Massive MIMO and Inter-Tier Interference Coordination

Ansuman Adhikary
University of Southern California
Los Angeles, CA 90089

Ebrahim Al Safadi
University of Southern California
Los Angeles, CA 90089

Giuseppe Caire
University of Southern California
Los Angeles, CA 90089

***Abstract*—We study the performance of a two-tier system where a large number of small cells is deployed under a macro-cellular "umbrella". The macro-cellular tier provides coverage and handles mobile users, while the small cell tier provides high rate locally to nomadic users. While the standard approach consists of operating the two tiers in different frequency bands, for various reasons (e.g., lack of licensed spectrum), it may be useful to operate both tiers in the whole available spectrum. Hence, we consider schemes for inter-tier interference coordination that do not assume any explicit data or channel state information sharing between tiers. In particular, we consider co-channel TDD and reverse TDD schemes, when the macro (tier-1) base station has a very large number of antennas and the tier-2 base stations have a moderately large number of antennas. We show that by exploiting the spatial directionality of the channel vectors, very efficient inter-tier interference management can be obtained with relatively low complexity. Our approach consists of a sort of "spatial blanking" of certain angle-of-departure of the tier-1 base station at given scheduled time-frequency slots, in order to create transmission opportunities for the corresponding tier-2 small cells. In particular, such "spatial blanking" is significantly more efficient than isotropic slot blanking (enhanced Inter-Cell Interference Coordination, eICIC) currently proposed in LTE standardization.**

I. Introduction

Dense spatial spectrum reuse has been widely recognized as the single most valuable resource for overcoming the wireless "spectrum crunch" [1]. A possible solution consists of deploying a large number of small cells that operate under a common macrocell umbrella. Tier-2 cells (small cells) have attracted a lot of attention both in cellular standards bodies such as 3GPP, as well as in academic research. A large body of literature has focused on mitigating the inter-tier interference, using schemes such as eICIC [2], [3] (and references therein), which involve orthogonalizing the time frequency resources allocated to both tiers.

A "cognitive" small cell approach was proposed in [4], where tier-2 base stations (BSs) have the ability to decode the tier-1 (macrocell) BS control channel and schedule their transmissions to deal with the inter-tier interference. It was demonstrated that with a simple power control approach and a moderate number of antennas at the tier-2 BSs, very high area spectral efficiency (bit/s/Hz/km^2) could be achieved in both tiers.

In this work, we focus on a scenario where the tier-1 BS is equipped with a very large number of antennas (massive MIMO). Since the tier-1 BS is typically located on an elevated position (e.g., tower-mounted, or deployed on a building roof), it "sees" both its own users and the tier-2 cells under a relatively narrow angular spread. This gives rise to highly directional channel vectors, which can be modeled as Gaussian random vectors with a small number of dominant eigenmodes (eigenvectors of their covariance matrix). As a result, the tier-1 BS can make use of directional beamforming, similar to the JSDM approach proposed in [5], in order to simultaneously achieve spatial multiplexing to its own tier-1 users as well as mitigate the inter-tier interference to the tier-2 cells. Inter-tier interference can be mitigated by nulling certain spatial directions, i.e., by transmitting in the orthogonal complement of the dominant eigenmodes of the channel vectors from the tier-1 BS to a subset of selected tier-2 cells. The selection of such directions can be allocated in the time-frequency resource such that each tier-2 cell has a fair share of transmission opportunities free from inter-tier interference. In analogy with the frame blanking approach of eICIC said above, we refer to this approach as "spatial blanking". As a result, the tier-2 throughput can be increased without significantly sacrificing the tier-1 throughput, as opposed to eICIC, which can only operate on the convex combination region of the individual throughput capacities of the tier-1 and tier-2 systems. A similar scenario has been considered in [6], but under different beamforming and power control strategies. The work reflects the effectiveness of "spatial blanking", but with higher complexity algorithms since the channel directionality properties are not taken into account, which can lead to efficient JSDM decomposition and dimensionality reduction. In addition, it was shown that reverse-TDD is not competitive, however, we find that reverse-TDD can yield significant advantages in some cases with respect to co-channel TDD.

This paper is organized as follows. We present the basic system model in Section II. We analyze the system performance in Sections III and IV, followed by some numerical results in Section V. Conclusions are pointed out in Section VI.

II. System model

We consider a macrocell (tier-1 system) comprising of a single BS having M antennas and containing F tier-2 small cells, each equipped with L antennas.

The system operates in TDD (Time Division Duplexing) mode, where both the downlink and uplink bands are accessed using OFDM/TDMA. For simplicity, we consider here a

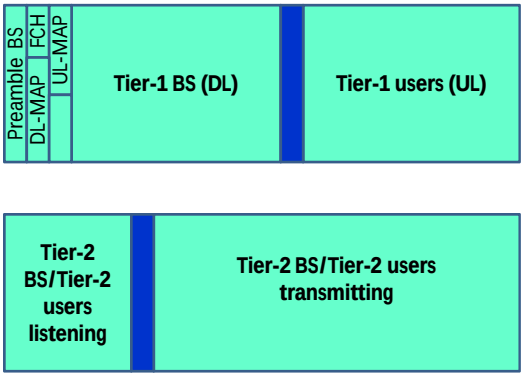


Fig. 1: Frame structure for the two-tier network

frequency flat channel corresponding to a set of adjacent subcarriers in the same channel coherence bandwidth. The TDD macrocell/femtocell frame structure is borrowed from [4], and is shown in Fig. 1. The tier-1 BS frame includes the control channel and the tier-1 uplink (UL) and downlink (DL) subframes, with a small guard interval in between. The tier-2 cells have two subframes. During the first one, overlapped with the tier-1 BS control channel, all the tier-2 nodes (both BS and users) listen to the tier-1 BS control channel. After a guard interval, during which the tier-2 nodes decode the control information in the tier-1 BS control channel and acquire the allocation of the tier-1 users on the tier-1 DL/UL subframes, all the tier-2 cells are active and transmit using TDD, both in the UL and in the DL (depending on the scenario). We specifically investigate two schemes: *reverse-TDD* (R-TDD), as proposed originally in [4], where the tier-1 DL is aligned with the tier-2 UL (and vice-versa), and *cochannel-TDD* (co-TDD), as examined in [6], where tier-1 UL is aligned with the tier-2 UL (and vice-versa). It is worthwhile to point out that all the tier-2 nodes can decode the tier-1 BS control signal and use it as common information for coordination. In particular, we assume that the power, rate and positioning (location) information of the active tier-1 users scheduled on the current frame are known to all tier-2 nodes in each frame. We examine the two tier system performance in terms of the achievable throughput tradeoff region between the tier-1 throughput (sum rate) and the tier-2 throughput (sum-rate).

A. Macrocell subsystem

In each slot, the tier-1 BS serves tier-1 users. These are divided into G groups. Tier-1 users are grouped according to their position in the cell. Since co-located users have the same scattering environment (see the discussion on the validity of this statement in [7], as well as efficient user grouping algorithms presented in [8]) but are separated by at least a few tens of wavelengths, their channel vectors are mutually independent but have the *same* second-order statistics, implying that the channel covariance matrix between tier-1 users and the tier-1 BS are identical within the same group. In contrast, provided that the different groups are widely separated in their scattering components, the dominant eigenspaces of the corresponding channel covariance matrix are linearly independent.

The instantaneous channel between a user k in group g (indicated as g_k, with some abuse of notation) and the tier-1 BS, at each time-frequency scheduling slot, is an $M \times 1$ Gaussian vector $\boldsymbol{h}_{g_k}$. Using the Karhunen-Loeve representation, we can write

$$\boldsymbol{h}_{g_k} = \boldsymbol{U}_g \boldsymbol{\Lambda}_g^{1/2} \boldsymbol{w}_{g_k}, \tag{1}$$

where $\boldsymbol{R}_g = \boldsymbol{U}_g \boldsymbol{\Lambda}_g \boldsymbol{U}_g^{\mathsf{H}}$ is the channel covariance matrix, common to all users in group g, $\boldsymbol{U}_g$ is the tap unitary matrix of eigenmodes, of dimension $M \times r_g$, and $\boldsymbol{\Lambda}_g$ is the $r_g \times r_g$ diagonal positive definite matrix of covariance eigenvalues (Karhunen-Loeve coefficients). Notice that these quantities are common to all users in group g. In contrast, the $r_g \times 1$ random vector $\boldsymbol{w}_{g_k} \sim \mathcal{CN}(0, \boldsymbol{I}_{r_g})$ is independent for different users and corresponds to the randomness due to the small-scale multipath fading components. The typical duration over which the channel covariances change is several orders of magnitude larger than the dynamics of small-scale fading. Therefore, for mathematical convenience, we assume $\boldsymbol{R}_g$ to be fixed in time and consider average rates (i.e., ergodic rates) with respect to the small-scale fading components. Notice that under the classical Wide-Sense Stationary Uncorrelated Scattering channel model [9], the channel process is wide-sense stationary and therefore its second-order statistics are constant in time, as we assume here. This assumption is valid "locally" when observing the system on the time-scale of a few tens of seconds. In practice, the channel covariance matrices must be adaptively learned and tracked in order to follow the non-stationary time-varying effects in the network (e.g., due to user mobility).

Following [8], we consider the one-ring scattering model in order to determine $\boldsymbol{R}_g$. Namely, for a user group located at an angle of arrival θ_g and having angular spread Δ_g, we have $\boldsymbol{R}_g = \boldsymbol{R}(\theta_g, \Delta_g)$ where, assuming a uniform linear array at the tier-1 BS, the element (m, n) of $\boldsymbol{R}(\theta_g, \Delta_g)$ is given by

$$[\boldsymbol{R}(\theta_g, \Delta_g)]_{m,n} = \frac{1}{2\Delta_g} \int_{\theta_g - \Delta_g}^{\theta_g + \Delta_g} e^{-j\pi(m-n)\sin(\alpha)} d\alpha. \tag{2}$$

and $a_{g,0}$ represents the The total tier-1 BS transmit power is denoted by P_0. For analytical simplicity, we assume that the aggregate sum power of all tier-1 users transmitting on the UL is also equal to P_0, such that the total radiated power of the tier-1 cell is the same in both UL and DL slots. Also for simplicity we consider equal power allocation, such that tier-1 BS DL data stream is allocated power $\frac{P_0}{S}$, where S is the total number of DL streams.

B. Small cell subsystem

Tier-2 cells operate in TDD with no constraint of aligning their DL and UL with the tier-1 DL/UL slots. We assume that tier-2 cells use intra-cell orthogonal access, such that only one tier-2 user per small cell is active over any time-frequency slot. Hence, as far as the tier-2 throughput is concerned, it is sufficient to consider a single user per tier-2 cell.

Using the information obtained from the control channel, the femtocells implement a simple power control strategy in order to mitigate the cross tier interference. In our proposed power control strategy, the tier-2 cells (for both UL and DL) adjust their transmission power levels such that the average interference power at the macrocell receiver is not larger than some threshold κ, which we call the "interference temperature" of tier-2 on tier-1 [4]. Furthermore, a peak transmit power P_1 is imposed to all tier-2 cells. In the tier-1 DL slot, only the tier-2 cells close to the active tier-1 users need to lower their transmission power below P_1, in order to satisfy the interference temperature constraint. Since the set of active tier-1 users changes randomly from slot to slot, depending on the DL scheduling of the tier-1 BS, the set of tier-2 cells that have to transmit at very low power also changes with time, thus obtaining a sort of statistical multiplexing in the spatial cell area. However, in the tier-1 UL, the tier-2 cells close to the tier-1 BS are required to significantly lower their transmission power on all slots, since the tier-1 BS does not change in time. Hence, the tier-2 cells close to the BS are permanently at a disadvantage. It is worthwhile to note that during tier-1 DL slots, when $G = 1$, only a small number of tier-2 cells around the (single) active tier-1 user group transmit at low power, but as G increases, more and more tier-2 cells need to lower their transmit power, resulting in a degradation of the tier-2 throughput. The same problem occurs in tier-1 UL slots, due to the presence of more user groups, which cause a lot of interference to their nearby tier-2 cells. Hence, we expect to see a tradeoff between the tier-1 and the tier-2 throughputs. In the following, we are interested in investigating this achievable throughput tradeoff region under a certain class of efficient MIMO precoding schemes that exploit the channel directionality properties said before.

III. System performance : Reverse TDD

We denote the set of user groups by $\mathcal{G}$ and the set of tier-2 cells by $\mathcal{C}$, of cardinality $|\mathcal{G}| = G$ and $|\mathcal{C}| = F$, respectively. For R-TDD we focus on the case of tier-1 DL and tier-2 UL only. In fact, for the other direction (tier-1 UL and tier-2 DL), we exploit uplink-downlink duality and achieve exactly the same rates with the same sum power (see [4] and references therein for details).

A. Tier-1 DL

Since the rank of the channel covariance matrix $\boldsymbol{R}_g$ is r_g, the number of simultaneously active users that can be served using multiuser MIMO (MU-MIMO) (i.e., the number of spatial multiplexing data streams in the tier-1 DL to users in group g) is $S_g \leq \min\{K_g, r_g\}$, where K_g is the number of users in group g. The covariance rank r_g is related to the angle of arrival θ_g and angular spread Δ_g for that particular group, and can be sharply characterized in closed form in the limit of large M (see [8]). The received signal at user k of group g (indicated as g_k, with some abuse of notation), is given by

$$y_{g_k} = \sqrt{a_{g,0}}\boldsymbol{h}_{g_k}^{\mathsf{H}}\boldsymbol{V}\boldsymbol{d} + \sum_{f\in\mathcal{C}}\sqrt{a_{g,f}}h_{g_k,f}x_f + z_{g_k}, \qquad (3)$$

where $\boldsymbol{V}$ is the $M \times S$ beamforming matrix of the tier-1 BS, where we let $S = \sum_{g=1}^{G} S_g$, $\boldsymbol{d}$ is the $S\times 1$ vector of data symbols, $h_{g_k,f} \sim \mathcal{CN}(0,1)$ is a complex coefficient representing the scalar channel between the tier-1 user g_k and the tier-2 user located in cell $f \in \mathcal{C}$, and $z_{g_k} \sim \mathcal{CN}(0,1)$ is the AWGN at the tier-1 user receiver. The path gain coefficients $a_{g,f}$ include log-normal shadowing, distance-dependent pathloss, and other possible geometric effects such as wall absorption (e.g., in the case where the tier-2 is an indoor femtocell and the tier-1 user is outdoor), between users in groups g and BS f. We use index 0 to indicate the tier-1 BS. Depending on $\boldsymbol{R}_g$, we consider two different beamforming schemes as outlined below:

Isotropic scattering (i.e., $\boldsymbol{R}_g = \boldsymbol{I}_M$, $r_g = M$): In this case, the tier-1 BS uses *zero-forcing beamforming*, such that $\boldsymbol{V}$ is the column-normalized Moore-Penrose pseudo inverse of the channel matrix $\boldsymbol{H} = [\boldsymbol{H}_1\boldsymbol{H}_2\ldots\boldsymbol{H}_G]$, where $\boldsymbol{H}_g$ denotes the $M \times S_g$ channel matrix formed by the channel vectors of the active users in group g. As a result, (3) becomes:

$$y_{g_k} = \sqrt{a_{g,0}}\left(\boldsymbol{h}_{g_k}^{\mathsf{H}}\boldsymbol{v}_{g_k}\right)d_{g_k} + \sum_{f\in\mathcal{C}}\sqrt{a_{g,f}}h_{g_k,f}x_f + z_{g_k}, \qquad (4)$$

where the intra-cell tier-1 multiuser interference disappears as a result of zero-forcing beamforming.

Directional scattering (i.e., $\boldsymbol{R}_g = \boldsymbol{U}_g\boldsymbol{\Lambda}_g\boldsymbol{U}_g^{\mathsf{H}}$, with $\boldsymbol{\Lambda}_g$ strongly skewed such that only $r_g < M$ eigenmodes collect significant energy): In this case, the tier-1 BS employs *JSDM with per group processing* [5] with $\boldsymbol{V} = [\boldsymbol{B}_1\ldots\boldsymbol{B}_G]\mathrm{diag}(\boldsymbol{P}_1,\ldots,\boldsymbol{P}_G)$, where $\boldsymbol{B}_g$ is the $M \times b_g$ pre-beamforming matrix associated with group g and $\boldsymbol{P}_g$ is the $b_g \times S_g$ precoding matrix obtaining by zero forcing on the effective channel $\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{H}_g$ of group g after pre-beamforming. Note that JSDM is a two stage processing technique, and has significant advantages in terms of pilot overhead and computational complexity. In addition, the pre-beamforming matrix $\boldsymbol{B}_g$ depends only on the channel second-order statistics $\boldsymbol{R}_g$, which can be learned accurately over a relatively long time interval, and the pre-beamforming transformation can be implemented in the RF analog domain (hybrid analog-digital beamforming), thus allowing the tier-1 BS to have a very large number of antennas M, while having a moderate number $b = \sum b_g$ of RF chains. The advantages of the JSDM architecture are discussed in [5], [8], [7]. The received signal (3) thus takes on the form:

$$\begin{aligned} y_{g_k} &= \sqrt{a_{g,0}}\left(\boldsymbol{h}_{g_k}^{\mathsf{H}}\boldsymbol{B}_g\mathbf{p}_{g_k}\right)d_{g_k} \\ &\quad + \sum_{g'\in\mathcal{G}:g'\neq g}\sum_{m=1}^{S'_g}\sqrt{a_{g,0}}\left(\boldsymbol{h}_{g_k}^{\mathsf{H}}\boldsymbol{B}_{g'}\mathbf{p}_{g'_m}\right)d_{g'_m} \qquad (5) \\ &\quad + \sum_{f\in\mathcal{C}}\sqrt{a_{g,f}}h_{g_k,f}x_f + z_{g_k}, \qquad (6) \end{aligned}$$

where the intra-group interference disappears due to zero-forcing beamforming, but we have an additional inter-group interference term (5) in addition to the inter-tier interference (6) due to the tier-2 cells.

a) Achievable rate with isotropic scattering: From (4), the received SINR at a user g_k is given as

$$\mathrm{SINR}_{g_k} = \frac{a_{g,0}|\boldsymbol{h}_{g_k}^{\mathsf{H}}\boldsymbol{v}_{g_k}|^2 P_0/S}{1+\sum_{f\in\mathcal{C}} a_{g,f}|h_{g_k,f}|^2 P_f}, \tag{7}$$

where P_f is the transmit power of the tier-2 user in cell f. Using Jensen's inequality, the achievable ergodic rate for user g_k can be lower-bounded by

$$\begin{aligned}&\mathbb{E}\left[\log_2\left(1+\mathrm{SINR}_{g_k}\right)\right]\\ &\geq \mathbb{E}\left[\log_2\left(1+\frac{a_{g,0}|\boldsymbol{h}_{g_k}^{\mathsf{H}}\boldsymbol{v}_{g_k}|^2 P_0/S}{1+\sum_{f\in\mathcal{C}} a_{g,f}P_f}\right)\right]\\ &\triangleq R_{g_k,\mathrm{IS}}^{\mathrm{DL}-1}\end{aligned} \tag{8}$$

Using the method of deterministic equivalents (see [10]), a simple approximation to (8) for $M\to\infty$ is given by

$$R_{g_k,\mathrm{IS-DA}}^{\mathrm{DL}-1} = \log_2\left(1+\frac{(M-S+1)a_{g,0}\frac{P_0}{S}}{1+\sum_{f\in\mathcal{C}} a_{g,f}P_f}\right), \tag{9}$$

where it can be shown that $R_{g_k,\mathrm{IS}}^{\mathrm{DL}-1} - R_{g_k,\mathrm{IS-DA}}^{\mathrm{DL}-1} \overset{M\to\infty}{\longrightarrow} 0$.

b) Achievable rate with directional scattering: Proceeding in a similar fashion, we have that the achievable ergodic rate for user g_k is lower bounded by

$$R_{g_k,\mathrm{DS}}^{\mathrm{DL}-1} = \mathbb{E}\left[\log_2\left(1+\frac{a_{g,0}|\boldsymbol{h}_{g_k}^{\mathsf{H}}\boldsymbol{B}_g\mathbf{p}_{g_k}|^2 P_0/S}{\left[1+\sum_{g'\neq g} a_{g,0}|\boldsymbol{h}_{g_k}^{\mathsf{H}}\boldsymbol{B}_{g'}\mathbf{p}_{g'_k}|^2 P_0/S + \sum_{f\in\mathcal{C}} a_{g,f}P_f\right]}\right)\right] \tag{10}$$

Following again the approach in [10], an easy to compute approximation of $R_{g_k,\mathrm{DS}}^{\mathrm{DL}-1}$ for $M\to\infty$ is given by (see [5] for details)

$$\begin{aligned}&R_{g_k,\mathrm{DS-DA}}^{\mathrm{DL}-1}\\ &= \log_2\left(1+\frac{b_g m_g a_{g,0}\frac{P_0}{S}}{1+\sum_{g'\neq g} I_{g,g'}a_{g,0}\frac{P_0}{S}+\sum_{f\in\mathcal{C}} a_{g,f}P_f}\right),\end{aligned} \tag{11}$$

where m_g is obtained as the solution of the fixed-point equation

$$\begin{aligned} m_g &= \frac{1}{b_g}\mathrm{tr}\left(\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_g\boldsymbol{T}_g^{-1}\right)\\ \boldsymbol{T}_g &= \boldsymbol{I}_{b_g}+\frac{S_g}{b_g}\frac{\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_g}{m_g}\end{aligned} \tag{12}$$

and $\{I_{g,g'} : \forall\ \ g,g'\in\mathcal{G}\}$ is the solution of the system of coupled fixed-point equations

$$\begin{aligned} I_{g,g'} &= S_{g'}\frac{n_{g,g'}}{m_{g'}}\\ n_{g,g'} &= \frac{\frac{1}{b_{g'}}\mathrm{tr}\left(\boldsymbol{B}_{g'}^{\mathsf{H}}\boldsymbol{R}_{g'}\boldsymbol{B}_{g'}\boldsymbol{T}_g^{-1}\boldsymbol{B}_{g'}^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_{g'}\boldsymbol{T}_g^{-1}\right)}{1-J_{g'}}\\ J_{g'} &= \frac{\frac{1}{b_{g'}}\mathrm{tr}\left(\boldsymbol{B}_{g'}^{\mathsf{H}}\boldsymbol{R}_{g'}\boldsymbol{B}_{g'}\boldsymbol{T}_g^{-1}\boldsymbol{B}_{g'}^{\mathsf{H}}\boldsymbol{R}_{g'}\boldsymbol{B}_{g'}\boldsymbol{T}_g^{-1}\right)}{b_{g'}m_{g'}^2}.\end{aligned} \tag{13}$$

B. Tier-2 UL

Tier-2 cells schedule their respective UL slots when the tier-1 cell is in the DL slot. As a result, a tier-2 BS receiving its desired user signal suffers from both intra-tier interference from the tier-2 users in neighboring cells and inter-tier interference from the tier-1 BS. The received signal at the receiver of tier-2 BS f is given by

$$\boldsymbol{y}_f = \boldsymbol{h}_{f,f}x_f + \sum_{f'\neq f}\sqrt{a_{f,f'}}\boldsymbol{h}_{f,f'}x_{f'} + \sqrt{a_{f,0}}\boldsymbol{H}_{f,0}^{\mathsf{H}}\boldsymbol{V}\boldsymbol{d} + \boldsymbol{z}_f, \tag{14}$$

where $\boldsymbol{h}_{f,f'}$ denotes the $L\times 1$ channel between the tier-2 BS f receiving antenna array and a tier-2 user in cell f', x_f denotes the scalar symbol transmitted by a tier-2 user in cell f, $\boldsymbol{H}_{f,0}$ is the $M\times L$ channel matrix between the tier-1 BS and the tier-2 BS f, and $\boldsymbol{z}_f$ is an AWGN vector. The channel matrix $\boldsymbol{H}_{f,0}$ depends on the type of scattering in the propagation from the tier-1 BS to the tier-2 BS. Instead, the channel between the tier-2 users and the tier-2 BS is assumed to have isotropic scattering, since the tier-2 BS is surrounded by scattering elements with uniform angle of arrivals.[1] Hence, we let $\boldsymbol{h}_{f,f'}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{I}_L)$. Finally, without loss of generality, we normalize the channel gain coefficient inside each tier-2 cell by $a_{f,f}=1$, since because of the small size of tier-2 cells, all users inside such cells have roughly the same path loss.

The tier-2 BSs use linear MMSE in order to detect their intended user. The resulting SINR is given by

$$\begin{aligned}\mathrm{SINR}_f &= \boldsymbol{h}_{f,f}^{\mathsf{H}}\left(\boldsymbol{I}_L+\sum_{f'\neq f}a_{f,f'}\boldsymbol{h}_{f,f'}\boldsymbol{h}_{f,f'}^{\mathsf{H}}P_{f'}\right.\\ &\qquad\left.+a_{f,0}\boldsymbol{H}_{f,0}^{\mathsf{H}}\boldsymbol{V}\boldsymbol{V}^{\mathsf{H}}\boldsymbol{H}_{f,0}\frac{P_0}{S}\right)^{-1}\boldsymbol{h}_{f,f}P_f.\end{aligned} \tag{15}$$

Recall that the transmit power P_f is controlled as a function of the interference temperature κ. In particular, we let:

$$\begin{aligned}P_f &= \min\left\{P_1,\frac{\kappa}{a_f^{\max}}\right\}\\ a_f^{\max} &= \max\left\{a_{g,f} : g\in\mathcal{G}\right\}\end{aligned} \tag{16}$$

Obviously, $a_{g,f}P_f \leq \frac{a_{g,f}}{a_f^{\max}}\kappa \leq \kappa$, such that the average inter-tier interference caused by a tier-2 user f to an active tier-1

[1] Think of an indoor femtocell, or a small BS at low elevation in an urban square.

user in group g is at most κ. The achievable ergodic rate for tier-2 cell f is given by $R_f^{\mathrm{UL}-2} = \mathbb{E}\left[\log_2(1+\mathrm{SINR}_f)\right]$ where the SINR is given by (15).

The statistics of $\boldsymbol{H}_{f,0}$ (i.e., the channel matrix between the tier-1 BS and a tier-2 BS) depends on the type of scattering landscape. As for the case of tier-1 users, we consider both isotropic and directional scattering.

Isotropic Scattering: In this case, $\boldsymbol{H}_{f,0}$ has i.i.d. elements $\sim \mathcal{CN}(0,1)$. As said before, the tier-1 BS beamforming matrix is given by $\boldsymbol{V} = \boldsymbol{H}\left(\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\right)^{-1}\boldsymbol{\Xi}$ where $\boldsymbol{\Xi}$ is a diagonal matrix that makes the columns of $\boldsymbol{H}\left(\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}\right)^{-1}$ to have unit norm. Letting $\boldsymbol{H} = \boldsymbol{U}_0\boldsymbol{\Sigma}_0^{1/2}\boldsymbol{V}_0^{\mathsf{H}}$ denote the SVD of $\boldsymbol{H}$, we have

$$\begin{aligned}\boldsymbol{H}_{f,0}^{\mathsf{H}}\boldsymbol{V}\boldsymbol{V}^{\mathsf{H}}\boldsymbol{H}_{f,0} &= \boldsymbol{H}_{f,0}^{\mathsf{H}}\boldsymbol{H}(\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H})^{-1}\boldsymbol{\Xi}^2(\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H}_{f,0}\\ &= \boldsymbol{H}_{f,0}^{\mathsf{H}}\boldsymbol{U}_0\boldsymbol{\Sigma}_0^{-1/2}\boldsymbol{V}_0^{\mathsf{H}}\boldsymbol{\Xi}^2\boldsymbol{V}_0\boldsymbol{\Sigma}_0^{-1/2}\boldsymbol{U}_0^{\mathsf{H}}\boldsymbol{H}_{f,0}\\ &= \tilde{\boldsymbol{H}}_{f,0}\boldsymbol{\Sigma}_0^{-1/2}\boldsymbol{V}_0^{\mathsf{H}}\boldsymbol{\Xi}^2\boldsymbol{V}_0\boldsymbol{\Sigma}_0^{-1/2}\tilde{\boldsymbol{H}}_{f,0}^{\mathsf{H}} \quad (17)\end{aligned}$$

where $\tilde{\boldsymbol{H}}_{f,0} = \boldsymbol{H}_{f,0}^{\mathsf{H}}\boldsymbol{U}_0$ is also Gaussian i.i.d., by the well-known unitarily invariant property of Gaussian i.i.d. matrices. Hence, we can write the tier-2 UL achievable rate in the form

$$\begin{aligned}&R_f^{\mathrm{UL}-2}\\ &= \mathbb{E}\left[\log_2\left(1+\boldsymbol{h}_{f,f}^{\mathsf{H}}\left(\boldsymbol{I}_L + \sum_{f'\neq f} a_{f,f'}\boldsymbol{h}_{f,f'}\boldsymbol{h}_{f,f'}^{\mathsf{H}}P_{f'}\right.\right.\right.\\ &\left.\left.\left. + a_{f,0}\tilde{\boldsymbol{H}}_{f,0}\boldsymbol{\Sigma}_0^{-1/2}\boldsymbol{V}_0^{\mathsf{H}}\boldsymbol{\Xi}^2\boldsymbol{V}_0\boldsymbol{\Sigma}_0^{-1/2}\tilde{\boldsymbol{H}}_{f,0}^{\mathsf{H}}\frac{P_0}{S}\right)^{-1}\boldsymbol{h}_{f,f}P_f\right)\right] \quad (18)\end{aligned}$$

Using the general results of [11], we can obtain a convergent approximation $R_{f,\mathrm{IS,DA}}^{\mathrm{UL}-2}$ such that $R_f^{\mathrm{UL}-2} - R_{f,\mathrm{IS,DA}}^{\mathrm{UL}-2} \overset{M\to\infty}{\longrightarrow} 0$, with

$$R_{f,\mathrm{IS,DA}}^{\mathrm{UL}-2} = \log_2\left(1+\zeta_f P_f\right), \quad (19)$$

where ζ_f is the solution of the following fixed-point equation

$$\begin{aligned}&\zeta_f\\ &= \left[1 + \sum_{f'\neq f}\frac{a_{f,f'}P_{f'}}{1+La_{f,f'}P_{f'}\zeta_f}\right.\\ &\left. \quad + \frac{S}{L\zeta_f}\left[1-\eta_{(\bar{\boldsymbol{H}}^{\mathsf{H}}\bar{\boldsymbol{H}})^{-1}}\left(\frac{L(M-S+1)}{S^2}a_{f,0}P_0\zeta_f\right)\right]\right]^{-1} \quad (20)\end{aligned}$$

where $\eta_{(\bar{\boldsymbol{H}}^{\mathsf{H}}\bar{\boldsymbol{H}})^{-1}}(\gamma)$ denotes the η-transform (see [12]) of the asymptotic eigenvalue distribution of the random matrix $\bar{\boldsymbol{H}} = \frac{1}{\sqrt{S}}\boldsymbol{H}$.

Directional Scattering: In this case, we resort to an accurate and easy to compute lower bound on the achievable rate, which significantly simplifies calculations. First, we need the following auxiliary lemma:

Lemma 1. *Let $\boldsymbol{A}$ take on values in the cone of positive definite Hermitian symmetric matrices, and $\boldsymbol{x}$ be a complex vector. Then, the function $f(\boldsymbol{A}) = \log(1+\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}^{-1}\boldsymbol{x})$ is convex in $\boldsymbol{A}$.*

Proof: The proof follows by showing that, for any positive definite matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, vector $\boldsymbol{x}$ and $0\leq\lambda\leq 1$, we have

$$\begin{aligned}&\log\left(1+\boldsymbol{x}^{\mathsf{H}}\left(\lambda\boldsymbol{A}_1+(1-\lambda)\boldsymbol{A}_2\right)^{-1}\boldsymbol{x}\right)\\ \leq\ &\lambda\log\left(1+\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}_1^{-1}\boldsymbol{x}\right)+(1-\lambda)\log\left(1+\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}_2^{-1}\boldsymbol{x}\right)\end{aligned} \quad (21)$$

Since $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are positive definite, we have from [13]

$$\boldsymbol{x}^{\mathsf{H}}\left(\lambda\boldsymbol{A}_1+(1-\lambda)\boldsymbol{A}_2\right)^{-1}\boldsymbol{x} \leq \left(\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}_1^{-1}\boldsymbol{x}\right)^{\lambda}\left(\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}_2^{-1}\boldsymbol{x}\right)^{1-\lambda} \quad (22)$$

Denoting $a = \left(\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}_1^{-1}\boldsymbol{x}\right)$ and $b = \left(\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}_2^{-1}\boldsymbol{x}\right)$, we have

$$\begin{aligned}&\log\left(1+\boldsymbol{x}^{\mathsf{H}}\left(\lambda\boldsymbol{A}_1+(1-\lambda)\boldsymbol{A}_2\right)^{-1}\boldsymbol{x}\right)\\ \leq\ &\log\left(1+a^{\lambda}b^{1-\lambda}\right)\\ \overset{(a)}{\leq}\ &\log\left((1+a)^{\lambda}(1+b)^{1-\lambda}\right)\\ =\ &\lambda\log(1+a)+(1-\lambda)\log(1+b)\\ =\ &\lambda\log\left(1+\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}_1^{-1}\boldsymbol{x}\right)+(1-\lambda)\log\left(1+\boldsymbol{x}^{\mathsf{H}}\boldsymbol{A}_2^{-1}\boldsymbol{x}\right),\end{aligned}$$

where (a) is due to the generalized Holder's inequality. ■

Applying Jensen's inequality, we can write

$$\begin{aligned}&R_f^{\mathrm{UL}-2}\\ &\geq \mathbb{E}\left[\log_2\left(1+\boldsymbol{h}_{f,f}^{\mathsf{H}}\left(\boldsymbol{I}_L + \sum_{f'\neq f} a_{f,f'}\boldsymbol{h}_{f,f'}\boldsymbol{h}_{f,f'}^{\mathsf{H}}P_{f'}\right.\right.\right.\\ &\left.\left.\left. + a_{f,0}\boldsymbol{H}_{f,0}^{\mathsf{H}}\mathbb{E}[\boldsymbol{V}\boldsymbol{V}^{\mathsf{H}}]\boldsymbol{H}_{f,0}\frac{P_0}{S}\right)^{-1}\boldsymbol{h}_{f,f}P_f\right)\right]\\ &\triangleq R_{f,\mathrm{LB}}^{\mathrm{UL}-2}\end{aligned} \quad (23)$$

Recall that in this case the tier-1 BS beamforming matrix $\boldsymbol{V}$ takes on the form

$$\boldsymbol{V} = [\boldsymbol{B}_1\ldots\boldsymbol{B}_G]\mathrm{diag}(\boldsymbol{P}_1,\ldots,\boldsymbol{P}_G),$$

with $\boldsymbol{P}_g$ given by the normalized Moore-Penrose pseudo inverse of $\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{H}_g$. Also, the channel matrix $\boldsymbol{H}_{f,0}$ is given by

$$\boldsymbol{H}_{f,0} = \boldsymbol{R}_{f,0}^{1/2}\boldsymbol{W}_{f,0} \quad (24)$$

where $\boldsymbol{R}_{f,0} = \boldsymbol{U}_{f,0}\boldsymbol{\Lambda}_{f,0}\boldsymbol{U}_{f,0}^{\mathsf{H}}$ is the $M\times M$ covariance matrix of the channels between the tier-1 BS array and each antenna of the tier-2 BS array, and $\boldsymbol{W}_{f,0}$ is an $M\times L$ matrix i.i.d. matrix with $\mathcal{CN}(0,1)$ elements. From a geometric viewpoint, the tier-2 BS antenna are analogous to L co-located users, such that the columns of $\boldsymbol{H}_{f,0}$ are mutually independent, assuming that the antenna elements are separated by more than one wavelength.

In order to evaluate (23), we first need to compute $\mathbb{E}[\boldsymbol{V}\boldsymbol{V}^{\mathsf{H}}]$.

Note that

$$\begin{aligned}
&\mathbb{E}[\boldsymbol{V}\boldsymbol{V}^{\mathsf{H}}] \\
&= \sum_{g=1}^{G} \mathbb{E}[\boldsymbol{B}_g \boldsymbol{P}_g \boldsymbol{P}_g^{\mathsf{H}} \boldsymbol{B}_g^{\mathsf{H}}] \\
&= \sum_{g=1}^{G} \mathbb{E}[\boldsymbol{B}_g \boldsymbol{B}_g^{\mathsf{H}} \boldsymbol{H}_g \left(\boldsymbol{H}_g^{\mathsf{H}} \boldsymbol{B}_g \boldsymbol{B}_g^{\mathsf{H}} \boldsymbol{H}_g\right)^{-2} \boldsymbol{H}_g^{\mathsf{H}} \boldsymbol{B}_g \boldsymbol{B}_g^{\mathsf{H}}] \\
&= \sum_{g=1}^{G} \mathbb{E}[\boldsymbol{B}_g \tilde{\boldsymbol{U}}_g^{\mathsf{H}} \tilde{\boldsymbol{\Lambda}}_g^{1/2} \tilde{\boldsymbol{W}}_g^{\mathsf{H}} \left(\tilde{\boldsymbol{W}}_g \tilde{\boldsymbol{\Lambda}}_g \tilde{\boldsymbol{W}}_g^{\mathsf{H}}\right)^{-2} \tilde{\boldsymbol{W}}_g \tilde{\boldsymbol{\Lambda}}_g \tilde{\boldsymbol{U}}_g \boldsymbol{B}_g^{\mathsf{H}}] \\
&= \sum_{g=1}^{G} \boldsymbol{B}_g \tilde{\boldsymbol{U}}_g^{\mathsf{H}} \mathbb{E}\left[\tilde{\boldsymbol{\Lambda}}_g^{1/2} \tilde{\boldsymbol{W}}_g^{\mathsf{H}} \left(\tilde{\boldsymbol{W}}_g \tilde{\boldsymbol{\Lambda}}_g \tilde{\boldsymbol{W}}_g^{\mathsf{H}}\right)^{-2} \tilde{\boldsymbol{W}}_g \tilde{\boldsymbol{\Lambda}}_g\right] \tilde{\boldsymbol{U}}_g \boldsymbol{B}_g^{\mathsf{H}}
\end{aligned} \tag{25}$$

where $\boldsymbol{B}_g^{\mathsf{H}} \boldsymbol{H}_g = \tilde{\boldsymbol{U}}_g^{\mathsf{H}} \tilde{\boldsymbol{\Lambda}}_g^{1/2} \tilde{\boldsymbol{W}}_g^{\mathsf{H}}$, with $\tilde{\boldsymbol{\Lambda}}_g$ diagonal $b_g \times b_g$ and $\tilde{\boldsymbol{W}}_g$ i.i.d. Gaussian $b_g \times K_g$. Letting $\boldsymbol{K}_g = \mathbb{E}\left[\tilde{\boldsymbol{\Lambda}}_g^{1/2} \tilde{\boldsymbol{W}}_g^{\mathsf{H}} \left(\tilde{\boldsymbol{W}}_g \tilde{\boldsymbol{\Lambda}}_g \tilde{\boldsymbol{W}}_g^{\mathsf{H}}\right)^{-2} \tilde{\boldsymbol{W}}_g \tilde{\boldsymbol{\Lambda}}_g\right]$, we can write the (i,j)-th element as

$$\begin{aligned}
&[\boldsymbol{K}_g]_{i,j} \\
&= \mathbb{E}\left[\sqrt{\tilde{\lambda}_{g,i}\tilde{\lambda}_{g,j}} \tilde{\boldsymbol{w}}_{g,i}^{\mathsf{H}} \left(\tilde{\boldsymbol{W}}_g \tilde{\boldsymbol{\Lambda}}_g \tilde{\boldsymbol{W}}_g^{\mathsf{H}}\right)^{-2} \tilde{\boldsymbol{w}}_{g,j}\right] \\
&= \mathbb{E}\left[\sqrt{\tilde{\lambda}_{g,i}\tilde{\lambda}_{g,j}} \tilde{\boldsymbol{w}}_{g,i}^{\mathsf{H}} \left(\sum_{k=1}^{b_g} \tilde{\lambda}_{g,k} \tilde{\boldsymbol{w}}_{g,k} \tilde{\boldsymbol{w}}_{g,k}^{\mathsf{H}}\right)^{-2} \tilde{\boldsymbol{w}}_{g,j}\right] \\
&\overset{(a)}{=} \mathbb{E}\left[\frac{\sqrt{\tilde{\lambda}_{g,i}} \tilde{\boldsymbol{w}}_{g,i}^{\mathsf{H}} \left(\sum_{k=1,k\neq i}^{b_g} \tilde{\lambda}_{g,k} \tilde{\boldsymbol{w}}_{g,k} \tilde{\boldsymbol{w}}_{g,k}^{\mathsf{H}}\right)^{-1}}{\left(1 + \tilde{\lambda}_{g,i} \tilde{\boldsymbol{w}}_{g,i}^{\mathsf{H}} \left(\sum_{k=1,k\neq i}^{b_g} \tilde{\lambda}_{g,k} \tilde{\boldsymbol{w}}_{g,k} \tilde{\boldsymbol{w}}_{g,k}^{\mathsf{H}}\right)^{-1} \tilde{\boldsymbol{w}}_{g,i}\right)}\right. \\
&\qquad \left. \times \frac{\left(\sum_{k=1,k\neq j}^{b_g} \tilde{\lambda}_{g,k} \tilde{\boldsymbol{w}}_{g,k} \tilde{\boldsymbol{w}}_{g,k}^{\mathsf{H}}\right)^{-1} \tilde{\boldsymbol{w}}_{g,j} \sqrt{\tilde{\lambda}_{g,j}}}{\left(1 + \tilde{\lambda}_{g,j} \tilde{\boldsymbol{w}}_{g,j}^{\mathsf{H}} \left(\sum_{k=1,k\neq j}^{b_g} \tilde{\lambda}_{g,k} \tilde{\boldsymbol{w}}_{g,k} \tilde{\boldsymbol{w}}_{g,k}^{\mathsf{H}}\right)^{-1} \tilde{\boldsymbol{w}}_{g,j}\right)}\right]
\end{aligned}$$

where $\tilde{\boldsymbol{w}}_{g,k}$ denotes the k-th column of $\tilde{\boldsymbol{W}}_g$, $\tilde{\lambda}_{g,k}$ the (k,k)-th entry of $\tilde{\boldsymbol{\Lambda}}_g$ and (a) follows by the matrix inversion lemma. We will evaluate $[\boldsymbol{K}_g]_{i,j}$ in the large-system limit when $S_g, b_g \longrightarrow \infty$ with a fixed ratio. After some rather heavy algebra, omitted here due to space limitations, it can be shown that the off-diagonal entries of $\boldsymbol{K}_g$ go to zero, and that the diagonal entries are given by

$$\begin{aligned}
&[\boldsymbol{K}_g]_{i,i} \\
&= \mathbb{E}\left[\frac{\tilde{\lambda}_{g,i} \tilde{\boldsymbol{w}}_{g,i}^{\mathsf{H}} \left(\sum_{k=1,k\neq i}^{b_g} \tilde{\lambda}_{g,k} \tilde{\boldsymbol{w}}_{g,k} \tilde{\boldsymbol{w}}_{g,k}^{\mathsf{H}}\right)^{-2} \tilde{\boldsymbol{w}}_{g,i}}{\left(1 + \tilde{\lambda}_{g,i} \tilde{\boldsymbol{w}}_{g,i}^{\mathsf{H}} \left(\sum_{k=1,k\neq i}^{b_g} \tilde{\lambda}_{g,k} \tilde{\boldsymbol{w}}_{g,k} \tilde{\boldsymbol{w}}_{g,k}^{\mathsf{H}}\right)^{-1} \tilde{\boldsymbol{w}}_{g,i}\right)^2}\right]
\end{aligned} \tag{26}$$

Using the method of [10], a convergent approximation $[\boldsymbol{K}_g^o]_{i,i}$ such that $[\boldsymbol{K}_g]_{i,i} - [\boldsymbol{K}_g^o]_{i,i} \overset{b_g, S_g \to \infty}{\longrightarrow} 0$ takes on the form

$$\begin{aligned}
[\boldsymbol{K}_g^o]_{i,i} &= \frac{\tilde{\lambda}_{g,i} S}{1 + \tilde{\lambda}_{g,i} f_g} \frac{1}{\sum_{k=1}^{b_g} \frac{\tilde{\lambda}_{g,k}}{1 + \tilde{\lambda}_{g,k} f_g}} \\
f_g &= \frac{S}{\sum_{k=1}^{b_g} \frac{\tilde{\lambda}_{g,k}}{1 + \tilde{\lambda}_{g,k} f_g}}
\end{aligned} \tag{27}$$

Using (27) in (23), and using Karhunen-Loeve expansion for the columns of $\boldsymbol{H}_{f,0}$, we obtain

$$\begin{aligned}
&R_{f,\mathrm{LB}}^{\mathrm{UL}-2} \\
&= \mathbb{E}\left[\log_2\left(1 + \boldsymbol{h}_{f,f}^{\mathsf{H}} \left(\boldsymbol{I}_L + \sum_{f'\neq f} a_{f,f'} \boldsymbol{h}_{f,f'} \boldsymbol{h}_{f,f'}^{\mathsf{H}} P_{f'}\right.\right.\right. \\
&\qquad \left.\left.\left. + a_{f,0} \tilde{\boldsymbol{W}}_{f,0}^{\mathsf{H}} \tilde{\boldsymbol{\Lambda}}_{f,0} \tilde{\boldsymbol{W}}_{f,0} \frac{P_0}{S}\right)^{-1} \boldsymbol{h}_{f,f} P_f\right)\right]
\end{aligned} \tag{28}$$

where we used the eigenvalue decomposition

$$\boldsymbol{\Lambda}_{f,0}^{1/2} \boldsymbol{U}_{f,0}^{\mathsf{H}} \left[\sum_{g=1}^{G} \boldsymbol{B}_g \tilde{\boldsymbol{U}}_g^{\mathsf{H}} \boldsymbol{K}_g^o \tilde{\boldsymbol{U}}_g \boldsymbol{B}_g^{\mathsf{H}}\right] \boldsymbol{U}_{f,0} \boldsymbol{\Lambda}_{f,0}^{1/2} = \tilde{\boldsymbol{U}}_{f,0} \tilde{\boldsymbol{\Lambda}}_{f,0} \tilde{\boldsymbol{U}}_{f,0}^{\mathsf{H}}$$

and we let $\tilde{\boldsymbol{W}}_{f,0} = \tilde{\boldsymbol{U}}_{f,0}^{\mathsf{H}} \boldsymbol{W}_{f,0}$, where $\tilde{\boldsymbol{W}}_{f,0}$ are also i.i.d. $\sim \mathcal{CN}(0,1)$. As $M \longrightarrow \infty$, a convergent approximation to $R_{f,\mathrm{LB}}^{\mathrm{UL}-2}$ in (28) is given by

$$R_{f,\mathrm{DS,DA}}^{\mathrm{UL}-2} = \log_2(1 + \xi_f P_f), \tag{29}$$

where ξ_f is the solution of the fixed-point equation

$$\begin{aligned}
\xi_f = \left[1 + \sum_{f'\neq f} \frac{a_{f,f'} P_{f'}}{1 + L a_{f,f'} P_{f'} \xi_f}\right. \\
\left. + \sum_k \frac{\tilde{\lambda}_{f,0,k} a_{f,0} \frac{P_0}{S}}{1 + L \tilde{\lambda}_{f,0,k} a_{f,0} \frac{P_0}{S} \xi_f}\right]^{-1}
\end{aligned} \tag{30}$$

IV. System performance : Co-channel TDD

In co-TDD, as the name suggests, the tier-2 cells align their UL slot with the UL slot of the tier-1 cell (and vice-versa). Again, we evaluate the performance in one direction only, since the results hold for the other direction via uplink-downlink duality. The preferred direction for the analysis purpose is tier-1 UL/tier-2 UL. Since the tier-1 cell operates in UL mode, the tier-2 cells close to the tier-1 BS must lower their transmit power in order to meet the interference temperature requirement. Unlike R-TDD, these tier-2 cells are permanently at a disadvantage, since they have to transmit with small power on all slots. Nevertheless, thanks to the very large number of antennas, interference can be handled effectively in both tiers.

A. Tier-1 UL

In this case, the received signal vector at the tier-1 BS is given by the superposition of the signals transmitted by both the tier-1 and the tier-2 users, such that we have

$$\boldsymbol{y} = \sum_{g=1}^{G} \sum_{k=1}^{S_g} \sqrt{a_{g,0}} \boldsymbol{h}_{g_k} d_{g_k} + \sum_{f\in\mathcal{C}} \sqrt{a_{f,0}} \boldsymbol{h}_{f,0} x_f + \boldsymbol{z} \tag{31}$$

The tier-1 BS makes use of a linear MMSE receiver in order to detect the symbols of the tier-1 users and eventually decode their codewords. As a result, the ergodic achievable rate for tier-1 user g_k takes on the form

$$\begin{aligned}
&\bar{R}_{g_k}^{\mathrm{UL}-1} \\
&= \mathbb{E}\left[\log_2(1+\mathrm{SINR}_{g_k})\right] \\
&= \mathbb{E}\left[\log_2\left(1+a_{g,0}\boldsymbol{h}_{g_k}^{\mathsf{H}}\left(\boldsymbol{I}_M+\sum_{(g',m)\neq(g,k)}a_{g',0}\boldsymbol{h}_{g'_m}\boldsymbol{h}_{g'_m}^{\mathsf{H}}\frac{P_0}{S}\right.\right.\right. \\
&\left.\left.\left.+\sum_{f\in\mathcal{C}}a_{f,0}\boldsymbol{h}_{f,0}\boldsymbol{h}_{f,0}^{\mathsf{H}}P_f\right)^{-1}\boldsymbol{h}_{g_k}\frac{P_0}{S}\right)\right] \qquad (32)
\end{aligned}$$

In the case of directional scattering, as $M \longrightarrow \infty$ we obtain the convergent approximation

$$\bar{R}_{g_k,\mathrm{DA}}^{\mathrm{UL}-1} = \log_2\left(1+\eta_g\right), \qquad (33)$$

where η_g is given by the solution of the fixed-point equation

$$\begin{aligned}
\eta_g &= a_{g,0}\frac{P_0}{S}\mathrm{tr}\left(\boldsymbol{R}_g\boldsymbol{T}^{-1}\right) &(34)\\
\eta_f &= a_{f,0}P_f\mathrm{tr}\left(\boldsymbol{R}_{f,0}\boldsymbol{T}^{-1}\right) &(35)\\
\boldsymbol{T} &= \boldsymbol{I}_M+\sum_{g'=1}^{G}\frac{S_{g'}\boldsymbol{R}_{g'}}{1+\eta_{g'}}\frac{a_{g',0}P_0}{S}+\sum_{f\in\mathcal{C}}\frac{\boldsymbol{R}_{f,0}a_{f,0}P_f}{1+\eta_f} &(36)
\end{aligned}$$

For isotropic scattering, i.e., when $\boldsymbol{R}_g = \boldsymbol{R}_{f,0} = \boldsymbol{I}_M$, we can simplify further to get

$$\bar{R}_{g_k,\mathrm{IS,DA}}^{\mathrm{UL}-1} = \log_2\left(1+a_{g,0}\frac{P_0}{S}\eta\right) \qquad (37)$$

where η is given by the solution of the fixed-point equation

$$\eta = \left[1+\sum_{g'=1}^{G}\frac{S_{g'}\frac{a_{g',0}P_0}{S}}{1+Ma_{g',0}\frac{P_0}{S}\eta}+\sum_{f\in\mathcal{C}}\frac{a_{f,0}P_f}{1+Ma_{f,0}P_f\eta}\right]^{-1} \qquad (38)$$

B. Practical considerations for tier-1 UL

The calculation of the the linear MMSE receiver requires the computation of the inverse of the $M\times M$ covariance matrix of the received signal vector. In practice, the tier-1 BS computes the sample sample covariance matrix of the received signal and uses this estimate *in lieu* of the true covariance matrix, which is obviously not known a priori in general. For $M \gg 1$, an accurate estimation of the sample covariance would require a large number of data samples, which may not be possible to obtain when the data frame length is limited. In addition, calculating the inverse of an $M \times M$ matrix at each data frame may be computationally too hard. Finally, conventional "tracking" techniques such as RLS, which update directly the inverse covariance matrix, cannot be applied in this context since the received signal covariance matrix changes abruptly from frame to frame, depending on which tier-1 users are scheduled in the UL. Hence, in order to make a realistic comparison, we need to look at alternative receivers, which do not suffer from these drawbacks.

In this work, we propose to use a JSDM-type receiver for the tier-1 UL, which is outlined next. The linear receiver for decoding user g_k is formed by two stages. First, we multiply the received signal (31) by $\boldsymbol{B}_g$, where $\boldsymbol{B}_g$ depends only on the channel second-order statistics, in order to mitigate the inter-group interference from tier-1 users of other groups $g' \neq g$. Then, we use a second linear transformation (zero-forcing) $\boldsymbol{P}_g$ in order to remove the intra-group interference from the users belonging to group g. $\boldsymbol{P}_g$ is given by the pseudo-inverse of the resulting effective channel $\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{H}_g$. Note that this approach is exactly the "dual" of what we have seen before for the tier-1 DL. The major disadvantage of this approach is that we have no control over mitigating the cross-tier interference coming from the tier-2 users. However, if the tier-2 cells have non-overlapping angular spectra (direction of arrival) with the eigenmodes of user group in consideration, the first stage inherently performs some form of inter-tier interference suppression, especially when the number of antennas M is large [5]. Using this approach, the received signal for user g_k at the tier-1 BS is given by

$$\begin{aligned}
\hat{y}_{g_k} &= \mathbf{p}_{g_k}^{\mathsf{H}}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{y} \\
&= \sqrt{a_{g,0}}\mathbf{p}_{g_k}^{\mathsf{H}}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{h}_{g_k}d_{g_k} \\
&\quad+\sum_{g'=1,g'\neq g}^{G}\sum_{m=1}^{S_{g'}}\sqrt{a_{g',0}}\mathbf{p}_{g_k}^{\mathsf{H}}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{h}_{g'_m}d_{g'_m} \\
&\quad+\sum_{f\in\mathcal{C}}\sqrt{a_{f,0}}\mathbf{p}_{g_k}^{\mathsf{H}}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{h}_{f,0}x_f+\mathbf{p}_{g_k}^{\mathsf{H}}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{z}, \qquad (39)
\end{aligned}$$

where $\mathbf{p}_{g_k}$ denotes the columns of $\boldsymbol{P}_g$ corresponding to user g_k and where the intra-group interference disappears due to zero forcing. The resulting achievable ergodic rate for user g_k can be accurately calculated in the limit of large M by the convergent approximation

$$\begin{aligned}
&\bar{R}_{g_k,\mathrm{DS,JSDM,DA}}^{\mathrm{UL}-1} \\
&= \log_2\left(1+\frac{a_{g,0}\frac{P_0}{S}}{\frac{1}{b_g\bar{m}_g}+\sum_{g'\neq g}\bar{I}_{g,g'}^{(1)}a_{g',0}\frac{P_0}{S}+\sum_{f\in\mathcal{C}}\bar{I}_{g,f}^{(2)}a_{f,0}P_f}\right) \qquad (40)
\end{aligned}$$

where $\bar{m}_g$ is obtained as the solution of the fixed-point equation

$$\begin{aligned}
\bar{m}_g &= \frac{1}{b_g}\mathrm{tr}\left(\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_g\boldsymbol{T}_g^{-1}\right) \\
\boldsymbol{T}_g &= \boldsymbol{I}_{b_g}+\frac{S_g}{b_g}\frac{\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_g}{\bar{m}_g} \qquad (41)
\end{aligned}$$

and where $\{\bar{I}_{g,g'} : \forall\ g,g' \in \mathcal{G}\}$ and $\{\bar{I}_{g,f} : \forall\ g \in \mathcal{G}, f \in \mathcal{F}\}$

are the solution of the system of coupled fixed-point equations

$$\begin{aligned}
\bar{I}_{g,g'}^{(1)} &= \frac{S_{g'}}{b_g}\frac{\bar{n}_{g,g'}^{(1)}}{\bar{m}_g^2} \\
\bar{n}_{g,g'}^{(1)} &= \frac{\frac{1}{b_g}\mathrm{tr}\left(\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_g\boldsymbol{T}_g^{-1}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_{g'}\boldsymbol{B}_g\boldsymbol{T}_g^{-1}\right)}{1-J_g} \\
\bar{I}_{g,f}^{(2)} &= \frac{1}{b_g}\frac{\bar{n}_{g,f}^{(2)}}{\bar{m}_g^2} \\
\bar{n}_{g,f}^{(2)} &= \frac{\frac{1}{b_g}\mathrm{tr}\left(\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_g\boldsymbol{T}_g^{-1}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_{f,0}\boldsymbol{B}_g\boldsymbol{T}_g^{-1}\right)}{1-J_g} \\
J_g &= \frac{\frac{1}{b_g}\mathrm{tr}\left(\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_g\boldsymbol{T}_g^{-1}\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{R}_g\boldsymbol{B}_g\boldsymbol{T}_g^{-1}\right)}{b_g\bar{m}_g^2}
\end{aligned} \tag{42}$$

C. Tier-2 UL

With co-TDD, the $L \times 1$ received signal vector at a tier-2 BS f is given by

$$\boldsymbol{y}_f = \boldsymbol{h}_{f,f}x_f + \sum_{f'\neq f}\sqrt{a_{f,f'}}\boldsymbol{h}_{f,f'}x_{f'} + \sum_{g=1}^{G}\sum_{k=1}^{S_g}\sqrt{a_{f,g}}\boldsymbol{h}_{f,g_k}d_{g_k} + \boldsymbol{z}_f \tag{43}$$

Assuming a linear MMSE receiver, it is immediate to write the corresponding ergodic achievable rate as

$$\begin{aligned}
&\bar{R}_f^{\mathrm{UL-2}} \\
&= \mathbb{E}\left[\log_2\left(1+\boldsymbol{h}_{f,f}^{\mathsf{H}}\left(\boldsymbol{I}_L + \sum_{f'\neq f}a_{f,f'}\boldsymbol{h}_{f,f'}\boldsymbol{h}_{f,f'}^{\mathsf{H}}P_{f'}\right.\right.\right. \\
&\left.\left.\left. + \sum_{g=1}^{G}\sum_{k=1}^{S_g}a_{f,g}\boldsymbol{h}_{f,g_k}\boldsymbol{h}_{f,g_k}^{\mathsf{H}}\frac{P_0}{S}\right)^{-1}\boldsymbol{h}_{f,f}P_f\right)\right].
\end{aligned} \tag{44}$$

Because of the interference temperature constraint, and since with co-TDD the tier-2 UL interfere with the tier-1 UL, the transmit power in tier-2 cell f is given by

$$P_f = \min\left\{P_1, \frac{\kappa}{a_{f,0}}\right\}. \tag{45}$$

The channel vectors between tier-2 users and tier-2 BSs as well as between tier-1 users and tier-2 BSs are i.i.d. Gaussian (i.e., $\boldsymbol{h}_{f,f'} \sim \mathcal{CN}(0,\boldsymbol{I}_L)$, $\boldsymbol{h}_{f,g_k} \sim \mathcal{CN}(0,\boldsymbol{I}_L)$). As $M \longrightarrow \infty$, a convergent approximation to (44) is given by

$$\bar{R}_{f,\mathrm{DA}}^{\mathrm{UL-2}} = \log_2(1+\bar{\zeta}_f P_f), \tag{46}$$

where $\bar{\zeta}_f$ is the solution of the fixed-point equation

$$\bar{\zeta}_f = \left[1 + \sum_{f'\neq f}\frac{a_{f,f'}P_{f'}}{1+La_{f,f'}P_{f'}\bar{\zeta}_f} + \sum_{g=1}^{G}\frac{S_g a_{f,g}\frac{P_0}{S}}{1+La_{f,g}\frac{P_0}{S}\bar{\zeta}_f}\right]^{-1} \tag{47}$$

V. Results

We present some numerical results in order to compare the various schemes treated in Sections III and IV. We consider a squared area of side 1km, one tier-1 BS in the center, 80 indoor tier-2 cells with radius of 40 m, wall absorption of 5 dB, located on a regular squared grid[2], no log-normal shadowing and distance dependent pathloss coefficient with $a = 1/(1+(d/d_0)^{3.5})$, where the 3 dB loss distance d_0 is 50 m. The tier-1 total transmit power is $P_0 = 43$ dB, and the tier-2 peak transmit power is $P_1 = 20$ dB (relative to a noise floor of 0 dB). Tier-1 users are assumed to be outdoor.[3]

A. Design of the pre-beamforming matrices

The design of pre-beamformers is relevant both in tier-1 UL and DL, when directional scattering is involved. We adopt an approximate block diagonalization approach similar to [5]. In the following, we consider two sets of results depending on the ability of the tier-1 BS to mitigate the inter-tier interference to the tier-2 cells.

Tier-2 unaware pre-beamforming: Here, the tier-1 BS calculates the pre-beamformers in order to minimize the inter-group interference, both in the DL as well as in the UL. Denoting by $\boldsymbol{U}_g$ the eigenvectors of the channel covariance matrix of group g, we need $\boldsymbol{B}_g^{\mathsf{H}}\boldsymbol{H}_{g'} \approx 0 \quad \forall g' \neq g$. This can be achieved by restricting $\boldsymbol{B}_g$ to the orthogonal complement of the span of $\{\boldsymbol{U}_{g'}^* : g' \neq g\}$(see [5] for details), where $\boldsymbol{U}_g^* \subset \boldsymbol{U}_g$ comprises of the eigenvectors corresponding to the dominant eigenvalues of $\boldsymbol{R}_g$.

Tier-2 aware pre-beamforming (spatial blanking): Here, the tier-1 BS constructs the pre-beamformers in order to minimize also the inter-tier interference to the tier-2 cells. In the UL, the tier-1 BS does not cause any interference to the tier-2 cells and the tier-2 cells use power control to keep the interference caused to tier-1 users (in R-TDD) or to the tier-1 BS (in co-TDD) below the target interference temperature. In the tier-1 DL, the idea is to construct pre-beamformers such that they are *approximately orthogonal* to the tier-2 cells closest to the tier-1 BS, since these are the ones that suffer the most inter-tier interference. Hence, we require that

$$\boldsymbol{B}_g \subset \mathrm{Span}^{\perp}\left[\{\boldsymbol{U}_{g'}^* : g' \neq g\}\bigcup\{\boldsymbol{U}_{f,0}^* : f \in \mathcal{C}_{\mathrm{BS}-1}\}\right],$$

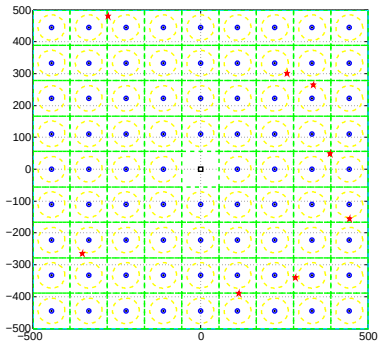
where $\mathcal{C}_{\mathrm{BS}-1}$ denotes the set of femtocells closest to the macro BS, and $\boldsymbol{U}_{f,0}^*$ contains the eigenvectors associated with the dominant eigenvalues of $\boldsymbol{R}_{f,0}$. Notice that this approach consists of creating transmission opportunities for the tier-2 cells by leaving "blank" slots in the spatial domain (in contrast to the eICIC strategy, based on blank time-frequency slots). For this reason, we refer to this approach as "spatial blanking".

B. Simulations

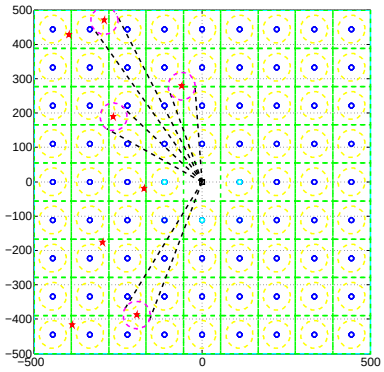
Figs. 2(a) and 2(b) show the two tier network consisting of a single tier-1 BS ("black" square) serving groups of users (a

[2]We consider an exclusion zone near the tier-1 BS, such that there are no tier-2 cells within a radius of 50m.

[3]Equivalently, we may think of an "open access" policy for the tier-2 cells, for which if a user enters a tier-2 cell, then it is offloaded automatically and is "swallowed" by the tier-2 cell.

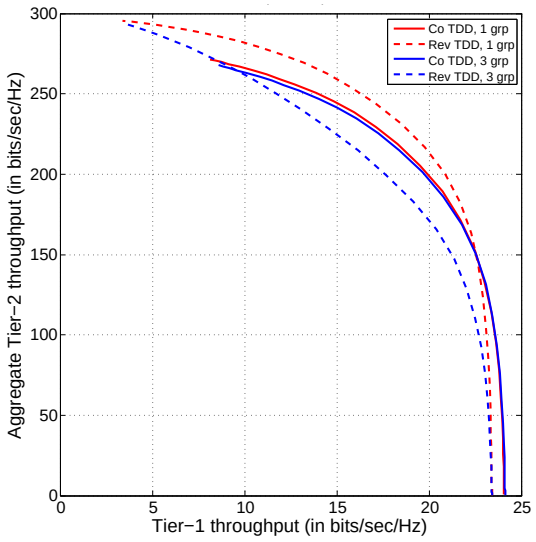


(a) Layout (Isotropic)

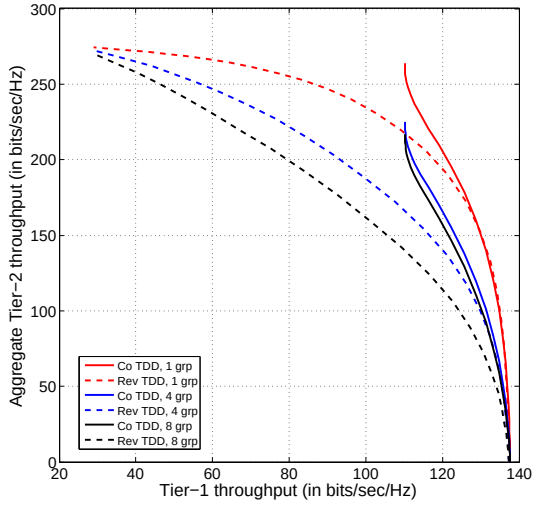


(b) Layout (Directional)

Fig. 2: Sample layout showing the two tier network.



(a) $M = 10, K = 6$



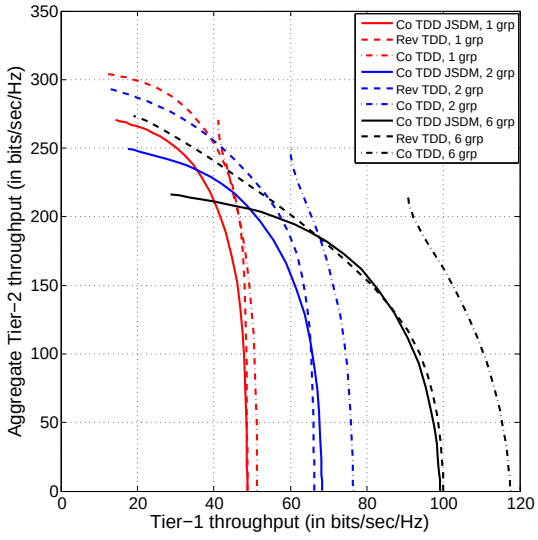
(b) $M = 100, K = 24$

Fig. 3: Reverse TDD vs Co TDD (Isotropic Scattering)

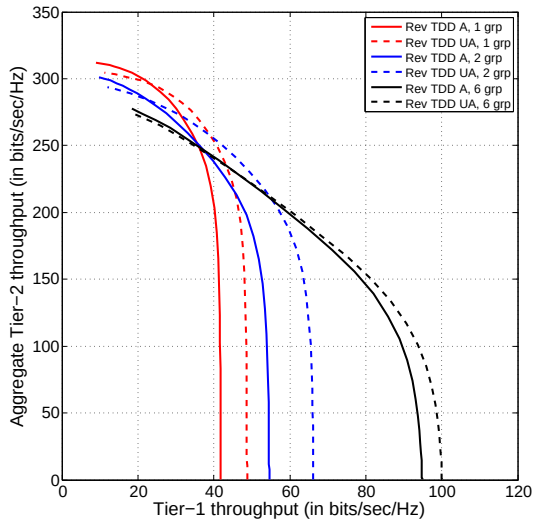
group of macro users is denoted by a "red" star) and several tier-2 cells ("blue" circles). With directional scattering, the tier-1 BS only serves groups of users with disjoint angular support, motivated by the result of [5] which states that the channel covariances of groups with disjoint angular support become orthogonal as the number of antennas M becomes very large. The tier-2 cells denoted as "cyan" circles are the ones that the tier-1 BS protects by spatial blanking. Note that these cells also have also a disjoint angular support with the groups of served tier-1 users, since trying to zero-force inter-tier interference with same or strongly overlapped angular support would yield a too large SNR penalty for the tier-1 user rates.

Fig. 3 shows a performance comparison between co-TDD and R-TDD for isotropic scattering. These schemes are compared in terms of the tradeoff curve between tier-1 and tier-2 aggregate throughput. This is obtained by varying the interference temperature level κ. In Fig. 3(a), we set the number of tier-1 BS antennas $M = 10$, and the number of tier-1 users $K = 6$. In Fig. 3(b), we set $M = 100$ and $K = 24$. In both these cases, the number of tier-2 BS antennas $L = 4$. We obtain different curves by varying the number of user groups G and let $K_g = \frac{K}{G}\ \forall\ g \in \mathcal{G}$. The "dashed" curves refer to R-TDD, and the "solid" curves to co-TDD. In the regime when all the tier-2 cells are transmitting with their peak power, we observe a higher tier-2 throughput for R-TDD because of the fact that only the tier-2 cells near the tier-1 BS suffer from inter-tier interference, and only these cells experience a rate degradation. However, in co-TDD, as the number of tier-1 user groups increases, more and more tier-2 cells are affected by the inter-tier interference, and this leads to lower tier-2 throughput. Co-TDD exhibits a higher tier-1 throughput in this regime because the tier-1 BS can effectively eliminate interference in the spatial domain. With R-TDD, the tier-1 users suffer from significant inter-tier interference since all the tier-2 cells are transmitting at their peak power, and this leads to performance degradation. In the other extreme, when the tier-2 cells are transmitting with low power, the tier-1 throughput in both the case remains almost the same while the tier-2 throughput is close to negligible.

Fig. 4(a) shows the same type of comparison for directional scattering. While the throughput tradeoff behavior may seem similar, there is one major difference with respect to the case of isotropic scattering. With directional scattering, the number of users that can be spatially multiplexed in a group



(a) Femtocell Unaware



(b) Femtocell Aware

Fig. 4: Reverse TDD vs Co TDD (Directional Scattering)

g is limited by the rank r_g of the group. Therefore, as the number of groups increases, we are able to serve more and more macro users, leading to increased macrocell throughput. Fig. 4(b) shows the performance comparison between tier-2 aware (spatial blanking) and tier-2 unaware pre-beamforming. The tier-2 unaware scheme is denoted by *Rev TDD UA* while the tier-2 aware scheme is denoted by *Rev TDD A*, where the tier-1 BS first chooses the user groups to serve and then decides to cancel inter-tier interference to (some) tier-2 cells with non overlapping directions.

VI. Conclusions

We have studied the performance of a two tier cellular network consisting of a single tier-1 BS and several tier-2 BSs sharing the same frequency channel. We considered the regime of large number of antennas at the tier-1 BS and moderately large number of antennas at the tier-2 BSs. We analyzed the system performance under two duplexing strategies: *co-TDD* and *R-TDD*. We derived closed form expressions and lower bounds to the achievable ergodic rates for different channel models using tools from random matrix theory and use these results to perform numerical comparisons in a scenario representative of a sub-urban area with indoor tier-2 cells. Our study shows that under isotropic scattering, R-TDD works well with a small number of antennas and few user groups, whereas co-TDD performs better for large user groups and large number of antennas. With directional channels, and insisting on using for both R-TDD and co-TDD a practical JSDM-like two-stage beamforming scheme, R-TDD yields better results when tier-2 cells make use of power control to minimize the inter-tier interference. Overall, under realistic directional scattering channel models, a combination of JSDM-like two-stage beamforming, interference temperature power control and R-TDD yields a tier-1/tier-2 throughput tradeoff far superior to eICIC, which can only operate on the "time-sharing line" (i.e., on the convex combination) of the individual tier-1 and tier-2 throughputs.

References

[1] V. Chandrasekhar, J. Andrews, and A. Gatherer, "Femtocell networks: a survey," *Communications Magazine, IEEE*, vol. 46, no. 9, pp. 59–67, 2008.

[2] D. Lopez-Perez, I. Guvenc, G. D. L. Roche, M. Kountouris, T. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogenous networks," *IEEE Wireless Commun. Mag.*, vol. 18, no. 3, pp. 22–30, June 2011.

[3] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. S. Jo *et al.*, "Heterogeneous cellular networks: From theory to practice," *Communications Magazine, IEEE*, vol. 50, no. 6, pp. 54–64, 2012.

[4] A. Adhikary, V. Ntranos, and G. Caire, "Cognitive femtocells: Breaking the spatial reuse barrier of cellular systems," in *Information Theory and Applications Workshop (ITA), 2011*. IEEE, 2011, pp. 1–10.

[5] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: The large-scale array regime," *Information Theory, IEEE Transactions on*, vol. 59, no. 10, pp. 6441–6463, 2013.

[6] J. Hoydis, K. Hosseini, S. t. Brink, and M. Debbah, "Making smart use of excess antennas: Massive mimo, small cells, and tdd," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 5–21, 2013.

[7] A. Adhikary, E. Al Safadi, M. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, "Joint spatial division and multiplexing for mm-wave channels," *arXiv preprint arXiv:1312.2045*, 2013.

[8] J. Nam, A. Adhikary, J. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *Submitted to IEEE Journal of Selected Topics in Signal Processing*, 2013.

[9] A. F. Molisch, *Wireless communications*. Wiley, 2010, vol. 15.

[10] S. Wagner, R. Couillet, M. Debbah, and D. T. Slock, "Large system analysis of linear precoding in correlated miso broadcast channels under limited feedback," *Information Theory, IEEE Transactions on*, vol. 58, no. 7, pp. 4509–4537, 2012.

[11] R. Couillet, M. Debbah, and J. W. Silverstein, "A deterministic equivalent for the analysis of correlated mimo multiple access channels," *Information Theory, IEEE Transactions on*, vol. 57, no. 6, pp. 3493–3514, 2011.

[12] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends® in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[13] W. Muir, "Inequalities concerning the inverses of positive definite matrices," *Proceedings of the Edinburgh Mathematical Society (Series 2)*, vol. 19, no. 02, pp. 109–113, 1974.