Aalto University

OPEN ACCESS

This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Lassila, Pasi; Gebrehiwot, Misikir Eyob; Aalto, Samuli

Optimal energy-aware load balancing and base station switch-off control in 5G HetNets

Published in:
Computer Networks

DOI:
10.1016/j.comnet.2019.05.001

Published: 04/08/2019

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Published under the following license:
CC BY-NC-ND

Please cite the original version:
Lassila, P., Gebrehiwot, M. E., & Aalto, S. (2019). Optimal energy-aware load balancing and base station switch-off control in 5G HetNets. *Computer Networks*, *159*, 10-22. https://doi.org/10.1016/j.comnet.2019.05.001

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Optimal energy-aware load balancing and base station switch-off control in 5G HetNets

Pasi Lassila*, Misikir Eyob Gebrehiwot, Samuli Aalto

Department of Communications and Networking, Aalto University, Finland

Abstract

We consider optimal energy-aware load balancing of elastic downlink data traffic inside a macrocell with multiple small cells within its coverage area. The system is modeled as a set of parallel queues. In particular, the model of the small cell includes the setup delay resulting from activating the base station after being placed in a low power off state and the idle timer controlling the amount of time to wait before being switched off. We apply the theory of MDPs to develop state-dependent dynamic policies for controlling both the routing of the arrivals as well as the length of the idle timer that minimizes the weighted sum of energy and performance. In particular, we show that in the optimal policy the idle timer control can be simplified to selecting a value arbitrarily close to zero or infinite. Additionally, by utilizing the first step of the well-known policy iteration method, we develop an explicit near-optimal dynamic policy for routing the arrivals and also for determining the idle timer configuration of the system, based on the expressions for the future marginal costs. The performance of the policy is illustrated through numerical examples.

Keywords: HetNets, load balancing, performance-energy tradeoff, parallel queues, Markov Decision Processes

1. Introduction

Energy-aware heterogeneous networks (HetNets) is one of the key enabling technologies for realizing the future 5G networks to allow user transmission speeds reaching several Gbit/s [2, 29]. Specifically, HetNets address the problem of spatially heterogeneous distribution of the traffic load within a cell by introducing inside the coverage region of the macrocell so-called small cells, sometimes also referred to as pico- or femtocells, with low-power base stations that are operating under the control of the macrocell. The small cells can offer at a traffic hot spot a high transmission rate to nearby users and thus some of

*Corresponding author

Email addresses: pasi.lassila@aalto.fi (Pasi Lassila), misikir.gebrehiwot@aalto.fi (Misikir Eyob Gebrehiwot), samuli.aalto@aalto.fi (Samuli Aalto)

the traffic can be off-loaded away from the macrocell to the respective small cell. Such load balancing clearly can benefit the performance of the users, for example, by minimizing the delay. However, in modern systems it is also important to consider possibilities for saving energy by switching off small cells when the load is low.

We study the following scenario. We consider a single macrocell with several small cells inside its coverage region serving downlink traffic. The user traffic consists of elastic flows, roughly corresponding to file transfers controlled by TCP, that are downloaded through the base stations. New flows arrive according to a Poisson process. Thus, the number of active flows varies randomly over time and the flow-level performance is represented by the mean flow-level delay, i.e., the mean file transfer delay. Upon the arrival of a new flow, a routing policy will decide whether to serve the flow through the local small cell or the macrocell. From the energy point of view, the macrocell is always on in order to provide coverage in the whole cell area. However, the small cells can be switched off to save energy during low load. Activating again a small cell that has been switched off incurs a performance cost in the form of a setup delay, thus giving rise to a performance-energy tradeoff in the system. Our objective is to develop energy- and delay-aware load balancing algorithms. There is already a vast literature on algorithms for controlling the performance-energy tradeoff in HetNets, and we review them in detail in the related work.

In our flow-level queuing model, the macrocell is represented by a multiclass M/M/1-PS queue, where the classes represent flows that arrived in a given small cell but are served by the macrocell. The small cells are modeled as single-class M/M/1-PS queues with a setup delay, and we additionally allow another control parameter, the so-called idle timer, which defines how long a small cell waits before switching off after the small cell becomes idle (i.e., becomes empty of flows). Overall, the system consists of a set of parallel queues. The control policy consists of two parts: the routing policy decides whether the arrival is routed to the small cell or the macrocell and the idle timer control policy decides the appropriate idle timer value to be applied in each small cell.

To characterize the performance-energy tradeoff, we represent the system cost as a weighted sum of the delay and energy, which is a popular cost metric for analyzing the tradeoff, see, e.g., [1, 4, 28]. To optimize the cost, we consider dynamic state-dependent policies and apply the theory of Markov Decision Processes (MDP) and the so-called policy iteration method [20]. In particular, in the first policy iteration (FPI) approach, the idea is to consider only the first step of the policy iteration. By using a probabilistic initial policy the future costs from the routing actions only depend on the so-called value functions of each small cell and macrocell independently. The policy is near-optimal in the sense that typically the largest improvements are obtained with the first step of the policy iteration. The MDP approch has been also applied to the HetNet load balancing problem earlier, as will be reviewed later in the related work. Typically, the MDP problems are only analyzed by using numerical methods and fundamental analytical explicit results on how the optimal (or even near-optimal) policies are scarce. Our novel contributions include explicit analytical insights on the structure of the optimal policy, in particular related to the idle timer control problem.

This paper is an extended version of our earlier conference paper [12][1], where we studied the same system model as considered here. The explicit forms of the value functions for the small cells and the macrocell were derived and the FPI policy was determined. However, in [12] only the routing decision was considered as part of the FPI optimization step. The idle timer was assumed to be fixed statically and we numerically investigated its impact on the performance energy-tradeoff. In this paper, we explicitly include the idle timer as part of the optimization in the MPD framework and study the joint dynamic control of routing decisions and idle timer values, which can be selected from an arbitrary set with given lower and upper bounds.

In more detail, our new technical contributions compared with [12] are the following. We obtain novel fundamental insights to the optimal control of the routing and the idle timer values. Based on the properties of the so-called optimality equation, we first show that in the optimal policy, the optimal timer value is always set equal to either the lower or upper bound, i.e., intermediate values do not give any additional benefit. The lower and upper bounds are arbitrary and can be selected to be arbitrarily close to zero and infinite, suggesting that it is optimal for the small cell to switch off immediately or never to switch off. We also show that in the optimal policy the optimization becomes separable with respect to all decisions, and there is no need to consider combinations of the possible actions in a given state. Additionally, for the FPI policy itself, based on the explicit expressions of the value functions, we show that the idle timer control action becomes equivalent to determining the optimal idle timer values based on the mean costs following from the initial probabilistic routing policy. These results significantly simplify the idle timer control problem in our case, and also generalize similar earlier results on the optimal idle timer selection in an energy-aware M/G/1-PS queue with Poisson arrivals, see [13], to queues with non-Poisson input. Finally, the near-optimal FPI policy then consists of a static selection of the idle timers and a dynamic state-dependent routing policy, which depends on the marginal cost of routing the flow to the local small cell or to the macrocell. The form of the FPI policy has a linear marginal performance cost and a constant energy cost with respect to the number of flows in the system. Our extensive simulations confirm that the FPI policy is able to achieve significant gains compared with the static initial policy as well as other dynamic heuristic policies based on the well-known JSQ-policy (Join-the-Shortest-Queue).

The paper is organized as follows. The system model is given in Section 2. The optimal routing and idle timer control problem is defined and analyzed in Section 3. The FPI approach is introduced and the FPI policy is developed in Section 4. Numerical results and conclusions are in Sections 5 and 6, respectively.

[1] available from: `http://www.netlab.tkk.fi/~pelassil/lassila-networking2018.pdf`

2. Related work

As mentioned earlier, energy-aware load balancing in HetNets has been studied considerably recently. Typically, the approaches include developing a static optimization problem formulation and solution algorithms for minimizing the energy consumption subject to constraints on the average traffic characteristics, such as, the average load in the system. The locations of the users can be either assumed to known in the problem, as in, e.g., [30] or there is an assumed given user density, as in, e.g., [6, 19, 31], based on which user traffic characteristics, e.g., the average load, can be determined. Also, by using similar system models more heuristic load balancing and on/off control algorithms without a rigorous mathematical optimization framework have been presented in [15, 21]. Finally, a more detailed recent survey on similar approaches can be found in [14]. However, while being insightful and allowing a more detailed modeling of the MAC layer and the wireless channel, the above mentioned works do not take into account the stochastically varying user population and the delay performance of the users. Our focus is to take these aspects explicitly into account by using a queuing theoretic approach.

There is a significantly smaller body of work applying stochastic models for the HetNet load balancing problem. In this context, as mentioned in Section 1, it is natural to apply the theory of MDPs for developing control algorithms. In particular, MDPs have been applied recently for various control problems in HetNets in [7, 8, 9, 16, 17, 22, 23, 24, 25, 26, 27]. Next we comment on these in more detail.

In [25], an MDP formulation is derived for a mobile user for the handover decision based on local information that the mobile user observes, and thus the modeling view point is very different from ours as the optimization is not concerning the behavior of the whole system. On the other hand, in [17] and [27] the optimization of the routing actions is studied, but the system model only considers a single small cell inside the macrocell and importantly no energy aspects are considered. In [24], the model includes many small cells inside the macrocell and applies to elastic traffic, but a significant difference in the model, which makes the model in [24] more complex than ours, is that in their model small cell base stations operate in the same frequency as the macrocell thus creating interference. We assume that small cells and the macrocell do not share the same frequency, but operate in the outband mode. However, in [24] the optimization does not include the routing actions at all, but only concerns the base station on/off actions, while we consider jointly optimizing both the routing and the base station on/off switching. In [7], both base station on/off control and routing actions are considered for a system with multiple small cells, as done in our paper. The system model in [7] applies to realtime (inelastic) traffic and also includes mobility (which affects the underlying queuing model). In [8], partly by the same authors as [24], a similar problem is studied but the network setting is slightly different and corresponds to a flat non-hierarchical sectorized system (i.e., not a HetNet), where again the objective is to optimize the base station sleep modes. The problem is cast as an MDP but also an online learning algorithm is proposed to allow an adaptive implementation of the optimized policy. Similarly, in [9] a learning algorithm is provided for a HetNet scenario to optimize the base station on/off switching and traffic load balancing. However, common to all papers [7, 8, 9, 17, 24, 25, 27] is that they do not contain explicit analytical results on the optimal policy. In these papers, the decision problem is formulated as an MDP and the solution to the problem is obtained only numerically, typically by applying the value iteration algorithm, or a learning algorithm (which can be viewed as an iterative solution to an underlying MDP) is proposed. Our contributions are, on the other hand, precisely in the structural explicit results related to the optimal (or near-optimal) policy, as detailed in Section 1.

Notably, in [22] (and its earlier version [23]), the authors consider an MDP problem for the load balancing in a model with only one macrocell base station and one small cell carrying both elastic and inelastic traffic, and structural results that the optimal load balancing policy is of threshold type have been obtained. However, their model is simpler than ours and does not consider the energy-aware features.

To develop an explicit form of the policy for routing the flows, we use the FPI approach, as discussed earlier. In our prior work, we have applied the FPI approach to load balancing in HetNets for elastic traffic in [16] and [26]. However, in [16] the system model was simpler and we only considered the routing actions for load balancing without any energy-aware features. In [26], our system model was similar to the one in this paper, but the idle timer control for base station on/off switching was not included in the MDP problem. More importantly, in [26] the FPI policy was determined numerically and no explicit results were given. Also, in a slightly different modeling context involving job dispatching among parallel queues, we developed the FPI policy for the routing decisions among a system of energy-aware queues with setup delays but without the idle timer control feature in [11]. In contrast to the above mentioned papers [11, 16, 26], here we obtain fundamental insights on the structure of the optimal policy, in particular related to the idle timer control problem.

3. Model

We consider a heterogeneous wireless system consisting of a single macrocell and K separate small cells located inside the coverage area of the macrocell. The macrocell is indexed by 0 and the small cells by $k = 1, \dots, K$. We assume that the small cells operate in an outband mode, i.e., they have their own radio resources and do not interfere with the macrocell. In addition, we assume that the small cells are far enough from each other so that they do not interfere with each other either.

Traffic consists of elastic downlink data flows (such as TCP file transfers). Let λ_k denote the arrival rate of new flows within the area of small cell k. Each such a flow can be served either by the small cell itself or the macrocell (but not by the other small cells). Upon the arrival, a routing decision must be made whether the flow is attached to the small cell or the macrocell. In addition, let

λ_0 denote the arrival rate of those flows (outside the "hotspot" areas covered by the small cells) that can only be served by the macrocell. All the arrival processes are assumed to be independent Poisson processes.

Each small cell k is modeled by a single server PS queue, which implies that flows are scheduled in each time slot so that all resources are given to one flow at a time and the flows are served in a round-robin manner. We assume that the service time S_k of an arbitrary flow in small cell k is exponentially distributed with mean $E[S_k] = 1/\mu_k$. The mean service time $E[S_k]$ is the average time needed to complete the transfer of a random flow if there were no other flows to be carried by the same cell. We note that the service time is affected at least by the size of the original flow, the location of the corresponding mobile terminal (within the small cell), and the radio channel conditions during the flow transfer. However, since the scheduler of the small cell does not utilize these features, we do not model them separately. PS queues have been applied as reasonable models for modern wireless systems in a large number of papers since the seminal paper [5].

For each small cell, we apply the DELAYEDOFF energy state control (according to the terminology of [10]). As long as there are flows in the small cell to be served, the state of the server is said to be BUSY, but as soon as the system becomes empty, the state changes to IDLE. The server remains IDLE as long as one of the following events take place. Either a new flow arrives, in which case the server becomes again BUSY and starts serving the new flow, or the idle timer (associated with the idle server) expires, in which case the server is immediately switched OFF. In our model, the length I_k of the idle timer of small cell k is assumed to be independently and exponentially distributed with mean $E[I_k] = 1/\omega_k$. In the latter case, the server remains OFF until a new flow is routed to small cell k and the server is put to the SETUP state. After a setup delay D_k, which in our model is assumed to be independently and exponentially distributed with mean $E[D_k] = 1/\delta_k$, the server becomes again BUSY and starts serving the waiting flows.

The state of small cell k at time t is described by the pair $(X_k(t), B_k(t))$, where $X_k(t)$ refers to the number of flows and $B_k(t)$ to the state of the server. Note that server k is

$$\begin{array}{ll} \text{BUSY}, & \text{if } X_k(t) > 0 \text{ and } B_k(t) = 1; \\ \text{IDLE}, & \text{if } X_k(t) = 0 \text{ and } B_k(t) = 1; \\ \text{OFF}, & \text{if } X_k(t) = 0 \text{ and } B_k(t) = 0; \\ \text{SETUP}, & \text{if } X_k(t) > 0 \text{ and } B_k(t) = 0. \end{array}$$

We denote the power consumption in these energy states of small cell k by $P_k^{\mathrm{E-STATE}}$, and we assume that

$$0 = P_k^{\mathrm{o}} < P_k^{\mathrm{i}} \leq \min\{P_k^{\mathrm{s}}, P_k^{\mathrm{b}}\},$$

where superscripts o, i, s, and b refer to OFF, IDLE, SETUP, and BUSY states, respectively.

The macrocell (0) serves $K+1$ different classes of flows. Class 0 refers to those flows that can only be served by the macrocell and class k to those flows that arrive in the area of small cell k but are routed to the macrocell. The macrocell itself is modeled by a single server multiclass PS queue, which again implies that flows are scheduled in each time slot so that all resources are given to one flow at a time and the flows are served in a round-robin manner. We assume that the service time $S_{0,k}$ of an arbitrary flow in class $k \in \{0, 1, \ldots, K\}$ is exponentially distributed with mean $1/\mu_{0,k}$.

For the macrocell, we do not apply any energy state control, but it is a NEVEROFF server (according to the terminology of [10]). Thus, the state of the macrocell at time t is described by the vector $X_0(t) = (X_{0,0}(t), X_{0,1}(t), \ldots, X_{0,K}(t))$, where $X_{0,k}(t)$ refers to the number of flows in class k. Note that server 0 is

$$\begin{array}{ll} \text{BUSY}, & \text{if } |X_0(t)| > 0; \\ \text{IDLE}, & \text{if } |X_0(t)| = 0, \end{array}$$

where

$$|X_0(t)| = X_{0,0}(t) + X_{0,1}(t) + \ldots + X_{0,K}(t)$$

denotes the total number of flows in macrocell. The power consumption of macrocell is denoted by $P_0^{\mathrm{E-STATE}}$, and we assume that

$$0 < P_0^{\mathrm{i}} \leq P_0^{\mathrm{b}}.$$

In this paper, we develop policies for dynamically controlling the routing of the arrivals and the idle timers. For any such policy, the necessary condition for the stability of the system is as follows. As in [26], the macrocell is the bottleneck in the system, because it serves as an overflow system for the arrivals in the small cells. Thus, the maximal stability condition for any policy is

$$\frac{\lambda_0}{\mu_{0,0}} + \sum_{k=1}^{K} \frac{(\lambda_k - \mu_k)^+}{\mu_{0,k}} < 1. \tag{1}$$

In (1), the terms of the sum correspond to the load caused by class k at the macrocell due to the excess arrival rate from small cell k. The macrocell is a multiclass PS queue, and the sum of the loads of all classes must be less than 1.

4. Optimal routing and idle timer control problem

Our objective is to develop a dynamic, state-dependent routing and idle timer control policy that simultaneously takes into account both the power consumption and the delay of the flows. The routing policy decides for each arriving flow in the small cells, whether the arrival is served by the local small cell or the macrocell, and the idle timer control policy determines the time to wait before switching off upon becoming idle. For our model, optimal policies can be developed in the framework of MDPs [20].

4.1. MDP formulation of the problem

For the optimization, the cost rates with respect to the performance (delay) and the energy need to be defined first. Let the vector $x_0 = (x_{0,0}, \ldots, x_{0,K})$ denote a given state of the macrocell. Similarly, we denote by $x = (x_1, \ldots, x_K)$ and $b = (b_1, \ldots, b_K)$ vectors for the number of flows and the state of the server in each small cell. Given that there are (x_o, x) flows in the system, the instantaneous cost rate for the performance, $c^{\mathrm{p}}(x_0, x)$, is given by

$$c^{\mathrm{p}}(x_0, x) = \sum_{k=0}^{K} x_{0,k} + \sum_{k=1}^{K} x_k, \tag{2}$$

i.e., it is the total number of flows in the system. For the energy, the instantaneous cost rate, $c^{\mathrm{e}}(x_0, x, b)$, is the total power in the given state and it equals

$$\begin{aligned} c^{\mathrm{e}}(x_0, x, b) &= 1_{|x_0|=0} P_0^{\mathrm{i}} + 1_{|x_0|>0} P_0^{\mathrm{b}} + \\ &\sum_{k=1}^{K} (1_{x_k>0, b_k=1} P_k^{\mathrm{b}} + 1_{x_k=0, b_k=1} P_k^{\mathrm{i}} + 1_{x_k>0, b_k=0} P_k^{\mathrm{s}}), \end{aligned} \tag{3}$$

where 1_A denotes the indicator function of the event A. To characterize the tradeoff between performance and energy, the total instantaneous cost rate in state (x_0, x, b), $c(x_0, x, b)$, is defined as the weighted sum of performance and energy,

$$c(x_0, x, b) = w_1 c^{\mathrm{p}}(x_0, x) + w_2 c^{\mathrm{e}}(x_0, x, b),$$

where $w_1, w_2 \geq 0$ are the weight parameters. We return to the interpretation of the weights later when defining the optimization objective (4).

As actions we consider the routing decisions and the idle-timer control. To explain the dynamics we introduce the vector e_k as a K-dimensional unit vector with zeros elsewhere except in the component k, $k = 1, \ldots, K$. Similarly, let $e_{0,k}$ denote a $(K+1)$-dimensional unit vector with zeros elsewhere except in the component k, $k = 0, \ldots, K$. The routing decisions relate to arrivals in each small cell k and the routing decision is denoted by $a_k^{\mathrm{r}}(x_0, x, b)$ in state (x_0, x, b). The decision affects where the arrival is routed and it is selected from the set $\{\mathrm{s}, \mathrm{m}\}$, where the actions 's' and 'm' denote that the arriving flow is routed to the local small cell k or the macro cell, respectively. If the action is to serve the arrival in the small cell locally, the process makes a transition at rate λ_k to state $(x_0, x + e_k, b)$. If the flow is routed to the macrocell, the process makes a transition at rate λ_k to state $(x_0 + e_{0,k}, x, b)$. Note that we do not consider the re-routing of on-going flows from small cells to the macrocell or vice versa. The idle-timer control is performed each time a small cell k becomes idle, i.e., reaches a state with $x_k = 0$ and with $b_k = 1$. The idle timer rate ω_k of small cell k can be chosen from an arbitrary discrete set with lower bound ω_k^{L} and upper bound ω_k^{H}, i.e., $\omega_k \in \{\omega_k^{\mathrm{L}}, \ldots, \omega_k^{\mathrm{H}}\}$. We denote the chosen value of the idle timer as $a_k^{\omega}(x_0, x, b)$. Given an idle-timer action, the process moves at rate $a_k^{\omega}(x_0, x, b)$ to state $(x_0, x, b - e_k)$, i.e., the small cell k goes to off state.

We denote by π the set of all possible routing and idle-time control policies that are stable under the condition (1). For a given policy π, let $E[X^\pi]$ and $E[P^\pi]$ denote the mean total number of flows in the system and the resulting mean power consumption, respectively. Our objective is to consider the following optimization problem,

$$\min_{\pi} w_1 E[X^\pi] + w_2 E[P^\pi]. \tag{4}$$

Note that by dividing (4) with the total arrival rate $\sum_k \lambda_k$, the objective is, by Little's law, equal to to the weighted sum of the mean flow delay and the mean energy per flow. In this case, the weights w_1 and w_2 can be interpreted such they convert the delay and energy into equivalent monetary costs, i.e., the units of w_1 and w_2 would be cost/s and cost/J, respectively. This weighted form of the cost function, also sometimes in the literature referred to as ERWS (Energy Response time Weighted Sum), is one commonly used way to characterize the performance energy tradeoff, see [1, 4, 28] and the references therein.

4.2. Optimal dynamic policy

The optimization problem (4) is an MDP and can be solved iteratively with the policy iteration method [20]. In a generic form, the policy iteration works as follows. Consider a given state of the system $y = (x_0, x, b)$. In the routing and timer control policy, associated with each state y there is a set of actions $\mathcal{A}(y)$ relating to the decisions where the arrivals are routed and what timer value to select, in case a small cell is idle. Let π_n denote the policy at the n^{th} iteration step. The iterated policy at the next step, π_{n+1}, is obtained by solving in each state y the following optimization

$$\pi_{n+1}(y) = \operatorname*{arg\,min}_{a \in \mathcal{A}(y)} \left(c(y) - \bar{c}^{\pi_n} + \sum_{y'} q_{y,y'}(a)(v^{\pi_n}(y') - v^{\pi_n}(y)) \right), \forall y, \tag{5}$$

where $c(y)$ is the instantaneous cost in state y, $\bar{c}^{\pi_n}$ is the mean cost under policy π_n, $q_{y,y'}(a)$ is the transition rate from state y to state y' when action a is taken and $v^{\pi_n}(y')$ is the so-called value function of state y' for policy π_n, characterizing the future costs from the actions.

The future costs are characterized by the value function of each state. More precisely, the value function of a state y under given policy π gives the mean difference in the cost when starting initially the process from the state y and the long term average cost $\bar{c}^\pi$. The value function of each state is, on the other hand, characterized by the following set of linear equations

$$c(y) - \bar{c}^\pi + \sum_{y'} q_{y,y'}(a^\pi(y))(v^\pi(y') - v^\pi(y)) = 0, \ \forall y, \tag{6}$$

where $a^\pi(y)$ is the action taken in state y under policy π.

The policy iteration step (5) is in a generic form and next we define the detailed form of the policy iteration adapted to our problem to obtain insight. In the joint routing and timer control policy, associated with each state (x_0, x, b) there is a set of actions relating to the decisions, where the arrivals are routed, $a_k^{\mathrm{r}} \in \{\mathrm{s}, \mathrm{m}\}$ and for the possible idle cells the timer values to be used in that state, $a_k^{\omega} \in \{\omega_k^{\mathrm{L}}, \ldots, \omega_k^{\mathrm{H}}\}$. We denote by $y_k(a_k^{\mathrm{r}}, x_0, x, b)$ the vector state resulting from the routing action, i.e.,

$$y_k(a_k^{\mathrm{r}}, x_0, x, b) = \begin{cases} (x_0, x + e_k, b), & \text{if } a_k^{\mathrm{r}} = \mathrm{s}, \\ (x_0 + e_{0,k}, x, b), & \text{if } a_k^{\mathrm{r}} = \mathrm{m}. \end{cases}$$

The policy iteration step (5) can then be expressed as

$$\begin{aligned} \pi_{n+1}(x_0, x, b) = & \operatorname*{arg\,min}_{a_k^{\mathrm{r}} \in \{\mathrm{s},\mathrm{m}\}, a_k^{\omega} \in \{\omega_k^{\mathrm{L}}, \ldots, \omega_k^{\mathrm{H}}\}, k=1,\ldots,K} \Bigg(c(x_0, x, b) - \bar{c}^{\pi_n} \\ & + \sum_{k=0}^{K} 1_{x_{0,k} > 0} \mu_{0,k} (v^{\pi_n}(x_0 - e_{0,k}, x, b) - v^{\pi_n}(x_0, x, b)) \\ & + \sum_{k=1}^{K} 1_{x_k > 0, b_k = 1} \mu_k (v^{\pi_n}(x_0, x - e_k, b) - v^{\pi_n}(x_0, x, b)) \\ & + \sum_{k=1}^{K} 1_{x_k > 0, b_k = 0} \delta_k (v^{\pi_n}(x_0, x, b + e_k) - v^{\pi_n}(x_0, x, b)) \\ & + \lambda_0 (v^{\pi_n}(x_0 + e_{0,0}, x, b) - v^{\pi_n}(x_0, x, b)) \\ & + \sum_{k=1}^{K} \lambda_k (v^{\pi_n}(y_k(a_k^{\mathrm{r}}, x_0, x, b)) - v^{\pi_n}(x_0, x, b)) \\ & + \sum_{k=1}^{K} 1_{x_k = 0, b_k = 1} a_k^{\omega} (v^{\pi_n}(x_0, x, b - e_k) - v^{\pi_n}(x_0, x, b)) \Bigg), \forall (x_0, x, b). \end{aligned} \tag{7}$$

In (7), it can be immediately observed that the instantaneous cost rate relative to the mean cost rate (first line) and the future costs from the departures (2nd and 3rd lines), expiration of the setup delay (4th line) and arrivals in the macrocell class 0 are not affected by the actions and therefore do not influence the minimization.

In the following, we state an important structural result for the optimal policy denoted by π^*. We prove that in the optimal policy π^*, the optimal timer value $a_k^{\omega^*}$ is either ω_k^{L} or ω_k^{H}, i.e., the intermediate values of the set $\{\omega_k^{\mathrm{L}}, \ldots, \omega_k^{\mathrm{H}}\}$ are never used. Additionally, the routing and idle-timer control decisions can be done independently of each other.

Proposition 1. *In the optimum policy π^*, the optimal timer value in any state (x_0, x, b) with $x_k = 0$ and $b_k = 1$ is chosen from $a_k^{\omega^*} \in \{\omega_k^{\mathrm{L}}, \omega_k^{\mathrm{H}}\}$ for all k. Furthermore, the optimization with respect to the routing actions a_k^{r} and timer control actions $a_k^{\omega^*}$ can be done independently for all k.*

Proof. The optimal policy π^* satisfies the following optimality equation, see [20],

$$
\begin{aligned}
0 = & \min_{a_k^{\mathrm{r}} \in \{\mathrm{s},\mathrm{m}\}, a_k^{\omega} \in \{\omega_k^{\mathrm{L}},\dots,\omega_k^{\mathrm{H}}\}, k=1,\dots,K} \Bigg(c(x_0, x, b) - \bar{c}^{\pi^*} \\
& + \sum_{k=0}^{K} 1_{x_{0,k}>0} \mu_{0,k} (v^{\pi^*}(x_0 - e_{0,k}, x, b) - v^{\pi^*}(x_0, x, b)) \\
& + \sum_{k=1}^{K} 1_{x_k>0, b_k=1} \mu_k (v^{\pi^*}(x_0, x - e_k, b) - v^{\pi^*}(x_0, x, b)) \\
& + \sum_{k=1}^{K} 1_{x_k>0, b_k=0} \delta_k (v^{\pi^*}(x_0, x, b + e_k) - v^{\pi^*}(x_0, x, b)) \\
& + \lambda_0 (v^{\pi^*}(x_0 + e_{0,0}, x, b) - v^{\pi^*}(x_0, x, b)) \\
& + \sum_{k=1}^{K} \lambda_k (v^{\pi^*}(y_k(a_k^{\mathrm{r}}, x_0, x, b)) - v^{\pi^*}(x_0, x, b)) \\
& + \sum_{k=1}^{K} 1_{x_k=0, b_k=1} a_k^{\omega} (v^{\pi^*}(x_0, x, b - e_k) - v^{\pi^*}(x_0, x, b)) \Bigg), \forall (x_0, x, b).
\end{aligned} \tag{8}
$$

Observing (8), it is seen that the first 5 terms do not affect the minimization. Also, in any state (x_0, x, b), the routing decision $a_k^{\mathrm{r}} \in \{\mathrm{s}, \mathrm{m}\}$ related to arrivals in cell k only concerns whether the process makes a transition at rate λ_k to state $(x_0, x + e_k, b)$ or $(x_0 + e_{0,k}, x, b)$. Similarly, if in a given state (x_0, x, b), small cell k is idle with $x_k = 0$ and $b_k = 1$, the choice of the timer value $a_k^{\omega} \in \{\omega_k^{\mathrm{L}}, \dots, \omega_k^{\mathrm{H}}\}$ only affects the rate at which the small cell k makes its transition to the off state. Thus, we can express (8) as

$$
\begin{aligned}
& 0 = (\dots) + \sum_{k=1}^{K} \min_{a_k^{\mathrm{r}} \in \{\mathrm{s},\mathrm{m}\}} \lambda_k (v^{\pi^*}(y_k(a_k^{\mathrm{r}}, x_0, x, b)) - v^{\pi^*}(x_0, x, b)) \\
& + \sum_{k=1}^{K} \min_{a_k^{\omega} \in \{\omega_k^{\mathrm{L}},\dots,\omega_k^{\mathrm{H}}\}} 1_{x_k=0, b_k=1} a_k^{\omega} (v^{\pi^*}(x_0, x, b - e_k) - v^{\pi^*}(x_0, x, b)), \\
& \forall (x_0, x, b),
\end{aligned} \tag{9}
$$

where $(\dots)$ refers to the first 5 terms of (8). From (9), it is seen that the optimization problem becomes separable with respect to each action and there is no need to consider all possible combinations of actions in a given state.

Now consider a given small cell k in state $x_k = 0$ and $b_k = 1$, i.e., the idle state. In that state, the action space allows to modify the duration of the idle timer depending on the future costs from the set $a_k^{\omega} \in \{\omega_k^{\mathrm{L}}, \dots, \omega_k^{\mathrm{H}}\}$. Again, from (9), we observe that selecting a_k^{ω} affects linearly the future cost of moving to the off state, and hence future costs are minimized by setting the timer equal to either of the extreme values $\{\omega_k^{\mathrm{L}}, \omega_k^{\mathrm{H}}\}$, i.e, there is no benefit from using any of the intermediate values. Which one is optimal depends then on the difference in the value functions between the neighboring off and idle states, $(v(x_0, x, b-e_k) - v(x_0, x, b))$. If this difference is positive, then the optimal timer value $a_k^{\omega^*}$ minimizing the costs in the optimal policy π^* becomes $a_k^{\omega^*} = \omega_k^{\mathrm{L}}$, and if the difference is negative, the cost is minimized by setting $a_k^{\omega^*} = \omega_k^{\mathrm{H}}$. □

As the lower bound ω_k^{L} and upper bound ω_k^{H} of the timer are arbitrary, it seems reasonable that the optimal selection of the lower bound and upper bounds are $\omega_k^{\mathrm{L}} = 0$ and $\omega_k^{\mathrm{H}} = \infty$, respectively. Thus, each small cell is in the optimal policy always operating either as a NEVEROFF queue with $\omega_k = 0$ or an INSTANTOFF queue with $\omega_k = \infty$. This would extend the known result, see [13], that for a single server M/G/1-DELAYEDOFF PS-queue the optimal timer value for a weighted objective function, such as (4), is always either zero or infinite into the present more complex scenario with non-Poisson arrivals. However, the MDP framework does not allow to consider infinite transition rates, see [20], and thus we can only conjecture this asymptotic property.

The structural property in Proposition 1 helps us in simplifying the optimization problem. However, the complete characterization of the optimal policy π^* including the routing actions is intractable. Thus, in the following we consider a near-optimal policy based on the policy iteration idea.

5. Near-optimal FPI policy

In the FPI (First Policy Iteration) approach, the idea is that by selecting the initial policy appropriately the first step of the optimization (7) can be carried out explicitly. Typically this first step already gives the largest improvement, which makes the FPI policy near optimal.

5.1. Static load balancing initial policy

The initial policy to be defined is a static policy that is independent of the state of the system. It consists of a probabilistic policy for the routing decisions and static timer values for each small cell.

Consider first the routing policy which is a probabilistic policy determined by the vector $p = (p_1, \ldots, p_K)$, where each component p_k gives the probability to route the incoming class-k flow to the small cell k and with probability $(1 - p_k)$ the flow is routed to the macrocell. A reasonable selection for the initial probabilistic policy is to balance the load in all the cells, as much as possible. We denote this policy by p^{LB}. Assume that the classes of the small cells, $k = 1, \ldots, K$, are ordered in a descending order according to the cell loads, i.e.,

$$\frac{\lambda_1}{\mu_1} > \cdots > \frac{\lambda_K}{\mu_K}.$$

Note that for all those small cells where the small cell load λ_k/μ_k is already less than the load of the macrocell when serving its own traffic, $\lambda_0/\mu_{0,0}$, no traffic can be moved to the macrocell, i.e., the corresponding $p_k^{\mathrm{LB}} = 1$. Thus, let k^* denote the index value of the last small cell from which traffic can be moved to the macrocell, i.e.,

$$k^* = \max\{k = 1, \ldots, K : \lambda_k/\mu_k > \lambda_0/\mu_{0,0}\}.$$

It is easy to see that the load is then equalized by setting

$$\begin{cases} p_k^{\mathrm{LB}} = \frac{\mu_k}{\lambda_k} \cdot \frac{\frac{\lambda_0}{\mu_{0,0}} + \sum_{j=1}^{k^*} \frac{\lambda_j}{\mu_{0,j}}}{1 + \sum_{j=1}^{k^*} \frac{\mu_j}{\mu_{0,j}}}, & k = 1, \ldots, k^*, \\ p_k^{\mathrm{LB}} = 1, & k = k^* + 1, \ldots, K. \end{cases}$$

Additionally, the initial policy defines the timer configuration of each small cell. Let the vector $\omega^{\mathrm{LB}} = \{\omega_1^{\mathrm{LB}}, \ldots, \omega_K^{\mathrm{LB}}\}$ denote the timer value configuration of each small cell. The routing policy p^{LB} splits probabilistically the offered Poisson arrival rate λ_k and renders each small cell stochastically independent of each other. Thus, each small cell behaves as an independent M/M/1 DELAYEDOFF queue. Now, let $\bar{c}_k^{\mathrm{LB}}(\omega)$ denote a function for the mean cost of a small cell k with routing policy p^{LB} and timer value ω, which is given by, see [13],

$$\begin{aligned} \bar{c}_k^{\mathrm{LB}}(\omega) =& w_1 \left(\frac{p_k^{\mathrm{LB}}\lambda_k}{\mu_k - p_k^{\mathrm{LB}}\lambda_k} + \frac{1 + p_k^{\mathrm{LB}}\lambda_k/\delta_k}{1 + \delta_k/(p_k^{\mathrm{LB}}\lambda_k) + \delta_k/\omega} \right) + \\ & w_2 \left(\frac{p_k^{\mathrm{LB}}\lambda_k}{\mu_k} P^{\mathrm{b}} + \left(1 - \frac{p_k^{\mathrm{LB}}\lambda_k}{\mu_k}\right) \frac{P^{\mathrm{s}} + \delta_k P^{\mathrm{o}}/(p_k^{\mathrm{LB}}\lambda_k) + \delta_k P^{\mathrm{i}}/\omega}{1 + \delta_k/(p_k^{\mathrm{LB}}\lambda_k) + \delta_k/\omega} \right). \end{aligned} \tag{10}$$

As mentioned already earlier, in an M/G/1-DELAYEDOFF PS-queue, the optimal timer selection is either 0 or infinite, see [13]. Thus, in our case the timer value is selected between the extreme values $\{\omega_k^{\mathrm{L}}, \omega_k^{\mathrm{H}}\}$, and the optimal timer configuration ω^{LB} under routing policy p^{LB} is determined independently for all k by

$$\omega_k^{\mathrm{LB}} = \begin{cases} \omega_k^{\mathrm{L}}, & \text{if } \bar{c}_k^{\mathrm{LB}}(\omega_k^{\mathrm{L}}) \le \bar{c}_k^{\mathrm{LB}}(\omega_k^{\mathrm{H}}), \\ \omega_k^{\mathrm{H}}, & \text{otherwise.} \end{cases}$$

5.2. The FPI policy

Next we consider the policy iteration step (7) for the LB initial policy. Since the initial policy renders the stochastic behavior of the macrocell and the small cells independent of each other, the relative value of the state (x_0, x, b) under the initial policy can be expressed as

$$v(x_0, x, b) = v_0(x_0) + \sum_{k=1}^{K} v_k(x_k, b_k), \tag{11}$$

where $v_0(x_0)$ and $v_k(x_k, b_k)$ are the value functions of the macrocell and small cell k, respectively. The value functions are defined by the so-called Howard equations (6) and they have been derived in [12]. However, we have included them in this extended version of the paper for completeness, and they can be found in Appendix A.

In (7), the optimization related to the routing decision to serve the arrival in small cell k or route it to the macrocell simply consists of evaluating the additional cost of adding the arrival to the small cell or to the macrocell. In the first iteration of the optimization (7), due to the separable form of the value function (11), the action to serve the arrival in the macrocell or the small cell k in any state (x_0, x, b) is given by

$$a_k^{\mathrm{r}}(x_0, x, b) = \tag{12}$$
$$\begin{cases} \mathrm{m}, & \text{if } v_0(x_{0,0}, \ldots, x_{0,k}+1, \ldots, x_{0,K}) - v_0(x_{0,0}, \ldots, x_{0,K}) < \\ & \quad v_k(x_k+1, b_k) - v_k(x_k, b_k) \\ \mathrm{s}, & \text{otherwise.} \end{cases}$$

Thus, the future cost of adding the arrival to the macro cell or the small cell can be evaluated readily by considering the local marginal cost of one new flow in the macro cell or the small cell.

Similarly, in any state with $x_k = 0$ and $b_k = 1$, i.e., with the small cell k in the idle state, the action related to the optimal timer value is determined by

$$a_k^{\omega}(x_0, x, b) = \begin{cases} \omega_k^{\mathrm{L}}, & \text{if } v_k(0,0) - v_k(0,1) \geq 0, \\ \omega_k^{\mathrm{H}}, & \text{otherwise.} \end{cases} \tag{13}$$

Next we will state an important further structural property related to the optimal decision for the timer in (13). Namely, below we show that in the FPI policy the optimal timer value in fact coincides with that of the solution from the static load balancing policy, in other words, the vector ω^{LB} provides also the optimal timer values to be used in the FPI policy.

Proposition 2. *The optimal action for the timer values in the FPI policy is given by*

$$a_k^{\omega}(x_0, x, b) = \omega_k^{\mathrm{LB}}, \quad k = 1, \ldots, K.$$

Proof. As the timer control action only depends on the value function of a given small cell k, we omit the explicit depence on k in this proof. The small cell is represented by a DELAYEDOFF queue with parameters λ, μ, δ and ω. Consider first the case where $\omega^{\mathrm{LB}} = \omega^{\mathrm{L}}$. Thus, we have $\bar{c}^{\mathrm{LB}}(\omega^{\mathrm{L}}) \leq \bar{c}^{\mathrm{LB}}(\omega^{\mathrm{H}})$, which from (10) leads to the following inequality

$$\alpha \left(-\frac{\lambda\mu(\delta+\lambda)w_1}{(\mu-\lambda)\delta} + ((\delta+\lambda)P^{\mathrm{i}} - \delta P^{\mathrm{o}} - \lambda P^{\mathrm{s}})w_2 \right) \leq 0,$$

where

$$\alpha = \frac{(\omega^{\mathrm{H}} - \omega^{\mathrm{L}})(\mu-\lambda)\delta\lambda}{(\delta\lambda + \delta\omega^{\mathrm{H}} + \lambda\omega^{\mathrm{H}})(\delta\lambda + \delta\omega^{\mathrm{L}} + \lambda\omega^{\mathrm{L}})} \geq 0.$$

From this it follows that

$$\frac{\lambda\mu(\delta+\lambda)w_1}{(\mu-\lambda)\delta} \geq ((\delta+\lambda)P^{\mathrm{i}} - \delta P^{\mathrm{o}} - \lambda P^{\mathrm{s}})w_2. \tag{14}$$

By utilizing the expressions for the marginal performance and energy costs in Corollary 1 and 2 in Appendix A, the explicit form of the marginal cost in the timer action in (13) when assuming that $\omega^{\mathrm{LB}} = \omega^{\mathrm{L}}$ becomes

$$v_k(0,0) - v_k(0,1) = \\ \frac{1}{(\delta\lambda + \delta\omega^{\mathrm{L}} + \lambda\omega^{\mathrm{L}})}\left(\frac{\lambda\mu(\delta+\lambda)w_1}{(\mu-\lambda)\delta} - ((\delta+\lambda)P^{\mathrm{i}} - \delta P^{\mathrm{o}} - \lambda P^{\mathrm{s}})w_2\right)$$

By (14) it follows that $v_k(0,0) - v_k(0,1) \geq 0$ and thus by (13) the optimal selection of the timer in the FPI policy is ω^{L}, which is the same as the original choice of ω^{LB}, in this case.

On the other hand, assuming the opposite that $\bar{c}^{\mathrm{LB}}(\omega^{\mathrm{L}}) > \bar{c}^{\mathrm{LB}}(\omega^{\mathrm{H}})$, i.e., $\omega^{\mathrm{LB}} = \omega^{\mathrm{H}}$, leads to

$$\frac{\lambda\mu(\delta+\lambda)w_1}{(\mu-\lambda)\delta} < ((\delta+\lambda)P^{\mathrm{i}} - \delta P^{\mathrm{o}} - \lambda P^{\mathrm{s}})w_2. \tag{15}$$

In this case, the marginal cost in the timer action (13) equals

$$v_k(0,0) - v_k(0,1) = \\ \frac{1}{(\delta\lambda + \delta\omega^{\mathrm{H}} + \lambda\omega^{\mathrm{H}})}\left(\frac{\lambda\mu(\delta+\lambda)w_1}{(\mu-\lambda)\delta} - ((\delta+\lambda)P^{\mathrm{i}} - \delta P^{\mathrm{o}} - \lambda P^{\mathrm{s}})w_2\right)$$

By (15) it follows that $v_k(0,0) - v_k(0,1) < 0$ and thus by (13) the optimal selection of the timer in the FPI policy is ω^{H}, which is the same as the original choice of ω^{LB}, in this case. This completes the proof. □

Finally, the routing actions defined by (12) and, by Proposition 2, the initial timer configuration ω^{LB} together define the FPI policy. In addition to the state of the system, the value functions needed in the routing action (12) depend on the parameters of the initial policy, i.e., the routing probabilities p^{LB} and the initial timer configuration ω^{LB}.

In summary, the main advantage in the FPI approach is that it allows us to systematically construct an optimized dynamic policy, which is near optimal and (almost) fully explicit, see Appendix A. Note that through our results, the FPI policy is also fully specified in the entire state space of the system, i.e., there is no need for any truncation to evaluate the mean cost under the FPI policy by simulation.

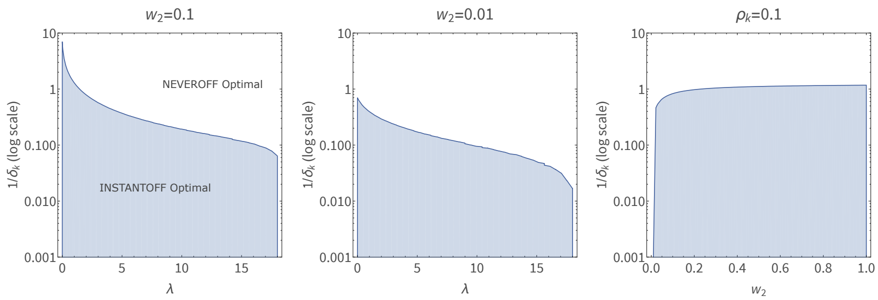

Figure 1: Region plots for the optimal configuration choices. The blue shaded region indicates the parameter space where INSTANTOFF is optimal while NEVEROFF is superior in the remaining parameter space.

6. Numerical results

Here, we study the performance of the proposed policies through simulation runs. We consider a system consisting of one macro and seven small cells. This is motivated by a hexagonal setting studied in [26], where six small cells are located at the edges of the macro cell's hexagonal coverage area while the seventh small cell resides at the center.

Small cells can be in an off state where they do not consume power. Otherwise, power consumption of small cell k in the busy, setup and idle states are $P_k^{\mathrm{b}} = P_k^{\mathrm{s}} = 100$ and $P_k^{\mathrm{i}} = 70$ W, respectively. The macro cell consumes 1000 W when it is busy and 700 W when idle. These values are within the range of power consumption values considered in [3, 26].

Service rate for a request originating from within a small cell is $\mu_k = 18.73$ s^{-1} if it is served by the small cell. For small cells at the edges, the traffic offloaded to the macro is served with a rate $\mu_{0,k} = 6.37$ s^{-1}, whereas that of the central small cell is served with a rate $\mu_{0,1} = 12.34$ s^{-1}. Additionally, the macro cell also serves users that are outside the coverage area of the small cells with a rate $\mu_0 = 12.34$ s^{-1}. The service rates have been obtained by assuming file sizes of 5 Mb and typical measured mean channel qualities, see [26] for more detailed justifications. In the entire simulation study we set arrival rate in the macro cell to $\lambda_0 = 1$ s^{-1} and $w_1 = 1$, for the system cost given by (4). Simulations are run as a function of the load. At each load point, each simulation run consists of 10^6 arrivals.

Considering the load balancing initial policy, the optimal configuration of a small cell is NEVEROFF when $\bar{c}_k^{LB}(0) \leq \bar{c}_k^{LB}(\infty)$, and INSTANTOFF otherwise, see (10). With the other system parameters fixed, the optimal configuration may change as a function of the mean setup delay ($1/\delta_k$), the energy weight (assuming w_1 is fixed), and the small cell load parameters. The shaded region in Figure 1 shows the parameter space in which $\bar{c}_k^{LB}(0) > \bar{c}_k^{LB}(\infty)$ when one of the parameters is fixed and the other two are varied. The figure clearly shows that the optimal configuration is largely influenced by the setup delay.

For example, when $1/\delta_k > 7$ s, NEVEROFF remains optimal even for a large energy weight of $w_2 = 0.1$ at a load value as low as 10^{-6}. The right most figure affirms this observation by showing that increasing w_2 has little effect on the choice of optimal configuration for $\rho_k = 0.1$. In our simulation study sufficiently large ($1/\omega = 10^6$) and sufficiently small ($1/\omega = 10^{-6}$) timer values are chosen to represent the NEVEROFF and INSTANTOFF configurations, respectively.

The five scenarios presented in Table 1 are chosen for the numerical illustration. The $1/\delta_k$ and w_2 values are selected so that representative points from the optimality regions of both NEVEROFF and INSTANTOFF are covered, see Figure 1. Three of the scenarios assume the arrival rate in all the small cells is the same (homogeneous), whereas in the remaining two scenarios the small cells have different arrival rates (heterogeneous).

Scenario	Parameters	Description
Hom1	$1/\delta_k = 10$ s $w_2 = 0.01$	Homogeneous system one
Hom2	$1/\delta_k = 1$ s, $w_2 = 0.1$	Homogeneous system two
Hom3	$1/\delta_k = 0.1$ s, $w_2 = 0.01$	Homogeneous system three
Het1	$1/\delta_k = 10$ s, $w_2 = 0.01$	Heterogeneous system one
Het2	$1/\delta_k = 0.1$ s, $w_2 = 0.01$	Heterogeneous system two

Table 1: Scenarios considered for numerical study.

In addition to the load balancing initial policy (RND), we also study the FPI and Join the Shortest Queue (JSQ). As dispatching decisions in JSQ are oblivious to the energy configuration of the small cell, two cases are considered: JSQ with all the small cells configured as NEVEROFF (which we refer to as JSQ-NO) and JSQ with optimally configured small cells. The performance of each of these policies is studied using the mean response time, mean power, and the weighted sum cost metrics.

6.1. Homogeneous load at small cells

Consider a case where the arrival rate in all the small cells is uniform. The arrival rate at each small cell is increased from 1 to 17 s^{-1} simultaneously and the resulting system performance is studied. In this case, the system stability limit is $\lambda_k < 19.6283$ s^{-1} for $1 \le k \le 7$, see (1).

Figure 2 shows the relevant metrics as a function of arrival rate at small cells when $1/\delta_k = 10$ s and $w_2 = 0.01$ as in scenario Hom1. The top two plots show mean response time and mean power consumption of the system. The bottom left figure depicts the weighted sum cost of the policies under study, whereas the bottom right figure shows the cost relative to the initial policy RND. The weighted sum cost of RND is represented by the dotted gray line at 1. A relative cost value less than 1 indicates that the respective policy yields lower cost than RND at that specific load value under the given system parameter values.

In this case, the optimal configuration is NEVEROFF for the LB initial policy over the entire load spectrum, see Figure 1. Thus, JSQ and JSQ-NO operate in an identical manner, which explains why they have overlapping curves in the figure. The relative cost plot shows that FPI saves up to 24% on weighted

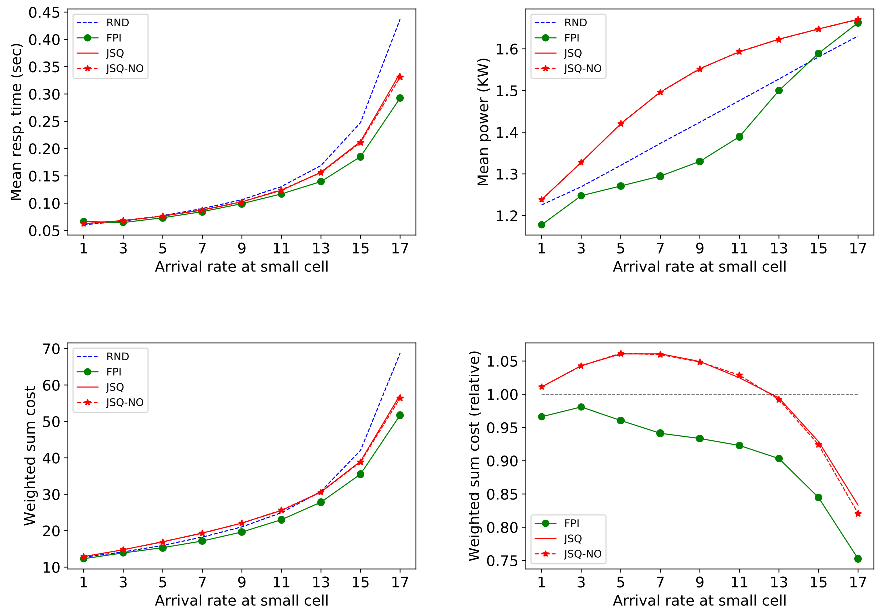

Figure 2: Mean response time, mean power consumption, weighted sum cost and relative weighted sum cost as a function of arrival rate for scenario Hom1. The bottom right plot shows the weighted sum cost of the FPI, JSQ and JSQ-NO policies relative to the load balancing initial policy (RND).

sum cost compared to the initial load balancing policy (RND). It achieves this by utilizing small cells as much as possible and avoiding the high busy power at the macro. Especially, when the load is light, FPI dispatches all incoming traffic to the small cells. RND also routes most of the traffic to the small cells at low load, but it starts using the macro more often as load increases. Due to this reason, the gain of FPI over RND increases with load. JSQ performs marginally better than RND with respect to response time but consumes more energy as it uses the macro more often than RND.

Figure 3 depicts the system cost metrics when $1/\delta_k = 1$ s and $w_2 = 0.1$ (scenario Hom2). The optimal configuration starts with being INSTANTOFF at light load and quickly changes to NEVEROFF as arrival rate reaches 1.4487 s^{-1}, see Figure 1. This is illustrated in the figure by the sharp drop in mean response time and the corresponding increase in mean power as arrival rate changes from 1 to 3 s^{-1}. Mean response time and mean power increase in a less steep manner thereafter. The FPI policy achieves the lowest weighted sum cost over RND with up to 10% improvement.

Figure 4 shows the cost metrics of a system with a very short setup delay $1/\delta_k = 0.1$ (Scenario Hom3). In this case, INSTANTOFF is optimal when the small cell traffic is low/moderate. For arrival rate $\lambda_k > 9.704$ s^{-1}, the optimal

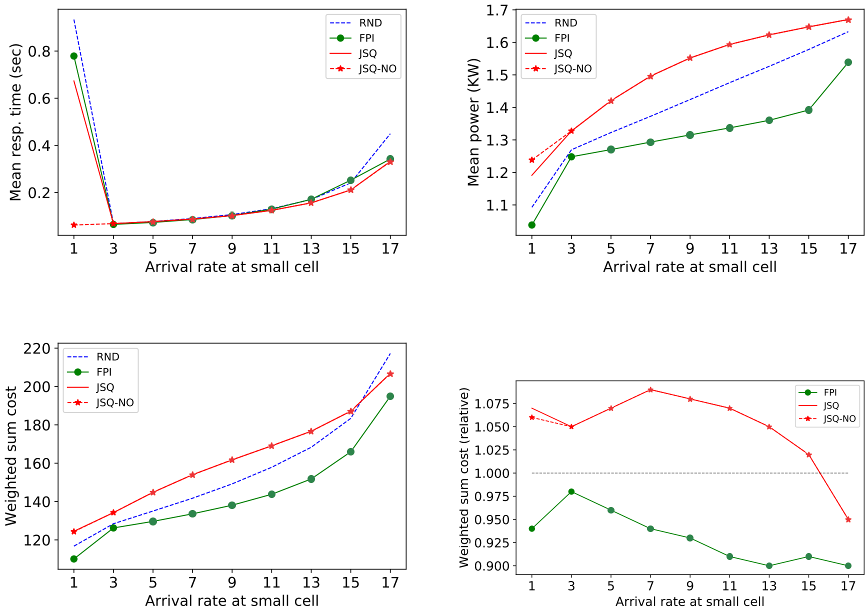

Figure 3: Mean response time, mean power consumption, weighted sum cost and relative weighted sum cost as a function of arrival rate for scenario Hom2. The bottom right plot shows the weighted sum cost of the FPI, JSQ and JSQ-NO policies relative to the load balancing initial policy (RND).

configuration changes to NEVEROFF, which explains the sharp decrease in mean response time around that region. FPI yields up to 24% improvement over RND. Figures 2 - 4 consistently illustrate that the gain of FPI over RND increases with load. More importantly, these gains are achieved by improving both the mean response time and mean power consumption. In most cases, FPI also provides lower cost relative to JSQ and JSQ-NO.

6.2. Heterogeneous load at small cells

This section applies the hexagonal setting introduced earlier and considers a specific setting where the arrival rate at the central small cell is fixed at $\lambda_1 = 9$ s^{-1}. resulting in a load value of $\rho_1 = 0.48$. Additionally, the load on two of the remaining six small cells is fixed at $\rho_2 = \rho_3 = 0.1$. The arrival rate at the other four small cells is gradually increased and the performance of the four dispatching policies is studied. The system remains stable when $\lambda_k < 20.1934$ s^{-1} for $4 \leq k \leq 7$.

Figure 5 shows the cost metrics of the heterogeneous system described above when $1/\delta_k = 10$ s and $w_2 = 0.01$ as in scenario Het1. Due to the high setup delay, the optimal configuration is NEVEROFF for all cells and load values. Once again, energy (and overall cost) savings of FPI over RND come from

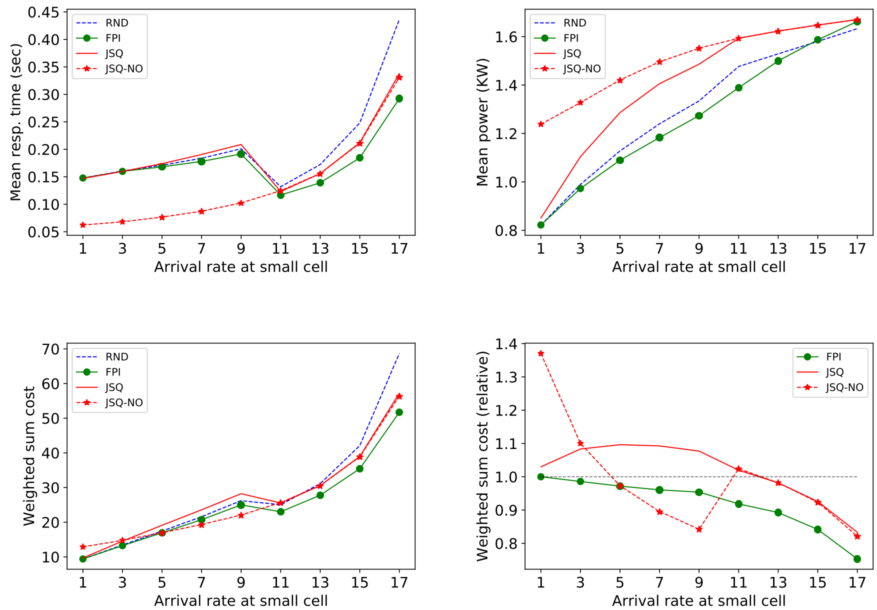

Figure 4: Mean response time, mean power consumption, weighted sum cost and relative weighted sum cost as a function of arrival rate for scenario Hom3.

routing as little traffic as possible to the macro. In this case, FPI achieves up to 40% cost reduction over RND. Recall that a smaller margin was achieved in the homogeneous setting. Heterogeneity of small cell load has possibly given the FPI policy more room for optimizing over the RND policy. Additionally, this improvement is achieved for the most part by improving both mean response time and mean power consumption.

Figure 6 shows the cost metrics of the heterogeneous system with $1/\delta_k = 0.1$ s and $w_2 = 0.01$ (scenario Het2). With the setup delay being very short, the optimal configuration remains INSTANTOFF until the arrival rate is increased up to 9 s^{-1} and switches to NEVEROFF afterwards. Note that this switch happens only for the four small cells. Since the load on the central and the other two cells is fixed, they remain as INSTANTOFF cells. This is also illustrated by the narrowing mean response time gap between JSQ and JSQ-NO at $\lambda_k = 11$ s^{-1}. Generally, we observe similar behavior as already seen earlier.

7. Conclusions

We have considered energy-aware joint control of routing and base station switch-off in a system consisting of a macrocell with several small cells inside its coverage area. The system is modeled as a set of parallel queues consisting of a multiclass M/M/1-PS queue, representing the always-on macrocell, and each

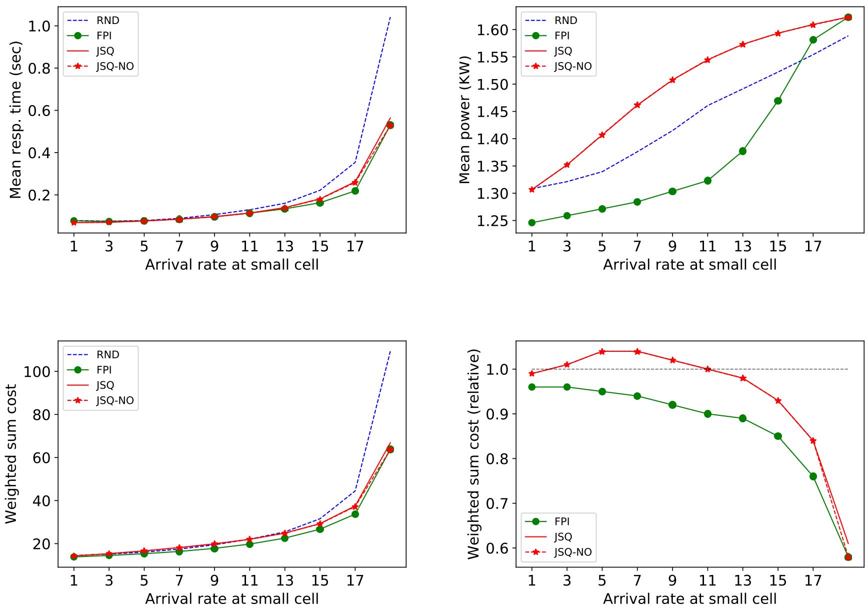

Figure 5: Mean response time, mean power consumption and weighted sum cost as a function of arrival rate at four of the small cells ($k = \{4, ...7\}$) for scenario Het1.

small cell is characterized by an energy-aware M/M/1-PS queue with an off state and setup delay. As an additional control feature, the model of the small cells included an idle timer, which is used for controlling how long the small cell waits after the end of a busy period until it falls into the off state.

Our main results relate to the fundamental structural properties of the optimal timer value selection. The joint control problem for routing and idle timer value selection was analyzed by applying the theory of MDPs. We showed that for a weighted sum of the performance and energy costs the routing and idle timer control actions can be optimized individually without the need to consider combinations of different actions in a state. More importantly, the optimal dynamic policy always selects the timer value to be either arbitrarily close to zero or infinite, suggesting that each individual small cell always behaves in our model as either a NEVEROFF or INSTANTOFF queue. Furthermore, by applying the policy iteration method once when starting from an initial probabilistic policy, we showed that the optimal timer configuration is in the dynamic FPI policy also static and coincides with the solution of the initial policy. These results significantly simplify the overall optimization problem. The FPI policy then finally consists of a static selection of the idle timer values and the dynamic policy for the routing decisions, for which explicit expressions were derived. Our numerical results demonstrated that the FPI policy is able to achieve considerable cost reductions compared with other reference policies.

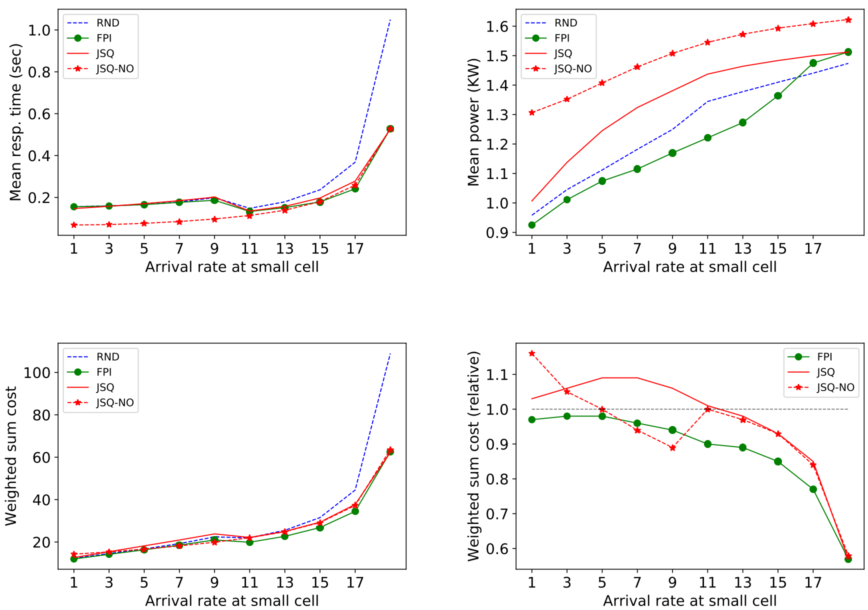

Figure 6: Mean response time, mean power consumption and weighted sum cost as a function of arrival rate at four of the small cells ($k = \{4, ...7\}$) for scenario Het2.

Possible future research topics include analyzing the impact of non-exponential service time distributions as well as interference between the base stations. However, these extensions are analytically very difficult to handle, but simulations can be used to this end.

Acknowledgements

This research has been partially supported by EIT Digital under the HII ACTIVE project and by the Academy of Finland under the ITTECH5G project (Grant No. 284735).

References

[1] S. Albers and H. Fujiwara. Energy-efficient algorithms for flow time minimization. *ACM Trans. Algorithms*, 3(4), Nov. 2007.

[2] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang. What will 5G be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, June 2014.

[3] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. Imran, D. Sabella, M. Gonzalez, O. Blume, et al. How much energy is needed to run a wireless network? *IEEE Wireless Communications*, 18(5):40–49, 2011.

[4] N. Bansal, H.-L. Chan, and K. Pruhs. Speed scaling with an arbitrary power function. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 693–701, 2009.

[5] T. Bonald and A. Proutière. Wireless downlink data channels: user performance and cell dimensioning. In *Proc. of ACM MobiCom*, pages 339–352, September 2003.

[6] S. Cai, Y. Che, L. Duan, J. Wang, S. Zhou, and R. Zhang. Green 5G Heterogeneous Networks Through Dynamic Small-Cell Operation. *IEEE Journal on Selected Areas in Communications*, 34(5):1103–1115, 2016.

[7] G. Carvalho, I. Woungang, A. Anpalagan, and E. Hossain. QoS-aware energy-efficient joint radio resource management in Multi-RAT heterogeneous networks. *IEEE Transactions on Vehicular Technology*, 65(8):6343–6365, 2016.

[8] R. Combes, S. E. Elayoubi, A. Ali, L. Saker, and T. Chahed. Optimal online control for sleep mode in green base stations. *Computer Networks*, 78:140 – 151, 2015.

[9] A. D. Domenico, V. Savin, D. Ktenas, and A. Maeder. Backhaul-aware small cell dtx based on fuzzy q-learning in heterogeneous cellular networks. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6, May 2016.

[10] A. Gandhi, V. Gupta, M. Harchol-Balter, and M. Kozuch. Optimality analysis of energy-performance trade-off for server farm management. *Performance Evaluation*, 67:1155–1171, 2010.

[11] M. Gebrehiwot, S. Aalto, and P. Lassila. Near-optimal policies for energy-aware task assignment in server farms. In *Proc. of CCGrid*, pages 1017–1026, May 2017.

[12] M. Gebrehiwot, P. Lassila, and S. Aalto. Dynamic load balancing in 5G HetNets for optimal performance-energy tradeoff. In *Proc. of IFIP NETWORKING*, pages 271–279, May 2018. available from: `http://www.netlab.tkk.fi/~pelassil/lassila-networking2018.pdf`.

[13] M. E. Gebrehiwot, S. Aalto, and P. Lassila. Energy-performance trade-off for processor sharing queues with setup delay. *Operations Research Letters*, 44(1):101–106, 2016.

[14] F. Han, S. Zhao, L. Zhang, and J. Wu. Survey of strategies for switching off base stations in heterogeneous networks for greener 5G systems. *IEEE Access*, 4:4959–4973, 2016.

[15] M. F. Hossain, K. S. Munasinghe, and A. Jamalipour. Distributed inter-BS cooperation aided energy efficient load balancing for cellular networks. *IEEE Transactions on Wireless Communications*, 12(11):5929–5939, 2013.

[16] A. Khalid, P. Lassila, and S. Aalto. Load balancing of elastic data traffic in heterogeneous wireless networks. In *Proc. of ITC*, September 2013.

[17] E. Khloussy, X. Gelabert, and Y. Jiang. Investigation on MDP-based radio access technology selection in heterogeneous wireless networks. *Computer Networks*, 91:57–67, 2015.

[18] J. Leino and J. Virtamo. Determining the moments of queue-length distribution of discriminatory processor-sharing systems with phase-type service requirements. In *Proc. of NGI*, pages 205–208, May 2007.

[19] E. Oh, K. Son, and B. Krishnamachari. Dynamic base station switching-on/off strategies for green cellular networks. *IEEE Transactions on Wireless Communications*, 12(5):2126–2136, 2013.

[20] M. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Wiley, 2005.

[21] A. Rangisetti, T. Pasca, and B. R. Tamma. QoS Aware load balance in software defined LTE networks. *Computer Communications*, 97:52–71, 2017.

[22] A. Roy, P. Chaporkar, and A. Karandikar. Optimal radio access technology selection algorithm for LTE-WiFi network. *IEEE Transactions on Vehicular Technology*, 67(7):6446–6460, July 2018.

[23] A. Roy and A. Karandikar. Optimal radio access technology selection policy for LTE-WiFi network. In *Proc. of WiOpt*, pages 291–298, May 2015.

[24] L. Saker, S. E. Elayoubi, R. Combes, and T. Chahed. Optimal control of wake up mechanisms of femtocells in heterogeneous networks. *IEEE Journal on Selected Areas in Communications*, 30(3):664–672, April 2012.

[25] Y. Song, P. Y. Kong, and Y. Han. Potential of network energy saving through handover in HetNets. *IEEE Transactions on Vehicular Technology*, 65(12):10198–10204, 2016.

[26] I. Taboada, S. Aalto, P. Lassila, and F. Liberal. Delay- and energy-aware load balancing in ultra-dense heterogeneous 5G networks. *Transactions on Emerging Telecommunications Technologies*, 28:e3170, 2017.

[27] P. K. Taksande, A. Roy, and A. Karandikar. Optimal traffic splitting policy in LTE-based heterogeneous network. In *Proc. of IEEE Wireless Communications and Networking Conference*, pages 1–6, April 2018.

[28] A. Wierman, L. L. H. Andrew, and A. Tang. Power-aware speed scaling in processor sharing systems. In *IEEE INFOCOM 2009*, pages 2007–2015, April 2009.

[29] Q. Wu, G. Y. Li, W. Chen, D. W. K. Ng, and R. Schober. An overview of sustainable green 5G networks. *IEEE Wireless Communications*, 24(4):72–80, Aug 2017.

[30] N. Yu, Y. Miao, L. Mu, H. Du, H. Huang, and X. Jia. Minimizing energy cost by dynamic switching ON/OFF base stations in cellular networks. *IEEE Transactions on Wireless Communications*, 15(11):7457–7469, Nov 2016.

[31] J. Zheng, Y. Cai, X. Chen, R. Li, and H. Zhang. Optimal base station sleeping in green cellular networks: A distributed cooperative framework based on game theory. *IEEE Transactions on Wireless Communications*, 14(8):4391–4406, 2015.

Appendix A. Value functions

Here we present the explicit value functions for the performance and energy for the M/M/1-PS DELAYDOFF and multiclass M/M/1-PS models. They appeared already in our earlier paper [12], but they are included here for completeness, and the present paper is an extended version of that article.

Appendix A.1. M/M/1-PS DELAYEDOFF

Consider a generic M/M/1-PS DELAYEDOFF queue with arrival rate λ, mean service time $E[S] = 1/\mu$, mean idle timer $E[I] = 1/\omega$, and mean setup delay $E[D] = 1/\delta$. Assume that the system is stable, i.e., $\rho < 1$, where load $\rho = \lambda E[S]$. In addition, let P^{o}, P^{i}, P^{s}, and P^{b} denote the power consumption in energy states OFF, IDLE, SETUP, and BUSY, respectively.

Proposition 3. *For a stable M/M/1-PS DELAYEDOFF queue, the relative* *value function with respect to performance in state (n, b) is given by*[2]

$$
\begin{aligned}
v^{\mathrm{p}}(n,1) - v^{\mathrm{p}}(0,0) &= \frac{E[S]n(n+1)}{2(1-\rho)} - \frac{E[D]n\rho(1+\lambda E[D])}{(1-\rho)(1+\lambda E[D]+\lambda E[I])} \\
&\quad - \frac{E[I]\lambda E[D](1+\lambda E[D])}{(1-\rho)(1+\lambda E[D]+\lambda E[I])}; \\
v^{\mathrm{p}}(n,0) - v^{\mathrm{p}}(0,0) &= \frac{E[S]n(n+1)}{2(1-\rho)} + \frac{E[D]n(1+\lambda E[D])}{1+\lambda E[D]+\lambda E[I]} \\
&\quad + \frac{E[I](n-1)\lambda E[D]}{(1-\rho)(1+\lambda E[D]+\lambda E[I])}.
\end{aligned}
$$

Proof. The Howard equations for the system are:

$$
\begin{aligned}
&-\bar{c}^{\mathrm{p}} + \lambda(v^{\mathrm{p}}(1,0) - v^{\mathrm{p}}(0,0)) = 0, \\
&n - \bar{c}^{\mathrm{p}} + \lambda(v^{\mathrm{p}}(n+1,0) - v^{\mathrm{p}}(n,0)) \ + \delta(v^{\mathrm{p}}(n,1) - v^{\mathrm{p}}(n,0)) = 0, \quad n \geq 1, \\
&n - \bar{c}^{\mathrm{p}} + \lambda(v^{\mathrm{p}}(n+1,1) - v^{\mathrm{p}}(n,1)) \ + \mu(v^{\mathrm{p}}(n-1,1) - v^{\mathrm{p}}(n,1)) = 0, \quad n \geq 1, \\
&-\bar{c}^{\mathrm{p}} + \lambda(v^{\mathrm{p}}(1,1) - v^{\mathrm{p}}(0,1)) \ + \omega(v^{\mathrm{p}}(0,0) - v^{\mathrm{p}}(0,1)) = 0,
\end{aligned}
$$

where $\bar{c}^{\mathrm{p}}$ denotes the mean number of flows given by

$$
\bar{c}^{\mathrm{p}} = E[X] = \lambda E[T] = \frac{\rho}{1-\rho} + \frac{\lambda E[D](1+\lambda E[D])}{1+\lambda E[D]+\lambda E[I]}.
$$

Now it is a straightforward task to verify that these equations are satisfied by the proposed relative value function. Note that $\bar{c}^{\mathrm{p}}$ appeared in a slightly different form already as part of the definition of $\bar{c}_k^{\mathrm{LB}}(\omega)$ in (10). It is restated here for completeness. □

Corollary 1. *For a stable M/M/1-PS DELAYEDOFF queue, the marginal performance cost in state (n, b) is given by*

$$
\begin{aligned}
&v^{\mathrm{p}}(n+1,1) - v^{\mathrm{p}}(n,1) = \frac{E[S](n+1)}{1-\rho} - \frac{E[D]\rho(1+\lambda E[D])}{(1-\rho)(1+\lambda E[D]+\lambda E[I])}, \quad n \geq 0; \\
&v^{\mathrm{p}}(1,0) - v^{\mathrm{p}}(0,0) = \frac{E[S]}{1-\rho} + \frac{E[D](1+\lambda E[D])}{1+\lambda E[D]+\lambda E[I]}; \\
&v^{\mathrm{p}}(n+1,0) - v^{\mathrm{p}}(n,0) = \frac{E[S](n+1)}{1-\rho} + \frac{E[D](1+\lambda E[D])}{1+\lambda E[D]+\lambda E[I]} \\
&\quad + \frac{E[I]\lambda E[D]}{(1-\rho)(1+\lambda E[D]+\lambda E[I])}, \quad n \geq 1; \\
&v^{\mathrm{p}}(0,0) - v^{\mathrm{p}}(0,1) = \frac{E[I]\lambda E[D](1+\lambda E[D])}{(1-\rho)(1+\lambda E[D]+\lambda E[I])};
\end{aligned}
$$

[2]Note that our earlier paper [12] has a misprint and is lacking the term $(1+\lambda E[D])$ in the numerator of the 3rd term in $v^{\mathrm{p}}(n,1) - v^{\mathrm{p}}(0,0)$.

Proposition 4. *For a stable M/M/1-PS DELAYEDOFF queue, the relative value function with respect to the energy in state (n, b) is given by*[3]

$$v^{\mathrm{e}}(n,1) - v^{\mathrm{e}}(0,0) = n \cdot \gamma + \frac{E[I]((P^{\mathrm{i}} - P^{\mathrm{o}}) + \lambda E[D](P^{\mathrm{i}} - P^{\mathrm{s}}))}{1 + \lambda E[D] + \lambda E[I]},$$
$$v^{\mathrm{e}}(n,0) - v^{\mathrm{e}}(0,0) = n \cdot \gamma + \frac{E[D](P^{\mathrm{s}} - P^{\mathrm{o}}) + E[I](P^{\mathrm{i}} - P^{\mathrm{o}})}{1 + \lambda E[D] + \lambda E[I]},$$

where

$$\gamma = \frac{E[S]((P^{\mathrm{b}} - P^{\mathrm{o}}) + \lambda E[D](P^{\mathrm{b}} - P^{\mathrm{s}}) + \lambda E[I](P^{\mathrm{b}} - P^{\mathrm{i}}))}{1 + \lambda E[D] + \lambda E[I]}$$

Proof. The Howard equations for the system are:

$$\begin{aligned}
&P^{\mathrm{o}} - \bar{c}^{\mathrm{e}} + \lambda(v^{\mathrm{e}}(1,0) - v^{\mathrm{e}}(0,0)) = 0,\\
&P^{\mathrm{s}} - \bar{c}^{\mathrm{e}} + \lambda(v^{\mathrm{e}}(n+1,0) - v^{\mathrm{e}}(n,0)) \ + \delta(v^{\mathrm{e}}(n,1) - v^{\mathrm{e}}(n,0)) = 0, \quad n \geq 1,\\
&P^{\mathrm{b}} - \bar{c}^{\mathrm{e}} + \lambda(v^{\mathrm{e}}(n+1,1) - v^{\mathrm{e}}(n,1)) \ + \mu(v^{\mathrm{e}}(n-1,1) - v^{\mathrm{e}}(n,1)) = 0, \quad n \geq 1,\\
&P^{\mathrm{i}} - \bar{c}^{\mathrm{e}} + \lambda(v^{\mathrm{e}}(1,1) - v^{\mathrm{e}}(0,1)) \ + \omega(v^{\mathrm{e}}(0,0) - v^{\mathrm{e}}(0,1)) = 0,
\end{aligned}$$

where $\bar{c}^{\mathrm{e}}$ denotes the mean power consumption given by

$$\bar{c}^{\mathrm{e}} = E[P] = \rho P^{\mathrm{b}} + (1-\rho)\frac{P^{\mathrm{o}} + \lambda E[D]P^{\mathrm{s}} + \lambda E[I]P^{\mathrm{i}}}{1 + \lambda E[D] + \lambda E[I]}.$$

It is again a straightforward task to verify that these equations are satisfied by the proposed relative value function. Note that $\bar{c}^{\mathrm{e}}$ appeared in a slightly different form already as part of the definition of $\bar{c}_k^{\mathrm{LB}}(\omega)$ in (10). It is restated here for completeness. □

Corollary 2. *For a stable M/M/1-PS DELAYEDOFF queue, the marginal energy cost in state (n, b) is given by*

$$\begin{aligned}
&v^{\mathrm{e}}(n+1,1) - v^{\mathrm{e}}(n,1) = \gamma, \quad n \geq 0;\\
&v^{\mathrm{e}}(1,0) - v^{\mathrm{e}}(0,0) = \gamma + \frac{E[D](P^{\mathrm{s}} - P^{\mathrm{o}}) + E[I](P^{\mathrm{i}} - P^{\mathrm{o}})}{1 + \lambda E[D] + \lambda E[I]},\\
&v^{\mathrm{e}}(n+1,0) - v^{\mathrm{e}}(n,0) = \gamma, \quad n \geq 1,\\
&v^{\mathrm{e}}(0,0) - v^{\mathrm{e}}(0,1) = -\frac{E[I]((P^{\mathrm{i}} - P^{\mathrm{o}}) + \lambda E[D](P^{\mathrm{i}} - P^{\mathrm{s}}))}{1 + \lambda E[D] + \lambda E[I]}.
\end{aligned}$$

where γ is defined in Proposition 4.

[3] Note that our earlier paper [12] has a misprint and has extra parentheses in the numerator of the 2nd term of $v^{\mathrm{e}}(n,0) - v^{\mathrm{e}}(0,0)$.

The results in Corollary 1 and 2 are used when evaluating the marginal cost of serving the flow in small cell k in (12) and the optimal timer action in (13). In Corollary 1 and 2, the first three marginal cost expressions are needed for the routing decisions while the last, fourth expression, is needed in the timer-control decision. Finally, in the marginal cost expressions of Corollary 1 and 2, when applied to small cell k with the initial LB policy, the parameters are set as $\lambda = p_k^{\mathrm{LB}}\lambda_k$, $E[S] = 1/\mu_k$, $E[D] = 1/\delta_k$ and the timer value $E[I] = 1/\omega_k^{\mathrm{LB}}$.

Also, observe that in Corollary 1, the marginal cost with respect to the performance is linear in the number of jobs, i.e., similarly as in the JSQ rule (Join-the-Shortest-Queue), but in addition there is a positive or negative constant factor depending on whether the server is in off/setup state or busy state. Thus, it is from the future cost better to keep an already busy server busy than to wake it up, which is also logical. On the other hand, by Corollary 2, the marginal energy cost is interestingly constant and does not depend on busy/setup state, unless the server is switched off.

Appendix A.2. MULTICLASS M/M/1-PS NEVEROFF

Consider a generic multiclass M/M/1-PS NEVEROFF queue with $K+1$ classes of customers indexed by $k = 0, 1, \ldots, K$. Let λ_k and $E[S_k] = 1/\mu_k$ denote the arrival rate and the mean service time for class k, respectively. In addition, let $\lambda = \lambda_0 + \lambda_1 + \ldots + \lambda_K$ denote the total arrival rate and

$$E[S] = \frac{1}{\lambda}(\lambda_0 E[S_0] + \lambda_1 E[S_1] + \ldots + \lambda_K E[S_K])$$

refer to the mean service time over all the customers. Assume that the system is stable, i.e., $\rho < 1$, where load $\rho = \lambda E[S]$. Finally, let P^{i} and P^{b} denote the power consumption in energy states IDLE and BUSY, respectively.

Proposition 5. *For a stable multiclass M/M/1-PS NEVEROFF queue, the relative value function with respect to performance in state* $\mathbf{n} = (n_0, n_1, \ldots, n_K)$ *is given by*

$$v^{\mathrm{p}}(\mathbf{n}) - v^{\mathrm{p}}(\mathbf{0}) = \sum_{k=0}^{K} a_k(n_k^2 + n_k) + \sum_{k=0}^{K-1}\sum_{\ell=k+1}^{K} 2a_{k,\ell}n_k n_\ell,$$

where the coefficients a_k *and* $a_{k,\ell}$ $(k < \ell)$ *are solved from the following system of linear equations:*

$$\begin{aligned}
&1 + 2\sum_{i=0}^{K}\lambda_i a_{k,i} - 2\mu_k a_k = 0, \quad 0 \le k \le K;\\
&1 + \sum_{i=0}^{K}\lambda_i(a_{k,i} + a_{\ell,i}) - (\mu_k + \mu_\ell)a_{k,\ell} = 0,\\
&\quad 0 \le k < \ell \le K,
\end{aligned}$$

with notations $a_{k,k} = a_k$ and $a_{k,\ell} = a_{\ell,k}$ for any k, ℓ.

Proof. The general result for a multiclass M/M/1-PS NEVEROFF queue was proved in [18]. $\square$

Proposition 6. *For a stable multiclass M/M/1-PS NEVEROFF queue, the relative value function with respect to energy in state $\mathbf{n} = (n_0, n_1, \dots, n_K)$ is given by*

$$v^{\mathrm{e}}(\mathbf{n}) - v^{\mathrm{e}}(\mathbf{0}) = \sum_{k=0}^{K} E[S_k] n_k (P^{\mathrm{b}} - P^{\mathrm{i}}),$$

where $\mathbf{0} = (0, 0, \dots, 0)$ is the null vector.

Proof. The Howard equations for the system read as follows:

$$\begin{aligned}
&P^{\mathrm{i}} - \bar{c}^{\mathrm{e}} + \sum_{k=0}^{K} \lambda_k (v^{\mathrm{e}}(\mathbf{e}_k) - v^{\mathrm{e}}(\mathbf{0})) = 0, \\
&P^{\mathrm{b}} - \bar{c}^{\mathrm{e}} + \sum_{k=0}^{K} \frac{n_k \mu_k}{n_0 + \ldots + n_K} (v^{\mathrm{e}}(\mathbf{n} - \mathbf{e}_k) - v^{\mathrm{e}}(\mathbf{n})) + \\
&\quad \sum_{k=0}^{K} \lambda_k (v^{\mathrm{e}}(\mathbf{n} + \mathbf{e}_k) - v^{\mathrm{e}}(\mathbf{n})) = 0, \quad \mathbf{n} \neq \mathbf{0},
\end{aligned}$$

where $\mathbf{e}_k$ is a unit vector into direction k and $\bar{c}^{\mathrm{e}}$ denotes the mean power consumption given by

$$\bar{c}^{\mathrm{e}} = E[P] = \rho P^{\mathrm{b}} + (1 - \rho) P^{\mathrm{i}}.$$

It is again a straightforward task to verify that these equations are satisfied by the proposed relative value function. $\square$

Corollary 3. *For a stable multiclass M/M/1-PS NEVEROFF queue, the marginal cost with respect to energy in state $\mathbf{n} = (n_0, n_1, \dots, n_K)$ is given by*

$$v^{\mathrm{e}}(\mathbf{n} + \mathbf{e}_k) - v^{\mathrm{e}}(\mathbf{n}) = E[S_k](P^{\mathrm{b}} - P^{\mathrm{i}}),$$

where $\mathbf{e}_k$ is the unit vector into direction k.

The results in Proposition 5 and Corollary 3 are used when evaluating the additional cost of serving the flow in the macrocell in (12). Note that after the initial policy each class $k = 1, \dots, K$ has arrival rate $(1 - p_k^{\mathrm{LB}})\lambda_k$.

As the value function for the performance has a quadratic form, see Proposition 5, it is clear that the marginal performance cost $v^{\mathrm{p}}(\mathbf{n} + \mathbf{e}_k) - v^{\mathrm{p}}(\mathbf{n})$ has in general a linear form. On the other hand, by Corollary 3, the marginal energy cost is constant.