Non-Orthogonal Unicast and Broadcast Transmission via Joint Beamforming and LDM in Cellular Networks

Junlin Zhao†, Osvaldo Simeone*, Deniz Gündüz†, and David Gómez-Barquero‡

†Department of Electrical and Electronic Engineering, Imperial College London, London, UK
*CWCSPR, ECE Department, New Jersey Institute of Technology, NJ, USA
‡Institute of Telecommunications and Multimedia Applications, Universitat Politècnica de València, València, Spain
Email: j.zhao15@imperial.ac.uk, osvaldo.simeone@njit.edu, d.gunduz@imperial.ac.uk, dagobar@iteam.upv.es

Abstract*—Research efforts to incorporate multicast and broadcast transmission into the cellular network architecture are gaining momentum, particularly for multimedia streaming applications. Layered division multiplexing (LDM), a form of non-orthogonal multiple access (NOMA), can potentially improve unicast throughput and broadcast coverage with respect to traditional orthogonal frequency division multiplexing (FDM) or time division multiplexing (TDM), by simultaneously using the same frequency and time resources for multiple unicast or broadcast transmissions. In this paper, the performance of LDM-based unicast and broadcast transmission in a cellular network is studied by assuming a single frequency network (SFN) operation for the broadcast layer, while allowing for arbitrarily clustered cooperation for the transmission of unicast data streams. Beamforming and power allocation between unicast and broadcast layers, and hence the so-called *injection level* in the LDM literature, are optimized with the aim of minimizing the sum-power under constraints on the user-specific unicast rates and on the common broadcast rate. The problem is tackled by means of successive convex approximation (SCA) techniques, as well as through the calculation of performance upper bounds by means of semidefinite relaxation (SDR). Numerical results are provided to compare the orthogonal and non-orthogonal multiplexing of broadcast and unicast traffic.

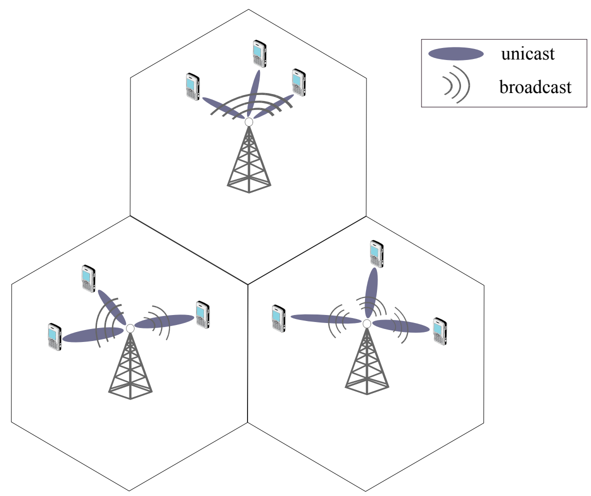

Fig. 1. Illustration of a multicell network with N=3 cells and $K = 3$ users in each cell with simultaneous unicast and broadcast transmission.

I. Introduction

Research efforts to incorporate multicast and broadcast transmission into the cellular network architecture have intensified in recent years, particularly targeting multimedia streaming applications. In 3G networks, multimedia broadcast multicast services (MBMS) was designed to introduce new point-to-multipoint (p-t-m) radio bearers and multicast support in the core network [1]. But it was never commercially deployed because of its limited potential to provide mass media services, mainly due to its reduced capacity. The broadcast extension of 4G LTE is named evolved MBMS (eMBMS), although it is commercially known as LTE Broadcast [2]. Following many field trials worldwide, the first commercial network was launched in South Korea in 2014. eMBMS provides full integration and seamless transition between broadcast and unicast modes [3] together with improved performance, thanks to the higher and more flexible LTE data rates. Furthermore, it also allows single frequency network (SFN) operation across different cells as in digital television broadcasting, since the LTE waveform is OFDM-based. Although eMBMS, as it exists today, needs further enhancements to be adopted for TV broadcasting [4], it has been proposed as a converged platform in the UHF band for TV and mobile broadband [5], [6]. There is ongoing work at 3GPP to study scenarios and use cases, and to identify potential requirements and improvements for eMBMS TV services [7].

LTE Broadcast entails reduction in system capacity for unicast services, since eMBMS and unicast services are multiplexed in time in different sub-frames. Superposition coding, a form of non-orthogonal multiple access (NOMA), was proposed in [8] to improve unicast throughput and broadcast coverage with respect to traditional orthogonal frequency division multiplexing (FDM) or time division multiplexing (TDM), by simultaneously using the same frequency and time resources for multiple unicast or broadcast transmissions.

Superposition coding has been adopted in the next-generation TV broadcasting US standard ATSC 3.0 [9] under the name layer division multiplexing (LDM) [10].

At the cost of an increased complexity at the receivers, which need to perform interference cancellation by decoding the generic broadcast content prior to decoding the unicast content, LDM may provide significant gains when the superposed signals exhibit large differences in terms of signal-to-noise-plus-interference ratio (SINR). This is expected to be the case for multiplexing broadcast and unicast services. In fact, the unicast throughput is limited by intercell interference, and hence, increasing the transmit unicast power across the network does not necessarily improve the unicast SINR. In contrast, broadcast does not suffer from intercell interference in an SFN, and increasing the broadcast power results in an increased SINR. A performance comparison of LDM with TDM/FDM for unequal error protection in broadcast systems in the absence of multicell interference from an information theoretic perspective can be found in [11].

In this paper, we study the performance of non-orthogonal unicast and broadcast transmission in a cellular network via LDM. We assume an SFN operation for the broadcast layer, while allowing arbitrarily clustered cooperation for the unicast data streams. Cooperative transmission for broadcast traffic, and potentially also for unicast data streams, takes place by means of beamforming at multi-antenna base stations. Beamforming and power allocation between unicast and broadcast layers, and hence the so-called injection level in the LDM literature (see, e.g., [11]), are optimized with the aim of minimizing the sum-power under constraints on the user-specific unicast rates and the common broadcast rate. The optimization of orthogonal transmission via TDM/FDM is also studied for comparison, and the corresponding nonconvex optimization problems are tackled by means of successive convex approximation (SCA) techniques [12], as well as through the calculation of performance upper bounds by means of semidefinite relaxation (SDR).

Related work includes [13][14], in which the optimization of beamforming for multicell systems was investigated for a scenario where the base station in each cell multicasts one or more data streams to given groups of in-cell users. The coexistence of broadcast and unicast traffic was studied in [15], in which the surplus of degrees of freedom provided by massive MIMO systems was leveraged to broadcast data to a group of users whose channel state information (CSI) is unavailable, without creating interference to conventional unicast users.

The rest of this paper is organized as follows. Section II introduces the system model and the problem formulation. In Section III, the characterized problem is tackled by using SDR and SCA techniques. Numerical results are presented in Section IV, followed by the conclusions in Section V.

II. System Model and Problem Formulation

A. System Model

We investigate downlink transmission in a cellular network that serves both unicast and broadcast traffic. Specifically, we focus on a scenario in which a dedicated unicast data stream is to be delivered to each user, while there is a common broadcast data stream intended for all the users. A more general broadcast traffic model, in which distinct data streams are sent to different subsets of users, could be included in the analysis at the cost of a more cumbersome notation, but will not be further pursued in this paper.

The network is comprised of N cells, each consisting of a base station (BS) with M antennas and K single-antenna mobile users. The notation (n, k) identifies the k-th user in cell n. All BSs cooperate via joint beamforming for the broadcast stream to all users, while an arbitrary cluster $\mathcal{C}_{n,k}$ of BSs cooperate for the unicast transmission to user (n, k) (see Fig. 1). The BSs in cluster $\mathcal{C}_{n,k}$ are hence informed about the unicast data stream to be delivered to user (n, k). Note that, non-cooperative unicast transmission, whereby each BS serves only the users in its own cell, can be obtained as a special case when $\mathcal{C}_{n,k} = \{n\}$, for all users (n, k). Similarly, fully cooperative unicast transmission is obtained when $\mathcal{C}_{n,k} = \{1, \ldots, N\}$, for all users (n, k). We denote the set of users whose unicast messages are available at BS i as

$$\mathcal{U}_i = \{(n, k) \mid i \in \mathcal{C}_{n,k}\}. \quad (1)$$

We assume frequency-flat quasi-static complex channels, and define $\boldsymbol{h}_{i,n,k} \in \mathbb{C}^{M\times 1}$ as the vector from the BS in cell i to user (n, k). The subscript i is used to denote the BS index in terms of transmission. Full CSI is assumed at all nodes. We use the notation $s^U_{n,k}$ to denote an encoded unicast symbol intended for user (n, k), and s^B to represent an encoded broadcast symbol.

The signal received by user (n, k) at any given channel use can be written as

$$y_{n,k} = \sum_{i=1}^{N} \boldsymbol{h}^H_{i,n,k} \boldsymbol{x}_i + n_{n,k}, \quad (2)$$

where $\boldsymbol{x}_i \in \mathbb{C}^{M\times 1}$ is the symbol transmitted by BS i, and $n_{n,k} \sim \mathcal{CN}(0, \sigma^2_{n,k})$ is the additive white Gaussian noise.

1) TDM: We first consider the standard TDM approach based on the orthogonal transmission of unicast and broadcast signals. Note that orthogonalization can also be realized by means of other multiplexing schemes such as FDM, yielding the same mathematical formulation. With TDM, each transmission slot of duration T channel uses, is divided into two subslots: a subslot of duration T_0 channel uses for unicast transmission, and a subslot of duration $T - T_0$ for broadcast transmission. Therefore, the signal $\boldsymbol{x}_i$ transmitted by cell i can be written as

$$\boldsymbol{x}_i = \begin{cases} \sum\limits_{(n,k)\in\mathcal{U}_i} \boldsymbol{w}^U_{i,n,k} s^U_{n,k} & \text{for} \quad 0 \le t < T_0 \\ \boldsymbol{w}^B_i s^B & \text{for} \quad T_0 \le t < T \end{cases}, \quad (3)$$

where $\boldsymbol{w}_{i,n,k}^U \in \mathbb{C}^{M\times 1}$ represents the unicast beamforming vector applied at the BS in cell i towards user (n,k), and $\boldsymbol{w}_i^B \in \mathbb{C}^{M\times 1}$ is the broadcast beamforming vector applied at the same BS.

The received signal $y_{n,k}$ at user (n,k) can be expressed as

$$y_{n,k} = \begin{cases} \Big(\sum\limits_{i\in\mathcal{C}_{n,k}} \boldsymbol{h}_{i,n,k}^H \boldsymbol{w}_{i,n,k}^U\Big) s_{n,k}^U + z_{n,k} + n_{n,k} & \text{for } 0 \le t < T_0 \\ \Big(\sum\limits_{i=1}^{N} \boldsymbol{h}_{i,n,k}^H \boldsymbol{w}_i^B\Big) s^B + n_{n,k} & \text{for } T_0 \le t < T \end{cases}, \tag{4}$$

where

$$z_{n,k} = \sum_{(p,q)\neq(n,k)} \Big(\sum_{i\in\mathcal{C}_{p,q}} \boldsymbol{h}_{i,n,k}^H \boldsymbol{w}_{i,p,q}^U\Big) s_{p,q}^U \tag{5}$$

denotes the interference at user (n,k).

2) LDM: With LDM, the transmitted signal $\boldsymbol{x}_i$ from the BS in cell i is the superposition of the broadcast and unicast signals for the entire time slot, which can be written as

$$\boldsymbol{x}_i = \boldsymbol{w}_i^B s^B + \sum_{(n,k)\in\mathcal{U}_i} \boldsymbol{w}_{i,n,k}^U s_{n,k}^U \quad \text{for } 0 \le t \le T, \tag{6}$$

for all channel uses in an entire time slot, i.e., for $0 \le t \le T$. We note that the power ratio between broadcast and unicast, which is referred to as the injection level (IL) in the literature (see, e.g., [11]), can be obtained as

$$\text{IL} = 10\,\log_{10}\frac{P^B}{P^U}, \tag{7}$$

where $P^B = \sum_{i=1}^N ||\boldsymbol{w}_i^B||^2$ is the total broadcast power, and $P^U = \sum_{i=1}^N \sum_{(n,k)\in\mathcal{U}_i} ||\boldsymbol{w}_{i,n,k}^U||^2$ is the total unicast power. The received signal at user (n,k) is hence given as

$$\begin{aligned} y_{n,k} \;\; = \;\; & \Big(\sum_{i=1}^{N} \boldsymbol{h}_{i,n,k}^H \boldsymbol{w}_i^B\Big) s^B + \Big(\sum_{i\in\mathcal{C}_{n,k}} \boldsymbol{h}_{i,n,k}^H \boldsymbol{w}_{i,n,k}^U\Big) s_{n,k}^U \\ & + z_{n,k} + n_{n,k} \quad \text{for } 0 \le t \le T, \end{aligned} \tag{8}$$

where $z_{n,k}$ is defined in (5).

B. Problem Formulation

Assuming that each user has a quality of service (QoS) requirement for both broadcast and unicast services, the optimization problem is formulated as the minimization of the sum-power under these QoS requirements. For both TDM and LDM systems, the optimization problem can be expressed in the following form:

$$\min \;\; \sum_{i=1}^{N} \Big(||\boldsymbol{w}_i^B||^2 + \sum_{(n,k)\in\mathcal{U}_i} ||\boldsymbol{w}_{i,n,k}^U||^2\Big) \tag{9a}$$

$$\text{s.t.} \;\; \text{SINR}_{n,k}^B \ge \gamma^B, \; \forall n,k \tag{9b}$$

$$\text{SINR}_{n,k}^U \ge \gamma_{n,k}^U, \; \forall n,k, \tag{9c}$$

where the explicit expressions for the SINRs at user (n,k) for broadcast and unicast transmissions, namely $\text{SINR}_{n,k}^B$ and $\text{SINR}_{n,k}^U$ will be given below for TDM and LDM. Note that, since all users receive the same broadcast signal, we have enforced a common broadcast QoS requirement. In contrast, the unicast SINR requirements are allowed to be user-dependent.

1) TDM: From the expression of the received signal in (4), we derive the SINR for the broadcast layer in TDM for user (n,k) as

$$\text{SINR}_{n,k}^{B\text{-TDM}} = \frac{|\boldsymbol{h}_{n,k}^H \boldsymbol{w}^B|^2}{\sigma_{n,k}^2}, \tag{10}$$

where $\boldsymbol{h}_{n,k} = [\boldsymbol{h}_{1,n,k}^T, \ldots, \boldsymbol{h}_{N,n,k}^T]^T \in \mathbb{C}^{NM\times 1}$ collects the channel vectors from all BSs to user (n,k). All broadcast beamforming vectors are similarly aggregated into the vector $\boldsymbol{w}^B = [\boldsymbol{w}_1^{B^T}, \ldots, \boldsymbol{w}_N^{B^T}]^T \in \mathbb{C}^{NM\times 1}$. The SINR for the unicast layer is instead given as

$$\text{SINR}_{n,k}^{U\text{-TDM}} = \frac{|\boldsymbol{h}_{n,k}^{(n,k)^H} \boldsymbol{w}_{n,k}^U|^2}{\sum\limits_{(p,q)\neq(n,k)} |\boldsymbol{h}_{n,k}^{(p,q)^H} \boldsymbol{w}_{p,q}^U|^2 + \sigma_{n,k}^2}, \tag{11}$$

where $\boldsymbol{h}_{n,k}^{(p,q)} = [\boldsymbol{h}_{i,n,k}^{U^T}]_{i\in\mathcal{C}_{p,q}}^T$ collects the channel vectors to user (n,k) from all the BSs in the cluster $\mathcal{C}_{p,q}$ of BSs that serve user (p,q), and $\boldsymbol{w}_{n,k}^U = [\boldsymbol{w}_{i,n,k}^{U^T}]_{i\in\mathcal{C}_{n,k}}^T$ is similarly defined as the aggregate unicast beamforming vector for user (n,k) from all the BSs in cluster $\mathcal{C}_{n,k}$.

We observe that the SINR targets $\gamma_{n,k}^{U\text{-TDM}}$ and $\gamma^{B\text{-TDM}}$ for unicast and broadcast traffic can be obtained from the corresponding transmission rates $R_{n,k}^U$ and R^B, respectively, as

$$\frac{T_0}{T}\log_2(1+\gamma_{n,k}^{U\text{-TDM}}) = R_{n,k}^U, \tag{12}$$

and

$$\frac{T-T_0}{T}\log_2(1+\gamma^{B\text{-TDM}}) = R^B. \tag{13}$$

2) LDM: With LDM, the broadcast layer, which is intended for all users and usually has a higher SINR, is decoded first by treating unicast signals as background noise, as in [8]. The users decode their unicast data streams after canceling the decoded broadcast message. The broadcast SINR in LDM for user (n,k) is hence obtained from the received signal (8) as follows

$$\text{SINR}_{n,k}^{B\text{-LDM}} = \frac{|\boldsymbol{h}_{n,k}^H \boldsymbol{w}^B|^2}{\sum\limits_{(p,q)} |\boldsymbol{h}_{n,k}^{(p,q)^H} \boldsymbol{w}_{p,q}^U|^2 + \sigma_{n,k}^2}, \tag{14}$$

while the unicast SINR is the same as TDM given in (11), i.e.,

$$\text{SINR}_{n,k}^{U\text{-LDM}} = \text{SINR}_{n,k}^{U\text{-TDM}}. \tag{15}$$

Similarly to TDM, SINR thresholds for unicast and broadcast can be obtained from the transmission rates $R_{n,k}^U$ and R^B, respectively, as

$$\log_2(1+\gamma_{n,k}^{U\text{-LDM}}) = R_{n,k}^U, \tag{16}$$

and

$$\log_2(1+\gamma^{B\text{-LDM}}) = R^B. \tag{17}$$

III. OPTIMIZATION

The optimization problem formulated in (9) is nonconvex due to the QoS constraints (9b) and (9c). In this section, we first consider the SDR approach in order to obtain a lower bound on the minimum sum-power. Then we leverage the SCA algorithm [12] in order to obtain a feasible solution, and hence, an upper bound on the minimum sum-power.

A. Lower Bound via Semidefinite Relaxation

In order to define an SDR of problem (9), we introduce the matrices $\boldsymbol{W}^B = \boldsymbol{w}^B \boldsymbol{w}^{B^H}$, $\boldsymbol{W}^U_{n,k} = \boldsymbol{w}^U_{n,k} \boldsymbol{w}^{U^H}_{n,k}$, $\boldsymbol{H}_{n,k} = \boldsymbol{h}_{n,k} \boldsymbol{h}^H_{n,k}$, and $\boldsymbol{H}^{(p,q)}_{n,k} = \boldsymbol{h}^{(p,q)}_{n,k} \boldsymbol{h}^{(p,q)^H}_{n,k}$ for all $i, n, p = 1, \ldots, N$ and $k, q = 1, \ldots, K$. With these definitions, the problem in (9) can then be transformed into a semidefinite program by dropping the rank constraints on matrices $\boldsymbol{W}^B$ and $\boldsymbol{W}^U_{n,k}$. Specifically, for TDM, the relaxed problem can be written as

$$\min \;\; \text{tr}(\boldsymbol{W}^B) + \sum_{(n,k)} \text{tr}(\boldsymbol{W}^U_{n,k}) \tag{18a}$$

$$\text{s.t.} \;\; \text{tr}(\boldsymbol{H}_{n,k} \boldsymbol{W}^B) \geq \gamma^B \sigma^2_{n,k}, \;\; \forall n, k \tag{18b}$$

$$\text{tr}(\boldsymbol{H}^{(n,k)}_{n,k} \boldsymbol{W}^U_{n,k}) \geq \gamma^U_{n,k} \sum_{(p,q) \neq (n,k)} \text{tr}(\boldsymbol{H}^{(p,q)}_{n,k} \boldsymbol{W}^U_{p,q}) + \gamma^U_{n,k} \sigma^2_{n,k}, \;\; \forall n, k \tag{18c}$$

$$\boldsymbol{W}^B \succeq 0 \tag{18d}$$

$$\boldsymbol{W}^U_{n,k} \succeq 0, \;\; \forall n, k, \tag{18e}$$

while, for LDM, the SDR formulation is

$$\min \;\; \text{tr}(\boldsymbol{W}^B) + \sum_{(n,k)} \text{tr}(\boldsymbol{W}^U_{n,k}) \tag{19a}$$

$$\text{s.t.} \;\; \text{tr}(\boldsymbol{H}_{n,k} \boldsymbol{W}^B) \geq \gamma^B \sum_{(p,q)} \text{tr}(\boldsymbol{H}^{(p,q)}_{n,k} \boldsymbol{W}^U_{p,q}) + \gamma^B \sigma^2_{n,k}, \;\; \forall n, k \tag{19b}$$

and Eq. (18c), (18d), and (18e).

We emphasize that, due to the relaxation of the rank-1 constraints on the matrices $\boldsymbol{W}^B$ and $\boldsymbol{W}^U_{n,k}$, the optimal values of the convex programs (18) and (19) yield lower bounds on the minimum sum-power obtained from problem (9). We also observe that, if the optimal solutions $\boldsymbol{W}^B$ and $\boldsymbol{W}^U_{n,k}$ of problems (18) and (19) happen to be all rank 1, i.e., $\boldsymbol{W}^B = \boldsymbol{w}^B \boldsymbol{w}^{B^H}$ and $\boldsymbol{W}^U_{n,k} = \boldsymbol{w}^B_{n,k} \boldsymbol{w}^{U^H}_{n,k}$, then the obtained solutions will indeed be the optimal solutions of the original problem.

B. Upper Bound via Successive Convex Approximation

In order to obtain an effective feasible solution of the problem (9), we now leverage the SCA algorithm outlined in [12]. In particular, given that the nonconvex QoS constraints can be rewritten as the difference of convex (DC) functions, the SCA algorithm reduces to the conventional Convex-Concave Procedure [16]. It is recalled that the SCA scheme is known to converge to a stationary point of the original problem [12]. In order to apply the SCA approach, each nonconvex constraint in (9) is expressed as

$$g(\boldsymbol{w}) = g^+(\boldsymbol{w}) - g^-(\boldsymbol{w}) \leq 0, \tag{20}$$

where $g^+(\boldsymbol{w})$ and $g^-(\boldsymbol{w})$ are both convex functions on the set of all beamforming vectors $\boldsymbol{w}$. Then a convex upper bound is obtained by linearizing the nonconvex part around any given vector $\boldsymbol{u}$, yielding the stricter constraint on the solution $\boldsymbol{w}$ as

$$\tilde{g}(\boldsymbol{w}; \boldsymbol{u}) \triangleq g^+(\boldsymbol{w}) - g^-(\boldsymbol{u}) - \bigtriangledown_{\boldsymbol{w}} g^-(\boldsymbol{u})^T (\boldsymbol{w} - \boldsymbol{u}) \leq 0. \tag{21}$$

For example, the constraint on unicast QoS in (9c) for both TDM and LDM can be replaced by

$$\gamma^U_{n,k} \Big(\sum_{(p,q) \neq (n,k)} |\boldsymbol{h}^{(p,q)^H}_{n,k} \boldsymbol{w}^U_{p,q}|^2 + \sigma^2_{n,k} \Big) - \Big(|\boldsymbol{h}^{(n,k)^H}_{n,k} \boldsymbol{w}^U_{n,k}(\nu)|^2 + 2\, \boldsymbol{w}^{U^H}_{n,k}(\nu) \boldsymbol{H}^{(n,k)}_{n,k} \left(\boldsymbol{w}^U_{n,k} - \boldsymbol{w}^U_{n,k}(\nu) \right) \Big) \leq 0, \;\; \forall n, k, \tag{22}$$

for a given beamforming vector $\boldsymbol{w}^U_{n,k}(\nu)$, which will correspond to the current iteration, indexed by an integer ν, of the SCA algorithm, as discussed below.

The SCA method is detailed in Table I. For TDM, the optimization problem to be solved in the ν-th iteration of the algorithm is

$$\min \;\; ||\boldsymbol{w}^B||^2 + \sum_{(n,k)} ||\boldsymbol{w}^U_{n,k}||^2 \tag{23a}$$

$$\text{s.t.} \; \boldsymbol{w}^{B^H}(\nu) \boldsymbol{H}_{n,k} \boldsymbol{w}^B \geq \frac{|\boldsymbol{h}^H_{n,k} \boldsymbol{w}^B(\nu)|^2 + \gamma^B \sigma^2_{n,k}}{2}, \;\; \forall n, k \tag{23b}$$

and Eq. (22).

Instead, the problem to be solved for LDM is

$$\min \;\; ||\boldsymbol{w}^B||^2 + \sum_{(n,k)} ||\boldsymbol{w}^U_{n,k}||^2 \tag{24a}$$

$$\text{s.t.} \; \sum_{(p,q)} |\boldsymbol{h}^{(p,q)^H}_{n,k} \boldsymbol{w}^U_{p,q}|^2 - \frac{2}{\gamma^B} \boldsymbol{w}^B(\nu)^H \boldsymbol{H}_{n,k} \boldsymbol{w}^B \leq -\sigma^2_{n,k} - \frac{1}{\gamma^B} |\boldsymbol{h}^H_{n,k} \boldsymbol{w}^B(\nu)|^2, \;\; \forall n, k \tag{24b}$$

and Eq. (22).

Due to the fact that the convexified constraints in (21) are stricter than the original constraints in (20), the solution obtained at each iteration is feasible for the original problem (9) as long as the initial point is available. When the stopping criterion is satisfied, we take the last iteration as the solution of the SCA algorithm.

TABLE I
SCA ALGORITHM

STEP 0:	Set $\nu = 1$. Set a step size μ. Initialize $\boldsymbol{w}^B(1)$ and $\boldsymbol{w}^U_{n,k}(1)$ with feasible values
STEP 1:	If a stopping criterion is satisfied, then STOP
STEP 2:	Set $\boldsymbol{w}^B(\nu+1) = \boldsymbol{w}^B(\nu) + \mu(\boldsymbol{w}^B - \boldsymbol{w}^B(\nu))$, $\boldsymbol{w}^U_{n,k}(\nu+1) = \boldsymbol{w}^U_{n,k}(\nu) + \mu(\boldsymbol{w}^U_{n,k} - \boldsymbol{w}^U_{n,k}(\nu))$, where $\{\boldsymbol{w}^B\}$ and $\{\boldsymbol{w}^U_{n,k}\}$ are obtained as solutions of problems (23) for TDM and (24) for LDM
STEP 3:	Set $\nu = \nu + 1$, and go to STEP 1

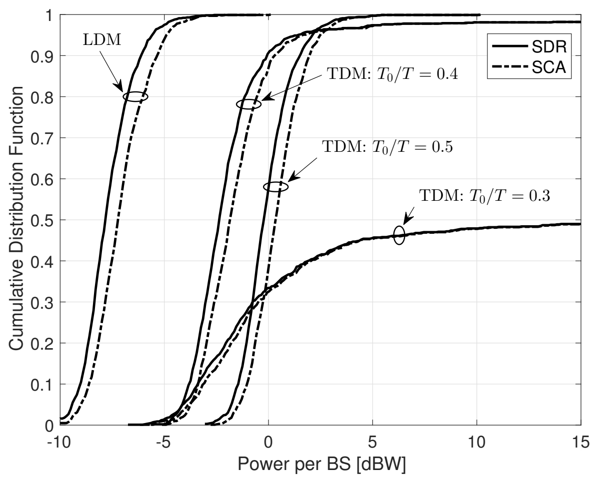

Fig. 2. The CDF of power consumption per BS with target rates R^B=3 bps/Hz and R^U=0.3 bps/Hz.

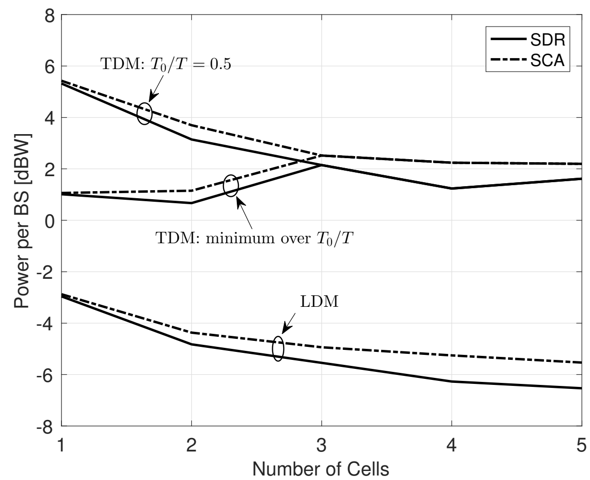

Fig. 3. Power consumption per BS as a function of the number of cells with target rates R^B=3 bps/Hz and R^U=0.3 bps/Hz.

When obtaining the numerical results in the next section, initialization of the SCA algorithm is carried out based on the solution $\{\boldsymbol{W}^B\}$ and $\{\boldsymbol{W}^U_{n,k}\}$ obtained by SDR. Specifically, we perform a rank-1 reduction of matrices $\{\boldsymbol{W}^B\}$ and $\{\boldsymbol{W}^U_{n,k}\}$, obtaining vectors $\{\boldsymbol{w}^B_i\}$ and $\{\boldsymbol{w}^U_{n,k}\}$, respectively, as the largest principal component. These vectors are then scaled with the smallest common factor t to satisfy constraints (9b) and (9c), yielding the initial points $\{\boldsymbol{w}^B(1)\}$ and $\{\boldsymbol{w}^U_{n,k}(1)\}$ for SCA. To this end, a line search is performed over t. If a feasible value for t is not found through a line search, then the SCA method is considered to be infeasible. Further discussion on this point can be found in the next section.

IV. Simulation Results

In this section, simulation results are presented to obtain insights into the performance comparison between LDM and TDM for the purpose of transmission of unicast and broadcast services in cellular systems. We consider a network comprised of macro-cells, each with $K = 3$ single-antenna active users. The radius of each cell is 500 m, and the users are located uniformly around the BS at a distance of 400 m. Each BS is equipped with $M = 3$ antennas. All channel vectors $\boldsymbol{h}_{i,n,k}$ are written as $\boldsymbol{h}_{i,n,k} = (1/\mathrm{PL}_0)^{\frac{1}{2}} (d_0/d_{i,n,k})^{\frac{\mathrm{PL}}{2}} \tilde{\boldsymbol{h}}_{i,n,k}$, where the path loss exponent $\mathrm{PL} = 3$, $d_{i,n,k}$ is the distance (in meters) between the i-th BS and user (n, k), $\tilde{\boldsymbol{h}}_{i,n,k}$ is an i.i.d. vector accounting for Rayleigh fading of unitary power, and d_0 is a reference distance at which a transmission power of 0 dBW yields an average SNR per antenna of $1/\mathrm{PL}_0 = -10$ dB, where the average is taken with respect to Rayleigh fading. The common noise variance is set to $\sigma^2_{n,k} = 1$ for all users (n, k). Unless stated otherwise, we assume non-cooperative unicast transmission, i.e., each BS is informed only about the unicast data streams of its own users.

We first plot the cumulative distribution function (CDF) of the transmission power per BS for LDM and TDM with $N = 3$ cells in Fig. 2. For the latter, we consider different values for the fraction of time T_0/T devoted to unicast traffic. Other values of T_0/T were seen not to improve the performance. The transmission power per BS is defined as the sum-power divided by the number of BSs. We observe that the curves may represent improper CDFs in the sense that there exists a gap between their asymptotic values with large powers and 1. This gap accounts for the probability of the set of channel realizations in which the problem is found to be infeasible. We refer to the previous section for the assumed definition of infeasibility for SCA, whereas the standard definition is used for the convex problem in (18) and (19) solved by SDR. Henceforth, we refer to the probability of an infeasible channel realization as the *outage probability*.

We can observe from Fig. 2 that LDM enables a significant reduction in the transmission power per BS as compared with TDM. In fact, even with an optimized choice of T_0/T, LDM can improve the 95th percentile of the transmitted power per BS by around 5 dB. Another observation is that SCA operates close to the lower bound set by SDR. Note also that LDM has a significantly lower outage probability than TDM. Finally, we remark that a large value of T_0/T is beneficial to obtain a lower outage probability in TDM, suggesting that the unicast constraints have more significant impact on the feasibility of the problem due to the need to cope with the mutual interference among unicast data streams. For the rest of this section, the displayed power values correspond to the 95th percentile of the corresponding CDF.

We now study the impact of the number of cells on the performance of the system. To this end, Fig. 3 and Fig. 4 show the power per BS as a function of the number of cells. Specifically, Fig. 3 shows the overall power per BS, while Fig. 4 illustrates separately the power per BS used for the broadcast and unicast layers. Note that in Fig. 4 we fixed $T_0/T = 0.5$, while in Fig. 3 we also show the power obtained by selecting, for any number of cells, the value of T_0/T that minimizes the overall sum-power consumption (obtained by a

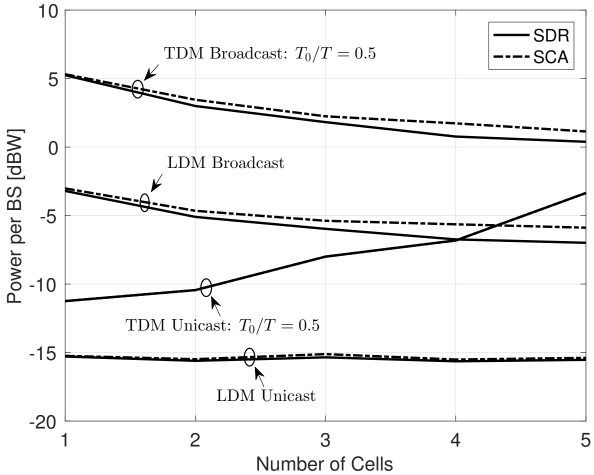

Fig. 4. Unicast and broadcast power consumption per BS as a function of the number of cells with target rates R^B=3 bps/Hz and R^U=0.3 bps/Hz.

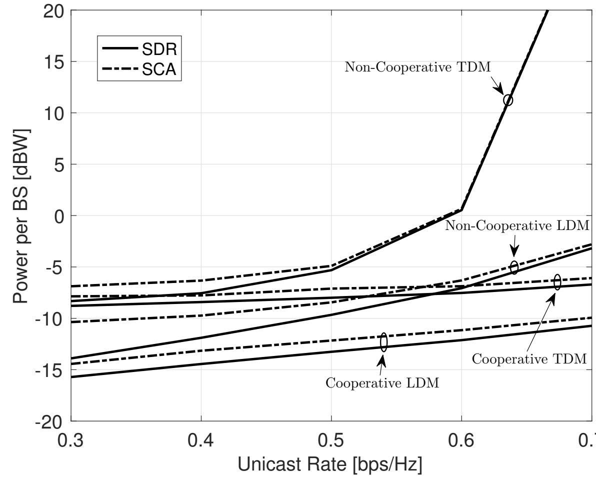

Fig. 5. Power consumption per BS for values of unicast rate with R^B=0.5 bps/Hz for non-cooperative and fully cooperative schemes.

line search). A key observation from Fig. 3 is that the power saving afforded by LDM increases with the number of cells. This gain can be attributed to the following two facts: (*i*) the optimal injection level is high (see Fig. 4), and hence the broadcast layer requires more power than unicast; and (*ii*) the performance of LDM is enhanced by the presence of more cells broadcasting the same message in the SFN, which increases the broadcast SINR and the broadcast layer can be more easily canceled by the users. The latter fact can be seen from Fig. 4, in which the required unicast power decreases with the number of cells when using LDM, unlike TDM. Furthermore, the optimal IL of TDM decreases significantly, also suggesting that TDM is more sensitive to the mutual interference introduced by unicast data streams.

Finally, Fig. 5 compares the required power per BS for non-cooperative unicast transmission and for fully cooperative unicast transmission, i.e., clusters $\mathcal{C}_{n,k} = \{1, \ldots, N\}$ for all users (n, k). Here we consider a network comprised of $N = 3$ cells, and set $T_0/T = 0.8$ for TDM. From Fig. 5, it can be concluded that a higher unicast rate entails larger power savings by means of cooperative unicast transmission, especially for TDM.

V. Conclusions

In this paper we have analyzed the performance gain of LDM over TDM/FDM as a potential non-orthogonal multiplexing approach for broadcast and unicast transmission. Joint beamforming design and power allocation was formulated as a sum-power minimization problem under QoS constraints. The resultant nonconvex optimization problem was tackled by means of SCA and SDR, which provide upper and lower bounds on the optimal solution, respectively. A general conclusion is that LDM has a significantly better performance than TDM with respect to power consumption and system robustness, particularly for dense deployments with a sufficiently large number of cells.

References

[1] F. Hartung, U. Horn, J. Huschke, M. Kampmann, T. Lohmar, and M. Lundevall, "Delivery of broadcast services in 3G networks," *IEEE Trans. Broadcast.*, vol. 53, no. 1, pp. 188–199, Mar. 2007.

[2] Qualcomm, "LTE broadcast," https://www.qualcomm.com/documents/lte-broadcast-white-paper-idc, Sep. 2014.

[3] J. F. Monserrat, J. Calabuig, A. Fernandez-Aguilella, and D. Gómez-Barquero, "Joint delivery of unicast and E-MBMS services in LTE networks," *IEEE Trans. Broadcast.*, vol. 58, no. 2, pp. 157–167, Jun. 2012.

[4] G. K. Walker, J. Wang, C. Lo, X. Zhang, and G. Bao, "Relationship between LTE broadcast/eMBMS and next generation broadcast television," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 185–192, Jun. 2014.

[5] L. Shi, E. Obregon, K. W. Sung, J. Zander, and J. Bostrom, "CellTV - on the benefit of TV distribution over cellular networks: A case study," *IEEE Trans. Broadcast.*, vol. 60, no. 1, pp. 73–84, Mar. 2014.

[6] L. Shi, K. W. Sung, and J. Zander, "Future TV content delivery over cellular networks from urban to rural environments," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6177–6187, Nov. 2015.

[7] 3GPP TR 22.816 V2.0.0, "3GPP enhancement for TV service," http://www.3gpp.org/DynaReport/22816.htm, Dec. 2015.

[8] D. Kim, F. Khan, C. V. Rensburg, Z. Pi, and S. Yoon, "Superposition of broadcast and unicast in wireless cellular systems," *IEEE Commun. Mag.*, vol. 46, no. 7, pp. 110–117, Jul. 2008.

[9] L. Fay, L. Michael, D. Gómez-Barquero, N. Ammar, and M. W. Caldwell, "An overview of the ATSC 3.0 physical layer specification," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 159–171, Mar. 2016.

[10] S. I. Park *et al.*, "Low complexity layered division multiplexing for ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 233–243, Mar. 2016.

[11] D. Gómez-Barquero and O. Simeone, "LDM versus FDM/TDM for unequal error protection in terrestrial broadcasting systems: An information-theoretic view," *IEEE Trans. Broadcast.*, vol. 61, no. 4, pp. 571–579, Dec. 2015.

[12] G. Scutari, F. Facchinei, L. Lampariello, and P. Song, "Parallel and Distributed Methods for Nonconvex Optimization-Part I: Theory," Available: http://arxiv.org/abs/1410.4754.

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

[14] H. Pennanen, D. Christopoulos, S. Chatzinotas, and B. E. Ottersten, "Distributed coordinated beamforming for multi-cell multigroup multicast systems," Available: http://arxiv.org/abs/1511.03561.

[15] E. G. Larsson and H. V. Poor, "Joint beamforming and broadcasting in massive MIMO," Available: http://arxiv.org/abs/1602.03648.

[16] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, Apr. 2003.