Pollakis *et al. EURASIP Journal on Wireless Communications and Networking* (2016) 2016:61
DOI 10.1186/s13638-016-0539-y

EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Traffic demand-aware topology control for enhanced energy-efficiency of cellular networks

Emmanuel Pollakis[1*], Renato L. G. Cavalcante[1] and Slawomir Stanczak[1,2]

Abstract

The service provided by current mobile networks is not adapted to spatio-temporal fluctuations in traffic demand, but such fluctuations offer opportunities for energy savings. In particular, significant gains in energy efficiency are realizable by disengaging temporarily redundant hardware components of base stations. We therefore propose a novel optimization framework that considers both the load-dependent energy radiated by the antennas and the remaining forms of energy needed for operating the base stations. The objective is to reduce the energy consumption of mobile networks, while ensuring that the data rate requirements of the users are met throughout the coverage area. Building upon sparse optimization techniques, we develop a majorization-minimization algorithm with the ability to identify energy-efficient network configurations. The iterative algorithm is load-aware, has low computational complexity, and can be implemented in an online fashion to exploit load fluctuations on a short time scale. Simulations show that the algorithm can find network configurations with the energy consumption similar to that obtained with global optimization tools, which cannot be applied to real large networks. Although we consider only one currently deployed cellular technology, the optimization framework is general, potentially applicable to a large class of access technologies.

Keywords: Energy efficiency, Optimization, Majorization-minimization algorithm

1 Introduction

The strive for ubiquitous connectivity and high throughput in the development of the fifth generation (5G) of mobile networks is envisioned to lead to highly dense network topologies providing the best possible service to users at all times. Currently, the network topology and also global network parameters are chosen to meet the quality of service (QoS) demand at peak hours and are largely static, but, as pointed out in many studies (see for instance [2–5]), the traffic load fluctuates significantly over time and space. Such spatio-temporal fluctuations create large capacity surpluses at times of low traffic demand, which in turn offers opportunities for energy savings through adaptation of the service supply to the actual demand. However, in order to utilize the capacity surpluses for significant energy savings, it is essential to reduce the energy consumed by hardware and auxiliary equipment (e.g., coolers), which is a dominant form of energy consumption in current mobile networks. In fact, for a typical network with today's technology, base stations consume over 50 % of the total network energy budget [6]. From this, we conclude that significant energy savings can be achieved only by temporarily disengaging redundant hardware components of base stations. Hereafter, we call this energy-saving mechanism network topology control.[1] Indeed, as pointed out by [4], reducing the number of active base stations in periods of low traffic load offers a huge potential for energy savings. This effect will become even more pronounced in 5G networks because of the envisioned densification of the networks [7].

*Correspondence: emmanuel.pollakis@gmail.com
Parts of the material in this paper were presented at the 2012 IEEE Signal Processing Advances Wireless Communications (SPAWC) Workshop, Cesme, Turkey, 17–20 June 2012 [1].
[1] Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, 10587 Berlin, Germany, Einsteinufer 37, 10587 Berlin, Germany
Full list of author information is available at the end of the article

1.1 Related work

Over recent years, some research effort has been devoted to exploiting temporal and spatial redundancies in wireless systems for energy savings. For instance, references [8–10] address the problem of finding an optimal number of base stations and cell site placements so as to minimize the overall energy consumption subject to QoS requirements of users. Assuming a wireless network based on time division multiple access (TDMA), the objective of the study in [8] is to minimize the overall expected energy consumption by optimizing the number of base stations and their locations. The authors formulate the problem as a mixed integer programming problem and suggest using a simplex method together with the branch and bound algorithm. The drawback of this approach is that, due to the TDMA assumption, the analysis does not carry over to systems with inter-cell interference, which is one of the major challenges faced by designers of modern wireless communication systems [11, 12]. Furthermore, branch and bound methods may be slow [13], which excludes an application of these methods to real-time scenarios, even if the underlying problem is of moderate size.

References [9, 14] propose centralized and decentralized algorithms for wireless communication networks to address the problem of base station selection in the presence of traffic load fluctuations. Although the proposed approach seems to provide good solutions in reasonable time, it does not allow incorporation of different sources of energy consumption, which is of utmost importance in modern networks consisting of hierarchical structures. In addition, the authors focus on numerical evaluations to justify the approach. No analytical justification for the performance of the proposed algorithms is given.

The authors of [10] argue in favor of sleep mode techniques coupled with various network planning schemes. A genetic algorithm is used to find energy-efficient network deployments, so the authors have developed a heuristic approach to put selected base stations into a sleep mode for energy-efficient network operation. In addition to the lack of a mathematical justification, the main shortcoming of this work is that the proposed approach cannot incorporate other radio technologies other than universal mobile telecommunications system (UMTS) terrestrial radio access network. In contrast, as mentioned before, our optimization framework is general enough to be applied to multi-radio access technology (RAT) scenarios, including the second, third, and fourth generations of cellular networks [15].

1.2 Our contribution

This paper deals with the problem of minimizing the overall energy consumption in the downlink channel of mobile (cellular) networks. By taking into account the energy consumed by hardware and auxiliary equipment, we address key shortcomings of most existing approaches to the challenge of boosting energy efficiency of cellular networks. The underlying problem is of combinatorial nature because it essentially amounts to selecting a subset of network elements corresponding to the most energy-efficient network configuration, while providing the desired network coverage. More precisely, motivated by [14, 16], we formulate a combinatorial optimization problem to find a network configuration that consumes the least amount of energy, while satisfying traffic demands expressed in terms of minimum data rate requirements. In doing so, we balance different forms of energy consumption in an optimal manner by taking into account both the load-dependent energy used for transmission and the static energy consumed by hardware regardless of the actual load. Similar to [14], the technology-specific constraints are defined to capture the QoS requirements of the users. Although our optimization framework is generic in the sense that it can be applied to multi-RAT systems by incorporating different RAT-specific constraints, owing to the lack of space, our focus is on a single RAT according to the long term evolution (LTE) standard.

In the following, we highlight the main contribution of the paper:

- In contrast to our previous work [1], we use a more detailed and broadly suitable energy consumption model that explicitly considers both the energy consumed by each cell (sector) at a base station and models the basic energy consumed if at least one cell is active at a base station. The model is based on the computation of the cell load, which is also used to account for the load-dependent energy consumption. Based on this energy consumption model, we derive an algorithm that is able to identify not only entire base stations for deactivation but also individual cells for base stations with multiple cells. We present an extensive evaluation showing the effect of the different energy consumption parts (static and dynamic) on the solution of the energy saving network topology.
- Starting from the worst-case interference assumption used in [1], we develop a novel algorithm that uses the framework of interference calculus [17, 18] to arrive at larger energy savings by calculating more accurate values for the spectral efficiency of links.
- We show how our algorithms can be applied to systems where coordinated multi-point (CoMP) strategies are employed. We also elaborate on the fact that the use of CoMP techniques render the application of some involved heuristics unnecessary.
- We complete the analysis of our algorithms by including short discussions about the convergence and the complexity of our proposed algorithms.

1.3 Notation and paper organization

For a vector $\boldsymbol{x} \in \mathbb{R}^N$, its ith component is $x_i \in \mathbb{R}$. Similarly, for a matrix $\mathbf{X} \in \mathbb{R}^{M\times N}$, its (i,j)-th component is $x_{i,j}$. Inequalities involving vectors, such as $\boldsymbol{x}_1 \geq \boldsymbol{x}_2$, are to be understood as component-wise inequalities. The set $\mathbb{R}_+$ denotes the set of non-negative real numbers, while $\mathbb{R}_{++} := \mathbb{R}_+ \backslash \{0\}$ is the set of positive real numbers.

Given a matrix $\mathbf{X} \in \mathbb{R}^{M\times N}$, we use $\tilde{\boldsymbol{x}} := \text{vec}(\mathbf{X}) \in \mathbb{R}^{MN}$ to denote the vector obtained by stacking the columns of $\mathbf{X}$. Note that the entries of $\tilde{\boldsymbol{x}}$ may be confined to take values on $[0,1]$ or $\{0,1\}$ depending on whether $\mathbf{X} \in [0,1]^{M\times N}$ or $\mathbf{X} \in \{0,1\}^{M\times N}$.

Definition 1 (l_0-norm)**.** *For any vector $\boldsymbol{x} \in \mathbb{R}^N$ and matrix $\mathbf{X} \in \mathbb{R}^{M\times N}$, their l_0-norms $|\boldsymbol{x}|_0$ and $|\boldsymbol{X}|_0$ are equal to the number of nonzero elements of $\boldsymbol{x}$ and $\mathbf{X}$, respectively. For a scalar $x \in \mathbb{R}$, $|x|_0 := 1$ if $x \neq 0$ and $|x|_0 := 0$ otherwise.*[2]

The remainder of this paper is organized as follows. Section 2 introduces the underlying system model, and in Section 3, we outline the general problem to solve. In Section 4, the proposed algorithm to find solutions for our optimization problem is derived based on a worst-case inter-cell interference assumption. Section 5 presents how to explicitly take into account a more realistic inter-cell interference model. We present empirical evaluations of the proposed algorithm in Section 6.

2 System model

We consider the downlink channel of a multi-cell LTE network with an established network topology and a central network controller. The central network controller is responsible for collecting measurements, executing the proposed algorithm, and propagating updated network configuration parameters throughout the network. In this work, as in [8–10], we assume that the mechanisms required to collect measurements and to determine when to execute our proposed algorithm are available at a central network controller.

We assume that the network consists of L base stations. Each base station has multiple sectors (called cells in the following), and we denote the set of cells belonging to base station l by $\mathcal{S}_l$. The set of all base stations is denoted by $\mathcal{L}$, and we use $\mathcal{M} := \cup_{l\in\mathcal{L}}\mathcal{S}_l$ to denote the set of all M cells in the network. The cell deployment is assumed to be dense enough so that coverage areas of different cells overlap. This implies that users can be served by different neighboring cells.

2.1 Ensuring coverage via test points

In order to ensure the desired coverage anytime and everywhere in the considered area, we impose coverage constraints by adopting the concept of test points, which is widely used in network planning and optimization [19, 20].

Definition 2 (Test point)**.** *A test point (TP) is a centroid of a pre-defined geographical subarea that represents an aggregated QoS requirement resulting from individual QoS demands of all potential users in this subarea.*[3] *Without loss of generality, we assume N TPs with the set of all TPs denoted by $\mathcal{N} := \{1, 2, \ldots, N\}$.*

An interpretation of this definition is depicted in Fig. 1 for a generic service area. A consequence of Definition 2 is that small-scale fluctuations in QoS demand at the user level are averaged out at the TPs. These small-scale fluctuations must be compensated by the lower layers of the protocol stack (e.g., through adaptive modulation or coding). On a large scale, the traffic demand is assumed to be static for a sufficiently large period of time for which we derive a feasible network configuration that supports this traffic demand. The duration of this period depends on the accuracy of the demand estimates and other factors such as security margins included in the optimization framework.

Assumption 1. *The QoS requirement for a TP corresponds to the aggregated expected traffic over the respective area per unit time. This traffic requirement is expressed in terms of the minimum required data rate per TP.*

Assumption 2. *If the minimum rate requirement of TP j is met, so are the requirements of the users in the associated subarea.*[4]

For services with no explicit data rate requirements (e.g., voice calls), we assume that they can be supported if a minimum data rate per service request is ensured. By Assumption 1, each TP $j \in \mathcal{N}$ is assigned rate requirement r_j, and we collect the rate requirements of all TPs in the vector $\boldsymbol{r} = [r_1, r_2, \ldots, r_N] \in \mathbb{R}^N_{++}$. In general, a TP can be assigned to any cell, and an assignment should be understood as follows. If TP $j \in \mathcal{N}$ is assigned to cell $i \in \mathcal{M}$, then all users in the respective subarea associated with TP j are served by cell j. The assignment of the TPs to the cells is subject to optimization in this paper. We use $\mathbf{X} = [x_{i,j}] \in \{0,1\}^{M\times N}$ to denote the assignment matrix where $x_{i,j} = 1$ if TP j is assigned to cell i and $x_{i,j} = 0$ otherwise.

Assumption 3. *While each TP is assigned to exactly one cell, each cell can serve multiple TPs, and the set of TPs served by cell i under assignment $\mathbf{X}$ is denoted by $\mathcal{N}_i(\mathbf{X}) \subset \mathcal{N}$.*

We point out that this assumption has been widely used in previous studies [9, 14, 20], and it is valid throughout

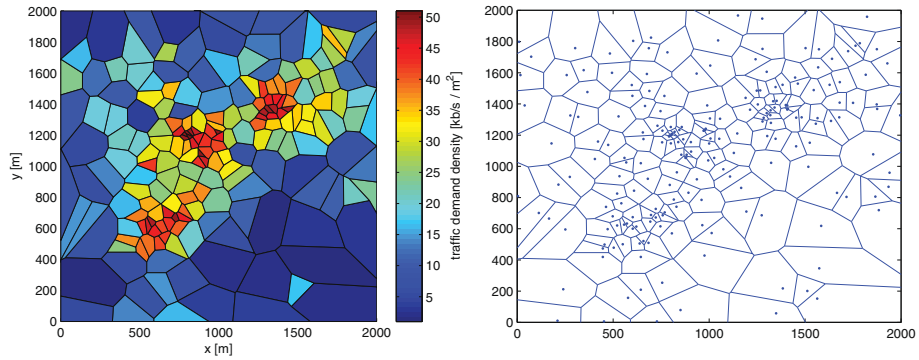

Fig. 1 Test point concept. Illustration of the test point concept. Generic traffic demand map (*left*) and exemplary test point layout (*right*)

the paper except for Section 4.3, where it is shown how to include scenarios in which each TP can be served by multiple cells. Note that if $\mathcal{N}_i(\mathbf{X}) = \emptyset$ for some $i \in \mathcal{M}$, then cell i can be deactivated for energy savings because no TP is assigned to cell i. In contrast, if $\mathcal{N}_i(\mathbf{X}) \neq \emptyset$, then cell i is active, and each TP connected to it induces some amount of *cell load*.

Definition 3 (Cell load)**.** *Given the assignment* $\tilde{\boldsymbol{x}} := vec(\mathbf{X})$*, the load of cell i, denoted by* $\rho_i(\tilde{\boldsymbol{x}}) \in [0, 1]$ *or simply* ρ_i *for notational simplicity, is defined to be the ratio of the number of resource blocks requested by TPs served by cell* $i \in \mathcal{M}$ *to the total number of resource blocks* B_i *available at this cell.*[5]

We use $\boldsymbol{\rho} := [\rho_1, \ldots, \rho_M]^T \in [0, 1]^M$ to denote the vector of all cell loads. From the definition of cell load, we have the following:

Fact 1. *The load at cell i satisfies* $\rho_i > 0$ *if and only if (iff) cell i serves at least one TP.*

2.2 Spectral efficiency and resource usage

The optimal assignment of TPs to cells is strongly influenced by the spectral efficiency of the corresponding links. For the analysis in this paper, we adopt an OFDMA-based (Orthogonal Frequency-Division Multiple Access (OFDMA)) model for the spectral efficiency that is widely used in the literature [11, 21, 22]. The spectral efficiency also depends on radio propagation properties. Therefore, we associate to each TP a path-loss vector and write the path-loss vectors of all TPs as columns of the path-loss matrix $\mathbf{G} = [g_{i,j}] \in \mathbb{R}_{++}^{M \times N}$, where $g_{i,j}$ captures the long-term path loss and shadowing effects for a radio link from cell i to TP j.

Assumption 4 (Reliable path-loss estimates)**.** *A reliable estimate of* $\mathbf{G}$ *is available at the central network controller.*

Remark 1. *The problem of reliable estimation and tracking of the path-loss matrix is out of the scope of the paper. However, the matrix captures only long-term fading effects, so reliable estimates of* $\mathbf{G}$ *can be obtained and tracked in practice. Promising algorithmic solutions to this estimation problem are for instance presented in [23]. Moreover, in network planning problems, knowledge of* $\boldsymbol{G}$ *is a very common assumption in the literature [11, 20, 24].*

Now, we are in a position to define the signal-to-interference-noise-ratio (SINR) $\gamma_{i,j} : \mathbb{R}_+^M \to \mathbb{R}_+$ between cell $i \in \mathcal{M}$ and TP $j \in \mathcal{N}$ by [11, 24]:

$$\gamma_{i,j}(\boldsymbol{\rho}) = \frac{P_i \, g_{i,j}}{\sum_{k \in \mathcal{M} \setminus \{i\}} P_k \, g_{k,j} \, \rho_k + \sigma^2}, \tag{1}$$

where $P_i > 0$ is the transmit power per resource block of cell i and $\sigma^2 > 0$ is the noise power per resource block. Accordingly, the link spectral efficiency $\omega_{i,j} : \mathbb{R}_+^M \to \mathbb{R}_+$ (in bits per resource block[6]) for the link from cell i to TP j is given by [21]

$$\omega_{i,j}(\boldsymbol{\rho}) = \eta_{i,j}^{\mathrm{BW}} \log_2\Big(1 + \frac{\gamma_{i,j}(\boldsymbol{\rho})}{\eta_{i,j}^{\mathrm{SINR}}}\Big) \tag{2}$$

where $\eta_{i,j}^{\mathrm{BW}} \in \mathbb{R}_{++}$ and $\eta_{i,j}^{\mathrm{SINR}} \in \mathbb{R}_{++}$ are suitably chosen constants, referred to as bandwidth and SINR efficiency, respectively. These constants depend on the overall system design, which includes the choice of scheduling protocols and multi-antenna techniques. The choice of these constants has no impact on our results, so they are assumed to be arbitrary and fixed throughout the paper. For realistic values of these constants, we refer the interested reader to [11, 21].

From (2), we can easily see that the necessary number of resource blocks $b_{i,j}$ at cell i to serve TP j with data rate r_j is equal to $b_{i,j} = \frac{r_j}{\omega_{i,j}(\boldsymbol{\rho})} > 0$. In addition, following

Definition 3, the load at cells can be computed by the following system of non-linear equations

$$\rho_i = \sum_{j \in \mathcal{N}_i(\mathbf{X})} \frac{b_{i,j}}{B_i} = \sum_{j \in \mathcal{N}_i(\mathbf{X})} \frac{r_j}{B_i \omega_{i,j}(\boldsymbol{\rho})},\ i \in \mathcal{M}. \tag{3}$$

Remark 2. *In practice, cells need to reserve some fraction of their resource blocks for signaling. If cell i has B_i^* resource blocks in total, and it needs to reserve $a_i > 0$ of its resource blocks for signaling, then the resource blocks at cell i available for allocation to TPs are $B_i = B_i^* - a_i$.*

For a fixed assignment, $\mathbf{X}$ cell load $\boldsymbol{\rho}$ in 3 can be efficiently computed by means of fixed-point algorithms (c.f. Section 5). However, the assignment of TPs to cells is the main subject of our optimization problem and thus we cannot evaluate (3) easily. In order to keep the complexity of the optimization problem tractable, we lower bound the spectral efficiency.

Assumption 5 (Worst-case interference)**.** *We have the worst-case interference scenario if all cells are fully loaded, i.e. $\boldsymbol{\rho} = \mathbf{1}$.*

Unless otherwise stated, we use the worst-case interference assumption, which results in a lower bound on the true link spectral efficiency $\omega_{i,j}(\boldsymbol{\rho}) \geq \tilde{\omega}_{i,j} := \omega_{i,j}(\mathbf{1})$ for every $\boldsymbol{\rho} \in [0,1]^M$. In general, this bound diminishes gains in energy savings when taking into account the energy consumption of hardware, and we show in Section 5 how to incorporate the actual link spectral efficiency to improve the energy savings. Nevertheless, having fully loaded cells as in Assumption 5 is desirable because it is an effective way to avoid coverage holes as a result of deactivating cells based, for instance, on imperfect information.

Remark 3. *The worst-case interference assumption cannot exploit the full potential for energy savings, but the assumption is of high practical relevance because it is an effective way to avoid coverage holes as a result of deactivating cells based, for instance, on imperfect information.*

2.3 Energy consumption model

In contrast to most works in literature, we consider a model for the energy consumption of a base station and its cells that takes into account not only the cell load-dependent transmit energy radiated by antennas but also the remaining sources of energy consumption that are independent of the cell load as long as the cell/base station is *active*.

Definition 4 (Active base station/cell)**.** *Consider a particular base station $l \in \mathcal{L}$ and its cells $i \in \mathcal{S}_l$. Let $\rho_i \in [0,1]$ be the load of cell i. We say that a cell i is active iff $\rho_i > 0$ and that base station l is active iff one of its cells is active, i.e., $\sum_{i \in \mathcal{S}_l} \rho_i > 0$. If a cell or base station is not active it is said to be inactive.*

With Definition 4, we are in the position to define the energy consumption of a base station.

Definition 5 (Energy consumption)**.** *Given a TP assignment $\mathbf{X}$ inducing a cell load $\boldsymbol{\rho}$, the energy consumption $E_l(\boldsymbol{\rho}) \geq 0$ of base station l is defined to be the power that the respective base station consumes per unit of time, where $E_l(\boldsymbol{\rho}) = 0$ iff base station l is inactive.*

The function $E_l(\boldsymbol{\rho})$ depends on the hardware setup of the base station, but it can be split into three parts in general:

(i) The static energy consumption of the base station $c_l > 0$ (due to shared hardware between sectors, e.g., cooling and power supply),
(ii) The static energy consumption $e_i > 0$ ($i \in \mathcal{S}_l$) of its *active* cells (e.g., due to power amplifiers and signal processing units), and
(iii) the load-dependent dynamic energy consumption of its *active* cells $f_i(\rho_i)$ ($i \in \mathcal{S}_l$), where $f_i : [0,1] \to \mathbb{R}_+$ is a given continuous function relating the energy consumption to the corresponding cell load.

By these definitions and Fact 1, $E_l(\boldsymbol{\rho})$ is a discontinuous function of the cell load, and we have

$$E_l(\boldsymbol{\rho}) = \begin{cases} 0 & \text{cells } i \in \mathcal{S}_l \text{ serve no TP} \\ c_l + \sum_{i \in \mathcal{S}_{l,\text{active}}} e_i + f_i(\rho_i) & \text{otherwise}, \end{cases}$$

where $\mathcal{S}_{l,\text{active}} \subset \mathcal{S}_l$ is the set of active cells of base station l. Therefore, the total energy consumption in a network, which is the accumulated energy consumption of all active base stations, yields

$$E(\boldsymbol{\rho}) = \sum_{l \in \mathcal{L}} E_l(\boldsymbol{\rho}) = \sum_{l \in \mathcal{L}} \left(c_l \left| \sum\nolimits_{i \in \mathcal{S}_l} \rho_i \right|_0 + \sum_{i \in \mathcal{S}_l} \left(e_i |\rho_i|_0 + f_i(\rho_i) \right) \right). \tag{4}$$

For concreteness, we make the following assumption throughout the paper (see also Remark 4)

Assumption 6 (Concave dynamic energy consumption)**.** *$f_i : [0,1] \to \mathbb{R}_+$ ($i \in \mathcal{M}$), is concave and continuously differentiable.*

In particular, this assumption is satisfied by a linear dependency of the base station energy consumption and the cell load reported in current studies such as [27].

Remark 4. *In fact, the load-dependent dynamic energy consumption can also be assumed to be a convex function of the load. Moreover, we could even assume that it is a sum*

of convex and concave functions. The optimization framework presented in this paper can be straightforwardly extended to cover these cases.

For convenience, we have summarized the main system variables in Table 1.

3 Problem statement

Spatio-temporal redundancies in coverage and capacity resulting from day-time fluctuations in traffic demand present great opportunities for energy savings by deactivating redundant cells at times of relatively low traffic demand. Indeed, if the traffic demand decreases, some or all entries of the rate requirement vector $\boldsymbol{r} \in \mathbb{R}^{N}_{++}$ become relatively small, which can be utilized to reduce the total energy consumption by minimizing the cost function in (4) subject to different constraints that follow from the system model and (3). Formally, the problem under consideration can be stated as follows (note that the complete set of equations is referred to as (5)):

$$\min . \sum_{l\in\mathcal{L}} \left(c_l \left| \sum_{i\in\mathcal{S}_l} \rho_i \right|_0 + \sum_{i\in\mathcal{S}_l} \left(e_i |\rho_i|_0 + f_i(\rho_i) \right) \right) \tag{5a}$$

$$\text{s. t.:} \sum_{j\in\mathcal{N}} \frac{r_j}{B_i\, \tilde{\omega}_{i,j}}\, x_{i,j} = \rho_i \qquad i \in \mathcal{M} \tag{5b}$$

$$\sum_{i\in\mathcal{M}} x_{i,j} = 1 \qquad j \in \mathcal{N} \tag{5c}$$

$$\rho_i \in [0,1] \qquad i \in \mathcal{M} \tag{5d}$$

$$x_{i,j} \in \{0,1\} \qquad i \in \mathcal{M}, j \in \mathcal{N}, \tag{5e}$$

Table 1 List of variables

Variable	Symbol
Set of all base stations	$\mathcal{L} = \{1, \ldots, L\}$
Set of all cells	$\mathcal{M} = \{1, \ldots, M\}$
Set of cells associated with base station l	$\mathcal{S}_l \subset \mathcal{M}$
Set of all test points	$\mathcal{N} = \{1, \ldots, N\}$
Set of test points served by cell i	$\mathcal{N}_i \subset \mathcal{N}$
Rate requirements of all test points	$\boldsymbol{r} = [r_1, r_2, \ldots, r_N] \in \mathbb{R}^{N}_{++}$
Assignment matrix	$\mathbf{X} = [x_{i,j}] \in \{0,1\}^{M\times N}$
Cell load vector	$\boldsymbol{\rho} := [\rho_1, \ldots, \rho_M]^T \in [0,1]^M$
Path-loss matrix	$\mathbf{G} = [g_{i,j}] \in \mathbb{R}^{M\times N}_{++}$
Link spectral efficiency in bits per resource block	$\omega_{i,j} : \mathbb{R}^{M}_{+} \to \mathbb{R}_{+}$
Worst-case link spectral efficiency in bits per resource block	$\tilde{\omega}_{i,j} \in \mathbb{R}_{+}$
Number of resource blocks available at cell i	$B_i \in \mathbb{R}_{++}$
Static energy consumption of base station l	$c_l \in \mathbb{R}_{+}$
Static energy consumption of cell i	$e_i \in \mathbb{R}_{+}$
Dynamic energy consumption of cell i	$f_i : [0,1] \to \mathbb{R}_{+}$

where the optimization variables are $x_{i,j}$ and ρ_i ($i \in \mathcal{M}, j \in \mathcal{N}$). In particular, Assumption 3 is captured by (5c) together with (5e). Constraints 5b and (5d), in contrast, ensure that the cell load is in accordance with Definition 3.

To ensure feasibility of the above problem and to show the effectiveness of our approach, we consider scenarios where the rate requirements of TPs are sufficiently low for a reasonable amount of redundancies that allow for deactivation of cells. Moreover, if the traffic requirements in the system are sufficiently low or the number of cells is sufficiently large, $\boldsymbol{\rho}^{\star}$ is expected to be *sparse* with zero entries specifying cells that can be deactivated.

4 Energy-efficiency optimization

The difficulty of problem (5) lies in its combinatorial nature. In fact, it can be shown that the problem is related to the classical bin-packing problem, which is known to be NP-hard (non-deterministic polynomial-time hard) [28]. Consequently, the complexity is expected to grow exponentially with the number of cells. On the positive side, problem (5) has a special structure that can be exploited by majorization-minimization techniques [29], which have been widely used in recent years to tackle various problems in compressed sensing [30] and machine learning [31].

Instead of finding a global solution to (5), we will pursue a less ambitious goal. We apply the majorization-minimization techniques mentioned above to develop a low-complexity anytime algorithm that has a strong analytical justification. This algorithm is expected to provide good results (in terms of low energy consumption) with low complexity. To this end, we reformulate problem (5) to pose it in a more tractable form. First, we observe that each load ρ_i is, in fact, a function of $\mathbf{X}$ (c.f. Definition 3 and 5b). We can therefore modify the problem to have only $\mathbf{X}$ as an optimization variable. The objective function in (5a) can be equivalently written as

$$\begin{aligned}
&\sum_{l\in\mathcal{L}} \left(c_l \left| \sum_{i\in\mathcal{S}_l} \rho_i \right|_0 + \sum_{i\in\mathcal{S}_l} \left(e_i\, |\rho_i|_0 + f_i(\rho_i) \right) \right) \\
&= \sum_{l\in\mathcal{L}} \left(c_l \left| \sum_{i\in\mathcal{S}_l} \sum_{j\in\mathcal{N}} x_{i,j} \right|_0 + \sum_{i\in\mathcal{S}_l} \left(e_i \left| \sum_{j\in\mathcal{N}} x_{i,j} \right|_0 + f_i(\rho_i) \right) \right) \\
&= \sum_{l\in\mathcal{L}} \left(c_l \left| \boldsymbol{t}_l^T \tilde{\boldsymbol{x}} \right|_0 + \sum_{i\in\mathcal{S}_l} \left(e_i \left| \boldsymbol{s}_i^T \tilde{\boldsymbol{x}} \right|_0 + f_i(\rho_i) \right) \right)
\end{aligned} \tag{6}$$

where $\boldsymbol{s}_i := \mathrm{vec}(\mathbf{S}_i)$ with $\mathbf{S}_i \in \{0,1\}^{M\times N}$ being a matrix of zeros, except for its ith row, which is a row of ones. Similarly, $\boldsymbol{t}_l := \mathrm{vec}(\mathbf{T}_l)$ with $\mathbf{T}_l \in \{0,1\}^{M\times N}$ is a matrix of zeros, except for its rows $i \in \mathcal{S}_l$, which are rows of

ones. The first equality in (6) follows from Fact 1 and the definition of the l_0-norm, which does not account for magnitudes. More precisely, if at least one TP is served by cell i (i.e., $\sum_{j\in\mathcal{N}} x_{i,j} \geq 1$), then the cell load at cell i is non-zero $\rho_i > 0$ and we have $|\rho_i|_0 = \left|\sum_{j\in\mathcal{N}} x_{i,j}\right|_0 = 1$. The second equality in (6) uses vector multiplication to represent the sums in a more compact way.

Definition 6. *Given the assignment $\tilde{\boldsymbol{x}}$ and the load dependent energy consumption $f_i(\rho_i(\tilde{\boldsymbol{x}}))$ of cell i with $\rho_i(\tilde{\boldsymbol{x}}) = \sum_{j\in\mathcal{N}} \frac{r_j}{B_i\tilde{\omega}_{i,j}} x_{i,j}$ (c.f. 5b), we define the function $\tilde{f}_i : [0,1]^{NM} \to \mathbb{R}_+ : \tilde{\boldsymbol{x}} \mapsto f_i(\sum_{j\in\mathcal{N}} \frac{r_j}{B_i\tilde{\omega}_{i,j}} x_{i,j})$.*

Considering Definition 6 and using $\rho_i \leq 1$ (see Definition 3) in (5b), we arrive at an equivalent problem given by

$$\min. \sum_{l\in\mathcal{L}} \left(c_l \left| \boldsymbol{t}_l^T \tilde{\boldsymbol{x}} \right|_0 + \sum_{i\in\mathcal{S}_l} \left(e_i \left| \boldsymbol{s}_i^T \tilde{\boldsymbol{x}} \right|_0 + \tilde{f}_i(\tilde{\boldsymbol{x}}) \right) \right) \tag{7a}$$

$$\text{s. t.:} \sum_{j\in\mathcal{N}} \frac{r_j}{B_i\,\tilde{\omega}_{i,j}} x_{i,j} \leq 1 \qquad i \in \mathcal{M} \tag{7b}$$

$$\sum_{i\in\mathcal{M}} x_{i,j} = 1 \qquad j \in \mathcal{N} \tag{7c}$$

$$x_{i,j} \in \{0,1\} \qquad i \in \mathcal{M}, j \in \mathcal{N}, \tag{7d}$$

where the assignment variables $x_{i,j}$ ($i \in \mathcal{M}, j \in \mathcal{N}$) are the only optimization variables.

4.1 Problem relaxation

To obtain an optimization problem that is computationally tractable, we first relax the binary constraint (7d) to[7]

$$x_{i,j} \in [0,1], \forall i \in \mathcal{M}, \forall j \in \mathcal{N}. \tag{8}$$

The above makes all constraints convex, so now the only problem is the objective function, which is not continuous due to the l_0-norm. We also note that by Assumption 6 and Definition 6, the load-dependent term $\tilde{f}_i(\tilde{\boldsymbol{x}})$ in the objective function (7a) is concave and continuously differentiable for $\tilde{\boldsymbol{x}} \in [0,1]^{NM}$ since these properties are preserved under a composition with a linear function [32, 33]. To address the non-continuity of the l_0-norm, we consider the following relation [30]:

$$\forall_{\boldsymbol{z}\in\mathbb{R}^K} \; |\boldsymbol{z}|_0 = \lim_{\epsilon\to 0} \sum_{k=1}^{K} \frac{\log\left(1 + |z_k|\,\epsilon^{-1}\right)}{\log\left(1+\epsilon^{-1}\right)}. \tag{9}$$

By using 9 and the non-negativity of $\boldsymbol{s}_i$, $\boldsymbol{t}_i$, $\tilde{\boldsymbol{x}}$, the cost function in 7a can be equivalently written as

$$\begin{aligned} &\sum_{l\in\mathcal{L}} \left(c_l \left| \boldsymbol{t}_l^T \tilde{\boldsymbol{x}} \right|_0 + \sum_{i\in\mathcal{S}_l} e_i \left| \boldsymbol{s}_i^T \tilde{\boldsymbol{x}} \right|_0 + \tilde{f}_i(\tilde{\boldsymbol{x}}) \right) \\ &= \lim_{\epsilon\to 0} \sum_{l\in\mathcal{L}} \left(c_l \frac{\log\left(1+\epsilon^{-1}\,\boldsymbol{t}_l^T\tilde{\boldsymbol{x}}\right)}{\log\left(1+\epsilon^{-1}\right)} \right. \\ &\qquad \left. + \sum_{i\in\mathcal{S}_l} \left(e_i \frac{\log\left(1+\epsilon^{-1}\,\boldsymbol{s}_i^T\tilde{\boldsymbol{x}}\right)}{\log\left(1+\epsilon^{-1}\right)} + \tilde{f}_i(\tilde{\boldsymbol{x}}) \right) \right). \end{aligned} \tag{10}$$

We can therefore obtain an approximation to problem (5) by replacing the objective function by the right-hand side of 10 for a sufficiently small but fixed $\epsilon > 0$. More precisely, for some $\epsilon > 0$, the objective is to find a matrix $\mathbf{X} \in [0,1]^{M\times N}$ or, equivalently, a vector $\tilde{\boldsymbol{x}} = \text{vec}(\mathbf{X}) \in [0,1]^{NM}$ that solves the following problem

$$\begin{aligned} \min. \sum_{l\in\mathcal{L}} &\left(c_l \frac{\log\left(1+\epsilon^{-1}\,\boldsymbol{t}_l^T\tilde{\boldsymbol{x}}\right)}{\log\left(1+\epsilon^{-1}\right)} \right. \\ &\left. + \sum_{i\in\mathcal{S}_l} \left(e_i \frac{\log\left(1+\epsilon^{-1}\,\boldsymbol{s}_i^T\tilde{\boldsymbol{x}}\right)}{\log\left(1+\epsilon^{-1}\right)} + \tilde{f}_i(\tilde{\boldsymbol{x}}) \right) \right) \end{aligned} \tag{11a}$$

$$\text{s. t.:} \sum_{j\in\mathcal{N}} \frac{r_j}{B_i\,\tilde{\omega}_{i,j}} x_{i,j} \leq 1 \qquad i \in \mathcal{M} \tag{11b}$$

$$\sum_{i\in\mathcal{M}} x_{i,j} = 1, \qquad j \in \mathcal{N} \tag{11c}$$

$$x_{i,j} \in [0,1] \qquad i \in \mathcal{M}, j \in \mathcal{N}. \tag{11d}$$

Solving problem (11) is not straightforward because we need to *minimize* a non-convex function over a convex set. Fortunately, reference [30] presents an optimization framework based on the majorization-minimization (MM) algorithm [29] to handle problems of this type. The framework can be used to decrease the value of the objective function in a computationally efficient way. For completeness, the reader can find some details of the MM algorithms in the Appendix.

4.2 Majorization-minimization (MM) algorithm

For notational convenience, we define $\hat{c}_l := \frac{c_l}{\log(1+\epsilon^{-1})}$ and $\hat{e}_i := \frac{e_i}{\log(1+\epsilon^{-1})}$, and we use these definitions in (11a) to simplify the objective function (ignoring unnecessary constants):

$$\begin{aligned} &h : \mathcal{X} \to \mathbb{R}, \\ &\qquad h(\tilde{\boldsymbol{x}}) := \sum_{l\in\mathcal{L}} \left(\hat{c}_l \log\left(\epsilon + \boldsymbol{t}_l^T\tilde{\boldsymbol{x}}\right) \right) \\ &\qquad\qquad + \sum_{i\in\mathcal{M}} \left(\hat{e}_i \log\left(\epsilon + \boldsymbol{s}_i^T\tilde{\boldsymbol{x}}\right) + \tilde{f}_i(\tilde{\boldsymbol{x}}) \right), \end{aligned} \tag{12}$$

where $\mathcal{X} \subset \mathbb{R}^{MN}$ is the closed convex set of points satisfying the constraints (11b)–(11d) and we have used the fact that $\mathcal{M} = \cup_{l\in\mathcal{L}}\mathcal{S}_l$. Since $\tilde{f}_i$ is concave and continuously differentiable by Assumption 6, so is the function in (12) for any $\epsilon > 0$. Therefore, according to the explanations in the Appendix, we can use the following function

$$g : \mathcal{X} \times \mathcal{X} \to \mathbb{R} : \ (\boldsymbol{x}, \boldsymbol{y}) \mapsto h(\boldsymbol{y}) + \nabla h(\boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y})$$

as a majorizing function of 12, where the gradient can be easily calculated:

$$\nabla h(\tilde{\boldsymbol{x}}) = \sum_{l\in\mathcal{L}} \hat{c}_l \frac{1}{\epsilon + \boldsymbol{t}_l^T\tilde{\boldsymbol{x}}} + \sum_{i\in\mathcal{M}} \left(\hat{e}_i \frac{1}{\epsilon + \boldsymbol{s}_i^T\tilde{\boldsymbol{x}}} + \nabla\tilde{f}_i(\tilde{\boldsymbol{x}}) \right). \tag{13}$$

Thus, updates of the MM algorithm take the form (see the Appendix)

$$\begin{aligned} \tilde{\boldsymbol{x}}^{(n+1)} &\in \arg\min_{\tilde{\boldsymbol{x}}\in\mathcal{X}} g\left(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{x}}^{(n)}\right) \\ &= \arg\min_{\tilde{\boldsymbol{x}}\in\mathcal{X}} \sum_{l\in\mathcal{L}} \hat{c}_l \frac{\boldsymbol{t}_l^T\tilde{\boldsymbol{x}}}{\epsilon + \boldsymbol{t}_l^T\tilde{\boldsymbol{x}}^{(n)}} \\ &\quad + \sum_{i\in\mathcal{M}} \left(\hat{e}_i \frac{\boldsymbol{s}_i^T\tilde{\boldsymbol{x}}}{\epsilon + \boldsymbol{s}_i^T\tilde{\boldsymbol{x}}^{(n)}} + \nabla\tilde{f}_i\left(\tilde{\boldsymbol{x}}^{(n)}\right)^T\tilde{\boldsymbol{x}} \right) \end{aligned} \tag{14}$$

for some feasible starting point[8] $\tilde{\boldsymbol{x}}^{(0)} \in \mathcal{X}$. In words, the MM algorithm solves iteratively a sequence of convex optimization problems. For the chosen majorizing function, the problem to be solved in every iteration is a linear programming problem (LP), which can be typically solved efficiently with standard optimization tools.

As discussed in the Appendix, the sequence $\{\tilde{\boldsymbol{x}}^{(n)}\}_{n\in\mathbb{N}} \subset \mathcal{X}$ for some $\tilde{\boldsymbol{x}}^{(0)} \in \mathcal{X}$ generated by (14) produces a non-increasing sequence $\{h(\tilde{\boldsymbol{x}}^{(n)})\}_{n\in\mathbb{N}}$ of objective values. Therefore, as $n \to \infty$, we expect the corresponding sequence of assignment matrices $\{\boldsymbol{X}^{(n)}\}_{n\in\mathbb{N}}$ (note that $\tilde{\boldsymbol{x}}^{(n)} =: \mathrm{vec}(\mathbf{X}^{(n)})$) to evolve towards network configurations with low energy consumption.

We stop the algorithm if the improvements in the objective value are small enough in the sense that for some sufficiently small $\epsilon^\star > 0$, the following condition is met

$$h\big(\tilde{\boldsymbol{x}}^{(n)}\big) - h\big(\tilde{\boldsymbol{x}}^{(n+1)}\big) \le \epsilon^\star. \tag{15}$$

Note that (12) is monotonically decreasing (c.f. Appendix) and bounded from below ($h(\tilde{\boldsymbol{x}}) \ge \sum_{l\in\mathcal{L}} \hat{c}_l \log\epsilon + \sum_{i\in\mathcal{M}} \hat{e}_i \log\epsilon, \forall\tilde{\boldsymbol{x}} \in \mathcal{X}$), so the sequence $\{h(\tilde{\boldsymbol{x}}^{(n)})\}_{n\in\mathbb{N}}$ converges by the monotone convergence theorem. We emphasize that this does not imply a convergence of $\{\tilde{\boldsymbol{x}}^{(n)}\}_{n\in\mathbb{N}}$. For properties of the sequence $\{\tilde{\boldsymbol{x}}^{(n)}\}_{n\in\mathbb{N}}$, we refer the reader to [34].

Upon termination, the resulting assignment matrix $\mathbf{X}^{(n)} \in [0,1]^{M\times N}$ needs to be mapped to a matrix $\mathbf{X}^\star \in \{0,1\}^{M\times N}$ in order to obtain a feasible point to the problem in (5). For this purpose, we use the heuristic described in Algorithm 1 (Fig. 2). The main idea is as follows. We start by rounding the entries $x_{i,j}^{(n)}$ to the closest integer, and then we check if the obtained assignment matrix is part of the set $\mathcal{X}$. Otherwise, we activate additional cells and connect TPs to them. By using the standard LP solver of CPLEX ("IBM ILOG CPLEX Optimization Studio" [35]) in our simulations, most entries of the matrix $\mathbf{X}^{(n)} \in [0,1]^{M\times N}$ are typically either zero or one, so the rounding operation rarely results in a violation of a constraint (but we emphasize that this is not guaranteed to be true in general).

For convenience, we summarize the complete approach in Algorithm 2 (Fig. 3).

4.3 Serving a test point with multiple cells

By Assumption 3, each TP is restricted to be served by exactly one cell. This strict limitation introduces the non-convex constraint (5e) to the optimization problem in (5), which motivates the relaxation (8) and the heuristic mapping introduced in Algorithm 1 (Fig. 2). To avoid these heuristic approaches for which we are not guaranteed to find solutions, we assume in this section that each TP can be served by multiple cells. This assumption is implemented by using 8 directly instead of 5e. As a result, there is no need for any relaxations of the constraints or the use of heuristic mappings such as that in Algorithm 1 (Fig. 2). We only need to approximate the cost function as done in (11a) and apply the MM algorithm to the resulting optimization problem, and we note that these operations have a strong analytical justification.

The assumption of multiple cells serving one TP has a practical interpretation when considering Definition 2. It means that cells can serve only a fraction of the traffic generated in the area corresponding to some TP. In other words, we do not use a all-or-nothing approach, where cells should serve either all users or no users in the area corresponding to a TP.

5 Load-aware energy-efficiency optimization

The model presented in Section 2 assumes the worst-case interference in a fully loaded system, which leads to a lower bound on the link spectral efficiency (c.f. Assumption 5). As pointed out in Remark 3, the main rationale behind this approach is the need for avoiding coverage holes when network elements are deactivated. The price is a sub-optimal performance in terms of energy efficiency because the interference is overestimated, and therefore, users may use more resource blocks than required to keep their minimum data rate requirements. An immediate consequence of this is that

Algorithm 1 Heuristic to map $[0,1]^{M\times N} \rightarrow \{0,1\}^{M\times N}$

Input: $\mathbf{X}^{(n)}$, $\mathcal{N}$, $\mathcal{M}$, set of constraints $\mathcal{X}_2$ representing (11b) and (5e)

Output: final assignment matrix $\mathbf{X}^{\star}$

initialize: set of assigned TPs $\mathcal{A}=\emptyset$ and final assignment matrix $\mathbf{X}^{\star}=\mathbf{0}$.
for all $i\in\mathcal{M}, j\in\mathcal{N}$ **do**
  if $x_{i,j}^{(n)}\in\{1\}$ **then**
    $x_{i,j}^{\star}=x_{i,j}^{(n)}$ and $\mathcal{A}=\mathcal{A}\cup\{j\}$.
  end if
end for
Define set $\mathcal{B}=\left\{x_{i,j}^{(n)}\in(0,1)\,|\forall i\in\mathcal{M},\forall j\in\mathcal{N}\backslash\mathcal{A}\right\}$.
while $\mathcal{B}\neq\emptyset$ **do**
  $(i,j)=\text{argmax}_{i,j}\{\mathcal{B}\}$
  if $x_{i,j}^{\star}:=1\rightarrow\mathbf{X}^{\star}\in\mathcal{X}_2$ **then**
    $x_{i,j}^{\star}=1$ and $\mathcal{A}=\mathcal{A}\cup\{j\}$.
    $\mathcal{B}=\mathcal{B}\backslash\{x_{i,j}^{(n)}\,|\forall i\in\mathcal{M}\}$
  else
    $\mathcal{B}=\mathcal{B}\backslash\{x_{i,j}^{(n)}\}$
  end if
end while
for all $j\notin\mathcal{A}$ **do**
  activate closest non-active cell i which yields $x_{i,j}^{\star}:=1\rightarrow\mathbf{X}^{\star}\in\mathcal{X}_2$ and assign $x_{i,j}^{\star}=1$.
  $\mathcal{A}=\mathcal{A}\cup\{j\}$.
end for

Fig. 2 Heuristic to map $[0,1]^{M\times N}\rightarrow\{0,1\}^{M\times N}$. Heuristic to map the assignment matrix $\mathbf{X}^{(n)}\in[0,1]^{M\times N}$ obtained by the proposed algorithm to a feasible assignment $\mathbf{X}^{\star}\in\{0,1\}^{M\times N}$ for the problem in (5)

more cells are activated than are necessary for meeting the minimum rate requirements at the TPs. In this section, we extend the optimization problem in (11) to incorporate more precise estimates of the load induced by a given user-cell assignment, which is not a trivial task because it involves load computation (with fixed assignments) that requires the solution of a system of nonlinear equations [11, 24, 36] (note that we can easily estimate the link spectral efficiency from the load by using 2).

Algorithm 2 Network reconfiguration for improved energy efficient operation

Input: set of TPs, set of cells, constraints

Output: optimized network configuration according to $\mathbf{X}^{\star}$.

initialize $\mathbf{X}^{(0)}$ with a feasible point.
repeat
  compute $\tilde{\boldsymbol{x}}^{(n)}$ by solving (14)
  increment n
until (15) is valid
use Alg. 1 to map $\mathbf{X}^{(n)}$ to $\mathbf{X}^{\star}\in\{0,1\}^{M\times N}$
connect the TPs to cells according to $\mathbf{X}^{\star}$.
deactivate all cells no TP is connected to.

Fig. 3 Network reconfiguration algorithm for improved energy-efficient operation. Algorithmic description of the main approach proposed in this paper. The algorithm iteratively solves and updates a problem closely related to the original assignment problem and maps its solution into the solution space of the original problem

In what follows, we propose an approach that typically yields good approximations of the true link spectral efficiencies. The idea is to use a two-step alternating iterative scheme:

Step 1 Compute the link spectral efficiency $\forall_{i\in\mathcal{M},j\in\mathcal{N}}\ \omega_{i,j}(\boldsymbol{\rho})$ defined in (2) for the load value obtained in the previous iteration of Step 2 of the algorithm (in the first iteration of the algorithm, we can use the worst-case spectral efficiency) and solve Problem (11) with these (fixed) link spectral efficiencies to obtain an TP-cell assignment $\mathbf{X}$.

Step 2 For the TP-cell assignment obtained in Step 1, compute the load induced by this assignment.

Regarding the load computation in Step 2, we use the fact that the load $\boldsymbol{\rho}$ induced by a given assignment $\mathbf{X}$ is a fixed point of the following standard interference mapping (see [36, 37] and the references therein for further details):

$$\mathcal{J}:\mathbb{R}_{+}^{M}\rightarrow\mathbb{R}_{++}^{M}:\ \boldsymbol{\rho}\mapsto[\,I_1(\boldsymbol{\rho})\ \ldots\ I_M(\boldsymbol{\rho})\,]^T\ ,$$

where

$$I_i(\boldsymbol{\rho}):=\min\left\{\sum_{j\in\mathcal{N}}\frac{\lambda_{i,j}\,x_{i,j}}{\log_2\left(1+\frac{1}{\eta_{i,j}^{\mathrm{SINR}}}\frac{P_i\,g_{i,j}}{\sum_{k\in\mathcal{M}\setminus\{i\}}P_k\,g_{k,j}\,\rho_k+\sigma^2}\right)},\Gamma\right\}.$$

Γ is a large constant and $\lambda_{i,j}:=\frac{r_j}{B_i\,\eta_{i,j}^{\mathrm{BW}}}$. Since $\mathcal{J}$ is a standard interference mapping and $I_i(\boldsymbol{\rho})$ is bounded above, we conclude that the fixed-point always exists and is *unique* [17, 18]. Moreover, efficient iterative methods are known to approach the fixed point with an arbitrary precision [17, 18]. We summarize the heuristic proposed in this section in Algorithm 3 (Fig. 4).

Remark 5. *The convergence of Algorithm 3 is established as follows. The convergence of Step 1 of the algorithm for a fixed load is outlined in Section 4.2, which results in a monotonically decreasing energy consumption. In Step 2, we obtain better estimates for the interference caused by active cells, which let us compute the spectral efficiency of links for the next iteration. The feasibility of the resulting optimization problem is guaranteed because we only allow for deactivation of cells (we do not allow reactivation of cells deactivated in a previous iteration step).*

6 Numerical evaluation

In the following, we present a numerical evaluation of the performance of the proposed algorithm in different networks. We start by outlining the basic simulation scenario followed by a comparison with two reference schemes with respect to the energy savings and computational time. Next, we present the ability of the proposed algorithm to incorporate a variety of different base station energy consumption models. Finally, we show the performance gains achieved by applying Algorithm 3 (Fig. 4) from Section 5.

6.1 Basic simulation scenario

The simulated network is located in a square-shaped area of size 2 km × 2 km, where L base stations are placed at locations chosen uniformly at random. Unless stated otherwise, each base station has three cells directed at 0°, 120°, and 240°, respectively. Traffic generated by users is represented by N TPs on an irregular grid. Hence, each TP represents the traffic requirements of an area of different sizes. To obtain spatially varying traffic requirements, we use the following traffic model in each run of the simulations. We define three circular hot-spot areas with centers chosen uniformly at random within the area. There are

Algorithm 3 Load-aware energy minimization

Input: Worst-case spectral efficiency $\boldsymbol{\omega}^{(-1)}=\boldsymbol{\omega}(\mathbf{1})$. Maximum number of iterations Z.

Output: Network configuration $\mathbf{X}^{(Z)}$ with low energy consumption.

1: **for** $n=0:Z$ **do**
2: Use $\boldsymbol{\omega}^{(n-1)}$ to construct Problem (5).
3: Use Alg. 2 to obtain $\mathbf{X}^{(n)}$ and remove deactivated cells from the set of cells to be considered in subsequent iterations.
4: Compute the new link spectral efficiency $\boldsymbol{\omega}^{(n)}$ for the assignment $\mathbf{X}^{(n)}$ by computing the fixed point of the standard interference mapping $\mathcal{J}$.
5: **end for**
6: Return the network configuration resulting from $\mathbf{X}^{(Z)}$.

Fig. 4 Load-aware network energy minimization algorithm. The algorithm alternately executes Algorithm 2, to obtain network configurations with low energy consumption, and updates the link spectral efficiency $\boldsymbol{\omega}$ by a fixed point algorithm

two types of TPs: "hot-spot TPs (HTP)" and "standard TPs (STP)". Each TP in the simulation has probability 0.3 of being a HTP and probability 0.7 of being a STP. While the position of STP is chosen uniformly at random within the whole area, a HTP can be assigned uniformly at random to one of three hot-spot areas. Its final position is determined in polar coordinates by sampling the distance from the hot-spot center from a normal distribution and the angle from a uniform distribution. We use a wrap around model to avoid boundary effects and determine the location of TPs to be placed outside the square-shaped area. The data rate requirements of TPs are derived from a normal distribution with $\mu_{\mathrm{d}} = 128$ kbps and variance $\sigma_{\mathrm{d}}^2 = 32$ kbps^2 with a lower bound of 1 kbps. The signal attenuation for links between cells and TPs follows the International Telecommunication Union (ITU) propagation model for urban macro cell environments with a horizontal antenna pattern for three-sector cell sites with fixed antenna patterns [38].

Unless otherwise stated, we use the following simulation parameters: $\epsilon^\star = 10^{-3}$, $\epsilon = 10^{-3}$, $B_i = 20$ MHz, $P_i = 40$ dB, $\eta_{\mathrm{SINR}} = 1$, $\eta_{\mathrm{BW}} = 0.83$, $c_i = 500$ W, and $e_i = 280$ W. The values of the last six parameters have been chosen to mimic the behavior of commercial LTE systems. Furthermore, we use $f_i(\rho_i) = 564\ \rho_i$ to model the load-dependent energy consumption, which is a value similar to the dynamic energy consumption of current macro cells with six transmit antennas [27].

The proposed algorithms are compared with a solution of the original problem in (5) and, where possible, with the centralized cell zooming approach from [9]. The solution to the problem in (5) is obtained by using Matlab 2013a in combination with IBM's CPLEX on a Intel Core i7 computer with four cores. As shown later in this section, the computational time to solve 5 grows fast with the problem size. Therefore, to solve the problem in (5) in a reasonable time for comparison purposes, we confine our attention to small networks with $M = 102$ cells ($L = 34$ base stations) and $N = 100$ TPs, unless otherwise stated. We obtained the 95 % confidence intervals depicted in the figures by applying the bias corrected and accelerated bootstrap method [39] to the outcome of 100 independent runs of the simulations. Results related to the overall network energy consumption will be normalized to the energy consumption of the network when all cells are active and fully loaded.

Definition 7 (Normalized network energy consumption)**.** *Given a TP assignment* $\mathbf{X}$ *inducing cell load* $\boldsymbol{\rho}$ *and given the resulting network energy consumption* $E(\boldsymbol{\rho})$*, the normalized network energy consumption is defined to be*

$$E_{norm}(\boldsymbol{\rho}) := \frac{E(\boldsymbol{\rho})}{E(\mathbf{1})} = \frac{E(\boldsymbol{\rho})}{\sum_{l\in\mathcal{L}} c_l + \sum_{i\in\mathcal{M}} \left(e_i + f_i(\mathbf{1})\right)},$$

where the term in the denominator is the energy consumption for a fully loaded system ($\boldsymbol{\rho} = \mathbf{1}$)*.*

We refer to the sparsity supporting majorization-minimization algorithm as "sMM" and to any algorithm that solves 5 directly as mixed-integer programming (MIP) algorithm. We refer to solutions obtained by the centralized cell zooming algorithm in [9] as "cCZ". The alternating approach proposed in Section 5 is referred to as "alternating sMM" algorithm.

6.2 Notes on the complexity

The complexity of the proposed algorithm is of the same order of solving iteratively LP problems, which is a class of problems that can be solved efficiently with many standard optimization tools [33]. In our simulations for this task, we use CPLEX, which implements the dual simplex algorithm to solve LPs [35]. Typically, our proposed algorithm terminates after a few iterations ($\ll 100$) [1]. The complexity of the proposed algorithm is linear in the complexity of the simplex method, which has a polynomial time complexity on average and an exponential time worst-case complexity. In contrast, integer programming problems are typically solved by branch and cut algorithms (also in CPLEX [35]), which have an upper bound on the number of nodes 2^{MN} and solves one LP per node resulting in an exponential complexity.

6.3 Computational performance comparison between sMM, cCZ, and MIP

The cCZ has limited capability to incorporate different energy consumption models and base stations with several sectors, so we confine ourselves to a simple base station model. We assume a homogeneous network model under which all base stations have only one omni-directional cell, and all base stations have the same energy consumption model. More precise, we use $|\mathcal{L}| = M = 100$, $|\mathcal{S}_l| = 1$ and (4) with $c_l = 500$, $e_i = 280$, $f_i(\rho_i) = 0$ ($l \in \mathcal{L}$, $i \in \mathcal{M}$).

To show trends, we start with the standard setup described above, and we gradually increase the number of TPs in the system. Figure 5 shows the *normalized network energy consumption*. As expected, the normalized network energy consumption for all three algorithms increase as the number of TPs increases. This is intuitive because additional TPs add extra rate requirements that increase the total system load, which in turn reduces the redundancy in the network to be exploited for energy savings. The proposed sMM algorithm as well as the MIP algorithm provide network configurations that exhibit much smaller normalized network energy consumption when compared with the network configurations obtained with the cCZ algorithm. The smallest energy consumptions are achieved with the MIP algorithm, which outperforms the

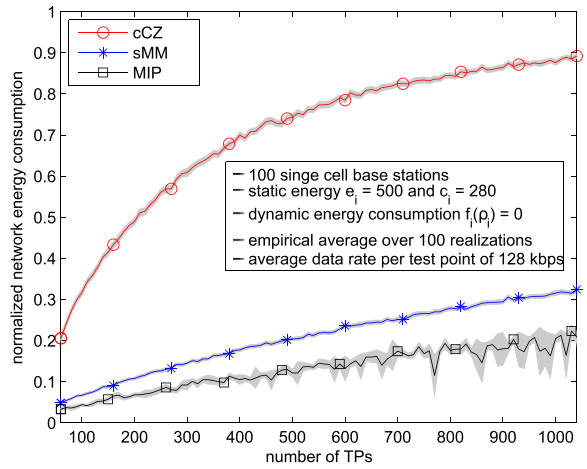

Fig. 5 Normalized network energy consumption. Comparison of normalized network energy consumption obtained with the sMM algorithm, the cCZ algorithm, and the solution of the MIP problem for increasing number of TPs. Normalization with respect to the network energy consumption for a fully loaded system ($\boldsymbol{\rho} = \mathbf{1}$) when all cells are active. Results are averaged over 100 different realizations of the network and the 95 % confidence intervals are provided in *gray*

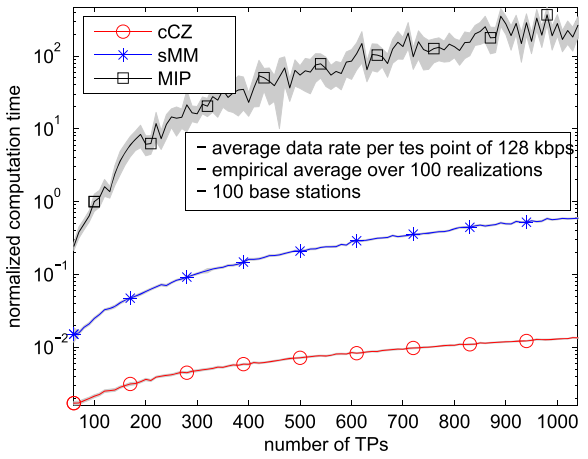

Fig. 6 Normalized computation time. Comparison of normalized computation time to obtain results with the sMM algorithm, the cCz algorithm, and the direct solution of the MIP problem. Normalization with respect to the empirical average of the MIP's computation time for 100 cells and 100 TPs over 100 realizations. The 95 % confidence intervals are provided in *gray*

proposed sMM algorithm. For the scenario with 200 TPs, the sMM algorithm results in normalized network energy consumption of 12 % on average. For the same number of TPs, the average normalized energy consumption under the cCZ and MIP algorithm are 49 and 7%, respectively. Similarly, for 1000 TPs, the resulting average normalized network energy consumption of 31 % for the sMM algorithm is still larger than the 21 % normalized energy consumption corresponding to the MIP solutions. However, it is still much smaller than cCZ with 88 % normalized energy consumption. These results emphasize that the sMM algorithm is a suboptimal heuristic, which is able to find network configurations consuming low energy. Even though the resulting network energy consumption is not globally optimal, it shows much larger energy savings than the comparison scheme cCZ. The main advantage of the proposed sMM algorithm is its fairly low computational complexity, which is directly affecting the time required to obtain an optimization result. Figure 6 depicts the normalized time needed to obtain the results of Fig. 5. This time is normalized with respect to the computation time of the MIP algorithm with 100 cells and 100 TPs. The sMM algorithm always provides results in a substantially shorter time than the MIP algorithm. Even for a relatively small scenario of 100 cells and 300 TPs, the computation time is already about 200 times larger for the MIP algorithm compared to the proposed sMM algorithm. For larger setups with 1000 TPs, the normalized time to solve the MIP was $\approx$ 237 compared to $\approx$ 0.49 for the sMM algorithm, which is an approximate 488-fold reduction in the computation time. We emphasize that the simulated scenarios are small and the computation of the MIP solution becomes infeasible in practical scenarios. Already for a network with 200 cells and 10,000 TPs, the sMM algorithm provided a solution in about 13 s, whereas the MIP algorithm could not find a solution within 1 h. Compared to the cCZ algorithm, the proposed sMM algorithm takes longer time due to the lower complexity heuristic used in the cCZ algorithm. For a scenario of 300 TPs, the average computation time is about 22 times larger for the sMM algorithm, and with 1000 TPs, it is about 43 times larger. However, with typical values of less than 1 s, the computation time is still reasonably small to allow for an online implementation. Considering the advantages in energy savings, as seen from Fig. 5, the proposed sMM algorithm presents a good trade-off between computation time and energy savings.

6.4 Cells with different sources of energy consumption

In contrast to other approaches to the problem of energy-efficient network topology control, our optimization framework can easily deal with heterogeneous networks in which cells have different static and load-dependent energy consumptions in (4). In other words, the proposed sMM algorithm can cope with different energy consumption models of cells. It can select those network configurations that exhibit as low overall energy consumption as possible. To illustrate the impact of different energy consumption models on the optimization result, we start by varying the static energy consumption of all cells, while keeping the load-dependent energy consumption fixed.

Later in this section, we show the impact of the load-dependent energy consumption by changing the weight of the load-dependent part relative to the static part.

To study the impact of the static energy consumption of cells e_i, in the following simulations, we use single-cell omni-directional base stations, and we set the load-dependent part for all cells and the common static part at base stations to zero $f(\rho_i) = 0$ and $c_l = 0$. The static energy consumption of half of the cells is varied, while the static energy consumption of the other half remains unchanged. We refer to the cells with standard fixed energy consumption as *type 1*, while *type 2* is used to refer to cells with a varying energy consumption. The energy consumption of *type 2* cells is specified relative to that of *type 1* cells. More precisely, an energy consumption relation of $\beta = 0.5$ means that if $c_i = 780$ W for *type 1* cells, then $c_i = 390$ W for *type 2* cells. The results for a scenario consisting of 100 cells and 100 TPs are shown in Fig. 7. The simulation confirms the ability of our optimization framework to incorporate different static energy consumptions. When all cells consume the same amount of energy ($\beta = 1$), the algorithm makes no difference between *type 1* and *type 2* cells. The energy consumption of *type 1* and *type 2* cells is roughly the same indicating that equally many *type 1* and *type 2* cells are active in the obtained solution. In contrast, if *type 2* cells consume less energy than *type 1* ($\beta < 1$), then the algorithm prefers to deactivate *type 1* cells, while attempting to keep *type 2* cells active. Obviously, if $\beta > 1$, the situation is reversed in the sense that, if possible, *type 2* cells are preferably selected for deactivation.

The differentiation becomes even more evident for cell deployments, where *type 1* and *type 2* cells are co-located. In such a case, two cells of different types are located at the same site and are "exchangeable" with respect to the service provided to the TPs (recall that we use omni-directional cells in these simulations). In other words, if a TP is assigned to a location with two co-located cells, then it does not matter which cell is used to provide the service to the TP. This implies that the decision whether to deactivate a cell or not should depend only on the energy consumption of this cell in relation to its co-located cell[9]. The simulations with such a deployment are shown in Fig. 8, where we see that, for $\beta < 1$, there is no active cell of *type 1*, while, for $\beta > 1$, *type 2* cells consume more energy and the simulations confirm that the algorithms clearly prefer to activate *type 1* cell.

To obtain insight into the impact of the load-dependent energy consumption, we fix the static energy consumption of a single-cell omni-directional base station to be $e_i = 780$ W and $c_l = 0$ W, and we vary the load-dependent energy consumption $f_i(\boldsymbol{\rho}) = 564\, c'\, \rho_i$ by letting c' take values on $c' \in \{0, 1, 10\}$. For an increasing number of TPs, Fig. 9 shows the fraction of active cells, while the normalized network energy consumption is shown in Fig. 10. First, we observe that the network energy consumption always increases with an increasing number of TPs, which is in fact no surprise. Moreover, the fraction of active cell increases with c' for both the sMM algorithm and the MIP algorithm. An examination of the objective function in (5a) shows that this is what we expect because if the ratio of the load-dependent energy consumption becomes larger relative to the static one, then the algorithm tends to increase the fraction of active cells for an improved load balancing in order to keep the load of each active cell at a relatively low level. In other words, instead of deactivating as many cells as possible to minimize the static energy consumption, the algorithm deactivates the cells to find the best possible balance between the static and load-dependent energy consumption. This can be observed

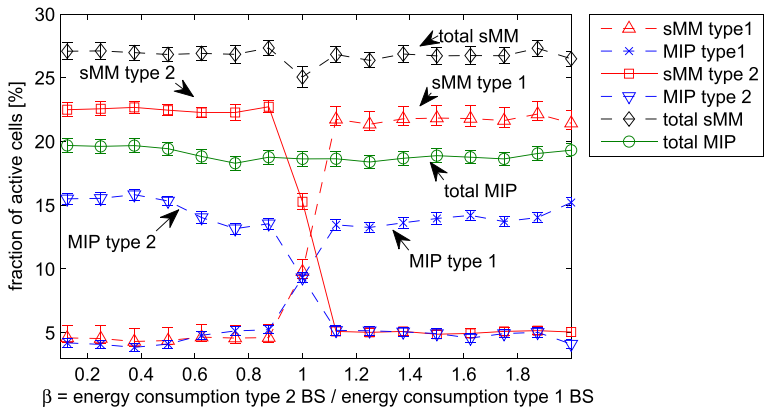

Fig. 7 Cell selection for heterogeneous energy consumption models—random deployment. Fraction of active *type 1* and *type 2* cells in the final solution obtained with sMM and MIP. Deployment uniformly at random for *type 1* and *type 2* cells. Results are averaged over 100 different realizations of the network and the 95 % confidence intervals are provided

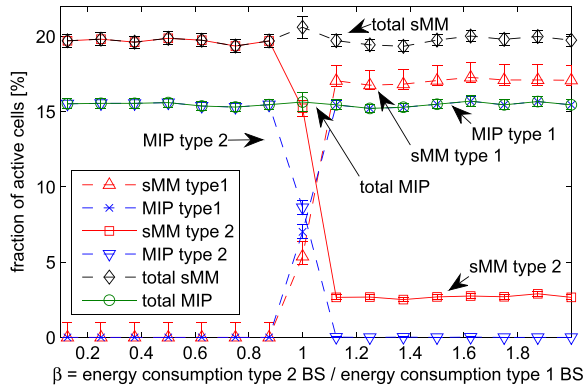

Fig. 8 Cell selection for a heterogeneous energy consumption model—co-located deployment. Fraction of *type 1* and *type 2* cells in the final solution obtained with sMM and MIP. Deployment of *type 1* cells uniformly at random and *type 2* cells are co-located with *type 1*. Results are averaged over 100 different realizations of the network and the 95 % confidence intervals are provided

in Fig. 9, where we can see that the higher is the load-dependent energy consumption (which is reflected by $c' \geq 0$), the more cells are activated under both the sMM algorithm and MIP algorithm. In particular, if $c' = 10$, then the fraction of active cells is significantly increased compared with the situation, in which the load-dependent energy consumption is negligible ($c' = 0$).

6.5 Alternating sMM algorithm

We now study the performance of the alternating sMM algorithm presented in Section 5. The standard simulation parameters are used with a total number of $Z = 10$ iterations. To show the effect of different TP requirements, we performed simulations under our standard simulation setup for different mean data rates μ_d at TPs and, for each mean data rate, we used 100 different realizations of the

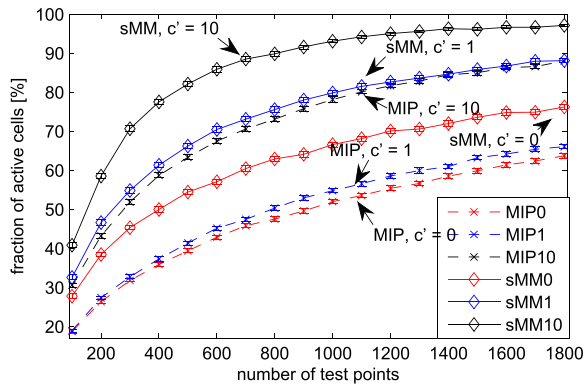

Fig. 9 Varying dynamic energy consumption—active cells. Fraction of active cells for different dynamic energy consumption c′ with increasing number of TPs. Results are averaged over 100 different realizations of the network and the 95 % confidence intervals are provided

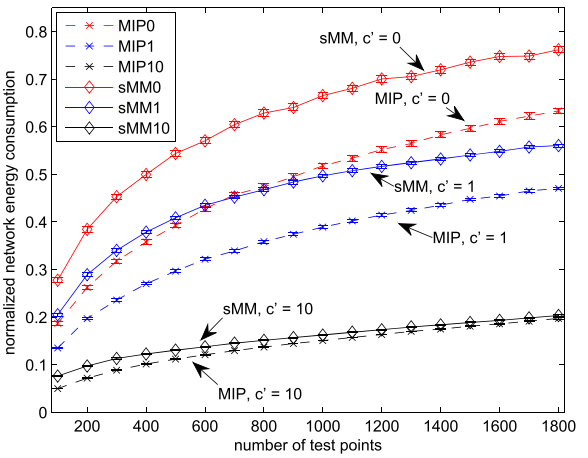

Fig. 10 Varying dynamic energy consumption—normalized network energy consumption. Normalized network energy consumption for different dynamic energy consumption weights c′ with increasing number of TPs. Normalization with respect to the energy consumption when all cells are active. Results are averaged over 100 different realizations of the network and the 95 % confidence intervals are provided

simulation scenario. The initial link spectral efficiency is computed based on the worst-case interference according to Eq. (2). Our goal is to show the huge potential for energy savings when the actual load is estimated as in Algorithm 3 (Fig. 4), instead of assuming the worst-case interference scenario, which corresponds to the full-loaded system (see Definition 5). The outcome of the simulation is depicted in Fig. 11, which includes the 95 % confidence level and shows the normalized network energy

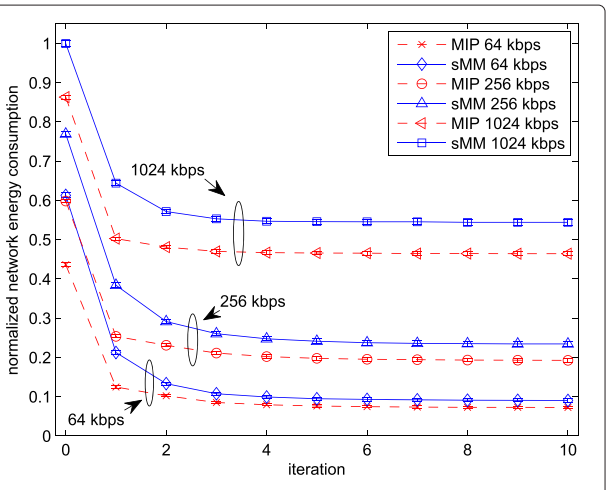

Fig. 11 Alternating sMM algorithm. Normalized network energy consumption when applying Algorithm 3 (Fig. 4) when using the sMM algorithm compared to the using the MIP solution in each iteration. Normalization with respect to the energy consumption when all cells are active. Results are averaged over 100 different realizations of the network and the 95 % confidence intervals are provided

consumption with respect to the energy consumption when all cells are active. We can see that the application of the techniques from Section 5 leads to a significant reduction of the normalized network energy consumption for both the sMM algorithm and the optimal MIP solution. Furthermore, the largest reduction was always observed after the first iteration, which shows that the worst-case interference assumption is very conservative and the load estimation may lead to considerable performance gains.

7 Conclusions

We have introduced an optimization framework for enhancing the energy efficiency of cellular networks. In wireless systems, problems of this type are hard to solve because they are combinatorial problems, and they have a complex interference coupling structure among cells. Indeed, even with a simplifying assumption of the worst-case interference, the energy saving problem is a mixed integer programming problem that is strongly related to the bin-packing problem, which in turn is known to be NP-hard. As a result, we cannot expect to find optimal solutions quickly, so we focused in this study on fast sub-optimal heuristics. Unlike many existing approaches in the literature, the proposed methods can naturally consider both the dynamic and static energy consumption of base stations with multiple cells in heterogeneous networks.

In the first proposed heuristic, we relaxed the mixed integer programming problem to a form suitable for the application of majorization-minimization techniques. The resulting algorithm requires the solution of a series of linear programming problems that can be efficiently solved with standard mathematical solvers. Therefore, it can be applied to large-scale problems, and it is also suitable for online operation. One limitation of this first method is that it uses the worst-case interference scenario, so it can be too conservative in terms of energy savings. To address this limitation, we also proposed a two-step alternating approach that obtain accurate values of the spectral efficiency of links by using the framework of standard interference functions. Simulations show that the proposed fast heuristics are able to obtain network configurations that are competitive in terms of energy consumption against optimal algorithms.

Endnotes

[1] We lend the term topology control from work in the field of ad hoc networks where it refers the task of generating a network with desired features by coordinating the nodes' transmitting range [40, 41].

[2] Although the l_0-norm is not a norm, we use the term "norm" as it is a common practice in literature.

[3] A test point becomes a user if it represents a QoS requirement of one particular user, in which case the subarea is a point corresponding to the position of this user.

[4] The smaller the area represented by each TP, the better is this approximation. However, smaller areas imply an increased number of TPs, and the computational complexity of the proposed algorithm grows.

[5] Note that B_i can also be interpreted as the total bandwidth available at cell i, in which case ρ_i is expressed in terms of the fraction of required and available bandwidth.

[6] A resource block is defined as a portion of the available time-frequency plane spanning a number of consecutive OFDM symbols in the time domain over a number of sub-carriers in the frequency domain.

[7] This relaxation together with 7c leads to a communication scenario where multiple cells serve one TP. A more detailed discussion on the implications is presented in Section 4.3

[8] In our experience a good starting point is derived from a feasible assignment matrix obtained by connecting each TP to the cell providing the strongest received signal strength.

[9] Even though such setups are unlikely in practice, we use it for reasons of illustration.

Appendix

Majorization-minimization (MM) algorithm

Here, we briefly summarize the majorization-minimization (MM) algorithm [29], which can be seen as a generalization of the well- known expectation-maximization (EM) algorithm. The presentation that follows is heavily based on that in the study in [42] (see also [1, 36]).

Suppose that the objective is to minimize a function $h : \mathcal{X} \rightarrow \mathbb{R}$, where $\mathcal{X} \subset \mathbb{R}^N$. Assume that there exists a solution to this optimization problem, and let $\mathbf{x}^\star \in \mathcal{X}$ be a global minimizer of h; i.e., $h(\mathbf{x}^\star) \leq h(\mathbf{x})$ for every $\mathbf{x} \in \mathcal{X}$. Unless h has a special structure that can be exploited (e.g., convexity), finding $\mathbf{x}^\star$ is computationally intractable in general [43]. Hence, we typically have to content ourselves with generating a sequence of vectors with non-increasing objective value. To this end, we can use the majorization-minimization (MM) technique, which drives h downhill with the help of a majorizing function $g : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$. In more detail, we say that g is majorizing function for h if it satisfies the following properties:

C.1 g majorizes h at every point in $\mathcal{X}$, i.e.,

$$h(\mathbf{x}) \leq g(\mathbf{x}, \mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}, \tag{16}$$

C.2 g and h coincide at $(\mathbf{x}, \mathbf{x})$ so that

$$h(\mathbf{x}) = g(\mathbf{x}, \mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{X}. \tag{17}$$

By starting from a feasible point $\mathbf{x}^{(0)} \in \mathcal{X}$, the MM algorithm generates a sequence $\left\{\mathbf{x}^{(n)}\right\}_{n\in\mathbb{N}} \subset \mathcal{X}$ with monotone decreasing function values $h(\mathbf{x}^{(n)})$ according to (we assume that the optimization problems have a solution)

$$\mathbf{x}^{(n+1)} \in \arg\min_{\mathbf{x}\in\mathcal{X}} g(\mathbf{x}, \mathbf{x}^{(n)}) . \tag{18}$$

Irrespective of the choice of g, we can easily verify monotonicity of the objective value with the help of (16), (17), and (18): $h(\mathbf{x}^{(n)}) = g(\mathbf{x}^{(n)}, \mathbf{x}^{(n)}) \geq g(\mathbf{x}^{(n+1)}, \mathbf{x}^{(n)}) \geq g\left(\mathbf{x}^{(n+1)}, \mathbf{x}^{(n+1)}\right) = h\left(\mathbf{x}^{(n+1)}\right)$. Therefore, since the function h is bounded below when restricted to $\mathcal{X}$ by assumption, we can conclude that $h\left(\mathbf{x}^{(n)}\right) \to c \in \mathbb{R}$ for some $c \geq h(\mathbf{x}^{\star})$ as $n \to \infty$. However, we emphasize that this in general does not imply the convergence of the sequence $\left\{\mathbf{x}^{(n)}\right\}$.

The choice of the function g is problem dependent, but it should be sufficiently structured in order to make the optimization problem in 18 tractable. In particular, in our study, we deal with concave and continuously differentiable functions h. In such cases, a natural choice for g satisfying (16) and (17) is

$$g(\mathbf{x}, \mathbf{y}) = h(\mathbf{y}) + \nabla h(\mathbf{y})^T(\mathbf{x} - \mathbf{y}). \tag{19}$$

This particular choice is common in, for example, sparse signal recovery [30].

Remark 6. *We note that, instead of solving the optimization problem in 18 exactly, it is sufficient for the monotonicity of the sequence* $\{h(\boldsymbol{x}^{(n)})\}$ *that* $g\left(\mathbf{x}^{(n+1)}, \mathbf{x}^{(n)}\right) \leq g\left(\mathbf{x}^{(n)}, \mathbf{x}^{(n)}\right)$ *for every* $n \in \mathbb{N}$. *This observation is relevant if the right-hand side of 18 can only be solved asymptotically, in which case the iteration can be truncated whenever the above inequality is satisfied.*

Competing interests
The authors declare that they have no competing interests.

Acknowledgments
This work has been partly supported by the framework of the research project ComGreen under the grant-number 01ME11010, which is funded by the German Federal Ministry of Economics and Technology (BMWi). Part of this work has been performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in METIS, although the views expressed are those of the authors and do not necessarily represent the project. The research leading to these results has received funding from the European Union's Seventh Framework Programme managed by REA - Research Executive Agency (FP7/2007-2013) under grant agreement no 286822.

Author details
[1]Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, 10587 Berlin, Germany, Einsteinufer 37, 10587 Berlin, Germany. [2]Communications and Information Theory Group, Technical University of Berlin, Einsteinufer 25, 10587 Berlin, Germany.

Received: 5 June 2015 Accepted: 26 January 2016
Published online: 25 February 2016

References

1. E Pollakis, RLG Cavalcante, S Stanczak, in *Signal Processing Advances in Wireless Communications (SPAWC), 2012 IEEE 13th International Workshop On*. Base station selection for energy efficient network operation with the majorization-minimization algorithm, (2012), pp. 219–223. doi:10.1109/SPAWC.2012.6292896
2. D Willkomm, S Machiraju, J Bolot, A Wolisz, Primary user behavior in cellular networks and implications for dynamic spectrum access. IEEE Commun. Mag. **47**(3), 88–95 (2009). doi:10.1109/MCOM.2009.4804392
3. A Corliano, M Hufschmid. Energieverbrauch der mobilen Kommunikation - Schlussbericht. Technical report, Bundesamt für Energie, Schweizerische Eidgenossenschaft, Bern, Swiss, (February 2008). (in German)
4. E Oh, B Krishnamachari, X Liu, Z Niu, Toward dynamic energy-efficient operation of cellular network infrastructure. Commun. Mag. IEEE. **49**(6), 56–61 (2011). doi:10.1109/MCOM.2011.5783985
5. Alcatel-Lucent, 9900 Wireless Network Guardian. Tech. White Paper (2008). http://www.alcatel-lucent.com. Accessed 2008
6. C Han, T Harrold, S Armour, I Krikidis, S Videv, PM Grant, H Haas, JS Thompson, I Ku, C-X Wang, TA Le, MR Nakhai, J Zhang, L Hanzo, Green radio: radio techniques to enable energy-efficient wireless networks. IEEE Commun. Mag. **49**(6), 46–54 (2011). doi:10.1109/MCOM.2011.5783984
7. RW Terry Norman, LTE infrastructure: Worldwide demand drivers and base station forecast 2012–2017. Technical report, Analysys Mason Group (May 2012)
8. P Gonzalez-Brevis, J Gondzio, Y Fan, HV Poor, J Thompson, I Krikidis, P-J Chung, in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*. Base station location optimization for minimal energy consumption in wireless networks, (2011), pp. 1–5. doi:10.1109/VETECS.2011.5956204
9. Z Niu, Y Wu, J Gong, Z Yang, Cell zooming for cost-efficient green cellular networks. IEEE Commun. Mag. **48**(11), 74–79 (2010). doi:10.1109/MCOM.2010.5621970
10. L Chiaraviglio, D Ciullo, G Koutitas, M Meo, L Tassiulas, in *Wireless On-demand Network Systems and Services (WONS), 2012 9th Annual Conference On*. Energy-efficient planning and management of cellular networks, pp. 159–166. doi:10.1109/WONS.2012.6152223
11. K Majewski, M Koonert, in *Telecommunications (AICT), 2010 Sixth Advanced International Conference On*. Conservative cell load approximation for radio networks with shannon channels and its application to LTE network planning, (2010), pp. 219–225. doi:10.1109/AICT.2010.9
12. D Lopez-Perez, I Guvenc, G de la Roche, M Kountouris, TQS Quek, J Zhang, Enhanced intercell interference coordination challenges in heterogeneous networks. Wireless Commun. IEEE. **18**(3), 22–30 (2011)
13. S Joshi, S Boyd, Sensor selection via convex optimization. IEEE Trans. Signal Process. **57**(2), 451–462 (2009)
14. S Zhou, J Gong, Z Yang, Z Niu, P Yang, in *Proc. ACM MobiCom*. Green mobile access network with dynamic base station energy saving, (Beijing, China, 2009), pp. 1–3
15. E Pollakis, RLG Cavalcante, S Stanczak, in *Signal Processing Advances in Wireless Communications (SPAWC), 2013 IEEE 14th Workshop On*. Enhancing energy efficient network operation in multi-rat cellular environments through sparse optimization, (2013), pp. 260–264. doi:10.1109/SPAWC.2013.6612052
16. I Yamada, M Yukawa, M Yamagishi, *Minimizing the Moreau envelope of nonsmooth convex functions over the fixed point set of certain quasi-nonexpansive mappings. Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. (Springer, New York, 2011), pp. 345–390
17. RD Yates, A framework for uplink power control in cellular radio systems. IEEE J. Select. Areas Commun. **13**(7), 1341–1348 (1995)
18. M Schubert, H Boche, *Interference Calculus - A General Framework for Interference Management and Network Utility Optimization*. (Springer, Berlin, 2012)
19. K Tutschku, in *INFOCOM '98. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*. Demand-based radio network planning of cellular mobile communication systems, vol. 3, (1998), pp. 1054–10613
20. E Amaldi, A Capone, F Malucelli, F Signori, in *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*. Optimization models and

algorithms for downlink umts radio planning, vol. 2, (2003), pp. 827–8312. doi:10.1109/WCNC.2003.1200478
21. P Mogensen, W Na, IZ Kovacs, F Frederiksen, A Pokhariyal, KI Pedersen, T Kolding, K Hugl, M Kuusela, in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*. LTE capacity compared to the shannon bound, (2007), pp. 1234–1238. doi:10.1109/VETECS.2007.260
22. AJ Fehske, GP Fettweis, in *Modeling Optimization in Mobile, Ad Hoc Wireless Networks (WiOpt), 2013 11th International Symposium On*. On flow level modeling of multi-cell wireless networks, (2013), pp. 572–579
23. M Kasparick, RLG Cavalcante, S Valentin, S Stanczak, M Yukawa, Kernel-based adaptive online reconstruction of coverage maps with side information. IEEE Trans. Veh. Technol. (2015). Available at http://arxiv.org/abs/1404.0979. Accessed May 2015
24. I Siomina, D Yuan, Analysis of cell load coupling for LTE network planning and optimization. IEEE Trans. Wireless Commun. **11**(6), 2287–2297 (2012)
25. CK Ho, D Yuan, L Lei, S Sun, Power and load coupling in cellular networks for energy optimization. Wireless Commun. IEEE Trans. **14**(1), 509–519 (2015)
26. RLG Cavalcante, E Pollakis, S Stanczak, in *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference On*. Power estimation in lte systems with the general framework of standard interference mappings, (2014), pp. 818–822
27. G Auer, O Blume, V Giannini, I Godor, MA Imran, Y Jading, E Katranaras, M Olsson, D Sabella, P Skillermark, W Wajda, D2.3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown. Technical report, INFSO-ICT-247733 EARTH (Energy Aware Radio and NeTwork TecHnologies) (December 2010)
28. R Cavalcante, E Pollakis, S Stanczak, S Stefanski, R Nowak, T Kürner, A Eisenblätter, D Montvila, Energy savings in cellular networks. COST IC1004 (2013). https://www.ict-earth.eu/publications/deliverables/deliverables.html
29. DR Hunter, K Lange, A tutorial on MM algorithms. Am. Stat. **58**(1), 30–37 (2004)
30. EJ Candes, MB Wakin, SP Boyd, Enhancing sparsity by reweighted l_1 minimization. J. Fourier Anal. Appl. **14**(5), 877–905 (2008)
31. BK Sriperumbudur, DA Torres, GRG Lackriet, A majorization-minimization approach to the sparse generalized eigenvalue problem. Mach. Learn. **85**(1–2), 3–39 (2011)
32. HH Bauschke, PL Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. (Springer, Springer, 2011)
33. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge Univ. Press, Cambridge, U.K., 2006)
34. C FJ Wu, On the Convergence Properties of the EM Algorithm. The Annals of statistics. **11**(1), 95–103 (1983). Institute of Mathematical Statistics: http://www.jstor.org/stable/2240463
35. IBM Corp, IBM ILOG CPLEX optimization studio - CPLEX user's manual. Tech. manual (2015). http://www.ibm.com/
36. RLG Cavalcante, S Stanczak, M Schubert, A Eisenblatter, U Turke, Toward Energy-Efficient 5G Wireless Communications Technologies: Tools for decoupling the scaling of networks from the growth of operating power. in Signal Processing Magazine, IEEE. **31**(6), 24–34 (2014)
37. AJ Fehske, H Klessig, J Voigt, GP Fettweis, Concurrent load-aware adjustment of user association and antenna tilts in self-organizing radio networks. Vehicular Technol. IEEE Trans. **62**(5), 1974–1988 (2013)
38. 3GPP, Further Advancements for EUTRA: Physical Layer Aspects (Release 9), TR 36.814 V2.0.1 (2010). http://www.3gpp.org/dynareport/36814.htm
39. B Efron, Better bootstrap confidence intervals. J. Am. Stat. Assoc. **82**(397), 171–185 (1987)
40. R Ramanathan, R Rosales-Hain, in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*. Topology control of multihop wireless networks using transmit power adjustment, vol. 2, (2000), pp. 404–4132
41. P Santi, Topology control in wireless ad hoc and sensor networks. ACM computing surveys (CSUR). **37**(2), 164–194 (2005)
42. B Sriperumbudur, D Torres, GG Lanckriet, A majorization-minimization approach to the sparse generalized eigenvalue problem. Mach. Learn. **85**(1–2), 3–39 (2011). doi:10.1007/s10994-010-5226-3
43. RT Rockafellar, *Convex Analysis,* Princeton Mathematical Series (University Press, Princeton, 1970). http://books.google.de/books?id=QTK3RwAACAAJ. Accessed 1970