Assessing the Performance of NOMA in a Multi-Cell Context: A General Evaluation Framework

Anthony Bardou*‡ Jean-Marie Gorce† Thomas Begin‡
*IC, EPFL, Lausanne, Switzerland (anthony.bardou@epfl.ch)
†INSA Lyon, Inria, CITI, UR3720, Villeurbanne, France (jean-marie.gorce@insa-lyon.fr)
‡ENS Lyon, Univ Lyon 1, CNRS, LIP, Lyon, France (thomas.begin@ens-lyon.fr)

***Abstract*—Non-Orthogonal Multiple Access (NOMA) is a resource-sharing mechanism (RSM) envisioned for 5G and beyond cellular networks. While its benefit at the level of a single cell has been properly established, assessing NOMA performance at the scale of a cellular network remains an open research problem. This is mainly due to the inter-dependencies between scheduling, power control and inter-cell interference. Some algorithms have been proposed to optimize resource allocation in a multi-cell network, but they require a perfect and unrealistic knowledge of the whole channel states.**

In this paper, we leverage black-box optimization techniques to build a versatile evaluation framework, able to assess the performance of multi-cell networks implementing a large variety of RSMs. Subsequently, we exploit this evaluation framework with a decentralized Bayesian optimization algorithm to assess the performance of several, well-known RSMs under various fairness requirements and beamforming efficiencies. Our results show that, among the studied RSMs, NOMA combined with a full reuse policy consistently emerges as the one able to achieve, at the cellular network scale, the highest end-users Shannon capacities under fairness constraints.

***Index Terms*—Cellular Networks, NOMA, Bayesian Optimization, Fairness**

I. Introduction

SHARING radio resources is a key problem in cellular networks since their infancy. Inside each cell, multiple access techniques exploit orthogonal resources (in time, frequency, or codes) while inter-cell interference (ICI) is controlled through sophisticated reuse policies. Typically, for 2G cellular networks, ICI was controlled through static resource allocation based on graph coloring. For 3G cellular networks, ICI was mitigated by the use of scrambling codes and by controlling the load factor of the cells. For 4G cellular networks exploiting OFDMA, graph coloring strategies were enhanced with fractional frequency reuse (FFR) [1]–[3] or even full-reuse policies [4]. Indeed, if graph coloring-based approaches provide high signal-to-interference-plus-noise-ratio (SINR) to edge users, they suffer from a strong reduction of the frequency bandwidth available for each cell. Conversely, full reuse consists in providing each cell with the full bandwidth in spite of the potential high interference generated for edge users. Full reuse has been shown to provide a higher network spectral efficiency despite the low SINR associated with edge users (typically below 0 dB TODO: Verify). FFR was elaborated as a tradeoff between both strategies, by considering the superposition of two classes of users: inner users in the center of the cell, and edge users, whose reuse strategy differs. Inner users exploit a fraction of the spectrum in a full reuse scheme whereas a graph coloring strategy on the remainder of the spectrum is applied for edge users [2], [3].

During the elaboration of the fifth generation (5G), enhancing 4G capabilities, resource-sharing strategies based on non-orthogonal multiple access (NOMA) techniques have been extensively investigated [5]. NOMA is a powerful technique that has the potential to increase the multiple access cellular network capacity [6] both in the downlink and the uplink. NOMA relies on the well-known principle of superposition coding (SC) elaborated for the Gaussian broadcast channel (G-BC) where two (or more) data streams are superposed [7]. In a 2-user NOMA, pairing a strong and a weak user together (with respect to their channel quality) greatly increases the cell's capacity. In such a setup, and unlike orthogonal multiple access (OMA), NOMA achieves the Shannon capacity limit for the G-BC. The expression of this capacity for a single cell under massive access is provided in [8] and shows the benefit of NOMA approaches even with dealing with a limited number of superposed data streams.

While NOMA has been shown to be a powerful strategy at the scale of a single cell, its benefits at the scale of a full network are not straightforward to establish. In particular, its impact on the ICI and on the optimal user equipment-base station (UE-BS) association has not yet been precisely evaluated. In [9], the authors analyzed the disincentives for the use of NOMA in 4G, and raised a few common myths about NOMA, showing that: (i) NOMA can achieve a good tradeoff between spectral efficiency and fairness; (ii) NOMA is not very complex to implement, at least when restricted to the superposition of pairwise users; (iii) NOMA can be combined with FFR strategies (contrary to a previous statement [10]).

Because of these observations, optimizing non-orthogonal resource-sharing mechanisms (RSMs) at the scale of a network is a complex matter. However, if the transmission powers at each BS and the UE/BS associations are efficiently performed, then NOMA can let the cells breathe, thus bringing more flex-

This work was supported in part by the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX- 0007) operated by the French National Research Agency (ANR).

ibility than classical intra-cell optimization while remaining less complex than a cell-free approach as currently investigated for 6G [11].

In this paper, we propose a general evaluation framework to assess the performance of an arbitrary RSM in a multi-cell context. In essence, this framework proposes to optimize the parameters of a cellular network implementing the RSM to be evaluated, before collecting the RSM's performance metrics. The framework formulates the optimization of the cellular network as a black-box optimization problem. This makes our framework versatile, as it is able to find efficient configurations in the cellular network under a minimal set of assumptions (e.g. regardless of the actual network topology or the interference model in use). Then, we apply this framework to compare the respective performance of several existing RSMs on a simulated, real-life-inspired cellular network. Our results consistently show the superiority of NOMA (used in conjunction with a full reuse policy) at finding the best tradeoffs between high end-users Shannon capacities, under various fairness requirements and beamforming efficiencies.

The remainder of the paper is organized as follows. Section II describes the methods previously proposed to evaluate NOMA in a multi-cell context and motivates the need for a general, agnostic evaluation framework for RSMs. Section III introduces the optimization framework and its parameters. Section IV describes the framework parameters chosen for the conducted study. Section V illustrates the abilities of the framework by comparing the performance of several RSMs and provides insights on the performance of NOMA at the cellular network scale. Finally, Section VI concludes this paper.

II. NOMA-Based Resource-Sharing Evaluation

Evaluating NOMA at the scale of a network is an important but hard problem if we consider simultaneously power transmission levels, UE/BS association, and NOMA encoding ordering.

In a couple of papers, such as [12] and more recently in [13], algorithmic solutions for individual power control for multi-cell networks with NOMA are investigated. In these papers, the user-BS association is fixed, and the proposed algorithms jointly optimize the resource and power allocation to each user. The proposed strategies rely on a multi-level NOMA over multiple resource blocks. The proposed approaches leverage the multiple degrees of freedom brought by NOMA. However, in practice, these solutions would require a huge amount of signaling between all users and BSs, making their implementation challenging. In [14], an optimal algorithm is studied for NOMA allocation in a single cell, on multiple sub-bands, with a weighted sum-rate. This cost function may consider fairness, but the choice of individual weights is complex. A less complex algorithm is proposed in [15].

In [16], the authors combine FFR and NOMA. They consider a classical FFR and they superpose, on top of it, a third category of users with low power. Unfortunately, this scheme induces a higher complexity and does not prevent the loss of performance already observed for FFR.

Another avenue to evaluate cellular networks is stochastic geometry [17]. Based on spatial randomness, this powerful tool enables the computation of the distribution of BS/UE radio links' quality (e.g. rate or outage). Stochastic geometry has been used to evaluate the impact of FFR mechanisms with OFDMA in [1], [18], [19]. However, stochastic geometry does not seem to be applicable to the case of NOMA mechanisms as the computation of sum-rates and other customary capacity metrics involves second-order statistics. In [20], the authors studied NOMA at a large scale but their results were limited to SINR meta-distributions. The authors were able to evaluate individual rates of UEs but not a global evaluation of the sum-rates, as this would require the distribution of the cell geometry as discussed in [21], which is generally hard to determine.

The mentioned papers provide different ways to address the hard problem of assessing the performance of NOMA at a multi-cell network scale. However, in addition to the difficulty of obtaining key performance metrics or the requirement of unrealistic knowledge about the channel states, the proposed evaluation frameworks cannot be applied to another RSM without major adjustments. This begs the question of the fair comparison between different RSMs. In our desire to fairly compare RSMs without the need of unrealistic knowledge, we introduce in the next section an evaluation framework capable of optimizing an arbitrary RSM on an arbitrary cellular network and of evaluating key performance metrics at the cellular network scale.

III. Unified Framework for the Evaluation of RSMs

A. Core Idea

A RSM can only be fairly evaluated on a cellular network which has been optimized for its needs. Therefore, an efficient evaluation framework must be able to optimize the performance of a cellular network given the RSM to evaluate. That is why, at the core of the proposed evaluation framework lies the following optimization problem:

$$\underset{\boldsymbol{x}\in\mathcal{D}}{\text{maximize}} \;\; f(\boldsymbol{x};\mathcal{T},\mathcal{P},\mathcal{R}). \tag{1}$$

The problem (1) comprises five core ingredients:

- the RSM $\mathcal{R}$ to evaluate (e.g. NOMA)
- the propagation model $\mathcal{P}$ (e.g. log-distance path loss, accounting for beamforming, shadowing...),
- the cellular network topology $\mathcal{T}$,
- the network parameters to optimize (e.g. the power levels of the BSs), specified by a search space $\mathcal{D}$,
- an objective function $f : \mathcal{D} \to \mathbb{R}$ collecting the performance of the cellular network specified by the triplet $(\mathcal{T},\mathcal{P},\mathcal{R})$ and parameterized by $\boldsymbol{x} \in \mathcal{D}$.

These five ingredients are set by the performance analyst and completely specify the optimization problem (1). The goal of the evaluation framework is to solve (1) by discovering $\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}\in\mathcal{D}} f(\boldsymbol{x})$[1] (note that the parameters $\mathcal{T}$, $\mathcal{P}$ and $\mathcal{R}$ have been omitted for the sake of brevity) with the

[1]In this paper, f is assumed to be scalar. We postpone the extension of the framework to vector-valued objective functions to future works.

procedure explained in the next section. Finally, now that the cellular network is adequately configured for the RSM $\mathcal{R}$, key performance metrics can be collected and sent to the performance analyst.

B. Bayesian Optimization

Getting a closed-form for the objective function f in (1) is challenging (if not impossible in the general case). In fact, the ingredients $\mathcal{D}, \mathcal{T}, \mathcal{P}$ and $\mathcal{R}$ have complex, intricate effects on the objective function f. As such, f can be seen as a black-box function, that is, a function with an unknown closed-form for which oracles providing higher-order information (e.g. $\nabla_{\boldsymbol{x}} f$, $\nabla^2_{\boldsymbol{x}} f$) are unknown or do not exist. As a direct consequence, classical first-order optimization algorithms (such as gradient descent [22]) cannot be applied to the problem (1). Since f is computed by collecting performance metrics on the cellular network, it can also be considered expensive to evaluate and noisy. Therefore, higher-order information cannot be approximated by Monte-Carlo-based estimators [23]. In this context, Bayesian Optimization (BO) appears as a natural candidate to solve (1). In fact, BO is widely used for black-box optimization purposes in a variety of domains [24]–[26]. In this section, we provide some background about BO and introduce its core assumption.

A BO algorithm often leverages a Gaussian process (GP) as a surrogate model [27] for the black-box objective function f. A GP is a stochastic process, that is, a collection of random variables $\{Y(\boldsymbol{x})\}_{\boldsymbol{x}\in\mathcal{D}}$ indexed by a set $\mathcal{D}$. As its name suggests, in a GP, any finite collection $\{Y(\boldsymbol{x}_1), \cdots, Y(\boldsymbol{x}_k)\}$ has a joint multivariate Gaussian distribution. As such, it is fully specified by its mean function $\mu(\boldsymbol{x}) = \mathbb{E}[Y(\boldsymbol{x})]$ (we assume $\mu(\boldsymbol{x}) = 0$ without loss of generality) and its covariance function $k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(Y(\boldsymbol{x}) - \mu(\boldsymbol{x}))(Y(\boldsymbol{x}') - \mu(\boldsymbol{x}'))]$.

Naturally, a BO algorithm requires a single, core assumption about its black-box objective function f.

Assumption 1. *f is a Gaussian process $\mathcal{GP}(0, k)$, with mean 0 and covariance function $k : \mathcal{D} \times \mathcal{D} \to \mathbb{R}^+$.*

Note that the covariance function k makes the GP a versatile surrogate model. In fact, a GP can be a satisfying model for rough functions (e.g. by setting k so that the GP is a Ornstein–Uhlenbeck process [28]) as well as smooth ones (e.g. by setting k to be a RBF kernel [29]). Covariance functions can also be combined with basic operators (e.g. summed or multiplied) to form richer covariance functions and make the GP model a better fit for the objective function.

Let $\boldsymbol{X}_t = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_t)$ be a $t \times d$ matrix of inputs (*i.e.* t inputs of d dimensions) and $\boldsymbol{y}_t = (f(\boldsymbol{x}_1) + \epsilon, \cdots, f(\boldsymbol{x}_t) + \epsilon)$ be a corresponding t-dimensional vector of observed noisy outputs, with $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Then, under Assumption 1, $f(\boldsymbol{x})|\boldsymbol{X}_t, \boldsymbol{y}_t \sim \mathcal{N}\left(\mu_t(\boldsymbol{x}), \sigma_t^2(\boldsymbol{x})\right)$ with

$$\mu_t(\boldsymbol{x}) = \boldsymbol{k}^\top(\boldsymbol{x}, \boldsymbol{X}_t)\boldsymbol{\Delta}_t^{-1}\boldsymbol{y}_t, \tag{2}$$

$$\sigma_t^2(\boldsymbol{x}) = k(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}^\top(\boldsymbol{x}, \boldsymbol{X}_t)\boldsymbol{\Delta}_t^{-1}\boldsymbol{k}(\boldsymbol{x}, \boldsymbol{X}_t) \tag{3}$$

where $\boldsymbol{k}(\boldsymbol{x}, \boldsymbol{X}_t) = (k(\boldsymbol{x}, \boldsymbol{x}_i))_{i\in[1,t]}$, $\boldsymbol{\Delta}_t = \boldsymbol{K}_t + \sigma^2\boldsymbol{I}_t$, $\boldsymbol{I}_t$ is the $t \times t$ identity matrix and $\boldsymbol{K}_t = (k(\boldsymbol{x}_i, \boldsymbol{x}_j))_{i,j\in[1,t]}$.

At time t, a BO algorithm must find the next query $\boldsymbol{x}_{t+1} \in \mathcal{D}$ such that $f(\boldsymbol{x}_{t+1})$ is likely to be large, and such that observing $f(\boldsymbol{x}_{t+1})$ brings information about the black-box f. This exploration-exploitation dilemma is usually solved by maximizing an acquisition function $\varphi_t(\boldsymbol{x})$. Many acquisition functions exist, such as GP-UCB [30], Knowledge Gradient [31], Probability of Improvement [32] and Expected Improvement [33].

C. Evaluation Framework Algorithm

Algorithm 1 Proposed Evaluation Framework

Input: RSM $\mathcal{R}$, propagation model $\mathcal{P}$, network topology $\mathcal{T}$, network parameters domain $\mathcal{D}$, black-box objective function f, BO algorithm $\mathcal{A}$.
Output: key performance metrics $\mathcal{M}$

```
t = 0
Initialize the datasets X_t = ∅, y_t = ∅
while A does not converge do
    Use A and the datasets (X_t, y_t) to find x_{t+1}
    Set the network parameters to x_{t+1}
    Collect y_{t+1} = f(x_{t+1}) + ε on the network
    Update the datasets X_{t+1} = X_t ∪ {x_{t+1}}, y_{t+1} = y_t ∪ {y_{t+1}}
end while
Collect key performance metrics M on the network with optimized parameters x* ∈ D
return M
```

The proposed evaluation framework is given in Algorithm 1. As previously described, Algorithm 1 solves the optimization problem (1) with a BO algorithm $\mathcal{A}$, before collecting the key performance metrics on the optimized network and returning them to the performance analyst.

In addition to the five main ingredients described in Section III-A, the performance analyst must also provide a BO algorithm $\mathcal{A}$. In fact, depending on the problem specified by the five main ingredients of the framework, some BO algorithms may be more suited than others (e.g. high-dimensional BO algorithms [34], batch BO algorithms [35]...)

In the following sections, we illustrate the capabilities of Algorithm 1 by providing valuable insights on the performance of NOMA-based RSMs. Section IV describes how the main ingredients of the framework (including the BO algorithm $\mathcal{A}$) are specified, while Section V provides and discusses the results of our study.

IV. Description of the Study

In this section, we describe how the evaluation framework proposed in Section III is used to produce key performance metrics enabling the comparison between five different RSMs. The study is conducted on a cellular network covering a service area Ω. It comprises a set of BSs $\mathcal{N} = [1, n]$ and a set of UEs $\mathcal{M} = [1, m]$ distributed over Ω. Each UE is associated with the BS that provides the highest SINR.

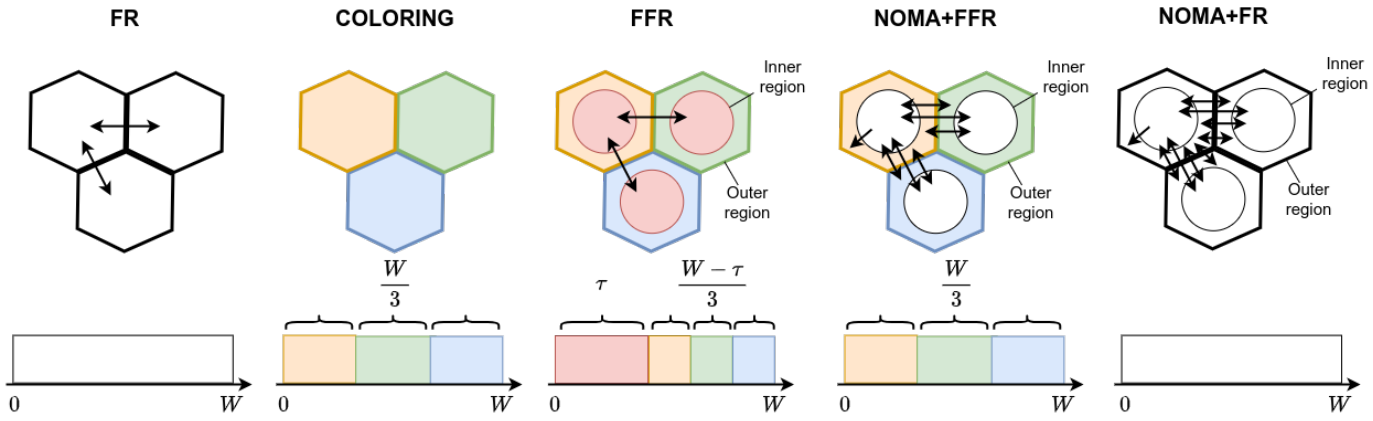

Fig. 1. The studied RSMs illustrated on three adjacent hexagonal cells and the corresponding splitting of the frequency band of size W. Interference between the top left cell's regions and other regions are represented by black arrows.

A. Resource-Sharing Mechanisms $\mathcal{R}$

In this study, we compare five different RSMs, illustrated by Figure 1. Let us briefly describe each of them.

The simplest studied RSM is the full reuse solution (FR) used for 4G networks, which does not introduce any additional mechanism to mitigate ICI. Another simple RSM is the classical graph-coloring solution (COLORING) used for 2G networks, which partitions the frequency band into chunks of equal size and allocates different chunks to neighboring cells. ICI is effectively mitigated, at the expense of a significant shrinkage of the frequency band for each UE. Next, we consider the fractional frequency reuse solution (FFR), which introduces an additional structure within each cell by splitting it into an inner and an outer region. It allocates a chunk of size τ in the frequency band for the cells' inner regions while the cells' outer regions share the remainder of the frequency band (of size $W-\tau$) with a color-based solution.

We compare these RSMs with two NOMA-based RSMs, where each BS exploits a 2-user NOMA technique.[2] Each cell is split into an inner and an outer region. SC is used on the two regions, and SIC is performed by the inner users to remove the interference due to the signals transmitted to the outer users. On the one hand, NOMA+FFR combines FFR and NOMA, following the concept proposed in [9]. The inner users adopt a full reuse strategy, while the outer users adopt a color-based strategy. On the other hand, NOMA+FR proposes that all users (inner and outer) adopt a full reuse strategy.

Note that the chosen procedure used to split a cell into an inner and an outer region is detailed in the Appendix.

B. Parameters $\mathcal{D}$ and Objective Function f

We consider a single transmission power level for each region of the cellular network. For the RSMs FR and COLORING, this naturally implies that each BS i ($i \in [1,n]$) has a single transmission power level x_i. We assume that each BS has a minimal transmission power of P_- and a maximal transmission power of P_+. This leads to a natural formulation for the cellular network parameters $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathcal{D}$ with

$$\mathcal{D} = \mathcal{D}^{(1)} \times \cdots \times \mathcal{D}^{(n)} = [P_-, P_+]^n. \tag{4}$$

For the RSMs that split each cell into two regions, we consider two transmission power levels $\boldsymbol{p}_i = \left(p_i^{(in)}, p_i^{(out)}\right)$ per BS i, $i \in [1,n]$. Despite the relative simplicity of this scenario, one can expect a significant capacity gain [8]. In this formulation, each $\boldsymbol{p}_i$ must also satisfy a power constraint, that is $P_- \leq ||\boldsymbol{p}_i||_1 \leq P_+$. Additionally, we enforce $p_i^{(in)} \leq p_i^{(out)}$. In order to embed these two constraints into a rectangular parameter space, we exploit the following bijection:

$$\underbrace{\left(p_i^{(in)}, p_i^{(out)}\right)}_{\boldsymbol{p}_i} \rightarrow \underbrace{\left(p_i^{(in)} + p_i^{(out)}, \frac{p_i^{(out)}}{p_i^{(in)} + p_i^{(out)}}\right)}_{\boldsymbol{x}_i}.$$

Consequently, we consider that the cellular network parameters are $\boldsymbol{x} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n) \in \mathcal{D}$ with

$$\mathcal{D} = \mathcal{D}^{(1)} \times \cdots \times \mathcal{D}^{(n)} = \left([P_-, P_+] \times \left[\frac{1}{2}, 1\right)\right)^n. \tag{5}$$

Having defined the parameter space $\mathcal{D}$ for the considered RSMs, we turn to the specification of the objective function f. Given a configuration $\boldsymbol{x}_t$ of the cellular network, we consider the achievable rate $c^{(j)}(\boldsymbol{x}_t)$ of each UE $j \in [1,m]$ at time t, with respect to the propagation model $\mathcal{P}$ (see Section IV-D), as the main performance metric to optimize at the UE scale. Consequently, the performance metrics at the cellular network scale are $\boldsymbol{c}(\boldsymbol{x}_t) = \left(c^{(1)}(\boldsymbol{x}_t), \cdots, c^{(m)}(\boldsymbol{x}_t)\right) \in \mathbb{R}^m$.

To turn these performance metrics into a scalar objective value, we consider the α-fairness, introduced by [36]:

$$F_\alpha(\boldsymbol{c}_t) = \begin{cases} \sum_{i=1}^m \log c_t^{(j)} & \text{if } \alpha = 1, \\ \sum_{j=1}^m \frac{\left(c_t^{(j)}\right)^{1-\alpha}}{1-\alpha} & \text{otherwise,} \end{cases} \tag{6}$$

where, by a slight abuse of notation, $c_t^{(j)}$ denotes $c^{(j)}(\boldsymbol{x}_t)$.

Observe that (6) leads to natural tradeoffs between the sum of the rates in $\boldsymbol{c}_t$ and the fairness of these rates. In fact, α controls the importance of fairness in the objective function. As examples, $\alpha = 0$ defines the sum-rate as the objective

[2] Although NOMA can deal with more superposition levels, this would increase the complexity of both resource allocation and decoding algorithms for a moderate additional gain [8].

function, $\alpha = 1$ boils down to the proportional fairness and, since $\lim_{\alpha \to +\infty} F_\alpha(\boldsymbol{c}_t) = \min_{j \in [1,m]} c_t^{(j)}$, $\alpha \to +\infty$ can be viewed as one of the most stringent definitions of fairness.

Because of its versatility, we use (6) as the objective function of the cellular network. To express its final form, we combine (6) with the optimal scheduling strategy within each region of each cell. Given a set $\mathcal{A}^{(k)}$ of $m^{(k)}$ UEs belonging to a region k, we define the optimal scheduling strategy $\boldsymbol{s}^{(k)} = \left(s_j^{(k)}\right)_{j \in \mathcal{A}^{(k)}}$ as the solution of

$$\begin{aligned} &\underset{\boldsymbol{s} \in [0,1]^{m^{(k)}}}{\text{maximize}} \; F_\alpha\left(\boldsymbol{s} \odot \boldsymbol{c}_t^{(k)}\right) \\ &\text{s.t. } \boldsymbol{s}^\top \mathbf{1} = 1 \end{aligned} \tag{7}$$

with $\boldsymbol{c}_t^{(k)} = \left(c_t^{(j)}\right)_{j \in \mathcal{A}^{(k)}}$, $\mathbf{1}$ a conformable vector of ones and $\odot$ the Hadamard product.

Proposition 1. *The optimal scheduling strategy for a set $\mathcal{A}^{(k)}$ of UEs is $\boldsymbol{s}^{(k)} = \left(s_j^{(k)}\right)_{j \in \mathcal{A}^{(k)}}$, with*

$$s_j^{(k)} = \begin{cases} \mathbb{1}_{j=j^*} & \text{if } \alpha = 0, \\ \dfrac{\left(c_t^{(j)}\right)^{(1-\alpha)/\alpha}}{\sum_{i \in \mathcal{A}^{(k)}} \left(c_t^{(i)}\right)^{(1-\alpha)/\alpha}} & \text{otherwise,} \end{cases} \tag{8}$$

with $j^* = \arg\max_{j \in \mathcal{A}^{(k)}} c_t^{(j)}$ and $\mathbb{1}_{j=j^*}$ the indicator function indicating if j is the index of the UE that has the largest rate among the UEs in $\mathcal{A}^{(k)}$.

Proof (sketch). The special case $\alpha = 0$ is trivial to solve as the solution to maximize the sum-rate is to find the UE that has the largest rate and let it use the full resource.

For $\alpha \neq 0$, the function in the constrained problem (7) is convex since it is a sum of convex functions. Therefore, it can be turned into an unconstrained problem by considering its Lagrangian relaxation $\mathcal{L}_\alpha(\boldsymbol{s}, \lambda)$, and its solution is one of the critical points of $\mathcal{L}_\alpha$, found by the criterion $\nabla \mathcal{L}_\alpha = 0$. The relaxation for $\alpha \neq 1$ is

$$\begin{aligned} \mathcal{L}_\alpha(\boldsymbol{s}, \lambda) &= F_\alpha\left(\boldsymbol{s} \odot \boldsymbol{c}_t^{(k)}\right) - \lambda\left(\boldsymbol{s}^\top \mathbf{1} - 1\right) \\ &= \frac{1}{1-\alpha} \sum_{j \in \mathcal{A}^{(k)}}^{k} s_j^{1-\alpha} \left(c_t^{(j)}\right)^{1-\alpha} \\ &\quad - \lambda \left(\sum_{j \in \mathcal{A}^{(k)}} s_j - 1 \right). \end{aligned} \tag{9}$$

Computing the gradient $\nabla \mathcal{L}_\alpha$ by differentiating (9) is trivial, as is solving the system of equations $\nabla \mathcal{L}_\alpha = 0$ for $\boldsymbol{s}$. It yields a single critical point, which is the solution of (7) given in (8). Applying the same technique to the special case $\alpha = 1$ yields the same solution. □

We can now derive the form of the objective function. Let $\mathcal{A} = \left\{\mathcal{A}^{(1)}, \cdots, \mathcal{A}^{(r)}\right\}$ be the set of all regions in the cellular network,[3] the objective function is

$$f_\alpha(\boldsymbol{x}_t) = \sum_{\mathcal{A}^{(k)} \in \mathcal{A}} \sum_{j \in \mathcal{A}^{(k)}} F_\alpha\left(\boldsymbol{s}^{(k)} \odot \boldsymbol{c}^{(j)}(\boldsymbol{x}_t)\right) \tag{10}$$

[3]Note that $r = n$ or $r = 2n$ depending on the considered RSM.

with F_α and $\boldsymbol{s}^{(k)}$ given by (6) and (8), respectively.

C. BO Algorithm $\mathcal{A}$

In the context of our study, the optimization problem (1) may have up to $2n$ parameters for a cellular network comprising n BSs (see Section IV-B). In the field of BO, this problem is considered high-dimensional even for medium-sized networks (say $n > 5$). Furthermore, we would like to address the problem (1) in a decentralized fashion, as it would ease the replication of this study in a real cellular network.

We rely on INSPIRE, the decentralized, high-dimensional BO algorithm proposed in [24]. INSPIRE is initially designed to optimize a $2n$-dimensional proportional fairness within a WLAN comprising n nodes. The problem is addressed in a decentralized fashion, under the assumption that the interference experienced by a user j associated with a node i is caused solely by a subset of nodes, defined as a neighborhood of node i and denoted $\mathcal{N}_i$.

Note that the proportional fairness optimized by INSPIRE is a special case ($\alpha = 1$) of the objective function (10). Moreover, because (10) is based on the achievable rates of the UEs (and hence, on their SINRs), assuming that only the neighborhood of BS i causes the interference of its associated UEs is mild. Therefore, INSPIRE is a good candidate for the black-box optimization of (10).

Assumption 2.

$$f_\alpha(\boldsymbol{x}_t) \approx \sum_{i=1}^{n} \sum_{\mathcal{A}^{(k)} \in \mathcal{A}_i} \sum_{j \in \mathcal{A}^{(k)}} F_\alpha(\boldsymbol{s}^{(k)} \odot \boldsymbol{c}^{(k)}(\boldsymbol{x}_t^{(\mathcal{N}_i)})) \tag{11}$$

with $\mathcal{A}_i$ the regions controlled by BS i, $\boldsymbol{x}_t^{(\mathcal{N}_i)} = \left(\boldsymbol{x}_t^{(i')}\right)_{i' \in \mathcal{N}_i}$ and $\mathcal{N}_i$ comprising all the BSs at radio range of BS i (including BS i itself).

To apply INSPIRE, we need an additive decomposition of the objective function. Let us denote by $|\mathcal{N}_i|$ the cardinality of $\mathcal{N}_i$ and consider $f_\alpha^{(i)} : \mathcal{D}^{(\mathcal{N}_i)} \to \mathcal{D}$, with $\mathcal{D}^{(\mathcal{N}_i)} = \prod_{i' \in \mathcal{N}_i} \mathcal{D}^{(i')}$, such that

$$f_\alpha^{(i)}\left(\boldsymbol{x}_t^{(\mathcal{N}_i)}\right) = \sum_{i' \in \mathcal{N}_i} \sum_{\mathcal{A}^{(k)} \in \mathcal{A}_{i'}} \sum_{j \in \mathcal{A}^{(k)}} \frac{F_\alpha(\boldsymbol{s}^{(k)} \odot \boldsymbol{c}^{(k)}(\boldsymbol{x}_t^{(\mathcal{N}_{i'})}))}{|\mathcal{N}_{i'}|}. \tag{12}$$

Observe that $f_\alpha^{(i)}$ is an additive decomposition of (11). That is, it is trivial to show that

$$f_\alpha(\boldsymbol{x}_t) = \sum_{i=1}^{n} f_\alpha^{(i)}\left(\boldsymbol{x}_t^{(\mathcal{N}_i)}\right). \tag{13}$$

We can now introduce the standard assumptions used in decentralized BO, starting with the Lipschitz continuity.

Assumption 3. *$\forall i \in [1, n]$, $f_\alpha^{(i)}$ is Lipschitz continuous with Lipschitz constant $L^{(i)}$, that is*

$$\forall \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}^{(\mathcal{N}_i)}, |f_\alpha^{(i)}(\boldsymbol{x}) - f_\alpha^{(i)}(\boldsymbol{x}')| \leq L^{(i)} ||\boldsymbol{x} - \boldsymbol{x}'||_2 \tag{14}$$

Intuitively, Assumption 3 implies that the rate of change of (12) is bounded by $L^{(i)}$ across its entire domain. Finally, we introduce GPs as surrogate models for f_α.

Assumption 4. *$\forall i \in [1,n]$, $f_\alpha^{(i)}$ described with (12) is a Gaussian process $\mathcal{GP}^{(i)}\left(0, k^{(i)}(\boldsymbol{x}, \boldsymbol{x}')\right)$, with $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}^{(\mathcal{N}_i)}$.*

With Assumption 4, each GP $\mathcal{GP}^{(i)}$ is a surrogate model for the corresponding term $f_\alpha^{(i)}$. Assumption 4 naturally implies that Assumption 1 holds, since a sum of GPs is also a GP.

Under Assumptions 2, 3 and 4, INSPIRE (as described in [24]) can be applied. Note that, under these assumptions, INSPIRE is minimax optimal (*i.e.* it minimizes a tight upper bound of its objective function).

D. Propagation Model $\mathcal{P}$

The radio spectrum bandwidth W is shared among all BSs while each BS shares its spectrum, divided into resource blocks (RBs), among all its associated UEs. The average radio link gain between one BS i and one UE j is given by:

$$g_{i,j} = g_0 \cdot s_{i,j} \cdot d_{i,j}^{-\alpha}, \tag{15}$$

where g_0 and α are the standard parameters of a distance-based propagation model, $s_{i,j}$ stands for shadowing and $d_{i,j}$ is the distance between the BS i and the UE j.

For each RSM $\mathcal{R}$, we compute the SINRs using power densities. Although this boils down to the well-known SINR expressions for three of the considered RSMs (namely `FR`, `COLORING` and `FFR` described in Section IV-A), let us derive the SINR expressions for the two NOMA-based RSMs.

Recall that, in these schemes, each BS has to classify its UEs into two classes: strong UEs and weak UEs (see the Appendix). Strong UEs are typically located around the BS whereas weak UEs tend to be located nearby the cell edge. Thus, the service area of a BS can be divided into two regions: the inner region containing all strong UEs and the outer region containing all weak UEs. Additionally, BS i must handle two transmission powers: $p_i^{(in)}$ for strong UEs and $p_i^{(out)}$ for weak UEs (see Section IV-B). Finally, observe that `NOMA+FR` is a special case of `NOMA+FFR` with a number of colors $K = 1$ (see Figure 1).

Following the SIC strategy, strong UEs decode first the signal for weak UEs and cancel their interference, whereas weak UEs decode only their own signal. Note that the latter do not experience strong interference from inner users signals thanks to the chosen parameters domain (5) enforcing $p_i^{(in)} \leq p_i^{(out)}$.

The power density for an inner UE (resp. outer UE) associated with a BS i is simply given by $dp_1^{(i)} = p_1^{(i)}/W$ (resp. $dp_2^{(i)} = Kp_2^{(i)}/W$). Table I indicates the expression of the interference power densities according to the UEs' locations. Observe that, in agreement with Figure 1, an inner UE does not suffer interference resulting from transmissions to outer UEs of the same cell (thanks to SIC). Conversely, an inner UE suffers interference resulting from transmissions to UEs associated to a different BS i (be they in the inner or outer region). In that case, the transmission in the outer region causes an interference of power density $\frac{g_{i,j}p_2^{(i)}}{W/K}$ mW/Hz, but only applied to a fraction W/K of the frequency band. Therefore, when considering the average of the power density over the whole frequency band of size W, it boils down to a power density of $\frac{1}{K}\frac{g_{i,j}p_2^{(i)}}{W/K} = \frac{g_{i,j}p_2^{(i)}}{W}$. Note that a more accurate allocation may be done by letting the BS allocate a RB in the sub-band with the lowest interference to each of its inner UEs. This is kept out of the scope of this paper.

TABLE I
POWER DENSITIES OF THE INTERFERENCE CAUSED BY AN UE j' (ASSOCIATED WITH A BS i) AND SUFFERED BY AN UE j (ASSOCIATED WITH A BS OF THE SAME COLOR).

Region of UE j	UE j and UE j' on same BS?	Region of UE j'	Interference (dBm/Hz)
Inner	Yes	Inner	0
		Outer	0
	No	Inner	$g_{i,j}p_1^{(i)}/W$
		Outer	$g_{i,j}p_2^{(i)}/W$
Outer	Yes	Inner	$g_{i,j}p_1^{(i)}/W$
		Outer	0
	No	Inner	$g_{i,j}p_1^{(i)}/W$
		Outer	$Kg_{i,j}p_2^{(i)}/W$

The SINRs resulting from this strategy for an inner UE j and an outer UE j' associated with a BS i, denoted by $\gamma_{i,j}^{(in)}$ and $\gamma_{i,j'}^{(out)}$, respectively, are:

$$\gamma_{i,j}^{(in)} = \frac{g_{i,j}p_1^{(i)}}{WN + \sum_{i' \in \mathcal{N}\setminus\{i\}} g_{i',j}(p_1^{(i')} + p_2^{(i')})} \tag{16}$$

$$\gamma_{i,j}^{(out)} = \frac{Kg_{i,j}p_2^{(i)}}{WN + g_{i,j}p_1^{(i)} + \sum_{i' \in \mathcal{N}\setminus\{i\}} g_{i',j}(p_1^{(i')} + \delta_{i,i'}Kp_2^{(i')})}, \tag{17}$$

where N is the noise power density and $\delta_{i,i'}$ is equal to 1 if i and i' use the same color for outer users, and 0 otherwise.

In our study, we consider both the case where BSs are equipped with single omnidirectional antennas (antenna gain of 0 dB) as well as the case where BSs exploit multiple antennas using a beamforming mode. We account for the benefit of using a antennas at a BS by adding a gain of $G_b(a) = 10\log_{10}(a)$ (dB) in the SINR of all UEs (see (16) and (17)). TODO: Est-ce qu'on présente toujours le cas où le beamforming ne réduit pas les interférences? Because the BSs do not coordinate for the beamforming action, we do not consider any interference reduction when using beamforming.

For a given SINR, the achievable rate of a given UE j associated with a BS i is given by the modified Shannon capacity [37]:

$$c^{(j)} = c(\gamma_{i,j}) = W_j \cdot \kappa \log_2(1 + \beta\gamma_{i,j}) \tag{18}$$

where W_j is the bandwidth used by the UE j, while κ and β represent the performance loss of a real system, due to imperfect modulation, coding schemes or fading impact. In our simulation, we set $\beta = \kappa = 1$ for the sake of simplicity.

E. Cellular Network Topology $\mathcal{T}$

Finally, we specify the cellular network topology considered in this study. We simulate the cellular network deployed in the city of Lyon, France by the main French mobile network

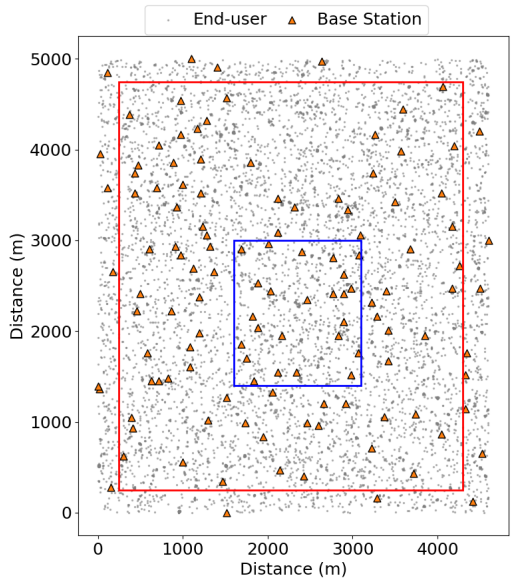

Fig. 2. Illustration of the simulated cellular network, inspired by Orange's cellular deployment in Lyon. The blue rectangle depicts the subset of BSs shown in Figure 6 while the red rectangle depicts the subset of the BSs used to collect our performance metrics (see Figures 3, 4 and 5 and Table III).

TABLE II
HOMEMADE SIMULATOR PARAMETERS.

Name	**Value**
Background noise N	-100 dBm/Hz
Path Loss	`LogDistance`$(d_{\texttt{ref}} = 1$ km, $L_{\texttt{ref}} = 128.1$ dBm, $n = 3.76)$
Bandwidth W	20 MHz
Beamforming gain	Within $[0, 6]$ dB
Optimization steps	110
GP covariance function k	Matérn ($\nu = \frac{3}{2}$) [29]
Acquisition function	Expected Improvement [33]

TABLE III
FIRST, SECOND AND THIRD QUARTILES FOR EUS' RATES DISTRIBUTIONS (IN MBPS) WHEN $\alpha = 2$.

	Beamforming Gain (dB)								
	0			**3**			**6**		
RSM	Q25	Q50	Q75	Q25	Q50	Q75	Q25	Q50	Q75
`FR`	0.19	0.30	0.55	0.38	0.65	1.03	0.56	0.87	1.75
`COL`	0.15	0.22	0.40	0.18	0.29	0.51	0.21	0.36	0.63
`FFR`	0.18	0.30	0.57	0.29	0.47	0.80	0.40	0.62	1.09
`NOMA+FFR`	**0.22**	0.33	0.65	0.37	0.68	1.15	0.55	0.99	1.83
`NOMA+FR`	0.21	**0.34**	**0.68**	**0.49**	**0.81**	**1.45**	**0.83**	**1.61**	**2.37**

operator (Orange). The precise locations of the BSs are made publicly available by a French national agency [38].

We consider a Log-Gaussian Cox process to generate a realistic spatial distribution of UEs that contains clusters and empty regions. Let us describe briefly how we implemented the process (for a thorough description, please refer to [39]). The domain containing the cellular network is split into tiles of 25 meters-long sides. For each tile k, a Gaussian random variable $X_k \sim \mathcal{N}(\mu, \sigma^2)$ is sampled. Then Y_k, the number of UEs within the tile k, is sampled from a Poisson distribution of intensity e^{X_k}, that is $Y_k \sim \text{Pois}(e^{X_k})$. Finally, Y_k UEs are uniformly generated within the tile. In our scenario, we set the parameters so that the process generated a set of UEs with a global spatial density similar to the one observed in Lyon ($\approx$ 11,000 persons / km^2). Then, a fraction of these UEs was randomly selected according to the market share of Orange in France ($\approx 38.5\%$). Finally, 10% of the UEs from this subset were randomly selected to account for active UEs only.

Figure 2 illustrates the obtained cellular network, with 125 BSs and around 9,000 UEs. Note that, in the next section, we collect and discuss performance metrics on the BSs (and their associated UEs) located within the red rectangle, to limit artifacts at the border of the simulated network.

V. NUMERICAL RESULTS

In this section, we illustrate the proposed evaluation framework (see Section III) by implementing all its ingredients according to the specifications described in Section IV in our homemade simulator (please see Table II for additional parameters) and we present insightful data about the performance of NOMA in a multi-cell context. More specifically, we study the distribution of the EUs' spectral efficiencies, the distribution of their achievable rates (see (18)) and the quality of the capacity-fairness tradeoff made by each RSM at the cellular network scale.

First, let us study the distribution of the EUs' spectral efficiencies. Figure 3 depicts the cumulative distribution functions (CDFs) of the UEs' spectral efficiencies, for each considered RSM, different fairness constraints α (see (6)) and beamforming gains. Although the fairness constraint α has virtually no effect on the CDFs, the benefits of using beamforming are substantial. In fact, the spectral efficiencies directly depend on the EUs' SINRs (see (18)), which get larger when the beamforming gain increases. This is illustrated by the right shift experienced by the CDFs as the beamforming gain grows. Finally, as one could expect, the RSMs exploiting orthogonal resources (`COLORING` and `FFR`) are the most spectrally efficient. This can be deduced by observing that the corresponding CDFs are consistently the lowest ones for a given spectral efficiency.

We now turn to the distribution of the EUs' achievable rates. Figure 4 depicts the CDFs of the UEs' rates, for each considered RSM, different fairness constraints α and beamforming gains. As one could expect, the distribution of rates among UEs is more spread when the fairness constraint α is low, and more concentrated when α is large. As an example, observe at the top row of Figure 4 that the distributions' supports spread over no less than eight orders of magnitude when $\alpha = 0.25$ (from 10^{-6} to 10^{2} Mbps) while they only cover 3 orders of magnitude when $\alpha = 2$ (from 10^{-1} to 10^{1} Mbps). Additionally, as in the case of the spectral efficiency (see Figure 3), increasing the beamforming gain significantly improves the rates. Finally, observe that the CDF of `NOMA+FULL_REUSE` is consistently the lowest CDF. This intuitively translates into the fact that `NOMA+FULL_REUSE` systematically ensures better (*i.e.* larger) rates for its users.

For a comparison between RSMs that is more quantitative

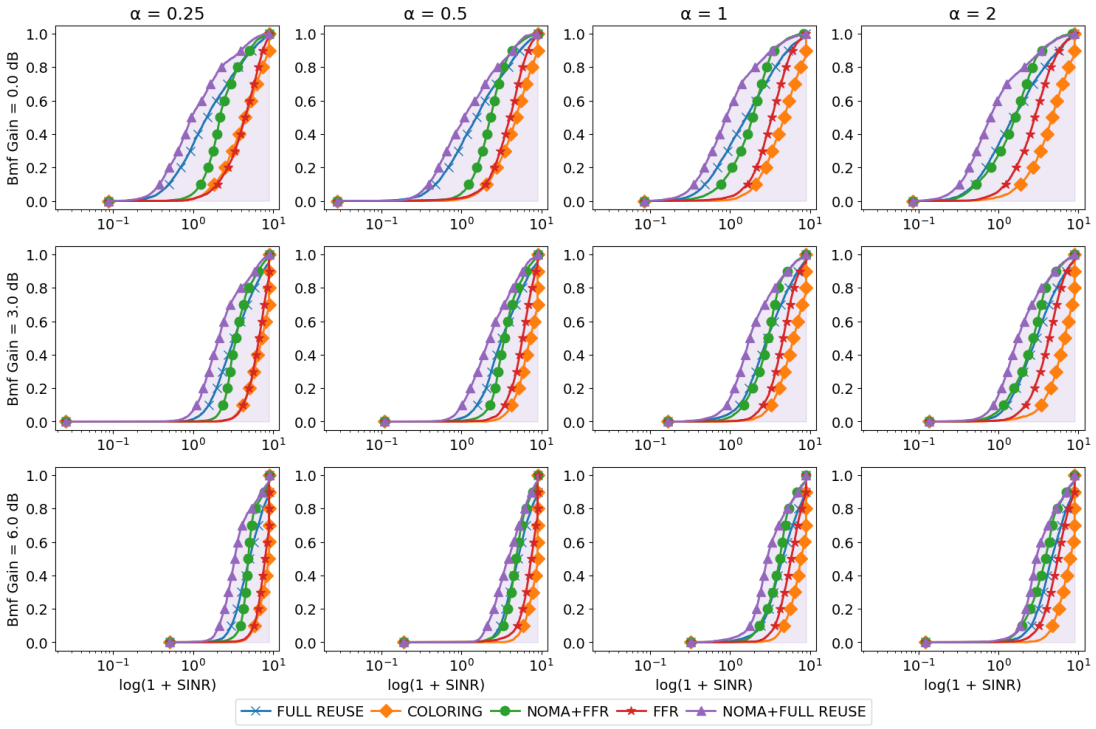

Fig. 3. The cumulative distribution functions of the UEs' spectral efficiencies (expressed as $\log_2(1 + \text{SINR})$) in the simulated cellular network, for each RSM, several fairness constraints α (column-wise) and multiple beamforming gains (row-wise). Note that for each plot, the x-axis is in log-scale.

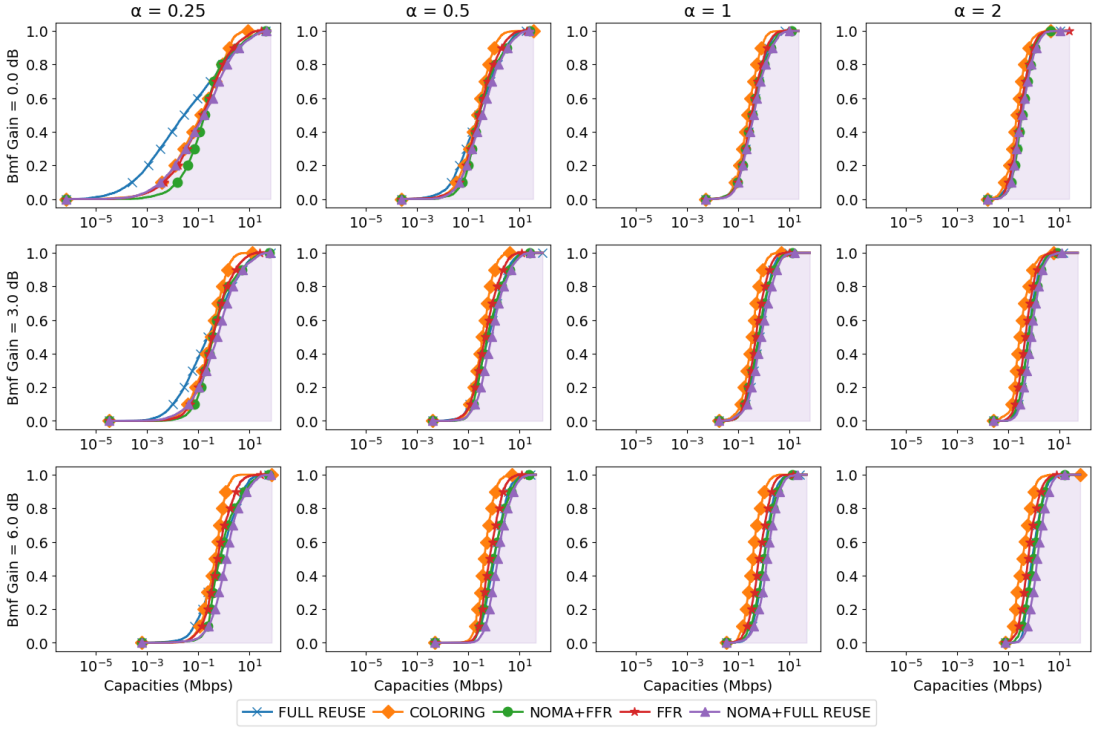

Fig. 4. The cumulative distribution functions of the UEs' Shannon capacities in the simulated cellular network (in Mbps), for each RSM, several fairness constraints α (column-wise) and multiple beamforming gains (row-wise). Note that for each plot, the x-axis is in log-scale.

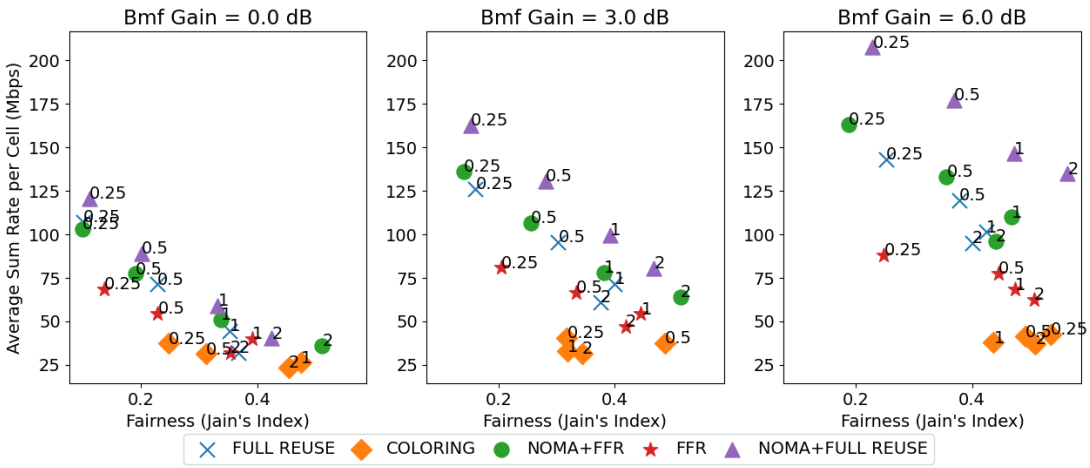

Fig. 5. The performance of each RSM in terms of sum rate and fairness (Jain). Each plot corresponds to a different value of the beamforming gain. By changing the value of α, Pareto fronts appear in the output space and allow the comparison between RSMs.

than what Figure 4 allows, we provide Table III. The 25%, 50% and 75% quantiles of the EUs' rates distributions are reported for each considered RSM, beamforming level and a fairness constraint $\alpha = 2$. The best quantile value among all RSMs is written in bold. `NOMA+FULL_REUSE` consistently appears as the RSM providing the best rate for all beamforming gains and all quantiles reported, sometimes by a considerable margin.

To further compare the RSMs in the quality of the rate-fairness tradeoff they achieve, we consider a 2-dimensional space, with the average sum rate per cell and the fairness (according to Jain's index [40]) as dimensions. We call it the *output space*. By varying α, one can compare how the achieved tradeoffs evolve when the incentive for a fair distribution of the rates changes. These trajectories draw Pareto fronts in the output space, and the concept of Pareto-dominance [41] can be used to rank the RSMs. Figure 5 depicts the performance of the RSMs in the output space, for several beamforming gains. Observe that the NOMA-based RSMs appear to find better tradeoffs than `FFR` and `COLORING` across all the considered values of α and all beamforming gains. They also appear equivalent to `FULL REUSE` when no beamforming gain is considered (except for $\alpha = 2$ where they manage to get better rates and fairness than all the other RSMs). With positive beamforming gains, `NOMA+FULL_REUSE` Pareto-dominates all the other RSMs (including `NOMA+FFR`), for almost all the considered values of α. In accordance to the CDFs observed in Figure 4 and to the values reported in Table III, `FFR` and `COLORING` obtain significantly lower performance metrics, while `FULL REUSE` gets slightly lower performance metrics than `NOMA+FFR`.

Finally, Figure 6 shows a subset of the cellular network corresponding to the blue rectangle in Figure 2, for the `NOMA+FULL_REUSE` mechanism. The impact of the fairness constraint α on the cells' geometry can be easily visualized: the larger the value of α, the bigger the inner regions of the cells. Note that, because the cells are depicted with the convex hulls of their associated UEs, some cells may overlap in the visualization.

Overall, our results suggest that NOMA combined with a full-reuse policy (denoted `NOMA+FULL_REUSE` in our experiments) consistently emerges as the RSM that finds the best tradeoff between a large sum of rates and a large fairness. This may be an important piece of information when designing the next-generation of cellular networks.

VI. Conclusion

In this paper, we proposed a general framework to fairly evaluate an arbitrary RSM in a complex, multi-cell context. Such evaluations cannot be handled by classical tools based on stochastic geometry, or on algorithmic designs, due to unrealistic knowledge requirements or prohibitive computational costs. On the opposite, our framework allows the evaluation of an arbitrary RSM by solving a difficult optimization problem under a minimal set of assumptions. To address this challenge, we leveraged recent advances in the black-box optimization literature by exploiting a distributed BO algorithm. The use of a BO algorithm ensures the desired versatility of our framework (*i.e.*, no specific assumptions about the cellular network or the interference model in use). In fact, the proposed evaluation framework could be paired to any simulator or implemented on a real-life network.

Using this evaluation framework, we compared five RSMs: full reuse, graph coloring, FFR, a combination of NOMA with FFR and a combination of NOMA with full reuse, under several levels of fairness (*i.e.* several values of α) and multiple beamforming gains on a real-life inspired scenario for the locations of BSs and UEs. Our results showed that the combination of NOMA with a full reuse policy is able to consistently find better tradeoffs (with respect to sum rate and fairness) than the other RSMs, sometimes by a comfortable margin. We believe that the proposed framework can be of interest to mobile network operators (MNOs) for comparing the efficiency of RSMs on their cellular network deployments. In particular, it is up to MNOs to decide whether the gains brought by NOMA are enough in light of its additional, though limited, complexity in the case of a 2-user NOMA.

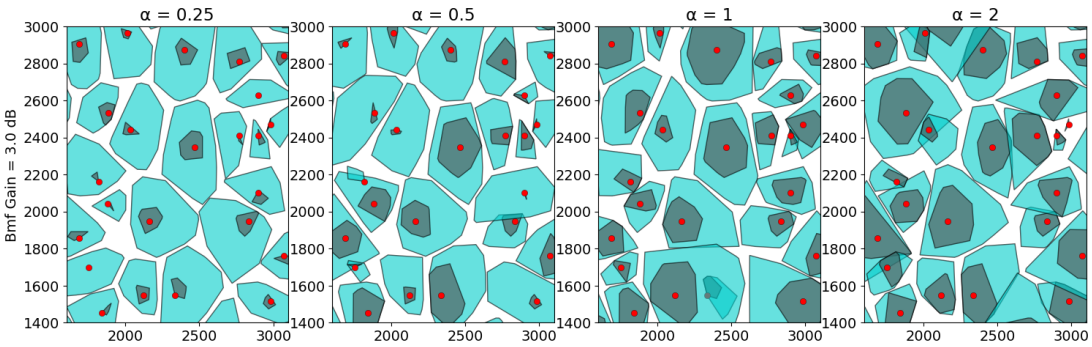

Fig. 6. A subset of the cellular network configured by the best configurations of NOMA+FULL_REUSE, found by INSPIRE for a beamforming gain of 3 dB and several values of α. Each BS is shown with a red dot. Each cell is depicted with the convex hull of its associated UEs. Finally, the shaded areas depict the inner regions of the cells.

As future work, we plan to extend our framework to handle dynamic scenarios where UEs are moving and subject to birth and death processes. The BO algorithm in use (denoted $\mathcal{A}$ in the previous sections) would need to be refined, in particular, to manage the potential staleness of previously collected data (e.g., via an elaborated sliding window technique). Compared to state-of-the-art algorithms, such an approach would allow the cells to breathe as a function of their local traffic load (an idea that was used for 3G networks), adapting over time the UEs/BSs association. This is an important feature to contribute to cell-free solutions with limited complexity, as currently investigated for 6G networks.

REFERENCES

[1] Zhikun Xu, Geoffrey Ye Li, Chenyang Yang, and Xiaolong Zhu. Throughput and optimal threshold for FFR schemes in OFDMA cellular networks. *IEEE Transactions on Wireless Communications*, 11(8):2776–2785, 2012.

[2] Mohamed Elwekeil, Masoud Alghoniemy, Osamu Muta, Adel B Abdel-Rahman, Haris Gacanin, and Hiroshi Furukawa. Performance evaluation of an adaptive self-organizing frequency reuse approach for OFDMA downlink. *Wireless networks*, 25:507–519, 2019.

[3] S Ezhilarasi and PTV Bhuvaneswari. Maximization of sum throughput in LTE heterogeneous network using region splitting-based resource partitioning scheme. *Wireless Personal Communications*, 121:905–938, 2021.

[4] Wonjae Shin, Mojtaba Vaezi, Byungju Lee, David J Love, Jungwoo Lee, and H Vincent Poor. Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges. *IEEE Communications Magazine*, 55(10):176–183, 2017.

[5] Refik Caglar Kizilirmak and Hossein Khaleghi Bizaki. Non-orthogonal multiple access (NOMA) for 5G networks. *Towards 5G Wireless Networks-A Physical Layer Perspective*, 83:83–98, 2016.

[6] Mojtaba Vaezi, Zhiguo Ding, and H Vincent Poor. *Multiple access techniques for 5G wireless networks and beyond*, volume 159. Springer, 2019.

[7] Souvik Sen, Naveen Santhapuri, Romit Roy Choudhury, and Srihari Nelakuditi. Successive interference cancellation: A back-of-the-envelope perspective. In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, pages 1–6, 2010.

[8] Jean-Marie Gorce, Philippe Mary, Dadja Anade, and Jean-Marc Kélif. Fundamental limits of non-orthogonal multiple access (noma) for the massive gaussian broadcast channel in finite block-length. *Sensors*, 21(3):715, 2021.

[9] Mojtaba Vaezi, Robert Schober, Zhiguo Ding, and H Vincent Poor. Non-orthogonal multiple access: Common myths and critical questions. *IEEE Wireless Communications*, 26(5):174–180, 2019.

[10] Kaidi Wang, Yuanwei Liu, Zhiguo Ding, Arumugam Nallanathan, and Mugen Peng. User association and power allocation for multi-cell non-orthogonal multiple access networks. *IEEE Transactions on Wireless Communications*, 18(11):5284–5298, 2019.

[11] Salah Elhoushy, Mohamed Ibrahim, and Walaa Hamouda. Cell-free massive mimo: A survey. *IEEE Communications Surveys & Tutorials*, 24(1):492–523, 2021.

[12] Zhaohui Yang, Cunhua Pan, Wei Xu, Yijin Pan, Ming Chen, and Maged Elkashlan. Power control for multi-cell networks with non-orthogonal multiple access. *IEEE Transactions on Wireless Communications*, 17(2):927–942, 2017.

[13] Subhankar Banerjee, Chung Shue Chen, Marceau Coupechoux, and Abhishek Sinha. Joint power and subcarrier allocation in multi-cell multi-carrier NOMA. In *2022 Thirteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 180–185. IEEE, 2022.

[14] Lou Salaün, Marceau Coupechoux, and Chung Shue Chen. Weighted sum-rate maximization in multi-carrier noma with cellular power constraint. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pages 451–459. IEEE, 2019.

[15] Lou Salaün, Marceau Coupechoux, and Chung Shue Chen. Joint subcarrier and power allocation in noma: Optimal and approximate algorithms. *IEEE Transactions on Signal Processing*, 68:2215–2230, 2020.

[16] Can Altay and Mutlu Koca. Fractional frequency reuse in non-orthogonal multiple access heterogeneous networks. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2018.

[17] Jeffrey G Andrews, François Baccelli, and Radha Krishna Ganti. A tractable approach to coverage and rate in cellular networks. *IEEE Transactions on communications*, 59(11):3122–3134, 2011.

[18] Jan García-Morales, Guillem Femenias, and Felip Riera-Palou. Analysis and optimization of ffr-aided ofdma-based heterogeneous cellular networks. *IEEE Access*, 4:5111–5127, 2016.

[19] Jan García-Morales, Guillem Femenias, and Felip Riera-Palou. On the design of ofdma-based ffr-aided irregular cellular networks with shadowing. *IEEE Access*, 6:7641–7653, 2018.

[20] Mohammad Salehi, Hina Tabassum, and Ekram Hossain. Meta distribution of sir in large-scale uplink and downlink noma networks. *IEEE Transactions on Communications*, 67(4):3009–3025, 2018.

[21] Lélio Chetot, Jean-Marie Gorce, and Jean-Marc Kélif. Fundamental limits in cellular networks with point process partial area statistics. In *2019 International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, pages 1–8. IEEE, 2019.

[22] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[23] Yi-An Ma, Tianqi Chen, and Emily Fox. A complete recipe for stochastic gradient mcmc. *Advances in neural information processing systems*, 28, 2015.

[24] Anthony Bardou and Thomas Begin. Inspire: Distributed bayesian optimization for improving spatial reuse in dense wlans. In *Proceedings of the 25th International ACM Conference on Modeling Analysis and Simulation of Wireless and Mobile Systems*, pages 133–142, 2022.

[25] James Bergstra, Daniel Yamins, and David Cox. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In *International conference on machine learning*, pages 115–123. PMLR, 2013.

[26] Daniel Lizotte, Tao Wang, Michael Bowling, and Dale Schuurmans. Automatic gait optimization with gaussian process regression. In *Proceedings of the 20th International Joint Conference on Artifical*

Intelligence, IJCAI'07, page 944–949, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc.
[27] Christopher K. I. Williams and Carl Edward Rasmussen. Gaussian processes for regression. In *Conference on Neural Information Processing Systems (NeurIPS'95)*, 1995.
[28] Joseph L Doob. The brownian movement and stochastic equations. *Annals of Mathematics*, pages 351–369, 1942.
[29] Marc G Genton. Classes of kernels for machine learning: a statistics perspective. *Journal of machine learning research*, 2(Dec):299–312, 2001.
[30] Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265, 2012.
[31] Shanti S Gupta and Klaus J Miescke. Bayesian look ahead one-stage sampling allocations for selection of the best population. *Journal of statistical planning and inference*, 54(2):229–244, 1996.
[32] Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.
[33] Jonas Mockus. Application of bayesian approach to numerical methods of global and stochastic optimization. *Journal of Global Optimization*, 4:347–365, 1994.
[34] Anthony Bardou, Patrick Thiran, and Thomas Begin. Relaxing the additivity constraints in decentralized no-regret high-dimensional bayesian optimization. *arXiv preprint arXiv:2305.19838*, 2023.
[35] Tinu Theckel Joy, Santu Rana, Sunil Gupta, and Svetha Venkatesh. Batch bayesian optimization using multi-scale search. *Knowledge-Based Systems*, 187:104818, 2020.
[36] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on networking*, 8(5):556–567, 2000.
[37] JHB Kemperman. On the shannon capacity of an arbitrary channel. In *Indagationes Mathematicae (Proceedings)*, volume 77, pages 101–115. North-Holland, 1974.
[38] Autorité de Régulation des Communications Électroniques des postes et de la distribution de la Presse (ARCEP). https://files.data.gouv.fr/arcep_donnees/mobile/sites/2022_T3/, 2022. [Online; accessed 27-July-2023].
[39] Jesper Møller, Anne Randi Syversveen, and Rasmus Plenge Waagepetersen. Log gaussian cox processes. *Scandinavian journal of statistics*, 25(3):451–482, 1998.
[40] Rajendra K Jain, Dah-Ming W Chiu, William R Hawe, et al. A quantitative measure of fairness and discrimination. *Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA*, 21, 1984.
[41] Mark Voorneveld. Characterization of pareto dominance. *Operations Research Letters*, 31(1):7–11, 2003.

APPENDIX
SPLITTING A CELL INTO TWO REGIONS

In this appendix, we describe how we compute the set of all regions $\mathcal{A}$, that is, how the UEs are associated with their BSs and distributed over the two regions of each cell.

First, recall that given a configuration $\boldsymbol{x}_t \in \mathcal{D}$, each UE associates with the BS that provides the best SINR.

Then, within each cell i, the optimal 2-partition of the set of associated UEs (*i.e.* the one that maximizes (10) within the cell) must be computed. Note that since each UE j knows its SINR with its BS, its achievable rate $c^{(j)}$ and, a fortiori, the objective function (10) within the cell can be computed. Because there exist $2^{m^{(i)}-1}-1$ partitions if the cell i contains $m^{(i)}$ UEs, using a greedy algorithm that explores all the partitions would be prohibitive.

To circumvent this combinatorial problem, all the associated UEs are sorted by SNR in descending order. Then, the set of UEs is partitioned based on a SNR threshold β: the UEs with a SNR larger than β are associated with the inner region, while the others are associated with the outer region of the cell. The objective function (10) is computed for all the $m^{(i)}-1$ 2-partitions of the list of ordered SNRs in order to determine β^*, which is the SNR threshold associated with the 2-partition that yields the largest value of (10). Figure 7 illustrates how β^* is computed.

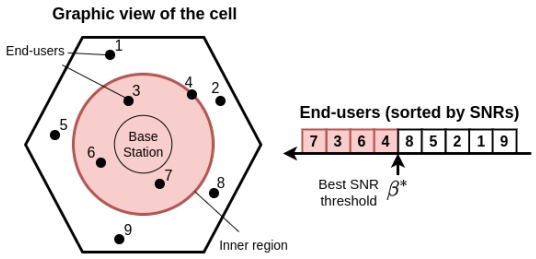

Fig. 7. An illustration of how a cell is split into two regions based on the SNRs of its associated UEs. The cell is shown as an hexagon, the inner region as a red circle and the outer region as the complementary of the circle in the hexagon. The BS is at the center of the cell, and each UE is shown as a black dot.

Anthony Bardou TODO: Insert Anthony's bio here

Jean-Marie Gorce TODO: Insert Jean-Marie's bio here

Thomas Begin is a Full Professor at University Lyon 1 in France since 2022. He holds a Ph.D. degree in Computer Science from University Sorbonnes at Paris. In 2009, he was a post-doctoral fellow at UC Santa Cruz. Between 2009 and 2022, he was an Associate Professor at University Lyon 1. During the academic year 2015-2016, he was on a research leave at the Univ. of Ottawa in Canada. His research interests are in performance evaluation, computer networks, clouds, and system modeling. His principal applications pertain to high-level modeling, wireless networks, resource allocation, and queueing systems. His work includes both theoretical and practical aspects. He has published over 100 papers in international refereed journals and conference proceedings and was awarded twice the best paper award. He received the ACM MSWiM Rising Star Award in 2020.