

Caching in Gaussian Interference Channel with QoS Constraints

Estefanía Recayte*[†], Giuseppe Cocco*, Alessandro Vanelli-Coralli[†]

* Institute of Communications and Navigation, German Aerospace Center (DLR), 82234 Weßling, Germany.

{estefania.recayte, giuseppe.cocco}@dlr.de

[†] University of Bologna, DEIS, Viale del Risorgimento, 2 - 40136 Bologna, Italy.

alessandro.vanelli@unibo.it

Abstract—Small cells will play an important role in the fifth generation mobile networks. As recent works pointed out, a significant improvement in energy efficiency can be obtained if small base stations (SBSs) are provided with storage capabilities. In this paper, we study the problem of caching optimization in the presence of interference with and without cooperation between the SBSs. We consider a setup in which two SBSs and one macro base station (MBS) are connected through a wireless backhaul link. Cooperation is applied following the Han-Kobayashi rate splitting approach and its variation including common information. Our results show that applying cooperation to caching systems yields significant gains in terms of power and provide indications on how much interference can be tolerated which, in turn, has impact on the network design both in terms of frequency reuse planning and SBS deployment.

I. INTRODUCTION

Achieving high spectral efficiency, increasing data rates and volume and decreasing power consumption are some of the main challenges of the next generation cellular networks, the 5G. A key instrument to meet such targets is the deployment of small cells placed a few hundreds of meters one from each other [1], which allows an efficient spatial reuse of the available spectrum. The benefits of this approach may increase if the small base stations (SBSs) serving such small cells have storage capabilities. This would allow the SBSs to store files during periods in which the network traffic is low and directly serve the users during high-load periods, thus sparing backhaul resources. Moreover, the energy consumption of the whole system decreases significantly when contents are placed closer to end users. Many proposed solutions focus on exploiting the statistics of past users file requests and filling up local caches with the most popular contents [2][3]. In [4] the authors aim at reducing the total energy cost in a heterogeneous cellular network by proposing a caching algorithm for multicast transmission. In [5] the

concept of “PHY-caching” is introduced where cooperation between several SBSs creates a MIMO broadcast topology to achieve higher spectral efficiency. In [6], authors study the problem in the backhaul link. They propose a transmission strategy at the MBS and caching policy at the SBS for optimizing the so called *caching gains*. Energy efficiency is also studied in [7], where authors recommend a proactive downloading strategy to exploit the good channel conditions and minimize the total energy.

Unlike previous works, in this paper we study caching strategies in heterogeneous networks. Specifically, the SBSs are equipped with storage capabilities and a minimum required rate for each file is assumed. The SBSs apply different cooperation strategies depending on the considered caching solution. Specifically, when the same file is requested by the two users, the SBSs implement an interference channel with common information which allows to reduce the transmission power. When different files are requested, the Han-Kobayashi approach is applied by splitting each SBS’s message into a common and a private message and then applying superposition. Power and transmission rates are optimized for both in order to allow users to achieve the QoS rate requirements even in the presence of interference. We compare this scenario with a non cooperative approach where caching is optimized considering the interference caused by the other transmitter as noise. We formulate and analyse the caching problem to determinate the cache placement which minimizes the total power consumption both in case of cooperation between SBSs and in case in which no cooperation is in place. Furthermore, our results show that memorizing the most popular files is not always the most convenient strategy when constraints on the minimum required rates are in place. Finally, we show that, already with a small number of files and only two SBSs, the optimization problem is not trivial and leads to interesting and in part counterintuitive results when fading is considered.

The remainder of the paper is organized as follows. In Section II, we introduce the system model. The problem formulation is presented in Section III. In Section IV the minimum power required for the optimization sub-problems are derived. The numerical results are given in Section V. Finally, the conclusions are presented in Section VI.

This work has been presented at IEEE ICC 2017

©2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting /republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

II. SYSTEM MODEL

We consider a heterogeneous 5G network composed by a macro base station (MBS), two small base stations (SBSs) and two users u_n , $n = \{1, 2\}$. We assume one user is connected to each SBS (SBS $_n$ refers the base station of user u_n). SBSs are connected to the MBS through a wireless backhaul link, as shown in Fig. 1. User terminals can communicate with the SBSs and the MBS. The MBS has access to a collection of N files $\mathcal{F} = \{f_1, \dots, f_N\}$. We assume that all files have the same size. This assumption is justifiable in practice since files can be divided into blocks of same size [8]. Each file i has minimum required transmission rate R_i , $i = 1, \dots, N$, i.e., the minimum rate measured in bits/s/Hz at which the file has to be transmitted to the user in order to satisfy the quality of service (QoS)¹ requirements. The probability that file i is requested by a user, usually referred to as *popularity ranking*, [9] is indicated as q_i . We assume that SBS $_n$ is equipped with a local cache of size M_n . We denote with c_S the power cost for transferring a file into a cache through the backhaul link from the MBS to each SBS. Such quantity also accounts for the power cost of storing the file in the SBS. Channel state information at the transmitter (CSIT) is assumed. The communication takes place in two phases. In the first phase, each u_n requests a file to the corresponding SBS $_n$. It is assumed that both SBSs and MBS can listen to the requests of both terminals and the link is error-free. In the second phase, the delivery of the content takes place: user u_n is served by SBS $_n$ if the file required is present in the cache and can be sent at the required transmission rate R_i . If the files are present in the caches of the two SBSs but the channel conditions (interference plus fading) do not allow reliable communication at the required rate, then users are served by the MBS. A user is also served by the MBS if the file requested is not present in the respective SBS.

The idea of having local demand served by SBSs aims at reducing transmissions from the MBS and, in turn, to reduce the overall energy cost and to leave the MBS available to users who are not associated to any SBS. We call $P_i^{(n)}$ the power used by the transmitter for sending f_i to user n .

Our aim is to cache those files which minimize the expected power consumption required to fulfil requests from both users at the required transmission rate. We formulate the optimization problem considering Gaussian codebooks and referring to achievable rates. Two different setups are considered: (i) minimization with cooperation between SBSs and (ii) minimization without cooperation between SBSs. Note that the cooperative approach does not necessarily rely on the knowledge at the SBSs of both requested files, but rather leverages on rate splitting and superposition [10]. In this way, in the cooperative

¹although we consider a constraint in terms of transmission rate over the channel, this can be easily translated into a constraint in terms of delay and file quality which is particularly suited to model video and image transmission.

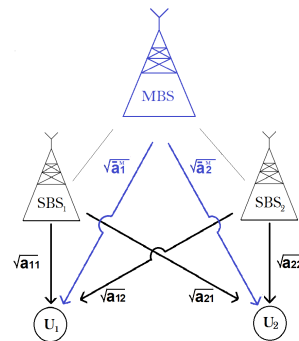


Fig. 1: System model of the heterogeneous network with respective normalized channel coefficient seen by the users. User u_n is associated to both MBS and SBS $_n$.

approach users can satisfy rate requirements even in the presence of interference. To this purpose, it is assumed that transmitters have knowledge of channel coefficients, that can be estimated during the file request phase. In the non cooperative approach, the interference caused by the other transmitter is considered as noise.

Let us introduce a binary variable x_{in} that indicates whether the file f_i is present in the memory of SBS $_n$ ($x_{in} = 1$) or not ($x_{in} = 0$). Without loss of generality, let us assume that f_i is requested by u_1 while f_j is requested by user u_2 (the case when $f_i = f_j$ is considered later on).

Let us denote with $Q(\mathbf{x}_1, \mathbf{x}_2)$ the expected power cost for sending the requested files f_i and f_j to u_1 and u_2 respectively. The expectation is taken over all possible file requests. Depending on the subset of base stations that are transmitting (that is (SBS $_1$, SBS $_2$), (SBS $_n$, MBS) or MBS), the fact that the files requested by u_1 and u_2 are different or not and whether cooperation between transmitters is in place or not, the communication channel can be modelled in the following ways:

- (i) *interference as noise (GIN)*: this channel is considered in the non cooperative case when $x_{i1} = 1$, $x_{j2} = 1$ and $\mathbb{1}_\alpha = 1$. Here, $\mathbb{1}_\alpha = 1$ indicates that it is possible to transmit the files at rates R_i and R_j even in the presence of mutual interference (Fig. 2-b). We will come back on the meaning of $\mathbb{1}_\alpha$ later on in the paper.
- (ii) *interference channel with common information (GICci)*: this channel is considered in the cooperative case when $f_i = f_j$ and $x_{i1} = 1$, $x_{i2} = 1$ (Fig. 2-b).

For the case of Gaussian interference channel it is assumed that the coefficient of the link between u_n and SBS $_n$ is normalized, i.e., $a_{nn} = 1$. The normalization is done as in [10]. Let us consider the physical channel model:

$$y_1 = ax_1 + bx_2 + z_1^*, \quad (1)$$

$$y_2 = cx_1 + dx_2 + z_2^*, \quad (2)$$

where z_1^* and z_2^* are independent Gaussian variables with zero mean and variance N_1 and N_2 , respectively.

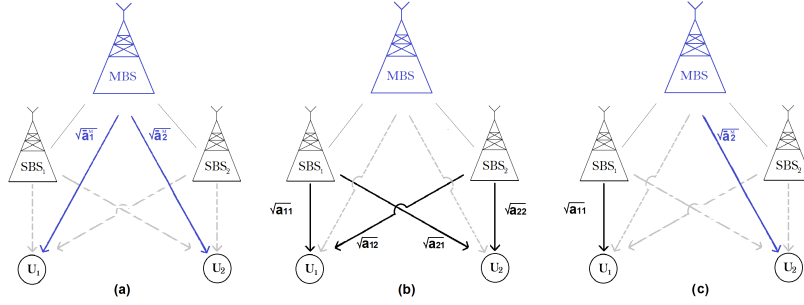


Fig. 2: (a) Broadcast or Multicast Channel, (b) Interference Channel, (c) Orthogonal Channels.

Using the normalization in [10], it can be shown that from the point of view of the achievable rates such model is equivalent to:

$$y_1 = \sqrt{a_{11}}x_1 + \sqrt{a_{12}}x_2 + z_1, \quad (3)$$

$$y_2 = \sqrt{a_{21}}x_1 + \sqrt{a_{22}}x_2 + z_2, \quad (4)$$

where $a_{11} = 1$, $a_{22} = 1$, $a_{12} = b^2 N_2 / d^2 N_1$, $a_{21} = c^2 N_1 / a^2 N_2$ while z_1 and z_2 are independent Gaussian additive noise with zero mean and unitary variance. Denoting with P_{x_1} (P_{x_2}) the physical power of x_1 (x_2) then in the new equivalent system the power P'_{x_n} of x_n is scaled as follow: $P'_{x_1} = a^2 P_{x_1} / N_1$ and $P'_{x_2} = d^2 P_{x_2} / N_2$.

- (iii) *interference channel without common information (GICnc)*: this channel is considered in the cooperative approach when $f_i \neq f_j$ and the two files are stored in the respective SBSs ($x_{i1} = 1$, $x_{j2} = 1$) (Fig. 2-b).
- (iv) *orthogonal channel*: such channel occurs when u_1 (u_2) request is satisfied by the SBS and u_2 (u_1) request is satisfied by the MBS ($x_{i1} = 1$, $x_{j2} = 0$) ($x_{i1} = 0$, $x_{j2} = 1$), (Fig. 2-c). In this case, given the physical channel model:

$$y_1 = ax_1 + z_1^*,$$

$$y_2 = dx_2 + z_2^*,$$

the following normalized model can be derived:

$$y_1 = \sqrt{\bar{a}_1}x_1 + z_1,$$

$$y_2 = \sqrt{\bar{a}_2}x_2 + z_2,$$

where $\bar{a}_1 = a^2 / N_1$ and $\bar{a}_2 = d^2 / N_2$ while z_1 and z_2 are independent Gaussian additive noise with zero mean and unitary variance. Note that since channels are orthogonal they do not produce mutual interference.

- (v) *broadcast channel*: this type of channel is considered when $f_i \neq f_j$ in the following cases, (i) user terminals are served by the MBS because files are not present in SBSs caches ($x_{i1} = 0$, $x_{j2} = 0$), (ii) in the non cooperative approach when ($x_{i1} = 1$, $x_{j2} = 1$, $\mathbb{1}_\alpha = 0$), where $\mathbb{1}_\alpha$ indicates that it is not possible to transmit the files at rates R_i and R_j because of the mutual interference (Fig. 2-a).

In the broadcast and the multicast Gaussian channel we have the following physical channel model

$$y_1 = c_1^M x_1 + z_1^*,$$

$$y_2 = c_2^M x_2 + z_2^*,$$

where we denote the channel coefficient between MBS and u_n as c_n^M and is equivalent to the normalized model:

$$y_1 = \sqrt{\bar{a}_1^M} x_1 + z_1,$$

$$y_2 = \sqrt{\bar{a}_2^M} x_2 + z_2,$$

where $\bar{a}_1^M = (c_1^M)^2 / N_1$ and $\bar{a}_2^M = (c_2^M)^2 / N_2$. In the following, we denote $\bar{a}_+^M = \max\{\bar{a}_1^M, \bar{a}_2^M\}$, $\bar{a}_-^M = \min\{\bar{a}_1^M, \bar{a}_2^M\}$ and R_+ (R_-) the rate related to the file required from the user with MBS channel coefficient \bar{a}_+^M (\bar{a}_-^M). Similarly P_+ and P_- are defined.

- (vi) *multicast channel*: this channel is considered when $f_i = f_j$ in the following cases: (i) requests are satisfied by the MBS ($x_{i1} = 0$, $x_{j2} = 0$), in this case, since both users require the same file, the MBS sends the file at the level of power needed to satisfy the u_n experiencing the worst channel condition and (ii) in the non cooperative approach when ($x_{i1} = 1$, $x_{j2} = 1$, $\mathbb{1}_\alpha = 0$) (Fig. 2-a).

III. OPTIMIZATION PROBLEM

A. Cooperative Case

Let us denote with \mathbf{x}_1 and \mathbf{x}_2 the allocation vector containing the values for x_{i1} and x_{i2} , $i = 1, \dots, N$ respectively. Then, the expected power cost in the cooperative case $Q_c(\mathbf{x}_1, \mathbf{x}_2)$ can be written as:

$$Q_c(\mathbf{x}_1, \mathbf{x}_2) = \sum_{f_i \in \mathcal{F}} \sum_{\substack{f_j \in \mathcal{F} \\ f_i \neq f_j}} q_i \cdot q_j \cdot [(1 - x_{i1})(1 - x_{j2}) \cdot c_{broad} + x_{i1}x_{j2} \cdot c_{GIC}^{nc} + (1 - x_{j2})x_{i1} \cdot c_{orth} + (1 - x_{i1})x_{j2} \cdot c_{orth}] + \sum_{f_i \in \mathcal{F}} q_{ii}^2 \cdot [(1 - x_{i1})(1 - x_{i2}) \cdot c_{multic} + x_{i1}x_{i2} \cdot c_{GIC}^{ci} + (1 - x_{i2})x_{i1} \cdot c_{orth} + (1 - x_{i1})x_{i2} \cdot c_{orth}],$$

where c_{broad} , c_{GIC}^{nc} , c_{GIC}^{ci} , c_{orth} , c_{multic} , denote the minimum power consumption ($P_i^{(1)} + P_j^{(2)}$) required in the

broadcast, GICnc, GINci, orthogonal and multicast channels, respectively. Denoting the Shannon capacity for a point to point channel with $C(x) = \frac{1}{2} \log_2(1+x)$, the minimization problem for the cooperative case can be formulated as follows:

$$\underset{P^{(1)}, P^{(2)}, \mathbf{x}_1, \mathbf{x}_2}{\text{minimize}} \quad \frac{c_S}{k} + Q_c(\mathbf{x}_1, \mathbf{x}_2), \quad (5)$$

$$\text{subject to: } \sum_{f_i \in \mathcal{F}} x_{in} \leq M_n, n = \{1, 2\}, \quad (6)$$

$$(1-x_{i1})(1-x_{j2})R_- \leq C\left(\frac{\bar{a}_-^M P_-}{1+\bar{a}_-^M P_+}\right), f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (7)$$

$$(1-x_{i1})(1-x_{j2})R_+ \leq C(\bar{a}_+^M P_+), f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (8)$$

$$(1-x_{i1})(1-x_{j2})R_i \leq C(\bar{a}_1^M P_i^{(1)}), f_i, f_j \in \mathcal{F}, f_i = f_j, \quad (9)$$

$$(1-x_{i1})(1-x_{j2})R_i \leq C(\bar{a}_2^M P_j^{(2)}), f_i, f_j \in \mathcal{F}, f_i = f_j, \quad (10)$$

$$(1-x_{j2}) \cdot x_{i1} \cdot R_i \leq C(\bar{a}_1 \cdot P_i^{(1)}), f_i, f_j \in \mathcal{F}, \quad (11)$$

$$(1-x_{j2}) \cdot x_{i1} \cdot R_j \leq C(\bar{a}_2^M \cdot P_j^{(2)}), f_i, f_j \in \mathcal{F}, \quad (12)$$

$$(1-x_{i1}) \cdot x_{j2} \cdot R_i \leq C(\bar{a}_1^M \cdot P_i^{(1)}), f_i, f_j \in \mathcal{F}, \quad (13)$$

$$(1-x_{i1}) \cdot x_{j2} \cdot R_j \leq C(\bar{a}_2 \cdot P_j^{(2)}), f_i, f_j \in \mathcal{F}, \quad (14)$$

$$x_{i1} \cdot x_{j2} \cdot R_i \leq \rho_1, f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (15)$$

$$x_{i1} \cdot x_{j2} \cdot R_j \leq \rho_2, f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (16)$$

$$x_{i1} \cdot x_{j2} \cdot (R_i + R_j) \leq \rho_{12}, f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (17)$$

$$x_{i1} \cdot x_{j2} \cdot (2R_i + R_j) \leq \rho_{10}, f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (18)$$

$$x_{i1} \cdot x_{j2} \cdot (R_i + 2R_j) \leq \rho_{20}, f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (19)$$

$$x_{i1}x_{j2}R_i \leq C\left(\left(\sqrt{P_i^{(1)}} + \sqrt{a_{12}P_j^{(2)}}\right)^2\right), f_i, f_j \in \mathcal{F}, f_i = f_j, \quad (20)$$

$$x_{i1}x_{j2}R_i \leq C\left(\left(\sqrt{P_j^{(2)}} + \sqrt{a_{21}P_i^{(1)}}\right)^2\right), f_i, f_j \in \mathcal{F}, f_i = f_j, \quad (21)$$

$$P_i^{(1)} > 0, f_i \in \mathcal{F}, \quad (22)$$

$$P_j^{(2)} > 0, f_j \in \mathcal{F}, \quad (23)$$

$$x_{in} \in \{0, 1\}, n = \{1, 2\}, f_i \in \mathcal{F}, \quad (24)$$

where the first term in the objective function is the power cost associated with the storage operation while the second is the servicing cost. The caching cost c_S is the initial cost of the storage process and accounts for the file transfer from the MBS to an SBS plus the power consumed to store the file in the SBS. Denoting with k the total number of requests between two successive storage process then each request costs c_S/k of the initial cost. Since c_S is constant and file independent, it will be neglected in the calculations.

Inequality in (6) ensures that the total amount of data stored in a cache will not exceed its size, (7)-(8) denote the capacity region of the Gaussian broadcast channel, (9)-(10) denote the capacity region of the multicast channel, (11)-(14) represent the conditions for the case of orthogonal channels, (15)-(19) are the Han-Kobayashi constraints

[10] and define an achievable rate of the GICnc (the explicit expressions for such constraints are given in the Appendix), (20)-(21) define an achievable rate region of the GICci when both transmitters have to transmit the same file and (22)-(23) are imposed to guarantee the non negativity of the transmit power while (24) represent the caching choice and account for the discrete nature of the optimization variable. Note that the power variables in constraints (15)-(21) refer to the normalized channel model $(P_i^{(1)}, P_j^{(2)})$. Since we are interested in evaluating the physical power consumption, then the power values in the calculation of the cost have to be scaled as $P_i^{(1)} = P_i^{(1)}/\bar{a}_1$ and $P_j^{(2)} = P_j^{(2)}/\bar{a}_2$.

B. Non cooperative case

Let us introduce the indication function $\mathbb{1}_\alpha$ defined as:

$$\mathbb{1}_\alpha := \begin{cases} 0 & \alpha \leq 0, \\ 1 & \alpha > 0, \end{cases}$$

where $\alpha = 1 + a_{12} \cdot a_{21} \cdot (1 - 2^{2R_i}) \cdot (2^{2R_j} - 1)$. It can be easily shown that when $\mathbb{1}_\alpha = 1$ the SBSs are able to successfully deliver their files to the own users with high probability. Instead, $\mathbb{1}_\alpha = 0$ means that do not exist $P_i^{(1)}$ and $P_j^{(2)}$ such that the minimum required rate can be achieved. In such case, the SBSs are not able to serve their users and MBS will serve both users either applying broadcast (if files are different) or multicast transmission (if files are the same).

The expected power cost in the non cooperative case $Q_{nc}(\mathbf{x}_1, \mathbf{x}_2)$ can be written as:

$$\begin{aligned} Q_{nc}(\mathbf{x}_1, \mathbf{x}_2) = & \sum_{f_i \in \mathcal{F}} \sum_{\substack{f_j \in \mathcal{F} \\ f_i \neq f_j}} q_i \cdot q_j \cdot [(1-x_{i1})(1-x_{j2}) \cdot c_{broad} + \\ & + \mathbb{1}_\alpha x_{i1}x_{j2}c_{GIN} + (1-x_{j2})x_{i1} \cdot c_{orth} + (1-x_{i1})x_{j2} \cdot c_{orth} + \\ & + (1-\mathbb{1}_\alpha)x_{i1}x_{j2} \cdot c_{broad} + (1-x_{i1})(1-x_{j2}) \cdot c_{broad}] \\ & + \sum_{f_i \in \mathcal{F}} q_{ii}^2 \cdot [(1-x_{i1})(1-x_{i2}) \cdot c_{multic} + \mathbb{1}_\alpha x_{i1}x_{i2} \cdot c_{GIN} + \\ & + (1-\mathbb{1}_\alpha)x_{i1}x_{i2} \cdot c_{multic} + (1-x_{i2})x_{i1} \cdot c_{orth} + \\ & + (1-x_{i1})x_{i2} \cdot c_{orth}]. \end{aligned}$$

The optimization problem in the non cooperative case can

be written as:

$$\min_{P^{(1)}, P^{(2)}, \mathbf{x}_1, \mathbf{x}_2} \frac{c_S}{k} + Q_{nc}(\mathbf{x}_1, \mathbf{x}_2), \quad (25)$$

subject to: (6)-(14)

$$\mathbb{1}_\alpha x_{i1} x_{j2} \cdot R_i \leq C \left(\frac{P_i^{(1)}}{1 + a_{12} P_j^{(2)}} \right), f_i, f_j \in \mathcal{F}, f_i, f_j, \quad (26)$$

$$\mathbb{1}_\alpha x_{i1} x_{j2} \cdot R_j \leq C \left(\frac{P_j^{(2)}}{1 + a_{21} \cdot P_i^{(1)}} \right), f_i, f_j \in \mathcal{F}, f_i, f_j, \quad (27)$$

$$(1 - \mathbb{1}_\alpha) x_{i1} x_{j2} \cdot R_- \leq C \left(\frac{\bar{a}_-^M \cdot P_-}{1 + \bar{a}_-^M \cdot P_+} \right), f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (28)$$

$$(1 - \mathbb{1}_\alpha) x_{i1} x_{j2} \cdot R_+ \leq C (\bar{a}_+^M \cdot P_+), f_i, f_j \in \mathcal{F}, f_i \neq f_j, \quad (29)$$

$$(1 - \mathbb{1}_\alpha) x_{i1} x_{j2} \cdot R_i \leq C (\bar{a}_+^M \cdot P_i^{(1)}), f_i, f_j \in \mathcal{F}, f_i = f_j, \quad (30)$$

$$(1 - \mathbb{1}_\alpha) x_{i1} x_{j2} \cdot R_i \leq C (\bar{a}_-^M \cdot P_j^{(2)}), f_i, f_j \in \mathcal{F}, f_i = f_j, \quad (31)$$

$$P_i^{(1)} > 0, f_i \in \mathcal{F}, \quad (32)$$

$$P_j^{(2)} > 0, f_j \in \mathcal{F}, \quad (33)$$

$$x_{in} \in \{0, 1\}, n = \{1, 2\}, f_i \in \mathcal{F}. \quad (34)$$

Conditions (6)-(14) are in common with the cooperative approach. Instead, conditions (28)-(29) denote the capacity region of the Gaussian broadcast channel, (30)-(31) the capacity region of the multicast channel. As mentioned previously, these channels are considered only when the SBSs are not capable to deliver the files at the requested rates. Conditions (26)-(27) represent the constraints of the GIN channel.

IV. POWER MINIMIZATION SUB-PROBLEMS

In this section we minimize the overall power $P_i^{(1)} + P_j^{(2)}$ for each channel required to transmit f_i to u_i and f_j to u_j at rate R_i and R_j , respectively.

A. Broadcast Channel

$$\begin{aligned} & \min_{P_-, P_+} P_- + P_+, \\ & \text{s.t.: } \frac{1}{2} \log_2 (1 + \bar{a}_+^M P_+) \geq R_+, \\ & \frac{1}{2} \log_2 \left(1 + \frac{\bar{a}_-^M P_-}{1 + \bar{a}_-^M P_+} \right) \geq R_-, \\ & P_-, P_+ \geq 0. \end{aligned} \quad (35)$$

Can be solved as follow:

$$P_+ \geq \frac{2^{2R_+} - 1}{\bar{a}_+^M}, \quad (36)$$

and:

$$P_- \geq \frac{(2^{2R_-} - 1)(1 + \bar{a}_-^M P_+)}{\bar{a}_-^M}.$$

Plugging P_+ in the expression of P_- we obtain $P_- = \frac{1}{\bar{a}_-^M} (2^{2R_-} - 1) \left(1 + \frac{\bar{a}_-^M}{\bar{a}_+^M} (2^{2R_+} - 1) \right)$.

B. Multicast channel

In the multicast channel the same file f_i has to be transmitted to both users at rate R_i . The problem to be solved is:

$$\begin{aligned} & \min_{P_i} P_i, \\ & \text{s.t.: } \frac{1}{2} \log_2 (1 + \bar{a}_1^M P_i) \geq R_i, \\ & \frac{1}{2} \log_2 (1 + \bar{a}_2^M P_i) \geq R_i, \\ & P_i \geq 0. \end{aligned} \quad (37)$$

It can be easily shown that the minimum power required for u_1 is $P_i^{(1)} = \frac{2^{2R_i} - 1}{\bar{a}_1^M}$ while for u_2 is $P_i^{(2)} = \frac{2^{2R_i} - 1}{\bar{a}_2^M}$. Thus we have:

$$\begin{aligned} P_i &= \max \left\{ \frac{2^{2R_i} - 1}{\bar{a}_1^M}, \frac{2^{2R_i} - 1}{\bar{a}_2^M} \right\} \\ &= \frac{2^{2R_i} - 1}{a_-^M}. \end{aligned} \quad (38)$$

C. Orthogonal channel

This channel is considered when one user is connected to the respective SBS while the other is connected to the MBS. Let us assume, without loss of generality, that u_1 is connected to the SBS₁ and u_2 to the MBS.

$$\begin{aligned} & \min_{P^{(1)}, P^{(2)}} P_i + P_j, \\ & \text{s.t.: } \frac{1}{2} \log_2 (1 + \bar{a}_1 P_i^{(1)}) \geq R_i, \\ & \frac{1}{2} \log_2 (1 + \bar{a}_2 P_j^{(2)}) \geq R_j, \\ & P^{(1)}, P^{(2)} \geq 0, \end{aligned} \quad (39)$$

where the solution is $P_i^{(1)} = \frac{2^{2R_i} - 1}{\bar{a}_1}$ and $P_j^{(2)} = \frac{2^{2R_j} - 1}{\bar{a}_2}$.

D. Interference as noise

$$\begin{aligned} & \min_{P^{(1)}, P^{(2)}} P_i + P_j, \\ & \text{s.t.: } \frac{1}{2} \log_2 \left(1 + \frac{P_i^{(1)}}{1 + a_{12} \cdot P_j^{(2)}} \right) \geq R_i, \\ & \frac{1}{2} \log_2 \left(1 + \frac{P_j^{(2)}}{1 + a_{21} \cdot P_i^{(1)}} \right) \geq R_j, \\ & P^{(1)}, P^{(2)} \geq 0, \end{aligned} \quad (40)$$

where $P_i^{(1)}$ is minimum when:

$$P_i^{(1)} = (2^{2R_i} - 1) \cdot (a_{12} \cdot P_j^{(2)} + 1), \quad (41)$$

and solving in $P_j^{(2)}$ we have that:

$$P_j^{(2)} \geq \frac{(2^{2R_j} - 1) \cdot (a_{21} \cdot (2^{2R_i} - 1) + 1)}{1 + a_{21} \cdot a_{12} \cdot (2^{2R_j} - 1) \cdot (1 - 2^{2R_i})}. \quad (42)$$

Note that inequality (42) does not always have a feasible solution since $P_j^{(2)}$ should be positive. Specifically, for this

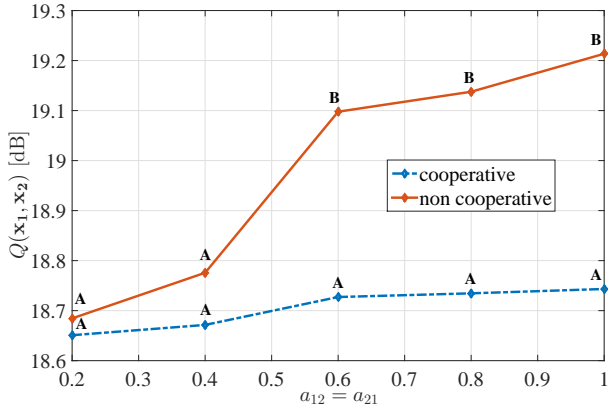


Fig. 3: Minimum average power cost in dB in the AWGN case plotted against the interference coefficient. The optimization is done for each average interference coefficient. In the simulation we set $\bar{a}_1^M = \bar{a}_2^M = 0.01$. “A” points represent the memory allocation: $M_1 = \{f_1, f_3\}, M_2 = \{f_1, f_3\}$ while “B” points represent $M_1 = \{f_2, f_3\}, M_2 = \{f_2, f_3\}$.

type of channel transmission can be successful only if: $1 + a_{21} \cdot a_{12} \cdot (2^{2R_j} - 1) \cdot (1 - 2^{2R_i}) > 0$ otherwise files are delivered by the MBS through a the broadcast channel.

Solving analytically the power minimization sub-problems for the case of GICnc and GICci is a challenging task due to the intricacy of the conditions (see formulation of $Qc(\mathbf{x}_1, \mathbf{x}_2)$ and Appendix), and thus they are addressed numerically.

V. NUMERICAL RESULTS

We consider the system depicted in Fig.1 where each SBS has a memory of size $M_n = 2$, the set of files is $\mathcal{F} = \{f_1, f_2, f_3\}$ and the rates of the files are $R_1 = 1.20$ bits/s/Hz, $R_2 = 0.40$ bits/s/Hz and $R_3 = 0.60$ bits/s/Hz. The probability that files f_1, f_2, f_3 are requested are $q_1 = 0.15$, $q_2 = 0.35$, $q_3 = 0.5$, respectively.

We consider two different link conditions, namely AWGN and fading. In the AWGN case the channel coefficients are fixed, while in the fading case we assume that each link is affected by an independent block fading process with Rayleigh distributed channel coefficients. Powers are exponentially distributed with mean values $\mathbb{E}\{\bar{a}_1\} = 1$, $\mathbb{E}\{\bar{a}_2\} = 1$, $\mathbb{E}\{\bar{a}_1^M\} = \mathbb{E}\{\bar{a}_2^M\} = 0.01$ while $\mathbb{E}\{\bar{a}_{12}\} = \mathbb{E}\{\bar{a}_{21}\}$ take values in $[0, 1]$, $\mathbb{E}\{\cdot\}$ being the expectation operator. Studying the behaviour of the system at different interference levels mutually induced by two SBSs can give hints on how to efficiently plan the frequency reuse and the SBS distance in next generation mobile network.

We start considering the system in the AWGN case. In Fig. 3 the minimum power cost $Q(\mathbf{x}_1, \mathbf{x}_2)$ is plotted versus the interference coefficients (assumed to be the same for both users) in an AWGN environment. Note that the optimization is done for each interference level. The letter “A” indicates that the optimal file caching for that specific interference level in the two SBSs are $\{f_1, f_3\}$ and $\{f_1, f_3\}$

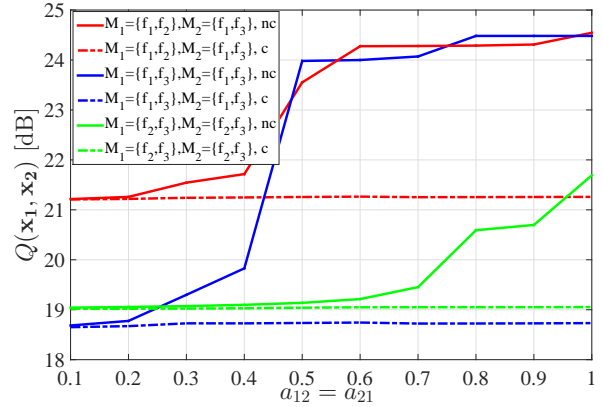


Fig. 4: Average power cost in dB plotted versus interference coefficients for different memory allocations in an AWGN environment $\bar{a}_1^M = \bar{a}_2^M = 0.01$.

while the letter “B” indicates the storage solution $\{f_2, f_3\}$ $\{f_2, f_3\}$. In order to see the behaviour of each cache allocation for different interference levels, in Fig. 4 the power cost $Q(\mathbf{x}_1, \mathbf{x}_2)$ versus the interference coefficients for the same channel condition is plotted for three different memory allocations. The cooperative case is represented with dashed curves while the non cooperative approach is represented with solid curves. It is interesting to see how the cooperative approach copes well with increasing interference levels, showing only a slight increase in the required power. On the contrary, in the non cooperative approach each memory allocation presents a critical interference value for which the power cost significantly increases. This is due to the fact that in correspondence with such points the parameter α introduced in Section III-A becomes negative, which implies that it is not possible for the SBSs to reliably transmit the files over the GIN channel at the required rates, and the system is forced to use the MBS. Interestingly, the plot shows also that the most power-efficient cache allocation (blue curve) is not the one including the files with the highest popularity levels (green curve).

In Fig. 5 the minimum power cost $\bar{Q}(\mathbf{x}_1, \mathbf{x}_2) = \mathbb{E}\{Q(\mathbf{x}_1, \mathbf{x}_2)\}$ in the case of fading channels is shown, where the average is taken over the fading distribution. As in the AWGN case, we note that different average interference levels have a different optimal file allocation. However, unlike the AWGN case, the curves are not monotonically increasing. In order to understand the reason for such behavior, in Fig. 6 the minimum average cost $\bar{Q}(\mathbf{x}_1, \mathbf{x}_2) = \mathbb{E}\{Q(\mathbf{x}_1, \mathbf{x}_2)\}$ for three different memories allocation is plotted for different average interference levels in a fading environment. Also here it can be seen that the curves are not monotonically increasing. This is due partly to the different rates and file popularity levels and partly to the different interference configuration generated by the random channel fluctuations, which determine radically different system behaviors. In particular, the fact that one or both interference links are stronger than the correspondent direct link appears to play a major role. This is in line

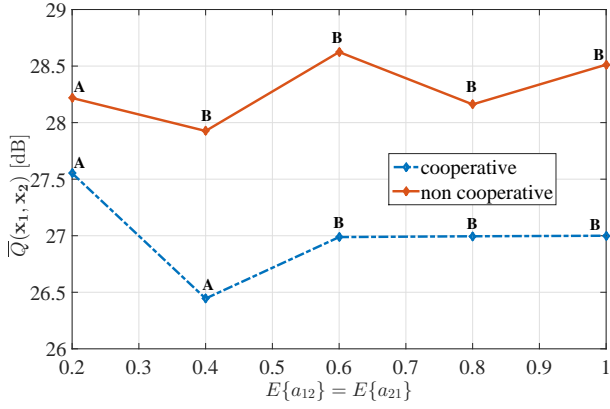


Fig. 5: Minimum average power cost $\bar{Q}(\mathbf{x}_1, \mathbf{x}_2) = \mathbb{E}\{Q(\mathbf{x}_1, \mathbf{x}_2)\}$ in dB in the case of fading plotted against average interference coefficient. The optimization is done for each average interference coefficient. $\mathbb{E}\{\bar{a}_1^M\} = \mathbb{E}\{\bar{a}_2^M\} = 0.01$. “A” points represent the memory allocation: $M_1 = \{f_1, f_3\}, M_2 = \{f_1, f_3\}$ while “B” points represent $M_1 = \{f_2, f_3\}, M_2 = \{f_2, f_3\}$.

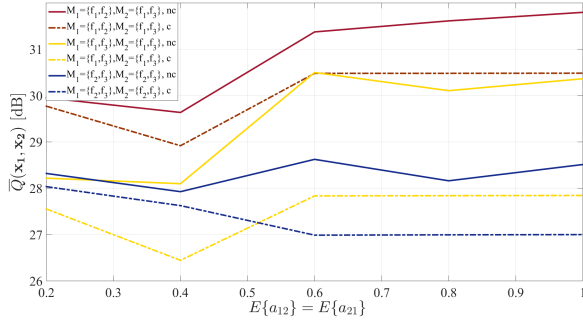


Fig. 6: Average power cost in dB in function of interference channels coefficient for different memory allocations in a fading environment. $\mathbb{E}\{\bar{a}_1^M\} = \mathbb{E}\{\bar{a}_2^M\} = 0.01$.

with what presented in [10], where a strong dependence of the capacity region from the relative interference levels is shown. To confirm such dependence, in Fig. 7, the average minimum power required for the SBSs to serve their users in a GIC without common information is shown. The same average channel coefficients as in Fig. 6 were used. It can be seen that the average power consumption does not increase monotonically with the interference level.

VI. CONCLUSION

We studied the problem of caching for power minimization in heterogeneous networks under quality of service constraints. We formulated and analysed the caching problem with the aim to minimize the total power consumption both in case of cooperation between SBSs as well as in the non cooperative case. Our results highlight that storing the most popular files is not always the most convenient strategy when transmitters have to respect rate requirements. We showed that already with a small number of files and only two SBSs each with only one associated user, the optimization problem is not trivial and

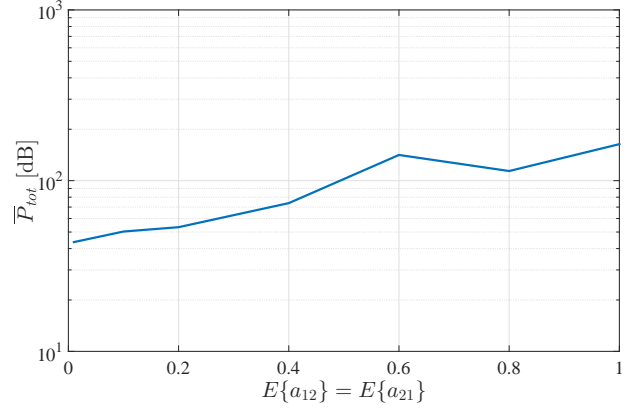


Fig. 7: Minimum power required at different level of interference for GIC without common information in fading environment. $R_i = 0.6$ bits/sec/Hz and $R_j = 1.2$ bits/sec/Hz were considered.

leads to interesting results. Surprisingly, we found that in the case of fading the minimum required power does not increase monotonically with the interference level, which is partly due to the fact that power fluctuations induced by the fading lead to different GIC configurations (strong, weak and intermediate interference channels) that present very different behaviors.

APPENDIX

$$\begin{aligned} \rho_1 &= \sigma_1^* + I(Y_1; u_1 | w_1 w_2), \\ \rho_2 &= \sigma_2^* + I(Y_2; u_2 | w_1 w_2), \\ \rho_{12} &= \sigma_{12} + I(Y_1; u_1 | w_1 w_2) + I(Y_2; u_2 | w_1 w_2), \\ \rho_{10} &= 2\sigma_1^* + 2I(Y_1; u_1 | w_1 w_2) + I(Y_2; u_2 | w_1 w_2) + \\ &\quad - [\sigma_1^* - I(Y_2; w_1 | w_2)]^+ + \min\{I(Y_2; w_2 | w_1), \end{aligned}$$

$$\begin{aligned} & I(Y_2; w_2) + [I(Y_2; w_1 | w_2) - \sigma_1^*]^+, I(Y_1; w_2 | w_1), \\ & I(Y_1; w_1 w_2) - \sigma_1^* \}, \\ \rho_{20} &= 2\sigma_2^* + 2I(Y_2; u_2 | w_1 w_2) + I(Y_1; u_1 | w_1 w_2) + \\ & - [\sigma_2^* - I(Y_1; w_2 | w_1)]^+ + \min\{I(Y_1; w_1 | w_2), \\ & I(Y_1; w_1) + [I(Y_1; w_2 | w_1) - \sigma_2^*]^+, I(Y_2; w_1 | w_1), \\ & I(Y_2; w_1 w_2) - \sigma_1^* \} \\ \sigma_1^* &= \min\{I(Y_1; w_1 | w_2), I(Y_2; w_1 | w_2 w_2)\} \\ \sigma_2^* &= \min\{I(Y_2; w_1 | w_1), I(Y_1; w_1 | w_1 w_1)\} \\ \sigma_{12} &= \min\{I(Y_1; w_1 w_2), I(Y_2; w_1 w_2); I(Y_1; w_1 | w_2) + \\ & I(Y_2; w_2 | w_1), I(Y_2; w_1 | w_2) + I(Y_1; w_2 | w_1)\}, \end{aligned}$$

where:

$$\begin{aligned}
I(Y_1; u_1 | w_1 w_2) &= C(\lambda_1 P'_1 / (1 + a_{12} \lambda_2 P'_2)), \\
I(Y_2; u_2 | w_1 w_2) &= C(\lambda_2 P'_2 / (1 + a_{21} \lambda_1 P'_1)), \\
I(Y_1; w_1 | w_2) &= C(\bar{\lambda}_1 P'_1 / (1 + \lambda_1 P'_1 + a_{12} \lambda_2 P'_2)), \\
I(Y_1; w_2 | w_1) &= C(a_{12} \bar{\lambda}_2 P'_2 / (1 + \lambda_1 P'_1 + a_{12} \lambda_2 P'_2)), \\
I(Y_1; w_1 w_2) &= C((\bar{\lambda}_1 P'_1 + a_{12} \bar{\lambda}_2 P'_2) / (1 + \lambda_1 P'_1 + a_{12} \lambda_2 P'_2)), \\
I(Y_2; w_2 | w_1) &= C(\bar{\lambda}_2 P'_2 / (1 + \lambda_2 P'_2 + a_{21} \lambda_1 P'_1)), \\
I(Y_2; w_1 | w_2) &= C(a_{21} \bar{\lambda}_1 P'_1 / (1 + \lambda_2 P'_2 + a_{21} \lambda_1 P'_1)), \\
I(Y_2; w_1 w_2) &= C((\bar{\lambda}_2 P'_2 + a_{21} \bar{\lambda}_1 P'_1) / (1 + \lambda_2 P'_2 + a_{21} \lambda_1 P'_1)), \\
I(Y_1; w_1) &= C(\bar{\lambda}_1 P'_1 / (1 + \lambda_2 P'_2 + a_{21} \lambda_1 P'_1)), \\
I(Y_2; w_2) &= C(\bar{\lambda}_2 P'_2 / (1 + \lambda_1 P'_1 + a_{12} \lambda_2 P'_2)), \\
I(Y_1; w_2 | u_1 w_1) &= C(a_{12} \bar{\lambda}_2 P'_2 / (1 + a_{12} \lambda_2 P'_2)), \\
I(Y_2; w_1 | u_2 w_2) &= C(a_{21} \bar{\lambda}_1 P'_1 / (1 + a_{21} \lambda_1 P'_1)).
\end{aligned}$$

REFERENCES

- [1] L. Wei, R. Q. Hu, Y. Qian, G. Wu. "Key Elements to Enable Millimeter Wave Communications for 5G Wireless System". *IEEE Wireless Comm.*, pp. 136-143, Dec. 2014.
- [2] G. Paschos, E. Bastug, I. Lang, G. Caire, M. Debbah, "Wireless Caching: Technical Misconceptions and Business Barriers", *IEEE Comm. Magazine*, vol. 54, pp. 16-22, Aug. 2016.
- [3] E. Bastug, J.-L. Guenego, and M. Debbah, "Proactive small cell networks," in *Proc. of the IEEE Int. Conf. on Telecom, Casablanca, Morocco*, pp. 1-5, May 2013.
- [4] K. Poularakis, G. Iosifidis, V. Sourlas, L. Tassiulas, "Exploiting Caching and Multicast for 5G Wireless Networks", *IEEE Trans. on Wireless Comm.*, vol. 15, pp. 2995-3007, Apr. 2016.
- [5] W. Han, A. Liu, V. Lau, "PHY-Caching in 5G Wireless Networks: Design and Analysis", *IEEE Comm. Magazine*, vol. 54, pp. 30-36, Aug. 2016.
- [6] M. Gregori, J. Gómez-Vilarbedò, J. Matamoros. "Joint Transmission and Caching Policy Design for Energy Minimization in the Wireless Backhaul Link", in *Proc. IEEE Int. Symp. on Inf. Theory (ISIT)*, Hong Kong, China, 2015, pp. 1004-1008.
- [7] A. Can Güngör, D. Gündüz. "Proactive wireless caching at mobile user devices for energy efficiency", *Proc. IEEE Int. Symp. Wireless Commun. Syst. (ISWCS)*, 2015.
- [8] L. Shanmugan, N. Golrezaei, A. Dimakis, A. Molisch, G. Caire, "FemtoCaching: Wireless Video Content Delivery through Distributed Caching Helpers", in *Proc. IEEE Transactions of Inf. Theory*, vol. 59, no. 12, pp. 8402-8413, Dec. 2013.
- [9] L. Breslau, "Web Caching and Zipf-like Distributions: Evidence and implications," in *Proc. IEEE INFOCOM*, vol. 1, 1999, pp. 126-134.
- [10] T. Han, K. Kobayashi. "A New Achievable Rate Region for the Interference Channel." *IEEE Trans. on Inf. Theory*, vol. 52, pp. 1176-1184, Mar. 2006.