Load Balancing in Heterogeneous LTE: Range Optimization via Cell Offset and Load-Coupling Characterization

Iana Siomina[1] and Di Yuan[1,2]

[1]Ericsson Research, Ericsson AB, Sweden

[2]Department of Science and Technology, Linköping University, Sweden

Emails: iana.siomina@ericsson.com, diyua@itn.liu.se

***Abstract*— Heterogeneous networks (HetNets) with a topology of mixed macro-cells and low-power nodes (LPNs) form an important step of capacity enhancement for LTE and LTE-A. In this paper, we present an optimization framework for load balancing in LTE HetNets, by means of cell range assignment using cell-specific offset. For any given offset setting, the resulting cell load is effectively approached via the solution of a system of non-linear equations characterizing the load-coupling relation between cells. We present a computationally efficient bounding scheme to approximate the solution of the non-linear system and provide theoretical insights into the monotonicity and convergence of the scheme. The bounding scheme is embedded into an algorithm based on the principle of design of experiments (DOE) for cell offset optimization. Simulation results demonstrate the effectiveness of the optimization process for LTE load balancing with HetNet elements.**

Index terms **– Heterogeneous networks, LTE, load balancing, radio resource management, optimization.**

I. Introduction

The interest in deploying low-power nodes, or LPNs, such as pico/micro base stations, home eNodeBs, relays, or remote radio heads, to enhance the macro-cell network performance in terms of coverage, capacity, and service experience of individual users has been growing over the last few years. Wireless networks where LPNs are deployed as an overlay to the macro-cell layer form an example deployment of heterogeneous networks (HetNets). HetNet is a promising approach for capacity extension in certain areas, e.g., the so-called traffic hotspots, where short-range cells created by LPNs may be efficiently used to provide additional resources for locally enhancing the network performance [5]. LPNs are becoming particularly attractive for handling the growing geographical diversity of traffic and the increasing variation in local densities of user distribution. HetNet deployments may thus be viewed as a way of densifying networks, in order to effectively adapt to the traffic demand and the radio environment [3], [4].

Along with the potential, HetNet deployments bring challenges to efficient network operation to ensure superior user experience. For example, network planning and optimization are becoming more challenging due to a more dynamic and potentially less coordinated nature of LPN deployment, especially in a co-channel scenario with macro-cells. Poorly planned LPN site location or configuration not only results in unsatisfactory performance in that cell, but may also cause a significant performance degradation and overload in the neighboring cells.

Another challenge is cell coverage control. The coverage area of LPNs is particularly sensitive to the relative location, radio propagation conditions, and received signal strength of neighboring cells. As a result, the cell size of LPNs may vary greatly over the network. Traditionally, cell coverage is determined by the strongest received signal strength. Due to the higher output power of neighboring cells, in particular macro-cells, the coverage area of LPNs may be too small to serve a closely located traffic hotspot. In this case, the coverage may be extended by means of a biased cell selection/reselection rule for some or all user equipments (UEs). That is, a UE is allowed to reselect or perform handover to a weaker cell, by configuring the cell reselection offset which is becoming a common cell parameter. However, a very large offset value is not desirable either, since the serving cell may become too weak in relation to the interference. Consequently LPN cells become easily overloaded in serving the high traffic demand in hotspot areas. To some extent, the issue can be alleviated by providing more resource to the LPNs, e.g., higher output power, which, however, in case of poor isolation will increase the interference and load in the neighboring cells.

Load balancing is a key aspect in radio resource management. For UMTS, load balancing, e.g., by adjusting pilot power levels, has been extensively studied [10], [12], [13]. For HetNet deployment, the topic is much more challenging, because of the asymmetric interference relation between the macro-cells and LPNs. To avoid overloading and radio resource exhaustion, approaches for effective load balancing are essential.

In this paper, we approach LTE HetNet load balancing via LPN range offset optimization. In contrast to, e.g., [2], [6], [11], our approach allows for non-uniform cell offsets. A feature of our approach is load assessment by cell load-coupling characterization. In Section II, we present the optimization framework for virtual cell range assignment by means of cell-specific offsets, as well as the cell-load coupling system model and its key properties. Next, in Section III, a cell load estimation scheme based on a bounding principle is developed, and theoretical insights into the monotonicity and convergence of the scheme are presented. In Section IV, a load balancing algorithm is developed for virtual range assignment of LPNs. We also provide a concept for improving algorithm efficiency, along with a theoretical proof of the concept. Simulation results, presented in Section V, demonstrate the effectiveness of the proposed optimization process in achieving significantly more balanced cell load in comparison to two baseline cell offset solutions. Conclusions are then given in Section VI.

II. System Model

A. Optimization Framework

Denote by $\mathcal{I}_1$ and $\mathcal{I}_2$ the sets of macro-cells and LPN cells, respectively. Let $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$, and $n = |\mathcal{I}|$. The set of UEs is denoted by $\mathcal{J}$. We use ρ_i as the generic notation of the load of cell i; the load represents the level of resource consumption of the cell. For LTE, the entity corresponds to the average fraction of the resource blocks (RBs) that are scheduled for data transmission, for the frequency-time domain in question. The network-wide load is given by the vector, $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_n)^T$.

We assume that cell-specific offset optimization applies to the LPN cells in set $\mathcal{I}_2$, and the cells in $\mathcal{I}_1$ use zero offset (set $\mathcal{I}_1$ may also be empty). The candidate offset values compose set $\mathcal{S}$.

978-1-4577-2053-6/12/$31.00 ©2012 IEEE

Only non-negative offset values in $\mathcal{S}$ are considered in Section V, which is, however, not a limitation of the algorithm. We use x_{is} to denote whether or not cell $i \in \mathcal{I}_2$ uses offset level $s \in \mathcal{S}$. The resulting network-wide range assignment of $\boldsymbol{x}$, that is, the allocation of UEs to cells, is denoted by $\boldsymbol{u}(\boldsymbol{x})$, with $u_{ij}(\boldsymbol{x}) = 1$, if cell $i \in \mathcal{I}$ is the serving cell of UE $j \in \mathcal{J}$. For each UE, the assignment follows the rule of selecting the cell (macro- or LPN cell) with the best received signal power plus the offset.

Load ρ_i is dependent on the UEs served by cell i, as well as the load of the other cells because of inter-cell interference. We use $\boldsymbol{\rho}^{\boldsymbol{u}(\boldsymbol{x})}$ as a compact notation for the load vector resulted by offset setting $\boldsymbol{x}$. The following optimization formulation is used for load balancing by virtual cell range assignment,

$$\max \quad F(\boldsymbol{\rho}^{\boldsymbol{u}(\boldsymbol{x})}) = \frac{\left(\sum_{i\in\mathcal{I}} \rho_i^{\boldsymbol{u}(\boldsymbol{x})}\right)^2}{n \cdot \sum_{i\in\mathcal{I}} \left(\rho_i^{\boldsymbol{u}(\boldsymbol{x})}\right)^2} \tag{1a}$$

$$\sum_{i\in\mathcal{I}} u_{ij}(\boldsymbol{x}) = 1, \qquad \forall j \in \mathcal{J} \tag{1b}$$

$$\sum_{s\in\mathcal{S}} x_{is} = 1, \qquad \forall i \in \mathcal{I}_2 \tag{1c}$$

$$x_{is} \in \{0,1\}, \qquad \forall i \in \mathcal{I}_2, s \in \mathcal{S} \tag{1d}$$

$$\rho_i^{\boldsymbol{u}(\boldsymbol{x})} \in [0,1], \qquad \forall i \in \mathcal{I}. \tag{1e}$$

The objective function in (1a) is the well-known Jain's fairness index. The function value has range $[\frac{1}{n}, 1.0]$. The maximum of 1.0 is attained if all cells have identical load. Thus the function reflects a load balancing strategy aiming at a "fair" load distribution among cells.

B. Load-Coupling Characterization

Solving problem (1) necessitates the estimation of cell load with given offset $\boldsymbol{x}$, i.e., the calculation of $\boldsymbol{\rho}^{\boldsymbol{u}(\boldsymbol{x})}$. Clearly, there is an inter-cell coupling relation between the load values. In fact, ρ_i corresponds to the activity factor, and hence the probability of interference originating from cell i in the time-frequency domain. We apply the cell load-coupling function developed in our previous works [8], [9]:

$$\rho_i = f_i(\boldsymbol{\rho}) = \sum_{j\in\mathcal{J}:u_{ij}(\boldsymbol{x})=1} \frac{r_j}{\log_2(1 + \frac{1}{\sum_{k\in\mathcal{I}\setminus\{i\}} b_{ikj}\rho_k + c_{ij}})}. \tag{2}$$

In (2), r_j is the normalized traffic demand of UE j by the amount of effective time-frequency resources within one resource unit. Parameter b_{ikj} denotes inter-cell coupling for UE j, namely, the ratio between the interfering signal of cell k, if active, and that of the serving cell i. Parameter c_{ij} is the noise effect over the received signal power of cell i by UE j. To simplify notation, superscript $\boldsymbol{u}(\boldsymbol{x})$ is omitted in (2). Note that ρ_k in (2) plays the role of interference scaling in the product $b_{ikj}\rho_k$. As a result, the second term in the logarithmic function is the SINR of UE j. For detailed derivation of (2), please see [8]. Here, one observes the intuitive result that ρ_i is strictly increasing in normalized demand r_j, inter-cell interference, and noise.

The load-coupling system can be written in a compact form as in Eq. (3). To enable to quantify scenarios of over-loaded network, we do not limit the cell load vector $\boldsymbol{\rho}$ to be at most 1.0 in this equation system [9]. However, constraints (1e) can be ensured while solving (1).

$$\boldsymbol{\rho} = \mathbf{f}(\boldsymbol{\rho}), \quad \boldsymbol{\rho} \geq \mathbf{0} \tag{3}$$

We use $\boldsymbol{\rho}^*$ to denote the solution to (3). In this paper, repeatedly solving (3) to evaluate candidate offset solutions forms the bulk of the computational effort of optimizing (1). Hence we are particularly interested in time-efficiently approaching a tight approximation of $\boldsymbol{\rho}^*$. To this end, we summarize some useful theoretical properties of (3). Formal proofs of the properties are provided in [9].

- Property 1 (*Function characterization*): Function $f_i, i \in \mathcal{I}$, is twice differentiable and strictly concave.
- Property 2 (*Solution uniqueness*): $\boldsymbol{\rho}^*$, if exists, is unique.
- Property 3 (*Solution existence*): System (3) has a solution *if and only if* the linear system $\boldsymbol{\rho} = \mathbf{H}^0\boldsymbol{\rho} + \mathbf{f}(\mathbf{0})$ has a solution in $\mathbb{R}^n_+$, where $\mathbf{H}^0$ is the $n \times n$ matrix with $H_{ii} = 0, i \in \mathcal{I}$, and $H_{ik} = \ln(2) \sum_{j\in\mathcal{J}:u_{ij}(\boldsymbol{x})=1} r_j b_{ikj}$ for $i \neq k$.
- Property 4 (*Lower bounding*): If $\boldsymbol{\rho} = \mathbf{H}^0\boldsymbol{\rho} + \mathbf{f}(\mathbf{0})$ is feasible, then its solution (which clearly is unique), denoted by $\boldsymbol{\rho}^0$, bounds $\boldsymbol{\rho}^*$ from below, that is, $\boldsymbol{\rho}^0 \leq \boldsymbol{\rho}^*$.
- Property 5 (*Upper bounding*): For any $\bar{\boldsymbol{\rho}} \in \mathbb{R}^n_+$, consider the $n \times n$ matrix $\bar{\mathbf{H}}$ in which $\bar{H}_{ii} = 0, i = \mathcal{I}$, and $\bar{H}_{ik} = \frac{\partial f_i}{\partial \rho_k}(\bar{\boldsymbol{\rho}})$ for $i \neq k$, that is, $\bar{\mathbf{H}}(\boldsymbol{\rho} - \bar{\boldsymbol{\rho}}) + \mathbf{f}(\bar{\boldsymbol{\rho}})$ is a linearization of $\mathbf{f}$ at $\bar{\boldsymbol{\rho}}$. If $\boldsymbol{\rho} = \bar{\mathbf{H}}(\boldsymbol{\rho} - \bar{\boldsymbol{\rho}}) + \mathbf{f}(\bar{\boldsymbol{\rho}})$ has a (unique) solution $\hat{\boldsymbol{\rho}} \in \mathbb{R}^n_+$, then $\hat{\boldsymbol{\rho}}$ bounds $\boldsymbol{\rho}^*$ from above, i.e, $\boldsymbol{\rho}^* \leq \hat{\boldsymbol{\rho}}$, because the load function is concave (Property 1).

III. A Bounding Scheme for Load Estimation

In this section, we focus on load estimation for a given offset setting $\boldsymbol{x}$. As in Section II-B, superscript $\boldsymbol{u}(\boldsymbol{x})$ is omitted for better readability. The feasibility of (3) is verified by Property 3. The non-linear system (3) has a solution if $\mathbf{I} - \mathbf{H}^0$ is invertible, i.e., has non-zero eigenvalues. If (3) is infeasible, the current offset setting can be discarded from further consideration. Assume that the system is feasible. One solution approach is the Newton method that uses the update formula $\boldsymbol{\rho} + (\mathbf{I} - H(\boldsymbol{\rho}))^{-1} \times (\mathbf{f}(\boldsymbol{\rho}) - \boldsymbol{\rho})$, with H being the Hessian. The method is typically fast. However, there is no convergence guarantee due to, e.g., ill-conditioned iteration matrix $\mathbf{I} - H(\boldsymbol{\rho})$, or poor initial solution. Newton-like methods with trust region may provide global convergence, at the price of additional complexity in parameter tuning. Further, except for the residual, the type of methods generally does not provide quality estimates in form of lower and upper bounds or uncertainty interval.

We develop a bounding scheme that effectively computes lower and upper bounds to estimate $\boldsymbol{\rho}^*$ with guaranteed convergence. The key elements of the scheme are:

1) a lower-bounding sequence by Gauss-Seidel updates;
2) an upper-bounding sequence computed via Property 5.

For both sequences, we prove their convergence to $\boldsymbol{\rho}^*$. In addition, we prove that the upper-bounding sequence is monotonically improving. (For the lower bounds, the monotonicity follows from the definition of Gauss-Seidel update.) The bounding scheme uses a tolerance threshold on the difference between the bounds as the termination criterion. This enables a precise control of solution quality in estimating $\boldsymbol{\rho}^*$.

Denote by $\check{\boldsymbol{\rho}}^k$ and $\hat{\boldsymbol{\rho}}^k$ the sequences of lower and upper bounds, respectively, with k being the iteration number. A natural starting point for lower bounding is to utilize Property 4, and set $\check{\boldsymbol{\rho}}^0 = \boldsymbol{\rho}^0$. After iteration k, there exists at least one cell, say i, of which the current load, $\check{\rho}_i^k$, is lower than the load function value evaluated for the current load of the other cells, that is, $\check{\rho}_i^k < f_i(\check{\boldsymbol{\rho}}^k)$, otherwise $\check{\boldsymbol{\rho}}^k = \boldsymbol{\rho}^*$. A Gauss-Seidel update increases the load of cell i to the function value, that is,

$$\check{\rho}_i^{k+1} = f_i(\check{\boldsymbol{\rho}}^k). \tag{4}$$

By (4), it is clear that $\check{\boldsymbol{\rho}}^{k+1} \leq \boldsymbol{\rho}^*$. Hence the Gauss-Seidel updates yield a sequence of lower bounds. Note also that the load vector elements may be updated in parallel or

sequentially (assumed in Algorithm 1), with the latter giving better performance. We denote by $\check{\epsilon}^k$ the Euclidean norm of the error in the lower bound, i.e., $\check{\epsilon}^k = ||\boldsymbol{\rho}^* - \check{\boldsymbol{\rho}}^k||$. The sequence $\check{\epsilon}^k, k \geq 0$, has the following properties.

- *Lemma 1 (Lower bound monotonicity):* $\check{\epsilon}^{k+1} < \check{\epsilon}^k$, $k \geq 0$.
  Proof: The result follows directly from formula (4). ■
- *Theorem 2 (Lower bound convergence):* $\lim_{k\to\infty} \check{\epsilon}^k = 0$.
  Proof: See Appendix A. ■

By Property 5, any non-negative point can be used to compute an upper bound via solving the corresponding linear system. Applying the result to $\check{\boldsymbol{\rho}}^k$ yields the sequence of upper bounds $\hat{\boldsymbol{\rho}}^k$, of which we prove the following properties.

- *Theorem 3 (Coherence in bound improvement):* For any two lower bounds $\check{\boldsymbol{\rho}}^\ell$ and $\check{\boldsymbol{\rho}}^k$, with $\check{\boldsymbol{\rho}}^\ell \prec \check{\boldsymbol{\rho}}^k$, that is, the former has at least one element being strictly smaller than the latter, the corresponding upper bounds, if exist, satisfy $\hat{\boldsymbol{\rho}}^k \prec \hat{\boldsymbol{\rho}}^\ell$.
  Proof: See Appendix B. ■
- *Corollary 4 (Upper bound monotonicity):* For any $k \geq 0$, if $\hat{\epsilon}^k$ exists, then $\hat{\epsilon}^{k+1} < \hat{\epsilon}^k$, where $\hat{\epsilon}^k = ||\hat{\boldsymbol{\rho}}^k - \boldsymbol{\rho}^*||$.
  Proof: Follows from Lemma 1 and Theorem 3. ■
- *Theorem 5 (Upper bound convergence):* $\lim_{k\to\infty} \hat{\epsilon}^k = 0$.
  Proof: See Appendix C. ■

Algorithm 1 Cell load estimation using bounding principle

Input: $\mathbf{f}$, t
Output: Lower and upper bounds to $\boldsymbol{\rho} = \mathbf{f}(\boldsymbol{\rho}), \boldsymbol{\rho} \geq 0$
1: $\check{\boldsymbol{\rho}} \leftarrow \boldsymbol{\rho}^0$
2: **if** $\exists\ (I - \mathbf{H}^0)^{-1}$ **then**
3: **repeat**
4: **for** $i = 1 : n$ **do**
5: **if** $\check{\rho}_i < f_i(\check{\boldsymbol{\rho}})$ **then**
6: $\check{\rho}_i \leftarrow f_i(\check{\boldsymbol{\rho}})$
7: **if** $\check{\rho}_i > 1.0$ **then**
8: **return** NULL //Reject allocation $\boldsymbol{u}(\boldsymbol{x})$
9: **end if**
10: **end if**
11: **end for**
12: **if** $\exists\ (I - \check{\mathbf{H}})^{-1}$ **then**
13: $\hat{\boldsymbol{\rho}} \leftarrow \check{\boldsymbol{\rho}} + (\mathbf{I} - \check{\mathbf{H}})^{-1}(\mathbf{f}(\check{\boldsymbol{\rho}}) - \check{\boldsymbol{\rho}})$
14: **end if**
15: **until** $\|\hat{\boldsymbol{\rho}} - \check{\boldsymbol{\rho}}\|/\|\check{\boldsymbol{\rho}}\| \leq t$
16: **end if**
17: **return** $\hat{\boldsymbol{\rho}}$, $\check{\boldsymbol{\rho}}$

The theoretical results together give a convergent bounding scheme for load estimation, if the load-coupling system passes the feasibility check (Property 1). The procedure is formalized in Algorithm 1, where t is the tolerance parameter, and $\check{\mathbf{H}}$ denotes the matrix defined in Property 5 at point $\check{\boldsymbol{\rho}}$. Iteration index k is omitted for compactness. Note that upper bound may not be available immediately (due to the feasibility requirement in Property 5), but will always become available in some iteration and remain available in all subsequent iterations by Theorem 2.

IV. A Load Balancing Algorithm

A. The Solution Approach

For load balancing with objective (1a), we apply the principle of Design of Experiments (DOE) to optimize the offset vector $\boldsymbol{x}$. DOE is a statistical approach for estimating the effects of multiple factors simultaneously and identifying the most important ones by factor screening. The approach is an iterative procedure. An iteration involves 1) designing a set of experiments, which, in our case, amounts to selecting a set of configurations to be evaluated, and 2) performing regression analysis for the evaluated set to determine the configurations in the next step (see, e.g., [7] and the references therein for details of the general methodology and other applications). The choice of the optimization approach is justified by that (1a) is concave in load and has no local optimas for continuous load variables.

Algorithm 2 formalizes the solution approach. The algorithm goes through $numIter$ iterations, each of which is associated with a design matrix $\mathbf{D}$. A column represents the offset selection of a specific cell. A row corresponds to a network-wide offset configuration. Thus matrix element D_{hi} indicates the offset of cell i in network configuration solution h. In this paper, two-level designs based on Hadamard matrices are used. Hence the matrix elements are either '-1' or '1', corresponding to the lowest and higher levels of the current offset range of the cell in question. The initial offset range is $\mathcal{S}$. It is then skunk successively over the iterations. The change in the offset range, which is cell-specific, is determined by the signs of the regression coefficients ($\boldsymbol{\beta}$ in Algorithm 2, where matrix inverse or pseudo inverse may be used to solve the system in line 14). The algorithm terminates when the offset range comprises a single number for all LPN cells. If the offset ranges of all LPN cells are changed in every iteration, the algorithm stops after at most $|\mathcal{S}|$ iterations. The design matrix $\mathbf{D}$ may be the same or different in each iteration. The output are the maximum value of the objective function (1a), and the corresponding offset setting $\boldsymbol{x}^*$ and load vector $\boldsymbol{\rho}^{\boldsymbol{u}(\boldsymbol{x}^*)}$.

Algorithm 2 Load balancing using DOE

Input: $\mathcal{S}$, $numIter$
Output: $\boldsymbol{x}^*$, F^*, $\boldsymbol{\rho}^{\boldsymbol{u}(\boldsymbol{x}^*)}$
1: $\mathcal{S}_i' \leftarrow \mathcal{S}, \forall i \in \mathcal{I}_2$
2: $F^* = 0$
3: **for** $k = 1 : numIter$ **do**
4: Generate design matrix $\mathbf{D}$
5: **for** $h = 1 : \text{size}(\mathbf{D}, 1)$ **do**
6: Find $\boldsymbol{x}$: apply D_{hi} to cell offset ranges $\mathcal{S}_i', \forall i \in \mathcal{I}_2$
7: Find $\boldsymbol{\rho}^{\boldsymbol{u}(\boldsymbol{x})}$: apply Algorithm 1 to $\boldsymbol{u}(\boldsymbol{x})$
8: $F_h = F(\boldsymbol{\rho}^{\boldsymbol{u}(\boldsymbol{x})})$
9: **if** $F_h > F^*$ **then**
10: $F^* = F_h$
11: $\boldsymbol{x}^* = \boldsymbol{x}$
12: **end if**
13: **end for**
14: Find $\boldsymbol{\beta}$ from equation system $[\mathbf{1}\ \mathbf{D}] \cdot \boldsymbol{\beta} = \mathbf{F}$
15: **for** $i = 1 : |\mathcal{I}_2|$ **do**
16: **if** $\beta_{i+1} > 0$ **then**
17: Increase the lower boundary of offset range $\mathcal{S}_i'$
18: **else**
19: Decrease the upper boundary of offset range $\mathcal{S}_i'$
20: **end if**
21: **end for**
22: **end for**
23: **return** $\boldsymbol{x}^*$, F^*, $\boldsymbol{\rho}^{\boldsymbol{u}(\boldsymbol{x}^*)}$

B. Efficiency Improvement

Cell offset optimization repeatedly calls Algorithm 1 (Line 7 of Algorithm 2). From a computational efficiency standpoint, it is desirable to have an initial lower bound better than $\boldsymbol{\rho}^0$. The challenge is that, after an adjustment in some cell's offset, the current lower and upper bounds are not valid any more. In this section, we demonstrate how to compute a valid lower bound that is better than $\boldsymbol{\rho}^0$. Let $\tilde{\boldsymbol{\rho}}$ denote the load estimation (e.g., any of $\hat{\boldsymbol{\rho}}$ and $\check{\boldsymbol{\rho}}$, or the average of the two) in the previous algorithm iteration. Assume that the current load-coupling system is feasible, and denote by $\boldsymbol{\rho}^0$ and $\boldsymbol{\rho}^*$ the corresponding lower bound from Property 4 and the system solution, respectively. The linear system defined in Property 5, $\boldsymbol{\rho} = \tilde{\mathbf{H}}(\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}) + \mathbf{f}(\tilde{\boldsymbol{\rho}})$, is most likely feasible. By the property, solving the system leads to a valid upper bound $\hat{\boldsymbol{\rho}} \geq \boldsymbol{\rho}^*$. Consider the $n \times n$ matrix $\hat{\mathbf{H}}$ with zero diagonal and $\hat{H}_{ik} = \frac{\partial f_i}{\partial \rho_k}(\hat{\boldsymbol{\rho}})$ for the non-diagonal element of row i and column k. In the following, we prove the feasibility of the linear system $\boldsymbol{\rho} = \hat{\mathbf{H}}\boldsymbol{\rho} + \mathbf{f}(\mathbf{0})$ and its usefulness in deriving a valid and better lower bound.

Theorem 6: There exists a solution $\check{\boldsymbol{\rho}}$ satisfying $\check{\boldsymbol{\rho}} = \hat{\mathbf{H}}\check{\boldsymbol{\rho}} + \mathbf{f}(\mathbf{0})$, with $\boldsymbol{\rho}^0 < \check{\boldsymbol{\rho}} \leq \boldsymbol{\rho}^*$.

Proof: See Appendix D. ■

By Theorem 6, we can use $\check{\boldsymbol{\rho}}$ in place of $\boldsymbol{\rho}^0$ in the offset optimization algorithm. The theorem also guarantees better bounding performance of the former.

V. Numerical Results

In this section, we numerically investigate the theoretical findings and evaluate performance for a representative LPN deployment scenario in LTE HetNets. The evaluation is performed for cell load on data channels, i.e., for any cell offset in $\mathcal{S}$, it is assumed that UE is able to synchronize to the serving cell, decode control channels, and perform measurement reporting.

A. Network Configuration

The LTE HetNet scenario consists of three macro-sites with three cells per site, assuming an inter-site distance of 500 m. There are 18 LPN sites/cells. Two LPN sites are randomly placed within each macro-cell area. All macro-sites are equipped with a three-sector downtilted directional antenna with 14 dBi antenna gain, whilst the LPN sites are equipped with an omni-directional antenna with 0 dBi antenna gain. The network layout, adopting the wrap-around technique, is illustrated in Figure 1, where the numbers indicate the cell IDs. For clarity, the IDs of the macro-cells are shown in squares. Two layers of users are generated, with 30 users per macro-cell area, out of which 2/3 (the dot markers) are in two randomly placed hotspots, and 1/3 (the x-markers) are distributed randomly and uniformly. The LPN sites are installed in the middle of the hotspots. Every UE has an omni-directional antenna with 0 dBi antenna gain. The traffic demand is 400 Kb/s for all UEs. The simulated system operates at 2 GHz with 10 MHz bandwidth. The propagation environment and user distribution follow the 3GPP specification in [1], assuming Propagation Model 1 (Okumura-Hata, urban, 8 dB standard deviation shadow fading) and User Generation Scenario 4b (two hotspots of 40 m radius per macro-cell). The output power of the macro base stations is 46 dBm, and that of LPNs is 30 dBm. For each LPN site, the candidate offsets range from 0 dB to 10 dB, with step of 1 dB, i.e., updating the offset ranges $\mathcal{S}_i'$ of all cells $i \in \mathcal{I}_2$ in each iteration k implies $numIter = 10$. Two baseline solutions, using 0 dB offset and the maximum of 10 dB offsets in all LPN cells, respectively, are used for performance comparison.

B. Results of Cell Load Estimation

Figure 2 shows the numerical results of the cell-load bounding scheme (Algorithm 1). In the figure, the first nine cells are the macro-cells, and the remaining ones are the LPN cells. The results are computed for the baseline configuration of 0 dB offset. For this configuration, the figure shows the lower and upper bounds of the first two iterations of Algorithm 1. The exact load solution $\boldsymbol{\rho}^*$, obtained by solving the non-linear system (3) directly, is also shown for reference. One can observe that, with just two iterations, the bounds have become tight for most cells. In particular, the upper bound after the second iteration virtually coincides with the exact load $\boldsymbol{\rho}^*$. Hence the results demonstrate the fast convergence of the algorithm. Note that two LPN cells, with cell IDs 16 and 25, respectively, are empty for the given offset configuration.

C. Results of Load Balancing

Figure 3 illustrates the cell load with optimized offset, obtained by the load balancing algorithm in Section IV-A, and compares the results to those of the two baseline solutions. For

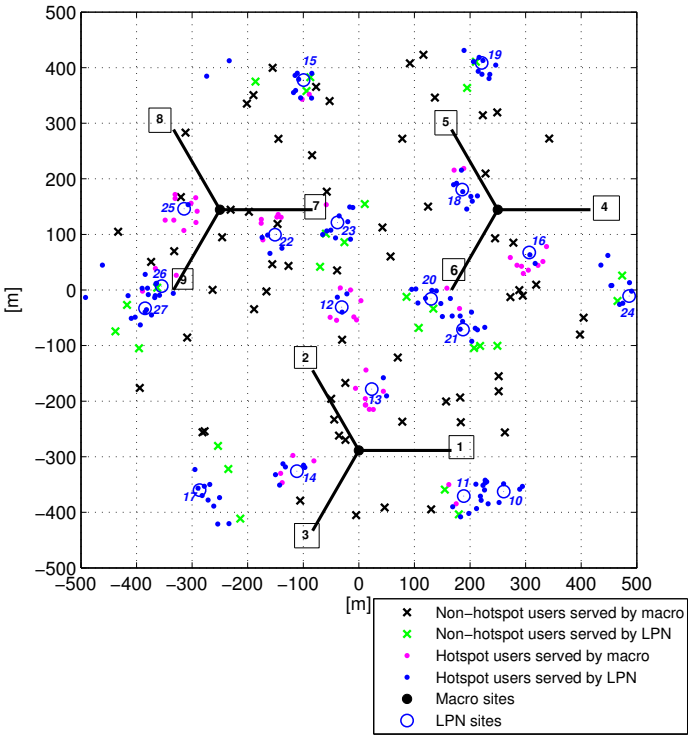

Fig. 1. Network layout with UE coverage for the optimized cell offsets. (Macro-cell IDs: 1–9, LPN cell IDs: 10–27.)

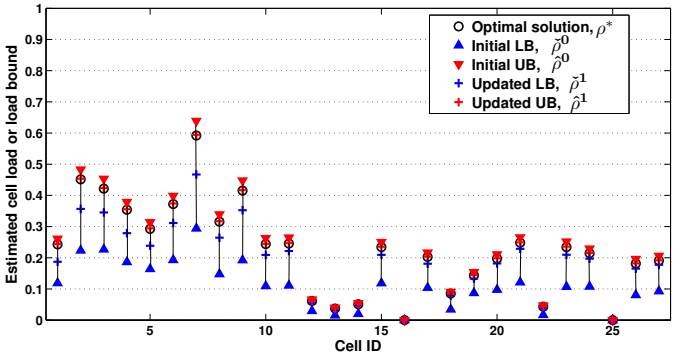

Fig. 2. Cell load estimation for the reference solution with zero cell offset based on the bounding scheme. (Macro-cell IDs: 1–9, LPN cell IDs: 10–27.)

the optimized solution, the cell offset values (in dB), returned by the algorithm, are shown to the right of the corresponding cell markers. In every iteration k, the searched cell offset range of each LPN cell is reduced by 1 dB by adjusting either the lower or the upper bound of the range (see Algorithm 2).

From the figure, one can observe that the outcome of optimization, represented by the dot markers, provides a much more uniform load among the cells in comparison to the two baseline solutions. Using 0 dB offset throughout, some macro-cells are quite heavily loaded, whereas a significant number of LPN cells are close to being empty. At the other extreme, the solution of 10 dB maximum offset reduces the macro-cell load, but at the same time leads to high load in some LPNs, in particular cell 23, because the output power of LPNs is low in relation to macro-cell interference. The optimized solution is able to offload macro-cells without overloading LPNs. As expected, high cell load does not necessarily imply a large offset or vice versa. For the objective function (1a), the values are 0.71 for 0 dB offset, 0.67 for 10 dB offset, and 0.90 for the optimized offset solution. Note that the last value is close to the theoretical maximum by the Jain's fairness criterion.

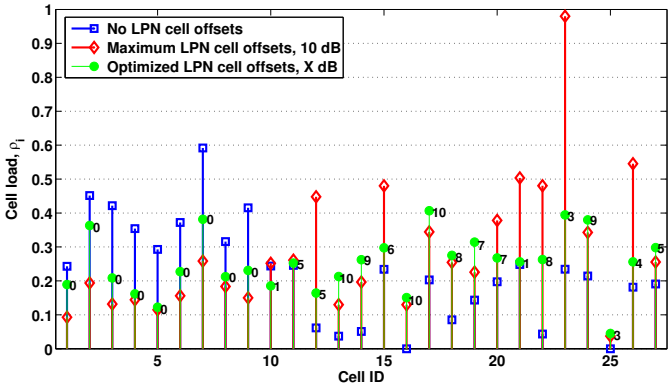

Fig. 3. Cell load levels for the optimized and two reference solutions. (Macro-cell IDs: 1–9, LPN cell IDs: 10–27.)

VI. Conclusions

We have proposed an optimization framework for load balancing in LTE HetNets by LPN range adjustment using cell-specific offsets. Performance evaluation of candidate offset solutions is based on a non-linear equation system characterizing the load-coupling relation between cells. An effective bounding scheme is developed for solving the load-coupling system. The theoretical findings, supported by formal proofs, are further demonstrated by simulation results. For a representative scenario of HetNet deployment, LPN range optimization leads to much more balanced cell load in comparison to the baseline solutions. The results also demonstrate the effectiveness of the proposed optimization procedures. As future work, one line of further investigations is more extensive simulation experiments to assess the benefits of optimizing offset. Another interesting topic is to extend the framework to other design decision, such as the power setting.

Acknowledgment

The work of Di Yuan is supported by the Swedish Foundation of Strategic Research, ELLIIT, and CENIIT.

References

[1] 3GPP TS 36.814. Evolved universal terrestrial radio access (E-UTRA); further advancements for E-UTRA physical layer aspects, v.9.0.0. `http://www.3gpp.org`

[2] R1-112429, Investigation on performance improvement by CRE, NTT DoCoMo, 3GPP TSG RAN WG1 Meeting 66, Aug. 2011.

[3] K. Hiltunen. Comparison of different network densification alternatives from the LTE downlink performance point of view. *Proc. of IEEE VTC '11*, Sep. 2011.

[4] K. Hiltunen. Comparison of different network densification alternatives from the LTE uplink point of view. *Proc. of IEEE PIMRC '11*, Sep. 2011.

[5] S. Landström, A. Furuskär, K. Johansson, L. Falconetti, and F. Kronestedt. Heterogeneous networks — increasing cellular capacity. *Ericsson Review*, vol. 1, 2011.

[6] K. Okino, T. Nakayama, C. Yamazaki, H. Sato, and Y. Kusano. Pico cell range expansion with interference mitigation toward LTE-advanced heterogeneous networks. *Proc. of ICC '11 Workshops*, June 2011.

[7] I. Siomina and S. Ahlinder. Lean optimization using supersaturated experimental design. *J. of Applied Numerical Mathematics*, vol. 58, pp. 1–15, 2008.

[8] I. Siomina, A. Furuskär, and G. Fodor. A mathematical framework for statistical QoS and capacity studies in OFDM networks, *Proc. of IEEE PIMRC '09*, Sep. 2009.

[9] I. Siomina and D. Yuan. Analysis of cell load coupling for LTE network planning and optimization. *Submitted to IEEE Transactions on Wireless Communications*, 2011. Available as *arXiv:1201.4116v1* at `http://arxiv.org/abs/1201.4116`.

[10] I. Siomina and D. Yuan. Optimization of pilot power for load balancing in WCDMA networks. *Proc. of IEEE GLOBECOM '04*, Nov./Dec. 2004.

[11] T. Qu, D. Xiao, D. Yang, W. Jin, and Y. He. Cell selection analysis in outdoor heterogeneous networks, *Proc. of IEEE ICACTE '10*, Aug. 2010.

[12] K. Valkealahti, A. Höglund, J. Pakkinen, and A. Hämäläinen. WCDMA common pilot power control for load and coverage balancing. *Proc. of IEEE PIMRC '02*, Sep. 2002.

[13] H. Zhu, T. Buot, R. Nagaike, and S. Harmen. Load balancing in WCDMA systems by adjusting pilot power. *Proc. of WPMC '02*, Oct. 2002.

Appendix

A. Proof of Theorem 2

By definition, (4) is a fixed-point iteration and $f_i(\boldsymbol{\rho})$ is Lipschitz continuous. The sequence of lower bounds generated by (4) and thus the corresponding fixed-point iterations converge to the unique fixed point, the optimal load solution $\boldsymbol{\rho}^*$, when there exists Lipschitz constant $L < 1$ such that $||\mathbf{f}(\check{\boldsymbol{\rho}}^{k+1}) - \mathbf{f}(\check{\boldsymbol{\rho}}^k)|| \leq L \cdot ||\check{\boldsymbol{\rho}}^{k+1} - \check{\boldsymbol{\rho}}^k||$ for any starting point $\check{\boldsymbol{\rho}}^0$ such that $\boldsymbol{\rho}^* - \mathbf{f}(\mathbf{0}) \leq \check{\boldsymbol{\rho}}^0 \leq \boldsymbol{\rho}^* + \mathbf{f}(\mathbf{0})$. The Lipschitz condition holds because $||\mathbf{f}(\check{\boldsymbol{\rho}}^{k+1}) - \mathbf{f}(\check{\boldsymbol{\rho}}^k)|| < ||\mathbf{f}(\check{\boldsymbol{\rho}}^k) - \mathbf{f}(\check{\boldsymbol{\rho}}^{k-1})|| = ||\check{\boldsymbol{\rho}}^{k+1} - \check{\boldsymbol{\rho}}^k||$ due to the monotonic increase and concavity properties of $\mathbf{f}(\cdot)$. Equivalently, there exists $L = 1-\varepsilon, \varepsilon \in (0,1)$, such that $||\mathbf{f}(\check{\boldsymbol{\rho}}^{k+1}) - \mathbf{f}(\check{\boldsymbol{\rho}}^k)|| \leq (1-\varepsilon) \cdot ||\check{\boldsymbol{\rho}}^{k+1} - \check{\boldsymbol{\rho}}^k||$. □

B. Proof of Theorem 3

The two upper bounds $\hat{\boldsymbol{\rho}}^k$ and $\hat{\boldsymbol{\rho}}^\ell$ are solutions to $\boldsymbol{\rho} = \check{\mathbf{H}}^k(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^k) + \mathbf{f}(\check{\boldsymbol{\rho}}^k)$ and $\boldsymbol{\rho} = \check{\mathbf{H}}^\ell(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^\ell) + \mathbf{f}(\check{\boldsymbol{\rho}}^\ell)$, respectively, with $\boldsymbol{\rho} \geq \mathbf{0}$. Here, $\check{\mathbf{H}}^k$ and $\check{\mathbf{H}}^\ell$ are defined for $\check{\boldsymbol{\rho}}^k$ and $\check{\boldsymbol{\rho}}^\ell$, respectively, using the construction in Property 5. Observe that solving the linear system $\boldsymbol{\rho} = \check{\mathbf{H}}^k(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^k) + \mathbf{f}(\check{\boldsymbol{\rho}}^k)$ is equivalent to the linear programming formulation $\{\max\ \mathbf{e}^T\boldsymbol{\rho}, \text{s.t. } \boldsymbol{\rho} \leq \check{\mathbf{H}}^k(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^k) + \mathbf{f}(\check{\boldsymbol{\rho}}^k)\}$, where $\mathbf{e}$ is the vector of ones. At the optimum $\hat{\boldsymbol{\rho}}^k$, which is clearly unique, the constraints are all binding. For $\boldsymbol{\rho} = \check{\mathbf{H}}^\ell(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^\ell) + \mathbf{f}(\check{\boldsymbol{\rho}}^\ell)$, an equivalent linear program is $\{\max\ \mathbf{e}^T\boldsymbol{\rho}, \text{s.t. } \boldsymbol{\rho} \leq \check{\mathbf{H}}^\ell(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^\ell) + \mathbf{f}(\check{\boldsymbol{\rho}}^\ell), \boldsymbol{\rho} \geq \boldsymbol{\rho}^k\}$. The validity of the last constraint follows from $\check{\boldsymbol{\rho}}^\ell \prec \check{\boldsymbol{\rho}}^k$.

Consider any cell i, and let $\check{h}_i^k$ and $\check{h}_i^\ell$ denote row i in matrices $\check{\mathbf{H}}^k$ and $\check{\mathbf{H}}^\ell$, respectively. Then $\check{h}_i^k(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^k) + f_i(\check{\boldsymbol{\rho}}^k) \leq \check{h}_i^\ell(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^\ell) + f_i(\check{\boldsymbol{\rho}}^\ell)$, for all $\boldsymbol{\rho} \geq \check{\boldsymbol{\rho}}^k$, by the matrix definition (see Property 5) and that function $\mathbf{f}$ is strictly concave (Property 1). Moreover, as $\check{\boldsymbol{\rho}}^\ell \prec \check{\boldsymbol{\rho}}^k$, the inequality is strict for at least one cell. Hence $\hat{\boldsymbol{\rho}}^\ell$ does not satisfy $\boldsymbol{\rho} \leq \check{\mathbf{H}}^k(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^k) + \mathbf{f}(\check{\boldsymbol{\rho}}^k)$, implying $\hat{\boldsymbol{\rho}}^k \neq \hat{\boldsymbol{\rho}}^\ell$. One can now conclude that $\hat{\boldsymbol{\rho}}^k \prec \hat{\boldsymbol{\rho}}^\ell$, because the rows in $\boldsymbol{\rho} \leq \check{\mathbf{H}}^k(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^k) + \mathbf{f}(\check{\boldsymbol{\rho}}^k)$ are either equal to, or strictly more stringent than the corresponding rows in $\boldsymbol{\rho} \leq \check{\mathbf{H}}^\ell(\boldsymbol{\rho} - \check{\boldsymbol{\rho}}^\ell) + \mathbf{f}(\check{\boldsymbol{\rho}}^\ell)$, for $\boldsymbol{\rho} \geq \check{\boldsymbol{\rho}}^k$. □

C. Proof of Theorem 5

By the construction in Property 5, the linear system $\mathbf{H}^*$ defined for $\boldsymbol{\rho}^*$ also has $\boldsymbol{\rho}^*$ as its solution, because the linear functions represent the tangent planes at $\boldsymbol{\rho}^*$. In other words, $\boldsymbol{\rho}^* = (\mathbf{I} - \mathbf{H}^*)^{-1}(\mathbf{f}(\boldsymbol{\rho}^*) - \mathbf{H}^*\boldsymbol{\rho}^*)$. As the sequence $\check{\boldsymbol{\rho}}^k$ approaches $\boldsymbol{\rho}^*$, the sequence of matrices $\check{\mathbf{H}}^k$, by the construction in Property 5, becomes arbitrarily close to $\mathbf{H}^*$ when k grows. Let $\boldsymbol{\epsilon}_{\check{\boldsymbol{\rho}}}^k = \boldsymbol{\rho}^* - \check{\boldsymbol{\rho}}^k$ and $\boldsymbol{\epsilon}_{\check{\mathbf{H}}}^k = \check{\mathbf{H}}^k - \mathbf{H}^*$. The upper bound $\hat{\boldsymbol{\rho}}^k$ is then given by $\hat{\boldsymbol{\rho}}^k = (\mathbf{I} - \mathbf{H}^* - \boldsymbol{\epsilon}_{\check{\mathbf{H}}}^k)^{-1}(\mathbf{f}(\boldsymbol{\rho}^* - \boldsymbol{\epsilon}_{\check{\boldsymbol{\rho}}}^k) - (\mathbf{H}^* + \boldsymbol{\epsilon}_{\check{\mathbf{H}}}^k)(\boldsymbol{\rho}^* - \boldsymbol{\epsilon}_{\check{\boldsymbol{\rho}}}^k))$. The first term can be approximated using elementary results of matrix approximation. Since $\lim_{k\to\infty} \boldsymbol{\epsilon}_{\check{\boldsymbol{\rho}}}^k = \mathbf{0}$ and $\lim_{k\to\infty} \boldsymbol{\epsilon}_{\check{\mathbf{H}}}^k = \mathbf{0}$, and function $\mathbf{f}$ is continuous, $\mathbf{f}(\boldsymbol{\rho}^* - \boldsymbol{\epsilon}_{\check{\boldsymbol{\rho}}}^k)$ approaches $\mathbf{f}(\boldsymbol{\rho}^*)$, and all the other terms involving $\boldsymbol{\epsilon}_{\check{\boldsymbol{\rho}}}^k$ and $\boldsymbol{\epsilon}_{\check{\mathbf{H}}}^k$ approach zeros. Hence $\hat{\boldsymbol{\rho}}^k$ converges to $\boldsymbol{\rho}^*$. □

D. Proof of Theorem 6

Consider any cell i, and let $\hat{h}_i$ denote row i in matrix $\hat{\mathbf{H}}$. Solution $\check{\boldsymbol{\rho}}$, if exists, is the unique optimum of the linear program $\{\max\ \mathbf{e}^T\boldsymbol{\rho}, \text{s.t. } \boldsymbol{\rho} \leq \hat{\mathbf{H}}^k\boldsymbol{\rho} + \mathbf{f}(\mathbf{0})\}$, c.f., the proof of Theorem 3. The linear program has a non-empty feasible region, because it is easily verified that $\boldsymbol{\rho}^0$ satisfies the constraints. Hence $\check{\boldsymbol{\rho}}$ exists. Because f_i is strictly concave (Property 1), $\hat{h}_i\boldsymbol{\rho} + f_i(0) \leq f_i(\boldsymbol{\rho})$, $\forall \boldsymbol{\rho} < \boldsymbol{\rho}^*$, and consequently $\check{\boldsymbol{\rho}} \leq \boldsymbol{\rho}^*$. In addition, $\mathbf{H}^0 < \hat{\mathbf{H}}$ in all the non-diagonal elements, because $\mathbf{H}^0$ in Property 4 is defined using the limit values of partial derivatives [9]. It follows then $\check{\boldsymbol{\rho}} > \boldsymbol{\rho}^0$. □