Massive MIMO Multicasting in Noncooperative Cellular Networks

Zhengzheng Xiang, Meixia Tao, *Senior Member, IEEE*, and Xiaodong Wang, *Fellow, IEEE*

Abstract—We study physical layer multicasting in cellular networks where each base station (BS) is equipped with a very large number of antennas and transmits a common message using a single beamformer to multiple mobile users. The messages sent by different BSs are independent and the BSs do not cooperate. We first show that when each BS knows the perfect channel state information (CSI) of its own served users, the asymptotically optimal beamformer at each BS is a linear combination of the channel vectors of its multicast users. Moreover, the optimal and explicit combining coefficients are obtained. Then we consider the imperfect CSI scenario where the CSI is obtained through uplink channel estimation in time-division duplex systems. We propose a new pilot scheme that estimates the composite channel which is a linear combination of the individual channels of multicast users in each cell. This scheme is able to completely eliminate pilot contamination. The pilot power control for optimizing the multicast beamformer at each BS is also derived. Numerical results show that the asymptotic performance of the proposed scheme is close to the ideal case with perfect CSI. Simulation also verifies the effectiveness of the proposed scheme with finite number of antennas at each BS.

Index Terms—Physical layer multicasting, massive MIMO, asymptotic orthogonality, pilot contamination.

I. INTRODUCTION

The array of emerging wireless data services, such as media streaming, cell broadcasting, and mobile TV have been propelling the development of new wireless communication technologies to make efficient use of the precious spectrum resources. One typical scenario is the delivery of common information, such as headline news, financial data, or hot video, to multiple mobile users. Enabling such service is the wireless multicasting technique that transmits the same information to a group of users simultaneously. Wireless multicasting has been included in the Third Generation Partnership Project (3GPP) LTE standards known as evolved Multimedia Broadcast Multicast Service (eMBMS) [1]. Recently, mobile operators Verizon and AT&T announced their plans to build LTE-multicast/broadcast networks for more efficient content delivery to a large number of users [2], [3].

By exploiting the channel state information (CSI) at the transmitter equipped with multiple antennas, physical layer multicasting in the form of beamforming is a promising way

to improve the wireless multicast throughput. The multicast beamforming was first considered in [4]–[6] for a single group of users. The similar problem of multicasting to multiple cochannel groups was then studied in [7], [8]. In [9], the coordinated multicast beamforming in multicell networks was investigated. The authors in [10] studied the multicast beamforming and resource allocation for orthogonal frequency division multiplexing (OFDM) systems. In [11], the multicast precoding was considered for a single-group network. Different from the traditional unicast beamforming problem in multi-user or multi-cell systems, the multicast beamforming problem is generally NP-hard [4] and hence more challenging. A key tool for finding near-optimal solutions in [4], [7]–[10] is semidefinite relaxation (SDR). Other approaches such as channel orthogonalization-based beamforming, stochastic beamforming and space-time coding-based beamforming were considered in [12]–[17].

Recently, the massive, or large-scale multiple-input multiple output (MIMO) technology has emerged as a promising and efficient approach to mitigating the severe cochannel interference caused by the ever-increasing mobile data traffic [18]. It was first shown in [19] that by applying a large number of antennas at each base station (BS) in a noncooperative cellular network, the intracell interference can be substantially reduced without sophisticated beamformer design. Moreover the transmit power at the BS can be made arbitrarily small. Other related works on the massive MIMO technology can be found, for example, in [20]–[23]. Massive MIMO provides a new dimension that can be exploited to further increase the throughput of wireless networks and has become an important candidate of key technologies for the future 5G wireless communication systems [24]–[26].

The goal of this paper is to investigate massive MIMO in the context of multicast transmission in cellular networks. In urban areas, where wireless multicasting finds applications, there are typically many high buildings with mobile users located at different floors. For such scenario, BSs with massive MIMO can efficiently shape the transmitted signals and significantly increase the quality of service compared with traditional BSs [27]. On the other hand, coordinated transmission among multiple cells can generally provide better performance [9], [28]–[30]. However, for massive MIMO systems, due to the large amount of CSI information associated with the large number of BS antennas, coordination among different cells will induce a huge amount of information exchange overhead, which is impractical for systems with limited-capacity backhaul links. Moreover, as shown in [19], [31], for massive MIMO networks, coordinated transmission becomes less necessary

compared with conventional systems since simple signal processing techniques can efficiently mitigate interference. Thus, in this paper, we will focus on noncooperative transmission in multicell multicast networks with massive MIMO. Note that the aforementioned works on MIMO multicast are all for traditional finite-size MIMO and they mainly used numerical optimization tools to obtain the beamformers except [12], [13].

One challenge associated with massive MIMO is the acquisition of CSI at the transmitter, especially for the downlink transmission. Currently, the common assumption is the TDD protocol where the BS estimates the uplink channels and obtains the downlink CSI by exploiting the channel reciprocity. However, this approach suffers from the so-called pilot contamination problem in the multicell scenario since the users in other cells may use the same (or nonorthogonal) pilot sequences causing non-vanishing interference to the channel estimate at the BS, and the performance of the network will be severely degraded. A few initial attempts have been made very recently to tackle this problem. For example, in [31], a time-shifted scheme was proposed to suppress the effects of pilot contamination. In [32], the coordinated channel estimation for massive MIMO networks was proposed by exploiting the second-order statistical information about the user channels. These works, however, are not specifically designed for multicast transmission.

The main contributions of this paper are as follows. First, we consider the multicast beamformer design with perfect CSI. We show that the asymptotically optimal beamforming vector for each BS is a linear combination of the channel vectors of the multicast users served in its cell. The optimal combining coefficients are obtained and they explicitly depend on the large-scale attenuation of the channels.

Second, we consider the imperfect CSI scenario using the conventional channel estimation method, and characterize the impact of pilot contamination on the signal-to-interference-plus-noise ratio (SINR) of each user.

Third and most importantly, we propose a novel pilot scheme for multicell multicasting to tackle the pilot contamination problem. The key idea is to assign one pilot sequence to one cell which is shared by all the multicast users served in the cell, and orthogonal pilot sequences are assigned to different cells. In this way, each BS estimates a composite channel which is a linear combination of the individual channels of its served multicast users. Such a composite channel has the same structure as the asymptotically optimal multicast beamformer and hence can serve as the beamfomer. By applying pilot power control among the users in each cell, we can optimize the composite channel and hence the multicast beamformer at each BS. The closed-form pilot power control rule is obtained and the SINR gaps between the perfect CSI case and the proposed pilot scheme with power control are analytically characterized. Compared with the traditional pilot scheme which has an SINR ceilling due to pilot contamination, our proposed pilot scheme completely eliminates pilot contamination and has a constant gap to the perfect CSI case. This gap is quite small and is independent of the BS transmitting power. Besides, we consider the synchronization issue for the pilot sequence transmission and analyze the effect of asynchrony on the performance of our proposed scheme.

The remainder of the paper is organized as follows. In Section II, the system model for multicell multicasting with massive MIMO is presented. Section III considers the beamformer design with perfect CSI. In Section IV, we consider the imperfect CSI scenario and analyze the negative effects of the pilot contamination. In Section V, a novel pilot scheme for multicast in multicell is proposed which completely eliminates pilot contamination and the optimal pilot power control is derived. Section VI provides numerical and simulation results. Concluding remarks are made in Section VII.

Notations: Boldface uppercase letters denote matrices and boldface lowercase letters denote vectors. $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Z}^+$ denote the sets of real numbers, complex numbers, and positive integers, respectively. $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ and $\mathrm{Tr}\{\cdot\}$ are the transpose, conjugate, Hermitian transpose and trace operators, respectively. $\mathbb{E}(\cdot)$ is the expectation operator. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ means $\mathbf{x}$ is a circular symmetric complex Gaussian random vectors whose mean vector is $\mathbf{0}$ and covariance matrix is $\mathbf{I}$. $f(y) \sim o(y)$ means that $\lim\limits_{y \to \infty} \frac{f(y)}{y} = 0$.

II. System Model

We consider the downlink transmission of a multicast cellular network comprising N cells and K mobile users per cell as shown in Fig. 1. The BS in each cell is equipped with M antennas where M is a very large number under massive MIMO, and every mobile user has a single antenna. The channel between each BS and each user is assumed to be flat and quasi-static fading. Let $\mathbf{g}_{i,j,k}^H$ denote the $1 \times M$ channel vector from the BS in the ith cell to the kth user in the jth cell. It can be written as

$$\mathbf{g}_{i,j,k}^H = \sqrt{\beta_{i,j,k}} \mathbf{h}_{i,j,k}^H, \qquad (1)$$

where $\{\beta_{i,j,k} \in \mathbb{R}\}$ are the large-scale channel attenuations which change slowly and can be tracked easily, and $\{\mathbf{h}_{i,j,k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)\}$ are the small-scale fading coefficients and modeled as independent and identically distributed (i.i.d.) random vectors. Let s_i denote the multicast information symbol for all the K users in the ith cell with $\mathbb{E}\left[|s_i|^2\right] = 1$ and $\mathbf{w}_i$ be the corresponding $M \times 1$ unit-norm multicast beamforming vector. The transmitted signal from BS i is then given by

$$\mathbf{x}_i = \sqrt{p_i} \mathbf{w}_i s_i, \qquad (2)$$

where p_i is the transmit power by BS i. We assume that all the BSs and users are perfectly synchronized in time and use the same spectrum band. The discrete-time baseband signal received by the kth user in the ith cell is given by, for $k = 1, ..., K$ and $i = 1, ..., N$,

$$y_{i,k} = \sum_{j=1}^{N} \mathbf{g}_{j,i,k}^H \mathbf{x}_j + z_{i,k}$$

$$= \sqrt{p_i} \mathbf{g}_{i,i,k}^H \mathbf{w}_i s_i + \sum_{j \neq i} \sqrt{p_j} \mathbf{g}_{j,i,k}^H \mathbf{w}_j s_j + z_{i,k} \qquad (3)$$

where $z_{i,k} \sim \mathcal{CN}(0, \sigma^2)$ is the additive noise. In (3), the second term on the right-hand side (RHS) is the interference
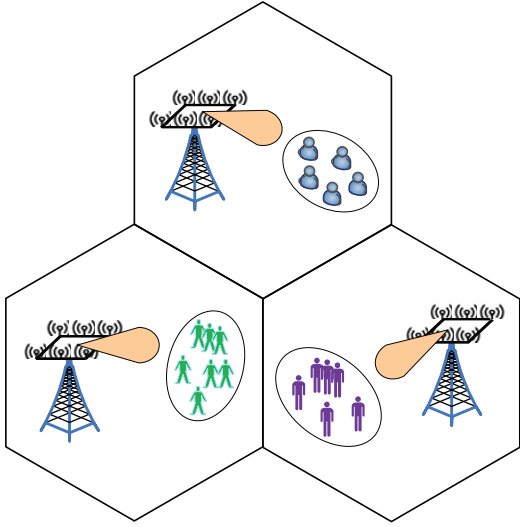
Fig. 1. Multicell multicast network with massive MIMO.

from the BSs in other cells. Based on the received signal model in (3), the performance of each user can be characterized by the output SINR, defined as

$$\text{SINR}_{i,k} = \frac{p_i|\mathbf{g}_{i,i,k}^H\mathbf{w}_i|^2}{\sum\limits_{j\neq i}^N p_j|\mathbf{g}_{j,i,k}^H\mathbf{w}_j|^2 + \sigma^2}. \tag{4}$$

Since a common message is transmitted to the K users in each cell, we characterize the performance of each cell by the user with the minimum SINR.

Before moving on to the next section, we review two basic results on random vectors [33] that will be useful later on. Let $\mathbf{x} = [x_1, ..., x_n]^T$ and $\mathbf{y} = [y_1, ..., y_n]^T$ be two mutually independent $n \times 1$ vectors whose elements are i.i.d zero-mean random variables with variances σ_x^2 and σ_y^2, respectively. Then from the law of large numbers, we have

$$\lim_{n\to\infty} \frac{\mathbf{x}^H\mathbf{x}}{n} \xrightarrow{\text{a.s.}} \sigma_x^2, \text{ and } \lim_{n\to\infty} \frac{\mathbf{x}^H\mathbf{y}}{n} \xrightarrow{\text{a.s.}} 0. \tag{5}$$

Here, $\xrightarrow{\text{a.s.}}$ denotes the almost sure convergence.

III. ASYMPTOTIC ANALYSIS WITH PERFECT CSI

In this section, we provide the asymptotic analysis for the multicell multicast network with massive MIMO when each BS knows the perfect CSI of its served users. Since for multicast transmission, the performance of the worst user in each cell is the bottleneck, our goal of the beamformer design for each BS is to maximize the minimum SINR of the K served users. This is formulated as the following optimization problem

$$\mathcal{P}: \max_{\{\mathbf{w}_i\}} \min_{\forall k} \frac{p_i|\mathbf{g}_{i,i,k}^H\mathbf{w}_i|^2}{\sum\limits_{j\neq i}^N p_j|\mathbf{g}_{j,i,k}^H\mathbf{w}_j|^2 + \sigma^2} \tag{6}$$
$$\text{s.t.} \quad \|\mathbf{w}_i\| = 1.$$

A. Asymptotically Optimal Beamformer Structure

The optimization problem $\mathcal{P}$ for each BS is equivalent to the conventional single-cell multicast beamforming problem in [4]. Due to the NP-hardness, its optimal solution cannot be obtained efficiently in general. In our case, since each BS has a very large number of antennas, the different channels become asymptotically orthogonal. Therefore, we are able to obtain the asymptotically optimal structure of the beamforming vector in the following theorem.

Theorem 1: The asymptotically optimal beamformer for each BS when $M \to \infty$ is a linear combination of the channels between the BS and its served users, i.e.,

$$\mathbf{w}_i = \sum_{k=1}^K \xi_{i,k}\mathbf{g}_{i,i,k}, \ \forall i \in \{1, ..., N\}, \tag{7}$$

where $\{\xi_{i,k}\}$ are the linear combination coefficients.

Proof: The proof is based on the asymptotic orthogonality of the channel vectors when M is large. We first show that with the beamformers in (7), the intercell interference that each user suffers from diminishes as $M \to \infty$:

$$y_{i,k} = \sqrt{p_i}\mathbf{g}_{i,i,k}^H\mathbf{w}_i s_i + \sum_{j\neq i}^N \sqrt{p_j}\mathbf{g}_{j,i,k}^H\mathbf{w}_j s_j + z_{i,k}$$
$$= \sqrt{p_i}\mathbf{g}_{i,i,k}^H \left(\sum_{k'=1}^K \xi_{i,k'}\mathbf{g}_{i,i,k'}\right) s_i$$
$$+ \sum_{j\neq i}^N \sqrt{p_j}\mathbf{g}_{j,i,k}^H \left(\sum_{k'=1}^K \xi_{j,k'}\mathbf{g}_{j,j,k'}\right) s_j + z_{i,k}$$
$$= \underbrace{\sqrt{p_i}\xi_{i,k}\mathbf{g}_{i,i,k}^H\mathbf{g}_{i,i,k}s_i}_{\Gamma} + o(\Gamma) + z_{i,k}. \tag{8}$$

Next if it does not have this structure, the beamformer of BS i can be written as

$$\mathbf{w}_i' = \sum_{k=1}^K \xi_{i,k}'\mathbf{g}_{i,i,k} + \sum_{t=1}^{M-K} \zeta_{i,t}\mathbf{u}_{i,t}, \tag{9}$$

where $\{\mathbf{u}_{i,t}\}_{t=1}^{M-K}$ is an orthonormal basis for the orthogonal complement of the space spanned by $\{\mathbf{g}_{i,i,k}\}_{k=1}^K$.

From (9) it is seen that the component $\sum\limits_{t=1}^{M-K} \zeta_{i,t}\mathbf{u}_{i,t}$ makes no contribution to the useful signals of the users in the ith cell but will cause extra interference to the users in the other cells. Hence, we can always reconstruct another beamformer for this BS without changing the transmit power as

$$\mathbf{w}_i'' = \sum_{k=1}^K \eta\xi_{i,k}'\mathbf{g}_{i,i,k}. \tag{10}$$

Here,

$$\eta = \sqrt{\left(\sum_{k=1}^K |\xi_{i,k}'|^2 + \sum_{t=1}^{M-K} |\zeta_{i,t}|^2\right) \Big/ \left(\sum_{k'=1}^K |\xi_{i,k}'|^2\right)} > 1.$$

Given that the beamformers of other BSs keep unchanged, the received power of the users in the ith cell under $\mathbf{w}_i''$ is larger than that under $\mathbf{w}_i'$, and the interference caused by $\mathbf{w}_i''$ to the users in other cells still vanishes. This completes the proof. ∎

From Theorem 1, we can see that the massive MIMO beamformer for the multicell multicast network has a rather simple structure. In this paper, we will exploit this structure to obtain the optimal multicast beamformers under both perfect and estimated CSI.

Next, we will derive the asymptotic SINR achieved at the user receiver for the considered network. We first normalize the beamforming vector for each BS i as

$$\mathbf{w}_i = \frac{\sum_{k=1}^{K} \xi_{i,k}\mathbf{g}_{i,i,k}}{\alpha_i\sqrt{M}} \tag{11}$$

where the normalization factor is

$$\alpha_i = \frac{\left\| \sum_{k=1}^{K} \xi_{i,k}\mathbf{g}_{i,i,k} \right\|}{\sqrt{M}}. \tag{12}$$

According to (5), we derive the asymptotic behavior of α_i^2 as below

$$
\begin{aligned}
\lim_{M\to\infty} \alpha_i^2 &= \lim_{M\to\infty} \frac{\left(\sum_{k=1}^{K} \xi_{i,k}\mathbf{g}_{i,i,k}\right)^H \left(\sum_{k=1}^{K} \xi_{i,k}\mathbf{g}_{i,i,k}\right)}{M} \\
&= \lim_{M\to\infty} \frac{\sum_{k=1}^{K}\sum_{k'=1}^{K} \left|\xi_{i,k}^*\xi_{i,k'}\right|\sqrt{\beta_{i,i,k}\beta_{i,i,k'}}\mathbf{h}_{i,i,k}^H\mathbf{h}_{i,i,k'}}{M} \\
&= \sum_{k=1}^{K} |\xi_{i,k}|^2 \beta_{i,i,k}. \tag{13}
\end{aligned}
$$

Assume that the transmit power of each BS p_i is scaled with M according to $p_i = \frac{E_i}{M}$ where E_i is fixed. Then, the asymptotic SINR of user k, for $k = 1, .., K$ at cell i is derived as

$$
\begin{aligned}
&\lim_{M\to\infty} \mathrm{SINR}_{i,k} = \lim_{M\to\infty} \frac{p_i|\mathbf{g}_{i,i,k}^H\mathbf{w}_i|^2}{\sum_{j\neq i}^{N} p_j|\mathbf{g}_{j,i,k}^H\mathbf{w}_j|^2 + \sigma^2} \\
&= \lim_{M\to\infty} \frac{\frac{E_i\beta_{i,i,k}}{\alpha_i^2}\left|\frac{\sum_{k'=1}^{K} \xi_{i,k'}\sqrt{\beta_{i,i,k'}}\mathbf{h}_{i,i,k}^H\mathbf{h}_{i,i,k'}}{M}\right|^2}{\sum_{j\neq i}^{N} \frac{E_j\beta_{j,i,k}}{\alpha_j^2}\left|\frac{\sum_{k'=1}^{K} \xi_{j,k'}\sqrt{\beta_{j,j,k'}}\mathbf{h}_{j,i,k}^H\mathbf{h}_{j,j,k'}}{M}\right|^2 + \sigma^2}. \tag{14}
\end{aligned}
$$

Applying (5) and plugging (13) into (14), we obtain

$$\lim_{M\to\infty} \mathrm{SINR}_{i,k} = \frac{\lambda_{i,k}E_i\beta_{i,i,k}}{\sigma^2}, \tag{15}$$

where

$$\lambda_{i,k} = \frac{|\xi_{i,k}|^2 \beta_{i,i,k}}{\sum_{k'=1}^{K} |\xi_{i,k'}|^2 \beta_{i,i,k'}}. \tag{16}$$

From (15) it is seen that with a large M, each BS just needs the power $\frac{E_i}{M}$ to achieve the same performance as that of a single-input single-output (SISO) system with transmit power $\lambda_{i,k}E_i$, without any interference and the fast-fading effect, which coincides with [20]. For each antenna, its transmit

power will be $\frac{E_i}{M^2}$. Thus, given the SINR target, we can reduce the power consumption at the BS by increasing the BS antennas. In the rest of this paper, we will assume that $p_i = \frac{E_i}{M}, \forall i$.

B. Optimal Beamforming Parameters

In this subsection, we will derive the optimal linear combination coefficients for the asymptotically optimal beamformer in Theorem 1 based on the max-min criterion in (6). Recall that the asymptotic SINR is directly related to $\{\lambda_{i,k}\}$ instead of $\{\xi_{i,k}\}$ as defined in (16). Moreover, an important observation is that

$$\sum_{k=1}^{K} \lambda_{i,k} = 1. \tag{17}$$

Therefore, optimizing the coefficients $\{\xi_{i,k}\}$ is equivalent to finding the optimal $\{\lambda_{i,k}\}$. This problem can be formulated as

$$\mathcal{Q}: \max_{\{\lambda_{i,k}\}} \min_{\forall k} \frac{\lambda_{i,k}E_i\beta_{i,i,k}}{\sigma^2} \tag{18}$$
$$\text{s.t.} \quad (17).$$

Lemma 2: The optimal solution of problem $\mathcal{Q}$ is obtained when all the K users have the same SINR, i.e.,

$$\frac{\lambda_{i,1}E_i\beta_{i,i,1}}{\sigma^2} = ... = \frac{\lambda_{i,K}E_i\beta_{i,i,K}}{\sigma^2} \tag{19}$$

and the optimal $\{\lambda_{i,k}^*\}$ are

$$\lambda_{i,k}^* = \frac{1}{\sum_{k'=1}^{K} \frac{\beta_{i,i,k}}{\beta_{i,i,k'}}}, \ \forall i, \ \forall k. \tag{20}$$

Proof: Please refer to Appendix A. ∎

According to Lemma 2 and (16) and after some manipulations, we can obtain the following theorem:

Theorem 3: The asymptotically optimal multicast beamforming vector is given by

$$\mathbf{w}_i^* = \mu_i \sum_{k=1}^{K} \frac{\mathbf{g}_{i,i,k}}{\beta_{i,i,k}}, \ \forall i \tag{21}$$

where the common factor μ_i is

$$\mu_i = \frac{1}{\sqrt{M \sum_{k=1}^{K} \frac{1}{\beta_{i,i,k}}}}, \tag{22}$$

and the resulting asymptotic SINR of each user is given by

$$\lim_{M\to\infty} \mathrm{SINR}_{i,k} = \frac{E_i}{\sigma^2 \sum_{k'=1}^{K} \frac{1}{\beta_{i,i,k'}}}, \ \forall i, k. \tag{23}$$

From Theorem 3 it is seen that the optimal combination coefficients in the beamformer depend only on the large-scale channel parameters $\{\beta_{i,i,k}\}$, and the user with the larger large $\beta_{i,i,k}$ will have lower weight in the beamformer. Also, when the user number K increases, the performance will decrease due to the multicast feature of the network.

Substituting the asymptotically optimal $\mathbf{w}_i^*$ (21) into the received signal model in (3), we obtain that

$$
\begin{aligned}
y_{i,k} =& \sqrt{\frac{E_i}{M}} \sum_{k'=1}^{K} \frac{\mu_i \mathbf{g}_{i,i,k}^H \mathbf{g}_{i,i,k'}}{\beta_{i,i,k'}} s_i \\
&+ \sum_{j \neq i}^{N} \sqrt{\frac{E_j}{M}} \sum_{k'=1}^{K} \frac{\mu_j \mathbf{g}_{j,i,k}^H \mathbf{g}_{j,j,k'}}{\beta_{j,j,k'}} s_j + z_{i,k} \\
=& \underbrace{\sqrt{\frac{E_i \mu_i^2}{M \beta_{i,i,k}}} \mathbf{g}_{i,i,k}^H \mathbf{g}_{i,i,k} s_i}_{\Gamma} + o(\Gamma) + z_{i,k}.
\end{aligned} \tag{24}
$$

For large M, the effective channel coefficient for the kth user in the ith cell converges to:

$$
\sqrt{\frac{E_i \mu_i^2}{M \beta_{i,i,k}}} \mathbf{g}_{i,i,k}^H \mathbf{g}_{i,i,k} \rightarrow \sqrt{\frac{E_i}{\sum\limits_{k'=1}^{K} \frac{\beta_{i,i,k}^2}{\beta_{i,i,k'}}}}.
$$

Thus, to decode its desired message, each mobile user does not need to know the small-scale fading channel vector which has a big size and changes rather quickly, but it only needs to learn the effective channel which is a scalar and only contains the slowly-changing large-scale fading parameters. Such feature is preferable for massive MIMO networks since the acquisition of CSI at the user end is highly challenging.

The asymptotic behavior of multicasting was considered in [34]–[36] when the number of transmit antennas goes to infinity. However, the scopes of those works are different from ours. In particular, in [34], [35], the capacity scaling of the multicast network is analyzed when the number of transmit antennas grows large, without considering the asymptotic beamformer structure. In [36], the optimal beamformer is given for the two-user case, which cannot be generalized to the case of an arbitrary number of users. In our work, we focus on multicell massive MIMO and derive the asymptotically optimal beamformer structure for multiuser multicast. In the special case with two users, our beamformer structure coincides with the solution of [36] with the same form of linear combination.

Thus far, we have obtained the asymptotically optimal beamformer for the massive MIMO multicasting when having perfect CSI. In the next section, we will consider the imperfect CSI scenario and investigate the pilot contamination issue for this network.

IV. ANALYSIS WITH IMPERFECT CSI AND PILOT CONTAMINATION

For the downlink transmission of cellular networks, there are two main approaches for the acquisition of CSI at the BS. In frequency-division duplex (FDD) systems, each user first estimates the downlink channel and then feeds back some quantized version to the BS. In TDD systems, the user transmits the pilot signal to the BS. The BS estimates the uplink channel and treats it as the downlink CSI according to channel reciprocity, and no CSI feedback is needed.

In the massive MIMO network, since the size of the channel vector is very large, both the channel estimation and CSI feedback at the user end become very difficult and expensive. However, for channel estimation at the BS end, the time required for pilot transmission is independent of the number of BS antennas and it is much easier for the BS to do channel estimation. Thus, we assume the TDD mode and use uplink pilots in this paper. We also assume that the channel varies slowly enough such that the channel reciprocity holds.

We consider the estimation of individual channel vectors. Specifically, the same set of K orthogonal pilot sequences is used in all the N cells. Every kth user in all the cells uses the same $1 \times \tau$ pilot sequence

$$
\sqrt{\tau}\boldsymbol{\phi}_k = \sqrt{\tau}\left[\phi_{k1}, ..., \phi_{k\tau}\right],
$$

where $\tau \geq K$ and $\boldsymbol{\phi}_k \boldsymbol{\phi}_{k'}^H = 1$ for $k = k'$ and 0 otherwise.

At the pilot transmission phase, all the users simultaneously transmit the pilot signals and the ith BS receives the pilot signal as

$$
\mathbf{Y}_{B_i} = \sum_{l=1}^{N} \sum_{k=1}^{K} \sqrt{p_u \tau}\, \mathbf{g}_{i,l,k} \boldsymbol{\phi}_k + \mathbf{Z}_{B_i}, \tag{25}
$$

where p_u is the average transmit power of the users and $\mathbf{Z}_{B_i} \in \mathbb{C}^{M \times \tau}$ is the additive noise matrix with each element being a symmetric complex Gaussian random variable with mean zero and variance σ_p^2. The BS then estimates the channel vector for the kth user as

$$
\begin{aligned}
\hat{\mathbf{g}}_{i,i,k} &= \mathbf{Y}_{B_i} \boldsymbol{\phi}_k^H \\
&= \sum_{l=1}^{N} \sum_{k'=1}^{K} \sqrt{p_u \tau}\, \mathbf{g}_{i,l,k'} \boldsymbol{\phi}_{k'} \boldsymbol{\phi}_k^H + \mathbf{Z}_{B_i} \boldsymbol{\phi}_k^H \\
&= \sum_{l=1}^{N} \sqrt{p_u \tau}\, \mathbf{g}_{i,l,k} + \mathbf{z}_{B_{ik}}, \tag{26}
\end{aligned}
$$

where $\mathbf{z}_{B_{ik}} = \mathbf{Z}_{B_i} \boldsymbol{\phi}_k^H \sim \mathcal{CN}(\mathbf{0}, \sigma_p^2 \mathbf{I}_M)$. It is seen that the channel estimate at each BS not only contains the desired CSI $\sqrt{p_u \tau}\, \mathbf{g}_{i,i,k}$, but also the undesired CSI $\sum_{l \neq i}^{N} \sqrt{p_u \tau}\, \mathbf{g}_{i,l,k}$. That is, the channel estimation is contaminated by the interference from other cells.

Based on the estimated CSI in (26) and the beamformer structure in Theorem 1, the normalized beamforming vector for BS i can be written as

$$
\mathbf{w}_i = \frac{\sum\limits_{k=1}^{K} \xi_{i,k} \hat{\mathbf{g}}_{i,i,k}}{\gamma_i \sqrt{M}}, \tag{27}
$$

where γ_i is a normalization factor such as that $\|\mathbf{w}_i\| = 1$.

As in Section III, we first derive the asymptotic behavior of γ_i^2 which will be used in the derivation of the asymptotic SINR:

$$
\begin{aligned}
&\lim_{M \to \infty} \gamma_i^2 \\
&= \lim_{M \to \infty} \frac{\left\| \sum\limits_{k=1}^{K} \xi_{i,k} \left(\sum\limits_{l=1}^{N} \sqrt{p_u \beta_{i,l,k} \tau}\, \mathbf{h}_{i,l,k} + \mathbf{z}_{B_{ik}} \right) \right\|^2}{M} \\
&= \sum_{k=1}^{K} \sum_{l=1}^{N} \xi_{i,k}^2 p_u \tau \beta_{i,l,k} + \sigma_p^2 \sum_{k=1}^{K} \xi_{i,k}^2 \triangleq \gamma_{i,\infty}^2. \tag{28}
\end{aligned}
$$

Then, the asymptotic SINR of each user for the imperfect CSI scenario is given by:

$$\lim_{M\to\infty} \text{SINR}_{i,k} = \lim_{M\to\infty} \frac{p_i|\mathbf{g}_{i,i,k}^H\mathbf{w}_i|^2}{\sum\limits_{j\neq i}^{N} p_j|\mathbf{g}_{j,i,k}^H\mathbf{w}_j|^2 + \sigma^2}$$

$$= \lim_{M\to\infty} \frac{p_i\left|\mathbf{g}_{i,i,k}^H\left(\frac{\sum\limits_{k'=1}^{K}\xi_{i,k'}\hat{\mathbf{g}}_{i,i,k'}}{\gamma_i\sqrt{M}}\right)\right|^2}{\sum\limits_{j\neq i}^{N} p_j\left|\mathbf{g}_{j,i,k}^H\left(\frac{\sum\limits_{k'=1}^{K}\xi_{j,k'}\hat{\mathbf{g}}_{j,j,k'}}{\gamma_j\sqrt{M}}\right)\right|^2 + \sigma^2}$$

$$= \frac{\frac{E_i}{\gamma_{i,\infty}^2}\beta_{i,i,k}^2\xi_{i,k}^2 p_u\tau}{\sum\limits_{j\neq i}^{N}\frac{E_j}{\gamma_{j,\infty}^2}\beta_{j,i,k}^2\xi_{j,k}^2 p_u\tau + \sigma^2} \triangleq \mathcal{A}. \qquad (29)$$

In the extreme case where $E_i = E \to \infty, \forall i$, we have

$$\lim_{E\to\infty} \mathcal{A} = \frac{\beta_{i,i,k}^2\xi_{i,k}^2}{\sum\limits_{j\neq i}^{N}\frac{\gamma_{i,\infty}^2}{\gamma_{j,\infty}^2}\beta_{j,i,k}^2\xi_{j,k}^2}.$$

That is, the asymptotic SINR reaches a ceiling and will not increase with the transmit power of BS. This clearly indicates the effect of pilot contamination.

Note that, based on the SINR expression in (29), it is difficult to obtain the asymptotically optimal parameters $\{\xi_{i,k}\}$ like the perfect CSI case for our considered noncooperative cellular network.

V. A CONTAMINATION-FREE PILOT SCHEME

In order to avoid pilot contamination, a simple and straightforward approach is to assign a different and orthogonal pilot sequence to each user in the whole network. However, this will need a very long pilot sequence which becomes prohibitive in practice. In this section, we propose a new pilot scheme for the multicell multicast network with large scale antenna arrays to eliminate the pilot contamination.

A. The New Pilot Scheme

The key idea of our proposed pilot scheme is that we apply a set of N orthogonal pilot sequences for the whole network and the K users in each cell share the same pilot sequence. Thus, each BS will obtain an estimated composite channel of the served K users.

For cell i, the K users transmit the same $1\times\omega$ pilot sequence

$$\sqrt{\omega}\boldsymbol{\psi}_i = \sqrt{\omega}\left[\psi_{i1}, ..., \psi_{i\omega}\right]$$

where $\omega \geq N$ and $\boldsymbol{\psi}_i\boldsymbol{\psi}_j^H = 1$ for $i = j$ and 0 otherwise. The received signal at BS i is then written as

$$\mathbf{Y}_{B_i} = \sum_{l=1}^{N}\sum_{k=1}^{K}\sqrt{p_u\omega}\mathbf{g}_{i,l,k}\boldsymbol{\psi}_l + \mathbf{Z}_{B_i}. \qquad (30)$$

BS i then estimates the composite channel vector for the users in cell i as

$$\hat{\mathbf{g}}_{i,c} = \mathbf{Y}_{B_i}\boldsymbol{\psi}_i^H$$

$$= \sum_{l=1}^{N}\sum_{k=1}^{K}\sqrt{p_u\omega}\mathbf{g}_{i,l,k}\boldsymbol{\psi}_l\boldsymbol{\psi}_i^H + \mathbf{Z}_{B_i}\boldsymbol{\psi}_i^H$$

$$= \sum_{k=1}^{K}\sqrt{p_u\omega}\mathbf{g}_{i,i,k} + \mathbf{z}_{B_i}. \qquad (31)$$

Clearly, the estimate of the composite channel vector is a linear combination of the channel vectors of the K users from the desired cell (plus the additive noise), and does not contain the channels of the users from other cells. According to Theorem 1, the asymptotically optimal beamfomer for each BS is also a linear combination of the channels between the BS and its served users. Therefore, the BS can design the beamforming vector as

$$\mathbf{w}_i = \frac{\hat{\mathbf{g}}_{i,c}}{\|\hat{\mathbf{g}}_{i,c}\|}$$

$$= \sum_{k=1}^{K}\frac{\sqrt{p_u\omega}}{\|\hat{\mathbf{g}}_{i,c}\|}\mathbf{g}_{i,i,k} + \frac{\mathbf{z}_{B_i}}{\|\hat{\mathbf{g}}_{i,c}\|}. \qquad (32)$$

As a result, there will be no intercell interference for the composite channel estimation scheme.

B. Pilot Power Control

From the previous subsection, it is seen that the proposed pilot scheme completely eliminates the pilot contamination. However, with the resulting composite channel estimation, the BS no longer has the flexibility to optimize the beamformer coefficients.

In this subsection, we propose power control among the K users of each cell in the pilot transmission phase to optimize the coefficients in each BS's beamformer, and to improve the network performance.

Let $p_{i,k}$ denote the transmit power of user k in BS i in the pilot transmission phase and be subject to the peak power constraint $p_{i,k} < p_u$. By replacing p_u with $p_{i,k}$, the composite channel vector in (31) can be rewritten as

$$\hat{\mathbf{g}}_{i,c} = \sum_{k=1}^{K}\sqrt{\omega\beta_{i,i,k}p_{i,k}}\mathbf{h}_{i,i,k} + \mathbf{z}_{B_i}. \qquad (33)$$

Then the normalized beamforming vector is

$$\mathbf{w}_i = \frac{\sum\limits_{k=1}^{K}\sqrt{\omega\beta_{i,i,k}p_{i,k}}\mathbf{h}_{i,i,k} + \mathbf{z}_{B_i}}{\mu_i\sqrt{M}}, \qquad (34)$$

where μ_i is a normalization factor and its limiting value is given by

$$\lim_{M\to\infty}\mu_i^2 = \omega\sum_{k=1}^{K}\beta_{i,i,k}p_{i,k} + \sigma_p^2 \triangleq \mu_{i,\infty}^2. \qquad (35)$$

With the above beamformer, we derive the asymptotic SINR of each user when $M \to \infty$ for the proposed composite channel

estimation scheme as follows

$$\lim_{M\to\infty} \mathrm{SINR}_{i,k}$$

$$= \lim_{M\to\infty} \frac{p_i\beta_{i,i,k} \left| \mathbf{h}_{i,i,k}^H \frac{\left(\sum\limits_{k'=1}^{K} \sqrt{\omega\beta_{i,i,k'}p_{i,k'}}\mathbf{h}_{i,i,k'}+\mathbf{z}_{B_i} \right)}{\mu_i\sqrt{M}} \right|^2}{\sum\limits_{j\neq i}^{N} p_j\beta_{j,i,k} \left| \mathbf{h}_{j,i,k}^H \frac{\left(\sum\limits_{k'=1}^{K} \sqrt{\omega\beta_{j,j,k'}p_{j,k'}}\mathbf{h}_{j,j,k'}+\mathbf{z}_{B_j} \right)}{\mu_j\sqrt{M}} \right|^2 + \sigma^2}$$

$$= \frac{E_i\beta_{i,i,k}^2\omega p_{i,k}}{\mu_{i,\infty}^2\sigma^2}$$

$$= \frac{E_i}{\sigma^2}\cdot\frac{\beta_{i,i,k}^2 p_{i,k}}{\sum\limits_{k'=1}^{K}\beta_{i,i,k'}p_{i,k'}+\frac{\sigma_p^2}{\omega}}. \tag{36}$$

Based on the max-min criterion, we can formulate the pilot power control problem for each cell as follows:

$$\mathcal{T}: \max_{\{p_{i,k}\}} \min_{\forall k} \frac{\beta_{i,i,k}^2 p_{i,k}}{\sum\limits_{k'=1}^{K}\beta_{i,i,k'}p_{i,k'}+\frac{\sigma_p^2}{\omega}} \tag{37}$$

$$\text{s.t.} \quad p_{i,k}\leq p_u, \ \forall k,$$

where p_u denotes the peak transmit power of each user.

Changing the objective function of $\mathcal{T}$, we can rewrite problem $\mathcal{T}$ as

$$\mathcal{T}: \min_{\{p_{i,k}\}} \max_{\forall k} \frac{\sigma_p^2}{\omega}\beta_{i,i,k}^{-2}p_{i,k}^{-1} + \sum_{k'=1}^{K}\beta_{i,i,k}^{-2}\beta_{i,i,k'}p_{i,k}^{-1}p_{i,k'} \tag{38}$$

$$\text{s.t.} \quad p_{i,k}\leq p_u, \ \forall k\in\{1,...,K\}.$$

To further simplify the objective function of the above optimization problem, we introduce a slack variable t and reformulate the problem as follows

$$\mathcal{O}: \min_{\{p_{i,k},t\}} t \tag{39}$$

$$\text{s.t.} \quad \frac{\sigma_p^2}{\omega}\beta_{i,i,k}^{-2}p_{i,k}^{-1}t^{-1} + \sum_{k'=1}^{K}\beta_{i,i,k}^{-2}\beta_{i,i,k'}p_{i,k}^{-1}p_{i,k'}t^{-1} \leq 1$$

$$p_{i,k}\leq p_u, \ \forall k\in\{1,...,K\}.$$

The above problem is a geometric programming problem. It can be transformed to a convex problem and solved optimally using the software package [37]. By exploiting the specific structure of the problem, we further show that a closed-form expression for the optimal solution can be obtained.

Lemma 4: At the optimal solution of problem (37), the pilot power of user $k^\star$ with the minimum channel gain $\beta_{i,i,k^\star}$ is p_u, i.e.,

$$p_{i,k^\star} = p_u. \tag{40}$$

Proof: Please refer to Appendix B. ∎

The above lemma tells us that the user with the minimum channel gain should transmit the pilot sequence with full power. Based on this result, we can derive the optimal closed-form solution of problem (37) in the following theorem

Theorem 5: The optimal solution of problem (37) is obtained when the achieved SINRs of all the users are the same, and the optimal pilot power for each user is

$$p_{i,k}^* = \frac{\beta_{i,i,k^\star}^2 p_u}{\beta_{i,i,k}^2}, \ \forall k, \tag{41}$$

where $\beta_{i,i,k^\star}$ is the minimum channel gain among the K users.

Proof: We first prove that the pilot power of the user $j^\star(\neq k^\star)$ with the second minimum fading ratio $\beta_{i,i,j^\star}$ is

$$p_{i,j^\star} = \frac{\beta_{i,i,k^\star}^2 p_u}{\beta_{i,i,j^\star}^2}. \tag{42}$$

by contradiction.

If the optimal solution satisfies $p_{i,j^\star} > \frac{\beta_{i,i,k^\star}^2 p_u}{\beta_{i,i,j^\star}^2}$, the SINR of user $j^\star$ will be larger than that of user $k^\star$ since the expression of each user's SINR has a common denominator $\sum\limits_{k'=1}^{K}\beta_{i,i,k'}p_{i,k'}+\frac{\sigma_p^2}{\omega}$. Then we can set the pilot power of user $j^\star$ as $p_{i,j^\star} = \frac{\beta_{i,i,k^\star}^2 p_u}{\beta_{i,i,j^\star}^2}$. Thus, it can be easily seen that the objective value of problem (37) will increase since the common denominator $\sum\limits_{k'=1}^{K}\beta_{i,i,k'}p_{i,k'}+\frac{\sigma_p^2}{\omega}$ decreases, which contradicts with the optimality.

Else if $p_{i,j^\star} < \frac{\beta_{i,i,k^\star}^2 p_u}{\beta_{i,i,j^\star}^2}$ holds in the optimal solution, we will have that

$$p_{i,k} < \frac{\beta_{i,i,k^\star}^2 p_u}{\beta_{i,i,j^\star}^2}, \forall k\in\mathcal{S}\backslash\{k^\star\} \tag{43}$$

where $\mathcal{S} = \{1,2,...,K\}$. The reason is similar to the proof of Lemma 4 and we omit the details here. Under such condition, the user with the minimum SINR must lie in the set $\mathcal{S}\backslash\{k^\star\}$ since the SINR of user $j^\star$ is smaller than that of user $k^\star$.

Now, based on (43), we can always scale up the pilot power of all the $K-1$ users with $\eta > 1$:

$$p_{i,k}' = \eta p_{i,k}, \forall k\in\mathcal{S}\backslash\{k^\star\} \tag{44}$$

while still maintaining that $p_{i,k}' < \frac{\beta_{i,i,k^\star}^2 p_u}{\beta_{i,i,j^\star}^2}, \forall k\in\mathcal{S}\backslash\{k^\star\}$. After this scaling, although the SINR of user $k^\star$ decreases, the SINR of user $j^\star$ is still smaller than that of user $k^\star$ since $p_{i,j^\star}' < \frac{\beta_{i,i,k^\star}^2 p_u}{\beta_{i,i,j^\star}^2}$, which means that the minimum SINR still lies in the set $\mathcal{S}\backslash\{k^\star\}$. On the other hand, it can be seen that the SINR of each user in $\mathcal{S}\backslash\{k^\star\}$ increases. Thus, the minimum SINR increases which contradicts with the optimality. Finally, we must have (42).

Using similar argument, we can continue the proof for the remaining users and the details are omitted. Thus, we complete the proof of this theorem. ∎

From the above theorem, we can see that the proposed pilot transmit power control can be performed by each individual user in a distributed manner. The only information that needs to be exchanged among the users is the minimum channel gain $\beta_{i,i,k^\star}$.

Substituting the optimal pilot power control (41) into (34), we can obtain the beamforming vector for the BS as follows:

$$\mathbf{w}_i = \sum_{k=1}^{K} \frac{\mathbf{g}_{i,i,k} + \frac{\mathbf{z}_{B_i}}{\sqrt{\omega p_{i,k}}}}{\sqrt{\sum_{k'=1}^{K} \beta_{i,i,k'} \frac{p_{i,k'}}{p_{i,k}} + \frac{\sigma_p^2}{\omega p_{i,k}}} \sqrt{M}}$$

$$= \sum_{k=1}^{K} \frac{\mathbf{g}_{i,i,k} + \frac{\mathbf{z}_{B_i}}{\sqrt{\omega p_{i,k}}}}{\sqrt{\sum_{k'=1}^{K} \frac{\beta_{i,i,k}^2}{\beta_{i,i,k'}} + \frac{\sigma_p^2}{\omega p_{i,k}}} \sqrt{M}}. \tag{45}$$

By comparing (45) with (21), one can see that the beamforming vector obtained by the optimal pilot power control is almost the same as the optimal beamforming vector of the perfect CSI case except for the noise terms $\frac{\mathbf{z}_{B_i}}{\sqrt{\omega p_{i,k}}}$ and $\frac{\sigma_p^2}{\omega p_{i,k}}$. Thus, effectively, the optimization of pilot power at each user side plays the similar role as the optimization of the linear combination coefficients at the BS side in the perfect CSI case.

In the following, we characterize the SINR gap between the proposed pilot scheme and the perfect CSI case. Plugging (41) into (36), we can obtain the asymptotic SINR for our proposed pilot scheme with optimal power control as

$$\lim_{M \to \infty} \text{SINR}'_{i,k} = \frac{E_i}{\sigma^2} \cdot \frac{1}{\sum_{k'=1}^{K} \frac{1}{\beta_{i,i,k'}} + \frac{\sigma_p^2}{\omega \beta_{i,i,k^\star}^2 p_u}}. \tag{46}$$

Comparing (23) and (46), we then obtain the SINR gap

$$\triangle\text{SINR} = 10 \log_{10} \left(\frac{E_i}{\sigma^2 \sum_{k'=1}^{K} \frac{1}{\beta_{i,i,k'}}} \right)$$

$$- 10 \log_{10} \left(\frac{E_i}{\sigma^2} \cdot \frac{1}{\sum_{k'=1}^{K} \frac{1}{\beta_{i,i,k'}} + \frac{\sigma_p^2}{\omega \beta_{i,i,k^\star}^2 p_u}} \right)$$

$$= 10 \log_{10} \left(1 + \frac{\sigma_p^2}{\omega p_u \beta_{i,i,k^\star}^2 \sum_{k'=1}^{K} \frac{1}{\beta_{i,i,k'}}} \right). \tag{47}$$

This gap decreases as the peak pilot power p_u increases, and is independent of the transmit power of the BS.

Remark 1: In the proposed pilot scheme for composite channel estimation, the length of the pilot should not be less than N. Recall that if the individual channel estimation scheme is applied, the pilot length will be at least K. Similar to the user number K, the cell number N (adjacent cells) will not be too large and is proper for practical implementations.

Remark 2: In the optimal pilot power control, the knowledge of large-scale channel attenuation factors of all users in the same cell, i.e. $\{\beta_{i,i,k}\}$ should be known. In practice, the channel attenuation changes very slowly compared with the small-scale fading, and hence can be estimated in advance using off-line approaches. On the other hand, if this knowledge is not available, we simply let each user transmit at its peak power, but the performance will decrease. However, from the simulation results, we will see that the proposed scheme without pilot power control still outperforms the individual channel estimation scheme.

C. Discussion On Synchronization

During the uplink pilot transmission phase, the users are not necessarily perfectly synchronized. In addition, the pilot signal coming from different users may arrive at the BS with different propagation delays. In this subsection, we discuss the synchronization and delay issues of the proposed composite channel estimation.

We first consider the case that the time difference between the pilot arrivals at the BS is so small and hence negligible compared with one symbol duration. The channel is then effectively phase shifted by the time delay which can be modeled as

$$\mathbf{Y}_{B_i} = \sum_{l=1}^{N} \sum_{k=1}^{K} \sqrt{p_{l,k}\omega} \mathbf{g}_{i,l,k} e^{j\theta_{i,l,k}} \boldsymbol{\psi}_l + \mathbf{Z}_{B_i}, \tag{48}$$

where $e^{j\theta_{i,l,k}}$ is the effectively shifted phase from the kth user in the lth cell to BS i. It can be easily to see that such phase offset can be absorbed in the instantaneous fading coefficient and hence does not affect the system performance.

Next, we analyze the case that the time difference between the pilot arrivals at the BS is considerable compared with one symbol duration. We define the pilot sequence stream function as below

$$\psi_i(m) = \psi_{im}, \ m = 1, ..., \omega \tag{49}$$

where $\psi_i(m)$, the mth element of the pilot sequence $\boldsymbol{\psi}_i$, denotes the pilot symbol transmitted by the users of cell i at time m. Let $\tau_{i,l,k}$ denote the propagation delay from BS i to the kth user in the lth cell. Further let $A(t)$ denote a unit-energy baseband rectangular pulse of duration T_p used by the users. Based on the signal model in [38], we can rewrite the equivalent received signal of BS i at time t during the pilot transmission stage as follows

$$\mathbf{r}_{B_i}(t)$$
$$= \sum_{m=1}^{\omega} \sum_{l=1}^{N} \sum_{k=1}^{K} \sqrt{p_{l,k}\omega} A\left[t - (m-1)T_p - \tau_{i,l,k}\right] \mathbf{g}_{i,l,k} \psi_l(m)$$
$$+ \mathbf{z}_{B_i}(t) \tag{50}$$

where $\mathbf{z}_{B_i}(t)$ is the additive white Gaussian noise vector.

At BS i, the received signal at time t, $\mathbf{r}_{B_i}(t)$, is passed through a filter matched to $A\left[t - (m-1)T_p - \bar{\tau}_{B_i}\right]$, where $\bar{\tau}_{B_i}$ can be chosen as the average delay from the users to BS i without loss of generality. Then, we obtain that

$$\mathbf{y}_{B_i}(m) = \sum_{l=1}^{N} \sum_{k=1}^{K} \sqrt{p_{l,k}\omega} \mathbf{g}_{i,l,k} \tilde{\psi}_{i,l,k}(m) + \mathbf{z}_{B_i}(m) \tag{51}$$

where $\tilde{\psi}_{i,l,k}(m)$ is the received pilot symbol from the kth user in the lth cell at time m, which is polluted by asynchrony. It can be seen that $\tilde{\psi}_{i,l,k}(m)$ is determined by the time difference $\tau_{i,l,k} - \bar{\tau}_{B_i} \neq 0$. Let $\delta_{i,l,k} \triangleq (\tau_{i,l,k} - \bar{\tau}_{B_i}) \mod T_p$ and the resulting quotient be $m_{i,l,k}$, we have

$$\tilde{\psi}_{i,l,k}(m) = \rho(\delta_{i,l,k})\psi_l(m + m_{i,l,k})$$
$$+ \rho(T_p - \delta_{i,l,k})\psi_l(m + m_{i,l,k} - 1) \tag{52}$$

where $\rho(\tau) = \int_0^{T_p} A(t)A(t-\tau)dt$. We can see that the polluted pilot symbol $\tilde{\psi}_{i,l,k}(m)$ is the linear combination of two consecutive symbols $\psi_l(m+m_{i,l,k})$ and $\psi_l(m+m_{i,l,k}-1)$. Collecting the received symbols for all the ω transmission intervals at BS i, we have

$$\mathbf{Y}_{B_i} = \sum_{l=1}^{N}\sum_{k=1}^{K}\sqrt{p_{l,k}\omega}\,\mathbf{g}_{i,l,k}\tilde{\psi}_{i,l,k} + \mathbf{Z}_{B_i} \qquad (53)$$

where $\tilde{\psi}_{i,l,k} = \left[\tilde{\psi}_{i,l,k}(1),...,\tilde{\psi}_{i,l,k}(\omega)\right]$ is the received polluted pilot sequence.

As in the perfect synchronization case, BS i tries to estimate the composite channel for the users in cell i as

$$\hat{\mathbf{g}}_{i,c} = \mathbf{Y}_{B_i}\boldsymbol{\psi}_i^H$$
$$= \sum_{k=1}^{K}\sqrt{p_{i,k}\omega}\,\mathbf{g}_{i,i,k}\tilde{\psi}_{i,i,k}\boldsymbol{\psi}_i^H$$
$$\qquad + \sum_{l\neq i}^{N}\sum_{k=1}^{K}\sqrt{p_{l,k}\omega}\,\mathbf{g}_{i,l,k}\tilde{\psi}_{i,l,k}\boldsymbol{\psi}_i^H + \mathbf{z}_{B_i}$$
$$= \underbrace{\sum_{k=1}^{K}\sqrt{p_{i,k}\omega}\,\kappa_{i,i,k}\mathbf{g}_{i,i,k}}_{\text{desired CSI}} + \underbrace{\sum_{l\neq i}^{N}\sum_{k=1}^{K}\sqrt{p_{l,k}\omega}\,\kappa_{i,l,k}\mathbf{g}_{i,l,k}}_{\text{undesired CSI}} + \mathbf{z}_{B_i} \quad (54)$$

where $\kappa_{i,i,k} = \tilde{\psi}_{i,i,k}\boldsymbol{\psi}_i^H$ and $\kappa_{i,l,k} = \tilde{\psi}_{i,l,k}\boldsymbol{\psi}_i^H$. Since $\tilde{\psi}_{i,l,k} \neq \boldsymbol{\psi}_l$, the orthogonality of the pilot sequences is destroyed, which implies that $\kappa_{i,i,k} \neq 1$ and $\kappa_{i,l,k} \neq 0$, $\forall l \neq i$. Thus, compared with the perfect synchronous case, the estimated composite channel suffers a *scaling loss* as well as *pilot contamination*.

Finally, we demonstrate the asymptotical SINR of each user when the above mismatched channel is used in the beamformer design. Specifically, BS i applies the following normalized beamformer

$$\mathbf{w}_i = \frac{\hat{\mathbf{g}}_{i,c}}{\|\hat{\mathbf{g}}_{i,c}\|}$$
$$= \frac{\sum_{l=1}^{N}\sum_{k=1}^{K}\sqrt{\omega\beta_{i,l,k}p_{l,k}}\,\kappa_{i,l,k}\mathbf{h}_{i,l,k} + \mathbf{z}_{B_i}}{\mu_i\sqrt{M}}$$

where $\mu_i = \frac{\|\hat{\mathbf{g}}_{i,c}\|}{\sqrt{M}}$ is the normalization factor. Similar to the analysis with pilot contamination in Section IV, we can compute the asymptotic SINR of each user for the non-orthogonality scenario as

$$\lim_{M\to\infty}\text{SINR}_{i,k} = \frac{\frac{E_i}{\mu_{i,\infty}^2}\beta_{i,i,k}^2\kappa_{i,i,k}^2}{\sum_{j\neq i}^{N}\frac{E_j}{\mu_{j,\infty}^2}\beta_{j,i,k}^2\kappa_{j,i,k}^2 + \frac{\sigma^2}{\omega p_{i,k}}} \triangleq \mathcal{B}$$

where $\mu_{i,\infty}^2 \triangleq \lim_{M\to\infty}\mu_i^2$. When $E_i = E \to \infty, \forall i$, we have

$$\lim_{E\to\infty}\mathcal{B} = \frac{\beta_{i,i,k}^2\kappa_{i,i,k}^2}{\sum_{j\neq i}^{N}\frac{\mu_{i,\infty}^2}{\mu_{j,\infty}^2}\beta_{j,i,k}^2\kappa_{j,i,k}^2},$$

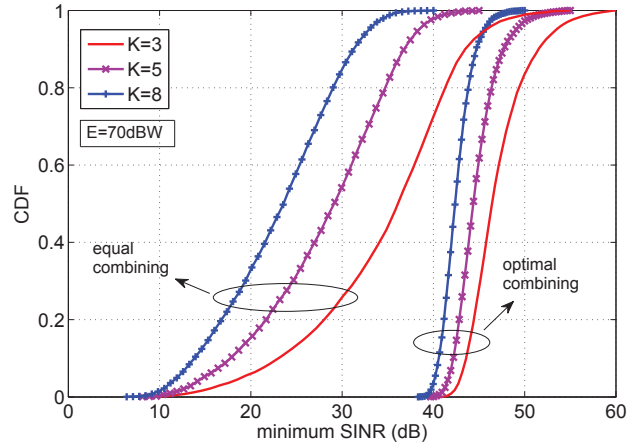which reaches a ceiling as in the pilot contamination case.

Fig. 2. The CDF of the minimum asymptotic SINR of the proposed beamformer under perfect CSI.

From the above analysis, we conclude that if the time difference between the pilot arrivals at the BS is not negligible compared with one symbol duration, the composite channel estimation suffers both scaling loss and pilot contamination and, therefore, the performance of the proposed pilot scheme will be degraded. How to overcome the effect of the non-orthogonality by asynchrony will be left for the future work.

VI. NUMERICAL AND SIMULATION RESULTS

In this section, we provide numerical and simulation results to illustrate the asymptotic as well as finite-size performance of the considered multicell multicast network with massive MIMO. The cells of the network are hexagonal with a radius (from center to vertex) of $r = 1000$ meters. The mobile users are randomly distributed in each cell with uniform distribution, excluding an inner circle of $r_0 = 100$ meters. The channel bandwidth is $20M$Hz. For the large-scale fading, the distance-dependent path loss in dB is modeled as $PL_{NLOS} = 128.1 + 37.6\log_{10}(d)$, where d is the distance from the user to the BS in kilometers. The log-normal shadowing is considered with $\sigma_{shadow} = 8$dB and the penetration loss is assumed to be 20dB. The small-scale fading is assumed as the normalized Rayleigh fading. The noise power spectral density is set to be -174dBm/Hz, and we set $\sigma_p^2 = 0.1\sigma^2$. Unless otherwise specified, throughout this section we consider the network containing $N = 7$ cells and choose the cell in the center for performance evaluation, which also represents the worst scenario. For each analytical result, 1000 independent large-scale channel realizations are generated. For simulations, the results for our proposed algorithm are obtained by averaging over 1000 independent small-scale fading parameters for each large-scale channel realization.

A. Asymptotic SINR Performance with Perfect CSI

In this subsection, we demonstrate the effectiveness of the asymptotically optimal beamformer when the perfect CSI is available at each BS.

Fig. 2 plots the CDF (Cumulative Distribution Function) curve of the minimum asymptotic SINR among all K users in the central cell. The SINR is computed according to (15). For
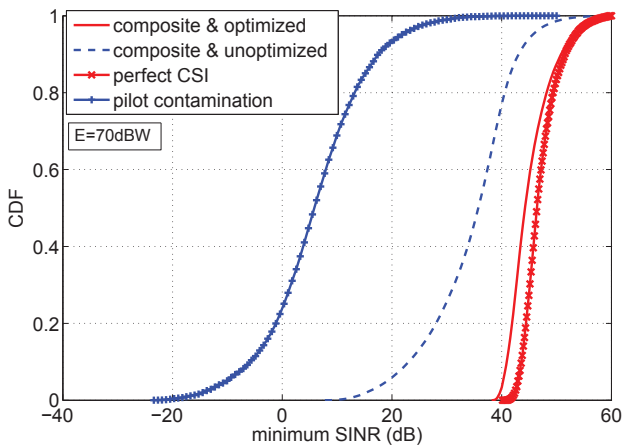
Fig. 3. The CDF of the minimum asymptotic SINR of the proposed beamformer with different CSI knowledge at $K = 3$ users per cell.
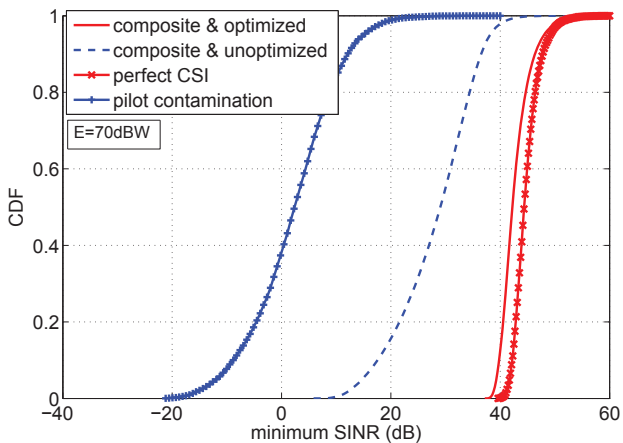

Fig. 4. The CDF of the minimum asymptotic SINR of the proposed beamformer with different CSI knowledge at $K = 5$ users per cell.
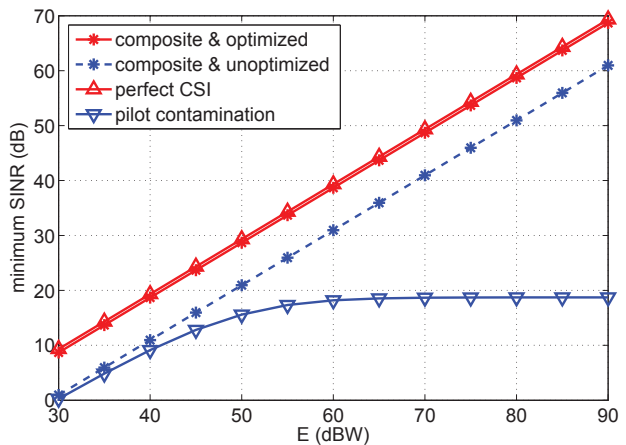

Fig. 5. The averaged minimum asymptotic SINR of the proposed beamformer with different CSI knowledge at $K = 3$ users per cell with varying E.
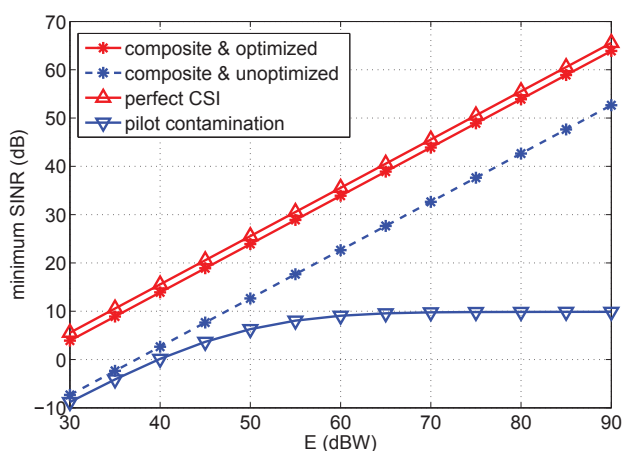

Fig. 6. The averaged minimum asymptotic SINR of the proposed beamformer with different CSI knowledge at $K = 5$ users per cell with varying E.

the "optimal combining" case, the parameters $\{\lambda_{i,k}\}$ are set as in (20); For the "equal combining" case, the parameters $\{\lambda_{i,k}\}$ are set as in (16) with $\{\xi_{i,k}\}$ being 1. We can see that by using the beamformer with the optimal combination coefficients, the 50-th percentile of SINR of the system can be improved by 10 dB for $K = 3$ over that with equal combination coefficients, and the gain increases when K increases. This clearly shows the significance of the linear combination coefficients in the proposed asymptotically optimal beamformer structure. From Fig. 2 we also see that the minimum asymptotic SINR decreases when the user number increases.

B. Asymptotic SINR Performance with Imperfect CSI

For the imperfect CSI scenario, each BS obtains the CSI through channel estimation. The length of the pilot sequence is set to be 8. The pilot peak transmit power is assumed as $p_u = 2\text{dB}W$ unless otherwise stated. We first investigate in Fig. 3 and Fig. 4 the asymptotic performance of the proposed beamformer with different CSI knowledge: the ideal case with perfect CSI, the conventional individual channel estimation with pilot contamination, and the proposed composite channel estimation using contamination-free pilot. The asymptotic S-INR of the "pilot contamination" case is computed according to (29) and we set $\xi_{i,k} = 1, \forall i, \forall k$ since we do not con-

sider parameter optimization. The asymptotic SINR of the composite channel estimation is computed according to (36). For the "composite & optimized" case, the pilot power is set according to (41); for the "composite & unoptimized" case, the pilot power of each user is set to be p_u. It is seen that the proposed composite channel estimation scheme can benefit significantly from the optimal pilot power control. In particular, the performance with optimal power control is close to the perfect CSI case (with optimal combining coefficients). It is also seen that the performance of the proposed pilot scheme for composite channel estimation without pilot power control is still much better than that of the conventional individual channel estimation scheme.

In Fig. 5 and Fig. 6, we plot the asymptotic SINR curves for different E's and similar phenomenon can be seen. We also observe that when increasing E, the average mininum SINR increases almost linearly. However, for the conventional individual channel estimation scheme, the asymptotic SINR reaches a ceiling when E becomes large due to the pilot contamination.

Next, we consider the effects of maximum pilot transmit power on the proposed composite channel scheme. In Fig. 7, we plot the asymptotic SINR of the proposed pilot scheme for different levels of pilot transmitting power. It can be seen
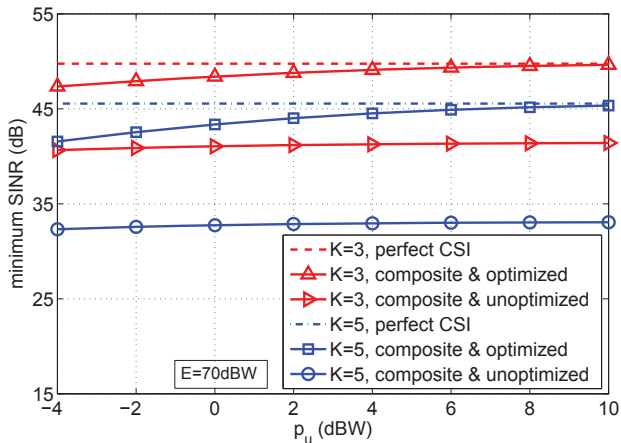
Fig. 7. The performance the proposed composite channel estimation with different pilot power levels.
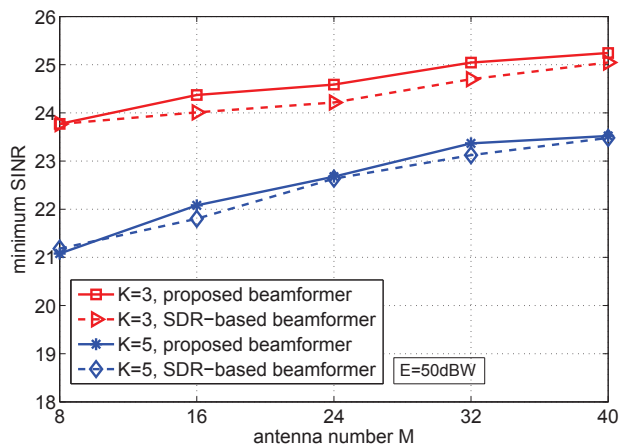

Fig. 8. Performance comparison between the proposed beamformer and the noncooperative SDR-based beamformer with perfect CSI.

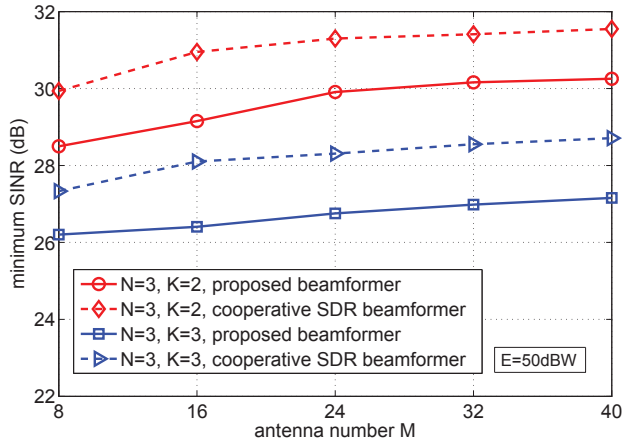
Fig. 9. Performance comparison between the proposed beamformer and the cooperative SDR-based beamformer with perfect CSI at $N = 3$ cells.

that its performance is closer to the perfect CSI case as the pilot transmitting power increases. This is because larger pilot transmitting power leads to smaller effect of the noise in the channel estimation.

C. Actual SINR Performance at Finite Number of Antennas

In this subsection, we show the actual SINR performance of the proposed scheme by simulation at finite number of antennas. We first compare the performance of the asymptotically optimal multicast beamformer in (21) with the traditional SDR-based multicast beamformer [4], [7]–[10][1] when perfect CSI is available. Then, we show the performance when using the new pilot scheme for composite channel estimation.

In the perfect CSI case, we first compare with the noncooperative SDR-based algorithm in which each BS individually designs their multicast beamformer to serve its users using its local CSI, and treats the intercell interference as noise. Fig. 8 shows the averaged minimum SINR performance. We can see that our proposed beamformer slightly outperforms the noncooperative SDR beamformer with the considered number of antennas. However, the complexity of the noncooperative SDR algorithm is much higher since it uses interior-point method to iteratively solve the semidefinite programming problem and applies bisection method for the max-min fairness design criterion. Our proposed beamformer, on the other hand, has closed-form solution and no numerical computation is needed.

We then compare with the cooperative SDR-based algorithm in which all the BSs coordinate with each other to design their multicast beamformers and the global CSI is needed for each BS. Note that the cooperative SDR beamformer has both much higher computational complexity and much higher signaling overhead. To lighten the computation burden of simulation, we consider the $N = 3$ case instead of $N = 7$ case in this comparison. The results are shown in Fig. 9. It can be seen that the cooperative SDR beamformer outperforms our

---

[1]For the SDR based algorithm, due to its high complexity, we generate 100 independent large-scale channel realizations and 100 independent small-scale fading parameters for each large-scale channel realization to maintain moderate calculation works in the simulation.

proposed beamformer with the considered number of antennas. This is reasonable since the cooperative SDR beamformer can be near-optimal for small number of users [9]. However, the cooperative SDR beamformer can hardly be used in the cellular network with massive MIMO since the computation complexity is too high and the global CSI is also very hard to be obtained at the BS end. Our proposed multicast beamformer is asymptotically optimal and only needs local CSI.

When the proposed pilot scheme for composite channel estimation is applied, the beamformer is obtained according to (34), where the pilot power is optimized according to (41). Fig. 10 shows the averaged minimum SINR of our proposed scheme at finite antenna number M and with different peak pilot power p_u. For comparison the asymptotic SINR obtained using (46) at infinite M is also plotted. It is seen that when $M = 300$, the actual achieved minimum SINR is only 1.5dB worse that the asymptotic result. The performance gap reduces to 1dB when $M = 500$.

VII. CONCLUSIONS

This paper considered the noncooperative multicell multicast network with massive MIMO. For the perfect CSI case, we obtained the asymptotically optimal beamformer for each BS in closed-form that is a function of the large-scale channel
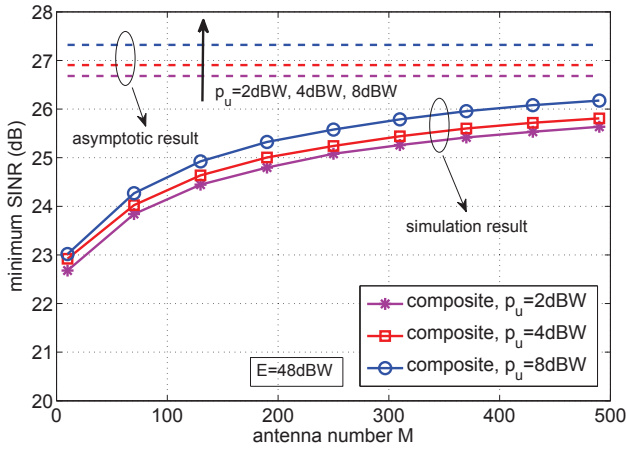
Fig. 10. The averaged minimum SINR of the proposed pilot scheme for composite channel estimation with finite M and at $K = 3$ users per cell.

attenuation $\{\beta_{i,j,k}\}$. For the imperfect CSI, we first considered the conventional individual channel estimation scheme and analyzed the effect of the pilot contamination. Then we proposed a novel pilot scheme for composite channel estimation that completely eliminates pilot contamination. By performing power control among users in the pilot transmission stage, our proposed scheme offers performance that is close to the performance under the perfect CSI scenario. Besides, we analyzed the effect of asynchrony on the performance of our proposed scheme.

APPENDIX A
PROOF OF LEMMA 1

We first prove (19) by contradiction. Let $\{\lambda^*_{i,k}\}$ be the optimal solution of problem $\mathcal{Q}$ but the condition (19) does not hold. Without loss of generality, we assume $\frac{\lambda^*_{i,1}E_i\beta_{i,i,1}}{\sigma^2}$ is the exclusive maximum one and $\frac{\lambda^*_{i,K}E_i\beta_{i,i,K}}{\sigma^2}$ is the exclusive minimum one[2], and $\frac{\lambda^*_{i,1}E_i\beta_{i,i,1}}{\sigma^2} > \frac{\lambda^*_{i,K}E_i\beta_{i,i,K}}{\sigma^2}$. We can then reset $\lambda^*_{i,1}$ and $\lambda^*_{i,K}$ as follows

$$\tilde{\lambda}^*_{i,1} = \lambda^*_{i,1} - \triangle\lambda,$$
$$\tilde{\lambda}^*_{i,K} = \lambda^*_{i,K} + \triangle\lambda,$$

where

$$\triangle\lambda = \frac{\lambda^*_{i,1}\beta_{i,i,1} - \lambda^*_{i,K}\beta_{i,i,K}}{\beta_{i,i,1} + \beta_{i,i,K}} > 0.$$

With the new parameters, we have that

$$\frac{\tilde{\lambda}^*_{i,1}E_i\beta_{i,i,1}}{\sigma^2} = \frac{\tilde{\lambda}^*_{i,K}E_i\beta_{i,i,K}}{\sigma^2} > \frac{\lambda^*_{i,K}E_i\beta_{i,i,K}}{\sigma^2}.$$

It means that the objective value of $\mathcal{Q}$ has been increased, which contradicts to the assumption that $\{\lambda^*_{i,k}\}$ is the optimal solution of problem $\mathcal{Q}$. Thus, we must have condition (19).

Combining (17) and (19), we can easily obtain (20) and thus the lemma is proven.

[2]The extension to the cases that there are multiple equal maximum SINR's and multiple equal minimum SINR's is straightforward and is omitted here.

APPENDIX B
PROOF OF LEMMA 2

We prove (40) by contradiction. If $p_{i,k^\star} < p_u$, we will have that

$$p_{i,k} < p_u, \ \forall k. \tag{55}$$

Supposing (55) does not hold, we can assume $p_{i,1} = p_u$ without loss of generality. Since $\beta^2_{i,i,1}p_{i,1} > \beta^2_{i,i,k^\star}p_{i,k^\star}$ and the expression of each user's SINR has a common denominator $\sum_{k'=1}^{K} \beta_{i,i,k'}p_{i,k'} + \frac{\sigma^2_p}{\omega}$, user 1's SINR is larger than that of user $k^\star$. Then, we can set user 1's pilot power as $p'_{i,1} < p_u$ to make that $\beta^2_{i,i,1}p'_{i,1} = \beta^2_{i,i,k^\star}p_{i,k^\star}$. Thus, the objective value of problem (37) will increase since the denominator $\sum_{k'=1}^{K} \beta_{i,i,k'}p_{i,k'} + \frac{\sigma^2_p}{\omega}$ decreases, which contradicts with the optimality.

Based on (55), we can always increase each user's pilot power as $p'_{i,k} = \eta p_{i,k}, \forall k$ where $\eta > 1$ without violating the power constraints. Then each item in the objective function of problem (37) will be

$$\frac{\beta^2_{i,i,k}\eta p_{i,k}}{\sum_{k'=1}^{K} \beta_{i,i,k'}\eta p_{i,k'} + \frac{\sigma^2_p}{\omega}} = \frac{\beta^2_{i,i,k}p_{i,k}}{\sum_{k'=1}^{K} \beta_{i,i,k'}p_{i,k'} + \frac{\sigma^2_p}{\eta\omega}} \tag{56}$$

which increases with $\eta > 1$. Then the objective value of problem (37) will increase which contradicts with the optimality. Thus, we must have (40).

REFERENCES

[1] D. Lecompte and F. Gabin, "Evolved multimedia broadcast/multicast service (eMBMS) in LTE-advanced: overview and Rel-11 enhancements," *IEEE Communications Magazine*, vol. 50, no. 11, pp. 68–74, November 2012.

[2] "LTE multicast: The future of mobile content delivery," Verizon Wireless News Center. http://www.verizonwireless.com/news/article/2014/01/lte-multicast-verizon-power-house.html.

[3] "AT&T to build LTE-broadcast network for video," USA Today. http://www.usatoday.com/story/tech/2013/09/25/att-lte-broadcast-network-video/2867475/.

[4] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[5] Y. Sun and K. Liu, "Transmit diversity techniques for multicasting over wireless networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, March 2004, pp. 593–598.

[6] M. Gans and B. Chen, "Beaconing in MIMO broadcast channels," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, March 2005, pp. iv/769–iv/772 Vol. 4.

[7] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1268–1279, 2008.

[8] Y. Gao and M. Schubert, "Group-oriented beamforming for multi-stream multicasting based on quality-of-service requirements," in *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec 2005, pp. 193–196.

[9] Z. Xiang, M. Tao, and X. Wang, "Coordinated multicast beamforming in multicell networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 12–21, Jan. 2013.

[10] B. Ozbek, D. L. Ruyet, and H. Khiari, "Adaptive resource allocation for multicast OFDM systems with multiple transmit antennas," in *IEEE International Conference on Communications (ICC)*, vol. 10, June 2006, pp. 4409–4414.

[11] H. Zhu, N. Prasad, and S. Rangarajan, "Precoder design for physical layer multicasting," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5932–5947, Nov 2012.

[12] R. Hunger, D. Schmidt, M. Joham, A. Schwing, and W. Utschick, "Design of single-group multicasting-beamformers," in *IEEE International Conference on Communications (ICC)*, June 2007, pp. 2499–2505.

[13] A. Abdelkader, A. Gershman, and N. Sidiropoulos, "Multiple-antenna multicasting using channel orthogonalization and local refinement," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3922–3927, July 2010.

[14] S. X. Wu and W.-K. Ma, "Multicast transmit beamforming using a randomize-in-time strategy," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 3376–3379.

[15] J. Wang, M. Zoltowski, and D. Love, "Improved space-time coding for multiple antenna multicasting," in *IEEE International Waveform Diversity Design Conference (WDDC)*, 2006, pp. 593–598.

[16] S. X. Wu, A. M.-C. So, and W.-K. Ma, "Rank-two transmit beamformed alamouti space-time coding for physical-layer multicasting," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 2793–2796.

[17] X. Wen, K. L. Law, S. J. Alabed, and M. Pesavento, "Rank-two beamforming for single-group multicasting networks using OSTBC," in *IEEE 7th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2012, pp. 69–72.

[18] 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 Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[19] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[20] H. Q. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser mimo systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[21] H. Yang and T. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[22] J. Hoydis, S. ten Brink, and M. Debbah, "Comparison of linear precoding schemes for downlink massive MIMO," in *IEEE International Conference on Communications (ICC)*, 2012, pp. 2135–2139.

[23] O. Alrabadi, E. Tsakalaki, H. Huang, and G. Pedersen, "Beamforming via large and dense antenna arrays above a clutter," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 314–325, Feb. 2013.

[24] RP-122015, "New SID proposal: Study on full dimension MIMO for LTE," *Samsung, 3GPP TSG RAN Meeting 58*, Barcelona, Spain, Dec. 4-7 2012.

[25] B. L. Ng, Y. Kim, J. Lee, Y. Li, Y.-H. Nam, J. C. Zhang, and K. Sayana, "Fulfilling the promise of massive MIMO with 2D active antenna array," in *Globecom Workshops, 2012 IEEE*. IEEE, 2012, pp. 691–696.

[26] Y.-H. Nam, B. L. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim, and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology." *IEEE Communications Magazine*, vol. 51, no. 6, Jun. 2013.

[27] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[28] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[29] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5450–5458, Oct. 2010.

[30] K. Wang, X. Wang, W. Xu, and X. Zhang, "Coordinated linear precoding in downlink multicell MIMO-OFDMA networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4264–4277, Aug. 2012.

[31] F. Fernandes, A. Ashikhmin, and T. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[32] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[33] H. Cramér, *Random variables and probability distributions*. Cambridge University Press, 2004, no. 36.

[34] N. Jindal and Z.-Q. Luo, "Capacity limits of multiple antenna multicast," in *IEEE International Symposium on Information Theory*, 2006, pp. 1841–1845.

[35] S. Park and D. Love, "Capacity limits of multiple antenna multicasting using antenna subset selection," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2524–2534, June 2008.

[36] I. H. Kim, D. Love, and S. Park, "Optimal and successive approaches to signal design for multiple antenna physical layer multicasting," *IEEE Transactions on Communications*, vol. 59, no. 8, pp. 2316–2327, August 2011.

[37] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming*. Available: http://stanford.edu.boyd/cvx.

[38] H. Zhang, N. Mehta, A. Molisch, J. Zhang, and H. Dai, "Asynchronous interference mitigation in cooperative base station systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 155–165, Jan 2008.

PLACE PHOTO HERE

Zhengzheng Xiang received the B.S. degree in electronic engineering from Shanghai Jiao Tong University, Shanghai, China, in 2010. He is currently working toward the Ph.D. degree with the Institute of Wireless Communication Technology, Shanghai Jiao Tong University.

His research interests include interference management in wireless networks, wireless relay technologies, and advanced signal processing for wireless cooperative communication.

PLACE PHOTO HERE

Meixia Tao (S'00-M'04-SM'10) received the B.S. degree in electronic engineering from Fudan University, Shanghai, China, in 1999, and the Ph.D. degree in electrical and electronic engineering from Hong Kong University of Science and Technology in 2003. She is currently a Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China. Prior to that, she was a Member of Professional Staff at Hong Kong Applied Science and Technology Research Institute during 2003-2004, and a Teaching Fellow then an Assistant Professor at the Department of Electrical and Computer Engineering, National University of Singapore from 2004 to 2007. Her current research interests include cooperative communications, wireless resource allocation, MIMO techniques, and physical layer security.

Dr. Tao is an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE WIRELESS COMMUNICATIONS LETTERS. She was on the Editorial Board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2011 and the IEEE COMMUNICATIONS LETTERS from 2009 to 2012. She also served as Guest Editor for IEEE COMMUNICATIONS MAGAZINE with feature topic on LTE-Advanced and 4G Wireless Communications in 2012, and Guest Editor for EURISAP J WCN with special issue on Physical Layer Network Coding for Wireless Cooperative Networks in 2010.

Dr. Tao is the recipient of the IEEE Heinrich Hertz Award for Best Communications Letters in 2013, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2009, and the International Conference on Wireless Communications and Signal Processing (WCSP) Best Paper Award in 2012.

PLACE PHOTO HERE

Xiaodong Wang (S'98-M'98-SM'04-F'08) received the Ph.D degree in Electrical Engineering from Princeton University.

He is a Professor of Electrical Engineering at Columbia University in New York. Dr. Wang's research interests fall in the general areas of computing, signal processing and communications, and has published extensively in these areas. Among his publications is a book entitled "Wireless Communication Systems: Advanced Techniques for Signal Reception", published by Prentice Hall in 2003. His current research interests include wireless communications, statistical signal processing, and genomic signal processing.

Dr. Wang received the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an Associate Editor for the *IEEE Transactions on Communications*, the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Signal Processing*, and the *IEEE Transactions on Information Theory*. He is a Fellow of the IEEE and listed as an ISI Highly-cited Author.