

Uplink Energy-Delay Trade-off under Optimized Relay Placement in Cellular Networks

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

Abstract—Relay nodes-enhanced architectures are deemed a viable solution to enhance coverage and capacity of nowadays cellular networks. Besides a number of desirable features, these architectures reduce the average distance between users and network nodes, thus allowing for battery savings for users transmitting on the uplink. In this paper, we investigate the extent of these savings, by optimizing relay nodes deployment in terms of uplink energy consumption per transmitted bit, while taking into account a minimum uplink average user delay that has to be guaranteed. A novel performance evaluation framework for uplink relay networks is first proposed to study this energy-delay trade-off. A simulated annealing is then run to find an optimized relay placement solution under a delay constraint; exterior penalty functions are used in order to deal with a difficult energy landscape, in particular when the constraint is tight. Finally, results show that relay nodes deployment consistently improve users uplink energy efficiency, under a wide range of traffic conditions and that relays are particularly efficient in non-uniform traffic scenarios.

Index Terms—Cellular networks, relay, relay placement, energy consumption, delay, simulated annealing.



1 INTRODUCTION

Heterogeneous network deployments in modern cellular networks are regarded as a promising solution to meet the ever-increasing demand of wireless data and voice traffic. They consist of installing a number of low-power nodes, possibly of different types (e.g., femtocells, Relay Nodes (RNs), etc.), inside the coverage area of macro base stations (also called eNodes-B (eNBs) in this paper), resulting in a more *dense* network architecture. Indeed, a higher density of nodes entails a number of benefits, e.g., coverage and capacity are boosted [1] and power efficiency improves [2]. This generates operational expenditures savings for operators, and lowers the environmental impact of their infrastructure. In this paper, we study the energy-delay trade-off that arises when relays are deployed in a cellular network. By reducing the distance between a UE and its serving station, relays indeed increase the uplink energy efficiency [3]. On the other hand however, the delay to reach the base station may be increased because of the two-hop communication.

1.1 Related Work

Results in [2], [4] et. al. show that the introduction of RNs and picocells can reduce downlink power consumption.

- M. Minelli is with the School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue Singapore 639798 and with Télécom ParisTech and CNRS LTCI, 46, rue Barrault, Paris, France.
E-mail: mattia.minelli@gmail.com
- M. Maode is with the School of Electrical and Electronic Engineering, Nanyang Technological University
E-mail: EMDMa@ntu.edu.sg
- M. Coupechoux, M. Sigelle and P. Godlewski are with Télécom ParisTech and CNRS LTCI.
E-mails: {first name.name}@telecom-paristech.fr
- J.-M. Kelif is with Orange Labs, Issy-Les-Moulineaux, France.
Email: jeanmarc.kelif@orange.com
- M. Coupechoux, J.-M. Kelif and P. Godlewski are partly supported by the french ANR project NETLEARN ANR-13-INFR-004.

Manuscript received xxx; revised xxx

However, this applies to the uplink as well (see e.g. [5], [6]), and the decrease in battery energy consumption for users can be consistent: [7], [8] indicate that uplink transmit power has a strong impact on users overall energy demand, especially for high transmit powers.

Hence, heterogeneous networks can be considered as an effective means to extend mobile users batteries duration, and several research works are dedicated to this topic. Authors of [5] and [9] propose a game theory-based framework for the maximization of femtocells uplink energy efficiency, by means of users transmit powers tuning [5] or radio resources assignment and power control parameters optimization [9]. In [10], closed-form expressions of the Signal to Interference Ratio (SIR) and of the outage probability are derived and used to measure the impact of femtocells coverage and users density on energy efficiency. Both [5] and [9] apply Quality of Service (QoS)-related constraints to the optimization problem, so as to avoid solutions with high energy efficiency but poor performance. Other works in the area of cellular networks, e.g. [11], address this issue by jointly optimizing power and user experienced QoS.

The close interaction between energy consumption and performance [11] makes the problem of energy efficiency maximization in relay-enhanced networks different from those addressed in [5], [9] et al. for the femtocells deployment case; this is because RNs communicate to their *donor* node through a wireless backhaul link [12], which represents a well-known performance bottleneck. On the contrary, small cells and femto cells benefit from a wired backhaul, where the delay issue is less crucial. Our work is also different from the literature on energy harvesting sensor networks, where energy-delay trade-off is also analyzed (see e.g. [13]). We indeed assume that relays are non-cooperative and fueled by time-invariant energy sources.

Surprisingly, uplink RN networks energy consumption has received limited attention in the literature, to the best of our knowledge. Reference [14] proposes an optimization

of the uplink power control in relay-based networks but ignores energy consumption. Authors of [6], [15] treat the maximization of uplink energy efficiency via either optimal assignment of subcarriers, users powers and bit allocations [6], or optimal radio resources allocation [15]. More than one order of magnitude can be achieved in user power consumption [15]. A drawback of [6] is that the decisive impact of co-channel interference is neglected. Also, the paper assumes that users are fixed in number and position, and always have data to transmit. Hence, the influence of the traffic intensity is not considered, while a number of studies (see e.g. [11], [16]) show the importance of the traffic load for energy efficiency evaluation. Finally, [6], [15] lack of a thorough theoretical framework for the analysis of relay-enhanced cellular networks uplink energy efficiency, and they are based on the sole minimization of users energy consumption, without investigating the necessary tradeoff between this consumption and uplink performance. Hence, the need arises for a finer model to study this tradeoff.

1.2 Contributions

The contributions of this paper are the following.

- We study the tradeoff between energy consumption and users experienced delay in uplink relays cellular networks. This is the first study of this type, to the best of our knowledge. We show that in many cases, relays can help both saving energy and reducing delays despite the constraint imposed by the wireless backhaul. We also highlight the interest of using relays in scenarios, where traffic is not uniform. In order to find upper bounds on the achievable gains in terms of energy saving and/or average delay, we formulate a constrained optimization problem. Its objective is to place the relays and tune the network parameters so as to minimize the energy consumption per transmitted bit under delay constraints. To solve this problem, we rely on a Simulated Annealing (SA) algorithm enhanced with exterior search and penalty functions. Exterior search and penalty functions are well-known in network modelization notably in the *deterministic* case but seldom in the *stochastic* case. Our choice of the penalty term based on the product of the objective function and of the constraint(s) to be satisfied is new to the best of our knowledge.
- We propose a framework for the analysis of energy efficiency in uplink relay networks. In order to study the delay constraint, we consider the dynamic nature of cellular traffic, by means of a model of UE arrivals and departures. This results in a hierarchical flow level analysis. The loads of eNBs and RNs are accounted for in the estimation of interference and transmission delays. Our propagation model includes shadowing and fast fading, and UEs power control is considered. It is the first time that these elements are jointly considered in the relay context. Overall, the proposed model is more comprehensive, compared to the existing literature on uplink energy efficiency in heterogeneous networks.

Table 1: Notations

Symbol	Definition
\mathcal{A}_{nw}	network area
$\mathcal{K}, \mathcal{K}_B, \mathcal{K}_R$	sets of stations, eNBs, RNs
$\mathcal{K}_c, \mathcal{A}_c$	stations and area of cell c
\mathcal{A}_k	set of UEs served by station k
n_{RN}	number of RNs
B	RN bias
β	RB quota for backhaul link
$\omega(s), \bar{\omega}, \phi(s)$	traffic density in s , average, normalized density
$\lambda(s), \xi$	flow arrival rate in s , average flow size
$\rho_k(s), \rho_k$	load generated by s on station k , load of k
$T(s), T_{max}$	power transmitted by s , max. transmit power
$G_k(s), \nu(s, \tau)$	path-loss, fast fading from s to station k
$P_k(s), P$	power received in k from s , target received power
$\gamma_k(s, \tau)$	SINR in k for UE in s on RN τ
$\Theta_j(\tau)$	is 1 if station j grants RB τ , 0 otherwise
$f_j(\tau)$	location of the UE scheduled by j on τ
$p_{\gamma,k}(s, z)$	PDF of $\gamma_k(s)$
$\pi_{\gamma,k}(s, z)$	SINR PDF of a scheduled UE in s
$C(\gamma)$	physical data rate
$R_{BL,j,k}$	harmonic average of rate bw. RN k and eNB j
$\rho_{BL,j}$	backhaul link load of station j
$\gamma_{BL,j,k}(\tau)$	SINR on backhaul link bw. RN k and eNB j
$D_k(s)$	average access delay in s
D_k	average delay for a flow served by k
$D_{BL,j,k}$	average delay on the backhaul link bw. k and j
$\epsilon(s)$	energy consumed by a UE in s to transmit 1 bit
Π	average uplink energy consumption per bit
\mathbf{x}	a configuration
$\Omega, \tilde{\Omega}$	configuration space, constrained subspace
\mathcal{F}, V	energy function, penalty function

- We adopt a non-trivial scheduling scheme, which represents an approximation of the Proportional Fair (PF) scheduler, and derive the probability density function of the Signal to Interference plus Noise Ratio (SINR) of a scheduled UE. The choice of the scheduling algorithm has indeed a decisive impact on the delay performance. This is a novel contribution, to the best of our knowledge.

The paper is organized as follows. Section 2 introduces our system model, while Section 3 describes the framework used for evaluating energy consumption per transmitted bit and delay. Section 4 is devoted to the optimization problem and our proposed algorithm, while Section 5 gives our numerical results. Finally, Section 6 concludes our work. Main notations are given in Tab. 1.

2 SYSTEM MODEL

We work on the uplink of a cellular network, where transmissions are performed on synchronized frames, and each network *station* (i.e. eNB or RN) is associated with an uplink frame. Frames are partitioned into Radio Blocks (RBs) of the same bandwidth and time duration, which are labeled with an index τ . Each RB τ in the frame of station k can be either granted to one of the UEs served by k , or be left unused (if there is no UE to be served).

The network is composed of a set \mathcal{K} of K stations, and divided into cells; let \mathcal{K}_B be the set of eNBs, and \mathcal{K}_R the set of RNs, so that $\mathcal{K} = \mathcal{K}_B \cup \mathcal{K}_R$. We focus on one given cell c containing one *donor* eNB [17] and n_{RN} RNs, which are connected to the eNB by means of a wireless *backhaul link* (Fig. 1). The set of stations of cell c is denoted \mathcal{K}_c . The

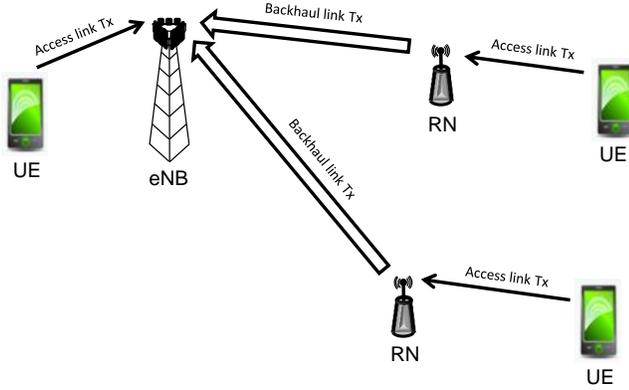


Figure 1: Uplink model: Backhaul and access links use orthogonal resources. RN transmissions towards a eNB are orthogonal. UE transmissions towards the same node (RN or eNB) are orthogonal. UE or RN transmissions towards different nodes (RN or eNB) interfere.

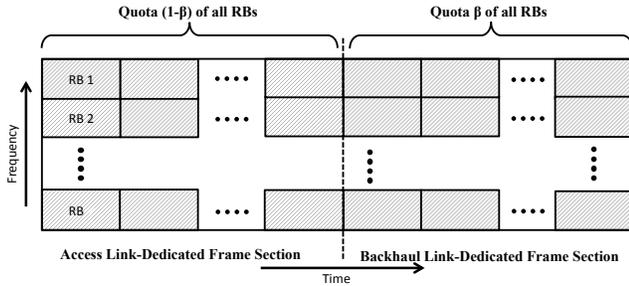


Figure 2: Frame structure.

surface \mathcal{A}_c of cell c is defined as the region where UEs are served by one of the stations of c , and its area is denoted A_c . Similarly, surface \mathcal{A}_k , of area A_k , is the region where UEs are served by station k .

UEs select their serving station based on the highest product of received downlink pilot power times a station specific *bias*. They are then served over the *access link*. The technique of biasing the user association is referred to in literature as *cell range expansion* [18], and it is deemed beneficial for network performance as it allows for load balancing. We assume that eNBs have a bias of 1 and RNs have a common bias \mathcal{B} .

2.1 Frame Structure

Access link and backhaul link transmissions are dedicated two orthogonal sections of the frame (Fig. 2). RN transmissions on the backhaul link, which uses a quota β of the RBs, do not interfere with UEs transmissions on the access link. This choice is widely adopted in the literature (see e.g. [14], [15]), as it avoids interference between RNs and UEs on the uplink. For the same reason, we assume that β is the same for all network stations. The value of β is supposed to be set by the operator and based on considerations on the overall network performance, while our analysis is focused on the performance related to one given network cell, *given* β .

2.2 Traffic Model

We assume that UEs arrive in the network according to a spatial Poisson point process of intensity $\lambda(s)$ [arrivals/s/m²], transmit a flow of average size ξ [bits] to their serving station, and leave the network. UEs are assumed to be static during their file transfer¹. Flows transmitted from UEs to a RN on the access link are then forwarded by the RN to its donor eNB on the backhaul link.

The traffic density $\omega(s)$ at a given location s is denoted by $\omega(s) = \lambda(s)\xi$ [bits/s/m²], while the average traffic density $\bar{\omega}$ in the network can be computed as $\bar{\omega} = (1/A_{nw}) \int_{\mathcal{A}_{nw}} \lambda(s)\xi ds$ [bits/s/m²], where \mathcal{A}_{nw} is the overall network surface, of area A_{nw} ². We define $\phi(s) \triangleq \omega(s)/\bar{\omega}, \forall s \in \mathcal{A}_{nw}$, to account for the local variation of the traffic density with respect to $\bar{\omega}$. The ratio $\phi(s)/A_{nw}$ can be seen as the spatial Probability Density Function (PDF) of UEs arrivals. We also define $\Phi_k \triangleq \int_{\mathcal{A}_k} \phi(s) ds$ for the sake of further developments.

The typical number of bits carried by a RB is assumed to be *much smaller* than ξ . Hence, access link buffer queue of all stations $k \in \mathcal{K}$ can be modeled as an M/G/1/PS (Processor Sharing) queue (Fig. 3), and its load is denoted with ρ_k . It is the sum of load contributions $\varrho_k(s)$ over \mathcal{A}_k , so that $\rho_k = \int_{\mathcal{A}_k} \varrho_k(s) ds$. Only *stable* scenarios, i.e., $\rho_k < 1, \forall k \in \mathcal{K}$ are considered in this work. We define the vectors $\boldsymbol{\rho} = [\rho_1 \cdots \rho_K]$ and $\boldsymbol{\rho}_{-k} = [\rho_1 \cdots \rho_{k-1} \rho_{k+1} \cdots \rho_K]$, for the sake of further developments. Let's introduce a binary random variable $\Theta_k(\tau)$, which is equal to one when RB τ in station k frame is granted to a UE, and equal to zero otherwise. We will assume that the probability for any station k to receive data on RB τ depends only on ρ_k , and is equal to $\mathbb{E}_\tau[\Theta_k(\tau)] = \rho_k$.

Backhaul link buffer queue is also modeled as a multi-class M/G/1/PS (Fig. 3). Each flow in the backhaul queue belongs to a class $k \in \{1 \cdots n_{RN}\}$, according to the transmitting RN. The model shown in Fig. 3 thus allows for a hierarchical flow level analysis.

2.3 Propagation Model

Consider a UE transmitting on RB τ , and located on s . We denote with $T(s)$ its transmit power, and with $G_k(s)$ the component of the channel gain towards station k due to distance dependent attenuation and shadowing. We assume $G_k(s)$ constant on all RBs³. The power $P_k(s)$ received by k from the considered UE is thus obtained as: $P_k(s, \tau) = T(s)G_k(s)\nu(s, \tau)$, where $\nu(s, \tau)$ is the variable component due to the fast fading. We adopt here a block Rayleigh fading model [21], i.e., fast fading is constant on a RB, and fast fading realizations associated to any two distinct RBs or locations are independent of one another. The Full Compensation Power Control (FCPC) scheme adopted by the LTE standard [17] is chosen to determine $T(s)$. We thus have: $T(s) = \min\{T_{\max}, \frac{\bar{P}}{G_k(s)}\}$, $s \in \mathcal{A}_k$, where T_{\max} is

¹Including mobility would require significant changes in the queuing model (see e.g. [19]) and is left for further study.

²We assume A_{nw} large enough so that interference in cell c is accurately computed.

³The underlying assumption is that shadowing does not significantly change in time, for a fixed location. This is consistent with, e.g., the measurements in [20].

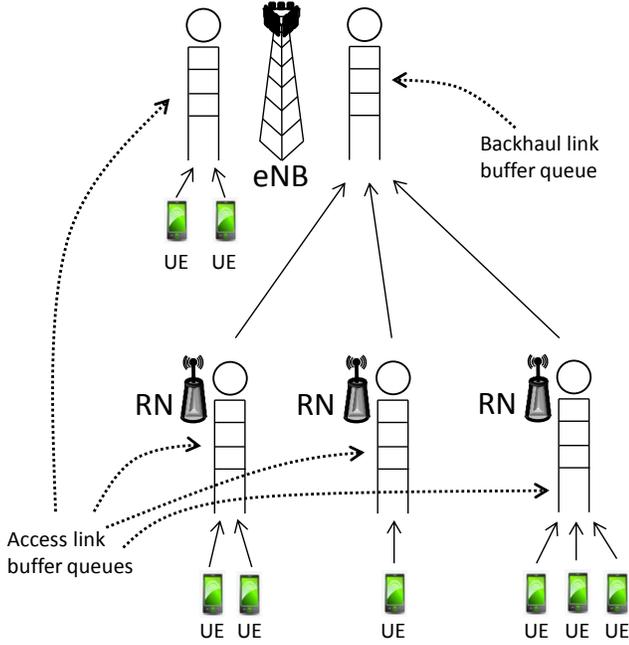


Figure 3: Queuing model.

the UE maximum transmit power and \bar{P} is a target received power broadcast by k to all UEs. We assume that \bar{P} has the same value $\bar{P} = \bar{P}_{eNB}$ for all eNBs, and $\bar{P} = \bar{P}_{RN}$ for all RNs, for the sake of simplicity.

2.4 SINR model

Consider a UE located in s and a RB τ . We define the instantaneous SINR $\gamma_k(s, \tau)$ experienced by station k for the considered UE as:

$$\gamma_k(s, \tau) = \frac{P_k(s, \tau)}{\sum_{j \in \mathcal{K}, j \neq k} \Theta_j(\tau) P_k(f_j(\tau), \tau) + N}, \quad (1)$$

where $f_j(\tau) \in \mathcal{A}_j$ yields the position of the UE scheduled by station j on RB τ , and N is the thermal noise power. Recall that $\Theta_j(\tau)$ captures the activity of station j on τ . Because of the block fading assumption, the received power depends on τ . Interferers of the UE located in s are the UEs scheduled by other stations, hence the variable $f_j(\tau)$. In the following, we drop the dependency of all variables on τ , for the sake of simplicity. Let $p_{\gamma,k}(s, z), z \geq 0$ be the PDF of $\gamma_k(s)$. Following [22], [23], we approximate $p_{\gamma,k}(s, z)$ by a lognormal distribution, i.e., $\gamma_k(s) \sim \text{Log-N}(\mu_{\gamma,k}, \sigma_{\gamma,k})$ (see App. B for the derivation of $\mu_{\gamma,k}$ and $\sigma_{\gamma,k}$).

2.5 Scheduling Policy

Let's assume that stations have perfect Channel State Information (CSI) about all served UEs and consider a given frame of station k . We label the UEs served by k during the considered frame with an index $i = 1 \dots U$, while their locations are denoted by $s_1 \dots s_U$ respectively. Every station implements a *Maximum Quantile Scheduling* [24], [25] (MQS) among served UEs. The MQS scheduler of station k assigns RB τ to the UE i that maximizes its instantaneous SINR $\gamma_k(s_i, \tau)$, with respect to $\gamma_k(s_i, \tau - h), h = 1 \dots W - 1$

over a window of W RBs (more details can be found in App. C). This policy has the property of being fair in RBs allocation between UEs, while maintaining a good throughput performance [25]. Moreover, [25] shows that the performance degradation introduced by an imperfect estimation of the SINR distribution can be lower than that incurred in practical implementations of other popular scheduling algorithms. The MQS scheduling differs from the well-known PF scheduler [26], because the latter allocates each radio resource to the user maximizing its scaled SINR $\gamma_k(s_i, \tau) / \bar{\gamma}_k(s_i)$ on τ , where $\bar{\gamma}_k(s_i)$ represents the average SINR experienced by k for user i over the last W RBs. However, for large values of W its behavior approximates that of a PF scheduler (see e.g. [24]).

Now, consider the distribution $\pi_{\gamma,k}(s_i, z) dz = \mathbb{P}(z \leq \gamma_k(s_i) \leq z + dz | f_k = s_i), s_i \in \mathcal{A}_k, z \geq 0$ of $\gamma_k(s_i)$ conditioned to $f_k = s_i$, i.e., the PDF of the SINR of a scheduled UE. Contrary to $p_{\gamma,k}(s_i, z)$, this distribution is not lognormal.

Theorem 1. The distribution $\pi_{\gamma,k}(s_i, z)$ can be expressed as:

$$\pi_{\gamma,k}(s_i, z) = p_{\gamma,k}(s_i, z) \sum_{n=1}^W Q_n(i, z) T_k(W, n), \quad (2)$$

where

$$Q_n(i, z) = \binom{W-1}{n-1} \left(\int_0^z p_{\gamma,k}(s_i, t) dt \right)^{W-n} \times \left(1 - \int_0^z p_{\gamma,k}(s_i, t) dt \right)^{n-1},$$

$$T_k(W, n) = \frac{W^2(1 - \rho_k)}{(W - \rho_k(W - n))^2}.$$

Proof. See App. C. □

2.6 UEs Physical Data Rate

We define the function $C(\gamma) > 0$, yielding the physical data rate achieved, when the receiver experiences SINR γ (or equivalently the throughput of a user, if it were alone in the serving area of its serving station). We assume that $C(\gamma)$ is non decreasing in γ .

3 PERFORMANCE EVALUATION

In this section, we define and derive our performance parameters in terms of delay and energy consumption per transmitted bit. We first obtain the expressions of access and backhaul link loads, and then we use them to get flow transmission delays.

3.1 Access Link Load

Let's focus on station k , and recall that $\varrho_k(s), s \in \mathcal{A}_k$ is the contribution of s to the load ρ_k of station k . We have the following Lemma:

Lemma 1. The local contribution $\varrho_k(s), s \in \mathcal{A}_k$ to the load of the access link buffer queue of station k is expressed by:

$$\varrho_k(s) = \frac{1}{1 - \beta} \mathbb{E}_\tau \left[\frac{\bar{\omega} \phi(s)}{C(\gamma_k(s, \tau))} \middle| f_k(\tau) = s \right]. \quad (3)$$

Proof. See Appendix A.1. □

Theorem 2. The load ρ_k of the access link of station k can be expressed as:

$$\rho_k = \frac{\bar{\omega}}{1-\beta} \int_{\mathcal{A}_k} \phi(s) \int_0^\infty \frac{\pi_{\gamma,k}(s,z)}{C(z)} dz ds. \quad (4)$$

Proof. See Appendix A.2. \square

Note first that if all UEs transmit at power $T(s) < T_{\max}, \forall s \in \mathcal{A}_k$, then $\pi_{\gamma,k}(s,z) = \pi_{\gamma,k}(z), \forall s \in \mathcal{A}_k$ and equation (4) reduces to: $\rho_k = \frac{\bar{\omega}\Phi_k}{1-\beta} \mathbb{E}_\gamma \left[\frac{1}{C(\gamma)} \right]$. Φ_k can thus be seen as the *effective area* covered by k , in terms of traffic density.

Note then that ρ_k is a function of ρ_{-k} , as the load depends on the interference from other stations. We define the operator $F_k(\rho_{-k}) = \bar{\omega}/(1-\beta) \int_{\mathcal{A}_k} \phi(s) \int_0^\infty \pi_{\gamma,k}(s,z)/C(z) dz ds$, which yields the value of ρ_k corresponding to a given ρ_{-k} .

Lemma 2. The operator $\mathbf{F}(\boldsymbol{\rho}) = (F_1(\rho_{-1}) \cdots F_K(\rho_{-K}))$ maps $\boldsymbol{\rho}$ to a finite K -dimensional interval.

Proof. See Appendix A.3. \square

Theorem 3. There exists at least one $\boldsymbol{\rho}^*$ such that $\mathbf{F}(\boldsymbol{\rho}^*) = \boldsymbol{\rho}^*, \boldsymbol{\rho}^* \in \prod_{k \in \mathcal{K}} [0, \rho_k^{\max}]$.

Proof. See Appendix A.4. \square

It is not feasible in general to draw any conclusion about the uniqueness of the fixed point. However, following the approach of [27], we start from a single cell without interference ($\boldsymbol{\rho} = \mathbf{0}$) and iterate function \mathbf{F} , so as to model a scenario of increasing traffic.

3.2 Backhaul Link Load

We now focus on the backhaul link queue of the eNB of cell c . Let $j \in \mathcal{K}_c$ be the index of the eNB and $\rho_{BL,j}$ the load of the queue. The probability for a given RN k served by j to be selected for backhaul transmission is equal to $\Phi_k / \sum_{k \in \mathcal{K}_c} \Phi_k$, where $\Phi_k = \int_{\mathcal{A}_k} \phi(s) ds$. Hence, $\rho_{BL,j}$ can be written as:

$$\rho_{BL,j} = \frac{\bar{\omega}}{\beta} \sum_{k=1}^{n_{RN}} \frac{\Phi_k}{R_{BL,j,k}}, \quad (5)$$

where the inverse of $R_{BL,j,k}$ is the average of the inverse of the rate on the backhaul link between RN k and eNB j , which depends on propagation conditions, and is equal to: $R_{BL,j,k} = 1/\mathbb{E}_\tau \left[\frac{1}{C(\gamma_{BL,j,k}(\tau))} \right]$, where $\gamma_{BL,j,k}(\tau)$ denotes the SINR on the backhaul of RN k served by eNB j , during RB τ . A detailed derivation of $R_{BL,j,k}$ is presented in App. D.

3.3 Flow Average Transmission Delay

The flow average transmission delay is an effective parameter to measure network uplink performance. It is the sum of the average access delay and the average backhaul delay. In this section, we derive its formulation. Consider a user located in s , served by RN k with donor eNB j . We denote by $D_k(s)$ the average access delay of a flow in s , by D_k the average delay of a flow served by k and by $D_{BL,j,k}$ the average delay on the backhaul link.

Lemma 3. The average access delay in s is expressed by:

$$D_k(s) = \frac{\xi}{1-\rho_k} \int_0^\infty \frac{\pi_{\gamma,k}(s,z)}{(1-\beta)C(z)} dz. \quad (6)$$

Proof. See Appendix A.5. \square

Corollary 1. The average access delay for a UE served by $k \in \mathcal{K}_R$ is:

$$D_k = \int_{\mathcal{A}_k} \frac{\phi(s)}{\Phi_k} D_k(s) ds. \quad (7)$$

Proof. See Appendix A.6. \square

Lemma 4. The average backhaul link delay is expressed by:

$$D_{BL,j,k} = \frac{\xi}{(1-\rho_{BL,j})\beta R_{BL,j,k}}, k \in \mathcal{K}_R, j \in \mathcal{K}_B. \quad (8)$$

Proof. See Appendix A.7. \square

Proposition 1. The average transmission delay \bar{D}_c for an uplink flow transmitted in cell c is equal to:

$$\bar{D}_c = \sum_{k \in \mathcal{K}_c} \frac{\Phi_k}{A_k} (D_k + D_{BL,j,k}). \quad (9)$$

3.4 Energy Consumption per Transmitted Bit

Let $\epsilon(s)$ be the energy consumed by a UE located in $s \in \mathcal{A}_k$ for transmitting one bit. This metric is sometimes called Energy Consumption Rating in the literature [28], [29]. We have:

$$\begin{aligned} \epsilon(s) &= \mathbb{E}_\tau \left[\frac{T(s)}{C(\gamma_k(s,\tau))} \middle| f_k(\tau) = s \right] \\ &= T(s) \int_0^\infty \frac{\pi_{\gamma,k}(s,z)}{C(\gamma_k(s,z))} dz, s \in \mathcal{A}_k \end{aligned} \quad (10)$$

which can be also expressed as a function of $D_k(s)$: $\epsilon(s) = T(s)D_k(s)(1-\rho_k)(1-\beta), s \in \mathcal{A}_k$. The **average uplink energy consumption per bit** Π associated to UEs in cell c is given by:

$$\begin{aligned} \Pi &= \frac{1}{A_c} \int_{\mathcal{A}_c} \phi(s) \epsilon(s) ds \\ &= \frac{1-\beta}{A_c} \sum_{k \in \mathcal{K}_c} (1-\rho_k) \int_{\mathcal{A}_k} T(s) D_k(s) \phi(s) ds. \end{aligned} \quad (11)$$

4 OPTIMIZATION

In this Section, we discuss the optimization of UEs uplink energy efficiency. The proposed algorithm aims at minimizing the UE average energy consumption per transmitted bit in (11), by optimizing P, \mathcal{B} and the RN placement in the cell of interest, for a given $\bar{\omega}, n_{RN}$ and traffic spatial profile $\phi(s)$. Our optimization is constrained to respect a maximum tolerable value of \bar{D}_c in (9), so as to take into account the users experienced quality of service.

RN placement is usually optimized with respect to the *downlink* performance of the network because of its larger traffic volume. On the contrary, we consider here the *uplink* performance because we aim at finding upper bounds on the possible gains that an operator can achieve in terms of energy and/or average delay on this link. Note however that with the growing traffic related to multimedia content sharing, social networks and other peer-to-peer

applications, uplink and downlink traffics tend to be more balanced⁴.

4.1 Problem Statement

We assume that RNs can be placed on a discrete and finite grid of candidate sites inside cell c^5 , while the positions of RNs outside cell c are assumed to be already set by the network operator, and not modifiable during the optimization. Position of a given RN i is denoted with $z(i)$. Moreover, we define $\bar{\mathbf{P}} = [\bar{P}_{eNB} \ \bar{P}_{RN}]$. Target received powers \bar{P}_{eNB} and \bar{P}_{RN} can vary between a minimum and a maximum value, denoted with \bar{P}_{\min} and \bar{P}_{\max} respectively. Similarly, we assume $\mathcal{B}_{\min} \leq \mathcal{B} \leq \mathcal{B}_{\max}$ (see Section 2). Now, the *configuration* \mathbf{x} of the network is defined as the set of positions of the n_{RN} RNs in cell c , plus the adopted $\bar{\mathbf{P}}$ and \mathcal{B} :

$$\begin{aligned} \mathbf{x} : \{ & z(1), \dots, z(n_{RN}), \bar{\mathbf{P}}, \mathcal{B} \}, \\ & z(i) \in \mathcal{A}_c, \forall i \in \{1 \dots n_{RN}\}, \\ & \bar{P}_{\min} \leq \bar{P} \leq \bar{P}_{\max}, \mathcal{B}_{\min} \leq \mathcal{B} \leq \mathcal{B}_{\max}. \end{aligned} \quad (12)$$

We name *configuration space* the set of all configurations, and denote it with Ω .

Our problem is to minimize energy consumption (11) under delay constraint (9):

$$\min_{\mathbf{x}} \Pi(\mathbf{x}) \quad (13)$$

$$\text{s.t. } \bar{D}_c(\mathbf{x}) \leq D_{\max}, \mathbf{x} \in \Omega, D_{\max} \in \mathbb{R}^+, \quad (14)$$

where constraint (14) restricts the domain of feasible solutions to the subspace:

$$\tilde{\Omega} = \{ \mathbf{x} \in \Omega \text{ s.t. } D_c(\mathbf{x}) \leq D_{\max} \} \quad (15)$$

Recall that the computation of the station loads via the fixed point iteration of Section 3.1 is required for the evaluation of the delay.

Now, the problem (13) is in general non-convex, and the cardinality of Ω , especially for high n_{RN} , makes it intractable via exhaustive search. Hence, we rely on stochastic optimization algorithms, and propose a customized version of the well-known Simulated Annealing algorithm to solve (13). In the following, we first briefly recall the generic SA, and then introduce our version.

4.1.1 Generic SA Algorithm

The SA is a metaheuristic aimed at solving large non-convex problems, which has been first proposed by Metropolis [31] and then applied on a wide range of optimization problems (see, e.g., [32], [33]). The literature shows that the SA is an effective algorithm, if its parameters are appropriately tuned (see, e.g., [34]). Let $\mathcal{F}(\mathbf{x})$ denote the *energy*⁶ associated with configuration \mathbf{x} , and consider the problem of minimizing $\mathcal{F}(\mathbf{x})$ over the configuration space Ω . The SA explores only a subset $\Psi \subset \Omega$ of the configuration space, where usually

$|\Psi| \ll |\Omega|$, and is able to find the optimal configuration by means of an *appropriate* selection of the analyzed configurations. At temperature T , the algorithm proceeds by assigning to each configuration \mathbf{x} an exponential probability, given by:

$$\mathbb{P}_T(\mathbf{x}) = \frac{\exp(-\frac{\mathcal{F}(\mathbf{x})}{T})}{Z_T}, \forall \mathbf{x} \in \Omega, \quad (16)$$

where Z_T is a normalizing constant. Hence, the solution is the \mathbf{x} that maximizes $\mathbb{P}_T(\mathbf{x})$. According to the Metropolis-Hastings variant [35] of the SA, the set Ψ of configurations to be analyzed is determined according to the following procedure:

- At step $m = 0$ a configuration $\mathbf{x}(0) \in \Omega$ is arbitrarily selected and designated as *current solution* $\mathbf{x}^*(0)$ for step 0: $\mathbf{x}^*(0) \leftarrow \mathbf{x}(0) \in \Omega$.
- At any step $m \geq 0$: assume $\mathbf{x}^*(m-1)$ is the current solution for step $m-1$. The SA will first draw a configuration $\mathbf{x}'(m) \in \Omega$ for step m from a given probability distribution $r(\mathbf{x}^*(m-1) \rightarrow \mathbf{x}'(m))$, named *proposal law* and defined by the algorithm designer. Then, assuming that r is symmetric, $\mathbf{x}'(m)$ will be accepted as current solution for step m with probability:

$$\mathbb{P}_{T_m}(\mathbf{x}^*(m) \leftarrow \mathbf{x}'(m)) = \min \left(1, \exp \left(-\frac{\mathcal{F}(\mathbf{x}'(m)) - \mathcal{F}(\mathbf{x}^*(m-1))}{T_m} \right) \right),$$

where T_m is a parameter called *temperature*, which decreases to zero *slowly enough* as $m \rightarrow +\infty$, and is such that $T_m \geq \frac{T_0}{1+\log(m+1)}$. If $\mathbf{x}'(m)$ is not accepted, then $\mathbf{x}^*(m) \leftarrow \mathbf{x}^*(m-1)$. In most practical implementations of the SA, the temperature is updated following a law of the kind $T_m = T_0 h^m$, $h < 1$, where h is *close* to 1.

4.1.2 Proposed Exterior Search Approach for SA

We seek for a minimizer of $\Pi(\mathbf{x})$ over $\tilde{\Omega}$, where the cardinality of the feasible configurations space depends on the value of the constraint: $|\tilde{\Omega}| = f(D_{\max})$. In particular, f is an increasing function of D_{\max} , because the number of feasible configurations increases as the constraint loosens.

When D_{\max} is small, we can reasonably expect to have $|\tilde{\Omega}| \ll |\Omega|$. This may reduce the effectiveness of the SA, as the energy surface that the algorithm explores could be fragmented into many isolated regions, some of which are unreachable for the algorithm. Moreover, we can expect the optimal solution to tightly respect the constraint, i.e., to lie close to the border of $\tilde{\Omega}$ [36]. Thus, the standard SA may not appropriately cover the region where the solution is, as many of its configurations are non feasible.

This problem can be solved by means of an *exterior search* approach [37], [38]. It consists in extending the search to configurations outside $\tilde{\Omega}$, while adding a *penalty* to the energy of those configurations that violate the constraint. This method presents analogies with the Lagrangian relaxation method (see the classic works [39], [40]), and it is regarded as a powerful instrument to deal with constrained optimization problems (see, e.g., [41], [42]). This idea is illustrated in Fig. 4, where we see on a fictitious example how exterior search can reduce the path to the optimal configuration.

⁴A joint uplink and downlink RN placement optimization is left for further work.

⁵This assumption is consistent with real deployment scenarios, where normally only a certain number of locations inside the cell are available for RN installation [30].

⁶Not to be confused with the energy in J involved in the energy consumption (11).

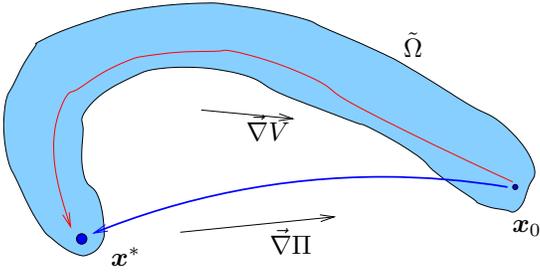


Figure 4: Exterior search principle: the search from the initial configuration x_0 to the optimal configuration x^* is extended to configurations outside $\tilde{\Omega}$ (in red: interior penalty, in blue: exterior penalty).

We thus adopt an exterior search approach, and reformulate (13) as:

$$\min_{\mathbf{x}} \mathcal{F}(\mathbf{x}), \mathbf{x} \in \Omega, \quad (17)$$

$$\text{where } \mathcal{F}(\mathbf{x}) = \Pi(\mathbf{x}) + \mathbb{1}_{D_c(\mathbf{x}) > D_{\max}} V(\mathbf{x}, D_{\max}),$$

and where $V(\mathbf{x})$ is a function $\Omega \mapsto \mathbb{R}$ accounting for constraint (14) and named *penalty function*, which plays a crucial role in the optimization [38].

Penalty functions are classified into *static*, *dynamic*, and *adaptive*. A static penalty function does not change during the optimization. It can be represented, e.g., by a fixed constant added to the energy of those configurations $\mathbf{x} \notin \tilde{\Omega}$, or by a function proportional to the Euclidean distance of the considered configuration to the feasible region [43]. On the contrary, dynamic penalty functions can be adjusted in accordance to the progress of the optimization. A common approach is to gradually increase the penalty with the number of explored configurations [44], so as to guarantee the convergence of the optimization towards a feasible solution. Finally, an adaptive penalty function (see, e.g., [42], [45]) considers further aspects of the search, such as steering the algorithm towards regions of the energy surface which are deemed promising for the search.

We propose the following exterior penalty function, which is dynamic and adaptive at the same time:

$$V(\mathbf{x}, D_{\max}) = \alpha(m-1) \frac{\Pi(\mathbf{x})}{D_{\max}} (D_c(\mathbf{x}) - D_{\max}), \quad (18)$$

where α is a constant and m is the SA step. The benefits of using function (18) are manifold. First, the penalty is proportional to the violation of constraint D_{\max} , favoring the exploration of configurations which are out of $\tilde{\Omega}$ but *close* to it, while penalizing more those which are *far*. At the same time the penalty is independent of the adopted value of D_{\max} , depending rather on the percentage of exceeding delay, with respect to D_{\max} . This represents an important feature when dealing with constraints of a different order of magnitude, compared to $\Pi(\mathbf{x})$. Moreover, the penalty is proportional to $\Pi(\mathbf{x})$, so as to favor the exploration of configurations outside $\tilde{\Omega}$ when they appear to be *promising*, i.e., their energy is sensibly lower to that of the current solution (hence the adaptive nature of (18)). Finally, acting on α we can modify the percentage of acceptance for configurations outside $\tilde{\Omega}$. We use here the works of Geman and Robini [46], [47]: Choosing $\alpha(m-1) > \log(m)$ is sufficient to guarantee

convergence of the SA algorithm towards a feasible minimizer. App. E explains why the convergence property of the SA is maintained with such a penalty function.

Surprisingly, stochastic constrained optimization based on exterior penalties has been seldom employed in wireless networks optimization. There are however three recent interesting references using this method [48], [49], [50]. In this context, two important problems arise:

- 1) The generic choice of the penalty term itself. To our knowledge, the most interesting analysis appears in [50], using the so-called logarithmic barrier. Our approach is somewhat similar: since our penalty term is the product of the objective function $\Pi(\mathbf{x})$ and the constraint $D_c(\mathbf{x}) - D_{\max}$, its gradient is a linear combination of the gradients of both previous expressions. Thus, in a deterministic, continuous framework, during a phase when the penalty is active ($D_c(\mathbf{x}) > D_{\max}$) it adapts the minimizing direction search to both the objective function and the constraint opposite gradients (see Fig. 4). We expect the stochastic, discrete SA to behave in a somewhat similar way.
- 2) The selection of the Lagrangian multipliers and their evolution during the algorithm. Apart from the usual choice of constant parameters, an interesting development can be found in [48]: at each step of the SA algorithm, each penalty weight itself either increases if the associated constraint is violated, or decreases (in a geometric way) if this constraint has been satisfied during several previous steps. In [49] all penalty weights increase regularly.

To the best of our knowledge, our approach is the first one to use stochastic constrained optimization for station placement in wireless networks. The originality of our work also lies in the multiplicative penalty function and the theoretical choice of the penalty coefficient $\alpha(m-1)$.

4.1.3 Initial Temperature Choice

Following [51], we address the problem of finding a good value for the initial temperature T_0 in each considered optimization problem via dichotomic search during a series of preliminary runs of the algorithm.

5 RESULTS

5.1 Considered Scenario and SA Parameters

The proposed theoretical framework is here applied on a test scenario drawn according to [12, case 1] for what concerns propagation, shadowing and stations transmit power. System bandwidth is 10 MHz. The network is composed of one central eNB surrounded by 6 eNBs regularly distributed around it on a circle of ray 500 m. Moreover, each eNB backhauls n_{RN} RNs. We optimize RN placements in the central cell only, while RNs in the surrounding cells are assumed to be regularly distributed on a circle of radius 160 m around their donor eNB. All stations have omnidirectional antennas. The same realization of shadowing (drawn according to [12]) is used for all simulations, so as to ensure that results be comparable. If not otherwise mentioned, we adopt a *uniform* traffic density spatial distribution, i.e., $\phi(s)$

is constant $\forall s \in \mathcal{A}_{nr}$, in order to draw general conclusions about energy efficiency. In Section 5.2.5 we show the performance under a *non-uniform* traffic distribution. We set $\beta=0.1$.

Function C is approximated by means of the Modulation and Coding Scheme (MCS) indicated in [52], so as to take into account the effect of an upper-bounded capacity function, which is the typical case in real deployments. Also, fast fading on the backhaul link is not considered. This choice is justified by the assumption that RNs do not move, and optimization of their positions is performed on a long-term basis (see, e.g. [53]). We have assumed $\mathcal{B}_{\min}=0$ and $\mathcal{B}_{\max}=15$ dB, with a step of 1 dB. The grid of candidate RN sites has a step of 50 m.

The fixed point of $\mathbf{F}(\boldsymbol{\rho})$ is found by iteratively computing stations loads. At iteration t , probabilities $\pi_{\gamma,k}(z, s)$ are computed according to $\boldsymbol{\rho}(t-1)$, obtained at iteration $t-1$, and then fed into (4) to get $\boldsymbol{\rho}(t)=\mathbf{F}(\boldsymbol{\rho}(t-1))$, starting from $\boldsymbol{\rho}(0) = \mathbf{0}$. Iterations are stopped when either $|\rho_k(t) - \rho_k(t-1)| < 0.01, \forall k \in \mathcal{K}$, or $\exists k : \rho_k(t) \geq 1, k \in \mathcal{K}$. In the latter case, the analyzed configuration is labeled as non-valid. Similarly, fixed point iterations are used to compute the loads on the backhaul link, after $\boldsymbol{\rho}$ is determined. We were not able to analytically prove the convergence of the fixed point iteration (by showing for example that \mathbf{F} is a contraction mapping). We have however numerically observed that the iteration always converges in less than 10 iterations.

SA algorithm is implemented as described in Section 4. Once again, we emphasize the need to optimize network parameters and in particular RN locations in order to obtain upper bounds on the achievable performance. The algorithm yields a solution after 45 temperature steps. For each tested temperature value, the energy of a certain number of network configurations is calculated. This number varies according to n_{RN} , e.g., for $n_{RN} = 2$, 400 network configurations are tested at each temperature step. For every optimization, the SA is repeated 4 times, and the best solution among the 4 obtained solutions is elected as a final result. The configuration corresponding to the final result is denoted with \mathbf{x}^* .

5.2 Numerical Results

We denote with \mathbf{x}_0^* the configuration with minimum energy when no RNs are deployed in the whole network, and with $\Pi_0(\mathbf{x}_0^*)$ its corresponding energy. Unless otherwise mentioned, results are expressed in terms of the energy consumption ratio $\Pi(\mathbf{x}^*)/\Pi_0(\mathbf{x}_0^*)$, so as to show the energy gain (or loss) resulting from RN deployment, and they are plotted against the normalized constraint D_{\max}/D_0 , where D_0 is the value of \bar{D}_c corresponding to \mathbf{x}_0^* , i.e., the average delay when no RNs are deployed. Hence, a point in the region $\Pi(\mathbf{x}^*)/\Pi_0(\mathbf{x}_0^*) < 1$ corresponds to an energy consumption per bit gain with respect to the network with no RNs, whereas a point in the region $D_{\max}/D_0 < 1$ corresponds to an average uplink transmission delay gain. All the curves that we obtained exhibit a hyperbolic-like shape, which follows from the nature of our constrained optimization. When D_{\max} is large, $|\tilde{\Omega}| \approx |\Omega|$, the delay doesn't play any role in the optimization and so the energy consumption gain

reaches its maximum. Instead, if the constraint is tight we have $|\tilde{\Omega}| \ll |\Omega|$, and it is unlikely that any configuration that performs *well* in terms of energy efficiency lies in $\tilde{\Omega}$.

5.2.1 Energy - Delay Trade-off

Fig. 5 shows the trade-off between energy consumption and delay, for a varying number of RNs. Note how RN deployment can bring consistent UEs energy savings: All curve points below the line $y = 1$ correspond to an energy consumption reduction. Now, adding RNs in a cell has two opposite effects: On the one hand, the average distance between UEs and their serving station is reduced, they can thus reduce their transmit power; On the other hand, backhaul delay is increased because of the increased relayed traffic. If the delay constraint is loose (in region where $D_{\max}/D_0 > 1$), increasing the number of RNs reduces energy consumption because the latter effect is not really an issue with respect to the constraint. If the delay constraint is tight (when $D_{\max}/D_0 < 0.5$), backhaul delay is crucial and UEs attached to a RNs will tend to boost their transmit power in order to reduce the access delay (\bar{P} is increased in the optimal solution). This results in an overall increase of the energy consumption. From an optimization point of view, the cardinality $|\tilde{\Omega}|$ of the valid configurations set tends to decrease with D_{\max}/D_0 , narrowing the space for network optimization and reducing the achievable energy consumption gains.

5.2.2 Effect of Offered Traffic

We analyze the effect of $\bar{\omega}$ on energy consumption in Fig. 6, where $\Pi(\mathbf{x}^*)$ and D_{\max} are scaled with the constants $\Pi_0(\mathbf{x}_0^*(\bar{\omega} = 5))$ and $D_0(\bar{\omega} = 5)$ resp., obtained in a network with no RN and a traffic density $\bar{\omega} = 5$ bits/s/m². This choice has been made in order to keep the same scaling constants for all curves. As we can see, the deployment of 1 RN is sufficient to reduce the average energy consumption of UEs to less than a half, compared to the eNBs-only network. Now, let's focus on the line $D_{\max}/D_0(\bar{\omega} = 5) = 1$ (same delay as in the eNB-only case). We note that adding a single RN can help the terminals to be more energy efficient without touching the quality of service and even if the load increases to $\bar{\omega} = 20$ bits/s/m². Nonetheless, beyond approximately $\bar{\omega} = 30$ bits/s/m², delay and interference negatively affect user performance. From this figure, we also observe that energy consumption is increasing with the traffic density for any delay constraint. This effect has been also observed for 2 and 3 relays in our simulations. This can be explained by the fact that when load increases, delay increases as well. To compensate for that, UEs have to increase their data rate and so their transmit power (\bar{P} is increased in the optimal solution). This results in an increase of the energy consumption.

5.2.3 Energy Efficiency of RNs vs Small Cells

Fig. 7 compares the results obtained with our system model, with those obtained using small cells instead of RNs (i.e., $\beta = 0, D_{BL,j,k} = 0, \forall j \in \mathcal{K}_B, \forall k \in \mathcal{K}_R$), highlighting the difference of our work with respect to those dedicated to devices with wired (or ideal) backhaul. As expected, RNs allow for a smaller energy consumption gain compared to

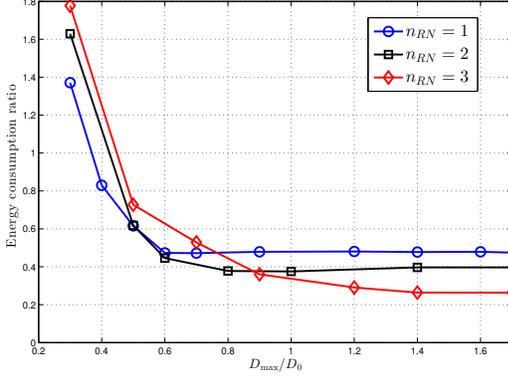


Figure 5: Energy consumption ratio $\Pi(\mathbf{x}^*)/\Pi_0(\mathbf{x}_0^*)$ vs delay D_{\max}/D_0 ($\bar{\omega} = 5$ bits/s/m²).

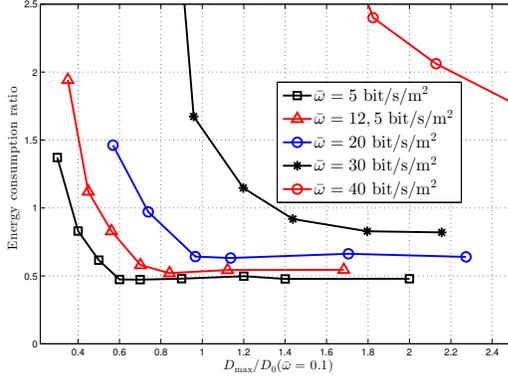


Figure 6: Energy consumption ratio $\Pi(\mathbf{x}^*)/\Pi_0(\mathbf{x}_0^*)$ ($\bar{\omega} = 5$) vs delay D_{\max}/D_0 ($\bar{\omega} = 5$) for varying $\bar{\omega}$ ($n_{RN} = 1$).

small cells, for a fixed D_{\max} . This is due to the transmission delay on the backhaul link, to the increased delay on the access link (due to $\beta < 1$) and to the constraints on RN placement related to backhaul path-loss and shadowing. Performance of RNs is more impaired when the traffic density or n_{RN} increase. However, RN deployment still yields consistent uplink energy consumption gains.

5.2.4 Effectiveness of Exterior Search with Penalty Function

Fig. 8 compares the results of the optimization using penalty function and exterior search, with those obtained by means of an interior search. Both the interior and the exterior search have been carried on with the same number of iterations, for an unbiased comparison. The exterior search proves to be more effective when the delay constraint is tight, as $\bar{\Omega}$ is expected to be fragmented and $|\bar{\Omega}| \ll |\Omega|$. No meaningful gain in terms of search effectiveness is observed when the constraint is loose, as in this case the two approaches tend to coincide.

5.2.5 Effect of Offered Traffic Spatial Distribution

Fig. 9 compares the results obtained using a uniform $\phi(s)$ with those obtained when $\phi(s)$ is the sum of a uniform traffic profile and a bi-dimensional gaussian function, centered at $s : \{x = 115; y = 143\}$ m, with the same average traffic density. We can see that RNs allow for larger energy

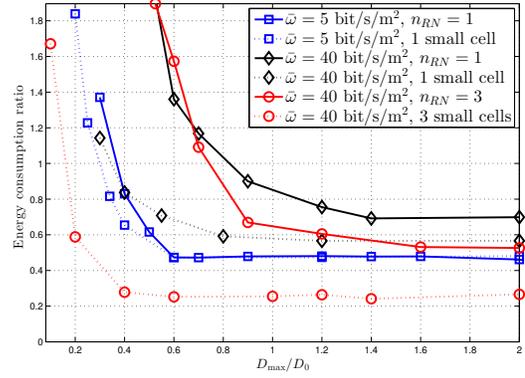


Figure 7: Comparison between wireless RNs and wired small cells, for varying $\bar{\omega}$ and n_{RN} ($\beta = 0.1$).

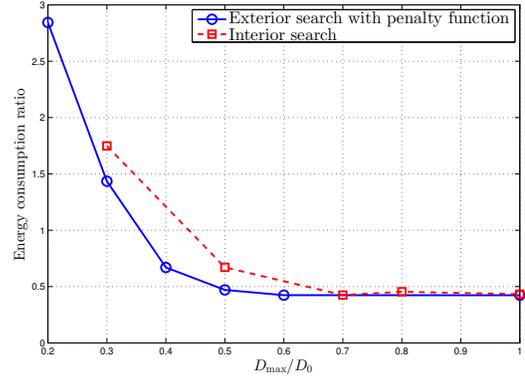


Figure 8: Comparison between exterior and interior search approach effectiveness ($\bar{\omega} = 5$ bits/s/m², $\beta = 0.1$, $n_{RN} = 1$).

consumption gains when traffic is non-uniform. This is due to the flexibility of the RN solution that allows relays to be placed close to the hot-spot center.

5.2.6 Effect of Optimal Relay Placement

Fig. 10 shows the effect of optimizing the relay placement. Optimal placement is compared to a scenario, where relays are regularly placed at a distance of 160 m around the cell center. The figure shows the great advantage of optimal placement. It also highlights the fact that erroneous conclusions can be drawn concerning the energy consumption gain if relays are arbitrarily placed in the cell.

6 CONCLUSION

We have proposed a comprehensive framework for the optimization and performance evaluation of uplink energy consumption per bit in relays-enhanced cellular networks. This framework considers the arrival and departure of users, and the loads of network stations. Moreover, shadowing and fast fading are both taken into account in the propagation model. A realistic radio resource scheduling scheme is assumed, and its impact on users performance is derived. A customized optimization algorithm based on exterior search with penalty functions is proposed for the optimization, which is carried on under quality of service

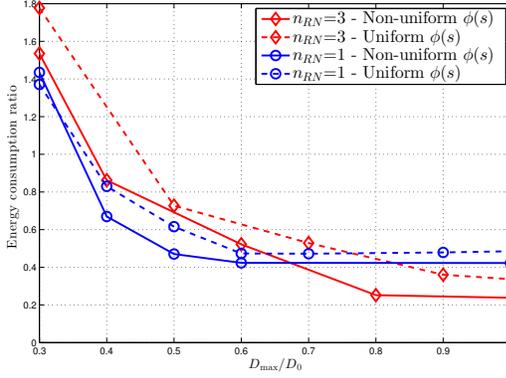


Figure 9: Energy consumption vs delay tradeoff: $\bar{\omega} = 0.1$, $\beta = 0.1$, comparison of uniform and non-uniform traffic profile $\phi(s)$ for varying n_{RN} .

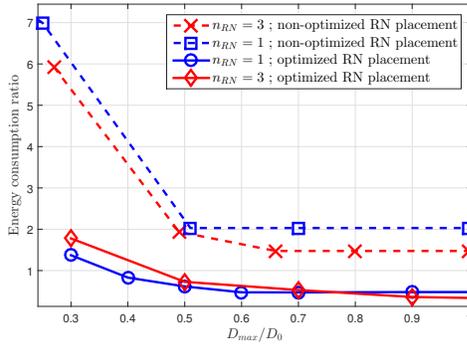


Figure 10: Energy consumption vs delay tradeoff: $\bar{\omega} = 5$, $\beta = 0.1$, comparison of optimized and non-optimized relay placement for varying n_{RN} .

constraints. Results show a meaningful boosting of users energy efficiency given by the deployment of RNs, even if the average flow delay is imposed to be the same as in the network without relays. Proposed exterior search approach is shown to be more effective than the traditional interior search. Further work would include the impact of mobility on the energy-delay tradeoff or the study of cooperative relays.

APPENDIX A SOME PROOFS

A.1 Proof of Lemma 1

The contribution to the load is expressed as the ratio of the traffic density $\bar{\omega}\phi(s)$ in s to the uplink rate achieved by a scheduled UE in s . The term $(1 - \beta)$ takes into account the quota of the frame RBs dedicated to the backhaul link.

A.2 Proof of Theorem 2

The load ρ_k is the sum of the contributions $\varrho_k(s)$ over \mathcal{A}_k : $\rho_k = \int_{\mathcal{A}_k} \varrho_k(s) ds$. Now, the term $\mathbb{E}_\tau[1/C(\gamma_k(s, \tau)) | f_k(\tau) = s]$ in (3) can be expressed according the law of total probability, obtaining (4).

A.3 Proof of Lemma 2

For all stations $k \in \mathcal{K}$, the corresponding $F_k(\boldsymbol{\rho}_{-k})$ is increasing with respect to every coordinate, as an increase in the load of any station $j \neq k, j \in \mathcal{K}$ produces an increase in the interference received at k . We define $\forall k, \rho_k^{\max} = F_k(1 \cdots 1)$, which represents an upper bound for the load of k . Then, $\forall k, \forall \rho_k, F_k(\boldsymbol{\rho}_{-k}) = F_k(\max\{\mathbf{1}, \boldsymbol{\rho}_{-k}\}) \leq F_k(1 \cdots 1) = \rho_k^{\max}$. Hence,

$$\mathbf{F}(\boldsymbol{\rho}) \in \prod_{k \in \mathcal{K}} [0, \rho_k^{\max}], \rho_k^{\max} \in \mathbb{R}^+, \forall k \in \mathcal{K}. \quad (19)$$

A.4 Proof of Theorem 3

This follows from Lemma 2 and (19): if \mathbf{F} is a continuous mapping from an K -dimensional closed interval to itself, then the Brouwer's fixed point theorem guarantees the existence of at least one fixed point.

A.5 Proof of Lemma 3

The delay to transmit a bit of information is the inverse of the UE rate, multiplied by the number x_k of UEs served by the same station during a frame (because each UE is scheduled on a fraction $1/x_k$ of the RBs with MQS). The average transmission delay for the whole flow is hence given by: $D_k(s) = \xi \mathbb{E}_\tau \left[\frac{x_k}{(1-\beta)C(\gamma_k(s, \tau))} | f_k(\tau) = s \right]$. Considering that $\mathbb{P}(x_k = U | x_k > 0) = \rho_k^{U-1} (1 - \rho_k)$, the law of total probability can be used to average $D_k(s)$ over $x_k > 0$: $D_k(s) = \sum_{U=1}^{\infty} \mathbb{P}(x_k = U) \frac{\xi}{1-\beta} U \mathbb{E}_\tau \left[\frac{1}{C(\gamma_k(s, \tau))} | f_k(\tau) = s \right]$, obtaining expression (6).

A.6 Proof of Corollary 1

The statement can be verified by means of the Little's law: $D_k = \frac{\mathbb{E}_\tau[x_k(\tau)]}{\bar{\omega}\Phi_k} = \frac{\rho_k}{1-\rho_k} \frac{1}{\bar{\omega}\Phi_k}$. Now, we can substitute (4) to the numerator and conclude the proof.

A.7 Proof of Lemma 4

Similarly to what performed for Lemma 3, we proceed by expressing $D_{BL,j,k}$ as: $D_{BL,j,k} = \frac{\xi}{\beta R_{BL,j,k}} \mathbb{E}_\tau[x_{BL,j}(\tau) | x_{BL,j}(\tau) > 0]$ and apply the law of total probability, obtaining:

$$D_{BL,j,k} = \frac{\xi}{\beta R_{BL,j,k}} \sum_{v=1}^{\infty} P(x_{BL,j}(\tau) = v | x_{BL,j}(\tau) > 0) \times \mathbb{E}_\tau[x_{BL,j}(\tau) | x_{BL,j}(\tau) = v],$$

where $x_{BL,j}$ is the number of flows in the backhaul queue of j during the considered frame. Conventionally, we set $D_{BL,j,j} = 0$. From this equation, we obtain (8).

APPENDIX B

DERIVATION OF $\mu_{\gamma,k}$ AND $\sigma_{\gamma,k}$

We use here a simple and classical approach consisting in introducing a auxiliary Random Variable (RV). (1) can indeed be rewritten as: $\gamma_k(s) = \frac{P_k(s)x}{\sum_{j \in \mathcal{K}, j \neq k} \Theta_j P_k(f_j)x + N_x} = P_k(s)x R_k$, where x is a realization of a lognormal RV X , and R_k^{-1} is the denominator. Note that R_k does not depend

on s . Several authors have shown that the interference term $I_k = \sum_{j \neq k} \Theta_j P_k(f_j)$ can be well approximated by a lognormal distribution $\text{Log-}\mathcal{N}(\mu_{I_k}, \sigma_{I_k}^2)$ even when the number of interferers is variable (see e.g. [22], [23]). In [22], it is also shown that the product of an exponential RV ($P_k(s)$) and a lognormal RV (X) as well as the sum of two lognormal RVs can be again approximated by a lognormal RV. Parameters $(\mu_{\gamma,k}, \sigma_{\gamma,k})$ are then obtained by moment matching. The derivation of $(\mu_{I_k}, \sigma_{I_k})$, which is less standard, is now detailed. The mean M_{1,I_k} of $I_k = \sum_{j \neq k} \Theta_j P_k(f_j)$ is equal to:

$$\begin{aligned} M_{1,I_k} &= \mathbb{E}_\tau \left[\sum_{j \neq k} \Theta_j P_k(f_j) \right] \\ &= \sum_{j \neq k} \rho_j \mathbb{E}_\tau [P_k(f_j)] \\ &\stackrel{(1)}{=} \sum_{j \neq k} \rho_j \int_{\mathcal{A}_j} \frac{\varrho_j(s)}{\rho_j} T(s) G_k(s) ds \\ &= \sum_{j \neq k} \rho_j Y_{j,k}, \end{aligned}$$

where $Y_{j,k} = \int_{\mathcal{A}_j} \frac{\varrho_j(s)}{\rho_j} T(s) G_k(s) ds$. (1) is obtained by weighting the power received from location s with the local load in s . Note that $T(s) G_k(s)$ are input parameters of the considered deployment scenario (propagation and transmit power assumptions). Now $\varrho_j(s)$ depends on the SINR distribution in s and hence on the first moment M_{1,I_k} of I_k . In order to avoid an additional complexity to our model, we make the approximation $\varrho_j(s)/\rho_j \approx \phi(s)/\Phi_j$. This is justified by the fact that in a urban environment, it is unlikely that UEs transmit at their maximum power [14] so that every user of j is received with the same average power. The approximation comes then from the expressions (3), (4) of $\varrho_j(s)$ and ρ_j . In the same way, we derive the variance M_{2,I_k} of I_k : $M_{2,I_k} = \sum_{j \neq k} 2\rho_j H_{j,k} - \rho_j^2 Y_{j,k}^2$, where $H_{j,k} = \int_{\mathcal{A}_j} \frac{\varrho_j(s)}{\rho_j} T^2(s) G_k^2(s) ds$. Again, we approximate $\varrho(s)/\rho_j \approx \phi(s)/\Phi_j$. Finally, μ_{I_k} and σ_{I_k} are found by matching M_{1,I_k} and M_{2,I_k} with the mean and variance of the approximating lognormal.

APPENDIX C

PROOF OF THEOREM 1: DERIVATION OF $\pi_{\gamma,k}(s, z)$ UNDER MQS SCHEDULING

The MQS scheduler orders the values of SINR $\gamma_k(s, \tau - j)$, $j \in \{0 \dots W - 1\}$ of each UE in ascending order. The ranking $r_i(\tau)$ of UE i located in s_i on RB τ is the ranking of $\gamma_k(s_i, \tau)$ in the ordered vector $\gamma_k(s_i, \tau - j)$, $j \in \{0 \dots W - 1\}$. The lower the ranking, the better the SINR on τ (wrt the SINR on previous blocks). Station k assigns RB τ to the UE with the lowest $r_i(\tau)$, $i \in \{1 \dots U\}$. We assume W large enough so that the probability for two or more UEs to have the same ranking is negligible. We have: $\mathbb{P}(r_i(\tau)=n)=1/W, \forall n \in \{1 \dots W\}$ ⁷. Moreover, the probability for any UE served by k to be scheduled is $\mathbb{P}(f_k(\tau)=s_i) = 1/U, \forall i \in \{1 \dots U\}$ (due to the fairness in

RBs allocation). Let $\Pi_{\gamma,k}(s_i, z) = \int_0^z p_{\gamma,k}(s_i, t) dt$ denote the Cumulative Distribution Function (CDF) of $\gamma_k(s_i, \tau)$. We have:

$$\begin{aligned} \mathbb{P}(r_i(\tau)=n \mid \gamma_k(s_i, \tau)=z) &= \\ &= \binom{W-1}{n-1} \Pi_{\gamma,k}(s_i, z)^{W-n} (1 - \Pi_{\gamma,k}(s_i, z))^{n-1}. \end{aligned} \quad (20)$$

We denote $Q_n(i, z)$ the right hand side, which does not depend on U . Now, we derive $\pi_{\gamma,k}(s, z)$, given U . Applying the Bayes theorem on $\pi_{\gamma,k}(s, z)$ we obtain:

$$\begin{aligned} \pi_{\gamma,k}(s, z) &= \frac{\mathbb{P}(f_k(\tau)=s_i \mid \gamma_k(s_i, \tau)=z) p_{\gamma,k}(s_i, z)}{\mathbb{P}(f_k(\tau)=s_i)} \\ &= \mathbb{P}(f_k(\tau)=s_i \mid \gamma_k(s_i, \tau)=z) p_{\gamma,k}(s_i, z) U. \end{aligned} \quad (21)$$

We then work on $\mathbb{P}(f_k(\tau)=s_i \mid \gamma_k(s_i, \tau)=z)$ applying the law of total probability, conditioning it with respect to the possible rankings $r_i(\tau)$:

$$\begin{aligned} \mathbb{P}(f_k(\tau)=s_i \mid \gamma_k(s_i, \tau)=z) &= \\ &= \sum_{n=1}^W \mathbb{P}(r_i(\tau)=n \mid \gamma_k(s_i, \tau)=z) \times \\ &= \mathbb{P}(f_k(\tau)=s_i \mid \gamma_k(s_i, \tau)=z, r_i(\tau)=n). \end{aligned} \quad (22)$$

The probability of being scheduled depends on the instantaneous SINR only through the ranking: once we know $r_i(\tau)$, knowing $\gamma_k(s_i, \tau)$ does not add any additional information regarding the probability of being scheduled by k . Hence,

$$\begin{aligned} \mathbb{P}(f_k(\tau)=s_i \mid \gamma_k(s_i, \tau)=z, r_i(\tau)=n) &= \\ \mathbb{P}(f_k(\tau)=s_i \mid r_i(\tau)=n) &= \left(\frac{W-n}{W} \right)^{U-1}, \end{aligned} \quad (23)$$

and (21) becomes

$$\pi_{\gamma,k}(s, z) = p_{\gamma,k}(s_i, z) U \sum_{n=1}^W Q_n(i, z) \left(\frac{W-n}{W} \right)^{U-1}. \quad (24)$$

So far, we have found $\pi_{\gamma,k}(s, z)$ given U . We now define $\Psi = (W-n)/W$, and use the law of total probability, summing up $\pi_{\gamma,k}(s, z)$ given U for all the possible values of U , and obtaining its general expression:

$$\begin{aligned} \mathbb{P}(\gamma_k(s_i, \tau) = z \mid f_k(\tau)=s_i, U > 0) &= \\ p_{\gamma,k}(s_i, z) \sum_{n=1}^W Q_n(i, z) (1 - \rho_k) \sum_{U=1}^{\infty} \rho_k^{U-1} U \Psi^{U-1}, \end{aligned} \quad (25)$$

where

$$(1 - \rho_k) \sum_{U=1}^{\infty} \rho_k^{U-1} U \Psi^{U-1} = \frac{W^2(1 - \rho_k)}{(W - \rho_k(W - n))^2}. \quad (26)$$

We name $T_k(W, n) \triangleq W^2(1 - \rho_k)/(W - \rho_k(W - n))^2$ and get the conclusion.

⁷as in our block Rayleigh fading environment each realization of $\gamma_k(s_i, \tau)$ is independent of the others.

APPENDIX D BACKHAUL RATE DERIVATION

We start from the definition of $R_{BL,j,k}$ as $R_{BL,j,k} = 1/\mathbb{E}_\tau[\frac{1}{C(\gamma_{BL,j,k}(\tau))}]$, where $\gamma_{BL,j,k}(\tau)$ is given by:

$$\gamma_{BL,j,k}(\tau) = \frac{P_{BL,j}(s_{RN,j,k})}{\sum_{h \in \mathcal{K}_B, h \neq j} \Theta_{BL,h}(\tau) P_{BL,j}(f_{BL,h}(\tau)) + N}, \quad (27)$$

where $\Theta_{BL,h}$ is a binary RV indicating whether the backhaul of h is active on τ , i.e., $P(\Theta_{BL,h} = 1) = \rho_{BL,h}$; $P_{BL,j}(s)$ is the power received by eNB j from a RN located in s ; $s_{RN,j,k}$ indicates the location of RN k served by eNB j and $f_{BL,h}(\tau)$ yields the location of the RN scheduled by h for backhaul transmission on τ .

During each RB τ a RN k in cell j can be scheduled for uplink transmission with probability $p_s(j, k) = \frac{\tilde{\rho}_{BL,j,k}}{\sum_t \tilde{\rho}_{BL,j,t}}$, where $\tilde{\rho}_{BL,j,t} = (\bar{\omega}\Phi_t/(\beta R_{BL,j,t}))$. Let $T = |\mathcal{K}_B|$ denote the total number of cells in our network, and $i_j(\tau), j \in \mathcal{K}_B, i_j(\tau) \in \{0 \dots n_{RN}\}$ denote the index of the RN scheduled on the backhaul of eNB j on RB τ , where $i_j(\tau) = 0$ means that no RN has been scheduled for transmission. Now, assuming that backhaul scheduling decisions in a given cell do not depend on those taken in other cells, we have that the probability of having a given set $\{i_1(\tau), \dots, i_T(\tau)\}$ of scheduled RNs on RB τ is equal to

$$\mathbb{P}(f_{BL,1}(\tau) = s_{RN,1,i_1}, \dots, f_{BL,T}(\tau) = s_{RN,T,i_T}) = \prod_j p_s(j, i_j(\tau)). \quad (28)$$

There are $(n_{RN} + 1)^T$ possible scheduling combinations in all cells. We assign an index $\delta, \delta = \{1 \dots (n_{RN} + 1)^T\}$ to each combination, and denote $\mathcal{V}(\delta) = \prod_j p_s(j, i_j(\delta))$. Finally, we apply the law of total probability to obtain

$$R_{BL,j,k} = \frac{1}{\sum_{\delta=1}^{(n_{RN}+1)^T} \mathcal{V}(\delta) \frac{1}{C\left(\frac{P_{BL,j}(s_{RN,j,k})}{\sum_{h \in \mathcal{K}_B, h \neq j} P_{BL,j}(s_{RN,h,i_h(\delta)}) + N}\right)}}, \quad (29)$$

where we assume $P_{BL,j}(s_{RN,h,0}) = 0$.

APPENDIX E SOME PROPERTIES OF GIBBS DISTRIBUTIONS FOR PENALIZED ENERGIES

The purpose of this Appendix is to give a hint to the "good" convergence of a SA process when an increasing exterior penalty such as (18) is added in the global energy, coupled with an adequate SA temperature scheme. Consider such an augmented energy:

$$\mathcal{F}(\mathbf{x}) = U(\mathbf{x}) + \mu \Phi(\mathbf{x}), \quad (30)$$

$$\Phi(\mathbf{x}) = \mathbb{1}_{D_c(\mathbf{x}) > D_{\max}} \frac{\Pi(\mathbf{x})}{D_{\max}} (D_c(\mathbf{x}) - D_{\max}) \quad (31)$$

with for instance $\mu = \mu(m) = \alpha(m-1) \geq 0$. We say that Φ is a **penalty** if the set of its minimizers

$$\Omega^* = \{x \in \Omega \text{ s.t. } \Phi(\mathbf{x}) = \Phi^* = \min_{\mathbf{y} \in \Omega} \Phi(\mathbf{y})\} \quad (32)$$

is exactly the feasible subspace $\tilde{\Omega}$ (this holds for (18)). Now let us in general define the following set of "iso-constrained"

subspaces: $\epsilon \in \mathbb{R} \mapsto \Omega_\epsilon = \{\mathbf{x} \in \Omega \text{ s.t. } \Phi(\mathbf{x}) = \epsilon\}$, and consider the exponential distribution given in (16) (see also [46]). For any value of ϵ such that $\Omega_\epsilon \neq \emptyset$, one has:

$$\begin{aligned} \forall \mathbf{x}, \mathbf{y} \in \Omega_\epsilon, \frac{P(X = \mathbf{y})}{P(X = \mathbf{x})} &= \frac{\exp -\mathcal{F}(\mathbf{y})}{\exp -\mathcal{F}(\mathbf{x})} \\ &= \frac{\exp - [U(\mathbf{y}) + \mu \Phi(\mathbf{y})]}{\exp - [U(\mathbf{x}) + \mu \Phi(\mathbf{x})]} \\ &= \frac{\exp -U(\mathbf{y})}{\exp -U(\mathbf{x})}. \end{aligned} \quad (33)$$

So one can safely write in such an iso-constrained subspace:

$$P(X = \mathbf{x} \mid \mathbf{x} \in \Omega_\epsilon) = \frac{\exp -U(\mathbf{x})}{\sum_{\mathbf{y} \in \Omega_\epsilon} \exp -U(\mathbf{y})}. \quad (34)$$

which is also an exponential distribution.

Two key points are now at stake in view of SA purposes [46], [47]:

- First, as $\mu \rightarrow +\infty$ (μ is similar to an inverse temperature associated to the constraint $\Phi(\mathbf{x})$), the global distribution (16) becomes concentrated on the subspace Ω^* defined in (32) while keeping the exponential form (34). The proof is classical: one can re-write (16) as

$$\begin{aligned} P(X = \mathbf{x}) &= \frac{\exp - [U(\mathbf{x}) + \mu \Phi(\mathbf{x})]}{\sum_{\mathbf{y} \in \Omega} \exp - [U(\mathbf{y}) + \mu \Phi(\mathbf{y})]} \\ &= \frac{\exp - [U(\mathbf{x}) + \mu (\Phi(\mathbf{x}) - \Phi^*)]}{\sum_{\mathbf{y} \in \Omega} \exp - [U(\mathbf{y}) + \mu (\Phi(\mathbf{y}) - \Phi^*)]} \end{aligned}$$

Now, either $\Phi(\mathbf{x}) > \Phi^*$ and $\exp(-\mu (\Phi(\mathbf{x}) - \Phi^*)) \rightarrow 0$ when $\mu \rightarrow +\infty$, or $\Phi(\mathbf{x}) = \Phi^*$ and in this case, $\exp(-\mu (\Phi(\mathbf{x}) - \Phi^*)) = 1 \forall \mu$.

- Then, consider distribution (16) with a temperature parameter $T > 0$. It can be written as:

$$\begin{aligned} P_T(X = \mathbf{x}) &= \frac{\exp - \left[\frac{U(\mathbf{x}) + \mu \Phi(\mathbf{x})}{T} \right]}{\sum_{\mathbf{y} \in \Omega} \exp - \left[\frac{U(\mathbf{y}) + \mu \Phi(\mathbf{y})}{T} \right]} \\ &= \frac{\exp - \left[\frac{(U(\mathbf{x}) - U^*)}{T} + \frac{\mu}{T} (\Phi(\mathbf{x}) - \Phi^*) \right]}{\sum_{\mathbf{y} \in \Omega} \exp - \left[\frac{(U(\mathbf{y}) - U^*)}{T} + \frac{\mu}{T} (\Phi(\mathbf{y}) - \Phi^*) \right]}, \end{aligned}$$

where U^* is the minimum of value of objective function $U(\cdot)$ on Ω^* . Now, if both $T \rightarrow 0^+$ and $\mu \rightarrow +\infty$ s.t. $\mu' = \frac{\mu}{T} \sim \log m$ (new inverse temperature), a similar analysis as before, now in two steps shows that for the penalty case where $\Omega^* = \tilde{\Omega}$, the distribution of interest $P_T(\cdot)$ concentrates on those configurations with minimal energy U^* on feasible subspace $\tilde{\Omega}$ (see rigorous proof in [46], [47]).

REFERENCES

- [1] R. Pabst, B. H. Walke, D. Schultz, P. Herhold, H. Yanikomeroğlu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. Falconer, and G. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Communications Magazine*, vol. 42, pp. 80–89, Sep. 2004.

- [2] Y. Hou and D. Laurenson, "Energy efficiency of high QoS heterogeneous wireless communication network," in *IEEE Vehicular Technology Conference Fall (VTC)*, pp. 1–5, Sep. 2010.
- [3] Z. Hasan, H. Boostanimehr, and V. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Communications Surveys Tutorials*, vol. 13, pp. 524–540, Nov. 2011.
- [4] A. Saleh, O. Bulakci, S. Redana, B. Raaf, and J. Hamalainen, "Evaluating the energy efficiency of LTE-Advanced relay and picocell deployments," in *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2335–2340, Apr. 2012.
- [5] J. Zhao, W. Zheng, X. Chu, X. Wen, H. Zhang, and Z. Lu, "Game theory based energy-aware uplink resource allocation in OFDMA femtocell networks," *International Journal of Distributed Sensor Networks*, Feb. 2014.
- [6] Y. Jiang, J. Zhang, X. Li, and W. Xu, "Energy-efficient resource optimization for relay-aided uplink ofdma systems," in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*, pp. 1–5, 2012.
- [7] B. Dusza, I. C., C. L., and W. C., "An accurate measurement-based power consumption model for LTE uplink transmission," in *IEEE International Conference on Computer Communications*, Apr. 2013.
- [8] F. Liu, K. Zheng, W. Xiang, and H. Zhao, "Design and performance analysis of an energy-efficient uplink carrier aggregation scheme," *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 197–207, Feb. 2014.
- [9] X. Chen, H. Zhang, T. Chen, and M. Lasanen, "Improving energy efficiency in green femtocell networks: A hierarchical reinforcement learning framework," in *IEEE International Conference on Communications (ICC)*, pp. 2241–2245, Jun. 2013.
- [10] J. Zhang, X. Ge, M. Chen, M. Jo, X. Yang, Q. Du, and J. Hu, "Uplink energy efficiency analysis for two-tier cellular access networks using kernel function," *Telecommunication Systems*, vol. 52, pp. 1305–1312, Feb. 2013.
- [11] K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1525–1536, Sep. 2011.
- [12] N. S. Networks, "Type-1 relay performance for downlink (R1-106216)." 3GPP TSG-RAN WG1, Nov. 2010.
- [13] C. Huang, R. Zhang, and S. Cui, "Throughput maximization for the gaussian relay channel with energy harvesting constraints," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 8, pp. 1469–1479, 2013.
- [14] O. Bulakci, A. Awada, A. Saleh, S. Redana, and J. Hamalainen, "Automated uplink power control optimization in LTE-Advanced relay networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, pp. 1–19, Jan. 2013.
- [15] J.-M. Liang, Y.-C. Wang, J.-J. Chen, J.-H. Liu, and Y.-C. Tseng, "Energy-efficient uplink resource allocation for IEEE 802.16j transparent-relay networks," *Computer Networks*, vol. 55, pp. 3705–3720, Nov. 2011.
- [16] K. Son, S. Nagaraj, M. Sarkar, and S. Dey, "QoS-aware dynamic cell reconfiguration for energy conservation in cellular networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2022–2027, Apr. 2013.
- [17] 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.
- [18] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 18, pp. 10–21, Jun. 2011.
- [19] M. Karray, "User's mobility effect on the performance of wireless cellular networks serving elastic traffic," *Wireless Networks*, vol. 17, no. 1, pp. 247–262, 2011.
- [20] 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.
- [21] S. Ohno and G. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 2138–2145, Sep. 2004.
- [22] Z. Kotic, I. Maric, and X. Wang, "Fundamentals of dynamic frequency hopping in cellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 2254–2266, Nov. 2001.
- [23] C. Fischione, F. Graziosi, and F. Santucci, "Approximation for a sum of on-off lognormal processes with wireless applications," *IEEE Transactions on Communications*, vol. 55, pp. 1984–1993, Oct. 2007.
- [24] T. Donald, "A score-based opportunistic scheduler for fading radio channels," *Proc. European Wireless*, Feb. 2004.
- [25] D. Park, H. Seo, H. Kwon, and B. G. Lee, "A new wireless packet scheduling algorithm based on the CDF of user transmission rates," in *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, pp. 528–532, Dec. 2003.
- [26] F. Kelly, "Charging and rate control for elastic traffic," *European Transactions on Telecommunications*, vol. 8, pp. 33–37, Jan. 1997.
- [27] L. Rong, S. E. Elayoubi, and O. B. Haddada, "Performance Evaluation of Cellular Networks Offering TV Services," *IEEE Trans. on Vehicular Technology*, vol. 60, pp. 644–655, Feb. 2011.
- [28] L. Suarez, L. Nuaymi, and J.-M. Bonnin, "An overview and classification of research approaches in green wireless networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, 2012.
- [29] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. Thompson, I. Ku, C.-X. Wang, T. A. Le, M. Nakhai, J. Zhang, and L. Hanzo, "Green radio: radio techniques to enable energy-efficient wireless networks," *Communications Magazine, IEEE*, vol. 49, pp. 46–54, June 2011.
- [30] O. Bulakci, A. Bou Saleh, J. Hamalainen, and S. Redana, "Performance analysis of relay site planning over composite fading/shadowing channels with cochannel interference," *IEEE Transactions on Vehicular Technology*, vol. 62, pp. 1692–1706, May 2013.
- [31] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, et al., "Equation of State Calculations by Fast Computing Machines," *The Journal of Chemical Physics*, vol. 21, pp. 1087–1092, Mar. 1953.
- [32] S. Kirkpatrick, C. Gelatt, and M. Vecchi, "Optimization by Simulated Annealing," *Science*, vol. 220, pp. 671–680, May 1983.
- [33] G. Keung, Q. Zhang, and B. 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.
- [34] M. Robini, T. Rastello, and I. Magnin, "Simulated Annealing, Acceleration Techniques, and Image Restoration," *IEEE Trans. on Image Processing*, vol. 8, pp. 1374–1387, Oct. 1999.
- [35] W. K. Hastings, "Monte Carlo Sampling Methods using Markov Chains and their Applications," *Biometrika*, vol. 57, pp. 97–109, Apr. 1970.
- [36] W. Siedlecki and J. Sklansky, "Constrained genetic optimization via dynamic reward-penalty balancing and its use in pattern recognition," in *Proceedings of the Third International Conference on Genetic Algorithms*, pp. 141–150, Jun. 1989.
- [37] H.-P. Schwefel, *Evolution and Optimum Seeking*. Wiley, Jan. 1995.
- [38] N.-H. Kim, "EAS 6939 Approximation and Optimization in Engineering. Constrained Optimization. Chp 5," 2010. www2.mae.ufl.edu/nkim/eas6939/ConstrainedOpt.pdf.
- [39] M. Held and R. M. Karp, "The travelling salesman problem and minimum spanning trees," *Operations Research*, vol. 18, pp. 1138–1162, 1970.
- [40] M. L. Fisher, "Optimal solution of scheduling problems using lagrange multipliers: Part i," *Operations Research*, vol. 21, no. 5, pp. 1114–1127, 1973.
- [41] Z. Michalewicz, "Genetic algorithms, numerical optimization, and constraints," in *Sixth International Conference on Genetic Algorithms*, pp. 151–158, Morgan Kaufmann, Jul. 1995.
- [42] D. W. Coit, A. E. Smith, and D. M. Tate, "Adaptive penalty methods for genetic optimization of constrained combinatorial problems," *INFORMS Journal on Computing*, vol. 8, pp. 173–182, May 1996.
- [43] J. T. Richardson, M. R. Palmer, G. E. Liepins, and M. Hilliard, "Some guidelines for genetic algorithms with penalty functions," in *Proceedings of the Third International Conference on Genetic Algorithms*, pp. 191–197, Morgan Kaufmann Publishers Inc., Jun. 1989.
- [44] J. Joines and C. Houck, "On the use of non-stationary penalty functions to solve nonlinear constrained optimization problems with ga's," in *IEEE World Congress on Computational Intelligence*, pp. 579–584 vol.2, Jun. 1994.
- [45] J. Bean and A. Hadj-Alouane, "A dual genetic algorithm for bounded integer programs," Tech. Rep. 92-53, The University of Michigan, Department of Industrial and Operations Engineering, Oct. 1992.

- [46] D. Geman, S. Geman, C. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 609–628, Jul. 1990.
- [47] M. C. Robini, Y. Bresler, and I. E. Magnin, "On the convergence of metropolis-type relaxation and annealing with constraints," *Probability in the Engineering and Informational Sciences*, vol. 16, pp. 427–452, Oct. 2002.
- [48] L. Yang, Y. Sagduyu, J. Zhang, and J. H. Li, "Distributed power control for ad-hoc communications via stochastic nonconvex utility optimization," in *Communications (ICC), 2011 IEEE International Conference on*, pp. 1–5, June 2011.
- [49] A. Saifullah, C. Wu, P. B. Tiwari, Y. Xu, Y. Fu, C. Lu, and Y. Chen, "Near optimal rate selection for wireless control systems," *ACM Trans. Embed. Comput. Syst.*, vol. 13, pp. 128:1–128:25, Apr. 2014.
- [50] M. Azim, Z. Aung, W. Xiao, V. Khadkikar, and A. Jamalipour, "Localization in wireless sensor networks by constrained simultaneous perturbation stochastic approximation technique," in *Signal Processing and Communication Systems (ICSPCS), 2012 6th International Conference on*, pp. 1–9, Dec 2012.
- [51] M. Minelli, M. Ma, M. Coupechoux, J.-M. Kelif, M. Sigelle, and P. Godlewski, "Optimal relay placement in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, pp. 998–1009, Feb. 2014.
- [52] 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.
- [53] A. Saleh, O. Bulakci, J. Hamalainen, S. Redana, and B. Raaf, "Analysis of the impact of site planning on the performance of relay deployments," *IEEE Transactions on Vehicular Technology*, vol. 61, pp. 3139–3150, Sep. 2012.



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.



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.



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 Télécommunications 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.



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*.



Philippe Godlewski received the Engineer's 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.