Uplink Power Control with MMSE Receiver in Multi-Cell MU-Massive-MIMO Systems

Kaifeng Guo*, Yan Guo*, Gábor Fodor†, and Gerd Ascheid*
*Institute for Communication Technologies and Embedded Systems, RWTH Aachen University, Germany.
Email: {kaifeng.guo, gerd.ascheid} @ice.rwth-aachen.de, yan.guo@rwth-aachen.de
†Ericsson Research, Sweden.
Email: gabor.fodor@ericsson.com

***Abstract*—In the current literature considering multi-cell multi-user massive multiple-input multiple-output (MU-Massive-MIMO) systems, equal uplink power allocation among users is typically assumed, which does not exploit the potential of per-user power control. By contrast, in this paper we apply multi-cell uplink power control, assuming the minimum mean-square-error receiver based on the pilot contaminated channel estimation and a very large but finite number of antennas at the base station. We derive the lower bound on the average post-processing uplink signal to interference-plus-noise ratio (SINR) with individual power assignment between pilot and data transmissions for each user, which facilitates a joint iterative uplink pilot and data power control strategy that minimizes the sum transmit power of all users subject to the per-user SINR and per-user power constraints. The convergence of the proposed algorithm to a unique fixed point optimal solution is discussed for both single- and multi-user scenarios. Numerical results indicate the significance of uplink power control which further improves the energy efficiency in MU-Massive-MIMO systems.**

I. Introduction

The breakthrough work in [1] has stimulated substantial research activities investigating various aspects of multi-user massive multiple-input multiple-output (MU-Massive-MIMO) systems, where hundreds of antennas are deployed at the base station (BS) serving a much smaller number of single-antenna users at the same time-frequency resource. Based on the favorable propagation [2], where any pair of distinct channel vectors to a given BS tends to become orthogonal as the number of BS antennas increases, the results in [1]-[3] have demonstrated that simple linear receivers with infinite number of BS antennas completely eliminate the intra-cell interference and noise. In particular, the linear minimum mean-square-error (MMSE) receiver is proved to be optimal among all linear detectors assuming perfect channel state information (CSI) and Gaussian data symbols, since it maximizes the achievable uplink rate in both single- and multi-cell scenarios [3] [4].

In practice, the CSI is obtained through uplink training under the assumption of time-division duplexing (TDD). Due to the inevitable reuse of pilot sequences among neighboring cells, the so-called pilot contamination [5] has been shown to limit the ultimate performance of the MMSE receiver (as well as other linear detectors) [6] [7]. Nevertheless, the uplink data transmit power of each user can be scaled down inversely proportional to the square-root of the number of BS antennas while maintaining the same performance [3], which leads to a significant energy efficiency.

On the other hand, multi-cell uplink power control has been illustrated to largely improve the energy efficiency in conventional MIMO systems [8] [9]. With perfect CSI and multi-cell coordination, the optimal solution of jointly assigning interference-plus-noise ratio (SINR) targets and transmit power can substantially reduce the uplink power consumption [8]; while with imperfect CSI but free of pilot contamination, the authors in [9] have exploited the inter-dependency between pilot and data transmissions to adjust pilot power in the context of existing data power control, such that total power saving is achieved subject to the per-user SINR constraint. However, these advantages of uplink power control have not been addressed in previous massive MIMO work [1]-[7], since equal uplink power allocation among users has always been assumed. Therefore, we are motivated to investigate the potential of uplink pilot and data power control in multi-cell MU-Massive-MIMO systems.

In this paper, the introduced MMSE receiver based on the practical channel estimates with pilot contamination is applied at the BS. In order to minimize the sum pilot and data transmit power of all users under the per-user SINR and per-user power constraints, we first derive the closed-form lower bound on the average uplink SINR (instead of directly computing the average SINR which in general has intractable form [6]) for a very large but finite number of BS antennas, which takes the individual power allocation among different users into account. Based on the lower bound, we then propose a joint iterative pilot and data power control algorithm which embeds the pilot power minimization into a *standard* data power control process [10]. The existence and uniqueness of the optimal solution are proved for single-user setup, while the convergence in multi-user case is guaranteed by an extra monotone decreasing constraint. Simulation results demonstrate the tightness of the lower bound as well as the large gain of applying multi-cell uplink power control together with massive MIMO technique to further improve the energy efficiency.

II. System Model

We consider a multi-cell MU-Massive-MIMO system with L cells. Each cell has K single-antenna users and a BS equipped with M ($M \gg K$) antennas. The system operates in TDD mode, and a frequency reuse of one is applied which

Part of this work has been performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in METIS, although the views expressed are those of the authors and do not necessarily represent the project.

allows all users in the system to share the same time-frequency resource while communicating with their own BSs. The k-th column vector of flat-fading channel matrix $\mathbf{G}_{li} = \mathbf{H}_{li}\mathbf{D}_{li}^{1/2}$ represents the propagation channel from the k-th user in cell i to the BS in cell l. In particular, the elements in $\mathbf{H}_{li} \in \mathbb{C}^{M\times K}$ are the fast fading coefficients and modelled (under favorable propagation [2]) as independent and identically distributed (i.i.d.) complex Gaussian random variables, i.e. $\mathcal{CN}(0,1)$, and the diagonal matrix $\mathbf{D}_{li} = \text{diag}\{[\beta_{li1},\ldots,\beta_{liK}]\} \in \mathbb{R}^{K\times K}$ describes the real-valued path-loss and large-scale fading which are assumed to be constant and known as a priori. Unlike the equal power allocation in [1]-[7], the diagonal matrix $\mathbf{P}_{\text{r},i} = \text{diag}\{[p_{\text{r},i1},\ldots,p_{\text{r},iK}]\} \in \mathbb{R}^{K\times K}$ consists of uplink data transmit power of each user in cell i. The transmitted vector $\mathbf{x}_i = [x_{i1},\ldots,x_{iK}]^T$ contains uplink data symbols from K users in cell i and $\mathbf{n}_l \sim \mathcal{CN}(\mathbf{0},\mathbf{I}_M)$ is the normalized additive noise vector. Therefore the received data vector at BS l is given as

$$\mathbf{y}_l = \sum_{i=1}^{L} \mathbf{G}_{li}\mathbf{P}_{\text{r},i}^{1/2}\mathbf{x}_i + \mathbf{n}_l. \tag{1}$$

A. Channel Estimation with Pilot Contamination

At the beginning of each channel coherence interval, the BSs have to estimate the propagation vectors of their serving users based on a standard uplink training process before starting the data transmission. The length of the pilot/training sequence τ should be much smaller than the coherence time, thus it is impossible in practice to provide orthogonal pilot sequences to all users in the system. As a compromise, only the orthogonality of pilot sequences assigned to the users within each cell is ensured with $\tau \geq K$, but the same set of pilots is reused among different cells, which leads to the pilot contamination problem [5]. Similar to the uplink data transmission in (1), the received signal of pilot transmission at BS l is given by

$$\mathbf{Y}_{\text{p},l} = \sum_{i=1}^{L} \mathbf{G}_{li}\mathbf{P}_{\text{p},i}^{1/2}\mathbf{\Phi}^T + \mathbf{N}_l \tag{2}$$

where $\mathbf{\Phi} \in \mathbb{C}^{\tau\times K}$ denotes the pilot matrix with orthonormal column vectors such that $\mathbf{\Phi}^H\mathbf{\Phi} = \mathbf{I}_K$, $\mathbf{P}_{\text{p},i} \in \mathbb{R}^{K\times K}$ is the diagonal pilot transmit power matrix with $[\mathbf{P}_{\text{p},i}]_{kk} = p_{\text{p},ik}$ assigned to the k-th user in cell i and $\mathbf{N}_l$ has i.i.d. $\mathcal{CN}(0,1)$ elements representing the normalized additive noise during the uplink training. The standard MMSE channel estimate is then given as [11]

$$\hat{\mathbf{G}}_{li} = \mathbf{Y}_{\text{p},l}\mathbf{\Phi}^*\left(\mathbf{I}_K + \sum_{j=1}^{L}\mathbf{D}_{lj}\mathbf{P}_{\text{p},j}\right)^{-1}\mathbf{D}_{li}\mathbf{P}_{\text{p},i}^{1/2}. \tag{3}$$

In principal, all cross channel gains $\mathbf{G}_{li}, \forall i \neq l$ could be estimated at BS l based on the received matrix in (2). However due to the high path-loss of cross channels, this benefit would be limited, hence only the estimation of own channel $\mathbf{G}_{ll}$ is performed. In addition, it is observed in (3) that the estimate of every channel at BS l is simply a scaled version of the same term $\mathbf{Y}_{\text{p},l}\mathbf{\Phi}^*\left(\mathbf{I}_K + \sum_{j=1}^{L}\mathbf{D}_{lj}\mathbf{P}_{\text{p},j}\right)^{-1}$, which is referred to in [5] as the pilot contamination. Thus we have the relation

$$\hat{\mathbf{G}}_{li} = \hat{\mathbf{G}}_{ll}\bar{\mathbf{D}}_{li} \quad\Longrightarrow\quad \hat{\mathbf{g}}_{lik} = \left[\bar{\mathbf{D}}_{li}\right]_{kk}\hat{\mathbf{g}}_{llk} \tag{4}$$

where $\bar{\mathbf{D}}_{li} = \text{diag}\left\{\left[\frac{\beta_{li1}}{\beta_{ll1}}\frac{p_{\text{p},i1}^{1/2}}{p_{\text{p},l1}^{1/2}},\ldots,\frac{\beta_{liK}}{\beta_{llK}}\frac{p_{\text{p},iK}^{1/2}}{p_{\text{p},lK}^{1/2}}\right]\right\}$. If we decompose $\hat{\mathbf{G}}_{li} \triangleq \mathbf{G}_{li} + \tilde{\mathbf{G}}_{li}$ with $\tilde{\mathbf{G}}_{li}$ denoting the estimation error matrix, from the property of MMSE estimation [11], the channel estimate $\hat{\mathbf{G}}_{li}$ is statistically independent of $\tilde{\mathbf{G}}_{li}$. Therefore, the distributions of the estimated channel vector $\hat{\mathbf{g}}_{lik}$ and error vector $\tilde{\mathbf{g}}_{lik}$ (which are the k-th column vectors of $\hat{\mathbf{G}}_{li}$ and $\tilde{\mathbf{G}}_{li}$ respectively) are given by

$$\hat{\mathbf{g}}_{lik} \sim \mathcal{CN}\left(\mathbf{0},\delta_{lik}\mathbf{I}_M\right) \tag{5}$$

$$\tilde{\mathbf{g}}_{lik} \sim \mathcal{CN}\left(\mathbf{0},\epsilon_{lik}\mathbf{I}_M\right) \tag{6}$$

where $\delta_{lik} \triangleq \frac{p_{\text{p},ik}\beta_{lik}^2}{1+\sum_{j=1}^{L}p_{\text{p},jk}\beta_{ljk}}$, $\epsilon_{lik} \triangleq \frac{\beta_{lik}(1+\sum_{j\neq i}^{L}p_{\text{p},jk}\beta_{ljk})}{1+\sum_{j=1}^{L}p_{\text{p},jk}\beta_{ljk}}$ and $\delta_{lik}+\epsilon_{lik} = \beta_{lik}$.

B. Linear MMSE Receiver

The BS in cell l has only the information of estimated channel matrix $\hat{\mathbf{G}}_{ll}$ which is treated as the true channel, thus (1) can be rewritten as

$$\mathbf{y}_l = \hat{\mathbf{G}}_{ll}\mathbf{P}_{\text{r},l}^{1/2}\mathbf{x}_l - \tilde{\mathbf{G}}_{ll}\mathbf{P}_{\text{r},l}^{1/2}\mathbf{x}_l + \sum_{i\neq l}^{L}\mathbf{G}_{li}\mathbf{P}_{\text{r},i}^{1/2}\mathbf{x}_i + \mathbf{n}_l. \tag{7}$$

Considering the last three terms in (7) as uncorrelated noise and assuming that the data symbols in $\mathbf{x}_i$ are i.i.d. Gaussian with zero mean and unit variance, the linear MMSE receiver $\hat{\mathbf{A}}_l$ at BS l is given by

$$\begin{aligned}\hat{\mathbf{A}}_l &= \arg\min_{\mathbf{A}_l}\mathbb{E}\left\{\left\|\mathbf{P}_{\text{r},l}^{1/2}\mathbf{x}_l - \mathbf{A}_l^H\mathbf{y}_l\right\|^2\Big|\hat{\mathbf{G}}_{ll}\right\} \\ &= \left(\hat{\mathbf{G}}_{ll}\mathbf{P}_{\text{r},l}\hat{\mathbf{G}}_{ll}^H + \left(E_{ll} + \sum_{i\neq l}^{L}B_{li} + 1\right)\mathbf{I}_M\right)^{-1}\hat{\mathbf{G}}_{ll}\mathbf{P}_{\text{r},l}\end{aligned} \tag{8}$$

where $E_{li} \triangleq \mathbb{E}\left\{\tilde{\mathbf{G}}_{li}\mathbf{P}_{\text{r},i}\tilde{\mathbf{G}}_{li}^H\right\} = \sum_{\kappa=1}^{K}p_{\text{r},i\kappa}\epsilon_{li\kappa}$ and $B_{li} \triangleq \mathbb{E}\left\{\mathbf{G}_{li}\mathbf{P}_{\text{r},i}\mathbf{G}_{li}^H\right\} = \sum_{\kappa=1}^{K}p_{\text{r},i\kappa}\beta_{li\kappa}$. Based on the relation of $\delta_{li\kappa}$, $\epsilon_{li\kappa}$ and $\beta_{li\kappa}$, it is easy to get $\Delta_{li} + E_{li} = B_{li}$ with $\Delta_{li} \triangleq \mathbb{E}\left\{\hat{\mathbf{G}}_{li}\mathbf{P}_{\text{r},i}\hat{\mathbf{G}}_{li}^H\right\} = \sum_{\kappa=1}^{K}p_{\text{r},i\kappa}\delta_{li\kappa}$. Denoting $\hat{\mathbf{\Lambda}}_{lk} \triangleq \sum_{\kappa\neq k}^{K}p_{\text{r},l\kappa}\hat{\mathbf{g}}_{ll\kappa}\hat{\mathbf{g}}_{ll\kappa}^H + \left(\sum_{i=1}^{L}E_{li} + \sum_{i\neq l}^{L}\Delta_{li} + 1\right)\mathbf{I}_M$, from Lemma 1 in the Appendix A, the k-th column vector of $\hat{\mathbf{A}}_l$ is given as

$$\hat{\mathbf{a}}_{lk} = \frac{p_{\text{r},lk}\hat{\mathbf{\Lambda}}_{lk}^{-1}\hat{\mathbf{g}}_{llk}}{1 + p_{\text{r},lk}\hat{\mathbf{g}}_{llk}^H\hat{\mathbf{\Lambda}}_{lk}^{-1}\hat{\mathbf{g}}_{llk}}. \tag{9}$$

III. Lower Bound on Average Uplink SINR

In this section, we compute the SINR of the uplink data transmission at BS l applying the introduced linear MMSE receiver under pilot contamination, and derive the lower bound on the average SINR over given (own) channel estimate for large M instead of the mean SINR which in general cannot be used for uplink power control [6]. The post-processing received signal $r_{lk} = \hat{\mathbf{a}}_{lk}^H\mathbf{y}_l$ for detecting x_{lk} at BS l is given by

$$\begin{aligned}r_{lk} = {}& p_{\text{r},lk}^{1/2}\hat{\mathbf{a}}_{lk}^H\hat{\mathbf{g}}_{llk}x_{lk} + \hat{\mathbf{a}}_{lk}^H\sum_{\kappa\neq k}^{K}p_{\text{r},l\kappa}^{1/2}\hat{\mathbf{g}}_{ll\kappa}x_{l\kappa} + \hat{\mathbf{a}}_{lk}^H\mathbf{n}_l \\ & - \hat{\mathbf{a}}_{lk}^H\sum_{i=1}^{L}\tilde{\mathbf{G}}_{li}\mathbf{P}_{\text{r},i}^{1/2}\mathbf{x}_i + \hat{\mathbf{a}}_{lk}^H\sum_{i\neq l}^{L}\underbrace{\hat{\mathbf{G}}_{li}\mathbf{P}_{\text{r},i}^{1/2}\mathbf{x}_i}_{\clubsuit}\end{aligned} \tag{10}$$

$$\hat{\theta}_{lk} = \frac{p_{\mathrm{r},lk}\left|\hat{\mathbf{a}}_{lk}^H\hat{\mathbf{g}}_{llk}\right|^2}{\sum_{\kappa\neq k}^{K} p_{\mathrm{r},l\kappa}\left|\hat{\mathbf{a}}_{lk}^H\hat{\mathbf{g}}_{ll\kappa}\right|^2 + \left(\sum_{i=1}^{L} E_{li}+1\right)\left\|\hat{\mathbf{a}}_{lk}^H\right\|^2 + \left(\sum_{i\neq l}^{L} p_{\mathrm{r},ik}\left[\bar{\mathbf{D}}_{li}\right]_{kk}^2\right)\left|\hat{\mathbf{a}}_{lk}^H\hat{\mathbf{g}}_{llk}\right|^2 + \sum_{i\neq l}^{L}\sum_{\kappa\neq k}^{K} p_{\mathrm{r},i\kappa}\left|\mathbf{a}_{lk}^H\hat{\mathbf{g}}_{li\kappa}\right|^2} \quad (11)$$

where only the first term in (10) is the desired information, while the other terms represent the intra-cell interference, noise, channel estimation error and inter-cell interference respectively. Since $\hat{\mathbf{G}}_{ll}$ and $\tilde{\mathbf{G}}_{li}$ are uncorrelated for all i, $\hat{\mathbf{A}}_l$ is also uncorrelated with $\tilde{\mathbf{G}}_{li}$. As a contrast, the term ♣ in (10) can be decomposed as $p_{\mathrm{r},ik}^{1/2}\hat{\mathbf{g}}_{lik}x_{ik} + \sum_{\kappa\neq k}^{K} p_{\mathrm{r},i\kappa}^{1/2}\hat{\mathbf{g}}_{li\kappa}x_{i\kappa}$, which is correlated with $\hat{\mathbf{a}}_{lk}$ and $\hat{\mathbf{g}}_{llk}$ because of the given relation in (4). As a result, the post-processing uplink SINR of user k in cell l is expressed in (11) shown at the top of the page, which can be used to obtain the ergodic achievable uplink rate by simply computing $\mathbb{E}\left\{\log_2\left(1+\hat{\theta}_{lk}\right)\right\}$ [6] [7].

Proposition 1. *In the asymptotic limit where $M\to\infty$, the deterministic equivalents of the following norms are identical for all $i\neq l$ and $\kappa\neq k$:*

$$\left|\mathbf{a}_{lk}^H\hat{\mathbf{g}}_{li\kappa}\right|^2 - \frac{\delta_{li\kappa}\delta_{llk}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-2}\right\}}{\left(p_{\mathrm{r},lk}^{-1}+\delta_{llk}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-1}\right\}\right)^2} \xrightarrow[M\to\infty]{\text{a.s.}} 0 \quad (12)$$

$$\delta_{li\kappa}\left\|\hat{\mathbf{a}}_{lk}^H\right\|^2 - \frac{\delta_{li\kappa}\delta_{llk}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-2}\right\}}{\left(p_{\mathrm{r},lk}^{-1}+\delta_{llk}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-1}\right\}\right)^2} \xrightarrow[M\to\infty]{\text{a.s.}} 0 \quad (13)$$

where "$\xrightarrow[M\to\infty]{\text{a.s.}}$" denotes almost sure convergence.

Proof: See Appendix B. ∎

We apply Proposition 1 where $\left|\mathbf{a}_{lk}^H\hat{\mathbf{g}}_{li\kappa}\right|^2 \approx \delta_{li\kappa}\left\|\hat{\mathbf{a}}_{lk}^H\right\|^2$ can be assumed for large M and the identity $\hat{\mathbf{a}}_{lk}^H\hat{\mathbf{\Lambda}}_{lk}\hat{\mathbf{a}}_{lk} = \sum_{\kappa\neq k}^{K} p_{\mathrm{r},l\kappa}\left|\hat{\mathbf{a}}_{lk}^H\hat{\mathbf{g}}_{ll\kappa}\right|^2 + \left(\sum_{i=1}^{L}E_{li} + \sum_{i\neq l}^{L}\Delta_{li}+1\right)\left\|\hat{\mathbf{a}}_{lk}^H\right\|^2$ to (11), the average of reciprocal uplink SINR is given as

$$\mathbb{E}\left\{\hat{\theta}_{lk}^{-1}\right\} \approx p_{\mathrm{r},lk}^{-1}\sum_{i\neq l}^{L} p_{\mathrm{r},ik}\left[\bar{\mathbf{D}}_{li}\right]_{kk}^2 + \underbrace{\mathbb{E}\left\{\frac{1}{p_{\mathrm{r},lk}\hat{\mathbf{g}}_{llk}^H\hat{\mathbf{\Lambda}}_{lk}^{-1}\hat{\mathbf{g}}_{llk}}\right\}}_{\mathrm{i}} - \left(p_{\mathrm{r},lk}^{-1}\sum_{i\neq l}^{L} p_{\mathrm{r},ik}\delta_{lik}\right)\underbrace{\mathbb{E}\left\{\frac{1}{\left\|\hat{\mathbf{g}}_{llk}\right\|^2}\right\}}_{\mathrm{ii}}. \quad (14)$$

Furthermore, Jensen's inequality on the convex function $1/x$ ($\forall x>0$) leads to

$$\hat{\theta}_{\mathrm{lb},lk} \triangleq \left(\mathbb{E}\left\{\hat{\theta}_{lk}^{-1}\right\}\right)^{-1} \leq \mathbb{E}\left\{\hat{\theta}_{lk}\right\} \quad (15)$$

where $\hat{\theta}_{\mathrm{lb},lk}$ is the lower bound on the average uplink SINR (over given $\hat{\mathbf{g}}_{llk}$) of user k in cell l. In the following we compute two expectations marked as i and ii in (14) separately.

i) We denote $z_k \triangleq p_{\mathrm{r},lk}\hat{\mathbf{g}}_{llk}^H\hat{\mathbf{\Lambda}}_{lk}^{-1}\hat{\mathbf{g}}_{llk}$. Following the same manipulations as in [4], we have

$$z_k = \left(\left[\left(\mathbf{I}_K + \frac{\mathbf{P}_{\mathrm{r},l}^{1/2}\hat{\mathbf{G}}_{ll}^H\hat{\mathbf{G}}_{ll}\mathbf{P}_{\mathrm{r},l}^{1/2}}{E_{ll}+\sum_{i\neq l}^{L}B_{li}+1}\right)^{-1}\right]_{kk}\right)^{-1} - 1. \quad (16)$$

As discussed in [12], z_k can be modelled as a Gamma distribution with the probability density function $p_{z_k}(z) = \frac{z^{\alpha_k-1}e^{-z/\vartheta_k}}{\Gamma(\alpha_k)\vartheta_k^{\alpha_k}}$, in which the shape parameter α_k and scale parameter ϑ_k are given as

$$\begin{cases} \alpha_k = & \frac{(M-K+1+(K-1)\mu)^2}{M-K+1+(K-1)\sigma^2} \\ \vartheta_k = & \frac{M-K+1+(K-1)\sigma^2}{M-K+1+(K-1)\mu}\Sigma_k \end{cases}. \quad (17)$$

With each row vector of $\sqrt{\rho}\hat{\mathbf{G}}_{ll}\mathbf{P}_{\mathrm{r},l}^{1/2} \sim \mathcal{CN}(\mathbf{0},\mathbf{R}_l)$, $\rho \triangleq \frac{1}{E_{ll}+\sum_{i\neq l}^{L}B_{li}+1}$ and $\mathbf{R}_l = \rho\frac{\mathbf{D}_{ll}^2\mathbf{P}_{\mathrm{p},l}}{\mathbf{I}_K+\sum_{i=1}^{L}\mathbf{D}_{li}\mathbf{P}_{\mathrm{p},i}}\mathbf{P}_{\mathrm{r},l}$, we define $\Sigma_k \triangleq \left(\left[\mathbf{R}_l^{-1}\right]_{kk}\right)^{-1}$ and $\zeta_\kappa \triangleq \left(1-\frac{K-1}{M}+\frac{K-1}{M}\mu\right)M\upsilon_\kappa+1$, in which υ_κ ($\forall\kappa\neq k$) is the κ-th eigenvalue of $\mathbf{R}_l$. The parameters μ and σ^2 in (17) are determined by solving the following equations

$$\begin{cases} \mu = & \frac{1}{K-1}\sum_{\kappa\neq k}^{K}\zeta_\kappa^{-1} \\ \sigma^2\left(1+\sum_{\kappa\neq k}^{K}\frac{\upsilon_\kappa}{\zeta_\kappa^2}\right) = & \frac{1}{K-1}\sum_{\kappa\neq k}^{K}\frac{(K-1)\mu\upsilon_\kappa+1}{\zeta_\kappa^2} \end{cases}. \quad (18)$$

Therefore the mean value of $1/z_k$ is computed as

$$\mathbb{E}\left\{\frac{1}{z_k}\right\} = \int_0^\infty \frac{1}{z}p_{z_k}(z)\,\mathrm{d}z = \frac{1}{(\alpha_k-1)\,\vartheta_k}. \quad (19)$$

Substituting (17) into (19), we obtain for large M

$$\mathbb{E}\left\{\frac{1}{p_{\mathrm{r},lk}\hat{\mathbf{g}}_{llk}^H\hat{\mathbf{\Lambda}}_{lk}^{-1}\hat{\mathbf{g}}_{llk}}\right\} = \frac{E_{ll}+\sum_{i\neq l}^{L}B_{li}+1}{Mp_{\mathrm{r},lk}\delta_{llk}}. \quad (20)$$

ii) As shown in (3) and (5), it is straightforward to get $\left\|\hat{\mathbf{g}}_{llk}\right\|^2 = \delta_{llk}\left\|\mathbf{h}_{llk}\right\|^2$ with $\left\|\mathbf{h}_{llk}\right\|^2 \sim \mathcal{W}_1(M,1)$ being a 1×1 central complex Wishart random variable of M degrees of freedom [3] [4]. As a result, we have for $M\geq2$

$$\mathbb{E}\left\{\frac{1}{\left\|\hat{\mathbf{g}}_{llk}\right\|^2}\right\} = \frac{1}{(M-1)\delta_{llk}}. \quad (21)$$

Inserting (20), (21) in (14) and using (15), we finally obtain the lower bound on the average uplink SINR for large M as

$$\hat{\theta}_{\mathrm{lb},lk} = \frac{Mp_{\mathrm{r},lk}\delta_{llk}}{M\sum_{i\neq l}^{L}p_{\mathrm{r},ik}\delta_{lik}+E_{ll}+\sum_{i\neq l}^{L}B_{li}+1}. \quad (22)$$

Remark 1. *The closed-form lower bound in* (22) *is applicable in general scenarios with random user locations while the deterministic equivalent SINR obtained in [6] has a tractable form only in special setups, e.g. in the Wyner-type model with the same path-loss for all interfering users.*

Remark 2. *In the asymptotic limit as $M\to\infty$, the ultimate SINR of the k-th user in cell l converges to $\frac{p_{\mathrm{r},lk}p_{\mathrm{p},lk}\beta_{llk}^2}{\sum_{i\neq l}^{L}p_{\mathrm{r},ik}p_{\mathrm{p},ik}\beta_{lik}^2}$, which reflects the effect of pilot contamination and is a generalized result with individual power allocation of each user compared to the observations in [5]-[7]. Moreover, if we let the pilot and data transmit power of all users be p_{p} and*

$p_{\mathrm{r}} = E_{\mathrm{r}}/\sqrt{M}$ *respectively, the SINR converges to* $\frac{E_{\mathrm{r}}\beta_{llk}^2}{E_{\mathrm{r}}\sum_{i\neq l}^{L}\beta_{lik}^2}$ *for user* k *at BS* l *as* $M \to \infty$*, which verifies the power-scaling law that the uplink data power of each user can be reduced proportional to* $\sqrt{M}$ *without performance degradation [3].*

IV. Uplink Pilot and Data Power Control

In this section, we formulate the multi-cell uplink power control problem and propose an iterative algorithm to find the optimal solution based on the obtained lower bound on the average SINR. As we aims at minimizing the sum transmit power of all users while meeting their pre-defined SINR targets, the optimization problem can be given as

$$\begin{aligned} \underset{p_{\mathrm{p},lk},\, p_{\mathrm{r},lk}}{\text{minimize}} \quad & \sum_{l=1}^{L}\sum_{k=1}^{K}\left(p_{\mathrm{p},lk} + p_{\mathrm{r},lk}\right) \\ \text{subject to} \quad & \hat{\theta}_{\mathrm{lb},lk} \geq \gamma_{lk} \\ & p_{\mathrm{p},lk} + p_{\mathrm{r},lk} \leq p_{lk} \quad \forall (l,k) \end{aligned} \tag{23}$$

where γ_{lk} and p_{lk} are the per-user SINR target and per-user transmit power constraint respectively. Since the objective in (23) involves the summation over $p_{\mathrm{p},lk}$ and $p_{\mathrm{r},lk}$ which appear in the expression of $\hat{\theta}_{\mathrm{lb},lk}$ in (22), it is very difficult to find the solution through a direct approach. As a result, we divide the optimization into two consecutive parts: In the part where the sum pilot transmit power $\sum_{l=1}^{L}\sum_{k=1}^{K} p_{\mathrm{p},lk}$ is minimized, we assume the data power $p_{\mathrm{r},lk}$, $\forall (l,k)$ is fixed; and vice versa for the other part.

A. Data Transmit Power Control with Fixed Pilot Power

For fixed pilot power, the objective function in (23) reduces to $\sum_{l=1}^{L}\sum_{k=1}^{K} p_{\mathrm{r},lk}$ with the per-user SINR constraint being

$$\begin{aligned} p_{\mathrm{r},lk} \geq \frac{\gamma_{lk}}{M\delta_{llk}} \Bigg[& \sum_{i=1}^{L}\sum_{\kappa=1}^{K} p_{\mathrm{r},i\kappa}\beta_{li\kappa} - \sum_{\kappa=1}^{K} p_{\mathrm{r},l\kappa}\delta_{ll\kappa} \\ & + M\sum_{i\neq l}^{L} p_{\mathrm{r},ik}\delta_{lik} + 1 \Bigg]. \end{aligned} \tag{24}$$

By defining $\mathbf{p}_{\mathrm{r}} \triangleq \left[p_{\mathrm{r},11},\ldots,p_{\mathrm{r},1K},\ldots,p_{\mathrm{r},L1},\ldots,p_{\mathrm{r},LK}\right]^T$, the interference function $\boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right)$ is expressed as

$$\boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right) = \boldsymbol{\Psi}\left(\mathbf{C}\mathbf{p}_{\mathrm{r}} + \mathbf{1}_{KL\times 1}\right) \tag{25}$$

where $\mathbf{C} \triangleq \mathbf{B} + \boldsymbol{\Delta}$ with $\mathbf{B}$ and $\boldsymbol{\Delta}$ being block matrices, $\boldsymbol{\Psi} \triangleq \mathrm{diag}\left\{\left[\frac{\gamma_{11}}{M\delta_{111}},\ldots,\frac{\gamma_{1K}}{M\delta_{11K}},\ldots,\frac{\gamma_{L1}}{M\delta_{LL1}},\ldots,\frac{\gamma_{LK}}{M\delta_{LLK}}\right]\right\}$ is a diagonal matrix, and $\mathbf{1}$ denotes a vector whose elements are all 1. According to (24), the l-th on-diagonal submatrix of $\boldsymbol{\Delta}$ is given as $[\boldsymbol{\Delta}]_{ll} = -\mathbf{1}_{K\times 1}\cdot[\delta_{ll1},\ldots,\delta_{llK}]$, while its off-diagonal submatrix has the form $[\boldsymbol{\Delta}]_{li} = \mathrm{diag}\left\{M\cdot[\delta_{li1},\ldots,\delta_{liK}]\right\}$; the other block matrix $\mathbf{B} = \left[\mathbf{B}_1^T,\ldots,\mathbf{B}_L^T\right]^T$ with its submatrix $\mathbf{B}_l = \mathbf{1}_{K\times 1}\cdot[\beta_{l11},\ldots,\beta_{l1K},\ldots,\beta_{lL1},\ldots,\beta_{lLK}]$. Therefore, the per-user SINR requirement in (24) can be rewritten in a vector inequality form

$$\mathbf{p}_{\mathrm{r}} \geq \boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right). \tag{26}$$

The uplink data transmit power control has a *feasible* solution $\mathbf{p}_{\mathrm{r}} \geq \mathbf{0}$ fulfilling the per-user SINR constraint in (26) if the maximum eigenvalue of $\boldsymbol{\Psi}\mathbf{C}$ is less than 1, i.e. $\lambda_{\max}\left(\boldsymbol{\Psi}\mathbf{C}\right) < 1$ [10]. Moreover, the optimal solution can be obtained directly from $\mathbf{p}_{\mathrm{r}}^* = \left(\mathbf{I}_{KL} - \boldsymbol{\Psi}\mathbf{C}\right)^{-1}\boldsymbol{\Psi}\cdot\mathbf{1}_{KL\times 1}$, which is a fixed point result as it satisfies $\mathbf{p}_{\mathrm{r}}^* = \boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}^*\right)$. However, as we consider massive MIMO systems in this paper, we should avoid computing the inverse of a matrix with large dimension. As a result, we apply an iterative method as

$$\mathbf{p}_{\mathrm{r}}(n_{\mathrm{r}}+1) = \boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\left(n_{\mathrm{r}}\right)\right) \tag{27}$$

where n_{r} denotes the iteration index. As discussed in [10], if the interference function $\boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right)$ is *standard*, the iterative approach in (27) converges for any initial power vector $\mathbf{p}_{\mathrm{r}}\left(n_{\mathrm{r}}=0\right)$ to a unique fixed point which is the optimal solution to the data power control as long as a feasible solution exists. As a result, we prove that $\boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right)$ in (25) is standard.

Proposition 2. *The interference function in* (25) *is standard for* $\mathbf{p}_{\mathrm{r}} \geq \mathbf{0}$*, as it satisfies the positivity, monotonicity and scalability.*

Proof: $\boldsymbol{\Psi}$ is a diagonal matrix with positive entries and $\mathbf{C}$ has only positive elements, thus for $\mathbf{p}_{\mathrm{r}} \geq \mathbf{0}$

- Positivity: $\boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right) = \boldsymbol{\Psi}\mathbf{C}\mathbf{p}_{\mathrm{r}} + \boldsymbol{\Psi}\cdot\mathbf{1}_{KL\times 1} > \mathbf{0}$;
- Monotonicity: $\boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}} + \mathbf{p}_{\mathrm{r}}'\right) = \boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right) + \boldsymbol{\Psi}\mathbf{C}\mathbf{p}_{\mathrm{r}}' > \boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right), \forall \mathbf{p}_{\mathrm{r}}' > 0$;
- Scalability: $\alpha\boldsymbol{I}_{\mathrm{r}}\left(\mathbf{p}_{\mathrm{r}}\right) = \alpha\boldsymbol{\Psi}\mathbf{C}\mathbf{p}_{\mathrm{r}} + \alpha\boldsymbol{\Psi}\cdot\mathbf{1}_{KL\times 1} > \boldsymbol{\Psi}\mathbf{C}\left(\alpha\mathbf{p}_{\mathrm{r}}\right) + \boldsymbol{\Psi}\cdot\mathbf{1}_{KL\times 1} = \boldsymbol{I}_{\mathrm{r}}\left(\alpha\mathbf{p}_{\mathrm{r}}\right), \forall \alpha > 1$. ∎

B. Pilot Transmit Power Control with Fixed Data Power

For fixed data power, the objective function in (23) becomes $\sum_{l=1}^{L}\sum_{k=1}^{K} p_{\mathrm{p},lk}$, and the per-user SINR constraint is

$$\begin{aligned} p_{\mathrm{p},lk} \geq \frac{\gamma_{lk}}{M\beta_{llk}^2 p_{\mathrm{r},lk}} \Bigg[& \left(1 + \sum_{i=1}^{L} B_{li}\right)\sum_{i=1}^{L} p_{\mathrm{p},ik}\beta_{lik} \\ & \underbrace{- \left(1 + \sum_{i=1}^{L} p_{\mathrm{p},ik}\beta_{lik}\right)\sum_{\kappa=1}^{K}\frac{p_{\mathrm{r},l\kappa}p_{\mathrm{p},l\kappa}\beta_{ll\kappa}^2}{1+\sum_{i=1}^{L} p_{\mathrm{p},i\kappa}\beta_{li\kappa}}}_{\spadesuit} \\ & + M\sum_{i\neq l}^{L} p_{\mathrm{r},ik}p_{\mathrm{p},ik}\beta_{lik}^2 + \left(1 + \sum_{i=1}^{L} B_{li}\right) \Bigg]. \end{aligned} \tag{28}$$

In contrast to the data power control, the per-user SINR requirement in this case cannot be formulated into a vector inequality due to the term ♠ in (28), which is non-linear with respect to the pilot power. Nevertheless, we still define $\mathbf{p}_{\mathrm{p}} \triangleq \left[p_{\mathrm{p},11},\ldots,p_{\mathrm{p},1K},\ldots,p_{\mathrm{p},L1},\ldots,p_{\mathrm{p},LK}\right]^T$ and stack accordingly the right-hand side of (28) into the interference vector $\boldsymbol{I}_{\mathrm{p}}\left(\mathbf{p}_{\mathrm{p}}\right)$. Similarly we propose an iterative approach for the uplink pilot transmit power control as

$$\mathbf{p}_{\mathrm{p}}(n_{\mathrm{p}}+1) = \boldsymbol{I}_{\mathrm{p}}\left(\mathbf{p}_{\mathrm{p}}\left(n_{\mathrm{p}}\right)\right) \tag{29}$$

where n_{p} denotes the iteration index. In a special case where there is only one user in each cell, i.e. $K = 1$, referring to Proposition 2, the interference function $\boldsymbol{I}_{\mathrm{p}}\left(\mathbf{p}_{\mathrm{p}}\right)$ is standard, thus the iterative algorithm in (29) results in a unique optimal solution. It is difficult to analytically evaluate the feasibility and convergence of (29) in a multi-user setup due to the interference function being non-linear over $\mathbf{p}_{\mathrm{p}}$, however, an extra requirement can be inserted in each iteration which ensures the monotone decreasing of the output.

C. Joint Iterative Algorithm for Pilot and Data Power Control

The general idea of the joint iterative algorithm is to perform the pilot power control with prior data power output, if the per-user power constraint is currently satisfied. Based on the obtained pilot power allocation, the standard data power control converges to a unique optimal solution which is used as updated input for the subsequent pilot power control. The algorithm is stated below:

Algorithm 1 Joint Pilot and Data Transmit Power Control

1: *Initialization:*
Let iteration indices $n = n_{\mathrm{p}} = n_{\mathrm{r}} = 0$; Generate inner loop power vectors $\mathbf{p}_{\mathrm{p}}(n_{\mathrm{p}})$ and $\mathbf{p}_{\mathrm{r}}(n_{\mathrm{r}})$ with an initial pilot and data power allocation; Generate global outer loop power vectors $\mathbf{p}'_{\mathrm{p}}(n) = \mathbf{p}_{\mathrm{p}}(n_{\mathrm{p}})$ and $\mathbf{p}'_{\mathrm{r}}(n) = \mathbf{p}_{\mathrm{r}}(n_{\mathrm{r}})$.
2: *Start outer loop:*
While $p'_{\mathrm{p},lk}(n) + p'_{\mathrm{r},lk}(n) \leq p_{lk}$, $\forall(l,k)$
Let $n = n + 1$;
Reset $n_{\mathrm{p}} = n_{\mathrm{r}} = 0$;
Reset $\mathbf{p}_{\mathrm{p}}(n_{\mathrm{p}}) = \mathbf{p}'_{\mathrm{p}}(n-1)$ and $\mathbf{p}_{\mathrm{r}}(n_{\mathrm{r}}) = \mathbf{p}'_{\mathrm{r}}(n-1)$;
Reset convergence coefficients $\eta_{\mathrm{p}} = \eta_{\mathrm{r}} = \eta$;
3: *Start inner loop for pilot power control:*
While $\eta_{\mathrm{p}} \geq \eta$
Let $n_{\mathrm{p}} = n_{\mathrm{p}} + 1$;
Compute $\mathbf{p}_{\mathrm{p}}(n_{\mathrm{p}}) = \boldsymbol{I}_{\mathrm{p}}(\mathbf{p}_{\mathrm{p}}(n_{\mathrm{p}}-1))$ for fixed $\mathbf{p}'_{\mathrm{r}}(n-1)$;
If $\sum_{(l,k)} p_{\mathrm{p},lk}(n_{\mathrm{p}}) > \sum_{(l,k)} p_{\mathrm{p},lk}(n_{\mathrm{p}}-1)$
If $n_{\mathrm{p}} = 1$ && $n > 1$
Return $\mathbf{p}'_{\mathrm{p}}(n-1)$ and $\mathbf{p}'_{\mathrm{r}}(n-1)$.
Else
Update outer loop vector $\mathbf{p}'_{\mathrm{p}}(n) = \mathbf{p}_{\mathrm{p}}(n_{\mathrm{p}}-1)$;
Go to 5;
Else
Compute $\eta_{\mathrm{p}} = \max_{(l,k)} \frac{|p_{\mathrm{p},lk}(n_{\mathrm{p}}) - p_{\mathrm{p},lk}(n_{\mathrm{p}}-1)|}{p_{\mathrm{p},lk}(n_{\mathrm{p}}-1)}$;
4: *After inner loop for pilot power control:*
Update outer loop vector $\mathbf{p}'_{\mathrm{p}}(n) = \mathbf{p}_{\mathrm{p}}(n_{\mathrm{p}})$;
5: *Start inner loop for data power control:*
While $\eta_{\mathrm{r}} \geq \eta$
Let $n_{\mathrm{r}} = n_{\mathrm{r}} + 1$;
Compute $\mathbf{p}_{\mathrm{r}}(n_{\mathrm{r}}) = \boldsymbol{I}_{\mathrm{r}}(\mathbf{p}_{\mathrm{r}}(n_{\mathrm{r}}-1))$ for fixed $\mathbf{p}'_{\mathrm{p}}(n)$;
Compute $\eta_{\mathrm{r}} = \max_{(l,k)} \frac{|p_{\mathrm{r},lk}(n_{\mathrm{r}}) - p_{\mathrm{r},lk}(n_{\mathrm{r}}-1)|}{p_{\mathrm{r},lk}(n_{\mathrm{r}}-1)}$;
6: *After inner loop for data power control:*
Update outer loop vector $\mathbf{p}'_{\mathrm{r}}(n) = \mathbf{p}_{\mathrm{r}}(n_{\mathrm{r}})$;

Provided the problem is feasible, Algorithm 1 converges to a unique fixed point optimal solution for any initial power allocation in single-user setup, since each inner loop itself is a fixed point optimization. In contrast, for the multi-user scenario, the constraint $\sum_{(l,k)} p_{\mathrm{p},lk}(n_{\mathrm{p}}) \leq \sum_{(l,k)} p_{\mathrm{p},lk}(n_{\mathrm{p}}-1)$ requires the outcome of each iteration in pilot power control to decrease monotonically, otherwise the smallest result so far from the previous iteration is immediately output to the next inner loop for data power control, where a fixed point optimal result based on the current pilot power allocation is ensured. Therefore, at the end of each outer loop iteration, we have

$$\begin{cases} \sum_{(l,k)} p'_{\mathrm{p},lk}(n) \leq \sum_{(l,k)} p'_{\mathrm{p},lk}(n-1) \\ \sum_{(l,k)} p'_{\mathrm{r},lk}(n) \leq \sum_{(l,k)} p'_{\mathrm{r},lk}(n-1) \end{cases} \tag{30}$$

which proves that the sum power of all users after each outer loop iteration is monotone decreasing and converges since bounded below.

In addition, Algorithm 1 can be reformulated easily in a distributed manner, where each BS performs the algorithm in turn based on the local information and decides the power allocation of its own users. However, due to space limitation, we omit the extension here.

V. Numerical Results

We consider a MU-Massive-MIMO system consisting of $L = 3$ hexagonal cells which have a radius of 1000 meters. Each BS located in the cell center serves $K = 5$ users at the same time-frequency resource. All users are distributed uniformly inside the cell and have at least a distance of 100 meters away from the BS. The path-loss exponent is selected to be 4 and the large-scale fading is modelled as a zero mean log-normal distribution with a standard deviation of 8 dB. Throughout the simulations in this section, normalized additive Gaussian noise with variance of 1 is assumed; the same target SINR and power allowed for uplink transmission are applied for all users in the system, i.e. $\gamma_{lk} = \gamma$, $p_{lk} = p = 200\,\mathrm{mW}$, $\forall(l,k)$; referring to [9], around $1/6$ of the transmit power is assigned initially to the uplink training[1].

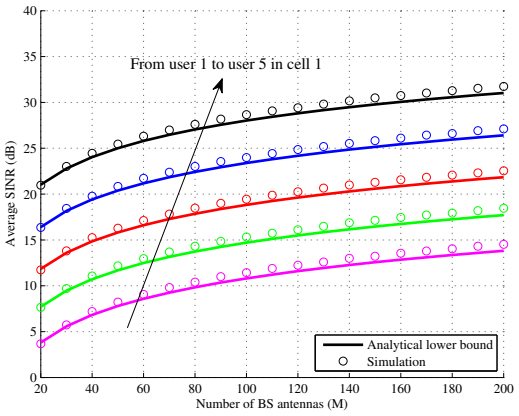

Fig. 1. Lower bound on average SINR over 10^5 channel realizations.

In Fig. 1 we depict the tightness of the lower bound on the average uplink SINR for all users in the 1-st cell based on a chosen user placement[2]. Without loss of generality, equal pilot and data power allocation among users is applied, i.e. $p_{\mathrm{p},lk} = \frac{1}{6}p$ and $p_{\mathrm{r},lk} = \frac{1}{2}p$, $\forall(l,k)$. Numerical results demonstrate that

[1]The sum power of all users is $10\log_{10}(K \times L \times p) = 34.8$ dBm and the initial sum pilot power is $10\log_{10}(K \times L \times \frac{1}{6}p) = 27$ dBm.

[2]One realization of user placement is chosen where the path-loss and large-scale fading matrices are given as

$$\mathbf{D}_{11} = \mathrm{diag}\{[0.8227, 0.3011, 0.1197, 0.0566, 0.0301]\}$$
$$\mathbf{D}_{12} = 10^{-4} \times \mathrm{diag}\{[0.1019, 0.1223, 0.0952, 0.0718, 0.1176]\}$$
$$\mathbf{D}_{13} = 10^{-4} \times \mathrm{diag}\{[0.1299, 0.0908, 0.0767, 0.0952, 0.0732]\}$$
$$\mathbf{D}_{21} = 10^{-4} \times \mathrm{diag}\{[0.1046, 0.1481, 0.1386, 0.0784, 0.1647]\}$$
$$\mathbf{D}_{22} = \mathrm{diag}\{[0.6830, 0.3501, 0.1526, 0.0767, 0.0357]\}$$
$$\mathbf{D}_{23} = 10^{-4} \times \mathrm{diag}\{[0.1343, 0.1148, 0.0955, 0.1133, 0.1901]\}$$
$$\mathbf{D}_{31} = 10^{-4} \times \mathrm{diag}\{[0.1321, 0.1601, 0.1603, 0.0834, 0.0681]\}$$
$$\mathbf{D}_{32} = 10^{-4} \times \mathrm{diag}\{[0.1475, 0.1619, 0.1013, 0.0696, 0.1027]\}$$
$$\mathbf{D}_{33} = \mathrm{diag}\{[0.4823, 0.0953, 0.0625, 0.0468, 0.0264]\}.$$

the lower bound for all users is very tight even for a moderate number of BS antennas. Moreover, as M grows, the SINRs start to saturate due to the pilot contamination which is in accordance with the discussions in Section III.

Next we illustrate the performance of the proposed joint iterative power control based on the previous user placement and initial power. In the upper part of Fig. 2, e.g. for $M = 100$, the optimal assigned sum pilot power is larger than the sum data power except for very low target SINRs, which coincides with the fact that only data transmission takes advantage of massive BS antennas (i.e. power-scaling law), while uplink training does not, since it is performed on a per-receive antenna basis [3]. Due to the same reason, for γ up to 10.7 dB, the obtained data power allocation leads to a sum data power varying from 12.6 to 17.4 dBm as the pursued SINR rises, which is much smaller than the increment from 11.4 to 27 dBm in sum pilot power. In contrast, in the high target SINR region with $\gamma \geq 10.7$ dB, the inner loop for pilot power control is terminated by the monotone decreasing constraint only after 1 iteration, because the pilot contamination confines the impact of pilot power control. As a result, the joint algorithm reduces to a simple data power control with fixed (initial) pilot power. On the other hand, as M grows, the lower limit of high target SINR region also increases, which extends the region where power saving can be obtained through joint power control. In addition, as depicted in both parts of Fig. 2, the joint algorithm becomes infeasible when the target SINR is above a certain value, e.g. 15 dB for $M = 100$, due to the per-user power constraint. However, as M becomes large, this upper limit increases as well, which reflects the potential of combining the joint uplink power control with massive MIMO technique to attain a higher SINR target with less sum transmit power, hence improve the energy efficiency.

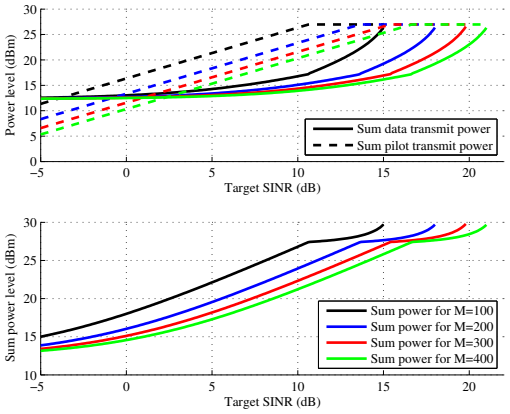

Fig. 2. Joint iterative pilot and data power control vs. target SINRs.

Moreover, as shown in Fig. 3, we demonstrate the advantage in power saving of the proposed joint algorithm comparing to a simple data power control where the pilot power allocation is fixed to the initial value as in the joint algorithm, i.e. $\frac{1}{6}p$ for all users. Here the y-axis represents $\frac{\sum_{(l,k)}\left(\frac{1}{6}p+p_{\mathrm{r},lk}^{(\mathrm{dpc})}\right)-\sum_{(l,k)}\left(p'_{\mathrm{p},lk}+p'_{\mathrm{r},lk}\right)}{\sum_{(l,k)}\left(\frac{1}{6}p+p_{\mathrm{r},lk}^{(\mathrm{dpc})}\right)} \cdot 100$, where $p_{\mathrm{r},lk}^{(\mathrm{dpc})}$ stands for the obtained data power after simple data power control. It is obvious that a power saving up to 95% can be achieved with joint algorithm in low target SINR region. The advantage disappears however when the target SINR exceeds the lower limit of high SINR region. It is especially worth noting that the benefit of deploying a large number of BS antennas tends to become marginal as M keeps growing, since the ultimate SINR performance is exclusively limited by the pilot contamination and transmit power control.

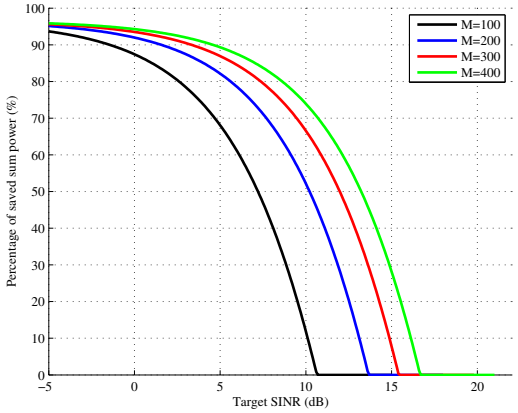

Fig. 3. Percentage of power saving vs. target SINRs.

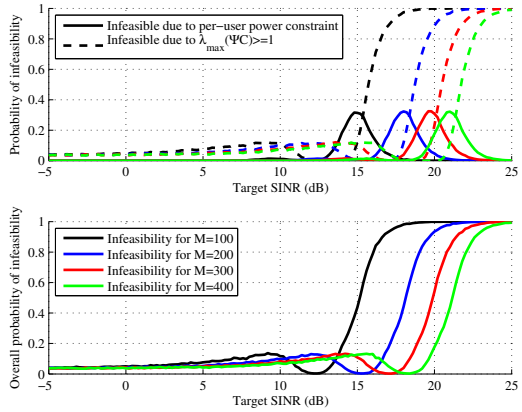

Fig. 4. Probability of infeasibility over 10^5 random user placements.

Finally we evaluate the infeasibility of the joint algorithm which is caused by either the per-user power constraint or the unattainable per-user SINR target. The latter is checked based on the obtained pilot power prior to the data power control to avoid unnecessary computations if the SINR target is infeasible. The results in the upper part of Fig. 4 verify that the significant infeasibility ($> 20\%$) first occurs due to the per-user power limit. However as the target SINR increases, it tends to be simply unachievable with the joint power control, thus $\lambda_{\max}(\mathbf{\Psi C}) \geq 1$ starts to become the dominant cause of infeasibility. In particular, the overall probability in the lower part of Fig. 4 has a decreasing behavior (e.g. around 10 dB for $M = 100$) due to the fact that the joint algorithm reduces to a simple data power control. Although the joint power control has a large gain in power saving, the inner loop of pilot power control may result in a case where some of the cell edge

users are infeasible to attain the target SINR, which can be however achieved at the expense of higher sum transmit power, if only the data power control with fixed initial pilot power is applied. Besides, Fig. 4 also illustrates that by deploying a larger number of BS antennas the achievable target SINR is increased, e.g. 3 dB gain from 100 to 200 BS antennas.

VI. Conclusion

In this paper we investigated the significance of uplink power control for energy-efficient communications in multi-cell MU-Massive-MIMO systems. The main contributions are twofold: i) The closed-form lower bound on the average uplink SINR with linear MMSE receiver has been derived, which is applicable with random user placement and takes the pilot contamination as well as the individual pilot and data power assignment of each user into account; ii) a joint iterative pilot and data power control has been proposed, which targets at minimizing the sum transmit power of all users while satisfying the per-user SINR and per-user power constraints. The existence of a unique fixed point optimal solution is proved straightforward for the single-user case, while for the multi-user setup, a monotone decreasing requirement has been applied to ensure that the joint algorithm converges. Numerical results have demonstrated the tightness of the obtained lower bound and the attainable power saving of the joint algorithm comparing to a simple data power control with fixed pilot power, which comes with the trade-off in a bit higher infeasibility. Nevertheless, if combined with a very large number of BS antennas, the joint uplink power control can achieve a higher target SINR with less sum transmit power and significantly improve the system energy efficiency.

Appendix A
Useful Lemmas

Lemma 1. *Let $\mathbf{A} \in \mathbb{C}^{M\times M}$ be Hermitian and invertible. For any vector $\mathbf{x} \in \mathbb{C}^{M\times 1}$ and any scalar $c \in \mathbb{C}$ such that $\mathbf{A} + c\mathbf{x}\mathbf{x}^H$ is invertible, Matrix Inversion Lemma yields*

$$\mathbf{x}^H \left(\mathbf{A} + c\mathbf{x}\mathbf{x}^H\right)^{-1} = \frac{\mathbf{x}^H\mathbf{A}^{-1}}{1 + c\mathbf{x}^H\mathbf{A}^{-1}\mathbf{x}}. \tag{31}$$

Lemma 2. *Consider $\mathbf{A} \in \mathbb{C}^{M\times M}$ with uniformly bounded spectral norm (with respect to M) and $\mathbf{x} \in \mathbb{C}^{M\times 1}$ where $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\mathbf{x})$. Then, for $\mathbf{x}$ being independent of $\mathbf{A}$,*

$$\frac{1}{M}\mathbf{x}^H\mathbf{A}\mathbf{x} - \frac{1}{M}\mathrm{tr}\left\{\mathbf{A}\mathbf{R}_\mathbf{x}\right\} \xrightarrow[M\to\infty]{\text{a.s.}} 0. \tag{32}$$

Appendix B
Proof of Proposition 1

Starting with (13) and denoting $a_M \asymp b_M$ as equivalent to $a_M - b_M \xrightarrow[M\to\infty]{\text{a.s.}} 0$ for two infinite sequences a_M and b_M, Lemma 2 and (9) lead directly to

$$\begin{aligned}\delta_{li\kappa}\left\|\hat{\mathbf{a}}_{lk}^H\right\|^2 &= \frac{\delta_{li\kappa}\hat{\mathbf{g}}_{llk}^H\hat{\mathbf{\Lambda}}_{lk}^{-2}\hat{\mathbf{g}}_{llk}}{\left(p_{\mathrm{r},lk}^{-1} + \hat{\mathbf{g}}_{llk}^H\hat{\mathbf{\Lambda}}_{lk}^{-1}\hat{\mathbf{g}}_{llk}\right)^2} \\ &\asymp \frac{\delta_{li\kappa}\delta_{llk}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-2}\right\}}{\left(p_{\mathrm{r},lk}^{-1} + \delta_{llk}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-1}\right\}\right)^2}.\end{aligned} \tag{33}$$

Next we proceed to $\left|\mathbf{a}_{lk}^H\hat{\mathbf{g}}_{li\kappa}\right|^2$. For all $i \neq l$ and $\kappa \neq k$, Lemma 2 and (9) yield

$$\left|\mathbf{a}_{lk}^H\hat{\mathbf{g}}_{li\kappa}\right|^2 \asymp \frac{\hat{\mathbf{g}}_{li\kappa}^H\hat{\mathbf{\Lambda}}_{lk}^{-1}\left(\delta_{llk}\mathbf{I}_M\right)\hat{\mathbf{\Lambda}}_{lk}^{-1}\hat{\mathbf{g}}_{li\kappa}}{\left(p_{\mathrm{r},lk}^{-1} + \delta_{llk}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-1}\right\}\right)^2} \tag{34}$$

where $\hat{\mathbf{g}}_{li\kappa}$ and $\hat{\mathbf{\Lambda}}_{lk}$ from the numerator in (34) are correlated. As a result, we define a Hermitian invertible matrix $\hat{\mathbf{\Lambda}}_{lk\kappa} \triangleq \hat{\mathbf{\Lambda}}_{lk} - p_{\mathrm{r},l\kappa}\hat{\mathbf{g}}_{ll\kappa}\hat{\mathbf{g}}_{ll\kappa}^H$ which is uncorrelated with $\hat{\mathbf{g}}_{li\kappa}$, and reformulate the numerator as

$$\begin{aligned}\delta_{llk}\hat{\mathbf{g}}_{li\kappa}^H\hat{\mathbf{\Lambda}}_{lk}^{-2}\hat{\mathbf{g}}_{li\kappa} &\overset{(a)}{=} \delta_{llk}\hat{\mathbf{g}}_{li\kappa}^H\left(\frac{\hat{\mathbf{\Lambda}}_{lk\kappa}^{-1}}{1 + p_{\mathrm{r},l\kappa}\hat{\mathbf{g}}_{ll\kappa}^H\hat{\mathbf{\Lambda}}_{lk\kappa}^{-1}\hat{\mathbf{g}}_{ll\kappa}}\right)^2\hat{\mathbf{g}}_{li\kappa} \\ &\overset{(b)}{\asymp} \delta_{li\kappa}\delta_{llk}\frac{\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk\kappa}^{-2}\right\}}{\left(1 + p_{\mathrm{r},l\kappa}\delta_{ll\kappa}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk\kappa}^{-1}\right\}\right)^2} \\ &\overset{(c)}{=} \delta_{li\kappa}\delta_{llk}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-2}\right\}\end{aligned} \tag{35}$$

where (a) is obtained according to Lemma 1, (b) follows from Lemma 2, and (c) is due to the fact that $\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk}^{-2}\right\} = \frac{\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk\kappa}^{-2}\right\}}{\left(1+p_{\mathrm{r},l\kappa}\hat{\mathbf{g}}_{ll\kappa}^H\hat{\mathbf{\Lambda}}_{lk\kappa}^{-1}\hat{\mathbf{g}}_{ll\kappa}\right)^2} \asymp \frac{\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk\kappa}^{-2}\right\}}{\left(1+p_{\mathrm{r},l\kappa}\delta_{ll\kappa}\mathrm{tr}\left\{\hat{\mathbf{\Lambda}}_{lk\kappa}^{-1}\right\}\right)^2}$. Substituting (35) into (34), we hence complete the proof.

References

[1] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[2] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[4] K. Guo and G. Ascheid, "Performance analysis of multi-cell MMSE based receivers in MU-MIMO systems with very large antenna arrays," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2013, pp. 3175–3179.

[5] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[6] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[7] H. Ngo, E. Larsson, and T. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

[8] G. Fodor, M. Johansson, and P. Soldati, "Near optimum power control and precoding under fairness constraints in network MIMO systems," *EURASIP International Journal of Digital Multimedia Broadcasting*, vol. 2010, pp. 1–17, Mar. 2010.

[9] Y.-K. Song, D. Kim, and J. Zander, "Pilot power adjustment for saving transmit power in pilot channel assisted DS-CDMA mobile systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 488–493, Feb. 2010.

[10] R. Chen, J. Andrews, R. Heath, and A. Ghosh, "Uplink power control in multi-cell spatial multiplexing wireless systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2700–2711, Jul. 2007.

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[12] P. Li, D. Paul, R. Narasimhan, and J. Cioffi, "On the distribution of SINR for the MMSE MIMO receiver and performance analysis," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 271–286, Jan. 2006.