The Value of Channel Prediction in CoMP Systems with Large Backhaul Latency

Liyan Su, Chenyang Yang, and Shengqian Han

Beihang University, Beijing, China

Email: liyansu@ee.buaa.edu.cn, cyyang@buaa.edu.cn and sqhan@ee.buaa.edu.cn

***Abstract*—The quality of channel state information (CSI) has large impact on the performance of coordinated multi-point (CoMP) systems. In this paper we study the impact of channel prediction on the performance of coherent CoMP transmission in time-varying channels with large backhaul latency. We resort to large system analysis to derive an explicit expression of the average user rate of the CoMP system when predicted channels are employed for downlink precoding. By comparing with the Non-CoMP systems, we find that channel prediction brings much higher performance gain to CoMP systems with respect to channel estimation. As a result, a CoMP system can perform well for mobile users even under large backhaul latency, if channel prediction will be used instead of channel estimation. Simulation results are provided to validate our analysis.**

I. Introduction

Inter-cell interference (ICI) is one of the major bottlenecks to improve spectral efficiency in future cellular networks, especially when multi-input multi-output (MIMO) techniques are applied. Except for various interference mitigation techniques, a concept of base station (BS) cooperative transmission, also known as coordinated multi-point transmission (CoMP) or network MIMO, has attracted much attention recently [1–4].

It is widely recognized that time division duplex (TDD) is more applicable for CoMP systems than frequency division duplex (FDD), because the latter needs large feedback overhead for providing channel state information at the transmitter (CSIT) [5]. In TDD systems, the downlink channels can be obtained by the BS via estimating the uplink channels exploiting the channel reciprocity [6].

A typical centralized CoMP system consists of a control unit (CU) that are connected with multiple BSs via backhaul links. Under such a framework, the CSIT is first estimated by each BS by uplink training, then the CU collects the CSIT from all BSs through backhaul links. With the CSIT of all BSs as well as the data intended to all active users at the CU, the centralized CoMP systems enable coherent cooperative transmission among multiple BSs (i.e., CoMP-JP in the context of 3GPP). This however involves the penalty of increasing the infrastructure costs for network upgrading. In currently deployed cellular systems and emerging mobile standards, the backhaul latency, e.g. X2 interface latency in 3GPP-LTE, is on the order of 10 to 20 milliseconds [7, 8]. The outdated CSIT leads to severe performance deterioration of multi-user MIMO CoMP systems even when the mobile stations (MSs) move in a low speed [8, 9].

In time-varying wireless channels, channel prediction is a popular approach to provide up-to-date channel information, which has been well investigated for traditional single-cell systems [10–12]. In fact, if we do not consider the correlation of the channels between different BSs and MSs, existing channel prediction methods can be extended straightforwardly to CoMP systems. However, this does not mean that the performance gain of a CoMP system attained from channel prediction will be the same as that of a Non-CoMP system.

In this paper, we intend to show that with channel prediction CoMP systems can perform surprisingly well in time-varying channels even under the large backhaul latency. To this end, we first derive an explicit expression of the average per-user data rate with large system analysis where the number of transmit antennas approaches infinity. Then, we analyze the impact of channel prediction errors on coherent CoMP systems using multicell zero-forcing beamforming (ZFBF), and compare with corresponding Non-CoMP systems. Analytical and simulation results show that CoMP systems can benefit more from channel prediction than Non-CoMP systems because outdated CSIT is more detrimental to CoMP transmission.

II. System Model

Consider a TDD downlink CoMP system, where B BSs each equipped with N_t antennas cooperatively serve M single-antenna MSs using multi-cell ZFBF. The CSIT is obtained through uplink training by exploiting channel reciprocity. The CU collects the CSIT from each BS via backhaul links, based on which the downlink precoders are computed. Finally, the CU sends the data and precoders to each BS via backhaul links and the coordinated BSs serve the MSs jointly.

A. Uplink Training

The uplink training signals from multiple MSs are assumed orthogonal, therefore we only need to consider the channel acquisition of a single MS in the following.

Denote $\mathbf{g}_{mb}(t) = \sqrt{\alpha_{mb}}\mathbf{h}_{mb}(t)$ as the channel vector between BS_b and MS_m, where α_{mb} is the large scale fad-

This work was supported in part by the national key project of next generation wideband wireless communication networks (2011ZX03003-001), by the International S&T Cooperation Program of China (ISCP) (2008DFA12100), and by the Fundamental Research Funds for the Central Universities.

978-1-4673-0437-5/12/$31.00 ©2012 IEEE

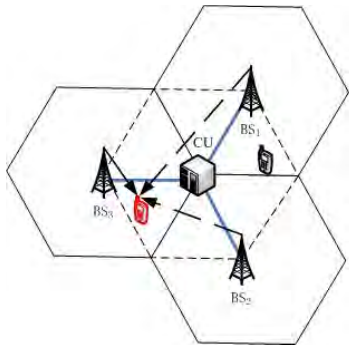

Fig. 1. An example of CoMP systems. BSs are connected to the CU via backhaul links. The solid lines denote the local channel while the dashed lines represent the cross channels. MSs are located within the dotted line zone.

ing factor including path loss and shadowing, $\mathbf{h}_{mb}(t)$ is the time-varying small scale fading channel vector of size N_t. Then the global channel vector of MS_m is $\mathbf{g}_m(t) = [\mathbf{g}_{m1}^H(t) \cdots \mathbf{g}_{mB}^H(t)]^H$, which is an aggregation of multiple single-cell channel vectors with different average energies, i.e., CoMP channel is asymmetric [6]. MS_m transmits K uplink training symbols, $x_k, k = 1, \cdots, K$ respectively located at the p_1th, $\cdots$, p_Kth symbols in the uplink frame.

The kth received training symbol from MS_m at the BS_b is

$$\mathbf{y}_{mb}(t - p_k T_s) = \mathbf{g}_{mb}(t - p_k T_s) x_k + \mathbf{n}, \tag{1}$$

where $\mathbf{n}$ is additive white Gaussian noises (AWGN) with zero mean and covariance matrix $\sigma_n^2 \mathbf{I}$, where $\mathbf{I}$ is an identity matrix, and T_s is the symbol duration.

Each BS can employ the received training symbols to estimate the uplink channel vector, then simply regards it as the downlink channel. Alternatively, each BS can employ the received training symbols to predict the downlink channel with channel statistics including spatial correlation matrix, doppler spectrum, and uplink signal-to-noise-ratio (SNR) [12]. To reveal the potential of using channel prediction to deal with the outdated CSIT, we assume that a *priori* information of these channel statistics is known, as is commonly assumed for channel prediction or estimation, e.g., as in [10–12].

Based on the minimum mean square error (MMSE) criterion, the channels at the qth downlink symbol can be predicted at BS_b as

$$\hat{\mathbf{g}}_{mb}(t + qT_s) = \mathbf{R}_{yg}^H \mathbf{R}_y^{-1} \tilde{\mathbf{y}}_{mb}(t), \tag{2}$$

where $\tilde{\mathbf{y}}_{mb}(t) = [\mathbf{y}_{mb}^H(t - p_1 T_s) \cdots \mathbf{y}_{mb}^H(t - p_K T_s)]^H$ is a vector of the received training symbols at the BS_b, $\mathbf{R}_y = \mathbb{E}\{\tilde{\mathbf{y}}_{mb}(t)\tilde{\mathbf{y}}_{mb}^H(t)\}$ is its autocorrelation matrix, $\mathbf{R}_{yg} = \mathbb{E}\{\tilde{\mathbf{y}}_{mb}(t)\mathbf{g}_{mb}^H(t + qT_s)\}$ is the cross-correlation between the received training symbols and the channel to be predicted.

Then the channel prediction can be modeled as

$$\mathbf{g}_{mb}(t + qT_s) = \hat{\mathbf{g}}_{mb}(t + qT_s) + \mathbf{e}_{mb}, \tag{3}$$

where $\mathbf{e}_{mb}$ is uncorrelated of the predicted channel and is Gaussian with covariance $\varepsilon_{mb}\mathbf{I}$ [13].

Define the normalized mean square error (NMSE) of the channel prediction as $\bar{\varepsilon}_{mb} = \varepsilon_{mb}/\alpha_{mb}$, then we have

$$\begin{aligned} \bar{\varepsilon}_{mb} &= \frac{\mathbb{E}\{\|\hat{\mathbf{g}}_{mb}(t + qT_s) - \mathbf{g}_{mb}(t + qT_s)\|^2\}}{\mathbb{E}\{\|\mathbf{g}_{mb}(t + qT_s)\|^2\}} \\ &= 1 - \frac{1}{\alpha_{mb}} \mathbf{R}_{yg}^H \mathbf{R}_y^{-1} \mathbf{R}_{yg}, \end{aligned} \tag{4}$$

which in fact is the mean square error (MSE) of the small scale fading channel prediction $\hat{\mathbf{h}}_{mb}(t + qT_s)$.

The prediction error of the global channel of the MS_m is

$$\mathbf{e}_m = \hat{\mathbf{g}}_m - \mathbf{g}_m, \tag{5}$$

where $\hat{\mathbf{g}}_m = [\hat{\mathbf{g}}_{m1}^H \cdots \hat{\mathbf{g}}_{mB}^H]^H$, $\hat{\mathbf{g}}_m$ is uncorrelated with $\mathbf{e}_m$, whose elements are zero mean and $\mathbb{E}\{\mathbf{e}_m \mathbf{e}_m^H\} = diag(\varepsilon_{m1}\mathbf{I} \cdots \varepsilon_{mB}\mathbf{I})$.

B. Downlink Transmission

The CU collects the predicted channels from all cooperative BSs to all MSs, which can be expressed as $\hat{\mathbf{G}} = [\hat{\mathbf{g}}_1 \cdots \hat{\mathbf{g}}_M]^H$. For CoMP transmission, the ZFBF precoder should be designed under per-BS power constraint (PBPC) [1], but its performance is hard to analyze. In this paper, we consider a suboptimal but more tractable power constraint, which is per-user power constraint (PUPC). It has been shown that the ZFBF precoder with different power constraints performs closely when the number of users is large [14]. Under PUPC where the transmit power of each MS is 1, the ZFBF precoder of MS_m is

$$\mathbf{w}_m = \frac{\hat{\mathbf{Q}}_m \hat{\mathbf{g}}_m}{\|\hat{\mathbf{Q}}_m \hat{\mathbf{g}}_m\|}, \tag{6}$$

where $\hat{\mathbf{Q}}_m = \mathbf{I} - \bar{\mathbf{G}}_m^H (\bar{\mathbf{G}}_m \bar{\mathbf{G}}_m^H)^{-1} \bar{\mathbf{G}}_m$, $\bar{\mathbf{G}}_m = [\hat{\mathbf{g}}_1 \cdots \hat{\mathbf{g}}_{m-1} \hat{\mathbf{g}}_{m+1} \cdots \hat{\mathbf{g}}_M]^H$, i.e., $\hat{\mathbf{Q}}_m$ is the null space of the predicted channels of all the MSs except MS_m.

The received signal of MS_m is

$$y_m = \mathbf{g}_m^T \mathbf{w}_m x_m + \sum_{j \neq m} \mathbf{g}_m^T \mathbf{w}_j x_j + z_m, \tag{7}$$

where z_m is AWGN with zero mean and variance σ^2.

III. Downlink Transmission Performance Analysis with Channel Prediction

In this section, we analyze the the average per-user data rate of downlink CoMP with the channel prediction, and compare it with that of Non-CoMP systems.

A. Average Per-user Rate of CoMP Systems

The average per-user rate of MS_m achieved by ZFBF is

$$\begin{aligned} \bar{R}_m^C &= \mathbb{E}\{\log_2(1 + \text{SINR}_m)\} \\ &= \mathbb{E}\{\log_2(1 + \frac{|\mathbf{g}_m^H \mathbf{w}_m|^2}{\sigma^2 + \sum_{j \neq m} |\mathbf{g}_m^H \mathbf{w}_j|^2})\}. \end{aligned} \tag{8}$$

In order to obtain an explicit expression of the average per-user rate, we resort to large system analysis where the number of transmit antennas approaches infinity. We begin with obtaining two propositions as follows.

Prop. 1: As $N_t \to \infty$, the random variable $|\mathbf{g}_m^H \mathbf{w}_m|^2/N_t$ converges to a deterministic constant, i.e.,

$$\frac{|\mathbf{g}_m^H \mathbf{w}_m|^2}{N_t} \to \sum_{b=1}^{B} \alpha_{mb}(1-\bar{\varepsilon}_{mb}), \qquad N_t \to \infty. \tag{9}$$

This proposition is not hard to prove by using the strong law of large numbers. It implies that when N_t is large enough, the signal power of the desired MS_m can be approximated as

$$|\mathbf{g}_m^H \mathbf{w}_m|^2 \approx N_t \sum_{b=1}^{B} \alpha_{mb}(1-\bar{\varepsilon}_{mb}). \tag{10}$$

Prop. 2: As $N_t \to \infty$, the interference of the co-scheduled MS_j to the desired MS_m, $\mathbf{g}_m^H \mathbf{w}_j$, is a complex Gaussian random variable with zero mean, i.e.,

$$\mathbf{g}_m^H \mathbf{w}_j \sim CN(0, \lambda_{mj}^C), \tag{11}$$

where the variance $\lambda_{mj}^C = \sum_{b=1}^{B} \alpha_{mb}\bar{\varepsilon}_{mb}\beta_{jb}$, and

$$\beta_{jb} = \frac{\mathbb{E}\{\|\hat{\mathbf{g}}_{jb}\|_2^2\}}{\sum_{b=1}^{B}\mathbb{E}\{\|\hat{\mathbf{g}}_{jb}\|_2^2\}} = \frac{\alpha_{jb}(1-\bar{\varepsilon}_{jb})}{\sum_{b=1}^{B}\alpha_{jb}(1-\bar{\varepsilon}_{jb})} \tag{12}$$

reflects the channel asymmetry of MS_j. This proposition is not hard to prove by using the central limit theorem.

Denote $\lambda_{mk}^C = \min\{\lambda_{m1}^C \cdots \lambda_{mm-1}^C \lambda_{mm+1}^C \cdots \lambda_{mM}^C\}$. Substituting (9) and (11) into (8), we obtain an upper bound of the average per-user rate of CoMP system with large enough N_t and high enough downlink SNR as follows,

$$\begin{aligned}
\bar{R}_m^C &\overset{(a)}{\approx} \mathbb{E}\{\log_2(\frac{|\mathbf{g}_m^H \mathbf{w}_m|^2}{\sigma^2 + \sum_{j\neq m}|\mathbf{g}_m^H \mathbf{w}_j|^2})\} \\
&\overset{(b)}{<} \mathbb{E}\{\log_2(|\mathbf{g}_m^H \mathbf{w}_m|^2\} - \mathbb{E}\{\log_2(\sum_{j\neq m}\frac{\lambda_{mk}^C}{\lambda_{mj}^C}|\mathbf{g}_m^H \mathbf{w}_j|^2)\} \\
&= \underbrace{\log_2(N_t \sum_{b=1}^{B}\alpha_{mb}(1-\bar{\varepsilon}_{mb}))}_{S_m^C\text{:contribution from signal}} \\
&\quad \underbrace{-(\log_2(\sum_{b=1}^{B}\alpha_{mb}\bar{\varepsilon}_{mb}\beta_{kb}) + Z^C)}_{I_m^C\text{:contribution from interference}} \\
&\triangleq R_C^{UB},
\end{aligned} \tag{13}$$

where (a) is because $\mathrm{SINR}_m \gg 1$ when N_t is large enough, (b) is because we reduce the sum of $M-1$ interference to $M-1$ times of its minimum λ_{mk}^C and ignore the noise, $Z^C = (\sum_{n=1}^{M-2} 1/n - 0.5772\ldots)/\ln 2$ is a constant, and I_m^C is the average per-user rate loss caused by the interference. As will be shown in simulations later, R_C^{UB} is close to $\bar{R}_C$ even with finite N_t when downlink SNR is high.

As will be explained in the sequel, R_C^{UB} depends on the channel asymmetry of both the desired user MS_m and the co-scheduled user MS_k.

1) Channel asymmetry of the desired MS: When MS_m moves from cell edge to approach its master BS, BS_{mb_m}, the large scale fading gain of its local channel α_{mb_m} will increase while those of its cross channels $\alpha_{mb}, b \neq b_m$ will decrease. Since $\alpha_{mb_m} \gg \alpha_{mb}, b \neq b_m$ when MS_m is in cell center, we can see from (13) that the values of S_m^C and I_m^C will be dominated by the NMSE of the local channel prediction, $\bar{\varepsilon}_{mb_m}$. In other words, given the location of the co-scheduled MSs, the reduction of NMSE of channel prediction of the desired MS, $\bar{\varepsilon}_{mb_m}$, will improve the performance more significantly when it is close to the cell center.

2) Channel asymmetry of the co-scheduled MSs: If the desired MS_m, is located at a specific position where $\alpha_{m1} = \cdots = \alpha_{mB} = \alpha_m$, then $\bar{\varepsilon}_{m1} = \cdots = \bar{\varepsilon}_{mB} = \bar{\varepsilon}_m$. In this case the rate loss of the desired user MS_m caused by the inter-user interference does not depend on the large scale fading gains of its co-scheduled MSs, because $\sum_{b=1}^{B}\beta_{kb} = 1$ and

$$\begin{aligned}
I_m^C &= \log_2(\alpha_m\bar{\varepsilon}_m\sum_{b=1}^{B}\beta_{kb}) + Z^C \\
&= \log_2(\alpha_m\bar{\varepsilon}_m) + Z^C.
\end{aligned} \tag{14}$$

When MS_m is located at any other positions, the rate loss I_m^C depends on β_{kb}, i.e., the rate loss of the desired MS will largely depend on the location of its co-scheduled users.

For the co-scheduled user MS_k, $\alpha_{kb_k}(1-\bar{\varepsilon}_{kb_k})$ is the power of its predicted local channel, $\alpha_{kb}(1-\bar{\varepsilon}_{kb}), b \neq b_k$ is the power of its predicted cross channel. The minimal and maximal values of β_{kb} are respectively achieved when MS_k is located at cell center and cell edge.

When MS_k is located very close to its master BS, i.e., BS_{b_k}, $\beta_{kb_k} \approx 1$ and $\beta_{kb} \approx 0, b \neq b_k$. The rate loss of MS_m caused by the interference from the co-scheduled user MS_k is

$$I_{center} \approx \log_2(\alpha_{mb_k}\bar{\varepsilon}_{mb_k}) + Z^C, \tag{15}$$

which only depends on the large scale fading gain and the NMSE of the cross channel between MS_m and BS_{b_k} $\mathbf{g}_{mb_k}$. In this case, the NMSE of the predicted channel vector $\hat{\mathbf{g}}_{mb_k}$ dominates the contribution to the rate loss of the desired MS_m.

When MS_k is located at the exact cell edge, $\beta_{k1} = \cdots = \beta_{kB} = \frac{1}{B}$. The rate loss of MS_m becomes

$$I_{edge} = \log_2(\frac{1}{B}\sum_{b=1}^{B}\alpha_{mb}\bar{\varepsilon}_{mb}) + Z^C, \tag{16}$$

which depends on the large scale fading gains and NMSE of the predicted channels between MS_m and all the BSs. In this case, the channel prediction errors of all links of the co-scheduled MS_k contribute to the rate loss of the desired MS_m.

Compare (15) with (16), the rate of MS_m is decreased more significantly when its co-scheduled user is located at cell-edge.

B. Average Per-user Rate of Non-CoMP Systems

In a Non-CoMP system., BS b_m serves p MSs in its own cell with ZF precoding, and the set of these MSs is P. ICI exists due to the concurrent downlink transmission of multiple BSs. We obtain the third proposition as follows.

Prop. 3: As $N_t \to \infty$, the interference from MS_j to the desired user MS_m, $\mathbf{g}_m^H \mathbf{w}_j$, is a complex Gaussian random variable with zero mean, i.e.,

$$\mathbf{g}_m^H \mathbf{w}_j \sim CN(0, \alpha_{mb_j}), \tag{17}$$

where MS_j and MS_m are located in different cells. This proposition is not hard to prove by using the central limit theorem.

Denote $\alpha_{mb} = \min\{\alpha_{m1} \cdots \alpha_{mb_m-1}\alpha_{mb_m+1} \cdots \alpha_{mB}\}$. Analogous to the derivation for CoMP systems, we obtain an upper bound of the average per-user rate of Non-CoMP systems as follows,

$$\begin{aligned} \bar{R}_m^{NC} &\overset{(c)}{<} \mathbb{E}\{\log_2(|\mathbf{g}_m^H \mathbf{w}_m|^2\} - \mathbb{E}\{\log_2(\sum_{j \notin P} \frac{\alpha_{mb}}{\alpha_{mb_j}} |\mathbf{g}_m^H \mathbf{w}_j|^2)\} \\ &= \underbrace{\log_2(N_t \alpha_{mb_m}(1-\bar{\varepsilon}_{mb_m}))}_{S_m^{NC}:\text{contribution from signal}} - \underbrace{(\log_2 \alpha_{mb} + Z^{NC})}_{I_m^{NC}:\text{contribution from interference}} \\ &\triangleq R_{NC}^{UB}, \end{aligned} \tag{18}$$

where (c) is because we reduce the sum of $M-p$ inter-cell interference to $M-p$ times of its minimum interference and ignore the noise as well as the intra-cell interference[1], and $Z^{NC} = (\sum_{n=1}^{M-p-1} 1/n - 0.5772\ldots)/\ln 2$ is a constant.

Note that in Non-CoMP systems, the rate loss contributed by the interference does not depend on the NMSE of channel prediction. The NMSE only affects the first term of the average per-user rate, which induces a power reduction of the desired signal and leads to a rate loss as shown in the S_m^{NC} part.

C. Impact of NMSE on CoMP and Non-CoMP Systems

Since the average SNR of the local channel of MS_m is higher than those of the cross channels, the NMSE of its local channel prediction is less than that of the cross channels, i.e.,

$$\bar{\varepsilon}_m^l \leq \bar{\varepsilon}_{mb}, b = 1, \cdots, B. \tag{19}$$

Substituting (19) into (13) we obtain an upper bound with a simplified expression

$$\begin{aligned} R_C^{UB} &< \log_2(N_t) + \log_2(\sum_{b=1}^{B} \alpha_{mb}) + \log_2(1-\bar{\varepsilon}_m^l) \\ &\quad -(\log_2(\sum_{b=1}^{B} \alpha_{mb}\beta_{kb}) + \log_2(\bar{\varepsilon}_m^l) + Z^C) \\ &= \underbrace{\log_2(1-\bar{\varepsilon}_m^l)}_{\text{contribution from signal}} - \underbrace{\log_2(\bar{\varepsilon}_m^l)}_{\text{contribution from interference}} + A_1 \\ &\triangleq \hat{R}_C^{UB}(\bar{\varepsilon}_m^l), \end{aligned} \tag{20}$$

where $A_1 = \log_2(N_t \sum_{b=1}^{B} \alpha_{mb}) - \log_2(\sum_{b=1}^{B} \alpha_{mb}\beta_{kb}) - Z^C$ does not depend on channel prediction NMSE of MS_m.

For comparison, we rewrite (18) as

$$R_{NC}^{UB}(\bar{\varepsilon}_m^l) = \underbrace{\log_2(1-\bar{\varepsilon}_m^l)}_{\text{contribution from signal}} + A_2, \tag{21}$$

where $A_2 = \log_2(N_t \alpha_m^l) - \log_2 \alpha_{mb} - Z^{NC}$.

Denote the NMSE of local channel estimation as $\bar{\varepsilon}$, and the NMSE of local channel prediction as $\bar{\varepsilon}'$, where $\bar{\varepsilon}' < \bar{\varepsilon}$. Here we omit the index of users for concise. We next show that the average per-user upper bound will be improved by using channel prediction over using channel estimation. For CoMP systems, the improvement is

$$\begin{aligned} \Delta R_m^C &= \hat{R}_C^{UB}(\bar{\varepsilon}') - \hat{R}_C^{UB}(\bar{\varepsilon}) \\ &= \underbrace{\log_2(1-\bar{\varepsilon}') - \log_2(1-\bar{\varepsilon})}_{\text{contribution from signal}} + \underbrace{\log_2(\bar{\varepsilon}) - \log_2(\bar{\varepsilon}')}_{\text{contribution from interference}} \\ &= \log_2(\frac{1-\bar{\varepsilon}'}{1-\bar{\varepsilon}}) + \log_2(\frac{\bar{\varepsilon}}{\bar{\varepsilon}'}). \end{aligned} \tag{22}$$

For Non-CoMP systems, the improvement is

$$\begin{aligned} \Delta R_m^{NC} &= R_{NC}^{UB}(\bar{\varepsilon}') - R_{NC}^{UB}(\bar{\varepsilon}) \\ &= \log_2(\frac{1-\bar{\varepsilon}'}{1-\bar{\varepsilon}}). \end{aligned} \tag{23}$$

Comparing (22) with (23), we can see that ΔR_m^{NC} is the first term of ΔR_m^C. It implies that CoMP systems benefit more significantly from channel prediction than Non-CoMP systems. This can be explained as follows. For Non-CoMP systems, we can observe from (18) or (21) that channel prediction can reduce the power loss of the desired signal, but cannot reduce the rate loss induced by the ICI. For CoMP systems, on the contrary, channel prediction reduces the rate loss caused both by the desired signal and the ICI, as shown in (20). Since ICI is the most detrimental interference in cellular systems, now we can understand why channel prediction can improve the performance of CoMP systems remarkably.

IV. Simulation and Numerical Results

In this section, we verify our previous analysis via simulations, and show the impact of channel prediction on CoMP systems with large backhaul latency.

We consider a CoMP system with BS-to-BS distance as 500 m. The BSs transmit with a maximal power of 40 W and with 10 MHz bandwidth, and the MSs have a receiver noise figure of 9 dB. The path loss exponent is 3.76, the average power loss at the reference distance of 1 m is 36.3 dB, and the minimum distance between user and BS is 35 m. This simulation setup is based on [15]. We consider a realistic time-varying channel model, Jakes Model, which is a widely applied channel model in various standardization organizations. Its temporal correlation function is $R_h(\tau) = J_0(2\pi f_d \tau)$, where $J_0(\cdot)$ is the zero-th order Bessel function of the first kind, and f_d is the Doppler spread. The carrier frequency is 2 GHz, the symbol duration is 1 ms and the speed of the MSs is 3 km/h.

To verify the results in section III obtained for $N_t \to \infty$, Fig. 2 shows the average per-user rate with different cell-edge SNRs of 10 dB or 30 dB obtained via simulation, the upper bounds of the average rate of (13) and (18), and the simplified upper bound in (20) with finite N_t. The two MSs are located on the line connecting the two BSs, the distance between each MS and its local BS is 200 m and the backhaul latency is 20

[1] Without the impact of the large latency backual, Non-CoMP systems have more accurate channels than CoMP systems. Therefore we can ignore the interference between multiple users in each cell.

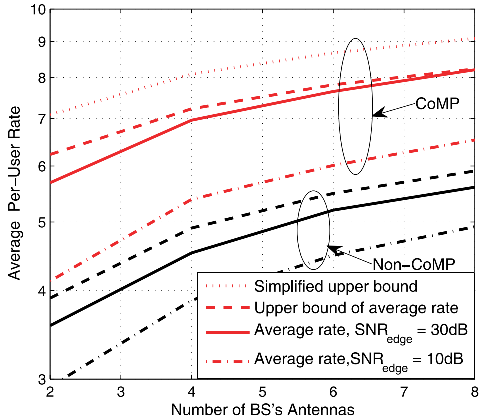

Fig. 2. The simulated average per-user rate, its upper bound and simplified upper bound of CoMP and Non-CoMP systems. $B = M = 2$

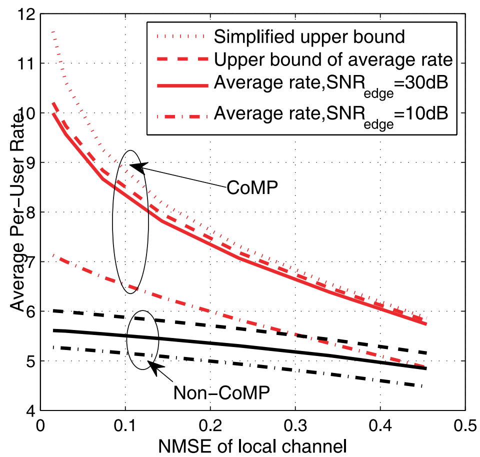

Fig. 3. Impact of channel prediction NMSE on the average per-user rate. $B = M = 2$, $N_t = 8$

ms. We can see that the upper bound of the average rate is close to the simulated average rate for high cell-edge SNR when $N_t = 4$. It indicates that the asymptotic results in large N_t are valid for the realistic antenna configuration.

In Fig. 3, we show the impact of the NMSE of channel prediction on the simulated average per-user rate and its upper bounds in CoMP and Non-CoMP systems.

From both Figs. 2 and 3, we can see that ignoring the noise is the most important factor which makes both the upper bound and the simplified upper bound not tight. Nevertheless, the slopes of the simulated average rate and the upper bounds are almost the same in each systems. Therefore this does not affect our comparison results.

Figure 4 shows the average per-user rate of a system with channel prediction or channel estimation, which is numerically obtained from the upper bounds of (13) and (18). Compared with the Non-CoMP system, the performance of the CoMP system is improved significantly by channel prediction when the co-scheduled user is located at cell-edge.

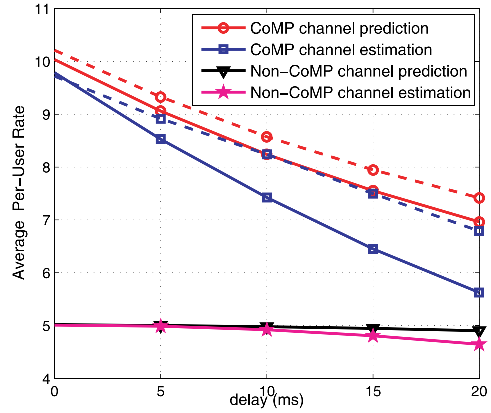

Fig. 4. Average per-user rate when either channel prediction or channel estimation is used. The CSIT delay caused by backhaul latency in CoMP systems is set equal to the CSIT delay induced by the uplink and downlink frame transmission in Non-CoMP systems. The distance between the desired user and local BS is 200 m. The solid and dashed lines represent that the distance between the co-scheduled the and its local BS is 240 m and 50 m, respectively. $B = M = 2$, $N_t = 4$

Finally, we simulate a more realistic setting, where each cell serves two randomly distributed MSs without user scheduling. With better scheduling algorithms, both CoMP and Non-CoMP systems will perform better, but the analytical results still hold. All the MSs are randomly distributed in a 'cell-edge region'. In particular, the ratio of the large scale fading gain of the local channel to the sum of those of the cross channels of the MSs in this region, e.g. $\alpha_{mb_m}/\sum_{b\neq b_m}\alpha_{mb}$ for MS_m, is less than a pre-defined value, e.g., 3 dB or 10 dB. The CSIT delay of Non-CoMP systems is 5 ms. Remind that the CSIT backhaul latency is 20 ms. We can see from Fig. 5 that the previous conclusions are still valid for the realistic systems with randomly dropped users. Comparing the results of the 3 dB cell-edge region and the whole cell region, we can find that the performance gain of channel prediction is more pronounced when the channels are more asymmetric, especially for CoMP systems. It is interesting to see that though the CSIT delay of CoMP system is much more larger than that of Non-CoMP systems, the CoMP system still outperforms Non-CoMP system when channel prediction is applied, even when the users are located in the 3 dB cell-edge region, which is usually supposed as a region for CoMP users.

V. Conclusion

In this paper, we investigated the impact of channel prediction on multiuser CoMP systems. Analysis and simulation results showed that the performance of coherent CoMP systems benefit more significantly from channel prediction than Non-CoMP systems. We found that the performance of CoMP systems will be improved remarkably even with the large backhaul latency when channel prediction is applied. It suggests that the current backhaul may be sufficient for CoMP in terms of the latency. This is very attractive since upgrading the deployed backhaul is hard and of high cost.

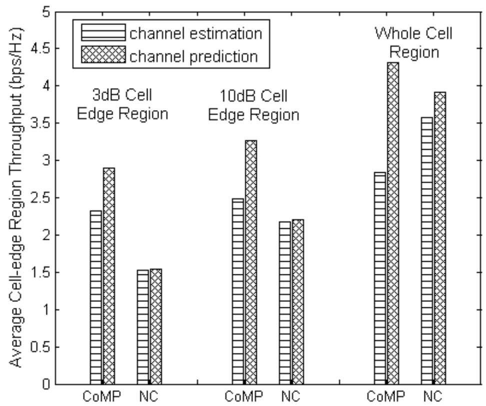

Fig. 5. Average cell-edge region per-user rate. The results are averaged over 1000 random drops. $B = 3$, $M = 6$, $N_t = 4$, cell-edge SNR is 10 dB.

REFERENCES

[1] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun. Mag.*, vol. 13, pp. 56–61, Aug. 2006.

[2] H. Zhang, N. B. Mehta, A. F. Molisch, J. Zhang, and H. Dai, "Asynchronous interference mitigation in cooperative base station systems," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 155–165, Jan. 2008.

[3] A. Tölli and M. Codreanu and M. Juntti, "Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1428–1440, Apr. 2008.

[4] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. Smith, and R. A. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 2983–2989, 2009.

[5] X. Hou and C. Yang, "How much feedback overhead is required for base station cooperative transmission to outperform non-cooperative transmission?" in *Proc. IEEE ICASSP*, May 2011.

[6] X. Hou, C. Yang, and B. K. Lau, "Impact of non-orthogonal training on performance of downlink base station cooperative transmission," *IEEE Trans. Vehi. Tech., to appear*.

[7] 3GPP TSG RAN and TS 36.420 v9.0.0, "Evolved universal terrestrial radio access network (e-utran): X2 general aspects and principles," Sep. 2009.

[8] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts performance, and field trial results," *IEEE Wireless Commun. Mag.*, vol. 49, pp. 102–111, Feb 2011.

[9] 3GPP Long Term Evolution (LTE), "Coordinated multi-point operation for LTE," *TSG RAN TR 36.819 v1.1.0*, Aug. 2011.

[10] C. Shin, J. G. Andrews, and E. J. Powers, "An efficient design of doubly selective channel estimation for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 585–588, Sep. 2007.

[11] S. Savazzi and U. Spagnolini, "Optimizing training lengths and training intervals in time-varying fading channels," *IEEE Trans. Signal Processing*, vol. 57, pp. 1098–1112, Mar. 2009.

[12] S. Han, Y. Tian, and C. Yang, "User-specified training symbol placement for channel prediction in TDD MIMO systems," *IEEE Trans. Vehi. Tech.*, vol. 60, pp. 2837 – 2843, July 2011.

[13] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Info. Theory*, vol. 56, Jun. 2010.

[14] J. Zhang, J. A. R. Chen, A. Ghosh, and R. W. H. Jr., "Networked MIMO with clustered linear precoding," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1910–1921, Apr. 2009.

[15] 3GPP TSG RAN and TR 25.814 v7.1.0, "Physical layer aspects for evolved UTRA," Sep. 2006.