International Journal of Digital Multimedia Broadcasting

Multicell Cooperation and MIMO Technologies for Broadcasting and Broadband Communications

Guest Editors: Hongxiang Li, Lingjia Liu, Guoqing Li, Younsun Kim, and Jinyun Zhang

Multicell Cooperation and MIMO Technologies for Broadcasting and Broadband Communications

International Journal of Digital Multimedia Broadcasting

Multicell Cooperation and MIMO Technologies for Broadcasting and Broadband Communications

Guest Editors: Hongxiang Li, Lingjia Liu, Guoqing Li, Younsun Kim, and Jinyun Zhang

Contents

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 848527, 2 pages
doi:10.1155/2010/848527

Editorial

Multicell Cooperation and MIMO Technologies for Broadcasting and Broadband Communications

Hongxiang Li,[1] Lingjia Liu,[2] Guoqing Li,[3] Younsun Kim,[4] and Jinyun Zhang[5]

[1] *Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND 58108-6050, USA*
[2] *Samsung Electronics, Richardson, TX 75082, USA*
[3] *Intel Corporation, Hillsboro, OR 97124, USA*
[4] *Samsung Electronics, Seoul, Republic of Korea*
[5] *Mitsubishi Electrical Research Labs (MERL), Cambridge, MA 02139, USA*

Correspondence should be addressed to Hongxiang Li, hongxiang.li@ndsu.edu

Received 6 April 2010; Accepted 6 April 2010

The wireless industry is experiencing an unprecedented increase in the number and sophistication of broadcasting and broadband communication systems. The growing diffusion of new services, like mobile television and multimedia communications, emphasizes the need of advanced transmission techniques that can fundamentally increase the system capacity. In this context, the multicell collaborative transmission is becoming a major subject of research in the wireless communication community as it has been identified as one of the underlying principles for future wireless communication systems. Further, if perfect cooperation is assumed, it allows the entire network to be viewed as a single "super-MIMO" system with distributed antenna array at the base station.

This special issue aims at promoting state-of-the-art research contributions from all research areas either directly involved in or contributing to improving the issues related to multicell cooperation and MIMO technologies for broadcasting and broadband communications.

The first paper "Multiantenna analog network coding for multihop wireless networks" by R. Annavajjala et al., proposes a two-phase minimum mean-square-error bidirectional amplify and forward (MMSE-BAF) relaying protocol to allow two sources exchange independent messages via a relay node equipped with multiple antennas. The proposed protocol extends upon the so-called analog network coding schemes in the literature in that it inherently exploits the multiple antennas at the relay station to reduce the noise enhancement typical of an AF protocol and can also compensate for link imbalances between the relay and the sources and is agnostic to sources' modulation and coding schemes.

The second paper "Near optimum power control and precoding under fairness constraints in network MIMO systems" by G. Fodor et al. considers the problem of setting the uplink signal-to-noise-and-interference (SINR) target and allocating transmit powers for mobile stations in multicell spatial multiplexing wireless systems. The authors develop a numerical technique for real-time optimization of SINR targets and transmit powers. The paper also studies the impact of near optimal precoding in a multicell MIMO environment.

The third paper "Broadcast network coverage with multicell cooperation" by H. Li et al. studies the benefits of multicell cooperation in broadcast TV network from an information theoretical perspective. The paper defines outage capacity as the figure of merit and evaluates the broadcast coverage area of multicell system where multiple base stations collaboratively transmit the broadcast signals. The results show that the coverage of a TV broadcast network can be significantly improved by multicell cooperation.

The fourth paper "An analytical multimodulus algorithm for blind demodulation in a time-varying MIMO channel context" by S. Daumont and D. Le Guennec addresses the issue of blind MIMO demodulation of communication signals, with time-varying channels and in an interception context. A new adaptive-blind source separation algorithm, which is based on the implementation of the multimodulus cost function, is proposed.

The fifth paper "A transmit beamforming and nulling approach with distributed scheduling to improve cell edge throughput" by W. C. Wong et al. proposes a transmit scheme for WiMAX systems, where multiple base stations (BSs) employ downlink transmit beamforming and nulling for interference mitigation, with minimal coordination amongst BSs. This scheme improves system throughput and robustness, by increasing cell edge and overall cell throughputs by 68% and 19%, respectively, and by delivering improvement for mobile speed up to 60 km/hour.

The sixth paper "Intercell interference coordination through limited feedback" by L. Liu et al. considers the applications of multicell transmission schemes to the downlink of future wireless communication networks. A MIMO-based scheme with limited coordination among neighboring base stations (BSs) is proposed to effectively combat the intercell interference by taking advantage of the degrees of freedom in the spatial domain.

The seventh paper "Phase-shift cyclic-delay diversity for MIMO OFDM systems" by Y. Nam et al. introduces phase-shift cyclic delay diversity (PS CDD) and space-frequency-block-code PS CDD schemes for MIMO OFDM system. The proposed PS CDD scheme preserves the diversity advantage of traditional CDD in uncorrelated multiantenna channels and furthermore removes frequency-selective nulling problem of the traditional CDD in correlated multiantenna channels.

The eighth paper "Spatial domain resource sharing for overlapping cells in indoor environment" by R. Kudo et al. presents the effectiveness of spatial resource sharing among two access points. The authors introduce the primary-secondary AP scenario based on zero forcing (PSZ) and cooperative AP scenario based on zero forcing (CZ), confirming that spatial resource sharing is very attractive for the overlapping of cell scenario.

The last paper of this special issue "Robust MMSE transceiver designs for downlink MIMO systems with multicell cooperation" by J. Li et al. proposes, for joint MMSE transceiver designs, three novel robust approaches: the *Robust-GIA* (the most general), *Robust-FIA* (the most efficient), and the *Robust-DCOA* (which guarantees the global optimality). When cell cooperation is available, the robust approaches provide a remedy for solving the cell edge problem without reducing the number of data streams.

Acknowledgments

The authors are grateful to the reviewers for their invaluable work and to the authors of the papers collected in this special issue.

Hongxiang Li
Lingjia Liu
Guoqing Li
Younsun Kim
Jinyun Zhang

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 368562, 10 pages
doi:10.1155/2010/368562

Research Article

Multiantenna Analog Network Coding for Multihop Wireless Networks

Ramesh Annavajjala, Amine Maaref, and Jinyun Zhang

Mitsubishi Electric Research Laboratories (MERL), 201 Broadway Cambridge, MA 02139, USA

Correspondence should be addressed to Ramesh Annavajjala, ramesh.annavajjala@gmail.com

Received 16 July 2009; Accepted 26 October 2009

Academic Editor: Guoqing Li

This paper proposes a two-phase minimum mean-square-error bidirectional amplify-and forward (MMSE-BAF) relaying protocol to allow two sources exchange independent messages via a relay node equipped with multiple antennas. MMSE-BAF performs a joint linear MMSE filtering of the received signal after the multiple access phase before amplifying and forwarding the filtered signal using a single transmit antenna, possibly through a specific antenna selection procedure, during the broadcast phase. The proposed protocol extends upon the so-called analog network coding schemes in the literature in that it inherently exploits the multiple antennas at the relay station to reduce the noise enhancement typical of an AF protocol, and can also compensate for link imbalances between the relay and the sources and is agnostic to sources' modulation and coding schemes. We derive the instantaneous signal-to-noise ratio expressions for the received signal by the sources in the downlink and provide extensive link-level simulations for the MMSE-BAF protocol subject to both frequency flat and selective fading. Furthermore, we pinpoint the modifications to be incorporated into the IEEE 802.16e orthogonal-frequency-division multiple access (OFDMA) cellular standard (mobile WiMax) to enable support of multiantenna bidirectional communications and show that MMSE-BAF is a viable solution within that framework.

1. Introduction

Half-duplex bidirectional relay systems in which two nodes S_1 and S_2 wish to exchange independent messages via a third node R, termed relay, give rise to some interesting challenges from a cooperative communications and information-theoretic points of view. This is especially true when the relay node R is equipped with multiple antennas. Such two-way relay channels have many applications in ad hoc and cellular networks in which all mobile-to-mobile communications have to pass through a common base station. Since full-duplex operation is of little practical interest given current state-of-the-art technology, our focus is on half-duplex nodes, where each active node can either transmit or receive an information message at a given point in time. In particular, without loss of generality, we are interested in the communications part of the problem in a cellular context where two mobile stations wish to exchange data simultaneously via a common base station.

The traditional baseline approach for bidirectional communications in half-duplex mode between two sources S_1 and S_2 via a relay station R consists of a 4-phase protocol with a completion time of 4-time-slots (TSs) whereby S_1 and S_2 send N-bit packets $\mathbf{b}_1$ and $\mathbf{b}_2 \in \{0,1\}^N$ to R during TS 1 and TS 2, respectively; R decodes the received packets and then sends $\mathbf{b}_2$ to S_1 and $\mathbf{b}_1$ to S_2 during TS 3 and TS 4, respectively. The gist of the 4-phase protocol is to avoid interference by preventing simultaneous transmissions from the sources to the relay and vice versa. However, it was shown in [1] that a three-phase protocol exploiting the so-called network coding concept by decoding packets $\mathbf{b}_1$ and $\mathbf{b}_2$ at the relay and then broadcasting a single packet $\mathbf{b}_1 \oplus \mathbf{b}_2$, where $\oplus$ denotes the bit-wise exclusive-or (XOR) operation, is actually more attractive in terms of achievable throughput, since the desired packet at S_1 can be decoded using another XOR operation (and similarly at S_1). Better still, a recent concept introduced in [2] and termed analog network coding (ANC) combines the first two phases of the conventional baseline protocol into a single

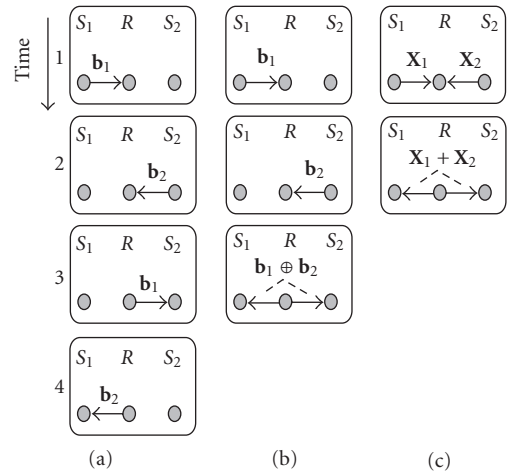


Figure 1: Bidirectional communication protocols: (a) 4-phase conventional protocol, (b) 3-phase protocol with decode-and-forward network coding, and (c) 2-phase BAF protocol with amplify-and-forward relaying.

multiple access (MA) phase with simultaneous transmissions from the sources to the relay; the received multiple access signal at the relay is then amplified and broadcast to S_1 and S_2, thereby yielding the so-called two-phase bidirectional amplify-and-forward (BAF) protocol. A similar concept to ANC, using estimate-and-forward relaying as opposed to AF relaying, has been proposed in [3] under the terminology of physical-layer network coding (PNC). A schematic diagram illustrating the aforementioned bidirectional protocols is illustrated in Figure 1.

Information-theoretic aspects such as bounds on the achievable throughput and the capacity region of the bidirectional relay channel have been investigated in [4–6]. A Markov-chain-based performance analysis for several variants of the BAF protocol was carried out in [7–9]. Linear beamforming filter designs for bidirectional communications with multiantenna relay stations are proposed in [10–12]. In this paper, we propose a simple two-phase minimum-mean-square-error- (MMSE-) BAF protocol which operates by filtering the received signal at the relay station after the MA phase using a specially designed joint linear MMSE filter before amplifying and forwarding the filtered signal during the broadcast phase. Whereas optimal relay beamforming structure for bidirectional multiantenna relay channels is sought in [12], our MMSE-BAF protocol is a simple low-complexity driven approach for bidirectional multiantenna relay channels which exploits linear signal processing on the uplink (from S_1 and S_2 to R) and transmit antenna selection (TAS) on the downlink (from R to S_1 and S_2). Besides, MMSE-BAF is a two-phase bidirectional relaying protocol, whereas the multiantenna relaying protocol put forward in [10] is a three-phase protocol which necessitates decoding and reencoding of the received signals at the RS prior to the broadcast phase. Finally, MMSE-BAF differs from the so-called spatial division duplex (SDD) bidirectional relaying scheme proposed in [11] in that it allows to bias the beamforming weights in favor of one of the two source nodes in as much as required to compensate for potential imbalance of the relay-to-source link channel gains or other parameters such as dissimilar signal constellations employed at S_1 and S_2.

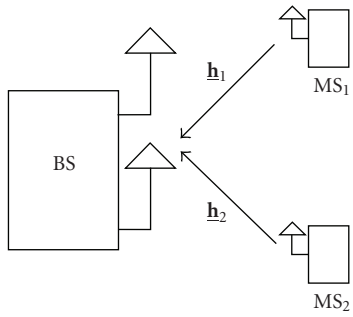
The remainder of this paper is structured as follows. Section 2 presents the system model for MMSE-BAF relaying. In Section 3, we describe the proposed MMSE-BAF protocol and derive the signal-to-noise ratio (SNR) expressions upon which our extensive simulation results provided in Section 4 are based. Finally, concluding remarks are drawn in Section 5.

2. System Setup and Signal Model

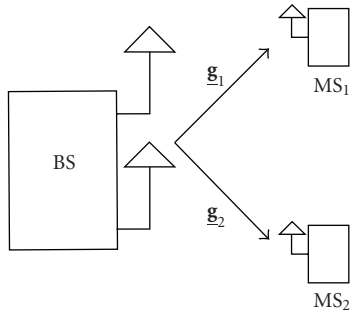
The following set of notations is employed throughout this paper: Boldface upper- and lower-case symbols are used to denote matrices and column vectors, respectively. $\mathbf{I}_m$ denotes the identity matrix of order m. Moreover, $(\cdot)^*, (\cdot)^T, (\cdot)^{\mathcal{H}}$, and $\mathrm{E}[\cdot]$ stand for conjugate, transpose, transpose-conjugate, and expectation operators, respectively.

Without loss of generality, and although our proposal and framework are applicable to any type of two-hop bidirectional relay setting, we focus our attention on cellular systems. For that purpose, we consider an infrastructure-based wireless communications system consisting of two mobile stations (MSs), MS_1 and MS_2, and one base station (BS). A block diagram of the system under consideration is depicted in Figure 2. Both MSs as well as the BS are equipped with multiple antennas for reception with the aim of canceling out potential other cell/user interference but are restricted to transmit through a single transmit antenna. This assumption is dictated by the need to reduce the transmit-power requirements for user terminals and to lower the complexity and cost of a transmission chain at the base station, generally higher than that of a reception chain, especially when accounting for high-cost radio-frequency amplifiers involved in the transmission chain. This is, for instance, the case in current cellular standards such as IEEE 802.16e [13]. Note that our proposed scheme works equally well with both time-division duplex (TDD) as well as frequency-division duplex (FDD) modes of operation. Without loss of generality and for the sake of notational brevity, we focus on the FDD mode in the following analysis. Performance results for both TDD and FDD will be presented in Section 4.

Complex baseband transmission is assumed throughout the paper. Let M_{bs} denote the number of receive antennas at the BS, and let $\mathbf{h}_1[n] = [h_1^1[n], h_1^2[n], \ldots, h_1^{M_{\mathrm{bs}}}[n]]^T$ and $\mathbf{h}_2[n] = [h_2^1[n], h_2^2[n], \ldots, h_2^{M_{\mathrm{bs}}}[n]]^T$ denote the $M_{\mathrm{bs}} \times 1$ uplink channels from MS_1 and MS_2 to the BS, respectively, where n is the discrete-time index. The corresponding time-varying channel vector elements $\{h_i^j[n]\}_{i=1,2;j=1,\ldots,M_{\mathrm{bs}}}$ are realizations of a zero-mean unit-variance Gaussian wide-sense stationary process. Assuming M_{ss} antennas at each of MS_1 and MS_2 for downlink reception, we define $\mathbf{g}_1[n] = [g_1^1[n], g_1^2[n], \ldots, g_1^{M_{ss}}[n]]^T$ and $\mathbf{g}_2[n] = [g_2^1[n], g_2^2[n], \ldots, g_2^{M_{ss}}[n]]^T$ as the downlink channels from



(a) Uplink multiple access phase



(b) Dowlink broadcast phase

Figure 2: Block diagram of the system setup.

BS to MS_1 and MS_2, respectively. For the special case of $M_{bs} = M_{ss} = 1$, TDD assumption allows us to set $\mathbf{g}_1[n] = \mathbf{h}_1^T[n]$ and $\mathbf{g}_2 = \mathbf{h}_2^T[n]$. Let $x_1[n]$ be the signal transmitted from MS_1 and intended for MS_2 and let $x_2[n]$ be the signal transmitted from MS_2 and intended for MS_1 at time n. Both $x_1[n]$ and $x_2[n]$ are drawn from two possibly different complex signal constellations with average energies $\sigma_1^2 = \mathrm{E}[|x_1[n]|^2]$ and $\sigma_2^2 = \mathrm{E}[|x_2[n]|^2]$, respectively. Prior to any signal processing at the BS, the $M_{bs} \times 1$-received signal at the end of the MA phase is given by

$$\mathbf{y}[n] = \mathbf{h}_1[n]x_1[n] + \mathbf{h}_2[n]x_2[n] + \mathbf{n}[n], \tag{1}$$

where $\mathbf{n}[n]$ is an $M_{bs} \times 1$ additive white Gaussian noise vector at the BS which is modeled as a zero-mean circularly symmetric Gaussian random vector with covariance matrix $\mathrm{E}[\mathbf{n}[n]\mathbf{n}[n]^{\mathcal{H}}] = \sigma_N^2\mathbf{I}_{M_{bs}}$. For notational simplicity, the discrete-time index n is henceforth omitted.

3. MMSE-BAF: Description and Analysis

3.1. Protocol Description. Figure 3 illustrates the physical-layer frame structure for enabling the MMSE-BAF protocol for the uplink MA and the downlink broadcast phases. A frame consists of packets originating from the link-layer whose size depends on the chosen MCS so that each downlink or uplink frame contains a fixed number of symbols. As can be seen in the left-hand side of Figure 3, the uplink frame structure is composed of two parts, one for pilot symbols which are chosen to be orthogonal for MS_1 and MS_2. Orthogonality of the pilot symbols can be maintained in the time, frequency, or two-dimensional time-frequency grid. The second part is for data symbols. The orthogonal pilot symbols are used to estimate the channels $\mathbf{h}_1$ and $\mathbf{h}_2$ corresponding to MS_1 and MS_2. A beamforming weight vector $\mathbf{w}_{opt}$ is then computed at the BS based on a joint MMSE criterion to be specified shortly. The BS then estimates an amplification factor β subject to an average power constraint. Likewise, the downlink frame structure contains pilot and data parts. Additionally, it contains a control part consisting of quantized versions of the amplification factor β (a positive scalar value) and two complex scalars values $\nu_1 := \mathbf{w}_{opt}^{\mathcal{H}}\mathbf{h}_1$ and $\nu_2 := \mathbf{w}_{opt}^{\mathcal{H}}\mathbf{h}_2$. The downlink data symbols consist of the beamformed amplified-and-forwarded symbols received on the uplink frame in FDD mode.

According to the MMSE-BAF protocol, the BS performs the following set of operations.

(1) Jointly minimize the MSE between the received signal $\mathbf{y}[n]$ and the transmitted signals $x_1[n]$ and $x_2[n]$, thus performing a joint linear-MMSE filtering of the received signal, using the following metric:

$$\mathbf{w}_{opt} = \underset{\mathbf{w}\in\mathbb{C}^{M_{bs}}}{\operatorname{argmin}}\left\{\delta_1 \mathrm{E}\left[\left|x_1 - \mathbf{w}^{\mathcal{H}}\mathbf{y}\right|^2 \mid \mathbf{h}_1, \mathbf{h}_2\right] + \delta_2 \mathrm{E}\left[\left|x_2 - \mathbf{w}^{\mathcal{H}}\mathbf{y}\right|^2 \mid \mathbf{h}_1, \mathbf{h}_2\right]\right\}, \tag{2}$$

where $\mathbb{C}$ is the field of complex numbers and $\delta_1 \geq 0$, $\delta_2 \geq 0$, $\delta_1 + \delta_2 = 1$, are the two design constants that control the relative weight assigned to the signals of MS_1 and MS_2. The minimization problem in (2) is a modified Wiener filtering problem whose solution can be easily found using the orthogonality principal in linear mean square estimation and is given by

$$\mathbf{w}_{opt} = \left(\sigma_1^2\mathbf{h}_1\mathbf{h}_1^{\mathcal{H}} + \sigma_2^2\mathbf{h}_2\mathbf{h}_2^{\mathcal{H}} + \sigma_N^2\mathbf{I}_{M_{bs}}\right)^{-1}(\delta_1\sigma_1^2\mathbf{h}_1 + \delta_2\sigma_2^2\mathbf{h}_2). \tag{3}$$

This minimization requires an estimation of both mobile stations' vector-valued channels $\mathbf{h}_1$ and $\mathbf{h}_2$.

(2) Amplify the linear MMSE-filter output to maintain a constant average transmit power P_T which leads to computing the amplification gain factor:

$$\beta = \sqrt{\frac{P_T}{\mathrm{E}\left[\left|\mathbf{w}^{\mathcal{H}}\mathbf{y}\right|^2 \mid \mathbf{h}_1, \mathbf{h}_2\right]}} = \sqrt{\frac{P_T}{\sigma_1^2\left|\mathbf{w}_{opt}^{\mathcal{H}}\mathbf{h}_1\right|^2 + \sigma_2^2\left|\mathbf{w}_{opt}^{\mathcal{H}}\mathbf{h}_2\right|^2 + \sigma_N^2\left\|\mathbf{w}_{opt}\right\|^2}}. \tag{4}$$

(3) Transmit the amplified signal back to the MSs on one of the antennas using an appropriate downlink transmit antenna-selection (TAS) algorithm, based on the uplink channel. One approach inherent to the MMSE-BAF protocol is to select the antenna that has the largest beamformer weight. We note that, owing to the channel reciprocity, TAS using the largest beamformer weight is applicable for TDD only.

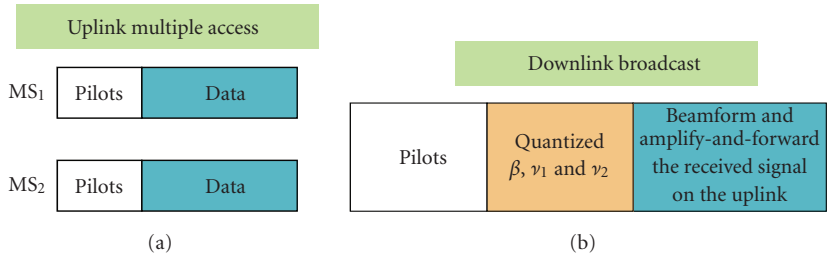


Figure 3: Block diagram of the system setup.

3.2. Performance Analysis. Define $z := \mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{y}$, where $\mathbf{y}$ is the uplink received signal (1), as the output of the MMSE filtering operation at the BS. The AF transmitted signal on the downlink is

$$x_r = \beta z = \beta \mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{y}. \tag{5}$$

The received signal on the downlink for MS_1 and MS_2 is therefore given by

$$\mathbf{y}_1 = \mathbf{g}_1 x_r + \mathbf{n}_1, \tag{6}$$

$$\mathbf{y}_2 = \mathbf{g}_2 x_r + \mathbf{n}_2, \tag{7}$$

where $\mathbf{n}_1$ and $\mathbf{n}_2$ are the zero-mean AWGN at MS_1 and MS_2, respectively, with covariance matrix $\sigma_N^2\mathbf{I}_{M_{ss}}$. Without loss of generality, let us focus on the signal received by MS_1 (6). A similar signal processing is required at MS_2. Incorporating (1) and (5) into (6) yields

$$\mathbf{y}_1 = \mathbf{g}_1\beta\mathbf{w}_{\text{opt}}^{\mathcal{H}}(\mathbf{h}_1 x_1 + \mathbf{h}_2 x_2 + \mathbf{n}) + \mathbf{n}_1. \tag{8}$$

Now, assuming that MS_1 (i) is able to perfectly estimate its own downlink channel vector $\mathbf{g}_1$ owing to the downlink pilot symbols sent by the BS, (ii) knows its own transmitted signal x_1, and (iii) is able to extract the value of the amplification factor β as well as the couple (ν_1, ν_2), all of which are sent on the downlink control channel as explained in Section 3.1, then MS_1 can subtract the so-called self-interference component of the received signal, that is, $\mathbf{g}_1\beta\nu_1 x_1$, hence winding up with a processed received signal of the form

$$\begin{aligned}\mathbf{r}_1 &= \mathbf{y}_1 - \mathbf{g}_1\beta\nu_1 x_1 \\ &= \mathbf{y}_1 - \mathbf{g}_1\beta\mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{h}_1 x_1 \\ &= \mathbf{g}_1\beta\mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{h}_2 x_2 + \underbrace{\mathbf{g}_1\beta\mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{n} + \mathbf{n}_1}_{:=\tilde{\mathbf{n}}_1}.\end{aligned} \tag{9}$$

Note that $\tilde{\mathbf{n}}_1$, defined in the previous equation, is a zero-mean colored noise vector with a conditional covariance matrix given by

$$\boldsymbol{\Sigma}_1 := \text{E}\left[\tilde{\mathbf{n}}_1\tilde{\mathbf{n}}_1^{\mathcal{H}} \mid \mathbf{g}_1, \beta, \mathbf{w}_{\text{opt}}\right] \tag{10}$$

$$= \beta^2\mathbf{g}_1\mathbf{w}_{\text{opt}}^{\mathcal{H}}\text{E}\left[\mathbf{n}\mathbf{n}^{\mathcal{H}}\right]\mathbf{w}_{\text{opt}}\mathbf{g}_1^{\mathcal{H}} + \text{E}\left[\mathbf{n}_1\mathbf{n}_1^{\mathcal{H}}\right] \tag{11}$$

$$= \sigma_N^2\left(\mathbf{I}_{M_{ss}} + \beta^2\mathbf{g}_1\mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{w}_{\text{opt}}\mathbf{g}_1^{\mathcal{H}}\right). \tag{12}$$

In the absence of knowledge of the conditional covariance matrix $\boldsymbol{\Sigma}_1$ at the receiver of MS_1, an estimate $\hat{x}_2$ of x_2 can be obtained as follows:

$$\hat{x}_2 = \frac{(\mathbf{g}_1\beta\nu_2)^{\mathcal{H}}}{(\mathbf{g}_1\beta\nu_2)^{\mathcal{H}}(\mathbf{g}_1\beta\nu_2)}\mathbf{r}_1, \tag{13}$$

which after simplification becomes

$$\hat{x}_2 = x_2 + \underbrace{\frac{1}{\beta\nu_2}\frac{\mathbf{g}_1^{\mathcal{H}}\tilde{\mathbf{n}}_1}{\mathbf{g}_1^{\mathcal{H}}\mathbf{g}_1}}_{:=\check{n}_2}. \tag{14}$$

The conditional variance of the newly defined scalar noise term $\check{n}_2$ can be expressed as

$$\begin{aligned}\boldsymbol{\Sigma}_1' = \text{E}\left[|\check{n}_2|^2\right] &= \frac{1}{\beta^2|\nu_2|^2}\frac{\mathbf{g}_1^{\mathcal{H}}\text{E}\left[\tilde{\mathbf{n}}_1^{\mathcal{H}}\tilde{\mathbf{n}}_1\right]\mathbf{g}_1}{\left(\mathbf{g}_1^{\mathcal{H}}\mathbf{g}_1\right)^2} \\ &= \frac{1}{\beta^2|\nu_2|^2}\frac{\mathbf{g}_1^{\mathcal{H}}\boldsymbol{\Sigma}_1\mathbf{g}_1}{\left(\mathbf{g}_1^{\mathcal{H}}\mathbf{g}_1\right)^2}.\end{aligned} \tag{15}$$

Now, we are in a position to determine the SNR γ_2 at MS_1 as

$$\gamma_2 = \frac{\text{E}\left[|x_2|^2\right]}{\text{E}\left[|\check{n}_2|^2\right]} \tag{16}$$

$$= \frac{\sigma_2^2\beta^2|\nu_2|^2\left(\mathbf{g}_1^{\mathcal{H}}\mathbf{g}_1\right)^2}{\mathbf{g}_1^{\mathcal{H}}\boldsymbol{\Sigma}_1\mathbf{g}_1}. \tag{17}$$

Note that after incorporating (10) into (17) and further simplification, (17) can be expressed as follows:

$$\gamma_2 = \frac{\sigma_2^2\beta^2|\nu_2|^2}{\sigma_N^2} \times \frac{||\mathbf{g}_1||^2}{1 + \beta^2\left|\left|\mathbf{w}_{\text{opt}}\right|\right|^2||\mathbf{g}_1||^2}. \tag{18}$$

Similarly, one can evaluate the SNR γ_1 for the signal x_1 received at MS_2 which is found to be

$$\gamma_1 = \frac{\sigma_1^2\beta^2|\nu_1|^2\left(\mathbf{g}_2^{\mathcal{H}}\mathbf{g}_2\right)^2}{\mathbf{g}_2^{\mathcal{H}}\boldsymbol{\Sigma}_2\mathbf{g}_2}, \tag{19}$$

where $\boldsymbol{\Sigma}_2$ is a noise covariance matrix (analogous to $\boldsymbol{\Sigma}_1$) defined as

$$\boldsymbol{\Sigma}_2 := \sigma_N^2\left(\mathbf{I}_{M_{ss}} + \beta^2\mathbf{g}_2\mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{w}_{\text{opt}}\mathbf{g}_2^{\mathcal{H}}\right). \tag{20}$$

Again, upon incorporation of (20) into (19), the latter can be simplified to

$$\gamma_1 = \frac{\sigma_1^2 \beta^2 |v_1|^2}{\sigma_N^2} \times \frac{||\mathbf{g}_2||^2}{1 + \beta^2 \left|\left|\mathbf{w}_{\text{opt}}\right|\right|^2 ||\mathbf{g}_2||^2}. \tag{21}$$

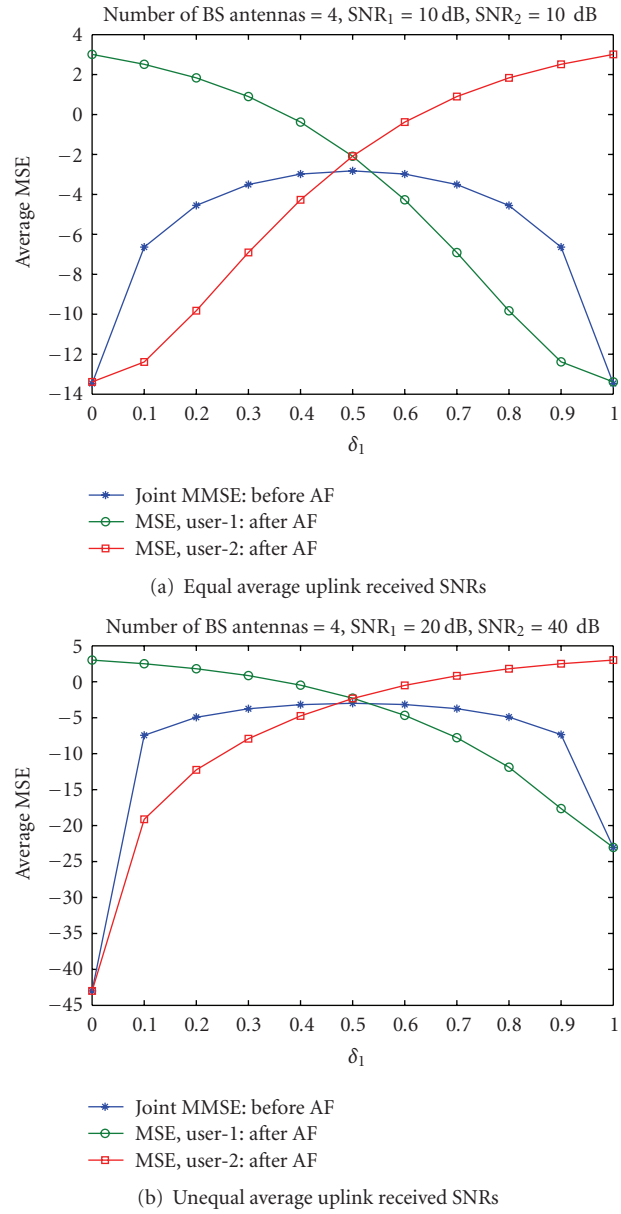
It is worthwhile to mention that the above SNR expressions for γ_2 (18) and γ_1 (21) have been obtained without exploiting the colored nature of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$. Somewhat surprisingly, and as indicated in the appendix, even through whitening of the colored noise, the SNR expressions for γ_1 and γ_2 remain the same, which is not so bad in itself, as it means that the signal processing cost associated with the whitening operation can be completed eliminated.

4. Simulation Results and Discussion

In this section, we present some simulation results on the performance of the proposed MMSE-BAF protocol. First, Figure 4 shows the MSE performance of MMSE-BAF as a function of the relative loading of user-1 over user-2, δ_1, with $M_{\text{bs}} = 4$ antennas at the base-station. Two scenarios are considered: in Figure 4(a), both users average received SNRs at the base-station are set to 10 dB, whereas in Figure 4(b) the average received SNR is set to 20 dB and 40 dB for user-1 and user-2, respectively. The weighted average MSE immediately after the application of MMSE beamforming and the average MSE of user-1 and user-2 after MMSE beamforming followed by AF are obtained by drawing independent channel realizations over 100000 trials. From Figures 4(a) and 4(b), we observe that increasing δ_1 minimizes the MSE of user-1 at the expense of an increase in MSE for user-2, whereas an optimum δ_1 exists that jointly minimizes the MSE of both users. Interestingly, from an implementation point of view, the range of δ_1 is broad to arrive at this optimum overall MSE. Comparing Figures 4(a) and 4(b), we notice that, due to unequal average received SNRs, the individual MSEs as well as the overall MSE are not symmetric functions of δ_1. Thus, one should take into account the knowledge of the average uplink SNRs to arrive at an appropriate δ_1 to maintain desired MSE levels for each of the two users.

In Figure 5, the empirical cumulative distribution function (CDF) of two-time-slots-based MMSE-BAF is compared against the four-time-slot-based baseline system. Here, the base-station and the mobile station receivers are each equipped with four receive antennas. The uplink average received SNRs, per antenna, of users 1 and 2 are set to 5 dB and 10 dB, respectively, and the average received SNR per antenna at each mobile station is set to 5 dB. We also assume $\delta_1 = \delta_2$. Figure 5 shows that the two-time-slots-based MMSE-BAF system outperforms the baseline system by an order of magnitude.

As argued earlier, the proposed MMSE-BAF protocol is equally attractive to both TDD and FDD systems. Figures 6 and 7 show uncoded symbol error rate (SER) performance of MMSE-BAF on block-fading TDD channels when the two users employ dissimilar modulation formats. In both figures, (i) user-1 employs QPSK modulation whereas user-2 employs 16-QAM modulation, (ii) a data frame contains 100 modulation symbols and 20 pilot symbols for channel estimation, and (iii) the channel remains constant over the duration of at least two frames (TDD assumption). In Figure 6, the base-station and the users all have single transmit/receive antennas. With equal average received SNRs at the base-station, Figure 6(a) shows that, at lower average received SNRs, pilot-based channel estimation matches very closely the performance achieved in case of perfect channel knowledge for both users. Since each user has to subtract its own channel-compensated transmitted symbol to decode the other user's modulation symbol, a user transmitting using a higher-order constellation has the potential to generate



(a) Equal average uplink received SNRs

(b) Unequal average uplink received SNRs

Figure 4: Average MSE of MMSE-BAF system with 4 antennas at the base-station.

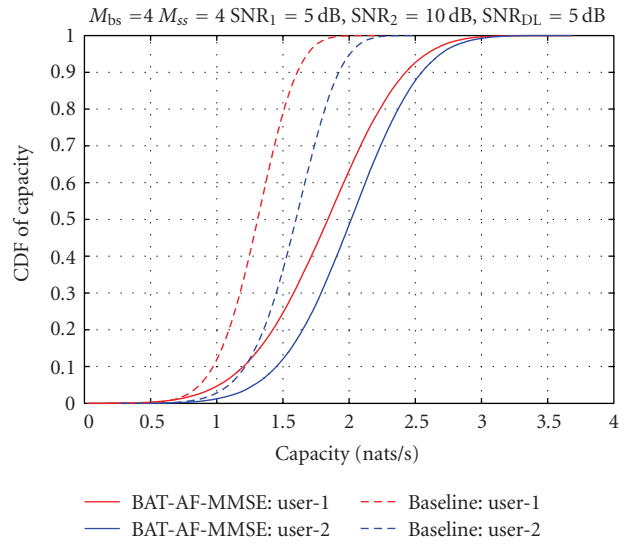
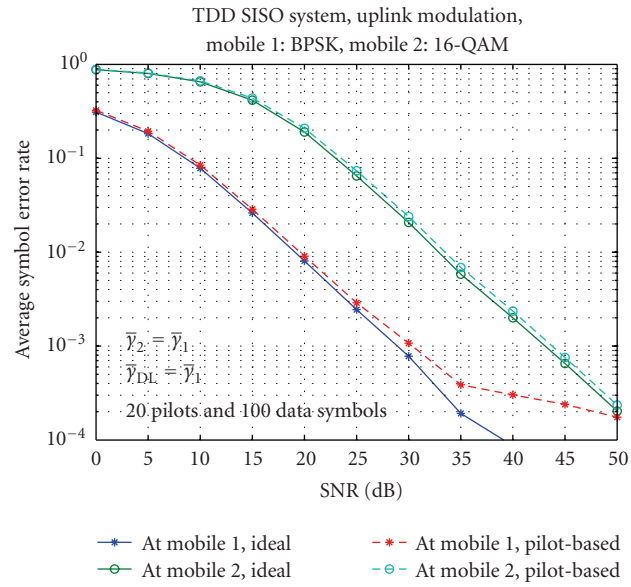
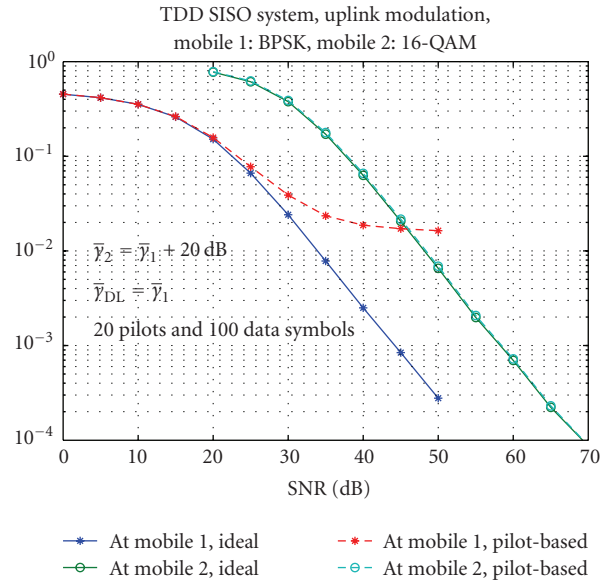


FIGURE 5: Comparison of empirical capacity CDFs of the MMSE-BAF system against the baseline system with 4 antennas at the base-station as well as the mobile receivers.



(a) Equal average uplink received SNRs



(b) Unequal average uplink received SNRs

FIGURE 6: Average uncoded SER of MMSE-BAF system with QPSK modulation for user 1 and 16-QAM modulation for user 2.

higher self-interference in the presence of channel estimation errors. Figure 6(a) shows that with equal average received SNRs, the average SER of BPSK exhibits an error floor for an average SNR higher than 30 dB. When user-2 transmits at an SNR that is 20 dB higher than user-1's SNR, Figure 6(b) shows that the error floor for the BPSK modulation occurs much earlier.

The advantages of transmit antenna selection over transmitting from an arbitrary antenna is investigated in Figure 7 when the base-station has four antennas for reception. Exploiting the channel reciprocity of TDD systems, we first compute the elementwise magnitude of the estimated beamformer and downlink transmission is directed from the antenna that has the highest magnitude. It is important to note that once the base-station computes the MMSE beamformer, no additional computation complexity for TAS is required. From Figure 7, we observe that the pilot-based channel estimation has excellent performance in comparison with the ideal performance and our proposed simple TAS yields an impressive gain of close to 3 dB at an average SER of 10^{-4}.

We have also investigated the feasibility of bidirectional relaying for OFDM/OFDMA-based 4G cellular standards such as IEEE 802.16e [13]. The IEEE 802.16e system is based on OFDMA physical layer for both uplink and downlink. Current mobile WiMax standard supports various subchannelization procedures, in both uplink and downlink directions, for data transmission in time (OFDM symbols) and frequency (OFDM subcarriers). One such uplink subchannelization procedure is termed partially utilized subchannelization (PUSC) wherein the modulation symbols of a given user are pseudorandomly spread over the frequency band to extract frequency diversity and to average interference across neighboring cells/sectors. Briefly, one slot in UL-PUSC is defined as 48 modulation symbols spanning over three consecutive OFDM symbols (which is a PUSC slot duration). The modulation symbols together with the pilot symbols needed to estimate the uplink channel are sent over 6 tiles distributed over frequency, where a tile is defined as four consecutive subcarriers over three consecutive OFDM symbols. Each tile contains 4 pilot symbols, placed at the corners of the tile, and 8 data symbols. An FEC block in WiMax comprises of a given number of slots and the maximum FEC block size is a function of the modulation order and channel coding rate. The WiMax standard supports 8 modulation order and coding rate combinations. These are QPSK modulation with

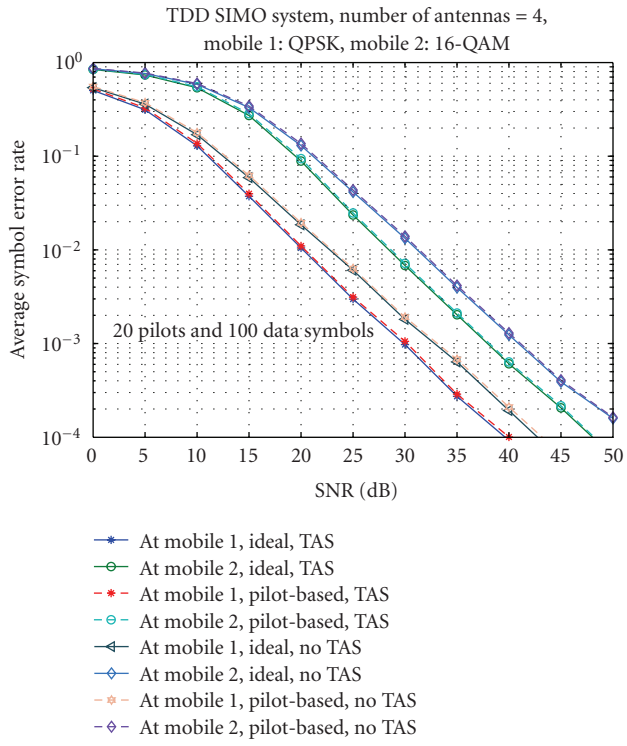


Figure 7: Performance of transmit antenna selection for TDD-based MMSE-BAF system with 4 antennas at the base-station. User-1 employs QPSK modulation whereas user-2 employs 16-QAM modulation. An uncoded system is considered with realistic channel estimation over block fading channels with 20 pilot and 100 data symbols per fading block.

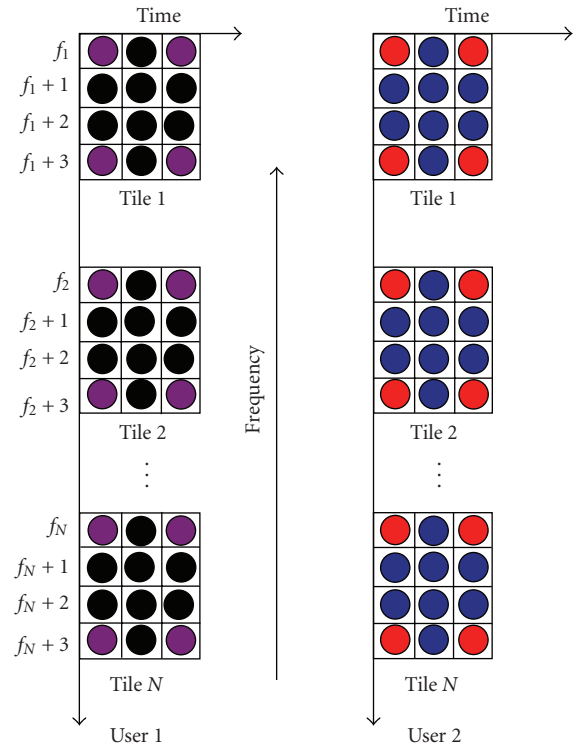


Figure 8: Subchannelization procedure for bidirectional communication using the IEEE 802.16e protocol. In uplink, each user is allocated K slots spanning N_f subchannels in frequency over N_t slot-durations. Each slot-duration comprises 3 consecutive OFDM symbols, whereas each subchannel contains 6 tiles distributed throughout of the useful frequency band. A tile contains 4 subcarries over a slot-duration with the 4 pilots at the corners of the tile and the remaining 8 tones used for data. In the above figure, there are N tiles with $N_t = 1$, $N_f = N/6$, and $K = N_t N_f = N/6$. The circles filled with violet and red colors indicate the pilot tones of users 1 and 2, respectively, whereas the circles filled with black and blue colors indicate the data tones of users 1 and 2, respectively.

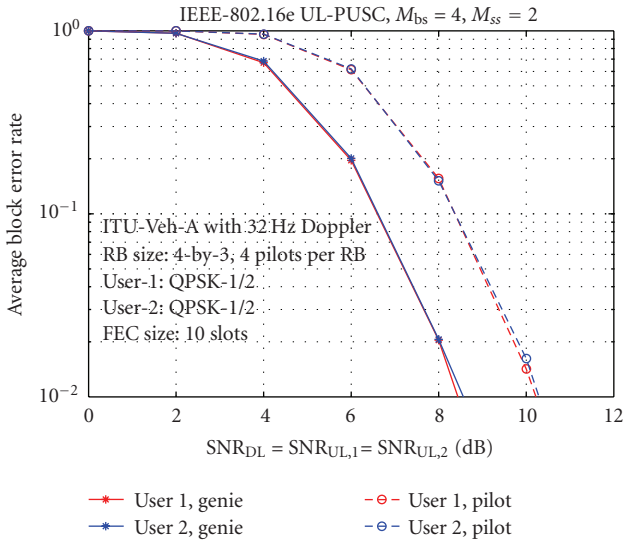
code rates 1/2 and 3/4, 16-QAM modulation with code rates 1/2 and 3/4, and 64-QAM modulation with rates 1/2, 2/3, 3/4, and 5/6. Figure 8 shows the modified UL-PUSC structure to support bidirectional communications. Each user employs Hadamard sequences as pilot symbols to enable the base-station to estimate the individual channels without interference. For downlink transmission, the base-station can use any subchannelization procedure. However, to render our proposal valid for FDD as well as TDD systems, the downlink subchannel structure is set identical to the uplink one and the broadcast pilots from each tile are used for channel estimation at the mobile stations. For efficient cancellation of self-interference, each mobile station requires the knowledge of uplink channel-related parameters β (a positive scalar value) and the two complex scalars values $v_1 := \mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{h}_1$ and $v_2 := \mathbf{w}_{\text{opt}}^{\mathcal{H}}\mathbf{h}_2$. Additional simulation parameters are listed in Table 1.

In Figures 9 and 10, we show the performance of MMSE-BAF when applied to an IEEE 802.16e system in an FDD mode of operation. These simulations are conducted for a base-station with four receive antennas and for a two-antenna mobile station receiver. In these plots, block error rate (BLER) performances of genie-aided perfect channel knowledge are compared against pilot-based realistic channel estimation schemes. For each tile, one channel estimate is obtained by sample averaging the received pilots over

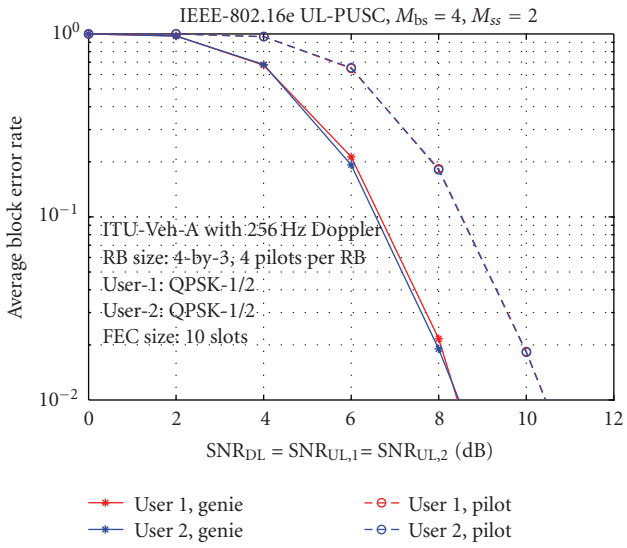
Table 1: Simulation parameters.

Parameter	Value
Bandwidth	10 MHz
Sampling rate	11.2 Msps
FFT size	1024
Subcarrier spacing	10.9375 KHz
Useful symbol duration (T_u)	91.4286 μsec
Cyclic prefix (T_G)	$T_u/8$
Useful subcarriers	840
Left guard subcarriers	92
Right guard subcarriers	91
Channel coding	Convolutional turbo coding (with 8 iterations)
Carrier frequency	2.0 GHz

that tile. Knowledge of neither the fading statistics nor the delay/Doppler spread is assumed at the receivers. In Figure 9, both users are assumed to encode their data using

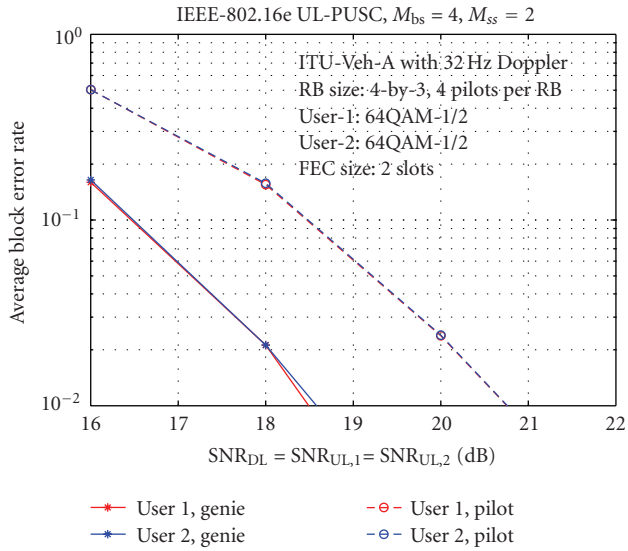


(a) With 32 Hz Doppler

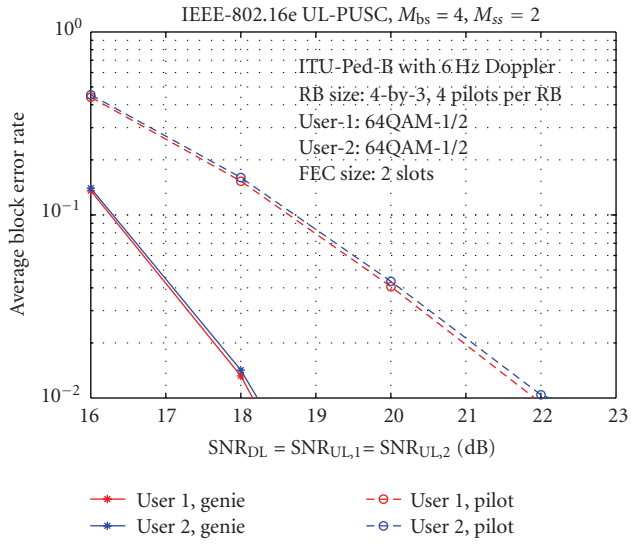


(b) With 256 Hz Doppler

Figure 9: Performance of MMSE-BAF on MIMO-OFDMA-based IEEE 802.16e system using uplink partial utilization of subchannelization (PUSC) permutation. Each user employs rate-1/2 convolutional turbo coding (CTC) with QPSK modulation. The FEC block length is 480 information bits which corresponds to 10 slots, with 48 modulation symbols per slot, as per the terminology in [13].



(a) ITU-Veh-A channel with 32 Hz Doppler



(b) ITU-Ped-B channel with 6 Hz Doppler

Figure 10: Performance of MMSE-BAF on MIMO-OFDMA-based IEEE 802.16e system using uplink partial utilization of subchannelization (PUSC) permutation. Each user employs rate-1/2 convolutional turbo coding (CTC) with 64QAM modulation. The FEC block length is 288 information bits which corresponds to 2 slots, with 48 modulation symbols per slot, as per the terminology in [13].

QPSK modulation with rate-1/2 convolutional turbo coding (CTC). For this MCS, the FEC block size is set to the maximum allowed, which is equal to 10 slots or $10 \times 48 \times 2 \times 1/2 = 480$ information symbols. We have considered ITU Vehicular-A channel model with both low and high Doppler spread values of 32 Hz and 256 Hz, respectively. Figure 9 shows that, under both high and low Doppler scenarios, pilot channel estimation incurs a loss of about 2 dB, and MMSE-BAF works extremely well in supporting data exchanges in high-mobile environments. Figure 10 shows the performance of MMSE-BAF with 64-QAM modulation and a CTC with code rate of 1/2. With this MCS level, supporting a per-user over-the-air spectral efficiency of 3 bits/sec/Hz, we have employed the maximum possible FEC block size of 2 slots, or $2 \times 48 \times 6 \times 1/2 = 288$ information bits. Figure 10(a) shows the block error performance over ITU-Vehicular-A channel with 32 Hz Doppler, whereas Figure 10(b) shows the performance on a Pedestrian-B channel with 6 Hz Doppler. Due to higher-frequency-selectivity of Pedestrian-B channel, compared with the Vehicular-A channel where the

channel estimation-based BLER performance (at 1 percent BLER) is about 2 dB away from the ideal performance, the simple sample-average-based channel estimation has a performance degradation of about 3.5 dB compared to the performance with perfect channel knowledge. It is expected that the BLER performance can be significantly improved by incorporating a more complex two-dimensional channel estimation scheme, such as Wiener filtering, which requires knowledge of fading statistics as well as delay and Doppler spread information.

5. Conclusion

In this paper, we have introduced a so-called MMSE-BAF protocol for bidirectional communications over two-way relay channels with multiantenna relay nodes. The features of this protocol include the usage of analog network coding at the relay node and the evaluation of a receive weight vector for the relay node using a joint linear MMSE filtering operation on the received uplink multiple access signal. Transmit antenna selection using the largest MMSE weight branch on the downlink is also an inherent feature of the proposed protocol in the TDD mode of operation. Extensive link-level simulations have been proposed for both TDD and FDD modes of operations and required modifications to the existing IEEE 802.16e standard have been proposed to accommodate the MMSE-BAF protocol. It has been shown through simulation results that the MMSE-BAF protocol is a simple yet efficient solution to the problem of bidirectional communications in two-way relay channels with multiantenna relays and half-duplex nodes.

Appendix

Received SNRs with Noise Whitening

We first rewrite the noise covariance matrices $\mathbf{\Sigma}_1$ (10) and $\mathbf{\Sigma}_2$ (20), using their eigenvalue decompositions, as

$$\begin{aligned}\mathbf{\Sigma}_1 &= \mathbf{P}_1^{\mathcal{H}}\mathbf{\Lambda}_1\mathbf{P}_1,\\ \mathbf{\Sigma}_2 &= \mathbf{P}_2^{\mathcal{H}}\mathbf{\Lambda}_2\mathbf{P}_2,\end{aligned} \tag{A.1}$$

where $\mathbf{P}_1$ and $\mathbf{P}_2$ are unitary matrices and $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are diagonal matrices containing the eigenvalues of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, respectively.

Let us now focus on demodulating x_2 from $\mathbf{r}_1$ by whitening the noise $\tilde{\mathbf{n}}_1$. Let

$$\begin{aligned}\mathbf{s}_1 &= \mathbf{\Lambda}_1^{-1/2}\mathbf{P}_1\mathbf{r}_1\\ &= \mathbf{\Lambda}_1^{-1/2}\mathbf{P}_1\left(\mathbf{g}_1\beta v_2 x_2 + \tilde{\mathbf{n}}_1\right).\end{aligned} \tag{A.2}$$

Since, conditioned on $\mathbf{\Lambda}_1$ and $\mathbf{P}$,

$$\mathrm{E}\left[\left(\mathbf{\Lambda}_1^{-1/2}\mathbf{P}_1\tilde{\mathbf{n}}_1\right)\left(\mathbf{\Lambda}_1^{-1/2}\mathbf{P}_1\tilde{\mathbf{n}}_1\right)^{\mathcal{H}}\right] = \mathbf{I}_{M_{ss}}, \tag{A.3}$$

it follows from (A.2) that the instantaneous received SNR of x_2 by whitening $\mathbf{r}_1$ is simply

$$\begin{aligned}\gamma_2' &= \sigma_2^2\left\|\mathbf{\Lambda}_1^{-1/2}\mathbf{P}_1\mathbf{g}_1\beta v_2\right\|^2\\ &= \sigma_2^2\beta^2|v_2|^2\mathbf{g}_1^{\mathcal{H}}\mathbf{P}_1^{\mathcal{H}}\mathbf{\Lambda}_1^{-1/2}\mathbf{\Lambda}_1^{-1/2}\mathbf{P}_1\mathbf{g}_1\\ &= \sigma_2^2\beta^2|v_2|^2\mathbf{g}_1^{\mathcal{H}}\mathbf{P}_1^{\mathcal{H}}\mathbf{\Sigma}_1^{-1}\mathbf{g}_1.\end{aligned} \tag{A.4}$$

In a similar manner, upon whitening $\mathbf{r}_2$ to demodulate x_1, the instantaneous received SNR of x_1 becomes

$$\gamma_1' = \sigma_1^2\beta^2|v_1|^2\mathbf{g}_2^{\mathcal{H}}\mathbf{\Sigma}_2^{-1}\mathbf{g}_2. \tag{A.5}$$

Using the following matrix inversion lemma (MIL) [14]

$$\left(\mathbf{I}_{M_{ss}} + \mathbf{x}\mathbf{x}^{\mathcal{H}}\right)^{-1} = \mathbf{I}_{M_{ss}} - \frac{\mathbf{x}\mathbf{x}^{\mathcal{H}}}{1+\|\mathbf{x}\|^2}, \tag{A.6}$$

where $\mathbf{x}$ is a column-vector of appropriate size, it is possible to further simplify (A.4) as

$$\begin{aligned}\gamma_2' &= \frac{\sigma_2^2\beta^2|v_2|^2}{\sigma_N^2}\times\mathbf{g}_1^{\mathcal{H}}\left(\mathbf{I}_{M_{ss}} - \frac{\beta^2\left\|\mathbf{w}_{\mathrm{opt}}\right\|^2\mathbf{g}_1\mathbf{g}_1^{\mathcal{H}}}{1+\beta^2\left\|\mathbf{w}_{\mathrm{opt}}\right\|^2\|\mathbf{g}_1\|^2}\right)\mathbf{g}_1\\ &= \frac{\sigma_2^2\beta^2|v_2|^2}{\sigma_N^2}\times\frac{\|\mathbf{g}_1\|^2}{1+\beta^2\left\|\mathbf{w}_{\mathrm{opt}}\right\|^2\|\mathbf{g}_1\|^2}.\end{aligned} \tag{A.7}$$

In a similar manner, application of MIL in (A.5) leads to

$$\gamma_1' = \frac{\sigma_1^2\beta^2|v_1|^2}{\sigma_N^2}\times\frac{\|\mathbf{g}_2\|^2}{1+\beta^2\left\|\mathbf{w}_{\mathrm{opt}}\right\|^2\|\mathbf{g}_2\|^2}. \tag{A.8}$$

As we mentioned earlier, expressions for γ_2' (A.7) and γ_1' (A.8) are, respectively, identical to SNRs γ_2 (21) and γ_1 (18) derived in Section 3.2 without performing noise whitening. This somewhat counterintuitive result leads us to conclude that SNR improvement is not an option with noise whitening when employing the MMSE-BAF protocol.

References

[1] Y. Wu, P. A. Chou, and S.-Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," in *Proceedings of the 39th Annual Conference on information Sciences and Systems (CISS '05)*, March 2005.

[2] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog netwok coding," Tech. Rep. MIT-CSAIL-TR-2007-012, Computer Science and Artificial Intelligence Laboratory, Cambridge, Mass, USA, February 2007.

[3] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: physical-layer network coding," in *Proceedings of the 12th Annual International Conference on Mobile Computing and Networking (MOBICOM '06)*, pp. 358–365, Los Angeles, Calif, USA, September 2006.

[4] S. J. Kim, P. Mitran, and V. Tarokh, "Performance bounds for bidirectional coded cooperation protocols," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5235–5241, 2008.

[5] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions for bi-directional relaying," http://www.citebase.org/abstract?id=oai:arXiv.org:0808.0954.

[6] S. J. Kim, N. Devroye, and V. Tarokh, "Bi-directional half-duplex protocols with multiple relays," http://www.citebase.org/abstract?id=oai:arXiv.org:0810.1268.

[7] P. Popovski and H. Yomo, "The anti-packets can increase the achievable throughput of a wireless multi-hop network," in *Proceedings of the IEEE International Conference on Communications (ICC '06)*, vol. 9, pp. 3885–3890, Istanbul, Turkey, June 2006.

[8] P. Popovski and H. Yomo, "Wireless network coding by amplify-and-forward for bi-directional traffic flows," *IEEE Communications Letters*, vol. 11, no. 1, pp. 16–18, 2007.

[9] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 707–712, Glasgow, Scotland, June 2007.

[10] I. Hammerström, M. Kuhn, C. Eşli, J. Zhao, A. Wittneben, and G. Bauch, "MIMO two-way relaying with transmit CSI at the relay," in *Proceedings of the 8th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC '07)*, pp. 1–5, Helsinki, Finland, June 2007.

[11] T. Unger and A. Klein, "Duplex schemes in multiple antenna two-hop relaying," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 128592, 14 pages, 2008.

[12] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 5, pp. 699–712, 2009.

[13] "IEEE Standard for Local and Metropolitan Area Networks, Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands," IEEE Std. 802.16-2005, 2005.

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

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 251719, 17 pages
doi:10.1155/2010/251719

Research Article

Near Optimum Power Control and Precoding under Fairness Constraints in Network MIMO Systems

Gábor Fodor,[1] Mikael Johansson,[2] and Pablo Soldati[2]

[1] *Ericsson Research, 164 80 Stockholm, Sweden*
[2] *Automatic Control Lab, Royal Institute of Technology, 100 44 Stockholm, Sweden*

Correspondence should be addressed to Gábor Fodor, gabor.fodor@ericsson.com

Received 1 July 2009; Accepted 25 October 2009

Academic Editor: Hongxiang Li

We consider the problem of setting the uplink signal-to-noise-and-interference (SINR) target and allocating transmit powers for mobile stations in multicell spatial multiplexing wireless systems. Our aim is twofold: to evaluate the potential of such mechanisms in network multiple input multiple output (MIMO) systems, and to develop scalable numerical schemes that allow real-time near-optimal resource allocation across multiple sites. We formulate two versions of the SINR target and power allocation problem: one for maximizing the sum rate subject to power constraints, and one for minimizing the total power needed to meet a sum-rate target. To evaluate the potential of our approach, we perform a semianalytical study in *Mathematica* using the augmented Lagrangian penalty function method. We find that the gain of the joint optimum SINR setting and power allocation may be significant depending on the degree of fairness that we impose. We develop a numerical technique, based on successive convexification, for real-time optimization of SINR targets and transmit powers. We benchmark our procedure against the globally optimal solution and demonstrate consistently strong performance in realistic network MIMO scenarios. Finally, we study the impact of near optimal precoding in a multicell MIMO environment and find that precoding helps to reduce the sum transmit power while meeting a capacity target.

1. Introduction

Recently, several works proposed and demonstrated various forms of tight network coordination as a means to provide high spectral efficiency in multicell multiple input multiple output (MIMO) cellular networks [1, 2]. Such coordination among the cells deployed over a certain geographical area has initially aimed at coordinating transmitter and receiver algorithms [3–5]. These promising results have triggered the interest of standardization bodies and industry players to investigate the architecture and protocol aspects of *network MIMO* systems employing multisite coordination of signal transmission and reception [6, 7].

Since network MIMO systems in general and multicell spatial multiplexing systems in particular inherently support the exchange of control information among multiple base stations, they can readily benefit of joint radio resource management functions, such as multicell scheduling, power control and precoding [8–10]. Multicell scheduling is concerned with assigning radio resources to users in multiple cells such that some utility function is maximized; see, for instance, [11, 12]. Multicell power control can be viewed as a finer granularity control which is concerned with allocating power to scheduled users. Specifically for the uplink, it has been shown that coordinated power control can minimize the overall transmit power so as to maintain a *predetermined* signal-to-noise-and-interference (SINR) target [13].

For multicell scenarios, Hande et al. have demonstrated significant advantages of optimizing the SINR targets according to some criterion set by the network operator [14]. That work used SINR expressions for single input single output (SISO) systems without spatial multiplexing and considered network utility maximization problems with fair user utility functions $u_i(\cdot)$, in the sense that if the SINR tends to zero, $u_i \to -\infty$. If proportionally fair rate allocations are desired, the SINR targets could also be set via optimal

distributed power control algorithms; see, for example, [15, 16]. However, none of these methods are easily extended to throughput maximization problems.

Similarly, we expect that setting the SINR targets in multiple network MIMO cells is an efficient means to control fairness and multicell throughput performance. Unlike traditional cellular networks, network MIMO systems allow the adjustment of these SINR targets on a time scale that is similar to scheduling and power control. Roughly speaking, manipulating the SINR targets can be seen as an extension of network MIMO power control algorithms (in the spirit of [14]) that adjust the individual power levels taking into account the channel variations such that predetermined SINR targets are reached and the overall power is minimized.

Building on this key observation, the first contribution of this paper is to evaluate the potential of such mechanisms in multicell spatial multiplexing systems. To this end, we develop a model that jointly optimizes the SINR targets and the power levels for the uplink of spatially multiplexing network MIMO systems. Our model can explicitly take into account fairness constraints by requiring that the ratio of the individual SINR (and thereby rate) targets for different mobile stations remain under some prescribed value, ranging from greedy throughput maximization ("no fairness") to equal rate allocation. Secondly, we develop scalable numerical schemes that allow real-time near-optimal resource allocation across multiple sites. Our major finding is that the new degree of freedom in network MIMO systems (i.e., multicell SINR target control) is an efficient tool to control the throughput performance and fairness in multicell systems. Finally, we find that when mobile stations employ fast, channel aware precoding, they can either significantly reduce their transmission power while maintaining a capacity target or enhance the throughput for a fixed power budget.

The paper is structured as follows. Section 2 models the uplink transmission of a network MIMO system employing minimum mean square error (MMSE) receiver. Section 3 formulates the SINR setting and power allocation problem. Sections 4 and 5 present a semianalytical and numerical solution approaches, respectively. Section 7 discusses numerical results and Section 8 concludes the paper.

2. System Model

2.1. Modelling the Received Signal. In order to establish the received signal model, we revise and merge the models of [13, 17, 18]. We consider the uplink transmission of a multicell system with K cells and assume that each cell consists of a base station (BS) with N_r being receive antennas and an active mobile station (MS) with N_t being transmit antennas and spatial multiplexing. The assumption of having a single MS in each cell is not limiting, because it includes time division, (orthogonal) frequency division and orthogonal code division, multiplexing systems which ensure (time, frequency, or code domain) orthogonality within a single cell [13]. A narrow-band quasistatic flat-fading channel is assumed, where the channel remains constant within several scheduling instances (frames).

The received signal at the kth BS is represented as

$$\mathbf{y}_k = \alpha_{k,k}\mathbf{H}_{k,k}\mathbf{T}_k\mathbf{x}_k + \sum_{j\neq k}\alpha_{k,j}\mathbf{H}_{k,j}\mathbf{T}_j\mathbf{x}_j + \mathbf{n}_k, \tag{1}$$

where

(i) $\alpha_{k,j} = \sqrt{P_j d_{k,j}^{-\rho}\chi_{k,j}/N_t}$ is a scalar coefficient depending on the total transmit power P_j for user j, the log-normal shadow fading $\chi_{k,j}$, and distance $d_{k,j}$ between the kth base station and the jth user with path loss exponent ρ;

(ii) $\mathbf{x}_k \in \mathbb{C}^{N_t\times 1}$ is the data vector that is assumed to be zero-mean, normalized, and uncorrelated, $\mathbb{E}(\mathbf{x}_k\mathbf{x}_k^\dagger) = \mathbf{I}_{N_t}$;

(iii) $\mathbf{H}_{k,j}$ denotes the $(N_r \times N_t)$ channel transfer matrix;

(iv) $\mathbf{T}_k$ is the MS-k $(N_t \times N_t)$ diagonal power loading matrix; to keep the total transmit power constant, $\mathbf{T}_k$ must satisfy

$$\mathbb{E}\left(\mathbf{T}_k^\dagger\mathbf{T}_k\right) = \operatorname{trace}\left(\mathbf{T}_k\mathbf{T}_k^\dagger\right) = \sum_{i=1}^{\mathrm{N_t}}\left|\mathbf{T}_k^{(i,i)}\right|^2 = N_t \quad \forall k; \tag{2}$$

(v) $\mathbf{n}_k$ is a $N_r \times 1$ additive white Gaussian noise vector at the kth base station with zero mean and covariance matrix $\mathbf{R}_{n_k} = \mathbb{E}(\mathbf{n}_k\mathbf{n}_k^\dagger) = \sigma_n^2\mathbf{I}_{N_r}$ for all k.

We note that the underlying assumption of the last bullet item on equal noise covariance for all base stations is reasonable for a set of base stations with the same antenna configuration and other physical and hardware characteristics within a limited geographical region and is often used both in the literature and in standardization [6, 7, 13, 19].

We rewrite the signal model (1) in a compact form as

$$\mathbf{y}_k = \alpha_{k,k}\mathbf{H}_{k,k}\mathbf{T}_k\mathbf{x}_k + \mathbf{z}_k + \mathbf{n}_k, \tag{3}$$

where $\mathbf{z}_k = \sum_{j\neq k}\alpha_{k,j}\mathbf{H}_{k,j}\mathbf{T}_j\mathbf{x}_j$ denotes the $(N_r \times 1)$ interference vector from users in other cells, with covariance matrix

$$\mathbf{R}_{z_k} = \mathbb{E}\left(\mathbf{z}_k\mathbf{z}_k^\dagger\right) = \sum_{j\neq k}\alpha_{k,j}^2\mathbf{H}_{k,j}\mathbf{T}_j\mathbf{T}_j^\dagger\mathbf{H}_{k,j}^\dagger. \tag{4}$$

For ease of notation, we define an *equivalent* noise vector that accounts both intercell interference and background noise

$$\mathbf{v}_k = \mathbf{z}_k + \mathbf{n}_k. \tag{5}$$

It is easy to show that $\mathbf{v}_k$ is zero-mean with covariance $\mathbf{R}_{v_k} = \mathbf{R}_{z_k} + \mathbf{R}_{n_k}$; see Appendix A.

2.2. MMSE Receiver Error Matrix and the Effective SINR. As we shall see, calculating the error matrix of the specific receiver that we employ in our system is a prerequisite for calculating the SINR. In this work we assume that the received signal is filtered through a linear MMSE receiver with weighting matrix $\mathbf{G}_k$ to obtain the following estimate:

$$\hat{\mathbf{x}}_k = \mathbf{G}_k\mathbf{y}_k. \tag{6}$$

Proposition 1. *For the linear MMSE receiver*

$$\mathbf{G}_k = \frac{1}{\alpha_{k,k}}\mathbf{T}_k^{\dagger}\mathbf{H}_{k,k}^{\dagger} \times \left(\mathbf{H}_{k,k}\mathbf{T}_k\mathbf{T}_k^{\dagger}\mathbf{H}_{k,k}^{\dagger} + \sum_{j\neq k}\frac{\alpha_{k,j}^2}{\alpha_{k,k}^2}\mathbf{H}_{k,j}\mathbf{T}_j\mathbf{T}_j^{\dagger}\mathbf{H}_{k,j}^{\dagger} + \frac{\sigma_n^2}{\alpha_{k,k}^2}\mathbf{I}_{N_r}\right)^{-1}. \tag{7}$$

In the special case of equal power distribution, that is, $\mathbf{T}_k = \mathbf{I}_{N_t}$*, the MMSE weighting matrix becomes*

$$\mathbf{G}_k = \frac{1}{\alpha_{k,k}}\mathbf{H}_{k,k}^{\dagger}\left(\mathbf{H}_{k,k}\mathbf{H}_{k,k}^{\dagger} + \sum_{j\neq k}\frac{\alpha_{k,j}^2}{\alpha_{k,k}^2}\mathbf{H}_{k,j}\mathbf{H}_{k,j}^{\dagger} + \frac{\sigma_n^2}{\alpha_{k,k}^2}\mathbf{I}_{N_r}\right)^{-1}. \tag{8}$$

Proof. See Appendix A. □

The $(N_t \times N_r)$ linear MMSE weighting matrix $\mathbf{G}_k$ can be expressed in an alternative, more compact, form as

$$\begin{aligned}\mathbf{G}_k &= \frac{1}{\alpha_{k,k}}\mathbf{T}_k^{\dagger}\mathbf{H}_{k,k}^{\dagger}\left(\mathbf{H}_{k,k}\mathbf{T}_k\mathbf{T}_k^{\dagger}\mathbf{H}_{k,k}^{\dagger} + \frac{1}{\alpha_{k,k}^2}\mathbf{R}_{\nu_k}\right)^{-1} \\ &= \left(\mathbf{I} + \mathbf{T}_k^{\dagger}\mathbf{R}_{H_k}\mathbf{T}_k\right)^{-1}\alpha_{k,k}\mathbf{T}_k^{\dagger}\mathbf{H}_{k,k}^{\dagger}\mathbf{R}_{\nu_k}^{-1},\end{aligned} \tag{9}$$

where $\mathbf{R}_{H_k} = \alpha_{k,k}^2\mathbf{H}_{k,k}^{\dagger}\mathbf{R}_{\nu_k}^{-1}\mathbf{H}_{k,k}$; see, for example, [20, Chapter 12]. To derive the streamwise SINRs at base station k, we will need the diagonal elements of the error matrix of the MMSE filtered signal. To this end, the following proposition is useful.

Proposition 2. *The MMSE estimation error matrix* $(N_r \times N_r)$ *for the kth base station is*

$$\mathbf{E}_k = \left(\alpha_{k,k}\mathbf{G}_k\mathbf{H}_{k,k}\mathbf{T}_k - \mathbf{I}\right)\left(\alpha_{k,k}\mathbf{T}_k^{\dagger}\mathbf{H}_{k,k}^{\dagger}\mathbf{G}_k^{\dagger} - \mathbf{I}\right) + \mathbf{G}_k\mathbf{R}_{\nu_k}\mathbf{G}_k^{\dagger}, \tag{10}$$

or, equivalently

$$\mathbf{E}_k = \left(\mathbf{I} + \mathbf{T}_k^{\dagger}\mathbf{R}_{H_k}\mathbf{T}_k\right)^{-1}. \tag{11}$$

Proof. The computation is derived in Appendix A. □

Note that with equal power distribution, that is, $\mathbf{T}_k = \mathbf{I}_{N_t}$ these results reduce to [13, Appendix A].

We are now in the position to calculate the SINR for the signal model (3) assuming a linear MMSE receiver. Using the linear MMSE weighting matrix $\mathbf{G}_k$, the MSE and SINR expressions can be rewritten, respectively, as

$$\begin{aligned}\mathrm{MSE}_{k,s} &\triangleq (\mathbf{E}_k)_{(s,s)} = \left\{\left(I + \mathbf{T}_k^{\dagger}\mathbf{R}_{H_k}\mathbf{T}_k^{\dagger}\right)^{-1}\right\}_{(s,s)}, \\ \gamma_{k,s} &\triangleq \frac{1}{\mathrm{MSE}_{k,s}} - 1.\end{aligned} \tag{12}$$

2.3. Summary. In this section we defined the multicell MIMO-received signal model (3) and, assuming a linear MMSE receiver, derived the associated effective SINR ($\gamma_{k,s}$) for each stream of the received signal. Equations (12) are important because they capture the dependence of the SINRs on the transmission powers of the own MS and the interfering MSs through the $\mathbf{R}_{H_k}$'s and the $\mathbf{R}_{\nu_k}$'s. Thus, these relations serve as the basis for the the optimization problems of the next section.

3. Problem Formulation

Our aim is to develop a mathematical framework for systematic optimization of SINR-targets, transmit powers, and precoding matrix to maximize a rate objective subject to power budget and fairness constraints (or to minimize power subject to rate constraints). To the best of the authors' knowledge, there are not efficient means for jointly optimizing all these variables. We build our theoretical developments on the following result from [13]: by assuming equal power allocation for all streams s (i.e., no uplink beam forming, $\mathbf{T}_k = \mathbf{I}_{N_t}$ for all k), the minimum stream SINR is lower bounded as

$$\min_{s\in[1,N_t]}\gamma_{k,s} \geq \underline{\gamma}_k(\mathbf{p}), \tag{13}$$

where $\mathbf{p} = (P_1 \cdots P_K)^T$ is the power allocation vector, and

$$\underline{\gamma}_k(\mathbf{p}) = \frac{P_k d_{k,k}^{-\rho}\chi_{k,k}}{\sum_{j\neq k}P_j d_{k,j}^{-\rho}\chi_{k,j}\mu_{\max}\left(\mathbf{\Omega}_{k,j,1}\right) + N_t\sigma_k^2\mu_{\max}\left(\mathbf{\Omega}_{k,j,2}\right)}. \tag{14}$$

Here, $\mu_{\max}(\cdot)$ is the maximum eigenvalue operator for a Hermitian matrix, while $\mathbf{\Omega}_{k,j,1}$ and $\mathbf{\Omega}_{k,j,2}$ are defined as

$$\begin{aligned}\mathbf{\Omega}_{k,j,1} &= \left(\mathbf{H}_{k,k}^{\dagger}\mathbf{H}_{k,k}\right)^{-1}\mathbf{H}_{k,k}^{\dagger}\mathbf{H}_{k,j}\mathbf{H}_{k,j}^{\dagger}\mathbf{H}_{k,k}\left(\mathbf{H}_{k,k}^{\dagger}\mathbf{H}_{k,k}\right)^{-1}, \\ \mathbf{\Omega}_{k,j,2} &= \left(\mathbf{H}_{k,k}^{\dagger}\mathbf{H}_{k,k}\right)^{-1}.\end{aligned} \tag{15}$$

This bound allows to associate a single SINR value

$$\gamma_k(\mathbf{p}) \triangleq \min_{s\in[1,N_t]}\gamma_{k,s} \tag{16}$$

with each MS-k. In what follows, we will search for SINR targets γ_k^{tgt} which are feasible for the lower-bound (and hence for each individual stream) and let $\mathbf{\Gamma} = \mathrm{diag}(\gamma_1^{\mathrm{tgt}} \cdots \gamma_K^{\mathrm{tgt}})$.

3.1. Minimizing Sum Power under Fixed SINR Target. The above result was used in [13] to design power control schemes which maintain a fixed minimum SINR target γ_k^{tgt} for every stream s by enforcing $\underline{\gamma}_k(\mathbf{p}) \geq \gamma_k^{\mathrm{tgt}}$ for each user. As shown in [13], the transmit power of MS-k must satisfy

$$P_k \geq \gamma_k^{\mathrm{tgt}} \cdot \left(\frac{\sum_{j\neq k}P_j \cdot d_{k,j}^{-\rho}\chi_{k,j}\mu_{\max}\left(\mathbf{\Omega}_{k,j,1}\right)}{d_{k,k}^{-\rho}\chi_{k,k}} + \frac{\sigma_n^2 N_t\mu_{\max}\left(\mathbf{\Omega}_{k,j,2}\right)}{d_{k,k}^{-\rho}\chi_{k,k}}\right) \tag{17}$$

Moreover, the power vector that satisfies this requirement and minimizes the sum power is

$$\mathbf{p}^{\star} = (\mathbf{I} - \boldsymbol{\Gamma}\mathbf{F})^{-1}\boldsymbol{\Gamma}\mathbf{n}, \tag{18}$$

where $\mathbf{n}$ is a K-dimensional *effective* noise variance vector whose kth element is $n_k = N_t\sigma_n^2\mu_{\max}(\boldsymbol{\Omega}_{k,j,2})/d_{k,k}^{-\rho}\chi_{k,k}$, and

$$\mathbf{F}_{k,j} = \begin{cases} \dfrac{d_{k,j}^{-\rho}\chi_{k,j}\mu_{\max}\left(\boldsymbol{\Omega}_{k,j,1}\right)}{d_{k,k}^{-\rho}\chi_{k,k}}, & k \neq j, \\ 0, & k = j. \end{cases} \tag{19}$$

Still, (18) requires feasibility of the SINR targets, which in practice cannot be guaranteed a priori. Precoding optimization was shown to be effective to balance the conservativeness of the bound (13) and increase feasibility; see [13].

3.2. Optimal SINR Target Selection. In this paper we take a step further and explore a key observation, not fully exploited in [13]. Since the minimum user-stream SINR bound (13) allows to associate a single SINR target per user, one can regard each MS-BS connection as an equivalent SISO system and model the minimum user-stream capacity as function of the power allocation with a Shannon-like expression (normalized to the bandwidth) as

$$c_k\left(\gamma_k^{\text{tgt}}\right) = \log_2\left(1 + \gamma_k^{\text{tgt}}\right) \quad \forall k, \tag{20}$$

where we enforce

$$\underline{\gamma}_k(\mathbf{p}) = \frac{P_k}{n_k + \sum_{j\neq k} G_{kj}P_j} \geq \gamma_k^{\text{tgt}} \quad \forall k \tag{21}$$

with $\mathbf{G} = \mathbf{I} + \mathbf{F}$. This observation is the basis for optimizing the minimum user-stream SINR targets.

In network MIMO, it is possible (and as we shall see beneficial) to exploit the possibility to set the SINR targets such that the sum power is kept at a minimum level and the overall system capacity (sum rate) target c_{m} is reached. This problem is formulated as follows:

$$\begin{aligned} \underset{\boldsymbol{\Gamma},\mathbf{p}}{\text{minimize}} \quad & \sum_k P_k \\ \text{subject to} \quad & \sum_k c_k\left(\gamma_k^{\text{tgt}}\right) \geq c_{\text{m}} \\ & \gamma_k^{\text{tgt}} = \underline{\gamma}_k(\mathbf{p}) \quad \forall k, \end{aligned} \tag{22}$$

in the optimization variables $\boldsymbol{\Gamma}$ (SINR targets) and $\mathbf{p}$ (power). We are also interested in the dual formulation of problem (22), that is, maximizing the multicell capacity (sum rate) subject to a total power budget:

$$\begin{aligned} \underset{\boldsymbol{\Gamma},\mathbf{p}}{\text{maximize}} \quad & \sum_k c_k\left(\gamma_k^{\text{tgt}}\right) \\ \text{subject to} \quad & \sum_k P_k \leq P_{\text{tot}} \\ & \gamma_k^{\text{tgt}} = \underline{\gamma}_k(\mathbf{p}) \quad \forall k. \end{aligned} \tag{23}$$

3.3. Enforcing Fairness Constraints. Fairness can be enforced in the above formulations by limiting the ratio between SINR targets, that is,

$$\gamma_k^{\text{tgt}} \leq \Phi_{kj}\gamma_j^{\text{tgt}} \quad \forall k, j \neq k. \tag{24}$$

The matrix $\boldsymbol{\Phi}$ collects the fairness ratios. These constraints are written more compactly as $\mathbf{a}(\boldsymbol{\Gamma}) \preceq \mathbf{b}(\boldsymbol{\Gamma})$, where $\mathbf{a} = \text{vec}(\mathbf{a}_1 \cdots \mathbf{a}_k)$ with $\mathbf{a}_k = (\mathbf{1} - \mathbf{e}_k)\gamma_k^{\text{tgt}}$, and $\mathbf{b}(\boldsymbol{\Gamma}) = \text{vec}(\boldsymbol{\Phi}\boldsymbol{\Gamma})$.(Here, $\mathbf{e}_k$ is the vector with 1 in the kth coordinate and 0's elsewhere). To account for fairness constraints, we include the inequalities $\mathbf{a}(\boldsymbol{\Gamma}) \preceq \mathbf{b}(\boldsymbol{\Gamma})$ in (22) and (23). In what follows, we develop a novel efficient SINR-target optimization procedure and combine it with iterative algorithms for power and precoding matrix optimization. As we will see, the minimum SINR bound (13) is quite conservative and including the precoding matrix $\mathbf{T}_k$ in the optimization is instrumental to enhance the performance.

4. A Semianalytical Solution Approach

We propose to solve the problems formulated in Section 3.2 through the augmented Lagrangian penalty function method [21]. In this method, the constrained nonlinear optimization task is transformed into an unconstrained problem by adding a penalty term to the Lagrangian function as follows:

$$\begin{aligned} L(\boldsymbol{\Gamma}, \mathbf{p}, \boldsymbol{\nu}, \mu, \varepsilon) = {} & \sum_{\text{k}} P_k + \mu\left(\sum_{\text{k}} c_k\left(\gamma_k^{\text{tgt}}\right) - c_{\text{m}}\right) + \boldsymbol{\nu}^T(\mathbf{a} - \mathbf{b}) \\ & + \varepsilon\left[\left(\sum_{\text{k}} c_k\left(\gamma_k^{\text{tgt}}\right) - c_{\text{m}}\right)^2 + \sum_{\text{n}}(a_n - b_n)^2\right]. \end{aligned} \tag{25}$$

Here, we present the method for the power minimization problem (22). The Lagrangian for problem (23) follows similarly. It can be shown that if the optimum Lagrange multipliers are known, the solution to this unconstrained problem corresponds to the solution of the original problem (22) regardless of the value of the penalty parameter ε, see, for example, [21, Chapter 9].

4.1. Solution of the Power Minimization Problem. For ease of presentation, we consider a three cell system, that is $K = 3$. First, we need to find the power vector as the function of the target multicell capacity (sum rate) c_{m} and the individual SINR targets (the γ_i^{tgt}'s):

$$\mathbf{p}^{\star}\left(c_{\text{m}}, \gamma_1^{\text{tgt}}, \gamma_2^{\text{tgt}}, \gamma_3^{\text{tgt}}\right) = \begin{bmatrix} \dfrac{M_{11} + M_{12} + M_{13}}{D_p} \\ \dfrac{M_{21} + M_{22} + M_{23}}{D_p} \\ \dfrac{M_{31} + M_{32} + M_{33}}{D_p} \end{bmatrix}, \tag{26}$$

where the parameters $M_{11}, \ldots, M_{33}$ and D_p are given in Appendix B. From the capacity constraint, it follows that

$(K-1)$ SINR values can be freely selected while the Kth SINR target value must be chosen such that the capacity constraint is fulfilled. In the case of $K = 3$,

$$\gamma_3^{\text{tgt}}\left(c_{\text{m}}, \gamma_1^{\text{tgt}}, \gamma_2^{\text{tgt}}\right) = e^{c_{\text{m}} - \log(1+\gamma_1^{\text{tgt}}) - \log(1+\gamma_2^{\text{tgt}})} - 1. \tag{27}$$

Using this relationship, the M_{ij} parameters are expressed as the functions of γ_1^{tgt} and γ_2^{tgt} (see Appendix B). That is, for a specific capacity target c_{m}, $\mathbf{p}^\star$ and the sum of its components are expressed as a two-variable function of γ_1^{tgt} and γ_2^{tgt}. Using (26), it is straightforward to find the stationary points of the unconstrained problem and, by establishing the second-order necessary conditions, to find the local optimum solutions (that is the local minimum points) of (22). In our *Mathematica* implementation, we found that in all considered practically relevant examples, a simple heuristic can then easily identify the global optimum solution (see also Section 7).

4.2. Solution of the Capacity Maximization Problem. For the capacity maximization case, we can freely choose $(K-1)$ SINR values, while the Kth SINR target needs to be selected to fulfil the constraint of (23). With an abuse of notation, let $P^\star = \sum_i P_i^\star$ denote the sum of the components of $\mathbf{p}^\star$ in (18). For the three cell system, summing the components of (18), it is straightforward to show that setting the SINR targets γ_1^{tgt} and γ_2^{tgt} implies setting γ_3^{tgt} as follows:

$$\begin{aligned}\gamma_3^{\text{tgt}}&\left(P^\star, \gamma_1^{\text{tgt}}, \gamma_2^{\text{tgt}}\right)\\ &\leq \frac{P^\star\left(1 - F_{12}F_{21}\gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}}\right)}{D_c}\\ &\quad - \frac{\gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}}(n_1F_{21} + n_2F_{12}) - n_1\gamma_1^{\text{tgt}} - n_2\gamma_2^{\text{tgt}}}{D_c},\end{aligned} \tag{28}$$

where D_c is

$$\begin{aligned}D_c = {}& P^\star\left(F_{13}F_{31}\gamma_1^{\text{tgt}} + F_{23}F_{32}\gamma_2^{\text{tgt}} + F_{12}F_{23}F_{31}\gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}}\right.\\ &\left.+F_{21}F_{13}F_{32}\gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}}\right)\\ &+\gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}}(\,n_1(F_{23}F_{31} - F_{23}F_{32} + F_{21}F_{32})\\ &\quad + n_2(F_{13}F_{32} - F_{13}F_{31} + F_{12}F_{31})\\ &\quad +n_3(F_{13}F_{21} + F_{12}F_{23} - F_{12}F_{21}))\\ &+\gamma_1^{\text{tgt}}(n_1F_{31} + n_3F_{13}) + \gamma_2^{\text{tgt}}(n_2F_{32} + n_3F_{23}).\end{aligned} \tag{29}$$

Similarly to the minimum $P^\star$ in problem (22), for a specific total power budget P_{tot}, (28) allows us to express the sum-rate $\sum c_k(\gamma_k^{\text{tgt}})$ as a two-variable function of γ_1^{tgt} and γ_2^{tgt}, which allows to derive the numerical results.

5. Scalable Near-Optimal SINR Target Setting

To design more scalable solutions and avoid the matrix inversion in (18), we make use of the model (20)-(21) to reformulate problem (23) as

$$\begin{aligned}\underset{\mathbf{p},\mathbf{r}}{\text{maximize}} \quad & \sum_k r_k\\ \text{subject to} \quad & r_k \leq c_k(\mathbf{p}) \quad \forall k\\ & \sum_k P_k \leq P_{\text{tot}}.\end{aligned} \tag{30}$$

This problem optimizes the minimum user-stream transmission rates $\mathbf{r}$ and powers $\mathbf{p}$, hence implicitly the minimum user-stream SINR. Similarly to the formulation (23), problem (30) is not convex due to the link rate constraints $r_k \leq c_k(\mathbf{p})$.

5.1. Monotonic Optimization. Through an exponential transform of the variables $P_k \leftarrow e^{\tilde{P}_k}$ and $r_k \leftarrow e^{\tilde{r}_k}$ and a log-transformation of the constraints, we rewrite problem (30) as

$$\begin{aligned}\underset{\tilde{\mathbf{p}},\tilde{\mathbf{r}}}{\text{maximize}} \quad & \sum_k e^{\tilde{r}_k}\\ \text{subject to} \quad & \tilde{r}_k \leq \log\left(c_k\left(e^{\tilde{\mathbf{p}}}\right)\right) \quad \forall k\\ & \sum_k e^{\tilde{P}_k} \leq P_{\text{tot}}.\end{aligned} \tag{31}$$

Since the objective function is convex and monotonically increasing in the variables $\tilde{\mathbf{r}}$ and the feasibility set is convex, problem (31) falls into the family of *monotonic optimization*, for which, unlike standard convex optimization problems, local optimality does not translate into global optimality. Only recently, Qian et al. [22] have shown the equivalence between the formulations (30) and (31) and have devised an algorithm, MAPEL, that finds the global optimum solution by constructing a series of polyblocks that approximate the SINR region with increasing precision (see [22] for details).

Proposition 3. *The MAPEL algorithm converges to the global optimal solution of problem* (31).

Proof. It follows analogously to [22, Theorem 2] by defining the feasibility set as

$$\mathcal{F} = \left\{\mathbf{p} \mid 0 \leq P_k \leq P_{\text{tot}},\ \sum_k P_k \leq P_{\text{tot}} f_i(\mathbf{p}) g_i(\mathbf{p})^{-1} \geq 1\right\}, \tag{32}$$

where $f_i(\mathbf{p})$ and $g_i(\mathbf{p})$ are defined as in [22]. In our case, we combine implicit peak-power constraints (i.e., $0 \leq P_k \leq P_{\text{tot}}$) with an explicit global power budget. □

The MAPEL algorithm allows to trade-off between accuracy and convergence time by tuning an approximation factor δ. Since the computation times drastically increase with

increasing accuracy and problem size, MAPEL is currently not feasible for real-time SINR target setting. Nevertheless, it is an excellent candidate for off-line benchmarking of the low-complexity schemes that we will develop next.

5.2. An Approximation of the Link Rate Constraint. To reduce problem complexity, we use an approximation to "convexify" the problem. Inspired by [15, 23], we use the relation $\theta\log(x)+\beta \le \log(1+x)$ with $\theta = x_0/(1+x_0)$ and $\beta = \log(1+x_0)-\theta\log(x_0)$ to approximate the link capacity. The approximation becomes exact for $x = x_0$. We replace the expression (20) with a more conservative one:

$$\widetilde{c}_k(\mathbf{p}) = \left[\theta_k \log_2\left(\underline{\gamma}_k(\mathbf{p})\right)+\beta_k\right] \le c_k(\mathbf{p}). \tag{33}$$

By applying the approximation (33) to the stream rate constraints of problem (30), we obtain the following approximation of problem (30)

$$\begin{aligned} \underset{\boldsymbol{\Gamma},\mathbf{p}}{\text{maximize}} \quad & \sum_k \left[\theta_k^{(t)} \log_2\left(\gamma_k^{\text{tgt}}\right)+\beta_k^{(t)}\right] \\ \text{subject to} \quad & \gamma_k^{\text{tgt}} \le \frac{P_k}{n_k+\sum_{j\neq k} G_{kj}P_j} \quad \forall k \\ & \sum_k P_k \le P_{\text{tot}}, \end{aligned} \tag{34}$$

which explicitly optimizes the SINR targets $\boldsymbol{\Gamma}$ and the transmit power $\mathbf{p}$. Here, the SINR expression (21) has been added to the constraint set to provide an explicit relationship between these variables. Similarly to [15], we propose to solve problem (30) through a sequence of convex approximations according to the iterative Algorithm 1. At the tth iteration of the algorithm, the following problem $\mathcal{P}^{(t)}$ is solved:

$$\begin{aligned} \underset{\widetilde{\boldsymbol{\Gamma}},\widetilde{\mathbf{p}}}{\text{maximize}} \quad & \sum_k \left[\theta_k^{(t)} \widetilde{\gamma}_k^{\text{tgt}}+\beta_k^{(t)}\right] \\ \text{subject to} \quad & \widetilde{\gamma}_k^{\text{tgt}} \le \widetilde{P}_k - \log\left(n_k+\sum_{j\neq k} G_{kj} e^{\widetilde{P}_j}\right) \quad \forall k \\ & \sum_k e^{\widetilde{P}_k} \le P_{\text{tot}}. \end{aligned} \tag{35}$$

The above formulation is obtained from problem (34) through the exponential change of variables $\gamma_k^{\text{tgt}} \leftarrow e^{\widetilde{\gamma}_k^{\text{tgt}}}$, $P_k \leftarrow e^{\widetilde{P}_k}$, and a log-transform of the constraints. Algorithm 1 iteratively solves the convex approximate problems $\{\mathcal{P}^{(t)}\}_t$ in the variables $\widetilde{\boldsymbol{\Gamma}}$ and $\widetilde{\mathbf{p}}$ and appropriately tunes θ and β to improve the objective function until convergence.

While the approximation was initially proposed in [23] to tune the transmission power in DSL systems and then applied in [15] to network utility maximization problems with concave utilities, our formulation is used to optimize the SINR targets. We prove that each iteration of Algorithm 1 consists of a convex problem (Proposition 4) and that the sequence of solutions is convergent (Proposition 5), which follow quite straightforward from [15]. In addition, we demonstrate that the sequence converges to a solution that satisfies the KKT optimality condition of both the monotonic optimization (31) (Theorem 5.1) and the original nonconvex problem (30) (Theorem 5.2).

Initialize $\mathbf{p}^{(0)}$, $\theta^{(0)}$, and $\beta^{(0)}$ to some feasible values for the original problem (30).
Start with iteration step $t = 1$.
(1) Solve the approximate optimization problem (35). Let $\{\boldsymbol{\Gamma}^{(t)}, \mathbf{p}^{(t)}\}$ denote the solution of the tth iteration.
(2) Update $\boldsymbol{\theta}^{(t+1)}$, $\boldsymbol{\beta}^{(t+1)}$ at $\mathbf{x}_0^{(t)} = \underline{\gamma}(\mathbf{p}^{(t)})$.
(3) Update $t = t+1$ and repeat until convergence

Algorithm 1: Series of convex approximations.

Proposition 4. *The approximating problem $\mathcal{P}^{(t)}$ is convex.*

Proof. The constraints contain a linear term in $\widetilde{\gamma}_k$ and $\widetilde{P}_k$ and a convex term (log-sum-exp) in $\widetilde{\mathbf{p}}$. The power budget is convex (sum-exp), and the objective is linear in $\widetilde{\gamma}_k$. □

Proposition 5. *The problem sequence $\{\mathcal{P}^{(t)}\}_t$ results in a series of monotonically improving objective values. The sequence always converges at which point the lower bound approximation* (33) *becomes exact.*

Proof. The proof details can be found in Appendix C. □

Theorem 5.1. *The problem sequence $\{\mathcal{P}^{(t)}\}_t$ converges to a KKT-point of the monotonic optimization problem* (31).

Proof. The proof follows from Proposition 5 and a direct inspection of the KKT optimality conditions for problems (C.1) and (31). The details can be found in Appendix C. □

Theorem 5.2. *The problem sequence $\{\mathcal{P}^{(t)}\}_t$ converges to a KKT-point of the original nonconvex problem* (30).

Proof. See Appendix C. □

6. Precoding Optimization

The mathematical framework devised in the previous sections allows to optimally select the SINR targets under equal power allocation for all streams (i.e., $\mathbf{T}_k = \mathbf{I}_{N_t}$) for two classes of problems: problem (22) minimizes the sum power while maintaining a fixed system capacity; problem (23) maximizes the multicell capacity subject to a fixed power budget. Both cases use the minimum per-stream SINR bound (13), that is,

$$\gamma_k(\mathbf{p}) \ge \frac{P_k d_{k,k}^{-\rho}\chi_{k,k}}{\sum_{j\neq k} P_j d_{k,j}^{-\rho}\chi_{k,j}\mu_{\max}\left(\boldsymbol{\Omega}_{k,j,1}\right)+N_t\sigma_k^2\mu_{\max}\left(\boldsymbol{\Omega}_{k,j,2}\right)}, \tag{36}$$

to formulate the SINR and power allocation problem. Originally proposed in [13, Lemma 1], this bound applies

Given $t = 0$, $\epsilon^{(0)} = 1$, P_{tot}, ε_{gap} and $\mathbf{T}_k^{(0)} = \mathbf{I}_{N_t}$ for all k.
Initialize SINR targets $\boldsymbol{\Gamma}^{(0)} = \text{diag}(\gamma_k^{\text{tgt}})$ and transmission powers $\mathbf{p}^{(0)}$ solving either problem (22) or (23).

Repeat:

(1) $t = t + 1$.

(2) For $k = 1$ to K

(a) Given $\{\mathbf{T}^{(t-1)}, \mathbf{P}^{(t-1)}\}$, compute the inter ference $\zeta_{k,s} = f(\mathbf{T}^{(t-1)}, \mathbf{P}^{(t-1)})$ as in (38).

(b) Calculate the optimum loading matrix $\mathbf{T}_k^{(t)}$ as

$$\left(\mathbf{T}_k^{(t)}\right)^{(s,s)} = \sqrt{\frac{\zeta_{k,s} N_t}{\sum_{j=1}^{N_t} \zeta_{k,j}}} \quad \forall s \in [1, N_t].$$

(c) Calculate the new SINR targets $\boldsymbol{\Gamma}^{(t)}$ and update the optimum transmit power $P_k^{(t)}$ as

$$\gamma_k^{(t)} = \gamma_k^{\text{tgt}} \cdot \epsilon^{(t-1)} \quad \forall k$$

$$P_k^{(t)} = \frac{\zeta_{k,s}}{\left|\left(\mathbf{T}_k^{(t)}\right)^{(ss)}\right|^2} (\gamma_k^{(t)} + 1) \quad \forall k, s$$

(3) Update the control parameter ϵ:

(a) If objective is power minimization: $\epsilon^{(t)} = \epsilon^{(t-1)}$;

(b) If objective is throughput maximization:

$$\epsilon^{(t)} = \left\{\epsilon^{(t-1)} - \kappa\left(\sum_k P_k^{(t)} - P_{tot}\right)\right\}^+.$$

Until $|P_k^{(t)} - P_k^{(t-1)}| \le \varepsilon_{\text{gap}}, \quad \forall k.$

Algorithm 2: Iterative SINR and precoding optimization.

to the minimum postprocessing user SINR with linear MMSE receiver with equal power allocation for all streams s. Unfortunately, the Rayleigh-Ritz Theorem [24] used in [13, Lemma 1] does not apply when the precoding matrix $\mathbf{T}_k$ is also included in the optimization framework. In what follows, we ask whether precoding optimization can bring an additional gain in network MIMO systems where the SINR targets are optimized based on this bound (without precoding).

From the signal model (1), when user kth uses a diagonal power loading matrix $\mathbf{T}_k \in C^{N_t \times N_t}$ with $\sum_{s=1}^{N_t} \left|\mathbf{T}_k^{(s,s)}\right|^2 = N_t$, the postprocessing SINR of its sth stream becomes

$$\gamma_{k,s} = \frac{P_k \left|\mathbf{T}_k^{(s,s)}\right|^2}{\zeta_{k,s}} - 1, \tag{37}$$

where

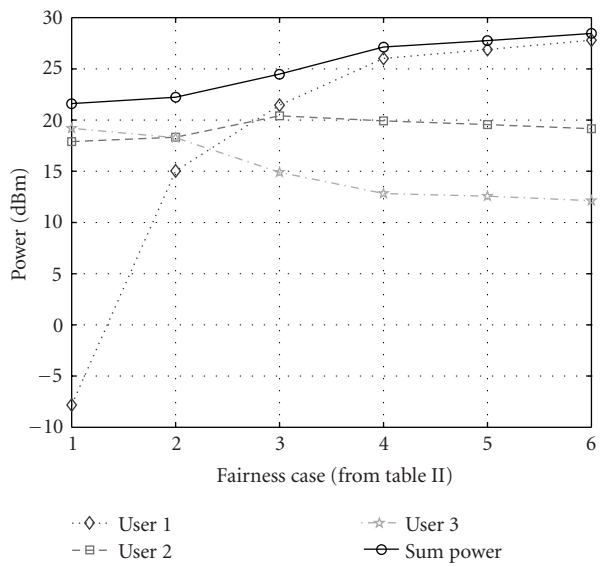
$$\zeta_{k,s} = \left\{\left(d_{k,k}^{-\rho} \chi_{k,k} \mathbf{H}_{k,k}^{\dagger} \left(\sum_{j \neq k} P_j d_{k,j}^{-\rho} \chi_{k,j} \mathbf{H}_{k,j} \mathbf{T}_j \mathbf{T}_j^{\dagger} \mathbf{H}_{k,j}^{\dagger} + N_t \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{H}_{k,k} + \frac{1}{P_k} \mathbf{I}\right)^{-1}\right\}^{(s,s)} \tag{38}$$

denotes the effective interference after MMSE processing. In [13], a heuristic algorithm for distributing the transmit power over different streams was presented. By inverting (37) for fixed SINR targets, the algorithm finds a near optimal (sum power minimizing) precoding matrix for uplink transmission. Precoding optimization is shown to enhance the feasibility space of a rough SINR targets selection with respect to the equal power allocation case. By applying this algorithm to our optimized SINR targets for problem (22), the total sum power can be reduced further. However, some modifications are necessary for the capacity maximization problem. At the optimal point of problem (23), the SINR targets will consume the entire power budget P_{tot}. In this case, by better distributing the power budget P_{tot}, precoding optimization allows to sustain higher SINR targets, thus yielding a throughput gain.

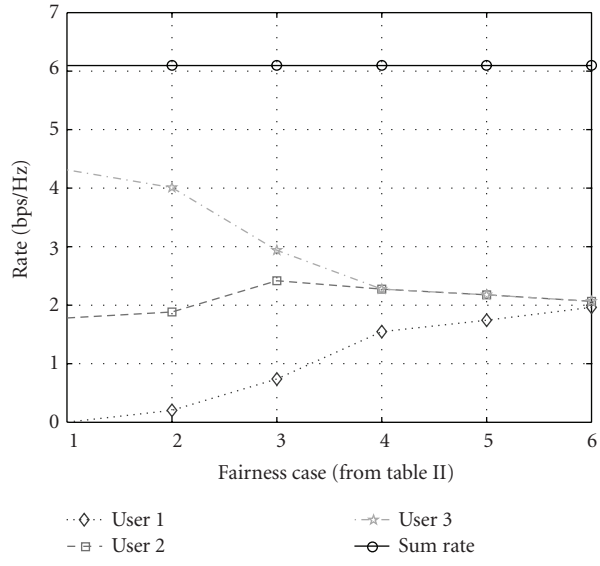
To capitalize on these gains, we modify the algorithm in [13] as in Algorithm 2. The SINR targets are initialized to the optimal values $\boldsymbol{\Gamma} = \text{diag}(\gamma_k^{\text{tgt}})$ yielded by either problem (22) or (23) without precoding, that is, with $\mathbf{T}_k = \mathbf{I}_{N_t}$ for all k. For sum-rate maximization, Algorithm 2 iteratively tunes the SINR targets, along with the precoding matrix and the transmission powers, until the entire power budget P_{tot} is spent. At every iteration, the effective interference and the new precoding matrix are computed as in steps (a) and (b), respectively; the control parameter $\epsilon^{(t)}$ is used to update the SINR targets as

$$\gamma_k^{(t)} = \gamma_k^{\text{tgt}} \cdot \epsilon^{(t)} \quad \forall k, \tag{39}$$

which become the new reference for the power control update in step (c). Finally, $\epsilon^{(t)}$ is tuned differently for the two problems in step (3): for problem (22), $\epsilon^{(t)}$ is kept



(a) The individual power levels and the sum power for the six cases we study. The sum power is significantly lower (21.6 dBm) without fairness than with fairness (≈ 28.45 dBm). In fact, in a real system Cases 4–6 would hardly be feasible for User-3 due to the typical power limitation of MS (≈ 24 dBm)



(b) The individual rates in the six fairness cases that we study. The sum rate is kept constant (2 bps/Hz/cell), but this sum rate is "distributed" unequally (Case 1) or nearly equally (Case 6) in the different cases

Figure 1: Sum power minimization subject capacity (sum rate) and fairness constraint.

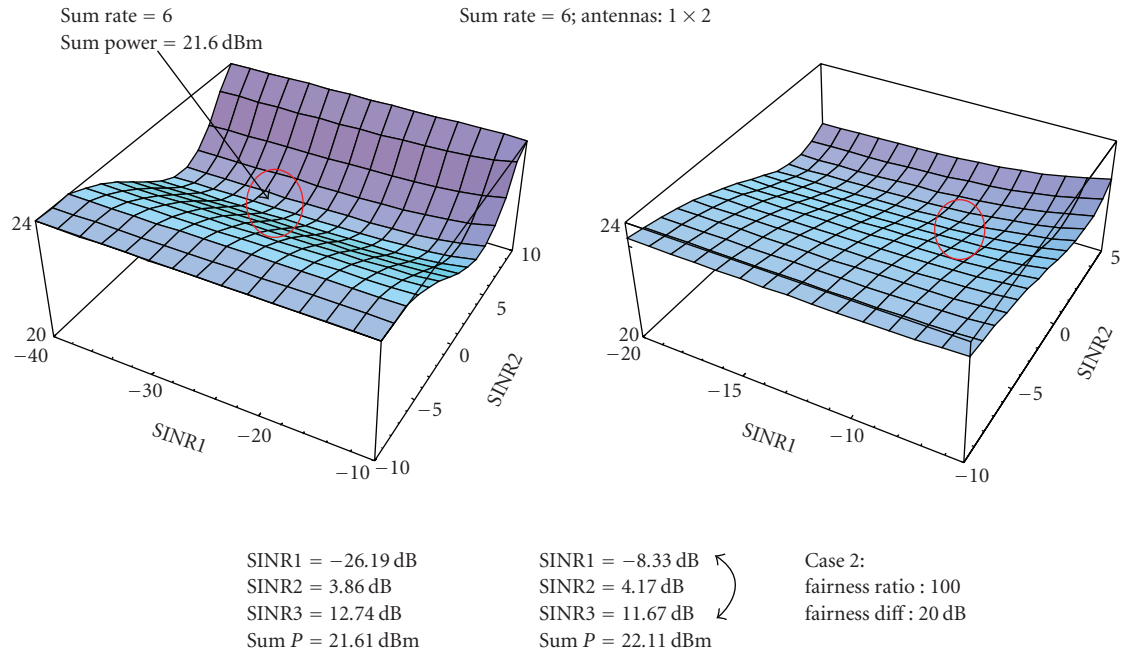
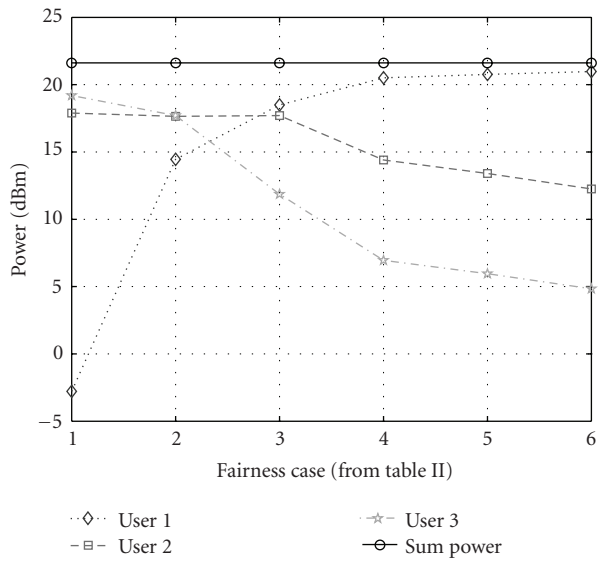


Figure 2: Sum power minimization. The figure shows the sum power as the function of SINR1 and SINR2 without fairness constraints (left) and for Case 2, when the ratios between the lowest and highest SINR must not be greater than 100. In this case the sum power is somewhat higher (22.11 dBm) than without fairness constraint (21.61 dBm).

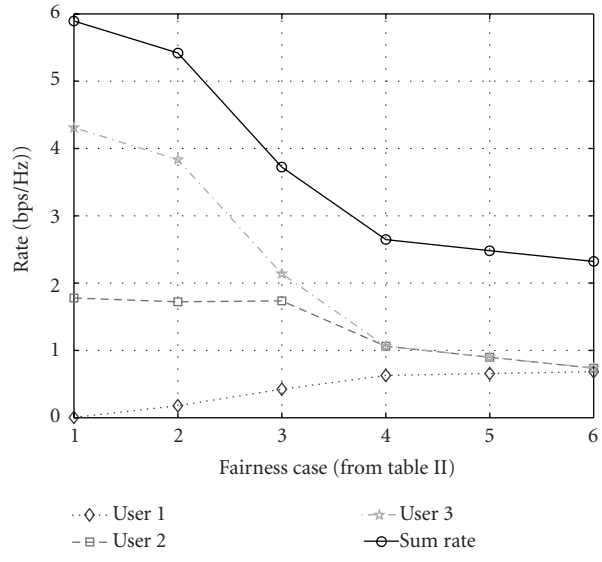
fixed to 1 so that the SINR targets remain unchanged, thus reflecting the original algorithm in [13]; for problem (23), $\epsilon^{(t)}$ is tuned with a subgradient-like step until the power budget constraint is met with equality (point $\epsilon^{(t)}$ will not change anymore).

7. Numerical Results

In this section we consider a three-cell system, each of which is serving a single MS. In an OFDM cellular network, for example, this setting corresponds to the situation in which



(a) The individual power levels for the six cases we study. The sum power is kept constant (21.6 dBm), but the "distribution" of this power budget changes in the six fairness cases. Note that MS-3 needs to drastically increase its power at the cost of MS-1 and MS-2 to reach similar SINR



(b) The individual rates and the sum rate in the six fairness cases we study. The sum rate decreases as the SINR distribution becomes more fair. The Case 6 sum rate is only about 40% of the highest (no fairness) sum rate

Figure 3: Sum rate maximization subject to power budget and fairness constraints.

a single MS is served on an OFDM resource block and interference is caused by MS served in other cells (i.e., assuming perfect intracell orthogonality). The main parameters of this system are summarized in Table 1. MS-1 is located at the cell edge, while MS-2 and MS-3 are close to their respective serving base stations. Table 2 reports six fairness ratios (in dB) between the best and the worst SINR targets, reflecting increasing fairness constraints from unfair allocation (case 1) to almost egalitarian SINR allocation (case 6).

7.1. Power Minimization under Rate and Fairness Constraints. We have implemented the augmented Lagrangian penalty function method in *Mathematica* [21]. Figures 1–6 are obtained by generating the optimum SINR targets and power allocations for 1000 independent channel realizations. Figures 1 and 2 refer to the power minimization task when the sum rate target is kept constant 6 bps (normalized per Hz) in the 3 cells.

Figure 1(b) illustrates the individual stream rates (hence implicitly the SINR targets) in each fairness case, Case 6 being the "most fair" SINR allocation at the expense of a total power increase at around 28.5 dBm (Figure 1(a)). Figure 1(b) confirms the fairness levels in the six cases in terms of the individual (per MS) rates. Figure 2 shows the sum power as a two-variable function of the SINR targets γ_1^{tgt} and γ_2^{tgt}. The sum power attains its minimum (21.6 dBm = 145 mW) when the SINR targets are set differently corresponding to an unfair rate allocation (Case 1). In Case 2, there is a (loose) constraint on the ratio of SINR targets which can only be fulfilled at a somewhat increased total power level (22.11 dBm).

7.2. Rate Maximization under Power and Fairness Constraints. Figures 3 and 4 show the results for the sum rate maximization task with the power budget set to 21.6 dBm (145 mW). In Figure 4 we see that the maximum sum rate is 6 bps; in other words this point is the same as the minimum power for the previous case. Similarly to the previous case, the unfair Case 1 provides the highest performance and enforcing more fair rate allocations reduces the achievable sum rate (Figures 3(a) and 3(b)). In Figure 3(a) we see that in Case 6, the cell edge user must take a lion share of the overall power budget. As Figure 3(b) clearly shows, the rate increase of the cell edge user is still much less than the rate loss of the cell center user, leading to an overall rate loss as compared with the unfair Case 1.

7.3. Accuracy and Computation Complexity. Next, we evaluate the technique based on the series convexifications described in Algorithm 1 and we compare it with the global optimization algorithm MAPEL. We consider a 5-cell 1×2 MIMO network and we solve the sum-rate maximization without fairness constraints for a set of 50 channel realizations with users placed at fixed positions in the plane (the same in all experiments). The power budget is set to 21.6 dBm. Figure 5(a) compares the CDF of the computation time for MAPEL and the iterative convexification procedure, while Figure 5(b) exhibits the their output in terms of achieved sum rate. For MAPEL, we select an accuracy δ between 0.1 and 0.07 (we found experimentally that $\delta > 0.1$ results in suboptimal points) and force the algorithm to stop after 2 hours. As we can see, the iterative convexification procedure converges to the optimal sum rate in only a couple of seconds while the computation time of MAPEL increases

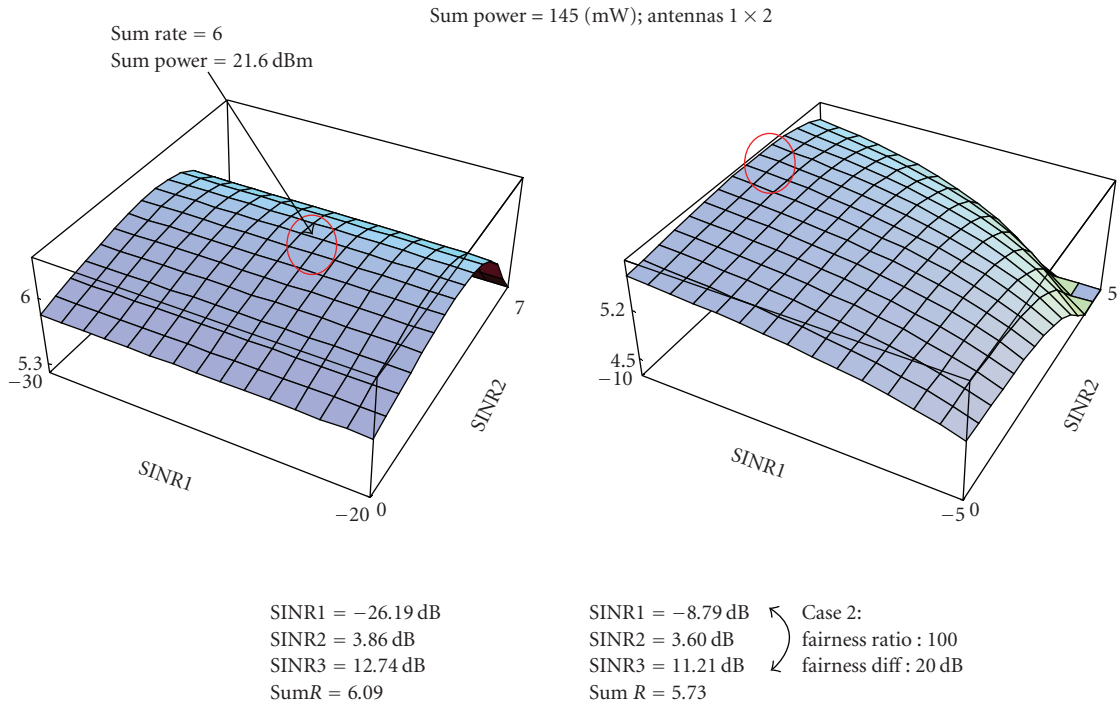
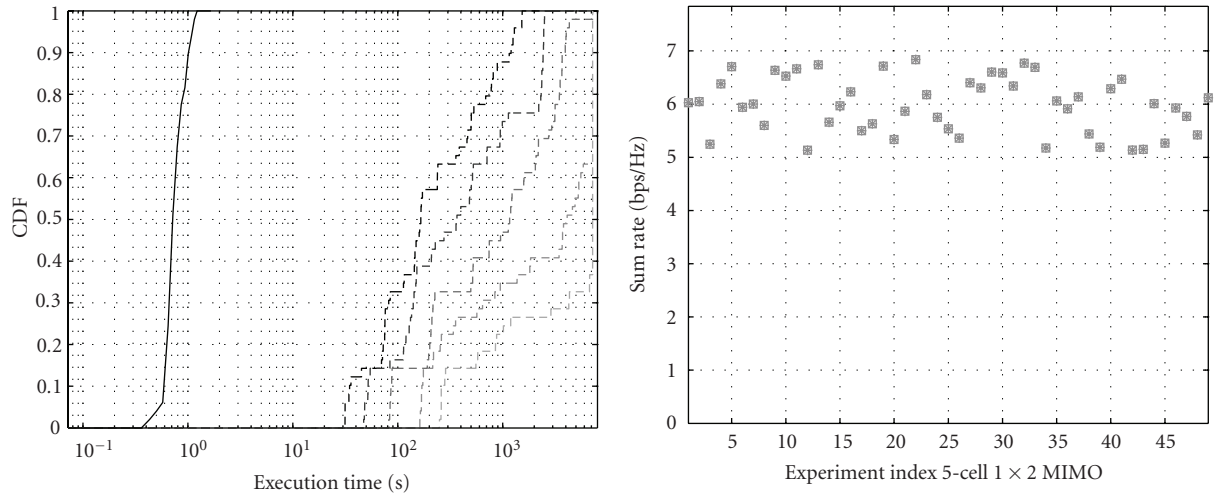


FIGURE 4: Capacity (sum rate) maximization while the total power is kept constant (here 21.6 dBm = 145 mW). Without fairness (Case 1, left) this sum power budget allows to reach 2 bps/Hz/cell (the same circled values as in 2), while in Case 2 (right), the overall sum rate is around 5% lower.



(a) Execution time CDF

(b) Achieved sum rate in each realization

FIGURE 5: The iterative convexification procedure in Algorithm 1 is compared aginst the global optimization algorithm MAPEL in terms of execution time and achieved sum rate for a set of 50 channel realizations in a 5-cell 1×2 MIMO network with power budget set to 21.6 dBm.

exponentially with the required accuracy. Similar results were observed for smaller networks, in which case the MAPEL algorithm takes several minutes to solve, even for the 3-cell scenario, while the iterative convexification procedure runs in less than 100 milliseconds.

7.4. MIMO Gains. Finally, Figures 6 and 7 show the power minimization and the capacity maximization results for the 1×4 SIMO case. Here we observe the impact of the increased receive diversity and array gains. For instance, in the power minimization case, roughly half of the power of that of

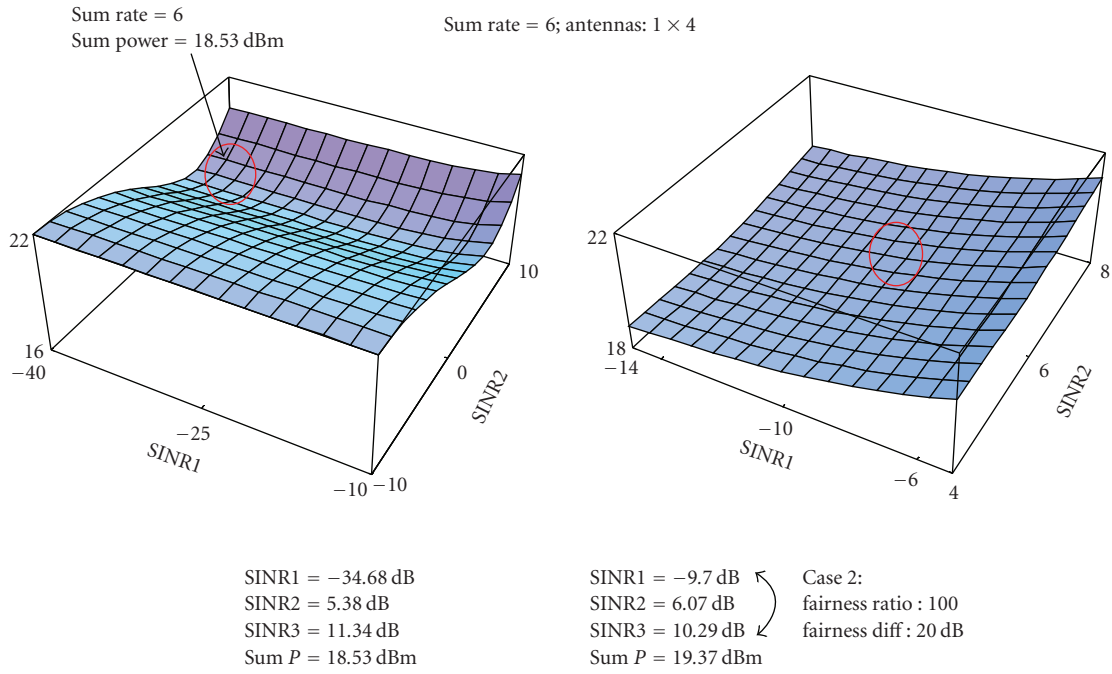


FIGURE 6: Sum power minimization for the 1 × 4 SIMO case. Compared with the 1 × 2 case (Figure 2), we notice that approximately half of the power is required to reach the same sum rate target (18.53 dBm = 71.2 mW).

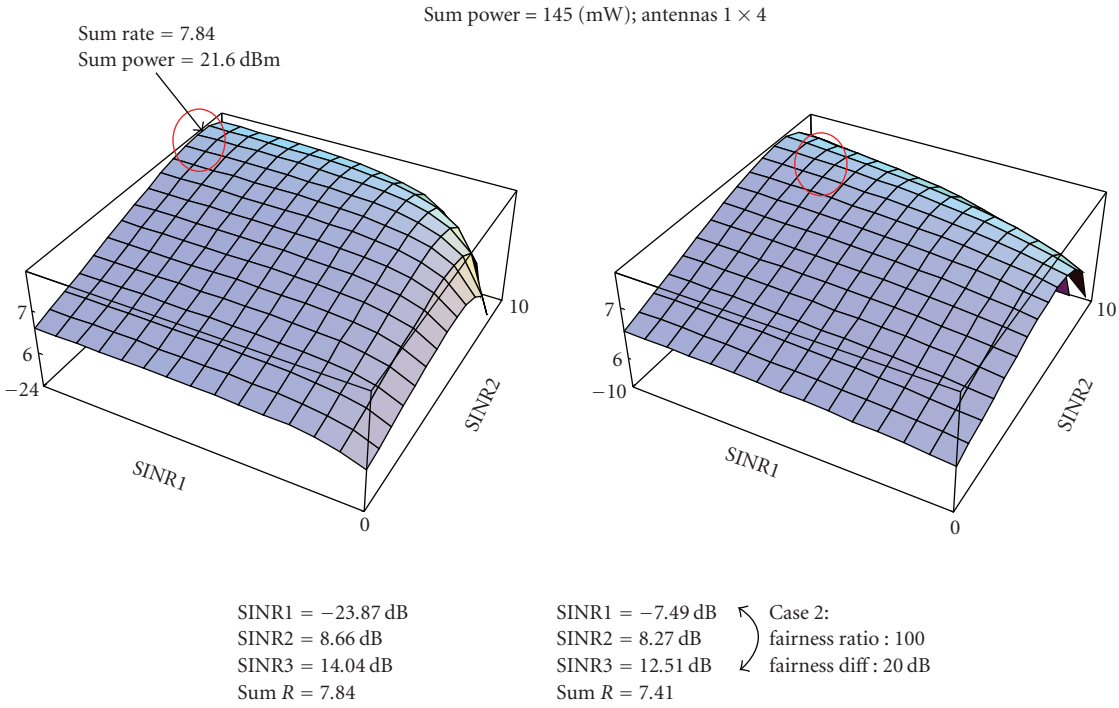


FIGURE 7: Capacity (sum rate) maximization with constant power budget (here 21.6 dBm 145 mW) for the 1 × 4 SIMO case. Compared with the 1 × 2 case (Figure 4), approximately 30% higher sum rate is reached in Case 1.

the 1 × 2 SIMO case is sufficient to maintain similar rate performance. In the rate maximization case, it is possible to increase the 1 × 2 capacity with approximately 30% using the same power budget.

7.5. Precoding Gains. Following the structure of Section 6, we first combine our results on finding the optimal SINR targets for sum-power minimization with the iterative channel inversion power control algorithm of [13]. To this end,

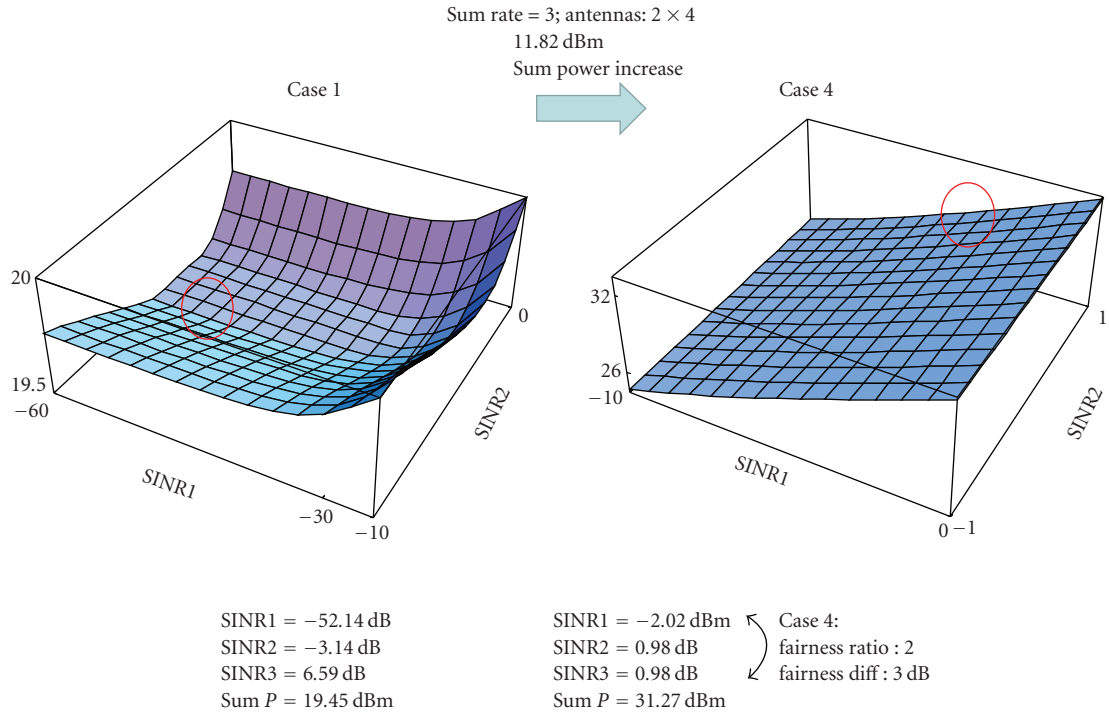
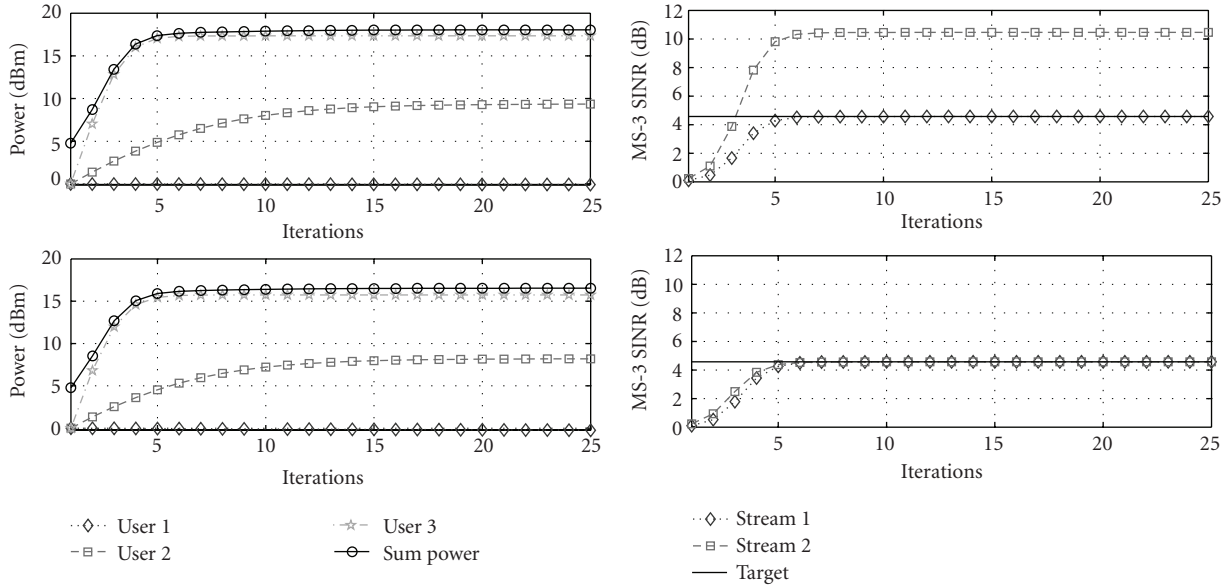


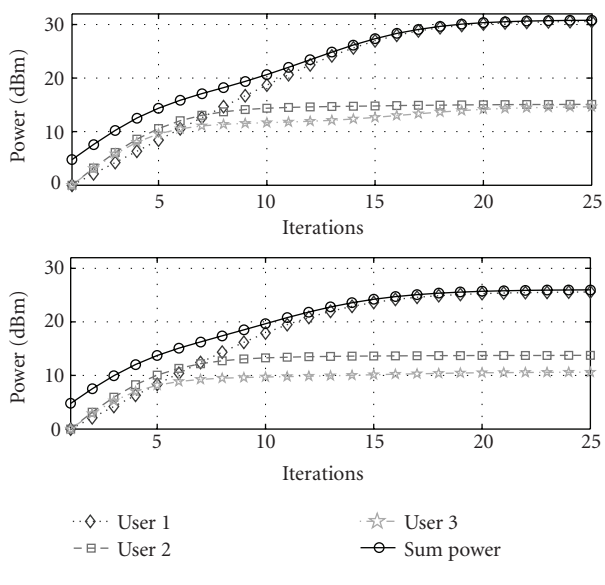
FIGURE 8: Power consumption without precoding in Case 1 and in Case 4. In Case 4, the lowest SINR must be at least half of the highest SINR MS. The point that satisfies this (and the capacity) constraint and minimizes the sum power is indicated on the right-hand side figure. With this constraint, the sum power is much higher than in the unconstraint case. Realistically, this fairness case is not feasible due to typical individual power limitations (24 dBm).



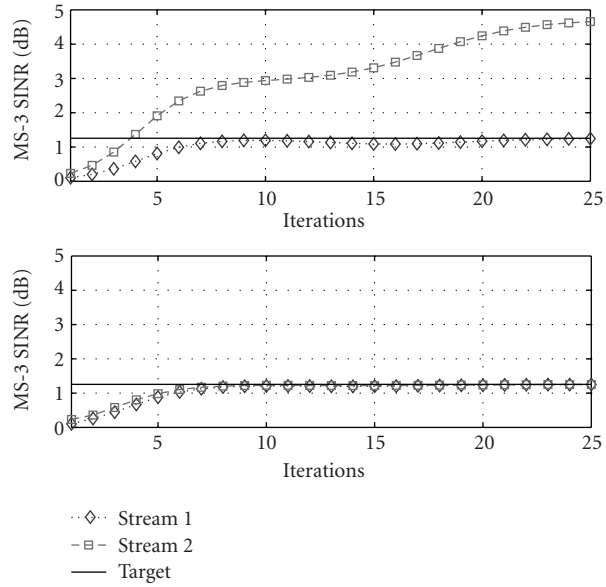
(a) Power consumption without (upper) and with (lower) precoding in Case 1. Precoding has a significant impact on the minimum sum power that reaches the capacity target. The sum power is dominated by MS-3 whose power can be reduced by more than 30% due to precoding gains

(b) SINR evolution for the two streams of MS-3 without (upper) and with (lower) precoding in Case 1. With precoding, the transmit weights keep the stream SINR values "together," while without precoding, one of the SINR values "overshoots"

FIGURE 9: The beneficial impact of transmission precoding in the sum power minimization problem (22) with optimal SINR target selection without fairness.



(a) Power consumption without (upper) and with (lower) precoding in Case 4. Case 4 requires much higher sum power than Case 1 due to the fairness constraints. However, precoding makes even Case 4 feasible, due to the 66% power saving (see lower figure)



(b) SINR evolution for the two streams of MS-3 in Case 4. The figure is similar to Figure , but now precoding helps to reduce the higher SINR streams SINR from around 5 (7 dB) to 1.2 (0.8 db)

Figure 10: The beneficial impact of transmission precoding in the sum power minimization problem (22) with optimal SINR target selection with fairness.

we consider a three-cell 2×4 MIMO system with the input parameters of Table 1 and a capacity target of 3 bps/Hz. For a given channel realization, we use the method of Section 4.1 to find the optimal SINR targets and use them as input values to the iterative channel inversion algorithm of [13] to find the associated power levels and precoding matrices for each mobile station.

The numerical results are shown in Figures 8–10. Figure 8 is obtained by the sum power minimization method of Section 4.1 without precoding. In Case 1, the required sum power is around 19.45 dBm, but in this case the SINR targets are drastically different (left-hand side). In Case 4, the imposed SINR fairness constraint dictates that the weakest SINR should be at least half of the strongest SINR, which leads to a dramatic sum power increase (right-hand side). In fact, this fairness case is not feasible in practical systems in which mobile stations are power limited to around 24 dBm.

Figure 9(a) shows the beneficial impact of using transmission precoding for the unconstrained case (Case 1). The horizontal axis shows the iteration steps of the iterative channel inversion algorithm, while the vertical axes show the individual and sum power levels. Here, the sum power is reduced more than 30% (on the linear scale) due to the power reduction of MS-3 from around 17 dBm to around 15 dBm. This is an impressive precoding gain considering the fact the power levels of Case 1 are low even without precoding thanks to the optimal SINR target setting. This demonstrates a twofold gain of our approach compared to the alternatives in [13]: unlike a rough SINR target selection, we have shown that optimizing the targets is an efficient tool to enhance the throughput and control fairness in network MIMO systems. Furthermore, since the approximation used to derive the optimal SINR targets may lead to high power consumption, optimizing the precoding matrix $\mathbf{T}_k$ based on the given optimal SINR target allows to further reduce system sum-power.

In Figure 9(b) we follow the evolution of the per-stream SINR levels of MS-3 in Case 1 as determined by the iterative channel inversion algorithm. Recall from Figure 8 that the SINR target for MS-3 in Case 1 is 4.5686 (6.6 dB). With equal power allocation, the stream with higher SINR is significantly overallocated (upper figure). This waste of transmission power is eliminated with optimal precoding setting (lower figure) which allows for a lower transmission power of MS-3. We also note that allocating less power for MS-3 reduces the interference caused to neighbor cells, which is a second contributing factor to an overall power decrease. Figure 10(a) compares the power levels without and with precoding for Case 4. Here, the power saving due to optimal precoding is even more pronounced (from 30.8 dBm to 25.6 dBm which is 66% reduction on the linear scale) than in Case 1. By equalizing the SINRs of the two streams as shown in Figure 10(b), the transmission power of MS-3 is drastically reduced, thus making Case 4 becomes feasible.

Finally, we consider the effect of precoding optimization for the second class of problems where the SINR targets are selected to maximize the system throughput for a fixed power budget. Figure 11 illustrates the results for the same three-cell 2×4 MIMO system with the input parameters of Table I and a power budget $P_{\text{tot}} = 21.6$ dBm. From the

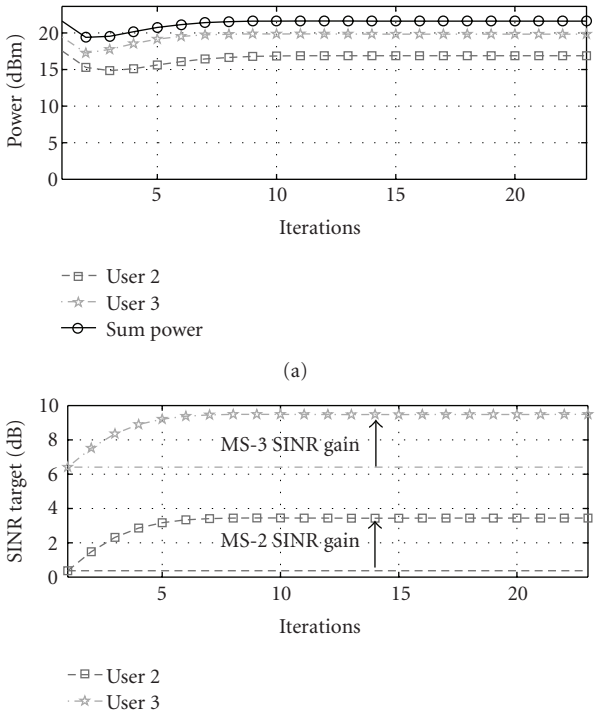


Figure 11: Precoding gain in SINR target selection. The figure shows that for the 2×4 MIMO case with a power budget 21.6 dBM, optimizing the precoder matrices allows to increase the original (without precoding) optimal SINR targets for the same power budget. This gain is more than 3 dB per user.

Table 1: Input parameters of the 3-cell OFDMA system.

Input parameters	
Inter Site Distance [m]	500
Distance of the MS-i's from their serving BS ($i = 1, \ldots, 3$)	0.45, 0.15, and 0.1 [ISD] respectively
Path loss exponent	3.07
Shadow fading	Lognormal; st. dev: 10 dB
Fast fading model	Rayleigh flat
AWGN noise variance	$\sigma_n^2 = 0.01$
Antenna configurations	1×2 and 1×4 SIMO

Table 2: Fairness ratio between the best and worst SINR targets.

Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
—	20 dB	10 dB	3 dB	1.76 dB	0.41 dB

initialization phase of Algorithm 2, the initial SINR targets derived with equal power allocation ($\mathbf{T}_k = \mathbf{I}_{N_t}$ for all k) are $\boldsymbol{\Gamma} = [-47.4\ 0.36\ 6.41]$ dB. By optimizing the precoding matrix, the same power budget allows to increase the SINR targets to the values $\boldsymbol{\Gamma} = [-44.3\ 3.44\ 9.48]$ dB with a throughput gain of roughly 42%. Interestingly, this approach yields the same gain margin to all user's SINR target (3.07 dB in this example) due to the multiplicative term $\epsilon^{(t)}$ used in (39).

Comparing Figures 2, 6, and 8, we observe that the optimal SINR targets are different for different antenna configurations. Our conjecture is that higher antenna systems, allowing for a higher overall capacity, may lead to greater differences in the SINR setting (and consequently a higher variation of the user bit rates) when no fairness constraint is imposed. The intuitive explanation is that there is a greater room for unequal SINR assignment in a system of higher capacity. Hence, in higher-order MIMO systems, dealing with fairness constraint may become increasingly important.

8. Conclusions

Tight coordination of network elements in cellular systems enables not only the introduction of network MIMO transmission and reception techniques but also the implementation of fast SINR target setting and tracking. We addressed the problem of optimally assigning SINR targets and transmission powers to mobile stations in tightly coordinated multicell spatial multiplexing systems. We considered two versions of the SINR target and power allocation problem: one that maximizes the sum rate subject to a power budget constraint, and another that minimizes the total power needed to meet a sum-rate target. Both formulations are constrained non-convex problems. We proposed a semi-analytical solution via the augmented Lagrangian penalty function method and developed a fast numerical technique for the joint optimization of SINR targets and transmit powers. Numerical results demonstrated significant gains of the joint SINR target and power optimization depending on the degree of fairness imposed. In realistic network MIMO scenarios, our method displayed strong performance on a par with the globally optimal, but computationally very expensive, solution. We also showed how the transmission power needed to maintain a given capacity target can be reduced even further by also optimizing the precoding. A natural extension of this work is to design distributed schemes (such as that in [14]) for spatial multiplexing systems and to analyze their robustness against limited and inaccurate channel knowledge.

Appendices

A. Derivations for the Linear MMSE

For the sake of simplicity we rewrite the system model (3) as

$$\mathbf{y}_k = \mathbf{A}_k \mathbf{x}_k + \mathbf{v}_k, \tag{A.1}$$

where $\mathbf{A}_k = \alpha_{k,k}\mathbf{H}_{k,k}\mathbf{T}_k$, the vector $\mathbf{x}_k$ is zero mean with covariance $\mathbf{R}_{x_k} = \mathbf{I}$, and $\mathbf{v}_k = \mathbf{z}_k + \mathbf{n}_k$ models the intercell interference plus noise with mean and covariance, respectively:

$$\begin{aligned} \boldsymbol{\mu}_{v_k} &= \mathbb{E}[\mathbf{v}_k] = \mathbb{E}[\mathbf{z}_k] + \mathbb{E}[\mathbf{n}_k] = 0, \\ \mathbf{R}_{v_k} &= \mathbb{E}\left[\mathbf{v}_k \mathbf{v}_k^{\dagger}\right] = \mathbb{E}\left[\mathbf{z}_k \mathbf{z}_k^{\dagger}\right] + \mathbb{E}\left[\mathbf{n}_k \mathbf{n}_k^{\dagger}\right] = \mathbf{R}_{z_k} + \mathbf{R}_{n_k}, \end{aligned} \tag{A.2}$$

where the intercell interference covariance matrix $\mathbf{R}_{z_k}$ is defined as in (4). Hence, $\mathbf{y}_k$ is zero-mean, and according to [25, 26] the linear MMSE receiver is given as

$$\mathbf{G}_k = \mathbf{R}_{x_k}\mathbf{A}_k^\dagger\left(\mathbf{A}_k\mathbf{R}_{x_k}\mathbf{A}_k^\dagger + \mathbf{R}_{\nu_k}\right)^{-1} = \alpha_{k,k}\mathbf{T}_k^\dagger\mathbf{H}_{k,k}^\dagger \cdot \left(\alpha_{k,k}^2\mathbf{H}_{kk}\mathbf{T}_k\mathbf{T}_k^\dagger\mathbf{H}_{k,k}^\dagger + \sum_{j\neq k}\alpha_{k,j}^2\mathbf{H}_{k,j}\mathbf{T}_j\mathbf{T}_j^\dagger\mathbf{H}_{k,j}^\dagger + \sigma_n^2\mathbf{I}\right)^{-1}. \tag{A.3}$$

Proposition 1 follows immediately by extracting $\alpha_{k,k}^2$.

A.1. Derivation of the MSE in Proposition 2. By applying the standard theory on linear MMSE computation to the model in (A.1), see, for example, [20, Chapter 12], the MMSE error covariance matrix for the kth base station is

$$\begin{aligned}\mathbf{E}_k &= \mathbb{E}\left[(\hat{\mathbf{x}}_k - \mathbf{x}_k)(\hat{\mathbf{x}}_k - \mathbf{x}_k)^\dagger\right]\\ &= \alpha_{k,k}^2\mathbf{G}_k\mathbf{H}_{k,k}\mathbf{T}_k\mathbf{T}_k^\dagger\mathbf{H}_{k,k}^\dagger\mathbf{G}_k^\dagger - 2\alpha_{k,k}\mathbf{G}_k\mathbf{H}_{k,k}\mathbf{T}_k\\ &\quad + \mathbf{G}_k\mathbf{R}_{z_k}\mathbf{G}_k^\dagger + \mathbf{G}_k\mathbf{R}_{n_k}\mathbf{G}_k^\dagger\\ &= \left(\alpha_{k,k}\mathbf{G}_k\mathbf{H}_{k,k}\mathbf{T}_k - \mathbf{I}\right)\left(\alpha_{k,k}\mathbf{T}_k^\dagger\mathbf{H}_{k,k}^\dagger\mathbf{G}_k^\dagger - \mathbf{I}\right) + \mathbf{G}_k\mathbf{R}_{\nu_k}\mathbf{G}_k^\dagger.\end{aligned} \tag{A.4}$$

Finally, by replacing the expression of $\mathbf{G}_k$ in (A.3) into $\mathbf{E}_k$ and using similar techniques as in [17] we obtain

$$\begin{aligned}\mathbf{E}_k &= \mathbf{I} - \alpha_{k,k}\mathbf{T}_k^\dagger\mathbf{H}_{k,k}^\dagger \cdot \left(\alpha_{k,k}^2\mathbf{H}_{k,k}\mathbf{T}_k\mathbf{T}_k^\dagger\mathbf{H}_{k,k}^\dagger + \mathbf{R}_{\nu_k}\right)^{-1}\alpha_{k,k}\mathbf{H}_{k,k}\mathbf{T}_k\\ &= \left(I + \mathbf{T}_k^\dagger\mathbf{R}_{H_k}\mathbf{T}_k\right)^{-1},\end{aligned} \tag{A.5}$$

where $\mathbf{R}_{H_k} = \alpha_{k,k}^2\mathbf{H}_{k,k}^\dagger\mathbf{R}_{\nu_k}^{-1}\mathbf{H}_{k,k}$.

B. Elements of the Sum Power Vector

The parameters of (26) are as follows:

$$\begin{aligned}M_{11} &= \gamma_1^{\text{tgt}}n_1\left(1 + (1-\kappa)F_{23}F_{32}\gamma_2^{\text{tgt}}\right),\\ M_{12} &= \gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}}n_2(F_{12} - (1-\kappa)F_{13}F_{32}),\\ M_{13} &= -(1-\kappa)\left(F_{13} + F_{12}F_{23}\gamma_2^{\text{tgt}}\right)\gamma_1^{\text{tgt}}n_3,\\ M_{21} &= \gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}}n_1(F_{21} - (1-\kappa)F_{23}F_{31}),\\ M_{22} &= \gamma_2^{\text{tgt}}n_2\left(1 + (1-\kappa)F_{13}F_{31}\gamma_1^{\text{tgt}}\right),\\ M_{23} &= -(1-\kappa)\left(F_{23} + F_{13}F_{21}\gamma_1^{\text{tgt}}\right)\gamma_2^{\text{tgt}}n_3,\\ M_{31} &= -\gamma_1^{\text{tgt}}n_1\left((1-\kappa)\left(F_{31} + F_{21}F_{32}\gamma_2^{\text{tgt}}\right)\right),\\ M_{32} &= -\gamma_2^{\text{tgt}}n_2\left((1-\kappa)\left(F_{32} + F_{12}F_{31}\gamma_1^{\text{tgt}}\right)\right),\\ M_{33} &= -(1-\kappa)\left(1 - F_{12}F_{21}\gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}}\right)n_3;\end{aligned} \tag{B.1}$$

where $\kappa = e^{c_m - \log(1+\gamma_1^{\text{tgt}}) - \log(1+\gamma_2^{\text{tgt}})}$, and

$$D_p = 1 - F_{12}F_{21}\gamma_1^{\text{tgt}}\gamma_2^{\text{tgt}} + (1-\kappa)\cdot\left(F_{31}\gamma_1^{\text{tgt}}\left(F_{13} + F_{12}F_{23}\gamma_2^{\text{tgt}}\right) + F_{32}\gamma_2^{\text{tgt}}\left(F_{23} + F_{13}F_{21}\gamma_1^{\text{tgt}}\right)\right). \tag{B.2}$$

C. Proofs

For ease of notation, we rewrite problem (34) as follows

$$\begin{aligned}\underset{\mathbf{p},\mathbf{r}}{\text{maximize}} \quad & \sum_k r_k\\ \text{subject to} \quad & r_k \le W\left[\theta_k^{(t)}\log_2\left(\underline{\gamma}_k\left(e^{\tilde{\mathbf{p}}}\right)\right) + \beta_k^{(t)}\right] \quad \forall k\\ & \sum_k e^{\tilde{P}_k} \le P_{\text{tot}},\end{aligned} \tag{C.1}$$

where we applied the approximation (33) to the link rate constraint and the change of variables $P_k \leftarrow e^{\tilde{P}_k}$. In what follows, we use this formulation to prove our theoretical achievements since it easily maps back to the original nonconvex problem (30) and to the monotonic optimization (31).

C.1. Proof of Proposition 4. Similarly to [15, Lemma 4], we show that

$$r_k^{(t)} = \tilde{c}_k\left(\mathbf{p}^{(t)}, \theta_k^{(t)}, \beta_k^{(t)}\right) \le c_k\left(\mathbf{p}^{(t)}\right) = \tilde{c}_k\left(\mathbf{p}^{(t)}, \theta_k^{(t+1)}, \beta_k^{(t+1)}\right). \tag{C.2}$$

We prove the first relationship by contradiction. Assume that at the optimal solution of $\mathcal{P}^{(t)}$ the rate r_k is strictly less than the approximate capacity. Then, we could increase r_k (while keeping $\mathbf{p}$ fixed) until we achieve equality. This would improve the objective function; thus the solution was not optimal. The other two relations follow from the approximation (33). The rest of the proof follows analogously to [15, Lemma 4].

C.2. Proof of Theorem 5.1. Let $\{\mathbf{r}^\star, \tilde{\mathbf{p}}^\star, \boldsymbol{\lambda}^\star, \omega^\star\}$ denote the primal-dual optimal solution for the series of convex problems $\{P^{(t)}\}_t$. The associated KKT optimality conditions can be written as

$$\lambda_k^\star\theta_k^\star - \omega^\star e^{\tilde{P}_k^\star} - \sum_{n\neq k}\lambda_n^\star\theta_n^\star\underline{\gamma}_n\left(e^{\tilde{\mathbf{p}}^\star}\right)\frac{G_{nk}e^{\tilde{P}_k^\star}}{G_{nn}e^{\tilde{P}_n^\star}} = 0, \tag{C.3}$$

$$1 - \lambda_k^\star = 0. \tag{C.4}$$

By Proposition 5, the problem sequence $\{P^{(t)}\}_t$ converges at which point the rate constraint becomes tight, that is,

$$\tilde{c}_k\left(e^{\tilde{\mathbf{p}}^{(t)}}, \theta_k^{(t)}, \beta_k^{(t)}\right) = c_k\left(e^{\tilde{\mathbf{p}}^{(t)}}\right) = \tilde{c}_k\left(e^{\tilde{\mathbf{p}}^{(t+1)}}, \theta_k^{(t+1)}, \beta_k^{(t+1)}\right) \quad \forall k. \tag{C.5}$$

Therefore, by replacing $\mathbf{p}^\star = e^{\tilde{\mathbf{p}}^{(t)}}$ and $\theta_k^\star = \underline{\gamma}_k(\mathbf{p}^\star)/(1+\underline{\gamma}_k(\mathbf{p}^\star))$ into (C.3), the KKT conditions become

$$\lambda_k^\star \frac{\underline{\gamma}_k(\mathbf{p}^\star)}{1+\underline{\gamma}_k(\mathbf{p}^\star)} - \omega^\star P_k^\star - \sum_{n\neq k}\lambda_n^\star \frac{\underline{\gamma}_n^2(\mathbf{p}^\star)}{1+\underline{\gamma}_n(\mathbf{p}^\star)}\frac{G_{nk}P_k^\star}{G_{nn}P_n^\star} = 0, \tag{C.6}$$

$$1-\lambda_k^\star = 0. \tag{C.7}$$

It is easy to recognize that (C.6)-(C.7) coincides with the KKT optimality conditions of the original nonconvex problem (30). Thus, the problem series $\{P^{(t)}\}_t$ converges to a point $\{\mathbf{r}^\star, \mathbf{p}^\star\}$ that satisfies the KKT conditions of problem (30).

C.3. Proof of Theorem 5.2. Let $\boldsymbol{\nu}$ denote the Lagrange multipliers for the rate constraint in the monotonic optimization problem (31), and let $\{\tilde{\mathbf{r}}^\star, \tilde{\mathbf{p}}^\star, \boldsymbol{\nu}^\star, \omega^\star\}$ denote the primal-dual optimal solution. The associated KKT conditions can be written as

$$\frac{\nu_k^\star}{c_k^\star}\frac{\underline{\gamma}_k^\star}{1+\underline{\gamma}_k^\star} - \omega^\star e^{\tilde{P}_k^\star} - \sum_{n\neq k}\frac{\nu_n^\star}{c_n^\star}\frac{\underline{\gamma}_n^\star}{1+\underline{\gamma}_n^\star}\frac{G_{nk}e^{\tilde{P}_k^\star}}{G_{nn}e^{\tilde{P}_n^\star}} = 0, \tag{C.8}$$

$$e^{\tilde{r}_k^\star} - \nu_k^\star = 0, \tag{C.9}$$

where $\underline{\gamma}_k^\star \triangleq \underline{\gamma}_k(e^{\tilde{\mathbf{p}}^\star})$ and $c_k^\star \triangleq c_k(e^{\tilde{\mathbf{p}}^\star})$. Moreover, (C.9) can be expressed equivalently as follows:

$$e^{\tilde{r}_k^\star} - \nu_k^\star = 0 \Longleftrightarrow r_k^\star - \nu_k^\star = 0 \Longleftrightarrow r_k^\star\left(1-\frac{\nu_k^\star}{r_k^\star}\right) = 0. \tag{C.10}$$

Let now $\{\hat{\mathbf{r}}^\star, \hat{\mathbf{p}}^\star, \boldsymbol{\lambda}^\star, \omega^\star\}$ denote the optimal solution of the problem series $\{\mathcal{P}^{(t)}\}_t$, along with the approximation vectors $\{\boldsymbol{\theta}^\star, \boldsymbol{\beta}^\star\}$. By Theorem 5.1, this solution satisfies the KKT optimality conditions (C.7), and by Proposition 5 all constraints are active and the approximation becomes exact, that is,

$$\hat{r}_k^\star = \tilde{c}_k\left(\hat{\mathbf{p}}^\star, \boldsymbol{\theta}^\star, \boldsymbol{\beta}^\star\right) = c_k(\hat{\mathbf{p}}^\star) \quad \forall k. \tag{C.11}$$

By constructing an auxiliary set of optimal dual variables as

$$\nu_k^\star = \lambda_k^\star c_k(\hat{\mathbf{p}}^\star) = \lambda_k^\star \hat{r}_k^\star \quad \forall k, \tag{C.12}$$

the proof follows immediately by replacing $\lambda_k^\star = \nu_k^\star / c_k(\hat{\mathbf{p}}^\star)$ and $\lambda_k^\star = \nu_k^\star/\hat{r}_k^\star$ in (C.6)-(C.7), respectively.

Acknowledgments

This work was part of the SERAN project and in part supported by VINNOVA. The authors thank Dr. Claes Tidestav, Dr. Mikael Prytz, and Mats Blomgren, from Ericsson Research, for their valuable comments throughout this work. They also thank the anonymous reviewers for the valuable comments that improved the contents and the presentation of the paper.

References

[1] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications*, vol. 13, no. 4, pp. 56–61, 2006.

[2] D. Samardzija, H. Huang, R. Valenzuela, and T. Sizer, "An experimental downlink multiuser MIMO system with distributed and coherently-coordinated transmit antennas," in *Proceedings of IEEE International Conference on Communications (ICC '07)*, pp. 5365–5370, Glasgow, UK, June 2007.

[3] T. Tamaki, K. Seong, and J. M. Cioffi, "Downlink MIMO systems using cooperation among base stations in a slow fading channel," in *Proceedings of IEEE International Conference on Communications (ICC '07)*, pp. 4728–4733, Glasgow, UK, June 2007.

[4] T.-D. Nguyen, O. Berder, and O. Sentieys, "Impact of transmission synchronization error and cooperative reception techniques on the performance of cooperative MIMO systems," in *Proceedings of IEEE International Conference on Communications (ICC '08)*, pp. 4601–4605, Beijing, China, May 2008.

[5] H. Zhang, G.-S. Kuo, and T. M. Bohnert, "Cooperative diversity for virtual MIMO system in the presence of spatial correlated fading model," in *Proceedings of the 68th IEEE Vehicular Technology Conference (VTC '08)*, Calgary, Canada, September 2008.

[6] P. Komulainen and M. Boldi, Eds., "Initial Report on Advanced Multiple Antenna Systems," Deliverable of the Wireless World Initiative New Radio - WINNER+, Januray 2009.

[7] "Deployment Aspects," Third Generation Partnership Project (3GPP), June, 2007.

[8] W. Choi and J. G. Andrews, "The capacity gain from base station cooperative scheduling in a MIMO DPC cellular system," in *Proceedings of IEEE International Symposium on Information Theory (ISIT '06)*, pp. 1224–1228, Seattle, Wash, USA, July 2006.

[9] H. Kim and R. M. Buehrer, "Power allocation strategies in cooperative MIMO networks," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC '06)*, vol. 3, pp. 1675–1680, Las Vegas, Nev, USA, April 2006.

[10] M. Nokleby, A. L. Swindlehurst, Y. Rong, and Y. Hua, "Cooperative power scheduling for wireless MIMO networks," in *Proceedings of the 50th Annual IEEE Global Telecommunications Conference (GLOBECOM '07)*, pp. 2982–2986, Washington, DC, USA, November 2007.

[11] G. Li and H. Liu, "Downlink radio resource allocation for multi-cell OFDMA system," *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3451–3459, 2006.

[12] C. Koutsimanis and G. Fodor, "A dynamic resource allocation scheme for guaranteed bit rate services in OFDMA networks," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 2524–2530, Beijing, China, May 2008.

[13] R. Chen, J. G. Andrews, R. W. Heath Jr., and A. Ghosh, "Uplink power control in multi-cell spatial multiplexing wireless systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2700–2711, 2007.

[14] P. Hande, S. Rangan, M. Chiang, and X. Wu, "Distributed uplink power control for optimal SIR assignment in cellular data networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 6, pp. 1420–1433, 2008.

[15] J. Papandriopoulos, S. Dey, and J. Evans, "Optimal and distributed protocols for cross-layer design of physical and transport layers in MANETs," *IEEE/ACM Transactions on Networking*, vol. 16, no. 6, pp. 1392–1405, 2008.

[16] P. Soldati and M. Johansson, "Reducing signaling and respecting timescales in cross-layer protocols design for wireless networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '09)*, Honolulu, Hawaii, USA, December 2009.

[17] D. P. Palomar, "Convex primal decomposition for multicarrier linear MIMO transceivers," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4661–4674, 2005.

[18] H. Dai, A. F. Molisch, and H. V. Poor, "Downlink capacity of interference-limited MIMO systems with joint detection," *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 442–453, 2004.

[19] J. G. Andrews, W. Choi, and R. W. Heath Jr., "Overcoming interference in spatial multiplexing mimo cellular networks," *IEEE Wireless Communications*, vol. 14, no. 6, pp. 95–104, 2007.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, vol. 1 of *Prentice Hall Signal Processing Series*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.

[21] M. A. Bhatti, *Practical Optimization Methods with Mathematica Applications*, Springer, Berlin, Germany, 2000.

[22] L. P. Qian, Y. J. Zhang, and J. Huang, "MAPEL: achieving global optimality for a non-convex wireless power control problem," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1553–1563, 2009.

[23] J. Papandriopoulos and J. S. Evans, "SCALE: a low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3711–3724, 2009.

[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[25] A. Paulraj, *Course Notes, EE 492m*, Stanford University, Stanford, Calif, USA, 2002.

[26] S. Boyd, *Course Notes, EE 363*, Stanford University, Stanford, Calif, USA, 2003.

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 218564, 7 pages
doi:10.1155/2010/218564

Research Article

Broadcast Network Coverage with Multicell Cooperation

Hongxiang Li,[1] Samee Ullah Khan,[1] and Hui Liu[2]

[1] *Electrical and Computer Engineering, North Dakota State University, Fargo, ND 58102, USA*
[2] *Electrical Engineering, University of Washington, Seattle, WA 98195, USA*

Correspondence should be addressed to Hongxiang Li, hongxiang.li@ndsu.edu

Received 17 August 2009; Accepted 22 October 2009

Academic Editor: Lingjia Liu

Multicell cooperation has been identified as one of the underlying principles for future wireless communication systems. This paper studies the benefits of multicell cooperation in broadcast TV network from an information theoretical perspective. We define outage capacity as the figure of merit and derive the broadcast coverage area to evaluate such system. Specifically, we calculate the broadcast coverage area with given common information rate and outage probabilities when multiple base stations collaboratively transmit the broadcast signals. For the general MIMO case where receivers have multiple antennas, we provide simulation results to illustrate the expanded coverage area. In all cases, our results show that the coverage of a TV broadcast network can be significantly improved by multicell cooperation.

1. Introduction

The wireless industry is experiencing an unprecedented increase in demand for better multimedia broadcasting systems. The growing diffusion of new services, like HDTV broadcasting and mobile television, emphasize the need of advanced transmission techniques that can fundamentally increase the system performance over the conventional broadcast network. Cooperative transmission is a key technology that has been identified to fulfill this demand.

Most of the studies on cooperative transmission so far focus on reducing intercell interference and increasing system capacity in two-way (uplink/downlike) cellular systems, where cooperation can be employed at either the base station (BS) or mobile station (MS). Multicell cooperation (MCC), sometimes also referred to as *multicell processing* or *distributed antenna system,* prescribes joint encoding/decoding of the signals transmitted/received at the BSs through the exploitation of the high-capacity backbone connecting the BSs. A survey on MCC can be found in [1–4]. Assume all the BSs in the network are connected to a central processor via links of unlimited capacity, the set of BSs effectively acts as a geographically distributed multiantenna system. Under this perfect MCC assumption, [5] investigated multicell downlink channel capacity with single-class network where the users are clustered at the cell edges. The analytical performance expression for dirty paper coding (DPC), cophasing, zero forcing (ZF), and MMSE precoders are derived in this paper. Recently [6] studied the capacity of a cellular system with a large number of BSs and MSs. The evaluation is based on an accurate modeling of the multicell deployment, different level of cooperation among BSs, and practical per-antenna transmit power constraint of a BS transmit antenna. The performance upper and lower bounds are also derive for different cooperation scenarios. On the other hand, the impact of limited-capacity backhaul on MCC (i.e., partial cooperation) has also been studied in the literature [7–14]. In [7], the authors studied achievable rates of a multicell network under the assumption that BS and MS cooperation are enabled by error-free but limited-capacity interbase station and interuser links, respectively. The analytical treatment enhances the insight into the potential and the limitations imposed by capacity constraints on the performance gains provided by cooperative techniques. Some recent results on intercell interference coordination through limited feedback can also be found in [14].

While there are many existing work showing the benefits of MCC in cellular networks, few studies can be found in the literature on cooperative broadcast network. Compared

with the cellular (i.e., voice and data) network, there are some fundamental differences in broadcast (i.e, TV) network.

(i) Communication is only one-way from base station to subscriber stations (SS); there is no reverse link available. Therefore, the channel state information (CSI) is not available at the transmitter. In cellular network, reverse link is available and the BS can have perfect or partial CSI.

(ii) BS sends the same common information to all the subscribers. In cellular network, BS sends different individual information to each subscriber.

(iii) SSs (TV receivers) can only receive but not transmit signals. Therefore, there is no possibility of cooperation (information exchange) among SSs. In cellular network, MSs (smart phones, laptops) can transmit signals and thus cooperation is possible among MSs.

Unlike the two-way cellular network, the concern of broadcast TV network is usually not the throughput. Rather, the design objective of such network is to maximize the broadcast coverage, with a predetermined transmission rate. It has been expected that MCC can extend the broadcast coverage area, forming an extended ellipse coverage beyond the superposition of individual cells [15]. However, to the best knowledge of the author of this article, there are no quantitative studies in the literature on how much the broadcast coverage can be extended by MCC. In this paper, we consider full MCC in broadcast network with cooperation among some small number of neighboring base stations. Specifically, we define the figure-of-merit of MCC in broadcast from an information theoretical perspective. Then we quantify the extended coverage areas under different MCC scenarios and derive the coverage gains over single-cell independent transmission. We also investigate the broadcast coverage with optimal base station separation. This paper is an extension of our earlier work [16], where we calculated the broadcast coverage in a broadcast and unicast hybrid network with three collaborating cells.

The paper is organized as the following. In Section 2, we define the performance metric and present the broadcast channel model. We then analyze and compute the broadcast coverage with MCC for single/multiple receiving antenna(s) in Section 3. Finally, a conclusion is drawn in Section 4.

To facilitate our discussion, we use the following notation conventions throughout this paper:

N_r: the number of antennas at each receiver

N_t: the number of collaborating base stations

$\mathbf{H}_k$: $N_r \times N_t$ (channel matrix associated with user k)

$\mathbf{Q}$: input spatial covariance matrix

P: the total transmit power

B: the total channel bandwidth

r_o: common information broadcast rate

q_o: outage Probability

N_0: white noise variance (10^{-9} w/Hz)

2. System Model

In wireless community, the term *broadcast* has been used for both TV broadcast and cellular downlink system. To avoid ambiguity, we define *broadcast network* as the delivery of *common information* from the base station or TV tower to silent users in multicast applications. As we discussed in Section 1, there is no reverse link and thus CSI is not available at the transmitter.

In order to quantify the performance of a broadcast system, we must first define the performance metric. In information theory, there are two channel capacity definitions that are relevant to the system design for a broadcast channel with an uninformed transmitter: the ergodic capacity (also called the Shannon capacity) and the outage capacity [17]. The ergodic capacity defines the maximum data rate that can be sent to the receiver with asymptotically small error probability through all the fading states. In most cases, TV receivers at fixed locations experience slow fading and the choose of ergodic capacity will incur intolerant decoding delays. On the other hand, the outage capacity defines the maximum data rate that can be transmitted with certain outage probability that the received data can not be decoded with negligible error probability. If the received SNR is above the threshold corresponding to the outage probability, the transmitted data can be decoded with negligible probability of error; otherwise, the transmission is in outage. By allowing some outage, the broadcast receiver can decode the message during each fading state and thus meet the delay requirement. In this paper, we use the more practical outage capacity as the figure of merit for evaluating the broadcast network.

In general, since broadcast receivers are geographically distributed over a large area, they decode the received signals with different outage probabilities depending on the fading statistics. Within a given coverage area A, We define the outage probability $q(W_A)$ as the outage probability (package loss rate) associated with the worst broadcast receiver W in this area. Note that the worst receiver W always experiences the highest outage in A. Literally, there are three interchangeable ways to optimize the broadcast performance.

(1) Given a broadcast rate r_o and an outage probability q_o, maximize the coverage area A such that $q(W_A) \leq q_o$.

(2) Given a coverage area A and an outage probability q_o, maximize the common information rate r_o such that $q(W_A) \leq q_o$.

(3) Given the broadcast rate r_o and coverage area A, minimize the worst user outage probability $q(W_A)$.

Note that the above three problems are equivalent in optimizing the common information transmission from BS to receiver W_A. In other words, we only need to focus on the performance of the worst receiver in the covered area. In practical TV broadcast networks, r_o and q_o are usually prefixed and thus our objective is to maximize broadcast coverage area with given quality of service (outage) requirement.

In conventional TV broadcasting, the terrestrial television stations are usually high power tall towers that cover a large area and the TV receivers are at fixed locations. With the increasing demand of mobile TV, many wireless carriers are offering TV services over their cellular base stations to mobile subscribers with smart phones or PDAs. In U.S., wireless carriers (Verizon, AT&T mobility and etc.) are using MediaFlo (forward link only) technology to deliver mobile TV service. In this paper, we are not limited to a particular broadcast network or technique, rather, we study the broadcast coverage from an information theory point of view. To ease our discussion, we simply use the term "base station" to refer any broadcast transmitter (radio mast, tower or cellular base station). In traditional broadcast network, base stations are operated independently. In this paper, we investigate the scenario where multiple base stations can collaborate with each other to transmit the common information.

To calculate the broadcast coverage, we have to assume the channel fading statistics. The same channel model in [16] is used. For free space propagation loss, Hata model is the most common model for signal prediction in large urban macro-cells. This model is applicable over distances of 1–100 km and frequency ranges of 150–1500 MHz. The standard formula for empirical path loss in urban areas under the Hata model is

$$\begin{aligned} P_{L,\text{urban}}(d)\,\text{dB} = {} & 26.16\log_{10}(f_c) - 13.82\log_{10}(h_t) \\ & - a(h_r) + \Big(44.9 - 6.55\log_{10}(h_t)\Big)\log_{10}(d) \\ & + 69.55, \end{aligned} \tag{1}$$

where f_c is the carrier frequency, h_t/h_r is the transmitter/receiver antenna height and d is the distance between transmitter and receiver. For larger cities at frequencies $f_c >$ 300 MHz, the correction factor $a(h_r)$ is given by

$$a(h_r) = 3.2\Big(\log_{10}(11.75h_r)\Big)^2 - 4.97\,\text{dB}. \tag{2}$$

For small scale fading, we assume Rayleigh flat fading, that is, the envelope of the complex channel gain for each spatial channel has the following distribution:

$$f(|h|) = \frac{|h|}{\sigma^2}\exp\left(-\frac{|h|^2}{\sigma^2}\right), \tag{3}$$

where $E[|h|^2] = 2\sigma^2$ is determined by path loss in (1).

Note that in a real urban environment with many high-rise buildings, the actual channel statistics is much more complicated. A combined path loss and shadowing model is needed to calculate an outage of a particular receiver. However, this fading statistics will highly depends on the actual urban terrain and structure and varies from city to city. Assume a flat open area, our simplified channel model can give enough insights on the benefit of MCC. We also assume slow flat fading throughout the paper.

3. Broadcast Coverage

In a single cell broadcast network with a fixed broadcast rate r_o, the information theoretical outage probability associated with receiver k is given by [17]

$$q_o(k) = \text{Prob}\left[h_k : B\log_2\left(1 + \frac{P|h_k|^2}{N_0 B}\right) < r_o\right]. \tag{4}$$

In continuous frequency-selective case, the outage probability $q_o(k)$ becomes

$$\begin{aligned} & \text{Prob}\left[h_k : \int_0^B \log_2\left(1 + \frac{P(f)\,|\,h_k(f)\,|^2}{N_0}\right)df < r_o\right] \\ & \text{s.t.} \int_0^B P(f)df = P. \end{aligned} \tag{5}$$

In multicarrier/OFDM system where the channel is block frequency-selective, (5) becomes

$$\begin{aligned} & \text{Prob}\left[h_k : \sum_{i=1}^{N} B_n\log_2\left(1 + \frac{P(i)|h_k(i)|^2}{N_0 B_n}\right) < r_o\right] \\ & \text{s.t.} \sum_{i=1}^{N} P(i) = P \end{aligned} \tag{6}$$

Obviously, the coverage area A determined by Equation (1)–(4) is a circle for any given outage requirement. Without loss of generality, we choose r_o = 515 Kbps and P = 1 watt such that the radius d_0 = 1 km is the benchmark distance with $q_o(d_0)$ = 5%. Figure 1 shows the broadcast coverage area with different outage probabilities, where the solid black curve (5% outage) is our benchmark. We see that the coverage increases with q_o as expected.

Next, we will discuss the coverage expansion by multicell cooperation.

3.1. MCC with Single Receiver Antenna (N_r = 1). We assume multiple base stations are wire connected through high-capacity backbone network and thus can fully collaborate with each other. In this case, each base station can be viewed as a transmitting antenna of a distributed multiple input single output (MISO) system. Also, because the distances among base stations are on the order of kilometers, spatial channels from different base stations are considered mutually independent. Let $\vec{\mathbf{x}} = [x(1), x(2), \ldots, x(N_t)]^H$ be the transmitted input vector, the outage probability associated with receiver k is given by [17]

$$\begin{aligned} q_o = \text{Prob}\left[\vec{\mathbf{h}_k} : B\log_2\det\left(\mathbf{I} + \frac{\vec{\mathbf{h}_k}\mathbf{Q}\vec{\mathbf{h}_k}^H}{N_0 B}\right) < r_o\right] \\ \text{s.t. } \mathbf{E}\Big[x(i)x(i)^*\Big] = \frac{P}{N_t}, \quad 1 \le i \le N_t, \end{aligned} \tag{7}$$

where $\vec{\mathbf{h}_k} = [h_k(1), h_k(2), \ldots, h_k(N_t)]$ is the channel vector and $\mathbf{Q} = \mathbf{E}[\vec{\mathbf{x}}\vec{\mathbf{x}}^H]$ is the spatial input covariance matrix.

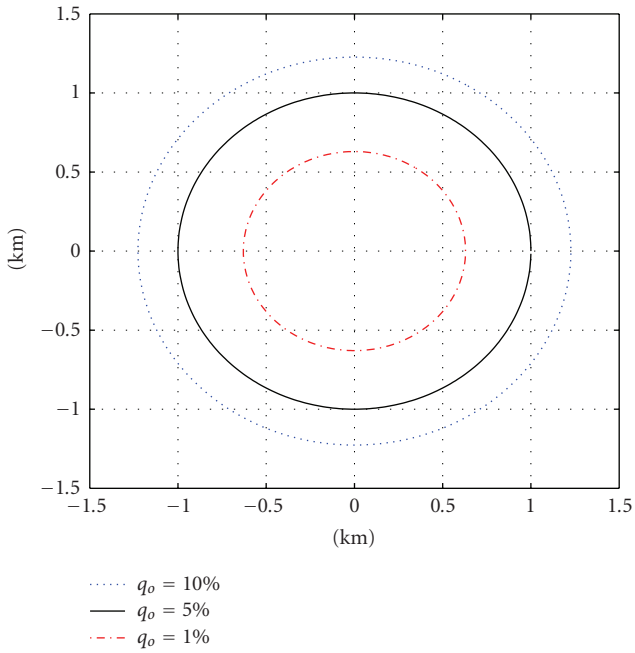


Figure 1: Single cell broadcast coverage.

Remark 1. In practice, each base station is subject to its own power constraint, which is dictated by the amplifier of the transmitting antenna. In this paper, we adopt this individual BS power constraint. However, for fair comparison of broadcast coverage between single cell and multicell cases, we put a total power cap in multicell transmission, that is, the maximum sum power over all base stations is P.

Because of the nature of the broadcast network, the BSs do not have the channel and location information of the receivers. Therefore, the transmitter cannot optimize its input covariance structure across antennas for every receiving point simultaneously. Assume the zero-mean spatially white (ZMSW) channel model, the optimal power allocation strategy is to allocate equal power to each base station [18], resulting $\mathbf{Q}$ equal to the scaled identity matrix: $\mathbf{Q} = (\mathbf{P}/\mathbf{N_t})\mathbf{I}$. Equation (7) thus becomes

$$q_o = \text{Prob}\left[\vec{\mathbf{h}}_k : B\log_2 \det\left(\mathbf{I} + \frac{P}{N_t}\frac{\left\|\vec{\mathbf{h}}_k\right\|^2}{N_0 B}\right) < r_o\right]. \tag{8}$$

Note that equal power distribution over BSs does not mean all the BSs transmit the same signal. With channel unknown at transmitter, the Alamouti Scheme is a technique to achieve the N_t fold diversity gain [17]. From (8), we can see that, given r_o and q_o, the coverage area depends solely on the distribution of $\|\vec{\mathbf{h}}_k\|^2$. Since the average path loss is simply a function of the distance between base station and a broadcast receiver, we know the worst receiver W_A is always on the edge of area A. Therefore, the coverage area can be computed by locating edge users in all directions. To do so, we first need to derive the distribution of $\|\vec{\mathbf{h}}_k\|^2/N_t$. In practice, it is unrealistic and unnecessary to do multicell cooperation for a large number of cells [6, 7]. Here we study two practical scenarios where $N_t = 2$ and $N_t = 3$.

Let $Y = \|\vec{\mathbf{h}}\|^2/N_t$ and $X_i = |h(i)|^2/N_t$. Note that X_i and X_j are i.i.d $(i \neq j)$ random variables with chi-square distribution, that is, the distribution of X_i is given by

$$\begin{aligned}\text{PDF} : f(x) &= \frac{1}{2\sigma^2}\exp\left(-\frac{x}{2\sigma^2}\right)\\ \text{CDF} : F(x) &= 1 - \exp\left(-\frac{x}{2\sigma^2}\right).\end{aligned} \tag{9}$$

For the case $N_t = 2$, the cdf of Y can be obtained by convoluting (9) as

$$\begin{aligned}F_Y(y) &= f_{X_1}(y) * F_{X_2}(y)\\ &= \begin{cases}\text{if } \sigma_1 \neq \sigma_2\\ 1 + \dfrac{1}{\sigma_2^2 - \sigma_1^2}\left[\sigma_1^2\exp\left(-\dfrac{y}{2\sigma_1^2}\right) - \sigma_2^2\exp\left(-\dfrac{y}{2\sigma_2^2}\right)\right]\\ \text{if } \sigma_1 = \sigma_2\\ 1 - \exp\left(-\dfrac{y}{2\sigma_1^2}\right) - \dfrac{y}{2\sigma_1^2}\exp\left(-\dfrac{y}{2\sigma_1^2}\right),\end{cases}\end{aligned} \tag{10}$$

where σ_i is determined by the pass loss from base station i to the receiver.

In the general case with arbitrary N_t, the PDF of Y can be obtained as

$$f_Y(y) = \begin{cases}\text{if } \sigma_1 \neq \sigma_2 \neq \cdots \neq \sigma_N\\ \displaystyle\sum_{i=1}^{N}\frac{\sigma_i^{2N-4}}{\prod_{j=1 j\neq i}^{N}\left(\sigma_i^2 - \sigma_j^2\right)}\exp\left(-\frac{y}{\sigma_i^2}\right)\\ \text{if } \sigma_1 = \sigma_2 = \cdots = \sigma_N\\ \dfrac{y^{N-1}}{(N-1)!\sigma^{2N}}\exp\left(-\dfrac{y}{\sigma_i^2}\right).\end{cases} \tag{11}$$

Equations (10) and (11) can be used to determine the multicell broadcast coverage area. For example, Figure 2 shows the two cell collaborating case, where D is the cell separation distance and (ρ, θ) is the polar coordinates of the receiver. Note that for path loss, σ_i is simply a function of the distance between cell i and the receiver. For each direction $(0 \leq \theta \leq 360)$ from the origin, we calculate a radius $\rho(\theta)$ such that the distribution of $\|\vec{\mathbf{h}}\|^2/N_t$ at (ρ, θ) satisfies (8). By locating the coverage edge in all directions, the maximum coverage area can be numerically computed.

Figures 3–5 show some numerical results of multicell broadcast coverage. For $N_t = 2$, Figure 3 plots the extended coverage with two base stations located at $(d_0/2, 0)$ and $(-d_0/2, 0)$. Compared to single cell transmission shown in Figure 1, the coverage gains with two collaborating cells are 3.5399, 2.0922 and 3.4522 for outage probability 1%, 5% and 10% respectively, where we define the coverage gain as the ratio of MCC coverage over single cell coverage.

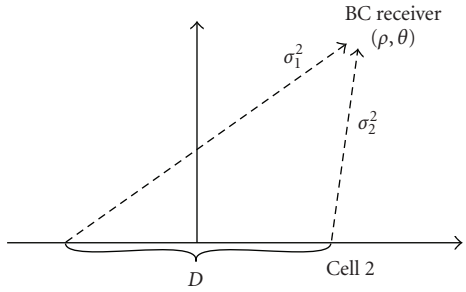


FIGURE 2: Multicell broadcast coverage.

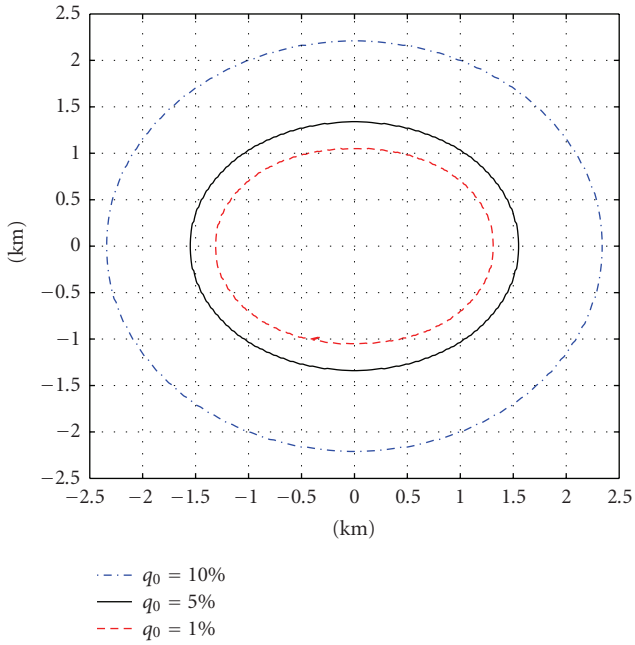


FIGURE 3: MISO with antenna location $(0.5, 0)$ and $(-0.5, 0)$.

Similarly, Figure 4 plots the two cell coverage where the distance between two base stations is doubled. Compared to single cell case, the coverage gains are 3.5245, 2.1469 and 3.5033 for outage probability 1%, 5% and 10% respectively. We observe that the two cell coverage gain is not sensitive to cell separation distance.

For $N_t = 3$, Figure 5 shows the broadcast coverage area for both single-cell and multicell transmissions with $q_o = 5\%$ [16]. The circle in the centre indicates the single cell coverage area. The outer region is the extended coverage area with multicell cooperation, and the three small circles around base stations are the multicell coverage area when they do not collaborate. Compared to the single cell case, the multicell cooperation coverage gain is 315% in this case.

To further find the optimal cell separation, we numerically calculate the coverage area as a function of cell separation distance for different q^os. Assume three collaborating cells are equally separated, Figure 6 shows the total coverage area as a function of BS separation distance [16]. The optimal cell separation is 1.8, 2.4 and 2.8 kms to achieve the maximum coverage area for outage probability 1%, 5% and 10%, respectively.

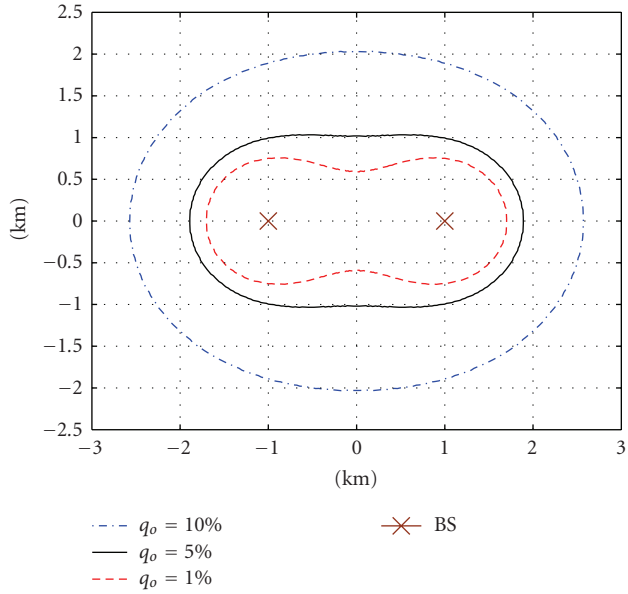


FIGURE 4: MISO with antenna location $(1, 0)$ and $(-1, 0)$.

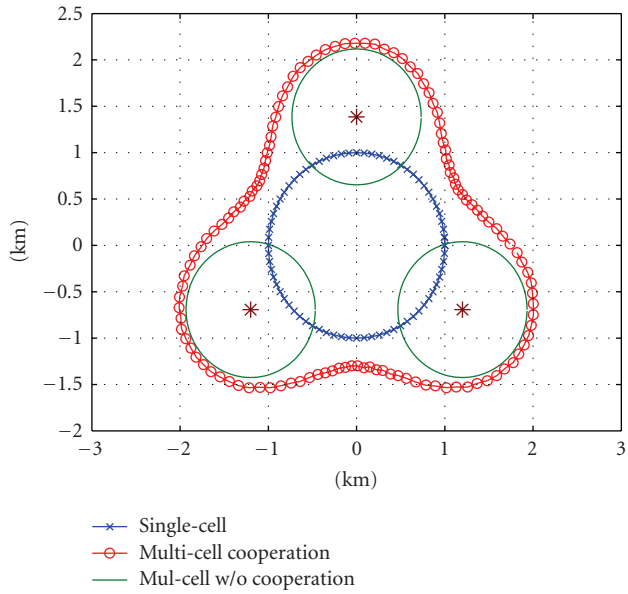


FIGURE 5: BC coverage area with $q_0 = 5\%$.

3.2. MCC with Multiple Receiver Antenna ($N_r = N_t = 2$). When the receiver has multiple antennas, the collaborative multicell broadcast network becomes a distributed MIMO system. While spatial channels are independent among base stations, they are correlated among receiving antennas of each receiver because multiple receiving antennas are closely located. In MIMO broadcast, the outage probability

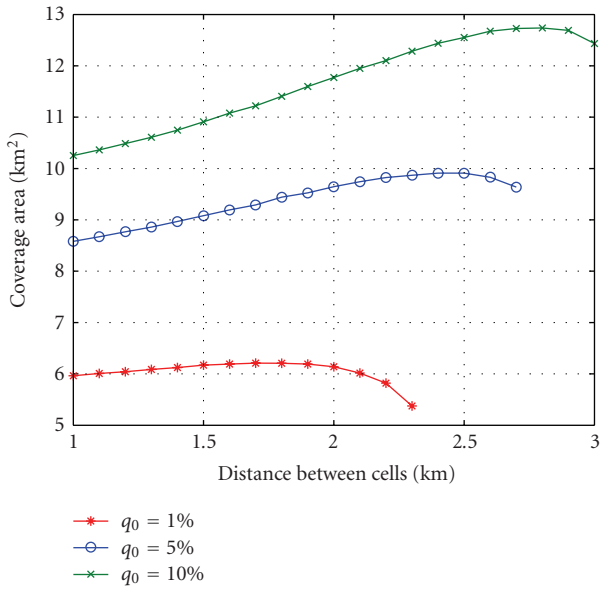


Figure 6: BC coverage area versus cell distance ($N_t = 3$).

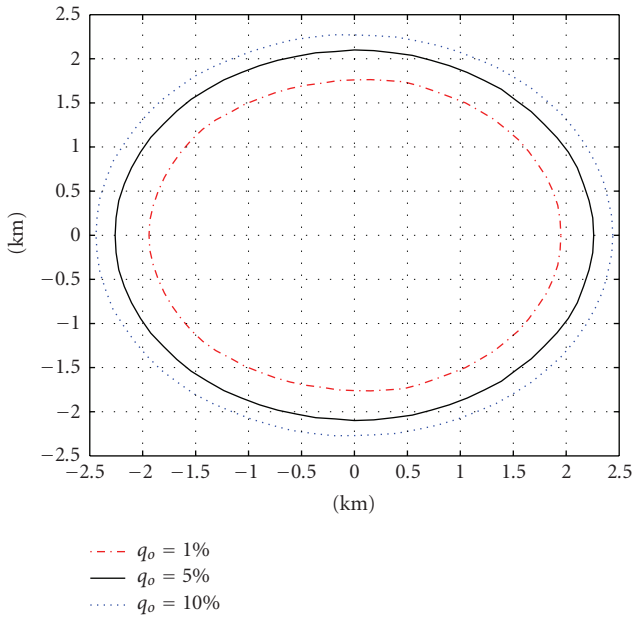


Figure 7: MIMO with antenna location $(0.5, 0)$ and $(-0.5, 0)$.

associated with receiver k is given by:

$$q_o = \text{Prob}\left[\mathbf{H}_k : B\log_2 \det\left(\mathbf{I} + \frac{\mathbf{H}_k\mathbf{Q}\mathbf{H}_k^H}{N_0 B}\right) < r_o\right] \quad \text{s.t. } \mathbf{E}\left[x(i)x(i)^*\right] = \frac{P}{N_t}, \quad 1 \le i \le N_t, \tag{12}$$

where $\mathbf{H}_k$ is the channel matrix associated with user k. Plug in $\mathbf{Q} = (\mathbf{P}/\mathbf{N_t})\mathbf{I}$, (12) becomes

$$q_o = \text{Prob}\left[\mathbf{H}_k : B\log_2 \det\left(\mathbf{I} + \frac{P\mathbf{H}_k\mathbf{H}_k^H}{N_t N_0 B}\right) < r_o\right]. \tag{13}$$

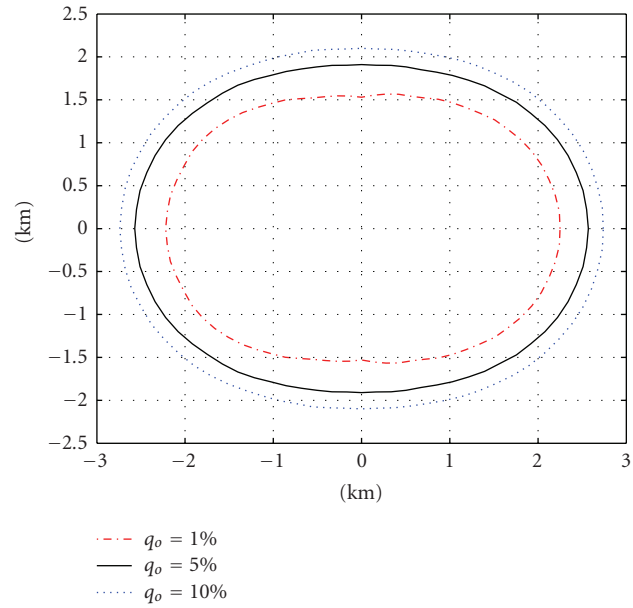


Figure 8: MIMO with antenna location $(1, 0)$ and $(-1, 0)$.

Due to the added spatial complexity, there is no close form for the distribution of $\det(\mathbf{I} + P\mathbf{H}_k\mathbf{H}_k^H/N_tN_0B)$ in (13). We thus use simulation to obtain the outage coverage numerically. In our simulation, the spatial channel correlation coefficient between any two receiving antennas is 0.5.

For the case of $N_r = N_t = 2$, Figure 7 shows the coverage gain with two base stations located at $(d_0/2, 0)$ and $(-d_0/2, 0)$. Compared to SISO broadcast network, the coverage gains are 8.8, 4.84 and 3.71 for outage probability 1%, 5% and 10% respectively. This huge gain is expected because of the added antenna at each receiver.

Similar results in Figure 8 show the coverage gain with base station locations $(d_0, 0)$ and $(-d_0, 0)$. The coverage gains in this case are 9.21, 5.13 and 3.95 for outage probability 1%, 5% and 10% respectively. Compared to Figures 7 and 8 yields a better gain due to a larger BS separation. However, the MCC gain will become zero when collaborating BSs are too far away. Therefore, the BSs separation distance should be carefully selected in order to optimize the network coverage. In all cases, we see the distributed MIMO broadcast network provides the maximum coverage.

4. Conclusion

In this paper, we studied the performance of multicell cooperation in broadcast network. Due to the unique nature of broadcast, we chose outage capacity and the coverage area as the figure of merit. In reality, cooperation is not feasible among a large number of cells. For the cases of two and three collaborating cells, we explicitly derived the distribution of the received power and calculated the broadcast coverage. We also evaluated the effect of base station separation on broadcast coverage gain. The results show that collaborating BSs should not be placed too close or two far way in order to achieve the maximum coverage. For the distributed

MIMO case where each receiver has multiple antennas, we provide numerical results to show the expanded coverage area. In all cases, our results show that the coverage area of broadcast network can be significantly increased by multicell cooperation.

References

[1] S. Shamai (Shitz) and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proceedings of the 53rd IEEE Vehicular Technology Conference (VTC '01)*, vol. 3, pp. 1745–1749, Rhodes, Greece, May 2001.

[2] S. Shamai (Shitz), O. Somekh, and B. M. Zaidel, "Multicell communications: an information theoretic perspective," in *Proceedings of the Joint Workshop on Communications and Coding (JWCC '04)*, Florence, Italy, October 2004.

[3] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), "Sum rate characterization of joint multiple cell-site processing," in *Proceedings of the 9th Canadian Workshop on Information Theory (CWIT '05)*, Montreal, Canada, June 2005.

[4] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, U. Spagnolini, and S. Shamai (Shitz), "An information theoretic view of distributed antenna processing in cellular systems," in *Distributed Antenna Systems: Open Architecture for Future Wireless Communications*, pp. 31–64, Auerbach, CRC Press, Boca Raton, Fla, USA, 2007.

[5] S. Jing, D. N. C. Tse, J. B. Soriaga, J. Hou, J. E. Smee, and R. Padovani, "Multicell downlink capacity with coordinated processing," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, Article ID 586878, 19 pages, 2008.

[6] Y. Kim, H. Li, and H. Liu, "Capacity of a multi-cell cooperative system," in *Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS '08)*, pp. 314–318, Reykjavik, Iceland, October 2008.

[7] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai (Shitz), "Distributed MIMO in multi-cell wireless systems via finite-capacity links," in *Proceedings of the 3rd IEEE International Symposium on Communications, Control, and Signal Processing (ISCCSP '08)*, pp. 203–206, March 2008.

[8] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), "Spectral efficiency of joint multiple cell-site processors for randomly spread DS-CDMA systems," *IEEE Transactions on Information Theory*, vol. 53, no. 7, pp. 2625–2637, 2007.

[9] E. Aktas, J. Evans, and S. Hanly, "Distributed decoding in a cellular multiple-access channel," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '04)*, p. 484, Chicago, Ill, USA, June-July 2004.

[10] A. Sanderovich, O. Somekh, and S. Shamai (Shitz), "Uplink macro diversity with limited backhaul capacity," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 11–15, Nice, France, June 2007.

[11] O. Shental, A. J. Weiss, N. Shental, and Y. Weiss, "Generalized belief propagation receiver for near-optimal detection of two-dimensional channels with memory," in *Proceedings of the IEEE Information Theory Workshop (ITW '04)*, pp. 225–229, San Antonio, Tex, USA, October 2004.

[12] P. Marsch and G. Fettweis, "A framework for optimizing the downlink performance of distributed antenna systems under a constrained backhaul," in *Proceedings of the 13th European Wireless Conference (EW '07)*, Paris, France, April 2007.

[13] P. Marsch and G. Fettweis, "A framework for optimizing the uplink performance of distributed antenna systems under a constrained backhaul," in *Proceedings of the IEEE International Conference on Communications (ICC '07)*, pp. 975–979, Glasgow, Scotland, June 2007.

[14] L. Liu, J. Zhang, J. Yu, and J. Lee, "Inter-cell interference coordination through limited feedback," accepted by *International Journal of Digital Multimedia Broadcasting*.

[15] M. Eriksson, "Dynamic single frequency networks," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 10, pp. 1905–1914, 2001.

[16] H. Li, B. Liu, and H. Liu, "Transmission schemes for multicarrier broadcast and unicast hybrid systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4321–4330, 2008.

[17] A. Goldsmith, *Wireless Communications*, Cambridge University Press, New York, NY, USA, 2005.

[18] E. Telatar, "Capacity of multi-antenna Gaussian channels," Internal Technical Memorandum, AT&T Bell Laboratories, June 1995.

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 307927, 11 pages
doi:10.1155/2010/307927

Research Article

An Analytical Multimodulus Algorithm for Blind Demodulation in a Time-Varying MIMO Channel Context

Steredenn Daumont[1,2] and Daniel Le Guennec[2]

[1] *IPSIS Company, 3 Square du chêne Germain, 35510 Cesson-Sévigné, France*
[2] *IETR/Supelec, Campus de Rennes, Avenue de la Boulaie, CS 47601, 35576 Cesson-Sévigné Cedex, France*

Correspondence should be addressed to Steredenn Daumont, steredenn.daumont@supelec.fr

Received 30 June 2009; Accepted 5 October 2009

Academic Editor: Hongxiang Li

This paper addresses the issue of blind multiple-input multiple-output (MIMO) demodulation of communication signals, with time-varying channels and in an interception context. A new adaptive-blind source separation algorithm, which is based on the implementation of the Multimodulus cost function by analytical methods, is proposed. First a batch processing analysis is performed; then an adaptive implementation of the (Analytical Multi-Modulus Algorithm) AMMA and its simplified version named (Analytical Simplified Constant Modulus Algorithm) ASCMA is detailed. These algorithms, named adaptive-AMMA and adaptive-ASCMA, respectively, are compared with the adaptive (Analytical Constant Modulus Algorithm) ACMA and the MMA (Multi-Modulus Algorithm). The adaptive-AMMA and adaptive-ASCMA achieve a lower residual intersymbol interference and bit error rate than those of the adaptive-ACMA and MMA.

1. Introduction

The context of this study is the Multiple-Input Multiple-Output (MIMO) interception, and more specifically blind demodulation in reception. The channel is supposed to be not stationary, so the methods used to demodulate the received signals must be fast. Such systems exist, among them are the Blind Source Separation (BSS) implemented with analytical methods.

BSS algorithms require few information about emission signals in order to process and recover the transmitted symbols (sources). This type of algorithms estimate an equalization matrix from received signals to obtain the transmitted sources. To find this matrix, different cost function can be minimized, like the Multimodulus (MM) [1, 2], the Constant Modulus (CM) [3, 4], the Simplified Constant Modulus (SCM) [5], and the Multiuser Kurtosis (MUK) [6]. The algorithms in [1–6] use the stochastic gradient to minimize these cost functions, so they are slow to converge and with time-varying environments they are less accurate. So, analytical methods such as the adaptive Analytical Constant Modulus Algorithm (ACMA) [7–10], which converges quickly, should be used with time-varying channels.

Another particularity of the BSS algorithms is their ability to restore, in the output, the sources with a certain permutation and a certain rotation or phase. Moreover, when a frequency offset is present, the constellations in output of adaptive-ACMA and ACMA spin. In order to avoid that and track the frequency offset jointly to the blind source separation, an MM or SCM cost functions can be used. To our knowledge, the MM and the SCM cost functions associated with analytical methods have never been tried so far (we have presented these works partly in conference [11]). So, in this paper an adaptive Analytical Multimodulus Algorithm and its simplified version named the adaptive Analytical Simplified Constant Modulus Algorithm are proposed.

The performances of the adaptive-AMMA and adaptive-ASCMA are compared with those of the adaptive-ACMA and MMA. With time-selective channels, the adaptive-AMMA and the adaptive-ASCMA allow to obtain a higher signal-to-interference noise ratio (SINR) and a lower bit error rate

(BER), than those obtained with the adaptive-ACMA and the MMA. Moreover, both proposed methods perform a source separation and a carrier phase recovery, unlike the adaptive-ACMA.

First of all, in this paper, the system model used here is described. Then a batch analysis of the blind source separation algorithm AMMA is performed, and an adaptive form of this algorithm and of its simplified version is presented. Finally, simulation results of the proposed algorithms are compared to those obtained with the adaptive-ACMA and MMA in time-varying envrionments.

The Kronecker and Khatri-Rao products have the following properties:

$$\begin{aligned}
&\mathrm{vec}(\mathbf{a}\mathbf{b}^H) = \mathbf{b}^* \otimes \mathbf{a},\\
&(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = \mathbf{AC}\otimes\mathbf{BD},\\
&\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T\otimes\mathbf{A})\,\mathrm{vec}(\mathbf{B}),\\
&\mathrm{vec}(\mathbf{A}\,\mathrm{diag}(\mathbf{b})\mathbf{C}) = (\mathbf{C}^T\circ\mathbf{A})\mathbf{b}.
\end{aligned}$$

2. System Model

A MIMO system with N_t transmitters and N_r receivers is considered. In this system, the information symbols $x(k)$ are transmitted through the time-varying MIMO $N_r \times N_t$ channel $\mathbf{H}(k)$ at the time k. Taking into consideration a carrier frequency offset δ_f, the $N_r \times 1$ vector of the received signals $\mathbf{y}(k)$ can be expressed as

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x}(k)e^{-j2\pi k\delta_f T_s} + \mathbf{b}(k), \tag{1}$$

where

$$\begin{aligned}
\mathbf{y}(k) &= \begin{pmatrix} y_1(k) \\ \vdots \\ y_{N_r}(k) \end{pmatrix}, \qquad \mathbf{b}(k) = \begin{pmatrix} b_1(k) \\ \vdots \\ b_{N_r}(k) \end{pmatrix},\\
\mathbf{H}(k) &= \begin{pmatrix} h_{11}(k) & h_{1N_t}(k) \\ \vdots & \vdots \\ h_{N_r1}(k) & h_{N_rN_t}(k) \end{pmatrix}, \qquad \mathbf{x}(k) = \begin{pmatrix} x_1(k) \\ \vdots \\ x_{N_t}(k) \end{pmatrix},
\end{aligned} \tag{2}$$

and

(i) $\mathbf{y}$ is an $(N_r \times 1)$ representing the baseband received symbols without any time oversampling,

(ii) $\mathbf{x}$ is an $(N_t \times 1)$ vector representing the transmitted signals (sources). The sources are independent, identically distributed (i.i.d.), and mutually independent with zero mean and unit variance,

(iii) $\mathbf{b}$ is an $(N_r \times 1)$ vector representing the additive Gaussian noise,

(iv) $\mathbf{H}$ is an $(N_r \times N_t)$ unknown, full column rank, instantaneous linear and time-selective MIMO channel,

(v) $1/T_s$ is the baud rate of the transmitter.

3. The AMMA: A BSS Analytical Method so as to Minimize the MM Cost Function

3.1. Hypotheses. The BSS can be performed under the following hypotheses.

(i) $N_r \geq N_t$.

(ii) $\mathbf{H}$ has i.i.d. complex components, is unitary unlike the received signals $\mathbf{y}(k)$, and must be prewhitened before applying the BSS.

(iii) The noise is additive, white, and Gaussian with a zero mean and a covariance matrix $\mathbf{C}_b = E[\mathbf{b}\mathbf{b}^H] = \sigma_b^2\mathbf{I}_{N_r}$.

(iv) The sources are zero mean discrete-time sequences, with a covariance matrix $\mathbf{C}_x = E[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_{N_t}$, and they must be mutually independent at a given time and identically distributed.

3.2. BSS's Principle. The BSS objective is to find a matrix $(N_r \times N_t)\mathbf{W}(k)$ such as

$$\mathbf{z}(k) = \mathbf{W}^H(k)\mathbf{y}(k), \tag{3}$$

where the output vector sequence $\mathbf{z}(k)$ contains accurate estimates of the N_t source signals $\mathbf{x}(k)$. Each column of $\mathbf{W}(k)$ allows to find one source. So, all the columns of $\mathbf{W}(k)$ must be orthogonal with one another, in order to obtain independent sources. The matrix channel $\mathbf{H}$ must be unitary, otherwise the received signals will have to be prewhitened. Generally $\mathbf{H}$ is not unitary, so we will consider a prewithening operation which also reduces the dimension of $\mathbf{x}(k)$ from N_r to N_t. An underscore is used here to denote pre-whitened variables. Thus, let $\underline{\mathbf{y}}(k) = \mathbf{F}^H\mathbf{y}(k)$ be the prewhitened received signals, where $\mathbf{F}$ is an $(N_r \times N_t)$ prefilter. Like in [8], the prewhitening algorithm in [12] is used. The initial problem to find $\mathbf{W}(k)$ is then replaced by finding a $(N_t \times N_t)$ matrix $\mathbf{T}(k)$:

$$\begin{aligned}
\mathbf{z}(k) &= \mathbf{T}^H(k)\underline{\mathbf{y}}(k)\\
&= \mathbf{T}^H(k)\mathbf{F}^H(k)\mathbf{y}(k)\\
&= \mathbf{W}^H(k)\mathbf{y}(k)\\
&= \mathbf{W}^H(k)(\mathbf{H}(k)\mathbf{x}(k) + \mathbf{b}(k))\\
&= \Xi^H(k)\mathbf{x}(k) + b'(k),
\end{aligned} \tag{4}$$

where $\mathbf{W} = \mathbf{FT}$ and $\Xi = \mathbf{H}^H(k)\mathbf{FT}$ are a global separation matrix. Unfortunately, the calculated matrix $\mathbf{T}(k)$ separates the sources, except for a possible permutation of $\mathbf{z}(k)$ and an arbitrary phase on each source signal. The different BSS cost functions, described in Section 3.3, are indeed minimized with any permutation and any rotation on sources. Thus after separation, separator outputs present the following form:

$$\mathbf{z}(k) = \Phi\Gamma\mathbf{x}(k), \tag{5}$$

where Φ is a diagonal matrix and Γ is a permutation matrix which represent the arbitrary phase and the permutation, respectively.

In order to find $\mathbf{T}(k)$, several cost functions can be used. The next section recalls the well-known (Constant Modulus) CM cost function and (Multimodulus) MM cost function.

3.3. Cost Function

3.3.1. CM Cost Function. Many telecommunication signals have a constant modulus (PSK, 4-QAM), normalized to one generally. If the samples of one source are shown on a complex plan, every samples are on an unitary circle, but it is not true if we show a linear combination of several sources. It is this property that cost function CM exploits. Thus, the algorithms using the CM cost function make use of the sources' constant-modulus property:

$$\begin{aligned} J(\mathbf{T}(k)) &= \sum_{n=1}^{N_t} E\left[\left(|z_n(k)|^2 - R\right)^2\right] \\ &= \sum_{n=1}^{N_t} E\left[\left(\left|\mathbf{t}_n^H(k)\underline{\mathbf{y}}(k)\right|^2 - R\right)^2\right], \end{aligned} \tag{6}$$

where $\mathbf{t}_n$ represents the nth row of $\mathbf{T}$ and $R = R_i = R_r = E[|\Re(\mathbf{x}(k))|^4]/E[|\Re(\mathbf{x}(k))|^2]$. This criterion forces BSS outputs to be on a circle with radius R.

3.3.2. MM Cost Function. The cost function is composed of two cost functions; one cost function for real part ($\Re$) and one for imaginary part ($\Im$) of the equalizer output $\mathbf{z}(k)$:

$$\begin{aligned} J(\mathbf{T}(k)) = \sum_{n=1}^{N_t} \Bigg(& E\left[\left(\Re^2\left(\mathbf{t}_n^H(k)\underline{\mathbf{y}}(k)\right) - R_r\right)^2\right] \\ & + E\left[\left(\Im^2\left(\mathbf{t}_n^H(k)\underline{\mathbf{y}}(k)\right) - R_i\right)^2\right]\Bigg). \end{aligned} \tag{7}$$

The error estimation for the real and imaginary parts is done jointly, and so the cost function uses implicitly the phase of the equalizer output. Thus, the carrier phase recovery is also accomplished with blind equalization, while the equalizer output signal constellation, obtained when using the CM cost function, suffers from an arbitrary phase rotation. So, the MM cost function is most interesting and used later in this paper.

3.3.3. SCM Cost Function. The SCM cost function [5] consists in projection of the equalizer outputs on one dimension (either real or imaginary part). This cost function presents a lower computational complexity compared to the CM and MM cost functions:

$$J(\mathbf{T}(k)) = \sum_{n=1}^{N_t} E\left[\left(\Re^2\left(\mathbf{t}_n^H(k)\underline{\mathbf{y}}(k)\right) - R_r\right)^2\right]. \tag{8}$$

3.4. Implementation of Cost Function.

Several methods exist to implement cost functions. Among them are the stochastic gradient descent and analytical methods.

The algorithms using stochastic gradient are adaptive algorithms which are slow to converge. Furthermore, in varying environments, subsequent signals tend to be less accurate because of a maladjustment. These algorithms, using a stochastic gradient in order to minimize the cost functions CM and MM, are named, respectively, (Constant Modulus Algorithm) CMA [3, 4] and (MultiModulus Algorithm) MMA [1].

Analytical methods exist in both the adaptive form and batch form. With the adaptive form, the algorithm converges quickly, so it is an interesting option in varying environments.

Our choice is to use, for varying environments, an adaptive analytical method so as to minimize the MM cost function. So, a new algorithm named AMMA and which uses this association, is proposed, once a batch analysis has been performed.

3.5. The Analytical Multimodulus Algorithm.

The following cost function should be minimized:

$$J\left(\widetilde{\mathbf{T}}\right) = \sum_{l=1}^{N_t} E\left[\left(\Re^2(z_l(k)) - R_r\right)^2 + \left(\Im^2(z_l(k)) - R_i\right)^2\right]. \tag{9}$$

In order to bring out the real and imaginary parts $z_l(k)$, this last term is written as

$$\begin{aligned} z_l(k) &= \mathbf{t}_l^H \underline{\mathbf{y}}(k), \quad l \in \{1, \ldots, N_t\}, \\ &= \widetilde{\mathbf{t}}_l^T \widetilde{\underline{\mathbf{y}}}(k) + j\widetilde{\mathbf{t}}_l^T \overline{\underline{\mathbf{y}}}(k), \end{aligned} \tag{10}$$

where $\widetilde{\mathbf{t}}_l = \begin{pmatrix} \Re(\mathbf{t}_l) \\ \Im(\mathbf{t}_l) \end{pmatrix}$, $\widetilde{\underline{\mathbf{y}}}(k) = \begin{pmatrix} \Re(\underline{\mathbf{y}}(k)) \\ \Im(\underline{\mathbf{y}}(k)) \end{pmatrix}$, $\overline{\underline{\mathbf{y}}}(k) = \begin{pmatrix} \Im(\underline{\mathbf{y}}(k)) \\ -\Re(\underline{\mathbf{y}}(k)) \end{pmatrix}$, $\overline{\underline{\mathbf{y}}}(k) = \mathbf{G}\widetilde{\underline{\mathbf{y}}}(k)$ with $\mathbf{G} = \begin{pmatrix} \mathbf{0}_{N_t} & \mathbf{I}_{N_t} \\ -\mathbf{I}_{N_t} & \mathbf{0}_{N_t} \end{pmatrix}$.

The matrix $\mathbf{G}$ has the following properties:

$$\begin{aligned} -\mathbf{G} &= \mathbf{G}^T = \mathbf{G}^{-1}, \\ \mathbf{G} \otimes \mathbf{G} &= (\mathbf{G} \otimes \mathbf{G})^T = (\mathbf{G} \otimes \mathbf{G})^{-1}. \end{aligned} \tag{11}$$

Using Kronecker product properties and taking into account [9], $\Re^2(z_n(k))$ and $\Im^2(z_n(k))$ can be written as

$$\begin{aligned} \Re^2(z_l(k)) &= \widetilde{\mathbf{t}}_l^T \left(\widetilde{\underline{\mathbf{y}}}(k)\widetilde{\underline{\mathbf{y}}}^T(k)\right)\widetilde{\mathbf{t}}_l \\ &= \left(\widetilde{\underline{\mathbf{y}}}(k) \otimes \widetilde{\underline{\mathbf{y}}}(k)\right)^T \left(\widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l\right), \\ \Im^2(z_l(k)) &= \widetilde{\mathbf{t}}_l^T \left(\overline{\underline{\mathbf{y}}}(k)\overline{\underline{\mathbf{y}}}^T(k)\right)\widetilde{\mathbf{t}}_l \\ &= \left(\overline{\underline{\mathbf{y}}}(k) \otimes \overline{\underline{\mathbf{y}}}(k)\right)^T \left(\widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l\right) \\ &= \left[(\mathbf{G} \otimes \mathbf{G})\left(\widetilde{\underline{\mathbf{y}}}(k) \otimes \widetilde{\underline{\mathbf{y}}}(k)\right)\right]^T \left(\widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l\right). \end{aligned} \tag{12}$$

3.5.1. The Batch Analysis. Let us denote N a fixed number of received signals, the N rows of $(\underline{\widetilde{\mathbf{y}}}(k) \otimes \underline{\widetilde{\mathbf{y}}}(k))^T$ and $(\underline{\overline{\mathbf{y}}}(k) \otimes \underline{\overline{\mathbf{y}}}(k))^T$ are then, respectively, stacked in the $N \times (2N_t)^2$ matrix $\widetilde{\mathbf{P}}$ and $\overline{\mathbf{P}}$.

Using the Khatri-Rao product ($\mathbf{A} \circ \mathbf{B} = (a_1 \otimes b_1\ a_2 \otimes b_2 \cdots)$), $\widetilde{\mathbf{P}}$ and $\overline{\mathbf{P}}$ can be written as $\widetilde{\mathbf{P}} = (\underline{\widetilde{\mathbf{Y}}} \circ \underline{\widetilde{\mathbf{Y}}})^T$ and $\overline{\mathbf{P}} = (\underline{\overline{\mathbf{Y}}} \circ \underline{\overline{\mathbf{Y}}})^T$. Let us define the $(2N_t)^2 \times 1$ vectors $\mathbf{d}_l = \widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l$ and $\mathbf{1} = (1, \ldots, 1)^T$, the cost function is then given by

$$J\left(\widetilde{\mathbf{T}}\right) = \frac{1}{N}\sum_{l=1}^{N_t}\left[\left|\left|\widetilde{\mathbf{P}}\mathbf{d}_l - R \cdot \mathbf{1}\right|\right|^2 + \left|\left|\overline{\mathbf{P}}\mathbf{d}_l - R \cdot \mathbf{1}\right|\right|^2\right] \tag{13}$$

with $R = R_i = R_r$.

The linear system can be rewritten by using a QR decomposition [7, 8].

We have $\mathbf{Q}$, an $N_s \times N_s$ unitary matrix, and such as $\mathbf{Q1} = \sqrt{N_s}\binom{1}{\mathbf{0}}$.

After doing the QR factorization of $(\mathbf{1}\ \widetilde{\mathbf{P}})$ and $(\mathbf{1}\ \overline{\mathbf{P}})$ and applying $\mathbf{Q}$ to $(\mathbf{1}\ \widetilde{\mathbf{P}})$ and $(\mathbf{1}\ \overline{\mathbf{P}})$, we obtain

$$\begin{aligned}\mathbf{Q}\left(\mathbf{1}\ \widetilde{\mathbf{P}}\right) &= \sqrt{N_s}\begin{pmatrix}1 & \widetilde{\boldsymbol{\gamma}}^T\\ \mathbf{0} & \widetilde{\mathbf{O}}\end{pmatrix},\\ \mathbf{Q}(\mathbf{1}\ \overline{\mathbf{P}}) &= \sqrt{N_s}\begin{pmatrix}1 & \overline{\boldsymbol{\gamma}}^T\\ \mathbf{0} & \overline{\mathbf{O}}\end{pmatrix}\end{aligned} \tag{14}$$

with $\mathbf{1}$ the column vector composed of N_s ones, $\mathbf{0}$ a column vector composed of $N_s - 1$ zeros, $\widetilde{\boldsymbol{\gamma}}^T$ and $\overline{\boldsymbol{\gamma}}^T$ two $1 \times (2N_t)^2$ rows vectors and $\widetilde{\mathbf{O}}$ and $\overline{\mathbf{O}}$ two $N_s - 1 \times (2N_t)^2$ matrices, then

$$\begin{aligned}\widetilde{\mathbf{P}}\mathbf{d}_l = R\mathbf{1} \Longleftrightarrow \mathbf{Q}\left(\mathbf{1}\ \widetilde{\mathbf{P}}\right)\begin{pmatrix}-R\\ \mathbf{d}_l\end{pmatrix} = \mathbf{0} \Longleftrightarrow \begin{cases}\widetilde{\boldsymbol{\gamma}}^T\mathbf{d}_l = R,\\ \widetilde{\mathbf{O}}\mathbf{d}_l = \mathbf{0},\end{cases}\\ \overline{\mathbf{P}}\mathbf{d}_l = R\mathbf{1} \Longleftrightarrow \mathbf{Q}(\mathbf{1}\ \overline{\mathbf{P}})\begin{pmatrix}-R\\ \mathbf{d}_l\end{pmatrix} = \mathbf{0} \Longleftrightarrow \begin{cases}\overline{\boldsymbol{\gamma}}^T\mathbf{d}_l = R,\\ \overline{\mathbf{O}}\mathbf{d}_l = \mathbf{0}.\end{cases}\end{aligned} \tag{15}$$

The cost function becomes

$$\begin{aligned}&J_{\mathrm{AMMA}}(\mathbf{d}_l)\\ &= \frac{1}{N_s}\sum_{l=1}^{N_t}\left[\left|\widetilde{\boldsymbol{\gamma}}^T\mathbf{d}_l - R\right|^2 + \left|\overline{\boldsymbol{\gamma}}^T\mathbf{d}_l - R\right|^2 + \left|\left|\widetilde{\mathbf{O}}\mathbf{d}_l\right|\right|^2 + \left|\left|\overline{\mathbf{O}}\mathbf{d}_l\right|\right|^2\right],\end{aligned} \tag{16}$$

where $|\widetilde{\boldsymbol{\gamma}}^T\mathbf{d}_l - R|^2$ and $|\overline{\boldsymbol{\gamma}}^T\mathbf{d}_l - R|^2$ allow to avoid the trivial solution $\mathbf{d}_l = \mathbf{0}$. By squaring (14), the expressions of vectors $\widetilde{\boldsymbol{\gamma}}$ and $\overline{\boldsymbol{\gamma}}$, and of matrices $\widetilde{\mathbf{C}} = \widetilde{\mathbf{O}}^T\widetilde{\mathbf{O}}$ and $\overline{\mathbf{C}} = \overline{\mathbf{O}}^T\overline{\mathbf{O}}$ are

$$\begin{aligned}\widetilde{\boldsymbol{\gamma}} &= \frac{1}{N_s}\widetilde{\mathbf{P}}^T\mathbf{1} = \frac{1}{N_s}\sum_{k=1}^{N_s}\underline{\widetilde{\mathbf{y}}}(k) \otimes \underline{\widetilde{\mathbf{y}}}(k),\\ \widetilde{\mathbf{C}} &= \widetilde{\mathbf{O}}^T\widetilde{\mathbf{O}} = \frac{1}{N_s}\widetilde{\mathbf{P}}^T\widetilde{\mathbf{P}} - \frac{1}{N_s}\widetilde{\mathbf{P}}^T\mathbf{1} \cdot \frac{1}{N_s}\mathbf{1}^T\widetilde{\mathbf{P}}\\ &= \frac{1}{N_s}\sum_{k=1}^{N_s}\left(\underline{\widetilde{\mathbf{y}}}(k) \otimes \underline{\widetilde{\mathbf{y}}}(k)\right)\left(\underline{\widetilde{\mathbf{y}}}(k) \otimes \underline{\widetilde{\mathbf{y}}}(k)\right)^T\\ &\quad - \frac{1}{N_s^2}\left(\sum_{k=1}^{N_s}\underline{\widetilde{\mathbf{y}}}(k) \otimes \underline{\widetilde{\mathbf{y}}}(k)\right)\left(\sum_{k=1}^{N_s}\underline{\widetilde{\mathbf{y}}}(k) \otimes \underline{\widetilde{\mathbf{y}}}(k)\right)^T,\\ \overline{\boldsymbol{\gamma}} &= \frac{1}{N_s}\overline{\mathbf{P}}^T\mathbf{1} = \frac{1}{N_s}\sum_{k=1}^{N_s}\underline{\overline{\mathbf{y}}}(k) \otimes \underline{\overline{\mathbf{y}}}(k),\\ \overline{\mathbf{C}} &= \overline{\mathbf{O}}^T\overline{\mathbf{O}} = \frac{1}{N_s}\overline{\mathbf{P}}^T\overline{\mathbf{P}} - \frac{1}{N_s}\overline{\mathbf{P}}^T\mathbf{1} \cdot \frac{1}{N_s}\mathbf{1}^T\overline{\mathbf{P}}\\ &= \frac{1}{N_s}\sum_{k=1}^{N_s}\left(\underline{\overline{\mathbf{y}}}(k) \otimes \underline{\overline{\mathbf{y}}}(k)\right)\left(\underline{\overline{\mathbf{y}}}(k) \otimes \underline{\overline{\mathbf{y}}}(k)\right)^T\\ &\quad - \frac{1}{N_s^2}\left(\sum_{k=1}^{N_s}\underline{\overline{\mathbf{y}}}(k) \otimes \underline{\overline{\mathbf{y}}}(k)\right)\left(\sum_{k}\underline{\overline{\mathbf{y}}}(k) \otimes \underline{\overline{\mathbf{y}}}(k)\right)^T.\end{aligned} \tag{17}$$

By using the definition of the estimated covariance matrices of $\underline{\widetilde{\mathbf{y}}}(k)$ and $\underline{\overline{\mathbf{y}}}(k)$, named, respectively, $\hat{\mathbf{R}}_{\underline{\widetilde{y}}}$ and $\hat{\mathbf{R}}_{\underline{\overline{y}}}$, and by using the property of the Kronecker product, we obtain the following equalities:

$$\begin{aligned}\widetilde{\boldsymbol{\gamma}} &= \mathrm{vec}\left(\frac{1}{N_s}\sum_{k=1}^{N_s}\underline{\widetilde{\mathbf{y}}}(k)\underline{\widetilde{\mathbf{y}}}^T(k)\right) = \mathrm{vec}\left(\hat{\mathbf{R}}_{\underline{\widetilde{y}}}\right),\\ \overline{\boldsymbol{\gamma}} &= \mathrm{vec}\left(\frac{1}{N_s}\sum_{k=1}^{N_s}\underline{\overline{\mathbf{y}}}(k)\underline{\overline{\mathbf{y}}}^T(k)\right) = \mathrm{vec}\left(\hat{\mathbf{R}}_{\underline{\overline{y}}}\right).\end{aligned} \tag{18}$$

Thus

$$\begin{aligned}\widetilde{\boldsymbol{\gamma}}^T\mathbf{d}_l &= \mathrm{vec}\left(\mathbf{R}_{\underline{\widetilde{y}}}\right)^T(\mathbf{t}_l \otimes \mathbf{t}_l) = \mathbf{t}_l^T\mathbf{R}_{\underline{\widetilde{y}}}\mathbf{t}_l,\\ \overline{\boldsymbol{\gamma}}^T\mathbf{d}_l &= \mathrm{vec}\left(\mathbf{R}_{\underline{\overline{y}}}\right)^T(\mathbf{t}_l \otimes \mathbf{t}_l) = \mathbf{t}_l^T\mathbf{R}_{\underline{\overline{y}}}\mathbf{t}_l.\end{aligned} \tag{19}$$

The criterion J_{AMMA} is minimized when $\widetilde{\boldsymbol{\gamma}}^T\mathbf{d}_l = R$, $\overline{\boldsymbol{\gamma}}^T\mathbf{d}_l = R$, $\|\widetilde{\mathbf{O}}\mathbf{d}_l\|^2 = 0$, and $\|\overline{\mathbf{O}}\mathbf{d}_l\|^2 = 0$. Afterwards, we consider $\widetilde{\boldsymbol{\gamma}}^T\mathbf{d}_l = R$, $\overline{\boldsymbol{\gamma}}^T\mathbf{d}_l = R$ as constraints. Since received signals are pre-whitened and $\mathbf{R}_{\underline{\widetilde{y}}} = \mathbf{R}_{\underline{\overline{y}}} = \mathbf{I}$ thus

$$\begin{aligned}\widetilde{\boldsymbol{\gamma}}^T\mathbf{d}_l = \overline{\boldsymbol{\gamma}}^T\mathbf{d}_l &= \mathbf{t}_l^T\mathbf{R}_{\underline{\widetilde{y}}}\mathbf{t}_l = \mathbf{t}_l^T\mathbf{R}_{\underline{\overline{y}}}\mathbf{t}_l\\ &= \mathbf{t}_l^T\mathbf{t}_l = \|\mathbf{t}_l\|^2 = R.\end{aligned} \tag{20}$$

Moreover, $\|\mathbf{d}_l\|^2 = (\mathbf{t}_l \otimes \mathbf{t}_l)^T(\widetilde{\mathbf{t}}_l \otimes \mathbf{t}_l) = \|\mathbf{t}_l\|^4 = R^2$. Thus, $\widetilde{\boldsymbol{\gamma}}^T\mathbf{d}_l = \overline{\boldsymbol{\gamma}}^T\mathbf{d}_l = R \Leftrightarrow \|\mathbf{d}_l\| = R$, and to a scaling which is not important the optimization problem becomes

$$\begin{aligned}\mathbf{t}_l &= \arg \min_{\substack{\mathbf{d}_l(k)=\mathbf{t}_l(k)\otimes\mathbf{t}_l(k)\\ \|\mathbf{d}_l(k)\|=R}} \sum_{l=1}^{N_t}\left(\left\|\widetilde{\mathbf{O}}\mathbf{d}_l\right\|^2 + \left\|\overline{\mathbf{O}}\mathbf{d}_l\right\|^2\right)\\ &= \arg \min_{\substack{\mathbf{d}_l(k)=\mathbf{t}_l(k)\otimes\mathbf{t}_l(k)\\ \|\mathbf{d}_l(k)\|=R}} \sum_{l=1}^{N_t}\left(\mathbf{d}_l^T\widetilde{\mathbf{O}}^T\widetilde{\mathbf{O}}\mathbf{d}_l + \mathbf{d}_l^T\overline{\mathbf{O}}^T\overline{\mathbf{O}}\mathbf{d}_l\right) \qquad (21)\\ &= \arg \min_{\substack{\mathbf{d}_l(k)=\mathbf{t}_l(k)\otimes\mathbf{t}_l(k)\\ \|\mathbf{d}_l(k)\|=R}} \sum_{l=1}^{N_t}\mathbf{d}_l^T(k)\left(\widetilde{\mathbf{C}}(k) + \overline{\mathbf{C}}(k)\right)\mathbf{d}_l(k).\end{aligned}$$

In the appendix, it is shown that vectors $\mathbf{d}_l$, for all $l \in \{1 \cdots N_t\}$, which minimize the function $\sum_{l=1}^{N_t}\mathbf{d}_l^T(\widetilde{\mathbf{C}} + \overline{\mathbf{C}})\mathbf{d}_l$ under the hypotheses $\mathbf{d}_l = \widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l$ and $\|\widetilde{\mathbf{d}}_l\| = R$, are in the image of $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$ (c.f. Theorem 2 and Lemma 2). Moreover, the minimization of $\sum_{l=1}^{N_t}\mathbf{d}_l^T(\widetilde{\mathbf{C}} + \overline{\mathbf{C}})\mathbf{d}_l$ is equivalent to minimize $\mathbf{d}_l^T(\widetilde{\mathbf{C}} + \overline{\mathbf{C}})\mathbf{d}_l$, for all $l \in \{1 \cdots 2\}$ (c.f. Theorem 1 and Lemma 1).

Now, we will see how to create vectors $\mathbf{d}_l$ that present a Kronecker structure, in the image of $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$.

3.5.2. The Batch Method. The vectors $\mathbf{d}_l$ which minimize $\mathbf{d}_l^T(\widetilde{\mathbf{C}} + \overline{\mathbf{C}})\mathbf{d}_l$ are eigenvectors associated with the smallest eigenvalues. However, these vectors have not necessarily a Kronecker structure. Thus, in order to obtain the vectors $\widetilde{\mathbf{t}}_l$ in the image and to have a Kronecker struture, we search a $N_t \times N_t$ full rank matrix $\boldsymbol{\Lambda}$ that relates the two bases such so that

$$\widetilde{\mathbf{T}} = \arg\min \left\|\mathbf{D} - \left(\widetilde{\mathbf{T}} \circ \widetilde{\mathbf{T}}\right)\boldsymbol{\Lambda}\right\|_F^2 \qquad (22)$$

with $\widetilde{\mathbf{T}} = (\widetilde{\mathbf{t}}_1, \ldots, \widetilde{\mathbf{t}}_{N_t})$ and $\mathbf{D} = (\mathbf{d}_1, \ldots, \mathbf{d}_{N_t})$ the $2N_t \times N_t$ and $(2N_t)^2 \times N_t$ matrix, respectively. Using the property of the Kronecker product, this minimization problem can be rewritten as

$$\begin{aligned}\left\|\mathbf{D} - \left(\widetilde{\mathbf{T}} \circ \widetilde{\mathbf{T}}\right)\boldsymbol{\Lambda}\right\|_F^2 &= \sum_{n=1}^{N_t}\left\|\mathbf{d}_n - \left(\widetilde{\mathbf{T}} \circ \widetilde{\mathbf{T}}\right)\boldsymbol{\lambda}_n\right\|_F^2\\ &= \sum_{n=1}^{N_t}\left\|\mathbf{D}_n - \widetilde{\mathbf{T}}\boldsymbol{\Lambda}_n\widetilde{\mathbf{T}}\right\|_F^2\end{aligned} \qquad (23)$$

with

(i) $\boldsymbol{\lambda}_n$ the nth vector of the $\boldsymbol{\Lambda}$ matrix,

(ii) $\boldsymbol{\Lambda}_n = \mathrm{diag}(\boldsymbol{\lambda}_n)$ is a diagonal matrix,

(iii) $\mathbf{D}_n = \mathrm{vec}(\mathbf{d}_n)$.

Please note that if the matrix $\widetilde{\mathbf{T}}$ was square, the minimization of (23) would be equivalent to a joint diagonalization. In order to obtain the $\widetilde{\mathbf{T}}$ matrix, let us suppose in first phase that it is indeed square. Thus, with the aid of joint diagonalization of $\mathbf{D}_n$ matrices, a $2N_t \times 2N_t\widetilde{\mathbf{T}}'$ matrix is obtained. This diagonalization can be done by using the algorithm described in [13]. The columns of the matrix $\widetilde{\mathbf{T}}'$ are dependent; the $2N_t \times N_t$ matrix $\widetilde{\mathbf{T}}$ is extracted by keeping the columns of $\widetilde{\mathbf{T}}'$ linearly independent.

Now, this algorithm is implemented thanks to an adaptive method. So, we must estimate in an adaptive manner, $\widetilde{\mathbf{C}}$, $\overline{\mathbf{C}}$, and the eigenvectors of $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$ associated with the smallest eigenvalues. The methods used are described in the next section. Then the adaptive tracking of the minor subspace of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$ is performed thanks to the use of the NOOJA algorithm. Finally the adaptive update of the joint diagonalization is realized.

3.5.3. The Adaptive Method. The function $\mathbf{d}_l^T(\widetilde{\mathbf{C}} + \overline{\mathbf{C}})\mathbf{d}_l$ will be minimized, at each time k, but iteratively

$$\widetilde{\mathbf{t}}_l(k) = \min_{\substack{\mathbf{d}_l=\widetilde{\mathbf{t}}_l\otimes\widetilde{\mathbf{t}}_l\\ \|\mathbf{d}_l\|=R}} \mathbf{d}_l^T(k)\left(\widetilde{\mathbf{C}}(k) + \overline{\mathbf{C}}(k)\right)\mathbf{d}_l(k), \quad \forall l \in \{1, \ldots, N_t\}. \qquad (24)$$

The matrices $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$ were defined by a batch of N samples. In this adaptive method they are converted into an exponential window (λ-scaling). The details are shown in [8].

The following method used to estimate iteratively the matrix $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$ is a generalization of the Van der Veen method [9]

$$\widetilde{\mathbf{C}}(k) = \lambda\widetilde{\mathbf{C}}(k-1) + \beta(k)\widetilde{\mathbf{c}}(k)\widetilde{\mathbf{c}}^T(k) = f\left(\widetilde{\mathbf{C}}(k-1), \widetilde{\mathbf{c}}(k)\right), \qquad (25)$$

$$\overline{\mathbf{C}}(k) = \lambda\overline{\mathbf{C}}(k-1) + \beta(k)\overline{\mathbf{c}}(k)\overline{\mathbf{c}}^T(k) = f\left(\overline{\mathbf{C}}(k-1), \overline{\mathbf{c}}(k)\right). \qquad (26)$$

Then, $\widetilde{\mathbf{C}}(k)$ and $\overline{\mathbf{C}}(k)$ are, respectively, the estimations of the autocorrelation matrices for the signals $\widetilde{\mathbf{c}}(k)$ and $\overline{\mathbf{c}}(k)$, with $0 < \lambda < 1$, $\beta(k) = (\alpha(k-1)/\alpha(k))\lambda(1-\lambda)$, $\alpha(k) = \lambda\alpha(k-1) + 1 - \lambda$ and

$$\begin{aligned}\widetilde{\mathbf{c}}(k) &= \underline{\widetilde{\mathbf{y}}}(k) \otimes \underline{\widetilde{\mathbf{y}}}(k) - \frac{\widetilde{\mathbf{p}}(k-1)}{\alpha(k-1)},\\ \overline{\mathbf{c}}(k) &= \underline{\overline{\mathbf{y}}}(k) \otimes \underline{\overline{\mathbf{y}}}(k) - \frac{\overline{\mathbf{p}}(k-1)}{\alpha(k-1)},\end{aligned} \qquad (27)$$

where

$$\begin{aligned}\widetilde{\mathbf{p}}(k) &= \lambda\widetilde{\mathbf{p}}(k-1) + (1-\lambda)\left(\underline{\widetilde{\mathbf{y}}}(k) \otimes \underline{\widetilde{\mathbf{y}}}(k)\right),\\ \overline{\mathbf{p}}(k) &= \lambda\overline{\mathbf{p}}(k-1) + (1-\lambda)\left(\underline{\overline{\mathbf{y}}}(k) \otimes \underline{\overline{\mathbf{y}}}(k)\right).\end{aligned} \qquad (28)$$

We saw that the solution of this minimization is in the image of $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$. In order to obtain these vectors in an adaptive manner, a subspace tracking algorithm is used, associated with an adaptive joint diagonalization.

Subspace Tracking. Used alone, the (Normalized Orthogonal Oja) NOOJA algorithm of [14] is utilized to track

minor subspace spanned by the eigenvectors relating to the smallest eigenvalues. But, when associated with an adaptive joint diagonalization, as described in the next section, the NOOJA algorithm tracks the minor subspace spanned by the eigenvectors in the image of $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$ relating to the smallest eigenvalues which are different from 0. This algorithm extracts adaptively the minor subspace $\mathbf{V}$ relating to the input signal $\mathbf{r}(k)$'s autocorrelation matrix $\mathbf{R}$, by maximization the following cost function:

$$\begin{aligned} J(\mathbf{V}) &= E\left\|\mathbf{r}(k) - \mathbf{V}\mathbf{V}^T\mathbf{r}(k)\right\|^2 \\ &= \operatorname{tr}(\mathbf{R}) - 2\operatorname{tr}\left(\mathbf{V}^T\mathbf{R}\mathbf{V}\right) + \operatorname{tr}\left(\mathbf{V}^T\mathbf{R}\mathbf{V}\mathbf{V}^T\mathbf{V}\right). \end{aligned} \tag{29}$$

In our problem, the minor subspace of $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$ is searched. So, in order to use the NOOJA algorithm, $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$ must be expressed as $\widetilde{\mathbf{C}}+\overline{\mathbf{C}} = E[\mathbf{c}\mathbf{c}^T]$. But since the function f (cf. (25) and (26)) is not bilinear, the vector c is unknown. Instead of searching the minor subspace of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$, the minor subspace of both $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$ will be searched. Since $\overline{\mathbf{C}}$ can not be expressed as

$$\overline{\mathbf{C}} = (\mathbf{G} \otimes \mathbf{G})\widetilde{\mathbf{C}}(\mathbf{G} \otimes \mathbf{G})^{-1}. \tag{30}$$

The matrices $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$ are similar. They have therefore the same image. In order to find the minor subspace $\mathbf{D} = (\mathbf{d}_1, \ldots, \mathbf{d}_{N_t})$ spanned by the eigenvectors and relating to the smallest eigenvalues different from 0 for both $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$, we propose to maximize the following cost function:

$$\begin{aligned} J(\mathbf{D}) &= E\left\|\widetilde{\mathbf{c}}(k) - \mathbf{D}\mathbf{D}^T\widetilde{\mathbf{c}}(k)\right\|^2 + E\left\|\overline{\mathbf{c}}(k) - \mathbf{D}\mathbf{D}^T\overline{\mathbf{C}}(k)\right\|^2 \\ &= \operatorname{tr}\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right) - 2\operatorname{tr}\left(\mathbf{D}^T\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right)\mathbf{D}\right) \\ &\quad + \operatorname{tr}\left(\mathbf{D}^T\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right)\mathbf{D}\mathbf{D}^T\mathbf{D}\right). \end{aligned} \tag{31}$$

Since $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$ have the same image, the first term of the cost function is sufficient to obtain convergence with a constant channel. Using only the matrix $\widetilde{\mathbf{C}}$ is equivalent to minimize the cost function Simplified CM [5] in analytical way. This simplification reduces the complexity of the adaptive AMMA and we name it the adaptive Analytical Simplified CMA (ASCMA). Please note that the ASCMA batch analysis is the same with the AMMA batch analysis performed in Section 3.5.1. However, the study is solely about $\widetilde{\mathbf{C}}$ rather than $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$.

The maximization of $J(\mathbf{D})$ may be achieved by using the gradient-descent technique:

$$\mathbf{D}(k+1) = \mathbf{D}(k) - \beta\nabla_J(k), \tag{32}$$

where

$$\nabla_J = \left(-2\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right) + \left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}(k)\right)\mathbf{D}\mathbf{D}^T + \mathbf{D}\mathbf{D}^T\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right)\right)\mathbf{D}. \tag{33}$$

The time instant k has been removed to simplify the equations.

By using $\mathbf{D}^T(k)\mathbf{D}(k) = \mathbf{I}$, the cost function can be simplified as

$$J(\mathbf{D}) = \operatorname{tr}\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right) - \operatorname{tr}\left(\mathbf{D}^T\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right)\mathbf{D}\right). \tag{34}$$

By replacing (32) in (34) and by doing the gradient of $J(\mathbf{D}(k+1))$ with respect to $\beta(k)$, the optimal stepsize $\beta_{\text{opt}}(k)$ is obtained

$$\beta_{\text{opt}}(k) = \frac{\operatorname{tr}\left(\mathbf{D}^T(k)\left(\widetilde{\mathbf{C}}(k) + \overline{\mathbf{C}}(k)\right)\nabla_J(k)\right)}{\operatorname{tr}\left(\nabla_J^T(k)\left(\widetilde{\mathbf{C}}(k) + \overline{\mathbf{C}}(k)\right)\nabla_J(k)\right)}. \tag{35}$$

Keeping in mind that $\mathbf{D}(k)$ is orthogonal at each iteration, that is, $\mathbf{D}^T(k)\mathbf{D}(k) = \mathbf{I}$, and replacing $\widetilde{\mathbf{C}}(k)$ and $\overline{\mathbf{C}}(k)$ by the estimate $\widetilde{\mathbf{c}}(k)\widetilde{\mathbf{c}}^T(k)$ and $\overline{\mathbf{c}}(k)\overline{\mathbf{c}}^T(k)$, respectively, we can write

$$\nabla_J(k) = \widetilde{\mathbf{p}}(k)\widetilde{\mathbf{z}}^T(k) + \overline{\mathbf{p}}(k)\overline{\mathbf{z}}^T(k), \tag{36}$$

where $\widetilde{\mathbf{z}}(k) = \mathbf{D}^T(k)\widetilde{\mathbf{c}}(k)$, $\overline{\mathbf{z}}(k) = \mathbf{D}^T(k)\overline{\mathbf{c}}(k)$, $\widetilde{\mathbf{p}}(k) = \widetilde{\mathbf{z}}(k) - \mathbf{D}(k)\widetilde{\mathbf{z}}(k)$, and $\overline{\mathbf{p}}(k) = \overline{\mathbf{z}}(k) - \mathbf{D}(k)\overline{\mathbf{z}}(k)$. The optimal step size becomes

$$\beta_{\text{opt}} = \frac{\left(\|\widetilde{\mathbf{c}}\|^2 - \|\widetilde{\mathbf{z}}\|^2\right)\|\widetilde{\mathbf{z}}\|^2 + \left(\|\overline{\mathbf{c}}\|^2 - \|\overline{\mathbf{z}}\|^2\right)\|\overline{\mathbf{z}}\|^2}{\left(\|\widetilde{\mathbf{c}}\|^2 - \|\widetilde{\mathbf{z}}\|^2\right)^2\|\widetilde{\mathbf{z}}\|^2 + \left(\|\overline{\mathbf{c}}\|^2 - \|\overline{\mathbf{z}}\|^2\right)^2\|\overline{\mathbf{z}}\|^2}. \tag{37}$$

The instant k has been removed in the $\beta_{\text{opt}}(k)$'s expression in order to simplify the equation.

The updating of $\mathbf{D}(k+1)$ is performed as

$$\mathbf{D}'(k+1) = \mathbf{D}(k) - \beta_{\text{opt}}(k)\left(\widetilde{\mathbf{p}}(k)\widetilde{\mathbf{z}}^T(k) + \overline{\mathbf{p}}(k)\overline{\mathbf{z}}^T(k)\right), \tag{38}$$

where the matrix $\mathbf{D}(k+1)$ is obtained thanks to a Gram-Schmidt orthogonalization of $\mathbf{D}'(k+1)$. The algorithm obtained is described in Algorithm 1 .

Once the eigenvectors relating to the smallest eigenvalues of both $\widetilde{\mathbf{C}}(k)$ and $\overline{\mathbf{C}}(k)$ are obtained, the constraint $\mathbf{d}_l = \widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l$ must be satisfied. In other words, $\mathbf{d}_l$ must have a Kronecker structure and must be in the image of $\widetilde{\mathbf{C}}(k) + \overline{\mathbf{C}}(k)$.

Adaptive Update of the Joint Diagonalization. The method used to satisfy the constraint $\mathbf{d}_l = \widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l$ is done as in [9]. $\widetilde{\mathbf{T}} \circ \widetilde{\mathbf{T}}$ is computed and regarded as the current estimate of the subspace basis ($\mathbf{D} = \widetilde{\mathbf{T}} \circ \widetilde{\mathbf{T}}$). Using this basis, the subspace update is performed by using the "Normalized Orthogonal OJA" giving $\mathbf{D}$. The last step of this algorithm is the mapping of the columns $\mathbf{d}_l$ of $\mathbf{D}$ to a Kronecker-product:

$$\begin{aligned} \mathbf{d}_l = \widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l &= \operatorname{vec}\left(\widetilde{\mathbf{t}}_l\widetilde{\mathbf{t}}_l^T\right), \\ \operatorname{vec}^{-1}(\mathbf{d}_l) = \mathbf{D}_l &= \widetilde{\mathbf{t}}_l\widetilde{\mathbf{t}}_l^T. \end{aligned} \tag{39}$$

Then a power iteration [15] is applied. This iteration can be written as

$$\widetilde{\mathbf{t}}_l(k+1) = \mathbf{D}_l\widetilde{\mathbf{t}}_l(k). \tag{40}$$

At the next iteration, the matrix $\mathbf{D}$ in the image, which has a Kronecker structure, is obtained with $\mathbf{D} = \widetilde{\mathbf{T}} \circ \widetilde{\mathbf{T}}$. Then, $\mathbf{D}$ is transmitted at the NOOJA algorithm, allowing to vectors $\mathbf{d}_l$ obtained in output NOOJA to be in the nullspace and to have a Kronecker struture.

Initialize $\mathbf{D}(0)$
for $k = 1, 2, \ldots$
$\widetilde{\mathbf{z}} = \mathbf{D}^T(k)\widetilde{\mathbf{c}},\ \overline{\mathbf{z}} = \mathbf{D}^T(k)\overline{\mathbf{c}},\ \widetilde{\mathbf{p}} = \widetilde{\mathbf{z}} - \mathbf{D}(k)\widetilde{\mathbf{z}},\ \overline{\mathbf{p}} = \overline{\mathbf{z}} - \mathbf{V}(k)\overline{\mathbf{z}}$
compute β_{opt} by using(37)
$\mathbf{D}'(k+1) = \mathbf{D}(k) - \beta_{\text{opt}}(\widetilde{\mathbf{p}}\widetilde{\mathbf{z}}^T + \overline{\mathbf{p}}\overline{\mathbf{z}}^T)$
$\mathbf{d}_1(k+1) = \mathbf{d}_1'(k+1)/\|\mathbf{d}_1'(k+1)\|$
for $q = 2$ to N_t
$\mathbf{s} = [0 \cdots 0]^T$
for $l = 1$ to $q-1$
$\mathbf{s} = \mathbf{s} + (\mathbf{d}_l^T(k+1)\mathbf{d}_q'(k+1))\mathbf{d}_l(k+1)$
end for
$\mathbf{d}_q(k+1) = (\mathbf{d}_q'(k+1) - \mathbf{s})/\|(\mathbf{d}_q'(k+1) - \mathbf{s})\|$
end for
end for

Algorithm 1: Adaptive nullspace of $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$ tracking by using NOOJA.

Initialize $\mathbf{F}, \widetilde{\mathbf{T}}, \overline{\mathbf{p}} = 0,\ \widetilde{\mathbf{p}} = 0,\ \alpha = 0$
for $k = 1, 2, \ldots$
Update F, the prewhitened filter,
by using $y(k)$ and [12] we obtain $\underline{\mathbf{y}}(k) = \mathbf{F}^H\mathbf{y}(k)$
Update real vectors
$\widetilde{\mathbf{y}}(k) = \begin{pmatrix} \Re(\mathbf{y}(k)) \\ \Im(\mathbf{y}(k)) \end{pmatrix},\ \overline{\mathbf{y}}(k) = \begin{pmatrix} -\Im(\mathbf{y}(k)) \\ \Re(\mathbf{y}(k)) \end{pmatrix},\ \widetilde{\mathbf{w}}_l = \begin{pmatrix} \Re(\mathbf{w}) \\ \Im(\mathbf{w}) \end{pmatrix}$
Update vector $\widetilde{\mathbf{c}}$ and $\overline{\mathbf{c}}$
$\widetilde{\mathbf{c}} = \widetilde{\mathbf{y}}(k) \otimes \widetilde{\underline{\mathbf{y}}}(k) - \widetilde{\mathbf{p}}/\alpha,\ \overline{\mathbf{c}} = \overline{\underline{\mathbf{y}}}(k) \otimes \overline{\underline{\mathbf{y}}}(k) - \overline{\mathbf{p}}/\alpha$
$\widetilde{\mathbf{p}} = \lambda\widetilde{\mathbf{p}} + (1-\lambda)(\widetilde{\underline{\mathbf{y}}}(k) \otimes \widetilde{\underline{\mathbf{y}}}(k)),\ \overline{\mathbf{p}} = \lambda\overline{\mathbf{p}} + (1-\lambda)(\overline{\underline{\mathbf{y}}}(k) \otimes \overline{\underline{\mathbf{y}}}(k))$
$\alpha = \lambda\alpha + 1 - \lambda$
$\mathbf{D} = \widetilde{\mathbf{T}} \circ \widetilde{\mathbf{T}}$
Regard $\mathbf{D}$ as a basis of the nullspace of $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$,
and update it using $\widetilde{\mathbf{c}}$ and $\overline{\mathbf{c}}$ (Algorithm 1)
for $l = 1, \ldots, N_t$
$\mathbf{D}_l = \text{vec}^{-1}1(\mathbf{d}_l),\ \widetilde{\mathbf{t}}_l = \mathbf{D}_l\widetilde{\mathbf{t}}_l,\ \widetilde{\mathbf{t}}_l = \widetilde{\mathbf{t}}_l/\|\widetilde{\mathbf{t}}_l\|$
Obtain complex $N_t \times N_t$ **matrix** $\mathbf{T}(k)$
for $m = 1, \ldots, 2N_t$
$\mathbf{t}_{lm} = \widetilde{\mathbf{t}}_{lm} + j\widetilde{\mathbf{t}}_{l(m+N_t)}$
end for
end for
$\mathbf{T} = \text{recond}(\mathbf{T}),\ \widetilde{\mathbf{T}} = (\Re(\mathbf{T})\ \Im(\mathbf{T}))^T$
Estimate the sources
$\widehat{\mathbf{s}}(k) = \mathbf{T}^H(k)\underline{\mathbf{y}}(k)$
end for

Algorithm 2: Adaptive Analytical-MMA.

Estimated Symbols. In order to obtain the transmitted symbols, the complex matrix $\mathbf{T}$ is rebuilt from $\widetilde{\mathbf{T}}$

$$\mathbf{t}_{lm} = \widetilde{\mathbf{t}}_{lm} + j\widetilde{\mathbf{t}}_{l(m+N_t)}, \quad l, m \in 1, \ldots, N_t. \tag{41}$$

In order to prevent the extraction of the same source many times, $\mathbf{T}$ is reconditioned after each update. The technique used here is the computation of a singular value decomposition of $\mathbf{T}$: $\mathbf{T} = \sum \sigma_j \mathbf{u}_j \mathbf{v}_j^H$. Then the singular values of $\mathbf{T}$, which are smaller than 0.5, are replaced by 1

$$\text{recond}(\mathbf{T}) = \sum \sigma_j' \mathbf{u}_j \mathbf{v}_j^H. \tag{42}$$

Then, the matrix $\mathbf{T}$ is applied on the pre-whitened received signals:

$$\widehat{\mathbf{s}}(k) = \mathbf{T}^H(k)\underline{\mathbf{y}}(k). \tag{43}$$

The suggested adaptive-AMMA algorithm is described in Algorithm 2 .

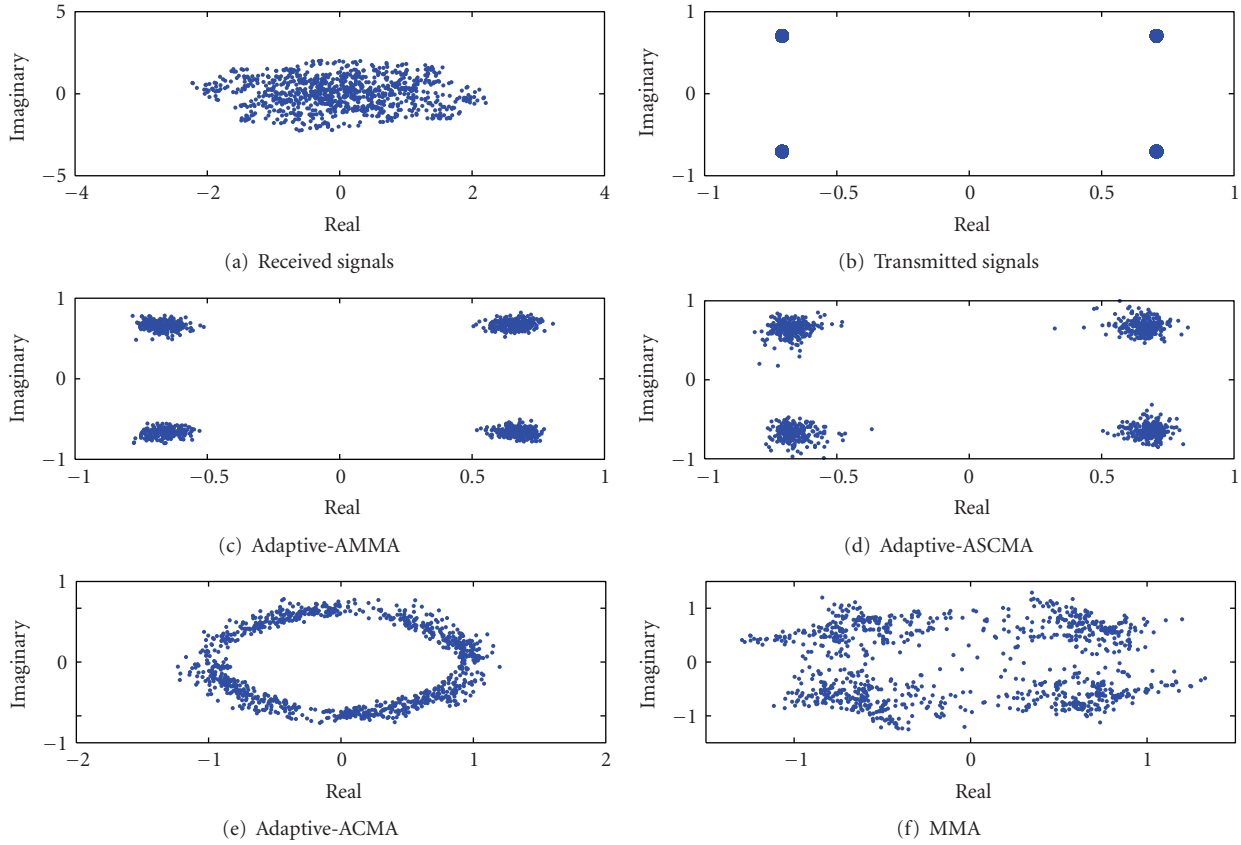


Figure 1: Constellations with a carrier frequency offset. SNR = 30 dB, $N_t = 2$, $N_r = 4$, $\delta_f T_s = 10^{-3}$, $f_d T_s = 0.001$.

4. Simulation Results

The channel used in simulations is a time-selective channel; that is, it is time varying but flat in frequency. The transmitted symbols, resulting from a 4-QAM constellation, are sent through a MIMO Rayleigh fading channel. Two transmitters and four receivers are used ($N_t = 2$ and $N_r = 4$). The performances of proposed adaptive-AMMA are compared with those of the adaptive-ASCMA, the adaptive-ACMA and the MMA, in terms of Bit Error Rate (BER) and Signal to Interference plus Noise Ratio (SINR).

In the simulation results of the Figure 1 , we considered the case of a carrier frequency offset ($\delta_f T_s = 10^{-3}$) with SNR = 30 dB. The equalizer output constellation of the adaptive-ACMA (Figure 1(e)) spins because of the carrier frequency offset. Normally, the MMA (Figure 1(f)) allows to track the carrier frequency offset, but in this case the carrier phase is too high to obtain a good tracking. Only the proposed adaptive-AMMA and adaptive-ASCMA can remove the frequency offset (Figures 1(c) and 1(d)) , without the use of a carrier tracking loop.

Figure 2 shows the average BER as a function of SNR (with the term "average" being here a time average), and this for two sources and 1000 runs, each one presenting different randomly varying channels and different noise sequences. The value of the Doppler shift $f_d T_s$ is $1 \cdot 1e^{-3}$. The simulation here was realized without any carrier frequency offset.

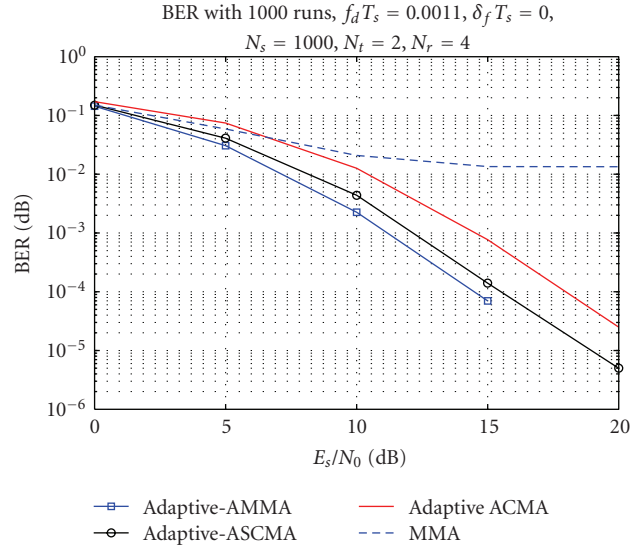


Figure 2: BER without carrier frequency offset versus SNR. $N_t = 2$, $N_r = 4$, $\delta_f T_s = 0$, $f_d T_s = 0.0011$.

The lowest BER on Figure 2 is obtained with the proposed adaptive-AMMA algorithm, followed by the adaptive-ASCMA, then by the adaptive-ACMA and the highest BER is obtained with the MMA.

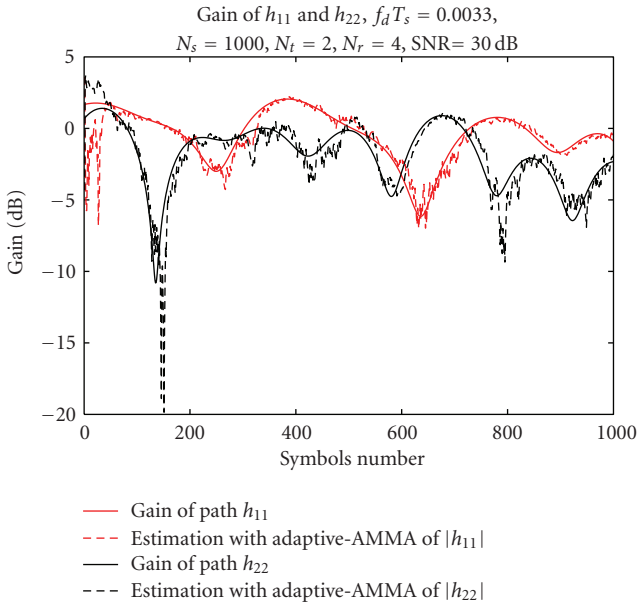


FIGURE 3: Gain of path h_{11} and h_{22} without carrier frequency offset versus symbols number. $N_t = 2$, $\delta_f T_s = 0$, $f_d T_s = 0.0033$ and SNR = 30 dB.

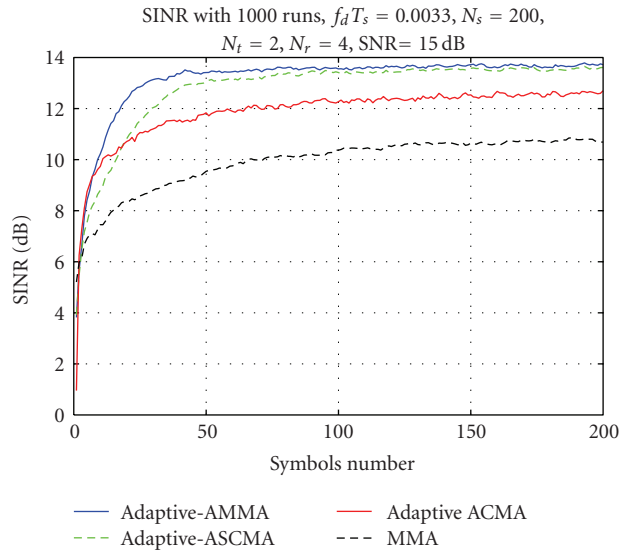


FIGURE 4: SINR without carrier frequency offset versus symbols number. $N_t = 2$, $N_r = 4$, $\delta_f T_s = 0$, $f_d T_s = 0.0033$ and SNR = 15 dB.

So the proposed adaptive-AMMA algorithm allows to obtain a 4 dB gain in SNR terms compared with that obtained by the adaptive-ACMA with a value of BER equal to 10^{-4}.

Figure 3 illustrates the convergence speed of the adaptive-AMMA. The path gain of the channel and its estimate obtained by the adaptive-AMMA are represented on the same figure. We can notice that estimation of adaptive-AMMA follows the variations of the channel.

From the results of the Figure 4 , we can see that the performances of the proposed adaptive-AMMA and the adaptive-ASCMA are very closed in residual SINR terms. Moreover, these two algorithms perform higher residual SINR than the adaptive-ACMA and the MMA.

5. Conclusion

In this paper, we have proposed two new BSS algorithms, named adaptive-AMMA and adaptive-ASCMA, build on the Multimodulus criteria, the Simplified Constant Modulus respectively, and an analytical processing method. The second algorithm is a simplified version of the adaptive-AMMA and it accomplishes performances results closed to the adaptive-AMMA. These algorithms perform better results in BER and SINR terms compared with the adaptive-ACMA and the MMA algorithms. In addition to the blind separation, they can perform a carrier phase recovery simultaneously unlike the adaptive-ACMA.

Appendix

Theorem 1. *The matrices* $\widetilde{\mathbf{C}}$ *and* $\overline{\mathbf{C}}$ *are symmetric and semidefine positive.*

Proof. Let us consider $\mathbf{A} = \mathbf{I} - (1/N)\mathbf{1}\mathbf{1}^T$ a symmetric matrix, so $\mathbf{A}^T = \mathbf{A}$ and $\mathbf{A}^2 = \mathbf{A}$ thus $\mathbf{A}$ is a projection. Then $\mathbf{A} = \mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A}$. Now, we have

$$\begin{aligned}\widetilde{\mathbf{C}} &= \frac{1}{N}\widetilde{\mathbf{P}}^T\widetilde{\mathbf{P}} - \frac{1}{N}\widetilde{\mathbf{P}}^T\mathbf{1}\frac{1}{N}\mathbf{1}^T\widetilde{\mathbf{P}} \\ &= \frac{1}{N}\widetilde{\mathbf{P}}^T\left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right)\widetilde{\mathbf{P}} \\ &= \frac{1}{N}\widetilde{\mathbf{P}}^T\mathbf{A}\widetilde{\mathbf{P}} \\ &= \frac{1}{N}\widetilde{\mathbf{P}}^T\mathbf{A}^T\mathbf{A}\widetilde{\mathbf{P}} \\ &= \frac{1}{N}\left[\mathbf{A}\widetilde{\mathbf{P}}\right]^T\left[\mathbf{A}\widetilde{\mathbf{P}}\right].\end{aligned} \tag{A.1}$$

Similarly

$$\begin{aligned}\overline{\mathbf{C}} &= \frac{1}{N}\overline{\mathbf{P}}^T\overline{\mathbf{P}} - \frac{1}{N}\overline{\mathbf{P}}^T\mathbf{1}\frac{1}{N}\mathbf{1}^T\overline{\mathbf{P}} \\ &= \frac{1}{N}[\mathbf{A}\overline{\mathbf{P}}]^T[\mathbf{A}\overline{\mathbf{P}}].\end{aligned} \tag{A.2}$$

Since $\widetilde{\mathbf{C}} = (1/N)[\mathbf{A}\widetilde{\mathbf{P}}]^T[\mathbf{A}\widetilde{\mathbf{P}}]$ and $\overline{\mathbf{C}} = (1/N)[\mathbf{A}\overline{\mathbf{P}}]^T[\mathbf{A}\overline{\mathbf{P}}]$, then $\widetilde{\mathbf{C}}$ and $\overline{\mathbf{C}}$ are symmetric, semidefine positive and so $\widetilde{\mathbf{C}} + \overline{\mathbf{C}}$ too. Furthermore, for all $\mathbf{y} \in \mathbb{R}^{(2N_t)^2}$, $\mathbf{y}^T(\widetilde{\mathbf{C}} + \overline{\mathbf{C}})\,\mathbf{y} \geq 0$. □

Lemma 1. $\mathbf{d}_l^T(\widetilde{\mathbf{C}} + \overline{\mathbf{C}})\mathbf{d}_l \geq 0$, *for all* $l \in \{1 \cdots N_t\}$ *and so the optimization problem becomes*

$$\arg \min_{\substack{\mathbf{d}_l = \widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l \\ \left\|\widetilde{\mathbf{d}}_l\right\| = R}} \sum_{l=1}^{N_t} \mathbf{d}_l^T\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right)d_l \Longleftrightarrow \arg \min_{\substack{\mathbf{d}_l = \widetilde{\mathbf{t}}_l \otimes \widetilde{\mathbf{t}}_l \\ \left\|\widetilde{\mathbf{d}}_l\right\| = R}} d_l^T\left(\widetilde{\mathbf{C}} + \overline{\mathbf{C}}\right)\mathbf{d}_l, \quad \forall l \in \{1 \cdots N_t\}. \tag{A.3}$$

Afterward, let us define the nullspace and the image of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$ as

$$\ker\left(\widetilde{\mathbf{C}}+\overline{\mathbf{C}}\right)=\left\{\mathbf{y}\in\frac{\mathbb{R}^{(2N_t)^2}}{\left(\widetilde{\mathbf{C}}+\overline{\mathbf{C}}\right)\mathbf{y}=0_{R^{(2N_t)^2}}}\right\},$$
$$\text{image}\left(\widetilde{\mathbf{C}}+\overline{\mathbf{C}}\right)=\left\{\mathbf{z}\in\frac{\mathbb{R}^{*(2N_t)^2}}{\mathbf{z}=\left(\widetilde{\mathbf{C}}+\overline{\mathbf{C}}\right)\mathbf{y}},\ \mathbf{y}\in\mathbb{R}^{(2N_t)^2}\right\}, \tag{A.4}$$

$\ker(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})$ and $\text{image}(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})$ are subspaces of $\mathbb{R}^{(2N_t)^2}$.

Property 1. $\ker(\mathbf{C})\oplus\text{image}(\mathbf{C})=\mathbb{R}^{(2N_t)^2}$ *where* $\oplus$ *denotes the direct sum, that is,* $\ker(\mathbf{C})\cap\text{image}(\mathbf{C})=\{0\}$ *and* $\ker(\mathbf{C})$ *and* $\text{image}(\mathbf{C})$ *are subspaces of* $\mathbb{R}^{(2N_t)^2}$.

Let us assume now, for the sake of simplicity, that $N_t=2$ and $Nr\geq 2$, $\dim(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})=(2N_t)^2\times(2N_t)^2=16\times 16$.

Afterward, the following affirmations will be proven. The vectors $\mathbf{d}_l$ which minimize $\mathbf{d}_l^T(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})\mathbf{d}_l$ are in the nullspace of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$, but the sole vectors $0_{\mathbb{R}^{16}}$ verify the constraint $\mathbf{d}_l=\widetilde{\mathbf{t}}_l\otimes\widetilde{\mathbf{t}}_l$, yet $0_{\mathbb{R}^{16}}$ is a trivial solution to be avoided. However, the vectors $\mathbf{d}_l$ are in the image of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$.

Definition 1. By construction, the sole dependence obtained on the columns of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$ is $C_2=C_5$, $C_3=C_9$, $C_4=C_{13}$, $C_7=C_{10}$, $C_8=C_{14}$, $C_{12}=C_{15}$, where C_n represents the nth column of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$. So, $\dim[\ker(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})]=6$. A basis B_{ker} of the nullspace is defined as

$$\begin{aligned}B_{\text{ker}}&=\{e_1',e_2',e_3',e_4',e_5',e_6'\}\\&=\{e_2-e_5,\ e_3-e_9,\ e_4-e_{13},\ e_7-e_{10},\ e_8-e_{14},\ e_{12}-e_{15}\},\end{aligned} \tag{A.5}$$

where $e_i,\ i\in\{1,\ldots,16\}$ are vectors in the standard basis of $\mathbb{R}^{16}$.

The e_i' are linearly independent and verify $(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})e_i'=0$.

Thanks to the following rank-nullity theorem:

$$\text{rank}\left(\widetilde{\mathbf{C}}+\overline{\mathbf{C}}\right)+\dim\left(\ker\left(\widetilde{\mathbf{C}}+\overline{\mathbf{C}}\right)\right)=\dim(\mathbb{R}^{16}), \tag{A.6}$$

and since the dimension of $\ker(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})$ is equal to 6, the dimension of $\text{Image}(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})$ is equal to 10. And B_{im} is a basis of image:

$$\begin{aligned}B_{im}&=\{e_1'',e_2'',e_3'',e_4'',e_5'',e_6'',e_7'',e_8'',e_9'',e_{10}''\}\\&=\{e_1,e_6,e_{11},e_{16},\ e_2+e_5,\ e_3+e_9,\ e_4+e_{13},\ e_7+e_{10},\\&\qquad e_8+e_{14},\ e_{12}+e_{15}\}.\end{aligned} \tag{A.7}$$

The e_i'' are linearly independent; verify $\mathbf{C}e_i''\neq 0$ and we can verify the Property 1. The e_i' and the e_i'' are linearly independent, then $\{e_1',\ldots,e_6',e_1'',\ldots,e_{10}''\}$ is a basis for $\mathbb{R}^{16}$.

Now, the following theorem can be proved.

Theorem 2. *In the nullspace of* $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$*, the vector* $0_{\mathbb{R}^{16}}$ *is the sole vector to have a Kronecker structure.*

Proof. Let $\mathbf{x}\in\mathbb{R}^{16}$, $\mathbf{x}=(x_1,\ldots,x_{16})^T$. Supposing $\mathbf{x}\in\ker(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})$,

$$\mathbf{x}=\sum_{i=1}^{6}\alpha_i e_i',\quad \alpha_i\in\mathbb{R}. \tag{A.8}$$

We have

(i) $x_1=x_6=x_{11}=x_{16}=0$,

(ii) $x_2=-x_5$, $x_3=-x_9$, $x_4=-x_{13}$, $x_7=-x_{10}$, $x_8=-x_{14}$, $x_{12}=-x_{15}$.

So, $\mathbf{X}=\text{vec}^{-1}(\mathbf{x})$ is a (4×4) skew-symmetric matrix.

$\mathbf{X}$ and $\text{vec}^{-1}(\mathbf{t}\otimes\mathbf{t})$ with $\mathbf{t}=(t_1\ t_2\ t_3\ t_4)^T\in\mathbb{R}^4$ are equal to

$$\mathbf{X}=\begin{pmatrix}0&-x_2&-x_3&-x_4\\x_2&0&-x_7&-x_8\\x_3&x_7&0&-x_{12}\\x_4&x_8&x_{12}&0\end{pmatrix},$$
$$\text{vec}^{-1}(\mathbf{t}\otimes\mathbf{t})=\begin{pmatrix}t_1^2&t_1t_2&t_1t_3&t_1t_4\\t_1t_2&t_2^2&t_2t_3&t_2t_4\\t_1t_3&t_2t_3&t_3^2&t_3t_4\\t_1t_4&t_2t_4&t_3t_4&t_4^2\end{pmatrix}. \tag{A.9}$$

$\text{vec}^{-1}(\mathbf{t}\otimes\mathbf{t})$ is a (4×4) symmetric matrix and its diagonal values are positive, while $\mathbf{X}$ is a skew-symmetric matrix. So, vectors $\mathbf{x}$ do not have a Kronecker structure, except $0_{\mathbb{R}^{16}}$. □

So, vectors $\mathbf{d}_l$ are not in the nullspace of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$, but vectors $\mathbf{d}_l$ are in the image of $\widetilde{\mathbf{C}}+\overline{\mathbf{C}}$. Using property 1, we deduce the following lemma.

Lemma 2. *Let us consider E, the set of vectors in* $\mathbb{R}^{16}$*, which have a Kronecker structure:* $E=\{\mathbf{p}\in\mathbb{R}^{16}/\mathbf{p}=\mathbf{t}\otimes\mathbf{t},\ \mathbf{t}\in\mathbb{R}^4\}$*, then* $E\subset\text{Image}(\widetilde{\mathbf{C}}+\overline{\mathbf{C}})$.

Notations

The following notations are adopted:

*:	Complex conjugation
^{T}:	Matrix or vector transpose
^{H}:	Matrix or vector complex conjugate transpose
$\mathfrak{R}$:	Real part of complex
$\mathfrak{I}$:	Imaginary part of complex
$\mathbf{1}$:	Vector of all 1s
$\mathbf{0}_N$:	An $N\times N$ matrix with zeros components
$\mathbf{I}_N$:	The $N\times N$ identity matrix
$E[\cdot]$:	Expectation operator

diag($\mathbf{a}$): A diagonal matrix constructed from a vector $\mathbf{a}$
vec($\mathbf{A}$): Stacking of the columns of $\mathbf{A}$ into a vector
$\otimes$: Kronecker product
$\circ$: Khatri-Rao product: $\mathbf{A} \circ \mathbf{B} = (\mathbf{a}_1 \otimes \mathbf{b}_1 \ \mathbf{a}_2 \otimes \mathbf{b}_2 \cdots)$
$\|\cdot\|_F^2$: The Frobenius norm

References

[1] P. Sansrimahachai, D. Ward, and A. Constantinides, "Blind source separation for BLAST," *Digital Signal Processing*, vol. 1, pp. 139–142, 2002.

[2] K. N. Oh and Y. O. Chin, "Modified constant modulus algorithm: blind equalization and carrier phase recovery algorithm," in *Proceedings of the IEEE International Conference on Communications*, vol. 1, pp. 498–502, 1995.

[3] L. Castedo, C. J. Escudero, and A. Dapena, "A blind signal separation method for multiuser communications," *IEEE Transactions on Signal Processing*, vol. 45, no. 5, pp. 1343–1348, 1997.

[4] R. Gooch and J. Lundell, "The CM array: an adaptive beamformer for constant modulus signals," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '86)*, vol. 11, pp. 2523–2526, Tokyo, Japan, April 1986.

[5] A. Ikhlef and D. Le Guennec, "A simplified constant modulus algorithm for blind recovery of MIMO QAM and PSK signals: a criterion with convergence analysis," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, Article ID 90401, 13 pages, 2007.

[6] C. B. Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3508–3519, 2000.

[7] A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, 1996.

[8] J. Li and P. Stoica, *Robust Adaptive Beamforming*, Wiley InterScience, New York, NY, USA, 2005.

[9] A.-J. van der Veen, "An adaptive version of the algebraic constant modulus algorithm," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 4, pp. 873–876, Philadelphia, Pa, USA, 2005.

[10] A.-J. van der Veen, "Asymptotic properties of the algebraic constant modulus algorithm," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1796–1807, 2001.

[11] S. Daumont and D. Le Guennec, "Adaptive analytical MMA with time-varying MIMO channels and diversity in interception context," in *Proceedings of the 16th International Conference on Digital Signal Processing (DSP '09)*, 2009.

[12] S. Douglas, "Combined subspace tracking, prewhitening, and contrast optimization for noisy blind signal separation," in *Proceedings of the 2nd International Workshop on Independent Component Analysis*, pp. 579–584, 2000.

[13] J.-F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 17, no. 1, pp. 161–164, 1996.

[14] S. Attallah and K. Abed-Meraim, "Fast algorithms for subspace tracking," *IEEE Signal Processing Letters*, vol. 8, no. 7, pp. 203–206, 2001.

[15] G. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Md, USA, 1996.

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 956847, 9 pages
doi:10.1155/2010/956847

Research Article

A Transmit Beamforming and Nulling Approach with Distributed Scheduling to Improve Cell Edge Throughput

Wendy C. Wong, Qinghua Li, and Shilpa Talwar

Intel Labs, Intel Corporation, 2200 Mission College Boulevard, Santa Clara, CA 95054, USA

Correspondence should be addressed to Wendy C. Wong, wendy.c.wong@intel.com

Received 16 September 2009; Accepted 16 December 2009

Academic Editor: Hongxiang Li

We propose a transmit scheme for WiMAX systems, where multiple base stations (BSs) employ downlink transmit beamforming and nulling for interference mitigation, with minimal coordination amongst BSs. This scheme improves system throughput and robustness, by increasing cell edge and overall cell throughputs by 68% and 19%, respectively, and by delivering improvement for mobile speed up to 60 km/h. First, cell edge users suffering from severe interferences are identified. Next, the RRM unit allocates resource to serving cell edge users only. BSs will schedule to serve their cell edge users independently using the allocated resources by the RRM. A special uplink sounding region is designed for BSs to learn the interference environment and form proper beams and nulls. The nulls formed towards users served by other BSs reduced interference from a BS towards these users and is the basic building block of our algorithm.

1. Introduction

In a cellular network with frequency reuse one, downlink (DL) performance is limited by cochannel interference. In the downlink of a cellular system, it is well known that BS with multiple transmit antennas can improve the desired signal power by transmit beamforming [1–3]. However, the performance gain generated by the nulls from the multiple antennas is much less studied.

In Figure 1, BS i helps BS j by forming a null towards MS j and BS j helps BS i by forming a null towards MS i. Hence, the SINR at MS i and MS j is increased by higher signal power from the beam of its own serving BS and reduced interference from the null of a nearby interfering BS. This is referred to as beamforming and nulling (BFaN) from hereon. All BSs with cell edge MSs must enable BFaN simultaneously in order to achieve throughput improvement for all cell edge MSs. For example, if BS i enables BFaN while BS j does not, only MS j benefits from the reduced interference. Note that MS i denotes the cell edge MS that is currently being scheduled to be served by BS i. Hence, in the next frame, another cell edge MS may be scheduled and MS i will refer to a different cell edge MS. However, we keep the same MS index to simplify our notations. In addition, BFaN at the BS is attractive since it moves the implementation and computation complexity from MSs (which is more cost sensitive and power limited) to BSs.

This paper studies the benefits of BFaN and has made two key contributions. First, a simple and effective BFaN scheme is proposed for cellular systems with multiple BS antennas. Unlike conventional downlink beamforming with nulling schemes, our BFaN scheme eliminates the restriction for equal UL and DL resource allocation by using a special UL sounding mechanism. Second, it quantifies the performance gains of BFaN in realistic cellular settings. It should be noted that the system throughput gains cannot be observed from link level simulations that are only suitable for a point to point setting due to the dynamic interference environment created by BS scheduling. System level simulations (SLSs) with multiple scheduled links in multi-cells and multi-sectors are required. This may be the reason for the lack of research results reported in the literature.

The paper is divided into literature survey, system model description, our proposed BFaN scheme, simulation results, and conclusions. The literature survey includes other interference mitigation techniques considered in IEEE 802.16m

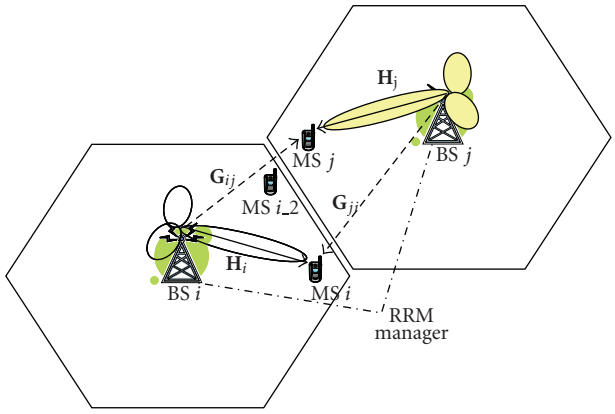


Figure 1: A simple 2 cell deployment showing BFaN.

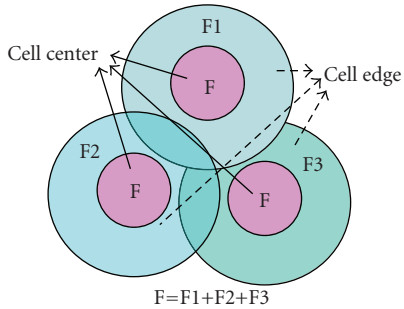


Figure 2: Vanilla FFR scheme.

study group. Most of these techniques are developed within the past year.

2. Literature Survey

In this section, we present a survey of schemes that were recently proposed by IEEE 802.16m task groups in [4] for interference mitigation in frequency reuse one systems. They include fractional-frequency reuse (FFR), precoding matrix index (PMI) restriction, multi-BS MIMO, interference mitigation using power control and scheduling, and multi-hop relay. These not only provide a background and reference for our scheme but are of interest to readers because most of them were published within the past year.

2.1. FFR. A system with FFR employs multiple frequency reuse factors within a cell, for example, reuse 1 and 3, as depicted in Figure 2. At the cell center where interference signal is weak, each BS can use frequency reuse 1. In contrast, at the cell edge where interference is strong, each BS only uses a portion of the bandwidth, for example, using reuse 3. Assuming that users are uniformly distributed throughout the network, the cumulative distribution function of users' average SINR shows that in a reuse 1 network more than 30% users' average SINR is below 0dB. Such users are typically at cell edge, and will likely experience a poor network connection, low downlink throughput, and high probability of outage. Higher frequency reuse factors, such as 3, can significantly reduce the cochannel interference amongst neighboring cells/sectors in that 2/3 of the cochannel interference sources are eliminated compared with reuse 1 networks. This leads to greatly improved coverage and average SINR for cell edge users. The gain from FFR comes from the reduced interference level to cell edge users. However, improvement of downlink average SINR by using higher reuse factors is achieved at the cost of system spectrum efficiency, defined as the ratio of system throughput to occupied spectrum bandwidth, since higher reuse also requires more spectrum bandwidth. One main advantage of FFR over other schemes is its ability to support users with high mobility. Hence, FFR is naturally complementary to our BFaN scheme, which primarily supports low to moderate speed mobile users.

2.2. PMI Restriction for Downlink Closed-Loop MIMO. The precoding matrix index (PMI) restriction scheme [5, 6] sacrifices cell center performance to reduce interference experienced by cell edge MSs. On a periodic basis, cell edge MSs measure the interference experienced from neighboring BSs transmission to their cell center MSs. The MSs report to their own serving BS the PMIs that cause high interference and these PMIs will be excluded from the original beamforming codebook. BSs will exchange the list of excluded PMIs over the backhaul. As a result, cell center MSs are served using a subset of the original beamforming codebook while cell edge MSs are served with the complete codebook. Namely, BSs may need to use precoding matrices that are not optimal for its transmission to cell center users. Hence, cell center throughput drops while cell edge throughput gains.

2.3. Multi-BS MIMO. In multi-BS MIMO, cell edge MSs in multiple cells can be served jointly and simultaneously by multiple BSs using the same channel resources. The multiple BSs act as a single BS with distributed antennas at the physical layer. However, this scheme increases system implementation cost significantly due to increased co-ordination. First, BSs need to have access to data destined for MSs of other BSs for macro-diversity. Second, a flexible pilot pattern design with a configurable pilot reuse factor depending on the number of BSs participating in the system is needed for accurate channel measurements and beamforming weight calculation. Third, scheduling decision will be jointly made at the BSs. Performance gain achieved by multi-BS MIMO in [7] for cell edge MSs at 30 km/h is modest at 26.14% while the overall gross throughput rate actually dropped by 1.5%. Hence, with all the increased complexities and BS cooperation, the gain of this scheme is moderate.

2.4. Interference Mitigation by Power Control and Scheduling. A base station can choose to boost its DL transmit power to selected cell edge MSs. However, the boosted signal transmissions will cause stronger interference to cell edge MSs of neighboring cells. As an alternative, a BS can coordinate with other BSs to schedule MSs with high mutual

interference potential on different channel resources [4]. However, this requires periodic BS coordination. Again, since BS operators may not have control over the backhaul latency, this scheme may not be implemented when the backhaul latency exceeds the latency requirement.

2.5. Multihop Relay. The original idea of multi-hop relay is to extend the BS coverage to isolated pockets within or outside the current BS coverage area on a temporary (for emergency response) or permanent basis. The presence of a relay can potentially help cell edge MSs. However, in-band relays operating in a time sharing fashion will decrease system peak throughput and the decrease is scaled down by the number of hops. Hence, relay deployment decreases spectrum efficiency and increases system overheads. For instance, additional preambles in the PHY layer and additional MAC signaling overheads (e.g., dedicated relay zone allocation) are needed to maintain the multi-hop links. Out-of-band relays will increase BS and MS production cost since these devices will need to support multiple radio bands in order to communicate with these relays. Finally, not all operators will deploy relays and hence we cannot count on relay as a solution to improve cell edge user throughput.

3. System Model and Beamforming Weight Calculation

We consider a time division duplex (TDD) system where the uplink (UL) and downlink (DL) channels are transposes of each other due to channel reciprocity. It is assumed that each BS has M antennae while each SS has N antennae.

The DL transmit and UL receive system models are introduced in Sections 3.1 and 3.2. Next, detailed beamforming weight calculation at BSs and MSs is described in Sections 3.3–3.5. Assume that the cell edge users transmitting information to their serving BSs in the UL will be allocated the same resources in the DL transmission from their BSs, UL beamforming weights applied at their BSs to receive information can be used in the DL to transmit information to the same set of cell edge users. Hence, our beamforming weight calculation process starts with looking at the UL beamforming weight calculation at each BS in Section 3.3. Next we show how to derive an optimal DL beamforming weight aiming to maximize cell edge user SINR and how this DL beamforming weight can be approximated by the normalized complex conjugate of the UL beamforming weight in Section 3.4. Last, the DL beamforming weight at a cell edge MS is shown in Section 3.5. Ideally, beamforming weight should be calculated for each subcarrier. However, to reduce overhead, pilots that are used for beamforming weight calculation are scattered over the whole allocated resources. Beamforming weights for intermediate subcarriers can be interpolated from beamforming weights calculated from these pilots in many different ways.

3.1. Data Model for BS DL Transmit. Consider the system model depicted in Figure 3 in the DL (from BSs to MSs)

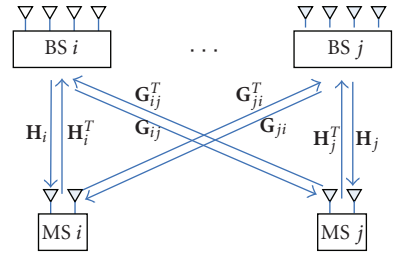


Figure 3: BFaN system model.

direction, the received signal at MS i is

$$\mathbf{x}_{\mathrm{MS}\,i} = \mathbf{H}_i \mathbf{w}_{\mathrm{DL}i} s_{\mathrm{BS}\,i} + \sum_{\substack{j, j \neq i \text{ all interfering} \\ \text{neighboring BS}}} \mathbf{G}_{ji} \mathbf{w}_{\mathrm{DL}j} s_{\mathrm{BS}\,j} + \mathbf{n}_{\mathrm{MS}}, \quad (1)$$

where $\mathbf{H}_i$ is the $N \times M$ channel matrix from serving BS i to MS i; $\mathbf{w}_{\mathrm{DL}i}$ and $\mathbf{w}_{\mathrm{DL}j}$s are DL transmit beamforming vectors applied by BS i and BS j, respectively; $\mathbf{G}_{ji}$ is the $N \times M$ interfering channel matrix from BS j to MS i; $\mathbf{n}_{\mathrm{MS}}$ is the noise vector at MSs. Note that (1) describe the received signal vector at MS i for each subcarrier. However, the subcarrier index is usually dropped to simplify notation.

3.2. Data Model for BS UL Receive. Consider the system model in the UL direction in Figure 3, the received signal at BS i is

$$\mathbf{x}_i = \mathbf{H}_i^T \mathbf{w}_{\mathrm{MS}i} s_i + \sum_{\substack{j, j \neq i, \text{cell edge users} \\ \text{of all interfering BSs}}} \mathbf{G}_{ij}^T \mathbf{w}_{\mathrm{MS}j} s_j + \mathbf{n}, \quad (2)$$

where $\mathbf{H}_i^T$ is the $M \times N$ channel matrix from MS i to its serving BS i; $\mathbf{w}_{\mathrm{MS}i}$ and $\mathbf{w}_{\mathrm{MS}j}$ are the transmit beamforming vectors at MSs i and j, respectively; s_i and s_j are the transmitted QAM symbols from MSs i and j, respectively; G_{ij}^T is the $N \times M$ interfering channel matrix for MS j to BS i; $\mathbf{n}$ is the noise vector at BSs.

If an MS can transmit using $N\,(N > 1)$ antennae in the UL, it can set its BF weight using various methods. For example, if an MS can only transmit using 1 antenna in the UL, $\mathbf{w}_{\mathrm{MS}i}(k) = [1, 0, \ldots, 0]^T$. If an MS can transmit using all N antennae, $\mathbf{w}_{\mathrm{MS}i} = \mathbf{v}_1$ where $\mathbf{v}_1(k)$ $(N \times 1)$ can be the right singular vector corresponding to the largest singular value of $\mathbf{H}_i^T$ [8].

3.3. UL BS Beamforming Weight Computation. Let $\mathbf{w}_i$ $(M \times 1)$ denote the UL beamforming weight vector at BS i. The

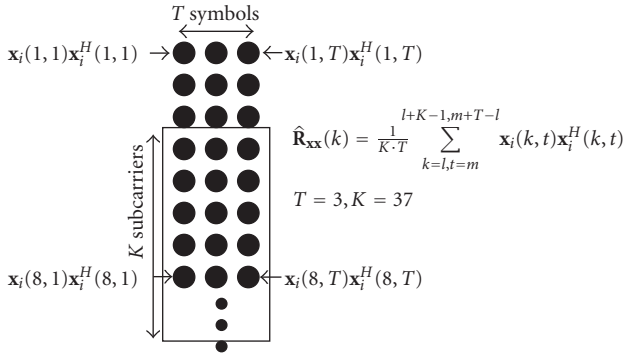


Figure 4: Moving average window.

estimated signal after applying UL BS beamforming weight on the UL receive signal in (2) becomes

$$\begin{aligned}\hat{s}_i &= \mathbf{w}_i^H \mathbf{x}_i \\ &= \mathbf{w}_i^H \mathbf{H}_i^T \mathbf{w}_{\mathrm{MS}i} s_i \\ &\quad + \sum_{\substack{j,j\neq i,\text{cell edge users} \\ \text{of all interfering BSs}}} \mathbf{w}_i^H \mathbf{G}_{ij}^T \mathbf{w}_{\mathrm{MS}\,j} s_j + \mathbf{w}_i^H \mathbf{n}.\end{aligned} \tag{3}$$

The beamforming weight vector $\mathbf{w}_i$ is derived by minimizing the mean squared error (MMSE) $E(|\hat{s}_{\mathrm{MS}i} - s_{\mathrm{MS}i}|^2)$ and is

$$\mathbf{w}_i = \mathbf{R}_{\mathbf{xx}_i}^{-1} \mathbf{H}_i^T \mathbf{w}_{\mathrm{MS}\,i}, \tag{4}$$

$$\mathbf{R}_{\mathbf{xx}_i} = E\left(\mathbf{x}_i \mathbf{x}_i^H\right). \tag{5}$$

At each BS, a moving window averaging technique in Figure 4 is used to generate a covariance matrix estimate $\hat{\mathbf{R}}_{\mathbf{xx}_i}(k)$ for each subcarrier k. The averaging is performed over frequency and time assuming the channel response during the frequency and time span is similar. We have investigated adding smoothing techniques [9] or Cholesky decomposition plus smoothing [10] for improving the accuracy of covariance matrix estimation. However, the performance gain is negligible.

Next, each BS can measure $\mathbf{H}_i^T \mathbf{w}_{\mathrm{MS}\,i}$ from its uplink received signal. Therefore, each BS will be able to calculate its UL beamforming weight using (4).

Assuming noise and data are uncorrelated and data from various MSs are uncorrelated, the covariance matrix in (5) can be expanded as

$$\begin{aligned}\mathbf{R}_{\mathbf{xx}_i} &= \sigma_s^2 \mathbf{H}_i^T \mathbf{w}_{\mathrm{MS}i} \mathbf{w}_{\mathrm{MS}i}^H \left(\mathbf{H}_i^T\right)^H \\ &\quad + \sigma_s^2 \sum_{\substack{i\neq j,\text{cell edge users} \\ \text{of all interfering BSs}}} \mathbf{G}_{ij}^T \mathbf{w}_{\mathrm{MS}j} \mathbf{w}_{\mathrm{MS}j}^H \left(\mathbf{G}_{ij}^T\right)^H + \sigma^2 \mathbf{I}.\end{aligned} \tag{6}$$

However, in the covariance matrix approximation process, there may be non zero cross terms shown in (A.1) in the appendix. These cross terms will cause inaccurate covariance matrix estimation and cause system degradation. Hence, to compare the performance degradation caused by inaccurate covariance matrix estimation, we have implemented *the reference scheme* in our simulation. In the reference scheme, BSs first schedule their cell edge MSs over the allocated resources independently and then calculate the covariance matrix using (6) and beamforming weights assuming perfect channel knowledge using (4). To mimic the delay in backhaul scheduling information exchange, the scheduling decision is made with perfect channel knowledge that exists 5 ms earlier.

3.4. DL BS Beamforming Weight Computation. Consider only the DL transmission from BS i to MS i, the receive signal at cell edge MS $1, 2, \ldots, i, \ldots, N_{\mathrm{BS}}$ is a vector. If the DL transmit beamforming weight vector at BS i and the MSs' receive beamforming vector ($\mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}}$) have been set correctly, the following expression will be minimized ignoring the noise term (since DL is interference limited):

$$\begin{aligned}&\mathrm{E}\left\| \underbrace{\begin{bmatrix} \mathbf{w}_{\mathrm{MS}\,1\,\mathrm{DL}}^H \mathbf{G}_{i1} \\ \vdots \\ \mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}}^H \mathbf{H}_i \\ \vdots \\ \mathbf{w}_{\mathrm{MS}\,N_{\mathrm{BS}}\,\mathrm{DL}}^H \mathbf{G}_{iN_{\mathrm{BS}}} \end{bmatrix}}_{\mathbf{A}} \mathbf{w}_{\mathrm{DL}i} s_{\mathrm{BS}i} - \underbrace{\begin{bmatrix} 0 \\ \vdots \\ s_{\mathrm{BS}i} \\ \vdots \\ 0 \end{bmatrix}}_{s_{\mathrm{BS}i}\mathbf{e}_i} \right\|^2 \\ &= \mathrm{E}\|\mathbf{A}\mathbf{w}_{\mathrm{DL}i} s_{\mathrm{BS}i} - s_{\mathrm{BS}i}\mathbf{e}_i\|^2 \\ &= \mathrm{E}\left(\left(\mathbf{w}_{\mathrm{DL}i}^H \mathbf{A}^H \mathbf{A} \mathbf{w}_{\mathrm{DL}i} - \mathbf{w}_{\mathrm{DL}i}^H \mathbf{A}^H \mathbf{e}_i - \mathbf{e}_i^H \mathbf{A} \mathbf{w}_{\mathrm{DL}i} + 1\right) s_{\mathrm{BS}i}^* s_{\mathrm{BS}i}\right),\end{aligned} \tag{7}$$

where $\mathbf{A}$ is $N_{\mathrm{BS}} \times M$ matrix, $\mathbf{e}_i(N_{\mathrm{BS}} \times 1)$ has all zero entries except a one at the ith position; N_{BS} is the number of BSs participating in the beamforming and nulling. Taking derivative of (7) with respect to $\mathbf{w}_{\mathrm{DL}i}$ and setting the derivative to zero, we have

$$\begin{aligned}\mathbf{w}_{\mathrm{DL}i}^H \mathrm{E}\left(\mathbf{A}^H \mathbf{A} s_{\mathrm{BS}i}^* s_{\mathrm{BS}i}\right) &= \mathrm{E}\left(\mathbf{e}_i^H \mathbf{A}\right) \\ &\Longrightarrow \mathrm{E}\left(\mathbf{A}^H \mathbf{A} s_{\mathrm{BS}i}^* s_{\mathrm{BS}i}\right) \mathbf{w}_{\mathrm{DL}i} = \mathrm{E}\left(\mathbf{A}^H \mathbf{e}_i\right),\end{aligned} \tag{8}$$

$$\begin{aligned}\mathbf{w}_{\mathrm{DL}i} &= \left(\mathrm{E}\left(\mathbf{A}^H \mathbf{A} s_{\mathrm{BS}i}^* s_{\mathrm{BS}i}\right)\right)^{-1} \mathrm{E}\left(\mathbf{A}^H \mathbf{e}_i\right) \\ &= \mathbf{R}_i^{-1} \mathbf{H}_i^H \mathbf{w}_{\mathrm{MS}i\,\mathrm{DL}},\end{aligned} \tag{9}$$

where

$$\begin{aligned}\mathbf{R}_i &= \mathrm{E}\left(\mathbf{A}^H s_{\mathrm{BS}i}^* s_{\mathrm{BS}i} \mathbf{A}\right) \\ &= \mathbf{H}_i^H \mathbf{w}_{\mathrm{MS}i\,\mathrm{DL}} \mathbf{w}_{\mathrm{MS}i\,\mathrm{DL}}^H \mathbf{H}_i \sigma_{\mathrm{BS}}^2 \\ &\quad + \sigma_{\mathrm{BS}}^2 \sum_{j,j\neq i} \mathbf{G}_{ij}^H \mathbf{w}_{\mathrm{MS}j\,\mathrm{DL}} \mathbf{w}_{\mathrm{MS}j\,\mathrm{DL}}^H \mathbf{G}_{ij}.\end{aligned} \tag{10}$$

Setting $\mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}} = \mathbf{w}_{\mathrm{MS}i}^{*}$ in (10), and take the complex conjugate of (10), we have

$$\begin{aligned}\mathbf{R}_i^{*} &= \mathbf{H}_i^{T}\mathbf{w}_{\mathrm{MS}i\,\mathrm{DL}}^{*}\mathbf{w}_{\mathrm{MS}i\,\mathrm{DL}}^{T}\mathbf{H}_i^{*}\sigma_{\mathrm{BS}}^{2} \\ &\quad + \sum_{j,j\neq i}\mathbf{G}_{ij}^{T}\mathbf{w}_{\mathrm{MS}j\,\mathrm{DL}}^{*}\mathbf{w}_{\mathrm{MS}j\,\mathrm{DL}}^{T}\mathbf{G}_{ij}^{*}\sigma_{\mathrm{BS}}^{2} \\ &= \mathbf{H}_i^{T}\mathbf{w}_{\mathrm{MS}i}\mathbf{w}_{\mathrm{MS}i}^{H}\mathbf{H}_i^{*}\sigma_{\mathrm{BS}}^{2} + \sum_{j,j\neq i}\mathbf{G}_{ij}^{T}\mathbf{w}_{\mathrm{MS}j}\mathbf{w}_{\mathrm{MS}j}^{H}\mathbf{G}_{ij}^{*}\sigma_{\mathrm{BS}}^{2}.\end{aligned} \tag{11}$$

Assuming $\sigma_s^2 = \sigma_{\mathrm{BS}}^2$ (since power of QAM symbols are usually normalized to 1) and equal UL and DL resource allocation (this implies that the same MSs that tranmit in the UL will be receiving from their serving BSs in the DL and the BS and MS indexes will be referring to the same BSs and MSs in both UL and DL), combining (6) and (11), we have $\mathbf{R}_{\mathbf{x}\mathbf{x}_i} = \mathbf{R}_i^{*} + \sigma^2\mathbf{I} \approx \mathbf{R}_i^{*}$ since BSs usually have low noise figure and low thermal noise. From (9), we have

$$\begin{aligned}\mathbf{w}_{\mathrm{DL}i} &= \mathbf{R}_i^{-1}\mathbf{H}_i^{H}\mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}} \\ &= \mathbf{R}_i^{-1}\mathbf{H}_i^{H}\mathbf{w}_{\mathrm{MS}i}^{*} = \left(\left(\mathbf{R}_i^{*}\right)^{-1}\mathbf{H}_i^{T}\mathbf{w}_{\mathrm{MS}i}\right)^{*} \\ &\approx \left(\left(\mathbf{R}_{\mathbf{x}\mathbf{x}_i}\right)^{-1}\mathbf{H}_i^{T}\mathbf{w}_{\mathrm{MS}i}\right)^{*} = \mathbf{w}_i^{*}.\end{aligned} \tag{12}$$

Hence, the optimal DL BS beamforming weight is derived to be the complex conjugate of the BS UL beamforming weight. The actual BS DL beamforming weight used is

$$\mathbf{w}_{\mathrm{DL}i} = \frac{\mathbf{w}_i^{*}}{\|\mathbf{w}_i\|}. \tag{13}$$

The normalization is used to ensure that applying the DL BS beamforming weight to the signal will not change its power level. For TDD systems, instead of calculating a separate DL BS beamforming weight, the DL BS beamforming weight can be set to be the normalized conjugated BS UL beamforming weight in (13) assuming equal UL and DL resource allocation.

3.5. Calculation of Receive Beamforming Weight at MS. If MS i has $N\,(N > 1)$ antennae, it can apply beamforming and nulling at its DL receiver to further improve its SINR. Note that this is a deviation from the assumption made in DL BS beamforming calculation. However, the degradation is minimum as shown by our simulation results reported in Section 4. At MS i, the estimated DL receive signal after applying beamforming and nulling to signal in (1) is

$$\begin{aligned}\hat{s}_{\mathrm{BS}\,i} &= \mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}}^{H}\mathbf{x}_{\mathrm{MS}i} \\ &= \mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}}^{H}\mathbf{H}_i\mathbf{w}_{\mathrm{DL}i}s_{\mathrm{BS}i} \\ &\quad + \sum_{\substack{j,j\neq i\,\text{all interfering}\\ \text{neighboring BS}}}\mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}}^{H}\mathbf{G}_{ji}\mathbf{w}_{\mathrm{DL}j}s_{\mathrm{BS}j} + \mathbf{w}_{\mathrm{MS}i\,\mathrm{DL}}^{H}\mathbf{n}_{\mathrm{MS}}.\end{aligned} \tag{14}$$

The MS beamforming weight $\mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}}\,(N\times 1)$ is derived by minimizing the mean squared error $E(|\hat{s}_{\mathrm{BS}\,i} - s_{\mathrm{BS}\,i}|^2)$ and is

$$\begin{aligned}\mathbf{w}_{\mathrm{MS}\,i\,\mathrm{DL}} &= \mathbf{R}_{\mathbf{x}_{\mathrm{MS}i}\mathbf{x}_{\mathrm{MS}i}}^{-1}\mathbf{H}_i\mathbf{w}_{\mathrm{DL}i} \\ &= \mathbf{R}_{\mathbf{x}_{\mathrm{MS}i}\mathbf{x}_{\mathrm{MS}i}}^{-1}\mathbf{H}_i\frac{\mathbf{w}_i^{*}}{|\mathbf{w}_i|}, \\ \mathbf{R}_{\mathbf{x}_{\mathrm{MS}i}\mathbf{x}_{\mathrm{MS}i}} &= E\left(\mathbf{x}_{\mathrm{MS}i}\mathbf{x}_{\mathrm{MS}i}^{H}\right).\end{aligned} \tag{15}$$

Again, $\mathbf{H}_i(\mathbf{w}_i^{*}/|\mathbf{w}_i|)$ can be measured by pilots in the DL transmission and $\mathbf{R}_{\mathbf{x}_{\mathrm{MS}i}\mathbf{x}_{\mathrm{MS}i}}$ can be estimated by using the received data depicted in (1) and using the same moving average method depicted in Figure 4.

4. Proposed BFaN Scheme

In conventional downlink transmit beamforming and nulling scheme for TDD systems, each BS derives a UL receive beamforming weight using signal received in the UL. Assume equal UL and DL resource allocation, each BS can use the normalized conjugated UL receive beamforming weight in (13) as its DL transmit beamforming weight. Compared to FDD or to system deploying DL beamforming alone, this saves channel feedback overhead and avoids channel information exchange among BSs/MSs. In addition, it is more responsive to channel exchange because it avoids the channel feedback latency. However, the restriction of equal UL and DL resource allocation is not practical and generates waste in using channel resource due to the uneven UL and DL traffic pattern. Traffic is much heavier in the DL than UL. It is highly likely to have DL traffic destined to a MS but the MS does not have any UL traffic. Hence, allocating resource for that user in the UL and ask it to just send some junk data to enable the beamforming weight calculation will deprive another user to use the UL transmission opportunity. To remove this restriction on resource allocation and eliminate waste while at the same time taking advantages offered by this scheme, we propose our BFaN to increase cell edge user throughput.

Our BFaN scheme is comprised of 3 steps:

(1) cell edge MS identification,

(2) resource allocation for cell edge MSs,

(3) distributed scheduling.

Our BFaN scheme provides two major advantages over previous interference mitigation schemes. First, it uses existing feedback and UL sounding mechanisms in existing IEEE 802.16 systems. Hence, this scheme can be implemented without additional standard changes. Second, no real time BS collaboration and real time information exchange over the backhaul network is required.

4.1. Cell Edge MS Identification. Cell edge MSs are defined as MSs that have moderate SNR ($>\mathrm{SNR}_{\mathrm{threshold}}$) but low SIR ($\leq \mathrm{SIR}_{\mathrm{threshold}}$) values. MSs with low SNRs are noise limited. For noise limited MSs, mitigating the interference will increase the SINR slightly and these MSs can be better

served by boosting the BS DL transmit power. Users that can receive signals from multiple BSs can measure the power level of the preambles transmitted by these BSs. IEEE 802.16e allows users to report their interference measurements to their serving BS for handover purposes.

After receiving MS interference reports, a BS determines which MSs would be cell edge MSs. The BS will report its total number of cell edge MSs to a RRM (radio resource management) unit over the backhaul. A BS only needs to update this information to the RRM when the number of its cell edge MSs changes significantly.

4.2. Resource Allocation for Cell Edge MSs. Upon receiving the number of cell edge users from all its BSs, the RRM unit determines how much resource to allocate across all BSs to serve cell edge MSs. For proportional fair type schedulers, the fraction of resource allocated to each MS in the long run is equal. Hence, we can allocate resources serving cell edge MSs to be proportional to the fraction of cell edge MSs present in the system. Let there be Q scheduling quanta available per frame, the number of scheduling quanta allocated to serving cell edge MSs only (U) can be determined as

$$U = \left\lfloor Q \times \frac{\text{number of cell edge MSs}}{\text{total number of MSs}} \right\rfloor. \tag{16}$$

The RRM unit allocates U scheduling quanta to serve cell edge MSs and sends the size and location of the allocation to all BSs. Again, this allocation only needs to be updated when the ratio U changes significantly.

4.3. Distributed Scheduling. Each BS schedules its cell edge MSs over the allocated resources in a distributed fashion with no coordination with other BSs. Conventional beamforming and nulling schemes proposed the restriction of equal UL and DL resource allocation. At the UL, each BS can calculate an approximation to the covariance matrix as $\hat{\mathbf{R}}_{\mathbf{xx}_i}$ from the UL receive signal and use it to calculate the UL beamforming weight as in (4) and the DL beamforming weight as in (13). To eliminate the restriction of equal UL and DL resource allocation and still be able to calculate the covariance matrix estimation $\hat{\mathbf{R}}_{\mathbf{xx}_i}$ and calculate the UL/DL beamforming weight, we propose a special uplink sounding mechanism for our BFaN scheme in the next section.

4.4. Special Uplink Sounding. Special uplink sounding is used by BSs to estimate the covariance matrices and calculate the UL/DL beamforming weight while eliminating the equal UL and DL resource allocation restriction. This sounding design enables each BS to estimate its covariance matrix $\mathbf{R}_{\mathbf{xx}_i}$ which is needed in calculating beamforming weight as shown in (4) and (13).

Cell edge MSs that are allocated same resources in DL perform UL sounding in the UL using the same resources. This enables the BSs to estimate the covariance matrices. The UL sounding zone allocation spans the same frequency range as the cell edge user resource allocations in the DL. This ensures that the BF weight calculation at each BS takes into account the channel selectivity over the DL frequency allocation. In addition, this special sounding should be performed in the UL subframe right before the DL subframe to minimize the channel difference between the UL (when the BF weights are calculated) and DL (when the BF weights are applied).

An example resource allocation is depicted in Figure 5 to illustrate how our special UL sounding design works. In Figure 5, there are 8 resource blocks that are allocated to serve cell edge MSs. The 8 blocks are spread out over 2 time units and 4 frequency units. Hence, 8 sounding regions (SR) are required in the special sounding zone to estimate covariance matrices for the 8 resource blocks (RB). First the special uplink sounding blocks span the same 4 frequency units as the resource blocks. Next, subcarriers within each frequency unit in the special sounding region are divided into two distinct sets (dotted line and solid line subcarriers) to support the sounding for the resource blocks in the two time units. Hence, there are 8 sounding regions. Current uplink sounding allocation in 802.16e supports our special uplink sounding allocation.

In Figure 5, BS 1 allocates resource blocks 1, 5, and 2 to MS 1 and BS 2 allocates resource blocks 1, 2, and 6 to MS 2. Hence, MS 1 will perform sounding in sounding regions 1, 5, and 2 and MS 2 will perform sounding in sounding regions 1, 2, and 6. After the special UL sounding, each BS should be able to calculate an approximation to the covariance matrix using the averaging scheme in Figure 4 and calculate the BS beamforming weights using (4) and (13).

5. Simulation Results

Link level simulation alone cannot accurately quantify the performance gain of an interference mitigation scheme. This is due to the fact that the actual interference present in the system is a function of the resource scheduling performed at each BS. Without the actual scheduling function, the actual interference environment cannot be captured and the performance gain cannot be measured. We implemented our BFaN scheme in a system level simulator developed at Intel and are compliant to the IEEE 802.16m Evaluation Methodology Document [11]. System level simulations with a proportional fair scheduler and full buffer traffic model at both the BSs and MSs are performed to investigate the system throughput improvement of our BFaN scheme. To reduce cost and power consumption at MSs, it is assumed that each MS uses 1 transmit in the UL and 2 receive antennae in the DL. A summary of simulation parameters can be found in Table 1.

To highlight the performance gain of our BFaN scheme, we compare the simulation result against a baseline system proposed in [11]. This baseline system is chosen due to the fact that all interference mitigation schemes proposed to the WiMAX standard compare their results to this baseline system. Hence, we can indirectly compare the performance of our scheme to all interference mitigation schemes in WiMAX if we use the same baseline system to showcase our results. In the baseline system, we switch between 2×2 STBC (Alamouti) or MIMO spatial multiplexing (SM) depending on channel conditions. In simulating the baseline system,

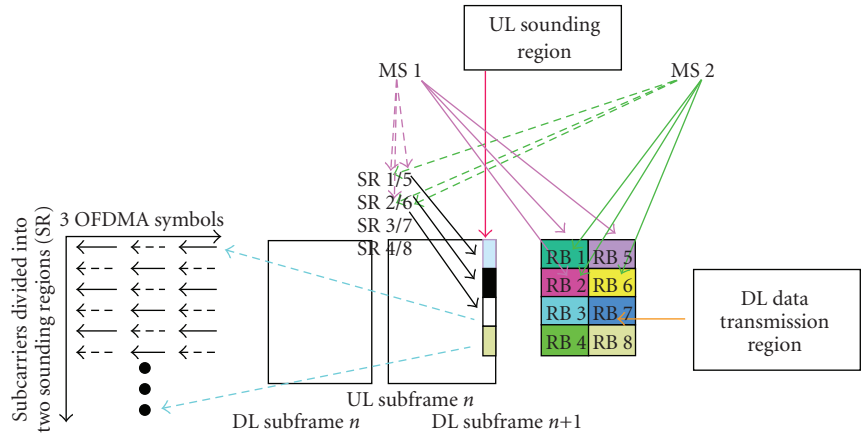


Figure 5: Special UL sounding design.

Table 1: Simulation parameters summary.

Parameters	Values
Cellular layout	7 cells or 19 cells, 3 sectors per cell
Inter-site distance	1500 m
Fading channel	IEEE 802.16m ITU-PEDB and IEEE 802.16m ITU-VEHA
MS speed	3 km/h, 10 km/h, 15 km/h, 30 km/h, 60 km/h
BFaNdwidth	10 MHz
Number of MS per sector	10
Subframe duration	2.5 ms
Antenna configuration	2×2 (2 BS Tx antennae, 2 MS Rx antennae) baseline, 4×2 (4 BS Tx antennas, 2 MS Rx antennas) for our BFaN

Table 2: Simulation result from BFaN and PMI restriction schemes.

Mode/performance	Average MS throughput increase % over baseline	Cell edge MS throughput increase % over baseline
2×2, reference scheme	15.97	34.38
2×2, BFaN scheme	9.69	15.63
2×2, PMI restriction	1.9	39.8
4×2, reference scheme	43.91	150
4×2, BFaN scheme	19.17	68.75
4×2, PMI restriction	−0.5	32

delayed perfect channel knowledge is assumed. This implies that BSs use perfect channel that exists 5 ms ago to calculate the scheduling metrics and make scheduling decisions. When our BFaN scheme is enabled, BSs will use DL transmit beamforming to communication with the cell edge users using 2 or 4 BS antennas while cell center MSs are served using 2×2 STBC/SM scheme.

5.1. Simulation Results and Comparison at Low Mobility. Simulation results at 3 km/h with 19 cells are presented in Table 2. A simple rule of thumb [12] shows that for $M \times N$ antenna configuration, our BFaN can null out $M + N - 2$ significant BS interferers. At the BS, it can form $M - 1$ nulls and point them to $M - 1$ cell edge MSs served by other BSs. This implies that the total number of interfering BSs experienced by any cell edge MS is reduced by $M - 1$. At the MS, it can form $N - 1$ nulls and point them to interfering BSs that do not point a null to the MS. Hence, using our BFaN at all BSs, the number of interfering BSs seen at a cell edge MS is reduced by $M + N - 2$.

For 2×2 antenna configuration, our BFaN scheme can null out only 2 significant BS interferers. In a system with frequency reuse 1, it is highly probable to find 3 or more significant BS interferers. Hence, our BFaN deliver modest gain at the cell edge without reducing cell center throughput rate as compared to the baseline system. In contrast, the PMI restriction promotes cell edge performance but sacrifices the cell center performance. In sum, our BFaN outperforms the PMI restriction at overall cell throughput but has a smaller cell edge improvement than the PMI restriction.

For 4×2 antenna configuration, our BFaN can null out approximately 4 significant BS interferers. With high probability, most significant interferers can be cancelled and high throughput rate is achieved by our 4×2 systems. Simulations show that our 4×2 scheme outperforms the PMI restriction in both the average user throughput rate and cell edge user throughput rate. A CDF of average MS throughput rate can be found in Figure 6. Cell edge user throughput rate is defined to be the throughput rate at 5% of the CDF curve. Average throughput rate is defined to be the throughput rate at 50% of the CDF curve.

The 4×2 reference scheme outperforms our 4×2 BFaN scheme by 81% at cell edge user throughput. This is due to the fact that the covariance matrix calculation in reference scheme is performed using perfect (but delayed) channel information. The covariance matrix calculation in our BFaN design is performed using UL sounding and can suffer from large cross terms (see the appendix) and approximation errors due to thermal noise. The effect of the cross terms can

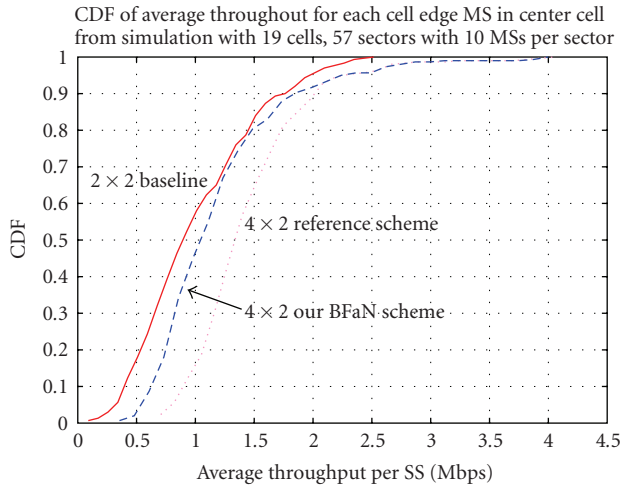


FIGURE 6: CDF of average MS throughput rate of BFaN.

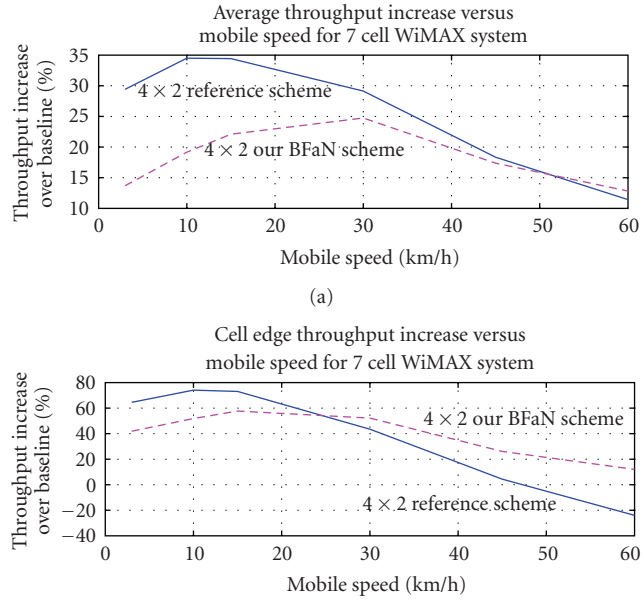


FIGURE 7: Throughput rate increase of BFaN over the baseline system with moving users.

be reduced by expanding the number of subcarriers in the averaging. However, this increases the system overhead.

5.2. Simulation Results for Higher Mobile Speeds. As discussed previously, feedback and computation delays of 5 ms is introduced between the time when beamforming weights are calculated and when the beamforming weights are actually applied. This causes channel mismatch (which increases with mobile speeds) between the time the BS beamforming weights are calculated to the time when BS beamforming weights are applied. From Figure 7, it is observed that gain in cell edge MS throughput rate drops as MS speed increases. However, if MS speed is kept under 60 km/h per hour, our BFaN scheme can still provide gain over the baseline system.

There are two other interesting observations. First, the reference scheme is more sensitive to the channel mismatch caused by user mobility than our BFaN scheme. This is due to the fact that the BFaN scheme forms wider beams and shallower nulls than the reference design. Hence, the channel mismatch causes greater error in the beamforming and null forming and pointing in the reference scheme at higher mobile speed and degrade the system performance more. Second, as the mobile speed increases, performance of the baseline system drops more than our BFaN schemes at moderate MS speed upto 30 km/h. It is only at higher MS speeds that our BFaN schemes start to drop more in system performance compared to the baseline system.

6. Conclusions

The optimality and efficiency of our BFaN scheme is proved and quantified by derivation and system level simulations, respectively. It increases the cell edge MS throughput rate compared to the baseline implementation upto MS speed of 60 km/h. Our special uplink sounding provides an elegant way to enable UL/DL BS beamforming weight calculation and eliminates the restriction of equal UL and DL resource allocation and real time information exchange over the backhaul. A survey of recent schemes for interference mitigation is presented and one of them is compared to our proposed scheme. Our scheme is simple to implement since it is already supported by the current 802.16e message exchange and sounding mechanisms.

Appendix

Assuming channel is flat over K subcarriers and T symbols, covariance matrix $\hat{\mathbf{R}}_{\mathbf{xx}_i}(k)$ is expanded as

$$\begin{aligned}
\hat{\mathbf{R}}_{\mathbf{xx_i}} &= \frac{1}{K \cdot T} \sum_{k=l,\,t=m}^{l+K-1,\,m+T-1} \mathbf{x}_i \mathbf{x}_i^H \\
&= \mathbf{H}_i^T \mathbf{w}_{\mathrm{MS}i} \mathbf{w}_{\mathrm{MS}i}^H \left(\mathbf{H}_i^T\right)^H \frac{1}{K \cdot T} \\
&\quad \times \sum_{k=l,\,t=m}^{l+K-1,\,m+T-1} s_i(k,t) s_i^*(k,t) \\
&\quad + \sum_{\substack{j, i \neq j, \text{ cell edge users of} \\ \text{all interfering BSs}}} \mathbf{G}_{ij}^T \mathbf{w}_{\mathrm{MS}j} \mathbf{w}_{\mathrm{MS}j}^H \left(\mathbf{G}_i^T\right)^H \frac{1}{K \cdot T} \\
&\quad \times \sum_{k=l,\,t=m}^{l+K-1,\,m+T-1} s_j(k,t) s_j^*(k,t) \\
&\quad + \frac{1}{K \cdot T} \sum_{k=l,\,t=m}^{l+K-1,\,m+T-1} \mathbf{n}(k,t) (\mathbf{n}(k,t))^H \\
&\qquad\qquad \text{(needed terms)}
\end{aligned}$$

$$
\begin{aligned}
&+\mathbf{H}_i^T\mathbf{w}_{\mathrm{MS}i}\left(\sum_{\substack{j,i\neq j,\text{ cell edge users of}\\ \text{all interfering BS}}}\mathbf{w}_{\mathrm{MS}j}^H\left(\mathbf{G}_{ij}^T\right)^H\frac{1}{K\cdot T}\right.\\
&\qquad\times\sum_{k=l,t=m}^{l+K-1,m+T-1}s_i(k,t)s_j^*(k,t)\\
&\qquad\left.+\frac{1}{K\cdot T}\sum_{k=l,t=m}^{l+K-1,m+T-1}s_i(k,t)(\mathbf{n}(k,t))^H\right)\\
&+\sum_{\substack{j,i\neq j,\text{ cell edge users of}\\ \text{all interfering BSs}}}\mathbf{G}_{ij}^T\mathbf{w}_{\mathrm{MS}j}\mathbf{w}_{\mathrm{MS}j}^H\left(\mathbf{H}_i^T\right)^H\frac{1}{K\cdot T}\\
&\times\sum_{k=l,t=m}^{l+K-1,m+T-1}s_j(k,t)\left(s_i^*(k,t)+(\mathbf{n}(k,t))^H\right)\\
&+\frac{1}{K\cdot T}\sum_{k=l,t=m}^{l+K-1,m+T-1}\mathbf{n}(k,t)\\
&\times\left(s_i^*(k,t)\mathbf{w}_{\mathrm{MS}i}^H\left(\mathbf{H}_i^T\right)^H+s_j^*(k,t)\right.\\
&\qquad\left.\times\sum_{\substack{j,i\neq j,\text{ cell edge users of}\\ \text{all interfering BSs}}}\mathbf{w}_{\mathrm{MS}j}^H\left(\mathbf{G}_{ij}^T\right)^H\right)\\
&\qquad\text{(cross-terms).}
\end{aligned}
\tag{A.1}
$$

References

[1] Q. Li, X. E. Lin, and J. Zhang, "MIMO precoding in 802.16e WiMAX," *Journal of Communications and Networks*, vol. 9, no. 2, pp. 141–149, 2007.

[2] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.

[3] L. C. Godara, "Application of antenna arrays to mobile communications—part II: beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

[4] S. Hamiti, Nokia, and SDD, Eds., "IEEE 802.16m system description document [Draft]," IEEE 802.16m-08/003r9a, 2009-05-31.

[5] D. Lim, S. Kim, J. Kim, and B.-C. Ihm, "PMI restriction for the downlink closed-loop MIMO," IEEE C802.16m-08_430r1.doc, July 2008.

[6] J. Kim, D. Lim, W. Lee, and B.-C. Ihm, "PMI restriction with adaptive feedback mode," IEEE C802.16m-09/0023, January 2009.

[7] S. Kim, J. Kim, D. Lim, B.-C. Ihm, and H. Cho, "Interference mitigation using FFR and multi-cell MIMO in downlink," IEEE C802.16m-08/783, July 2008.

[8] I. Emre Telatar, "Capacity of multi-antenna gaussian channels," in *Proceedings IEEE International Symposium on Information Theory*, 2001.

[9] A. Maltsev, R. Maslennikov, and A. Khoryaev, "Comparative analysis of spatial covariance matrix estimation methods in OFDM communication systems," in *Proceedings of the 6th IEEE International Symposium on Signal Processing andInformation Technology (ISSPIT '06)*, pp. 551–555, August 2006.

[10] Q. Li, J. Zhu, Q. Li, and C. Georghiades, "Efficient spatial covariance estimation for asynchronous co-channel interference suppression in MIMO-OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4849–4853, 2008.

[11] IEEE 802.16m-08/004r5, "Project 802.16m Evaluation Methodology Document (EMD)," January 2009.

[12] L. C. Godara, "Applications of antenna arrays to mobile communications—part I: performance improvement, feasibility, and system considerations," *Proceedings of the IEEE*, vol. 85, no. 7, pp. 1031–1060, 1997.

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 134919, 7 pages
doi:10.1155/2010/134919

Research Article

Intercell Interference Coordination through Limited Feedback

Lingjia Liu,[1] Jianzhong (Charlie) Zhang,[1] Jae-Chon Yu,[2] and Juho Lee[2]

[1] *Dallas Telecommunications R&D Center, Samsung Telecommunications America, Richardson, TX 75082, USA*
[2] *Standards and Industry Initiative (SII), Samsung Electronics, 416, Maetan-3dong, Yeongtong-gu, Suwon-si, Gyeonggi-do 443-742, South Korea*

Correspondence should be addressed to Lingjia Liu, lingjialiu@gmail.com

Received 3 August 2009; Accepted 2 November 2009

Academic Editor: Hongxiang Li

We consider the applications of multicell transmission schemes to the downlink of future wireless communication networks. A multicell multiple-input multiple output-(MIMOs) based scheme with limited coordination among neighboring base stations (BSs) is proposed to effectively combat the intercell interference by taking advantage of the degreesoffreedom in the spatial domain. In this scheme, mobile users are required to feedback channel-related information to both serving base station and interfering base station. Furthermore, a chordal distance-based compression scheme is introduced to reduce the feedback overhead. The performance of the proposed scheme is investigated through theoretical analysis as well as system level simulations. Both results suggest that the so-called "intercell interference coordination through limited feedback" scheme is a very good candidate for improving the cell-edge user throughput as well as the average cell throughput of the future wireless communication networks.

1. Introduction

Recent years have been marked by a soaring demand for network access. This trend is exemplified by the constant growth of wireless communication systems. The strong demand for network connectivity is partially fueled by new software applications and a widespread desire for real-time information access. Hence, future wireless communication networks will face the dual challenge of supporting large traffic volumes and providing reliable service to delay-sensitive applications such as voice over IP (VoIP), video-conferencing, and online gaming. There are two performance measures that are crucial for wireless systems: average cell throughput and cell-edge user throughput [1]. Improving both of the performance measures becomes one of the major tasks of the next generation wireless communication systems. However, it is important to note that improving average cell throughput is a relatively easy task, while improving cell-edge user throughput becomes extremely demanding. This is because the average cell throughput can be improved using simple methods such as transmission power boosting. However, for cell-edge user throughput, these simple methods are not valid any more. Cell-edge users usually have relatively low received signal strength; furthermore, they do suffer from strong inter-cell interference. Transmission power boosting may increase the received signal strength, but it will also create stronger inter-cell interference to other cell's cell-edge users and hence reduce their throughput. Therefore, improving cell-edge user throughput becomes highly nontrivial. This is also part of the reasons why interference mitigation technologies for next generation wireless systems receive enormous attention in the standardization societies as well as in the research community [1–4].

In the wireless systems equipped with multiple transmit antennas, each cell applies a precoding vector on the transmit antennas to form a beam pointing towards targeted mobile stations (MSs). Current design of the wireless systems requires the scheduler at each cell to choose the precoding vector for beam-forming purely based on the wireless channel between the BS and the targeted MS [5]. Without taking into account which precoding vectors are used in

the neighboring cells, the beams formed by different cells may randomly collide with each other, which results in substantial inter-cell interference for the cell-edge users.

In order to mitigate the interference to the cell-edge users and increase the system spectral efficiency, multi-cell MIMO is proposed as an enabling technology for future wireless systems [6, 7]. In multi-cell MIMO, the network is required to process and transmit the data for an intended mobile user jointly from multiple geographically separated cells. This technology can greatly enhance the performance of cell-edge mobile users since it effectively changes the interference into useful signals. However, it requires the network to have access to the full channel station information (CSI) and requires the data for the intended mobile user to be available at all base stations. These two assumptions seem to be pretty restrictive in the practical wireless systems. Currently, only codebook-based feedback of CSI is widely adopted in the standards to reduce the uplink overhead [5, 8]. Furthermore, practical issues such as backhaul delay and cost will limit the possibility of having one mobile user's data delivered to multiple base stations to perform joint processing. Therefore, it is also interesting to investigate interference mitigation schemes where the data for the intended mobile user is transmitted from a single serving cell. However, as opposed to the single cell operation, the scheduler should choose a precoding vector based on the link between the serving cell and the targeted MS together with the interference the serving cell may cause to the other cell's cell-edge users. Accordingly, we propose to jointly choose the precoding vectors among different cells to mitigate the inter-cell interference taking advantage of the spatial domain degrees of freedom introduced by MIMO systems. Only limited overhead control information is needed to enable this technology and each cell is able to choose his/her precoding vector in a distributed fashion. In other words, a central scheduler is not necessary for the proposed scheme. Both the analytical and simulation results suggest that the proposed scheme can significantly improve average cell throughput as well as the throughput of the cell-edge users.

The paper is structured as follows. Section 2 contains the system model. The theoretical foundation of the proposed inter-cell interference coordination schemes is illustrated in Section 3. Based on the derivation, we propose two different inter-cell interference coordination schemes in Section 4. We detail the simulation results in Section 5 and conclude in Section 6.

2. System Model

In this section, we start to analyze the throughput performance of cell-edge users. The typical scenario of two cell-edge users interfering with each other is illustrated in Figure 1. The corresponding system setup is that both of the base stations are communicating to the two corresponding serving mobile users simultaneously in the same frequency band. In Figure 1, BS1 is the serving cell for MS1 while BS2 is the serving cell for MS2. In this simple wireless system, assume that both MS1 and MS2 are cell-edge users and they are geometrically close to each other. The system described in Figure 1. is actually one of the worst interference cases for the cell-edge users because both users' performance are limited by the strong interference from the interfering cells. This fact can be seen most clearly from the expression of the received signal strength at each mobile user. To be specific, the received signals, Y_1 and Y_2, of MS1 and MS2 can be written as

$$\begin{aligned} Y_1 &= H_{11}w_1X_1 + H_{21}w_2X_2 + N_1, \\ Y_2 &= H_{12}w_1X_1 + H_{22}w_2X_2 + N_2, \end{aligned} \tag{1}$$

where H_{ij} denotes the channel gain from the ith BS to the jth MS, w_i is the precoding vector used at BS i, X_i is the vector of transmitted signal at BS i, and N_i is the additive white Gaussian noise (AWGN) vector at MS i. The received signal of MS1, Y_1, suffers from the interference from BS2 ($H_{21}w_2X_2$) and Y_2 suffers from the interference from BS1 ($H_{12}w_1X_1$). The received Signal to Interference-plus-Noise Ratio (SINR) for MS1 and MS2 can then be expressed as

$$\begin{aligned} \mathrm{SINR}_1 &= \frac{\|H_{11}w_1\|^2 P_1}{\|H_{21}w_2\|^2 P_2 + N}, \\ \mathrm{SINR}_2 &= \frac{\|H_{22}w_2\|^2 P_2}{\|H_{12}w_1\|^2 P_1 + N}, \end{aligned} \tag{2}$$

where P_i is the transmitted power of X_i at BS i, and N is the noise power. In current LTE (Long Term Evolution) system [5], scheduler at BS1 chooses the precoding vector, w_1, purely based on the wireless channel between the BS1 and the targeted MS1, that is, H_{11}; while scheduler at BS2 chooses w_2 purely based on the channel from BS2 to MS2, that is, H_{22}. Since MS1 and MS2 are geographically close to each other, the channel gains from the BSs to the MSs are usually correlated. That is, H_{11} and H_{12} are correlated, and H_{21} and H_{22} are correlated. Therefore, the precoding vector, w_1, which maximizes $H_{11}w_1X_1$ may also produce large magnitude of $H_{12}w_1X_1$ which is the interference from BS1 to MS2. Increasing the transmission power will also increase the interference to other cell's cell-edge users in a linear way. Since MS1 and MS2 are both cell-edge users, the received signal strength will be comparable to the received interference strength. Therefore, SINR_1 and SINR_2 will be normally below 0 dB. The fact that both the mobile users experience very low SINR limits the performance of the whole system and cannot be resolved by simply increasing the transmit power of BS1 and BS2.

3. Theoretical Foundation

In the previous section, we have developed some critical understandings of the interference for the cell-edge users. In this section, we will analyze fundamentals of "Inter-cell Interference Coordination through limited feedback" and show how it will improve the throughput of cell-edge users.

Even though the inter-cell interference cannot be effectively eliminated by increasing or reducing the total transmission power, it is interesting to note that it can actually be greatly reduced through optimizing over the precoding

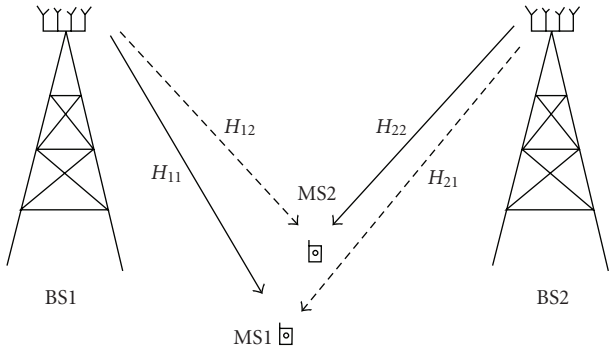


Figure 1: System model of two interfering cell-edge users.

vectors in the spatial domain. For the wireless system shown in Figure 1, it can be seen from (2) that the SINR_1 and SINR_2 are functions of w_1 and w_2. In other words, we can optimize over w_1 and w_2 to improve both SINR_1 and SINR_2. Furthermore, for a wireless system equipped with multiple transmit antennas the inter-cell interference can be partially or completely cancelled by applying different precoding vectors at different base stations. This can be achieved by exploring the additional degrees of freedom offered by multiple transmit antennas in the spatial domain.

In [9], an optimal noncooperative zero-forcing beamforming is proposed. A mobile user is required to feedback the precoding vector to the serving cell taking into account the effects of the interference channel. In this way, the transmitted signal from the serving cell can effectively "avoid" the interference from other cells. Assume that MS1 has the ability to estimate the interference channel $(H_{21}w_2)$ from base Station 2; mathematically, MS1 will compute the precoding vector based on

$$w_1' = \arg\max_{w_1 \in \Gamma} \frac{\|H_{11}w_1\|^2 P_1}{\|H_{21}w_2\|^2 P_2 + N}, \tag{3}$$

where Γ stands for the codebook. This scheme performs well under the assumption that there will be no communication between the cells. However, for the wireless system where the channel-related information can be exchanged over the network, it is strictly suboptimal. Accordingly, the optimal way is to jointly choose the precoding vectors in (2).

Before going to the details of the proposed communication scheme let us take a deeper look at the interference of MS1 from BS2 in the system depicted in Figure 1. Assume that the number of transmit antennas at the BS is N_T and the number of receive antennas at the MS is N_R; the channel gain matrix H_{21} then becomes an N_R by N_T matrix. In the wireless systems the number of transmit antennas at the base stations is always greater than or equal to the number of receive antennas at the mobile users; therefore, we can safely assume $N_R \le N_T$ throughout the paper. Applying the singular value decomposition (SVD) [10] to H_{21}, we have

$$H_{21} = U\Lambda V = U \times \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \vdots \\ & & \lambda_{N_R} \quad 0 \end{bmatrix}_{N_R \times N_T} \times V, \tag{4}$$

where U is an N_R by N_R unitary matrix, V is an N_T by N_T unitary matrix, and λ_1 through λ_{N_T} are the singular values of the channel gain matrix H_{21}. After applying the precoding vector, the interference seen at MS1 can be expressed as

$$\begin{aligned} H_{21}w_2X_2 &= U\Lambda V w_2 X_2 \\ &= U \times \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \vdots \\ & & \lambda_{N_R} \quad 0 \end{bmatrix}_{N_R \times N_T} \times V \times w_2 \times X_2. \end{aligned} \tag{5}$$

Let

$$X_2' = V \times w_2 \times X_2 \tag{6}$$

which is an N_T by 1 vector. Accordingly, the interference from BS2 which is seen at MS1 can be rewritten in the form of

$$H_{21}w_2X_2 = U \times \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \vdots \\ & & \lambda_{N_R} \quad 0 \end{bmatrix}_{N_R \times N_T} \times X_2'. \tag{7}$$

As long as we choose the precoding vector, w_2, such that X_2' satisfies the condition,

$$X_2' = V \times w_2 \times X_2 = \begin{bmatrix} 0 & \cdots & 0 & x_1 & \cdots & x_{N_T - N_R} \end{bmatrix}^T, \tag{8}$$

the interference seen at MS1, $H_{21}w_2X_2$, will be strictly zero. That is,

$$\begin{aligned} H_{21}w_2X_2 &= U \times \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \vdots \\ & & \lambda_{N_R} \quad 0 \end{bmatrix}_{N_R \times N_T} \times \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_1 \\ \vdots \\ x_{N_T - N_R} \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{N_R \times 1}. \end{aligned} \tag{9}$$

Note that x_1 and x_2 can be arbitrary values and the precoding vectors satisfying (8) are not unique. This result indicates

that there exists a set of precoding vectors at BS2 which will cause no interference to MS1 if we are allowed to choose precoding vectors freely. Furthermore, this set of precoding vectors lies in the null space of the interfering channel matrix. Similarly, the same result will apply to the received signal of MS2; that is, a set of precoding vectors at BS1 will cause no interference to MS2. This result is true as long as the number of transmit antennas at BS is larger than that at the MS which means that the null space of the interfering channel matrix is not empty. For the case where we have to select the precoding vectors from a predetermined set like in the LTE systems [5], there might be no precoding vectors satisfying the condition in (8). In this situation there always exists a precoding vector which creates least interference among all the available precoding vectors. By using this precoding vector, we can make sure that the interference created to the other cell is minimal within the predetermined precoding vector set.

Forcing the inter-cell interference to be zero or minimal is a very restrictive condition and greatly reduces the choice of the precoding vectors. For example, for an N_T by N_R wireless system, the precoding vectors satisfying (8) only spans $N_T - N_R$ dimensions of the overall spatial domain which has a total dimension of N_T. Therefore, we introduce a parameter, SINR_thd, to relax the requirements of the interference seen by each MS. To be specific, w_2 is chosen to satisfy the following condition:

$$w_2 = \arg \frac{\|H_{11}w_1\|^2 P_1}{\|H_{21}w\|^2 P_2 + N} \geq \text{SINR\_thd}, \tag{10}$$

where Ω is the set of all precoding vectors. Equation (10) means that w_2, when used at BS2, will introduce tolerable interference to MS1. Note that SINR_thd plays a crucial role in the precoding vector selection. When this threshold is large, more restrictive constraints are put on BS2's interference to MS1, which means that less number of precoding vectors will be used for BS2. In this scenario, interference can be greatly reduced but the multi-user diversity also reduced due to the restrictive selection of the precoding vectors at BS2. When this threshold is small, BS2 will have more freedom to choose the precoding vectors thus increasing multi-user diversity. However, the interference from BS2 to MS1 can still be large due to the loose condition of the SINR threshold. In a way, this threshold triggers a trade-off between multi-user diversity and interference mitigation. Interestingly, the condition expressed in (8) is actually the special case when

$$\text{SINR\_thd} = \frac{\|H_{11}w_1\|^2 P_1}{N}. \tag{11}$$

4. Intercell Interference Coordination through Limited Feedback

Motivated by the elegant results shown in Section 3, we start to investigate on practical interference mitigation schemes through limited coordination. From the analysis we know that each BS has a set of precoding vectors that will cause controlled interference to the cell-edge users in the adjacent cells through parameter SINR_thd. Throughout this paper, we call this set of precoding vectors as "the recommended set." Therefore, if a BS can choose a precoding vector within this recommended set to maximize his/her SINR to the targeted MS, the inter-cell interference will be greatly mitigated. Accordingly, the cell-edge user throughput will be significantly improved. However, one question remains: how does the BS know about the set of recommended precoding vectors. It is interesting to note that the condition shown in (10) can actually be tested at each MS. Therefore, each MS can feedback the recommended set of precoding vectors to the interfering cell which will cause tolerable interference to the interfering cells.

Feeding back the whole set of recommended precoding vectors will cause too much signaling overhead for the system. Therefore, we must further optimize the feedback information to reduce the system overhead. Note that the recommended set of precoding vectors contains all the precoding vectors satisfying (10). That is, w_2 belongs to the recommended set if and only if

$$\|H_{21}w_2\|^2 \leq \left(\frac{\|H_{11}w_1\|^2 P_1}{\text{SINR\_}thd} - N \right) / P_2 = \alpha. \tag{12}$$

In order to reduce the feedback overhead of the coordination scheme, we can take a deeper look at the necessary and sufficient condition of the recommended precoding vector in (12). The left-hand side (LHS) of above inequality is actually related to a distance measure between H_{21} and w_2. Therefore, (12) suggests that a distance measure threshold together with a reference precoding vector can be used to completely characterize the set of recommended precoding vectors. This result can be seen most clearly through a simple example. Assume that we have an N_T by 1 wireless system, that is, N_T transmit antennas at the BS and 1 receive antenna at the MS. For this simple system, the channel matrix H_{21} becomes a 1 by N_T vector which can be written as the hermitian of a N_T by 1 vector. That is, $H_{21} = w^*$, where w is a N_T by 1 vector. Therefore, the LHS of (12) can be rewritten as

$$\|H_{21}w_2\|^2 = \left|\left|w^* w_2\right|\right|^2 \leq \alpha. \tag{13}$$

The above expression is actually the cross-correlation between two N_T by 1 vectors. Since both w and w_2 are unitary, we can further rewrite (13) into

$$d_{\text{chordal}}(w, w_2) = \sqrt{1 - \|w^* w_2\|^2} \geq \sqrt{1 - \alpha}, \tag{14}$$

where $d_{\text{chordal}}(w, w_2)$ stands for the chordal distance between w and w_2 [11]. In this example, the distance measure is the chordal distance and the reference precoding vector is the precoding vector $w = H_{12}^*$.

In the case where one particular cell receives multiple recommended sets from various cells, the scheduler should be able to choose one of the requests based on overall system throughput. To facilitate the choice at the scheduler, each MS should also report the SINR or channel quality improvement when the recommended set of precoding vectors is used

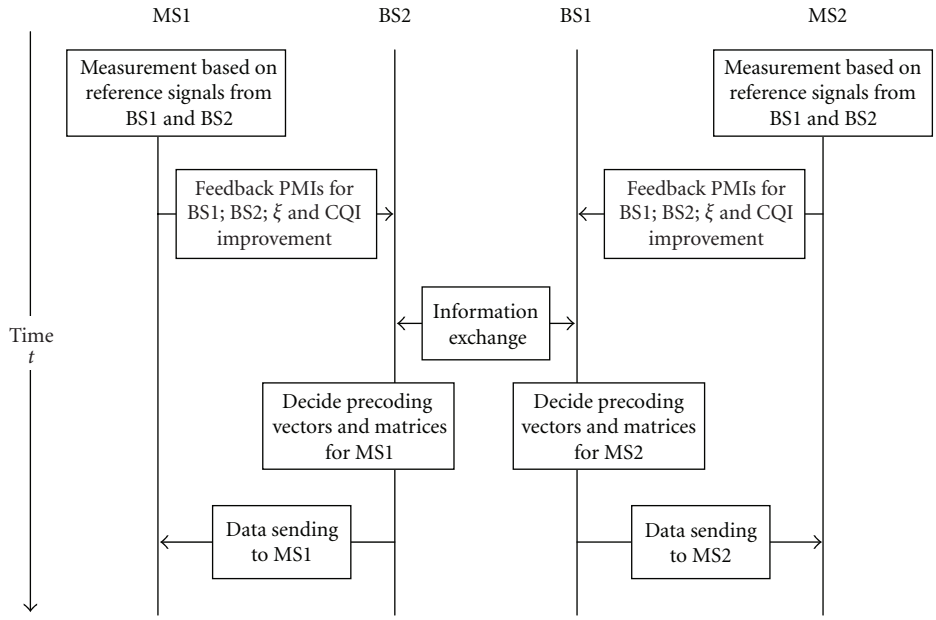


Figure 2: Time line of proposed inter-cell interference coordination through limited feedback.

at the interfering cells. Accordingly, when facing multiple requests at one cell, the scheduler should accept the request with highest the SINR or channel quality improvement so that the system performance improves most.

Based on all the results and understandings, the overall procedure of the proposed "Inter-cell interference coordination through limited feedback" scheme is listed as follows.

Step 1. Each MS measures the channel from the serving cell as well as the interfering cells.

Step 2. Each MS obtains the feedback information for interfering links. The feedback information contains what follows:

(i) reference precoding vectors (PMI) from the interfering cells,

(ii) a distance measure threshold indicating the sets of precoding vectors,

(iii) precoding vector and channel quality index (CQI) for the serving cell.

Step 3. Each MS obtains the measure for performance improvement:

(i) SINR improvement when the recommended set of precoding vectors is used at the interfering cells.

Step 4. Each MS feeds back the information to serving cell as well as interfering cells.

Note that in this mode of operation, the MS can send all the feedback information to the serving cell relying on the serving cell to relay all the related information to the interfering cells. Also, the MS can choose to feedback the information to the intended destination directly. That is, the reference precoding vector together with the threshold can be sent back to the interfering cells from the MS directly.

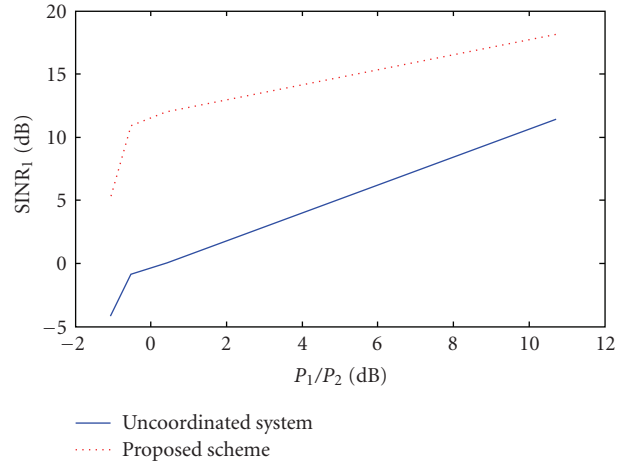


Figure 3: SINR improvement.

Step 5. Serving cell the interfering cells choose corresponding precoding vectors to serve their targeted users.

In this mode of operation, interfering cells are suggested to choose the precoding vector which maximizes his/her own serving MS's throughput within the recommended set if no central scheduler is present. In the case where a central scheduler is present, the precoding vectors for the serving MS are decided jointly across all the serving cells by the central scheduler.

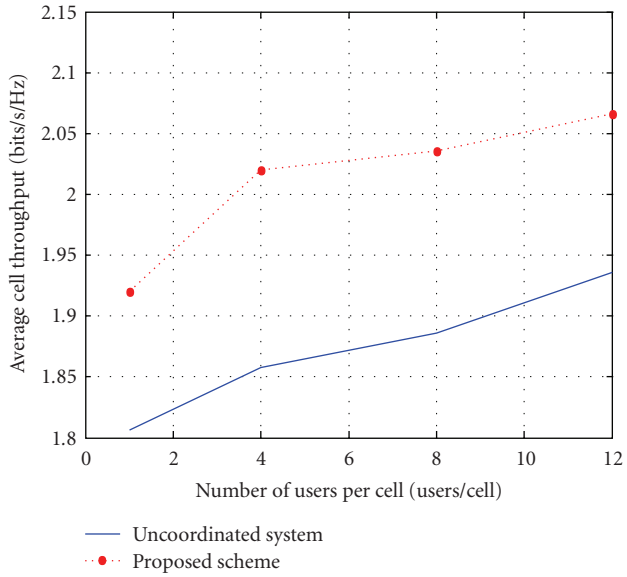


Figure 4: Average cell throughput comparison.

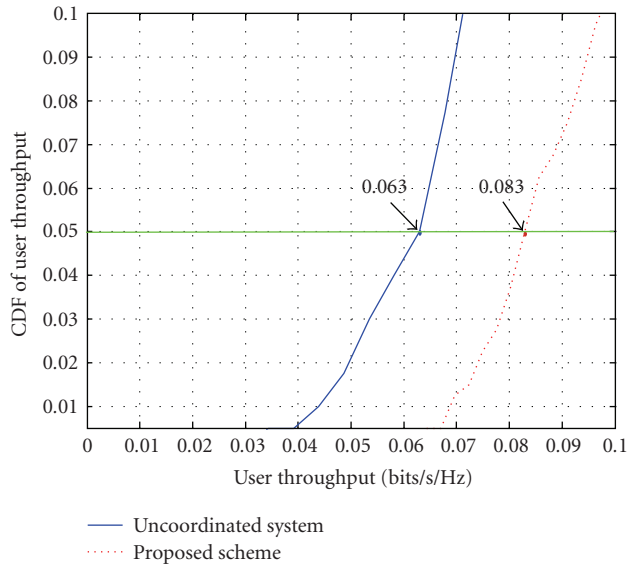


Figure 5: Cell-edge user throughput comparison.

The timeline of the proposed interference coordination scheme can be shown in Figure 2. In the proposed block diagram, ξ stands for the distance measure threshold summarizing the recommended precoding vector sets.

5. Simulation Results

The performance of the proposed coordination scheme can be evaluated through link level simulation on the SINR improvement as well as the system level simulation on average cell throughput together with cell-edge user throughput (5% user throughput). The system parameters for the simulations strictly follow the evaluation methodology proposed by the 3GPP community [12]. Furthermore, according to the current LTE specification, we assume that there are 4 transmit antennas at the base station and 2 receive antennas at the mobile user.

The link level simulation result is contained in Figure 3. Figure 3 compares the SINR of the "Inter-cell Interference coordination with limited feedback" and that of the uncoordinated system. P_1 is the average transmission power of base station 1 and P_2 is the average transmission power of base station 2. In this simulation, we assume that the interfering cell always accepts the recommendation from MS1 and the channel feedback is based on LTE codebook. It can be seen that there is a large improvement in terms of SINR gains of the cell-edge users. This performance gain is achieved by adding a little overhead (a message contains the set information) compared to current LTE system.

The SINR improvement shown in the link level simulation is somewhat biased in the sense that the hit of MS2's throughput is not shown. Since the coordination will limit the choice of precoding vectors at BS2, the throughput of MS2 will be affected. In order to take a more complete picture of the system, we conduct system level simulation. The system level simulation results for average cell throughput and cell-edge user throughput based on SINR_thd = 0 dB are shown in Figures 4 and 5, respectively.

In the system level simulation, we assume that all the base stations accept the recommended set and choose the precoding vectors within the set. Figures 4 and 5 suggest that the improvement in average cell throughput can be as large as 10% while the improvement in 5% sector throughput (cell-edge user throughput) can be as large as 30%. This is because by adopting the recommended set, both received signal strength and received interference strength reduced. The overall SINR again is not significant for the cell-center users while it is huge for the cell-edge users (the limiting factor in SINR for cell-edge UE is interference). Both results suggest that the proposed interference mitigation scheme is extremely efficient for combating the inter-cell interference especially for cell-edge users.

6. Conclusion

Multi-cell MIMO is believed to be one of the enabling technologies in next generation wireless systems. To be specific, the downlink multi-cell MIMO transmission is mainly characterized into two classes [13] in the LTE-A standards: coordinated scheduling and joint transmission.

In the class of joint transmission, data to single MS is simultaneously transmitted from multiple BSs to improve the received signal quality. It has been shown in [6] that this operation mode can significantly increase average cell throughput as well as cell-edge user throughput. However, this scheme requires data to be shared among various cells and requires the network to have the full CSI. In this paper, we investigate schemes falling in the class of coordinated scheduling where data to single MS is instantaneously transmitted from one BS. It is shown that a huge SINR improvement as well as a large throughput increase can be achieved through the proposed scheme. The gains are

achieved by using simple codebook-based channel feedback schemes and are crucial for cell-edge users. This gain is realized through taking advantage of the additional degrees of freedom from the spatial domain. Furthermore, a distance measure threshold-based technology is applied to further reduce the signaling overhead of the proposed scheme. Since the proposed scheme does not need to share data among different BSs and hence reduces the cost of coordination, we do believes this is a promising technology for interference mitigation in future wireless systems such as LTE-Advanced.

Acknowledgment

The authors would like to thank Dr. Farooq Khan, Mr. Zhouyue (Jerry) Pi, and Dr. Donghee Kim for useful discussions.

References

[1] 3GPP, "Feasibility study for further advancements for E-UTRA (LTE-Advanced)," Tech. Rep. TR36.912, 3GPP, Valbonne, France.

[2] A. Jovicic, H. Wang, and P. Viswanath, "On network interference management," submitted to *IEEE Transactions on Information Theory*.

[3] H. Li, B. Liu, and H. Liu, "Transmission schemes for multicarrier broadcast and unicast hybrid systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4321–4330, 2008.

[4] H. Li, S. U. Khan, and H. Liu, "Broadcast network coverage with multi-cell cooperation," to appear in *The International Journal of Digital Multimedia Broadcasting*.

[5] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); physical channels and modulation," Tech. Rep. TS36.211. v 8.6.0, 3GPP, Valbonne, France.

[6] S. Jing, D. N. C. Tse, J. B. Soriaga, J. Hou, J. E. Smee, and R. Padovani, "Multicell downlink capacity with coordinated processing," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, Article ID 586878, 19 pages, 2008.

[7] S. A. Jafar and S. Shamai, "Degrees of freedom region of the MIMO X channel," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 151–170, 2008.

[8] "IEEE 802.16 WirelessMAN Standard: Myths and Facts," ieee802.org.

[9] J. Kotecha and J. Mundarath, "Non-collaborative zerofForcing beamforming in the presence of co-channel interference and spatially correlated channels," in *Proceedings of the 66th IEEE Vehicular Technology Conference (VTC '07)*, pp. 591–595, Baltimore, Md, USA, 2007.

[10] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1994.

[11] D. J. Love and R. W. Heath Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2967–2976, 2005.

[12] 3GPP, "Physical layer aspect for evolved Universal Terrestrial Radio Access (UTRA)," Tech. Rep. TR25.814. v7.1.0, 3GPP, Valbonne, France.

[13] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); further advancements for E-UTRA Physical layer aspects," Tech. Rep. TR36.814. v1.0.0, 3GPP, Valbonne, France.

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 684967, 5 pages
doi:10.1155/2010/684967

Research Article

Phase-Shift Cyclic-Delay Diversity for MIMO OFDM Systems

Young-Han Nam, Lingjia Liu, and Jianzhong (Charlie) Zhang

Wireless Solutions Lab, Samsung Telecommunications America, Richardson, TX 75082, USA

Correspondence should be addressed to Young-Han Nam, ynam@sta.samsung.com

Received 1 September 2009; Accepted 29 November 2009

Academic Editor: Hongxiang Li

Phase-shift cyclic-delay diversity (PS CDD) scheme and space-frequency-block-code (SFBC) PS CDD are developed for multiple-input-multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) systems. The proposed PS CDD scheme preserves the diversity advantage of traditional CDD in uncorrelated multiantenna channels, and furthermore removes frequency-selective nulling problem of the traditional CDD in correlated multiantenna channels.

1. Introduction

It is well known that a multiple-input-multiple-output (MIMO) transmission system can provide benefits on throughput and reliability in multipath fading channels over single-antenna systems [1]. These benefits are now being realized; MIMO transmission schemes have been adopted in most wireless standards including 3GPP long-term evolution (LTE), 3GPP LTE-Advanced, WiMax, and IEEE 802.16 m, where these standards are based on orthogonal frequency division multiple access (OFDMA). OFDMA became popular partly because it is sum-rate optimal for general single-input-single-output (SISO) channels, and the probability of OFDMA being optimal is nonnegligible for MIMO channels [2].

In particular, transmit diversity (TxD) schemes are utilized to realize the reliability benefits of multi-antenna systems in slow-fading environment without channel state information (CSI) available at the transmitter side, by providing multiple signals conveying the same information over different spatial channels. In OFDM-based systems, Alamouti space-frequency block code (SFBC) [3] and cyclic delay diversity (CDD) [4] are widely adopted TxD schemes for 2-transmit antenna diversity (2-TxD) systems. The SFBC scheme has advantages over the other TxD shemes; it is simple to encode at the transmitter side and easy to decode at the receiver side while still achieving the optimal uncoded diversity gain in 2×1 Rayleigh fading channels (i.e., with two transmit and one receive antennas). On the other hand, CDD with small delay is an attractive diversity scheme in OFDM systems, in a sense that it requires only one set of pilot signals, as opposed to the SFBC scheme where two sets of pilot signals are required. However, CDD has some drawbacks such as that it does not have uncoded diversity and thus the block-error rate (BLER) performance is worse than the uncoded diversity schemes like SFBC, and that it has frequency-selective nulls in antenna-correlated channels.

In this paper, we develop and analyze a new TxD scheme for OFDM systems, phase-shift CDD (PS CDD), which takes advantages from both SFBC and CDD, and at the same time mitigates the issue of frequency-selective nulls. The performance of the introduced transmit diversity scheme is evaluated through numerical simualtion results.

2. System Model

We consider a MIMO OFDM system, with M transmit antennas and N receive antennas, where $M = 1, 2, \ldots$ and $N = 1, 2, \ldots$. For each subcarrier $k = 0, 1, \ldots, N_{\text{FFT}} - 1$, where N_{FFT} is the FFT size for the OFDM system, a received signal is described as in the following equation:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k, \tag{1}$$

where $\mathbf{y}_k \in C^{N\times 1}$ is a received vector, $\mathbf{H}_k \in C^{N\times M}$ is a MIMO channel matrix, $\mathbf{x}_k \in C^{M\times 1}$ is a transmitted vector, and $\mathbf{w}_k \in C^{N\times 1}$ is an additive white Gaussian noise vector with mean 0, covariance matrix $\text{diag}(\sigma^2)$. Here, C is the set of complex numbers.

In this paper, we focus on TxD schemes used for robust transmissions in various channel conditions, such as high-Doppler channels and highly frequency-selective channels. Such TxD schemes transmit only one channel-coded stream to ensure maximum reliability, while sacrificing spectral efficiency. We note that general TxD schemes may transmit multiple streams [1] and can also be used for multi-user setting [5] when spatial degrees of freedom of a MIMO channel are greater than one. Among this class of TxD schemes, Alamouti SFBC and CDD are two popular TxD schemes in OFDM-based wireless transmission systems.

3. Background: 2-TxD Schemes

Alamouti SFBC can be described in a 2-transmit and 1-receive antenna system. For an SFBC transmission, two transmit signals at two adjacent subcarriers are paired, which we denote as $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$. These two vectors are constructed in such a way that

$$\mathbf{x}_k = \sqrt{\frac{P}{2}}\begin{bmatrix} s_k \\ s_{k+1} \end{bmatrix}, \qquad \mathbf{x}_{k+1} = \sqrt{\frac{P}{2}}\begin{bmatrix} -s_{k+1}^* \\ s_k^* \end{bmatrix}, \tag{2}$$

where s_k and s_{k+1} are modulated symbols with variance 1 and P is total transmit power at each subcarrier. Under this construction, we are able to obtain an orthogonal system representation at the receiver side, that is, the system transfer matrix is orthogonal:

$$\begin{bmatrix} y_k \\ y_{k+1}^* \end{bmatrix} = \sqrt{\frac{P}{2}}\begin{bmatrix} h_{11} & h_{12} \\ h_{12}^* & -h_{11}^* \end{bmatrix}\begin{bmatrix} s_k \\ s_{k+1} \end{bmatrix} + \begin{bmatrix} w_k \\ w_{k+1}^* \end{bmatrix}, \tag{3}$$

where h_{11} and h_{12} are the channel coefficients between the receive antenna and each of the transmit antennas at subcarriers k and $k+1$, with an assumption that the channel does not vary in the two subcarriers. Due to the orthogonal system transfer matrix, Alamouti SFBC is called *an orthogonal TxD scheme*. Utilizing the orthogonality property, we can detect s_k and s_{k+1}, from the two Alamouti receiver equations as follows:

$$\begin{aligned} h_{11}^* y_k + h_{12} y_{k+1}^* &= \sqrt{\frac{P}{2}}\left(|h_{11}|^2 + |h_{12}|^2\right)s_k + h_{11}^* w_k + h_{12} w_{k+1}^*, \\ h_{12}^* y_k - h_{11} y_{k+1}^* &= \sqrt{\frac{P}{2}}\left(|h_{11}|^2 + |h_{12}|^2\right)s_{k+1} + h_{12}^* w_k - h_{11} w_{k+1}^*. \end{aligned} \tag{4}$$

From the receiver equations, one can easily verify that the received signal-to-noise ratios (SNRs) of the two modulated symbols are the same and equal to $(|h_{11}|^2 + |h_{12}|^2)P/(2\sigma^2)$, which reveals the uncoded diversity gain of the SFBC scheme.

When the number of receive antennas is greater that 1, maximal ratio combining (MRC) would produce a received SNR of

$$\sum_n \left(|h_{n1}|^2 + |h_{n2}|^2\right)P/(2\sigma^2), \tag{5}$$

where h_{n1} and h_{n2} are the channel coefficients between receive antenna n and each of the transmit antennas at subcarriers k and $k+1$. Note that when the number of transmit antennas is larger than 2, no orthogonal TxD schemes have been found achieving the full rate [6]. One popular extension of SFBC for 4-Tx antenna transmitter is frequency-switched transmit diversity (FSTD) [4]. For an SFBC-FSTD transmission, four adjacent subcarriers, k, $k+1$, $k+2$, and $k+3$ are grouped. On the first two subcarriers, that is, k and $k+1$, one SFBC pair is transmitted on the first and the second antennas, while the third and the fourth antennas are turned off. On the third and the fourth subcarriers, that is, $k+2$ and $k+3$, another SFBC pair is transmitted on the third and the fourth antennas, while the first and the second antennas are turned off. SFBC-FSTD is easy to code and decode since it keeps the orthogonality property and achieves coded diversity across four transmit antennas. However, we need four pilot signals for demodulation of SFBC-FSTD, which may increase pilot overhead of a system.

On the other hand, CDD is a coded TxD scheme in an OFDM system, which can be designed for arbitrary number of Tx antennas. In two-transmit and one-receive antenna system, at subcarrier k, a transmit signal $\mathbf{x}_k$ coded with CDD is

$$\mathbf{x}_k = \sqrt{\frac{P}{2}}\begin{bmatrix} s_k \\ e^{jk\delta}s_k \end{bmatrix}, \tag{6}$$

where δ is a positive number called CDD delay (e.g., $\delta = 2\pi/N_{\mathrm{FFT}}$). Then, a received signal at subcarrier k is written as

$$y_k = \begin{bmatrix} h_{11} & h_{12} \end{bmatrix}\sqrt{\frac{P}{2}}\begin{bmatrix} s_k \\ e^{jk\delta}s_k \end{bmatrix} + w_k = \sqrt{\frac{P}{2}}\left(h_{11} + h_{12}e^{jk\delta}\right)s_k + w_k. \tag{7}$$

As we can see from the receiver equation, CDD does not give uncoded diversity, as received SNR is $|h_{11} + h_{12}e^{jk\delta}|^2 P/(2\sigma^2)$. However, CDD in combination of channel coding across modulation symbols $\{s_k\}$ mapping to multiple subcarriers increases a coded diversity gain, as CDD increases the frequency selectivity of the composite channel: $h_{11} + h_{12}e^{jk\delta}$. We also note that with CDD, a receiver needs to know only the composite channel $h_{11} + h_{12}e^{jk\delta}$ for demodulation, especially when the delay δ is small so that the channel does not vary abruptly over subcarriers. This is one benefit of CDD over SFBC which requires knowledge of two channels.

CDD can be easily extended to cases where the number of Tx antennas is grater than 2. For example, when the number of Tx antennas is 4, we have a transmit signal $\mathbf{x}_k$ coded with CDD at subcarrier k as

$$\mathbf{x}_k = \sqrt{\frac{P}{4}}\begin{bmatrix} s_k \\ e^{jk\delta}s_k \\ e^{j2k\delta}s_k \\ e^{j3k\delta}s_k \end{bmatrix}. \tag{8}$$

A well-known drawback of CDD is frequency-selective nulling. As CDD artificially increases frequency selectivity, in some subcarriers, the two component channels of the composit channel, $h_{11} + h_{12}e^{jk\delta}$, coherently add, while in some other subcarriers, they destructively add. This problem becomes severer when the two channels h_{11} and h_{12} are correlated, which occurs when the two transmit antennas are geometrically close. This particular issue of CDD prevented it from being accepted as robust TxD schemes in wireless communication implementations, despite the benefits of CDD.

Reviewing these two TxD schemes of SFBC and CDD, in summary, we realize that SFBC is robust but not extendable to systems with large number of transmit antennas, while CDD is easily extendable and requires only one pilot signal but not robust in correlated channels. In the sequel, we develop new TxD schemes taking the advantages of both schemes while still ensuring robustness in correlated channels.

4. Design of 4-Tx Diversity Schemes

4.1. Phase-Shift Cyclic Delay Diversity (PS CDD). We recall that a major problem of CDD is nonrobustness in correlated channels when two terms from $h_{11}+h_{12}e^{jk\delta}$ destructively add. When the number of Tx antennas is greater than or equal to 4, a simple variation of CDD may prevent frequency nulling from occuring. Instead of giving the same phase component on the signals transmitted in the four transmit antennas as in (8), we attempt to apply a phase shift of ϕ in the signal transmitted in the fourth antenna. With this phase shift CDD (PS CDD), transmit signal $\mathbf{x}_k$ coded at subcarrier k is

$$\mathbf{x}_k = \sqrt{\frac{P}{4}}\begin{bmatrix} s_k \\ e^{jk\delta}s_k \\ e^{j2k\delta}s_k \\ e^{j(3k\delta+\phi)}s_k \end{bmatrix}. \tag{9}$$

In thise case, the composite channel at the receiver's point of view is $h_{11} + h_{12}e^{jk\delta} + h_{13}e^{2jk\delta} + h_{14}e^{j(3k\delta+\phi)}$.

To facilitate the analysis of performance of (9) in strongly correlated channels, we assume

$$h_{11} = h_{12} = h_{13} = h_{14} = h. \tag{10}$$

Accordingly, the composite channel becomes

$$h\left(1 + e^{jk\delta} + e^{j2k\delta} + e^{j(3k\delta+\phi)}\right). \tag{11}$$

To gain some insights on this approach, let $\delta = 2\pi/N_{\text{FFT}}$, $\phi = \pi$, and $N_{\text{FFT}} = 1024$ and compare the composite channel powers of CDD scheme (8) and PS CDD (9) normalized by $|h|^2$, as shown in Figure 1. In the figure, we can see that CDD suffers from frequency-selective nulls at $N_{\text{FFT}}/4, N_{\text{FFT}}/2$ and $3N_{\text{FFT}}/4$, while PS CDD does not have frequency nulls. Recalling that CDD intentionally introduces frequency selectivity for additional frequency diversity, we want to have PS CDD that has a property of having good frequency selectivity, while not suffering from frequency nulls. We may characterize this goal by considering the following objective function:

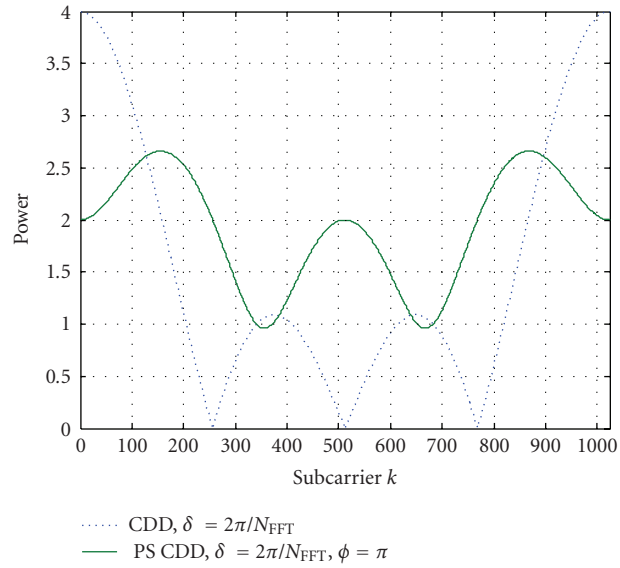


Figure 1: Comparison of powers of composite channels with CDD and PS CDD.

$$\underset{\phi}{\text{maximize}}\ (P_{\max} - P_{\min})P_{\min}, \tag{12}$$

where $P_{\min}$ and $P_{\max}$ are the minimum power and the maximum power, respectively, of a scheme across all the subcarriers. We find that $(P_{\max} - P_{\min})P_{\min}$ of the CDD and the PS CDD in Figure 1 are 0 and 1.6323, and hence under this objective function, PS CDD is better than CDD, in terms of both introducing frequency selectivity and keeping minimum power large. We note that the optimal ϕ with (9) can be found with numerical method.

The most general form of PS CDD can be written as

$$\mathbf{x}_k = \sqrt{\frac{P}{4}}\begin{bmatrix} s_k \\ e^{j(k\delta_1+\phi_1)}s_k \\ e^{j(k\delta_2+\phi_2)}s_k \\ e^{j(k\delta_3+\phi_3)}s_k \end{bmatrix}, \tag{13}$$

where we may optimize the performance with choosing parameters δ's and ϕ's. We also note that for demodulation PS CDD signal, we need only one pilot signal for the composite channel $h_{11} + h_{12}e^{j(k\delta_1+\phi_1)} + h_{13}e^{j(k\delta_2+\phi_2)} + h_{14}e^{j(k\delta_3+\phi_3)}$.

4.2. Space-Frequency Block-Code with Phase-Shift Cyclic Delay Diversity (SFBC PS CDD). In Section 4.1, we have introduced PS CDD that does not suffer from frequency-selective nulls, while keeping the coded diversity benefit of CDD and maintaining the required number of pilot signals to be one. In this section, we combine Alamouti SFBC and PS CDD, so that a new TxD scheme can enjoy uncoded diversity while keeping some benefits of PS CDD.

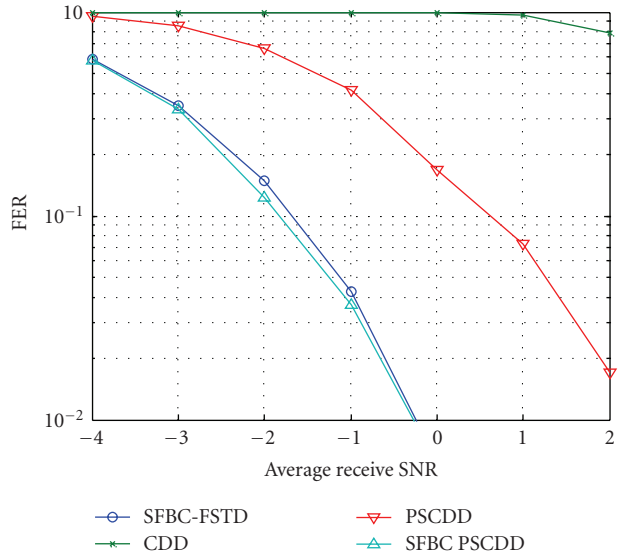


Figure 2: BLER performance in correlated channels.

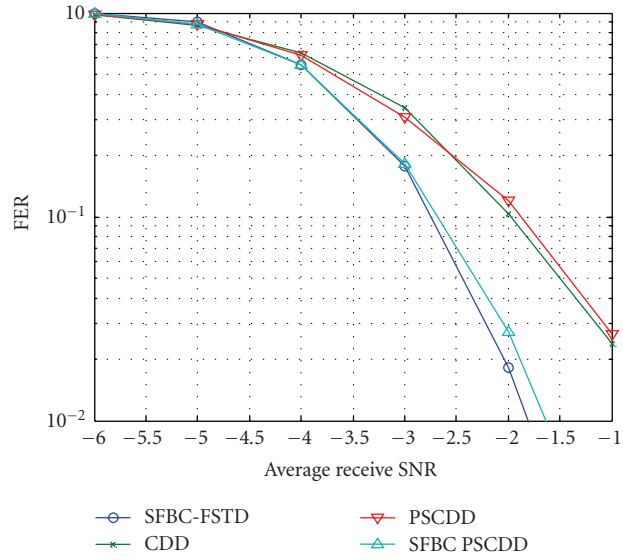


Figure 3: BLER performance in uncorrelated channels.

For an SFBC PS CDD transmission, two transmit signals at two adjacent subcarriers are paired, which we denote $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$. These two vectors are constructed in such a way that

$$\mathbf{x}_k = \sqrt{\frac{P}{4}}\begin{bmatrix} s_k \\ e^{j(k\delta_1+\phi_1)}s_{k+1} \\ e^{j(k\delta_2+\phi_2)}s_k \\ e^{j(k\delta_3+\phi_3)}s_{k+1} \end{bmatrix}, \qquad \mathbf{x}_{k+1} = \sqrt{\frac{P}{4}}\begin{bmatrix} -s_{k+1}^* \\ e^{j(k\delta_1+\phi_1)}s_k^* \\ -e^{j(k\delta_2+\phi_2)}s_{k+1}^* \\ e^{j(k\delta_3+\phi_3)}s_k^* \end{bmatrix}. \tag{14}$$

With this construction, we obtain an orthogonal system of equations at the receiver side:

$$\begin{bmatrix} y_k \\ y_{k+1}^* \end{bmatrix} = \sqrt{\frac{P}{2}}\begin{bmatrix} \widetilde{h}_{11} & \widetilde{h}_{12} \\ \widetilde{h}_{12}^* & -\widetilde{h}_{11}^* \end{bmatrix}\begin{bmatrix} s_k \\ s_{k+1} \end{bmatrix} + \begin{bmatrix} w_k \\ w_{k+1}^* \end{bmatrix}, \tag{15}$$

where

$$\begin{aligned} \widetilde{h}_{11} &= \left(h_{11} + e^{j(k\delta_2+\phi_2)}h_{13}\right), \\ \widetilde{h}_{12} &= \left(e^{j(k\delta_1+\phi_1)}h_{12} + e^{j(k\delta_3+\phi_3)}h_{14}\right). \end{aligned} \tag{16}$$

From (14), (15), and (16), we see that SFBC PS CDD construction reduces to a 2-Tx SFBC scheme, which ensures easy decodability relying on orthogonal structure and allows us to have uncoded diversity. Furthermore, each of the effective channels $\widetilde{h}_{11}$ and $\widetilde{h}_{12}$ is constructed with CDD, which allows us to have coded diversity with intentionally introduced frequency selectivity and facilitates demodulation of SFBC PS CDD using two pilot signals for $\widetilde{h}_{11}$ and $\widetilde{h}_{12}$.

5. Numerical Results

In this section, we present numerical simulation results of block-error rate (BLER, or frame error rate, FER) performance comparing the TxD schemes introduced in this paper and some existing schemes such as CDD and SFBC-FSTD. For the simulation, ITU typical urban 6-path channel model (TU-6) has been used and 120 km/hr is assumed for terminal speed. Furthermore, we consider a MIMO channel with 4-transmit and 2-receive antennas, where 4 transmit antennas are correlated with correlation coefficients 0.9 and 0, while 2 receive antennas are uncorrelated. For channel coding, 3GPP Turbo code [7] is used with code rate 1/3, and QPSK modulation is used. Channel-coded and modulated signals go through 6 distributed sets of 12 subcarriers in each time slot (or per block of time). For demodulation, perfect (or ideal) channel estimation is assumed. At the receiver, maximal ratio combing is used followed by per-receive-antenna SFBC decoder.

Figure 2 shows BLER curves obtained with various TxD schemes under highly correlated channels whose correlation coefficient is 0.9. As we discussed earlier, CDD performs worse than the others. As PS CDD removes frequency nulls, the performance is better than CDD. Both SFBC-FSTD and SFBC-PSCDD perform the best among these four schemes. Considering the fact that SFBC-PSCDD requires only two pilot signals, SFBC-PSCDD can potentially achieve larger spectral efficiency than SFBC-FSTD. On the other hand, Figure 3 shows the BLER curves obtained under uncorrelated channels. In the uncorrelated case, SFBC-FSTD and SFBC-PSCDD show similar performance and outperform PSCDD and CDD.

6. Conclusion

In this paper, we have introduced phase-shift cyclic delay diversity PS CDD and SFBC PS CDD schemes. The proposed schemes treat frequency-selective nulling problem of traditional CDD. In particular, SFBC PS CDD takes benefits of both SFBC and PS CDD, and achieves robust block-error

rate performance in both highly correlated and uncorrelated channels, while requiring only two pilot signals, as opposed to the well-known SFBC-FSTD requiring four pilot signals.

References

[1] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, 2003.

[2] H. Li and H. Liu, "On the optimality of OFDMA MIMO channels," in *Proceedings of the 40th Asilomar Conference on Signals, Systems, and Computers (ACSSC '06)*, pp. 1757–1761, Pacific Grove, Calif, USA, October 2006.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[4] F. Khan, *LTE for 4G Mobile Broadband*, Cambridge University Press, Cambridge, UK, 1st edition, 2009.

[5] Y.-H. Nam and H. E. Gamal, "On the optimality of lattice coding and decoding in multiple access channels," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '07)*, pp. 211–215, Nice, France, 2007.

[6] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[7] 3GPP TS 36.212 ver 8.7.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding," June 2009.

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 642542, 9 pages
doi:10.1155/2010/642542

Research Article

Spatial Domain Resource Sharing for Overlapping Cells in Indoor Environment

Riichi Kudo,[1] Yasushi Takatori,[1] Kentaro Nishimori,[1] Takeo Ichikawa,[1] Tomoki Murakami,[1] Masato Mizoguchi,[1] and Masahiro Morikura[2]

[1] *NTT Network Innovation Laboratories, NTT Corporation, Yokosuka 239-0847, Japan*
[2] *Graduate School of Informatics, Kyoto University, Kyoto 606-8317, Japan*

Correspondence should be addressed to Riichi Kudo, kudo.riichi@lab.ntt.co.jp

Received 1 July 2009; Accepted 25 December 2009

Academic Editor: Hongxiang Li

As microcell wireless systems become more widespread, intercell interference among the access points will increase due to the limited frequency resource. In the overlapping cell scenario, radio resources should be shared by multiple cells. Although time and frequency resource sharing has been described in many papers, there is no detailed report on dynamic spatial resource sharing among multiple cells for microcell wireless systems. Thus, we present the effectiveness of spatial resource sharing among two access points. We introduce two scenarios based on the zero forcing method; one is the primary-secondary AP scenario and the other is the cooperative AP scenario. To evaluate the transmission performance of spatial resource sharing, channel matrices are measured in an indoor environment. The simulation results using the measured channel matrices show the potential of spatial resource sharing.

1. Introduction

The rapid increase in data traffic has led to strong demands for large-capacity transmission systems. Wireless local area networks (WLAN) and cellular systems are two major wireless access systems that are being targeted to achieve even higher data rates. Orthogonal frequency division multiplexing (OFDM) and multiple input multiple output (MIMO) have been recognized as effective ways to attain higher throughput for IEEE 802.11n [1] and 3G Long term evolution (LTE) [2]. Unfortunately, the recent advances in these techniques still lag the demands being made, so bandwidth of each carrier is being increased from one generation to the next [1, 2]. Another trend in wireless access systems is the rapid increase in the number of wireless devices. The total number of cellular phone subscribers reached 100 million at the end of December in 2007 in Japan and exceeded the number of land phone subscribers [3]. The number of WLAN chip shipments in each year is also increasing and it reached to 200 million in 2008 [4]. It is thought that people will start putting wireless devices on machines in the near future. This changes the usage of wireless access from people to people (P-to-P) or people to machine (P-to-M) communication to machine to machine (M-to-M) communication and a huge growth in this market can be anticipated [1, 2].

The above trends bring new challenges for wireless technologies because both the larger bandwidth and the larger number of wireless devices may cause significant shortages in the frequency channels. One example is the ongoing standardization for very high throughput (VHT) wireless LANs in bands below 6 GHz, that is, IEEE802.11ac, which is expected to be the next standard after IEEE802.11n. It has been virtually agreed that IEEE802.11ac will double the maximum bandwidth of IEEE802.11n from 40 MHz to 80 MHz to achieve 1 Gbps throughput at the median access control (MAC) service access point (SAP) [5]. Since the microwave frequency band has been already assigned to various radio systems, it is difficult to expand the frequency bands allocated for wireless LANs into the microwave band. Thus, the number of available frequency channels would be half that of IEEE802.11n and is more likely that a frequency channel will be used at a nearby basic service set (BSS), which consists of access point (AP) and associated stations (STAs).

Therefore, how APs and STAs behave in the overlapping cell among BSS is important [6].

In the overlapping cell scenario, radio resources should be shared by multiple cells. There are three resources to be shared, time, frequency, and spatial resources. The first one, time resource sharing is realized by carrier sense multiple access collision avoidance (CSMA/CA) or request to send/clear to send (RTS/CTS) protocols [7, 8]. It allocates different time durations to each access to avoid inter-cell-interference (ICI). Dynamic resource sharing in the time domain is actualized by making transmission opportunity proportional to traffic load at each cell. The second one, frequency resource sharing, allocates a different frequency band to each cell to prevent intercell interference. The dynamic channel assignment (DCA) scheme was proposed to improve the spectrum efficiency [9]. The DCA concept has been extended to support the heterogeneous multiple radio system scenario and is considered as a type of cognitive radio technology. The drawback, the frequency domain method is the complexity of the resource allocation procedure. In the following, we will focus on the last radio resource, spatial resource, to improve the spectrum efficiency further.

Compared to time and frequency resource sharing, spatial resources sharing research is relatively immature especially for the overlapping cell scenario. MIMO technology which uses spatial signal processing is employed by the latest wireless access standards such as 3 G long term evolution (LTE) [10] and IEEE802.11n [11]. It increases the channel capacity between one transmitter and one receiver by using spatial domain. MIMO technology has been extended to multiuser MIMO (MU-MIMO) to attain high channel capacity even for multiple simple wireless terminals [12, 13]. Many spatial resource sharing methods for MU-MIMO were proposed to take into account various QoS conditions for isolated cells but not for overlapping cells [14, 15]. On the other hand, ICI suppression using spatial domain technology, for example, beamforming or smart antennas, was extensively studied to shorten the frequency reuse distance [16, 17]. It improves the spectrum efficiency in total system level and can be interpreted as one approach for spatial resource sharing. Recently, femto-cell technology was developed to actualize umbrella cells where multiple femto-cells are created inside a macrocell by transmission power adjustments [18]. Although these techniques realize spatial resource sharing even in autonomous systems, they are effective only for static resource allocation, and it remains hard to respond to dynamic traffic changes in each cell. The cooperation of multiple APs was proposed to attain the highest spectrum efficiency for a multiple overlapping cell scenario [19, 20]. In this method, all APs are connected to an access controller which manages all spatial resources using the channel state information (CSI) of all AP- station (STA) links. Unfortunately, this approach cannot be applied to the general case, for example, overlapping cells in a WLAN, where multiple APs are connected to different wired networks. To the best of our knowledge, no detailed report has examined dynamic spatial resource sharing among multiple cells where each cell is connected to a different network. Thus, this paper focuses on this scenario and presents the throughput improvement obtained by spatial resource sharing for overlapping cells. In spatial resource sharing, AP allocates spatial resources to STAs if the spatial resources are not consumed by the other AP. The conventional AP waits for another time duration even when only one spatial stream is used at the overlapped cells. Therefore, spatial resource sharing for overlapped cell enables dynamic spatial resource allocation and further improvements in spectrum efficiency.

In this paper, two scenarios (both based on the zero forcing method) for spatial resource sharing are considered. One is the primary-secondary AP scenario and the other is the cooperative AP scenario. To evaluate the effectiveness of these methods, a channel state information measurement experiment is conducted in a large office and 8×4 MIMO OFDM channel matrices for 48 subcarriers are obtained using the short and long preambles based on the IEEE802.11a standard. Although the measurement environment is a large room, we consider the two-room scenario by separating two APs with a "virtual" wall. The simulation results confirm the potential of spatial resource sharing.

This paper is organized as follows. Section 2 describes eigenvector transmission and the impact of spatial resource sharing in channel capacity. In Section 3, we introduce the experimental setup and the measurement environment. Section 4 presents simulation results. Section 5 summarizes the paper.

Throughout the paper, superscript $*$ and superscript H, denote the complex conjugate, and the Hermitian transposition, respectively. Term $|\mathbf{A}|$ is the determinant of $\mathbf{A}$.

2. System Model

We consider a downlink MIMO system for a single AP and STA in the overlapping cell scenario. To evaluate the effectiveness of spatial resource sharing, we compare the two overlapping cell scenario to the single cells scenario using time resource sharing. In this section, we describe the channel conditions and two scenarios of spatial domain resource sharing; the primary-secondary AP scenario and the cooperative AP scenario. Furthermore, we focus on the CSI error and show the relationship between CSI and ICI.

2.1. Channel Model for Time Resource Sharing. It is assumed that the number of APs is two, each hosts U mobile stations (STAs) and each AP and STA have N transmit antennas and M receive antennas, respectively. We denote this as the $N > M$ system. Here, transmission from AP k, $k = 1, 2$, is considered. The channel state between AP k and STA i (being hosted by AP k) is represented by channel matrix $\mathbf{H}_{k,i} \in \mathbb{C}^{M\times N}$, $k = 1, 2$, and $i = 1, 2, \ldots, U$. The $L_{k,i} \times 1$ transmit symbol vector for STA i, $\mathbf{x}_{k,i}$, is transmitted using transmission weight $\mathbf{W}_{k,i} \in \mathbb{C}^{N\times L_{k,i}}$. $L_{k,i}$ is the number of data streams transmitted simultaneously to STA i from AP k. The signal vector received at STA i is given as

$$\mathbf{y}_{k,i} = \mathbf{H}_{k,i}\mathbf{W}_{k,i}\mathbf{x}_{k,i} + \mathbf{n}_{k,i}, \tag{1}$$

where $\mathbf{n}_{k,i} \in \mathbb{C}^{M\times 1}$ is the additive white Gaussian noise vector and the elements of $\mathbf{n}_{k,i}$ are assumed to have a variance, σ^2. In this paper, σ^2 is set to be one for all STAs. Even if every STA has a different noise variance, we can consider the same situation by normalizing the norm of the channel matrix without loss of generality.

2.2. Channel Model for Spatial Resource Sharing. When spatial resource sharing is active, the two APs can transmit in the same time slot. The received signal vectors at STA i, which is communicating with AP 1, and at STA j, which is communicating with AP 2, are given as

$$\begin{aligned}\mathbf{y}_{1,i} &= \mathbf{H}_{1,i}\mathbf{W}_{1,i}\mathbf{x}_{1,i} + \mathbf{H}_{\mathrm{I}2,i}\mathbf{W}_{2,j}\mathbf{x}_{2,j} + \mathbf{n}_{1,i},\\ \mathbf{y}_{2,j} &= \mathbf{H}_{2,j}\mathbf{W}_{2,j}\mathbf{x}_{2,j} + \mathbf{H}_{\mathrm{I}1,j}\mathbf{W}_{1,i}\mathbf{x}_{1,i} + \mathbf{n}_{2,i},\end{aligned} \tag{2}$$

where $\mathbf{H}_{\mathrm{I},l,i} \in \mathbb{C}^{N\times M}$ is the interference channel matrix between AP l and STA i of AP k, where l is 1 or 2 and $l \neq k$. Therefore, $\mathbf{H}_{\mathrm{I}1,i}$ and $\mathbf{H}_{\mathrm{I}2,j}$ are the channel matrices between AP 1 and STA j of AP 2, and between AP 2 and STA i of AP 1, respectively.

2.3. Eigenvector Transmission. As the transmission weight for the time resource sharing scenario, we employ eigenvector transmission [21] and the water pouring strategies [22]. This is because these techniques maximize the mutual information when the channel state information is available at AP. In eigenvector transmission, the transmission weight at AP k for STA i is calculated as the eigenvectors of the correlation matrix, $\mathbf{H}_{k,i}^H\mathbf{H}_{k,i}$. By using singular value decomposition (SVD), the eigenvectors are obtained as the right singular vectors of $\mathbf{H}_{k,i}$ as

$$\mathbf{H}_{k,i} = \mathbf{U}_{k,i}\left(\mathbf{\Sigma}_{k,i} \;\; \mathbf{0}\right)\left(\mathbf{V}_{k,i}^{(\mathrm{s})} \;\; \mathbf{V}_{k,i}^{(\mathrm{n})}\right)^H, \tag{3}$$

where $\mathbf{V}_{k,i}^{(\mathrm{s})} \in \mathbb{C}^{N\times M}$ and $\mathbf{V}_{k,i}^{(n)} \in \mathbb{C}^{N\times(N-M)}$ represent the right singular vectors for the signal space and the null space, respectively. The diagonal elements of $\mathbf{\Sigma}_{k,i}$ are the square roots of $\lambda_{k,1}, \lambda_{k,2}, \ldots, \lambda_{k,M,}$ which are the eigenvalues of $\mathbf{H}_{k,i}^H\mathbf{H}_{k,i}$. In eigenvector transmission, transmission weight $\mathbf{W}_{k,i}$ can be chosen as $L_{k,i}$ vectors of $\mathbf{V}_{k,i}^{(\mathrm{s})}$ corresponding to higher eigenvalues. $L_{k,i}$ and the norm of each vectors of $\mathbf{W}_{k,i}$ are determined by water pouring strategies subject to power constraint, that is, trace$(\mathbf{W}_{k,i}^H\mathbf{W}_{k,i})$ is one. By using the right singular vectors as the transmission weight, high SNRs corresponding to the eigenvalues can be obtained. We define the transmission weight in eigenvector transmission as $\mathbf{W}_{\mathrm{E},k,i}$. In (3), $\mathbf{U}_{k,i}$ denotes the left singular vectors and corresponds to the reception weight at STA i in eigenvector transmission.

2.4. ZF Transmission for Interference AP. When there are two APs using the same frequency and the same time slot, the interference from the other AP degrades the transmission performance. Thus, we consider the transmission weight that prevents the interference to the other AP. In AP k, the interference channel matrix for STA j of the other AP, AP l, is expressed as

$$\mathbf{H}_{\mathrm{I},k,j} = \mathbf{U}_{\mathrm{I},k,j}\left(\mathbf{\Sigma}_{\mathrm{I},k,j} \;\; \mathbf{0}\right)\left(\mathbf{V}_{\mathrm{I},k,j}^{(\mathrm{s})} \;\; \mathbf{V}_{\mathrm{I},k,j}^{(\mathrm{n})}\right)^H, \tag{4}$$

where $\mathbf{V}_{\mathrm{I},k,j}^{(\mathrm{s})} \in \mathbb{C}^{N\times M}$ and $\mathbf{V}_{\mathrm{I},k,j}^{(\mathrm{n})} \in \mathbb{C}^{N\times(N-M)}$ represent the right singular vectors for the signal space and the null space, respectively. The diagonal elements of $\mathbf{\Sigma}_{\mathrm{I},k,j}$ are the square roots of $\lambda_{k,1}, \lambda_{k,2}, \ldots, \lambda_{k,M,}$ which are the null space eigenvalues. To obtain the transmission weight that does not interfere with the transmission of AP l, the null space channel matrix, $\mathbf{H}_{k,i}\mathbf{V}_{k,j}^{(\mathrm{n})}$ is decomposed by SVD to yield

$$\mathbf{H}_{k,i}\mathbf{V}_{\mathrm{I},k,j}^{(\mathrm{n})} = \mathbf{U}_{\mathrm{N},k,i}\left(\mathbf{\Sigma}_{\mathrm{N},k,j} \;\; \mathbf{0}\right)\left(\mathbf{V}_{\mathrm{N},k,j}^{(\mathrm{s})} \;\; \mathbf{V}_{\mathrm{N},k,j}^{(\mathrm{n})}\right)^H, \tag{5}$$

where $\mathbf{V}_{\mathrm{N},k,j}^{(s)} \in \mathbb{C}^{(N-M)\times M}$ and $\mathbf{V}_{\mathrm{N},k,j}^{(n)} \in \mathbb{C}^{N-M\times(N-2M)}$ represent the null space right singular vectors for the signal space and the null space, respectively. In ZF transmission, the transmission weight $\mathbf{W}_{k,i}$ can be chosen as $L_{k,i}$ vectors of $\mathbf{V}_{\mathrm{I},k,i}^{(n)}\mathbf{V}_{\mathrm{N},k,i}^{(s)}$ corresponding to the higher eigenvalues of $\lambda_{\mathrm{N},k,i}$. $L_{k,i}$ and the norm of each vector of $\mathbf{W}_{k,i}$ is determined by water pouring strategies using $\lambda_{\mathrm{N},k,i}$. The transmission weight is expressed as $\mathbf{W}_{\mathrm{Z},k,i}$.

2.5. Channel State Informationat AP And STA. The CSI is obtained at the transmitter and the receiver. We assume that the STA holds the perfect CSI while the CSI at the AP has some estimation error. The channel matrix obtained at AP k and is expressed as

$$\begin{aligned}\overline{\mathbf{H}}_{k,i} &= \mathbf{H}_{k,i} + \mathbf{E}_{k,i},\\ \overline{\mathbf{H}}_{\mathrm{I},k,j} &= \mathbf{H}_{\mathrm{I},k,j} + \mathbf{E}_{\mathrm{I},k,j},\end{aligned} \tag{6}$$

where $\mathbf{E}_{k,i}$ and $\mathbf{E}_{\mathrm{I},k,j}$ are all i.i.d zero-mean complex Gaussian with variance of σ_e^2. Thus, the transmission weight is calculated using the imperfect channel matrix, $\overline{\mathbf{H}}_{k,i}$ and $\overline{\mathbf{H}}_{\mathrm{I},k,j}$.

2.6. Achievable Bit Rate. When two APs transmit signals in different time slots, that is, time resource sharing systems, the mutual information between AP k and STA i is expressed using the eigenvector transmission weight as

$$C_{\mathrm{S},k,i} = \log_2\left|\mathbf{I}_M + \mathbf{H}_{k,i}\mathbf{W}_{\mathrm{E},k,i}\mathbf{W}_{\mathrm{E},k,i}^H\mathbf{H}_k^H\right|, \tag{7}$$

where k is 1 or 2 and $i = 1, 2, \ldots, U$. In this scenario, the simultaneous transmission at both APs is avoided. CSMA/CA and RTS/CTS/ACQ in MAC enables this scenario.

We describe the channel capacity in a spatial resource sharing system with two APs as

$$\begin{aligned}C_{\mathrm{M},i,j} = {}& \log_2\left|\mathbf{I}_M + \frac{\mathbf{H}_{1,i}\mathbf{W}_{1,i}\mathbf{W}_{1,i}^H\mathbf{H}_1^H}{\mathbf{I}_M + \mathbf{H}_{\mathrm{I}2,i}\mathbf{W}_{2,j}\mathbf{W}_{2,j}^H\mathbf{H}_{\mathrm{I}2,i}^H}\right|\\ &+ \log_2\left|\mathbf{I}_M + \frac{\mathbf{H}_{2,j}\mathbf{W}_{2,j}\mathbf{W}_{2,j}^H\mathbf{H}_{2,j}^H}{\mathbf{I}_M + \mathbf{H}_{\mathrm{I}2,i}\mathbf{W}_{1,i}\mathbf{W}_{1,i}^H\mathbf{H}_{\mathrm{I}2,j}^H}\right|,\end{aligned} \tag{8}$$

where $i = 1, 2, \ldots, U$, and $j = 1, 2, \ldots, U$. We consider two scenarios for spatial resource sharing. One is primary-secondary AP method based on zero forcing (PSZ) and the other is cooperative AP method based on zero forcing (CZ). In PSZ mode, AP k determines the transmission weight based on the eigenvector transmission using $\mathbf{V}_{k,i}^{(s)}$ while AP l ($l \neq k$) uses the transmission weight based on the ZF using $\mathbf{V}_{\mathrm{I},k,i}^{(n)}\mathbf{V}_{\mathrm{N},k,i}^{(s)}$. Thus, the transmission of AP l suffers interference from the transmission of AP k and the signal power is degraded due to the null space beamforming, although the transmission of AP k is expected to have the high channel capacity, see (7), due to the suppressed interference at AP l. The achievable transmission bit rate in the PSZ method is given as

$$C_{\mathrm{PSZ},k,i,j} = \log_2 \left| \mathbf{I}_M + \frac{\mathbf{H}_{k,i}\mathbf{W}_{\mathrm{E},k,i}\mathbf{W}_{\mathrm{E},1,i}^H\mathbf{H}_1^H}{\mathbf{I}_M + \mathbf{H}_{\mathrm{I},2,i}\mathbf{W}_{\mathrm{Z},2,j}\mathbf{W}_{\mathrm{Z},2,j}^H\mathbf{H}_{\mathrm{I},2,i}^H} \right| + \log_2 \left| \mathbf{I}_M + \frac{\mathbf{H}_{l,j}\mathbf{W}_{\mathrm{Z},2,j}\mathbf{W}_{\mathrm{Z},2,j}^H\mathbf{H}_{l,j}^H}{\mathbf{I}_M + \mathbf{H}_{\mathrm{I},l,i}\mathbf{W}_{\mathrm{E},k,i}\mathbf{W}_{\mathrm{E},k,i}^H\mathbf{H}_{\mathrm{I},l,j}^H} \right|, \tag{9}$$

where AP k ($k = 1, 2$) is the primary AP. If there is no channel estimation error at the AP, (9) is rewritten as

$$C_{\mathrm{PSZ},k,i,j} = \log_2 \left| \mathbf{I}_M + \mathbf{H}_{k,i}\mathbf{W}_{\mathrm{E},k,i}\mathbf{W}_{\mathrm{E},1,i}^H\mathbf{H}_1^H \right| + \log_2 \left| \mathbf{I}_M + \frac{\mathbf{H}_{l,j}\mathbf{W}_{\mathrm{Z},2,j}\mathbf{W}_{\mathrm{Z},2,j}^H\mathbf{H}_{l,j}^H}{\mathbf{I}_M + \mathbf{H}_{\mathrm{I},l,i}\mathbf{W}_{\mathrm{E},k,i}\mathbf{W}_{\mathrm{E},k,i}^H\mathbf{H}_{\mathrm{I},l,j}^H} \right|. \tag{10}$$

In the cooperating mode, both APs uses the transmission weight based on ZF. This case is expected to yield the high capacity among two APs although the signal powers at the destination STAs are degraded due to ZF beamforming. The achievable transmission bit rate in CZ method is given as

$$C_{\mathrm{CZ},i,j} = \log_2 \left| \mathbf{I}_M + \frac{\mathbf{H}_{k,i}\mathbf{W}_{\mathrm{Z},k,i}\mathbf{W}_{\mathrm{Z},k,i}^H\mathbf{H}_{k,i}^H}{\mathbf{I}_M + \mathbf{H}_{\mathrm{I},k,i}\mathbf{W}_{\mathrm{Z},l,j}\mathbf{W}_{\mathrm{Z},l,j}^H\mathbf{H}_{\mathrm{I},k,i}^H} \right| + \log_2 \left| \mathbf{I}_M + \frac{\mathbf{H}_{l,j}\mathbf{W}_{\mathrm{Z},l,j}\mathbf{W}_{\mathrm{Z},l,j}^H\mathbf{H}_{l,j}^H}{\mathbf{I}_M + \mathbf{H}_{\mathrm{I},l,i}\mathbf{W}_{\mathrm{Z},k,i}\mathbf{W}_{\mathrm{Z},k,i}^H\mathbf{H}_{\mathrm{I},l,j}^H} \right|. \tag{11}$$

If there is no channel estimation error at the AP, (11) is given as

$$C_{\mathrm{CZ},i,j} = \log_2 \left| \mathbf{I}_M + \mathbf{H}_{k,i}\mathbf{W}_{\mathrm{Z},k,i}\mathbf{W}_{\mathrm{Z},k,i}^H\mathbf{H}_{k,i}^H \right| + \log_2 \left| \mathbf{I}_M + \mathbf{H}_{l,j}\mathbf{W}_{\mathrm{Z},l,j}\mathbf{W}_{\mathrm{Z},\mathrm{l},j}^H\mathbf{H}_{l,j}^H \right|. \tag{12}$$

2.7. Intercell Interference. In spatial resource sharing, the ICI determines the channel capacity in (9) and (11). Therefore, the ICI is analyzed by numerical approach. We define the variance of elements of the interference channel matrix, $\mathbf{H}_{\mathrm{I},k,i}$, as γ_{I} to show the characteristics of the ICI. When the eigenvector transmission is applied, the ICI power, I_E, is expressed as

$$\begin{aligned} I_{\mathrm{E}} &= \mathrm{trace}\left(\mathbf{W}_{\mathrm{E},k,i}^H\mathbf{H}_{\mathrm{I},l,j}^H\mathbf{H}_{\mathrm{I},l,j}\mathbf{W}_{\mathrm{E},k,i}\right) \\ &= \mathrm{trace}\left(\mathbf{W}_{\mathrm{E},k,i}^H\mathbf{V}_{\mathrm{Z},l,j}^{(s)}\boldsymbol{\Sigma}_{\mathrm{I},l,j}^{(s)}\mathbf{U}_{\mathrm{I},l,j}^H\mathbf{U}_{\mathrm{I},l,j}\boldsymbol{\Sigma}_{\mathrm{I},l,j}^{(s)}\mathbf{V}_{\mathrm{Z},l,j}^{(s)H}\mathbf{W}_{\mathrm{E},k,i}\right) \\ &= \mathrm{trace}\left(\mathbf{W}_{\mathrm{E},k,i}^H\mathbf{V}_{\mathrm{Z},l,j}^{(s)}\boldsymbol{\Sigma}_{\mathrm{I},l,j}^{(s)2}\mathbf{V}_{\mathrm{Z},l,j}^{(s)H}\mathbf{W}_{\mathrm{E},k,i}\right) \\ &= \mathrm{trace}\left(\mathbf{A}^H\boldsymbol{\Sigma}_{\mathrm{I},l,j}^{(s)2}\mathbf{A}\right), \end{aligned} \tag{13}$$

where $\boldsymbol{\Sigma}_{\mathrm{I},l,j}^{(s)} \in \mathbb{C}^{M\times M}$ is the diagonal matrix for the signal space and $\mathbf{A} = \mathbf{V}_{\mathrm{Z},l,j}^{(s)H}\mathbf{W}_{\mathrm{E},k,i}$. Since the $\mathbf{V}_{\mathrm{Z},l,j}^{(s)}$ is not correlate to $\mathbf{W}_{\mathrm{E},k,i}$, I_{E} is determined by γ_{I} regardless of the channel estimation error.

Then, the ICI power for ZF weight, I_{Z}, is expressed as

$$\begin{aligned} I_{\mathrm{Z}} &= \mathrm{trace}\left(\mathbf{W}_{\mathrm{Z},l,j}^H\mathbf{H}_{\mathrm{I},l,j}^H\mathbf{H}_{\mathrm{I},l,i}\mathbf{W}_{\mathrm{Z},l,j}\right) \\ &= \mathrm{trace}\left(\overline{\mathbf{V}}_{\mathrm{N},l,i}^{(s)H}\overline{\mathbf{V}}_{\mathrm{Z},l,i}^{(n)H}\mathbf{V}_{\mathrm{Z},l,i}^{(s)}\boldsymbol{\Sigma}_{\mathrm{I},l,i}^{(s)2}\mathbf{V}_{\mathrm{Z},l,i}^{(s)H}\overline{\mathbf{V}}_{\mathrm{Z},l,i}^{(n)}\overline{\mathbf{V}}_{\mathrm{N},l,i}^{(s)}\right) \\ &= \mathrm{trace}\left(\mathbf{B}^H\boldsymbol{\Sigma}_{\mathrm{I},l,i}^{(s)2}\mathbf{B}\right), \end{aligned} \tag{14}$$

where $\overline{\mathbf{V}}_{\mathrm{Z},k,j}^{(n)} \in \mathbb{C}^{N\times(N-M)}$ and $\overline{\mathbf{V}}_{\mathrm{N},k,j}^{(s)} \in \mathbb{C}^{(N-M)\times L_{K,j}}$ are the right singular vectors for null space and null space right singular vectors for signal space calculated by using the channel matrices with channel estimation error. $\mathbf{B} = \mathbf{V}_{\mathrm{Z},l,i}^{(s)H}\overline{\mathbf{V}}_{\mathrm{Z},l,i}^{(n)H}\overline{\mathbf{V}}_{\mathrm{N},l,i}^{(s)}$. In the case without the channel estimation error, $I_{\mathrm{Z}} = 0$ since $\mathbf{V}_{\mathrm{Z},k,j}^{(s)H}\overline{\mathbf{V}}_{\mathrm{Z},k,j}^{(n)} = \mathbf{0}$. Since the norm of $\mathbf{B}$ increases as the channel estimation error becomes large, I_{Z} is affected by γ_{I} and channel estimation error, σ_e^2. The ratio of channel estimation error, σ_e^2, to γ_{I} increases as γ_{I} decreases. Therefore, the norm of $\mathbf{B}$ and γ_{I} have inverse correlation. Figure 1 shows I_{Z} versus γ_{I} when the estimation error, σ_e^2, is set to be 0.5, 1.0, and 5.0 times the noise variance, σ^2, of one. We can see that the local maximal value and γ_{I} of the local maximal become large the channel estimation error increases. When $\gamma_{\mathrm{I}} \gg 0$ dB,

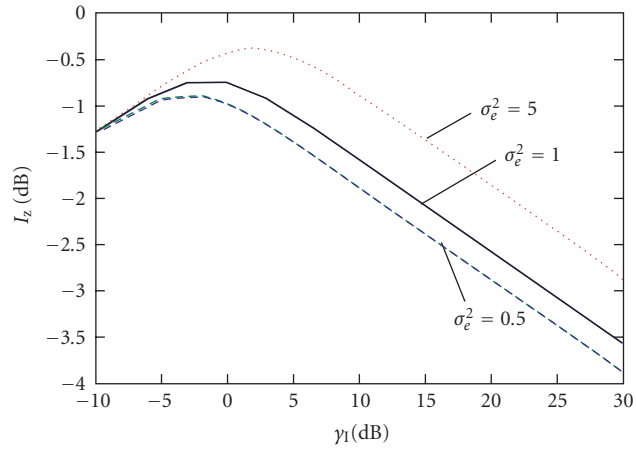


Figure 1: Intercell interference power versus γ_{I}.

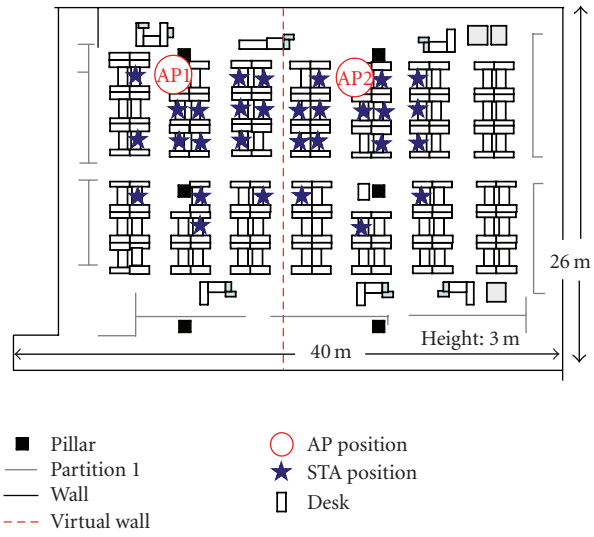


Figure 2: Experimental environment.

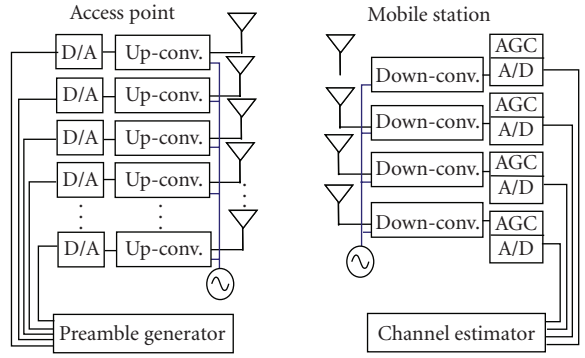


Figure 3: Block diagram of the MIMO testbed.

ZF works well since I_Z mitigation is proportional to the interference channel norm, γ_I. However, when γ_I is less than 5 dB, the transmission performance of spatial resource sharing based on ZF may be degraded. To estimate the realistic transmission performance, we measured the channel matrices, $\mathbf{H}$, and the interference channel matrices, $\mathbf{H}_I$, in an actual indoor environment and evaluated the effectiveness of space resource sharing by using the measured channel.

3. Channel Condition

3.1. Measured Channel Matrix. We measured the channel matrices in an actual indoor environment using the testbed for MIMO-OFDM transmission. The measurement environment is the $40 \times 26 \times 3$ [m] room shown in Figure 2. We used the channel matrices measured in an actual environment to evaluate the transmission performance of spatial resource sharing systems.

In the experiment, we used a circular array of eight antennas with element spacing of $1.0\,\lambda$ and a four-antenna linear array with element spacing of $0.5\,\lambda$ as the transmit antennas and the receive antennas, respectively. Therefore, we obtained 8×4 MIMO channel matrices. We consider 8×4 and 8×2 MIMO systems and 8×2 MIMO channel matrices are obtained using the center two receive antennas of the linear array. The AP was established on a 2.5 m pole and located at points AP 1, and AP 2 in Figure 2. The receive antenna array was attached to a laptop computer on a desk 0.7 m high that was placed at the star marks in Figure 2. The partitions are made of metal and partition height is 1.9 m. The each desk is separated by 1.4 m metal barriers.

Table 1: Parameters of MIMO-OFDM testbed.

Number of antennas	8 (Tx), 4 (Rx)
Radio frequency	4.85 GHz
Bandwidth	20 MHz
Total transmission power	10 dBm
Sensitivity	−20 to −70 dBm
Sampling rate	20 MHz (A/D), 80 MHz (D/A)
Number of FFT points	64
Number of subcarriers	48 (Information) + 4 (Pilot)
Short preamble length	0.8 μsec × 10 = 8 μsec
OFDM symbol length	3.2 × 2 μs + 1.6 μs (GI)

A block diagram of the experimental testbed is shown in Figure 3 and the main parameters for this testbed are given in Table 1. The center radio frequency is 4.85 GHz and the OFDM signal is transmitted in 20 MHz bandwidth. The AP transmits the preamble based on the short and long preambles in the IEEE 802.11a standard [10]. The STA estimates the channel matrices of the 48 sub-carriers from the received signals at 30 STA positions. We assume that the left 15 STAs and right 15 STAs communicate with AP 1 and AP 2, respectively. Thus, the measured channel matrices at AP 1 corresponding to the right 15 STAs and the left 15 STAs are $\mathbf{H}_{1,1}, \mathbf{H}_{1,2}, \ldots, \mathbf{H}_{1,15}$, and $\mathbf{H}_{\mathrm{I},1,1}, \mathbf{H}_{\mathrm{I},1,2}, \ldots, \mathbf{H}_{\mathrm{I},1,15}$, respectively. In AP 2, $\mathbf{H}_{2,1}, \mathbf{H}_{2,2}, \ldots, \mathbf{H}_{2,15}$ and $\mathbf{H}_{\mathrm{I},2,1}, \mathbf{H}_{\mathrm{I},2,2}, \ldots, \mathbf{H}_{\mathrm{I},2,15}$, are obtained between AP 2 and the right STAs and left STAs, respectively.

In this experiment, a virtual wall is considered between the AP 1's STAs and AP 2's STAs. Thus, we define the interference channel matrices as $\alpha\mathbf{H}_{\mathrm{I},2,1}$. α [dB] is the attenuation level of the virtual wall. When $\alpha = 0$ dB, we use measured channel matrix, $\mathbf{H}_{\mathrm{I},k,i}$ as the interference channel matrix. By using α, we can evaluate the performance of spatial resource sharing corresponding to the distance between two APs or the attenuation level of the wall. The averaged SNRs at STA corresponding to the two APs are shown in Figures 4(a) and 4(b). We assume that the measured channel matrices are accurate regardless of their SNR since the measured channel matrices are obtained by averaging a sufficient number of channel matrices.

4. Numerical Results

4.1. Achievable Transmission Bit Rate for i.i.d Channel. First, we consider that $\mathbf{H}_{k,i}$ is a complex Gaussian variable with zero mean (random i.i.d.). The variance of elements of $\mathbf{H}_{k,i}$, γ_S, is set to be 30 or 10 dB, and the variance of elements of

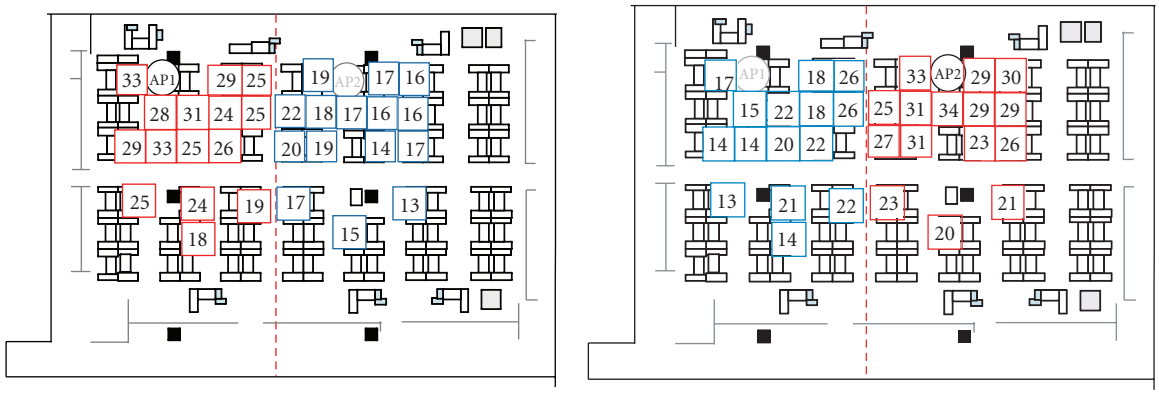


FIGURE 4: SNR distribution in an experimental environment corresponding to (a) AP 1 and (b) AP 2.

$\mathbf{H}_{\mathrm{I},k,i}$, γ_{I}, is determined from -5 to 30 dB. Since we assumed a noise variance of one, γ_{S} and γ_{I} express the expectation of the received signal to noise ratio (SNR) and interference to noise ratio (INR). Hereinafter, we take γ_{s} and γ_{I} as the SNR and INR, respectively. The achievable transmission bit rate corresponding to γ_{I} is calculated as the median value of 1,000 trials.

Figure 5(a) plots the median value of the achievable transmission bit rate for 8×4 MIMO systems versus the INR when the SNR is set to be 30 dB. The achievable transmission bit rate for the eigenvector transmission in the time resource sharing scenario and two spatial resource sharing scenarios, PSZ and CZ methods, are shown as C_{S}, C_{PSZ}, and C_{CZ}, respectively. The solid, dashed, and dotted lines represent the achievable transmission bit rate without channel estimation error and with channel estimation error, $\sigma_e^2 = 1$ and 5, respectively. It is found that the spatial resource sharing systems have much higher achievable transmission bit rate than that in time resource sharing systems. When the INR is 30 dB, C_{CZ} is 1.72 times greater than C_S. When the INR is less than 2.8 dB, C_{PSZ} without channel estimation error outperforms C_{CZ} without channel estimation error. When there are channel estimation error, the achievable transmission bit rates in spatial resource sharing degrade while the degradation by the channel estimation error in C_{S} is negligible. We can see that the intersection of C_{PSZ} and C_{CZ} becomes larger as channel estimation error increases. The range where C_{PSZ} outperforms C_{CZ} becomes large with large channel estimation error. When there is no channel estimation error, the achievable transmission bit rate of the primary AP in PSZ equals that in C_{S}. Therefore, the difference between C_{PSZ} and C_{S} denotes the achievable transmission bit rate of the secondary AP in PSZ. Since the interference from the primary AP decreases as the INR becomes small, the achievable transmission bit rate in PSZ increases.

Figure 5(b) plots the median value of the achievable transmission bit rate when SNR is 10 dB. In Figure 5(b), C_{CZ} is 1.42 times greater than C_{S} for the INR are 30 dB. The intersection of C_{PSZ} and C_{CZ} with channel estimation error decreases compared to that in Figure 5(a). We can see that the local minimal value of C_{CZ} with channel estimation error in Figures 5(a) and 5(b) corresponds to the local maximal of I_z in Figure 1 and that the degradation is large when the SNR is small.

Figures 6(a) and 6(b) plot that the median values of the achievable transmission bit rate for 8×2 MIMO systems for SNR values of 30 dB and 10 dB, respectively. Figures 6(a) and 6(b) shows that C_{CZ} is 1.91 times and 1.80 times greater than C_{S} when the INR is 30 dB and 10 dB, respectively. Spatial resource sharing is more effective in 8×2 MIMO systems than in 8×4 MIMO systems. This is because the fewer receive antennas of the STA corresponding to the interference channel matrix increases the dimension of $\mathbf{V}_{\mathrm{I},k,j}^{(n)} \in \mathbb{C}^{N \times (N-M)}$ in (4). $\mathbf{V}_{\mathrm{I},k,j}^{(n)}$ with large dimension values prevents the null space eigenvalues in (5) from decreasing compared to the eigenvalues in (3).

4.2. Achievable Transmission Bit Rate for Indoor Environment. The achievable transmission bit rate is calculated using 2 (AP positions) $\times$ 48 (subcarriers) $\times$ 10 (MS positions) measured channel matrices. Figures 7 and 8 show the cumulative probability of the achievable transmission bit rate in 8×4 and 8×2 MIMO systems. The solid, dashed, and dotted lines denote the achievable transmission bit rate with channel estimation error, $\sigma_e^2 = 0$, $\sigma_e^2 = 1$, and $\sigma_e^2 = 5$, respectively. In this section, we does not use the virtual wall attenuation, $\alpha = 0$ [dB]. In 8×4 MIMO systems without the channel estimation error, the median values of C_{CZ} and C_{SPZ} are 1.59 and 1.38 times greater than that of C_{S}, respectively. The median values of C_{CZ} and C_{SPZ} in 8×2 MIMO systems without channel estimation error are 1.90 and 1.51 times greater than that of C_{S}, respectively. The degradation in the achievable transmission bit rate due to the channel estimation error in 8×4 MIMO systems is larger than that in 8×2 MIMO systems. Figures 7 and 8 show that the median values of C_{CZ} and C_{PSZ} in 8×4 (8×2) MIMO systems with σ_e^2 of one and five are 1.41 (1.74) and 1.19 (1.51) times greater than that of C_{S}. The decrease in C_{CZ} due to the channel

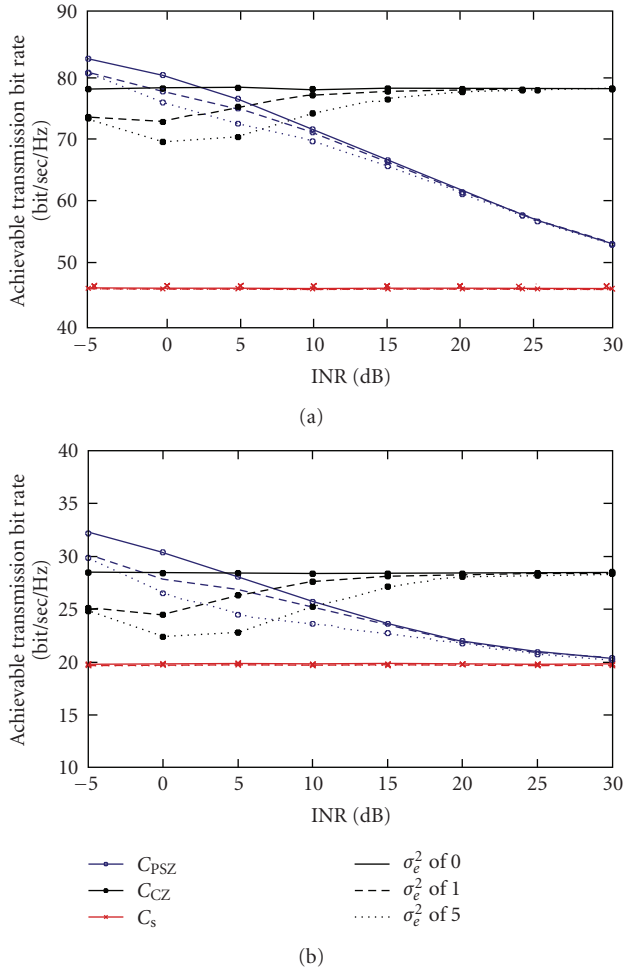


Figure 5: Median value of achievable transmission bit rate versus INR for 8 × 4 MIMO systems (a) when SNR is 30 dB and (b) when SNR is 10 dB.

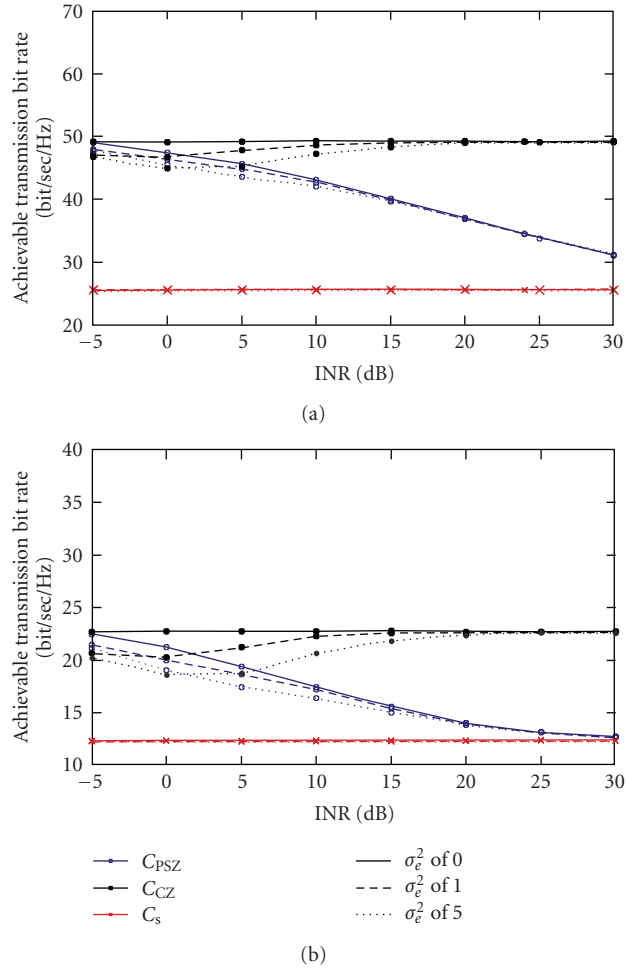


Figure 6: Median value of achievable transmission bit rate versus INR for 8 × 2 MIMO systems (a) when SNR is 30 dB and (b) when SNR is 10 dB.

estimation error is large compared to the results in Figures 5 and 6. It is found that the spatial resource sharing with large channel estimation error, σ_e^2 of one or five, high achievable transmission bit rates are obtained compared to C_S.

4.3. Virtual Wall Attenuation. Figures 9 and 10 show the median value of the achievable transmission bit rate versus the virtual wall attenuation factor, α, for 8×4 and 8×2 MIMO systems, respectively. The tendencies of Figures 9 and 10 are similar to C_{PSZ} and C_{CZ} of Figures 5 and 6. C_{PSZ} becomes large as the virtual wall attenuation factor decreases. In 8 × 4 MIMO systems without channel estimation error, C_{PSZ} outperforms C_{CZ} when α is less than −10 dB. As the channel estimation error increases, the intersection of C_{PSZ} and C_{CZ} increases. When there are large channel estimation error, PSZ method becomes effective compared to CZ method.

In 8 × 2 MIMO systems, CZ method is effective. When the estimation error, σ_e^2, is one, C_{CZ} is greater than C_{PSZ} when α is larger than −15 dB. Although the spatial resource sharing is affected by channel estimation error, it is found that CZ and PSZ method outperforms time resource sharing method even though there are channel estimation error whose variance, σ_e^2, is less than five.

Although Figures 5 and 6 show the local minimum points of C_{CZ} corresponding to the local maximum interference power (see Figure 1), C_{CZ} in Figures 9 and 10 has the local minimum points around $\alpha = 0$. This is because there are gaps between the expectations of the channel matrix norm and interference channel matrix norm as shown in Figures 4(a) and 4(b), and the eigenvalue distribution in an actual environment channel is different from that in random i.i.d. channel. It is found that the primary secondary AP is robust against CSI error compared with the cooperative AP scenario, and C_{PSZ} and C_{CZ} outperform C_{CZ} even when $\alpha = 0$.

5. Conclusion

We evaluated the transmission performance in spatial resource sharing. This paper introduced the primary-secondary AP scenario based on zero forcing (PSZ) and

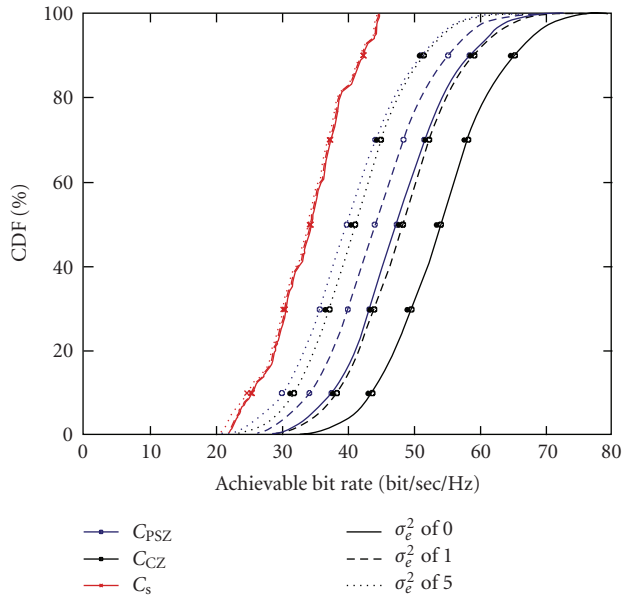


FIGURE 7: CDF of achievable bit rate using measured 8×4 channel matrix with $\alpha = 0$.

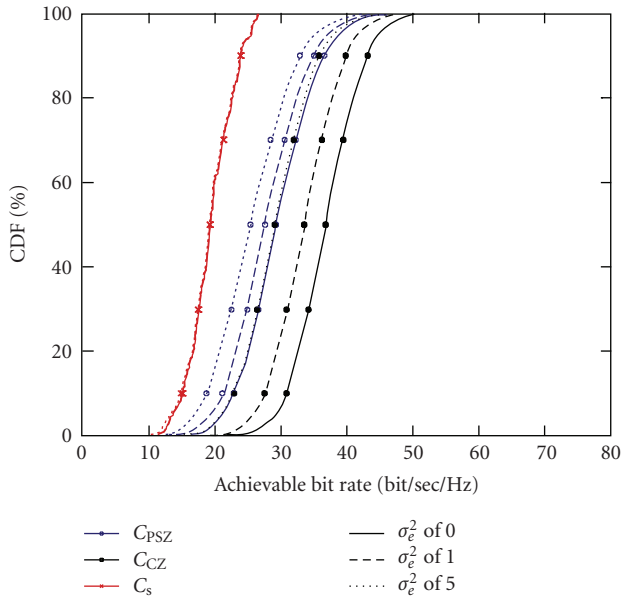


FIGURE 8: CDF of achievable bit rate using measured 8×2 channel matrix with $\alpha = 0$.

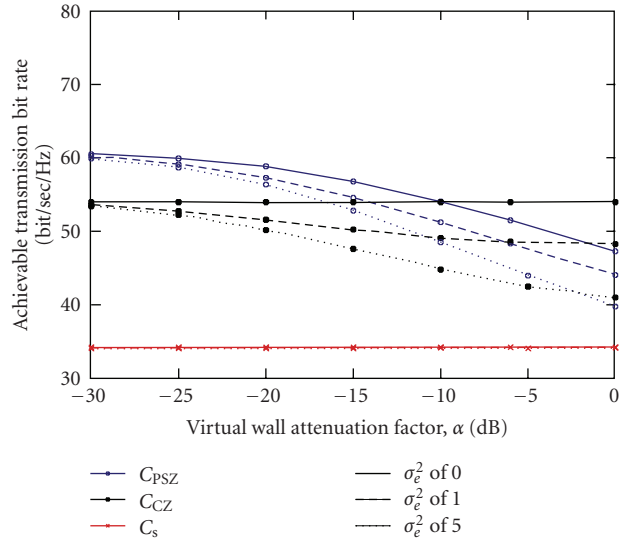


FIGURE 9: Median value of achievable bit rate using measured 8×4 channel matrices.

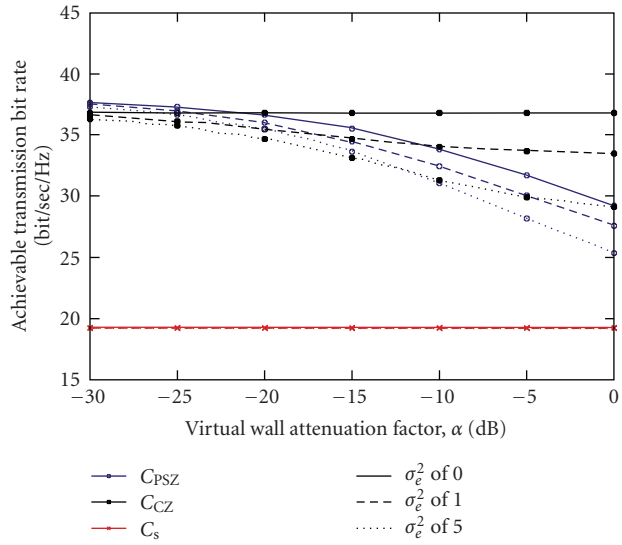


FIGURE 10: Median value of achievable bit rate using measured 8×2 channel matrix.

cooperative AP scenario based on zero forcing (CZ). Simulation results clarify the difference of applicability domain of these scenarios and the achievable transmission bit rate improvement compared to time domain resource sharing scenario. By using the channel matrices measured in an actual indoor environment, CZ and PSZ methods showed that they yielded 1.59 (1.90) times and 1.38 (1.51) times the achievable transmission bit rate of the time domain resource sharing system for 8×4 MIMO (8×2 MIMO) systems when there is no channel estimation error. Even though we considered large estimation error, σ_e^2 of one, the median values of C_{CZ} and C_{PSZ} in 8×4 MIMO (8×2 MIMO) systems with σ_e^2 of one and five were 1.41 (1.74) and 1.19 (1.51) times greater than that of C_S. By using a virtual wall attenuation factor, it is found that CZ method is effective in the case where mobile stations (STAs) have small number of the receive antennas and the norm of interference channel matrix is large while PSZ method outperforms CZ method when the norm of interference channel matrix becomes small and CSI error increases. We confirmed that spatial resource sharing is very attractive for the overlapping cell scenario.

Acknowledgments

The authors thank Mr. Kazuyasu Okada, the Executive Manager of NTT Network Innovation Laboratories, for his constant encouragement. The authors would also like to thank Mr. Yoshinobu Makise and Mr. Masaaki Ida of NTT Advanced Technology Corporation for their support.

References

[1] E. Perahia and R. Stacey, *Next Generation Wireless LANs: Throughput, Robustness, and Reliability in 802.11n*, Cambridge University Press, Cambridge, UK, 2008.

[2] L. M. Correia, Ed., *Mobile Broadband Multimedia Networks Techniques, Models and Tools for 4G*, Elsevier, New York, NY, USA, 2006.

[3] *Information and Communications in Japan*, White Paper, Ministry of Internal Affairs and Communications, Tokyo, Japan, 2008.

[4] A. Adler and M. Salhov, "A carrier-crade wireless LAN network implementation," *IEEE Microwave Magazine*, vol. 9, no. 4, pp. 108–119, 2008.

[5] G. Breit, et al., "TGac Channel Model Addendum," Doc. IEEE802.11-09/0308r5, May 2009.

[6] Y. Takatori, "Importance of Overlapped BSS issue in 802.11ac," Doc. IEEE802.11-09/0630r1, May 2009.

[7] E. Ziouva and T. Antonakopoulos, "CSMA/CA performance under high traffic conditions: throughput and delay analysis," *Computer Communications*, vol. 25, no. 3, pp. 313–321, 2002.

[8] P. Karn, "MACA: a new channel access method for packet radio," in *Proceedings of the Computer Networking Conference*, pp. 134–140, September 1990.

[9] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey," *IEEE Personal Communications*, vol. 3, no. 3, pp. 10–31, 1996.

[10] H. Ekström, A. Furuskär, J. Karlsson, et al., "Technical solutions for the 3G long-term evolution," *IEEE Communications Magazine*, vol. 44, no. 3, pp. 38–45, 2006.

[11] E. Perahia, "IEEE 802.11n development:history, process, and technology," *IEEE Communications Magazine*, vol. 46, no. 7, pp. 48–55, 2008.

[12] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 60–67, 2004.

[13] K. Nishimori, R. Kudo, N. Honma, Y. Takatori, O. Atsushi, and K. Okada, "Experimental evaluation using 16×16 multiuser MIMO testbed in an actual indoor scenario," in *Proceedings of IEEE International Symposium on Antennas and Propagation (APS '08)*, pp. 1–4, July 2008.

[14] Y. J. Zhang and K. B. Letaief, "An efficient resource-allocation scheme for spatial multiuser access in MIMO/OFDM systems," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 107–116, 2005.

[15] K. B. Letaief and Y. J. Zhang, "Dynamic multiuser resource allocation and adaptation for wireless systems," *IEEE Wireless Communications*, vol. 13, no. 4, pp. 38–47, 2006.

[16] S. Bellofiore, C. A. Balanis, J. Foutz, and A. S. Spanias, "Smart-antenna systems for mobile communication networks—part 1: overview and antenna design," *IEEE Antennas and Propagation Magazine*, vol. 44, no. 3, pp. 145–154, 2002.

[17] H. Dahrouj and W. Yu, "Coordinated beamforming for the multi-cell multi-antenna wireless system," in *in Proceedings of the 42nd Annual Conference on Information Sciences and Systems (CISS '08)*, pp. 429–434, Princeton, NJ, USA, March 19–21, 2008.

[18] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, 2008.

[19] Y. Takatori, F. Fitzek, K. Tsunekawa, and R. Prasad, "Channel capacity of TDD-OFDM-MIMO for multiple access points in a wireless single-frequency-network," *Wireless Personal Communications*, vol. 35, no. 1-2, pp. 19–33, 2005.

[20] R. Kudo, Y. Takatori, K. Nishimori, A. Ohta, and S. Kubota, "New downlink beamforming method for cooperative multiple access point systems," *IEICE Transactions on Communications*, vol. E90-B, no. 9, pp. 2303–2311, 2007.

[21] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, Cambridge, UK, 2003.

[22] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

Hindawi Publishing Corporation
International Journal of Digital Multimedia Broadcasting
Volume 2010, Article ID 815704, 19 pages
doi:10.1155/2010/815704

Research Article

Robust MMSE Transceiver Designs for Downlink MIMO Systems with Multicell Cooperation

Jialing Li, I-Tai Lu, and Enoch Lu

Department of ECE, Polytechnic Institute of NYU, 6 Metrotech Center, Brooklyn, NY 11201, USA

Correspondence should be addressed to Jialing Li, jialing.li.phd2@gmail.com

Received 1 July 2009; Accepted 16 December 2009

Academic Editor: Hongxiang Li

The *robust-generalized iterative approach* (*Robust-GIA*), *robust-fast iterative approach* (*Robust-FIA*), and *robust-decoder covariance optimization approach* (*Robust-DCOA*) are proposed for designing MMSE transceivers of downlink multicell multiuser MIMO systems with per-cell and per-antenna power constraints and possibly imperfect channel state information. The *Robust-DCOA* is the most restrictive but is always optimum, the *Robust-GIA* is the most general, and the *Robust-FIA* is the most efficient. When the *Robust-DCOA* is applicable and the decoder covariance matrices are full rank, the three proposed approaches are equivalent and all provide the optimum solution. Numerical results show that the proposed robust approaches outperform their non-robust counterparts in various single-cell and multicell examples with different system configurations, channel correlations, power constraints, and cooperation scenarios. Moreover, performances of the robust approaches are insensitive to estimation errors of channel statistics (correlations and path loss). With cell-cooperation, cell edge interference problems can be remedied without reducing the number of data streams by using the proposed robust approaches.

1. Introduction

Joint transceiver designs with criteria such as minimum mean square error (MMSE), maximum sum capacity, and minimum bit error rate (BER), and so forth, for multiple-input-multiple-output (MIMO) systems, with both uplink and downlink configurations, have been studied intensively in recent literature (e.g., see [1, 2]). Discussed in this paper is the robust MMSE transceiver design with respect to channel estimation errors for downlink multicell multiuser MIMO systems.

Assuming perfect channel state information (CSI), joint MMSE transceiver design has been studied by many researchers. A closed form design subject to the total power constraint for a single-user MIMO system is derived in [3]. Unfortunately though, this closed form design cannot be extended either to the multiuser case or to the per-antenna power constraint. For multiuser uplink MIMO problems subject to per-user power constraint, numerical solutions are provided mainly by the *transmit covariance optimization approach* (*TCOA*) [4, 5] and iterative approaches such as in [4]. We have developed a *generalized iterative approach* (*GIA*) for the uplink to deal with arbitrary linear power constraints (including the more practical per-antenna power constraint) [6]. Recently, we have also extended the *TCOA* to deal with arbitrary linear power constraints and have shown that the *GIA* and the *TCOA* are equivalent and optimum when the source covariance matrices are all projection matrices multiplied by the same constant and the transmit covariance matrices are all full rank [7].

For the downlink configuration, iterative approaches such as in [8] and a dual uplink approach in [9, 10] are employed to provide numerical MMSE solutions for multiuser MIMO systems subject to the total power constraint. The extension to deal with the per-antenna and per-cell power constraints for the downlink scenario is achieved by an iterative approach using a second-order cone programming (SOCP) [11] and by our *GIA* for the downlink [12]. Recently, we have also developed the *decoder covariance optimization approach* (*DCOA*) [13] to deal with arbitrary linear power constraints (including the per-antenna and per-cell power constraints). Furthermore, we have shown that the *GIA* and the *DCOA* are equivalent and optimum when the source covariance matrices are all identity matrices multiplied by the same constant and the decoder covariance matrices are all full rank [13].

All of the above mentioned MMSE transceiver designs are based on perfect CSI. However, the CSI is usually estimated in practice and is therefore subject to CSI estimation errors and possibly quantized CSI feedback errors. Hence, in practice, joint transceiver design has to be based on imperfect CSI. One option is to ignore that the CSI is imperfect. This type of approach is herein called non-robust. Unfortunately, the system performances derived from the non-robust approaches depend strongly on the quality of the available CSI (performances get worse quickly if the CSI quality deteriorates). Moreover, an optimum design based on poor CSI could be worse than suboptimum designs using the same CSI. Therefore, a more appealing option is to model the CSI error and to incorporate the error model into the transceiver design. This type of approach is herein called robust. The robust approaches can better mitigate the degradation of system performances due to imperfect CSI than the non-robust approaches if the CSI error is modeled correctly. Two classes of imperfect CSI models are usually employed: the stochastic model for the CSI estimation errors and the deterministic model for the CSI feedback errors. If a stochastic model is used, a statistically robust design is usually performed to optimize some system performance functions. If a deterministic model is used, a minimax or maximin design aiming at optimizing the worst-case system performance is usually carried out.

To cope with CSI estimation errors, closed form solutions for the robust joint MMSE transceiver design subject to the total power constraint are developed for single-user MIMO systems in [14–16]. But, similar to the perfect CSI case, no closed form solution is found when the problem is extended to deal with either multiuser applications or the per-antenna power constraint. For multiuser uplink MIMO problems subject to the inequality per-user and sum power constraints, the *robust transmit covariance optimization approach* (*Robust-TCOA*) is developed for independent identically distributed (i.i.d.) MIMO channels with CSI estimation errors in [17]. For multiuser downlink MIMO problems, when the CSI errors are bounded, the worst-case design under arbitrary power constraints is made based on SOCP in [18]; when the CSI errors are statistical errors, the robust design under the total power constraint is solved numerically by a dual uplink approach in [18].

So far, no statistically robust approach has been shown optimum in the MMSE sense for the downlink MIMO systems (either single-cell or multicell) under the per-antenna power constraint. Proposed in this paper is the robust MMSE transceiver design with respect to CSI estimation errors for downlink multicell MIMO systems subject to arbitrary linear power constraints. Specifically, the per-antenna and per-cell power constraints are considered. The work is relevant to frequency division duplex (FDD) systems where channel estimation is done at each user equipment (UE) and then fed back to the base station(s) (denoted as evolved Node B or eNB) via a zero-delay and error free communication link. Note that CSI feedback errors are not considered in this paper. The work may possibly also be extended to time division duplex (TDD) systems where channel estimation is done at the eNBs.

We first extend the statistical model of imperfect CSI in [16] to take into account the path loss effects. This extension is very crucial for practical multicell systems because the variances of CSI estimation errors depend on the distances between the UEs and the eNBs. The CSI estimation error of a UE near the eNB is much smaller than the CSI estimation error of a UE that is far away from the eNB. With the extended imperfect CSI model in hand, we herein propose three robust approaches to deal with arbitrary linear equality power constraints. The first, the *robust-generalized iterative approach* (*Robust-GIA*), is an extension of the *GIA* [12, 13] to the imperfect CSI case. The second, the *robust-decoder covariance optimization approach (Robust-DCOA)*, is an extension of the *DCOA* [13]. The third, the *robust-fast iterative approach (Robust-FIA)* is completely new. Though the first two are both extensions, their complexities are still similar to those of their predecessors!

The *DCOA* requires that the numbers of data streams are not prespecified and that all the source covariance matrices are identity matrices multiplied by the same constant. The *Robust-DCOA* is even more restricted since not only it requires all of the conditions of the *DCOA* but also that the transmit correlation matrix for each user is an identity matrix. The statistics of the CSI estimation error also need to be the same for all users if the power constraints of the users are interdependent. The *GIA* and the *Robust-GIA*, on the other hand, do not require any of the above mentioned conditions. The *Robust-FIA* has the same requirements as the *DCOA,* but not the additional restrictions of the *Robust-DCOA*.

The relationships between the *Robust-GIA,* the *Robust-FIA,* and the *Robust-DCOA* are very interesting. The *Robust-GIA* is the most general and can provide tradeoff between diversity and multiplexing gains. The *Robust-FIA* is the most efficient. Even though the *Robust-DCOA* is the most restricted, it always gives the optimum solution when it is applicable. But whenever the *Robust-DCOA* is applicable and all the decoder covariance matrices are full rank, the solutions obtained by the three robust approaches are actually equivalent (i.e., the *Robust-FIA* and the *Robust-GIA* are also optimum)! Interestingly, the *Robust-GIA* and the *Robust-DCOA* actually become the *GIA* and the *DCOA*, respectively, when the CSI is perfectly known, thereby providing a unified framework to take care of both perfect and imperfect CSI cases! We also denote the special case of the *Robust-FIA* when the CSI is perfectly known as the *fast iterative approach* (*FIA*) for convenience.

MMSE transceiver designs using the proposed robust approaches are performed for various single-cell and multicell examples with different system configurations, power constraints, channel correlations, and cooperation scenarios. System performances in terms of MSE and BER of various numerical examples are compared. Computational efficiency for various approaches is studied. Sensitivity studies with respect to channel statistics (channel correlations and path loss, estimated independently from channel estimation) are also investigated. The numerical results show that the proposed robust approaches are indeed superior to the non-robust approaches. Moreover, accurate channel correlations

and path loss are not required in the robust approaches. With cell cooperation, the cell edge UEs perform as well as those UEs without inter-cell interferences.

Notations are as follows. All boldface letters indicate vectors (lower case) or matrices (upper case). $\mathbf{A}'$, $\mathbf{A}^*$, $\mathbf{A}^{-1}$, $\mathrm{tr}(\mathbf{A})$, $\langle\mathbf{A}\rangle$, and $\|\mathbf{A}\|$ stand for the transpose, Hermitian, inverse, trace, expectation, and Frobenius norm of $\mathbf{A}$, respectively. Matrix $\mathbf{I}_a$ signifies an identity matrix with rank a. diag[...] denotes the diagonal matrix with elements [...] on the main diagonal. $\mathbf{A} \geq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semidefinite. $\mathbf{A} * \mathbf{B}$ denotes the Schur product of $\mathbf{A}$ and $\mathbf{B}$ (elementwise product of $\mathbf{A}$ and $\mathbf{B}$). $\mathrm{CN}(\mu, \sigma^2)$ denotes a complex normal random variable with mean μ and variance σ^2.

2. Formulation

2.1. Downlink Multicell Multiuser MIMO Systems. Consider the downlink of a multicell multiuser MIMO system with C eNBs and K UEs. Denote the number of transmit antennas at the cth eNB by t_c and the total number of transmit antennas by t, that is, $t = \sum_{c=1}^{C} t_c$. Also denote the number of receive antennas at the ith UE by r_i and the number of data streams intended for the ith UE by m_i.

In this system, there may be multiple groups where each group jointly designs its precoders and decoders but does so independently of the other groups. In the with-cooperation scenario (there is full cooperation among all eNBs), system-wide design is performed and there is only one group. In the without-cooperation scenario (there is no cooperation among eNBs), the eNB and UEs in a cell are one group. Let D_a (D_b) and S_a (S_b) define one such group in the with-cooperation (without-cooperation) scenario, D_a (D_b) being the set of indices of all eNBs in the group and S_a (S_b) being the set of indices of all UEs in the group.

At the ith UE, let $\mathbf{y}_i$ and $\mathbf{a}_i$ denote the received signal and noise, respectively. At eNB c, let $\mathbf{s}_i$ and $\mathbf{F}_{ic}$ denote the data and the precoder for the ith UE, respectively. Also let $\mathbf{H}_{ic}$ denote the channel matrix from eNB c to the ith UE. In the with-cooperation scenario, the data $\mathbf{s}_i$ for the ith UE are jointly transmitted by *all* eNBs. Thus, the received signal at the ith UE is:

$$\mathbf{y}_i = \sum_{c\in D_a} \mathbf{H}_{ic} \sum_{j\in S_a} \mathbf{F}_{jc}\mathbf{s}_j + \mathbf{a}_i, \tag{1a}$$

where $D_a = \{1, \ldots, C\}$ and $S_a = \{1, \ldots, K\}$.

In the without-cooperation scenario, let the eNB serving the ith UE be denoted by the index c_i where $c_i \in \{1, 2, \ldots, C\}$. Thus, the data $\mathbf{s}_i$ for the ith UE are only transmitted by eNB c_i. Therefore, a system with C eNB's can be decoupled into C single-cell downlink groups. In the c_ith group, the received signal vector $\mathbf{y}_i$ at the ith UE becomes

$$\mathbf{y}_i = \sum_{c\in D_b} \mathbf{H}_{ic} \sum_{j\in S_b} \mathbf{F}_{jc}\mathbf{s}_j + \mathbf{b}_i + \mathbf{a}_i, \quad \mathbf{b}_i = \sum_{c_j \notin D_b} \sum_{j\notin S_b} \mathbf{H}_{ic_j}\mathbf{F}_{jc_j}\mathbf{s}_j, \tag{1b}$$

where $D_b = \{c_i\}$ consists of the index of eNB c_i and $S_b = \{j \mid c_j = c_i\}$ consists of the indexes of all UEs served by eNB c_i.

Since there are multiuser precodings at eNB c_i, the data transmitted by eNB c_i to the UEs other than the ith UE are not considered as interference to the ith UE. But, the data transmitted from the eNBs other than eNB c_i result in the inter-cell interference (denoted as $\mathbf{b}_i$) to the ith UE.

In order to unify (1a) and (1b), let D and S define a group in the system, D being the set of indices of all eNBs in the group and S being the set of indices of all UEs in the group. For the ith UE, $i \in S$, the received signal vector, $\mathbf{y}_i$, can thus be expressed as

$$\mathbf{y}_i = \sum_{c\in D} \mathbf{H}_{ic} \sum_{j\in S} \mathbf{F}_{jc}\mathbf{s}_j + \mathbf{n}_i = \mathbf{H}_i \sum_{j\in S} \mathbf{F}_j\mathbf{s}_j + \mathbf{n}_i. \tag{2}$$

When there is full eNB cooperation,

$$\mathbf{H}_i = [\mathbf{H}_{i1} \cdots \mathbf{H}_{iC}], \qquad \mathbf{F}_j = \left[\mathbf{F}'_{j1} \cdots \mathbf{F}'_{jC}\right]', \qquad \mathbf{n}_i = \mathbf{a}_i, \qquad D = D_a, \qquad S = S_a. \tag{2a}$$

When there is no eNB cooperation,

$$\mathbf{H}_i = \mathbf{H}_{ic_i}, \qquad \mathbf{F}_i = \mathbf{F}_{ic_i}, \qquad \mathbf{n}_i = \mathbf{b}_i + \mathbf{a}_i, \qquad D = D_b, \qquad S = S_b. \tag{2b}$$

In (2b), $\mathbf{H}_i$ denotes the channel matrix from the eNB c_i to the ith UE, $\mathbf{F}_i$ denotes the precoder for the ith UE at eNB c_i, and $\mathbf{n}_i$ is the interference plus noise vector at the ith UE. But, in (2a), $\mathbf{H}_i$ denotes the composite channel matrix from all eNBs to the ith UE, $\mathbf{F}_i$ denotes the composite precoder for the ith user at all eNBs, and $\mathbf{n}_i$ is the noise vector. Equation (2) is essentially the same as the formulation in [11].

2.2. Extended Imperfect CSI Model [16]. In order to account for path loss and spatial correlation, the channel $\mathbf{H}_{ic}$ from eNB c to the ith UE is modeled as

$$\mathbf{H}_{ic} = \alpha d_{ic}^{-\beta} \mathbf{R}_{R,i}^{1/2} \mathbf{H}_{W,ic} \mathbf{R}_{T,c}^{1/2}. \tag{3}$$

Here, α is a constant, d_{ic} denotes the distance between the ith UE and eNB c, and 2β is the path loss exponent. In (3), $\mathbf{R}_{R,i}$ and $\mathbf{R}_{T,c}$ are known, normalized (unit diagonal entries), and full rank. They represent receive and transmit correlation matrices, respectively. The entries of $\mathbf{H}_{W,ic}$ are i.i.d. $\mathrm{CN}(0, 1)$. Here, the subscript W represents spatially white.

In practice, the CSI is estimated, resulting in estimation error. Thus,

$$\mathbf{H}_{ic} = \hat{\mathbf{H}}_{ic} + \mathbf{E}_{ic}, \tag{4}$$

where $\hat{\mathbf{H}}_{ic}$ is the channel estimate and $\mathbf{E}_{ic}$ is the CSI estimation error. By using an orthogonal training method and MMSE channel estimation, $\hat{\mathbf{H}}_{ic}$ and $\mathbf{E}_{ic}$ have been shown in [16] to be independent and

$$\mathbf{E}_{ic} = \alpha d_{ic}^{-\beta} \mathbf{R}_{E,ic}^{1/2} \mathbf{E}_{W,ic} \mathbf{R}_{T,c}^{1/2}, \qquad \mathbf{R}_{E,ic} = \left[\mathbf{I}_{r_i} + \sigma_{E,ic}^2 \cdot \mathbf{R}_{R,i}^{-1}\right]^{-1}, \tag{4a}$$

$$\sigma_{E,ic}^2 = \alpha^{-2} d_{ic}^{2\beta}\, \mathrm{tr}\left(\mathbf{R}_{T,c}^{-1}\right) \cdot \frac{\sigma_{\mathrm{a}}^2}{T_c}, \tag{4b}$$

where the entries of $\mathbf{E}_{W,ic}$ are i.i.d. CN$(0, \sigma^2_{E,ic})$. Here, σ^2_{a} is the noise variance at each of the receive antennas and T_c is the total training power transmitted from eNB c. Note that there is no inter-cell interference when the channel is estimated. Also note that the estimated channels $\{\widehat{\mathbf{H}}_{ic}\}_{i\in S}$ are fed back by the UEs to a central processing unit in the with-cooperation scenario and to eNB c in the without-cooperation scenario via a zero-delay and error free communication link. As to be expected, $\widehat{\mathbf{H}}_i$, $i \in S$, is equal to $[\widehat{\mathbf{H}}_{i1} \cdots \widehat{\mathbf{H}}_{iC}]$ in the with-cooperation scenario and to $\widehat{\mathbf{H}}_{ic}$, $c \in D$, in the without-cooperation scenario. Also, when perfect CSI is available, $\widehat{\mathbf{H}}_{ic} = \mathbf{H}_{ic}$ and $\sigma^2_{E,ic} = 0$.

2.3. Joint MMSE Precoder and Decoder Design Formulation. For a given group and thus a given D and S, the following is the problem formulation. Define the mean square error (MSE) of the data streams intended for the ith UE, $i \in S$, as

$$\eta_i(\mathbf{F}_i, \mathbf{G}_i) = \mathrm{tr}\left\langle (\mathbf{G}_i\mathbf{y}_i - \mathbf{s}_i)(\mathbf{G}_i\mathbf{y}_i - \mathbf{s}_i)^* \right\rangle, \tag{5a}$$

where $\mathbf{G}_i$ is the decoder at the ith UE, $\mathbf{F}_i$ is the precoder in (2a) or (2b), and $\mathbf{y}_i$ and $\mathbf{s}_i$ are given in (2). Equation (5a), using the actual channel $\mathbf{H}_i$ and actual noise vector $\mathbf{n}_i$, is the metric for MSE evaluation for the perfect-CSI, non-robust, and robust approaches. However, in the robust MMSE transceiver design, the following conditional MSE

$$\begin{aligned} \eta_i\left(\mathbf{F}_i, \mathbf{G}_i \mid \widehat{\mathbf{H}}_i\right) &= \mathrm{tr}\left\langle (\mathbf{G}_i\mathbf{y}_i - \mathbf{s}_i)(\mathbf{G}_i\mathbf{y}_i - \mathbf{s}_i)^* \right\rangle, \\ \mathbf{y}_i = \widehat{\mathbf{H}}_i \sum_{j\in S} \mathbf{F}_j\mathbf{s}_j + \widetilde{\mathbf{n}}_i, &\quad \widetilde{\mathbf{n}}_i = \mathbf{n}_i + \mathbf{E}_{ic} \sum_{j\in S} \mathbf{F}_j\mathbf{s}_j \end{aligned} \tag{5b}$$

is used. Here, when perfect CSI is available, $\widehat{\mathbf{H}}_i$ and $\widetilde{\mathbf{n}}_i$ represent the actual channel and actual noise vector, that is, $\widehat{\mathbf{H}}_i = \mathbf{H}_i$ and $\widetilde{\mathbf{n}}_i = \mathbf{n}_i$, respectively; otherwise, $\widehat{\mathbf{H}}_i$ represents the channel estimate, that is, $\widehat{\mathbf{H}}_i \neq \mathbf{H}_i$, and $\widetilde{\mathbf{n}}_i$ represents the equivalent interference plus noise vector. We will jointly choose the decoders and precoders corresponding to all UEs in S and all eNBs in D to minimize the sum MSE η:

$$\{\mathbf{F}_i, \mathbf{G}_i\}_{\mathrm{MMSE}} = \underset{\{\mathbf{F}_i, \mathbf{G}_i \mid i\in S\}}{\mathrm{argmin}} \{\eta\}, \quad \eta = \sum_{j\in S} \eta_j\left(\mathbf{F}_j, \mathbf{G}_j \mid \widehat{\mathbf{H}}_j\right). \tag{6}$$

Define the positive definite source covariance matrix as $\mathbf{\Phi}_{si} = \langle \mathbf{s}_i\mathbf{s}_i^* \rangle$ for the ith UE. The eNB(s) are subject to either the per-antenna or per-cell power constraints. For the per-antenna power constraint, the dth antenna of the cth eNB, $c \in D$, has power

$$P_{ad} = \mathbf{e}_d' \left(\sum_{j\in S} \mathbf{F}_j \mathbf{\Phi}_{\mathbf{s}j} \mathbf{F}_j^* \right) \mathbf{e}_d. \tag{7a}$$

For the per-cell power constraint, the cth eNB, $c \in D$, has power

$$P_{bc} = \mathrm{tr}\left(\mathbf{Q}_c' \left(\sum_{j\in S} \mathbf{F}_{jc} \mathbf{\Phi}_{\mathbf{s}j} \mathbf{F}_{jc}^* \right) \mathbf{Q}_c \right). \tag{7b}$$

Here, $\mathbf{e}_d$ in (7a) are an $l \times 1$ unit vector with the dth entry equal to one and all other entries equal to zero, $\mathbf{Q}_c$ in (7b) is an $l \times l$ matrix whose entries are all equal to zero except for the diagonal elements corresponding to the antennas of the cth eNB which are equal to one, and $l = \sum_{c\in D} t_c$.

2.4. Augmented Cost Function. To solve (6) subject to (7a) or (7b), one can use the *method of Lagrange multipliers* to set up the augmented cost function:

$$\xi = \eta + \mathrm{tr}\left(\mathbf{\Lambda} \left(\sum_{j\in S} \mathbf{F}_j \mathbf{\Phi}_{\mathbf{s}j} \mathbf{F}_j^* - \mathbf{P} \right) \right), \tag{8}$$

where $\mathbf{\Lambda}$ is an unknown diagonal matrix, representing the Lagrange multipliers. For the per-antenna power constraint in (7a),

$$\mathbf{\Lambda} = \mathrm{diag}[\lambda_1, \ldots, \lambda_l], \qquad \mathbf{P} = \mathrm{diag}[P_{a1}, \ldots, P_{al}]. \tag{9a}$$

For the per-cell power constraint in (7b), define $\mathbf{\Lambda}_k = \mathbf{I}_{t_k}\lambda_k$ and $\mathbf{P}_k = \mathbf{I}_{t_k} P_{bk}/t_k$, k=1,...,C. Thus,

$$\begin{aligned} &\mathbf{\Lambda} = \mathbf{\Lambda}_c, \quad \mathbf{P} = \mathbf{P}_c, \quad c \in D, \text{ without cooperation;} \\ &\mathbf{\Lambda} = \mathrm{diag}[\mathbf{\Lambda}_1 \ldots \mathbf{\Lambda}_C], \quad \mathbf{P} = \mathrm{diag}[\mathbf{P}_1 \ldots \mathbf{P}_C], \text{ with cooperation.} \end{aligned} \tag{9b}$$

2.5. Robust Design. The central processing unit is assumed to have knowledge about the channel estimate(s), $\widehat{\mathbf{H}}_j$, $j \in S$, and the channel statistics. Define the interference-plus-noise covariance matrix at the ith UE as $\mathbf{\Phi}_{\mathbf{n}i} = \langle \mathbf{n}_i\mathbf{n}_i^* \rangle$. Thus,

$$\mathbf{\Phi}_{\mathbf{n}i} = \begin{cases} \mathbf{\Phi}_{\mathbf{a}i} = \langle \mathbf{a}_i\mathbf{a}_i^* \rangle \text{ with cooperation;} \\ \mathbf{\Phi}_{\mathbf{b}i} + \mathbf{\Phi}_{\mathbf{a}i} = \langle \mathbf{b}_i\mathbf{b}_i^* \rangle + \langle \mathbf{a}_i\mathbf{a}_i^* \rangle \text{ without cooperation.} \end{cases} \tag{10}$$

Also define the equivalent interference-plus-noise covariance matrix at the ith UE as $\widetilde{\mathbf{\Phi}}_{\mathbf{n}i} = \langle \widetilde{\mathbf{n}}_i\widetilde{\mathbf{n}}_i^* \rangle$. After some math manipulations, (5b) for the ith UE, $i \in S$, becomes

$$\begin{aligned} \eta_i\left(\mathbf{F}_i, \mathbf{G}_i \mid \widehat{\mathbf{H}}_i\right) = &-\mathbf{G}_i\widehat{\mathbf{H}}_i\mathbf{F}_i\mathbf{\Phi}_{\mathbf{s}i} - \mathbf{\Phi}_{\mathbf{s}i}\mathbf{F}_i^*\widehat{\mathbf{H}}_i^*\mathbf{G}_i^* \\ &+ \mathbf{\Phi}_{\mathbf{s}i} + \mathbf{G}_i \left[\widehat{\mathbf{H}}_i \left(\sum_{j\in S} \mathbf{F}_j \mathbf{\Phi}_{\mathbf{s}j} \mathbf{F}_j^* \right) \widehat{\mathbf{H}}_i^* + \widetilde{\mathbf{\Phi}}_{\mathbf{n}i} \right] \mathbf{G}_i^*, \end{aligned} \tag{11}$$

where $\widetilde{\mathbf{\Phi}}_{\mathbf{n}i} = \mathbf{\Xi}_i + \mathbf{\Phi}_{\mathbf{n}i}$ and

$$\mathbf{\Xi}_i = \begin{cases} \sum_{c\in D} \sigma_c^2 P_{bc} \cdot \mathbf{R}_{E,ic}, & \text{if } \mathbf{R}_{T,c} = \mathbf{I}_{t_c},\ c \in D, \\ \sum_{c\in D} \sigma_c^2 \sum_{j\in S} \mathrm{tr}\left(\mathbf{F}_{jc}\mathbf{\Phi}_{\mathbf{s}j}\mathbf{F}_{jc}^*\mathbf{R}_{T,c}\right)\mathbf{R}_{E,ic}, & \text{if } \mathbf{R}_{T,c} \neq \mathbf{I}_{t_c},\ c \in D, \end{cases} \tag{12}$$

$$\sigma_c^2 = \mathrm{tr}\left(\mathbf{R}_{T,c}^{-1}\right) \cdot \frac{\sigma^2_{\mathrm{a}}}{T_c}. \tag{13}$$

For a given set of precoders $\{\mathbf{F}_j\}_{j\in S}$, setting the gradient of ξ in (8) with respect to $\mathbf{G}_i$ equal to zero, we yield the MMSE decoder for the ith UE, $i \in S$:

$$\mathbf{G}_i = \mathbf{\Phi}_{\mathbf{s}i}\mathbf{F}_i^*\hat{\mathbf{H}}_i^*\mathbf{M}_i, \quad \mathbf{M}_i = \left(\hat{\mathbf{H}}_i\left(\sum_{j\in S}\mathbf{F}_j\mathbf{\Phi}_{\mathbf{s}j}\mathbf{F}_j^*\right)\hat{\mathbf{H}}_i^* + \tilde{\mathbf{\Phi}}_{\mathbf{n}i}\right)^{-1}. \tag{14}$$

Substituting (11) and (14) into (8), the augmented cost function ξ in (8) is reduced to

$$\xi_1 = \sum_{j\in S}\mathrm{tr}\left(-\mathbf{\Phi}_{\mathbf{s}j}\mathbf{F}_j^*\hat{\mathbf{H}}_j^*\mathbf{M}_j\hat{\mathbf{H}}_j\mathbf{F}_j\mathbf{\Phi}_{\mathbf{s}j} + \mathbf{\Phi}_{\mathbf{s}j}\right) + \mathrm{tr}\left(\mathbf{\Lambda}\left(\sum_{j\in S}\mathbf{F}_j\mathbf{\Phi}_{\mathbf{s}j}\mathbf{F}_j^* - \mathbf{P}\right)\right). \tag{15}$$

Note that the ξ_1 in (15) no longer depends on $\{\mathbf{G}_j\}_{j\in S}$ explicitly.

On the other hand, for a given set of decoders $\{\mathbf{G}_j\}_{j\in S}$ and Lagrange multipliers $\mathbf{\Lambda}$, setting the gradient of ξ in (8) with respect to $\mathbf{F}_i$ equal to zero, we have the MMSE precoder for the ith UE, $i \in S$:

$$\mathbf{F}_i = \mathbf{N}\hat{\mathbf{H}}_i^*\mathbf{G}_i^*, \quad \mathbf{N} = \left(\sum_{j\in S}\hat{\mathbf{H}}_j^*\mathbf{G}_j^*\mathbf{G}_j\hat{\mathbf{H}}_j + \mathbf{\Lambda} + \mathbf{\Theta}\right)^{-1}, \tag{16}$$

where

$$\mathbf{\Theta} = \begin{cases} 0, & \text{if } \mathbf{R}_{T,c} = \mathbf{I}_{t_c},\ c \in D, \\ \Upsilon\mathbf{R}_T, & \text{if } \mathbf{R}_{T,c} \neq \mathbf{I}_{t_c},\ c \in D, \end{cases}$$

$$\Upsilon = \begin{cases} \Upsilon_c, & c \in D \text{ without cooperation}, \\ \mathrm{diag}[\Upsilon_1, \Upsilon_2, \ldots, \Upsilon_C] \text{ with cooperation}, \end{cases}$$

$$\mathbf{R}_T = \begin{cases} \mathbf{R}_{T,c}, & c \in D \text{ without cooperation}, \\ \mathrm{diag}[\mathbf{R}_{T,1}, \mathbf{R}_{T,2}, \ldots, \mathbf{R}_{T,C}] \text{ with cooperation}. \end{cases} \tag{17}$$

Here, $\Upsilon_c = \sigma_c^2 \sum_{i\in S}\mathrm{tr}(\mathbf{G}_i^*\mathbf{G}_i\mathbf{R}_{E,ic})\mathbf{I}_{t_c}$ and $\sigma_c^2$is given in (13). Substituting (11) and (16) into (8), the augmented cost function ξ in (8) is reduced to

$$\xi_2 = \mathrm{tr}\left(-\left(\sum_{j\in S}\hat{\mathbf{H}}_j^*\mathbf{G}_j^*\mathbf{\Phi}_{\mathbf{s}j}\mathbf{G}_j\hat{\mathbf{H}}_j\right)\mathbf{N} - \mathbf{\Lambda}\mathbf{P}\right) + \sum_{j\in S}\mathrm{tr}\left(\mathbf{\Phi}_{\mathbf{s}j} + \mathbf{G}_j\overline{\mathbf{\Phi}}_{\mathbf{n}j}\mathbf{G}_j^*\right), \tag{18}$$

where

$$\overline{\mathbf{\Phi}}_{\mathbf{n}j} = \begin{cases} \tilde{\mathbf{\Phi}}_{\mathbf{n}j}, & \text{if } \mathbf{R}_{T,c} = \mathbf{I}_{t_c},\ c \in D, \\ \mathbf{\Phi}_{\mathbf{n}j}, & \text{if } \mathbf{R}_{T,c} \neq \mathbf{I}_{t_c},\ c \in D. \end{cases} \tag{19}$$

The ξ_2 in (18) no longer depends on $\{\mathbf{F}_j\}_{j\in S}$ explicitly.

3. Robust Iterative Approaches

3.1. Robust-Generalized Iterative Approach (Robust-GIA). By setting the gradients of (15) with respect to $\mathbf{F}_j$ equal to zero, left multiplying the resulting equation with $\mathbf{F}_j$, and summing up the resulting equation over j, we obtain

$$\sum_{j\in S}\left(\mathbf{F}_j\mathbf{\Phi}_{\mathbf{s}j}\mathbf{F}_j^*\right)\mathbf{\Lambda} = \mathbf{B}, \tag{20}$$

where

$$\mathbf{B} = \sum_{j\in S}\mathbf{F}_j\mathbf{\Phi}_{\mathbf{s}j}^2\mathbf{F}_j^*\hat{\mathbf{H}}_j^*\mathbf{M}_j\hat{\mathbf{H}}_j - \mathbf{\Omega} - \sum_{j\in S}\left(\mathbf{F}_j\mathbf{\Phi}_{\mathbf{s}j}\mathbf{F}_j^*\right)\sum_{k\in S}\left(\hat{\mathbf{H}}_k^*\mathbf{M}_k\hat{\mathbf{H}}_k\mathbf{F}_k\mathbf{\Phi}_{\mathbf{s}k}^2\mathbf{F}_k^*\hat{\mathbf{H}}_k^*\mathbf{M}_k\hat{\mathbf{H}}_k\right), \tag{21}$$

$$\mathbf{\Omega} = \begin{cases} 0, & \text{if } \mathbf{R}_{T,c} = \mathbf{I}_{t_c},\ c \in D, \\ \left(\sum_{j\in S}\mathbf{F}_j\mathbf{\Phi}_{\mathbf{s}j}\mathbf{F}_j^*\right)\mathbf{R}_T\sum_{j\in S}\Sigma_j, & \text{if } \mathbf{R}_{T,c} \neq \mathbf{I}_{t_c},\ c \in D, \end{cases} \tag{22}$$

$$\Sigma_j = \begin{cases} \Sigma_{jc}, & c \in D \text{ without cooperation}, \\ \mathrm{diag}\left[\Sigma_{j1}, \Sigma_{j2}, \ldots, \Sigma_{jC}\right] \text{ with cooperation} \end{cases} \tag{23}$$

with $\Sigma_{jc} = \sigma_c^2\,\mathrm{tr}(\mathbf{R}_{E,jc}\mathbf{M}_j\hat{\mathbf{H}}_j\mathbf{F}_j\mathbf{\Phi}_{\mathbf{s}i}^2\mathbf{F}_j^*\hat{\mathbf{H}}_j^*\mathbf{M}_j)\mathbf{I}_{t_c}$. Utilizing (9a) and (9b), we can obtain explicit expressions for the Lagrange multipliers $\mathbf{\Lambda}$ as follows. For the per-antenna power constraint in (7a), we have

$$\mathbf{\Lambda} = \mathbf{P}^{-1}(\mathbf{I}_l * \mathbf{B}), \tag{24a}$$

and for the per-cell power constraint in (7b), we have

$$\lambda_c = P_{bc}^{-1}\,\mathrm{tr}[\mathbf{Q}_c'(\mathbf{I}_l * \mathbf{B})\mathbf{Q}_c]. \tag{24b}$$

With the explicit expression for the Lagrange multipliers $\mathbf{\Lambda}$ in (24a) or (24b) in hand, a *Robust-GIA* can be developed using the MMSE decoder in (14) and MMSE precoder in (16). There are three steps in each iteration of the *Robust-GIA*.

Step 1. Given $\{\mathbf{F}_i\}_{i\in S}$, obtain $\{\mathbf{G}_i\}_{i\in S}$ by (14).

Step 2. Given $\{\mathbf{F}_i\}_{i\in S}$, obtain $\mathbf{\Lambda}$ using (24a) or (24b).

Step 3. Given $\{\mathbf{G}_i\}_{i\in S}$ and $\mathbf{\Lambda}$, obtain $\{\mathbf{F}_i\}_{i\in S}$ by (16).

Note that the *Robust-GIA* can allow tradeoff between diversity and multiplexing gains because it can deal with various sets of prespecified numbers of data streams intended for the UEs.

3.2. Robust-Fast Iterative Approach (Robust-FIA) When $\mathbf{\Phi}_{\mathbf{s}i} = \sigma^2\mathbf{I}_{m_i}$. The *Robust-FIA* can be developed based on the *Robust-GIA* when the source covariance matrices are all identity matrices multiplied by the same constant, that is, $\mathbf{\Phi}_{\mathbf{s}i} = \sigma^2\mathbf{I}_{m_i}$, $i \in S$. For convenience and without loss of generality, we assume $\sigma^2 = 1$. Define the transmit covariance matrices as

$$\mathbf{U}_i = \mathbf{F}_i\mathbf{F}_i^*, \tag{25}$$

and the decoder covariance matrices as

$$\mathbf{V}_i = \mathbf{G}_i^*\mathbf{G}_i. \tag{26}$$

Substituting (14) into (26) and using (25), we obtain

$$\mathbf{V}_i = \mathbf{M}_i\hat{\mathbf{H}}_i\mathbf{U}_i\hat{\mathbf{H}}_i^*\mathbf{M}_i, \quad \mathbf{M}_i = \left(\hat{\mathbf{H}}_i\left(\sum_{j\in S}\mathbf{U}_j\right)\hat{\mathbf{H}}_i^* + \tilde{\mathbf{\Phi}}_{\mathbf{n}i}\right)^{-1}. \tag{27}$$

Similarly, substituting (16) into (25) and using (26), we obtain

$$\mathbf{U}_i = \mathbf{N}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i\mathbf{N}, \quad \mathbf{N} = \left(\sum_{j\in S}\hat{\mathbf{H}}_j^*\mathbf{V}_j\hat{\mathbf{H}}_j + \mathbf{\Lambda} + \mathbf{\Theta}\right)^{-1}, \tag{28}$$

where $\mathbf{\Theta}$ is given in (17). Substituting (26) into Υ_c, we can express $\mathbf{\Theta}$ in (17) in terms of $\{\mathbf{V}_i\}_{i\in S}$ in (26). To remove the dependence of $\{\mathbf{V}_i\}_{i\in S}$ on $\{\mathbf{U}_i\}_{i\in S}$, substitute (28) into (27) to yield

$$\begin{aligned}\mathbf{V}_i &= \mathbf{M}_i\hat{\mathbf{H}}_i\mathbf{N}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i\mathbf{N}\hat{\mathbf{H}}_i^*\mathbf{M}_i,\\ \mathbf{M}_i &= \left(\hat{\mathbf{H}}_i\mathbf{N}\left(\sum_{j\in S}\hat{\mathbf{H}}_j^*\mathbf{V}_j\hat{\mathbf{H}}_j\right)\mathbf{N}\hat{\mathbf{H}}_i^* + \tilde{\mathbf{\Phi}}_{\mathbf{n}i}\right)^{-1},\end{aligned} \tag{29}$$

where $\mathbf{N}$ is given in (28). Similarly, using the fact that $\mathbf{\Phi}_{\mathbf{s}i} = \sigma^2\mathbf{I}_{m_i}$ and substituting (25) and (28) into (21)–(23), we can express the Lagrange multipliers $\mathbf{\Lambda}$ in (24a) or (24b) in terms of $\{\mathbf{V}_i\}_{i\in S}$.

With (24a), (24b), and (29) being available, the *Robust-FIA* can be readily developed. There are two steps in each iteration of the *Robust-FIA*.

Step 1. Given $\{\mathbf{V}_i\}_{i\in S}$, obtain $\mathbf{\Lambda}$ using (24a) or (24b).

Step 2. Given $\mathbf{\Lambda}$, use (29) to obtain $\{\mathbf{V}_i\}_{i\in S}$ for the next iteration.

Note that the number of data streams intended for the UE's $\{m_i\}_{i\in S}$, has to be equal to the ranks of $\{\mathbf{V}_i\}_{i\in S}$ that the *Robust-FIA* returns and thus cannot be prespecified when using the *Robust-FIA*. When the *Robust-FIA* converges, the decoders $\{\mathbf{G}_i\}_{i\in S}$ can be obtained by the decomposition defined in (26) and the precoders $\{\mathbf{F}_i\}_{i\in S}$ can then be obtained from (16). Note that the decomposition in (26) is not unique: $\mathbf{V}_i = \mathbf{G}_i^*\mathbf{G}_i = \mathbf{G}_i^*\mathbf{A}_i\mathbf{A}_i^*\mathbf{G}_i$ where $\mathbf{A}_i$ is an arbitrary unitary matrix. One can easily show that when $\mathbf{\Phi}_{\mathbf{s}i} = \sigma^2\mathbf{I}_{m_i}$, $i \in S$, if $(\mathbf{F}_i, \mathbf{G}_i)$ is a pair of joint MMSE precoder and decoder, so is $(\mathbf{F}_i\mathbf{A}_i, \mathbf{A}_i^*\mathbf{G}_i)$. Both $(\mathbf{F}_i,\mathbf{G}_i)$ and $(\mathbf{F}_i\mathbf{A}_i, \mathbf{A}_i^*\mathbf{G}_i)$ give the same MMSE η. However, different choices of $\{\mathbf{A}_i\}$ matrices may lead to different BER results.

Note that when the CSI is perfectly known, the *Robust-FIA* is reduced to the *FIA* by replacing $\{\hat{\mathbf{H}}_i\}_{i\in S}$ by $\{\mathbf{H}_i\}_{i\in S}$ in (24a), (24b), and (29), setting $\tilde{\mathbf{\Phi}}_{\mathbf{n}i} = 0$ in $\mathbf{M}_i$ in (29) and setting $\mathbf{\Theta} = 0$ in $\mathbf{N}$ in (28).

4. Robust-Decoder Covariance Optimization Approach (*Robust-DCOA*)

When the source covariance matrices are all identity matrices multiplied by the same constant, that is, $\mathbf{\Phi}_{\mathbf{s}i} = \sigma^2\mathbf{I}_{m_i}$, $i \in S$, and when the transmit correlation matrices are all identity matrices, that is, $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$, $c \in D$, a *robust-decoder covariance optimization approach* (*Robust-DCOA*) can be used for jointly designing the MMSE transceivers. For convenience, we assume $\sigma^2 = 1$. The augmented cost function in (18) becomes

$$\begin{aligned}\xi_2 = \operatorname{tr}&\left[\left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \mathbf{\Lambda}\right)^{-1}\mathbf{\Lambda}\right]\\ &+ \sum_{i\in S}\operatorname{tr}\left(\mathbf{V}_i\tilde{\mathbf{\Phi}}_{\mathbf{n}i}\right) - \operatorname{tr}[\mathbf{\Lambda}\mathbf{P}] + \sum_{i\in S}m_i - l.\end{aligned} \tag{30}$$

The robust MMSE transceiver design problem becomes

$$\begin{aligned}&\min_{\{\mathbf{V}_i\}_{i\in S}}\max_{\mathbf{\Lambda}}\xi_2,\\ &\text{subject to } \mathbf{V}_i \geq 0,\ \operatorname{rank}(\mathbf{V}_i) = m_i,\ i \in S,\\ &\qquad\mathbf{\Lambda} \geq 0.\end{aligned} \tag{31a}$$

The problem in (31a) is not convex because of the implicit rank constraints dealing with the numbers of data streams, that is, $\operatorname{rank}(\mathbf{V}_i) = m_i$. Allowing $\{m_i\}_{i\in S}$s to be unspecified and noting that l is a known constant, we obtain the rank-relaxed decoder covariance optimization problem:

$$\begin{aligned}&\min_{\mathbf{V}_i\geq 0, i\in S}\max_{\mathbf{\Lambda}\geq 0}\xi_{2,\text{rel}},\\ \xi_{2,\text{rel}} = \operatorname{tr}&\left[\left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \mathbf{\Lambda}\right)^{-1}\mathbf{\Lambda} - \mathbf{\Lambda}\mathbf{P}\right] + \sum_{i\in S}\operatorname{tr}\left(\mathbf{V}_i\tilde{\mathbf{\Phi}}_{\mathbf{n}i}\right).\end{aligned} \tag{31b}$$

The cost function $\xi_{2,\text{rel}}$ in (31b) is convex with respect to $\{\mathbf{V}_i\}_{i\in S}$ and concave with respect to $\mathbf{\Lambda}$. Define $\min_{\{\mathbf{V}_i\geq 0, i\in S\}}\max_{\{\mathbf{\Lambda}\geq 0\}}\xi_{2,\text{rel}}$ as the primal problem and $\max_{\{\mathbf{\Lambda}\geq 0\}}\min_{\{\mathbf{V}_i\geq 0, i\in S\}}\xi_{2,\text{rel}}$ as the dual problem. Since both the primal problem and the dual problem are convex and strictly feasible, strong duality holds; that is, the optimum values of $\{\mathbf{V}_i\}_{i\in S}$, $\mathbf{\Lambda}$, and $\xi_{2,\text{rel}}$ obtained from the primal problem are the same as those obtained from the dual problem.

4.1. Primal-Dual Algorithm. We propose a novel *primal-dual algorithm* to solve the rank-relaxed decoder covariance optimization problem in (31b). Denote the feasible set of values for $\{\mathbf{V}_i\}_{i\in S}$ as the primal domain and the feasible set of values for $\mathbf{\Lambda}$ as the dual domain. In short, the approach consists of iterating between a primal domain step (looking in the primal domain for the best $\{\mathbf{V}_i\}_{i\in S}$ for a given $\mathbf{\Lambda}$) and a dual domain step (looking in the dual domain for the best $\mathbf{\Lambda}$ for a given $\{\mathbf{V}_i\}_{i\in S}$). The iterative procedure stops when the $\xi_{2,\text{rel}}$'s corresponding to the primal domain step and the dual domain step converge to the same value and when $\{\mathbf{V}_i\}_{i\in S}$ converge and $\mathbf{\Lambda}$ converge. The two steps of the $(j+1)$th iteration are as follows.

Step 1 (Primal domain step). Given $\mathbf{\Lambda} = \mathbf{\Lambda}^{(j)}$, find the $\{\mathbf{V}_i\}_{i\in S}$ which solves (32). Denote them as $\{\mathbf{V}_i^{(j+1)}\}_{i\in S}$:

$$\min_{\{\mathbf{V}_i\}_{i\in S}} \operatorname{tr}\left[\left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \mathbf{\Lambda}\right)^{-1}\mathbf{\Lambda}\right] + \sum_{i\in S}\operatorname{tr}\left(\mathbf{V}_i\tilde{\mathbf{\Phi}}_{\mathbf{n}i}\right), \quad \text{subject to } \mathbf{V}_i \geq 0,\ i \in S. \tag{32}$$

Step 2 (Dual domain step). Given $\{\mathbf{V}_i\}_{i\in S} = \{\mathbf{V}_i^{(j+1)}\}_{i\in S}$, find the $\mathbf{\Lambda}$ which solves (33). Denote it as $\mathbf{\Lambda}^{(j+1)}$:

$$\max_{\mathbf{\Lambda}} \operatorname{tr}\left[\left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \mathbf{\Lambda}\right)^{-1}\mathbf{\Lambda} - \mathbf{\Lambda}\mathbf{P}\right], \quad \text{subject to } \mathbf{\Lambda} \geq 0. \tag{33}$$

Both subproblems, defined in (32) and (33), are convex because their cost functions are convex and concave, respectively, and their constraints are all linear matrix inequalities. Therefore the resulting solution of each subproblem is optimum. Furthermore, the convexity of the rank-relaxed decoder covariance optimization problem guarantees that the resulting solution provided by the *primal-dual algorithm* is global optimum. Once the optimal $\mathbf{\Lambda}$ and $\{\mathbf{V}_i\}_{i\in S}$ are obtained, the optimum numbers of data streams $\{m_i\}_{i\in S}$, the optimum decoders $\{\mathbf{G}_i\}_{i\in S}$, and the corresponding optimum precoders $\{\mathbf{F}_i\}_{i\in S}$ are obtained in the same way as in the *Robust-FIA*. And, in all this, the power constraints have been accounted for by the Lagrange multipliers.

In practice, the *Robust-DCOA* given by solving (32) and (33) is considered to have converged at the $(j+1)$th iteration when $\{\|\mathbf{V}_i^{(j+1)} - \mathbf{V}_i^{(j)}\|\}_{i\in S}$, $\|\mathbf{\Lambda}^{(j+1)} - \mathbf{\Lambda}^{(j)}\|$, and the duality gap of the values of $\xi_{2,\text{rel}}$ derived from the two steps

$$\text{gap}^{(j+1)} = \xi_{2,\text{rel}}\left(\left\{\mathbf{V}_i^{(j+1)}\right\}_{i\in S}, \mathbf{\Lambda}^{(j+1)}\right) - \xi_{2,\text{rel}}\left(\left\{\mathbf{V}_i^{(j+1)}\right\}_{i\in S}, \mathbf{\Lambda}^{(j)}\right), \tag{34}$$

is less than some prespecified thresholds.

4.2. Semidefinite Programming (SDP) Procedure. Similar to the *TCOA* [5, 7] in uplink MIMO systems, (32) and (33) can be reformulated as SDP formulations, which can be solved numerically using existing codes such as SeDuMi [19] and Yalmip [20]. Equation (32) can be reformulated as SDP formulation:

$$\min_{\mathbf{W}_p, \{\mathbf{V}_i\}_{i\in S}} \operatorname{tr}\left[\mathbf{W}_p\mathbf{\Lambda}\right] + \sum_{i\in S}\operatorname{tr}\left(\mathbf{V}_i\tilde{\mathbf{\Phi}}_{\mathbf{n}i}\right), \quad \text{subject to } \mathbf{V}_i \geq 0,\ i \in S, \quad \begin{bmatrix} \mathbf{W}_p & \mathbf{I} \\ \mathbf{I} & \sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \mathbf{\Lambda} \end{bmatrix} \geq 0. \tag{35}$$

Since

$$\left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \mathbf{\Lambda}\right)^{-1}\mathbf{\Lambda} = \mathbf{I}_l - \left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \mathbf{\Lambda}\right)^{-1}\left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i\right), \tag{36}$$

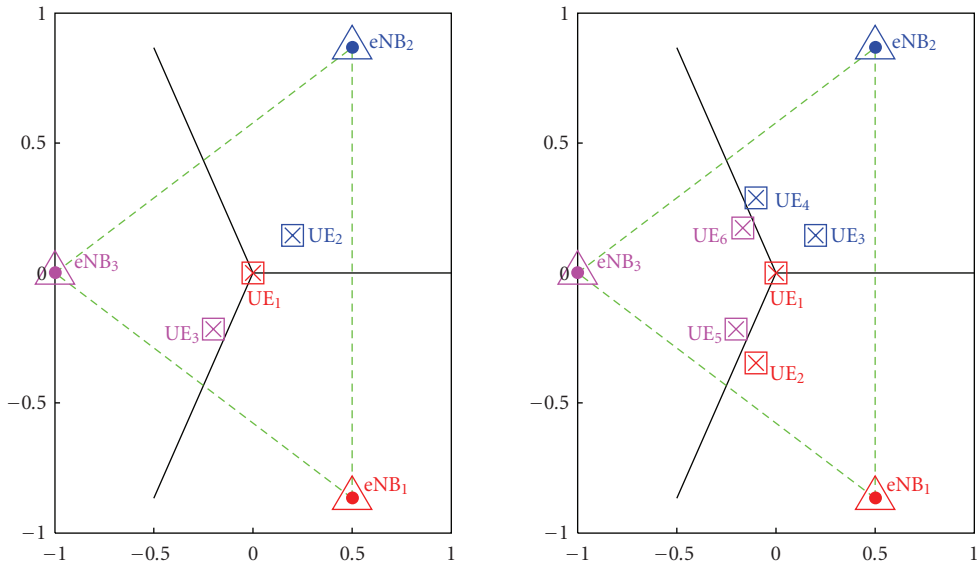
(33) can also be reformulated as SDP formulation:

$$\min_{\mathbf{W}_d, \mathbf{\Lambda}} \operatorname{tr}\left[\mathbf{W}_d\left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i\right)\right] + \operatorname{tr}(\mathbf{\Lambda}\mathbf{P}), \quad \text{subject to } \mathbf{\Lambda} \geq 0, \quad \begin{bmatrix} \mathbf{W}_d & \mathbf{I} \\ \mathbf{I} & \sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \mathbf{\Lambda} \end{bmatrix} \geq 0. \tag{37}$$

4.3. Numerically Efficient Procedure. We observe poor convergence behavior of the SDP procedure for the *Robust-DCOA* at high SNRs due to numerical errors introduced by SDP solvers. We therefore use the explicit closed form expression of $\mathbf{\Lambda}$ given in (24a) and (24b) in the *Robust-FIA* for the dual domain step in (33). The SDP procedure in (35) is still employed for the primal domain step in (32). This improves the convergence of the *Robust-DCOA* greatly.

4.4. Equivalence of the Robust-DCOA, the Robust-FIA, and the Robust-GIA. When the *Robust-DCOA* is applicable and the decoder covariance matrices $\{\mathbf{V}_i\}_{i\in S}$ are full rank, we claim that the *Robust-DCOA*, the *Robust-FIA,* and the *Robust-GIA* are equivalent. Thus, the solution of the *Robust-FIA* or the *Robust-GIA* is optimum under the above mentioned conditions because the solution given by the *Robust-DCOA* is always optimal (due to convexity).

Note that the *Robust-FIA* is equivalent to the *Robust-GIA* because the *Robust-FIA* is a special case of the *Robust-GIA* when $\mathbf{\Phi}_{\mathbf{s}i} = \sigma^2\mathbf{I}_{m_i}$, $i \in S$. To prove the equivalence between the *Robust-DCOA* and the *Robust-FIA*, it suffices to show that the KKT conditions of the two approaches are equivalent. The KKT conditions common to both approaches are (16) and the power constraint, (7a) or (7b).



(a) Example 3 (b) Example 4

FIGURE 1: System configurations of two multicell examples: (a) 3 eNBs and 3 UEs. (b) 3 eNBs and 6 UEs. The coordinates of relevant eNBs and UEs shown here are employed for simulations. Note that UE_1 is right on the edge of three cells in both systems. In example 3, UE_3 is near the cell edge between cells 1 and 3. In example 4, UE_2 and UE_2 are near the cell edge between cells 1 and 3, and UE_4 and UE_6 are near the cell edge between cells 2 and 3. *(Legends: Each UE is associated with the eNB in the same color.)*

For the *Robust-DCOA*, we set up the following augmented cost function from (31b) to include the nonnegative constraint on $\{\mathbf{V}_i\}_{i\in S}$:

$$\zeta = \mathrm{tr}\left[\left(\sum_{i\in S}\hat{\mathbf{H}}_i^*\mathbf{V}_i\hat{\mathbf{H}}_i + \boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda} - \boldsymbol{\Lambda}\mathbf{P}\right] + \sum_{i\in S}\mathrm{tr}\left(\mathbf{V}_i\tilde{\boldsymbol{\Phi}}_{\mathbf{n}i} - \boldsymbol{\Psi}_i\mathbf{V}_i\right), \tag{38}$$

where $\{\boldsymbol{\Psi}i\}_{i\in S}$ are the Lagrange Multipliers satisfying

$$\mathrm{tr}(\boldsymbol{\Psi}_i\mathbf{V}_i) = 0, \quad \boldsymbol{\Psi}_i \geq 0,\ i \in S. \tag{39}$$

Setting the gradient of (38) with respect to $\mathbf{V}_i$ equal to zero, we have

$$\hat{\mathbf{H}}_i\mathbf{N}\boldsymbol{\Lambda}\mathbf{N}\hat{\mathbf{H}}_i^* = \tilde{\boldsymbol{\Phi}}_{\mathbf{n}i} - \boldsymbol{\Psi}_i, \quad i \in S. \tag{40a}$$

When $\{\mathbf{V}_i\}_{i\in S}$ are full rank, the Lagrangian variable $\{\boldsymbol{\Psi}_i\}_{i\in S}$ is zero matrices and (40a) becomes

$$\hat{\mathbf{H}}_i\mathbf{N}\boldsymbol{\Lambda}\mathbf{N}\hat{\mathbf{H}}_i^* = \tilde{\boldsymbol{\Phi}}_{\mathbf{n}i}, \quad i \in S. \tag{40b}$$

The task of showing the equivalence of the KKT conditions of the two approaches which boils down to showing the above KKT condition of the *Robust-DCOA* (40b), can be derived from (and can be used to derived to) the KKT condition unique to the *Robust-FIA*, which are (14), (24a), and (24b). Substitute (16) and (28) into (14) to obtain

$$\mathbf{G}_i\hat{\mathbf{H}}_i\mathbf{N}\hat{\mathbf{H}}_i^*\left(\hat{\mathbf{H}}_i\mathbf{N}\left(\sum_{j\in S}\hat{\mathbf{H}}_j^*\mathbf{V}_j\hat{\mathbf{H}}_j\right)\mathbf{N}\hat{\mathbf{H}}_i^* + \tilde{\boldsymbol{\Phi}}_{\mathbf{n}i}\right)^{-1} = \mathbf{G}_i. \tag{41a}$$

Then left multiply (41a) by $\mathbf{V}_i^{-1}\mathbf{G}_i^*$ to get

$$\hat{\mathbf{H}}_i\mathbf{N}\hat{\mathbf{H}}_i^*\left(\hat{\mathbf{H}}_i\mathbf{N}\left(\sum_{j\in S}\hat{\mathbf{H}}_j^*\mathbf{V}_j\hat{\mathbf{H}}_j\right)\mathbf{N}\hat{\mathbf{H}}_i^* + \tilde{\boldsymbol{\Phi}}_{\mathbf{n}i}\right)^{-1} = \mathbf{I}_{r_i}. \tag{41b}$$

Summing up (41b) over $i \in S$ and using some matrix manipulations, we can show that the resulting equation and (40b) are equivalent. To get (24a) and (24b) from (40b), note that (24a) and (24b) can be obtained by using (20) and the power constraints. In turn, (20) can be obtained by substituting (14) and (26) into (16). Since (14) and (40b) can be derived from each other, this proof is complete. The above proof is done assuming $\boldsymbol{\Phi}_{\mathbf{s}i} = \mathbf{I}_{m_i}$, $i \in S$. It is also applicable when $\boldsymbol{\Phi}_{\mathbf{s}i} = \sigma^2\mathbf{I}_{m_i}$, $i \in S$, with $\sigma^2 \neq 1$.

5. Numerical Results

Without loss of generality, let $\alpha = 1$ and $\beta = 3.5$ (i.e., $2\beta = 7$) in the path loss model. Each cell is a hexagon with sides normalized to be 1 in length. The noise and source covariance matrices $\boldsymbol{\Phi}_{\mathbf{a}i}$ and $\boldsymbol{\Phi}_{\mathbf{s}i}$ are all identity matrices of dimension r_i and m_i, respectively. Let the per-antenna power constraint P_{ad} for antenna d in cell c ($d = 1, 2, \ldots\ t_c$) be equal to P (see (7a)) and let the per-cell power constraint P_{bc} for cell c, $c = 1, 2, \ldots, C$, be equal to t_cP (see (7b)). Therefore, the total transmission power from eNB c is $P_{bc} = t_cP$ under both per-antenna and per-cell power constraints.

Four examples (two single-cell and two 3-cell examples) will be considered. Their system parameters are shown in Table 1. The configurations of the two single-cell examples will be detailed later on while the configurations of the two 3-cell examples are shown here in Figure 1. In example 3

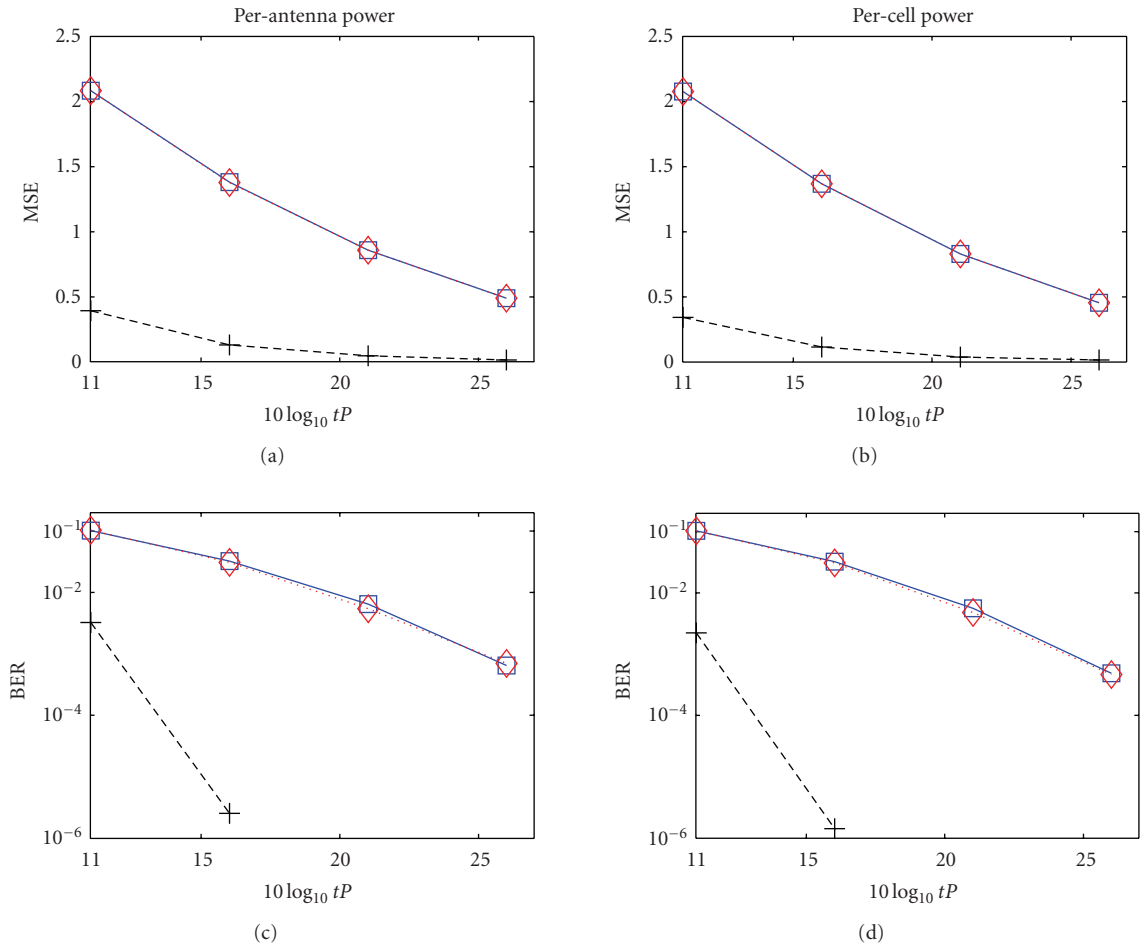


FIGURE 2: MSE and BER as functions of $10\log_{10}tP$ under per-antenna and per-cell power constraints. System parameters are shown in examples 1 and 2 in Table 1. (*Legends: Solid square lines: the DCOA with* $(m_1, m_2) = (2, 2)$*; Dotted diamond lines: the GIA* (*FIA*) *with* $(m_1, m_2) = (2, 2)$*; Dash "+" lines: the GIA with* $(m_1, m_2) = (1, 1)$.)

TABLE 1: System parameters.

Example	1	2	3	4
No. of cells (eNBs): C	1	1	3	3
No. of Tx Antennas per eNB: t_c, $c = 1, \ldots, C$	4	4	2	4
Total No. of Tx Antennas of the System: t	4	4	6	12
Number of UEs associated per eNB	2	2	1	2
Total No. of UEs of the System: K	2	2	3	6
No. of Rx Antennas per UE: r_j $(j = 1, \ldots, K)$	2	2	2	2
No. of data streams per UE: m_j $(j = 1, \ldots, K)$	2	1	2	2

(shown in Figure 1(a)), only one UE is associated with each eNB, and therefore, there are three UEs in total ($K = 3$). In example 4 (shown in Figure 1(b)), two UEs are associated with each eNB and there are 6 UEs in total ($K = 6$).

In the simulation, no CSI feedback error is assumed. The only CSI error is the CSI estimation error. $\sigma_a^2 = 1$ and T_c, the total transmission power of the cth eNBs training signal used for channel estimation, is the same as the total transmission power of the data signal, t_cP. Three types of designs (perfect-CSI, robust, and non-robust) will be performed. Take the family of *generalized iterative approaches* as an example. For the perfect-CSI design (denoted as the *GIA*), there is no CSI estimation error and the perfect CSI is employed for the joint MMSE design of precoders and decoders. On the other hand, there are CSI estimation errors for the non-robust design, *Non-robust-GIA*, and the robust design, *Robust-GIA*; only an estimated CSI is available to them. The difference between the non-robust and robust designs is simple; the non-robust design is unaware that the CSI it has is estimated and thus treats it as if it were perfect while the robust design is aware and thus incorporates the statistics of the CSI estimation error and the CSI into its design.

5.1. Equivalence of the Various Proposed Approaches. Without loss of generality, we will numerically show the equivalence of the *Robust-GIA*, *Robust-FIA*, and *Robust-DCOA* when the CSI is perfect (recall that the *Robust-GIA*, *Robust-FIA*, and *Robust-DCOA* are actually the *GIA*, *FIA*, and *DCOA*, resp., when the CSI is perfect!). To this end, consider two single-cell examples: examples 1 and 2 of Table 1. Also, for convenience, consider $d_{ic} = 1$, $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$, and $\mathbf{R}_{R,i} = \mathbf{I}_{r_i}$ for $i = 1, 2$ and $c = 1$.

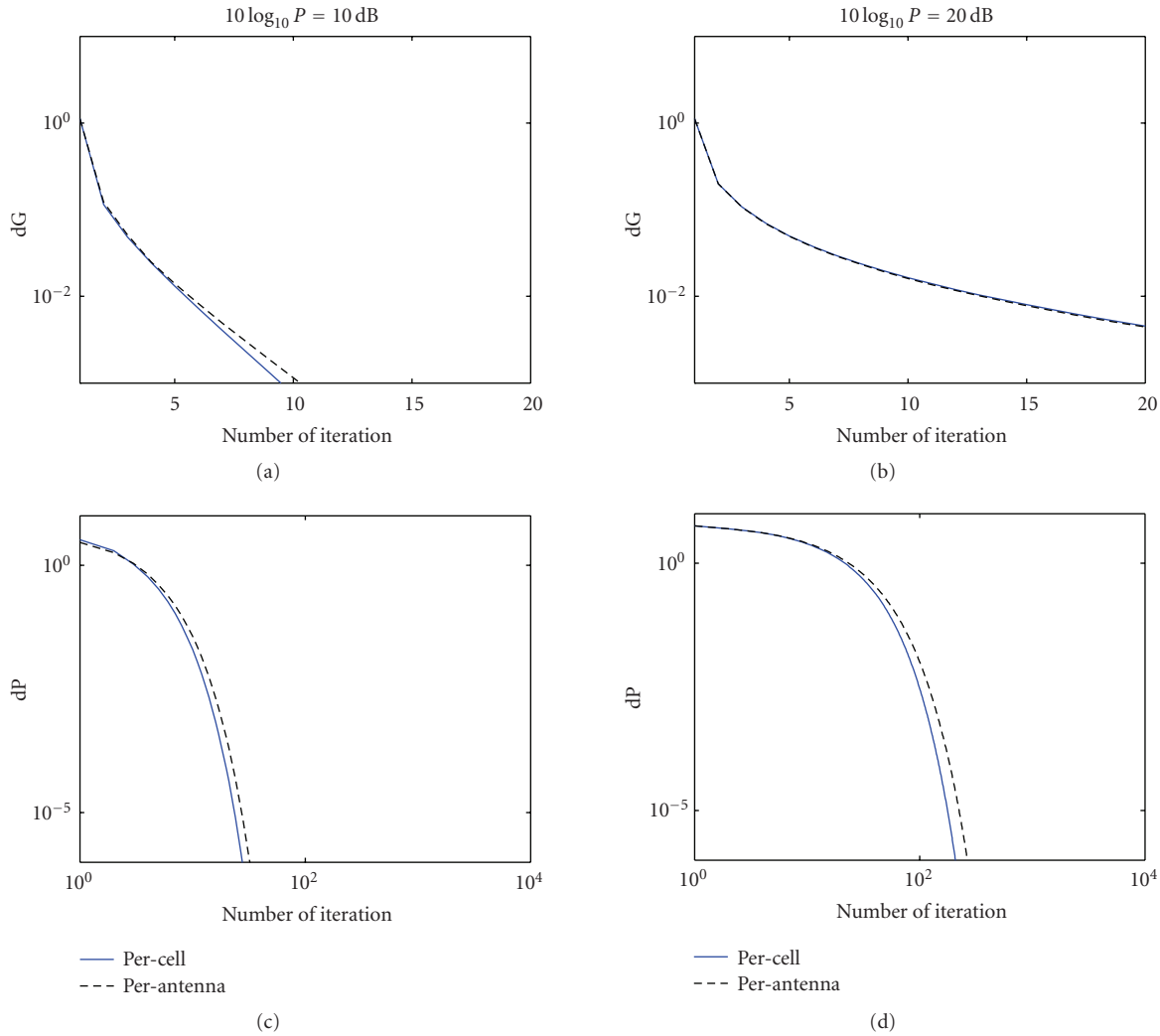


Figure 3: Convergence of the *GIA* (*FIA*): the differences dG and dP, defined in (42), as functions of the number of iterations under the per-antenna and per-cell power constraints. System parameters are given in example 1 of Table 1.

Figure 2 shows the MSE's and BER's as functions of the sum power of the system, that is, $10\log_{10} tP$ where $t = 4$. The results are obtained by averaging over 20 channel realizations. When two data streams are transmitted for both users (i.e., example 1 in Table 1 where $(m_1,m_2) = (2,2)$ and $\{\mathbf{V}_i\}$ are full rank), both the *GIA* (or the *FIA*) and the *DCOA* can be employed to find the globally optimum precoders and decoders. Comparing the two 2-data-stream curves in each subplot of Figure 2, we observe that the *GIA* (or the *FIA*) indeed has the same globally optimum performance as the *DCOA*. It is remarkable that the performances for the per-cell and per-antenna power constraints are similar to each other.

For the 1-data-stream scenario (i.e., example 2 in Table 1 where $(m_1, m_2) = (1,1)$ and $\{\mathbf{V}_i\}$ are not full rank), only the *GIA* can be employed because both the *DCOA* and the *FIA* result in $(m_1, m_2) = (2,2)$ and thus are not applicable here. Comparing the 1-data-stream curve against the 2-data-stream curves in each subplot of Figure 2, the MSE and BER performances, as predicted, are improved by transmitting fewer data streams than transmit antennas. But the increase in diversity gain is accompanied by a reduction in the multiplexing gain. For the 1-data-stream scenario, only the local optimality of the *GIA* can be guaranteed.

5.2. Computational Efficiency: The GIA (FIA) versus the DCOA. Without loss of generality, we will compare the computational efficiency of the various proposed approaches with perfect CSI. Consider example 1 in Table 1. The number of data streams is two for each of the two UEs so that the *DCOA* can be applicable. For convenience, we also choose $d_{ic} = 1$, $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$, $\mathbf{R}_{R,i} = \mathbf{I}_{r_i}$ for $i = 1,2$ and $c = 1$.

Note that the *GIA* and the *FIA* have the same convergence property because the *FIA* is a special case of the *GIA* when the source covariance matrices are all identity matrices multiplied by the same constant. The *FIA* is slightly more efficient than the *GIA* because it combines, into one step, two of the three steps in each iteration of the *GIA*. The convergence property (expressed as dG and dP) of the *GIA* (or the *FIA*) for both per-antenna and per-cell power

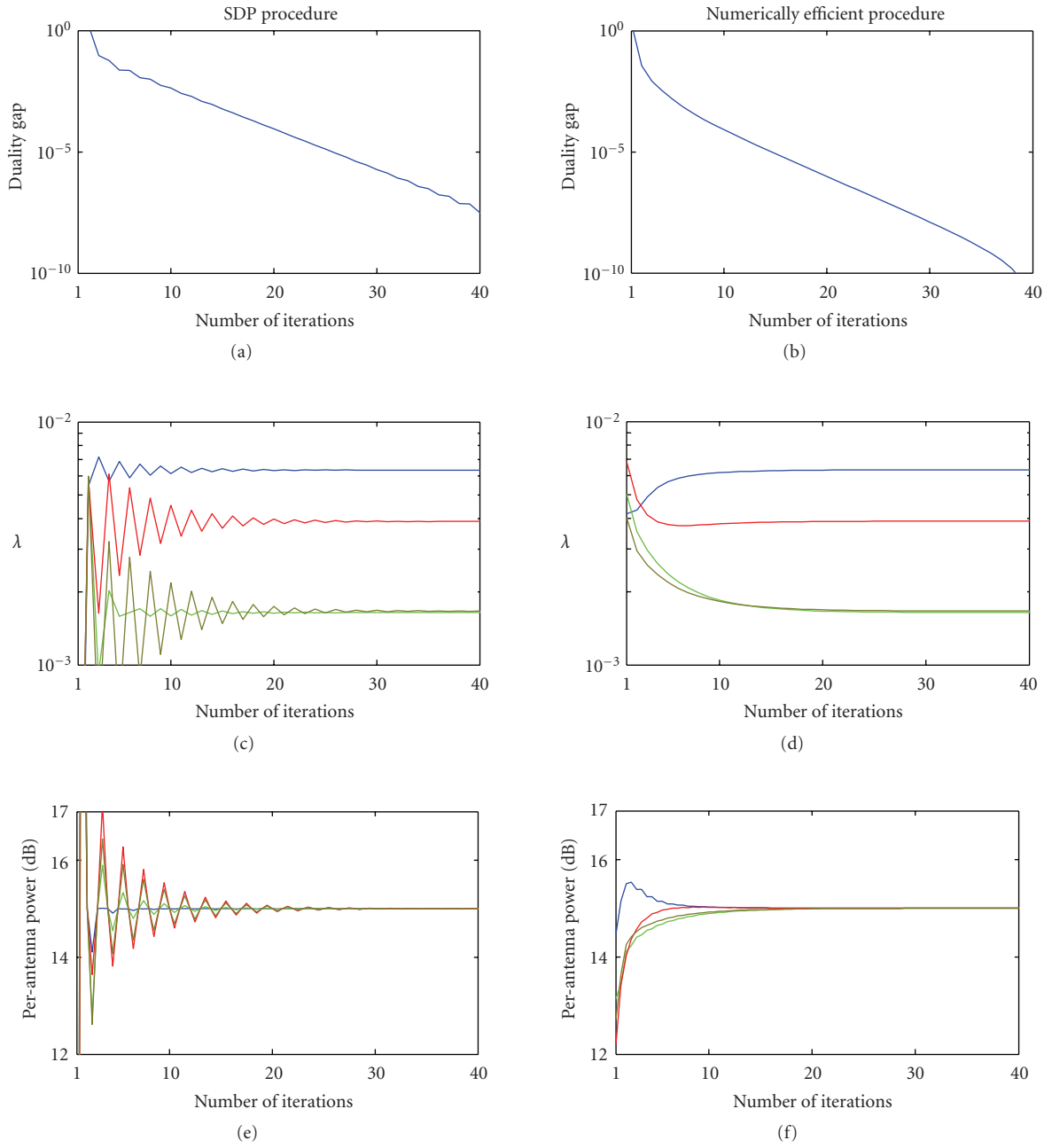


FIGURE 4: Convergence of the *DCOA*: duality gap, four Lagrange multipliers, and four per-antenna transmission power, as functions of the number of iterations under the per-antenna power constraint with $10\log_{10}P = 15\,\text{dB}$. (*Legend: the 4 different colors in the two middle subplots and two lower subplots correspond to the 4 transmit antennas.*)

constraints is shown in Figure 3. The difference in the decoders between the jth iteration and the $(j+1)$th iteration and the "distance" from the power constraints at the jth iteration are defined as, respectively,

$$\begin{aligned} \mathrm{dG}^{(j)} &= \sum_{i=1}^{2} \left\| \mathbf{G}_i^{(j+1)} - \mathbf{G}_i^{(j)} \right\|, \\ \mathrm{dP}^{(j)} &= \frac{1}{tP} \operatorname{tr}\left(\left| \sum_{i=1}^{2} \mathbf{F}_i^{(j)} \left(\mathbf{F}_i^{(j)}\right)^{*} - \mathbf{P} \right| \right). \end{aligned} \tag{42}$$

In Figure 3, the convergence rates for both power constraints are similar. It is remarkable that the *GIA* (or the *FIA*) converges much slower in higher power. This is due to the fact that, when P increases, the Lagrange multipliers λ_1 decrease quickly. For large P's, the Lagrange multipliers are very small. For example, when $10\log_{10}P = 30\,\text{dB}$, the Lagrange multipliers can be as small as 10^{-10}. Under such a situation, the equality power constraints in (7a) and (7b) are difficult to be met because the usage of (24a) or (24b) merely enforces the corresponding complementary slackness

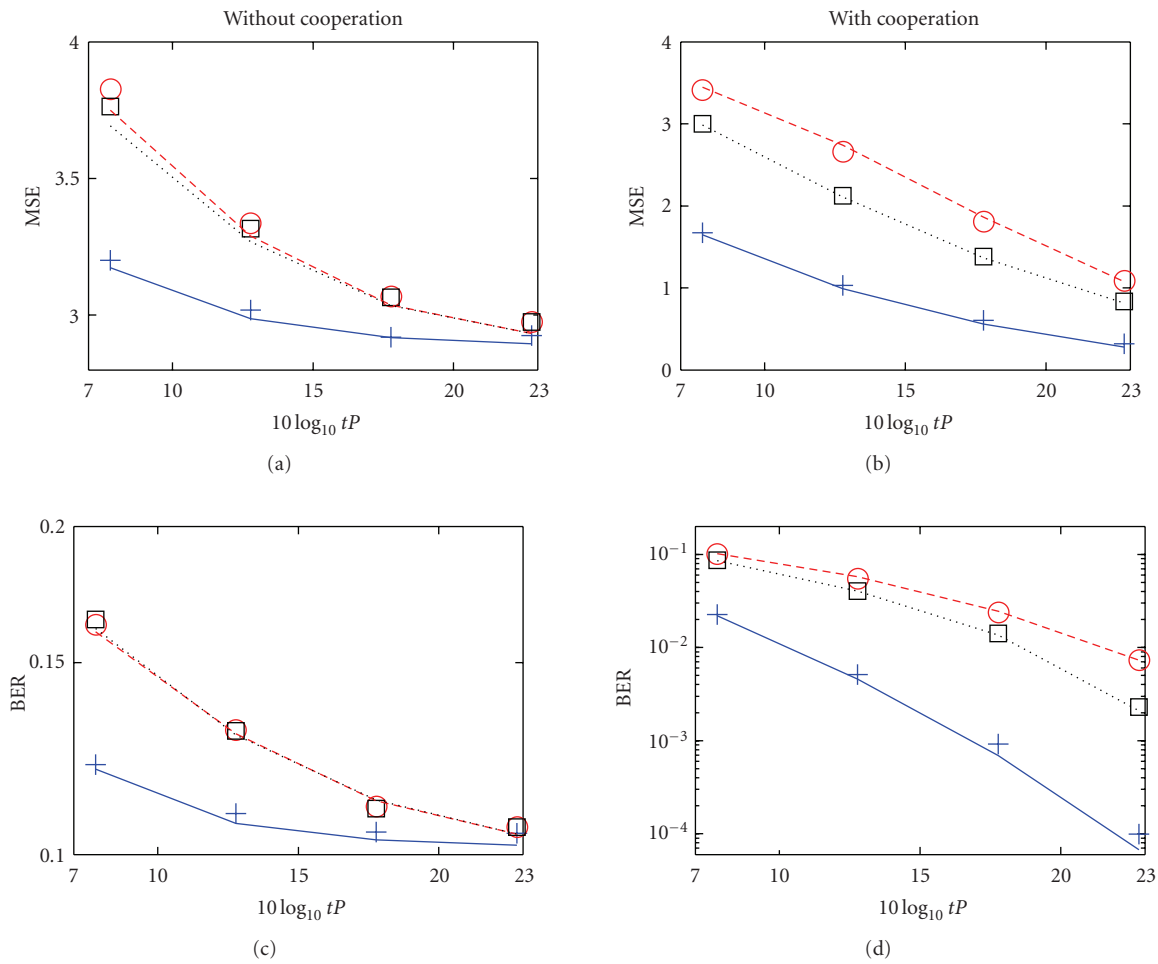


FIGURE 5: MSE and BER as functions of $10\log_{10} tP$ under per-antenna and per-cell power constraints. System parameters and configurations are shown in example 3 of Table 1 and Figure 1(a). (*Legends: the blue solid lines, red dashed lines, and black dotted lines represent, resp., the GIA, the Non-robust-GIA, and the Robust-GIA results under the per-cell power constraint. And the blue plus markers, red circle markers, and black square markers represent, resp., the GIA, the Non-robust-GIA, and the Robust-GIA results under the per-antenna power constraint.*)

conditions:

$$\mathbf{\Lambda}\left[\mathbf{I}_t * \left(\sum_{j=1}^{2} \mathbf{F}_j \mathbf{\Phi}_{\mathbf{s}j} \mathbf{F}_j^*\right) - \mathbf{P}\right] = 0, \tag{43a}$$

$$\lambda_1\left[\mathrm{tr}\left(\sum_{j=1}^{2} \mathbf{F}_j \mathbf{\Phi}_{\mathbf{s}j} \mathbf{F}_j^*\right) - P_{b1}\right] = 0. \tag{43b}$$

Thus, the number of iteration increases drastically as P increases if equality in the power constraints in (7a) or (7b) is insisted. However, if the equality constraints are relaxed and only inequality constraints (the per-antenna or per-cell transmission powers are allowed to be less than the corresponding power constraints) are required, the convergence rate at high power will be improved greatly.

Using the same single-cell example, the convergent properties of the SDP Procedure and the Numerically Efficient Procedure of the *DCOA* are shown in Figure 4. Here, $10\log_{10} P = 15\,\mathrm{dB}$ and, for convenience, only the per-antenna power constraint is considered. Observing the convergence rates of the duality gap in (34), the Lagrange multipliers in (9a), and the per-antenna transmission power from Figure 4, we conclude that the Numerical Efficient Procedure converges faster than the SDP procedure.

Comparing the *DCOA* with the *GIA* (or the *FIA*), the *GIA* (or the *FIA*) is numerically more efficient than the *DCOA*. This is because, for the *GIA* (or the *FIA*), closed form expressions are available for the precoders, decoders and Lagrange multipliers; but for the *DCOA*, a numerical optimization procedure has to be carried out to find the decoder covariance matrices in the primal step. Note that, just like the *GIA* (or the *FIA*), the number of iterations between the primal and dual steps of the *DCOA* increases drastically as P increases. This is because the convergence problem due to very small Lagrange multipliers at high power exists for both the *DCOA* and the *GIA* (or the *FIA*). In fact, the *DCOA* does not even converge at times due to the lack of numerical precision of the numerical solvers used. Thus, both the *DCOA* and the *GIA* (or the *FIA*) have difficulty in convergence at high power. Fortunately, within the practical power range, both the *DCOA* and the *GIA*

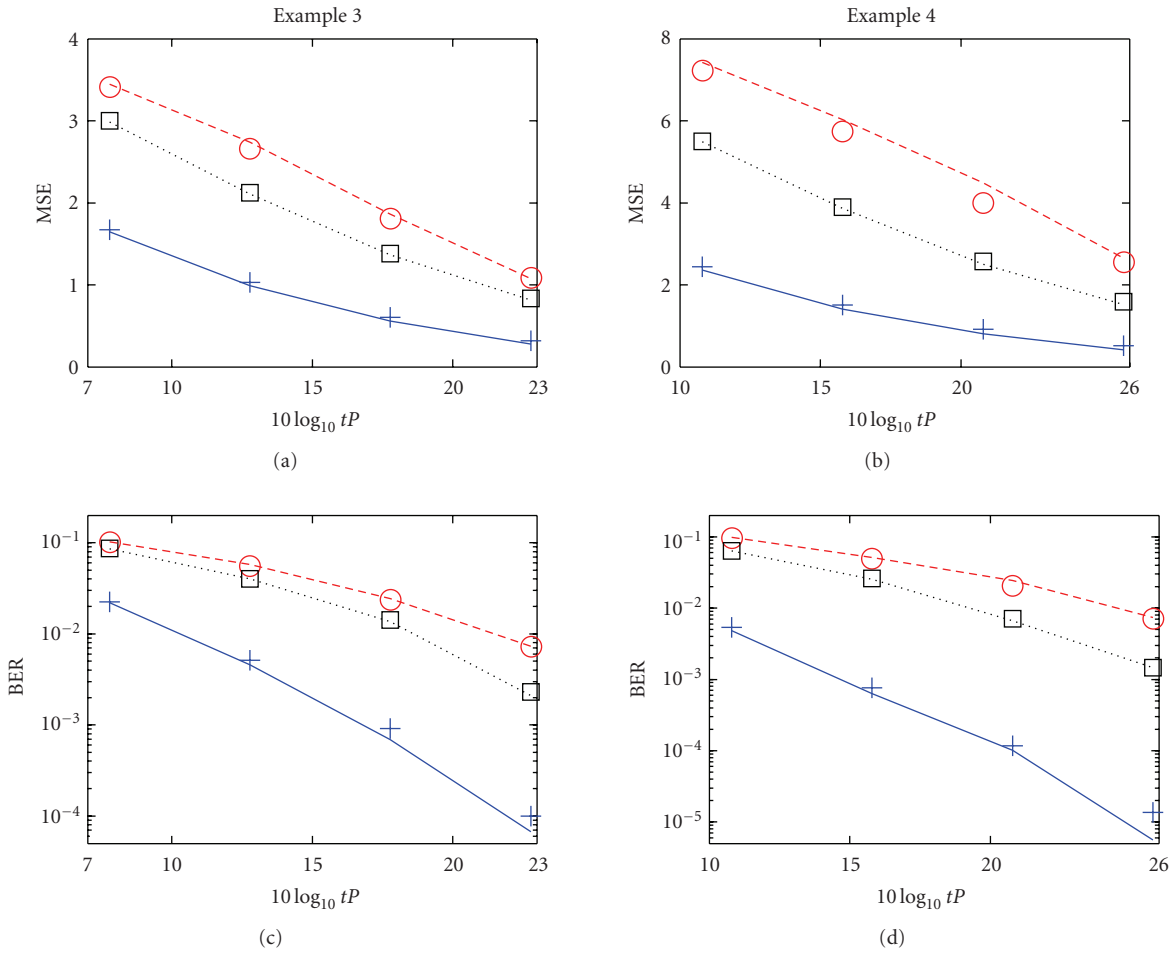


FIGURE 6: MSE and BER as functions of $10\log_{10} tP$ under per-antenna and per-cell power constraints for examples 3 and 4 in Table 1 and Figure 1. (*Legends are the same as in Figure 5.*)

(or the *FIA*) worked fine as long as some attention was paid to the selection of the initial values of the iteration process at high power. Note that the robust approaches have only a small increase in complexity compared to their perfect-CSI counterparts and the conclusion made here for complexity analysis is also applicable to the robust approaches.

In the following sections, we will consider the situation where $\{\mathbf{\Phi}_{\mathbf{s}i}\}$ are all identity matrices and $\{\mathbf{V}_i\}$ are full rank. Under such a situation, the *DCOA*, the *FIA*, and the *GIA* are equivalent. Moreover, the *Non-robust-DCOA*, the *Non-robust-FIA*, and the *Non-robust-GIA* are equivalent, and the *Robust-FIA*, and the *Robust-GIA* are also equivalent. If the *Robust-DCOA* is applicable, the *Robust-DCOA*, the *Robust-FIA*, and the *Robust-GIA* are equivalent. Thus, only the *GIA*, the *Robust-GIA*, and the *Non-robust-GIA* results are presented for convenience.

5.3. Multicell: With Cooperation versus without Cooperation. Using the 3-cell configuration in Figure 1(a) and the system parameters of example 3 in Table 1, two different cooperation scenarios will be simulated. In the first scenario, there is no cooperation among the eNBs. In the second scenario, there is full cooperation among the three eNBs. Note that the interference-plus-noise covariance matrix in (10) needs to be estimated in the without-cooperation scenario. With some derivations, we can show that (10) can be approximated as

$$\mathbf{\Phi}_{\mathbf{n}i} \approx \sum_{c \notin D_b} \alpha^2 d_{ic}^{-2\beta} P_{bc} \mathbf{R}_{R,i} + \mathbf{\Phi}_{\mathbf{a}i} \tag{44}$$

for the without-cooperation scenario. For convenience, we choose $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$ and $\mathbf{R}_{R,i} = \mathbf{I}_{r_i}$ for $i = 1, 2, 3$ and $c = 1, 2, 3$. Channel matrices are estimated and $\widehat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma^2_{E,ic} \neq 0$. Figure 5 shows the MSE and BER results derived with and without eNB cooperation. All the MSE and BER results are obtained by averaging over 30 channel realizations.

It is not surprising to see that the BER and the MSE of the without-cooperation scenario are much larger (worse) than the BER and the MSE of the with-cooperation scenario, respectively. Even with perfect CSI, the without-cooperation BER is larger than 10% even at high power. It is obvious that some kinds of time/frequency scheduling or code spreading are needed in order to reduce the cell edge interferences if no cooperation among eNBs is available. On the other hand, in the with-cooperation scenario, the BER of the *GIA* is below

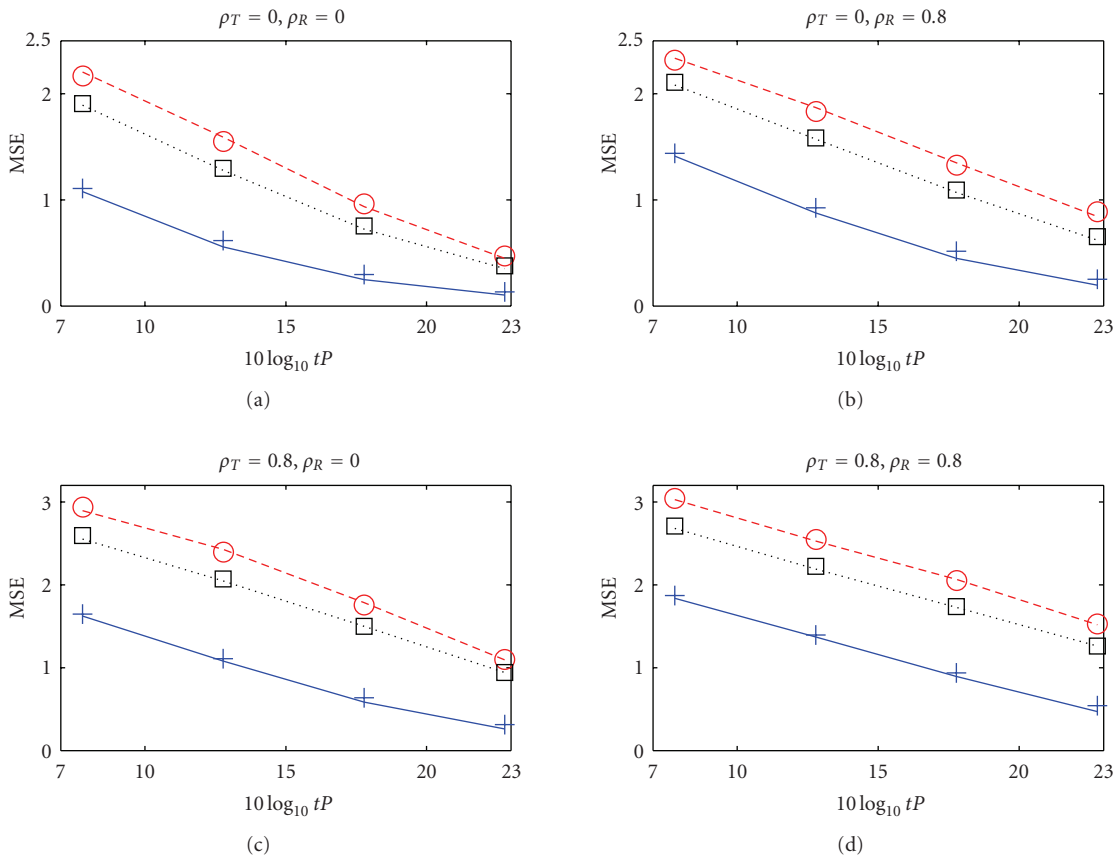


FIGURE 7: MSE's as functions of $10\log_{10}tP$ under per-antenna and per-cell power constraints for different values of ρ_T and ρ_R. System parameters are shown in example 1 of Table 1. (*Legends are the same as in Figure 5.*)

1% at the low transmission power ($10\log_{10}tP = 10\log_{10}6 + 5$ dB) when the perfect CSI is available. When the perfect CSI is not available, the *Robust-GIA* result is decent. Even the *Non-robust-GIA* result in the with-cooperation scenario is better than the *GIA* result in the without-cooperation scenario. The *Robust-GIA* result loses around 8 dB in SNR with respect to the *GIA* result and has around a 3 dB gain in SNR with respect to the *Non-robust-GIA*. Note that results obtained from both per-antenna and per-cell power constraints are similar.

5.4. Multicell: Example 3 versus Example 4. We now compare the results of example 3 with the results of example 4 in Table 1. The system configurations of examples 3 and 4 are shown in Figures 1(a) and 1(b), respectively. Note that there is one UE per cell in example 3 but there are two UEs per cell in example 4. For convenience, we choose $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$ and $\mathbf{R}_{R,i} = \mathbf{I}_{r_i}$ for $i = 1, 2, \ldots, K$ and $c = 1, 2, 3$. Channel matrices are estimated and $\widehat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma^2_{E,ic} \neq 0$. Note that the coordinates of the eNBs and UEs are shown in Figures 1(a) and 1(b); UE_1 is right on the 3-cell edge and each of the other UEs is near at least one of the 2-cell edges.

Figure 6 shows the MSE and BER results of examples 3 and 4 with full cooperation among 3 eNBs. All the MSE and BER results are obtained by averaging over 25 channel realizations. Note that the average per-antenna power P in examples 3 and 4 is the same. But the total power of example 4 is twice of (3 dB larger than) the total power of example 3 since the t_c in example 4 is twice of the t_c in example 3. Therefore, there is a 3 dB difference in the scales of the x-axes of examples 3 and 4 in Figure 6.

We make four main observations. First, the results for the per-cell and per-antenna power constraints are more or less the same for all of the approaches (the *GIA*, the *Robust-GIA*, and the *Non-robust-GIA*) in both examples 3 and 4. This is remarkable because the per-antenna power constraint, though more practical, is much stricter than the per-cell power constraint.

Secondly, as expected, the *Robust-GIA* yields better MSE and BER performances than the *Non-robust-GIA*. In the power ranges shown in Figure 6, the performance gain of the *Robust-GIA* over the *Non-robust-GIA* for the MSE results is around 5 dB for example 4 and 3 dB for example 3. The performance gain of the *Robust-GIA* over the *Non-robust-GIA* for the BER results is around 2–5 dB for example 4 and 0–3 dB for example 3. The performance gain for the MSE results decreases as power P increases. This is due to the fact that CSI estimation errors decrease as P increases ($T_c = t_cP$). However, the performance gain for the BER results increases as P increases. This is because the BER is expressed in log scale. We conclude that the robust approach is more crucial

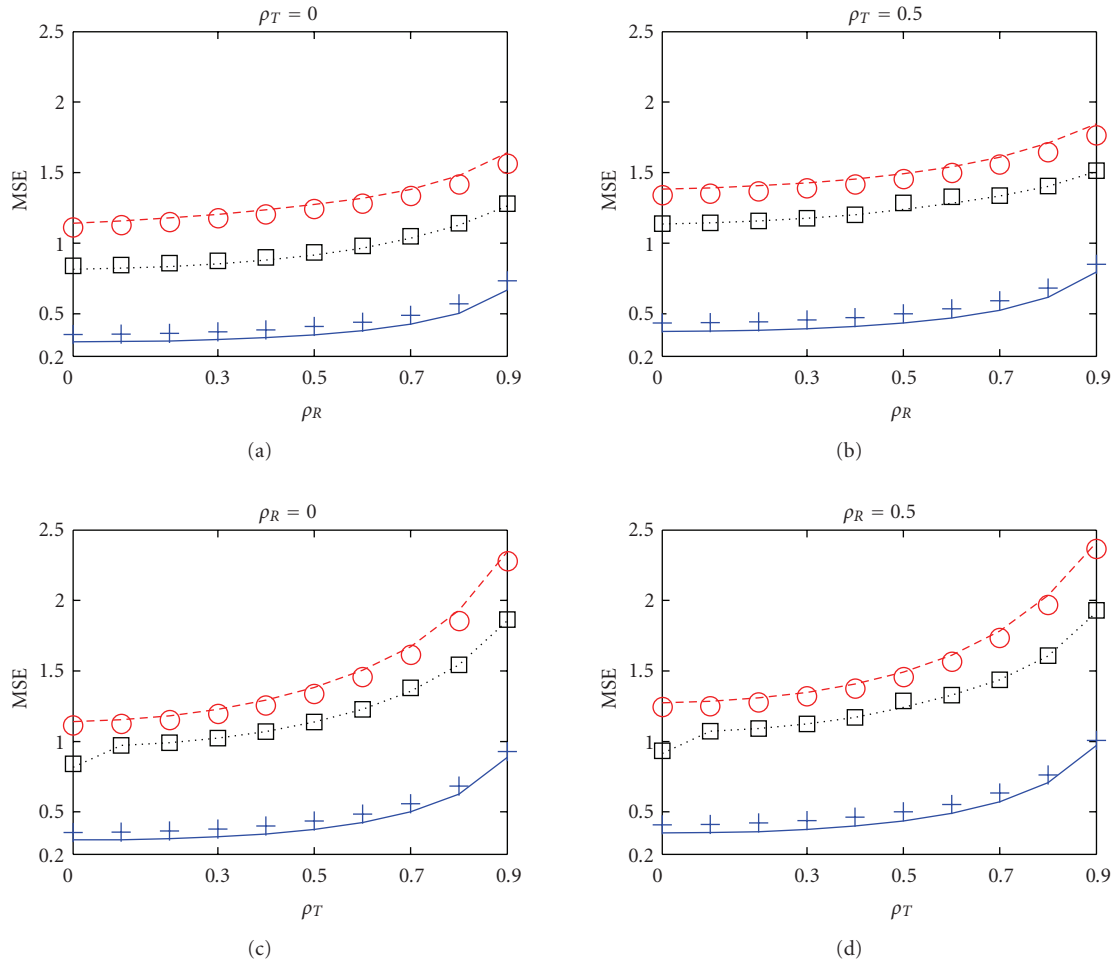


FIGURE 8: MSE as functions of ρ_T (with fixed ρ_R) and functions of ρ_R (with fixed ρ_T) under per-antenna and per-cell power constraints. The total transmission power is 16 dB. System parameters are shown in example 1 of Table 1. (*Legends are the same as in Figure 5.*)

to larger MIMO systems such as example 4 than smaller MIMO systems such as example 3.

Thirdly, also as expected, the *Robust-GIA* yields larger (worse) MSE and BER than the *GIA*. In the power ranges shown in Figure 6, the performance degradation of the *Robust-GIA* with respect to the *GIA* for the MSE results is around 9–10 dB for example 4 and around 7-8 dB for example 3. The performance degradation of the *Robust-GIA* with respect to the *GIA* for the BER results is around 9–12 dB for example 4 and 9 dB for example 3. For the same reasons stated before, the performance degradation of MSE results decreases as power P increases, but the performance degradation of BER results increases as power P increases. We conclude that larger MIMO systems such as example 4 are more sensitive to the CSI estimation errors than smaller MIMO systems such as example 3.

Lastly, compared to the results in example 3, the MSE results for all the approaches are noticeably higher in example 4, but the degradation of BER results in example 4 compared to example 3 is not significant if the per-antenna power P is the same in both examples. We conclude that cooperation among the eNBs is very effective in mitigating inter-cell interferences at cell edges. And, increasing the antenna numbers is an effective way to increase the system capacity even at cell edges as long as full eNB cooperation is allowed for the joint design of robust precoders and decoders.

5.5. Spatial Channel Correlations. Using the example 1 in Table 1, system performances of various approaches under different antenna correlation conditions are studied. The channel correlation matrices are defined as

$$\mathbf{R}_{T,c} = \begin{bmatrix} 1 & \rho_T \\ \rho_T & 1 \end{bmatrix}, \quad \mathbf{R}_{R,i} = \begin{bmatrix} 1 & \rho_R \\ \rho_R & 1 \end{bmatrix}, \quad i = 1, 2, \ c = 1. \tag{45}$$

We choose $d_{1c} = 1$ and $d_{2c} = 0.78$. Channel matrices are estimated and $\hat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma^2_{E,ic} \neq 0$.

Figures 7 and 8 show the MSE results for various values of ρ_T and ρ_R. In Figure 7, the MSE is plotted against the sum power; in Figure 8, the MSE is plotted against either ρ_T or ρ_R. The MSE results are obtained by averaging over 20 channel realizations. Again, we observe that the results for the per-cell and per-antenna power constraints are more or less the same for all of the approaches (the *GIA*, the *Robust-GIA*,

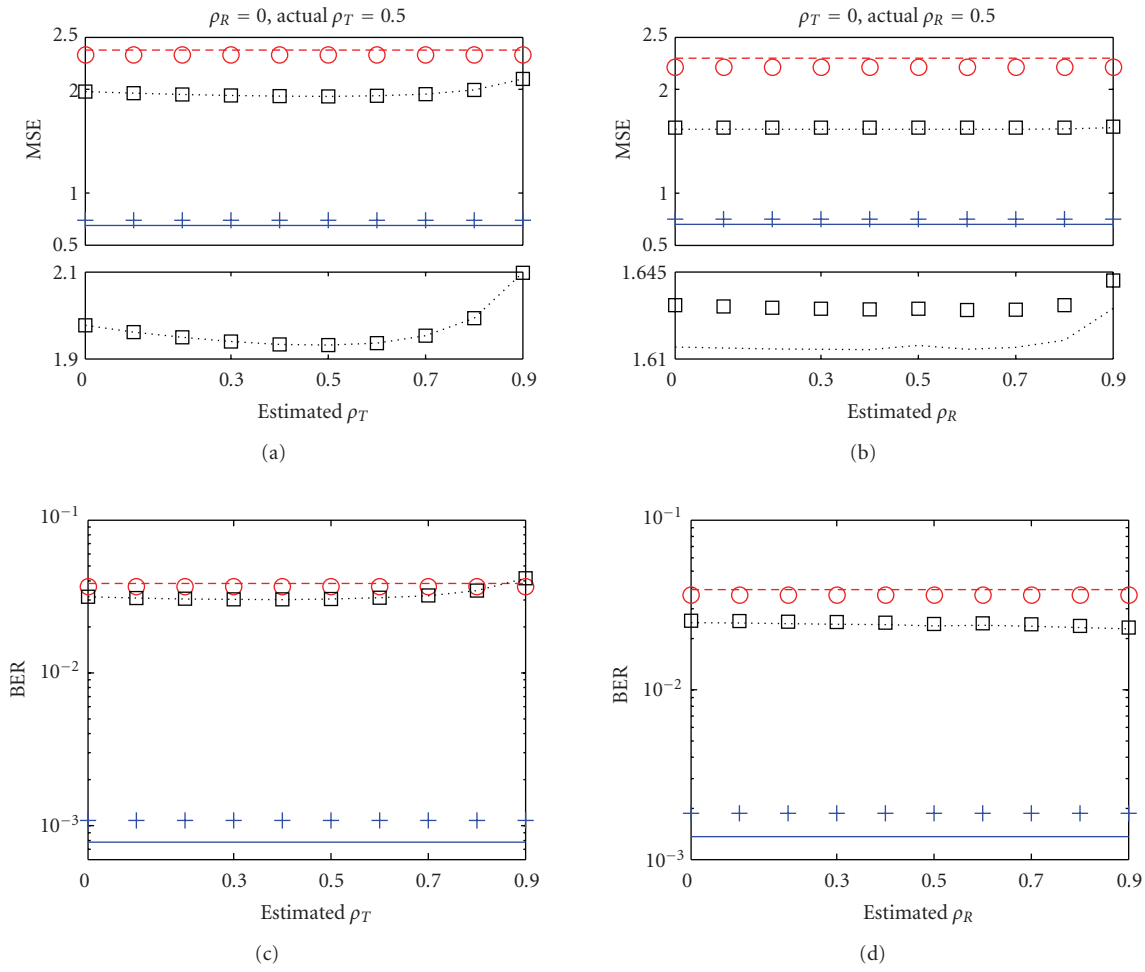


FIGURE 9: MSE and BER as functions of estimated ρ_T or ρ_R under per-antenna and per-cell power constraints ($10\log_{10}tP = 10\log_{10}3 + 10$ dB, averaging over 20 channel realizations). (*Legends are the same as in Figure 5.*)

and the *Non-robust-GIA*). As ρ_T and/or ρ_R increase, the MSE increases, the performance gain of the *Robust-GIA* over the *Non-robust-GIA* decreases, and the performance degradation of the *Robust-GIA* with respect to the *GIA* increases. The effect due to increasing ρ_T is more profound than the effect due to increasing ρ_R. We conclude that the robust approaches work satisfactorily in wireless channels with high channel correlations.

5.6. Sensitivity with Respect to Estimation Errors of ρ_T or ρ_R. Using the 3-cell configuration in Figure 1(a) and the system parameters of example 3 in Table 1, the sensitivity of MSE and BER performances with respect to estimation errors of ρ_T or ρ_R is studied. Channel matrices are estimated and $\hat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma^2_{E,ic} \neq 0$. Full cell cooperation is assumed. In Figure 9, MSE and BER are plotted against the estimated ρ_T for a fixed ρ_R or against the estimated ρ_R for a fixed ρ_T. The enlarged MSE results of the *Robust-GIA* are shown in the middle two subplots. First of all, the *GIA* results are independent of estimation errors of ρ_T and ρ_R because the perfect CSI is employed in the design. Similarly, the *Non-robust-GIA* results are also independent of estimation errors of ρ_T and ρ_R because channel correlation statistics are not needed in estimating the instantaneous channel matrices in practice. Secondly, the *Robust-GIA* outperforms the *Non-robust-GIA* in terms of both MSE and BER regardless of the estimation error in ρ_T or ρ_R. Thirdly, the performance degradation due to the estimation error in ρ_R (for a fixed ρ_T) is less profound than that due to the estimation error in ρ_T (for a fixed ρ_R). This is because the variance $\sigma^2_{E,ic}$ (see (4b)) of $\mathbf{E}_{W,ic}$ depends only on ρ_T and the accuracy of $\mathbf{R}_{E,ic}$ in (4a) is not significantly affected by the estimation error in ρ_R if SNR is sufficiently large. Lastly, the *Robust-GIA* is less sensitive to underestimates of ρ_T or ρ_R than overestimates of ρ_T or ρ_R. The same observations as above are made from sensitivity studies for various nonzero values of actual ρ_T or ρ_R. We conclude that effects of estimation errors (especially underestimates) of channel correlations ρ_T or ρ_R on the system performances of the *Robust-GIA* are very small.

5.7. Sensitivity with Respect to Estimation Errors of Path Loss. Using the 3-cell configuration in Figure 1(a) and the system parameters of example 3 in Table 1, the sensitivity of MSE and BER performances with respect to estimation

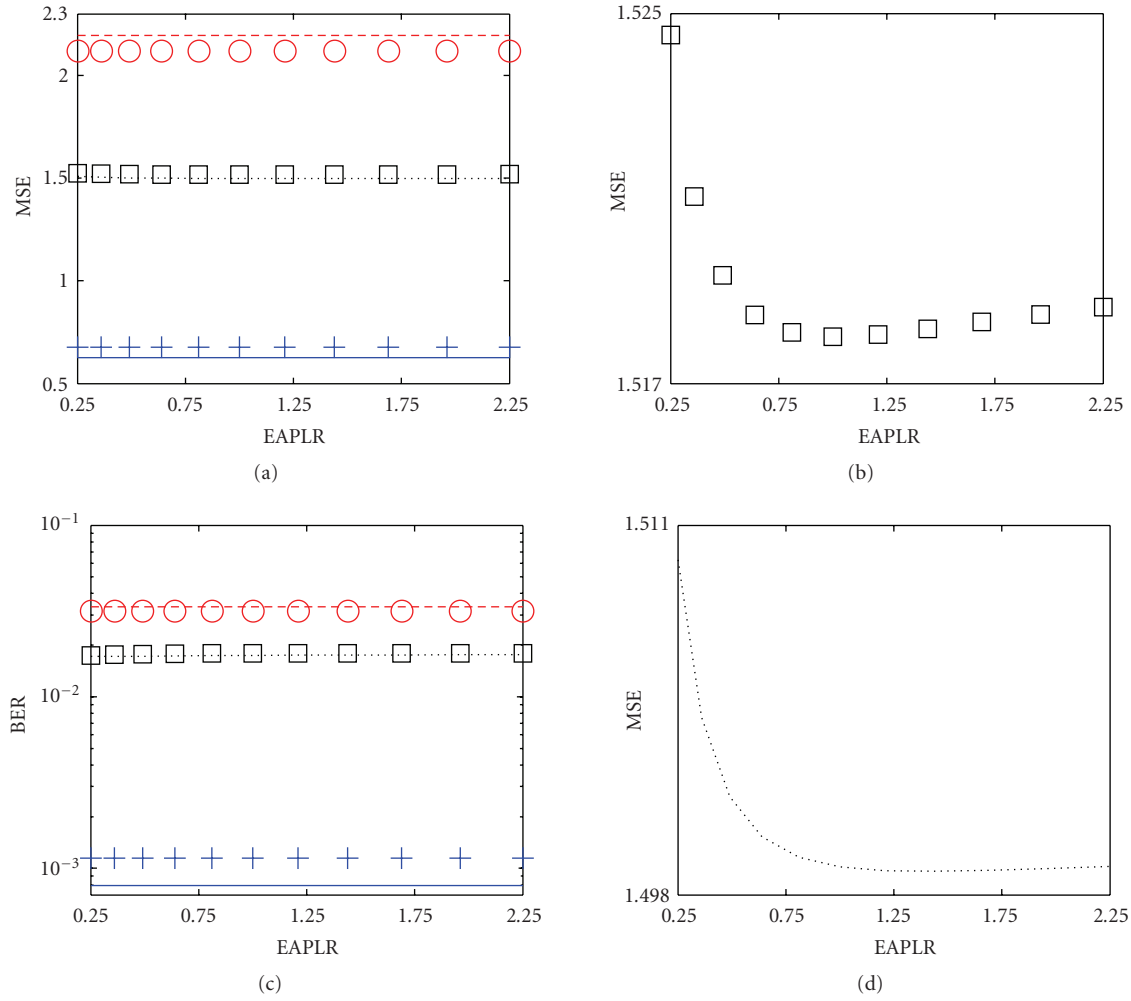


FIGURE 10: MSE and BER as functions of the estimated-to-actual-path loss ratio (EAPLR) under the per-antenna and per-cell power constraints ($10\log_{10} tP = 10\log_{10} 3 + 10$ dB, averaging over 20 channel realizations). (*Legends are the same as in Figure 5.*)

errors of path loss (PL $= \alpha d_{ic}^{-2\beta}$ in (4a)) is studied. Full cell cooperation is assumed. For convenience, we choose $\mathbf{R}_{T,c} = \mathbf{I}_{t_c}$ and $\mathbf{R}_{R,i} = \mathbf{I}_{r_i}$ for $i = 1, 2, 3$ and $c = 1, 2, 3$. Channel matrices are estimated and $\hat{\mathbf{H}}_{ic} \neq \mathbf{H}_{ic}$ and $\sigma_{E,ic}^2 \neq 0$. Define the estimated-to-actual-path-loss ratio (EAPLR) as $\text{PL}_{\text{estimated}}/\text{PL}_{\text{actual}}$. In Figure 10, MSE and BER are plotted against the EAPLR ranging from 0.25 to 2.25. The enlarged MSE results of the *Robust-GIA* under the per-antenna and per-cell power constraints are shown in the two right subplots, respectively. First of all, the *GIA* results are independent of estimation errors of path loss because the perfect CSI is employed in the design. Similarly, the *Non-robust-GIA* results are also independent of estimation errors of path loss because channel statistics are not used in estimating the instantaneous channel matrices. Secondly, the *Robust-GIA* outperforms the *Non-robust-GIA* in terms of both MSE and BER regardless of estimation errors of path loss. Thirdly, the degradation of MSE due to estimation errors of path loss is negligible. This is due to the fact that σ_c^2 in (13) is independent of path loss, and the accuracy of $\mathbf{R}_{E,ic}$ in (4a) is not significantly affected by path loss errors if SNR is sufficiently large. We conclude that the effects of estimation errors of path loss on the system performances of the *Robust-GIA* are negligible.

6. Conclusion

Three robust approaches, the *Robust-GIA*, the *Robust-FIA*, and the *Robust-DCOA*, are proposed for designing MMSE transceivers in the downlink of multicell multiuser MIMO Systems under general linear equality power constraints and with CSI estimation errors. The *GIA*, the *FIA*, and the *DCOA* (the perfect CSI approaches) are special cases of the *Robust-GIA*, the *Robust-FIA*, and the *Robust-DCOA* thus giving a general framework to deal with both perfect and imperfect CSI! Note that the robust approaches have only a small increase in complexity compared to their perfect-CSI counterparts.

The *Robust-DCOA* always gives optimum solutions but is only applicable when the rank constraints on the precoders are relaxed, the transmit correlation matrix of each user is an identity matrix, and the source covariance matrices are all identity matrices multiplied by the same constant. The statistics of the CSI estimation error also need to be the

same for all users if the power constraints of the users are interdependent. The *Robust-GIA*, on the other hand, has no such restrictions and is the most general among the three proposed robust approaches. It allows tradeoff between diversity and multiplexing gains, which is not possible in the *Robust-DCOA* or the *Robust-FIA*. The multiplexing gains of the *Robust-DCOA* or the *Robust-FIA* are determined by the ranks of the decoder covariance matrices. The *Robust-FIA* is a special case of the *Robust-GIA*. It, requiring that the source covariance matrices are identity matrices multiplied by the same constant, is a bit less flexible than the *Robust-GIA*. But, it is much more flexible than the *Robust-DCOA* since it does not require all of the transmit correlation matrices to be identity matrices. Both the *Robust-GIA* and the *Robust-FIA* are numerically more efficient than the *Robust-DCOA*. The *Robust-FIA* is slightly more efficient than the *Robust-GIA* because it combines two of the three steps in each iteration of the *Robust-GIA* into one step. All approaches show difficulties in convergence when the transmit power is very high. Relaxing the equality power constraints will improve the numerical efficiency greatly. Both the *Robust-GIA* and the *Robust-FIA* can only guarantee local optimality. But, whenever the *Robust-DCOA* is applicable and all the decoder covariance matrices are full rank, the three robust approaches are actually equivalent (i.e., the *Robust-GIA* and the *Robust-FIA* are also optimum).

MMSE transceiver designs using the three proposed approaches are performed for various single-cell and multicell examples with different system configurations, power constraints, channel spatial correlations, and cooperation scenarios. System performances in terms of MSE and BER are investigated. Important concluding remarks made from these numerical examples are list below. First of all, the robust approaches outperform their non-robust counterparts in most of the numerical simulations (even when the channel is highly correlated, when the CSI estimation errors are large, and when there exist estimation errors in statistics of channel parameters). Secondly, the performance of the with-cooperation scenario is much better than that of the without-cooperation scenario. With cell cooperation, the cell edge UEs perform as well as those UEs without inter-cell interferences and therefore the cell edge difficulties can be remedied. Thus, with full cell cooperation, the system throughput can increase linearly with the numbers of antennas for both transmission and reception. Thirdly, the robust approaches are insensitive to the estimation errors of the channel statistics (e.g., to channel correlations and path loss). This important feature makes robust approaches practical. Fourthly, the system performances derived under the more practical per-antenna power constraint are very similar to those with the per-cell power constraint. Thus, the practical per-antenna power constraint inflicts little performance losses compared to the optimum per-cell power constraint. Fifthly, the performance gain of the robust approaches over the non-robust approaches is more profound in larger MIMO systems. Sixthly, the performance gain of the robust approaches over the non-robust approaches is reduced if the channel correlations increase.

In short, we have herein proposed, for joint MMSE transceiver designs, three novel robust approaches: the *Robust-GIA* (the most general), *Robust-FIA* (the most efficient), and the *Robust-DCOA* (which guarantees the global optimality). The proposed approaches are indeed robust with respect to different system configurations, CSI estimation errors, channel correlations, and channel modeling errors. When cell cooperation is available, the robust approaches provide a remedy for solving the cell edge problem without reducing the number of data streams.

Acknowledgment

The authors would like to thank InterDigital Communications Corporation for its financial support.

References

[1] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, 2003.

[2] M. Codreanu, A. Tölli, M. Juntti, and M. Latva-aho, "Joint design of Tx-Rx beamformers in MIMO downlink channel," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4639–4655, 2007.

[3] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1051–1064, 2002.

[4] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 214–226, 2004.

[5] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 4, pp. 1037–1052, 2004.

[6] I.-T. Lu, "Joint MMSE precoder and decoder subject to arbitrary power constraints for uplink multiuser MIMO systems," in *Proceedings of the 70th IEEE Vehicular Technology Conference (VTC '09)*, Anchorage, Alaska, USA, September 2009.

[7] I.-T. Lu, J. Li, and E. Lu, "Novel MMSE precoder and decoder designs subject to per-antenna power constraint for uplink multiuser MIMO systems," in *Proceedings of the 3rd International Conference on Signal Processing and Communication Systems (ICSPCS '09)*, Omaha, Neb, USA, 2009.

[8] J. Zhang, Y. Wu, S. Zhou, and J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Communications Letters*, vol. 9, no. 11, pp. 991–993, 2005.

[9] M. Schubert, S. Shi, E. A. Jorswieck, and H. Boche, "Downlink sum-MSB transceiver optimization for linear multi-user MIMO systems," in *Proceedings of Annual Asilomar Conference on Signals, Systems and Computers*, pp. 1424–1428, Pacific Grove, Calif, USA, October 2005.

[10] G. Zheng, T.-S. Ng, and K.-K. Wong, "Optimal beamforming for sum-MSE minimization in MIMO downlink channels," in *Proceedings of the 63rd IEEE Vehicular Technology Conference (VTC '06)*, vol. 4, pp. 1830–1834, Melbourne, Canada, May 2006.

[11] S. Shi, M. Schubert, N. Vucic, and H. Boche, "MMSE optimization with per-base-station power constraints for

network MIMO systems," in *Proceedings of IEEE International Conference on Communications (ICC '08)*, pp. 4106–4110, Beijing, China, May 2008.

[12] I.-T. Lu, "Joint MMSE precoder and decoder design for downlink multiuser MIMO systems with arbitrary transmit power constraints," in *Proceedings of IEEE Sarnoff Symposium (SARNOFF '09)*, pp. 1–5, Princeton, NJ, USA, March 2009.

[13] J. Li, I.-T. Lu, and E. Lu, "Optimum MMSE transceiver designs for the downlink of multicell MIMO systems," in *Proceedings of Military Communications Conference (Milcom '09)*, Boston, Mass, USA, October 2009.

[14] N. Khaled, G. Leus, C. Desset, and H. De Man, "A robust joint linear precoder and decoder MMSE design for slowly time-varying MIMO channels," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 4, pp. 485–488, Montreal, Canada, May 2004.

[15] S. Serbetli and A. Yener, "MMSE transmitter design for correlated MIMO systems with imperfect channel estimates: power allocation trade-offs," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2295–2304, 2006.

[16] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Transactions on Signal Processing*, vol. 57, no. 3, pp. 1141–1150, 2009.

[17] X. Zhang, D. P. Palomar, and B. Ottersten, "Robust MAC MIMO transceiver design with partial CSIT and CSIR," in *Proceedings of the 41st Asilomar Conference on Signals, Systems and Computers*, pp. 324–328, Pacific Grove, Calif, USA, November 2007.

[18] M. B. Shenouda and T. N. Davidson, "On the design of linear transceivers for multiuser systems with channel uncertainty," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 1015–1024, 2008.

[19] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1–4, pp. 625–653, 1999.

[20] J. Löfberg, "YALMIP: a toolbox for modeling and optimization in MATLAB," in *Proceedings of the IEEE International Symposium on Computer-Aided Control System Design (CACSD '04)*, pp. 284–289, Taipei, Taiwan, September 2004.