Online Learning for Interference Coordination in Heterogeneous Networks

Jose A. Ayala-Romero, Juan J. Alcaraz, Javier Vales-Alonso, Esteban Egea-López
Universidad Politécnica de Cartagena
TIC Department
Plaza del Hospital, 1, Cartagena, Spain
Email: {josea.ayala, juan.alcaraz, javier.vales, esteban.egea}@upct.es

***Abstract*—This paper focuses on interference coordination between the small cell and macro cell tiers of a wireless access network. We propose a new perspective based on a model-free learning strategy, not requiring any previous knowledge about the network (e.g., topology, interference graph, scheduling algorithms). Our approach is based on a stochastic optimization algorithm known as Response Surface Methodology, that we use to learn the optimal parameter configuration during network operation (online learning) adapting to changes on network conditions (e.g., traffic, user positions). The result is a simple, effective and flexible mechanism that outperforms previous proposals. As a case study we apply our scheme to the dynamic adjustment of LTE-A eICIC parameters (CRE bias and ABS ratio).**

***Index Terms*—Interference coordination, heterogeneous networks, small cells, online learning**

I. Introduction

Small cells are considered one of the key technologies in current and future mobile access networks, since they increase the spacial spectral reuse, enhance the network coverage, and reduce the load of the macro cells (offloading) [1]. In a typical scenario, multiple small cells overlay each macro cell, resulting in a multi-layer deployment often referred to as heterogeneous network (HetNet) [1]. With a typical frequency reuse factor of 1, interference management in HetNets is a crucial and challenging feature, especially in dense scenarios, and has been a hot research topic for the last few years. In this paper we propose a new strategy to address this problem. In contrast to previous works, our proposal is based on a model-free online learning formulation. To illustrate our approach, we describe its application to the enhanced Inter Cell Interference Coordination (eICIC) functionality of Long Term Evolution Advanced (LTE-A) networks.

The interference between macro cells (macro eNBs in 3GPP terminology) and small cells (pico eNBs) is determined by two eICIC parameters: *Cell Range Extension* (CRE) bias, which controls the proportion of User Equipments (UEs) offloaded from macro eNBs to pico eNBs; and *Almost Blank Subframe* (ABS) ratio, which indicates the proportion of subframes that the macro eNB *mutes* in its radio frame in order to reduce the interference at UEs associated to pico eNBs. The most usual limitation of previous works is neglecting the dynamic nature of the network [2]–[5]. Figure 1 shows the 5^{th} percentile throughput (a usual metric for LTE-A [2] [6] [7]) obtained by simulating a scenario defined according to 3GPP guidelines [8]. Both surfaces were obtained for the same network layout, but using two different traffic demands. The optimal configurations of (ABS ratio, CRE bias) for 10 Mbps and 75 Mbps per macro sector are $(3/8,\ 9\ \text{dB})$, $(4/8,\ 12\ \text{dB})$, respectively. Therefore, as traffic conditions change, the eICIC parameters should be dynamically reconfigured.

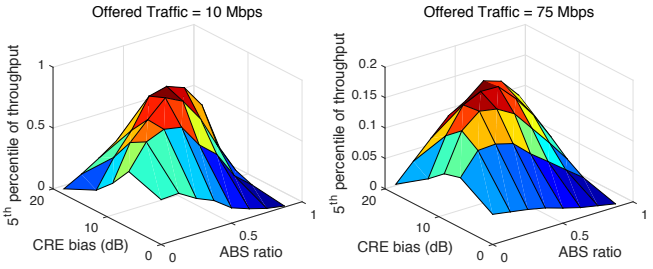

Fig. 1. 5^{th} percentile of throughput (in Mbps) as a function of eICIC parameters for two traffic intensities (given in Mbps per macro sector).

Motivated by this requirement, this paper proposes a mechanism based on a stochastic optimization algorithm called Response Surface Methodology (RSM) [9]–[11], which learns the optimum configuration online, adapting to the network changes. The rest of the paper is organized as follows. The related works and contributions are discussed in Section II. The problem is formulated in Section III, and Section IV presents our RSM-based mechanism. In Section V we describe the simulation model used to evaluate our proposal, and numerical results and benchmark comparisons are provided in Section VI.

II. Related work and contribution

A. Related Work

There is a broad literature on interference management for HetNets. However, we adopt a new perspective. We consider the *variable nature of the network* as a basic element of the problem formulation, in contrast to [2]–[5] that consider a static network situation, and do not address explicitly the real-time adaptation to dynamic network conditions (traffic intensity, number of users, user positions).

While some works address the problem considering dynamic network conditions as well [6], [12], their approach

		[6]	[2]	[3]	[4]	[12]	This work
Model	Dynamic/Static	Dynamic	Static	Static	Static	Dynamic	Dynamic
	Traffic Model	FTP	Full Buffer	Full Buffer	Full Buffer	Both	FTP
Algorithm	Variables	ABS	ABS, CRE	ABS, CRE	Power allocation, CRE	ABS	ABS, CRE
	Online/Offline	Online	Offline	Offline	Online	Online	Online
	Local/Global solution	Local	Global	Global	Local	Local	Global
	Method	Heuristic	Optimization	Optimization	Learning	Heuristic	Learning
Requirements	Topology information	Yes	Yes	Yes	Yes	Yes	No
	Interference graph	No	No	Yes	No	No	No

TABLE I
COMPARISON TO RELATED WORKS

differs to ours in some crucial aspects. First, they propose heuristic mechanisms that are very scenario-dependent. For example, although [6] performs well in the usual LTE-A scenarios for which it is designed, it is not adequate for highly dense scenarios. Second, these schemes only focus on optimizing the ABS ratio while, as shown in the introduction, both eICIC parameters have a joint impact on performance. Our mechanism can *optimize two or more parameters simultaneously*, *independently on the scenario*.

Therefore, we essentially follow a *model-free* optimization approach. This also contrasts with many previous works relying on *mathematical models* of the system [2] [3] [5]. For mathematical tractability, even the most complex models imply simplifications and assumptions that limit their application in real operating networks. For example, stochastic geometry [13] [14] is an excellent tool for high-level evaluation like determining the capacity bounds of a network. But it assumes that nodes are deployed according to spatial distributions like Poisson point process, while a fine-tuned parameter configuration should be done for the real network layout.

Finally, our approach requires low *overhead in terms of computation and signaling*. For example, some mechanisms require retrieving information about the network topology [2], and the interference graph of the network [3]. Our proposal only collects performance samples. Table I summarizes the characteristics of others works in comparison to ours.

B. Contribution

This paper proposes a novel self-organizing strategy for HetNets, capable of learning the optimal configuration of interference coordination parameters. Our mechanism is built upon a stochastic optimization algorithm known as Response Surface Methodology. The main novelty is the adoption of a model-free learning approach in interference management, not requiring any previous knowledge about the network topology, interference graph, radio propagation conditions, and so forth. As a result we obtain a simple, effective and adaptive algorithm that outperforms previous proposals. As a case study we apply our proposal to the dynamic adjustment of LTE-A eICIC parameters (CRE bias and ABS ratio).

III. MOTIVATIONAL SCENARIO AND PROBLEM FORMULATION

A. Application Framework: eICIC in LTE-A

The eICIC functionality was introduced by the 3GPP in Release 10 (LTE-A) and comprises two mechanisms for controlling the coexistence of pico and macro eNBs: CRE and ABS.

CRE allows the UEs to associate to a pico eNB when the Reference Signal Received Power (RSRP) from the pico eNB is lower than the RSRP from the macro eNB. This mechanism avoids that the lower transmission power of the pico eNBs cause the underuse of their resources. To select an eNB to associate with, the UE adds the CRE bias to the pico RSRP but not to the macro RSRP, and then selects the eNB with maximum (corrected) RSRP. Thus, the higher the CRE bias, the larger the downlink footprint of the pico eNBs. However, due to the high macro eNB interference, the UEs located at the extended region (CRE region) experience a poor Signal to Interference and Noise Ratio (SINR).

Hence, **ABS** is motivated by the need to improve the performance of UEs located at the CRE regions. This eICIC mechanism allows the macro eNBs to mute all the data symbols in some subframes, referred to as *Almost Blank Subframes*. In these protected subframes, the SINR of downlink pico transmission is notably improved because the macro interference is removed. The protected subframes are inserted following a periodic pattern lasting 8 subframes. Therefore, it is necessary to configure the ratio of ABS subframes over conventional subframes $(0/8, \ldots, 8/8)$ within the ABS pattern. We consider a cluster of macro eNBs sharing the same ABS pattern (synchronized muting) and the same CRE bias [15], [7], [14]. The cluster is assumed to cover a geographical area with a homogeneous traffic profile.

B. Stochastic Optimization Problem Formulation

Let us consider an operator who wants to optimize certain performance metric, referred to as F, depending on the interference coordination parameters of the system. In eICIC these parameters are the ABS ratio (γ) and CRE bias (ϕ). Because of the random nature of the system (number and position of the terminals, traffic intensity, propagation condition of the channels, etc), $F(\gamma, \phi)$ is in general a random variable. Therefore, our goal is to find a configuration γ and ϕ that maximizes $E\left[F(\gamma, \phi)\right]$.

Let $x = (\gamma, \phi)^T$ denote the eICIC configuration. For the ABS ratio we allow the system to randomize between contiguous ABS patterns, so that γ can take continuous values. The convex set $\mathcal{P} = \left\{x = (\gamma, \phi)^T : 0 \leq \gamma \leq 1, 0 \leq \phi \leq \phi_{\max}\right\}$ contains all the feasible values of x. We can formulate the main problem as

$$\max_{x \in \mathcal{P}} E\left[F(x)\right]. \quad (1)$$

We are assuming that the operator has full freedom to select the performance metric F. It can be associated to the physical layer (e.g. signal to interference and noise ratio, SINR) or to upper layers (e.g. data throughput). Given the amount and the complexity of the stochastic processes involved, a mathematical expression to estimate $E[F(x)]$ accurately for a real operating network is, in general, intractable.

C. Online Learning Approach

Because the objective function $E[F(x)]$ is unknown, our proposal relies on *learning* the optimal configuration by obtaining samples of $F(x)$ from the operating network itself. This data-driven approach is known as *online learning*. Let us reformulate (1) within this framework.

An online learning algorithm operates in countable time stages $k = 0, 1, \ldots$ At each stage k, the algorithm selects a configuration x_k and obtains an observation of the performance $y_k = F(x_k, \omega_k)$ where ω_k represents the (random) state of the system at stage k, comprising all the random variables affecting F (e.g. traffic demands, UEs positions, propagation conditions). The configuration x_k has to be selected according to the history of the system up to $k-1$, defined as the set of past configurations and observations: $\mathcal{H}_{k-1} = \{x_0, y_0, \ldots, x_{k-1}, y_{k-1}\}$. The objective, up to a given stage N, is to solve

$$\max_{x_0,\ldots,x_N \in \mathcal{P}} E\left[\sum_{k=0}^{N} F(x_k, \omega_k)\right] \tag{2}$$

Because no previous knowledge about the system is assumed, and ω_k is not included in the observations, the algorithm addressing the above problem follows a *black box* optimization approach. This type of algorithm needs to *explore* the objective function F, that is, to select configurations x_k that might not be considered as the best option given the history of the system, but that could provide useful knowledge to find even better configurations. Balancing exploitation (i.e. selecting the best option given the current knowledge) and exploration, is a central issue in online learning algorithms.

In our case, we face an additional challenge. Note that w_k is a stochastic process that is not stationary in general. This implies that, even if the learning algorithm eventually converges to an optimal configuration x^*, this configuration would not be optimal in future stages because the statistical conditions for which x^* is optimal are no longer present. Therefore, the algorithm should be able to adapt the system configuration to long term changes in user activity.

IV. Stochastic Optimization Algorithm

A. Response Surface Methodology

Let us consider an algorithm to find the optimal solution of the original problem (1) by means of a stochastic steepest-ascent search. Such type of algorithm operates iteratively, with iterations denoted by $n = 0, 1, 2, ...$, generating a sequence of configurations $\{x^{(n)}\} = x^{(0)}, x^{(1)}, \ldots x^{(n)}$, that should converge to an optimal solution as $n \to \infty$. In particular, starting from an initial point $x^{(0)}$, the sequence $\{x^{(n)}\}$ is generated according to

$$x^{(n+1)} = x^{(n)} + \alpha d(x^{(n)}) \tag{3}$$

where $d(x^{(n)})$ is the steepest-ascent search direction at point $x^{(n)}$, and α is the step-size weighting factor.

Let us define $f(x) = E[F(x)]$. When f is a differentiable function, the steepest-ascent direction at point $x^{(n)}$ is given by the gradient of f, i.e. $d(x^{(n)}) = \nabla f(x^{(n)})$. However, since f is unknown we can neither assess if it is differentiable nor compute $\nabla f(x^{(n)})$. We can overcome this issue by generating an approximate steepest ascent direction $\hat{d}(x(n))$ such that the iteration (3) provides, with high probability, better configurations as n increases. In order to compute $\hat{d}(x(n))$, the algorithm must take samples of the performance metric during network operation. Response Surface Methodology (RSM) [9]–[11] provides an efficient way to sampling the system and computing $\hat{d}(x(n))$. In this section we detail how we can use RSM as an online learning algorithm.

Let k denote the number of stages that the system has been operating when the gradient-ascent algorithm obtains the n-th iterate $x^{(n)}$ (note that, in general, it takes several stages to perform one gradient-ascent iteration). Therefore, $x^{(n)}$ is the outcome of the learning algorithm given the current system history $\mathcal{H}_k = \{x_0, y_0, \ldots, x_k, y_k\}$.

In order to obtain next iterate $x^{(n+1)}$ using (3), RSM generates a local approximation of the objective function f at the vicinity of $x^{(n)}$. The approximation consists of a linear model characterized by a parameter vector θ. For $x \in \mathbb{R}^2$, the parameter vector is $\theta = (\theta_0, \theta_1, \theta_2)^T$, and the approximation of f is given by

$$\hat{f}(x|\theta) = \theta_0 + x^T \theta_{-0} \tag{4}$$

where $\theta_{-0} = (\theta_1, \theta_2)^T$.

Given $x^{(n)}$, the RSM mechanism computes the corresponding parameter vector $\theta^{(n)}$ according to the following procedure:

1) Select a set $\mathcal{D}^{(n)}$ of *design points* in the vicinity of $x^{(n)}$.
2) Take a (possibly random) number p of samples using the points in $\mathcal{D}^{(n)}$. The set of samples $\mathcal{S}^{(n)}$ comprises points taken from $\mathcal{D}^{(n)}$ and their corresponding observations: $\mathcal{S}^{(n)} = \{x_{k+1}, y_{k+1}, \ldots, x_{k+p}, y_{k+p}\}$. The set $\mathcal{S}^{(n)}$ is therefore the incremental history between two consecutive gradient-ascent iterations, $\mathcal{S}^{(n)} = \mathcal{H}_{k+p} - \mathcal{H}_k$.
3) Using $\mathcal{S}^{(n)}$, construct matrix $W^{(n)}$ and vector $y^{(n)}$ as follows

$$W^{(n)} = \begin{pmatrix} 1 & x_{k+1}^T \\ 1 & x_{k+2}^T \\ \vdots & \vdots \\ 1 & x_{k+p}^T \end{pmatrix}, \quad y^{(n)} = \begin{pmatrix} y_{k+1} \\ y_{k+2} \\ \vdots \\ y_{k+p} \end{pmatrix}. \tag{5}$$

4) Compute the least squares estimation of $\theta^{(n)}$:

$$\theta^{(n)} = \left(\left(W^{(n)}\right)^T W^{(n)}\right)^{-1} \left(W^{(n)}\right)^T y^{(n)}. \tag{6}$$

Once $\theta^{(n)}$ is obtained, iteration (3) is applied with the following estimation of the steepest-ascent direction

$$\hat{d}(x^{(n)}) = \nabla \hat{f}(x^{(n)}|\theta^{(n)}) = \theta_{-0}^{(n)}. \quad (7)$$

B. Sampling Strategy

Let us see how to determine, at each iteration, the set $\mathcal{D}^{(n)}$ and the number of samples p with statistical guarantees. Since we aim at approximating f at current iterate $x^{(n)}$, the design points should be small perturbations of $x^{(n)}$. Let $\delta_i > 0$ be the size of the perturbation on the i-th component of $x^{(n)}$. For low-dimensional vectors, we can use a *full factorial design* [10], which implies considering all possible combinations of (positive and negative) perturbations on each component of $x^{(n)}$. In particular, for $x^{(n)} \in \mathcal{P} \subset \mathbb{R}^{(2)}$, we have the following set

$$\begin{aligned} \mathcal{D}^{(n)} \quad = \quad & \{x^{(n)}, \\ & x^{(n)} + (\delta_1, \delta_2)^T, \\ & x^{(n)} + (-\delta_1, \delta_2)^T, \\ & x^{(n)} + (\delta_1, -\delta_2)^T, \\ & x^{(n)} + (-\delta_1, -\delta_2)^T\} \end{aligned} \quad (8)$$

Thus, $\delta = (\delta_1, \delta_2)$ defines the local area where the first order approximation will be fitted. This area should be sufficiently large to show changes in the response, but not too large so as to provide biased gradient estimations. In our case study, δ configuration is based on experience (previous works and simulations) [9] [10]. An automatic adjustment of δ and α is left for further study.

If we take the same number of samples from each point in $\mathcal{D}^{(n)}$, the columns of $W^{(n)}$ will be orthogonal and will have the same number of positive and negative perturbations. This property guarantees that $\left(W^{(n)}\right)^T W^{(n)}$ in (6) is invertible.

The search direction is determined by the regression coefficients $\theta^{(n)}$ obtained from noisy samples of f. We can characterize statistically these coefficients to evaluate the accuracy of the search direction. Specifically, given $\theta^{(n)}$, we can obtain the angle φ within which the real steepest-ascent direction is contained with a given confidence degree C. The algorithm can take samples until φ is smaller than a desired target angle φ^*. We propose to use φ^* and C as design parameters that dynamically determine the amount of samples p taken during each iteration n.

Let $d = |\mathcal{D}^{(n)}|$ (= 5 in our case study). After obtaining $x^{(n)}$, the algorithm takes $p = d$ samples, one at each point in $\mathcal{D}^{(n)}$, builds $W^{(n)}$ and $y^{(n)}$ with these samples, and then uses (6) to obtain $\theta^{(n)}$. The variance of $\hat{f}(x|\theta^{(n)})$ is estimated by

$$\left(\hat{\sigma}^{(n)}\right)^2 = \frac{1}{p-2}\|y^{(n)} - W^{(n)}\theta^{(n)}\|^2 \quad (9)$$

and the covariance matrix of $\theta^{(n)}$ is given by

$$\hat{\chi}^{(n)} = \left(\left(W^{(n)}\right)^T W^{(n)}\right)^{-1} \left(\hat{\sigma}^{(n)}\right)^2. \quad (10)$$

Here, the j-th element on the diagonal of $\hat{\chi}^{(n)}$ is the variance of the j-th element of $\theta^{(n)}$ [16]. Let us denote by s_θ^2 the maximum variance of θ_{-0}.

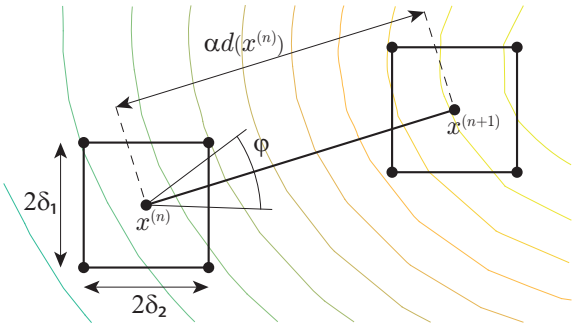

Fig. 2. Illustration of one steepest-ascent iteration of RSM with confidence in the search direction.

In order to obtain the angle φ for a confidence degree C, we can use the following expression from [9]

$$\frac{\varphi}{2} = \arcsin\left(\frac{s_\theta^2 \mathcal{F}_C(1, p-2)}{\sum_{j=1}^{2}\left(\theta_j^{(n)}\right)^2}\right)^{1/2} \quad (11)$$

where $\mathcal{F}_C(1, p-2)$ is the C-th percentile of the F-distribution.

Therefore, if $\varphi > \varphi^*$, then d new samples are taken from each point in $\mathcal{D}^{(n)}$ and incorporated to $W^{(n)}$and $y^{(n)}$, p is updated by $p \leftarrow p + d$, and $\theta^{(n)}$ is recomputed with (6). The resulting RSM algorithm with confidence in the search direction is shown in Algorithm 1, and illustrated in Figure 2. Note that, when $x^{(n)}$ is close to the optimum, $||\theta_{-0}^{(n)}||$ approaches 0 and subsequent iterations remain within the region of the optimum (exploitation) while still monitoring the gradient. When the optimum drifts to a different location, then $||\theta_{-0}^{(n)}||$ increases, making the algorithm re-start learning.

Algorithm 1 RSM with confidence in the search direction

```
Select initial point x^(0)
for n = 0, 1, 2, ... do
    Generate D^(n), set p = 0, and set φ to a value greater than φ*
    while φ > φ* do
        Obtain one sample at each point of D^(n) (d new samples)
        Insert the samples in W^(n) and y^(n)
        p ← p + d
        Estimate θ^(n) using (6)
        Estimate φ using (9), (10), and (11)
    end while
    x^(n+1) = x^(n) + αθ_{-0}^(n)
end for
```

V. Description of the Simulated Scenario

This section describes the LTE-A scenario where we have evaluated our mechanism. It is based on the 3GPP guidelines [8]. Regarding the *access network deployment and physical aspects* the scenario consists of a hexagonal grid of 19 three-sectorial macro eNBs (120 degrees per sector), that is, 57

directional macro cells. All the macro eNB sectors contain the same number of pico eNBs. These pico eNBs are uniformly distributed within each sector. We simulate the 7 central macro eNBs (21 sectors) while the external ones act as an inter-cell interference source, replicating the effect of a larger network. The inter-site distance is 500 m and the total bandwidth of the system is 10 MHz. Table II summarizes the remaining physical parameters.

The *traffic model* is defined by the following elements:

- At the user level we apply the FTP traffic model [8]. According to this model each incoming user has to download one file and stays in the system until this file is fully downloaded. The file size is 0.5 Mbytes.
- Users arrive according to a Poisson process. The parameter λ provides the average number of arrivals per second (UE/s or equivalently files/s).
- Each incoming user is dropped either at a uniformly distributed position over the macro eNB area, or in a cluster around a pico eNB. The first case occurs with probability P_{macro}, and the second one with $P_{\text{pico}} = 1 - P_{\text{macro}}$.

Although Proportional Fair (PF) is used as the *scheduling policy* in some related works [6] [4] [7], other works like [17], have shown that PF is not the best choice when eICIC is used, because ABS subframes should be preferably assigned to users in the CRE regions, while PF cannot discriminate among UEs. For this reason we adopt the modification proposed in [17] for eICIC.

VI. Numerical Results

A. Learning Rate and Adaptability

Let us consider a *plug-and-play* situation where the learning algorithm starts to operate with an initial eICIC configuration $x^{(0)}$, possibly far from the optimal one. The algorithm parameters are set to $\delta = (0.035, 0.5)$, $\alpha = 0.035$, $\varphi^* = 8^{\text{o}}$, $C = 0.95$. The performance metric y_k is the 5^{th} percentile throughput. Figure 3 shows the eICIC parameters (γ and ϕ) generated by the algorithm over consecutive steepest-ascent iterations, and the evolution of the average performance obtained with a sliding window of 60 samples. Around the average performance we depict the region containing 90% of the performance samples. The width of this region reflects the variance of the samples and the effects of the exploration of the algorithm.

We see that the algorithm approaches the optimum performance in roughly 25 steepest-ascent iterations. The number of samples per iteration depends directly on the setting of φ^* and C. For the selected values ($\varphi^* = 8^{\text{o}}$, $C = 0.95$), each steepest-ascent iteration needs 17.5 samples on average. For $\varphi^* = 15^{\text{o}}$ this number reduces to 6.6, although less precision on the angle estimation could increase the number of required iterations.

To illustrate how the algorithm adapts to network variations, Figure 4 shows a change on the traffic conditions from $\lambda = 18.75$ UE/s, $P_{\text{pico}} = 9/10$ to $\lambda = 7.5$ UE/s, $P_{\text{pico}} = 1/10$, the corresponding adaptation of $x^{(n)}$, and the evolution of the performance measurements.

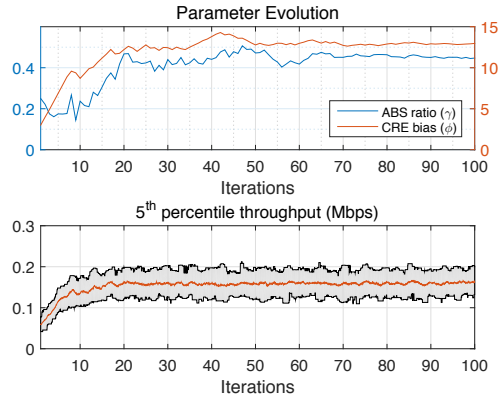

Fig. 3. Evolution of the configuration vector $x^{(n)}$ and the performance measures y_k for a system with $\lambda = 15$ UE/s, 4 pico eNB per macro sector, and $P_{\text{pico}} = 2/3$.

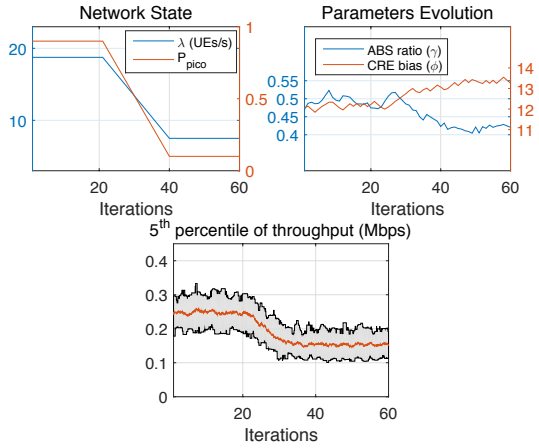

Fig. 4. Algorithm response under variations on traffic conditions (4 pico eNB per macro sector).

B. Comparison to benchmark solutions

We consider the following three benchmarks:

- **No eICIC**. The eICIC functionality is simply switched off in the network. This benchmark is intended to illustrate the necessity of this feature and provides a baseline reference for eICIC self-configuration schemes.
- **Optimal configuration for fixed traffic conditions**. We consider the case where the operator configures its eICIC parameters assuming static traffic conditions, following for example a worst-case approach or assuming the most usual situation.
- **Proportional-Fair-Based ABS (PF-ABS)**. This state-of-the-art heuristic proposed in [6] modifies γ according to the average of the Proportional Fair indexes at each eNB. This average, updated on a frame-by-frame basis, is used as an indicator of whether the eNB performance is improving or deteriorating, allowing PF-ABS to equalize the performance by redistributing the resources (subframes) among the macro and pico eNBs. In contrast to our mechanism, PF-ABS does not include CRE bias control.

Frame structure	Subframe 1 ms, Protected-subframe pattern 8 ms, Frame 10 ms
Transmit power	Macro eNB: 46 dBm, pico eNB: 30 dBm
Macro eNB antenna	Pattern: $A_H(\phi) = -\min\left[12\left(\frac{\phi}{70^0}\right)^2, 25dB\right]$, Gain: 14 dBi
Pico eNB antenna	Pattern: Omindirectional, Gain: 5 dBi
Macro to UE path loss	128.1 + 37.6 · $log_{10}(R[\text{Km}])$ where R is the macro eNB to UE distance
Pico to UE path loss	149.7 + 36.7 · $log_{10}(R[\text{Km}])$ where R is the pico cell to UE distance
Shadow fading	Lognormal distribution with 10 dB standard deviation

TABLE II
SIMULATION PARAMETERS

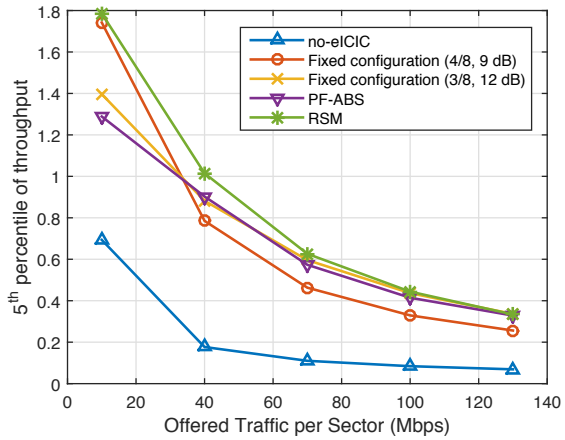

Fig. 5. 5^{th} percentile throughput as a function of offered traffic load for the proposed benchmarks with 10 pico eNBs per macro sector, and $P_{\text{pico}} = 2/3$

For the sake of comparison we assume that PF-ABS operates with optimal CRE.

Figure 5 shows the average performance for traffic intensities ranging from 10 to 130 Mbps per macro sector. One of the reasons why our scheme outperforms PF-ABS is that in PF-ABS each macro eNB has its own ABS pattern (unsynchronized muting) as a result of their local PF indexes, while our mechanism maintains a global ABS pattern (synchronized muting). Under unsynchronized muting, the ABS subframes present higher interference from other macro eNBs which may deteriorate the performance of the UEs located at CRE regions. For higher P_{pico} or more pico eNBs per sector (dense scenarios), this effect becomes more relevant because more users fall in the CRE regions.

VII. CONCLUSIONS

We have presented a novel perspective to address interference coordination between macro and small cells in HetNets. Our premises are that a real operating network is too complex to predict its performance accurately, and its response to control parameters changes over time. Therefore, we propose a model-free scheme capable of learning the optimal configuration online. Because our approach is data-driven instead of model-based, our mechanism can optimize any performance metric independently of the network layout and state. We have evaluated our scheme for the self-configuration of LTE-A eICIC parameters, obtaining promising results.

VIII. ACKNOWLEDGEMENTS

This work was supported by project grant MINECO/FEDER AIM TEC2016-76465-C2-1-R. Jose Antonio Ayala-Romero also acknowledges personal grant FPU14/03701.

REFERENCES

[1] J. Acharya, L. Gao, and S. Gaur, *Heterogeneous Networks in LTE-advanced*. John Wiley & Sons, 2014.
[2] S. Deb, P. Monogioudis, J. Miernik, and J. P. Seymour, "Algorithms for enhanced inter-cell interference coordination (eicic) in lte hetnets," *IEEE/ACM Transactions on Networking (TON)*, vol. 22, no. 1, pp. 137–150, 2014.
[3] A. Liu, V. K. Lau, L. Ruan, J. Chen, and D. Xiao, "Hierarchical radio resource optimization for heterogeneous networks with enhanced inter-cell interference coordination (eicic)," *Signal Processing, IEEE Transactions on*, vol. 62, no. 7, pp. 1684–1693, 2014.
[4] M. Simsek, M. Bennis, and I. Guvenc, "Learning based frequency-and time-domain inter-cell interference coordination in hetnets," *IEEE Trans. Vehicular Technol*, 2015.
[5] A. Bin Sediq, R. Schoenen, H. Yanikomeroglu, and G. Senarath, "Optimized distributed inter-cell interference coordination (icic) scheme using projected subgradient and network flow optimization," *Communications, IEEE Transactions on*, vol. 63, no. 1, pp. 107–124, 2015.
[6] B. Soret, K. Pedersen *et al.*, "Centralized and distributed solutions for fast muting adaptation in lte-advanced hetnets," *Vehicular Technology, IEEE Transactions on*, vol. 64, no. 1, pp. 147–158, 2015.
[7] B. Soret and K. I. Pedersen, "Macro transmission power reduction for hetnet co-channel deployments," in *Global Communications Conference (GLOBECOM), 2012 IEEE*. IEEE, 2012, pp. 4126–4130.
[8] 3GPP TR 36.814, "Evolved universal terrestrial radio access (eutra); further advancements for e-utra physical layer aspects," 3rd Generation Partnership Project (3GPP), Tech. Rep., 2010.
[9] R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook, *Response surface methodology: process and product optimization using designed experiments*. John Wiley & Sons, 2016.
[10] J. P. Kleijnen, *Design and analysis of simulation experiments*. Springer, 2008, vol. 20.
[11] K. Marti, *Stochastic optimization methods*. Springer, 2005.
[12] M. Al-Rawi, J. Huschke, and M. Sedra, "Dynamic protected-subframe density configuration in lte heterogeneous networks," in *Computer Communications and Networks (ICCCN), 2012 21st International Conference on*. IEEE, 2012, pp. 1–6.
[13] Y. Wang, H. Ji, and H. Zhang, "Spectrum-efficiency enhancement in small cell networks with biasing cell association and eicic: An analytical framework," *International Journal of Communication Systems*, vol. 29, no. 2, pp. 362–377, 2016.
[14] M. Cierny, H. Wang, R. Wichman, Z. Ding, and C. Wijting, "On number of almost blank subframes in heterogeneous cellular networks," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 10, pp. 5061–5073, 2013.
[15] 3GPP R1-100142, "System performance of heterogeneous networks with range expansion," 3rd Generation Partnership Project (3GPP), Tech. Rep., 2010.
[16] S. A. V. D. Geer, "Least squares estimation," *Encyclopedia of Statistics in Behavioral Science*, vol. 2, pp. 1041–1045, 2005.
[17] K. Min and J. So, "Scheduling and positioning for the expanded region of an indoor cell in heterogeneous networks," in *Indoor Positioning and Indoor Navigation (IPIN), 2014 International Conference on*. IEEE, 2014, pp. 685–692.