Self-Optimizing Energy Management in Heterogeneous Cellular Networks

Majid Ghaderi and Mohammad Naghibi
Department of Computer Science, University of Calgary
{mghaderi,snaghibi}@cs.ucalgary.ca

Abstract—**In this paper, we develop and evaluate a distributed algorithm to efficiently balance the trade-off between network throughput and energy consumption in a heterogeneous cellular network. We formulate the problem as a joint optimization of base station activation, power control and user association. To solve the problem, which is a non-convex optimization problem, we design a self-optimizing algorithm based on Gibbs sampling in which each base station individually optimizes its configuration without the involvement of any central controller. In our algorithm, base stations only need to exchange information in a locally defined neighborhood, yet the network state eventually converges to the global optimal. Simulation results are also provided, which show that, i) the proposed algorithm indeed converges to a state that is close to optimal, and ii) by dynamically activating base stations, we see about 10% reduction in network energy consumption without penalizing the network throughput.**

I. Introduction

A. Motivation

Heterogeneous cellular networks (HetNets) are introduced as a solution to cope with the ever-rising data demand, which is expected to increase by tenfold from 2014 to 2019 [1], [2]. In HetNets, in addition to the traditional macrocells, a large number of low-power small-cells, such as picocells, are deployed in geographic areas characterized by high traffic or poor coverage [3]. The increased capacity brought about by HetNets relies on the deployment of more base stations (BSs), specifically a large number of small-cells, which creates several problems for network operators. First, it increases the network management complexity. With many more base stations to manage [4], it is cumbersome and inefficient to manually set up and optimize the network. Automatic and intelligent mechanisms are needed to control network operations such as interference coordination and power control in a *self-optimizing* manner [5].

Second, deploying many small-cells imposes higher capital and operational expenditure on mobile network operators. In particular, increasing the number of base stations, increases the network energy consumption. As reported in [6] and [7], base stations account for 60-80% of the total network energy consumption. Current base stations are not *energy-proportional* and consume about 50-90% of their peak energy even in idle or low traffic state [8]. A notable observation is that while mobile traffic exhibits periodic behavior, the energy consumption of a cellular network remains approximately the same [9]. This can be attributed to the fact that cellular operators often deploy as many base stations as necessary to satisfy the peak traffic demand. Powering off underutilized base stations when their traffic load could be handled by nearby base stations can help in reducing the network energy consumption.

Third, when deploying both macrocells and small-cells, the strong signals from adjacent macrocells may overpower small-cell signals. In order for macro and small cells to coexist peacefully, some form of *interference coordination* [10] is required to protect small-cell users from high power macrocell transmissions. To this end, the concept of Almost Blank Subframes (ABS) is introduced in LTE-Advanced networks [11]. The idea of ABS subframes is to dedicate a portion of each radio frame for small-cell transmissions, and have the macrocells fully (or partially) restrained from transmitting data signals during those subframes to reduce the interference on small-cell users.

B. Our Work

In this paper, we consider a HetNet composed of macrocells and picocells. Picocells are small-cells typically with ranges from a few hundred meters to two kilometers [3]. Pico base stations are capable of entering the sleep mode in which they consume very little energy [12]. In active mode, pico base stations transmit at their maximum power, while in sleep mode they do not serve any user. Macro base stations, on the other hand, are always active to provide some level of coverage in the network even in areas where picocells are switched to sleep mode. However, they are capable of transmitting at various power levels (by means of ABS subframes) to mitigate the interference on other cells and lower their energy consumption [10].

We consider a generalized network model in which base station configurations, namely the operating mode of pico-BSs (*i.e.*, active or sleep) and ABS configuration for macro-BSs, in addition to user association and scheduling is considered. The resulting problem is *non-convex* due to combinatorial nature of BS activation and user association as well as complex interaction of interfering BS transmissions. Thus, it does not conform to conventional duality methods used typically to devise distributed algorithms for networking problems [13]. Therefore, to tackle the problem, we utilize Markov Chain Monte Carlo (MCMC) methods [14]. These methods are attractive as they are not limited to convex problems. In addition, convergence to the *optimal* solution is guaranteed if the algorithm runs for a sufficiently long time. Furthermore, we use an MCMC method known as Gibbs Sampling [14] that

978-3-901882-85-2 © 2016 IFIP

allows implementing the solution in a fully distributed manner. To the best of our knowledge, this is the first work that uses Gibbs Sampling to dynamically control HetNets.

The main advantage of our algorithm is that it works in a distributed self-optimizing fashion. That is, base stations work independently from each other while continuously optimize their configurations without the involvement of a central controller. Consequently, every base station decides on its own configuration, which could be different from that of other base stations (as opposed to some existing work, *e.g.*, [15], that force a uniform configuration for all base stations). In our algorithm, base stations exchange information with each other only in a *local neighborhood*, and yet the system is guaranteed to achieve *global optimality* [14] (in LTE networks, the X2 interface is introduced in base stations to allow them exchange messages with each other directly [16]).

C. Related Work

While the literature on energy management in cellular networks is vast, the following works are more closely related to HetNets, which are the focus on this work.

User Association and Power Control: The problem of user association and power control in HetNets has been somewhat studied in the literature (*e.g.*, see [17]–[22], and references therein). This problem is generally non-convex and involves solving an NP-hard integer program, which is difficult to efficiently solve except for very small networks [22]. Thus, one has to rely on approximations and heuristic techniques to obtain solutions that can be applied to networks of practical size.

Interference Coordination: A few works have recently considered the problem of interference coordination in HetNets [23]–[25]. In particular, the work in [23] is a simulation-based study of the problem. Using stochastic geometry techniques, [24] develops a model to determine the optimal number of ABS sunfarmes. The authors in [25] formulate the problem as a non-linear optimization problem and then develop a two-step greedy algorithm based on integer relaxation and rounding to solve the problem. These works neither consider dynamic base station activation nor distributed solutions.

Base Station Activation: Using measurement studies, it has been shown that $23 - 53\%$ energy saving is achievable via dynamic base station activation in cellular networks [9]. Dynamic base station activation in homogeneous cellular networks is considered in [26]–[28]. The problem of base station activation and interference coordination in a HetNet is considered in [15]. However, the solution proposed in [15] relies on a central controller, while our goal is to develop a distributed solution that does not need such a controller.

Gibbs Sampling: Designing distributed algorithms using Gibbs sampling has been recently investigated in wireless literature [29]. Self-optimization techniques for user association and power control based on this approach are discussed in [30] and [20]. The algorithm in [31] uses Gibbs Sampling to find the optimal location for deployment of small cells in HetNets. From a methodological point of view, the work in [32] is the closets to our work. This paper investigates dynamic activation of base stations to reduce energy consumption in traditional homogeneous cellular networks using a solution based on Gibbs sampling. Our work, on the other hand, considers a HetNet with ABS-based interference coordination, which is quite different from the problem considered in [32].

II. System Model

We consider the donwlink of an OFDM-based HetNet [11]. For ease of exposition, we consider only one channel though our model can be extended to multiple channels by simply replicating all variables for every channel. The system works in discrete time where time is divided into frames of equal length. Each frame is further divided into smaller subframes.

A. Network Model

We consider a HetNet with a set of base stations $\mathcal{B}$ consisting of macro-BSs $\mathcal{B}^{(\mathcal{M})}$ and pico-BSs $\mathcal{B}^{(\mathcal{P})}$. Each base station $b \in \mathcal{B}$ has a maximum transmit power P_b^{Max}, and periodically transmits a reference signal at this power. We denote by $\mathcal{U}$ the set of user devices of the network. The strength of the reference signal received by user u from base station b is given by

$$P_{u,b}^{(R)} = P_b^{Max} H_{b,u}, \tag{1}$$

where $H_{b,u}$ denotes the channel gain between BS b and user u, which includes all propagation impairments such as path loss and fading. Each user u receives reference signals from possibly several base stations. Typically, a receiver has a sensitivity threshold θ and cannot detect signals that are too weak [18], [25]. Let $\mathcal{B}_u$ denote the set of base stations that can serve or impact the rate of user u,

$$\mathcal{B}_u = \{b \in \mathcal{B} | P_{u,b}^{(R)} \geq \theta\}. \tag{2}$$

We assume that the base stations are deployed in such a way that $\mathcal{B}_u$ is non-empty for every user. Similarly, define $\mathcal{U}_b$ for $b \in \mathcal{B}$ as the set of users that can be served by base station b at the minimum power threshold θ. In practice, a user u is served by the base station b_u that can provide the strongest signal to u, that is,

$$b_u = \arg\max_{b \in \mathcal{B}_u} P_{u,b}^{(R)}, \tag{3}$$

where b_u can be either a macro-BS or a pico-BS.

B. Macro Base Stations

The number of macro base stations is usually much lower than the number of pico base stations. Each macro-BS has a high transmission power and covers a large area. In our model, we always keep the macro base stations on to ensure that every user in the system has at least a minimum signal reception quality. For each macro-BS $b \in \mathcal{B}^{(\mathcal{M})}$, let x_b denote the fraction of ABS subframes in a frame. The variable x_b takes its value from a discrete set $\mathcal{X}$ of candidate fractions,

$$\mathcal{X} = \{X_0, X_1, \ldots, X_{|\mathcal{X}|-1}\}, \tag{4}$$

where $|\mathcal{X}|$ is the cardinality of $\mathcal{X}$, and X_is are real numbers between 0 and 1. This set always includes zero as its first element X_0. This is to ensure that ABS subframes can be disabled for a macro-BS if it is not required, *e.g.*, when there is no pico-BS operating close to the macro-BS. The sequence of X_is forms an increasing series, that is,

$$\forall X_i, X_j \in \mathcal{X} : i < j \Rightarrow X_i < X_j . \tag{5}$$

The difference between ABS and regular subframes lies in the amount of transmission power allocated to them. In regular subframes, each macro-BS b transmits data signals to its associated users at its maximum available power P_b^{Max}. In ABS subframes however, macro-BSs use a lower power level to mitigate the interference on picocells. For a macro-BS b, this reduced power level is represented by a real number y_b that is used to scale the transmission power of the base station b as,

$$P_b = y_b \cdot P_b^{Max} . \tag{6}$$

Similar to x_b variables, y_bs are also selected from a discrete set $\mathcal{Y}$ of candidate power level ratios,

$$\mathcal{Y} = \{Y_0, Y_1, \ldots, Y_{|\mathcal{Y}|-1}\}, \tag{7}$$

where, Y_is are real numbers between 0 and 1, and zero is included as Y_0 to be able to have zero-power ABS subframes when it is necessary to mute all data transmissions from a macro-BS during ABS subframes. The sequence of Y_is forms an increasing sequence, that is,

$$\forall Y_i, Y_j \in \mathcal{Y} : i < j \Rightarrow Y_i < Y_j . \tag{8}$$

The variables x_b and y_b are assigned per macro-BS and may differ from one BS to another. This adds another degree of freedom to our model, compared to the approaches with a network-wide fraction of ABS subframes, as in [15].

C. Pico Base Stations

Pico base stations are less expensive and have lower transmit powers compared to macro ones, and are situated in coverage holes or densely populated areas. In our model, pico base stations can be in one of two modes. If having a pico-BS in active mode is not advantageous, we switch it to standby (sleep) mode. In standby mode, a pico-BS is almost shutdown, except for a small processor and accessories that allow it to contact other base stations and see whether it should wake up or not. The energy cost of switching a pico-BS on or off is negligible and not considered in our model (as in [17], [33]). If, on the other hand, the density of users close to a pico-BS is high, it is powered on. In this mode, in contrast with macro base stations, a pico-BS b always transmits at its maximum power level P_b^{Max} and does not have ABS subframes. We denote by a binary variable z_b the state of a pico-BS b:

$$z_{b\in\mathcal{B}^{(\mathcal{P})}} = \begin{cases} 1 & b \text{ is in active mode,} \\ 0 & b \text{ is in standby mode.} \end{cases} \tag{9}$$

III. Problem Formulation

A. Interference Characterization

The amount of interference on a user depends on the status of its neighboring base stations and varies even during one frame. For instance, a user of a macrocell suffers from less interference when a neighboring pico-BS is powered off. Moreover, because different macro-BSs may have different fractions and powers for ABS subframes, the amount of interference changes within a single frame. We denote by T_i the sequence of subframes between two consecutive candidate ABS fractions X_i and X_{i+1}, where $0 \le i < |\mathcal{X}|$. The interference on a user does not change within the subframes of T_i, however, it is likely to change from T_i to T_{i+1}. A user may experience a maximum of $|\mathcal{X}|$ different interference levels.

To formulate the total interference on user u, we first define P_b^i as the transmit power level of BS b during period T_i. One can simply observe that for a pico-BS, this value only depends on whether it is on or off. Therefore,

$$P^i_{b\in\mathcal{B}^{(\mathcal{P})}} = \begin{cases} P_b^{Max} & z_b = 1, \\ 0 & z_b = 0 . \end{cases} \tag{10}$$

For a macro-BS, P_b^i depends on the ratio of ABS subframes x_b and transmit power of the BS during ABS and regular subframes,

$$P^i_{b\in\mathcal{B}^{(\mathcal{M})}} = \begin{cases} P_b^{Max} & x_b \le X_i, \\ y_b \cdot P_b^{Max} & x_b > X_i . \end{cases} \tag{11}$$

Using the above notations, we can calculate the total interference on user u during each period T_i as follows,

$$Int_u^i = \sum_{b \ne b_u} P_b^i H_{b,u} . \tag{12}$$

Note that, the amount of interference increases from the beginning of the frame towards the end, that is,

$$Int_u^0 \le Int_u^1 \le Int_u^2 \le \cdots . \tag{13}$$

This happens because as time passes within a frame, more neighboring macro-BSs make a transition from ABS subframes to regular subframes.

B. Achievable User Rates

Because of the difference in both interference and signal power, the signal to interference and noise ratio (SINR) of each user u varies in different periods T_i of a frame, and can be expressed as,

$$SINR_u^i = \frac{P^i_{b_u}}{Int_u^i + N_0} . \tag{14}$$

As a result, base stations schedule each period separately. Each user gets a portion of subframes of each period. We denote by w_u^i the fraction of subframes of a frame that BS b allocates to user $u \in \mathcal{U}_b$ during period T_i,

$$\forall b \in \mathcal{B}, 0 \le i < |\mathcal{X}| : \sum_{u\in\mathcal{U}_b} w_u^i = X_{i+1} - X_i . \tag{15}$$

The right-hand side of the above equation is the ratio of duration of T_i to the frame length. The following equation is derived from the definition of w_u^i variables,

$$\forall b \in \mathcal{B} : \sum_{i=0}^{|\mathcal{X}|-1} \sum_{u\in\mathcal{U}_b} w_u^i = 1 . \tag{16}$$

To map SINRs to transmission rates, we use the Shannon capacity formula as follows,

$$c_u^i = B \log_2(1 + SINR_u^i), \tag{17}$$

where c_u^i is the total rate that can be achieved by user u during T_i if all subframes in T_i are allocated to u. The parameter B is the bandwidth of the channel. Thus, the rate of user u during the time period T_i can be obtained by the following formula,

$$r_u^i = w_u^i c_u^i \,. \tag{18}$$

Finally, the average rate of user u during a frame can be written as the sum of the rates over all the periods in the frame,

$$R_u = \sum_{i=0}^{|\mathcal{X}|-1} r_u^i \,. \tag{19}$$

C. Base Station Power Consumption

For the energy consumption of base stations, we follow the model proposed in [34] and adopted by [17] and [33]. According to this model, the energy consumption of a BS is given by:

$$W = \begin{cases} N_{TX}(P_0 + \Delta_P \cdot P_{TX}) & P_{TX} > 0, \\ N_{TX} \cdot P_{idle} & P_{TX} = 0, \end{cases} \tag{20}$$

where N_{TX} is the number of antennas, P_0 is the power consumption at zero RF output power, P_{TX} is the RF output power, P_{idle} is the power consumption of the base station in standby mode, and Δ_P is the slope of the load-dependent power consumption.

For simplicity, we assume the number of antennas N_{TX} equals 1. Note that P_0, P_{idle} and Δ_P are predefined parameters. For a pico-BS b, RF output power is either zero or P_b^{Max}. So we have

$$W_{b \in \mathcal{B}^{(\mathcal{P})}} = \begin{cases} P_0 + \Delta_P \cdot P_b^{Max} & z_b = 1, \\ P_{idle} & z_b = 0 \,. \end{cases} \tag{21}$$

Macro-BSs have different RF powers during a frame. Thus, we calculate the average RF power in a frame in order to find the energy consumption of a macro-BS. We have,

$$P_{TX} = x_b \cdot (y_b \cdot P_b^{Max}) + (1 - x_b) \cdot P_b^{Max} \,. \tag{22}$$

By substituting (22) in (20), it is obtained that,

$$W_{b \in \mathcal{B}^{(\mathcal{M})}} = P_0 + \Delta_P \cdot P_b^{Max} \Big(x_b \cdot y_b + 1 - x_b \Big) \,. \tag{23}$$

D. Optimization Objective

We define a utility function $U(\cdot)$ over user rates, and a cost function $C(\cdot)$ over power consumption of base stations. Any continuously differentiable concave function is suitable as the utility function. A commonly used utility function that results in *proportionally fair* user rates is the logarithmic function, that is,

$$U(\mathbf{R}) = \sum_{u \in \mathcal{U}} \log(R_u), \tag{24}$$

where $\mathbf{R} = [R_u]_{1 \times |\mathcal{U}|}$ is the vector of transmission rates of all users. For the power cost function, we simply add up the energy consumed by all the base stations:

$$C(\mathbf{W}) = \sum_{b \in \mathcal{B}} W_b, \tag{25}$$

where $\mathbf{W} = [W_b]_{1 \times |\mathcal{B}|}$ denotes the vector of BS energy consumptions. Finally, usong the weight factor λ, the optimization objective is expressed as,

$$F(\mathbf{R}, \mathbf{W}) = U(\mathbf{R}) - \lambda \cdot C(\mathbf{W}), \tag{26}$$

which is to be maximized. A larger value of λ increases the negative impact of energy consumption on the objective function, and favors low-throughput and low-energy states over high-throughput and high-energy states.

We summarize the optimization objective and the constraints in (27):

$$\begin{aligned} \underset{x_b, y_b, z_b, w_u^i}{\textit{maximize}} \quad & F(\mathbf{R}, \mathbf{W}) \\ \textit{subject to} \quad & x_b \in \mathcal{X}, \quad \forall b \in \mathcal{B}^{(\mathcal{M})} \\ & y_b \in \mathcal{Y}, \quad \forall b \in \mathcal{B}^{(\mathcal{M})} \\ & z_b \in \{0, 1\}, \quad \forall b \in \mathcal{B}^{(\mathcal{P})} \\ & \sum_{u \in \mathcal{U}_b} w_u^i = X_{i+1} - X_i, \quad \forall b \in \mathcal{B} \\ & w_u^i \in [0, 1], \quad \forall u \in \mathcal{U}, 0 \le i < |\mathcal{X}| \,. \end{aligned} \tag{27}$$

The objective of the problem (27) is non-convex. To solve the problem, we employ an MCMC method called Gibbs sampling to be described in the following section.

IV. Solution Method

A. Overview

The optimization problem in (27) contains a mixture of real variables w_u^i, discrete variables x_b and y_b, and binary variables z_b. Here we use an optimization technique known as *meta-optimization*, in which one optimization problem is solved and the results are fed into another one. In our case, each base station locally solves a convex optimization problem, and uses its solution in a larger-scale distributed algorithm in order to optimize the global objective function as defined in (26).

We begin by breaking the global utility and cost functions for the network into the sum of local utility and cost functions for each BS. According to (25), the local cost function can be easily written for each BS. Let C_b denote the cost of base station b. We have,

$$C_b = W_b \,. \tag{28}$$

The global utility function (24) is defined over user rates. We know that each user is associated to one base station, which is the one with the strongest received signal. Thus, if we write the local utility function for each base station as a function of the rates of the users associated to it, we obtain the global utility function as the sum of local ones over the base stations. Let U_b denote the local utility of base station b. We have,

$$U_b = \sum_{u \in \mathcal{U}_b} \log(R_u) \,. \tag{29}$$

Therefore, the local objective function can be written as,

$$F_b = U_b - \lambda \cdot C_b, \tag{30}$$

where F_b denotes the local objective function at base station b. Now the global objective function can be written as the sum of local ones, that is,

$$F(\mathbf{R}, \mathbf{W}) = \sum_{b \in \mathcal{B}} F_b. \tag{31}$$

B. Gibbs Sampler

Consider a graph G that represents the system, where the nodes of G are the set of base stations $\mathcal{B}$, both macro-BSs and pico-BSs. In this graph, two base stations are adjacent if they have at least one common user in their range. In line with the literature, we refer to this graph as the *Interference Graph*, since connected base stations in this graph directly interfere with each other's users. The set of the neighbors of node b in this graph is denoted by $\mathcal{N}_b$.

For each node b of the graph, define the state variable s_b. Macro-BSs and pico-BSs have different state spaces. For macro-BSs, the states are ordered pairs of the form (x_b, y_b), whereas for pico-BSs the state only contains the on/off state (z_b). Thus, the state spaces are given by,

$$\mathcal{S}_b = \begin{cases} \mathcal{X} \times \mathcal{Y} & b \in \mathcal{B}^{(\mathcal{M})}, \\ \{0, 1\} & b \in \mathcal{B}^{(\mathcal{P})}. \end{cases} \tag{32}$$

To calculate the local objective function F_b, each base station only needs to know its own state, and the state of its neighbors. Using this information, the only missing variables to calculate the local objective function of base station b in (30) are the shares $w_u^i, u \in \mathcal{U}_b$ of each user associated to b. To find the user shares, each base station b should solve the following optimization problem *locally*,

$$\max_{w_u^i} F_b, \tag{33}$$

given the values of variables z_b, x_b, and y_b for b and its neighbors, and the last two constraints in the main optimization problem (27). It is straightforward to verify that the problem in (33) is a convex optimization problem. Thus, standard methods in the field of convex optimization can be used to solve the problem efficiently.

Next, to compute the value of the local objective function of each node b in graph G (that is, F_b), the following values are used as input:

- states z_b, x_b, and y_b of b and its neighbors in $\mathcal{N}_b$,
- the optimal solution of (33).

We define the *local energy* of each state $s_b \in \mathcal{S}_b$, denoted by $\varepsilon_b(s_b)$, as follows,

$$\varepsilon_b(s_b) = \sum_{b' \in \mathcal{N}_b^+} F_{b'}, \tag{34}$$

where $\mathcal{N}_b^+$ denotes the set containing b and its neighbors,

$$\mathcal{N}_b^+ = \mathcal{N}_b \cup \{b\}. \tag{35}$$

To find the local energy, each node b needs to calculate its own local objective function and those of its neighbors $\mathcal{N}_b$. To find the latter, a node needs to have the states of its two-tier neighbors. We construct the graph G', which we call the *two-tier neighborhood* graph. Then each node b obtains the states of its neighbors in G' denoted by $\mathcal{N}_b^2$ via communication among neighboring nodes, as described next.

C. Distributed Algorithm

We break the algorithm into two phases, as described next.

1) Initialization Phase: Initially, all the picocells are on and all macro-BSs are in full power mode. Users connect to the BS with the highest received reference signal power. Each BS b requests each of its users u to report the list of all BSs they can hear (denoted by $\mathcal{B}_u$), and their channel gains (denoted by $H_{b,u}$). Then each BS sends the information gathered from its users to each of the neighboring BSs that it has identified so far. Next, they exchange their maximum RF output power P_b^{Max}, their initial state, and their neighbor list with their immediate neighbors $\mathcal{N}_b$. By processing these received lists, BSs can construct their two-tier neighborhood list $\mathcal{N}_b^2$.

2) Iterative Update Process: In this phase, each BS is updated independently of the other BSs. At the start of this phase, each BS b randomly chooses a timeout between 0 and τ, and triggers an update when the timeout expires. In an update process, a BS b selects a new state $s \in \mathcal{S}_b$ from its state space, independent of its current state.

To select a new state, BS b computes the local energy $\varepsilon_b(s')$ for each potential next state $s' \in \mathcal{S}_b$ by calculating the local objective function $F_{b'}$ of every BS $b' \in \mathcal{N}_b^+$ (b and its neighbors) given the state s'. Then it chooses its new state according to the following probability distribution, which is known as the Gibbs distribution [14],

$$\mathbb{P}\{s_b = s\} = \frac{e^{\frac{1}{T}\varepsilon_b(s)}}{\sum_{s' \in \mathcal{S}_b} e^{\frac{1}{T}\varepsilon_b(s')}}. \tag{36}$$

The variable T is a free parameter called the temperature, and is a decreasing function of the time t elapsed from the beginning of the iterative phase. An example of the temperature function is the following quadratic function with zero temperature at time $t = t_{end}$,

$$T(t) = T_0 \times (1 + \frac{t-1}{1-t_{end}})^2. \tag{37}$$

By repeating this process until convergence, the system reaches the optimal global state where each BS selects the same state at each update.

V. Simulation Results

A. Simulation Environment

We simulate a heterogeneous network to study the performance and utility of our algorithm. Since macro-BSs have a large state space of the size $|\mathcal{X} \times \mathcal{Y}|$ and all of the base stations are simulated on one machine (contrary to the reality where each BS has its own processor), we limit the number of cells to 19. It should be noted that this network size is adequate for prototyping purposes, and much smaller networks have been studied in the literature, for example in [31]. We have also simulated larger networks consisting of 400 picocells, and the

TABLE I: Default Simulation parameters.

Parameter	Value	
	Pico-BS	Macro-BS
N_0 (dBm)	-120	
θ (dBm)	-90	
B (Hz)	5×10^3	
P^{Max} (W)	0.5	20
P_{idle} (W)	4.3	75
P_0 (W)	6.8	130
Δ_P	4	4.7
T_0	5	
τ	10	
t_{end}	10^3	

results are consistent with (or, in some cases, outperform) the ones presented here [35]. Due to space limitation, only the results for the 19 cell network are presented here.

As depicted in Fig. 1, the network is composed of 7 macro-BSs located on a hexagonal layout and 12 pico-BSs located on cell edges. As in LTE networks, the frequency reuse factor is 1, *i.e.*, all BSs and users operate on the same frequency. One hundred users are spread over the network using a Poisson Point Process (which is widely used for modeling user locations [24]). We make sure that all users are in the covered area and receive a signal above the threshold θ from at least one macro or pico base station. **The choice of models for user locations, channel gains, *etc*. have no effect on our model.**

B. Simulation Parameters

For algorithm-independent parameters, we have used the fixed values reported in Table I throughout the simulations. To estimate the channel gains, we used a standard distance-dependent path loss model with the path-loss exponent $\gamma = 3.5$ [36]. The parameters in Table I are similar to the ones used in [33] and mostly aligned with HetNet specific parameters considered in [37]. To deal with numerical issues when computing the objective function, we use a scaling factor as follows,

$$F_b = 10^{-4}(U_b - \lambda \cdot C_b),$$

where, 10^{-4} is the scaling factor and $\lambda = 200$ is chosen to balance the throughput and energy consumption. The set of ABS duration and power ratios are set to $\mathcal{X} = \{0, \frac{1}{10}, \frac{2}{10}, \frac{3}{10}\}$ and $\mathcal{Y} = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$, respectively.

C. Converged Network State

Fig. 1 illustrates the state of the network after the convergence of the algorithm. For the sake of reporting, we number the cells as follows. The macrocell in the area **i** of the Fig. 1 is referred to as **Mi**. The picocell that is encompassed by **Mi** is called **Pi**. Also, the picocell that overlaps Mi and Mj is referred to as **Pij**.

Table II reports the final values of the performance measures after the convergence of the system. Table III shows the breakdown of measures and state of macro-BSs. Table IV reports the statistics of the pico-BSs, except for the ones that are in sleep mode after convergence.

The energy saving (of about 10%) in Table II comes from both macro-BSs and pico-BSs. All macro-BSs have chosen to have ABS duration ratio of $1/3$ (the largest value allowed). Macro-BS M7 is consuming less energy compared to the other ones, because it remains fully silent during the ABS subframes, while others have decided to transmit at quarter or half the maximum power. The reason could be the fact that M7 is interfering with more picocell users compared to other macro-BSs. Six of the pico-BSs are in standby mode. P1, P4 and P56 have no user in their coverage, and there is no reason for them to stay on. P3, P5 and P6 are in sleep mode because the users in their range can get a much better throughput from macro-BSs. Other pico-BSs are operating because the users they are serving are either not covered by macro-BSs, or are very close to macrocell edges.

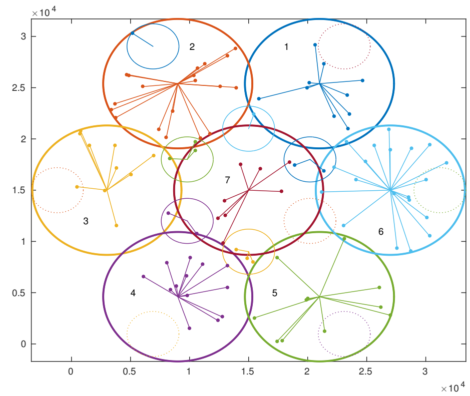

Fig. 1: The network after convergence. Lines indicate associations. Dotted picocells are in sleep mode.

TABLE II: Global performance measures.

Total Throughput	866705
Total energy consumption	1519.7
Energy saving	153.9
Global objective	56.3

TABLE III: Statistics of macro-BSs.

	M1	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**
ABS time	1/3	1/3	1/3	1/3	1/3	1/3	1/0
ABS power	1/4	1/4	1/2	1/2	1/2	1/2	0
Tput ($\times 10^2$)	963.8	829.2	1395.6	1010.3	953.3	1170.1	529.2
Energy	202.9	202.9	209.9	209.9	209.9	209.9	195.8
Objective	5.6	4.2	9.8	5.9	5.3	7.5	1.4

TABLE IV: Statistics of pico-BSs.

	P16	**P45**	**P34**	**P23**	**P12**	**P2**
Throughput	24273.8	79093.2	25484.6	31992.1	13669.4	7036.9
Energy	8.8	8.8	8.8	8.8	8.8	8.8
Objective	2.3	7.7	2.4	3	1.2	0.5

VI. Conclusion

In this work, we developed a distributed self-optimizing algorithm based on Gibbs sampling to balance the trade-off between energy consumption and user rates in a heterogeneous cellular network. Our algorithm achieves this by putting underutilized pico base stations in standby mode and adjusting the ratio of almost blank subframes in individual macro base stations. The simulation results show that enabling ABS subframes and deactivating underutilized picocells can reduce the network energy consumption by almost 10% without penalizing user rates.

References

[1] Cisco Visual Networking Index, "Global mobile data traffic forecast update, 2014–2019," 2015. [Online]. Available: http://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html

[2] "Ericsson mobility report: On the pulse of the networked society," Ericsson, Tech. Rep., 2015. [Online]. Available: http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf

[3] "Femtocell and microcells," May 2016. [Online]. Available: https://www.repeaterstore.com/pages/femtocell-and-microcell

[4] M. Bilgic, "Who needs LTE small cells first?" EXFO, Tech. Rep., 2013. [Online]. Available: http://www.exfo.com/corporate/blog/2013/needs-lte-small-cells

[5] O. Østerbø and O. Grøndalen, "Benefits of self-organizing networks (SON) for mobile operators," *Computer Networks and Communications*, 2012.

[6] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *IEEE ICC Workshops*, Jun. 2009.

[7] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 14, 2011.

[8] M. A. Marsan and M. Meo, "Green wireless networking: Three questions," in *IFIP Med-Hoc-Net*, Jun. 2011.

[9] C. Peng, S.-B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power saving in operational 3G cellular networks," in *ACM MobiCom*, Sep. 2011.

[10] D. Lopez-Perez *et al.*, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun. Mag.*, vol. 19, no. 3, 2011.

[11] J. Wannstrom and K. Mallinson, "Heterogeneous networks in LTE," May 2016. [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/1576-hetnet

[12] I. Ashraf, F. Boccardi, and L. Ho, "Sleep mode techniques for small cell deployments," *IEEE Communications Magazine*, vol. 49, no. 8, 2011.

[13] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, 2006.

[14] P. Brémaud, *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Springer Science & Business Media, 1999.

[15] Q. Ye, M. Al-Shalashy, C. Caramanis, and J. G. Andrews, "On/off macrocells and load balancing in heterogeneous cellular networks," in *IEEE Globecom*, Dec. 2013.

[16] 3GPP, "UTRAN Iub interface Node B Application Part (NBAP) signalling," 3rd Generation Partnership Project (3GPP), TS 25.433, 06 2011. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/25433.htm

[17] C. Bottai, C. Cicconetti, A. Morelli, M. Rosellini, and C. Vitale, "Energy-efficient user association in extremely dense small cells," in *European Conference on Networks and Communications*, Jun. 2014.

[18] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous LTE-A cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, 2010.

[19] V. N. Ha and L. B. Le, "Distributed base station association and power control for heterogeneous cellular networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, 2014.

[20] C. S. Chen and F. Baccelli, "Self-optimization in mobile cellular networks: Power control and user association," in *IEEE ICC*, May 2010.

[21] J. Ghimire and C. Rosenberg, "Resource allocation, transmission coordination and user association in heterogeneous networks: A flow-based unified approach," *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, 2013.

[22] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, 2013.

[23] M. I. Kamel and K. Elsayed, "Performance evaluation of a coordinated time-domain eICIC framework based on ABSF in heterogeneous LTE-advanced networks," in *IEEE Globecome*, Dec. 2012.

[24] M. Cierny, H. Wang, R. Wichman, Z. Ding, and C. Wijting, "On number of almost blank subframes in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, 2013.

[25] S. Deb, P. Monogioudis, J. Miernik, and J. P. Seymour, "Algorithms for enhanced inter-cell interference coordination (eICIC) in LTE HetNets," *IEEE/ACM Trans. Netw.*, vol. 22, no. 1, 2014.

[26] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in *IEEE Globecome*, Dec. 2010.

[27] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular network," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, 2011.

[28] A. Abbasi and M. Ghaderi, "Energy cost reduction in cellular networks through dynamic base station activation," in *IEEE Secon*, Jul. 2014.

[29] S. C. Borst, M. G. Markakis, and I. Saniee, "Nonconcave utility maximization in locally coupled systems, with applications to wireless and wireline networks," *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, 2014.

[30] B. Kauffmann, F. Baccelli, A. Chaintreau, V. Mhatre, K. Papagiannaki, and C. Diot, "Measurement-based self organization of interfering 802.11 wireless access networks," in *IEEE Infocom*, May 2007.

[31] X. Li, X. Tang, C.-C. Wang, and X. Lin, "Gibbs-sampling-based optimization for the deployment of small cells in 3g heterogeneous networks," in *IEEE WiOpt*, May 2013.

[32] A. Abbasi and M. Ghaderi, "Online algorithms for energy cost minimization in cellular networks," in *IEEE IWQoS*, May 2014.

[33] A. Prasad, A. Maeder, and C. Ng, "Energy efficient small cell activation mechanism for heterogeneous networks," in *IEEE Globecom Workshops*, Dec. 2013.

[34] M. Imran *et al.*, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," Tech. Rep. ICT-EARTH deliverable, Tech. Rep., 2011.

[35] M. Naghibi, "Distributed energy minimization in heterogeneous cellular networks." [Online]. Available: http://www.ucalgary.ca/~mghaderi/docs/naghibi.pdf

[36] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[37] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Mobility enhancements in heterogeneous networks," 3rd Generation Partnership Project (3GPP), TR 36.839, 06 2011. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36839.htm