

# Asynchronous Interference Mitigation in Cooperative Base Station Systems

Hongyuan Zhang, *Member, IEEE*, Neelesh B. Mehta, *Senior Member, IEEE*,  
Andreas F. Molisch *Fellow, IEEE*, Jin Zhang, *Senior Member, IEEE*, and Huaiyu  
Dai *Member, IEEE*

## Abstract

Cooperative transmission by base stations (BSs) can significantly improve the spectral efficiency of multiuser, multi-cell multiple input multiple output (MIMO) systems. We show that contrary to what is often assumed in the literature, the multiuser interference in such systems, is fundamentally asynchronous. Intuitively, perfect timing-advance mechanisms can at best only ensure that the desired signal components – but not also the interference components – are perfectly aligned at their intended mobile stations (MSs). We develop an accurate mathematical model for the asynchronicity, and show that it leads to a significant performance degradation of existing designs that ignore the asynchronicity of interference. Using three previously proposed linear precoding design methods for BS cooperation, we develop corresponding algorithms that are better at mitigating the impact of the asynchronicity of the interference. Furthermore, we also address timing-advance inaccuracies (jitter), which are inevitable in a practical system. We show that using jitter-statistics-aware precoders can mitigate the impact of these inaccuracies as well. The insights of this paper are critical for the practical implementation of BS cooperation in multiuser MIMO systems, a topic that is largely ignored or oversimplified in the literature.

## Index Terms

Base station cooperation, Multiuser MIMO, linear precoding, interference leakage, timing-advance, mean square error, spectral efficiency, jitter.

H. Zhang is with Marvell Semiconductor Inc, Santa Clara, CA, USA, e-mail: hongyuan@marvell.com. N. B. Mehta, A. F. Molisch, and J. Zhang are with the Mitsubishi Electric Research Labs (MERL), Cambridge, MA, USA, e-mails: molisch, mehta, jzhang@merl.com. H. Dai is with the Dept. of Electrical and Computer Eng., NC State Univ., Raleigh, NC, USA, e-mail: hdai@ncsu.edu. A. F. Molisch is also at the Dept. of Electroscience, Lund Univ., Lund, Sweden.

This work was done when H. Zhang was at MERL.

# Asynchronous Interference Mitigation in Cooperative Base Station Systems

## I. INTRODUCTION

While the spectral efficiency gains of multiple input multiple output (MIMO) systems are significant for point-to-point links [1], they are limited in multi-user cellular networks by inter-cell co-channel interference (CCI) [2], [3]. In conventional cellular systems, CCI is reduced by careful radio resource management techniques such as power control, frequency reuse, and spreading code assignments [4]. However, these techniques limit the achievable spectral efficiency gains and/or lead to insufficient suppression of CCI. Recently, it has been shown that base station (BS) cooperation, in which different BSs together transmit signals for different mobile stations (MSs), can significantly improve spectral efficiency.

The theoretical analyses of BS cooperation often assume that the multiple BSs can be modeled as a single giant BS with more antennas. This assumption was implicitly used in [5]–[10]. Specifically, [5]–[7] extend dirty paper coding for a single cell broadcast channel to the case of multiple cells with cooperative base stations. References [9], [10] look at sub-optimal but simplified precoding schemes for cooperative base stations. The advantage of the “single giant BS” model is that it enables the well-studied single cell downlink transmission model to be applied in a straightforward manner. This assumption also implies that both the desired and the interfering signals from different BSs arrive at each MS simultaneously. Thus, all the above papers assume that the interference is always synchronous.

However, as we show in this paper, the interference is inherently asynchronous. While a detailed and rigorous mathematical model is developed in Sec. II, this can be understood intuitively by the following argument. Perfect timing-advance mechanisms can ensure that the signals from the BSs arrive at their *intended* recipients synchronously. However, the BSs cannot also align all the interfering signals at each MS because of the different propagation times between the BSs and MSs. Thus, the simultaneous arrival of both the desired and interfering signals at all the MSs is fundamentally unrealizable.

As we shall see, ignoring this asynchronicity can significantly degrade the performance of the

BS cooperative schemes proposed in the literature. To the best of our knowledge, this problem of asynchronous MIMO interference has not been addressed in the literature. This paper develops a framework for BS cooperation – in a multi-user multi-cell MIMO cellular network – that explicitly accounts for the asynchronous interference. It then uses this framework to analyze the detrimental impact of asynchronism on existing precoding algorithms, and suggests how to mitigate it by adapting the precoding design methods.

BS cooperation can be implemented in multiple ways such as Dirty paper coding [5], [6] or Tomlinson-Harashima precoding [11], or multi-user detection in MSs [3]. However, the above promising solutions are prohibitively complicated. We shall therefore focus on linear precoding designs, which have relatively lower complexity requirements at both the BSs and MSs [8]–[10]. They mitigate inter-cell interference, exploit macro-diversity, and can avoid capacity bottlenecks in severely spatially correlated channels [5]–[10].

Given the complexity of the problem, various design methods have been used in the literature to arrive at the optimal linear precoding matrices. These include minimizing the mean square error (MSE) [12] or maximizing the signal to leakage plus noise ratio (SLNR) [13], [14] or maximizing the sum rate [5], [6], [10], which, arguably, is the ultimate metric that determines spectrum utilization. For each of these methods, we take into account the asynchronicity of the interference and show that doing so mitigates its impact. Furthermore, we also address the problem of timing-advance inaccuracies (jitter), which are inevitable in a practical system. We show that using jitter-statistics-aware precoders can mitigate the impact of these inaccuracies as well.

The rest of the paper is organized as follows. Section II develops a detailed model for the asynchronous interference. Section III develops the three linear precoding algorithms. Section IV extends the algorithms to the case in which timing errors lead to imperfect synchronization even for the desired signals. The numerical results are presented in Sec. V, and are followed by our conclusions and discussion in Sec. VI.

## II. SYSTEM MODEL

We consider a cellular system with  $B$  BSs (each with  $N_T$  antennas) and  $K$  MSs/users (each with  $N_R$  antennas). The cooperative BSs together transmit  $L_k$  data streams to MS  $k$ . The different links are independent and undergo frequency-flat Rayleigh fading. Therefore,  $\mathbf{H}_k^{(b)}$ , the baseband

matrix representation of the channel from BS  $b$  to MS  $k$ , has complex Gaussian elements. Let  $b_k$  denote the index of the BS closest to MS  $k$ . For any MS, the BSs cooperate and jointly transmit the signals intended for it. The transmit vector for MS  $k$  from BS  $b$  is linearly precoded by the  $N_T \times L_k$  matrix  $\mathbf{T}_k^{(b)}$  as  $\mathbf{x}_k^{(b)}(m) = \mathbf{T}_k^{(b)} \mathbf{s}_k(m)$ , where  $\mathbf{s}_k(m)$  denotes the zero-mean data vector, of size  $L_k \times 1$  at time  $m$ , meant for MS  $k$ . As in [5]–[10], we assume that each BS has complete channel state information (CSI) for all the channels to all the MSs. We also assume a block-fading channel model with a large enough coherence time so that the channel fading remains the same over the duration in which  $\mathbf{T}_k^{(b)}$  is used. Given current CSI, in order to maximize the per-user transmission information rate, a Gaussian code book is used for the transmit data vectors, with normalized power such that  $\mathbb{E} [\mathbf{s}_k(m) \mathbf{s}_k(m)^\dagger] = \mathbf{I}_{L_k}$ . Furthermore, the code books for different users are independent of each other, i.e.,  $\mathbb{E} [\mathbf{s}_k(m) \mathbf{s}_l(m)^\dagger] = \mathbf{0}$ , for  $k \neq l$ .

#### A. Asynchronous Interference Despite Perfect Synchronization

The CSI available at each BS also includes the knowledge of the propagation delay from each BS to each of the MSs. We allow for perfect timing synchronization among cooperative BSs, which can be realized by GPS or by a wired backbone. Such infrastructure is already in place in current CDMA2000 and IS-95 cellular networks to facilitate soft handoffs [4, Chp. 18]. We first assume that the timing-advance mechanisms can ensure that the *desired signals* for an MS that are transmitted from multiple BSs reach the MS at exactly the same time. Such timing-advance mechanisms are employed currently in the uplink of GSM and 3G cellular networks [15].<sup>1</sup> (We shall relax this assumption in Sec. IV.)

Specifically, let the propagation delay from BS  $b$  to MS  $k$  be denoted by  $\tau_k^{(b)}$ , as illustrated in Fig. 1 for two BSs and two MSs. To guarantee simultaneous reception of  $\left\{ \mathbf{x}_k^{(b)}(m) \right\}_{b=1}^B$  at MS  $k$ , the BS  $b$  advances the time when  $\mathbf{x}_k^{(b)}(m)$  is transmitted by  $\Delta\tau_k^{(b)} = \tau_k^{(b)} - \tau_k^{(b_k)}$  so that  $\left\{ \mathbf{x}_k^{(b)}(m) \right\}_{b=1}^B$  all arrive at MS  $k$  with the same delay,  $\tau_k^{(b_k)}$ . The equivalent received baseband signal at MS  $k$  when a linear modulation with a unit energy baseband signature waveform  $g(t)$

<sup>1</sup>Note that the timing-advance values are typically much smaller than the packet durations. The one difference between the standardized set up and the one in this figure is that a BS now needs to track this for every UE it is transmitting to (even if the UE is in a different cell.)

of duration  $T_S$  is used is given by

$$\mathbf{r}_k(t) = \sum_{m=0}^{\infty} g(t - mT_S - \tau_k^{(b_k)}) \mathbf{H}_k \mathbf{x}_k(m) + \mathbf{n}_k(t) + \sum_{m=0}^{\infty} \left\{ \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b=1}^B g(t - mT_S - \tau_k^{(b)} + \Delta\tau_j^{(b)}) \mathbf{H}_k^{(b)} \mathbf{x}_j^{(b)}(m) \right\}, \quad (1)$$

where  $\mathbf{n}_k(t)$  is the additive white Gaussian noise vector,  $\mathbf{H}_k = [\mathbf{H}_k^{(1)}, \dots, \mathbf{H}_k^{(B)}]$ , and  $\mathbf{x}_k(m) = [\mathbf{x}_k^{(1)}(m)^\dagger, \dots, \mathbf{x}_k^{(B)}(m)^\dagger]^\dagger$ .

At MS  $k$ , the received signal at time  $t$ ,  $\mathbf{r}_k(t)$ , is passed through a filter matched to  $g(t - mT_S - \tau_k^{(b_k)})$  – which is also delayed by  $\tau_k^{(b_k)}$  – to generate the following sufficient statistic  $\mathbf{y}_k(m)$ :

$$\mathbf{y}_k(m) = \mathbf{H}_k \mathbf{T}_k \mathbf{s}_k(m) + \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b=1}^B \mathbf{H}_k^{(b)} \mathbf{T}_j^{(b)} \mathbf{i}_{jk}^{(b)}(m) + \mathbf{n}_k(m), \quad (2)$$

where  $\mathbf{T}_k = [\mathbf{T}_k^{(1)\dagger}, \dots, \mathbf{T}_k^{(B)\dagger}]^\dagger$ ,  $\mathbf{n}_k(m)$  is the discrete noise vector at the  $m$ th interval satisfying  $\mathbb{E}[\mathbf{n}_k(m)\mathbf{n}_k(m)^\dagger] = N_0 \mathbf{I}_{N_R}$ , and  $\mathbf{i}_{jk}^{(b)}(m)$  is the asynchronous interference at MS  $k$  from the signal transmitted by BS  $b$  for MS  $j$ . It depends on the difference,  $\tau_{jk}^{(b)}$ , between the timing-advances used by BS  $b$  for MSs  $j$  and  $k$ :

$$\tau_{jk}^{(b)} = (\tau_k^{(b)} - \Delta\tau_j^{(b)}) - \tau_k^{(b_k)} = \Delta\tau_k^{(b)} - \Delta\tau_j^{(b)}. \quad (3)$$

In (2), the asynchronous interference term at MS  $k$ ,  $\mathbf{i}_{jk}^{(b)}(m)$ , arises from two consecutive symbols, say with indices  $m_{jk}^{(b)}$  and  $m_{jk}^{(b)} + 1$ , that are transmitted to MS  $j$  from BS  $b$ . This is illustrated in Fig. 2. Let  $\delta_{jk}^{(b)} = \tau_{jk}^{(b)} \bmod T_S$ . Then,

$$\mathbf{i}_{jk}^{(b)}(m) = \rho(\delta_{jk}^{(b)} - T_S) \mathbf{s}_j(m_{jk}^{(b)}) + \rho(\delta_{jk}^{(b)}) \mathbf{s}_j(m_{jk}^{(b)} + 1), \quad (4)$$

where  $\rho(\tau) = \int_0^{T_S} g(t)g(t - \tau)dt$  with  $\rho(0) = 1$ .

Only if the asynchronous nature of interference is neglected, does (2) simplify to the following form used in [5]–[10]:

$$\begin{aligned} \mathbf{y}_k(m) &= \mathbf{H}_k \mathbf{T}_k \mathbf{s}_k(m) + \sum_{\substack{j=1 \\ (j \neq k)}}^K \left( \sum_{b=1}^B \mathbf{H}_k^{(b)} \mathbf{T}_j^{(b)} \right) \mathbf{s}_j(m) + \mathbf{n}_k(m), \\ &= \mathbf{H}_k \mathbf{T}_k \mathbf{s}_k(m) + \sum_{\substack{j=1 \\ (j \neq k)}}^K \mathbf{H}_k \mathbf{T}_j \mathbf{s}_j(m) + \mathbf{n}_k(m). \end{aligned} \quad (5)$$

## B. Statistics of Asynchronous Interference

We now derive the second-order statistics of the asynchronous interference, which will come in handy later.

From (4), we have  $\mathbb{E} [\mathbf{i}_{jk}^{(b)}(m)] = \mathbf{0}$ , for all  $j$ ,  $k$ , and  $b$ , and  $\mathbb{E} [\mathbf{i}_{j_1k}^{(b_1)}(m)\mathbf{i}_{j_2k}^{(b_2)}(m)^\dagger] = \mathbf{0}$ , for  $j_1 \neq j_2$ ,  $j_1 \neq k$ , and  $j_2 \neq k$ . It can be shown that for  $j \neq k$ , the correlation between  $\mathbf{i}_{jk}^{(b_1)}(m)$  and  $\mathbf{i}_{jk}^{(b_2)}(m)$  is

$$\mathbb{E} [\mathbf{i}_{jk}^{(b_1)}(m)\mathbf{i}_{jk}^{(b_2)}(m)^\dagger] = \beta_{jk}^{(b_1,b_2)} \mathbf{I}_{L_j}, \quad (6)$$

where the asynchronous interference correlation,  $\beta_{jk}^{(b_1,b_2)}$ , for  $j \neq k$ , has the following properties:

$$\beta_{jk}^{(b_1,b_2)} = \begin{cases} 0, & \text{if } |m_{jk}^{(b_2)} - m_{jk}^{(b_1)}| > 1 \\ \rho(\delta_{jk}^{(b_1)})\rho(\delta_{jk}^{(b_2)} - T_S), & \text{if } m_{jk}^{(b_2)} = m_{jk}^{(b_1)} + 1 \\ \rho(\delta_{jk}^{(b_1)})\rho(\delta_{jk}^{(b_2)}) + \rho(\delta_{jk}^{(b_1)} - T_S)\rho(\delta_{jk}^{(b_2)} - T_S), & \text{if } m_{jk}^{(b_2)} = m_{jk}^{(b_1)} \\ \rho(\delta_{jk}^{(b_2)})\rho(\delta_{jk}^{(b_1)} - T_S), & \text{if } m_{jk}^{(b_2)} = m_{jk}^{(b_1)} - 1 \end{cases}. \quad (7)$$

Also,  $\beta_{kk}^{(b_1,b_2)} = 1$ , for all BSs  $b_1$  and  $b_2$ . Since all the  $K$  users use the same waveform, the asynchronous interference correlation values corresponding to different timing parameters can be pre-calculated and stored in a look-up table.

## III. JOINT LINEAR PRECODING BY COOPERATIVE BASE STATIONS

Our goal is to jointly optimize the transmitter precoding matrices,  $\{\mathbf{T}_k\}_{k=1}^K$ , subject to set of MS-specific power constraints:

$$\text{Tr} \left\{ \mathbf{T}_k^\dagger \mathbf{T}_k \right\} \leq P_k^{\text{tx}}, \quad 1 \leq k \leq K. \quad (8)$$

An additional constraint,  $BN_T \geq \sum_{k=1}^K L_k$ , follows from dimensionality arguments. For notational simplicity, we drop the symbol index  $m$  henceforth.

*Note:* A uniform per-MS power constraint,  $P_k^{\text{tx}} = P_T$ , for all  $k$ , was also assumed in [6], [8], [16] to ensure “power fairness” for the different users. This MS-specific power constraint is different from the per-BS power constraint, which was used, for example, in [5], [10]. While the per-BS power criterion makes more physical sense, the advantage of the MS-specific power criterion is that it leads to analytically tractable solutions (for further discussion see [10]). Most importantly, other, more general power constraints can now be obtained numerically. For example, this can be done by an “outer loop” that adjusts  $P_k^{\text{tx}}$  iteratively until certain criteria such as per-BS

power constraints or MS-specific quality-of-service constraints are fulfilled. This also facilitates dynamic radio resource allocation.

As mentioned, determining the linear precoding matrices, even for the synchronous scenario, is a hard and computationally involved problem. Therefore, various design methods have been proposed in the literature to reduce the complexity of determining them. The nullification method [8], [10], which forces the precoding matrices to satisfy the constraint,  $\mathbf{H}_k \mathbf{T}_j = \mathbf{0}$ , for all  $k \neq j$ , is one such method. However, in the presence of asynchronous interference, this constraint can no longer annul all the interference terms in (2). Another option, as put forth in [17], is to force a stronger per-BS constraint  $\mathbf{H}_k^{(b)} \mathbf{T}_j^{(b)} = \mathbf{0}$ , for all pairs of  $k$  and  $j$  such that  $k \neq j$ . While this constraint does ensure that the interference (even the asynchronous one) gets completely canceled, it can be shown to support only  $K \leq N_T/N_R$  users, which is a severe and undesirable limitation.

Methods for selecting the precoding matrices, which strive to minimize CCI to the extent required, instead of canceling it out completely at the expense of severe transmit power inefficiency, have also been proposed. In this paper, we study three such methods (metrics) that have been previously proposed in the literature, and see how the asynchronous interference changes the corresponding optimal linear precoding solutions.

#### A. Design Methods and Metrics of Interest

The following three design methods have been considered for optimizing linear precoding in the literature:

1) *Overall Normalized Mean Square Error (NMSE)*: In this method, the goal is to optimize the transmitter precoders  $\{\mathbf{T}_k\}_{k=1}^K$  to minimize the overall MSE between a *desired* form of the signal and the received signal over all the  $K$  users. The functional form of the metric and solution for optimizing it are derived in Sec. III-B.

2) *Signal to Leakage plus Noise Ratio (SLNR)*: An alternative method considers the signal-to-leakage-plus-noise-ratio (SLNR). More precisely, for MS  $k$  the precoding matrix  $\mathbf{T}_k$  is designed to maximize the SLNR, which is the ratio of the power of the desired signal received at MS  $k$  and the sum of the noise and the total interference power (leakage) due to  $\mathbf{x}_k$  at other MSs. This approach minimizes the interference that *stems from* the data streams intended for one user instead of the interference that *arrives at* that MS. We note that while using the SLNR for

precoding design was first suggested in [13], [14], these papers only cover the simple case of one data stream per user, and do not model asynchronous interference. The solution for it is derived in Sec. III-C.

3) *Sum of Information Rates*: Arguably, the most relevant metric from a system-wide spectral efficiency standpoint is the sum of the information rates over all users that is achieved by the precoding designs [3], [5], [6], [10]. However, the optimization of the sum rate is a non-linear and non-convex problem even in the synchronous scenario, which makes it difficult to find analytical solutions. Brute-force numerical optimization involves searching over an extremely large space of dimension  $BN_T \sum_{k=1}^K L_k$ , and is practically infeasible. Therefore, we develop in Sec. III-D an alternate, albeit sub-optimal, algorithm to determine the precoding matrices.

The advantage of the first two methods is that they are amenable to analysis. While the discussion below highlights the intuitive basis for the form that their metrics take, it must be noted that the methods are essentially ad hoc in nature. We shall therefore compare the methods both on the basis of how they improve their respective metrics, and also how they improve the overall spectral efficiency of the system.

### B. Joint Wiener Filtering (JWF) to Minimize Overall Normalized MSE

In the JWF method, our aim is to optimize the transmitter precoders  $\{\mathbf{T}_k\}_{k=1}^K$  to minimize the overall MSE between the received signal and the ‘desired’ signal for all the  $K$  users. To mitigate interference, we strive to make the actual input signal at the receiver of MS  $k$  as close as possible to a desired (virtual) ‘cleaned’ signal that mimics a single-user MIMO environment that is free of multi-user interference (MUI) and noise.

In such a clean environment, the desired signal input to the receiver would be  $\mathbf{z}_k = \mathbf{H}_k \mathbf{V}_k \mathbf{s}_k$ , where the matrix  $\mathbf{V}_k$  is only determined by the composite channel  $\mathbf{H}_k$  and the power constraint in (8). The linear precoding matrix  $\mathbf{V}_k$  is taken to be the eigen-beamforming matrix with water-filling power allocation over the channel  $\mathbf{H}_k$  since it maximizes the information rate in the clean interference-free environment [4, Chp. 20]. The metric is also normalized so as to emphasize the contribution of all the users. Therefore, the overall normalized MSE (NMSE) metric gets defined as

$$\text{NMSE} = \sum_{k=1}^K \frac{\mathbb{E} [\|\mathbf{y}_k - \mathbf{z}_k\|^2]}{\Omega_k} = \sum_{k=1}^K \text{NMSE}_k, \quad (9)$$

where  $\text{NMSE}_k = \frac{\mathbb{E}[\|\mathbf{y}_k - \mathbf{z}_k\|^2]}{\Omega_k}$  is the NMSE of MS  $k$ ,  $\Omega_k = \mathbb{E}[\text{Tr}\{\mathbf{z}_k \mathbf{z}_k^\dagger\}] = \text{Tr}\{\mathbf{H}_k \mathbf{V}_k \mathbf{V}_k^\dagger \mathbf{H}_k^\dagger\}$  is the average received power of its ‘‘desired’’ signal, and the expectation is over the random data vectors,  $\{\mathbf{s}_k\}_{k=1}^K$ , and the noise,  $\{\mathbf{n}_k\}_{k=1}^K$ . Note that the optimization criterion defined here is generic enough to be independent of the receiver design. On the other hand, [12] defined a receiver post-processing MSE, in which the transmitter and receiver designs were both optimized.

The NMSE optimization problem is then

$$\begin{aligned} \{\mathbf{T}_k^{\text{opt}}\}_{k=1}^K &= \arg \min_{\{\mathbf{T}_k\}_{k=1}^K} \sum_{k=1}^K \text{NMSE}_k, \\ \text{s. t. } \quad \text{Tr}\{\mathbf{T}_k^\dagger \mathbf{T}_k\} &= \text{Tr}\left\{\sum_{b=1}^B \mathbf{T}_k^{(b)\dagger} \mathbf{T}_k^{(b)}\right\} \leq P_k^{\text{tx}}, \quad \text{for } k = 1, \dots, K. \end{aligned} \quad (10)$$

As shown in Appendix A, the following closed-form solution for the optimal linear precoding matrices follows:

$$\mathbf{T}_k = \frac{1}{\Omega_k} [\mathbf{C}_k + \kappa_k \mathbf{I}_{N_T B}]^{-1} \mathbf{H}_k^\dagger \mathbf{A}_k. \quad (11)$$

Here,  $\mathbf{A}_k = \mathbf{H}_k \mathbf{V}_k$  and  $\mathbf{C}_k = \begin{bmatrix} \mathbf{C}_k^{(1,1)} & \mathbf{C}_k^{(1,2)} & \dots & \mathbf{C}_k^{(1,B)} \\ \mathbf{C}_k^{(2,1)} & \mathbf{C}_k^{(2,2)} & \dots & \mathbf{C}_k^{(2,B)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_k^{(B,1)} & \mathbf{C}_k^{(B,2)} & \dots & \mathbf{C}_k^{(B,B)} \end{bmatrix}$ , where the sub-matrices  $\mathbf{C}_k^{(b_1, b_2)}$  are given by

$$\mathbf{C}_k^{(b_1, b_2)} = \sum_{j=1}^K \frac{\beta_{kj}^{(b_1, b_2)}}{\Omega_j} \mathbf{H}_j^{(b_1)\dagger} \mathbf{H}_j^{(b_2)} \quad (12)$$

And,  $\kappa_1, \dots, \kappa_K$  are the Lagrange multipliers that are chosen to meet the power constraints for MSs  $1, \dots, K$ , respectively.

### C. Joint Leakage Suppression (JLS) to Maximize SLNR

In the JLS method, for each MS  $k$ , we design the precoding matrices to maximize the ratio of the power of the desired signal received by it and the sum of the noise and the total interference power (leakage) due to  $\mathbf{x}_k$  at all the other MSs. We limit the search space to scaled semi-unitary matrices of the form  $\mathbf{T}_k = \sqrt{\frac{P_k^{\text{tx}}}{L_k}} \mathbf{Q}_k$ , where the columns of the  $N_T B \times L_k$  matrix  $\mathbf{Q}_k$  are orthonormal, i.e.,  $\mathbf{Q}_k^\dagger \mathbf{Q}_k = \mathbf{I}_{L_k}$ . While this limitation is sub-optimal, it makes the optimization below analytically feasible and does lead to a considerable performance improvement as orthonormality eliminates cross-talk among the data streams that an MS receives. (Note that

the power constraints are now trivially satisfied with equality.) Clearly, optimizing the linear precoding matrices to maximize SLNRs is decoupled for different MSs.

Therefore, the optimization problem is

$$\mathbf{Q}_k^{\text{opt}} = \arg \min_{\mathbf{Q}_k} \text{SLNR}_k, \quad 1 \leq k \leq K. \quad (13)$$

As derived in Appendix B, the expression for  $\text{SLNR}_k$  is given by

$$\text{SLNR}_k = \frac{\text{Tr} \left\{ \mathbf{Q}_k^\dagger \mathbf{M}_k \mathbf{Q}_k \right\}}{\text{Tr} \left\{ \mathbf{Q}_k^\dagger \mathbf{N}_k \mathbf{Q}_k \right\}} = \frac{\sum_{l=1}^{L_k} \mathbf{q}_{kl}^\dagger \mathbf{M}_k \mathbf{q}_{kl}}{\sum_{l=1}^{L_k} \mathbf{q}_{kl}^\dagger \mathbf{N}_k \mathbf{q}_{kl}}, \quad (14)$$

where  $\mathbf{q}_{kl}$  is the  $l$ th column of  $\mathbf{Q}_k$ ,  $\mathbf{M}_k = P_k^{\text{tx}} \mathbf{H}_k^\dagger \mathbf{H}_k$ ,  $\mathbf{N}_k = N_0 N_R \mathbf{I}_{BN_T} + \sum_{\substack{j=1 \\ (j \neq k)}}^K P_k^{\text{tx}} \mathbf{W}_{kj}$ , and

$$\mathbf{W}_{kj} = \begin{bmatrix} \beta_{kj}^{(1,1)} \mathbf{H}_j^{(1)\dagger} \mathbf{H}_j^{(1)} & \cdots & \beta_{kj}^{(1,B)} \mathbf{H}_j^{(1)\dagger} \mathbf{H}_j^{(B)} \\ \beta_{kj}^{(2,1)} \mathbf{H}_j^{(2)\dagger} \mathbf{H}_j^{(1)} & \cdots & \beta_{kj}^{(2,B)} \mathbf{H}_j^{(2)\dagger} \mathbf{H}_j^{(B)} \\ \vdots & \ddots & \vdots \\ \beta_{kj}^{(B,1)} \mathbf{H}_j^{(B)\dagger} \mathbf{H}_j^{(1)} & \cdots & \beta_{kj}^{(B,B)} \mathbf{H}_j^{(B)\dagger} \mathbf{H}_j^{(B)} \end{bmatrix}. \quad (15)$$

Despite the decoupling in optimization, finding the optimal  $\mathbf{q}_{kl}, \dots, \mathbf{q}_{kL_k}$  is still analytically intractable. We therefore derive and maximize the following lower bound for  $\text{SLNR}_k$ , which follows from (14):

$$\text{SLNR}_k \geq \min_{l=1, \dots, L_k} \frac{\mathbf{q}_{kl}^\dagger \mathbf{M}_k \mathbf{q}_{kl}}{\mathbf{q}_{kl}^\dagger \mathbf{N}_k \mathbf{q}_{kl}}. \quad (16)$$

Thus, maximizing the above lower bound on the SLNR for an MS maximizes the smallest generalized Rayleigh quotient among all its data streams. Therefore, the optimal precoding matrix is the solution to the following max-min problem:

$$\mathbf{Q}_k^{\text{opt}} = \arg \max_{\mathbf{Q}_k: \mathbf{Q}_k^\dagger \mathbf{Q}_k = \mathbf{I}_{L_k}} \min_{l=1, \dots, L_k} \frac{\mathbf{q}_{kl}^\dagger \mathbf{M}_k \mathbf{q}_{kl}}{\mathbf{q}_{kl}^\dagger \mathbf{N}_k \mathbf{q}_{kl}}. \quad (17)$$

The following Lemma reveals the structure of the optimal precoding matrix.

**Lemma 1:** The lower bound of  $\text{SLNR}_k$  in (16) is maximized when:

$$\mathbf{q}_{kl}^{\text{opt}} = \mathbf{v}_l(\mathbf{N}_k^{-1} \mathbf{M}_k), \quad \text{for } 1 \leq l \leq L_k, \quad (18)$$

where  $\mathbf{v}_l(\mathbf{N}_k^{-1} \mathbf{M}_k)$  denotes the eigenvector of the matrix  $\mathbf{N}_k^{-1} \mathbf{M}_k$  corresponding to its  $l$ th largest eigenvalue.

*Proof:* Since  $\mathbf{q}_{k1}, \dots, \mathbf{q}_{kL_k}$  are orthonormal vectors, the vector space  $\mathcal{V} = \text{span}\{\mathbf{q}_{k1}, \dots, \mathbf{q}_{kL_k}\}$  has a dimension  $\dim(\mathcal{V}) = L_k$ . Therefore, (17) can be written as

$$\max_{\mathbf{Q}_k: \mathbf{Q}_k^\dagger \mathbf{Q}_k = \mathbf{I}_{L_k}} \min_{l=1, \dots, L} \frac{\mathbf{q}_{kl}^\dagger \mathbf{M}_k \mathbf{q}_{kl}}{\mathbf{q}_{kl}^\dagger \mathbf{N}_k \mathbf{q}_{kl}} = \max_{\mathcal{V}: \dim(\mathcal{V})=L_k} \min_{\mathbf{q} \in \mathcal{V}} \frac{\mathbf{q}^\dagger \mathbf{M}_k \mathbf{q}}{\mathbf{q}^\dagger \mathbf{N}_k \mathbf{q}}. \quad (19)$$

Since  $\mathbf{M}_k$  is Hermitian and  $\mathbf{N}_k$  is positive-definite, the Courant-Fischer Max-Min Theorem [18, Chp. 4] applies. Therefore, the right hand side is maximized when  $\mathcal{V}$  has as its basis the eigenvectors that correspond to the  $L_k$  eigenvalues  $\{\lambda_1(\mathbf{N}_k^{-1} \mathbf{M}_k), \dots, \lambda_{L_k}(\mathbf{N}_k^{-1} \mathbf{M}_k)\}$ , where  $\lambda_i(\mathbf{N}_k^{-1} \mathbf{M}_k)$  is the  $i$ th largest eigenvalue of  $\mathbf{N}_k^{-1} \mathbf{M}_k$ . And, the maximum is achieved when (18) is satisfied. ■

It is interesting to note that the single closed-form solution in (18) is less complex than the JWF solution in Sec. III-B.

The scenario of  $L_k = 1$  in [14] is a special case of the above Lemma. Unlike the general  $L_k > 1$  case, the Rayleigh-Ritz quotient theorem [18, Chp. 4] can be directly applied to maximize the ratio in (14) for  $L_k = 1$ .

#### D. Controlled Iterative Singular Value Decomposition (CISVD) to Maximize Sum Rate

In the CISVD method, we strive to maximize the sum rate over all users. The optimization problem is then:

$$\begin{aligned} \{\mathbf{T}_k^{\text{opt}}\}_{k=1}^K &= \arg \max_{\{\mathbf{T}_k\}_{k=1}^K} \sum_{k=1}^K R_k, \\ \text{s. t. } \quad \text{Tr}\{\mathbf{T}_k^\dagger \mathbf{T}_k\} &\leq P_k^{\text{tx}}, \quad \text{for } k = 1, \dots, K. \end{aligned} \quad (20)$$

From (2), the bandwidth-normalized information rate,  $R_k$ , of MS  $k$  is given by [6][11]

$$R_k = \log \left| \mathbf{I}_{N_R} + \mathbf{\Phi}_k^{-1} \mathbf{H}_k \mathbf{T}_k \mathbf{T}_k^\dagger \mathbf{H}_k^\dagger \right|, \quad (21)$$

where  $\mathbf{\Phi}_k$  is the covariance of noise plus interference for MS  $k$ .<sup>2</sup> It is given by

$$\mathbf{\Phi}_k = N_0 \mathbf{I}_{N_R} + \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b_1=1}^B \sum_{b_2=1}^B \beta_{jk}^{(b_1, b_2)} \mathbf{H}_k^{(b_1)} \mathbf{T}_j^{(b_1)} \mathbf{T}_j^{(b_2)\dagger} \mathbf{H}_k^{(b_2)\dagger}.$$

Given the non-linear and non-convex nature of the problem, we propose an iterative optimization ‘‘hill-climbing’’ algorithm to maximize the spectral efficiency. In each step we optimize the

<sup>2</sup>Treating the asynchronous interference term as noise is, in effect, a lower bound on the information rate [19].

precoding matrix for MS  $k$ ,  $\mathbf{T}_k$ , by keeping the other precoding matrices,  $\mathbf{T}_j$  ( $j \neq k$ ) fixed. The optimal  $\mathbf{T}_k$  is then obviously the water-filling power allocation on the equivalent MIMO channel  $\Phi_k^{-1/2}\mathbf{H}_k$  with unit additive noise power. The iterations are initialized using the JLS solution in (18) given its simplicity, and are continued only when the target sum-rate increases by at least a certain threshold amount.

The pseudo-code for the algorithm is as follows:

- 1) For  $k = 1, \dots, K$ , calculate  $\mathbf{T}_1, \dots, \mathbf{T}_K$  from (18), i.e., use JLS solution as the starting point.
- 2) For each  $k = 1, \dots, K$ , fix  $\mathbf{T}_j$  ( $j \neq k$ ) and update  $\mathbf{T}_k$  to the eigen-beamforming and water-filling power allocation for the MIMO channel  $\Phi_k^{-1/2}\mathbf{H}_k$  (with unit noise power).
- 3) Repeat previous step until the sum rate target function in (20) increases by less than a pre-defined threshold.

Compared with random or exhaustive search algorithms, this method iteratively optimizes one precoder in each step to improve the corresponding MS's performance, while maintaining a relatively low level of interference imposed on other users. (Otherwise, the iteration terminates.) While this procedure is simple and sub-optimal, we shall see that it provides good results. This algorithm falls under the general class of greedy "alternate & maximize" algorithms, e.g., [20], and is similar to the iterative water-filling algorithm in [21], which dealt with the sum rate over different orthogonal sub-carriers in DSL systems with cross-talk.

#### IV. GENERALIZATION TO IMPERFECT TIMING-ADVANCE CASE

The three joint BS precoding designs of the previous section were derived assuming perfect timing-advance, i.e., the desired signal components are assumed to arrive synchronously. In practical systems, imperfect timing-advance (jitter) is inevitable because of imperfect delay estimation, user mobility, inaccurate cross-BS synchronization, time synchronization granularity, and MS synchronization errors. As we shall see, the jitter affects both the desired signal and the asynchronicity of the interference. We now extend the three design methods to cover this case, as well.

Let  $\epsilon_j^{(b)}$  denote the timing-advance error (jitter) of BS  $b$  in transmitting the signal for MS  $j$ . Therefore, the BS  $b$  now advances the time for transmitting the signal,  $\mathbf{x}_j^{(b)}$ , by  $\Delta\tilde{\tau}_j^{(b)} =$

$\Delta\tau_j^{(b)} + \epsilon_j^{(b)}$ . At the MS  $k$ , the received signal at time  $t$ ,  $\mathbf{r}_k(t)$ , now takes the form

$$\begin{aligned} \mathbf{r}_k(t) = & \sum_{m=0}^{\infty} \sum_{b=1}^B g(t - mT_S - \tau_k^{(b_k)} + \epsilon_k^{(b)}) \mathbf{H}_k^{(b)} \mathbf{x}_k^{(b)}(m) + \mathbf{n}_k(t) \\ & + \sum_{m=0}^{\infty} \left\{ \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b=1}^B g(t - mT_S - \tau_k^{(b)} + \Delta\tilde{\tau}_j^{(b)}) \mathbf{H}_k^{(b)} \mathbf{x}_j^{(b)}(m) \right\}. \end{aligned} \quad (22)$$

After passing it through a filter matched to  $g(t - mT_S - \tau_k^{(b_k)})$ , the sufficient statistic gets modified to

$$\begin{aligned} \tilde{\mathbf{y}}_k(m) = & \sum_{b=1}^B \gamma_k^{(b)} \mathbf{H}_k^{(b)} \mathbf{T}_k^{(b)} \mathbf{s}_k(m) + \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b=1}^B \mathbf{H}_k^{(b)} \mathbf{T}_j^{(b)} \tilde{\mathbf{i}}_{jk}^{(b)}(m) + \sum_{b=1}^B \alpha_k^{(b)} \mathbf{H}_k^{(b)} \mathbf{T}_k^{(b)} \mathbf{s}_k(m_k^{(b)}) + \mathbf{n}_k(m), \\ = & \mathbf{H}_k \mathbf{\Gamma}_k \mathbf{T}_k \mathbf{s}_k(m) + \mathbf{J}_k(m) + \mathbf{O}_k(m) + \mathbf{n}_k, \end{aligned} \quad (23)$$

where  $\gamma_k^{(b)} = \rho(\epsilon_k^{(b)}) \leq 1$ ,  $\mathbf{O}_k(m) = \sum_{b=1}^B \alpha_k^{(b)} \mathbf{H}_k^{(b)} \mathbf{T}_j^{(b)} \mathbf{s}_k(m_k^{(b)})$  is the new ISI term with  $\alpha_k^{(b)}$  and  $m_k^{(b)}$  jointly given by

$$\left( \alpha_k^{(b)}, m_k^{(b)} \right) = \begin{cases} \left( \rho(T_S - \epsilon_k^{(b)}), m + 1 \right), & \text{if } \epsilon_k^{(b)} > 0 \\ \left( \rho(T_S + \epsilon_k^{(b)}), m - 1 \right), & \text{if } \epsilon_k^{(b)} < 0 \\ (0, m), & \text{if } \epsilon_k^{(b)} = 0 \end{cases}. \quad (24)$$

Here,  $\mathbf{\Gamma}_k = \text{blockdiag} \left\{ \gamma_k^{(1)} \mathbf{I}_{N_T}, \dots, \gamma_k^{(B)} \mathbf{I}_{N_T} \right\}$  is the power degradation due to imperfect timing-advance (it equals  $\mathbf{I}_{BN_T}$  for perfect timing-advance). As before, the asynchronous interference is  $\tilde{\mathbf{i}}_{jk}^{(b)}(m) = \rho(\tilde{\delta}_{jk}^{(b)} - T_S) \mathbf{s}_j(\tilde{m}_{jk}^{(b)}) + \rho(\tilde{\delta}_{jk}^{(b)}) \mathbf{s}_j(\tilde{m}_{jk}^{(b)} + 1)$ , where  $\tilde{m}_{jk}^{(b)}$  and  $\tilde{m}_{jk}^{(b)} + 1$  are the two consecutive symbols transmitted from BS  $b$  for MS  $j$  that overlap with the  $m$ th symbol of MS  $k$ ,  $\tilde{\delta}_{jk}^{(b)} = \tilde{\tau}_{jk}^{(b)} \bmod T_S$ , and  $\tilde{\tau}_{jk}^{(b)} = \Delta\tilde{\tau}_k^{(b)} - \Delta\tilde{\tau}_j^{(b)}$  is the difference between the timing-advances used by BS  $b$  for MSs  $j$  and  $k$ . The modified asynchronous interference coefficients  $\tilde{\beta}_{jk}^{(b_1, b_2)}$  are determined by  $\tilde{\tau}_{jk}^{(b)}$  in the same way that  $\beta_{jk}^{(b_1, b_2)}$  was determined by  $\tau_{jk}^{(b)}$  in (7).

Thus, imperfect timing-advance degrades performance in three ways: through the power degradation term  $\mathbf{\Gamma}_k$ , the additional ISI term,  $\mathbf{O}_k$ , and the imperfect knowledge of  $\tilde{\beta}_{jk}^{(b_1, b_2)}$ . Given that the errors are unknown,  $\tilde{\beta}_{jk}^{(b_1, b_2)}$  or  $\gamma_k^{(b)}$  are also unknown when determining the optimal linear precoding matrices. However, its statistics can certainly be determined and exploited in the designs, as done below.

### A. Modified Linear Precoding Designs

1) *Modified JWF Design Method:* The aim is now to minimize the jitter-averaged NMSE error. The new JWF design that incorporates the timing inaccuracy and asynchronous interference is derived in Appendix C. The modified joint precoder of MS  $k$  takes the form:

$$\mathbf{T}_k = \frac{1}{\Omega_k} [\bar{\mathbf{C}}_k + \kappa_k \mathbf{I}_{N_TB}]^{-1} \bar{\mathbf{\Gamma}}_k \mathbf{H}_k^\dagger \mathbf{A}_k, \quad (25)$$

where  $\bar{\mathbf{\Gamma}}_k = \text{blockdiag} \left\{ \bar{\gamma}_k^{(1)} \mathbf{I}_{N_T}, \dots, \bar{\gamma}_k^{(B)} \mathbf{I}_{N_T} \right\}$  and  $\bar{\gamma}_k^{(b)} = \mathbb{E}_\epsilon \left[ \rho(\epsilon_k^{(b)}) \right]$ , and  $\mathbb{E}_\epsilon [\cdot]$  denotes averaging over the jitter statistics. The matrix  $\bar{\mathbf{C}}_k$  has as its sub-matrices:

$$\bar{\mathbf{C}}_k^{(b_1, b_2)} = \frac{1}{\Omega_k} \left( \mathbb{E}_\epsilon \left[ \alpha_k^{(b_1)} \alpha_k^{(b_2)} \right] p_k^{(b_1, b_2)} + \mathbb{E}_\epsilon \left[ \gamma_k^{(b_1)} \gamma_k^{(b_2)} \right] \right) \mathbf{H}_k^{(b_1)\dagger} \mathbf{H}_k^{(b_2)} + \sum_{\substack{j=1 \\ (j \neq k)}}^K \frac{\bar{\beta}_{kj}^{(b_1, b_2)}}{\Omega_j} \mathbf{H}_j^{(b_1)\dagger} \mathbf{H}_j^{(b_2)}, \quad (26)$$

where  $p_k^{(b_1, b_2)} = \Pr(\text{sgn}(\epsilon_k^{(b_1)}) = \text{sgn}(\epsilon_k^{(b_2)}))$  and  $\bar{\beta}_{kj}^{(b_1, b_2)} = \mathbb{E}_\epsilon \left[ \tilde{\beta}_{kj}^{(b_1, b_2)} \right]$ .

Given the knowledge of jitter statistics, the BSs can pre-calculate the jitter-averaged asynchronous leakage  $\bar{\beta}_{kj}^{(b_1, b_2)}$ . Determining the first moment of  $\tilde{\beta}_{kj}^{(b_1, b_2)}$  is difficult due to the modulo  $T_S$  operation on  $\tilde{\tau}_{kj}^{(b)}$  (since  $\tilde{\delta}_{kj}^{(b)} = \tilde{\tau}_{kj}^{(b)} \bmod T_S$ ). However, the jitters are typically considerably smaller than the symbol duration. Therefore, we can assume that the symbol index differences do not change due to jitter, i.e.,  $\tilde{m}_{kj}^{(b)} = m_{kj}^{(b)}$ . We then have  $\tilde{\delta}_{kj}^{(b)} \approx \delta_{kj}^{(b)} + \epsilon_k^{(b)}$ , which simplifies the calculation of  $\bar{\beta}_{kj}^{(b_1, b_2)}$ .

2) *Modified JLS Design Method:* The aim of the modified method is to maximize the jitter-averaged signal to jitter-averaged leakage plus noise ratio. To determine the optimal solution, Lemma 1 still holds, but with the following modified expressions for  $\mathbf{M}_k$  and  $\mathbf{N}_k$  (denoted by  $\tilde{\mathbf{M}}_k$  and  $\tilde{\mathbf{N}}_k$ , respectively):

$$\tilde{\mathbf{M}}_k = P_k^{\text{tx}} \begin{bmatrix} \mathbb{E}_\epsilon \left[ \gamma_k^{(1)2} \right] \mathbf{H}_k^{(1)\dagger} \mathbf{H}_k^{(1)} & \dots & \mathbb{E}_\epsilon \left[ \gamma_k^{(1)} \gamma_k^{(B)} \right] \mathbf{H}_k^{(1)\dagger} \mathbf{H}_k^{(B)} \\ \vdots & \ddots & \vdots \\ \mathbb{E}_\epsilon \left[ \gamma_k^{(B)} \gamma_k^{(1)} \right] \mathbf{H}_k^{(B)\dagger} \mathbf{H}_k^{(1)} & \dots & \mathbb{E}_\epsilon \left[ \gamma_k^{(B)2} \right] \mathbf{H}_k^{(B)\dagger} \mathbf{H}_k^{(B)} \end{bmatrix}, \quad (27)$$

$$\tilde{\mathbf{N}}_k = N_0 N_R \mathbf{I}_{BN_T} + \sum_{j=1}^K P_k^{\text{tx}} \tilde{\mathbf{W}}_{kj}, \quad (28)$$

where  $\tilde{\mathbf{W}}_{kj}$  bears the same form as  $\mathbf{W}_{kj}$  in (15), with  $\bar{\beta}_{kj}^{(b_1, b_2)}$  replacing  $\beta_{kj}^{(b_1, b_2)}$ , for  $j \neq k$ , and  $\mathbb{E}_\epsilon \left[ \alpha_k^{(b_1)} \alpha_k^{(b_2)} \right] p_k^{(b_1, b_2)}$  replacing  $\beta_{kk}^{(b_1, b_2)}$  when  $j = k$ . The modified JLS design is derived in Appendix D.

3) *Modified CISVD Design Method*: The aim is to maximize the jitter-averaged sum rate. The jitter-averaged information rate of MS  $k$  is now  $R_k = \mathbb{E}_\epsilon \left[ \log_2 \left| \mathbf{I}_{N_R} + \tilde{\Phi}_k^{-1} \mathbf{H}_k \mathbf{\Gamma}_k \mathbf{T}_k \mathbf{T}_k^\dagger \mathbf{\Gamma}_k^\dagger \mathbf{H}_k^\dagger \right| \right]$ , where the covariance of the noise plus interference terms, for a given jitter, takes the form

$$\begin{aligned} \tilde{\Phi}_k = N_0 \mathbf{I}_{N_R} + \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b_1=1}^B \sum_{b_2=1}^B \tilde{\beta}_{jk}^{(b_1, b_2)} \mathbf{H}_k^{(b_1)} \mathbf{T}_j^{(b_1)} \mathbf{T}_j^{(b_2)\dagger} \mathbf{H}_k^{(b_2)\dagger} \\ + \sum_{b_1=1}^B \sum_{b_2=1}^B \alpha_k^{(b_1)} \alpha_k^{(b_2)} \mathbf{H}_k^{(b_1)} \mathbf{T}_k^{(b_1)} \mathbf{T}_k^{(b_2)\dagger} \mathbf{H}_k^{(b_2)\dagger} 1 \left[ \text{sgn}(\epsilon_k^{(b_1)}) - \text{sgn}(\epsilon_k^{(b_2)}) \right]. \end{aligned} \quad (29)$$

Here, the indicator function  $1[\cdot]$  equals 1 if the input argument is 0, and equals 0 otherwise.<sup>3</sup> It must be noted that as no closed-form exists for the jitter-averaged sum rate, it needs to be evaluated using Monte Carlo simulations. This makes the computational complexity of the modified CISVD algorithm burdensome. A suboptimal solution would be to run the CISVD iterations with  $\tilde{\Phi}_k$  being replaced by  $\mathbb{E}_\epsilon \left[ \tilde{\Phi}_k \right]$  and  $\mathbf{H}_k$  being replaced by  $\mathbf{H}_k \mathbb{E}_\epsilon \left[ \mathbf{T}_k \right]$  in (21).

## V. NUMERICAL RESULTS

We simulate the downlink of urban micro-cellular network that consists of two or three cells, each with 1 BS and 1 MS. The inter-BS distance is 500 m. As BS cooperation results in performance gains when the signal from one BS does not completely dominate the signal from the other BS, we consider scenarios in which the MSs are uniformly distributed in a limited cell area so that any MS is at least 150m from its nearest BS. (This is shown as the shaded area in Fig. 1 for the 2 cell case.) The path-loss coefficient for all the BS-MS channels is 2.0 (free-space propagation) up to a distance of 30 m, and increases to 3.7 thereafter. Without loss of generality, the channel path-loss values are normalized with respect to the largest in-cell path-loss in the cell. The symbol pulse is rectangular and has a duration,  $T_S$ , of 1  $\mu$ s. In all the considered scenarios we assume that  $L_k = L = 2$  and  $P_k^{\text{tx}} = P^{\text{tx}}$  for any MS  $k$ , and  $N_T = 3$  and  $N_R = 2$ .

### A. Perfect Timing-Advance

Figure 3 considers the average NMSE per user (in linear scale) and compares JWF when it takes the asynchronicity of the interference into account and when it incorrectly ignores it

<sup>3</sup>It arises because of the term  $\mathbb{E}_\epsilon \left[ \mathbf{s}_k(m_k^{(b_1)}) \mathbf{s}_k(m_k^{(b_2)})^\dagger \right]$  in  $\mathbb{E}_\epsilon \left[ \mathbf{O}_k \mathbf{O}_k^\dagger \right]$ .  $\mathbb{E}_\epsilon \left[ \mathbf{s}_k(m_k^{(b_1)}) \mathbf{s}_k(m_k^{(b_2)})^\dagger \right]$  equals  $\mathbf{I}_{N_T}$ , if  $m_k^{(b_1)} = m_k^{(b_2)}$ , and equals  $\mathbf{0}$ , if  $m_k^{(b_1)} \neq m_k^{(b_2)}$ . From (24),  $m_k^{(b_1)} = m_k^{(b_2)}$  if and only if  $\text{sgn}(\epsilon_k^{(b_1)}) = \text{sgn}(\epsilon_k^{(b_2)})$ .

despite it being present. Also shown is the NMSE achieved by conventional nullification [8]–[10]. Accounting for asynchronous interference significantly improves the NMSE of JWF at all SNRs.

Figure 4 compares the average SLNR per user for JLS when it takes the asynchronicity of the interference into account and when it incorrectly ignores it. Also shown is the SLNR achieved by conventional nullification. Accounting for the asynchronous interference significantly improves the SLNR achieved by JLS; ignoring it reduces JLS’s performance to that of conventional nullification.

Figure 5 compares the sum rate achieved by CISVD when it takes asynchronous interference into account and when it incorrectly ignores it. As the sum rate is arguably the most relevant metric to compare the various BS cooperation precoding schemes, the sum rates achieved by JWF and JLS is also shown for comparison purposes. For all the three designs, accounting for the asynchronous interference improves performance by, for example, about 1 bps/Hz at an SNR of 10 dB. In our simulations, we have observed that CISVD typically terminates within 5 to 7 iterations. Its initial starting point also has an impact on its performance: it is preferable to use as the starting point the JLS solution instead of a random choice.

The relative sum rate performance of the three methods is delved into further in Figs. 6 and 7. These figures, which are from simulations of a larger 3 BS, 3 MS system, compare the sum rates of JWF, JLS, and CISVD with the following benchmarks: (i) conventional eigen-beamforming, where an MS’s signal is transmitted only by its serving BS, which treats all interference as additive noise [1], [3], [10]; (ii) ideal point-to-point MIMO in an interference-free single cell, and (iii) conventional nullification, which ignores the asynchronous interference. Figure 6 pertains to the case when asynchronous interference is present, while Fig. 7 pertains to the idealized ‘single giant BS’ scenario when asynchronous interference is absent.

As expected, in both scenarios, MIMO in a single interference-free cell achieves the highest rate, while eigen-beamforming that treats interference as noise delivers the lowest rate. All the three methods outperform conventional nullification because they account for asynchronous interference. Of the three methods, CISVD has the best performance at all the SNRs. At low to medium SNRs, CISVD even outperforms single cell interference-free point-to-point MIMO. JWF outperforms JLS at all SNRs. This is aligned with the observations we made earlier for NMSE (Fig. 3) and SLNR (Fig. 4). Despite its simplicity, JLS outperforms conventional nullification.

Note that when redundant spatial dimensions for diversity are available, i.e., when  $N_T B > \sum_{k=1}^K L_k$ , as in our simulated scenario, the proposed BS cooperative precoding designs achieve considerable gains. As we can see in Fig. 6, over a large SNR range, BS cooperation using CISVD even outperforms single-cell interference-free MIMO.

### B. Imperfect Timing-Advance

Figure 8 considers imperfect timing-advance (in addition to asynchronous interference) and the performance of the modified JWF and JLS methods when they compensate for it using its statistics.<sup>4</sup> Each BS's timing-advance jitter is taken to be uniformly distributed in the interval  $[-0.1T_S, 0.1T_S]$ , and is independent of the jitters of other BSs. The figure shows that JLS's performance is better than that of JWF in the presence of jitter. At an SNR of 15 dB, modifying the JWF design leads to a marginal improvement in its sum rate by 0.3 bps/Hz, modifying the JLS design significantly improves its sum rate by 1.6 bps/Hz. (The performance of the conventional nullification method is not shown here as timing-advance inaccuracy is not modeled by it.)

Not shown here, due to space constraints, are results with other parameter settings. We have observed that the JWF and CISVD methods perform best when redundant spatial dimensions are present ( $N_T B > \sum_{k=1}^K L_k$ ), while JLS is the simplest method.

## VI. CONCLUSIONS AND DISCUSSION

In this paper, we investigated the impact of asynchronous interference on the downlink performance of MIMO systems with BS cooperation. We developed a detailed mathematical model that showed that when cooperative BSs jointly transmit to multiple users, the data streams intended for the multiple users inevitably interfere asynchronously with each other. This is so even when perfect timing-advance is used to synchronize the reception of the desired signal components. We looked at three linear precoding design methods, previously considered in the literature, and came up with three corresponding new precoding methods – JWF, JLS, and CISVD. All the three markedly outperformed conventional methods that did not account for the asynchronous nature of the interference. CISVD and JWF realized significant gains in spectral efficiency, while JLS achieved a good trade-off between asynchronous interference mitigation and

<sup>4</sup>The performance of modified CISVD is not shown given its complexity.

algorithmic complexity. All three methods perform well in the channels with redundant spatial dimensions. Essentially, the paper moves a step closer to realizing the great potential of BS cooperation in practical implementations of interference-limited multi-user MIMO systems. Part of our future work involves extending the analysis of this paper to directly optimize the linear precoding designs based on a per-BS power constraint. Also, exploiting the correlation between the interference observed in adjacent symbols can lead to further improvement in performance. Lower complexity solutions that can enable cooperation between more base stations are also another area of interest.

While this paper focused on single carrier communication over flat-fading channels, extending it to delay-dispersive (frequency-selective) channels and, in particular, the popular Orthogonal-Frequency-Division-Multiplexing (OFDM) systems, is of great interest. In frequency-selective channels, the BS cooperation schemes need to mitigate both multi-user interference (MUI) and inter-symbol interference (ISI). In OFDM systems, asynchronous interference can be greatly alleviated by accommodating the delay offset in the cyclic prefix (CP).<sup>5</sup> So long as the CP prefix is longer than the sum of the maximum delay offset and the channel delay dispersion, BS cooperation can be implemented on a sub-carrier basis using the approaches in [5]–[10]. However, increasing the CP length reduces the spectral efficiency of the system. Furthermore, in cellular systems with large cell sizes, a sufficiently long CP might not be practical. In this case, MUI, ISI, and inter-carrier interference (ICI) will occur simultaneously and will need to be tackled by the BS cooperation algorithms. (ICI occurs because MUI – from other subcarriers – is not orthogonal to the desired signals.). This results in a much more complicated system model; optimizing it is a subject to future research.

## APPENDIX

### A. Optimal Linear Precoding for JWF: Derivation

Denoting the MUI term in (2) as  $\mathbf{J}_k = \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b=1}^B \mathbf{H}_k^{(b)} \mathbf{T}_j^{(b)} \mathbf{i}_{jk}^{(b)}(m)$ ,  $\text{NMSE}_k$  takes the form

$$\text{NMSE}_k = \frac{1}{\Omega_k} \mathbf{E} \left[ (\mathbf{H}_k \mathbf{T}_k \mathbf{s}_k - \mathbf{A}_k \mathbf{s}_k + \mathbf{J}_k + \mathbf{n}_k)^\dagger (\mathbf{H}_k \mathbf{T}_k \mathbf{s}_k - \mathbf{A}_k \mathbf{s}_k + \mathbf{J}_k + \mathbf{n}_k) \right]. \quad (30)$$

<sup>5</sup>We differentiate between orthogonal-frequency-division-multiple-access (OFDMA) systems and OFDM systems. In an OFDMA downlink, different users are transmitted to using different small frequency chunks. On the other hand, in an OFDM system, each downlink transmission is to one user, it occupies the entire bandwidth, and it uses OFDM.

Using the results in Sec. II-B, the above equation, after considerable simplification, becomes

$$\begin{aligned} \text{NMSE}_k &= \frac{1}{\Omega_k} \text{Tr} \left\{ \sum_{b_1=1}^B \sum_{b_2=1}^B \mathbf{H}_k^{(b_1)} \mathbf{T}_k^{(b_1)} \mathbf{T}_k^{(b_2)\dagger} \mathbf{H}_k^{(b_2)\dagger} - \sum_{b=1}^B \mathbf{H}_k^{(b)} \mathbf{T}_k^{(b)} \mathbf{A}_k^\dagger \right\} + \frac{N_0 N_R}{\Omega_k} \\ &+ \frac{1}{\Omega_k} \text{Tr} \left\{ -\mathbf{A}_k \sum_{b=1}^B \mathbf{T}_k^{(b)\dagger} \mathbf{H}_k^{(b)\dagger} + \mathbf{A}_k \mathbf{A}_k^\dagger \right\} + \frac{1}{\Omega_k} \text{Tr} \left\{ \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b_1=1}^B \sum_{b_2=1}^B \beta_{jk}^{(b_1, b_2)} \mathbf{H}_k^{(b_1)} \mathbf{T}_j^{(b_1)} \mathbf{T}_j^{(b_2)\dagger} \mathbf{H}_k^{(b_2)\dagger} \right\}, \end{aligned}$$

where we used the identity  $\mathbb{E} [\mathbf{J}_k^\dagger \mathbf{J}_k] = \text{Tr} \left\{ \mathbb{E} [\mathbf{J}_k \mathbf{J}_k^\dagger] \right\}$ .

To solve (10) in closed-form, we minimize the following Lagrange objective function:

$$f \left( \{\mathbf{T}_k\}_{k=1}^K \right) = \sum_{k=1}^K \text{NMSE}_k + \sum_{k=1}^K \kappa_k \left( \text{Tr} \left\{ \sum_{b=1}^B \mathbf{T}_k^{(b)\dagger} \mathbf{T}_k^{(b)} \right\} - P_k^{\text{tx}} \right), \quad (31)$$

where  $\kappa_k$  are the Lagrange multipliers associated with the power constraints for MSs  $1, \dots, K$ , respectively. Equating to zero the general complex derivative of  $f$  with respect to the matrix  $\mathbf{T}_k^{(b)}$  results in<sup>6</sup>

$$\frac{\partial f}{\partial \mathbf{T}_k^{(b)}} = \sum_{j=1}^K \frac{1}{\Omega_j} \sum_{b_1=1}^B \beta_{kj}^{(b, b_1)} \mathbf{T}_k^{(b_1)\dagger} \mathbf{H}_j^{(b_1)\dagger} \mathbf{H}_j^{(b)} - \frac{1}{\Omega_k} \mathbf{A}_k^\dagger \mathbf{H}_k^{(b)} + \kappa_k \mathbf{T}_k^{(b)\dagger} = 0, \text{ for } b = 1, \dots, B. \quad (32)$$

Doing so leads to  $\mathbf{C}_k \mathbf{T}_k + \kappa_k \mathbf{T}_k = \frac{1}{\Omega_k} \mathbf{H}_k^\dagger \mathbf{A}_k$ , and thus (11). Substituting this solution in the power constraint shows that  $\kappa_k$  is one of the roots of the equation  $\sum_{i=1}^{N_T B} \frac{b_{ki}}{(x + \lambda_{ki})^2} = P_k^{\text{tx}} \Omega_k^2$ , where  $b_{ki} = \left[ \mathbf{U}_k^\dagger \mathbf{H}_k^\dagger \mathbf{A}_k \mathbf{A}_k^\dagger \mathbf{H}_k \mathbf{U}_k \right]_{ii}$ ,  $\mathbf{C}_k$  has the eigenvalue decomposition  $\mathbf{U}_k \mathbf{\Lambda}_k \mathbf{U}_k^\dagger$ , and  $\mathbf{\Lambda}_k = \text{diag} \{ \lambda_{k1}, \lambda_{k2}, \dots, \lambda_{k(N_T B)} \}$ . Since these equations may have multiple solutions,  $\kappa_1, \dots, \kappa_K$  are jointly chosen to minimize the overall NMSE. Note that the optimal solution satisfies power constraints with equality at all the MSs.

### B. Derivation of SLNR expression

The desired signal component received by MS  $k$  is  $\mathbf{x}_k = \sqrt{\frac{P_k^{\text{tx}}}{L_k}} \mathbf{H}_k \mathbf{Q}_k \mathbf{s}_k$ , and has a power  $P_k^{\text{rx}} = \frac{P_k^{\text{tx}}}{L_k} \text{Tr} \left\{ \mathbf{Q}_k^\dagger \mathbf{H}_k^\dagger \mathbf{H}_k \mathbf{Q}_k \right\}$ . From (2), the (asynchronous) interference leakage at MS  $j$  from

<sup>6</sup>We use the following definition of the matrix derivative:  $\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}} = \left[ \frac{\partial f}{\partial x_{ji}} \right]$ . From this it follows that,  $\frac{\partial \text{Tr}\{\mathbf{A}\mathbf{X}\}}{\partial \mathbf{X}} = \mathbf{A}$  and  $\frac{\partial \text{Tr}\{\mathbf{A}\mathbf{X}\mathbf{B}\mathbf{X}^\dagger\}}{\partial \mathbf{X}} = \mathbf{B}\mathbf{X}^\dagger \mathbf{A}$ .

the signal  $\mathbf{x}_k$  (meant for MS  $k$ ) that is transmitted by the  $B$  BSs is  $\sum_{b=1}^B \mathbf{H}_j^{(b)} \mathbf{T}_k^{(b)} \mathbf{i}_{kj}^{(b)}$ . Its power,  $P_{kj}^{\text{leak}}$ , is therefore given by

$$P_{kj}^{\text{leak}} = \frac{P_k^{\text{tx}}}{L_k} \sum_{b_1=1}^B \sum_{b_2=1}^B \beta_{kj}^{(b_1, b_2)} \text{Tr} \left\{ \mathbf{Q}_k^{(b_1)\dagger} \mathbf{H}_j^{(b_1)\dagger} \mathbf{H}_j^{(b_2)} \mathbf{Q}_k^{(b_2)} \right\}. \quad (33)$$

Here, the sub-matrix  $\mathbf{Q}_k^{(b)}$  collects the rows in the precoder  $\mathbf{Q}_k$  associated with the  $b$ th BS. Finally, the noise power at MS  $k$  is  $N_0 N_R$ . Therefore,  $\text{SLNR}_k = \frac{P_k^{\text{tx}}}{\sum_{\substack{j=1 \\ (j \neq k)}}^K P_{kj}^{\text{leak}} + N_0 N_R}$ , takes the form in (14).

### C. Modified JWF with Imperfect Timing-Advance

From (23), the mean value of  $\text{NMSE}_k$ , obtained by averaging over  $\epsilon_k^{(b)}$ , can be shown to be:

$$\begin{aligned} \mathbb{E}_\epsilon [\text{NMSE}_k] &= \frac{1}{\Omega_k} \text{Tr} \left\{ \sum_{b_1=1}^B \sum_{b_2=1}^B \mathbb{E}_\epsilon \left[ \gamma_k^{(b_1)} \gamma_k^{(b_2)} \right] \mathbf{H}_k^{(b_1)} \mathbf{T}_k^{(b_1)} \mathbf{T}_k^{(b_2)\dagger} \mathbf{H}_k^{(b_2)\dagger} \right\} + \frac{1}{\Omega_k} \text{Tr} \left\{ \mathbf{A}_k \mathbf{A}_k^\dagger \right\} + \frac{N_0 N_R}{\Omega_k} \\ &+ \frac{1}{\Omega_k} \text{Tr} \left\{ \sum_{\substack{j=1 \\ (j \neq k)}}^K \sum_{b_1=1}^B \sum_{b_2=1}^B \bar{\beta}_{jk}^{(b_1, b_2)} \mathbf{H}_k^{(b_1)} \mathbf{T}_j^{(b_1)} \mathbf{T}_j^{(b_2)\dagger} \mathbf{H}_k^{(b_2)\dagger} \right\} \\ &- \frac{1}{\Omega_k} \text{Tr} \left\{ \sum_{b=1}^B \bar{\gamma}_k^{(b)} \mathbf{A}_k \mathbf{T}_k^{(b)\dagger} \mathbf{H}_k^{(b)\dagger} + \sum_{b=1}^B \bar{\gamma}_k^{(b)} \mathbf{H}_k^{(b)} \mathbf{T}_k^{(b)} \mathbf{A}_k^\dagger \right\} \\ &+ \frac{1}{\Omega_k} \text{Tr} \left\{ \sum_{b_1=1}^B \sum_{b_2=1}^B \mathbb{E}_\epsilon \left[ \alpha_k^{(b_1)} \alpha_k^{(b_2)} \right] p_k^{(b_1, b_2)} \mathbf{H}_k^{(b_1)} \mathbf{T}_k^{(b_1)} \mathbf{T}_k^{(b_2)\dagger} \mathbf{H}_k^{(b_2)\dagger} \right\}. \end{aligned}$$

In a manner similar to (11) and (12), minimizing the following Lagrange objective function:

$$f \left( \{\mathbf{T}_k\}_{k=1}^K \right) = \sum_{k=1}^K \mathbb{E}_\epsilon [\text{NMSE}_k] + \sum_{k=1}^K \kappa_k \left( \text{Tr} \left\{ \sum_{b=1}^B \mathbf{T}_k^{(b)\dagger} \mathbf{T}_k^{(b)} \right\} - P_k^{\text{tx}} \right), \quad (34)$$

results in (25).

### D. Modified JLS with Imperfect Timing-Advance

From (23), the power of received desired signal component at MS  $k$ , averaged over  $\epsilon_k^{(b)}$ , is given by  $\mathbb{E}_\epsilon [P_k^{\text{rx}}] = \frac{P_k^{\text{tx}}}{L_k} \mathbb{E}_\epsilon \left[ \text{Tr} \left\{ \mathbf{Q}_k^\dagger \mathbf{\Gamma}_k^\dagger \mathbf{H}_k^\dagger \mathbf{H}_k \mathbf{\Gamma}_k \mathbf{Q}_k \right\} \right] = \frac{1}{L_k} \text{Tr} \left\{ \mathbf{Q}_k^\dagger \tilde{\mathbf{M}}_k \mathbf{Q}_k \right\}$ . Similarly, its leakage power to other users, averaged over  $\epsilon_k^{(b)}$ , is

$$\mathbb{E}_\epsilon [P_{kj}^{\text{leak}}] = \frac{P_k^{\text{tx}}}{L_k} \sum_{b_1=1}^B \sum_{b_2=1}^B \bar{\beta}_{kj}^{(b_1, b_2)} \text{Tr} \left\{ \mathbf{Q}_k^{(b_1)\dagger} \mathbf{H}_j^{(b_1)\dagger} \mathbf{H}_j^{(b_2)} \mathbf{Q}_k^{(b_2)} \right\}, \text{ and the averaged ISI (or "self-leakage")} is given by  $\mathbb{E}_\epsilon [P_{kk}^{\text{leak}}] = \frac{P_k^{\text{tx}}}{L_k} \sum_{b_1=1}^B \sum_{b_2=1}^B \mathbb{E}_\epsilon \left[ \alpha_k^{(b_1)} \alpha_k^{(b_2)} \right] p_k^{(b_1, b_2)} \text{Tr} \left\{ \mathbf{Q}_k^{(b_1)\dagger} \mathbf{H}_k^{(b_1)\dagger} \mathbf{H}_k^{(b_2)} \mathbf{Q}_k^{(b_2)} \right\}$ .$$

Deriving a corresponding lower bound (as in (16)) for the modified expression  $SLNR_k = \frac{\mathbb{E}_\epsilon [P_k^{\text{rx}}]}{\sum_{j=1}^K \mathbb{E}_\epsilon [P_{kj}^{\text{leak}}] + N_0 N_R}$  and applying Lemma 1 to it yields the solution in (27) and (28).

#### ACKNOWLEDGMENT

We wish to thank the reviewers for pointing out the similarity between CISVD and the algorithm in [21].

#### REFERENCES

- [1] G. J. Foschini and M. J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, 1998.
- [2] S. Catreux, P. Driessen, and L. Greenstein, "Attainable throughput of an interference-limited multiple-input-multiple-output (MIMO) cellular system," *IEEE Trans. Commun.*, vol. 49, pp. 1307–1311, Aug. 2001.
- [3] H. Dai, A. F. Molisch, and H. V. Poor, "Downlink capacity of interference-limited MIMO systems with joint detection," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 442–453, Mar. 2004.
- [4] A. F. Molisch, *Wireless Communications*. Wiley-IEEE Press, 2005.
- [5] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 684–702, Jun. 2003.
- [6] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmission end," in *Proc. VTC*, (Rhodes, Greece), pp. 1745–1749, May 2001.
- [7] S. Jafar, G. Foschini, and A. Goldsmith, "Phantomnet: Exploring optimal multicellular multiple antenna systems," *EURASIP J. App. Sig. Proc.*, pp. 591–604, 2004.
- [8] G. J. Foschini, H. Huang, K. Karakayali, R. A. Valenzuela, and S. Venkatesan, "The value of coherent base station coordination," in *Proc. Conf. on Inf. Sci. and Sys. (CISS)*, Mar. 2005.
- [9] B. L. Ng, J. S. Evans, S. V. Hanly, and D. Aktas, "Transmit beamforming with cooperative base stations," in *Proc. IEEE Intl. Symp. Inf. Th.*, (Adelaide, Australia), pp. 1431–1435, 2005.
- [10] H. Zhang and H. Dai, "Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," *European J. Wireless Commun. and Networking*, pp. 222–235, 2004.
- [11] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1305–1316, Jul. 2004.
- [12] J. Zhang, Y. Wu, S. Zhou, and J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Commun. Lett.*, vol. 9, pp. 991–993, Nov. 2005.
- [13] H. Dai, L. Mailaender, and H. V. Poor, "CDMA downlink transmission with transmit antenna arrays and power control in multipath fading channels," *European J. Wireless Commun. and Networking*, 3rd quarter 2004.
- [14] A. Tarighat, M. Sadek, and A. H. Sayed, "A multiuser beamforming scheme for downlink MIMO channels based on maximizing signal-to-leakage ratios," in *Proc. ICASSP*, (Philadelphia, PA), pp. 1129–1132, 2005.
- [15] "Technical specification group GSM/EDGE, radio access network; radio subsystem synchronization," Tech. Rep. 45.010 (v6.6.0), 3rd Generation Partnership Project (3GPP), 2005.

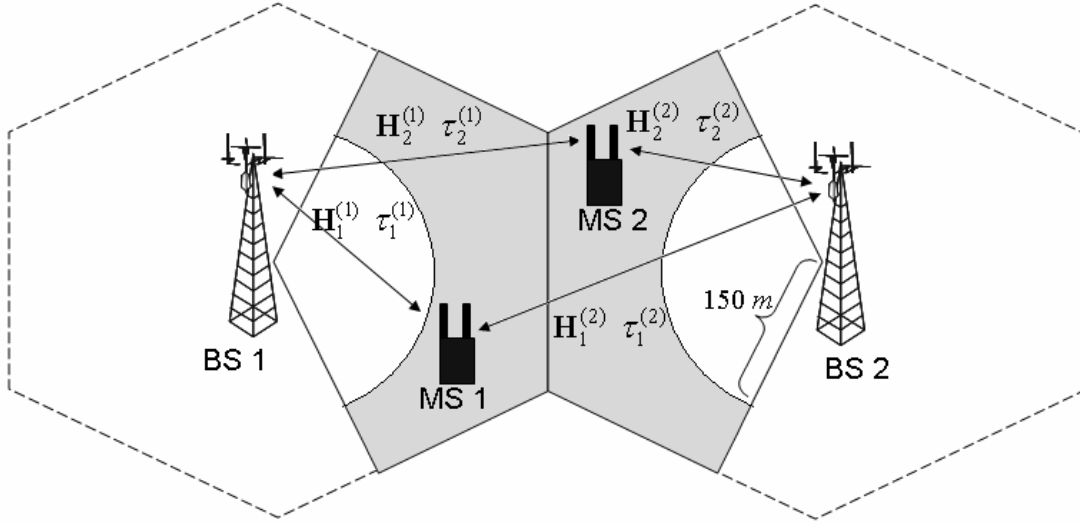


Fig. 1. A Simple BS Cooperation Scenario with 2 BSs and 2 MSs. The MSs are placed in the shaded area in the numerical simulations.

- [16] K.-K. Wong, R. S.-K. Cheng, K. B. Letaief, and R. D. Murch, "Adaptive antennas at the mobile and base stations in an OFDM/TDMA system," *IEEE Trans. Commun.*, vol. 49, pp. 195–206, Jan. 2001.
- [17] L. Shao and S. Roy, "Downlink multicell MIMO-OFDM: An architecture for next generation wireless networks," in *Proc. WCNC*, (Los Angeles, CA, USA), Mar. 2005.
- [18] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1996.
- [19] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, pp. 933–946, 2000.
- [20] J.-H. Chang, L. Tassiulas, and F. Rashid-Farrokhi, "Joint transmitter receiver diversity for efficient space division multiaccess," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 16–27, Jan. 2002.
- [21] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1105–1115, Jun. 2002.

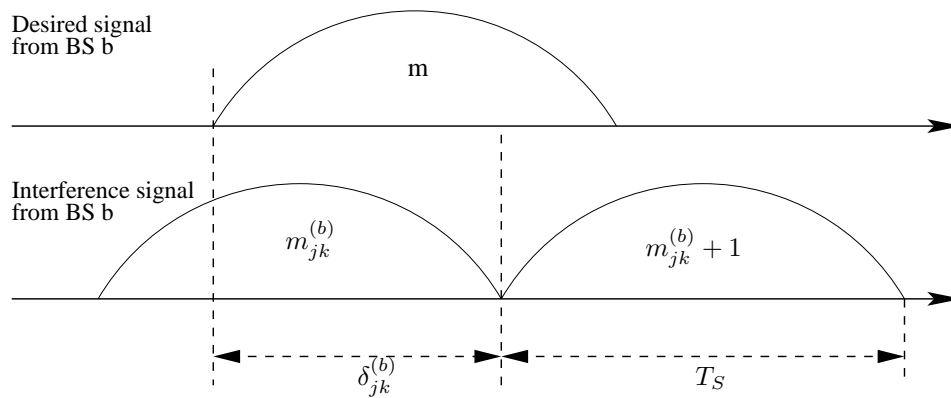


Fig. 2. Asynchronous timing difference illustration: Symbol with time index  $m_{jk}^{(b)}$ , transmitted by BS  $b$  to MS  $j$ , is the first symbol that overlaps with the desired signal symbol with index  $m$

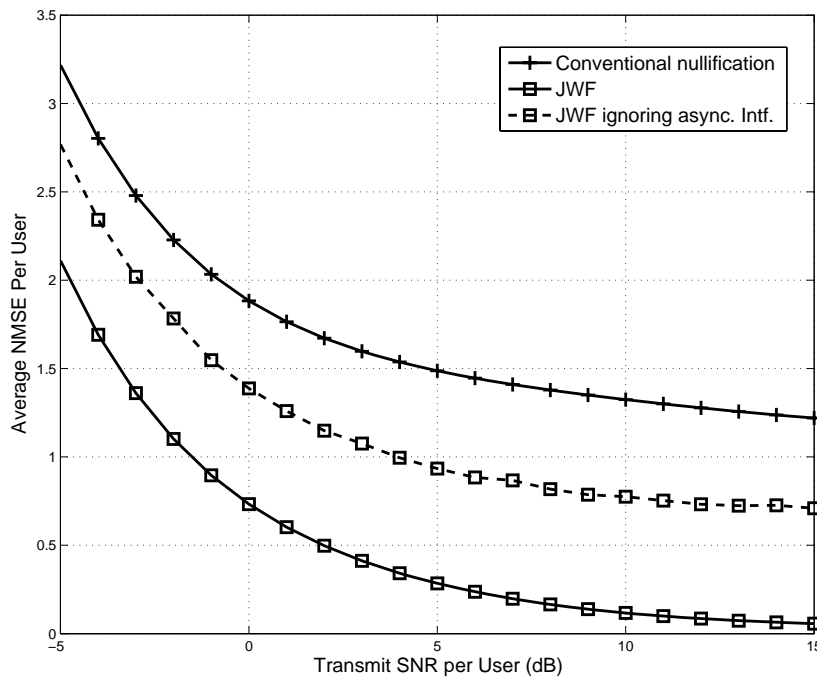


Fig. 3. Normalized MSE comparison of JWF and conventional nullification methods in the presence of asynchronous interference, when it is accounted for or neglected ( $K = 2$ ,  $B = 2$ ,  $N_T = 3$ ,  $N_R = 2$ ,  $L_k = 2$ )

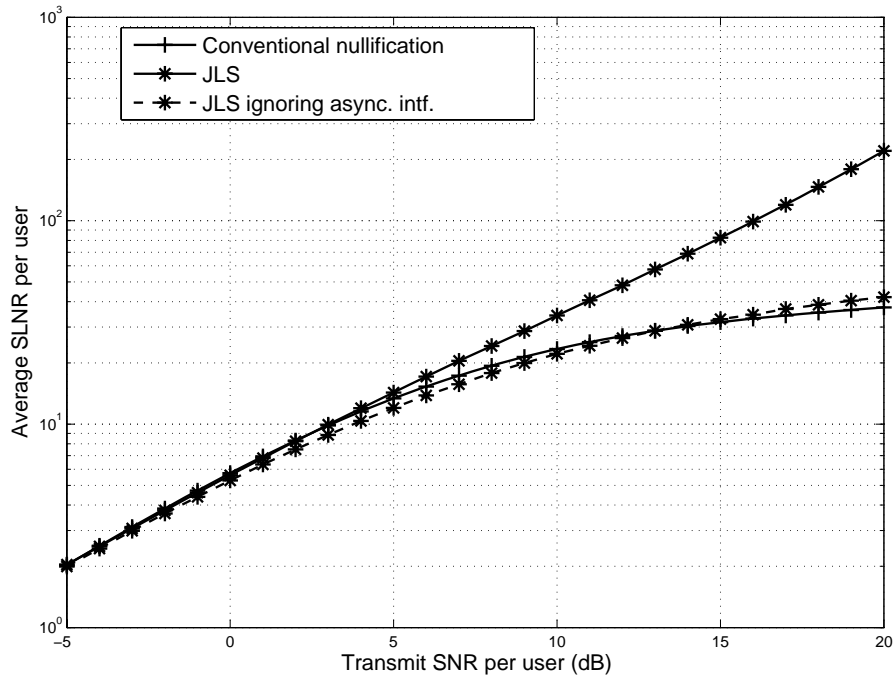


Fig. 4. SLNR comparison of JLS and conventional nullification methods in the presence of asynchronous interference, when it is accounted for or neglected ( $K = 2$ ,  $B = 2$ ,  $N_T = 3$ ,  $N_R = 2$ ,  $L_k = 2$ )

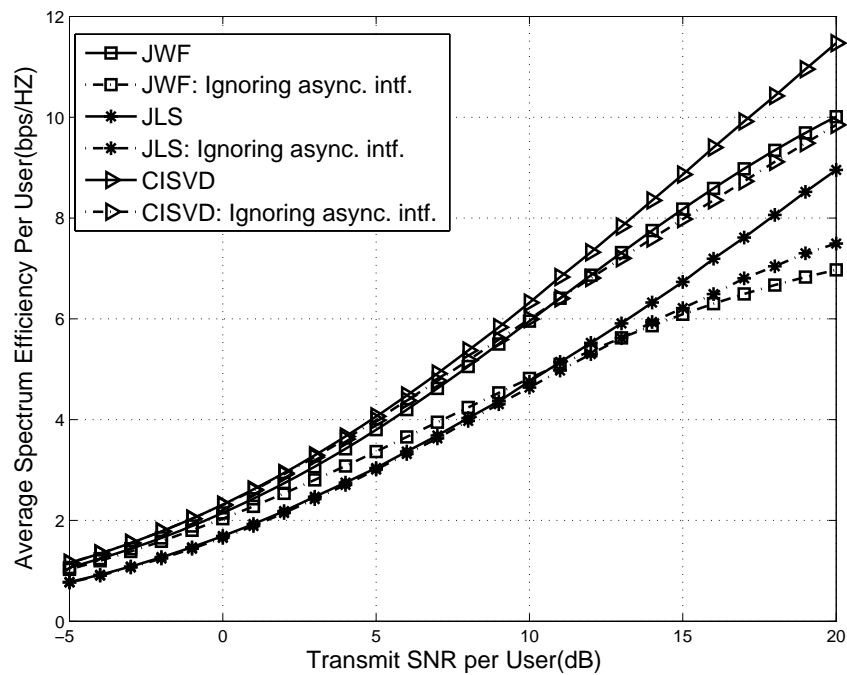


Fig. 5. Sum rates of JWF, JLS and CISVD when asynchronous interference is accounted for or neglected ( $K = 2$ ,  $B = 2$ ,  $N_T = 3$ ,  $N_R = 2$ ,  $L_k = 2$ )

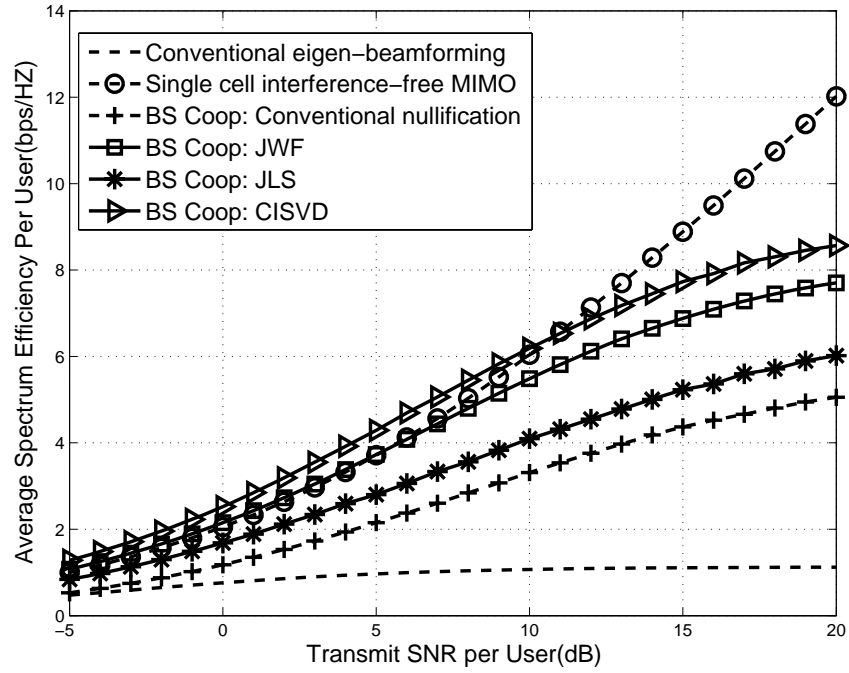


Fig. 6. Sum rate comparison of JWF, JLS, and CISVD in a 3-cell system with conventional benchmark schemes in the presence of asynchronous interference ( $K = 3$ ,  $B = 3$ ,  $N_T = 3$ ,  $N_R = 2$ ,  $L_k = 2$ )

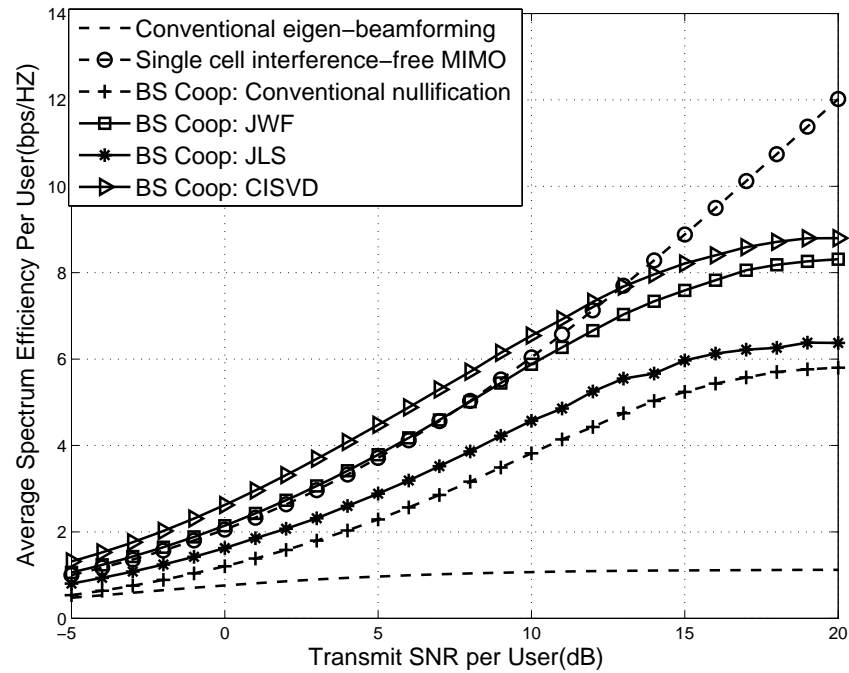


Fig. 7. Sum rate comparison of JWF, JLS and CISVD in a 3-cell system with conventional benchmark schemes in an idealized synchronous interference environment ( $K = 3$ ,  $B = 3$ ,  $N_T = 3$ ,  $N_R = 2$ ,  $L_k = 2$ )

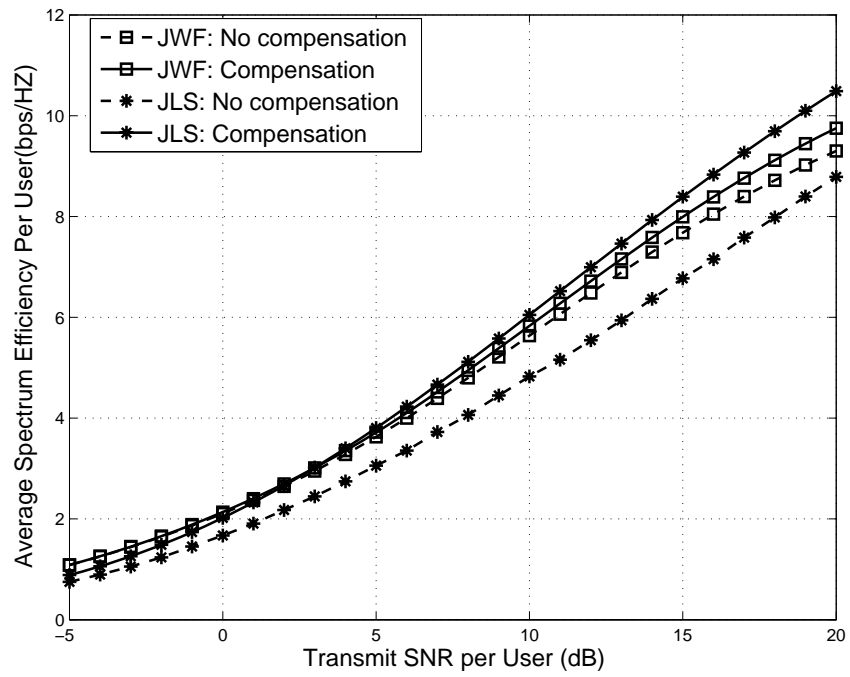


Fig. 8. Sum rate comparison of modified JLS and JWF in the presence of imperfect timing-advance ( $K = 2$ ,  $B = 2$ ,  $N_T = 3$ ,  $N_R = 2$ ,  $L_k = 2$ ). Modified CISVD is not evaluated given that no closed-form expression is available for the jitter-averaged sum rate.