE optimization becomes an unconstraint minimization problem expressed as

$$\min_{\{\mathbf{R}_k\}} \Xi_\Sigma. \qquad (12)$$

Since this objective function is convex and differentiable, a necessary and sufficient condition for $\mathbf{R}_k$ to become optimal is $\nabla_{\mathbf{R}_k^*}\Xi_\Sigma = 0$ [9]. As a result, the optimal receive filter for user k is given by

$$\mathbf{R}_k = \mathbf{\Lambda}_k\mathbf{\Omega}_{kk}^H\Big(\sum_{l=1}^{K}\mathbf{\Omega}_{kl}\mathbf{\Omega}_{kl}^H + \sigma_n^2\mathbf{I}_N\Big)^{-1}. \qquad (13)$$

The overall proposed WS-MSE scheme is summarized below.

---

1) Initialize $\{\mathbf{T}_k^{(i)}\}$ with arbitrary precoding matrices and compute $\{\mathbf{\Lambda}_k\}$ using (5).
2) **for** $i = 1 : B$
 Update $\{\mathbf{R}_k\}$ using (13).
 Compute λ_i based on (11).
 Update $\mathbf{T}_k^{(i)}$ using (10), $k = 1,\cdots,K$.
end
3) Go back to step 2) until convergence.

Although the WS-MSE function in (6) is not jointly convex on $\{\mathbf{T}_k^{(i)}\}$ and $\{\mathbf{R}_k\}$, it is convex over each precoder and receive filter. Obviously, the optimal precoder in (10) minimizes the WS-MSE when other filters are fixed. Similarly, the WS-MSE is minimized using the optimized receive filter in (13). This leads $\Xi_\Sigma(\{\mathbf{T}_k^{(i)}(n+1)\},\{\mathbf{R}_k(n+1)\}) \leq \Xi_\Sigma(\{\mathbf{T}_k^{(i)}(n)\},\{\mathbf{R}_k(n)\})$ where $\{\mathbf{T}_k^{(i)}(n)\}$ and $\{\mathbf{R}_k(n)\}$ represent precoders and received filters at the n-th iteration, respectively. Thus, at each iteration, updating $\{\mathbf{T}_k^{(i)}\}$ and $\{\mathbf{R}_k\}$ monotonically reduces the WS-MSE Ξ_Σ which is lower bounded by zero. As a result, our WS-MSE algorithm guarantees convergence at least to a local minimum solution.

IV. Modified WS-MSE Minimization Algorithm with Imperfect CSI

In practice, due to an estimation error and feedback error, the mismatch between the true channel $\mathbf{H}_k^{(i)}$ and the estimated channel $\hat{\mathbf{H}}_k^{(i)}$ is inevitable [10] [11]. In this section, we propose a modified WS-MSE transceiver design with imperfect CSI. We assume that $\hat{\mathbf{H}}_k^{(i)}$ is related to $\mathbf{H}_k^{(i)}$ as $\hat{\mathbf{H}}_k^{(i)} = \mathbf{H}_k^{(i)} + \mathbf{E}_k^{(i)}$ where the elements of $\mathbf{E}_k^{(i)}$ are i.i.d. complex Gaussian random variables with zero mean and variance σ_e^2. Then, the received signal (4) can be rewritten as

$$\hat{\mathbf{s}}_k = \mathbf{R}_k\sum_{i=1}^{B}\Big(\hat{\mathbf{H}}_k^{(i)} - \mathbf{E}_k^{(i)}\Big)\sum_{m=1}^{K}\mathbf{T}_m^{(i)}\mathbf{s}_m + \mathbf{R}_k\mathbf{n}_k$$

where $\mathbf{R}_k$ and $\mathbf{T}_k^{(i)}$ are computed only from $\hat{\mathbf{H}}_k^{(i)}$'s regardless of $\mathbf{E}_k^{(i)}$'s.

In order to mitigate the effect of the channel error, the WS-MSE metric is averaged over $\{\mathbf{E}_k^{(i)}\}$ as

$$\sum_{k=1}^{K}E_{\{\mathbf{E}_k^{(i)}\}}\left[\|\hat{\mathbf{s}}_k - \mathbf{\Lambda}_k\mathbf{s}_k\|^2\right] = \bar{\Xi}_\Sigma + \sum_{k=1}^{K}\Psi_k \qquad (14)$$

where the expectation with respect to $\mathbf{s}_k$ and $\mathbf{n}_k$ is implicitly included. Here, $\bar{\Xi}_\Sigma$ is computed by replacing $\mathbf{H}_k^{(i)}$ with $\hat{\mathbf{H}}_k^{(i)}$ in (6) and $\Psi_k = \sigma_e^2\mathrm{tr}\left(\mathbf{\Sigma}_{\mathbf{R}_k}\right)\sum_{l=1}^{K}\sum_{i=1}^{B}\mathrm{tr}(\mathbf{T}_l^{(i)}\mathbf{T}_l^{(i)H}) = \big(\sigma_e^2\sum_{l=1}^{K}\sum_{i=1}^{B}\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{T}_l^{(i)}})\big)\mathrm{tr}(\mathbf{R}_k\mathbf{R}_k^H)$. When obtaining (14), we have used the following equalities:

$$
\begin{aligned}
\mathbf{T}_l^{(i)}\mathbf{T}_l^{(i)H} &= \mathbf{U}_{\mathbf{T}_l^{(i)}}\mathbf{\Sigma}_{\mathbf{T}_l^{(i)}}\mathbf{U}_{\mathbf{T}_l^{(i)}}^H,\\
\mathbf{R}_k\mathbf{R}_k^H &= \mathbf{U}_{\mathbf{R}_k}\mathbf{\Sigma}_{\mathbf{R}_k}\mathbf{U}_{\mathbf{R}_k}^H,\\
E\Big[\mathbf{E}_k^{(i)}\mathbf{T}_l^{(i)}\mathbf{T}_l^{(i)H}\mathbf{E}_k^{(i)H}\Big] &= \sigma_e^2\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{T}_l^{(i)}})\mathbf{I}_N,\\
E\Big[\mathbf{E}_k^{(i)H}\mathbf{R}_k^H\mathbf{R}_k\mathbf{E}_k^{(i)}\Big] &= \sigma_e^2\mathrm{tr}\left(\mathbf{\Sigma}_{\mathbf{R}_k}\right)\mathbf{I}_M.
\end{aligned}
$$

Since the only difference between (6) and (14) is the last term $\sum_{k=1}^{K}\Psi_k$, the robust transceiver can be calculated in a similar way of Section III. Using KKT conditions, we have

$$\mathbf{T}_k^{(i)} = \Big(\sum_{n=1}^{K}\hat{\mathbf{H}}_n^{(i)H}\mathbf{R}_n^H\mathbf{R}_n\hat{\mathbf{H}}_n^{(i)} + \alpha_k^{(i)}\mathbf{I}_M\Big)^{-1}\mathbf{\Upsilon}_k^{(i)}, \qquad (15)$$

$$\mathbf{R}_k = \mathbf{\Lambda}_k\hat{\mathbf{\Omega}}_{kk}^H\Big(\sum_{l=1}^{K}\hat{\mathbf{\Omega}}_{kl}\hat{\mathbf{\Omega}}_{kl}^H + \beta\mathbf{I}_N\Big)^{-1} \qquad (16)$$

where $\alpha_k^{(i)} = \left(\lambda_i + \sigma_e^2\mathrm{tr}\left(\mathbf{\Sigma}_{\mathbf{R}_k}\right)\right)$, $\hat{\mathbf{\Omega}}_{km} = \Big(\sum_{i=1}^{B}\hat{\mathbf{H}}_k^{(i)}\mathbf{T}_m^{(i)}\Big)$, $\beta = \Big(\sigma_n^2 + \sigma_e^2\sum_{l=1}^{K}\sum_{i=1}^{B}\mathrm{tr}(\mathbf{\Sigma}_{\mathbf{T}_l^{(i)}})\Big)$, and the Lagrangian multiplier λ_i can be optimized in a similar fashion as described in (11). Finally, the overall iterative algorithm can be carried out by using (15) and (16) instead of (10) and (13) in the WS-MSE minimization algorithm, respectively.

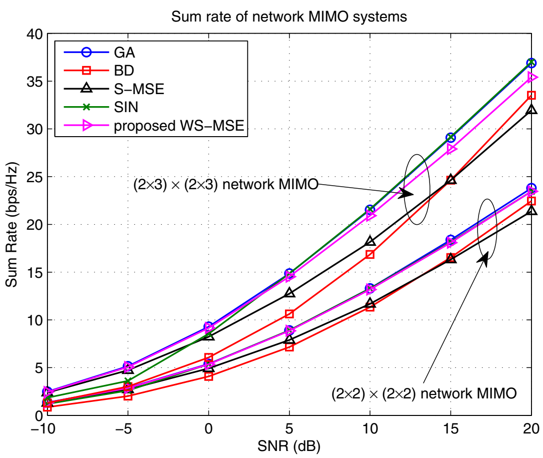

Fig. 1. Sum rate comparison for $(2\times2)\times(2\times2)$ and $(2\times3)\times(2\times3)$ network MIMO systems.

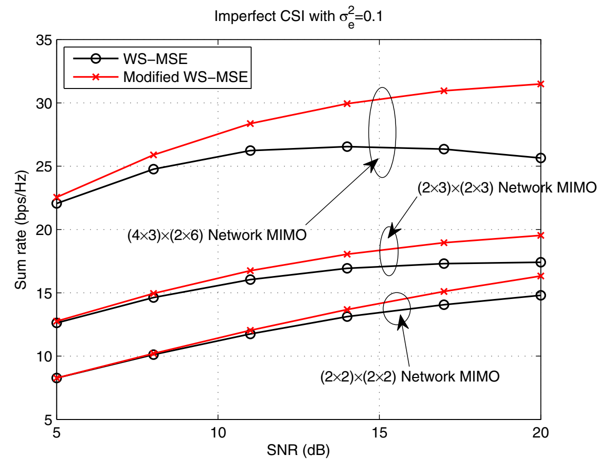

Fig. 2. Sum rate performance with imperfect channel knowledge.

V. Simulation Results

In this section, we provide simulation results to evaluate the sum rate performance of the proposed schemes. We present the average sum rate of the following linear precoding schemes as a function of SNR = $\frac{P_{\max}}{\sigma_n^2}$.

- GA: the gradient ascent algorithm with 5 initial points is performed for the sum rate maximization problem and the best local optimal solution is chosen [12].
- BD: the ZF precoding matrices are optimized with per-BS power constraint [3].
- S-MSE: the joint transceiver is optimized with the criterion of minimizing the S-MSE with per-BS power constraint [4], where the optimal precoding matrices are computed using standard convex tools such as SEDUMI.
- SIN: the soft interference nulling technique is performed to maximize the first order Taylor series expansion of a sum rate expression [5], where the SEDUMI is also utilized.

In Fig. 1, the average sum rate performance is presented for $(2\times2)\times(2\times2)$ and $(2\times3)\times(2\times3)$ network MIMO systems. From this plot, we can see that the proposed WS-MSE minimization scheme shows the sum rate performance almost identical to the near-optimal GA method and outperforms the conventional BD and S-MSE minimization method. It is observed that the sum rate gap between the WS-MSE minimization and the S-MSE minimization scheme increases as the SNR grows. This is due to a use of weight matrices $\{\mathbf{\Lambda}_k\}$ in the cost function in (7). Also, the cross-over point between the S-MSE and the BD scheme occurs at higher SNR as the number of antennas increases. These observations mean that the effect of the regularization tends to be more important when the number of data stream grows.

In Fig. 2, we demonstrate the effectiveness of the modified WS-MSE minimization algorithm compared to the original WS-MSE scheme in Section III with various system configurations when $\sigma_e^2 = 0.1$. It is clear that the channel mismatch compensation becomes more important for a larger system.

From the complexity perspective, the WS-MSE algorithm requires 1194 floating point (flop) operations in each iteration for $(2\times2)\times(2\times2)$ network MIMO systems. On the other hand, the required flops for GA and BD schemes are 3458 and 726, respectively. The flop evaluation is performed according to [13] as

- Multiplication of $m\times n$ and $n\times p$ matrices: $2mnp$.
- Inversion of an $m\times m$ with Gauss-Jordan elimination method: $4m^3/3$.
- SVD of an $m\times n$ matrix ($m\le n$): $4n^2m+13m^3$.

In addition, the GA and BD algorithms need a step size control for provable convergence. Thus, the complexity of the WS-MSE is significantly lower compared to the GA algorithm.

References

[1] G. J. Foschini, K. Karakayali, and R. A. Valenzuela, “Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency,” *IEE Proc. Commun.*, vol. 153, pp. 548–555, Aug. 2006.

[2] J. Zhang and J. G. Andrews, “Adaptive spatial intercell interference cancellation in multicell wireless networks,” *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1455–1468, Dec. 2010.

[3] R. Zhang, “Cooperative multi-cell block diagonalization with per-base-station power constraints,” *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1435–1445, Dec. 2010.

[4] S. Shi, M. Schubert, N. Vucic, and H. Boche, “MMSE optimization with per-base-station power constraints for network MIMO systems,” in *Proc. 2008 IEEE ICC*.

[5] C. T. K. Ng and H. Huang, “Linear precoding in cooperative MIMO cellular networks with limited coordiniation clusters,” *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1446–1452, Dec. 2010.

[6] J. Zhang, Y. Wu, S. Zhou, and J. Wang, “Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems,” *IEEE Commun. Lett.*, vol. 9, pp. 991–993, Nov. 2005.

[7] J. Joung and Y. H. Lee, “Regularized channel diagonalization for multiuser MIMO downlink using a modified MMSE criterion,” *IEEE Trans. Signal Process.*, vol. 55, pp. 1573–1579, Apr. 2007.

[8] H. Sung, S.-R. Lee, and I. Lee, “Generalized channel inversion methods for multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 57, pp. 3489–3499, Nov. 2009.

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

[10] H. Shen, B. Li, M. Tao, and X. Wang, “MSE-based transceiver designs for the MIMO interference channel,” *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3480–3489, Nov. 2010.

[11] S.-H. Park, H. Park, Y.-D. Kim, and I. Lee, “Regularized interference alignment based on weighted sum-MSE criterion for MIMO interference channels,” in *Proc. 2010 IEEE ICC*.

[12] H. Sung, S.-H. Park, K.-J. Lee, and I. Lee, “Linear precoder designs for K-user interference channels,” *IEEE Trans. Wireless Commun.*, vol. 9, pp. 291–301, Jan. 2010.

[13] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd edition. The Johns Hopkins University Press, 1996.