

Two-Stage Channel Quantization for Scheduling and Beamforming in Network MIMO Systems: Feedback Design and Scaling Laws

Behrouz Khoshnevis, Wei Yu, *Senior Member, IEEE* and Yves Lostanlen, *Senior Member, IEEE*

Abstract—This paper proposes an efficient two-stage channel quantization and feedback scheme for the downlink limited-feedback network multiple-input multiple-output (MIMO) system. In the first stage, the users report their best set of base-station antenna and physical resource block combinations, and the base-stations schedule the best user for each antenna in each resource block. The scheduled users are then polled in the second stage to feedback their quantized channel vectors. This paper proposes an analytical framework to show that, under a total feedback budget of B bits, the number of bits assigned to the second feedback stage should scale as $\log B$, and in quantizing channel vectors from different base-stations, each user should allocate feedback bits in proportion to the channel magnitudes in dB scale. Under these optimized bit allocations, the overall sum rate of the system is shown to scale double-logarithmically with B , linearly with the total number of antennas, and logarithmically with transmit power, thus achieving both multiuser diversity and spatial multiplexing gains under limited feedback. Finally, realistic wireless propagation model of an urban small-cell deployment is used to show that the proposed scheme can approach the performance of a network MIMO system with full channel state information with only modest amount of channel feedback.

Index Terms—Beamforming, channel quantization and feedback, coordinated multipoint (CoMP), limited-feedback systems, multi-input multi-output (MIMO), multiuser MIMO, network MIMO, scheduling

I. INTRODUCTION

Network multi-input multi-output (MIMO), also known as coordinated multipoint (CoMP), is a promising future architecture for wireless cellular systems with densely deployed base-stations [1]. By coordinating multiple cells via high-speed backhaul, a downlink network MIMO system can be thought of as a virtual MIMO array capable of jointly transmitting to remote users simultaneously across the multiple base-stations.

Manuscript submitted on October 12, 2012, revised on March 21, 2013, accepted on May 2, 2013. Materials in this paper have been presented in part at IEEE Global Communications Conference (Globecom), Houston, TX, U.S.A., Dec 2011, in part at IEEE International Conference on Communications (ICC), Ottawa, Canada, June 2012, and in part at IEEE Workshop on Signal Processing Advances for Wireless Communications (SPAWC), Cesme, Turkey, June 2012. This work was supported by the MITACS Elevate program and by SIRADEL Canada.

Behrouz Khoshnevis and Wei Yu are with The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario M5S 3G4, Canada (e-mails: bkhoshnevis@comm.utoronto.ca; weiyu@comm.utoronto.ca). Phone: 416-946-8665. FAX: 416-978-4425.

Yves Lostanlen is with SIRADEL Canada, Suite 500, 120 Eglinton Ave. East, Toronto, Ontario M4P 1E2, Canada (e-mail: yves.lostanlen@ieee.org).

Kindly address correspondence to Wei Yu.

This virtual MIMO architecture can theoretically eliminate intercell interference completely, thus significantly improving the overall network throughput.

The implementation of network MIMO systems, however, is also quite challenging due to its high demand on the network-wide channel state information (CSI) acquisition. CSI can be obtained in time-division-duplex (TDD) systems based on channel reciprocity. But for frequency-division-duplex (FDD) systems, downlink CSI must be estimated at the user terminals then quantized and fed back to the base-stations. Further, this has to be done not only for the direct channel between each user and its own base-station, but also for all the indirect channels from the interfering base-stations. The feasibility of CSI estimation and feedback is limited in practice either by the feedback channel capacity constraint or by the intrinsic delay associated with feedback leading to outdated CSI [2], [3]. Compounding this with the fact that the user scheduling step for a network MIMO system involves a larger pool of potential users, and that the CSI of all these users is needed for scheduling purposes, a conventional implementation of a channel quantization feedback scheme can quickly overwhelm the benefit of network MIMO.

The central idea of this paper is based on the following observation. In a limited-feedback network MIMO system involving many users, on one hand a large number of users need to estimate and feedback their CSI for scheduling purpose in order to benefit from multiuser diversity. Yet on the other hand, only a small number of users are eventually scheduled; hence accurate CSI (which is needed for achieving spatial diversity) is needed only from a few. For this reason, it is inefficient to design a feedback system in which every user quantizes its channel with equal degree of accuracy. Rather, an optimally designed limited-feedback system should acquire coarsely quantized channel feedback from many users in order to benefit from multiuser diversity, then finely quantized channel feedback from only a selected few in order to benefit from spatial diversity for the scheduled users.

The above observation, which is already made for single-cell multiuser MIMO systems [4], [5] (but equally applicable in the network MIMO context) leads to the idea of *two-stage* feedback, first proposed in [4]. This paper proposes a particular two-stage feedback design, named *agile scheduling-beamforming* (ASB), for a network MIMO system employing spatial multiplexing with orthogonal frequency division multiple access (OFDMA). In the first stage of the proposed scheme, coarse channel information is acquired from a large

pool of users by the feedback of each user's best set of spatial-frequency resource blocks for scheduling. The subsequent second stage acquires finer CSI from the scheduled users for the purpose of designing network beamformers. Intuitively, this scheme works because by feeding back each user's best set of resource blocks, the scheduler is already able to select a set of users that form a well-conditioned diagonally dominant channel matrix. Upon acquiring finer CSI in the second stage, the base-stations can then proceed with zero-forcing precoding without excessive power penalty. Note that the remaining multiuser interference due to quantized CSI in zero-forcing precoding can be controlled by allocating sufficient number of feedback bits in the second stage, while multiuser diversity is retained by allocating sufficient number of feedback bits in the first stage.

This paper presents a design, analysis, and performance evaluation framework for the proposed two-stage channel quantization and feedback scheme. Our main contributions are two fold. First, on a theoretical front, this paper makes the design intuition mentioned above precise by showing that the two-stage channel quantization scheme enables a network MIMO system to achieve both multiuser diversity and spatial diversity under limited feedback. This conclusion is derived based on an analytical framework for optimizing the allocation of quantization bits between the two stages and among the different base-station-user pairs in the two-stage feedback scheme under a flat-fading spatial multiplexed network MIMO model. The analysis shows that asymptotically as the total number of quantization bits B and the total number of users tend to infinity, most of the quantization bits should be allocated to the first stage and only $\log B$ bits should be allocated to the second stage. Moreover, the allocation of quantization bits among the different base-station-user pairs should be proportional to the channel magnitude in the dB scale. Using these optimized bit allocation laws, the system sum rate can be shown to scale double logarithmically with B , linear with the number of antennas, and logarithmically with SNR, thus achieving both multiuser diversity and spatial diversity gains.

Second, this paper considers a practical spatial multiplexing and frequency-selective setting under a network MIMO-OFDMA model in which the users are scheduled across both the spatial and the frequency dimensions using a proportional fairness scheduler to allow the maximization of network utility. This paper makes design choices under practical deployment scenarios, and evaluates the overall system using realistic ray-tracing-based wireless propagation models in actual urban environments [6]. The use of these realistic models is a distinct feature of this paper, and it enables the generation of rate maps that illustrate the impact of limited feedback on practical network MIMO deployment. The simulation results using realistic channel models show that the proposed two-stage feedback scheme can already approach the performance of network MIMO with perfect CSI using only modest amount of channel feedback.

A. Literature Review

The acquisition of CSI is a challenging issue for multiuser limited-feedback systems. The design of beamforming and scheduling algorithms that can efficiently utilize the rate-limited feedback channel has attracted a great deal of research [4], [5], [7]–[17]. Many of these works deal with the single-cell multiuser MIMO setting, but are equally applicable to the multicell network MIMO model.

Most of the multiuser scheduling and beamforming algorithms in the literature fall into one of the following two categories. The first line of work, as in [7]–[9], assumes fixed orthogonal beamforming codebooks. The users feedback the index of the beam with the highest signal-to-interference-plus-noise ratio (SINR) along with the corresponding SINR value. The base-station then selects the user with the highest SINR on each beam and eventually uses the same orthogonal beams for downlink transmission. Because of the opportunistic nature of this approach, this paper refers these schemes as *orthogonal/opportunistic beamforming* (OBF).

In the second approach, as in [5], [10], [12], the users explicitly quantize and feedback the channel direction information (CDI) along with certain channel quality indicators (CQI). The base-station then uses this information for scheduling and beamforming. One of the well-known and practically feasible scheduling-beamforming algorithms is the greedy user selection with zero-forcing beamforming [10]. This combined scheduling-beamforming approach is referred to as the *zero-forcing beamforming* (ZFBF) approach in this paper.

Each of these two schemes have their merits and disadvantages. The main advantage of the OBF approach lies in the simplicity of its scheduling algorithm. The OBF scheme, however, suffers from the low accuracy of the quantized channel information. The accuracy of the quantized information can be improved using a variation of the OBF approach as presented in [11], where a collection of orthogonal codebooks is used instead of a single codebook, but it still tends to have inferior performance as compared to ZFBF. The authors of [5] numerically compare the performance of ZFBF with the performance of OBF proposed by [11] in term of the downlink sum rate under a total feedback rate constraint. The comparison reveals that ZFBF achieves a higher sum rate than OBF for realistic systems with finite number of users. As OBF is easier to implement than ZFBF, there appears to be a tradeoff between the superior performance of ZFBF and the lower computational complexity of OBF.

This paper is inspired by both the OBF and the ZFBF approaches. We aim to retain the simplicity of OBF while improving its performance using the ZFBF ideas. Recall that the OBF scheme in [11] involves a library of orthogonal beamforming codebooks constructed at the base-station; each user reports its best beamformers across the codebooks; the best orthogonal codebook among all codebooks in the library is then chosen. The present paper also utilizes a fixed set of orthogonal beamformers just as in OBF. But unlike the one-stage feedback scheme of [8], [11], this paper suggests that a subsequent second stage that refines the channel quantization can significantly improve the overall performance.

More specifically, this paper proposes a two-stage feedback mechanism that achieves a performance comparable to the ZFBF scheme with an implementation complexity comparable to the OBF scheme. The main idea is to decompose the feedback process into two stages that are separately used for scheduling and beamforming. The first stage is similar to the OBF scheme, where the base-station schedules users based on their SINR feedback values. In the second stage, the scheduled users are asked to explicitly quantize and feedback their channel directions. The base-station then uses the quantized directions to form zero-forcing beamforming vectors that are eventually used for downlink transmission.

The idea of two-stage beamforming and scheduling has been previously considered in [4], but the first-stage feedback proposed in [4] consists of a coarse channel quantization by every user, rather than the more efficient scheme proposed here where each user reports only the best set of spatial-frequency resource blocks. Thus, the algorithm in [4] has a user-scheduling stage which has the same computational complexity as ZFBF. In contrast, our approach offers an OBF-like scheduling complexity in the first stage, which is much more efficient than the scheduling process in ZFBF. Yet, simulations show that the performance of the proposed scheme is as good as ZFBF, making the proposed scheme a viable candidate for practical implementation.

This paper analyzes the performance of the proposed channel feedback scheme for the network MIMO systems by assuming a limit on the total number of available feedback bits throughout the network. An analytical framework is developed for optimizing the feedback bit allocation between the two stages and also the bit allocation associated with quantizing direct vs. interference channels. The latter problem can be considered as an extension of the work [18], which considers a single-user per-cell model. The feedback bit allocation problem for the multiuser MIMO system has also been studied in [16], [17], [19].

The beamforming-scheduling algorithms discussed here are deterministic in nature, i.e., there is no probabilistic contention between users in accessing the feedback channel. In particular, in the OFDMA context, this paper treats the frequency resource the same way as the spatial resource, and applies the two-stage scheme to the entire spatial-frequency grid. The approach of this paper is different from and is easier to implement than the contention and threshold-based schemes of [13]–[15], [20] in which the users vie for the scheduling slots, because the creation of a contention channel is not necessarily straightforward. When restricted to OFDMA systems, the first stage of the approach proposed in this paper can be seen to be similar to that in [21].

Finally, there are also many works in the literature considering the interpolation, clustering, and grouping of beamformers across the OFDM tones for channel feedback purposes [22]–[24]. Many of these techniques can be applied on top of the schemes proposed here, so they are not explicitly considered in the present paper. Further, this paper considers the problem of channel quantization only. Channel estimation error and the issue of channel feedback delay [25] are not explicitly treated for simplicity purposes.

B. Organization of the Paper

The rest of this paper is organized as follows. The limited-feedback network MIMO channel model is presented in Section II. The proposed two-stage feedback scheme is first presented for the flat-fading case in Section III, and subsequently optimized for sum-rate maximization in Section IV, where analytical results on the scaling laws of the optimal allocation of feedback bits are also derived. In Section V, the proposed feedback scheme is generalized to the frequency-selective OFDMA setup under weighted sum-rate maximization. Simulation results based on both idealized channel models and on a realistic urban deployment are presented in Section V. The paper concludes with Section VI.

II. SYSTEM MODEL

The system model consists of I base-stations each with M antennas and a total of J single-antenna users uniformly distributed within the network boundaries. We assume a typical deployment with $J \gg MI$. All the users within the network share a common CSI feedback channel with a total capacity of B bits per fading block.

Let us define $\mathcal{I} \stackrel{\text{def}}{=} \{1, 2, \dots, I\}$ and $\mathcal{J} \stackrel{\text{def}}{=} \{1, 2, \dots, J\}$ as the index sets for the base-stations and users. The channel vector from base-station i to user j , denoted by $\mathbf{g}_{ij} \in \mathbb{C}^M$, is defined as

$$\mathbf{g}_{ij} = \sqrt{\alpha_{ij}} \mathbf{h}_{ij}, \quad (1)$$

where $\mathbf{h}_{ij} \in \mathbb{C}^M$ has i.i.d. $\mathcal{CN}(0, 1)$ entries and models the Rayleigh fading component and

$$\alpha_{ij} = \frac{G_0}{1 + (d_{ij}/\delta)^\zeta} \quad (2)$$

denotes the path-loss component [2]. Here, d_{ij} is the distance between base-station i and user j , G_0 is the transmission gain, δ is the 3dB breakpoint distance, and ζ is the path-loss exponent. Finally, the collective channel vector of user j from all base-stations is denoted by $\underline{\mathbf{g}}_j \in \mathbb{C}^{MI}$:

$$\underline{\mathbf{g}}_j = [\mathbf{g}_{1j}^T, \mathbf{g}_{2j}^T, \dots, \mathbf{g}_{Ij}^T]^T. \quad (3)$$

This paper first considers the flat-fading channel model. The frequency-selective case under an OFDMA framework is treated in the second part of this paper.

Consider the system downlink in a network MIMO mode, where the I base-stations cooperatively transmit to multiple users at the same time. With a collective number of MI antennas, a maximum of MI users can be supported in the downlink. The exact method of choosing those users depends on the scheduling algorithm and is described in the next section. Let us denote the scheduling function by $\pi(\cdot) : \mathcal{M} \times \mathcal{I} \rightarrow \mathcal{J}$ and the scheduled users by $\pi(mi)$, where the double index mi belongs to $\mathcal{M} \times \mathcal{I}$ and $\mathcal{M} \stackrel{\text{def}}{=} \{1, 2, \dots, M\}$. The intrinsic receiver noise power is normalized to 1. With this notation, the received signal-to-interference-plus-noise ratio (SINR) at user $\pi(mi)$ can be expressed as $\text{SINR}_{\pi(mi)} = \mathcal{N}_{\pi(mi)} / \mathcal{D}_{\pi(mi)}$,

where

$$\mathcal{N}_{\pi(mi)} \stackrel{\text{def}}{=} \rho \left| \sum_{k \in \mathcal{I}} \mathbf{g}_{k\pi(mi)}^\dagger \mathbf{v}_{k\pi(mi)} \right|^2 \quad (4)$$

$$\mathcal{D}_{\pi(mi)} \stackrel{\text{def}}{=} \rho \sum_{j \neq \pi(mi)} \left| \sum_{k \in \mathcal{I}} \mathbf{g}_{k\pi(mi)}^\dagger \mathbf{v}_{kj} \right|^2 + 1. \quad (5)$$

Here, $\rho \stackrel{\text{def}}{=} \frac{P}{MI}$, P is the transmit power at each base-station, and \mathbf{v}_{kj} is the beamforming vector used at base-station $k \in \mathcal{I}$ to transmit the data corresponding to user $j \in \mathcal{J}$.

Note that for simplicity, we adopt a network-wide power constraint, rather than a more realistic per-base-station or per-antenna power constraint, as otherwise the beamforming design problem would become considerably more complicated. Note also that we do not employ downlink power control, and implicitly assume a fully loaded system that schedules as many users as there are antennas. These are reasonable assumptions for a fully coordinated network MIMO system. Finally, this paper restricts attention to the case of single antenna at the user side. The generalization to the case where the users have multiple antennas is straightforward only if fixed receive beamforming is used with only a single data stream per user. The impact of the optimization of receive beamformers on channel feedback design is not trivial and is left for future studies.

With the above notations, the achievable downlink rate for the scheduled user is given by

$$R_{\pi(mi)} = \log(1 + \text{SINR}_{\pi(mi)}). \quad (6)$$

The objective is to design an efficient user scheduling and beamforming scheme to maximize some utility function of the user rates under the rate constraints on the feedback link. We begin with a sum-rate objective, then subsequently generalize to account for user fairness.

III. TWO-STAGE ASB FOR SUM-RATE MAXIMIZATION

We begin by describing the two-stage channel quantization and feedback scheme for a network MIMO system under a sum-rate maximization objective and on flat-fading channels. The reason for focusing on the sum rate and on flat-fading channels is that the resulting ASB scheme is more amenable to scaling law analysis. Subsequent generalization to the OFDMA setting and to account for fairness among the users is presented in Section V.

Prior to the CSI feedback, the base-stations are assumed to transmit pilot signals that allow the users to estimate their channels from all base-stations. This paper primarily focuses on the design of channel feedback mechanism, so issues such as pilot signal design, pilot bandwidth consumption, and pilot contamination are not treated directly here. For simplicity, the rest of this paper assumes that the channels are estimated perfectly at the mobile users for each fading block, and focuses attention on how to best utilize the rate-limited channel feedback link.

All the users in the network are assumed to share a common feedback channel with a capacity of B bits per fading block. The proposed two-stage feedback mechanism is named ASB,

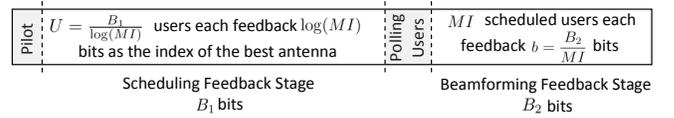


Fig. 1. Two-stage feedback mechanism.

because it is able to schedule the right set of users then feedback their channels in an agile manner. As shown in Fig. 1, the feedback process is divided into two stages, each consuming B_1 and B_2 bits respectively such that $B_1 + B_2 = B$. The two stages are used separately for scheduling and beamforming purposes as described below.

A. Scheduling Process

In the first feedback stage, each user estimates its channel from each antenna of each base-station and feeds back $\log(MI)$ bits as the index of the antenna with the largest channel magnitude along with the magnitude itself. The channel magnitude information is assumed unquantized here¹. When there are a large number of users in the network, each antenna is likely to have multiple users reporting it as their best antenna. In this case, the base-station schedules the user with the largest channel magnitude on each of its antennas.

The proposed per-antenna scheduling scheme is motivated by the orthogonal codebook approach in the literature [8], [11], where each user reports the best beamforming codeword within a library of codebooks. The proposed scheme here essentially adopts a set of trivial orthogonal beamforming codewords (i.e., columns of an identity matrix), with the rationale that the initial choice of orthogonal beamformers is not crucial, since a subsequent second-stage will refine the transmit beamformer design. The use of orthogonal set of beamformers ensures that the scheduled users (each of whom reports one of the orthogonal beamformers as its best) are likely to be almost orthogonal themselves.

It should be noted that the two-stage feedback scheme proposed in this paper differs from that in [4] only in this first stage. The first stage of [4] implements a direct quantization of the channel then searches over the best set of users under ZFBF. One of the main points of this paper is that the simple ranking-based per-antenna scheduling scheme described above can already perform very well, at a considerable complexity saving as compared to [4].

The proposed channel feedback scheme for scheduling is reminiscent of the general distributed source coding setup in which multiple correlated sources are encoded at multiple independent encoders and reconstructed at a centralized decoder, i.e., lossy version of Slepian-Wolf coding. This connection was first made in [26] where the channel quantizer design problem for maximizing multiuser diversity is first posed. In contrast to a typical Slepian-Wolf setup where the sources are correlated, in the multiuser channel quantization problem, the channels for the multiple users are independent, but the

¹This is a reasonable assumption, especially when MI is large, because the magnitudes are scalar, so only a small number of bits are required to quantize the channel magnitude information as compared to the channel direction. This assumption is commonly adopted in the feedback literature; see e.g., [11].

distortion function (which aims to select the best set of users to schedule) is highly coupled across the users. The ranking-based per-antenna feedback scheme is well suited for the scheduling problem, because in some sense the eventual distortion function (which aims to select the best user) is “matched” by each user’s quantization scheme (which selects the best antenna).

Given B_1 feedback bits in the first stage, the total number of users that can participate in this process is given by

$$U = \frac{B_1}{\log(MI)}, \quad (7)$$

which can be randomly chosen from the pool of J users in the network ensuring that each user gets an equal share. The reason for choosing only a subset of users to participate in the first stage is because the incremental benefit of selecting from a larger set of users eventually diminishes. We later optimize B_1 to ensure multiuser diversity is achieved, while still leaving sufficient number of bits for the second stage. Assuming that the user channels are drawn from an i.i.d. distribution, exactly which subset of users is chosen is not important, as far as sum-capacity scaling is concerned.

Denote the set of such users by \mathcal{U} . Let $\mathcal{U}_{mi} \subset \mathcal{U}$ be the set of users who report the m th antenna of base-station i as having the strongest channel. Let $g_{ij,m}$ denote the scalar channel magnitude from the m th antenna of base-station i to user j , i.e., the m th entry of the channel vector \mathbf{g}_{ij} . The user scheduled for the m th antenna of base-station i is determined as follows:

$$\pi(mi) = \arg \max_{j \in \mathcal{U}_{mi}} |g_{ij,m}| \approx \arg \max_{j \in \mathcal{U}} |g_{ij,m}|. \quad (8)$$

We denote the set of MI scheduled users by $\mathcal{U}^* \subseteq \mathcal{U}$. Clearly $|\mathcal{U}^*| = MI$.

Note that the scheduling algorithm described above is based on a sum-rate maximization objective (assuming that interference can be properly controlled in the second stage). We deal with the sum-rate objective first, because it is more amenable to analysis. The user scheduling rule can be modified to account for additional system objectives such as fairness, as developed in Section V of the paper.

B. Beamforming Process

After completing the scheduling process, the MI scheduled users are then polled by the base-stations to explicitly quantize and feedback their CSI in the second stage. This information is used for calculating a zero-forcing downlink network-wide beamformer. Assuming equal bit allocation among the users, given total of B_2 bits in the second stage, each of the MI scheduled users receives a share of

$$b \stackrel{\text{def}}{=} \frac{B_2}{MI} \quad (9)$$

bits for quantizing its channel vectors from all base-stations.

Now, consider an arbitrary user $j \in \mathcal{U}^*$. In general, it should use different codebooks (of potentially different sizes) to quantize its channel from different base-stations. Let \mathcal{C}_{kj} denote the codebook used for quantizing the channel direction $\hat{\mathbf{g}}_{kj} = \mathbf{g}_{kj}/\|\mathbf{g}_{kj}\|$ from base-station k to user j and let b_{kj} be the corresponding number of quantization bits, i.e.,

$b_{kj} \stackrel{\text{def}}{=} \log |\mathcal{C}_{kj}|$, we have

$$\sum_{k \in \mathcal{I}} b_{kj} = b. \quad (10)$$

Also let \mathbf{u}_{kj} denote the quantized direction corresponding to $\hat{\mathbf{g}}_{kj}$:

$$\mathbf{u}_{kj} \stackrel{\text{def}}{=} \arg \max_{\mathbf{w} \in \mathcal{C}_{kj}} |\mathbf{w}^\dagger \hat{\mathbf{g}}_{kj}|, \quad (11)$$

where $(\cdot)^\dagger$ denotes conjugate transpose. With this notation, the quantized version of the overall channel vector $\underline{\mathbf{g}}_j$, denoted by $\tilde{\underline{\mathbf{g}}}_j$, is formed as follows:

$$\tilde{\underline{\mathbf{g}}}_j \stackrel{\text{def}}{=} [\tilde{\mathbf{g}}_{1j}^T, \tilde{\mathbf{g}}_{2j}^T, \dots, \tilde{\mathbf{g}}_{Ij}^T]^T, \quad (12)$$

where

$$\tilde{\mathbf{g}}_{kj} \stackrel{\text{def}}{=} r_{kj} e^{i\theta_{kj}} \mathbf{u}_{kj}, \quad (13)$$

and $r_{kj} \stackrel{\text{def}}{=} \|\mathbf{g}_{kj}\| \cos \phi_{kj}$, $\phi_{kj} \stackrel{\text{def}}{=} \angle(\hat{\mathbf{g}}_{kj}, \mathbf{u}_{kj})$, and $\theta_{kj} \stackrel{\text{def}}{=} \text{phase}(\hat{\mathbf{g}}_{kj}^\dagger \mathbf{u}_{kj})$. It is easy to verify that as the codebook sizes increase and ϕ_{kj} ’s diminish to zero, $\tilde{\underline{\mathbf{g}}}_j$ approaches the actual channel $\underline{\mathbf{g}}_j$.

It should be noted that the mere feedback of the channel directions $\hat{\mathbf{g}}_{kj}$ is not sufficient by itself. The users should also feedback the magnitude and phase information, r_{kj} and θ_{kj} , so that the base-stations can form the quantized vector $\tilde{\underline{\mathbf{g}}}_j$. However, as magnitude and phase are scalar variables and can be quantized and reconstructed using only a few quantization bits, we can assume in the analysis that this information is unquantized and focus on the quantization bit allocation for the channel directions $\hat{\mathbf{g}}_{kj}$. Again, this is a commonly adopted assumption in the literature (see e.g., [10]), and it can be justified by the fact that an optimal quantizer for beamforming vectors of dimension n should allocate most of the bits to quantizing the direction, and only roughly $1/n$ of the total bits to quantizing the magnitude; see [16], [17] for a more detailed analysis.

Note that although each of the channel directions $\hat{\mathbf{g}}_{kj} = \mathbf{g}_{kj}/\|\mathbf{g}_{kj}\|$ is uniformly distributed on the complex unit hypersphere, the direction of the overall channel $\hat{\underline{\mathbf{g}}}_j = \underline{\mathbf{g}}_j/\|\underline{\mathbf{g}}_j\|$ is not uniform. This is due to the fact that the channels from different base-stations have different path-loss components, therefore the entries of the overall channel vector $\underline{\mathbf{g}}_j$ are not identically distributed. As a result one cannot use a single uniform codebook for quantizing $\underline{\mathbf{g}}_j$; instead we should use independent uniform codebooks for quantizing \mathbf{g}_{kj} ’s and form the quantized $\underline{\mathbf{g}}_j$ by stacking up the quantized \mathbf{g}_{kj} ’s.

With the quantized vector channel information available at the base-stations, the unit-norm downlink beams \mathbf{v}_j ’s for the scheduled users $j \in \mathcal{U}^*$ can now be formed as the zero-forcing directions for the quantized channel vectors $\tilde{\underline{\mathbf{g}}}_j$. We write $\mathbf{v}_j = [\mathbf{v}_{1j}^T, \mathbf{v}_{2j}^T, \dots, \mathbf{v}_{Ij}^T]^T$, where \mathbf{v}_{kj} denotes the beam used at base-station k for transmitting user j ’s data.

Consider a user $s \in \mathcal{U}^*$ among the scheduled users. Due to the imperfect CSI, the designed zero-forcing beams cannot completely cancel out the multiuser interference. User s will therefore experience interference from the signals intended for users $j \neq s$. By applying the zero-forcing principle and after some calculations, one can bound such interference in the

average sense as follows:

$$\mathbb{E} \left[\left| \sum_{k \in \mathcal{I}} \mathbf{g}_{ks}^\dagger \mathbf{v}_{kj} \right|^2 \right] \leq \mathbb{E} \left[\sum_{k \in \mathcal{I}} \|\mathbf{g}_{ks}\|^2 \sin^2 \phi_{ks} \right] \quad (14)$$

$$\leq \mathbb{E} \left[\sum_{k \in \mathcal{I}} \|\mathbf{g}_{ks}\|^2 2^{-b_{ks}/(M-1)} \right], \quad (15)$$

where in deriving (15), we have used the bound for the quantization error with respect to the quantization codebook size as described in [12], and the fact that the channels from each base-station are quantized independently, so the contributions from the quantization errors of different channels are added together. The bound in (15) is used in the next section to analyze and optimize the system performance.

IV. ANALYSIS AND OPTIMIZATION OF ASB

The scheduling rule of the ASB scheme discussed in the previous section is designed with a sum-rate objective and for flat-fading channels. This allows analytical scaling laws to be derived for the optimal feedback bit allocation among the two stages and for the overall sum rate. The derivation of the scaling laws depends on an approximation to the sum rate (called virtual sum rate) and a further optimization of ASB under the virtual sum rate objective.

A. Virtual Sum Rate

Based on the rate expression for the individual users (6), the expected sum rate for the multicell system can be expressed as:

$$R = MI \mathbb{E} \log(1 + \text{SINR}_{\pi(mi)}). \quad (16)$$

This sum-rate expression is not so amenable to analysis. For analytic tractability, this paper first upper bounds this sum-rate expression using Jensen's inequality as:

$$R \leq MI \log(1 + \mathbb{E}[\text{SINR}_{\pi(mi)}]), \quad (17)$$

then uses a modified performance measure, referred to as the *virtual sum rate*. First, define the *virtual SINR* as

$$\widetilde{\text{SINR}} = \frac{\mathbb{E}[\mathcal{N}_{\pi(mi)}]}{\mathbb{E}[\mathcal{D}_{\pi(mi)}]}, \quad (18)$$

where $\mathcal{N}_{\pi(mi)}$ and $\mathcal{D}_{\pi(mi)}$ are defined in (4) and (5). Then, the virtual sum rate can be defined as

$$\widetilde{R} = MI \log(1 + \widetilde{\text{SINR}}). \quad (19)$$

Clearly \widetilde{R} is only an approximation to the actual sum rate. We provide a justification for the approximation in the Appendix.

We further approximate and bound the expressions for $\mathcal{N}_{\pi(mi)}$ and $\mathcal{D}_{\pi(mi)}$ as follows. First, we argue that the scheduling rule in (8) guarantees the scheduled users' channels to be almost orthogonal in \mathbb{C}^{MI} . Assuming reasonably accurate CSI feedback in the second stage, the corresponding zero-forcing beams are therefore expected to be almost aligned with the users' channels. We can therefore use the following

approximation for $\mathcal{N}_{\pi(mi)}$:

$$\mathcal{N}_{\pi(mi)} = \rho \left| \sum_{k \in \mathcal{I}} \mathbf{g}_{k\pi(mi)}^\dagger \mathbf{v}_{k\pi(mi)} \right|^2 = \rho \left| \mathbf{g}_{\pi(mi)}^\dagger \mathbf{v}_{\pi(mi)} \right|^2 \quad (20)$$

$$\approx \rho \left\| \mathbf{g}_{\pi(mi)} \right\|^2 = \rho \sum_{k \in \mathcal{I}} \left\| \mathbf{g}_{k\pi(mi)} \right\|^2. \quad (21)$$

Therefore,

$$\mathbb{E}[\mathcal{N}_{\pi(mi)}] \approx \rho \sum_{k \in \mathcal{I}} \mathbb{E}[\lambda_{k\pi(mi)}], \quad (22)$$

where

$$\lambda_{k\pi(mi)} \stackrel{\text{def}}{=} \left\| \mathbf{g}_{k\pi(mi)} \right\|^2. \quad (23)$$

Also, considering (15) and (5) and summing over $MI - 1$ scheduled indices $j \neq \pi(mi)$, we have

$$\mathbb{E}[\mathcal{D}_{\pi(mi)}] < 1 + \rho MI \mathbb{E} \left[\sum_{k \in \mathcal{I}} \left\| \mathbf{g}_{k\pi(mi)} \right\|^2 2^{-b_{k\pi(mi)}/(M-1)} \right]. \quad (24)$$

Combining (22) and (24), we achieve the following lower bound for the virtual SINR:

$$\widetilde{\text{SINR}} \geq \frac{\rho \sum_{k \in \mathcal{I}} \mathbb{E}[\lambda_{k\pi(mi)}]}{1 + \rho MI \mathbb{E} \left[\sum_{k \in \mathcal{I}} \lambda_{k\pi(mi)} 2^{-b_{k\pi(mi)}/(M-1)} \right]}. \quad (25)$$

We can now use the virtual SINR lower bound in (25) as the objective for optimizing ASB. In particular, we optimize the feedback bit allocation in ASB in two levels:

- 1) CSI quantization bit allocation problem: optimizing the number of bits $b_{k\pi(mi)}$ used in the second stage to quantize the channels from different base-stations to each user as a function of channel magnitudes $\left\| \mathbf{g}_{k\pi(mi)} \right\|$ subject to the constraint $\sum_{k \in \mathcal{I}} b_{k\pi(mi)} = b = B_2/(MI)$.
- 2) Beamforming-scheduling bit allocation problem: optimizing the bit allocation between the first and the second stages subject to the constraint $B_1 + B_2 = B$.

The optimized bit allocation will also reveal the scaling of the sum-rate with respect to the total feedback bits.

B. CSI Quantization Bit Allocation

The optimal bit allocation problem in the second stage, which can be stated as a maximization of the SINR lower bound (25) subject to a sum bit constraint, can also be equivalently stated as a minimization of the residual interference in the denominator of (25):

$$\min_{\substack{b_{k\pi(mi)} \\ 1 \leq k \leq I}} \sum_{k \in \mathcal{I}} \lambda_{k\pi(mi)} 2^{-b_{k\pi(mi)}/(M-1)} \quad (26)$$

$$\text{s.t.} \quad \sum_{k \in \mathcal{I}} b_{k\pi(mi)} = b, \quad b_{k\pi(mi)} \geq b_0, \quad (27)$$

where $b_0 \stackrel{\text{def}}{=} \log M$ is the minimum number of bits required for quantizing an M -dimensional channel vector. Using a Lagrangian technique, it is easy to see that the above optimization problem has a waterfilling-type solution as follows:

$$b_{k\pi(mi)} = b_0 + (M-1) \left(\log \frac{\lambda_{k\pi(mi)}}{\mu} \right)^+, \quad (28)$$

where the Lagrange multiplier $\mu > 0$ is chosen such that $\sum_{k \in \mathcal{I}} b_{k\pi(mi)} = b$. The bit allocation rule (28) shows that

the scheduled user $\pi(mi)$ should allocate its CSI quantization bits in proportion to the logarithm of the channel magnitudes $\lambda_{k\pi(mi)}$.

If B_2 and therefore b are large enough, this bit allocation can be approximated as follows:

$$b_{k\pi(mi)} \approx \frac{b}{I} + (M-1) \log \frac{\lambda_{k\pi(mi)}}{\mathcal{G}(\lambda_{k\pi(mi)}|_{k \in \mathcal{I}})}, \quad (29)$$

where $\mathcal{G}(\lambda_{k\pi(mi)}|_{k \in \mathcal{I}})$ denotes the geometric mean of $\{\lambda_{k\pi(mi)}\}_{k \in \mathcal{I}}$. Thus, in the second stage, the number of bits devoted to quantizing the channels to different base-stations should be a constant plus a term which is proportional to the expected channel magnitude in dB scale.

C. Approximating the Virtual SINR

We now use a series of approximations to simplify the virtual SINR expression in order to derive the optimal bit allocation between the scheduling and beamforming stages. First, by substituting the optimal bit allocation (29) back into the objective function in (26), we obtain the following expression for residual interference:

$$\begin{aligned} & \sum_{k \in \mathcal{I}} \lambda_{k\pi(mi)} 2^{-b_{k\pi(mi)}/(M-1)} \\ &= I \mathcal{G}(\lambda_{k\pi(mi)}|_{k \in \mathcal{I}}) 2^{-b/((M-1)I)} \end{aligned} \quad (30)$$

Substitute the above into the denominator of the virtual SINR lower bound in (25), we see that the denominator satisfies

$$\begin{aligned} & 1 + \rho MI^2 \mathbb{E} \left[\mathcal{G}(\lambda_{k\pi(mi)}|_{k \in \mathcal{I}}) \right] 2^{-b/((M-1)I)} \\ & \leq 1 + \rho MI^2 \mathcal{G}(\mathbb{E}[\lambda_{k\pi(mi)}] |_{k \in \mathcal{I}}) 2^{-b/((M-1)I)}, \end{aligned} \quad (31)$$

where we have used the Jensen's inequality and the concavity of the geometric mean with respect to its arguments. If we define $\beta_{ki} \stackrel{\text{def}}{=} \mathbb{E}[\lambda_{k\pi(mi)}]$, and combine (31) and (25), we obtain the following lower bound for the virtual SINR:

$$\widetilde{\text{SINR}} \geq \frac{\rho \sum_{k \in \mathcal{I}} \beta_{ki}}{1 + \rho MI^2 \bar{\beta}_i 2^{-b/((M-1)I)}}, \quad (32)$$

where $\bar{\beta}_i \stackrel{\text{def}}{=} \mathcal{G}(\beta_{ki}|_{k \in \mathcal{I}})$ for notation convenience.

Further, since

$$\lambda_{k\pi(mi)} = \|\mathbf{g}_{k\pi(mi)}\|^2 \geq |\mathbf{g}_{k\pi(mi),m}|^2. \quad (33)$$

we have $\beta_{ki} = \mathbb{E}[\lambda_{k\pi(mi)}] \geq \chi_{ki}$, where $\chi_{ki} \stackrel{\text{def}}{=} \mathbb{E}[|\mathbf{g}_{k\pi(mi),m}|^2]$. Considering this notation and noting that the lower bound in (32) is monotonic in β_{ki} , we have

$$\widetilde{\text{SINR}} \geq \frac{\rho \sum_{k \in \mathcal{I}} \chi_{ki}}{1 + \rho MI^2 \bar{\chi}_i 2^{-b/((M-1)I)}}, \quad (34)$$

where $\bar{\chi}_i$ is the geometric mean of $\{\chi_{ki}\}_{k \in \mathcal{I}}$. Finally, according to the scheduling rule in (8), we have $\chi_{ii} \geq \chi_{ki}$, $\forall k \in \mathcal{I}$. Therefore, $\bar{\chi}_i \leq \chi_{ii}$ and clearly $\sum_{k \in \mathcal{I}} \chi_{ki} \geq \chi_{ii}$. Combining these with the expression in (34), we obtain a simplified lower bound for the virtual SINR:

$$\widetilde{\text{SINR}} \geq \frac{\rho \chi_{ii}}{1 + \rho MI^2 \chi_{ii} 2^{-b/((M-1)I)}}. \quad (35)$$

We plan to maximize this lower bound (35) with respect to the bit allocation between the scheduling and beamforming

stages subject to the constraint $B_1 + B_2 = B$. For this purpose, we go back to the scheduling rule in (8) and note that since the scheduled user is chosen among $|\mathcal{U}|$ users, i.e., it is approximately true that

$$|\mathbf{g}_{i\pi(mi),m}| = \max_{j \in \mathcal{U}} |\mathbf{g}_{ij,m}|, \quad (36)$$

where \mathcal{U} is the set of users participating in the first feedback stage with $|\mathcal{U}| = B_1/\log(MI)$, then χ_{ii} , being the maximum among $|\mu U|$ random variables, must scale as $\ln(B_1/\log(MI))$.

But, χ_{ii} appears in both the numerator and the denominator of (35). To obtain a lower bound on $\widetilde{\text{SINR}}$, we note that the function $\frac{x}{1+ax}$ is monotonic in x , and bound χ_{ii} as follows. Define \mathcal{U}_i as the subset of users in \mathcal{U} that reside in cell i , then clearly

$$|\mathbf{g}_{i\pi(mi),m}| \geq \max_{j \in \mathcal{U}_i} |\mathbf{g}_{ij,m}| \geq \sqrt{\alpha_d} \max_{j \in \mathcal{U}_i} |\mathbf{h}_{ij,m}|, \quad (37)$$

where

$$\alpha_d = \frac{G_0}{1 + (C/\delta)^\zeta} \quad (38)$$

is the highest path-loss possible within the boundaries of cell i , where C is the cell radius. By taking the expectation of both sides of (37), we obtain

$$\chi_{ii} = \mathbb{E}[|\mathbf{g}_{i\pi(mi),m}|^2] \geq \alpha_d \mathbb{E} \left[\max_{j \in \mathcal{U}_i} |\mathbf{h}_{ij,m}|^2 \right] \quad (39)$$

$$\stackrel{(a)}{\approx} \alpha_d \ln |\mathcal{U}_i| \stackrel{(b)}{=} \alpha_d \ln \frac{B_1}{I \log(MI)}, \quad (40)$$

where (b) is due to the fact that $|\mathcal{U}_i| = |\mathcal{U}|/I$ and the approximation (a) uses the result in [7] regarding the behavior of the maximum of multiple independent exponential random variables.

Now, noting that the lower bound in (35) is monotonic in χ_{ii} and using (40) and (9), we obtain the following lower bound on the virtual SINR in terms of B_1 and B_2 :

$$\widetilde{\text{SINR}} \geq \frac{\rho \alpha_d \ln \left(\frac{B_1}{I \log(MI)} \right)}{1 + \rho MI^2 \alpha_d \ln \left(\frac{B_1}{I \log(MI)} \right) 2^{-\frac{B_2}{M(M-1)I^2}}}. \quad (41)$$

The above expression illustrates the tradeoff between allocating feedback bits to the first stage vs. the second stage. More bits for the first-stage enlarges the set of candidate users, resulting in a higher expected signal power due to multiuser diversity (although it also increases the expected interference power). More bits for the second stage improves orthogonality, hence it reduces interference. The former scales as $\ln(B_1)$ due to the scaling behavior in selecting the best channel among a set of independent random variables. The latter scales as $2^{-\alpha B_2}$ due to the behavior of the rate-distortion tradeoff in quantization theory. The above expression suggests that an optimal (B_1, B_2) exists for maximizing the approximate virtual SINR.

D. Optimizing Scheduling-Beamforming Bit Allocation

We can now directly maximize the SINR lower bound (41) under a total feedback bit constraint B :

$$\begin{aligned} & \max \quad \widetilde{\text{SINR}} \\ & \text{s.t.} \quad B_1 + B_2 = B. \end{aligned} \quad (42)$$

Using the Lagrangian multiplier method, we see that the solution to this optimization problem is at where

$$\frac{\partial \widetilde{\text{SINR}}}{\partial B_1} = \frac{\partial \widetilde{\text{SINR}}}{\partial B_2}. \quad (43)$$

After some algebraic manipulation, it is not difficult to arrive at the following condition for the optimal values of B_1 and B_2 :

$$\begin{cases} \frac{\rho\alpha_d \ln 2}{M-1} B_1 (\ln B_1 - \kappa)^2 = 2^{\frac{B_2}{M(M-1)I^2}} \\ B_1 + B_2 = B \end{cases} \quad (44)$$

where $\kappa = \ln(I \log(MI))$.

We can further simplify the bit allocation results by assuming an asymptotic regime where $B \rightarrow \infty$. In doing so, we get the following asymptotic bit allocations:

$$B_1 \doteq B \quad (45)$$

$$B_2 \doteq M(M-1)I^2 \log B, \quad (46)$$

where the notation $x \doteq y$ means $\lim_{B \rightarrow \infty} x/y = 1$. These results show that as the total feedback rate increases, higher percentage of bits should be used for the scheduling stage. In particular, the ratio B_2/B behaves as $\log B/B$ as $B \rightarrow \infty$.

The above result essentially states that one only needs $B_2 \doteq M(M-1)I^2 \log B$ bits to drive the interference term in the SINR expression to zero. This can be seen by substituting the asymptotic bit allocations into the lower bound (41). It is easy to verify that the interference term in (41) goes down to zero as $B \rightarrow \infty$, since

$$\rho M I^2 \alpha_d \ln \left(\frac{B_1}{I \log(MI)} \right) 2^{-\frac{B_2}{M(M-1)I^2}} \doteq \rho M I^2 \alpha_d \frac{\log B}{B}. \quad (47)$$

All the rest of the feedback bits should be devoted to the first stage.

With this asymptotically optimal bit allocation, the SINR now scales as

$$\widetilde{\text{SINR}} \doteq \rho \alpha_d \ln \left(\frac{B}{I \log(MI)} \right). \quad (48)$$

Substituting this into the virtual rate expression (19), we obtain

$$\widetilde{R} \doteq MI \log(\rho \alpha_d) + MI \log \log B. \quad (49)$$

This \widetilde{R} is proportional to MI but also scales double-logarithmically with B . The linear scaling with MI suggests that the proposed scheme achieves the spatial multiplexing gain of the network MIMO system. The double-logarithmic scaling with B suggests that the multiuser diversity gain, corresponding to selecting the best user among B candidate users, is also realized under the proposed limited-feedback scheme. Note that since each user requires at least one bit feedback, the $\log \log B$ is the best possible scaling in B under limited feedback.

E. Summary of Scaling Law Analysis

To summarize the scaling law process, we start with the virtual SINR bound (25) and optimize the per-channel quantization bit allocation in order to minimize the residual interference in the denominator of (25). The resulting interference is expressed in terms of the geometric mean of the direct and

interfering channel magnitudes in (30), which is then upper-bounded by the magnitude of the direct channel in (34) and (35). We then use the scheduling rule to express the direct channel magnitude in terms of the number of scheduling feedback bits in (40). This leads to the bound in (41), which in turn is used to optimize the bit allocation between the two feedback stages and to derive the overall system performance.

Our main results are:

- 1) We derive asymptotic scheduling-beamforming bit allocation laws (45) and (46) showing that most of the feedback bits should be assigned to the scheduling stage. The fraction of bits assigned to the beamforming feedback stage should scale as $\log B/B$.
- 2) We derive the CSI bit allocation rules (28) and (29), which show that in the second stage, each scheduled user should allocate its feedback bit budget in proportion to the logarithm of the channel magnitudes from base-stations.
- 3) Finally, we use the optimized bit allocations to show that the overall sum-rate is proportional to the total number of base-station antennas, and scales double-logarithmically with B as $B \rightarrow \infty$.

These conclusions are reached using a series of approximations on the SINR expression and via asymptotic analysis. They are meant to convey the scaling law only. In spite of the many simplifications and approximate bounds used in the derivation process, simulation results presented in Section VI show that these scaling laws give remarkably accurate predictions of the optimal bit allocations.

V. ASB FOR NETWORK MIMO-OFDMA SYSTEMS

The preceding section focuses on the theoretical performance scaling for the proposed ASB scheme. In order to derive tractable results, simplifying assumptions (along with quite a few approximations) have been made. In particular, the sum rate is used as the system objective, so no fairness is considered. Also, a flat fading channel model is assumed. Further, the optimal bit allocation law is a consequence of the statistics of the maximum of a set of independent channel magnitudes, so it highly depends on the fading statistics of the underlying channel.

The goal of this section is to generalize the ASB algorithm to a practical network MIMO setup with frequency-selective channels, which also requires fair rate provisioning among the users. Toward this end, a combined spatial multiplexing and OFDMA setting for the MIMO frequency-selective channel is considered in which the users are scheduled across both the spatial and frequency dimensions. Moreover, a proportional fairness scheduler is incorporated to allow the maximization of a suitable network utility. Given the complexity of such a system, we no longer hope to derive analytical results but instead provide design guidelines for the optimal allocation of quantization bits under practical deployment scenarios. A feature of this paper is that our overall design is evaluated using realistic wireless propagation models in actual urban environments [6] to assess the impact of limited feedback on practical network MIMO-OFDMA deployments.

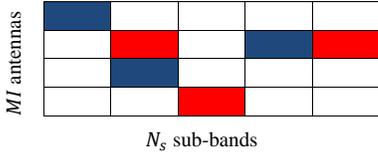


Fig. 2. CSI matrix. Two different colors show the best three (antenna,sub-band) pairs for two different users.

A. Network MIMO-OFDMA Model

Consider a MIMO-OFDMA transmission system across multiple base-stations, in which the total system bandwidth is divided into sub-carriers, and sub-carriers within the coherence bandwidth are grouped into physical resource blocks (PRBs). To reduce the channel feedback rate, every K adjacent PRBs are further grouped into what we call a *sub-band*. Assuming that there are a total number of N PRBs, this results in $N_s = N/K$ sub-bands in the system. The sub-band size determines the resolution of the users' CSI feedback over the frequency domain. We now propose the generalization of the two-stage agile scheduling beamforming algorithm for the network MIMO-OFDMA system as follows.

B. First-Stage CSI Feedback

The goal of first-stage CSI feedback is to obtain a coarse quantization of the channels from a large pool of users for scheduling purpose. As before, we adopt a mechanism in which a set of orthogonal beamforming vectors are fixed across the entire network MIMO system, then each user reports its best set of sub-channels over all spatial and frequency dimensions. Again, the initial set of orthogonal beamformers can be arbitrary, so a simple identity matrix can serve as the initial beamforming matrix, in which case each user's spatial-frequency CSI is simply a two-dimensional matrix consisting of the channel responses from each of the antennas at each of the base-stations on each of the sub-bands. This is an $MI \times N_s$ matrix as shown in Fig. 2.

To efficiently send a coarse quantization of each user's sub-channels to the base-stations, we propose a first-stage quantization mechanism in which U users in the network are asked to rank their spatial-frequency sub-channels in the CSI matrix according to some scheduling criterion, then send back the indices and some measure of qualities of the S best sub-channels. This is in contrast to the previous flat-fading case where $S = 1$. This feedback design is justified by the fact that at most one user will eventually be scheduled in each of the spatial-frequency sub-channels, so for each user only the CSI from the best few sub-channels would likely be useful. For example, Fig. 2 shows the best $S = 3$ entries for two different users marked by different colors. With this setup, the first stage consumes

$$B_1 = US \log_2(MIN_s) \quad \text{bits} \quad (50)$$

for feeding back the sub-channel indices. This information is used for the scheduling of users on the available spatial-frequency resource blocks. The scheduling procedure needs to take fairness and user priorities into account, and is described in detail below.

C. Scheduling Criterion

Most practical scheduling algorithms are based on selecting the user with the maximum weighted achievable rate, where the weights w_u are chosen according to some higher level protocols. For example, in a proportional fairness scheduler, the weights are set to be reciprocal to each user's respective long-term average rate. The weights can also depend on the backlog queue length of each user. Thus, an efficient use of the first-stage feedback is to let each user rank sub-channels in the CSI matrix according to an estimate of the weighted rate.

The proposed scheduling mechanism uses the following quantity as an estimate of weighted rate. Let h_{misu} denote the scalar channel from antenna m on base-station i to user u on sub-band s . The approximate weighted rate for the scheduled user at the antenna-sub-band pair (mi, s) is:

$$\gamma_{misu} = w_u \log \left(1 + \frac{\rho |h_{misu}|^2}{1 + \rho \sum_{(nj) \neq (mi)} |h_{njsu}|^2} \right) \quad (51)$$

where w_u is the weight for user u . In essence, this surrogate rate expression assumes that useful information is carried only on one antenna, and all other antennas contribute to the interference. However imperfect such an estimate of the weighted rate may be, it nonetheless gives some *relative* measure of the desirability of scheduling a particular user on the sub-channel in question. Given that h_{misu} is the only information each user has access to, this choice of channel quality measure is reasonable.

With the surrogate rate defined in (51), the scheduling operation can now be carried out as follows:

$$\pi(mi, s) = \arg \max_{u \in \mathcal{U}_{mi,s}} \gamma_{misu} \quad (52)$$

where $\mathcal{U}_{mi,s}$ is the set of users who have reported on the (antenna, sub-band) pair (mi, s) . This is in contrast to the previous sum-rate maximization case where the scheduler simply selects the best antenna.

D. Second-Stage CSI Feedback

After the users are scheduled in the first stage, the scheduled users are polled for the quantized values of their channel vectors from all the base-stations in the second stage. For this purpose, the total number of bits available for the second stage, $B_2 = B - B_1$, is divided equally among the N_s sub-bands. The B_2/N_s bits on each sub-band is then divided equally among the users scheduled on that sub-band. For clarity, assume an arbitrary sub-band with U_s users scheduled on it (note that $|U_s| < MI$ as we have a total of MI antennas on each sub-band). Each user on this sub-band can therefore use a total of $b = B_2/(N_s U_s)$ bits for quantizing its I channel vectors from each base-station, each of which is an $M \times 1$ complex vector. Then, the quantized channel information is used to design the zero-forcing beamforming vectors for the scheduled users by assuming that the quantized channel vectors in the second stage are accurate.

To complete the specifications of the second stage, we need to determine how each user allocates its b bits for quantizing the I channel vectors from different base-stations. For this

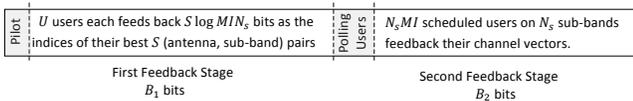


Fig. 3. Two-stage CSI feedback scheme

purpose, we use the bit allocation rule (29) derived in Section III, where the quantization bits are allocated in proportion to the norm of the channel vector in logarithmic scale. As shown before, this approach minimizes the residual interference caused by the inevitable inaccuracy of the beamforming vector design due to limited CSI.

Finally, we need to devise an optimal allocation of feedback bits between the first and the second stages. The theoretical scaling law of Section IV shows that most bits should be allocated to the first stage, and the second stage should use only a portion of $\log(B)/B$ bits. But this derivation does not account for the need to feedback along the frequency dimensions, and it is an asymptotic result in the limit as $B \rightarrow \infty$. A practical implementation likely needs to search for the optimal bit allocation heuristically.

Fig. 3 summarizes the proposed two-stage channel feedback scheme for agile scheduling-beamforming in network MIMO-OFDMA systems.

VI. NUMERICAL RESULTS

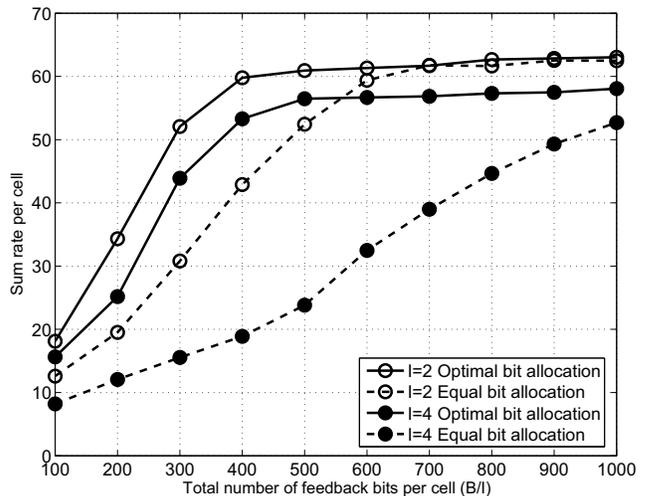
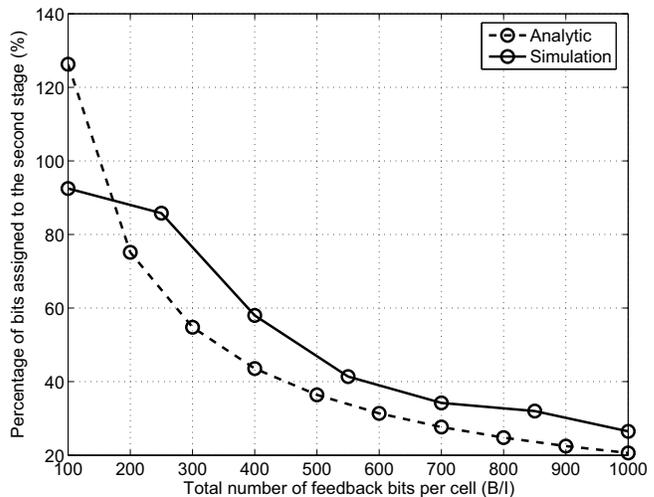
We now present numerical simulation to validate the asymptotic scaling law, and to evaluate the performance of the ASB algorithm in practical settings.

A. Sum Rate Performance in the Flat-Fading Case

We begin by simulating a multicell deployment with a cell radius of $C = 0.5$ km, a 3dB breakpoint distance of $\delta = 0.1$ km, a path-loss exponent $\zeta = 3.8$, a transmission gain of $G_0 = 20$ dB, and $M = 4$ antennas per base-station. The number of cells, the number of users per cell, and the transmit SNR vary throughout the simulations and are specified separately for each simulation result.

We start with the CSI quantization bit allocation problem discussed in Section IV-B. Fig. 4 compares the downlink sum rate per cell as a function of the total number of feedback bits per cell, where we compare the optimal CSI quantization bit allocation rule (28) with the case where one allocates equal number of bits for quantizing vector channels from different base-stations. As the figure shows, the optimized bit allocation provides significant gains in the sum rate. Furthermore, the gain in sum rate increases with the number of cells.

The next simulation result corresponds to the beamforming-scheduling bit allocation problem addressed in Section IV-D. In this regard, Fig. 5 plots the optimal percentage of bits that should be allocated to the second feedback stage for maximizing the sum rate, i.e., B_2/B , and compares it with the analytic bit allocation result in (46). As the figure verifies, the ratio B_2/B behaves as $\log B/B$ as B increases. Thus, in spite of the number of approximations used in deriving our main theoretical result, the resulting scaling law agrees remarkably well with the practical simulation. Note that a system with 300

Fig. 4. CSI quantization bit allocation: optimal vs. equal bit allocation for $I = 2$ and 4 cells, SNR = 15dB, and 300 users per cell.Fig. 5. Beamforming-scheduling bit allocation: percentage of bits assigned to the beamforming feedback stage, i.e. B_2/B , for $I = 2$ cells, SNR = 15dB, and 300 users per cell.

users per cell is simulated here in order to validate the result in the asymptotic regime.

Next, we address the system performance in terms of the downlink sum rate. Fig. 6 compares the performance of the proposed two-stage feedback mechanism (with optimized bit allocations) with the performance of the conventional ZFBF scheme with greedy user selection (with optimized number of users participating in the feedback process). The figure shows a significant improvement in the downlink sum rate. The reason is that the two-stage scheme allows more users to participate in the feedback process and is therefore much more efficient in finding users with better channel conditions, which in this case are the users closer to the base-stations.

It is important to note that all the rate improvements offered by the two-stage scheme are achieved in spite of its significantly lower computational complexity. Our simulation experience indicates that the scheduling part of the two-stage scheme can be performed about 8-10 times faster than that of the one-stage ZFBF scheme. Note that the complexity

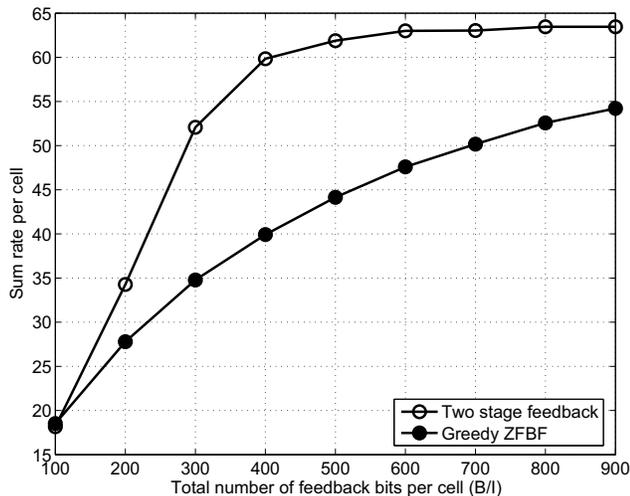


Fig. 6. Two-stage beamforming-scheduling algorithm vs. ZFBF with greedy user selection for $I = 2$ cells, SNR = 15dB, and 300 users per cell.

advantage of ASB over ZFBF also implies that the two-stage feedback scheme proposed in this paper is much more efficient than the two-stage scheme of [4]. This is because the scheduling stage of [4] essentially uses ZFBF.

B. Network MIMO-OFDMA Under Proportional Fairness

The effectiveness of the two-stage channel quantization and feedback scheme is further evaluated for an OFDMA system for a macrocell network MIMO deployment with 7 cells and 30 users per cell over 20MHz bandwidth. The available bandwidth is divided into 18 sub-bands, each of size 1.08 MHz, corresponding to about $1 \mu\text{s}$ delay spread. The base-stations are equipped with two antennas each, and are 1km apart. In this simulation, each user feeds back its best two subbands, i.e., the parameter $S = 2$. Figs. 7 and 8 show the achievable downlink sum rate per cell as a function of total number of feedback bits per cell and transmit power. Fig. 9 shows the cumulative density function of the achieved user rates under proportionally fair scheduling. The conventional scheme here refers to the ZFBF scheme with greedy user selection in which each user quantizes its channel with a fixed number of bits. The perfect CSI scheme refers to the scheme in which perfect CSI is known in both stages of the two-stage ASB process.

It is clear that the two-stage ASB algorithm significantly outperforms the conventional ZFBF scheme. With 3000 feedback bits per cell per coherence time and the transmit power of 39dBm per base-station as shown in Fig. 8, (which at 1ms coherence time, corresponds to a very small fraction of the downlink sum rate), the network MIMO with imperfect CSI can already approach the perfect CSI case. Note that the feedback bits here do not account for channel magnitude/phase quantization for either ASB or ZFBF.

C. Network MIMO in a Realistic Urban Environment

Finally, to validate the effectiveness of the two-stage ASB algorithm on realistic deployments, we use a ray-tracing-based model [6] to simulate an urban small-cell deployment

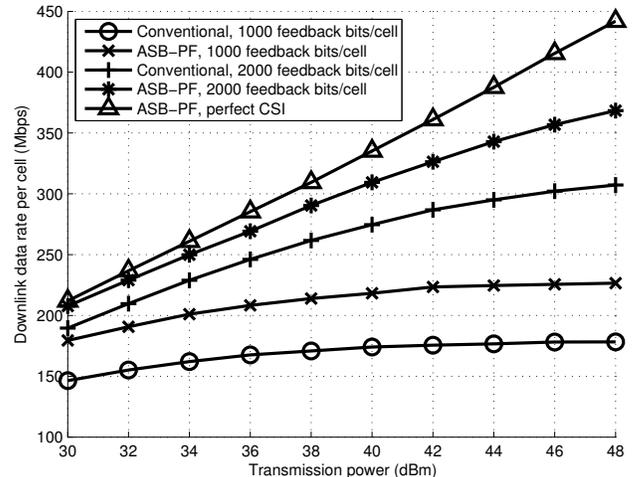


Fig. 7. Sum rate vs transmit power under, proportional fairness scheduling for $I = 7$ cells, and 30 users per cell.

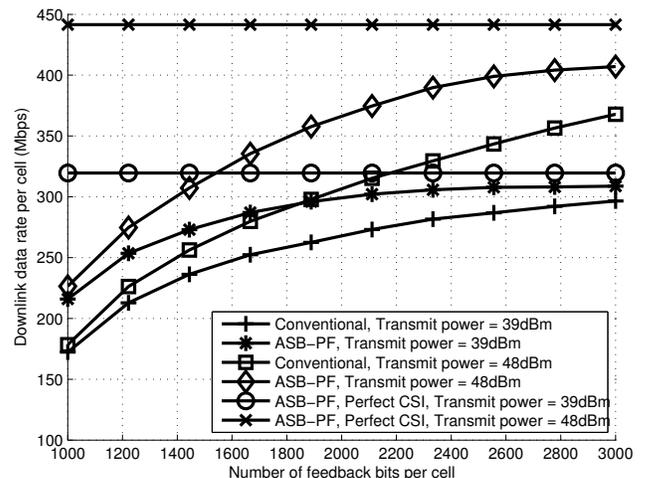


Fig. 8. Sum rate vs number of feedback bits, under proportional fairness scheduling for $I = 7$ cells, and 30 users per cell.

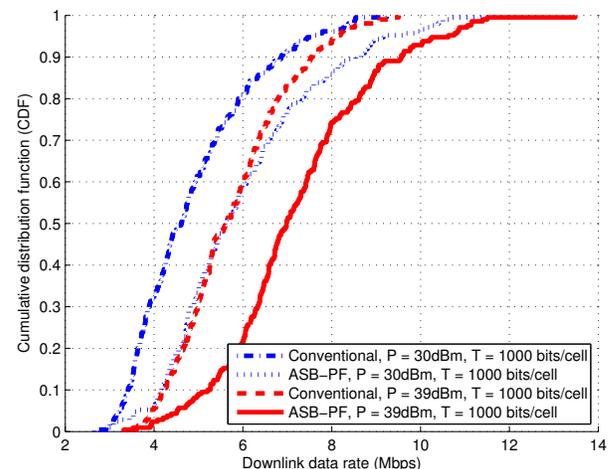


Fig. 9. Cumulative density function of user rates under proportional fairness scheduling for $I = 7$ cells, and 30 users per cell.

TABLE I
PARAMETERS OF A REALISTIC URBAN SMALL-CELL DEPLOYMENT

Layout	240m×240m
Number of Cells	7
Number of Users/Cell	7
No. of Antennas per Base-station	2
No. of Antennas per Remote User	1
Antenna Height at Base-station	8m
Antenna Height at Remote User	1.5m
Base-to-Base Distance	100m
Carrier Frequency	763MHz
Frequency Reuse	1
Bandwidth	10 MHz
OFDM Subcarrier	15 kHz
PRB Bandwidth	180 kHz
Transmit Power	20 dBm
Antenna Gain	15 dBi
Background Noise	-170dBm/Hz

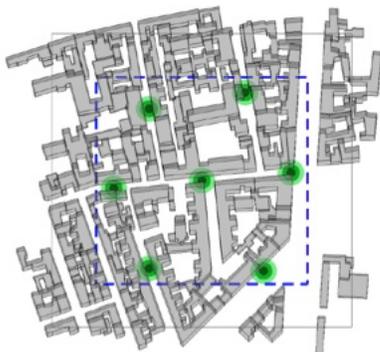


Fig. 10. A section of downtown Paris with small-cell deployment.

with 7 base-stations with two antennas per base-station and 7 users per cell over an 240m by 240m area. The simulation parameters are listed in Table I. As shown in the SINR map in Fig. 11, without network MIMO only the areas with a line of sight from the base-station get appreciable SINR, whereas much of elsewhere experiences poor SINR.

The situation improves significantly with network MIMO. As shown in Fig. 12, with 8000 feedback bits per cell, most of the areas in the cell are now capable of supporting 10-20Mbps per user. This can be compared visually with the perfect-CSI case, which is shown in Fig. 13. In fact, with perfect CSI, the simulated per-cell sum rate is about 245Mbps, and the two-stage ASB algorithm with 8000 feedback bits already achieves 88% of the sum rate with perfect CSI, as shown in Table II. The cumulative density functions of rates plotted in Fig. 14 further show that the loss from the perfect CSI case is small and is uniformly distributed across all users. If we assume a coherence time of 1ms (roughly corresponding to vehicular speed), the total feedback overhead of 8000 bits is below 5% of the downlink total rate (before accounting for channel magnitude/phase quantization). This particular ASB scheme is optimized with sub-band spacing of about 180KHz, and where each user feeds back its best $S = 21$ spatial-frequency sub-channels. About 20% of the total feedback bits are devoted to the first stage in this case. (The value 20% is chosen with a numerical search.) Table II also lists results for other feedback settings. These results show that the proposed two-stage ASB scheme is an attractive solution for practical

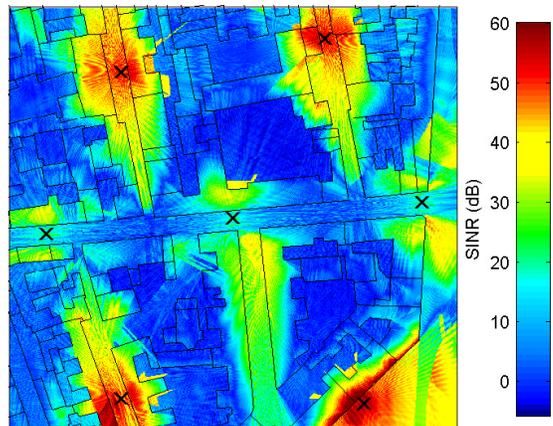


Fig. 11. SINR map of the small-cell deployment in Paris.

TABLE II
PERFORMANCE OF A URBAN SMALL-CELL DEPLOYMENT WITH LIMITED FEEDBACK

Feedback Bits/Cell	Best K	Best S	Feedback Overhead	Sum Rate (Mbps)	% of Perfect CSI Case
2000	3	7	1.5%	142	58%
4000	2	11	2.4%	169	69%
8000	1	21	3.7%	216	88%

system implementation.

As mentioned throughout this paper, the feedback bits here do not include channel magnitude/phase feedback. Here, we explicitly account for this overhead. Using the 8000-bit-per-cell feedback as an example, with 55 PRB's, 7 cells, and 2 antenna per cell, the indexing of each PRB in the first stage requires about 9-10 bits. The first stage of ASB is similar to OBF for which it has been shown in [11] that 3 additional bits of channel magnitude quantization is sufficient. Thus, accounting for channel magnitude quantization in the first stage results in at most 30% overhead. Likewise in the second stage, the optimal number of bits for channel direction quantization is about 8-9 bits per scheduled user. Assuming 3-4 bits for quantizing each of the magnitude and the phase, plus 2-3 bits for selecting the appropriate channel direction quantizer for each scheduled user (based on the channel magnitude), this results in an overall about doubling of the feedback overhead. Nevertheless, since 8000 feedback bits represent less than 4% of the total downlink data rate, the total amount of feedback is quite reasonable (and in any case less than 10%) for the overall system.

Finally, we note that the rate computation here is based on the achieved SINR. An adaptive modulation stage subsequent to scheduling and beamforming is assumed here for achieving these rates.

VII. CONCLUSIONS

This paper shows that it is possible to design efficient CSI feedback scheme for the network MIMO system to approach the perfect CSI performance with only modest amount of feedback. The key idea is to use a two-stage feedback mechanism

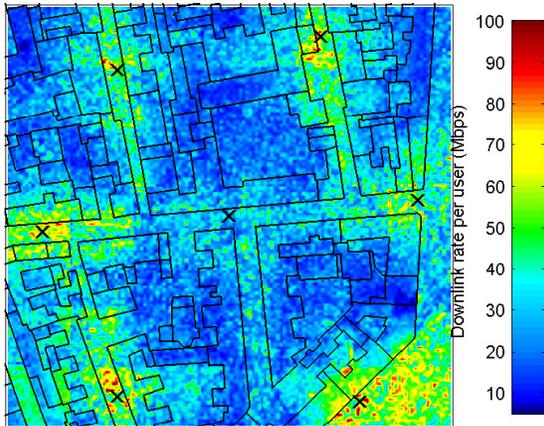


Fig. 12. Network MIMO rate map with 8000-bit-per-cell limited feedback for the small-cell deployment.

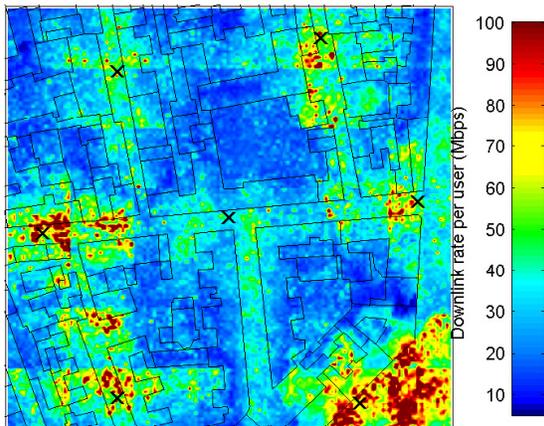


Fig. 13. Network MIMO rate map with perfect CSI for the small-cell deployment.

together with an agile scheduling and beamforming scheme to capture the benefit of multiuser diversity in the first stage and that of spatial multiplexing gains in the second stage.

On a theoretical front, this paper derives the performance scaling law for the proposed two-stage feedback mechanism. It is shown that the number of bits reserved for the second beamforming feedback stage should scale logarithmically with the total number of feedback bits. Further, in quantizing CSI from different base-stations in the second stage, each user should allocate its feedback bit budget in proportion to the corresponding channel magnitudes in dB scale. The bit allocation results are then used to show that the overall system performance scales double-logarithmically with the total feedback rate, linearly with the total number of antennas, and logarithmically with the transmit SNR, thus achieving both spatial and multiuser diversity gains.

On a practical front, this paper presents numerical results to show that the proposed scheme, in comparison to conventional

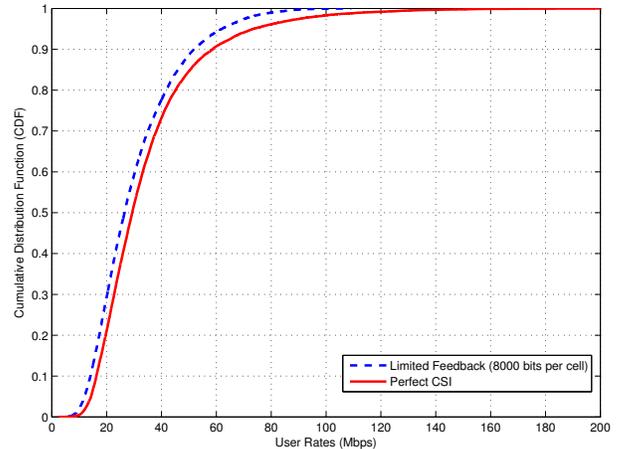


Fig. 14. Cumulative density function of user rates in an urban small-cell deployment: perfect CSI vs. limited feedback with 8000 feedback bits.

feedback mechanisms, improves both the downlink sum rate and the network utility. A feature of this paper is in its use of realistic deployment channel model to generate the rate maps for network MIMO systems under limited feedback. Finally, the advantages of the proposed ASB scheme are achieved in spite of a significant reduction in the scheduling computational complexity, which makes the proposed scheme an attractive solution for practical system implementation.

APPENDIX

This appendix provides a justification for the approximation process in defining the virtual SINR (18) in Section IV-A, where we replace the numerator and the denominator in the SINR expression with their expected values. To this end, define

$$X_{\pi(m),j} = \left| \sum_{k \in \mathcal{I}} \mathbf{g}_{k\pi(m)}^\dagger \mathbf{v}_{kj} \right|^2. \quad (53)$$

The sum rate upper bound (17) can then written as:

$$R = MI \log \left(\mathbb{E} \left[1 + \frac{\rho X_{\pi(m),\pi(m)}}{1 + I_{\pi(m)}} \right] \right), \quad (54)$$

where $I_{\pi(m)}$ is the interference power:

$$I_{\pi(m)} = \rho \sum_{j \neq \pi(m)} X_{\pi(m),j}. \quad (55)$$

In the asymptotic regime, where $B \rightarrow \infty$, any efficient feedback mechanism should force the interference power to zero; otherwise, the rate expression would saturate in high SNR regime. For example, for the feedback scheme proposed in this paper, if we consider the interference term in the denominator of (41) and apply the bit allocation rules in (45) and (46), we see that the interference power diminishes as $\ln B/B$ as $B \rightarrow \infty$.

We therefore expect $X_{\pi(m),j} \rightarrow 0$, for $j \neq \pi(m)$, as $B \rightarrow \infty$. Furthermore, our numerical results suggest that, for the scheduling-beamforming scheme in this paper, the ratio

$$\frac{\sigma_{X_{\pi(m),j}}^2}{\mathbb{E}[X_{\pi(m),j}^2]} = \frac{\mathbb{E} \left[|X_{\pi(m),j} - \mathbb{E}[X_{\pi(m),j}]|^2 \right]}{\mathbb{E}[X_{\pi(m),j}^2]} \quad (56)$$

also diminishes as $B \rightarrow \infty$. This result, although difficult to prove due to the complexity of $X_{\pi(m_i),j}$'s probability distribution function, suggests that one can safely ignore the difference term ($X_{\pi(m_i),j} - \mathbb{E}[X_{\pi(m_i),j}]$) in comparison with $\mathbb{E}[X_{\pi(m_i),j}]$ and therefore safely use the approximation $X_{\pi(m_i),j} \approx \mathbb{E}[X_{\pi(m_i),j}]$.

Using this justification, we can approximate the SINR terms in (54) as follows:

$$\begin{aligned} & \mathbb{E} \left[1 + \frac{\rho X_{\pi(m_i),\pi(m_i)}}{1 + \rho \sum_{j \neq \pi(m_i)} X_{\pi(m_i),j}} \right] \\ & \approx \mathbb{E} \left[1 + \frac{\rho X_{\pi(m_i),\pi(m_i)}}{1 + \rho \sum_{j \neq \pi(m_i)} \mathbb{E}[X_{\pi(m_i),j}]} \right] \\ & = 1 + \frac{\rho \mathbb{E}[X_{\pi(m_i),\pi(m_i)}]}{1 + \rho \sum_{j \neq \pi(m_i)} \mathbb{E}[X_{\pi(m_i),j}]} \end{aligned}$$

which justifies replacing each random variable in the SINR expression with its expected value as in (18).

REFERENCES

- [1] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, "Multicell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.
- [2] G. Caire, A. Ramprashad, and C. Papadopoulos, "Rethinking network MIMO: Cost of CSIT, performance analysis, and architecture comparisons," in *Inf. Theory Appl. Workshop (ITA)*, vol. II, San Diego, CA, Jan. 2010, pp. 803–806.
- [3] J. Zhang, M. Kountouris, J. G. Andrews, and R. W. Heath, Jr., "Multi-mode transmission for the MIMO broadcast channel with imperfect channel state information," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 803–814, Mar. 2011.
- [4] R. Zakhour and D. Gesbert, "A two-stage approach to feedback design in multi-user MIMO channels with limited channel state information," in *Proc. IEEE Int. Symp. Personal, Indoor, Mobile Radio Commun. (PIMRC)*, Athens, Greece, Sep. 2007.
- [5] N. Ravindran and N. Jindal, "Multi-user diversity vs. accurate channel feedback for MIMO broadcast channels," in *IEEE Int. Conf. Commun. (ICC)*, May 2008, pp. 3684–3688.
- [6] Y. Corre and Y. Lohan, "Three-dimensional urban EM wave propagation model for radio network planning and optimization over large areas," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3112–3123, Sep. 2009.
- [7] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.
- [8] K. Huang, R. W. Heath, Jr., and J. G. Andrews, "Space division multiple access with a sum feedback rate constraint," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3879–3891, Jul. 2007.
- [9] J. S. Kim, H. Kim, and K. B. Lee, "Limited feedback signaling for MIMO broadcast channels," in *Proc. IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, New York, NY, Jun. 2005, pp. 855–859.
- [10] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, Sep. 2007.
- [11] K. Huang, J. G. Andrews, and R. W. Heath, Jr., "Performance of orthogonal beamforming for SDMA with limited feedback," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 152–164, Jan. 2009.
- [12] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.
- [13] R. Agarwal, C. Hwang, and J. M. Cioffi, "Opportunistic feedback protocol for achieving sum-capacity of the MIMO broadcast channel," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Sep. 2007, pp. 606–610.
- [14] R. Agarwal, V. Majjigi, Z. Han, R. Vannithamby, and J. Cioffi, "Low complexity resource allocation with opportunistic feedback over downlink OFDMA networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1462–1472, Oct. 2008.
- [15] T. Tang, R. W. Heath, Jr., S. Cho, and S. Yun, "Opportunistic feedback for multiuser MIMO systems with linear receivers," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 1020–1032, May 2007.
- [16] B. Khoshnevis and W. Yu, "Bit allocation laws for multi-antenna channel feedback: single-user case," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2270–2283, May 2011.
- [17] —, "Bit allocation laws for multi-antenna channel feedback: multiuser case," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 367–382, Jan. 2012.
- [18] R. Bhagavatula and R. W. Heath, Jr., "Adaptive limited feedback for sum-rate maximizing beamforming in cooperative multicell systems," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 808–811, Feb. 2011.
- [19] J. H. Lee and W. Choi, "Optimal feedback rate sharing strategy in zero-forcing mimo broadcast channels," *IEEE Trans. Wireless Commun.*, to appear, 2013.
- [20] J. So and J. M. Cioffi, "Multiuser diversity in a MIMO system with opportunistic feedback," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 4909–4918, Nov. 2009.
- [21] J. Leinonen, J. Hamalainen, and M. Juntti, "Performance analysis of downlink OFDMA resource allocation with limited feedback," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2927–2937, Jun. 2009.
- [22] J. Choi and R. W. Heath, Jr., "Interpolation based transmit beamforming for MIMO-OFDM with limited feedback," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 4125–4135, Nov. 2005.
- [23] T. Pande, D. J. Love, and J. V. Krogmeier, "Reduced feedback MIMO-OFDM precoding and antenna selection," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2284–2293, May 2007.
- [24] J. Chen, R. Berry, and M. Honig, "Limited feedback schemes for downlink OFDMA based on sub-channel groups," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1451–1461, Oct. 2008.
- [25] R. Bhagavatula and R. W. Heath, Jr., "Adaptive bit partitioning for multicell intercell interference nulling with delayed limited feedback," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3824–3836, Aug. 2011.
- [26] A. P. T. Lau and F. R. Kschischang, "Feedback quantization strategies for multiuser diversity," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2072–2087, Jun. 2007.



Behrouz Khoshnevis received B.Sc. and M.Sc. degrees in Electrical Engineering from Sharif University of Technology, Tehran, Iran, in 2003 and 2005 respectively, and Ph.D. degree in Electrical and Computer Engineering from University of Toronto, Toronto, Ontario, Canada, in 2010. He is the recipient of Connaught Scholarship, Edward Rogers Sr. Graduate Scholarship, and Shahid U.H. Qureshi Memorial Prize in recognition of academic excellence from University of Toronto.



Wei Yu (S'97-M'02-SM'08) received the B.A.Sc. degree in Computer Engineering and Mathematics from the University of Waterloo, Waterloo, Ontario, Canada in 1997 and M.S. and Ph.D. degrees in Electrical Engineering from Stanford University, Stanford, CA, in 1998 and 2002, respectively. Since 2002, he has been with the Electrical and Computer Engineering Department at the University of Toronto, Toronto, Ontario, Canada, where he is now Professor and holds a Canada Research Chair in Information Theory and Wireless Communications.

His main research interests include information theory, optimization, wireless communications and broadband access networks.

Prof. Wei Yu currently serves as an Associate Editor for IEEE TRANSACTIONS ON INFORMATION THEORY. He was an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS (2009-2011), an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2004-2007), and a Guest Editor for a number of special issues for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and the EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING. He is member of the Signal Processing for Communications and Networking Technical Committee of the IEEE Signal Processing Society. He received the IEEE Signal Processing Society Best Paper Award in 2008, the McCharles Prize for Early Career Research Distinction in 2008, the Early Career Teaching Award from the Faculty of Applied Science and Engineering, University of Toronto in 2007, and the Early Researcher Award from Ontario in 2006.



Yves Lostanlen (SM98-M01-SM09) received the Dipl.-Ing (MSc EE) magna cum laude in 1996 from National Institute for Applied Sciences (INSA) in France. After three years of research at University College London and INSA he accomplished a European PhD summa cum laude in 2000. In 2009, he obtained a Sc.D. (HDR) from University of Rennes, France and in 2013 he graduated from MIT Sloan School of Management, USA (Executive MBA.)

Yves Lostanlen is an experienced Technology Expert (CTO) and International Business Executive

in the ICT Industry with a focus on Wireless Telecommunications. His current interests are innovative technologies and services enabling energy-efficient infrastructure networks (telecom, energy, water) in under-developed regions in order to catalyze competitive advantage, productivity and growth.

His entrepreneurial experience and expertise for 15+ years lies in delivering highly praised sophisticated software and expert consultancies on wireless system and mobile technologies to many top-tier companies in the ICT, Energy, Healthcare, Broadcast & Media Industries: Government, policy makers, regulators, infrastructure operators and equipment manufacturers.

Yves Lostanlen is also an Adjunct Professor of Wireless Communication Systems at the University of Toronto, Canada.

Yves Lostanlen has written more than 100 papers for international conferences, periodicals, book chapters and has been technical committee chairman, invited keynote speaker and session chairman at several international conferences. He received a "Young Scientist" Award" for two papers at the EuroEM 2000 conference.