Dynamic Enhanced Inter-Cell Interference Coordination for Realistic Networks

Pedersen, Klaus I.; Alvarez, Beatriz Soret; Barcos, Sonia; Gerardino, Guillermo Andrés Pocovi; Wang, Hua

Published in:
IEEE Transactions on Vehicular Technology

DOI (link to publication from Publisher):
10.1109/TVT.2015.2451212

Publication date:
2016

Document Version
Accepted author manuscript, peer reviewed version

Link to publication from Aalborg University

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
? You may not further distribute the material or use it for any profit-making activity or commercial gain
? You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: November 07, 2019
Dynamic Enhanced Inter-Cell Interference Coordination for Realistic Networks

Klaus I. Pedersen, Member IEEE, Beatriz Soret, Member IEEE, Sonia Barcos, Guillermo Pocovi, Student Member IEEE, Hua Wang, Member IEEE

Abstract—Enhanced Inter-Cell Interference Coordination (eICIC) is a key ingredient to boost the performance of co-channel Heterogeneous Networks (HetNets). eICIC encompasses two main techniques: Almost Blank Subframes (ABS), during which the macro cell remains silent to reduce the interference, and biased user association to offload more users to the picocells. However, its application to realistic irregular deployments opens a number of research questions. In this paper, we investigate the operation of eICIC in a realistic deployment based on three-dimensional data from a dense urban European capital area. Rather than the classical semi-static and network-wise configuration, the importance of having highly dynamic and distributed mechanisms that are able to adapt to local environment conditions is revealed. We propose two promising cell association algorithms: one aiming at pure load balancing and an opportunistic approach exploiting the varying cell conditions. Moreover, an autonomous fast distributed muting algorithm is presented, which is simple, robust, and well suited for irregular network deployments. Performance results for realistic network deployments show that the traditional semi-static eICIC configuration leads to modest gains, whereas the set of proposed fast dynamic eICIC algorithms result in capacity gains on the order of 35-120% depending on the local environment characteristics. These attractive gains together with the simplicity of the proposed solutions underline the practical relevance of such schemes.

Index Terms—Heterogeneous Networks, eICIC, Almost Blank Subframes, muting ratio, load balancing, realistic network scenarios.

I. INTRODUCTION

Heterogeneous Networks (HetNets) with co-channel deployment of high-power macrocells and low-power picocells using Long Term Evolution (LTE) are the focus of this paper [1] [2]. Although LTE HetNets offer attractive benefits, they also come with a number of challenges that must be carefully addressed in order to fully unleash the potential of such network deployments. Inter-cell co-channel interference is an old acquaintance for cellular networks, with increased significance for co-channel HetNet scenarios. In particular, the downlink interference from high-power macrocells is known to be a potential strong aggressor for users served by low-power picocells (victim users). If left untreated, the macrocell interference can seriously limit the performance benefits of the picocells.

To address the outlined interference problem, the Third Generation Partnership Project (3GPP) has standardized Enhanced Inter-Cell Interference Coordination (eICIC) for LTE networks. In a nutshell, the fundamental principle of eICIC is to perform muting of certain subframes at the macro-layer to reduce interference towards the pico-users. One of the main advantages of eICIC is that it offers increased flexibility for load balancing between macros and picos. Both features – macro muting and load balancing – should therefore be jointly addressed.

A. Related studies

Despite the relatively short time since the introduction of eICIC in 3GPP for LTE, the open literature includes an impressive number of related studies, so we only provide pointers to some of those in the following. An overview of the eICIC scheme is available in [3] - [5]. Performance results of eICIC for the standard 3GPP simulation scenarios are presented in [6] and [7] for uniform and non-uniform deployments, under the assumption of different commonly accepted stochastic radio propagation models. In addition to network-based coordination of subframe muting at the macro-cells, the eICIC framework also includes explicit terminal assistance in the form of dedicated measurement feedback and non-linear interference cancellation of certain physical-layer control channels and reference signals [8] [9]. Aspects of scheduling resource allocation are reported in [10] and [11]. Being an enabler for more extensive macro-pico load balancing, it is also worth highlighting the numerous studies related to user association or biased cell selection – see e.g. [12] [13]. The load balancing algorithms are typically considered part of the set of self-organizing network (SON) features [14], and hence assumed to operate in a semi-static fashion on a slow time scale. In [15], finding the optimal macrocell muting pattern is formulated as a utility maximization problem, and dynamic programming is used to solve it. In [16], the joint optimization of the muting configuration and assignment of users to macrocells and picocells is studied as a network-wise utility maximization problem, subject to using same muting pattern on all macrocells. As assumed for the majority of eICIC studies and also suggested in [17], first implementations of eICIC will likely adjust the muting configuration and load balancing on time-scales of hours or days based on collection of the networks coverage map and spatio-temporal traffic density profile. For example, a centralized SON-based scheme...
for eICIC parameter optimization was presented in [18], along with demonstration of the benefits by using input from a radio network planning tool and data from a HetNet deployment in New York City. To the best of our knowledge, only two recent studies propose using fast eICIC muting adaptation in coherence with rapid traffic variations to further boost the performance [19] [20].

B. Our contribution

The contribution of this article is a dynamic framework with fast adjustment of the key performance parameters of eICIC, as well as realistic performance evaluation using an advanced network model. To be more specific, the proposed solution encompasses two steps:

- A simple and robust algorithm for fast dynamic adjustment of the macrocell muting that aims at tracking fast variations in the traffic load. The solution is an extension of the schemes in [19] [20]. The proposed solution is fully distributed, where each of the macrocells perform autonomous muting decisions on a fast basis. The muting decisions by the macrocells are based on enhanced load measures from the picos that include the victim users having a given macrocell as their aggressor.
- Two different algorithms for user serving cell selection; one simple load-aware algorithm and an opportunistic rule. This is in contrast to the majority of eICIC studies that assume a simple power-based user cell association scheme with a single common value of cell range extension per picocell.

The study is conducted for LTE with Frequency Division Duplexing (FDD). It is shown that our solution offers significant benefits over schemes known from the open literature (Section I-A). The rationale behind these gains is that the proposed framework is able to more efficiently self-adjust in coherence with local environment conditions, and is robust to various imperfections and time-variant behavior of the system. A realistic network deployment is selected for the performance evaluation, based on three-dimensional (3D) topography map data from a dense urban European capital area. This ensures a high degree of realism and practical relevance of the results. Environment features such as buildings, streets, open squares, etc. are explicitly represented in the network model, and naturally influencing the radio propagation characteristics and experienced interference. Furthermore, the network model includes a time-variant traffic behavior with dynamic arrival and departure of users, as well as non-uniform spatial traffic distribution. The use of such a realistic network model forms a solid basis for evaluating if the proposed dynamic algorithms for eICIC parameter adaptation are attractive for implementation in the field. The complexity of the network model prevents the derivation of purely analytical expressions without discarding important performance-determining factors, so system level simulations are chosen for the performance assessment. The underlying simulation model assumptions are carefully addressed and validated in coherence with fundamental research methodology principles to ensure trustworthy and statistically reliable performance results. In essence, our reported findings shed additional light on the required network algorithms to harvest the full potential of eICIC in a practical setting.

C. Outline of the paper

The rest of the paper is organized as follows: Section II provides a background on eICIC, while Section III outlines the network model and considered performance metrics. The algorithms for fast dynamic adjustment of muting patterns and load balancing are presented in Section IV. Section V presents the corresponding performance results, followed by related discussions in Section VI. Concluding remarks appear in Section VII.

II. BACKGROUND: eICIC

eICIC is a time-domain resource division scheme designed to lower the co-channel interference from an aggressor macrocell towards victim users served by the picocells. This is achieved by assuming a time-synchronized network, where certain subframes of 1 ms duration at the macro-layer are muted. The muted subframes are denoted Almost Blank Subframes (ABS), and characterized by having no data transmission, while transmission of system information and common reference signals (CRS) is still allowed [4]. As shown in Figure 1, a macrocell transmitting ABS is not able to schedule any of its users during that subframe. During ABS transmission, it is possible for the picocell to serve distant users, which would suffer too high interference from the high-power macrocell(s) during normal macro transmission. Therefore, the use of eICIC enables enlargement of the footprint of the picocells, allowing the offloading of more users to the pico layer. In the classical configuration of eICIC, the pattern of normal subframes and ABS is periodically repeated\(^3\), where \(0 \leq \beta \leq 1\) expresses the muting ratio. For example, \(\beta = 0.5\) means that the macrocell uses ABS in half of the subframes as shown in Figure 1.

Due to the toggling between normal transmissions and ABS at the macrocells, users served by the picocells are exposed to relatively severe time-variant interference fluctuations, which could make it problematic to perform accurate link adaptation and radio channel aware scheduling decisions. In order to overcome this challenge, the eICIC concept allows the network to configure users with time-domain measurement restrictions of channel state feedback in coherence with the muting pattern of ABS at the macrocells. Furthermore, during the course of eICIC development, 3GPP also standardized non-linear interference cancellation of residual interference from ABS for pico-users [8]- [9]. For additional concept details on the 3GPP standardized eICIC scheme, we refer to the following overview papers [3]- [5].

\(^3\)Notice that although the LTE frame length counts 10 subframes, the ABS muting pattern is often set to be periodic every 8 subframes to protect uplink hybrid automatic repeat request operation as described in [4].
terminal measurements of received power levels [33]. Thus, the selected serving cell for a user is

$$i^* = \arg \max_i \{ P_i + RE_i \}, \tag{1}$$

where $P_i$ is the Reference Signal Received Power (RSRP) from cell $i$ (in dBm) while $RE_i$ is the Range Extension (RE) for cell $i$ (in dB). Assuming $RE_i = 0$ dB for macrocells and $RE_i \geq 0$ dB for picocells, the setting of the RE parameter essentially determines the inter-layer load balancing between macrocells and picocells, while neglecting potential benefits of intra-layer load balancing between cells of the same type. Without eICIC, it is typically only advisable to use values of the RE offset for the picocells of up to approximately 3 dB, while eICIC and ABS at the macro-layer allows using RE values of 10-15 dB. However, a picocell with a large RE value is typically only able to schedule its users in the extended cell range area during subframes where the potential strong interfering macrocell is using ABS. Hence, there exists an inherent dependency between the best setting of RE and ABS muting ratio, and thus user association and eICIC are intimately related. The value of the RE offset, which is signalled to the UEs using the Radio Resource Control (RRC) protocol, is adjusted on a relatively slow time scale in order to avoid excessive RRC signaling overhead at the air-interface [21].

In this study we consider the following cases for coordinated load balancing and eICIC parameter adaptation:

- Network-wise semi-static: The same ABS muting pattern is applied to all macrocells, as well as the same value of RE for all picocells, assuming serving cell selection according to (1). The ABS muting pattern and RE are adjusted on a slow time scale to maximize the system performance.
- Area-wise semi-static: The network is divided into subareas, and the ABS muting pattern is semi-statically adjusted per subarea, subject to using the same muting pattern at all the macrocells and the same RE for the picocells belonging to the same subarea.
- Network-wise RE and fast ABS: Cell selection is performed according to (1), assuming the same setting of RE for all the picos. The value of the RE is semi-statically adjusted. Macrocells autonomously adjust their ABS muting pattern on a fast basis using load information from the associated picocells.
- Fully dynamic: Dynamic load balancing is conducted at each connection set-up, and macrocells autonomously adjust their ABS muting pattern on a fast basis using load information from the associated picocells.

The semi-static options are the baseline cases in this work and are applied in the majority of the existing eICIC papers, where the desired ABS muting pattern and RE is found either from brute-force parameter sweeping to determine the best setting, or via application of SON-based algorithms with long-term adaptation. Although sometimes referred to as dynamic eICIC in the literature, the adaptation of the muting pattern or the RE in a slow basis does not exploit the instantaneous variations in the traffic load and network conditions. For the fully dynamic case, load balancing is performed at each connection set-up, enabling both inter-layer and intra-layer (i.e., among the cells in the same layer) load balancing, as well as fast adjustment of the macrocells ABS muting pattern. The proposed algorithms for fast dynamic adaptation are outlined in Section IV after having introduced the network model and performance metrics.

III. NETWORK MODEL AND PERFORMANCE METRICS

The underlying network modeling assumptions naturally influence the performance results, and the associated conclusions. It is therefore important to have the network model reflect the performance-determining factors of the real environment, such as base station placement, radio propagation, interference pattern and traffic behavior.

A. Modeling methodologies

There are basically two different modeling methodologies, namely (i) Generic models and (ii) Site-specific models. The generic models are commonly used, and designed to reproduce general effects that would typically occur when observing a large number of environments within the same category (e.g. dense urban environments). The static Wyner model is among the simplest generic network models, while more advanced models use random placement of base station nodes according to certain point processes [22] [23]. Moreover, generic models use commonly accepted stochastic radio propagation models.

One clear benefit of using commonly accepted generic network models is that the results produced under the same assumptions are comparable. It is for this reason that 3GPP has defined a number of generic network modeling cases for different types of environments, including also co-channel HetNet cases with random deployment of picocells [24] [25].

However, the family of generic network models also has its limitations. For example, due to the homogeneous characteristics of the 3GPP models, a network-wise configuration of the eICIC parameters is often found to provide good performance gains. However, real deployments are far from homogeneity, where e.g. the non-uniform spatial traffic and the three-dimensional (3D) feature characteristics of the environment are
important performance-determining factors. We have therefore chosen to quantify the performance of eICIC based on a 3D site-specific modeling approach. As the name indicates, the site specific model aims at accurately reproducing the considered environment, and hence at producing performance results that are highly realistic and therefore have practical relevance.

B. Site-specific network model

A detailed 3D topography map is used for the considered dense urban European capital area. The map contains 3D building data (i.e. building footprint and height) as well as information on streets, open squares, parks, etc. The area is dominated by multi-floor buildings. The building height varies, with an average height of 14 meters. Macro base stations are placed according to typical operator deployment, taking the local environment characteristics into account in order to have good wide area coverage [26]. The average macro antenna height is in the order of 30 meters, using a few degrees of antenna down-tilt. The considered network area includes hundreds of macro-sites with sectorized antennas. However, our performance analysis is based on data-collection only from a 1.2 km² segment of the area as pictured in Figure 2a, consisting of a number of macro-sites as well as 30 picocells with omni-directional antennas. eNBs outside the considered data-collection area act as interference sources. For comparison purposes, an example of a 3GPP simulation scenario as defined in [25] is shown in Figure 2b, where a cluster of 4 picocells is randomly placed in each macro cell. The picocells are deployed outdoor at 5 meters height in street canyons or at open squares according to the algorithm in [26] to improve the overall fifth-percentile (5%-ile) outage throughput of the network. Both macros and picos are transmitting at the same 10 MHz carrier at 2.6 GHz. Notice from Figure 2a that the considered network topology is highly irregular with a few selected areas of dense pico deployments. The radio propagation characteristics are obtained by using state-of-the-art ray-tracing techniques based on the Dominant Path Model (DPM) [27] [28]. In short, the first step of the DPM is to determine the dominant path between each transmitter and the receiver point. Each dominant path consists of multiple sub-paths, where the connection between two sub-paths is a result of e.g. diffraction around a corner of a building or over a rooftop. Afterwards, the integral path loss along the dominant path is calculated according to the DPM, including both effects of distance dependent attenuation and so-called shadowing caused by diffraction. Parameters of the DPM are calibrated against available data. In coherence with observations from field measurements, outdoor-to-indoor propagation is modelled by including an additional 20 dB wall penetration loss as well as 0.6 dB per meter that the users are placed inside the building [29]. For users placed at elevated floors, we subtract 3.4 dB per floor from the path loss towards macrocells to reproduce the so-called height gain effect, where users at higher floors tend to experience stronger received signal strength from macrocells [30] [31]. Small-scale fading (also known as fast fading) is modelled according to the commonly accepted stochastic typical urban model. Figure 3a shows the dominance area for the macrocells, where each colour represents different cells. Notice that the coverage area of each macrocell varies significantly, and deviates from the assumption of regular hexagonal cells as typically assumed for the generic 3GPP network models [24] [25].

C. Traffic modeling

A dynamic birth-death traffic model is assumed, where the generation of new users is according to a Poisson point process with arrival rate \( \lambda \). A fixed payload of \( B \) bits is assumed for each user. Once the payload has been successfully delivered to the user, the call is terminated and the user is removed. The average offered traffic load equals \( \mathcal{L} = \lambda \cdot B \). Whenever a new user is generated, the spatial location of the user in the horizontal plane is chosen randomly according to a discrete two-dimensional probability mass function. The probability mass function expresses the average traffic density for each (horizontal) pixel of 10 x 10 meters for the considered area in line with [26]. For users that are placed at locations that coincide with multi-floor buildings, there is an equal probability of placement per floor. The spatial traffic distribution is non-uniform with high variability of the traffic density per pixel. In fact, 80% of the users are indoor, although only 40% of the pixels represent areas with buildings. Furthermore, the 10% of the pixels with the highest traffic density account for nearly 50% of the total offered traffic.

D. Performance metrics

The primary performance metric is the downlink end-user experienced data rate. For each call, we monitor the time, \( T \), it takes to successfully transfer the \( B \) bits, and obtain the average user’s data rate as, \( \bar{R} = B / T \). The experienced data rate naturally varies among the users. Based on a large number of samples of users data rate, the empirical cumulative distribution function (cdf) is built, enabling us to compute the 5%-ile outage and 50-percentile (50%-ile) experienced user throughput. The system capacity is defined as the maximum offered load that can be tolerated while still being able to serve at least 95% of the users with a given minimum data rate. Unless otherwise mentioned, we use a minimum data rate constraint of 2 Mbps for determining the system capacity. As a final performance measure, we also use Jain’s fairness index, defined as [35]

\[
J(R_1, R_2, ..., R_U) = \frac{(\sum_{u=1}^{U} R_u)^2}{U \cdot \sum_{u=1}^{U} R_u^2}, \tag{2}
\]

where \( R_u \) is the throughput of user \( u \) and \( U \) is the total number of users. For the case with equal throughput for all users, Jain’s fairness index equals one, while it takes lower values if the users experience larger variance in the data rate.

As the considered dense urban capital area is highly irregular, one has to be careful if samples from the entire network area are used for computing the defined performance metrics. A cells dominance area is defined as the locations where a UE would receive the strongest power from that cell.
Fig. 2. Spatial location of macro and picocells in (a) Site-specific scenario and (b) a generic 3GPP simulation scenario. The macrocells are marked with circles and a line pointing in the direction of the main lobe of the antenna (a.k.a. broadside). The triangular symbols mark picocells with omni-directional antennas.

Fig. 3. (a) Dominance area of the macrocells. Each colour represents different macrocells. (b) Sub-division of the considered network area into separate areas for localized performance analysis.

The following two approaches are therefore considered: (i) Global performance based on samples from users in the entire considered area, and (ii) Localized performance based on samples from users positioned in a certain confined area (i.e. a sub-region of the considered area). Figure 3b pictures the five considered sub-regions, denoted Area 1 to 5, for extracting the localized performance. The areas have been selected to represent segments of the city with different characteristics. To give a few examples, Area 5 only contains macrocells and therefore there is obviously no gain from eICIC in this particular sub-region, and hence only a potential gain from load balancing between the macros. Area 4 is having the highest density of picos per macrocell area, while Area 1 is having picos deployed further away from the dominant macro. Analyzing the performance per area provides additional insight on the eICIC performance, as compared to only considering the global performance statistics that essentially represent the combined performance over the entire considered network area. In fact, given that Areas 1-5 have different characteristics, results from these areas enables us to analyze how the performance of eICIC varies depending on the local environment features, i.e. it allows us to present a more exhaustive eICIC performance sensitivity analysis.

IV. Dynamic eICIC Strategies

In contrast to previous eICIC studies, our framework includes different degrees of load balancing via dynamic cell association assignment principles, as well as fast autonomous ABS adaptation. We call this fast dynamic eICIC.

A. Dynamic user association

When a new user arrives (connection set-up), the first step is to determine which cell shall serve the user. As our system model does not include movement of users, algorithms for serving cell change (also known as handover) are not included.
in the analysis. This essentially means that load balancing is modeled as part of the serving cell selection each time a new user is created. In addition to the reference case based on received power measures in (1) with \( RE_i = 0 \) dB for all the macrocells and same value of \( RE_i \geq 0 \) dB for all the picocells, we consider two additional methods as outlined in the following.

1) Load-based cell selection: The first class of serving cell selection algorithms is primarily based on cell load measures. For the sake of simplicity, we measure the load per cell as the number of users, denoted by \( u_i \) for cell \( i \), although more elaborate cell load measures could also be applied, e.g. the composite available capacity measure [21]. In this paper we consider the two least loaded candidate cells as well as the received signal strength from those. The two least loaded cells are

\[
c_i^* = \arg \min_{i \in F} \{u_i\} \tag{3}
\]

and

\[
c_2^* = \arg \min_{i \in [F - c_i^*]} \{u_i\} , \tag{4}
\]

where \( F \) is the set of feasible candidate cells. In order to avoid users connecting to a cell that is too weak compared to the strongest received cell, the set of feasible candidate cells for the user is limited to cells fulfilling the following condition,

\[
F \triangleq \{c_i \mid \forall i \in M \cup S : \left( \frac{P_i}{P_{\text{max}}} > \Delta \right) \}, \tag{5}
\]

where \( M \) and \( S \) denote the set of macrocells and picocells, respectively, \( P_i \) is the received power (expressed in watts) by the UE from cell \( i \), \( P_{\text{max}} = \max \{P_i\} \), and \( \Delta \) is the maximum tolerable received power difference between cells. In LTE, \( P_i \) corresponds to the RSRP. By default, we set \( \Delta \) such that cells in the feasible set are those with a received power difference of maximum -10 dB as compared to the cell with the strongest received power.

Instead of just selecting the least loaded cell \((c_i^*)\), we select cell \( c_2^* \) if the second least loaded cell only has one more user, but has stronger received power than the least loaded cell, i.e.,

\[
i^* = \begin{cases} c_2^* & \text{if } (P_{c_2^*} > P_{c_i^*}) \land (u_{c_2^*} \leq u_{c_i^*} + 1) \\ c_i^* & \text{otherwise} \end{cases} \tag{6}
\]

where \( i^* \) is the selected cell. In the following, we refer to this scheme as load-based cell selection. Note that due to the assumed dynamic traffic, the load per cell is a time-variant process, and therefore the cell selection outcome is not only a function of the location of the user.

2) Opportunistic cell selection: Since Knopp and Humblet [41] showed that time-varying transmission conditions in a wireless system can be exploited by opportunistic scheduling, these kinds of schedulers have been widely applied. Inspired by that concept, we propose an equivalent method for selecting the serving cell that offers the highest throughput to the user. We refer to this criterion as opportunistic cell selection, and it can be formally written as

\[
i^* = \arg \max_{i \in F} \{\hat{R}_i\}, \tag{7}
\]

where \( i^* \) is the selected cell, \( \hat{R}_i \) is the estimate throughput for the user in cell \( i \) and \( F \) is the previously defined set of feasible candidate cells fulfilling (5), with a received power difference of maximum -15 dB as compared to the strongest received cell.

The estimated throughput for the user in cell \( i \) is obtained by using the Shannon capacity formula and assuming equal resource sharing between the users per cell, i.e.,

\[
\hat{R}_i = \frac{1}{u_i + 1} W \log_2(1 + \Gamma_i), \tag{8}
\]

where \( W \) is the carrier bandwidth, and \( \Gamma_i \) is the conditional estimated wideband signal-to-interference-plus-noise-ratio (SINR) for the UE if served by cell \( i \). The value of \( \Gamma_i \) is obtained as

\[
\Gamma_i = \frac{P_i}{\sum_{n \neq i} P_n + N_0} \tag{9}
\]

for all macrocells, where \( N_0 \) is the thermal noise power. If cell \( i \) is a picocell, the expression of \( \Gamma_i \) is modified to exclude the interference from the dominant macrocell in the denominator of (9), i.e. assuming ABS in the dominant macro. Note that the throughput estimation in (8) is rather simple as it does not fully account for the dynamic system behavior: the estimate is optimistic for macrocells, as the cost in terms muted subframes is not included; and optimistic for the picocells, too, where the dominant macro is assumed to mute all the time.

B. Generalized fast ABS

The distributed algorithm for fast autonomous ABS adjustment by each of the macrocells is based on the work originally presented in [19] [20]. The scheme relies on semi-static configuration of few subframes as normal transmission and ABS - also referred to as normal and mandatory ABS. During those semi-statically configured subframes, users are requested to perform time-domain restricted channel quality feedback measurements to reflect the quality depending on whether the macrocell is muting or not [20]. The remaining majority of subframes are called optional ABS. The optional ABS can be used as normal subframe or ABS in accordance with the desired optimization criterion. The macrocells decide shortly before the beginning of the optional ABS if it shall be used for normal transmission or ABS. It is assumed that each macrocell makes such decisions autonomously without any explicit coordination with its neighboring macrocells. The macrocell bases its decisions on load measurements from the picocells within its geographical coverage area, as well as knowledge of its own carried load.

In the majority of the published eICIC studies (including [19] and [20]), it is assumed that all users served by the same picocell perceive the same macrocell as their main interference aggressor. However, this is generally not the case for real deployments such as the irregular scenario studied in this paper. Here, it is likely that different pico-users served by the same cell perceive different macrocells as their strongest interferer. An example of this is depicted in Figure 4, where it is observed that the two pico-users associated to pico eNB 4 perceive
two different macrocell aggressors, namely macro eNB 1 and eNB 2. Consequently, one of them will benefit from the muting in macro eNB 1, whereas the other one can improve its performance when macro eNB 2 is muting. In any case, there are also pico-users not identified as victim users as it is the case of user D that would not benefit from macro muting.

In further outlining the algorithm, we use the following notation: $z$ and $n$ denotes the number of subframes used as ABS and normal in the current ABS period, respectively. Furthermore, $T_{ABS} \in \mathbb{N}$ is the repetition period of the ABS pattern, $u_{\text{aggressor}}$ is the total number of users served by the aggressor macrocell, $u_{\text{victim}}$ is the sum of users from the surrounding picocells that have the macrocell as their dominant aggressor, and $U = u_{\text{aggressor}} + u_{\text{victim}}$ is the total amount of affected users. For each optional ABS, the algorithm checks a double condition: first to ensure that the ratio of macro users is served with an appropriate ratio of normal transmissions subframes, and secondly to check whether the ratio of ABS resources assigned so far is lower than the ratio of victim users, i.e.

$$
\frac{u_{\text{aggressor}}}{U} < \frac{n}{T_{ABS}} \land \frac{u_{\text{victim}}}{U} > \frac{z}{T_{ABS}}.
$$

If (10) is true, the current subframe is muted, and the victim user will have an opportunity to be scheduled in the next subframe. Otherwise, normal transmission is applied in the subframe. Each of the macrocells acquire the value of $u_{\text{victim}}$ from load reports coming from the individual picocells, i.e. via inter-cell signaling.

V. PERFORMANCE RESULTS AND ANALYSIS

A. Simulation methodology and assumptions

We evaluate the fast dynamic eICIC framework for the network model in Section III, using a system-level LTE simulator following the 3GPP specifications, including detailed modeling of major RRM functionalities [43]. The system-level simulator has been used to generate a large variety of LTE performance results in coherence with various 3GPP simulation assumptions and in line with results from other 3GPP contributors. The basic methodology of the simulation tool is outlined in the following. For each 1 ms subframe, the experienced SINR for each scheduled user is calculated per sub-carrier, assuming an interference rejection combining (IRC) receiver [37]. Given the SINR per subcarrier, the effective exponential SINR model [38] for link-to-system-level mapping is applied to determine if the transmission was successfully decoded. Failed transmissions are retransmitted using hybrid ARQ with ideal Chase Combining (CC) [39]. For the latter case, the effect of hybrid ARQ with CC is captured by the link-to-system model by linearly adding the SINRs for the different hybrid ARQ transmissions. The modulation and coding scheme for first transmissions is determined by the link adaptation functionality based on frequency selective channel quality measurement feedback from the users [43]. The pico-users are configured to report separate channel quality measurement feedback when the macrocell is using normal subframe and ABS (and no measures taken during optional ABS). Closed loop 2x2 single user MIMO with pre-coding and rank adaptation is assumed for each link [42]. Ideal cancellation of CRS interference during protected subframes is further assumed for the pico-users [9]. Each cell schedules its users according to the Proportional Fair algorithm. The management and prioritization of scheduling new transmissions and hybrid ARQ retransmissions is according to [40].

The time-variant traffic model outlined in Section III-C is applied, assuming a payload size of $B = 4$ Mbit for each call. In order to obtain statistically reliable results for the end-user throughput, simulations are run for a time-duration corresponding to at least 2800 completed calls. This is sufficient to have a reasonable confidence level for the considered performance metrics. The default assumptions for the simulations are summarized in Table I.

<table>
<thead>
<tr>
<th>TABLE I. SUMMARY OF DEFAULT SIMULATION ASSUMPTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmit power</td>
</tr>
<tr>
<td>Bandwidth</td>
</tr>
<tr>
<td>Subframe duration</td>
</tr>
<tr>
<td>Modulation schemes and coding</td>
</tr>
<tr>
<td>HARQ modeling</td>
</tr>
<tr>
<td>Transmission mode</td>
</tr>
<tr>
<td>Antenna gain</td>
</tr>
<tr>
<td>Antenna pattern</td>
</tr>
<tr>
<td>eNB packet scheduler</td>
</tr>
<tr>
<td>UE capabilities</td>
</tr>
</tbody>
</table>

B. Global performance statistics

We first present performance results based on a collection of statistics from the entire network, i.e. global performance. Figure 5 shows the 5%-ile and 50%-ile user throughput as a function of the average offered load in the whole network. The selected offered load range corresponds to a system in equilibrium, where the carried traffic equals the offered load. Higher load values would lead to congestion. The solid line is the baseline case with no eICIC and no biased user association. With pentagrams, the result of using the classical semi-static eICIC is shown, where the RE giving the best global coverage performance is selected for each value of average offered load (i.e. it is only changed in a slow basis). The cross refers to a static network-wise configuration of the cell association and fast ABS locally applied per macrocell. Finally, the two fully dynamic eICIC curves use fast ABS and the two cell association methods discussed in Section IV: load balancing with circles and opportunistic cell association with diamonds.

As expected, the user throughput decreases as the load in the network increases. When using network-wise RE and semi-static muting, the optimal value of the parameters varies with the offered load. At low offered load, a small RE value is better, and there is little gain from applying macro muting. This is due to the fact that there is only marginal other-cell interference, and the gain in these low-loaded cases comes from the application of a small RE offset at the picocells. As the offered load increases, both macrocells and picocells start having higher probability of transmitting (and thus causing interference), and the system converges to using more ABS at
the macrocells and higher RE at the picocells. In contrast to previous studies using 3GPP simulation scenarios, semi-static eICIC only provides gains for the 5%-ile user throughput, but not for the 50%-ile user throughput. Instead, the importance of fast ABS in highly irregular networks is clearly observed, leading to significant gains for both indicators at the full range of considered offered traffic loads. Based on a closer inspection of the global user throughput statistics, it is found that the 5%-ile user throughput is primarily dominated by users positioned in Area 2 (see Figure 3b). This emphasizes the highly irregular nature of the considered network.

Furthermore, the two proposed dynamic cell selection methods at connection set-up are proven to improve the performance as the offered load increases. For low load, the best option is simply to connect to the cell offering the best received signal strength, since the probability of having simultaneous transmission with users in neighbor cells or in the serving cell is low. As the load increases, active load balancing becomes increasingly important to facilitate more efficient resource sharing and allocation for the users. Finally, the opportunistic cell selection scheme outperforms the simpler load-based strategy, owing to the addition of the channel conditions in the decision.

Figure 6 shows the empirical cdf of the number of users in the macro-layer for the different cell selection mechanisms and an average offered load of 420 Mbps (high load conditions). It is observed that the network-wise RE selection leads to long tails in the distribution function, and consequently saturation of the macro layer. The classical load-based cell selection and the opportunistic cell selection significantly reduce the long tail of number of users in a macrocell, giving the opportunistic cell selection the lowest average number of users. Similar results are observed in the pico-layer, where the number of pico users is in general lower compared to the macrocells. In an irregular network, it is very likely that the global performance is limited by few cells that first saturate, while other cells are operating at low or medium load. Thus, dynamic mechanisms for load balancing and resource management are crucial.

For a better understanding of the dynamics of the generalized fast ABS algorithm, the empirical cdf of the macrocell muting ratio is plotted in Figure 7 for four exemplary cells at high offered load (420 Mbps). As compared to the results reported in [20], the muting patterns are here more heterogeneous. It is observed that two of the eNBs (A and D) are using normal transmission most of the time. This is because they have on average a large number of connected users that cannot be offloaded to the picocell layer since there are no deployed picocells in that location. The other two eNBs (B and C) are aggressors for a large number of pico-users and therefore they apply much more muting in order to improve the service to victim users. Similar trends were observed for
the rest of macrocells in the network. Figure 7 proves the importance of a local and autonomous configuration of the macro muting in order to get the most out of the interference coordination mechanism.

Finally, Figure 8 shows Jain’s fairness index as defined in (2), as a function of the network offered load and for different eICIC configurations. When the offered load increases, the fairness diminishes. More importantly, it is also observed how the use of fast ABS (with fixed RE allocation or with opportunistic cell selection) results in improved fairness as compared to the semi-static muting configuration, being especially visible at the higher offered traffic loads.

C. Local performance statistics

To gain further insight into the performance of dynamic eICIC, we next analyze the local performance statistics collected for each of the defined network areas as illustrated in Figure 3b. Figure 9 shows the gain in user throughput performance of the eICIC schemes for the different areas. These results are obtained by loading the network up to the maximum tolerable load for each area without using eICIC, followed by enabling eICIC to quantify the experienced increase in end-user throughput. For the dynamic eICIC case, the best cell selection method for the area is chosen, i.e. using either the best area-wise RE, load-based cell selection, or opportunistic cell selection. Notice that Area 5 has no picocells and consequently there is no reason for applying eICIC (not shown in the Figure). In general, the eICIC performance gain is higher for 5%-ile user, although worthwhile gains are also observed for the 50%-ile user throughput metric.

In Figure 10 the capacity gain for the different areas is plotted. Recall from Section III-D that capacity is defined as the maximum tolerable offered load for each area without eICIC, followed by enabling eICIC to quantify the experienced increase in end-user throughput. For the dynamic eICIC case, the best cell selection method for the area is chosen, i.e. using either the best area-wise RE, load-based cell selection, or opportunistic cell selection. Notice that Area 5 has no picocells and consequently there is no reason for applying eICIC (not shown in the Figure). In general, the eICIC performance gain is higher for 5%-ile user, although worthwhile gains are also observed for the 50%-ile user throughput metric.
order of 35% to 120%. The main observations from the local performance statistics can be summarized as follows:

- Fast ABS is responsible for most of the gains from using dynamic algorithms. With fast ABS, each macro eNB dynamically changes its muting pattern based on the local traffic and interference conditions.
- Dynamic cell selection techniques also provide benefits over a network- or area-wise RE in cases of load imbalance, as it is the case in Area 2, thanks to their inter-cell load balancing capabilities. On the other hand, Areas 1, 3 and 4 do not significantly improve their performance by using dynamic cell selection algorithms as the load in these areas is more evenly distributed. In these cases, a network- or area-wise RE value achieves approximately the same performance.
- The highest gain from applying eICIC is observed for Area 2 and 4 as these regions are characterized by a relative high number of picocells that are exposed to dominant macrocell interference.
- On the contrary, the lowest eICIC gain is observed for Area 3. This is mainly contributed by two factors; (i) the picocells in this area are exposed to less macrocell interference and (ii) several of the traffic hotspots in Area 3 are not covered by a picocell.

In general, the relative large variability of the eICIC performance gain from one local area to another emphasizes the importance of basing the conclusions not only on the global performance statistics, but also to analyze the local performance statistics. As the local performance statistics present the performance for areas of different characteristics, it shows how sensitive the eICIC performance is depending on the local environment features. It is worth noticing that despite the relatively simple algorithms for dynamic ICIC operation, promising gains are observed for Areas 1-4, which are highly robust and able to efficiently adapt to the local time-variant conditions. As these results are obtained for a realistic 3D network model, it is fair to conclude that simple dynamic eICIC algorithms like the ones presented here have a high potential in practice.

VI. DISCUSSION

The presented results show promising gains from applying eICIC with fast dynamic ABS adjustment and load balancing. We next discuss how to realize such techniques in a practical implementation. The two considered load balancing techniques upon connection set-up rely on knowledge of the cell load, while the opportunistic cell selection requires also the estimation of the throughput; see (7). The throughput estimate used in this study for the opportunistic cell selection is rather simple, and could be further improved, e.g. by exploiting additional a priori knowledge of the terminal measurements and information of ABS muting ratios at different macrocells. The exchange of various load measures between eNBs via the X2 interface is already supported by the current LTE specifications [36], so this is in place for enabling the network to conduct the proposed load-aware cell selection algorithms, as well as the corresponding mobility load balancing for cases where users are moving (although not included in this study). The required exchange of information between the eNBs for the algorithms is rather modest, and therefore considered doable.

The algorithms for fast dynamic ABS adjustment also requires exchange of load information between each macrocell and the underlying picocells. As studied in greater details in [20], the rate of fast dynamic ABS adaption is therefore limited by how fast such load information is exchanged between cells. For cases where the picocells are implemented as remote radio heads with zero-latency fronthaul connections to the macros, the ABS adaptation can be on a subframe basis (as assumed in the presented results). For cases with traditional X2 backhauls, the maximum recommended rate of ABS adaptation depends on the X2 latency. From control theory, we know that adjustments should be slower than the loop delays to avoid instability. With X2 delays of e.g. 5-10 ms, it is possible to perform fast ABS approximately every 40 ms. This is also in line with the X2 specifications for signaling the used ABS
pattern from a macrocell to its underlying picocells. For more details on X2 signaling for ABS adjustment, see [4] and [36].

VII. CONCLUSIONS

In this paper we have presented a fully dynamic framework with fast adjustment of the key performance parameters of eICIC – namely the ABS muting ratio and the cell association for facilitating efficient load balancing. The performance of the proposed dynamic ICIC algorithms is evaluated in a realistic setting, using an advanced 3D network model based on data from a dense urban area. As compared with semi-static SON-based schemes, the fast dynamic algorithms can track rapid traffic fluctuations and autonomously adapt to the local network conditions. The joint use of opportunistic cell selection and generalized fast ABS provides a significant improvement in the studied KPIs (5%- and 50%-ile users throughput, fairness index and capacity gain). Thus, despite the simplicity of the proposed algorithms for dynamic eICIC operation, rewarding benefits are observed. As promising performance gains are observed for network areas of different characteristics, it suggests that the proposed scheme is generally applicable for real-life dense urban environments. This is a consequence of using simple and generic dynamic eICIC algorithms that are not based on numerous underlying assumptions that could vary from area to area. Finally, the proposed dynamic algorithms have the advantage of being implementable without excessive inter-eNB information exchanges.

REFERENCES


