Downlink Scheduling in Network MIMO Using Two-Stage Channel State Feedback

Li-Chuan Tseng[†,*], Xin Jin[†], Abdelwaheb Marzouki[†], ChingYao Huang[*]
[†]Institut Mines-Telecom, Telecom SudParis, CNRS Samovar UMR 5157, France
[*]National Chiao-Tung University, Hsinchu, Tawian
{li.tseng, xin.jin, abdelwaheb.marzouki}@it-sudparis.eu, cyhuang@mail.nctu.edu.tw

***Abstract*—In this paper, we propose a user scheduling scheme which is based merely on the ergodic rate. We derive the ergodic rate of each user for our network MIMO system from the average Signal-to-Noise Ratio (SNR) and the covariance of prediction error. Further more, we design a joint precoding for the selected users. The precoding matrices are contrived according to the predicted Channel State Information at Transmitter (CSIT) in the future time slot to reduce the deviation caused by delay. Such a network MIMO system is under consideration, where the users are cooperatively served by multiple Base Stations (BSs). User scheduling and joint precoding are executed by a central unit connected to all BSs. Both the above functions require the CSIT which is acquired by an uplink feedback overhead. Simulation results demonstrate a conspicuous improvement in user spectral efficiency with reduced feedback.**

***Index Terms*—Network MIMO, Channel Prediction, Channel State Feedback, Scheduling.**

I. Introduction

Frequency reuse among neighboring cells in conventional cellular networks can provoke Inter-Cell Interference (ICI) that leads to performance degradation. Base Station (BS) cooperation becomes a necessity to overcome the above problem. Multiple Input Multiple Output (MIMO) technique has been in a starring role to improve the throughput of broadband wireless networks. Assuming perfect back-haul connection, the network consisting of multiple cells can be viewed as a virtual MIMO system, and the users are jointly served by multiple BSs.

In each resource unit, the cooperating BSs in one cluster jointly select a set of active users. For the selected users, precoding is applied for the purpose that the received signal for the intended user is acquired with lowest interference. Linear precoding for Multi-User (MU) MIMO based on block diagonalization (BD) was proposed in [1], in which the signal for each user is projected onto the nullspace of the augmented channel matrix of other users. The authors in [3] considered precoding for clustered network MIMO with inter-cluster coordination. In [4], the authors proposed a greedy user selection algorithm for single-cell MU-MIMO networks based on BD precoding. Scheduling in network MIMO based on Successive Zero-Forcing (SZF) was discussed in [5], where the Multi-User Interference (MUI) is partially canceled.

The above works implicitly assumed that perfect Channel State Information at the Transmitter (CSIT) is available, which may not hold in practical systems. The ergodic capacity in single-cell MIMO system under delayed CSIT with BD precoding was analyzed in [6], where the authors pointed out that under certain conditions it is preferred to adopt Single-User (SU)-MIMO technique over MU-MIMO scenario. The authors in [7] addressed the scheduling problem considering different channel variation among users, and it classified the users into two categories: predictable and non-predictable. The users belong in the first category are served in MU mode, while the other users are served in SU mode.

The overhead introduced by CSIT feedback limits the performance of MIMO systems. For multi-cell networks with large number of users, further feedback reduction is desired. In this paper, we propose a framework that performs scheduling based on the asymptotic ergodic rate which is a function of large-scale channel behavior and the maximum Doppler shift of users. Then channel prediction is applied at the transmitter side to provide a predicted CSIT for selected users, based on which the precoding matrix can be determined. Our main contributions are as follows. First, we analyze the asymptotic ergodic rate in network MIMO systems with imperfect predicted CSIT. We extend the work for single-cell MIMO systems by introducing the concept of equivalent Gaussian channels. It is shown that under low SNR or high channel variation, selecting less users may improve the throughput due to the reduced MUI. Second, we propose a user selection method based on a two-stage feedback mechanism. The *long-term* feedback contains the first and second order statistics of all users, and are used to select a set of users to be served together in one frame, while the *short-term* feedback is applied only to the selected users on each resource unit. In this way the CSIT required for precoding is available, while the amount of uplink overhead for feedback is reduced.

The rest of the paper is organized as follows: In Section II, the system model and channel prediction method are described. In Section III, the ergodic rate analysis with imperfect CSIT is presented. In Section IV, we propose user selection algorithm based on the derived ergodic rate. Numerical results are provided in Section V, with the conclusion drawn in Section VI.

The following notations are used throughout this paper. Normal letters represent scalar quantities; uppercase and lower case boldface letters denote matrices and vectors, respectively. $(\cdot)^H$ stands for the conjugate transpose. We use $[\mathbf{T}]_{(\mathbf{r},\mathbf{c})}$ to denote the sub-matrix of $\mathbf{T}$ generated by selecting the rows and columns indexed by vectors $\mathbf{r}$ and $\mathbf{c}$ respectively, and

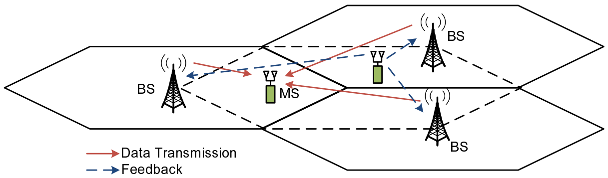

Fig. 1. Three-cell Cluster

the symbol : is used to represent all columns or rows. $\mathbf{I}_N$ represents the identity matrix with size N.

II. System Model

We consider a three-cell cluster layout depicted in Fig. 1. With sectoring, users in the dashed hexagonal area are jointly served by three BSs. The aggregated channel coefficient matrix of user u from all the B cooperating BSs is written as

$$\mathbf{H}_u = [\rho_{u,1}\mathbf{H}_{u,1}, \cdots, \rho_{u,B}\mathbf{H}_{u,B}] \tag{1}$$

where $\mathbf{H}_{u,b} \in \mathbb{C}^{N_r \times N_t}$ represents the small-scale fading channel matrix associated with user u and BS b, where N_t and N_r are the number of antennas of one BS and one user, respectively. $\rho_{u,b}$ captures the large-scale fading behavior including pathloss and shadowing. The coefficients in $\mathbf{H}_{u,b}$ are modeled as wide-sense stationary (WSS), narrow-band complex Gaussian processes. The received signal of user u can be written as

$$\mathbf{y}_u = \mathbf{H}_u \mathbf{T}_u \mathbf{x}_u + \mathbf{H}_u \sum_{u' \neq u} \mathbf{T}_{u'} \mathbf{x}_{u'} + \mathbf{n}_u \tag{2}$$

where $\mathbf{T}_u = [\mathbf{T}_{u,1}, \cdots, \mathbf{T}_{u,B}]$ is the aggregated precoder matrix with orthonormal columns, $\mathbf{x}_u$ is the transmitted signal and $\mathbf{n}_u$ is a white complex Gaussian noise vector with covariance matrix $\sigma^2\mathbf{I}_{N_r}$. The second term in (2) represents the multi-user interference (MUI). With BD precoding and perfect CSIT, the MUI can be eliminated by transmitting the signal of one user on the nullspace of the augmented channel matrix of other simultaneously selected users. However, the MUI is not perfectly canceled in our system due to the inaccuracy of channel prediction.

To track the channel variation, here we adopt the channel prediction method proposed in [8], in which the channel coefficients in future slots are predicted based on previous channel feedback, by applying the auto-regressive (AR) model and Kalman filter. The actual channel matrix $\mathbf{H}_u$ can be written as

$$\mathbf{H}_u = \mathbf{H}_u^{(P)} + \mathbf{E}_u \tag{3}$$

where $\mathbf{H}_u^{(P)}$ and $\mathbf{E}_u$ are the predicted channel and channel prediction error matrices, respectively. Similar to (1), these two matrices can be expressed as

$$\begin{aligned} \mathbf{H}_u^{(P)} &= [\rho_{u,1}\mathbf{H}_{u,1}^{(P)} \quad \cdots \quad \rho_{u,B}\mathbf{H}_{u,B}^{(P)}] \\ \mathbf{E}_u &= [\rho_{u,1}\mathbf{E}_{u,1} \quad \cdots \quad \rho_{u,B}\mathbf{E}_{u,B}] \end{aligned}$$

The elements in $\mathbf{H}_u^{(P)}$ and $\mathbf{E}_u$ are zero mean Gaussian random variables with variance $1-\epsilon_u^2$ and ϵ_u^2, respectively.

III. Asymptotic Ergodic Rate Analysis

In this section, based on the approximated channel, we provide asymptotic ergodic capacity analysis for network MIMO systems for single-user (SU) and multi-user (MU) cases. For better tractability, total power constraint (TPC) in each cluster is applied instead of per BS power constraint (PBPC). When m users are served simultaneously, with TPC and equal power allocation for each stream, the transmission power for one stream is given by $\gamma_m = P/(mN_r)$ for all users where P is the cluster power constraint, and assume N_r streams are assigned to each user. The asymptotic rates are given as functions of the large-scale fading behavior and the Doppler shift.

A. The Equivalent Channel Matrix

For single-cell MIMO systems with independent identically distributed (i.i.d.) Gaussian channel coefficients, the asymptotic ergodic rate is analyzed in [9]. To find the asymptotic ergodic rate for network MIMO systems where the elements of the aggregated channel matrix are not always i.i.d. Gaussian random variables, we propose a method to transform the network MIMO channel matrix into a virtual single cell MIMO channel. We first introduce the following definition [10]:

Definition 1: Let $\mathbf{z}_i \in \mathbb{C}^{q\times 1}$, $i = 1, ..., l$ be multivariate normal distributed vectors with zero mean and covariance matrix $\mathbf{C}$, and $\mathbf{Z}$ denotes the $q \times l$ matrix composed of the column vectors $\mathbf{z}_i$, then the matrix $\mathbf{Z}\mathbf{Z}^H$ has a central Wishart distribution with covariance matrix $\mathbf{C}$ and l degrees of freedom, denoted as $\mathbf{Z}\mathbf{Z}^H \sim \mathcal{CW}_l(0, \mathbf{C})$.

We observe that $\mathbf{H}_u\mathbf{H}_u^H = \sum_{b=1}^{B} \rho_{u,b}^2 \mathbf{H}_{u,b}\mathbf{H}_{u,b}^H$ is a linear combination of central Wishart matrices, and the row vectors of have identical covariance matrix. From [11], the distribution can be approximated $\mathbf{H}_u\mathbf{H}_u^H \sim \mathcal{CW}_{\hat{N}_{t,u}}(0, \rho_u \mathbf{I}_{N_r})$ where $\hat{N}_{t,u} = N_t \left(\frac{(\sum_b^B \rho_{u,b})^2}{\sum_b^B \rho_{u,b}^2} \right)$ and $\rho_u = \left(\frac{\sum_b^B \rho_{u,b}^2}{\sum_b^B \rho_{u,b}} \right)$.

The above approximation can be interpreted as if the user is communicating with a virtual BS with $\hat{N}_{t,u}$ transmitting antennas, and the channel is modeled as an equivalent $N_r \times \hat{N}_{t,u}$ channel matrix $\rho_u \hat{\mathbf{H}}_u$, where $\hat{\mathbf{H}}_u$ is the equivalent small-scale fading matrix and ρ_u is the equivalent large-scale fading parameter. Since ρ_u is determined by the large-scale fading gain from different BSs, it can be viewed that the user is actually served by the antennas from BSs with larger $\rho_{u,b}$'s.

B. Asymptotic Rate for SU Network MIMO

We first consider the case where only a single user is served in the multi-cell network.

Proposition 1: Under single user network MIMO transmission, as $BN_t, N_r \to \infty$ with $BN_t/N_r = \beta$, the asymptotic ergodic rate with normalized thermal noise power can be approximated by

$$\begin{aligned} \frac{C_{su}(\beta,\gamma)}{N_r} &= \log\left[1 + \hat{\beta}\hat{\gamma} - F(\hat{\beta},\hat{\gamma})\right] + \\ &\hat{\beta}\log_2\left[1 + \hat{\beta}\hat{\gamma} - F(\hat{\beta},\hat{\gamma})\right] - \log_2(e)F(\hat{\beta},\hat{\gamma}) \end{aligned} \tag{4}$$

where

$$F(x,y)=\frac{1}{4}\left[\sqrt{1+y(1+\sqrt{x})^2}-\sqrt{1+y(1-\sqrt{x})^2}\right]^2$$
$$\hat{\beta}=(N_{t,u}/\hat{N}_{t,u})\beta=\hat{N}_{t,u}/N_r,\ \hat{\gamma}=\rho_u^2\gamma$$

The proposition can be proved as in [6] by simple substitutions of parameters using the equivalent channel formation.

C. Asymptotic Rate for MU Network MIMO

Denote the BD precoder for user u designed based on the predicted CSIT as $\mathbf{T}_u^{(P)}$, the received signal can be rewritten as

$$\mathbf{y}_u=\mathbf{H}_u\mathbf{T}_u^{(P)}\mathbf{x}_u+\mathbf{E}_u\sum_{u'\neq u}\mathbf{T}_{u'}^{(P)}\mathbf{x}_{u'}+\mathbf{n}_u \tag{5}$$

The achievable rate of user u is given by

$$R_u^{(P)}=\mathbb{E}\left[\log_2\det\left(\mathbf{I}+\gamma\mathbf{H}_u\mathbf{T}_u^{(P)}\mathbf{T}_u^{(P)^H}\mathbf{H}_u^H\mathbf{R}_u^{-1}\right)\right] \tag{6}$$

where $\mathbf{R}_u^{-1}$ is the interference plus noise covariance matrix given by

$$\mathbf{R}_u=\mathbf{E}_u\left(\sum_{u'\neq u}\gamma_m\mathbf{T}_{u'}^{(P)}\mathbf{T}_{u'}^{(P)^H}\right)\mathbf{E}_u^H+\sigma^2\mathbf{I}_{N_r} \tag{7}$$

Before deriving the asymptotic rate for the MU case, we need the following lemma.

Lemma 1: Consider N complex random variables $x_i, i=1,...,N$ with $\sum_{i=1}^N\|x_i\|^2=1$. If $\|x_i\|^2$'s are independent uniformly distributed, as $N\to\infty$, the distribution of the summation of K randomly selected $\|x_k\|^2$ concentrates toward K/N.

The distribution can be found by a procedure similar to that in [12]. We skip the proof here due to space limitation.

Proposition 2: For a network MIMO system with imperfect CSIT, the asymptotic results for the achievable rate in MU mode with m simultaneously served users is approximated as

$$\frac{R_u(m)}{N_r}\approx(m-1)\log_2\left(\frac{1+N_r\rho_u^2\gamma_m\kappa\epsilon_u^2\eta_1}{1+N_r\rho_u^2\gamma_m\kappa\epsilon_u^2\eta_2}\right)+$$
$$\log_2(1+N_r\gamma_m\rho_u^2\kappa\eta_1)+\log_2\frac{\eta_2}{\eta_1}+(\eta_2-\eta_1)\log_2(e) \tag{8}$$

where κ, η_1 and η_2 are given in the proof.

Proof: Using the Wishart matrix approximation, rewrite the received signal for user u as

$$\mathbf{y}_u=\hat{\mathbf{H}}_u\hat{\mathbf{T}}_u\mathbf{x}_u+\hat{\mathbf{E}}_u\sum_{u'\neq u}\mathbf{T}_{u'}^{(P)}\mathbf{x}_{u'}+\mathbf{n}_u \tag{9}$$

where $\hat{\mathbf{H}}_u$ is the equivalent channel matrix with size of $N_r\times\hat{N}_{t,u}$, $\hat{\mathbf{T}}_u$ is the $\hat{N}_{t,u}\times N_r$ equivalent precoding matrix. Let $\mathbf{g}=[g_1,\cdots,g_{\hat{N}_{t,u}}]$ be the column index of consisting of biggest large-scale fading coefficients, the equivalent precoding matrix is given by selecting the corresponding rows from the $N_r\times BN_t$ precoding matrix $\mathbf{T}_u^{(P)}$, that is, $\hat{\mathbf{T}}_u=[\mathbf{T}_u^{(P)}]_{(\mathbf{g},:)}$. With approximation method similar to that applied to channel matrix, the equivalent CSIT error matrix $\hat{\mathbf{E}}_u\in\mathbb{C}^{N_r\times\hat{N}_{t,u}}$ has zero mean Gaussian distributed elements with variance ϵ_u^2. The achievable rate of user u is rewritten as

$$R_u=\mathbb{E}\left[\log_2\det\left(\hat{\mathbf{R}}_u+\gamma\rho_u^2\mathbf{H}_{eff,u}\mathbf{H}_{eff,u}^H\right)\right]$$
$$-\mathbb{E}\left[\log_2\det(\hat{\mathbf{R}}_u)\right] \tag{10}$$

where $\mathbf{H}_{eff,u}=\hat{\mathbf{H}}_u\hat{\mathbf{T}}_u$ is the $N_r\times N_r$ effective small-scale channel matrix for user u, and $\hat{\mathbf{R}}_u$ is the equivalent interference-plus-noise covariance matrix:

$$\hat{\mathbf{R}}_u=\mathbf{I}_{N_r}+\hat{\mathbf{E}}_u\left[\sum_{u'\neq u}\gamma_m\rho_u^2\hat{\mathbf{T}}_{u'}\hat{\mathbf{T}}_{u'}^H\right]\hat{\mathbf{E}}_u^H \tag{11}$$

Since $\hat{\mathbf{T}}_u$ is independent of $\hat{\mathbf{H}}_u$, the elements in $\mathbf{H}_{eff,u}$ are linear combinations of i.i.d. standard Gaussian random variables, which are zero mean Gaussian distributed with variance equal to the summation of the square of the coefficients in linear combination. Let $\mathbf{t}_i$ be the ith column of $\mathbf{T}_u^{(P)}$, since we have no knowledge of the small-scale fading behavior, a simple way is to assume the elements of $\mathbf{t}_i$ are independent uniformly distributed. Now pick the corresponding rows from $\mathbf{T}_u^{(P)}$ to form the equivalent precoder $\hat{\mathbf{T}}_u$. Let $\hat{\mathbf{t}}_i$ be the ith column of $\hat{\mathbf{T}}_u$, from *Lemma 1*, we have $\|\hat{\mathbf{t}}_i\|^2=\hat{t}_{i,g_1}\hat{t}_{g_1,i}^H\cdots+\hat{t}_{g_{\hat{N}_{t,u}},i}\hat{t}_{g_{\hat{N}_{t,u}},i}^H\approx\hat{N}_{t,u}/(BN_t)=\kappa$, and thus $\mathbf{H}_{eff,u}\sim\mathcal{CN}(\mathbf{0}_{N_r},\kappa\mathbf{I}_{N_r})$. Similarly, the effective channel error matrix $\mathbf{E}_{eff,u'}=\hat{\mathbf{E}}_u\hat{\mathbf{T}}_{u'}$ distributes as $\mathbf{E}_{eff,u'}\sim\mathcal{CN}(\mathbf{0}_{N_r},\kappa\epsilon_u^2\mathbf{I}_{N_r})$.

Interpreting the system in (9) as a MIMO channel under interference, the achievable rate in (8) can be obtained using the approximation in [9], where η_1 and η_2 are solutions to

$$\eta_1+\frac{N_r\gamma_m\rho_u^2\kappa\eta_1}{N_r\gamma_m\rho_u^2\kappa\eta_1+1}+(m-1)\frac{N_r\gamma_m\rho_u^2\kappa\epsilon_u^2\eta_1}{N_r\gamma_m\rho_u^2\kappa\epsilon_u^2\eta_1+1}=1$$

$$\eta_2+(m-1)\frac{N_r\gamma_m\rho_u^2\kappa\epsilon_u^2\eta_2}{N_r\gamma_m\rho_u^2\kappa\epsilon_u^2\eta_2+1}=1$$

■

To verify the accuracy of the equivalent channel matrix, consider the system model in Fig. 1 with the number of antennas given as $N_t=4$ and $N_r=2$. Six users are randomly place in the area covered by the three sectors surrounding the cluster center. The users are moving at the speed of 10 km/h, and channel prediction with AR model of order 2 is adopted. The cell edge signal-to-noise-ratio (SNR) is defined as the received SNR at the cell edge when one BS transmits at full power and other BSs are off. The asymptotic results compared with simulations are shown in Fig. 2. The largest difference between asymptotic and simulation results is around 20%. Fig. 2 also implies the SU/MU mode-switching for varying SNR.

IV. Two-stage Feedback and User Scheduling

The proposed two-stage feedback method is discussed in this section. At the beginning of each scheduling interval, each user reports its average SNR from all cooperating BSs and the maximum Doppler shift through the *long-term* feedback. Based on this information, the asymptotic ergodic rate is

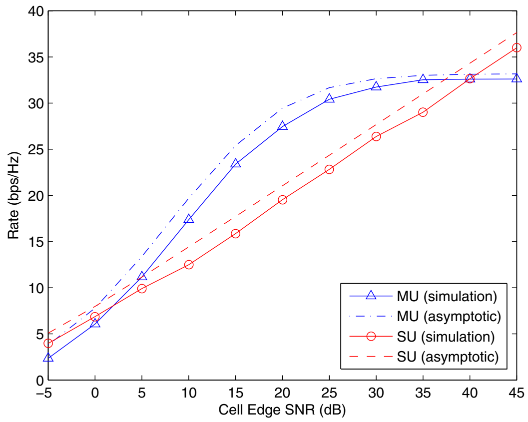

Fig. 2. Sum rate versus cell edge SNR with comparison of simulation and approximation where $N_t = 4$ and $N_r = 2$

evaluated for each user, and a set of users are selected on each resource units. This allocation is repeated on every frames before the next long-term feedback. In each frame, the small-scale channel coefficients are periodically sent to the central unit through *short-term* feedback. Channel prediction is applied to generate hypothesized channel coefficients for future frames which are used to calculate the precoder. After receiving the downlink pilot symbol, the channel coefficients are updated based on the measurements. The feedback of channel coefficients occupies limited uplink bandwidth since the number of selected users is much smaller as compared to the total number of users.

Let U and G be the set of all users and RBs, respectively. On each resource unit $j \in G$, a subset of users $U_j \subseteq U$ is selected, where the cardinality of U_j is less than or equal to $M = \lfloor BN_t/N_r \rfloor$ if the BD precoding is applied. Denote the achievable rate of user u on one resource unit as $R_u(n_j)$, where n_j is the number of simultaneously served users (service mode), the sum rate of user u is given by

$$\tilde{R}_u = \sum\nolimits_{j:u \in U_j} R_u(n_j) \tag{12}$$

For each user, the *utility* is defined as a function of its sum ergodic rate. The utility function affects the behavior of the scheduler. We consider two special cases of utility and scheduling. To perform maximum sum rate scheduling (MSRS), the utility equals to the ergodic rate,

$$v_u^{MSRS}(\tilde{R}_u) = \tilde{R}_u \tag{13}$$

To consider fairness among users, the proportional fair scheduling (PFS) is introduced with the utility given by

$$v_u^{PFS}(\tilde{R}_u) = \log(\tilde{R}_u) \tag{14}$$

In Orthogonal Frequency Division Multiple Access (OFDMA) systems, the radio resource in each frame is divided into units of a specific number of subcarriers (combined into a subchannel) for a predetermined amount of time (a slot), referred to as resource blocks (RBs). Then the goal of the scheduler is to assign the RBs to users so as to maximize the total utility based on the asymptotic ergodic rates for each user with different service mode, which can be calculated using the formulas given in Section III and stored in a lookup table. A simple user scheduling algorithm for each frame is presented in Algorithm 1. In brief, the algorithm assigns the RBs to a set of users sequentially. In each loop, for different number of simultaneously served users, it selects the user set that provides the highest marginal utility. Then the service mode with largest utility improvement is chosen and corresponding set of users are scheduled on this RB.

Algorithm 1 User Scheduling Algorithm

Initialization: $U_j \leftarrow \emptyset, \forall j \in G, \tilde{R}_u \leftarrow 0, \forall u \in U$
for each RB j **do**
  for each service mode m **do**
    for each user u **do**
      $v'_u \leftarrow v_u(\tilde{R}_u + R_u(m)) - v_u(\tilde{R}_u)$
    end for
    $S_m \leftarrow \arg\max_{S \subseteq U:|S|=m} \sum_{u \in S} v'_u$
    $r_m \leftarrow \sum_{u \in S_m} v'_u$
  end for
  $m^* = \arg\max_m r_m,\ U_j = S_{m^*}$.
  $\tilde{R}_u \leftarrow \tilde{R}_u + R_u(m^*), \forall u \in U_j$
end for

V. Numerical Results

In this section, we present some numerical results to evaluate the performance of different scheduling methods. The assumptions for cell layout, pathloss, and number of antennas are the same as that in previous sections. According to the antenna number setting, at most $M = 6$ users are selected for each frame. We consider the frame structure of the 3GPP-LTE standard with bandwidth of 1.25 MHz partitioned into 6 subchannels each with 180 kHz bandwidth. In this work we assign the slots in one subchannel to the same set of users for simplicity. The users are initial generated at random location with equal number of users in each sector, and randomly move at speeds of 10 to 30 km/h.

We adopt a drop-based simulation method for performance evaluation. In this approach, each drop correspond to a realization of user location, and the large-scale fading parameter as well as the velocity of user movement are assumed to be constant during one drop. Thus, the only varying parameter is the small-scale fading experienced by users. For a given simulation setting (e.g., number of users), 100 drops are run and each drop consists of 1000 frames. In practice, one drop corresponds to one scheduling period, and user location and channel implementation are independent among drops. The final results are averaged over drops.

A. Cluster Throughput

In order to evaluate the effect of scheduling based on asymptotic rate, we also consider the traditional per-frame scheduling methods. Specifically, we simulate the MSRS with

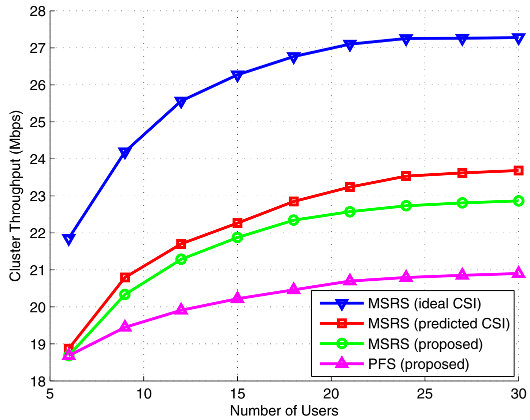

Fig. 3. Cluster throughput versus number of users

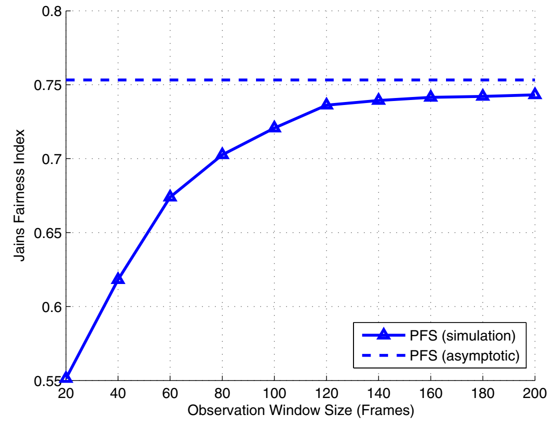

Fig. 4. Jain's fairness index versus observation window size

ideal CSI and predicted CSI, where the user set in each RB is selected in a greedy manner as in [4]. Fig. 3 depicts the cluster throughput of MSRS and PFS strategies versus different number of users. For all schedulers, as the number of user increases, the cluster throughput increases and then saturates due to the effect of multi-user diversity. The per-frame scheduler with ideal CSI outperforms the proposed algorithm by 17% to 30% with the cost of huge overhead. On the other hand, the scheduler based on predicted CSI brings limited improvement even the (predicted) CSI is fed back for all users. That is, with imperfect CSIT, it is efficient enough to schedule the users based only on the first and second order statistics. The effect of multi-user diversity is weaker for PFS as compared to that of MSRS, since it does not always choose users under best channel condition.

B. Fairness

In addition to the throughput maximization, fairness among users is an important concern for scheduler design. The Jain's fairness index defined in [13] is adopted as the performance metric for the proposed asymptotic-rate-based PFS. Here we consider the average fairness index within an observation window consisting of several frames. For an arbitrary drop, the Jain's fairness index versus varying observation window size in frames is depicted in Fig. 4 with 15 users in total. We compare the actual fairness with the constant fairness index calculated using the asymptotic rate of users ($\tilde{R}_u$) after scheduling. It can be observed that the average fairness is relatively low for small window size, since some users may be scheduled on subchannels with poor channel condition. However, as the observation window grows, the fairness index approaches the value "promised" by the asymptotic-rate-based scheduler. That is, the proposed method achieves a good long-term fairness.

VI. Conclusion

In this paper, we address the user scheduling and CSIT feedback problem in network MIMO systems. We propose a two-stage CSIT feedback mechanism and user scheduling methods based on the ergodic rate taking into account the large-scale fading and channel variation. The numerical results indicate that the proposed scheduling methods achieve high user spectral efficiency and fairness with much lower feedback overhead.

References

[1] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser mimo channels," *Signal Processing, IEEE Transactions on*, vol. 52, no. 2, pp. 461–471, 2004.

[2] Y. Hadisusanto, L. Thiele, and V. Jungnickel, "Distributed base station cooperation via block-diagonalization and dual-decomposition," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, 2008, pp. 1–5.

[3] J. Zhang, R. Chen, J. Andrews, A. Ghosh, and R. Heath, "Networked mimo with clustered linear precoding," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 4, pp. 1910–1921, 2009.

[4] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, and B. L. Evans, "Low complexity user selection algorithms for multiuser mimo systems with block diagonalization," *Signal Processing, IEEE Transactions on*, vol. 54, no. 9, pp. 3658–3663, 2006.

[5] S. Sigdel and W. Krzymien, "User scheduling for network mimo systems with successive zero-forcing precoding," in *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, 2010, pp. 1–6.

[6] J. Zhang, J. Andrews, and R. Heath, "Block diagonalization in the mimo broadcast channel with delayed csit," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, 2009, pp. 1–6.

[7] H. Shirani-Mehr, G. Caire, and M. J. Neely, "Mimo downlink scheduling with non-perfect channel state knowledge," *Communications, IEEE Transactions on*, vol. 58, no. 7, pp. 2055–2066, 2010.

[8] C. Min, N. Chang, J. Cha, and J. Kang, "Mimo-ofdm downlink channel prediction for ieee802.16e systems using kalman filter," in *Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE*, 2007, pp. 942–946.

[9] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures," *Information Theory, IEEE Transactions on*, vol. 48, no. 12, pp. 3117–3128, 2002.

[10] J. Wishart, "The generalised product moment distribution in samples from a normal multivariate population," *Biometrika 20A (1-2): 32-52*, 1928.

[11] A. Gupta and D. Nagar, *Matrix variate distributions*. Chapman & Hall/CRC, 2000, vol. 104.

[12] H. G. Borgos, "Partitioning of a line segment," 1997.

[13] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *DEC Research Report TR-301*, 1984.