Linear Precoding for Multiuser MIMO Systems with Multiple Base Stations

Imad H. Azzam
Research in Motion,
305 Phillip Street, Waterloo, ON, Canada N2L 3W8
Email: iazzam@rim.com

Raviraj S. Adve
Dept. of Elec. and Comp. Eng., University of Toronto,
10 King's College Road, Toronto, ON, Canada M5S 3G4.
Email: rsadve@comm.utoronto.ca

***Abstract*—Linear precoding for multiuser multiple input multiple output (MIMO) cellular systems has generally focused on a single isolated cell. A crucial tool in algorithm development has been a downlink/uplink duality. Here we investigate downlink multiuser communications across multiple cells. Previous work identified the resulting asynchronous interference as a key issue. We prove the existence of a downlink/uplink duality requiring time reversal. This duality is used to develop an effective linear precoding algorithm for multiuser, multi-cell, MIMO systems. Simulation results illustrate the importance of accounting for, and to quantify the loss due to, asynchronous interference.**

I. Introduction

The use of multiple antennas at the transmitters and/or receivers in a multiple input, multiple output (MIMO) system enables communication with multiple users over the same time/frequency channel [1]. A recent research theme in MIMO multiuser cellular communications is linear precoding - using channel state information at the transmitter to linearly predistort the users' signals at the transmitter to help decode them at the receivers [2–7]. Almost all the research in this area has focused on a single cell in a cellular network. To maximize frequency reuse, the goal of this paper is to develop linear precoding techniques for multiuser multi-cell communication. The major contributions of this paper are:

- developing downlink and virtual uplink system models for a multiuser, multi-BS, MIMO system that accounts for asynchronous interference;
- developing a useful downlink/uplink duality for such system;
- presenting a multiuser linear precoding algorithm that minimizes the SMSE of this multiuser system.

We focus on a system where multiple basestations (BSs) jointly communicate with multiple users distributed across the entire service area. The BSs and, potentially, the mobile users possess multiple antennas. The BSs are assumed to know the channel state information (CSI) and data to all users, an assumption common to all works on precoding [3–7]. One difference is the need for additional timing information. While the early works focused on minimizing the sum of the mean squared error (SMSE) across all users' signals [3–5], linear precoding to maximize sum data rate is also possible [6], [7]. In this work we use the SMSE as our figure of merit.

Efficient precoding algorithms are based on a duality between the multiuser downlink and a virtual multiuser uplink [3–5]. The duality states that, under the same sum power constraint and set of precoders and decoders, the downlink and uplink both have the same achievable signal-to-interference-plus-noise ratio (SINR) and MSE regions [3–5]. In other words, a certain set of MSEs can be achieved in the downlink and the virtual uplink given the same sum power constraint. This duality can be thought of as providing two perspectives of the multiuser system. If the solution of a certain problem can be found using one perspective, then this solution can be transformed to suit the other perspective. The duality allows for the precoding and decoding matrices to be obtained via receive processing only, i.e., the Wiener filter. Power allocation across users is then, provably, a convex optimization problem.

In attempting to generalize this duality, and linear precoding, to multiple cells in a cellular network, we are faced with an interesting fundamental problem: the interference at individual users is asynchronous. This in itself necessitates reexamining the linear precoding approach. Duality, as developed earlier, does not hold and the power allocation problem is not provably convex. One key contribution in this paper is a new virtual uplink system model, based on time-reversal, that restores duality.

The work in [8], with multiple interfering source-destination pairs, is probably the first to discuss multiuser communications with multiple BSs. The goal there is to minimize the transmit power while meeting quality of service requirements at the receivers. In a recent study, Tölli et al. [9] treat a problem similar to ours with a metric of sum-rate. The resulting algorithm is similar to that of [5] with a single 'super BS' using as many transmit antennas as all the cooperating BSs collectively possess. In [10], Tamakai et al. study the achievable sum rates for a MIMO downlink system with multiple BSs cooperating at three different levels. However, they assume that a variant of Time Division Multiple Access is used in the best results they achieve. In [11], Dahrouj and Yu consider the downlink of a multiuser, multi-BS system and propose an algorithm that jointly optimizes the beamformers used by the cooperating BSs to minimize the total transmit power while satisfying SINR constraints for each users. Note that all these works ignore the issue of asynchronous interference.

In [12] the authors illustrate the fundamental nature, and impact, of asynchronous interference. The work in [12] amends some existing algorithms to account for this asynchronous interference, but not for the scenario under consideration. Several other works consider the problem of asynchronous interference and propose different methods for mitigating its effect, mainly by using the cyclic redundancy provided by orthogonal frequency division multiplexing (OFDM) [13], [14] while in [15] the authors use an interference subtraction approach. Note that [13–15] all assume single-BS systems and do not address the multi-BS scenario, neither in terms of the difficulties that may arise nor in terms of any advantages.

This paper is organized as follows: Section II presents the downlink and virtual uplink models, including the crucial issue of timing. Section III presents the main theoretical contribution in this paper, the downlink/uplink duality in multi-BS systems. Section IV uses this duality to develop a linear precoding algorithm which is tested in Section V. The paper ends with some discussion and conclusions in Section VI.

II. Downlink and Virtual Uplink System Models

This section describes the system models of the multiuser downlink and a virtual multiuser uplink. The system comprises B BSs and K users randomly distributed across the B cells. Each BS has M antennas, while user k has N_k antennas ($N = \sum_{k=1}^{K} N_k$). Moreover, each user receives L_k data streams, where $L_k \leq \min\{M, N_k\}$, i.e., there are a total of $L = \sum_{k=1}^{K} L_k$ streams that share the same time/frequency channel. All BSs are assumed to know the CSI to all users perfectly. Importantly, this CSI also includes the propagation delays between all the BSs and all the users. How this information is received at the BSs is outside the scope of this paper.

A. Downlink System Model

Let $\mathbf{x}_k$ ($L_k \times 1$) be the data vector containing the independent data streams of user k to be transmitted from all BSs. The data symbols have unit average power. These vectors are independent over time and across users, i.e., $\mathbf{E}[\mathbf{x}_k(m)\mathbf{x}_j^H(n)] = \mathbf{I}_{L_k}\delta[k-j]\delta[m-n]$ where $\delta[\cdot]$ represents the discrete delta-function, $(\cdot)^H$ is the Hermitian operator, $\mathbf{E}[\cdot]$ is the expectation operator, and m and n are discrete time indices. Unless required, we will drop the time index.

The signal vector transmitted from BS b to user k is $\mathbf{t}_k^{(b)} = \mathbf{U}_k^{(b)}\sqrt{\mathbf{P}_k}\mathbf{x}_k$ where $\mathbf{P}_k$ is a diagonal power allocation matrix and $\mathbf{U}_k^{(b)}$ is the linear precoder for user k at BS b. The $N_k \times 1$ signal received by user k is

$$\mathbf{y}_k = \sum_{b=1}^{B} \mathbf{H}_k^{(b)^H} \mathbf{U}_k^{(b)} \sqrt{\mathbf{P}_k}\mathbf{x}_k + \text{interference} + \mathbf{n}_k, \quad (1)$$

where $\mathbf{H}_k^{(b)}$ denotes the flat channel and $\mathbf{n}_k$ denotes the additive white Gaussian noise (AWGN). The interference term will be explained in more detail below. Finally, user k processes its received signal by multiplying it by a decoding matrix $\mathbf{V}_k^H$ ($L_k \times N_k$) to estimate its own data vector $\hat{\mathbf{x}}_k = \mathbf{V}_k^H \mathbf{y}_k$.

Timing Issues: In the first term of Eqn. (1), it is assumed that the transmissions from all BSs meant for user k ($\mathbf{t}_k^{(b)}$, for $b = 1, \ldots, B$) are received by user k synchronously. Therefore, BS b must advance the time at which it transmits $\mathbf{t}_k^{(b)}$ by $\tau_k^{(b)} - \tau_k^{(b_k)}$, where $\tau_k^{(b)}$ is the propagation delay from BS b to user k, and $\tau_k^{(b_k)}$ is the propagation delay from user k to the nearest BS. Due to the random distribution of users, this implies that the interference at user k from other users in the system be is *asynchronous*. This asynchronism is significant is the key challenge in designing precoding for multiple BSs.

The misaligned interference can be expressed as

$$\text{interference} = \sum_{b=1}^{B} \mathbf{H}_k^{(b)^H} \sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{U}_j^{(b)} \sqrt{\mathbf{P}_j}\mathbf{i}_{jk}^{(b)}, \quad (2)$$

where, defining $g(t)$ as the unit power pulse shape used,

$$\mathbf{i}_{jk}^{(b)}(m) = \rho(\delta_{jk}^{(b)} - T_S)\mathbf{x}_j(m_{jk}^{(b)}) + \rho(\delta_{jk}^{(b)})\mathbf{x}_j(m_{jk}^{(b)} + 1), \quad (3)$$

$$\tau_{jk}^{(b)} = \tau_k^{(b)} - \tau_k^{(b_k)} - (\tau_j^{(b)} - \tau_j^{(b_j)}), \quad \delta_{jk}^{(b)} = \tau_{jk}^{(b)} \bmod T_S,$$

and T_s is the symbol period and $\rho(\tau) = \int_0^{T_S} g(t)g(t-\tau)dt$. The interference caused on user k when BS b transmits to user j, $\mathbf{i}_{jk}^{(b)}$, is a linear combination of two consecutive data vectors being transmitted to user j, at times $m_{jk}^{(b)}$ and $m_{jk}^{(b)}+1$. Note that in our convention, in the downlink, $m_{jk}^{(b)}$ is the index given to the first interfering symbol and $m_{jk}^{(b)}+1$ is that given to the second. The estimated data vector is, therefore,

$$\hat{\mathbf{x}}_k = \underbrace{\mathbf{V}_k^H \sum_{b=1}^{B} \mathbf{H}_k^{(b)^H} \mathbf{U}_k^{(b)} \sqrt{\mathbf{P}_k}\mathbf{x}_k}_{\text{desired signal}} + \underbrace{\mathbf{V}_k^H \sum_{b=1}^{B} \mathbf{H}_k^{(b)^H} \sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{U}_j^{(b)} \sqrt{\mathbf{P}_j}\mathbf{i}_{jk}^{(b)}}_{\text{asynchronous multiuser interference}} + \mathbf{V}_k^H \mathbf{n}_k. \quad (4)$$

Using Eqn. (3), it is easy to show that

$$\mathbf{E}\left[\mathbf{i}_{j_1 k}^{(b_1)} \mathbf{i}_{j_2 k}^{(b_2)^H}\right] = \begin{cases} \mathbf{0}, & \text{for } j_1, j_2, k \text{ distinct} \\ \beta_{jk}^{(b_1,b_2)}\mathbf{I}_{L_j}, & \text{for } j_1 = j_2 = j \neq k \\ \mathbf{I}_{L_j}, & \text{for } j_1 = j_2 = j = k \end{cases} \quad (5)$$

$$\beta_{jk}^{(b_1,b_2)} = \begin{cases} 0, & |m_{jk}^{(b_2)} - m_{jk}^{(b_1)}| > 1 \\ \rho(\delta_{jk}^{(b_1)})\rho(\delta_{jk}^{(b_2)} - T_s), & m_{jk}^{(b_2)} = m_{jk}^{(b_1)} + 1 \\ \rho(\delta_{jk}^{(b_1)})\rho(\delta_{jk}^{(b_2)}) + \\ \quad \rho(\delta_{jk}^{(b_1)} - T_s)\rho(\delta_{jk}^{(b_2)} - T_s), & m_{jk}^{(b_2)} = m_{jk}^{(b_1)} \\ \rho(\delta_{jk}^{(b_2)})\rho(\delta_{jk}^{(b_1)} - T_s), & m_{jk}^{(b_2)} = m_{jk}^{(b_1)} - 1 \end{cases} \quad (6)$$

This result will prove useful later.

B. Virtual Uplink System Model

In precoding for a single BS, the efficient algorithms use a SMSE downlink/uplink duality. One key contribution of this paper is a SMSE duality for asynchronous interference based on time reversal. This section presents the relevant *virtual* uplink model. The uplink system model is similar to that in the downlink, except that the K users are now communicating to B BSs with a power allocation matrix $\mathbf{Q}_k$ and matrices $\mathbf{U}_k$ and $\mathbf{V}_k$ switch roles to decoding and precoding matrices respectively. The signal transmitted by user k to all BSs is $\mathbf{t}_k = \mathbf{V}_k\sqrt{\mathbf{Q}_k}\mathbf{x}_k$. The channel matrices in the uplink are the Hermitian of those in the downlink. Therefore, the $M \times 1$ signal received by BS b from user k is

$$\mathbf{y}_k^{(b)} = \mathbf{H}_k^{(b)}\mathbf{V}_k\sqrt{\mathbf{Q}_k}\mathbf{x}_k + \text{ interference } + \mathbf{n}^{(b)}, \quad (7)$$

where $\mathbf{n}^{(b)}$ denotes the AWGN at BS b. The interference term is explained below. We assume that each BS can process its received signal with different delays for each of the K users.

The BSs use the decoding matrix $\mathbf{U}_k$ $(BM \times L_k)$ to estimate the data vector of user k; $\hat{\mathbf{x}}_k = \mathbf{U}_k^H\mathbf{y}_k = \sum_{b=1}^{B}\mathbf{U}_k^{(b)^H}\mathbf{y}_k^{(b)}$, where $\mathbf{U}_k = \left[\mathbf{U}_k^{(1)^T}, \ldots, \mathbf{U}_k^{(B)^T}\right]^T$ is the decoding matrix used by BS b for user k and ^{T} denotes the transpose operator.

One crucial difference between the work in [4], [5] and the multi-BS case is the issue of timing. It is fairly easy to show that if all users are assumed to transmit simultaneously, duality *does not hold*. On the other hand, if the users transmit at different times such that their signals arrive at the same time at the BS *closest to them*, say at $t = 0$, $\tau_{jk}^{(b)}$ is given by

$$\tau_{jk}^{(b)} = \tau_j^{(b)} - \tau_j^{(b_j)} - (\tau_k^{(b)} - \tau_k^{(b_k)}). \quad (8)$$

Eqn. (8) can be explained as follows. Since each BS should receive the signals transmitted by the users that have it as the closest BS at the same time, $t = 0$, user k transmits its signal at $t = -\tau_k^{(b_k)}$ and this signal arrives at BS b at $t = -\tau_k^{(b_k)} + \tau_k^{(b)}$. Similarly, the signal from user j arrives at BS b at $t = -\tau_j^{(b_j)} + \tau_j^{(b)}$.

The interference caused by user j on user k at BS b and the estimated data for user k, $\hat{\mathbf{x}}_k$ are given by

$$\hat{\mathbf{x}}_k = \sum_{b=1}^{B}\mathbf{U}_k^{(b)^H}\mathbf{H}_k^{(b)}\mathbf{V}_k\sqrt{\mathbf{Q}_k}\mathbf{x}_k + \sum_{b=1}^{B}\mathbf{U}_k^{(b)^H}\sum_{\substack{j=1\\j\neq k}}^{K}\mathbf{H}_j^{(b)}\mathbf{V}_j\sqrt{\mathbf{Q}_j}\mathbf{e}_{jk}^{(b)} + \sum_{b=1}^{B}\mathbf{U}_k^{(b)^H}\mathbf{n}^{(b)}, \quad (9)$$

$$\mathbf{e}_{jk}^{(b)}(m) = \rho(\delta_{jk}^{(b)})\mathbf{x}_j(m_{jk}^{(b)}) + \rho(\delta_{jk}^{(b)} - T_S)\mathbf{x}_j(m_{jk}^{(b)} + 1), \quad (10)$$

where $\delta_{jk}^{(b)} = \tau_{jk}^{(b)} \mod T_S$. Similar to Eqn. (5),

$$\mathbf{E}\left[\mathbf{e}_{j_1k}^{(b_1)}\mathbf{e}_{j_2k}^{(b_2)^H}\right] = \begin{cases} \mathbf{0}, & \text{for } j_1, j_2, k \text{ distinct} \\ \gamma_{jk}^{(b_1,b_2)}\mathbf{I}_{L_j}, & \text{for } j_1 = j_2 = j \neq k \\ \mathbf{I}_{L_j}, & \text{for } j_1 = j_2 = j = k \end{cases}, \quad (11)$$

where

$$\gamma_{jk}^{(b_1,b_2)} = \begin{cases} 0, & |m_{jk}^{(b_2)} - m_{jk}^{(b_1)}| > 1 \\ \rho(\delta_{jk}^{(b_2)})\rho(\delta_{jk}^{(b_1)} - T_s), & m_{jk}^{(b_2)} = m_{jk}^{(b_1)} + 1 \\ \rho(\delta_{jk}^{(b_1)})\rho(\delta_{jk}^{(b_2)}) + \rho(\delta_{jk}^{(b_1)} - T_s)\rho(\delta_{jk}^{(b_2)} - T_s), & m_{jk}^{(b_2)} = m_{jk}^{(b_1)} \\ \rho(\delta_{jk}^{(b_1)})\rho(\delta_{jk}^{(b_2)} - T_s), & m_{jk}^{(b_2)} = m_{jk}^{(b_1)} - 1 \end{cases} \quad (12)$$

We now prove a downlink/uplink duality that will lead to an efficient algorithm for multiuser linear precoding with multiple BSs.

III. Downlink/Uplink Duality

In general, proving a MSE downlink/uplink duality requires two steps. First, SINR targets are set on each data stream and the proof requires showing that the same total power is needed to meet these targets in the downlink and uplink. The proof then uses this SINR duality to, in turn, derive a MSE duality.

A. First Step: SINR Targets

Downlink: Let the target SINR for data stream j of user k be Γ_{kj}. Let $\mathbf{p}$ be a column downlink power allocation vector whose elements are the powers allocated to all data streams across all users, i.e., the diagonal elements of the matrices $\mathbf{P}_k$ for $k = 1, \ldots, K$. We would like to find the vector $\mathbf{p}$ that maximizes the minimum of the ratios $\text{SINR}_{kj}^{DL}/\Gamma_{kj}$ over all values of k and j while $||\mathbf{p}||_1 = \mathbf{1}^T\mathbf{p} \leq B \times P_{\max}$, where $\mathbf{1}$ is an all-ones $L \times 1$ vector and $P_{\max}$ is the power allocated to an individual base station. It was shown in [3] that the solution of this problem makes all these ratios equal:

$$C^{DL} = \frac{\text{SINR}_{kj}^{DL}}{\Gamma_{kj}}, ||\mathbf{p}||_1 \leq B \times P_{\max}, \quad (13)$$

for $1 \leq k \leq K$, $1 \leq j \leq L_k$. This is true because the SINR for any stream is increasing in the power of that data stream, and decreasing in the power of any other data stream.

In the uplink, define $\mathbf{q}$ and C^{UL} to be the power allocation vector and SINR ratio analogous to $\mathbf{p}$ and C^{DL}. The key to proving the first step is to show that $C^{DL} = C^{UL}$ *while* $||\mathbf{p}||_1 = ||\mathbf{q}||_1$, i.e., the same SINR targets are achievable with the same total available power.

Given the space constraints, a detailed proof is omitted. The key step is showing that $\mathbf{E}\left[\mathbf{i}_{j_1k}^{(b_1)}\mathbf{i}_{j_2k}^{(b_2)^H}\right] = \mathbf{E}\left[\mathbf{e}_{j_1k}^{(b_1)}\mathbf{e}_{j_2k}^{(b_2)^H}\right]$. This is made possible by showing that, *using time reversal*, $\delta_{jk}^{(b)^{DL}} = \delta_{kj}^{(b)^{UL}}$, i.e., $\beta_{jk}^{(b_1,b_2)} = \gamma_{kj}^{(b_1,b_2)}$.

B. Second Step: Equating MSEs

After establishing the first step to duality, showing that the downlink and uplink MSEs of a data stream are equal can be done along the lines of the proofs presented in [5]. Let $\mathbf{E}_k = \mathbf{E}\left[(\hat{\mathbf{x}}_k - \mathbf{x}_k)(\hat{\mathbf{x}}_k - \mathbf{x}_k)^H\right]$ denote the error covariance matrix of user k. The diagonal elements of $\mathbf{E}_k$ are the mean squared errors (MSEs) of the data streams of user k. In the uplink, the estimated data vector is given by Eqn. (9). Accordingly, the

MSE error matrix for user k in the uplink can be expressed as follows.

$$\mathbf{E}_k^{UL} = \mathbf{U}_k^H \left[\mathbf{H}_k \mathbf{V}_k \mathbf{Q}_k \mathbf{V}_k^H \mathbf{H}_k^H + \sum_{\substack{c=1 \\ c \neq k}}^{K} \mathbf{A}_{ck} + \sigma^2 \right] \mathbf{U}_k - \sqrt{\mathbf{Q}_k} \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{U}_k - \mathbf{U}_k^H \mathbf{H}_k \mathbf{V}_k \sqrt{\mathbf{Q}_k} + \mathbf{I}_{L_k}, \quad (14)$$

where $\mathbf{U}_k = [\mathbf{U}_k^{(1)^T} \ldots \mathbf{U}_k^{(B)^T}]^T$, $\mathbf{A}_{ck} = [\mathbf{A}_{ck}^{(1)^T} \ldots \mathbf{A}_{ck}^{(B)^T}]^T$, $\mathbf{A}_{ck}^{(b)} = \mathbf{H}_c^{(b)} \mathbf{V}_c \mathbf{Q}_c \mathbf{V}_c^H \mathbf{H}_c^H \mathbf{\Gamma}_{ck}^{(b)}$, and $\mathbf{\Gamma}_{ck}^{(b)} = diag[\gamma_{ck}^{(b,1)} \mathbf{I}_M, \ldots, \gamma_{ck}^{(b,B)} \mathbf{I}_M]$.

Let $\tilde{\mathbf{v}}_{kl}$ and $\tilde{\mathbf{u}}_{kl}$ be the MMSE decoding vectors for data stream l of user k in the downlink and uplink, respectively. The expressions for these vectors will be derived in detail later. Let $\mathbf{v}_{kl} = \tilde{\mathbf{v}}_{kl}/||\tilde{\mathbf{v}}_{kl}||$ and $\mathbf{u}_{kl} = \tilde{\mathbf{u}}_{kl}/||\tilde{\mathbf{u}}_{kl}||$. From Eqns. (4) and (14), the MSEs of this stream in the uplink and downlink are

$$\begin{aligned} \epsilon_{kl}^{UL} &= \tilde{\mathbf{u}}_{kl}^H \left[\mathbf{H}_k \mathbf{V}_k \mathbf{Q}_k \mathbf{V}_k^H \mathbf{H}_k^H + \sum_{\substack{c=1 \\ c \neq k}}^{K} \mathbf{A}_{ck} + \sigma^2 \right] \tilde{\mathbf{u}}_{kl} \\ &\quad - \sqrt{\mathbf{q}_{kl}} \mathbf{v}_{kl}^H \mathbf{H}_k^H \tilde{\mathbf{u}}_{kl} - \tilde{\mathbf{u}}_{kl}^H \mathbf{H}_k \mathbf{v}_{kl} \sqrt{\mathbf{q}_{kl}} + 1, \\ \epsilon_{kl}^{DL} &= \tilde{\mathbf{v}}_{kl}^H \left[\mathbf{H}_k^H \mathbf{U}_k \mathbf{P}_k \mathbf{U}_k^H \mathbf{H}_k + \mathbf{H}_k^H \sum_{\substack{c=1 \\ c \neq k}}^{K} \mathbf{B}_{ck} \mathbf{H}_k + \sigma^2 \right] \tilde{\mathbf{v}}_{kl} \\ &\quad - \sqrt{\mathbf{p}_{kl}} \mathbf{u}_{kl}^H \mathbf{H}_k \tilde{\mathbf{v}}_{kl} - \tilde{\mathbf{v}}_{kl}^H \mathbf{H}_k^H \mathbf{u}_{kl} \sqrt{\mathbf{p}_{kl}} + 1, \end{aligned}$$

where

$$\mathbf{B}_{ck} = [\mathbf{B}_{ck}^{(1)^T} \ldots \mathbf{B}_{ck}^{(B)^T}]^T, \quad \mathbf{B}_{ck}^{(b)} = \mathbf{U}_c^{(b)} \mathbf{P}_c \mathbf{U}_c^H \boldsymbol{\beta}_{ck}^{(b)}, \quad (15)$$

$$\boldsymbol{\beta}_{ck}^{(b)} = \text{diag}[\beta_{ck}^{(b,1)} \mathbf{I}_M, \ldots, \beta_{ck}^{(b,B)} \mathbf{I}_M]. \quad (16)$$

Using the previous step to set $\text{SINR}_{kl}^{DL} = \text{SINR}_{kl}^{UL}$ and simple mathematical manipulations, we get $\epsilon_{kl}^{DL} = \epsilon_{kl}^{UL}$.

In summary, in this section we have developed an downlink/uplink duality for the SINR and MSE of each individual data stream. Given a set of precoders, decoders and *sum* power allocation, the achievable SINR and MSE regions are the same in the virtual uplink and the downlink. This is possible because the uplink uses time reversal. We now use this duality to develop an effective algorithm to obtain the required precoders, decoders and power allocation to minimize the sum mean squared error across all data streams.

Before proceeding, it is worth discussing why we could not use OFDM to deal with asynchronism. A MIMO-OFDM transmitter has as many inverse Fast Fourier Transform (IFFT) modules as there are antennas; in a multiuser system all data for all users is applied at the input of the IFFT modules and processed in one go. As we have seen, in a multiuser multi-BS case, each BS will be required to transmit to different users separately *at different times* that differ by *non-integer* multiples of the symbol period. Accordingly, a BS would require a set of IFFT modules for every user and each antenna in the system, which is obviously impractical given the high number of users. One might attempt performing the IFFT operation on all the users' data all at once, but then individual timing advances and delays are not possible.

IV. Minimizing the Sum Mean Squared Error

In this section, we present an algorithm that minimizes the SMSE of the multiple BS system. With duality established for both power and MSE, it is possible to minimize the SMSE in the uplink (which is mathematically more tractable) and convert the solution obtained into the downlink. The solution consists of the precoding and decoding matrices and the power allocation that minimize the SMSE of the system. The algorithm is based on that of [5], which minimizes the SMSE of a system with only one BS and, hence, synchronous interference. The basic idea of both algorithms is to alternate between the downlink and virtual uplink, each time deriving the precoding matrices, decoding matrices, or power allocation that minimize the SMSE (equal in the downlink and virtual uplink), assuming all other variables are constant.

We begin by deriving an expression of the SMSE in the virtual uplink. Using Eqn. (14) and keeping all other matrices constant, the optimal matrix $\mathbf{U}_k$ is the Wiener solution:

$$\mathbf{U}_k^{MMSE} = \Big(\mathbf{H}_k \mathbf{V}_k \mathbf{Q}_k \mathbf{V}_k^H \mathbf{H}_k^H + \sum_{\substack{c=1 \\ c \neq k}}^{K} \mathbf{A}_{ck} + \sigma^2 \mathbf{I}_{N_k} \Big)^{-1} \times \mathbf{H}_k \mathbf{V}_k \sqrt{\mathbf{Q}_k}. \quad (17)$$

The columns of $\mathbf{U}_k^{MMSE}$ are normalized to be unit norm. Substituting this solution into Eqn. (14), we get

$$\mathbf{E}_k^{UL,MMSE} = \mathbf{I}_{L_k} - \sqrt{\mathbf{Q}_k} \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{J}_k^{-1} \mathbf{H}_k \mathbf{V}_k \sqrt{\mathbf{Q}_k}, \quad (18)$$

where

$$\mathbf{J}_k = \mathbf{H}_k \mathbf{V}_k \mathbf{Q}_k \mathbf{V}_k^H \mathbf{H}_k^H + \sum_{\substack{c=1 \\ c \neq k}}^{K} \mathbf{A}_{jk} + \sigma^2 \mathbf{I}. \quad (19)$$

The diagonal elements of $\mathbf{E}_k^{UL,MMSE}$ are the MSEs of the data streams of user k. Therefore:

$$\begin{aligned} SMSE &= \sum_{k=1}^{K} SMSE_k = \sum_{k=1}^{K} \text{tr}(\mathbf{E}_k^{UL,MMSE}) \\ &= L - \sum_{k=1}^{K} \text{tr} \left(\mathbf{H}_k \mathbf{V}_k \mathbf{Q}_k \mathbf{V}_k^H \mathbf{H}_k^H \mathbf{J}_k^{-1} \right). \end{aligned} \quad (20)$$

In the downlink, the optimal matrix $\mathbf{V}_k$ to minimize SMSE is, again, the Wiener solution:

$$\mathbf{V}_k^{MMSE} = \Big(\mathbf{H}_k^H \mathbf{U}_k \mathbf{P}_k \mathbf{U}_k^H \mathbf{H}_k + \mathbf{H}_k^H \sum_{\substack{c=1 \\ c \neq k}}^{K} \mathbf{B}_{ck} \mathbf{H}_k + \sigma^2 \mathbf{I}_{N_k} \Big)^{-1} \times \mathbf{H}_k^H \mathbf{U}_k \sqrt{\mathbf{P}_k}. \quad (21)$$

The final step is a power allocation that minimizes the SMSE.

The optimization problem is:

$$\min_{\substack{\mathbf{Q}_k\\k=1,\ldots,K}} \sum_{k=1}^{K} \text{tr}\left(\mathbf{E}_k^{UL,MMSE}\right) \text{ s.t.} \sum_{k=1}^{K} \text{tr}(\mathbf{Q}_k) \leq B \times P_{max}, \tag{22}$$

where P_{max} is the average transmit power of a single BS. In the single-BS case ($B = 1$), this is a convex problem [5]; unfortunately, Eqn. (22) is not provably so. To explain, if $B = 1$ and all interference is synchronous, $\mathbf{\Gamma}_{ck}^{(b)} = \mathbf{I}_{BM}$, and $\mathbf{A}_{ck} = \mathbf{H}_c\mathbf{V}_c\mathbf{Q}_c\mathbf{V}_c^H\mathbf{H}_c^H$. Consequently, the matrix $\mathbf{J}_k$ becomes common to all users and, hence, independent of k. It is this simplification that helps prove the convexity of the power allocation problem in [5]. In our multi-BS case, we are not able to guarantee convexity of this power allocation problem. It is only simulations that suggest that convexity in the uplink powers still holds. In our simulations, we solve the optimization problem in Eqn. (22) using sequential quadratic programming [16], an effective scheme to solve non-convex optimization problems (sometimes to a local minimum).

The steps given below detail the algorithm to minimize the SMSE in the multiuser multi-BS case. The matrices $\mathbf{P}$ and $\mathbf{Q}$ are the global downlink and uplink power matrices respectively; q_{kj} is the power allocated to data stream j of user k. The matrices $\mathbf{V}_k$ are initialized with the the right singular vectors of the channel matrices $\mathbf{H}_k$. The algorithm:

Initialization: $\mathbf{V}_k = \text{SVD}(\mathbf{H}_k)$, $\mathbf{Q} = (P_{max}/L)\mathbf{I}$

Iteration:

1) Find virtual uplink power allocation to minimize SMSE $\mathbf{Q} = \arg\min_{\mathbf{Q}} SMSE$, such that $q_{kj} > 0$, $tr(\mathbf{Q}) \leq BP_{max}$
2) Find downlink precoding matrices and normalize their columns: for $k = 1, \ldots, K$, $j = 1, \ldots, L_k$
$\mathbf{U}_k = \mathbf{J}_k^{-1}\mathbf{H}_k\mathbf{V}_k\sqrt{\mathbf{Q}_k}, \qquad \mathbf{u}_{kj} = \mathbf{u}_{kj}/||\mathbf{u}_{kj}||$
3) Set the target SINRs to the actual SINRs, for
$k = 1, \ldots, K$, $j = 1, \ldots, L_k$ $\Gamma_{kj} = SINR_{kj}^{UL}$
4) Find the downlink power allocation
$\mathbf{P} = diag\left(\sigma^2\left(\mathbf{D}^{-1} - \mathbf{\Psi}^{UL^T}\right)^{-1}\mathbf{1}\right)$
5) Find uplink precoding matrices and normalize their columns: for $k = 1, \ldots, K$, $j = 1, \ldots, L_k$
$\mathbf{V}_k = \mathbf{G}_k^{-1}\mathbf{H}_k^H\mathbf{U}_k\sqrt{\mathbf{P}_k}, \qquad \mathbf{v}_{kj} = \mathbf{v}_{kj}/||\mathbf{v}_{kj}||$
6) Repeat steps 1 to 5 until old SMSE - new SMSE $< \epsilon$

V. Numerical Simulations

This section presents the results of simulations to illustrate the efficacy of the algorithm described. The values used for all parameters are found in Table I. The SNR is defined as $SNR = P_{max}/\sigma^2$ and the system uses rectangular pulses. When path loss is taken into consideration, it is assumed that the power of the signal is proportional to the inverse of the distance raised to a path loss exponent of 3.5. The BSs are placed 500 meters apart, i.e., the cell radius is 250 meters. The minimum distance of any user to a BS was set to 150 meters. The K users are uniformly distributed in the rectangular area of width 200 meters and height $500/\sqrt{3}$ meters centered between the 2 BSs. The path loss is normalized to the path loss experienced 150 meters from a BS.

Number of BSs	B	2
Number of Tx. antennas per BS	M	4
Number of users	K	4
Number of Rx. antennas per user	N_k	1
Number of data streams per user	L_k	1
AWGN average power	σ^2	1
Symbol Period	T_S	$1\mu s$

TABLE I
Parameters for simulations

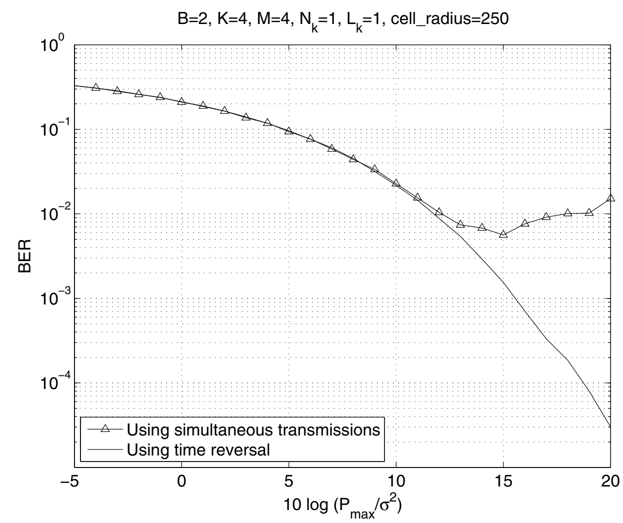

Fig. 1. Linear precoding with and without asynchronous interference

Figure 1 plots the error rate when the proposed algorithm is used. The simulated data includes path loss. In the first curve, we ignore the fact that duality does not exist and derive the needed matrices and power allocation using the uplink delays. As a result, the downlink power allocation was found to not always satisfy the power constraint; in such cases it was linearly scaled to meet the constraint. We can see that the error rate hits a floor before worsening, underlining the need to account for asynchronous interference. When time reversal is used to model the virtual uplink, duality exists, and the performance improves with increasing SNR.

In order to evaluate the proposed algorithm further, we compare it to the case where a super-BS with $B \times M$ antennas is used and interference is synchronous. To remain consistent with previous work in this scenario, in this example we ignore path loss. Figure 2 shows the results obtained. Again, when duality is assumed to exist, but does not, the error rate hits a floor and worsens. When time reversal is used, the performance is close to the super-BS case with synchronous interference. The gap between the curves demonstrates the loss due to asynchronous interference.

Figure 3 presents results of simulations suggesting that the

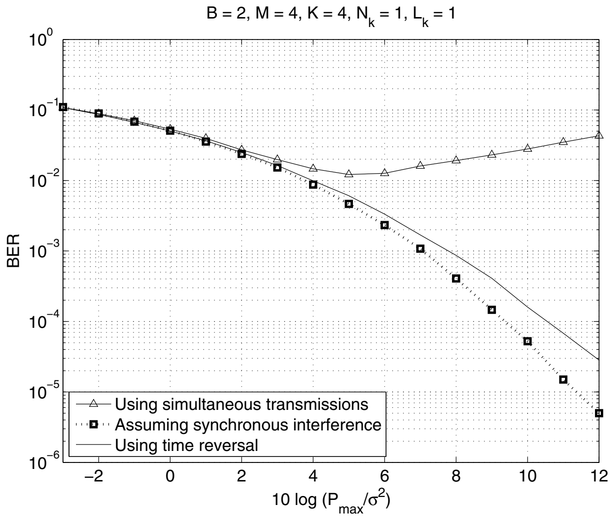

Fig. 2. Linear precoding with and without asynchronous interference, no path loss

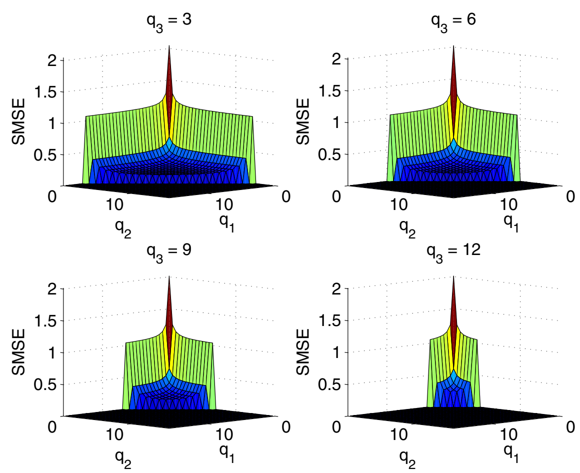

Fig. 3. SMSE versus power of three users

SMSE power allocation problem in Eqn. (20) is convex. The simulated scenario has $B = 2, K = 3, M = 3, N_k = 2$, and $L_k = 1$. The power of the third user q_3 is set to 4 different values, and for each value the SMSE is plotted versus the powers of the first two users, q_1 and q_2. The 3D plots suggest that the SMSE function is indeed convex.

VI. Conclusion

This paper has considered the downlink of a wireless cellular system wherein multiple base stations jointly communicate with multiple users simultaneously on the same time/frequency channel. The system assumes multiple antennas at BSs and potentially at the users. The key contributions here are the detailed model for asynchronous interference in both the downlink and a virtual uplink and the development of a duality between the two. It is assumed that perfect CSI is available at the multiple BSs and joint processing is possible. How this is implemented is beyond the scope of this paper.

Essentially, this paper accounts for the multiple BSs, the possibility of cooperation and asynchronous inter-user interference. We presented a detailed model for the system described above that takes into account the required timing advances for synchronous reception at the users, based on the model in [12]. We showed that, under time reversal, a duality exists between the downlink and uplink, despite the presence of asynchronous interference. This proof generalizes the downlink/uplink duality to the multiuser, multi-BS, MIMO case. With duality in hand, we extended an existing single-BS linear precoding algorithm to accommodate the presence of multiple BSs and asynchronous interference. In our case, the power allocation step was not provably convex; however, we provided simulations that suggested that the SMSE was still convex in the powers of the data streams. Simulations showed the performance of the extension and compared it to the performance assuming all reception was synchronous. This quantified the loss due to asynchronous interference.

References

[1] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[2] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. on Sig. Proc.*, vol. 52, no. 2, pp. 461–471, February 2004.

[3] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. on Veh. Tech.*, vol. 53, no. 1, pp. 18–28, January 2004.

[4] S. Shi and M. Schubert, "MMSE transmit optimization for multi-user multi-antenna systems," in *Proc. IEEE ICASSP 05*, March 2005.

[5] A. M. Khachan, A. J. Tenenbaum, and R. S. Adve, "Linear processing for the downlink in multiuser MIMO systems with multiple data streams," in *Proc. IEEE ICC 06*, June 2006.

[6] M. Codreanu, A. Tölli, M. Juntti, and M. Latva-aho, "Joint design of Tx-Rx beamformers in MIMO downlink channel," *IEEE Trans. on Signal Proc.*, vol. 55, no. 9, pp. 4639–4655, September 2007.

[7] A. J. Tenenbaum and R. S. Adve, "Improved sum-rate optimization in the multiuser MIMO downlink," in *Proc. CISS*, March 2008.

[8] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE VTC Spring*, May 1999, pp. 851–856.

[9] A. Tolli, M. Codreanu, and M. Juntti, "Linear cooperative multiuser MIMO transciever design with per BS power constraints," in *Proc. IEEE ICC 07*, Glasgow, Scotland, June 2007.

[10] T. Tamaki, K. Seong, and J. M. Cioffi, "Downlink MIMO systems using cooperation among base stations in a slow fading channel," in *Proc. IEEE ICC 07*, Glasgow, Scotland, June 2007.

[11] H. Dahrouj and W. Yu, "Coordinated beamforming for the multi-cell multi-antenna wireless system," in *Proc. CISS*, March 2008.

[12] H. Zhang, N. B. Mehta, A. F. Molisch, J. Zhang, and H. Dai, "On the fundamentally asynchronous nature of interference in cooperative base station systems," in *Proc. IEEE ICC 07*, Glasgow, Scotland, June 2007.

[13] T. A. Thomas and F. W. Vook, "Asynchronous interference suppression in broadband cyclic-prefix communications," in *Proc. IEEE WCNC*, New Orleans, LA, March 2003.

[14] K. Yano and M. Taromaru, "Pre-FFT type MMSE adaptive array antenna to suppress asynchronous interference for OFDM packet transmission," in *Proc. IEEE WCNC*, Hong Kong, March 2007.

[15] J. Hyejung and M. D. Zoltowski, "On the equalization of asynchronous multiuser OFDM signals in fading channels," in *Proc. IEEE ICASSP*, May 2004.

[16] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," in *Acta Numerica*. Cambridge University Press, 1995, pp. 1–51.