HAL
archives-ouvertes.fr

# Optimal Relay Placement in Cellular Networks

Mattia Minelli, Maode Ma, Marceau Coupechoux, Jean-Marc Kélif, Marc Sigelle, Philippe Godlewski

▶ **To cite this version:**

Mattia Minelli, Maode Ma, Marceau Coupechoux, Jean-Marc Kélif, Marc Sigelle, et al.. Optimal Relay Placement in Cellular Networks. IEEE Transactions on Wireless Communications, 2014, 13 (2), pp.998-1009. 10.1109/TWC.2013.010214.130814 . hal-00958679

**HAL Id: hal-00958679**

**https://hal-imt.archives-ouvertes.fr/hal-00958679**

Submitted on 13 Mar 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Optimal Relay Placement in Cellular Networks

Mattia Minelli[1,2], Maode Ma[1], Marceau Coupechoux[2], Jean-Marc Kelif[3], Marc Sigelle[2], and Philippe Godlewski[2]

[1] Nanyang Technological University, Nanyang Avenue, Singapore
[2] Télécom ParisTech and CNRS LTCI, 46, rue Barrault, Paris, France
[3] Orange Labs, Issy-Les-Moulineaux, France

***Abstract*—In this paper, we address the problem of optimally placing relay nodes in a cellular network with the aim of maximizing cell capacity. In order to accurately model interference, we use a dynamic framework, in which users arrive at random time instants and locations, download a file and leave the system. A fixed point equation is solved to account for the interactions between stations. We also propose an extension of a fluid model to relay based cellular networks. This allows us to obtain quick approximations of the Signal to Interference plus Noise Ratio (SINR) that are very close to 3GPP LTE-A guideline results in terms of SINR distribution. We then use these formulas to develop a dedicated Simulated Annealing (SA) algorithm, which adapts dynamically the temperature to energy variations and uses a combination of coarse and fine grids to accelerate the search for an optimized solution. Simulations results are provided for both in-band and out-of-band relays. They show how relays should be placed in a cell in order to increase the capacity in case of uniform and non-uniform traffic. The crucial impact of the backhaul link is analyzed for in-band relays. Insights are given on the influence of shadowing.**

***Index Terms*—Cellular network; relay; optimal placement; fluid model; simulated annealing; processor sharing.**

## I. Introduction

Relaying is a promising feature of future cellular networks. The scenarios envisioned by the two standards IEEE 802.16j (for WiMAX networks) and 3GPP Release 10 and 11 (LTE-A) are the following: (a) coverage extension: relays should increase user experience in indoor or allow connection in shadowed zones; (b) group mobility: relays can aggregate the traffic related to a group of users within a train or a bus; (c) capacity boost: by deploying low-cost relay stations, a cellular operator can densify its network and increase its capacity. In this paper, we tackle the problem of optimal relay placement for capacity increase in an LTE-A-like cellular network.

### A. Related Work

The relay placement problem arises in various contexts: wireless sensor networks (see e.g. [1] and references therein), Wireless Local Area Networks (WLAN), WiMAX networks and cellular networks (like LTE-A); relays represent also a paradigm in information theory (see e.g. [2] and references therein). In this short literature review, we focus on the three first cases characterized by a hierarchical structure made of Relay Nodes (RN) served by Base Stations (BS) or access points. The RNs location problem can be seen as a sub-case of the well studied facility location problem, which has been proved to be NP-hard [3]. It deserves however some attention because the objective functions and the interdependence between transmitting stations through interference make it very specific.

Several papers in the literature focus on a single BS controlling several RNs [4], [5], [6], [7], [8]. In [4], authors consider the problem of optimal relay placement in a single WLAN cell. The problem consists in minimizing the average packet transmission time. The method is based on a discretization of the possible relay locations, a Lagrangian relaxation and an iterative algorithm. In [5], RNs are constrained to be on a circle around the BS, so that authors come up with a single variable optimization problem. The same approach is taken in [6] except that cooperative strategies (Decode and Forward, DF and Amplify and Forward, AF) are assumed. In [7], a region controlled by a BS is divided into sub-regions characterized by their traffic requirements. The goal of the authors is to place relays such that capacity is maximized and user minimum bandwidth requirements are satisfied. Although useful, for example in rural areas, these approaches cannot be directly applied in a cellular network made of a dense network of BSs and RNs.

References [9], [10], [11], [12], [13], [14], [15] consider several BSs and several RNs per BS. A hierarchical optimization problem is formulated in [9] for WiMAX networks: authors first focus on short term call admission control decisions (they use here the Markov Decision Process framework) and then, on the long term, on network planning (a binary integer linear problem, ILP, is solved with standard methods). Authors of [10] consider the joint deployment of BSs and RNs and try to maximize network capacity with a fixed budget. They formulate the problem as an ILP and propose a two stage RN and BS deployment algorithm to obtain sub-optimal strategies. A similar idea is used in [11]. Wang et al. [15] try to minimize the installation cost for serving a given demand or to maximize the served demand for a fixed budget. Approximate algorithms are proposed and related approximation ratios are computed or bounded. Recently, [12] has also tackled the joint problem of BSs and RNs placement, user allocation and transmit power setting. The goal is to maximize the network sum capacity while minimizing the installation cost. Sub-optimal solutions to the resulting mixed integer non linear program are obtained

*Email addresses*: mattia1@e.ntu.edu.sg (Mattia Minelli); emdma@ntu.edu.sg (Maode Ma); marceau.coupechoux@telecom-paristech.fr (Marceau Coupechoux); jeanmarc.kelif@orange.com (Jean-Marc Kelif); marc.sigelle@telecom-paristech.fr (Marc Sigelle); philippe.godlewski@godlewski@telecom-paristech.fr (Philippe Godlewski).

thanks to an iterative algorithm. These papers suffer from two main drawbacks. First, interference is never accurately modeled. It is out of the scope of [9], [12], [15]; [10] considers only noise for the calculation of data rates; [11] tries to minimize the sum of path-losses, whereas data rate is related to the Signal to Interference plus Noise Ratio (SINR); [13] assumes a single-interferer model. Second, user locations are often supposed to be known from the network planner (see e.g. [15]). This assumption is realistic in a WiMAX network but cannot be considered for a cellular network.

Three recent papers take into account co-channel interference in their problem formulation. In [16], authors propose a fixed point algorithm for the RN placement problem. RNs are however constrained to be on a circle around their serving BS. [17] relies on extensive system level simulations and constraints also RNs to be placed according to a predefined pattern. Sambale and Walke [18] propose a Simulated Annealing (SA) algorithm based on Monte Carlo simulations compliant with 3GPP LTE-A guidelines (although shadowing is ignored in the optimization). [19] deals with RN positioning for network planning and optimization. Authors set up an analytic performance evaluation model, based on simple path-loss based SINR one-dimensional calculations with a single interferer. [14], which is an extension of [18], also copes with the problem of RN placement, taking into account physical and MAC layer in their model (shadowing is taken into account), and optimizing RNs placement in terms of cell spectral efficiency. However, this paper, as well as [16], [17], [18], [19] and other works, assumes that RNs and BSs transmit at every time instant (*full buffer* traffic). Hence, it does not model buffers loads and is equivalent to a *static* approach. This often results in misleading performance evaluation, as this model tends to overestimate interference.

### B. Contributions

The contributions of this paper are the following:

- We introduce a "dynamic" framework for relays placement performance evaluation based on traffic analysis. This model overcomes the limits of widely used static models by considering users arriving in the system at random time instants and locations, downloading a file and leaving the system, rather than a fixed set of known user locations. RNs and BSs are modeled as M/G/1/PS queue and a fixed point iteration captures the interactions between transmitting nodes. Contrary to most of the literature, our framework takes into account non uniform traffic patterns.
- We propose an extension of the fluid model, developed in [20] for interference modeling, to relay-based cellular networks in order to obtain quick calculations of the SINR. We thus take into account co-channel interference of the whole network in our optimization. Numerical results show that we are very close to 3GPP LTE-A guidelines results in terms of SINR distribution.
- Based on these quick SINR calculations, we develop a dedicated SA algorithm for the relay placement problem. Our algorithm enjoys some enhancements with respect to the standard SA. These enhancements have been already investigated in the field of image processing but have not been yet considered for wireless network optimization, to the best of authors' knowledge. More specifically, we propose a method to dynamically adapt temperature to energy variations and a combination of coarse and fine grids to accelerate the search for an optimal solution.

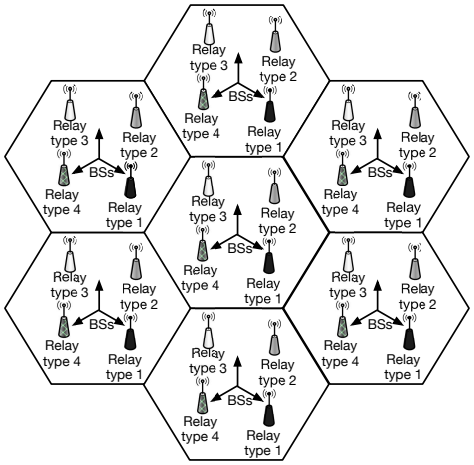


Fig. 1. Example of relay deployment with $n = 4$ relays. The cell range is $R$. Sites are tri-sectorized and boresight directions of sector BSs are shown with arrows.

The paper is organized as follows. In Section II, we present the system model, while in Section III, we deal with the performance of RNs placement, by formulating cell capacity. The SA algorithm is presented in Section IV and simulation results in Section V. Section VI concludes the paper.

## II. System Model

### A. Network Topology and Serving Station Assignment

We consider a single frequency cellular network (all stations use the same frequency) consisting of tri-sectorized hexagonal cells. Every sector is controlled by a sector BS. In every hexagonal cell, $n$ non-cooperative DF RNs [19] are deployed. Every *station* (sector BS or RN) can be either *active* (i.e., transmitting) or *idle* (i.e., not transmitting), at any given time instant. We denote with $P$ the transmit power of an active sector BS, and with $P_R$ the transmit power of an active RN. Let $\mathcal{B}$ and $\mathcal{R}$ be respectively the set of sector BSs and the set of RNs. We focus on capacity evaluation for the downlink.

The generic relay deployment is illustrated in Fig. 1. The deployment pattern is identical in all cells[1] (the relative locations of the relays w.r.t. the cell center are constant across the cells). We label RNs of each cell with indices $1 \cdots n$, and define as *type $i$* relays those RNs labeled with $i$. The set of type $i$ relays forms a regular stations pattern.

Each User Equipment (UE) is connected to its *best server*, i.e., the station which provides the highest signal power. We define $\mathcal{S}_i$ as the region, of surface $S_i$, where station $i$ is the best server. Moreover, denoting with $\mathcal{K}_c$ the set of

[1]The possibility to change the deployment pattern from cell to cell could improve performance. This aspect is left for further studies.

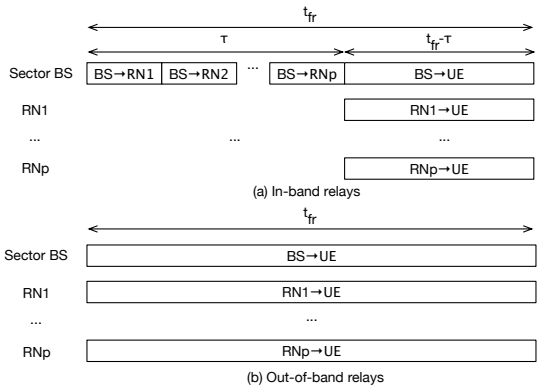


Fig. 2. Frame structure on the downlink for (a) in-band relays and (b) out-of-band relays when BS sector controls $p \leq n$ relays.

stations belonging to a given hexagonal cell $c$, we define $\mathcal{A}_c \triangleq \bigcup_{i \in \mathcal{K}_c} \mathcal{S}_i$ as the region, of surface $A_c$, where users are served by a station of $c$.

### B. Resource Organization

We consider in-band and out-of-band half-duplex relays. We assume a time division access between sector BSs and RNs illustrated in Fig. 2. Without loss of generality, the frame duration and the overall band are set to 1 time unit. When in-band relays are considered, a sector BS transmits data to the RNs it controls over the Backhaul Link (BL) during a time $\tau$. Then, during $1-\tau$ and simultaneously, the BS transmits over the Direct Link (DL, BS-UE link) and RNs transmit over the Relay Link (RL, RN-UE link) to their respective attached UEs (Fig. 3). When out-of-band relays are considered, the BL is using another frequency band (e.g. over a microwave link) or a dedicated narrow beam, so that BL radio resources are not subject to the cellular network planning. In this case, $\tau = 0$. These definitions of in-band and out-of-band relays are in accordance with the definitions of the 3GPP [21].

### C. Propagation Model

Let consider a transmitting station $j$ (RN on the RL, sector BS on the DL or BL) and a receiver $u$ (UE on the DL or RL, RN on the BL) located in $(r_j, \theta_j)$, where $r_j$ is the receiver-station distance and $\theta_j$ is defined as the angle between the receiver-station direction and the station antenna boresight direction ($\theta_j = 0$ for omnidirectional antennas). We denote with superscripts $L$ and $N$ the propagation parameters referred to Line of Sight (LOS) and Non Line of Sight (NLOS) propagation respectively, and with subscripts $B$, $R$ and $D$ the propagation parameters referred to BL, RL and DL respectively. According to the 3GPP guidelines for relay performance evaluation [22], the path-gain $g_l$ between a station and a receiver location can be written as follows:

$$g_l(r_j,\theta_j) = \delta_l(r_j) h_l^L(r_j,\theta_j) + (1-\delta_l(r_j)) h_l^N(r_j,\theta_j), \quad (1)$$
$$= \delta_l(r_j) A_l(\theta_j) \frac{K_l^L}{r_j^{\eta_l^L}} + (1-\delta_l(r_j)) A_l(\theta_j) \frac{K_l^N}{r^{\eta_l^N}} X_l^{(u,j)},$$

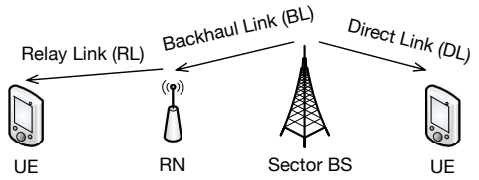


Fig. 3. Relay Link (RL), Backhaul Link (BL), and Direct Link (DL).

where $l \in \{B, R, D\}$ depends on the link type, $K_l^L$ and $K_l^N$ are propagation constants, $\eta_l^L$ and $\eta_l^N$ are path-loss exponents, $\delta_l(r_j)$ is a Bernoulli random variable (RV), which is equal to 1 when the considered link is LOS. The probability for a given link to be LOS depends on $l$ and $r_j$ (for details see [22]). Antenna pattern is given by $A_l(\theta)$. RNs are equipped with omnidirectional antennas, hence $A_R(\theta_j) = 1\ \forall \theta_j$. A sector BS uses the same antenna for BL and DL, so that $A_B = A_D \triangleq A$. Shadowing is modeled by a log-normal RV $X_l^{(u,j)}$ with standard deviation $\sigma_l$. Note that the LOS or NLOS condition is supposed to be constant in time. Shadowing is assumed to change *slowly* in time, this notion will be detailed later in the paper.

### D. Traffic Model

In this paper, we assume that each station is equivalent to an M/G/1/PS queue [23], [24]. This corresponds to a fair radio resource scheduling policy. Flow calls from users arrive according to a Poisson process of intensity $\lambda(s)ds\ [flows/s]$ in the location $s$ of surface $ds$ and flow calls sizes are i.i.d. with mean $\pi\ [bit/flow]$. We suppose for simplicity that the function $\lambda(s)$ follows the same pattern in every cell. Now, let decompose $\lambda(s) \triangleq \bar{\lambda}\phi(s)$, where the constant $\bar{\lambda} = \int_{\mathcal{A}_c} \lambda(s)ds/A_c\ [flows/s/m^2]$ is the average flow call intensity in $\mathcal{A}_c$, and $\phi(s) = \lambda(s)/\bar{\lambda}$ is named *normalized flow intensity*. The term $\phi(s)$ allows us to consider spatially non-uniform traffic and can be seen as a traffic profile. Moreover, we define $\omega(s) = \lambda(s)\pi\ [bit/s/m^2]$ as the *traffic density* in $s$. Finally, $\bar{\omega} = \bar{\lambda}\pi\ [bit/s/m^2]$ denotes the average traffic density in $\mathcal{A}_c$. The latter can be decomposed according to the surface controlled by each station: $\bar{\omega} = \sum_{i=1}^{n+3} \bar{\omega}_i \frac{S_i}{A_c}$, where $\bar{\omega}_i$ is the average traffic density on the surface $S_i$ controlled by device $i$: $\bar{\omega}_i \triangleq \frac{\pi}{S_i} \int_{\mathcal{S}_i} \lambda(s)ds$.

We denote with $C(s) > 0$ the *user spectral efficiency* (in bits/s/Hz) in location $s$, and with $\bar{\rho}_i$ the *load* of the M/G/1/PS associated to station $i$. If $\bar{\rho}_i > 1$, station $i$ is said to be *saturated*. According to the M/G/1/PS results, when $\bar{\rho}_i < 1$, $\bar{\rho}_i$ is also the probability that station $i$ is active (and thus interfere). By extension, if a station is saturated, it is always active. If $\alpha_i$ is the Bernoulli RV, which is equal to 1 when station $i$ is active and 0 otherwise, we have $\mathbb{P}[\alpha_i = 1] = \mathbb{E}[\alpha_i] = \min\{\bar{\rho}_i, 1\}$. By symmetry, all stations of the same type (sector BSs with the same antenna boresight direction or type $i$ RNs) have the same load. At a given time instant $t$, the spectral efficiency in $s$, which depends on the SINR, is an explicit function of the $\alpha_j$, $j \neq i$, where $i$ is the serving station, because of the interference term. We thus

write $C(s, \boldsymbol{\alpha}_{-i}(t))$, where $\boldsymbol{\alpha}_{-i}(t)$ is the vector of the activity variables of all stations in the network at time $t$, except $i$.

## III. Relays Placement Performance

Performance of RNs placement is measured with the corresponding cell capacity. This section is devoted to the assessment of the capacity of a cell made of three sectors and $n$ RNs. Cell capacity (derived in III-C) depends on the BL capacity (III-A), on the spectral efficiency (III-D) and on the maximum traffic intensity stations can sustain, which is obtained by deriving stations loads (III-B). Subsection III-E introduces a method to speed up capacity evaluation, by approximating interference from *far* stations.

### A. Backhaul Link Dimensioning

We derive here the proportion $\tau$ of radio resources dedicated to the BL, based on average RN throughputs, supposing that stations are not saturated. Let us consider a sector BS $b$ and the set of $p_b$ RNs controlled by $b$ (see Fig. 2). Denote $\tau_{i,b}$ the proportion of resources needed by $b$ to serve RN $i$ and $\tau_b = \sum_{i=1}^{p_b} \tau_{i,b}$. The average throughput of RN $i$ is given by $\int_{\mathcal{S}_i} \omega(s)ds = \bar{\omega}_i S_i$. Let $C_{BL}(i,b)$ be the average throughput on the BL between $b$ and $i$. The link between $b$ and RN $i$ is not overloaded if $C_{BL}(i,b)\tau_{i,b} > \bar{\omega}_i S_i$. Hence, a lower bound on $\tau_b$ is given by: $\tau_b > \sum_{i=1}^{p_b} \frac{\bar{\omega}_i S_i}{C_{BL}(i,b)}$. Note that $\tau_b$ is the load of the backhaul link between $b$ and its corresponding RNs. In order to ensure a time synchronization between all sectors of a cell, we impose $\tau = \max_b\{\tau_b\}$, so that in the best case:

$$\tau \triangleq \max_b \sum_{i=1}^{p_b} \frac{\bar{\omega}_i S_i}{C_{BL}(i,b)}. \tag{2}$$

Suppose that the flow call intensity is uniform, i.e., $\omega(s) = \bar{\omega}\ \forall s$. In this case (2) simplifies to:

$$\tau(\bar{\omega}) \triangleq \bar{\omega} \max_b \sum_{i=1}^{p_b} \frac{S_i}{C_{BL}(i,b)}. \tag{3}$$

In the following, we write $\tau(\bar{\omega})$ to designate (2) or (3) in order to show the dependency of $\tau$ on the input traffic. The value of $\tau$ is set to zero when out-of-band relaying is adopted.

### B. Stations Loads

The stations loads, which are related to their probability to be active, are coupled through the spectral efficiency function, which in turn is a function of the stations activities. We solve this problem using a fixed point iteration.

**Lemma 1.** *The load of a station $i$ is expressed by:*

$$\bar{\rho}_i = \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \int_{\mathcal{S}_i} \phi(s) \mathbb{E}_{\boldsymbol{\alpha}} \left[ \frac{1}{C(s, \boldsymbol{\alpha}_{-i})} \middle| \alpha_i = 1 \right] ds. \tag{4}$$

*Proof:*

$$\begin{aligned}
\bar{\rho}_i &\overset{(1)}{=} \lim_{T\to\infty} \frac{1}{1-\tau(\bar{\omega})} \frac{\int_0^T \int_{\mathcal{S}_i} \frac{\omega(s)}{C(s,\boldsymbol{\alpha}_{-i}(t))}\ \mathbb{1}_{\alpha_i(t)=1} ds\ dt}{\int_0^T \mathbb{1}_{\alpha_i(t)=1}\ dt}, \\
&\overset{(2)}{=} \lim_{T\to\infty} \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \frac{\frac{1}{T}\int_0^T \int_{\mathcal{S}_i} \frac{\phi(s)}{C(s,\boldsymbol{\alpha}_{-i}(t))}\ \mathbb{1}_{\alpha_i(t)=1} ds\ dt}{\frac{1}{T}\int_0^T \mathbb{1}_{\alpha_i(t)=1}\ dt}, \\
&\overset{(3)}{=} \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \frac{\mathbb{E}_{\boldsymbol{\alpha}}\left[\int_{\mathcal{S}_i} \frac{\phi(s)}{C(s,\boldsymbol{\alpha}_{-i})}\ \mathbb{1}_{\alpha_i=1}\ ds\right]}{\mathbb{P}_{\boldsymbol{\alpha}}(\alpha_i=1)}, \\
&\overset{(4)}{=} \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \mathbb{E}_{\boldsymbol{\alpha}}\left[\int_{\mathcal{S}_i} \frac{\phi(s)}{C(s,\boldsymbol{\alpha}_{-i})}\ ds \middle| \alpha_i = 1\right], \\
&\overset{(5)}{=} \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \int_{\mathcal{S}_i} \mathbb{E}_{\boldsymbol{\alpha}}\left[\frac{\phi(s)}{C(s,\boldsymbol{\alpha}_{-i})} \middle| \alpha_i = 1\right] ds.
\end{aligned}$$

(1) comes from the fact that at a given instant $t$, the load generated in a small area $ds$ in $s$ is $\frac{\omega(s)}{C(s,\boldsymbol{\alpha}_{-i}(t))}ds$ when the serving station $i$ is active. Averaging is done on the time activity of station $i$. (2) comes from $\omega(s) = \bar{\omega}\phi(s)$. (3) comes from the assumption that the process $\boldsymbol{\alpha}(t)$ has a time-stationary limit $\boldsymbol{\alpha}$. (4) is by definition of the conditional expectation. (5) comes from the assumption that the shadowing changes *slowly* in time (and hence with respect to the realizations of $\boldsymbol{\alpha}(t)$)[2]. Hence, we assume that the expectation is taken over a period of time large enough for the steady-state of the M/G/PS/1 queues to be reached, but sufficiently small to keep the channel constant at each location. Recall moreover that the LOS and NLOS conditions are fixed in time. As a consequence, $\mathcal{S}_i$ is fixed with respect to the realizations of $\boldsymbol{\alpha}(t)$. ■

We denote $\boldsymbol{\rho} \triangleq (\bar{\rho}_1, ..., \bar{\rho}_{n+3})$ the vector of loads corresponding to the $n+3$ stations ($n$ relays and three sector BSs) in the central hexagonal cell. For a given normalized flow intensity $\phi$, let define the operator $F(\boldsymbol{\rho}, \bar{\omega}) = (F_1(\boldsymbol{\rho}, \bar{\omega}), ..., F_{n+3}(\boldsymbol{\rho}, \bar{\omega}))$ as follows:

$$F_i(\boldsymbol{\rho}, \bar{\omega}) = \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \int_{\mathcal{S}_i} \phi(s) \mathbb{E}_{\boldsymbol{\alpha}} \left[ \frac{1}{C(s, \boldsymbol{\alpha}_{-i})} \middle| \alpha_i = 1 \right] ds, \tag{5}$$

where $\forall i,\ \mathbb{P}_{\boldsymbol{\alpha}}[\alpha_i = 1] = \mathbb{E}_{\boldsymbol{\alpha}}[\alpha_i] = \min\{\bar{\rho}_i, 1\}$. Let also define $\bar{\omega}_i^{max} \triangleq \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \int_{\mathcal{S}_i} \frac{\phi(s)ds}{C(s,\mathbf{1})}$.

**Theorem 1.** *If $\boldsymbol{\alpha} \mapsto C(s,\boldsymbol{\alpha})$ is a continuous mapping and $1/C(s,\boldsymbol{\alpha})$ is non-decreasing in $\boldsymbol{\alpha}$, $F : \prod_{i=1}^{n+3}[0; \bar{\omega}_i^{max}] \to \prod_{i=1}^{n+3}[0; \bar{\omega}_i^{max}]$ has at least one fixed point.*

*Proof:* We have for any $\boldsymbol{\rho} \in \prod_{i=1}^{n+3}[0; \bar{\omega}_i^{max}]$ and any

[2]Measurements for outdoor static scenarios reported in [25] justify our hypothesis, explaining this result with the fact that most of physical elements causing shadowing outdoor are fixed.

station $i$, $0 \leq F_i(\boldsymbol{\rho}, \bar{\omega})$ and:

$$\begin{aligned} F_i(\boldsymbol{\rho}, \bar{\omega}) &\overset{(1)}{=} \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \int_{\mathcal{S}_i} \phi(s) \lim_{N\to\infty} \frac{1}{N} \sum_{h=1}^{N} \frac{1}{C(s, \boldsymbol{\alpha}_{-i}(h))} ds, \\ &\overset{(2)}{\leq} \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \int_{\mathcal{S}_i} \phi(s) \lim_{N\to\infty} \frac{1}{N} \sum_{h=1}^{N} \frac{1}{C(s, \mathbf{1})} ds, \\ &\overset{(3)}{=} \frac{\bar{\omega}}{1-\tau(\bar{\omega})} \int_{\mathcal{S}_i} \frac{\phi(s)}{C(s, \mathbf{1})} ds, \\ &\overset{(4)}{=} \bar{\omega}_i^{max}, \end{aligned}$$

where $\boldsymbol{\alpha}_{-i}(h)$ indicates the $h$-th realization of RV $\boldsymbol{\alpha}_{-i}$. (1) comes from the definition of average. (2) results from the fact that, for each realization $h$ of $\boldsymbol{\alpha}_{-i}$, we have $1/C(s, \boldsymbol{\alpha}_{-i}(h)) \leq 1/C(s, \mathbf{1})$. (3) follows from the consideration that all terms $(1/C(s, \mathbf{1}))$ in the summation have the same value. Finally, (4) is by definition of $\bar{\omega}_i^{max}$. Now, using the Brouwer's fixed point theorem, we conclude the proof. ∎

A similar approach is considered in [26] (Corollary 1) for a cellular network without relays. We extend here the result to relay-based sectorized networks and we adopt a different mapping $F$ that guarantees the existence of a fixed point (which is possibly outside the $[0,1)^{n+3}$ interval). As in [26], we cannot conclude on the uniqueness of the fixed point. Starting from a central isolated cell (i.e., without interference) in the fixed point iteration makes however sense in a context of increasing traffic.

The fixed point of (5) yields the loads $\bar{\rho}_i$ of all stations $i = 1 \cdots n+3$, for a given RN placement, a given normalized flow call intensity $\phi(s)$, and a given traffic density $\bar{\omega}$.

### C. Cell Capacity

The *capacity* of the cell is defined as

$$C_{cell} \triangleq \bar{\omega}^{max} A_c \; [bit/sec/Hz/cell], \tag{6}$$

where $\bar{\omega}^{max}$ is the maximum average traffic density that can be supported by a RN placement (without any station being saturated):

$$\bar{\omega}^{max} = \max\{\bar{\omega} \in \mathbb{R}_+ : F(\boldsymbol{\rho}, \bar{\omega}) = \boldsymbol{\rho} \text{ and } \boldsymbol{\rho} \in [0,1)^{n+3}\} \tag{7}$$

The value of $\bar{\omega}^{max}$ can be found for any given normalized flow call intensity $\phi(s)$, by solving the fixed point of (5) for several values of $\bar{\omega}$ and choosing the highest $\bar{\omega}$, for which no station is saturated, according to the desired accuracy. This can be done e.g. with a dichotomic search over $\bar{\omega}$.

### D. Spectral Efficiency and SINR

We will now assume that the spectral efficiency in $s$ is derived by means of a saturated Shannon formula:

$$C(s, \boldsymbol{\alpha}_{-i}) = \min\left\{\log_2(1 + \gamma_s(\boldsymbol{\alpha}_{-i})), \tilde{C}\right\}, \tag{8}$$

where $\tilde{C}$ is the maximum achievable spectral efficiency and $\gamma_s(\boldsymbol{\alpha}_{-i})$ is the SINR in $s$. Note that this function fulfills the conditions of Theorem 1.

Now, consider a UE $u$ located in $s$ and receiving from its serving station $i$ a useful signal power $P_i(s) = \max_{j \in \mathcal{B}\cup\mathcal{R}} P_j(s)$. The SINR $\gamma_s(\boldsymbol{\alpha}_{-i})$ experienced by $u$ is:

$$\gamma_s(\boldsymbol{\alpha}_{-i}) = \frac{P_i(s)}{I_s(\boldsymbol{\alpha}_{-i}) + N_{th}}, \tag{9}$$

where $N_{th}$ is the thermal noise power and $I_s(\boldsymbol{\alpha}_{-i}) = \sum_{j \in \mathcal{B}\cup\mathcal{R},\, j\neq i} \alpha_j P_j(s)$.

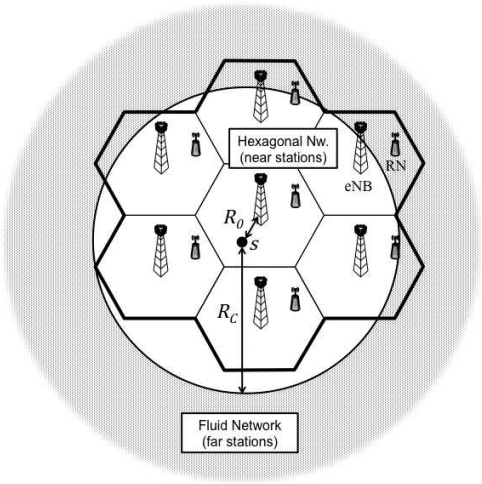


Fig. 4. Interference is exactly computed for near stations, while it is approximated for far stations by means of a fluid model of far stations network.

### E. Interference Computation with Fluid Model of Far Network Stations

Relays placements performance assessment through (6) and (8) involves intensive use of SINR evaluation: interference sum and SINR must be evaluated on the whole cell surface for each realization of $\boldsymbol{\alpha}$, at each iteration of the fixed point and for each proposed $\bar{\omega}$.

In this subsection, we propose a fast methodology for the computation of SINR, which should be simplified as much as possible in order to reduce the computation time of the optimization, while remaining accurate. Our approach decomposes the interference into two parts: the contribution of the near stations and that of the far stations. Far interference is approximated thanks to a fluid model [27] adapted to relay-based cellular networks, while interference from near stations is exactly computed.

Let define the *near BSs* as the set $\mathcal{B}_c$ of BSs in the central cell and in the first cells ring around the central cell, and *far BSs* as the set $\mathcal{B}_f$ of BSs which are not near, so that $\mathcal{B} = \mathcal{B}_c \cup \mathcal{B}_f$ and $\mathcal{B}_c \cap \mathcal{B}_f = \varnothing$. Sets $\mathcal{R}_c$ and $\mathcal{R}_f$ are similarly defined for RNs (see Figure 4).

Now, the overall interference $I_s(\boldsymbol{\alpha}_{-i})$ can be decomposed into two contributions: $I_s(\boldsymbol{\alpha}_{-i}) = I_{s,c}(\boldsymbol{\alpha}_{-i}) + I_{s,f}(\boldsymbol{\alpha}_{-i})$, where

$$I_{s,c}(\boldsymbol{\alpha}_{-i}) = \sum_{j \in \mathcal{B}_c \cup \mathcal{R}_c,\, j\neq i} \alpha_j P_j(s) \tag{10}$$

is the power received from near stations, while

$$I_{s,f}(\boldsymbol{\alpha}_{-i}) = \sum_{j \in \mathcal{B}_f \cup \mathcal{R}_f, j\neq i} \alpha_j P_j(s) \tag{11}$$

is the power received from far stations.

In our study, we compute $I_{s,c}(\boldsymbol{\alpha}_{-i})$ explicitly, while $I_{s,f}(\boldsymbol{\alpha}_{-i})$ is approximated with its average value over activity (cf. $\boldsymbol{\alpha}$) and shadowing (cf. $X$) variations:

$$\begin{aligned} I_{s,f}(\boldsymbol{\alpha}_{-i}) &\approx \mathbb{E}_{\boldsymbol{\alpha},X}\left[I_{s,f}(\boldsymbol{\alpha}_{-i})\right], \\ &= \sum_{j\in\mathcal{B}_f}\mathbb{E}_{\boldsymbol{\alpha},X}\left[\alpha_j P_j(s)\right] + \sum_{k\in\mathcal{R}_f}\mathbb{E}_{\boldsymbol{\alpha},X}\left[\alpha_k P_k(s)\right]. \end{aligned} \tag{12}$$

In the above equation, expectations should be taken *knowing* $\alpha_i = 1$. However, we make the approximation that the activities of the far stations are independent on $\alpha_i$.

It is reasonable to assume that a UE $u$ in the central cell be always served by a near station, and that the propagation on all links between $u$ and far stations be NLOS. This is valid considering commonly used shadowing standard deviations and LOS probability expressions (see e.g. [21], [22]). Our simulations, based on the assumptions suggested in [22], confirm these hypothesis, showing that less than 1% of users are connected to far stations, and LOS probability between $u$ and a far station is close to zero. Hence, we assume $I_{s,f}(\boldsymbol{\alpha}_{-i})$ to be composed by NLOS interferers, and express it as

$$\begin{aligned} &I_{s,f}(\boldsymbol{\alpha}_{-i}) \\ &\approx \sum_{j\in\mathcal{B}_f} e^{a\sigma_D^2/2}\mathbb{E}_{\boldsymbol{\alpha}}\left[\alpha_j P K_D^N r_{s,j}^{-\eta_D^N} A(\theta)\right] \\ &\quad + \sum_{k\in\mathcal{R}_f} e^{a\sigma_R^2/2}\mathbb{E}_{\boldsymbol{\alpha}}\left[\alpha_k P_R K_R^N r_{s,k}^{-\eta_R^N}\right] \\ &= e^{a\sigma_D^2/2}\sum_{k=1}^{3}\min\{1,\bar{\rho}_k\}\sum_{j\in\mathcal{B}_{f,k}} P K_D^N r_{s,j}^{-\eta_D^N} A(\theta) \\ &\quad + e^{a\sigma_R^2/2}\sum_{h=1}^{n}\min\{1,\bar{\rho}_h\}\sum_{k\in\mathcal{R}_{f,h}} P_R K_R^N r_{s,k}^{-\eta_R^N}, \end{aligned} \tag{13}$$

where $r_{s,j}$ is the distance between station $j$ and $s$, $\mathcal{B}_{f,k}$ is the set of far BSs of type $k$, $\mathcal{R}_{f,h}$ is the set of RNs of type $h$ and $a = \ln(10)/10$. Note that parameters $K_D^N$, $\eta_D^N$, $K_R^N$ and $\eta_R^N$ in (13) are referred to NLOS propagation.

Sums $\sum_{j\in\mathcal{B}_{f,k}} P K_D^N r_{s,j}^{-\eta_D^N} A(\theta)$ and $\sum_{k\in\mathcal{R}_{f,h}} P_R K_R^N r_{s,k}^{-\eta_R^N}$ can be approximated by adopting a fluid model [20] for the far stations networks $\mathcal{B}_{f,k}, k \in \{1\cdots 3\}$ and $\mathcal{R}_{f,h}, h \in \{1\cdots n\}$. The fluid model is a powerful tool to simplify interference computation in hexagonal [27] and dense Poisson [28] cellular networks. It is based on approximating a discrete set of network stations, lying on a given region, with a continuum of stations. The *density* (measured in $[stations/m^2]$) of the continuum is set to be equal to the density of discrete stations in the original network. Interference sum is then approximated by integrating received power from the continuum of stations, over the considered area. The main advantage of this approach is that it allows us to obtain approximate closed-form interference expressions, which solely depend on the distance between $s$ and the network center. We refer the reader to [27] for a detailed explanation and validation through Monte Carlo simulations.

Let now focus on $\sum_{j\in\mathcal{B}_{f,k}} P K_D^N r_{s,j}^{-\eta_D^N} A(\theta), k \in \{1\cdots 3\}$. We adopt the fluid model, substituting the hexagonal network of type-$k$ far BSs with a continuum of BSs of the same type, lying on a ring centered at $s$. Following the approach presented in [29], interference sum can be approximated as

$$\begin{aligned} &\sum_{j\in\mathcal{B}_{f,k}} P K_D^N r_{s,j}^{-\eta_D^N} A(\theta) \\ &\approx \rho_{BS}\int_{R_c-r_0}^{\infty}\int_0^{2\pi}\left(P K_D^N r^{-\eta_D^N} A(\theta)\right) r dr d\theta, \\ &= \rho_{BS} b \frac{P K_D^N}{\eta_D^N - 2}(R_c - r_0)^{2-\eta_D^N}, \end{aligned} \tag{14}$$

where $\rho_{BS}$ is the BS sites density, $b = \int_0^{2\pi} A(\theta)d\theta$, $r_0$ is the distance between the location $s$ and the cell center and $R_c$ is the distance between the closest far BS and the cell center.

Terms $\sum_{k\in\mathcal{R}_{f,h}} P_R K_R^N r_{s,k}^{-\eta_R^N}, h \in \{1\cdots n\}$ can also be approximated adopting a fluid network model, obtaining

$$\begin{aligned} &\sum_{k\in\mathcal{R}_{f,h}} P_R K_R^N r_{s,k}^{-\eta_R^N} \\ &\approx \rho_{BS}\int_{R_c-r_h}^{\infty}\int_0^{2\pi}\left(P_R K_R^N r^{-\eta_R^N}\right) r dr d\theta, \\ &= 2\pi\rho_{BS}\frac{P_R K_R^N}{\eta_R^N - 2}(R_c - r_h)^{2-\eta_R^N}, \end{aligned} \tag{15}$$

where $r_h$ is the distance between $s$ and the RN of type $h$ in the central cell.

Using (14) and (15), $I_s(\boldsymbol{\alpha}_{-i})$ is finally approximated as

$$\begin{aligned} I_s(\boldsymbol{\alpha}_{-i}) \approx &\sum_{j\in\mathcal{B}_c\cup\mathcal{R}_c,\, j\neq i}\alpha_j P_j(s) \\ &+ e^{a\sigma_D^2/2}\sum_{k=1}^{3}\min\{1,\bar{\rho}_k\}\rho_{BS} b\frac{P K_D^N}{\eta_D^N - 2}(R_c - r_0)^{2-\eta_D^N} \\ &+ e^{a\sigma_R^2/2}\sum_{h=1}^{n}\min\{1,\bar{\rho}_h\} 2\pi\rho_{BS}\frac{P_R K_R^N}{\eta_R^N - 2}(R_c - r_h)^{2-\eta_R^N}. \end{aligned} \tag{16}$$

The use of a fluid model allows to approximate the sum of interference from all far stations, at each considered cell point, by only computing $n+1$ closed-form formulae, while the hexagonal model requires to calculate the sum of $(n+3)\sum_{i=N_f}^{N_R} 6i$ far stations received powers, where $N_R$ denotes the number of cells rings around the central cell and $N_f$ the first ring of far cells. This considerably reduces computational burden, considering that interference calculation is inserted in multiple nested computation cycles, as mentioned above.

## IV. Optimal Placement

In this section, we study the optimization of relays placement, by means of a dedicated Metropolis-Hastings Simulated Annealing (SA) optimization algorithm, with the target of maximizing cell capacity $C_{cell}$, defined in Section III (see (6)). We first detail the configuration space, i.e., the set of variables to be sought, and then describe the used SA algorithm, with several enhancements which have been implemented.

In this work, a configuration is given by the positions of all relays in a cell, which we assume to lay on an hexagonal grid, due to the symmetry of the problem (see Fig. IV-1). As the RN location problem includes the capacitated facility location

problem as a special case, it is NP-hard [3]. For a typical grid of size $N = 1024$ measurements points per cell and $n = 6$ relays, the cardinality of the configuration space $\Omega$ is: $|\Omega| = C_n^N = C_6^{1024} > (1000)^6/6! > 10^{15}$. Hence, exhaustive search is infeasible in practice and we have to address other optimization techniques. Simulated Annealing (SA) is a well-known stochastic technique for solving such large combinatorial optimization problems endowed with fairly non-trivial energy landscapes. It originates to [30] but was rediscovered later [31], and with great success in network optimization up to now [32], [33]. It leads to efficient optimization if its parameters (especially the *temperature schedule*) are well set-up [34]. Recall that $\Omega$ is a finite configuration space and that we consider a cost energy function $U(x) : \Omega \mapsto \mathbb{R}$ to be minimized. In this work, the energy of a candidate configuration $x$ is the inverse of its related cell capacity (see (6)):

$$U(x) = -C_{cell}(x). \tag{17}$$

The principle of SA lays in assigning the following exponential probability to any configuration:

$$\mathbb{P}(x) = \frac{e^{-U(x)}}{Z} \quad \forall x \in \Omega \quad (\text{with } Z = \sum_{x \in \Omega} e^{-U(x)}). \tag{18}$$

A minimizer of $U(.)$ possesses thus maximal probability $\mathbb{P}(.)$ and can be found as follows (using Metropolis-Hastings variant [35]):

• Initialization: assign an arbitrary initial configuration $x_0 \in \Omega$.

• At step $m \geq 0$: let $x = x_m$ be the current configuration. Apply the following procedure: pick up a candidate configuration $x' \in \Omega$ according to a *user-specified* proposal law $r(x \to x')$ and compute then the following acceptance rate:

$$\begin{aligned} \Xi(x \to x') &= \min\left(1, \left(\frac{\mathbb{P}(x')}{\mathbb{P}(x)}\right)^{\frac{1}{T_m}} . \frac{r(x' \to x)}{r(x \to x')}\right), \\ &= \min\left(1,\ e^{-\frac{U(x')-U(x)}{T_m}} . \frac{r(x' \to x)}{r(x \to x')}\right). \end{aligned} \tag{19}$$

Assign $x_{m+1} = x'$ with probability $p = \Xi(x \to x')$ (and $x_{m+1} = x$ with probability $1 - p$).

Here, $T_m$ is a *positive temperature* parameter depending on step $m$ required to slowly decrease to 0 as $m \to +\infty$ (and rigorously to satisfy $T_m \geq \frac{T_0}{1+\log(m+1)}$). Usually a *geometric* law is adopted $T_m = T_0 \ . \ \beta^m$ with $\beta < 1$ but close to 1. Hereafter an adaptive temperature schedule is investigated. Notice also that when the proposal law is uniform or more generally symmetric, i.e., $r(x' \to x) = r(x \to x')$, the acceptance rate boils down to the usual Metropolis form:

$$\Xi(x \to x') = \min\left(1,\ e^{-\frac{U(x')-U(x)}{T_m}}\right). \tag{20}$$

This generic method enjoys a number of extensions and variants that are implemented here:

*1) "Restricted Image Spaces" [34]:* It is indeed preferable to draw at each step a configuration $x'$ which is close to $x$, for instance by varying one variable only (say, the position $\xi_i$ of the relay of type $i$) and in a restricted range around its current value. This is known to increase the global algorithm search speed [34]. To this aim we adopt a *continuous* framework, with symmetric gaussian proposal probability distribution function (PDF) given by:

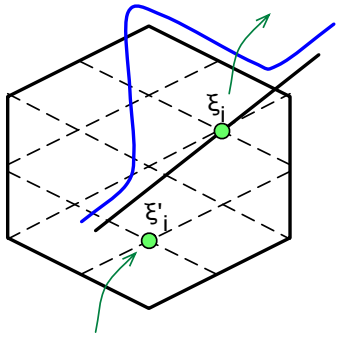


Fig. 5. Typical hexagonal cell with a hexagonal grid of possible relay locations. The current location is $\xi_i$ and the acceptance law is Gaussian along a random direction. If the candidate $\xi_i'$ falls outside the hexagon, it is "periodized" inside the cell.

$$r(\xi_i,\ \xi_i') = \frac{1}{2\pi}\ \frac{1}{\zeta^2}\ \mathcal{N}(\parallel \xi_i' - \xi_i \parallel\ ;\ \zeta^2). \tag{21}$$

Here, $\zeta$ controls the average distance between sites $\xi_i$ and $\xi_i'$ (Fig. IV-1). Finally the closest discrete grid site to $\xi_i'$ is selected. This is justified (mostly at fine resolution) for reasonable values of $\zeta$ w.r.t. the lattice step: first, there is low probability to reject proposed grid node, i.e., to find an already busy site. Then, the overall periodicity of RN lattice is injected by periodizing the proposed site itself into the whole lattice (Fig. IV-1). Thus, with this approximation the proposal laws cancels down in Hastings-Metropolis, and we obtain the usual Metropolis acceptance ratio.

*2) Adaptive temperature scheme:* In this work, we try to automatically (and adaptively) estimate both initial temperature $T_0$ and the cooling coefficient $\beta$. This has been previously done in the literature [34] provided that one knows the energy landscape (which is not the case here). This is done in two steps:

- Initial temperature setting: it is adjusted by imposing that the initial average acceptance ratio $\Xi_0 = <\Xi(x \to x')> \in [a, b]$ (in practice, we select $[a, b] = [0.5, 0.8]$). To do this, $T_0$ is first kept constant during some number of steps (depending on the size of $\Omega$). Then, a dichotomic-type update process is applied:

$$T_0 \leftarrow \beta\ T_0 \ \text{ if } \ \Xi_0 > b \ \text{ and } \ T_0 \leftarrow \frac{0.5}{\beta}\ T_0 \ \text{ if } \ \Xi_0 < a.$$

- Cooling schedule: after the previous initialization phase, a fixed number of simulated annealing steps $M^l$ and a final temperature ${T_M}^l$ [34] are assigned at each scale $l = 1, 2$ (see below), with related temperature schedule:

$${T_m}^l = {T_0}^l \ \times \ ({T_M}^l/{T_0}^l)^{m/M^l}. \tag{22}$$

*3) Multiscale implementation:* Multiscale algorithms have been employed long ago in various branches of Applied Mathematics. In image processing several multiscale strategies

have been outlined in [36], and fast algorithms for movement detection/segmentation of video sequences can be found in [37]. For instance a hand movement can be decomposed first in a large range displacement (translation-rotation) of the hand, followed by detection of fingers moves at a finer scale. Similarly here we use a two-step setup with a first optimization on a coarse grid, then refinement of this solution at a finer level. Both steps employ SA with their own adapted parameters (initial temperature, temperature step and spatial grid resolution). The advantage of such a multiscale strategy seems twofold here. First, it allows, at lower level, to avoid spurious transitions between (high level) symmetry-invariant configurations, which can arise due to the complexity of the global energy landscape. Then, it enables to check the robustness at both coarse and fine level of the placement solution to shadowing, LOS and NLOS conditions.

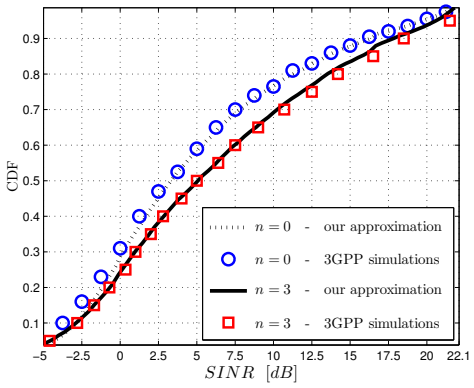


Fig. 6. SINR distribution, our SINR approximation vs. 3GPP results [22] (static case, $n = 0$ or 3 relays).

## V. Results

### A. Simulation Assumptions

*1) Spectral Efficiency Computation:* Simulation assumptions related to network layout and propagation (see (1)) follow the setup described in [22, Appendix: Simulation Assumptions] (case 3) for LTE-Advanced. In particular, the sector BS and RN transmit powers are resp. 46 dBm and 30 dBm and the cell range is 1 km. The shadowing standard deviations are 10, 8 and 6 dB on the RL, DL, and BL resp. for NLOS propagation. Shadowing RVs between cell sites and between sector BSs are correlated (see [13]). The correlation distance is 50 m (we use the Cholesky decomposition as in [38]).

The SINR is computed on a regular hexagonal grid of Measurement Points (MP) in the cell (with 25 m spacing), in order to adapt to the problem geometry. For a given realization of the propagation random variables, a given value of $\bar{\omega}$, a given vector of station loads $\boldsymbol{\rho}$ and a given realization of the activity vector $\boldsymbol{\alpha}$, the SINR is computed according to (9).

Instead of computing spectral efficiency using (8), we rely on the Modulation and Coding Schemes (MCS) indicated in [39] in order to have more realistic results. According to [39], below a certain SINR threshold, a MP is in *outage*, i.e., this position cannot be served by any station. In this case, the contribution of such MP to stations loads is zero. Any placement with an outage probability greater than 1% is rejected by the optimization process (see [40], [41]).

*2) Fixed Point Iteration:* The expectation of the spectral efficiency is computed on every MP over hundred realizations of $\boldsymbol{\alpha}$ and a new vector $\boldsymbol{\rho}$ is deduced from (4). We iteratively solve the fixed point equation $F(\boldsymbol{\rho}, \bar{\omega}) = \boldsymbol{\rho}$ (about six iterations ensures convergence in our context). If $\exists i$, s.t. $\bar{\rho}_i \geq 1$, this means that $\bar{\omega} > \bar{\omega}^{max}$, otherwise $\bar{\omega} < \bar{\omega}^{max}$, see (7). The values of $\bar{\omega}$ to be used in the fixed point iteration are selected based on a dychotomous update process, which ensures a maximum error for $C_{cell}$ equal to $\pm 2.3 \times 10^{-3}$ $[bit/s]$ (this corresponds to $\pm 23$ $[kbit/(s \times cell)]$ for a system bandwidth of 10 MHz).

*3) Simulated Annealing:* Optimization of RN positioning is performed by evaluating the energy of a number of candidate RN placements during the execution of the SA. Each RN is allowed to be located on a hexagonal grid spanning the whole cell. Two RNs are not allowed to be located on the same spot, or in the cell center. During the first optimization phase, RNs can be located on a coarse hexagonal grid (see Section IV), where the spacing between 2 neighboring MPs is 200 m. During the second phase, the spacing becomes 50 m. In this phase, candidate RN locations cannot be farther than 300 m from the location found at the end of the first phase. The number of states analyzed during each phase varies according to $n$. For example, if $n = 3$, the first phase of the SA is composed of $M^1 = 30$ temperatures steps. During each temperature step, the energy of 250 candidate placements are evaluated. The second phase of the SA is characterized by $M^2 = 20$ and 150 placements for each step. The SA is stopped if the acceptance rate $\Xi$ is zero for two consecutive temperature steps. In this case, we assume that the algorithm has already reached a 'stable' solution, and we elect as final placement the one with the lowest energy among those analyzed up to the algorithm stop. Note that the implemented SA provides *optimized* solutions and not necessarily *optimal* solutions that can be obtained theoretically after an infinite number of iterations.

### B. Model Validation

We first show the accuracy of SINR approximation introduced in Section III-E. Fig. 6 plots the Cumulative Distribution Function (CDF) of the SINR for the static case ($\boldsymbol{\alpha}(t) = \mathbf{1}, \forall t$) with $n = 3$ relays and without relays. We have compared the curves obtained in [22] with those derived by means of our SINR approximation under the same assumptions. The results show that there is a good match between our fluid model approach and the results obtained by 3GPP. This can be explained by the shadowing model (the standard deviation is relatively low and RVs are correlated) and the resulting low probability for a UE to be attached to a far station.

### C. Simulation Results

Unless specified, simulation assumptions are taken from [22] (case 3). In Fig. 7, we first show the influence of the

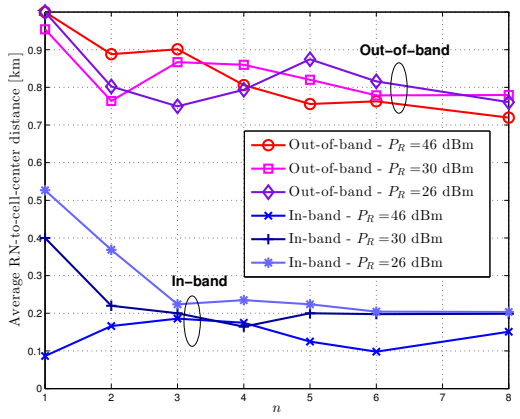


Fig. 7. Average RN to cell center distance vs. $n$, the number of RNs per cell (in-band and out-of-band, $P_R = 26$, 30 and 46 dBm).

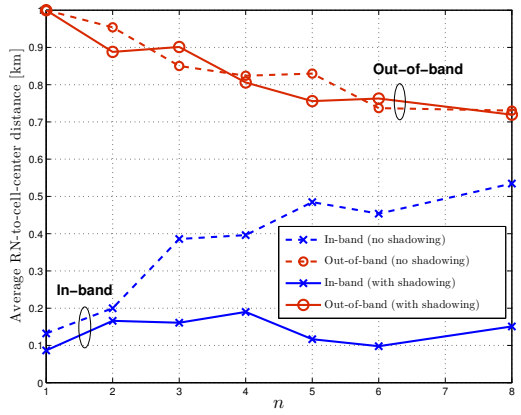


Fig. 8. Average RN to cell center distance vs. $n$ (in-band and out-of-band, $P_R = 46$ dBm, with and without shadowing).

number of RNs and of their transmit power on their distance to the cell center. Let us first consider out-of-band RNs. The absence of backhaul constraints allows the RNs to cover the most interfered regions of the cell, i.e., the cell edge. Increasing $n$ tends to decrease the average distance to the cell center because of a repulsion effect: RNs interfere more and tends to move away from each other. While a set of RNs is remaining close to cell edge, another set moves closer to the BS site so that several rings of RNs may appear. Increasing $P_R$ induces a similar but smaller repulsion effect for $n \geq 5$. When $n$ is small, $P_R$ has however less influence because the inter-RN distances are bigger and thus their mutual influence is lower. The in-band case is characterized by a trade-off between the advantage of covering the cell edge and the price to be paid on the backhaul in terms of capacity. When $P_R$ is increased, a RN controls a bigger region (especially when $n$ is small) and thus require a higher data rate on the backhaul. Consequently, RNs have to be placed closer to the sector BSs.

In Fig. 8, we can observe that the shadowing has a negligible effect on the location of out-of-band RNs. The best server mechanism assumed on the DL and the RL indeed compensates the effect of the channel variations. The impact is however decisive for in-band RNs. The best server policy is indeed not assumed on the BL[3]. This means that the BL is possibly highly interfered and its capacity is degraded. This explains that RNs have to move closer to the sector BS in order to benefit from higher MCSs and LOS propagation.

Fig. 9 and 10 show some examples of optimized RN placements. In the out-of-band case, we see how RNs are preferably placed on the cell edge. With one or two RNs (not shown on the figure), the RNs cover corners of the hexagon. When the number of RNs increases, e.g. with 6 RNs, we see how three of them move closer to the cell center. If $n$ still increases a second ring of RNs around the BS site appears. It is also noticeable from these figures that the RN placement in a given cell is coherent with the RN placement in neighboring cells. With four RNs for example, RNs attached to different cell sites form a regular pattern around the edge of the central cell. In the in-band case, RNs are much closer to the BS site because of the BL influence. The boresight direction of the sector BSs, which has almost no influence on the out-of-band case, plays now an important role. On the one hand, RNs tends to be in the boresight direction in order to benefit from a better backhaul. On the other hand, most interfered regions lies on the frontier between two sectors. For two RNs for example, the first effect is preponderant.

Fig. 11 shows the cell capacity as a function of $n$. In the out-of-band case, increasing $n$ or $P_R$ leads to a capacity increase. When there is no backhaul constraint, relaying is indeed equivalent to a classical *network densification*, which clearly increases the network capacity [42]. Higher is $P_R$, higher is the offload of the sector BSs towards RNs. As they control bigger regions, sector BSs are indeed the limiting stations: their load reaches 1 well before the RN loads do. According to (7), this limits in turn the cell capacity. Increasing $P_R$ consequently balances the traffic among RNs and BSs, which results in a higher capacity. For in-band RNs, there are two contradictory effects: increasing $P_R$ offloads the sector BSs but increases also the proportion of resources dedicated to the BL (RNs control bigger regions). Numerical results show that the first effect is slightly preponderant. The increase of capacity observed for out-of-band as well as in-band RNs contradicts the effect observed in [19], which suggests a capacity decrease after 3 to 7 RNs (depending on $P_R$). This can be explained by the fact that [19] assumes a full buffer traffic model (stations are always active) and supposes that RNs are placed on a circle around the BS. Fig. 12 confirms the influence of shadowing: it has a relative small impact on out-of-band RNs but greatly degrades the in-band RNs performance.

These results clearly show the advantage of deploying out-of-band relays, especially in terms of cell capacity. The benefit of in-band relays in a scenario similar to the one defined by the 3GPP is less obvious. First, the increase of capacity is small. Second, RNs are placed close to the BS because of the backhaul constraints, so that the capacity increase benefits mainly to UEs having already good radio conditions without relays. The deployment of in-band RNs can thus be interesting only if the BL benefits from a much higher capacity either

[3]Reference [13] proposes to attach a RN to the best BS, it is however not an option offered by the standard so far.

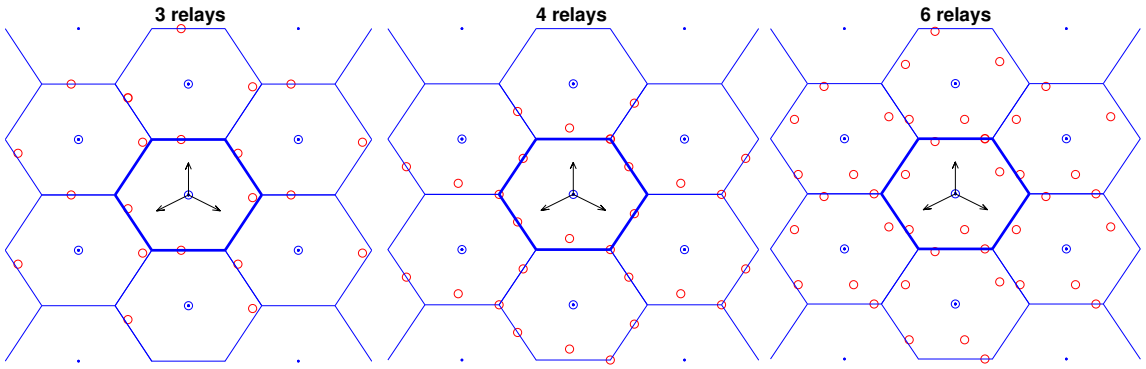


Fig. 9. Optimized out-of-band relay locations ($P_R = 30$ dBm, $n = 2$, 4 and 6, RNs are empty circles, BS sites are filled circles).

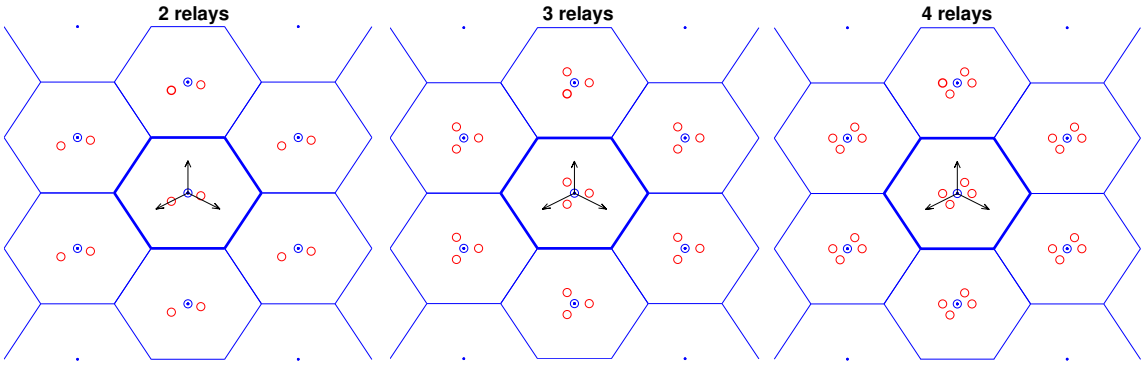


Fig. 10. Optimized in-band relay locations ($P_R = 30$ dBm, $n = 2$, 3 and 4, RNs are empty circles, BS sites are filled circles).

because of a good radio propagation (as in [43]) or because RN cell selection is adopted (as in [13]).

Finally, Fig. 13 shows an example of non-uniform traffic pattern inside the cell. The normalized traffic intensity $\phi(s)$ follows a bivariate normal distribution centered in $(0.3,\ 0.4)$ km and with standard deviation $0.3$ km. Iso-flow intensity levels are shown with concentric circles. As expected, the relays are placed closed to the hot spot. In the out-of-band case, the relays are also close to the cell border in order to cover the most interfered region. On the contrary, in-band relays are between the BS site and the hot spot in order to benefit from a good BL.

Now the question arises of the accuracy of the solutions obtained by SA. There are few results on this problem in the literature and they are generally based on some consideration over the energy surface. For example, the work of Catoni (see e.g. [44]) is based on large deviations theory, and deals with the evaluation (in a *probabilistic* framework) of the distance of the final attained configuration to the optimal one. However, the performed analysis seems not to be applicable in our case, since we have a very weak knowledge of the energy landscape of this complex problem.

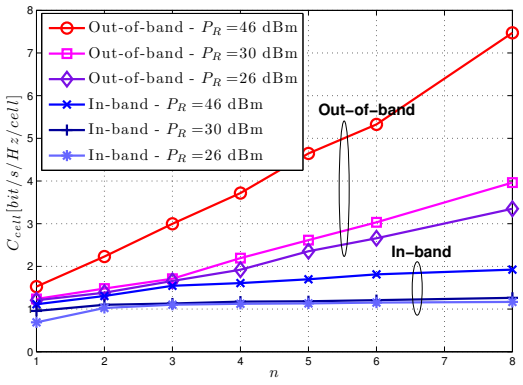


Fig. 11. Cell capacity vs. $n$ (in-band and out-of-band, $P_R = 26$, 30 and 46 dBm).

## VI. Conclusion

In this paper, we have addressed the problem of the optimal placement of relays in a cellular network with the aim of increasing the downlink cell capacity. Compared to other works on this subject, traffic and SINR computation models are more realistic: we have set up a dynamic traffic model, where each station is equivalent to a M/G/1/PS queue. The interaction between stations is captured by a fixed point equation. The system is stable if none of the stations is overloaded. This approach accurately models the activity of the stations and the relative weight of most interfered regions. Uniform and non-uniform traffic can be analyzed. The optimization of the placement is done using a dedicated Simulated Annealing algorithm. In order to speed up the search for an optimized solution, we have developed an extension of a fluid model

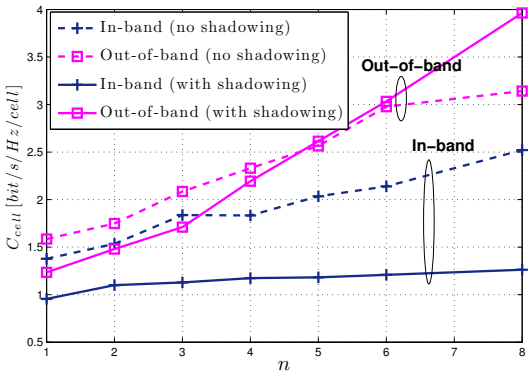


Fig. 12. Cell capacity vs. $n$ (in-band and out-of-band, $P_R = 30$ dBm, with and without shadowing).

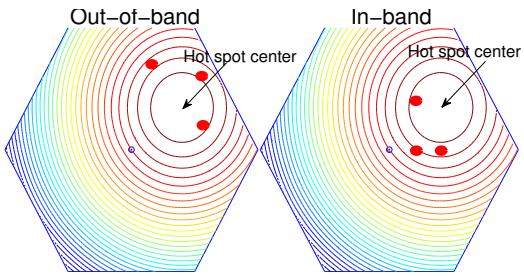


Fig. 13. Optimized out-of-band and in-band relay locations with a non-uniform traffic pattern ($P_R = 30$ dBm, $n = 3$, the BS site is in the center of the hexagons, RNs are filled circles).

to relay based cellular networks. This approach yields very good approximations of the SINR at every location. Our SA algorithm adapts dynamically the temperature to the energy variations and uses a combination of coarse and fine grids. Simulation results shows that out-of-band relays are preferably placed on the cell edge and are arranged in rings around the BS, when their number increases. In-band relays suffer from the poor quality of the backhaul link, especially in presence of shadowing and tends to be much closer to the BS. In both cases, cell capacity increases with the number of relays. The benefit is however small with in-band relays.

## Acknowledgment

The authors would like to thank the anonymous reviewers for their comments and suggestions that helped improve the quality of this paper. M. Minelli is supported by the A-STAR Singapore International Graduate Award. This work is also supported by the french ANR project NETLEARN (ANR-13-INFR-004).

## References

[1] K. Akkaya, M. Younis, and W. Youssef. Positioning of Base Stations in Wireless Sensor Networks. *IEEE Communications Magazine*, 45(4):96–102, Apr. 2007.

[2] A. Chattopadhyay, A. Sinha, M. Coupechoux, and A. Kumar. Optimal Capacity Relay Node Placement in a Multi-hop Network on a Line. In *Proc. of Int. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2012.

[3] Z. Drezner and H. W. Hamacher. *Facility Location: Applications and Theory*. Springer-Verlag, 2004.

[4] A. So and B. Liang. Enhancing WLAN Capacity by Strategic Placement of Tetherless Relay Points. *IEEE Trans. on Mobile Computing*, 6(5):487–500, May 2007.

[5] L.-C. Wang, W.-S. Su, J.-H. Huang, A. Chen, and C.-J. Chang. Optimal Relay Location in Multi-Hop Cellular Systems. In *Proc. of IEEE Wireless Communications and Networking Conf. (WCNC)*, Mar. 2008.

[6] Y. Dong, Y. Zhang, M. Song, Y. Teng, and Y. Man. Optimal Relay Location in OFDMA Based Cooperative Networks. In *Proc. of IEEE Int. Conf. on Wireless Communications, Networking and Mobile Computing (WiCom)*, Sept. 2009.

[7] C.-Y. Chang, C.-T. Chang, M.-H. Li, and C.-H. Chang. A Novel Relay Placement Mechanism for Capacity Enhancement in IEEE 802.16j WiMAX Networks. In *Proc. of IEEE Int. Conf. on Communications (ICC)*, June 2009.

[8] D. Yang, X. Fang, G. Xue, and J. Tang. Relay Station Placement for Cooperative Communications in WiMAX Networks. In *Proc. of IEEE Global Conf. on Communications (Globecom)*, Dec. 2010.

[9] D. Niyato, E. Hossain, D. I. Kim, and Z. Han. Relay-Centric Radio Resource Management and Network Planning in IEEE 802.16j Mobile Multihop Relay Networks. *IEEE Trans. on Wireless Communications*, 8(12):6115–6125, Dec. 2009.

[10] H.-C. Lu and W. Liao. Joint Base Station and Relay Station Placement for IEEE 802.16j Networks. In *Proc. of IEEE Global Conf. on Communications (Globecom)*, Nov. 2009.

[11] Y. Yu, S. Murphy, and L. Murphy. Planning Base Station and Relay Station Locations for IEEE 802.16j Network with Capacity Constraints. In *Proc. of IEEE Consumer Communications and Networking Conf. (CCNC)*, Jan. 2010.

[12] M. H. Islam, Z. Dziong, K. Sohraby, M. F. Daneshmand, and R. Jana. Capacity-Optimal Relay and Base Station Placement in Wireless Networks. In *Proc. of IEEE Int. Conf. on Information Networking (ICOIN)*, Feb. 2012.

[13] A. Bou Saleh, O. Bulakci, J. Hämäläinen, S. Redana, and B. Raaf. Analysis of the Impact of site Planning on the Performance of Relay Deployments. *IEEE Trans. on Vehicular Technology*, 61(7):3139–3150, Sept. 2012.

[14] K. Sambale and B. Walke. Cell Spectral Efficiency Optimization in Relay Enhanced Cells. In *IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2012.

[15] S. Wang, W. Zhao, and C. Wang. Approximation Algorithms for Cellular Networks Planning with Relay Nodes. In *Proc. of IEEE Wireless Communications and Networking Conf. (WCNC)*, Apr. 2013.

[16] S. Khakurel, M. Mehta, and A. Karandikar. Optimal Relay Placement for Coverage Extension in LTE-A Cellular Systems. In *Proc. of IEEE National Conf. on Communications (NCC)*, Feb. 2012.

[17] A. Hamdi, M. El-Khamy, and M. El-Sharkawy. Optimized Dual Relay Deployment for LTE-Advanced Cellular Systems. In *Proc. of IEEE Wireless Communications and Networking Conf. (WCNC)*, Apr. 2012.

[18] K. Sambale and B. Walke. Decode-and-Forward Relay Placement for Maximum Cell Spectral Efficiency. In *Proc. of IEEE European Wireless*, Apr. 2012.

[19] W. Guo and T. O'Farrell. Relay Deployment in Cellular Networks: Planning and Optimization. *IEEE Journal on Selected Areas in Communications*, 2013. to appear.

[20] J.-M. Kelif and E. Altman. Downlink Fluid Model of CDMA Networks. In *Proc. of IEEE Vehicular Technology Conf. (VTC Spring)*, May 2005.

[21] 3GPP. TR 36.814 v9.0.0 - Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9) - (2010-03), Mar. 2010.

[22] 3GPP TSG-RAN WG1 Meeting N63. Type-1 Relay Performance for Downlink. Technical report, Nokia Siemens Networks, Nokia, november 2010.

[23] F. P. Kelly. *Reversibility and Stochastic Networks*. Cambridge University Press, New York, NY, USA, 2011.

[24] T. Bonald and A. Proutière. Wireless Downlink Data Channels: User Performance and Cell Dimensioning. In *Proc. of ACM Int. Conf. Mobile Computing and Networking (Mobicom)*, Sept. 2003.

[25] A. Seetharam, J. Kurose, D. Goeckel, and G. Bhanage. A Markov Chain Model for Coarse Timescale Channel Variation in an 802.16e Wireless Network. In *Proc. of IEEE Int. Conf. on Computer Communications (Infocom)*, Mar. 2012.

[26] L. Rong, S. E. Elayoubi, and O. Ben Haddada. Performance Evaluation of Cellular Networks Offering TV Services. *IEEE Trans. on Vehicular Technology*, 60(2):644–655, Feb. 2011.

[27] J.-M. Kelif, M. Coupechoux, and P. Godlewski. A Fluid Model for Performance Analysis in Cellular Networks. *EURASIP Journal on Wireless Communications and Networking*, 2010(435189), Aug. 2010.

[28] R. Combes and J-M. Kelif. A Justification of the Fluid Network Model Using Stochastic Geometry. In *IEEE Int. Conf. on Communications (ICC)*, Jun. 2013.

[29] J.-M. Kelif, M. Coupechoux, and P. Godlewski. Fluid Model of the Outage Probability in Sectored Wireless Networks. In *Proc. of IEEE Wireless Communications and Networking Conf. (WCNC)*, Apr. 2008.

[30] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, et al. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, Mar. 1953.

[31] P.J.M. Laarhoven and E.H.L. Aarts. *Simulated Annealing: Theory and Applications*. Mathematics. Springer, 1987.

[32] G.Y. Keung, Qian Zhang, and Bo Li. The Base Station Placement for Delay-Constrained Information Coverage in Mobile Wireless Networks. In *Proc. of IEEE Int. Conf. on Communications (ICC)*, May 2010.

[33] C. Chen and F. Baccelli. Gibbsian Method for the Self-Optimization of Cellular Networks. *EURASIP Journal on Wireless Communications and Networking*, 2012(273), Aug. 2012.

[34] M.C. Robini, T. Rastello, and I.E. Magnin. Simulated Annealing, Acceleration Techniques, and Image Restoration. *IEEE Trans. on Image Processing*, 8(10):1374 – 1387, Oct. 1999.

[35] W. K. Hastings. Monte Carlo Sampling Methods using Markov Chains and their Applications. *Biometrika*, 57(1):97–109, Apr. 1970.

[36] C. Graffigne, F. Heitz, P. Pérez, F. Prêteux, M. Sigelle, and J. Zerubia. Hierarchical Markov Random Field Analysis Models Applied to Image Analysis: a Review. In *Proc. of SPIE Conf. on Neural, Morphological and Stochastic Methods in Image Processing*, July 1995.

[37] P. Pérez and F. Heitz. Restriction of a Markov Random Field on a Graph and Multiresolution Statistical Image Modeling. *IEEE Trans. on Information Theory*, 42(1):180–190, Jan. 1996.

[38] T. Klingenbrunn and P. Mogensen. Modelling Cross-Correlated Sahdowing in Network SImulations. In *Proc. of IEEE Vehicular Technology Conf. (VTC Fall)*, Sept. 1999.

[39] 3GPP. TR 36.942 v11.0.0 - Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios (Release 11) - (2012-09), Sep. 2012.

[40] C. Fischione, M. Butussi, K. H. Johansson, Kungliga Tekniska Hgskolan, C. Fischione, M. Butussi, and K. H. Johansson. Power and Rate Control with Outage Constraints in CDMA Wireless Networks. *IEEE Trans. on Communications*, 57(8):225–229, Aug. 2007.

[41] A. Alsawah and I. Fijalkow. Base-Station and Subcarrier Assignment in Two-Cell OFDMA Downlink under QoS Fairness. In *Proc. of IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2008.

[42] J.-M. Kelif and M. Coupechoux. Cell Breathing, Sectorization and Densification in Cellular Networks. In *Proc. of IEEE Int. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, June. 2009.

[43] R. Combes, Z. Altman, and E. Altman. Self-Organizing Relays: Dimensioning, Self-Optimization, and Learning. *IEEE Trans. on Network and Service Management*, 9(4):487–500, Dec. 2012.

[44] O. Catoni. Sharp Large Deviations Estimates for Simulated Annealing Algorithms. In *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, volume 27, pages 291–383. Gauthier-Villars, 1991.


**Mattia Minelli** received his Master degree in telecommunications engineering from Politecnico di Milano in Italy, in 2009. He is currently participating in a Joint PhD Program between the School of Electrical and Electronic Engineering at Nanyang Technological University in Singapore and the Computer and Network Science Department of Telecom ParisTech in Paris. His research interests include wireless networks, propagation and radio resource management issues, and relaying in OFDMA networks.

**Ma Maode** received his PhD degree in computer science from Hong Kong University of Science and Technology in 1999. Now, Dr. Ma is an Associate Professor in the School of Electrical and Electronic Engineering at Nanyang Technological University in Singapore. He has extensive research interests including wireless networking and network security. Dr. Ma has more than 250 international academic publications including over 100 journal papers and more than 130 conference papers. He currently serves as the Editor-in-Chief of *International Journal of Electronic Transport*. He also serves as an Associate Editor for other five international academic journals. Dr. Ma is an IET Fellow and a senior member of *IEEE Communication Society* and *IEEE Education Society*. He is the vice Chair of the *IEEE Education Society*, Singapore Chapter. He is also an *IEEE Communication Society Distinguished Lecturer*.

**Marceau Coupechoux** has been working as an Associate Professor at Telecom ParisTech since 2005. He obtained his Masters' degree from Telecom ParisTech in 1999 and from University of Stuttgart, Germany, in 2000, and his Ph.D. from Institut Eurecom, Sophia-Antipolis, France, in 2004. From 2000 to 2005, he was with Alcatel-Lucent. He was a Visiting Scientist at the Indian Institute of Science, Bangalore, India, during 2011–2012. Currently, at the Computer and Network Science department of Telecom ParisTech, he is working on cellular networks, wireless networks, ad hoc networks, cognitive networks, focusing mainly on radio resource management and performance evaluation.

**Jean-Marc Kelif** received the Licence of Fundamental Physics Sciences from the Louis Pasteur University of Strasbourg (France), and the Engineer degree in Materials and Solid State Physics Sciences from the University of Paris XIII in 1984. He worked as engineer in materials, and specialized in telecommunications. Since 1993, he has been with Orange Labs in Issy-les-Moulineaux, France. His current research interests include the performance evaluation, modeling, dimensioning and optimization of wireless telecommunication networks.

**Marc Sigelle** graduated as an Engineer from Ecole Polytechnique in 1975 and from Telecom ParisTech in 1977. He obtained a PhD from Telecom ParisTech in 1993. He worked first at Centre National d'Etudes des Tlcommunications in physics and computer algorithms. Since 1989 he has been working at Telecom ParisTech in image processing, his main topics of interests being the restoration and segmentation of images using stochastic/statistical models and methods (Markov Random Fields, Bayesian networks, Graph Cuts in relationships with Statistical Physics). In the past few years he has been involved in the transfer of this knowledge to the modeling and optimization (namely, stochastic) of wireless networks.

**Philippe Godlewski** received the Engineers degree from Telecom Paris (formerly ENST) in 1973 and the PhD in 1976 from University Paris 6. He is Professor at Telecom ParisTech in the Computer and Network Science department. His fields of interest include Cellular Networks, Air Interface Protocols, Multiple Access Techniques, Error Correcting Codes, Communication and Information Theory.