

# Joint Channel Estimation for Downlink Base Station Cooperative Transmission Exploiting Channel Asymmetry

Xueying Hou, *Student Member, IEEE*, and Chenyang Yang, *Senior Member, IEEE*  
School of Electronics and Information Engineering, Beihang University, Beijing, China

**Abstract**—In this paper, we study uplink joint channel estimation for downlink base station cooperative transmission orthogonal frequency division multiplexing systems. We present three linear estimators which jointly estimates the channel coefficients from users in different cells with minimum mean square error (MSE), robust design and least square criterion, then we analyze their performance. We find that due to the asymmetric feature of the multi-cell channels in cooperative transmission systems, the MSE for different links with three joint estimators differ. When the channel energy information is exploited for channel estimation, the cross channels that experience large attenuation have smaller estimation MSEs than the local channels, and using non-orthogonal training sequences among users in different cells leads to minor performance loss. This is in contrast to the traditional understanding. To reduce the complexity of the estimators when non-orthogonal training sequences are used, we also present an interference cancelation based estimator. Both the channel estimation errors and the throughput of the cooperative transmission system are simulated to verify our analysis.

## I. INTRODUCTION

Spectral efficiency of universal frequency reuse cellular systems is severely limited by the inter-cell interference (ICI). To avoid the interference, coherent cooperative transmission is a promising strategy, where multiple base stations (BSs) behave like a super-BS [1]. For downlink cooperative transmission, multi-user precoding can provide the highest throughput, while all the channel state information (CSI) among the coordinated BSs and the mobile stations (MSs) need to be acquired at the super-BS [2].

In traditional non-cooperative cellular systems, each BS only needs to estimate CSI of the *local channels*, *i.e.*, the channels between BS and MS who are in the same cell. In coherence cooperative transmission systems, the CSI of *cross channels*, *i.e.*, the channels between BS and MS who are in different cells, need to be acquired as well. Assuming that the multipath delays of channels from the serving BS and the interference BS do not overlap, [3] develops a joint channel estimation method for the desired and interference channels. [4] suggests to spread the orthogonal sequences in time domain from slot to slot to support the estimation of channels from multiple BSs. Considering the propagation delay differences in multicell channels, [5] designs orthogonal training sequences that are robust to the propagation delay. [6] designs the training pilots for users in different cells to mitigate the ICI.

Asymmetry and asynchrony are inherent features of multi-cell channels in cooperative transmission systems. Although

channel estimators and the training sequences designed for single user multi-input multi-output (MIMO) systems [7] or multiple user non-cooperative cellular systems [6] can be applied, they do not perform well in the cooperative transmission scenarios. First, the channels from different BSs to one user have different energies since those between the user and its non-serving BSs experience large attenuation. For the first glance, this will lead to poor estimation performance for the channels with large attenuation [8]. Second, since it is impossible to synchronize all MSs to all cooperative BSs [9], we need to consider the various propagation delays when constructing the orthogonal training sequences. This will introduce extra overhead if the training sequences among the users in different cells are designed orthogonal.

In this paper, we consider joint channel estimation for time division duplexing (TDD) downlink cooperative transmission Orthogonal Frequency Division Multiplexing (OFDM) systems using multi-user precoding. We analyze the Mean Square Error (MSE) for three joint estimators that need different channel statistical information, which are the minimum MSE (MMSE) estimator, a robust estimator and the least square (LS) estimator. The robust estimator is derived from MMSE estimator by assuming that the large scale channel energy is perfectly known and the power delay profile (PDP) of the small scale fading is uniform. Our analysis and simulation results show that both the MMSE estimator and the robust estimator exhibit small MSE for estimating cross channels and provides fairly good performance for downlink cooperative transmission, even when the training sequences for the users in different cells are non-orthogonal.

## II. SYSTEM MODEL

Consider a coordinated cellular system, where  $B$  BSs each equipped with  $N_T$  antennas cooperatively transmit to  $M$  single antenna MSs in  $B$  cells using multi-user precoding such as Zero Forcing Beamforming (ZFBBF).

We assume channel reciprocity for downlink and uplink. We further assume block fading channel, which means the channel remains constant during uplink channel estimation and downlink transmission. Each BS can estimate the CSIs from all MSs in the coordinated cells to itself through uplink training. Then the central processing unit can obtain all the CSIs from the coordinated BSs via high speed backhaul links. An example of cooperative transmission system is shown in

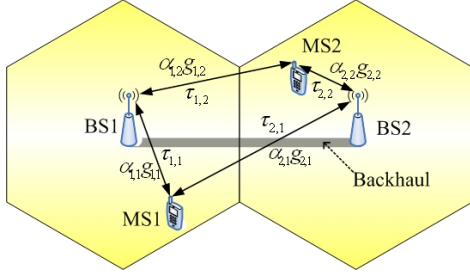


Fig. 1. An example of downlink BS cooperation transmission with 2 single antenna BSs and 2 single antenna MSs.

Fig. 1, where either BS1 or BS2 can serve as the central processing unit. Due to the asymmetric and asynchronous feature of the channels, both the large scale channel fading gains and the propagation delays of different BS-MS links are different.

We consider OFDM system since it is not only a popular candidate for cellular systems but also a promising scheme to deal with the asynchronous interference appeared in cooperative transmission systems.

Consider the composite channels from all users to all antennas at multiple BSs, which consist of both large scale fading and small scale fading. The channel frequency response (CFR) from MS  $m$  to antenna  $a$  of BS  $b$  at the subcarrier  $k$  can be expressed as

$$H_{b,a,m}(k) = \sum_{l=0}^{L-1} h_{b,a,m}(l) \exp(-j \frac{2\pi}{K} lk)$$

where  $h_{b,a,m}(l) = \alpha_{b,m} g_{b,a,m}(l)$  is the composite channel impulse response (CIR),  $\alpha_{b,m}$  is the large scale fading gain,  $g_{b,a,m}(l)$  is the small scale fading coefficient of the  $l$ th resolvable path, which is a complex Gaussian random variable with zero mean and variance  $\sigma_{gl}^2$ ,  $\sum_{l=0}^{L-1} \sigma_{gl}^2 = 1$ ,  $L$  is the number of resolvable paths,  $K$  is the number of subcarriers of the OFDM system.

Incorporating the propagation delay from MS  $m$  to BS  $b$ ,  $\tau_{b,m}$ , we can express an equivalent composite CFR as

$$\bar{H}_{b,a,m}(k) = H_{b,a,m}(k) \exp(-j \frac{2\pi}{K} \tau_{b,m} k)$$

where  $T_s$  is the sampling interval of the system.

Consider that all  $M$  MSs transmit training sequences in an uplink symbol. The frequency domain received signal vector at the antenna  $a$  of BS  $b$  can be expressed as

$$\begin{aligned} \mathbf{y}_{b,a} &= \sum_{m=1}^M \mathbf{P}_m \bar{\mathbf{H}}_{b,a,m} + \mathbf{N} \\ &= \sum_{m=1}^M \mathbf{P}_m \mathbf{F}_m \mathbf{h}_{b,a,m} + \mathbf{N} \\ &= \mathbf{X} \mathbf{h}_{b,a} + \mathbf{N} \end{aligned} \quad (1)$$

where  $\bar{\mathbf{H}}_{b,a,m} = [\bar{H}_{b,a,m}(0), \dots, \bar{H}_{b,a,m}(K-1)]^T$  is the equivalent composite CFR between MS  $m$  and antenna  $a$  of BS  $b$ ,  $\mathbf{h}_{b,a} = [(\mathbf{h}_{b,a,1})^T, \dots, (\mathbf{h}_{b,a,M})^T]^T$  is the composite CIR vector from all users to antenna  $a$  of BS  $b$  and  $\mathbf{h}_{b,a,m} =$

$[h_{b,a,m}(0), \dots, h_{b,a,m}(L-1)]^T$ ,  $(\cdot)^T$  is the transpose operator.  $\mathbf{P}_m = \text{diag}(p_{m1}, \dots, p_{mK})$ , whose diagonal elements are the frequency domain training sequence transmitted by MS  $m$ . The  $(k, l)$ th element of  $\mathbf{F}_m$  is  $[\mathbf{F}_m]_{k,l} = \exp(-j \frac{2\pi}{K} (l + \frac{\tau_{b,m}}{T_s}) k)$ ,  $k = 0, \dots, K-1$  and  $l = 0, \dots, L-1$ . When  $\tau_{b,m}$  and  $L$  are known,  $\mathbf{X} = [\mathbf{P}_1 \mathbf{F}_1, \dots, \mathbf{P}_M \mathbf{F}_M]$  is an equivalent training matrix.  $\mathbf{N}$  is the additive white Gaussian noise vector with zero mean and covariance matrix  $\sigma_n^2 \mathbf{I}_{K \times K}$ .

Since  $L \ll K$  in practical systems, we can improve the channel estimation performance significantly by exploiting the frequency correlation of channels [7]. If all the propagation delays and the number of resolvable paths are known, this can simply be implemented by first estimating the CIR, and then obtaining the CFR by Fourier transformation. In the following sections, we will present channel estimators for CIR and analyze their performance.

### III. JOINT CHANNEL ESTIMATION

When all MSs transmit training sequences at the same time, it is nature to estimate the CIRs from all MSs jointly at each BS. The performance of channel estimation depends both on the channel features and on the *a priori* information. In this paper, we assume that the propagation delays,  $\tau_{b,m}$ , and the number of resolvable paths,  $L$ , can be estimated perfectly [3].

When the channel covariance matrix,  $\mathbf{R}_{b,a} = E\{\mathbf{h}_{b,a} \mathbf{h}_{b,a}^H\}$ , is known, the MMSE estimator for the CIR vector from all users to the antenna  $a$  of BS  $b$  is optimal [10], which is,

$$\hat{\mathbf{h}}_{b,a}^{MMSE} = (\mathbf{X}^H \mathbf{X} + \sigma_n^2 \mathbf{R}_{b,a}^{-1})^{-1} \mathbf{X}^H \mathbf{y}_{b,a} \quad (2)$$

In the multi-cell scenarios, it is reasonable to assume that the channels among different BS-MS links are un-correlated. Then  $\mathbf{R}_{b,a} = \text{diag}(\alpha_{b,1}^2 \mathbf{R}_{b,a,1}, \dots, \alpha_{b,M}^2 \mathbf{R}_{b,a,M})$ , where  $\mathbf{R}_{b,a,m}$  is the covariance matrix of the small scale fading channel from MS  $m$  to antenna  $a$  of BS  $b$ . Although both  $\mathbf{R}_{b,a,m}$  and  $\alpha_{b,m}^2$  vary slowly and can be estimated in practice,  $\alpha_{b,m}^2$  is a scalar parameter that can be estimated more accurately. If we assume that  $\alpha_{b,m}^2$  can be estimated perfectly but  $\mathbf{R}_{b,a,m}$  is unknown, then by further assuming uniform PDP for small scale fading channels similar to the traditional robust channel estimation algorithms [10],  $\mathbf{R}_{b,a,m}$  reduces to  $\frac{1}{L} \mathbf{I}_{L \times L}$  for  $m = 1, \dots, M$ , and we obtain a robust estimation of the CIR vector as

$$\hat{\mathbf{h}}_{b,a}^{Robust} = (\mathbf{X}^H \mathbf{X} + \sigma_n^2 \mathbf{D}_{b,a}^{-1})^{-1} \mathbf{X}^H \mathbf{y}_{b,a} \quad (3)$$

where  $\mathbf{D}_{b,a} = \text{diag}(\alpha_{b,1}^2 \frac{1}{L} \mathbf{I}_{L \times L}, \dots, \alpha_{b,M}^2 \frac{1}{L} \mathbf{I}_{L \times L})$ .

When we know nothing more than  $\tau_{b,m}$  and  $L$ , we can obtain the LS estimation of the CIR vector as

$$\hat{\mathbf{h}}_{b,a}^{LS} = (\mathbf{X}^H \mathbf{X})^{-1} \mathbf{X}^H \mathbf{y}_{b,a} \quad (4)$$

According to [7], the training sequences in (2), (3) and (4) should be orthogonal to minimize the MSE. For BS cooperative transmission, this indicates that the training sequences should be orthogonal among different cells, different users and different antennas. Training sequences for users in different cells can be designed orthogonal in the same way as in traditional single cell systems. However, this leads to huge

overhead. Allocating different time or frequency resources to users in different cells can ensure the orthogonality among their training signals, but it is neither flexible nor spectrum efficient. Allocating orthogonal code sequences is not spectrum efficient as well. This is because it is impossible to synchronize all MSs to all cooperative BSs [9], we need to consider various propagation delays when constructing the orthogonal code sequences.

Nonetheless, we will show in next section that considering the asymmetric channel feature of cooperative transmission systems, the performance of MMSE and robust estimators degrade little when the training sequences among MSs in different cells are not orthogonal. This will lead to high spectrum efficiency when the system is not fully loaded.

It is worth to note that when non-orthogonal training sequences are used, all three estimators need to compute the inverse of a matrix of size  $ML \times ML$ . To reduce the resulting huge complexity, we can use the principle of serial interference cancelation (SIC), which is originally proposed in CDMA systems [11]. Taking the robust joint estimator as an example, we present a SIC based channel estimator, which can be directly extended to the joint MMSE channel estimator.

The difference between the presented SIC based channel estimator and that in [11] lies in that in each step our channel estimator jointly estimate all resolvable paths in CIR from a MS by using the robust single user estimator [10], rather than estimate each path by applying sliding correlation.

The estimation procedure consists of multiple iterations and each iteration proceeds  $M$  steps. Firstly, BS  $b$  ranks the MSs in descending order according to the large scale fading gains. Denote the sorted MS number set as  $\{\mathcal{S}_m\}$ . The BS applies the robust single user channel estimator [10] at the first step of the first iteration to estimate the CIR from the  $\mathcal{S}_1$ th MS, *i.e.*, the strongest MS. Next, the signal contributed by this MS is subtracted from the received signal. In the following steps, the CIR from each MS is estimated sequentially in the same way until the CIR of the MS with minimal large scale fading gain is obtained. Then another iteration starts to refine the estimation using the CIR estimated in the previous iteration. In the  $m$ th step of the  $i$ th iteration, the residual received signal after subtracting all the interference from MSs in  $\mathcal{S}_j$  becomes

$$\hat{\mathbf{y}}_{b,a}^{SIC} = \mathbf{y}_{b,a} - \sum_{j=1}^{m-1} \mathbf{P}_{\mathcal{S}_j} \mathbf{F}_{\mathcal{S}_j} \hat{\mathbf{h}}_{b,a,\mathcal{S}_j}^{iter(i)} - \sum_{j=m+1}^M \mathbf{P}_{\mathcal{S}_j} \mathbf{F}_{\mathcal{S}_j} \hat{\mathbf{h}}_{b,a,\mathcal{S}_j}^{iter(i-1)}$$

where  $\hat{\mathbf{h}}_{b,a,\mathcal{S}_j}^{iter(i)}$  is the estimated CIR from MS  $\mathcal{S}_j$  at the  $i$ th iteration.

Denote the number of iteration as  $N_{iter}$ . When the training sequences are ideally autocorrelated, the complexity of the SIC estimator is on the order of  $MN_{iter}$ . Otherwise, it is on the order of  $MN_{iter}L^3$  since the inverse of a matrix with size  $L \times L$  is required for each step. We will show in the simulation that the performance of the SIC estimator converges to the joint estimator rapidly, thus the SIC estimator is of lower complexity.

#### IV. PERFORMANCE ANALYSIS

In this section we analyze the impact of the asymmetric channels on the MSE of the MMSE, robust and LS channel estimators. For mathematic tractability, we consider a simple but fundamental scenario, where  $B$  multiple antenna BSs cooperatively serving two single antenna MSs.

Denote the channel estimation error vector of CIR from all MSs to the  $a$ th antenna of BS  $b$  as  $\Delta \mathbf{h}_{b,a} = \mathbf{h}_{b,a} - \hat{\mathbf{h}}_{b,a}$ . Then its covariance matrix is  $\mathbf{R}_{\Delta \mathbf{h}_{b,a}} = E\{\Delta \mathbf{h}_{b,a} \Delta \mathbf{h}_{b,a}^H\}$ . Substitute (2), (3) and (4) into  $\mathbf{R}_{\Delta \mathbf{h}_{b,a}}$ , respectively. Then the covariance matrix of channel estimation errors for the three joint estimators can be derived as follows by applying the Woodbury matrix identity [12],

$$\mathbf{R}_{\Delta \mathbf{h}_{b,a}}^{LS} = \sigma_n^2 \mathbf{B}^{-1} \quad (5a)$$

$$\mathbf{R}_{\Delta \mathbf{h}_{b,a}}^{MMSE} = \sigma_n^2 (\sigma_n^2 \mathbf{R}_{b,a}^{-1} + \mathbf{B})^{-1} \quad (5b)$$

$$\mathbf{R}_{\Delta \mathbf{h}_{b,a}}^{Robust} = \mathbf{\Delta} + \mathbf{R}_{\Delta \mathbf{h}_{b,a}}^{MMSE} \quad (5c)$$

where  $\mathbf{B} = \mathbf{X}^H \mathbf{X}$ ,  $\mathbf{\Delta} = \sigma_n^2 (\mathbf{B} + \sigma_n^2 \mathbf{D}_{b,a}^{-1})^{-1} \mathbf{B} (\mathbf{I} - \mathbf{R}_{b,a} \mathbf{D}_{b,a}^{-1}) (\mathbf{B} + \sigma_n^2 \mathbf{D}_{b,a}^{-1})^{-1} + \mathbf{R}_{b,a} \mathbf{B} [(\mathbf{B} + \sigma_n^2 \mathbf{R}_{b,a}^{-1})^{-1} - (\mathbf{B} + \sigma_n^2 \mathbf{D}_{b,a}^{-1})^{-1}]$ .

For simplicity, we assume that the training sequences are constant-modulus, *i.e.*,  $\mathbf{P}_m^H \mathbf{P}_m = \mathbf{I}_{K \times K}$ ,  $m = 1, 2$ . Then from (1), we have

$$\mathbf{B} = \mathbf{X}^H \mathbf{X} = \begin{pmatrix} K \mathbf{I}_{L \times L} & \mathbf{Q}_{21}^H \\ \mathbf{Q}_{21} & K \mathbf{I}_{L \times L} \end{pmatrix} \quad (6)$$

where  $\mathbf{Q}_{21} = \mathbf{F}_2^H \mathbf{P}_2^H \mathbf{P}_1 \mathbf{F}_1$ .

Substituting (6) into (5a), we can derive the MSE of the LS estimator after some manipulations as

$$MSE_{b,a,m}^{LS} = \sum_{i=(m-1)L+1}^{mL} \mathbf{R}_{\Delta \mathbf{h}_{b,a}}^{LS}(i, i) = \frac{\sigma_n^2}{K} \sum_{l=1}^L f_{LS}(\lambda_l) \quad (7)$$

where  $f_{LS}(\lambda_l) = \frac{1}{1-\lambda_l^2}$ ,  $\lambda_l^2$  is the  $l$ th eigenvalue of  $\frac{\mathbf{Q}_{21}^H \mathbf{Q}_{21}}{K^2}$ ,  $0 \leq \lambda_l^2 < 1$ .

For mathematic tractability, we assume uniform PDP for the small scale fading channel, *i.e.*,  $\mathbf{R}_{b,a} = \text{diag}(\alpha_{b,1}^2 \frac{1}{L} \mathbf{I}_{L \times L}, \alpha_{b,2}^2 \frac{1}{L} \mathbf{I}_{L \times L})$ . In this case, the MMSE estimator degenerates to the robust estimator. Substituting  $\mathbf{R}_{b,a}$  and (6) into (5b), after some manipulations we can derive the MSE of the robust estimator as

$$MSE_{b,a,m}^{MMSE} = MSE_{b,a,m}^{Robust} = \eta_{b,m} \cdot \frac{\sigma_n^2}{K} \sum_{l=1}^L f_{Robust}(\lambda_l) \quad (8)$$

where  $\eta_{b,m} = \frac{\alpha_{b,m}^2}{\alpha_{b,m}^2 + \frac{L\sigma_n^2}{K}}$ ,  $f_{Robust}(\lambda_l) = \frac{1}{1-\beta\lambda_l^2}$ ,  $\beta = \prod_{j=1}^2 \frac{\alpha_{b,j}^2}{\alpha_{b,j}^2 + \frac{L\sigma_n^2}{K}}$ .

Now we analyze the impact of channel asymmetry on the performance of the estimators by comparing (7) and (8).

##### A. Estimation MSE of Cross Channels

One difference between (7) and (8) comes from the weighting coefficient  $\eta_{b,m}$ , whose value is less than 1. We can observe that the performance of the LS estimator for channels

from different MSs to different BSs are identical, whereas the performance of the robust (MMSE) estimator for these channels depends on the large scale channel energy. The cross channels that experience large attenuation exhibit small estimate errors. This is surprising since it is in contradict with the traditional understanding [8]. Nonetheless, such a result is reasonable due to the fact that we estimate the composite channels for downlink cooperative transmission, and the signal-to-noise-ratio (SNR) for estimating the composite cross channels is in fact not low.

### B. Impact of Non-Orthogonal Training Signals

It is not hard to show that the training sequences of users in different cells should be orthogonal to minimize the MSE. If the training sequences are not orthogonal, i.e.,  $\lambda_l^2 \neq 0$  for  $l = 1, \dots, L$ , both  $f_{Robust}(\lambda_l)$  and  $f_{LS}(\lambda_l)$  in (7) and (8) will be larger than 1.

From (7) we can see that if  $\lambda_l^2$  is close to 1 for any  $l$ , the value of  $f_{LS}(\lambda_l)$  will be extremely large and the estimation performance will be severely degraded. This means that the LS estimator is quite sensitive to the orthogonality of the training sequences. From (8) we can see that the MSE of the robust (MMSE) estimator will not become very large due to the weighting coefficient  $\beta$ , whose value is always less than 1. When the large scale fading channel energy of the composite channel from a MS is small enough,  $\beta$  will be a small value. In this case, even if  $\lambda_l^2$  is close to 1,  $f_{Robust}(\lambda_l)$  will not be too large. This indicates that the performance loss of robust (MMSE) estimator will be minor when using non-orthogonal training sequences due to the severe energy attenuation of the channels from users to their non-serving BSs.

When the PDP is not uniform, which is the case in practice, the robust estimator will be inferior to the MMSE estimator. Nevertheless, we will show through simulations that the same conclusion can be drawn when the training sequences are not orthogonal.

## V. SIMULATION RESULTS

In this section, we simulate the channel estimation errors and the throughput of a cooperative transmission system to compare the channel estimators.

The urban macro-cellular system parameters are considered. The cell radius is 250 m, the MS transmit power  $P_t$  is 24 dBm, the noise power at BS is -100 dBm, the path loss is calculated by  $L_{PL} = 36.3 + 37.6 \times \log_{10}(d)$ , where  $d \geq 35$  m.

We set the system bandwidth as 5 MHz, which is divided into 256 subcarriers. The sampling rate  $1/T_s = 5$  MHz. The Spatial Channel Model (SCM) [13] is used. The training sequences for MSs within each cell are orthogonal, i.e., there are no intra-cell interference. The training sequences for MSs in different cells can be orthogonal or non-orthogonal, which are respectively constructed from the cyclic shift of one Constant Amplitude Zero Autocorrelation Code (CAZAC) [14] or from multiple CAZACs with different mother indexes.

### A. MSEs of the different estimators

In this subsection, we consider that 6 MSs are served cooperatively by two BSs. The received SNRs of the MSs, defined as  $SNR_{b,m} = P_t \alpha_{b,m}^2 / \sigma_n^2$ , range from -10 dB to 15 dB with 5 dB as the step size. According to the path loss model we used, the SNR is generally greater than 0 dB when the MS and the BS locate in the same cell, otherwise the SNR is usually smaller than 0 dB. Note that the received SNR defined in this section is for estimating small scale channels, which differs from that for the composite channels.

We use normalized MSE (NMSE) as the performance metric, which is defined as  $NMSE_{b,a,m} = MSE_{b,a,m} / \alpha_{b,m}^2$ . It can reflect the estimation performance of the small scale fading channels since we assume that the large scale fading energy is perfectly known for both robust and MMSE estimators.

In Fig.2, the  $NMSE_{b,a,m}$  versus  $SNR_{b,a,m}$  of different estimators are shown. It is shown that when non-orthogonal training sequences are used, the performance of LS estimator degrades severely, while the performance of both the joint MMSE and the joint robust estimator are almost the same as that employing orthogonal training sequences. This agrees well with our theoretical analysis.

Fig.3 compares the performance of joint robust and SIC-based robust estimator when the training sequences of MSs in different cells are not orthogonal. It is shown that the performance of SIC-based channel estimator converge to that of the joint estimator only with 4 ~ 5 iterations.

It is worth to note again that Fig.2 and Fig.3 show the NMSE performance. Even though the NMSE of the MMSE and the robust estimation is high at low SNR, this does not necessarily induce worse throughput for downlink BSs cooperative transmission. A low SNR here means a low channel energy. As we have stated earlier, it is the MSE of the composite channel that directly affects the multi-cell cooperative transmission performance. Since the MSE is the value of NMSE weighted by the large scale channel energy, the MSE shall be small at low SNR.

### B. System throughputs of different estimators

In the following system level simulation, we consider fixed clustering, where adjacent three BSs each with four antennas cooperatively transmit to twelve single antenna MSs in the three cells. Each BS estimates the CSIs from all users in the cooperative cells using the training sequences. Then the central processing unit uses all CSIs to compute the ZFBF with equal power allocation for downlink transmission. We use the cell average throughput and cell edge throughput as the performance metrics. The results are shown in Fig.4.

When the LS estimator with non-orthogonal training sequences is applied, the cell average throughput degrades severely compared with that under orthogonal training. By contrast, the corresponding performance losses of joint MMSE estimator, joint robust estimator and SIC-based estimator are minor. The throughput achieved by using the robust estimator with non-orthogonal training sequences even exceeds that by using the LS estimator with orthogonal training sequences.

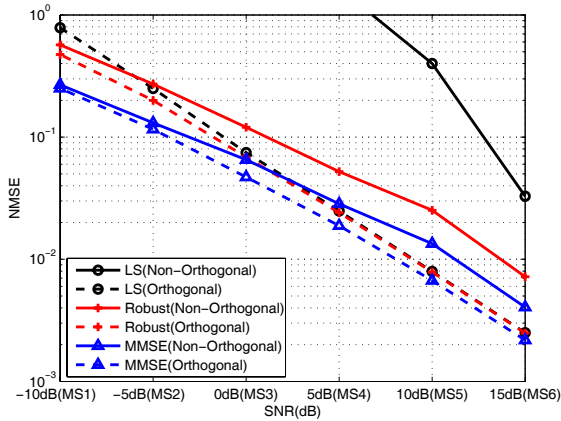


Fig. 2. NMSEs of different estimators versus SNR when orthogonal and non-orthogonal training sequences of users in different cells are considered.

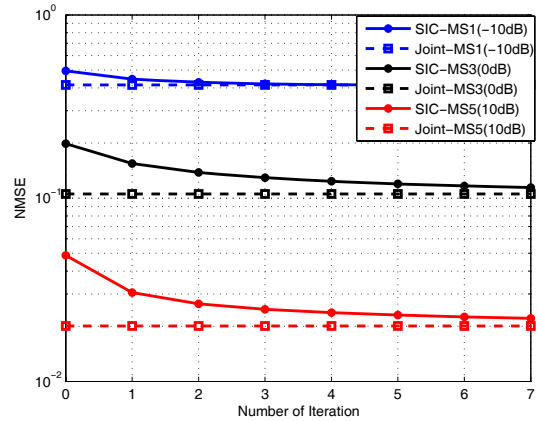


Fig. 3. NMSEs of SIC-based channel estimation for different number of iteration. The performance of channels from three MSs are shown as an example, whose receive SNR are  $-10\text{dB}$ ,  $0\text{dB}$  and  $10\text{dB}$  respectively.

This can be intuitively explained as follows. Since the downlink ZFBF is used for canceling the interference from other cells, the estimation errors of the composite cross channels play a critical role. We have shown through both analysis and simulation that the composite channels that experience large attenuation have small estimation errors when MMSE and robust estimator are used. This leads to better interference cancelation performance.

## VI. CONCLUSION

In this paper, we have studied three joint channel estimators for downlink BS cooperative transmission, the MMSE estimator, a robust estimator and the LS estimator. Due to the inherent asymmetric channel feature in the cooperative transmission systems, we have found that when the energies of large scale channels are exploited for channel estimation, the cross channels experiencing severe path loss exhibit small estimation errors, and non-orthogonal training sequences of users from different cells lead to minor performance loss. We validated our analysis by simulating both the channel estimation errors and the throughput of a cooperative transmission system.

## 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, no. 4, pp. 56-61, Aug. 2006.
- [2] L. Shao and S. Roy, "Downlink Multicell MIMO-OFDM- an Architecture for Next Generation Wireless Networks," *IEEE Wireless Commun. and Net. Conf.*, pp. 1120-1125, Mar. 2005.
- [3] R. Raghavendra, S. Bhashyam and K. Giridhar, "Interference Rejection for Parametric Channel Estimation in Reuse-1 Cellular OFDM Systems," *IEEE Trans. on Veh. Technol.*, vol.58, no.8, pp.4342-4352, Oct. 2009.
- [4] L. Thiele, M. Schellmanns, S. Schiffermuller, V. Jungnickel and W. Zirwas, "Multi-Cell Channel Estimation using Virtual Pilots," *IEEE Vehi. Tech. Conf.*, pp. 1211-1215, May 2008.
- [5] T. Kwon, H. Song and D. Hong, "Robust Channel Estimation in Multicell OFDM(A) Downlink Systems With Propagation Delay," *IEEE Vehi. Tech. Conf.*, pp. 1450-1454, April 2007.
- [6] G. Kang, P. Hasselbach, Y. Yang, P. Zhang and A. Klein, "Pilot design for inter-Cell interference mitigation in MIMO OFDM system," *IEEE Commun. Letters*, vol. 11, no. 3, pp. 237-239, Mar. 2007.

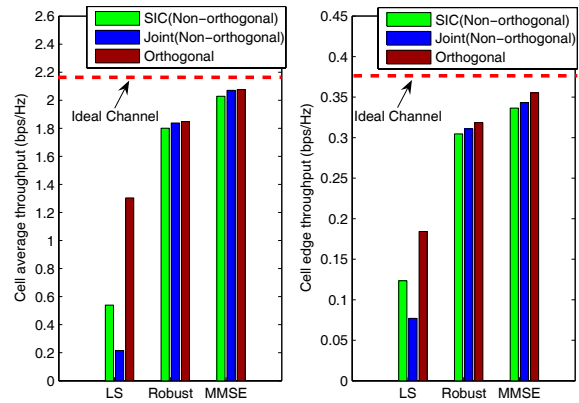


Fig. 4. Cell average and cell edge throughput when different channel estimators are used.

- [7] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. on Commun.*, vol. 1, no. 1, pp. 67-75, Jan. 2002.
- [8] A. Tolli and M. Codreanu and M. Juntti, "Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas," *IEEE Trans. on Wireless Commun.*, vol.7, no. 4, pp.1428-1440, Apr. 2008.
- [9] H. Zhang, N. B. Mehta, A. F. Molisch, J. Zhang, and H. Dai, "Asynchronous interference mitigation in cooperative base station systems," *IEEE Trans. on Wireless Commun.*, vol. 7, no. 1, pp. 155-165, Jan. 2008.
- [10] Y. Li, L. Cimini, Jr and N. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. on Commun.*, vol. 46, no. 7, pp. 902-915, Jul. 1998.
- [11] R. Fantacci and D. Marabissi, "A low complexity pilot aided channel estimation method for WCDMA communication systems," *IEEE Trans. on Vehic. Tech.*, vol.54., no.5, pp.1739 - 1746, Sept. 2005.
- [12] G. H. Golub and C. F. van Loan, *Matrix Computations*, Baltimore, MD: Johns Hopkins Univ. Press, 1996.
- [13] *Spatial channel models for MIMO simulations*, 3GPP std. TR 25.996, 2006.
- [14] D. C. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. on Inf. Theory*, vol. 18, no. 4, pp. 531-532, July 1972.