On Power and Load Coupling in Cellular Networks for Energy Optimization

Chin Keong Ho, Di Yuan, Lei Lei, and Sumei Sun

***Abstract*—We consider the problem of minimization of sum transmission energy in cellular networks where coupling occurs between cells due to mutual interference. The coupling relation is characterized by the signal-to-interference-and-noise-ratio (SINR) coupling model. Both cell load and transmission power, where cell load measures the average level of resource usage in the cell, interact via the coupling model. We present fundamental insights on the optimization problem, where power and load are to be set for meeting the target amount of data of users. We prove analytically that operating at full load is optimal in minimizing sum energy, and provide theoretical characterizations to enable the computation of the power solution for a given load vector. The insights lead to an iterative power adjustment algorithm with guaranteed convergence. We present numerical results illustrating the theoretical findings for a real-life and large-scale cellular network, showing the advantage of our solution compared to the conventional solution of deploying uniform power for base stations.**

***Index Terms*—Cellular networks, energy minimization, load coupling, power coupling, power adjustment allocation, standard interference function.**

I. Introduction

Data traffic is projected to grow at a compound annual growth rate of 78% from 2011 to 2016 [1], fueled mainly by multimedia mobile applications. This growth will lead to rapidly rising energy cost [2]. In recent years, information communication technology (ICT) has become the fifth largest industry in power consumption [3]. In cellular networks, in particular, base stations consume a significant fraction of the total end-to-end energy [4], of which 50%–80% of the power consumption is due to the power amplifiers [5], [6]. This observation has motivated green communication techniques for cellular networks [7]–[14]. These technologies include adaptive approaches such as switching off power amplifiers to provide a tradeoff of energy efficiency and spectral efficiency [7], [8], selectively turning off base stations [9], as well as energy minimization approaches for relay systems [10], OFDMA systems [11]–[13], and SC-FDMA systems [14]. Extensive survey of other saving-energy approaches are highlighted in [2], [15], [16].

In this paper, we focus on the important problem of minimizing the sum energy used for transmission in cellular networks. Besides reducing the energy cost for transmission, minimizing the transmission energy may lead to selection of power amplifiers with lower power rating, hence further reducing the overhead cost involved in turning on power amplifiers.

In a cellular network where base stations are coupled due to mutual interference, the problem of energy minimization is challenging, as each cell has to serve a target amount of data to its set of users, so as to maintain an appropriate level of service experience, subject to the presence of the coupling relation between cells. To tackle this energy minimization problem, we employ an analytical signal-to-interference-and-noise-ratio (SINR) model that takes into account the load of each cell [17]–[19], where a load of a cell translates into the average level of usage of resource (e.g., resource units in OFDMA networks) in the cell. This load-coupling equation system has been shown to give a good approximation for more complicated load models that capture the dynamic nature of arrivals and service periods of data flows in the network [20], especially at high data arrival rates. Further comparison of other approximation models concluded that the load-coupled model is accurate yet tractable [21]. By using this tractable model, useful insights can then be developed for the design of practical cellular systems. In our recent works [22], [23], we have used the load coupling equation to maximize sum utility that is an increasing function of the users' rates.

Previous works [17]–[20], [22], [23] using the load-coupling model all assume given and fixed transmission power. For transmission energy minimization, both power and load become variables and they interact in the coupling model, making the analysis more challenging. In fact, the coupling relation between cell powers cannot be expressed in closed form even for given cell loads. The key aspects motivating our theoretical and algorithmic investigations are as follows. First, is there an insightful characterization of the operating point in terms of load that minimizes the sum transmission energy? Second, given a system operating point in load, what are the properties of the coupling system in power? Third, even if power coupling cannot be expressed in closed form, is there some algorithm that converges to the power solution for given cell load?

Toward these ends, our contributions are as follows. We show that if full load is feasible, i.e., the users' data requirements can be satisfied, then operating at full load is optimal in minimizing sum transmission energy (Section IV-C, Theorem 1). If full load is not feasible, however, then no feasible solution exists (Section IV-C, Corollary 1). Thus, full load is necessary and sufficient to achieve the minimum transmission energy. Moreover, the optimal power allocation for all base stations is unique (Section V-B, Theorem 2),

C. K. Ho and S. Sun are with the Institute for Infocomm Research, A*STAR, 1 Fusionopolis Way, #21-01 Connexis, Singapore 138632 (e-mail: {hock, sunsm}@i2r.a-star.edu.sg).

D. Yuan and L. Lei are with the Department of Science and Technology, Linköping University, Sweden. (e-mail: {di.yuan, lei.lei}@liu.se)

and can be numerically computed based on an iterative algorithm that can be implemented iteratively at each base station (Section V-D, Algorithm 1). To prove the algorithmic result, we make use of the properties of the so-called standard interference function [24]; the proof is however non-standard, because the function of interest does not have a closed-form expression, and hence we use an implicit method verify its properties. We also characterize the load region given target data requirements over all possible power allocation (Section V-C, Theorems 3–4). Finally, we obtain numerical results to illustrate the optimality of the full-load solution on a cellular network based on a real-life scenario [25]. Compared with the conventional solution where the uniform power is used for base stations, we show the significant advantage of the power-optimal solution in terms of meeting user demand target and reducing the energy consumption.

The rest of the paper is organized as follows. Section II gives the system model of the load-coupled network. Section III formulates the energy minimization problem. Section IV characterizes the optimality of full load, while Section V derives properties of the power-coupling system and an iterative power allocation algorithm that achieves the power solution. Numerical results are given in Section VI. Section VII concludes the paper.

Notations: We denote a (tall) vector by a bold lower case letter, say $\boldsymbol{a}$, a matrix by a bold capital letter, say $\boldsymbol{A}$, and its (i, j)th element by its lower case a_{ij}. We denote a *positive* matrix as $\boldsymbol{A} > \mathbf{0}$ if $a_{ij} > 0$ for all i, j. Similarly, we denote a *non-negative* matrix as $\boldsymbol{A} \geq \mathbf{0}$ if $a_{ij} \geq 0$ for all i, j. Similar conventions apply to vectors. Finally, $\mathbf{0}$ and $\mathbf{1}$ denote the all-zeros and all-ones vectors of suitable lengths.

II. System Model

A. Preliminaries

We consider a cellular network consisting of n base stations that interfere with each other due to resource reuse. We focus on the downlink communication scenario where base station $i \in \mathcal{N} \triangleq \{1, \cdots, n\}$ transmits with power $p_i \geq 0$ per resource unit (in time and frequency). We refer to cell i interchangeably with base station i. For notational convenience, we collect all power $\{p_i\}$ as vector $\boldsymbol{p} \geq \mathbf{0}$.

We assume a given association of the users to the base stations. In this association, each base station i serves one unique group of users in set $\mathcal{J}_i$, where $|\mathcal{J}_i| \geq 1$. User $j \in \mathcal{J}_i$ is served in cell i at rate r_{ij} that has to be at least a rate demand of $d_{ij,\min} \geq 0$ nats. Thus, $d_{ij,\min}$ relates to a quality-of-service (QoS) constraint. We collect all the rates as vector $\boldsymbol{r}$ and the corresponding minimum demands as $\boldsymbol{d}_{\min} \geq \mathbf{0}$. Thus, a rate vector meets the QoS constraints if $\boldsymbol{r} \geq \boldsymbol{d}_{\min}$.

B. Load Coupling

We first consider the load coupling model for the cellular network. We denote by $\boldsymbol{x} = [x_1, \cdots, x_n]^T$ the load in the network, where $\mathbf{0} \leq \boldsymbol{x} \leq \mathbf{1}$. the fractional usage of resource units in cell i. In LTE systems, the load can be interpreted as the fraction of the time-frequency resources that are scheduled to deliver data. We model the SINR of user j in cell i as [17]–[20]

$$\mathsf{SINR}_{ij}(\boldsymbol{x}, \boldsymbol{p}) = \frac{p_i g_{ij}}{\sum_{k \in \mathcal{N} \setminus \{i\}} p_k g_{kj} x_k + \sigma^2} \tag{1}$$

where σ^2 represents the noise power and g_{ij} is the channel power gain from base station i to user j; note that $g_{kj}, k \neq i$, represents the channel gain from the interfering base stations. The SINR model (1) gives a good approximation of more complicated cellular network load models [20]. Intuitively, x_k can be interpreted as the likelihood of receiving interference from cell k on all the resource units. Thus, the combined term $(p_k g_{kj} x_k) \in [0, p_k g_{kj}]$ is interpreted as the average interference taken over time and frequency for all transmissions.

Given the SINR, we can transmit reliably at the maximum rate $\tilde{r}_{ij} = B \log(1 + \mathsf{SINR}_{ij})$ nat/s per resource block, where B is the bandwidth and $\log$ is the natural logarithm. To deliver a rate of r_{ij} nat for user j, the ith base station thus requires $x_{ij} \triangleq r_{ij}/\tilde{r}_{ij}$ resource units. We assume that M resource units are available. Thus, we get the load for cell i as $x_i = \sum_{j \in \mathcal{J}_i} x_{ij}/M$, i.e.,

$$x_i = \frac{1}{MB} \sum_{j \in \mathcal{J}_i} \frac{r_{ij}}{\log\left(1 + \mathsf{SINR}_{ij}(\boldsymbol{x}, \boldsymbol{p})\right)} \triangleq f_i(\boldsymbol{x}) \tag{2}$$

for $i \in \mathcal{N}$. Without loss of generality, we normalize r_{ij} by MB in (2) and so we set $MB = 1$. Let $\boldsymbol{f}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), \cdots, f_n(\boldsymbol{x})]^T$. In vector form, we obtain the *non-linear load coupling equation* (NLCE)

$$\text{NLCE}: \quad \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}; \boldsymbol{r}, \boldsymbol{p}) \tag{3}$$

for $\mathbf{0} \leq \boldsymbol{x} \leq \mathbf{1}$, where we have made the dependence of the load $\boldsymbol{x}$ on the rate $\boldsymbol{r}$ and power $\boldsymbol{p}$ explicit. We note that the load $\boldsymbol{x}$ appears in both sides of the equation and cannot be readily solved in closed form.

We collect the QoS constraints as $\boldsymbol{r} \geq \boldsymbol{d}_{\min}$. Without loss of generality, we assume $\boldsymbol{d}_{\min}$ is strictly positive, as those users with zero rate can be excluded from further consideration. Hence the power vector satisfies $\boldsymbol{p} > \mathbf{0}$ so as to serve all the users. Consequently, the load must be strictly positive, i.e., $\mathbf{0} < \boldsymbol{x} \leq \mathbf{1}$.

III. Energy Minimization Problem

Our objective is to minimize the sum transmission energy given by $\sum_{i=1}^{n} x_i p_i$. We note that the product $(x_i p_i)$ measures the transmission energy used by base station i, because the load x_i reflects the normalized amount of resource units used (in time and frequency) while the power p_i is the amount of energy used per resource unit.

The energy minimization problem is given by Problem $P0$.

$$P0: \quad \min_{\boldsymbol{p}>0, \boldsymbol{r}>0} \quad \boldsymbol{x}^T \boldsymbol{p} \tag{4a}$$

$$\text{s.t.} \quad \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}; \boldsymbol{r}, \boldsymbol{p}), \quad \mathbf{0} < \boldsymbol{x} \leq \mathbf{1} \tag{4b}$$

$$\boldsymbol{r} \geq \boldsymbol{d}_{\min}. \tag{4c}$$

As was mentioned earlier, the power and rate variables are strictly positive to satisfy the non-trivial QoS constraint. The first constraint (4b) is imposed to satisfy the NLCE. The second constraint (4c) is imposed so that the rate $\boldsymbol{r}$ satisfies the QoS constraint.

We denote an optimal solution to Problem $P0$ as $\boldsymbol{p}^\star, \boldsymbol{r}^\star$ and the corresponding load as $\boldsymbol{x}^\star$. We observe that the load is an implicit variable which depends on the power and rate vectors. A key challenge of Problem $P0$ is that a positive solution pair $(\boldsymbol{p}, \boldsymbol{r})$ is considered feasible only if there exists a load such that (4b) holds. Whether this existence holds is not obvious due to the non-linearity of the NLCE. As such, the convexity of the optimization problem cannot be readily established, and hence standard convex optimization techniques do not apply readily.

IV. Optimality of Full Load

In Section IV-A and Section IV-B, we consider fundamental properties of rate and load, respectively, such that there exists a power satisfying the NLCE. To study the fundamental properties, we consider the existence of a positive load satisfying $\boldsymbol{x} > \boldsymbol{0}$. The additional constraint that $\boldsymbol{x} \leq \boldsymbol{1}$ is taken into account in Section IV-C, in which we prove the key result that full load, i.e., $\boldsymbol{x} = \boldsymbol{1}$, is a necessary and sufficient condition for the solution in Problem $P0$ to be optimal.

A. Satisfiability of Rate

We first establish conditions on rate vector $\boldsymbol{r}$ such that a load $\boldsymbol{x}$ exists that satisfies the NLCE, possibly with $\boldsymbol{x} > \boldsymbol{1}$. We denote the spectral radius of matrix $\boldsymbol{A}$ as $\rho(\boldsymbol{A})$, defined as the absolute value of the largest eigenvalue of $\boldsymbol{A}$.

Lemma 1: For any power $\boldsymbol{p} > 0$, there exists a unique load $\boldsymbol{x} > \boldsymbol{0}$ satisfying the NLCE if and only if

$$\rho(\boldsymbol{\Lambda}(\boldsymbol{r})) < 1 \tag{5}$$

where the (i,k)th element of $\boldsymbol{\Lambda}(\boldsymbol{r})$ is given by

$$\lambda_{ik} = \begin{cases} 0, & \text{if } i = k; \\ \sum_{j \in \mathcal{J}_i} g_{kj} r_{ij} / g_{ij}, & \text{if } i \neq k \end{cases} \tag{6}$$

which is a function of $\boldsymbol{r}$ (but not $\boldsymbol{p}$).

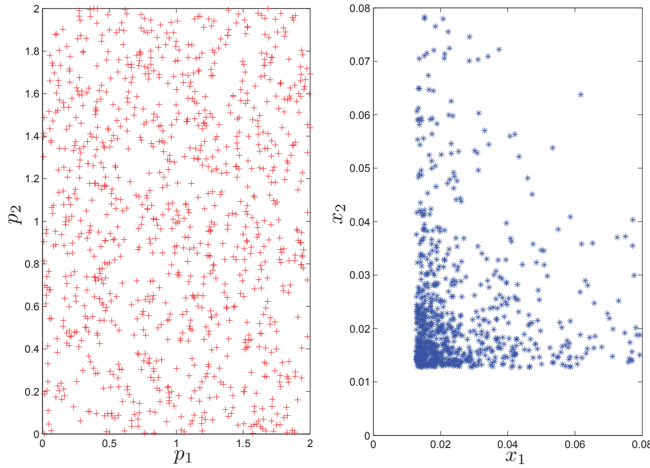
Proof: Follows directly from [23, Theorem 1]. ■

Due to Lemma 1, we say that the rate vector $\boldsymbol{r}$ is *satisfiable* if $\rho(\boldsymbol{\Lambda}(\boldsymbol{r})) < 1$. If $\boldsymbol{r}$ is not satisfiable, then there does not exist any power $\boldsymbol{p} > \boldsymbol{0}$ that results in a load satisfying constraint (4b) (which further requires the load vector to be less than one). We note that even if $\boldsymbol{r}$ is satisfiable, it is still possible that the load does not satisfy $\boldsymbol{x} \leq \boldsymbol{1}$ and hence violates (4b). Thus, satisfiability is a necessary condition for a feasible solution to exist in Problem $P0$, but it may not be sufficient.

Henceforth, we assume that a rate $\boldsymbol{r}$ is satisfiable; otherwise no feasible solution exists in Problem $P0$. Given $\boldsymbol{p}$, we can then numerically obtain $\boldsymbol{x}$ by the *iterative algorithm for load* (IAL) [23, Lemma 1], as follows. Specifically, starting from an arbitrary initial load $\boldsymbol{x}^0 > 0$, define the solution of the ℓth iteration as

$$\boldsymbol{x}^\ell = \boldsymbol{f}(\boldsymbol{x}^{\ell-1}; \boldsymbol{r}, \boldsymbol{p}) \tag{7}$$

for $\ell = 1, 2, \cdots, L$, where L is the total number of iterations. Then $\boldsymbol{x}^L$ converges to the fixed-point solution $\boldsymbol{x}$ of the NLCE as L goes to infinity. The IAL is derived using [24] by showing that $\boldsymbol{f}$ is a so-called standard interference function.



(a) Power region. (b) Load region.

Fig. 1. Corresponding power region and load region satisfying the NLCE.

B. Implementability of Load

Although Lemma 1 states that a given power vector $\boldsymbol{p}$ always corresponds to a load vector $\boldsymbol{x}$ that satisfies the NLCE, the reverse is not true. To obtain some intuition why this inverse mapping may fail, let us consider the special case of $n = 2$ base stations with channel gain $g_{ij} = 1$ for all i, j, rate $\boldsymbol{r} = \boldsymbol{1}$, and noise variance $\sigma^2 = 1$. We randomly choose the power $\boldsymbol{p} = [p_1, p_2]^T$ using a uniform distribution over $0 \leq p_i \leq 2, i = 1, 2$, which is plotted in Fig. 1(a). The corresponding load $\boldsymbol{x} = [x_1, x_2]^T$ obtained using the IAL is shown in Fig. 1(b). We see that indeed there is a load region that does not appear to correspond to any power $\boldsymbol{p} \geq \boldsymbol{0}$.

Given that $\boldsymbol{r}$ is satisfiable, we say that the load $\boldsymbol{x}$ is *implementable* if there exists power $\boldsymbol{p}$ such that the NLCE is satisfied. The following toy scenario shows that $\boldsymbol{x} = \boldsymbol{1}$ (full load) is not implementable. We assume $n = 2$ cells with one user per cell, each with rate $r = 2.1$ nat. The channel gains from a base station to the user it served and the user it does not serve is set as $g = 1$ and $g' = 1/3$, respectively. Note that the rate is satisfiable since we get $\boldsymbol{\Lambda} = \begin{bmatrix} 0 & g'r/g \\ g'r/g & 0 \end{bmatrix}$ and hence $\rho(\boldsymbol{\Lambda}) = 0.49$ and hence satisfies (5). By symmetry, the power allocated for all cells must be the same with $p_1 = p_2 = p$ and must thus satisfy (2), i.e., $\log(1 + gp/(g'p + \sigma^2)) = r$. But for any noise variance σ^2 and power p, the left hand side cannot exceed 2 nat and so the equation cannot hold. Hence, full load is not implementable in this case.

C. Main Result: Full Load is Optimal

Our first main result is given by Theorem 1, which states that full load, if implementable, is optimal to minimize the sum energy in Problem $P0$.

Theorem 1: Suppose full load, i.e., $\boldsymbol{x} = \boldsymbol{1}$, is implementable. Then the optimal solution for Problem $P0$ is as follows: $\boldsymbol{r}^\star = \boldsymbol{d}_{\min}$, and $\boldsymbol{p}^\star$ is such that $\boldsymbol{x}^\star = \boldsymbol{1}$. The optimal power vector $\boldsymbol{p}^\star$ is thus given implicitly by the NLCE as

$$\mathbf{1} = \boldsymbol{f}(\mathbf{1}; \boldsymbol{d}_{\min}, \boldsymbol{p}^\star). \tag{8}$$

Proof: The proof follows from Lemma 5 and Lemma 7 given in the Appendix, which state that $\boldsymbol{x}^\star = \mathbf{1}$ and $\boldsymbol{r}^\star = \boldsymbol{d}_{\min}$ are the optimal solutions, respectively. Substituting the optimal solutions into the NLCE results in (8). ∎

From Theorem 1, serving the minimum required rate is optimal. This observation is intuitively reasonable as less resources are used and hence less energy is consumed. Interestingly, Theorem 1 states that having full load is optimal. This second observation is not as intuitive, since it is not immediately clear the effect of using higher load on both the sum energy and interference. This is because using a high load may lead to more interference to neighbouring cells, which may then require other cells to use more energy to serve their users' rates. Mathematically, the reason can be attributed to the proof of Lemma 5, which shows that, as the power decreases, the energy as well as the interference for each cell decreases, while concurrently the load increases. Thus, by using full load, the energy is minimized.

Next, Corollary 1 provides a converse type of result to Theorem 1. The result follows from a theorem with a generalized statement, which we defer to Section V-B because the proof requires the use of algorithmic notions for finding power given load.

Corollary 1: If full load $\boldsymbol{x} = \mathbf{1}$ is not implementable, then there is no other load $\boldsymbol{x} \leq \mathbf{1}$ with $\boldsymbol{x} \neq \mathbf{1}$ that is implementable. Thus, there is no feasible solution for Problem $P0$.

Proof: The result follows as a special case of Theorem 3 later in Section V-B. ∎

Remark 1: Theorem 1 and Corollary 1 together thus show that full load is both necessary and sufficient to achieve the minimum energy in Problem $P0$.

Remark 2: It can be easily checked that Theorem 1 and Corollary 1 continue to hold even if we generalize the objective function to any function $c(x_1p_1, x_2p_2, \cdots, x_np_n)$ that is increasing in each of its argument. For example, $c(y_1, \cdots, y_n) = \sum w_i y_i$ gives the weighted sum energy with positive weights $\{w_i, i = 1, \cdots, n\}$.

V. Optimal Power Solution

Although full load is optimal for Problem $P0$, it is still not clear if the optimal power $\boldsymbol{p}^\star$ is unique and how to numerically compute $\boldsymbol{p}^\star$ in (8). Theorem 2 shall answer both questions, but in a more general setting. Namely, we provide theoretical and algorithmic results for finding power $\boldsymbol{p}$ given arbitrary load $\boldsymbol{x}$ (not necessarily all ones) and arbitrary rate $\boldsymbol{r}$ that is satisfiable (not necessarily equal to $\boldsymbol{d}_{\min}$), so as to satisfy the NLCE.

A. Standard Interference Function

Before we state the main result of the section, we introduce the standard interference function and an iterative algorithm. The algorithm shall be used to obtain the optimal power $\boldsymbol{p}^\star$, and is also a key step in the proof of the implementability of load.

We say a function $\boldsymbol{I} : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ is a standard interference function if it satisfies the following properties for all input $\boldsymbol{p} \geq \mathbf{0}$ [24]:

1) Positivity: $\boldsymbol{I}(\boldsymbol{p}) > \mathbf{0}$;
2) Monotonicity: If $\boldsymbol{p} \geq \boldsymbol{p}'$, then $\boldsymbol{I}(\boldsymbol{p}) \geq \boldsymbol{I}(\boldsymbol{p}')$.
3) Scalability: For all $\alpha > 1$, $\alpha \boldsymbol{I}(\boldsymbol{p}) > \boldsymbol{I}(\alpha \boldsymbol{p})$.

Next, we consider both the synchronous and asynchronous versions of the *iterative algorithm for power* (IAP). Assume an arbitrary initial power $\boldsymbol{p}^0 > \mathbf{0}$. In iteration $\ell = 1, \cdots, L$, of the *synchronous IAP*, where L is the total number of iterations, we obtain

$$\boldsymbol{p}^\ell = \boldsymbol{I}(\boldsymbol{p}^{\ell-1}). \tag{9}$$

Clearly, the entire power vector is used and updated iteratively, and any power element of $\boldsymbol{p}^\ell$ is solely determined by $\boldsymbol{p}^{\ell-1}$. In contrast, the *asynchronous IAP* iterates power element by power element, of which any update comes immediately into effect for the remaining elements. Specifically, for each *outer* iteration $\ell = 1, \cdots, L$, we perform n *inner* iterations. In the ith inner iteration, $i = 1, \cdots, n$, we obtain

$$p_i^\ell = \boldsymbol{I}(\boldsymbol{p}_{i-1}^{\ell-1}) \tag{10}$$

where $\boldsymbol{p}_{i-1}^{\ell-1}$ represents the power vector containing the most updated elements after $(\ell - 1)$ outer iterations and $(i - 1)$ inner iterations (during the ℓth outer iteration). After L outer iterations are fully completed, each with n inner iterations, we obtain $\boldsymbol{p}_n^L$ as the final power vector solution.

The following result is due to [24, Theorems 2,4]; we omit the proof.

Lemma 2: Suppose a fixed-point solution $\boldsymbol{p}$ exists for $\boldsymbol{p} = \boldsymbol{I}(\boldsymbol{p})$. If $\boldsymbol{I}$ is a standard interference function, then starting from any initial power vector, both the synchronous and asynchronous IAP algorithms converge to the fixed-point solution $\boldsymbol{p}$, which is unique.

B. Main Result: Existence and Computation of Power Solution

Before proving the main result, we present and prove some properties on how the elements of the power vector relate to each other in NLCE. The properties will then be used to establish that the results in [24] with the notion of standard interference function can be applied.

Let $\bar{\boldsymbol{p}}_i$ be the vector of length $(n-1)$ that contains all elements in $\boldsymbol{p}$ except for p_i. Lemma 3 shows that given $\boldsymbol{x}$ and $\boldsymbol{r}$, the dependency of p_i on $\bar{\boldsymbol{p}}_i$ (such that the NLCE holds) qualifies as a function, even if the function is not in closed form.

Lemma 3: Let $\boldsymbol{p}, \boldsymbol{x}, \boldsymbol{r}$ satisfy the NLCE, where the vectors are strictly positive. Then there exists function $h_i : \mathbb{R}^n_{++} \to \mathbb{R}^n_{++}$ satisfying $p_i = h_i(\bar{\boldsymbol{p}}_i; \boldsymbol{x}, \boldsymbol{r})$ for all $i = 1, \cdots, n$. Writing p_i's and h_i's in vector form, we get $\boldsymbol{p} = \boldsymbol{h}(\boldsymbol{p}; \boldsymbol{x}, \boldsymbol{r})$.

Proof: We fix $\boldsymbol{x}, \boldsymbol{r}$ and drop these notations in the function $h_i(\cdot)$ for simplicity. To prove the existence of the function $h_i(\cdot)$, we need to show that given $\bar{\boldsymbol{p}}_i$, there exists a unique p_i for $i = 1, \cdots, n$. First, we write the NLCE in (2) as

$$1 = \sum_{j \in \mathcal{J}_i} \frac{a_{ij}}{\log\left(1 + p_i b_{ij}(\bar{\boldsymbol{p}}_i, \sigma^2)\right)} \triangleq \eta_i(p_i) \tag{11}$$

where

$$a_{ij} \triangleq r_{ij}/x_i \tag{12}$$

$$b_{ij}(\bar{\boldsymbol{p}}_i, \sigma^2) \triangleq \frac{g_{ij}}{\sum_{k\in\mathcal{N}\setminus\{i\}} p_k g_{kj} x_k + \sigma^2} \tag{13}$$

are both independent of p_i. We fix $\bar{\boldsymbol{p}}_i > 0$ and $\sigma^2 \geq 0$. It follows that $b_{ij}(\bar{\boldsymbol{p}}_i, \sigma^2) > 0$ and so $\eta_i(p_i) > 0$. Observe that $\eta_i(p_i)$ is a strictly decreasing function of p_i. Since $\eta_i(p_i) \to \infty$ as $p_i \to 0$, and $\eta_i(p_i) \to 0$ as $p_i \to \infty$, there exists a unique $p_i > 0$ such that $\eta_i(p_i) = 1$, and thus satisfies (11). Hence there exists a function of the form $p_i = h_i(\bar{\boldsymbol{p}}_i)$, for any i. ■

Remark 3: The function $h_i(\cdot)$ does not submit to a closed-form solution. For example, consider expressing p_i in terms of $\bar{\boldsymbol{p}}_i$ in (11) where the number of summands is $|\mathcal{J}_i| > 1$. Because each of them is non-linear in p_i, the dependency of p_i on $\bar{\boldsymbol{p}}_i$ is not explicit.

Remark 4: Although $h_i(\cdot)$ cannot be expressed in closed form, we can numerically obtain the output p_i of the function h_i given the input $\bar{\boldsymbol{p}}_i$. Equivalently, this means that we want to obtain the value of p_i such that (11) holds. This is computed, for example, by a bisection search on $\eta_i(p_i) = 1$, making use of the property that $\eta_i(p_i)$ is a strictly decreasing function. Specifically, we first choose an arbitrary but small power p' such that $\eta_i(p') > 1$ and an arbitrary but large power p'' such that $\eta_i(p'') < 1$. Next we use the new power $p = (p' + p'')/2$ and evaluate if $g(p)$ is greater or smaller than one, then replace p' or p'' by p, respectively. By performing this procedure iteratively, we have guaranteed convergence to the desired p that satisfies $\eta_i(p_i) = 1$. This forms the basis for the proposed algorithm later in Section VI.

We observe that $\boldsymbol{h}(\cdot)$ is to some extent similar to $\boldsymbol{f}(\cdot)$ in the NLCE (3). From Remark 3, however, the function $\boldsymbol{h}(\cdot)$ cannot be readily written as a closed-form expression. Thus, proving properties related to $\boldsymbol{h}(\cdot)$ is more challenging, as compared to the case of $\boldsymbol{f}(\cdot)$ for which a closed-form solution is available. Nevertheless, Lemma 4 states that $\boldsymbol{h}(\cdot)$ qualifies as a standard interference function as defined in [24].

Lemma 4: Given load $\boldsymbol{x} \geq \boldsymbol{0}$ and rate $\boldsymbol{r} \geq \boldsymbol{0}$, $\boldsymbol{h}(\boldsymbol{p}; \boldsymbol{x}, \boldsymbol{r})$ is a standard interference function in $\boldsymbol{p}$.

Proof: For notational convenience, we drop the dependence of these entities in the notation of $\boldsymbol{h}(\cdot)$. We consider an arbitrary i and refer to $\eta_i(p_i), a_{ij}, b_{ij}$ as defined in (11), (12) and (13), respectively, throughout the proof. For this proof, it is useful to denote the function $\eta_i(p_i)$ explicitly as $\eta_i(p_i, \bar{\boldsymbol{p}}_i, \sigma^2)$ to ease the discussion. We prove each of the three properties required for standard interference function below.

Positivity: From the proof of Lemma 3, there exists a unique $p_i > 0$ that satisfies (11), i.e., $h_i(\bar{\boldsymbol{p}}_i) > 0$. This holds for all i, thus $\boldsymbol{h}(\boldsymbol{p}) > \boldsymbol{0}$.

Monotonicity: From (11), we observe that $\eta_i(p_i, \bar{\boldsymbol{p}}_i, \sigma^2)$ strictly increases as p_i decreases, or as any element of $\bar{\boldsymbol{p}}_i$ increases. Hence, to satisfy $\eta_i(p_i, \bar{\boldsymbol{p}}_i, \sigma^2) = 1$, p_i strictly increases if any element of $\bar{\boldsymbol{p}}_i$ increases. We note that an equivalent representation of $\eta_i(p_i, \bar{\boldsymbol{p}}_i, \sigma^2) = 1$ is $p_i = h_i(\bar{\boldsymbol{p}}_i)$. It follows that $h_i(\bar{\boldsymbol{p}}_i)$ is increasing in any of the arguments.

Scalability: Let $q_1 = h_i(\bar{\boldsymbol{p}}_i)$ and $q_2 = h_i(\alpha\bar{\boldsymbol{p}}_i)$, where $\alpha > 1$. Observe that

$$\eta_i(q_1, \bar{\boldsymbol{p}}_i, \sigma^2) = \eta_i(q_2, \alpha\bar{\boldsymbol{p}}_i, \sigma^2) \tag{14}$$

since both equal one according to (11). It is easy to check that $q_1 b_{ij}(\bar{\boldsymbol{p}}_i, \sigma^2) = \alpha q_1 b_{ij}(\alpha\bar{\boldsymbol{p}}_i, \alpha\sigma^2)$. That is, multiplying all the terms in the triplet $(q_1, \bar{\boldsymbol{p}}_i, \sigma^2)$ by a positive constant still allows (11) to be satisfied. Thus we get from (14)

$$\eta_i(\alpha q_1, \alpha\bar{\boldsymbol{p}}_i, \alpha\sigma^2) = \eta_i(q_2, \alpha\bar{\boldsymbol{p}}_i, \sigma^2). \tag{15}$$

With the second argument in $\eta_i(y, \cdot, z)$ fixed, we note that the output of the function strictly decreases with y and strictly increases with z. By the equality in (15), it follows that $\alpha q_1 > q_2$ because $\alpha\sigma^2 > \sigma^2$. Taking into account of the definition of q_1 and q_2, we have proved $\alpha h_i(\bar{\boldsymbol{p}}_i) > h_i(\alpha\bar{\boldsymbol{p}}_i)$. ■

Using Lemma 3 and Lemma 4, we are ready to provide the main result, stating that NLCE can be expressed in an alternative form with the power taken as the subject of interest. The proof is non-standard, because the relations among the power elements do not submit to a closed form (Remark 3). Hence, it has been necessary to first establish that the relation between one power element and the others qualifies as a function (Lemma 3). Next, we have used an implicit method to prove that $\boldsymbol{h}(\cdot)$ is indeed an interference function (Lemma 4).

Theorem 2: Given load $\boldsymbol{x}$ and rate $\boldsymbol{r}$, the power $\boldsymbol{p}$ that satisfies the NLCE can be represented equivalently in the form of a non-linear *power* coupling equation (NPCE) given by

$$\text{NPCE}: \quad \boldsymbol{p} = \boldsymbol{h}(\boldsymbol{p}; \boldsymbol{x}, \boldsymbol{r}) \tag{16}$$

where $\boldsymbol{h}(\cdot)$ is a standard interference function. Given that a solution $\boldsymbol{p}$ exists, then $\boldsymbol{p}$ is unique and can be obtained numerically by the IAP.

Proof: Lemma 3 states the existence of the function $\boldsymbol{h}(\cdot)$, and hence allows us to obtain the NPCE. Lemma 4 states that $\boldsymbol{h}(\cdot)$ satisfies all the properties required for a standard interference function. The uniqueness and iterative computation of $\boldsymbol{p}$ then follows from Lemma 2 with the standard interference function $\boldsymbol{h}(\cdot)$. ■

Remark 5: So far we have assumed that there is no maximum power constraint imposed for any element of power $\boldsymbol{p}$. If such power constraints are imposed, then a so-called standard constrained interference function defined in [24] can be used instead to perform the IAP, in which the output of each iteration is set to the maximum power constraint value, if that returned from $\boldsymbol{h}$ is higher. This type of iteration converges to a unique fixed point [24, Corollary 1].

C. Characterization on Implementability of Load

Theorem 3 provides a monotonicity result for load implementability. We recall that a load vector $\boldsymbol{x}$ is said to be implementable if there exists power $\boldsymbol{p}$ such that the NLCE holds.

Theorem 3: Consider two load vectors with $\boldsymbol{x}' \geq \boldsymbol{x}$ and $\boldsymbol{x}' \neq \boldsymbol{x}$. If $\boldsymbol{x}$ is implementable, then $\boldsymbol{x}'$ is implementable. Moreover, the respective corresponding powers $\boldsymbol{p}$ and $\boldsymbol{p}'$ satisfy $\boldsymbol{p}' < \boldsymbol{p}$.

Proof: Suppose $\boldsymbol{x}$ is implementable, i.e., there exists power $\boldsymbol{p}$ such that the NPCE (or equivalently the NLCE) holds. From Theorem 3, $\boldsymbol{h}(\cdot)$ is a standard interference function. We shall prove that $\boldsymbol{x}'$ is also implementable, i.e., $\boldsymbol{p}'$ exists.

Before we consider the general case of $\boldsymbol{x}' \geq \boldsymbol{x}$, we first focus on the special case that strict inequality holds only for the first element (with re-indexing if necessary), i.e., $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]^T$ and $\boldsymbol{x}' = [x_1', x_2, \cdots, x_n]^T$ with $x_1' > x_1$. We now use the asynchronous IAP (10) with load $\boldsymbol{x}'$, and we set the initial power as $\boldsymbol{p}^0 = \boldsymbol{p}$. Our objective is to show that the power converges to $\boldsymbol{p}'$ that satisfies the NPCE.

Consider the asynchronous IAP (10) with outer iteration $\ell = 1$ and inner iteration $i = 1, 2, \cdots, n$:

- For $i = 1$: Consider the NLCE for cell 1 with the original load $\boldsymbol{x}$ and power $\boldsymbol{p}$:

$$x_1 = \sum_{j \in \mathcal{J}_1} \frac{r_{1j}}{\log\left(1 + \frac{p_1 g_{1j}}{\sum_{k \geq 2} p_k g_{kj} x_k + \sigma^2}\right)}. \tag{17}$$

  In the first iteration, x_1 and p_1 are updated by the actual load of interest x_1' and the iterated power p_1^1, respectively, with other load and power unchanged. Since $x_1' > x_1$, we must have $p_1^1 < p_1$.
  From the proof in Lemma 5, the energy $e_1 \triangleq p_1 x_1$ with p_1, x_1 given by (17) satisfies $\partial e_1/\partial p_1 > 0$. Since $\partial e_1/\partial x_1 = \partial e_1/\partial p_1 \cdot \partial p_1/\partial x_1$ and clearly $\partial p_1/\partial x_1 < 0$, we get $\partial e_1/\partial x_1 < 0$. Thus, $p_1^1 x_1' < p_1 x_1$.
- For $i = 2$: We shall show that the iterated power satisfies $p_2^1 < p_2^0 = p_2$. The NLCE for cell 2 with the original load $\boldsymbol{x}$ and power $\boldsymbol{p}$ can be written as:

$$x_2 = \sum_{j \in \mathcal{J}_2} \frac{r_{2j}}{\log\left(1 + \frac{p_2 g_{2j}}{p_1 g_{1j} x_1 + \sum_{k \geq 3} p_k g_{kj} x_k + \sigma^2}\right)}$$

  Upon updating cell 2, we have updated x_1, p_1 to the newly iterated x_1', p_1^1, respectively. Since $p_1^1 x_1' < p_1 x_1$ as mentioned earlier, $p_2^1 < p_2$.
- For $i \geq 3$: For subsequent iterations, it can be shown similarly that $p_i^1 < p_i^0 = p_i$ for $i = 3, \cdots, n$. This completes the first outer iteration.

At this point, we get $\boldsymbol{p}_n^1 < \boldsymbol{p}$. It can be similarly shown that $\boldsymbol{p}_n^{\ell+1} < \boldsymbol{p}_n^{\ell}$ for $\ell > 1$.

For large number of iterations L, the decreasing sequence $\boldsymbol{p}_n^0, \boldsymbol{p}_n^1, \cdots$ must converge since $\boldsymbol{p}_n^{\ell} \geq 0$ (i.e., it is bounded from below) for any ℓ due to the positivity of the standard interference function. Thus, the power solution exists, i.e., $\boldsymbol{x}'$ is implementable.

At convergence, we have $\lim_{L \to \infty} \boldsymbol{p}_n^L = \boldsymbol{p}' < \boldsymbol{p}$. So far we have assumed that only one element of the load is strictly increased. In general, if more than one load element is increased, repeating the argument sequentially for every such element proves that power exists and is decreased. Thus, in general $\boldsymbol{x}'$ is implementable for $\boldsymbol{x}' \geq \boldsymbol{x}$, where $\boldsymbol{p}' < \boldsymbol{p}$. ■

From Theorem 3, we also obtain the equivalent result that $\boldsymbol{x}$ is not implementable if $\boldsymbol{x}'$ is not implementable. Moreover, by using Theorem 3, we conclude immediately that full load minimizes the sum energy (or minimizes the more general objective function such as sum weighted energy as given in Remark 2). Thus, an alternative proof for Theorem 1 is obtained by applying Theorem 3; the proof in Theorem 3 however requires the concept of standard interference function which is not needed in the proof of Theorem 1.

The next theoretical characterization is on the implementable load region $\mathcal{L}$ over all non-negative power vectors for any given satisfiable rate $\boldsymbol{r}$, i.e., $\mathcal{L} \triangleq \{\boldsymbol{x} \geq \boldsymbol{0} : \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x}; \boldsymbol{r}, \boldsymbol{p}), \boldsymbol{p} \geq \boldsymbol{0}\}$. Theorem 4 states that the boundary of this region is open.

Theorem 4: Suppose load $\boldsymbol{x}$ is implementable with power $\boldsymbol{p}$ and rate $\boldsymbol{r}$. Then there exists $\delta > 0$, such that any load vector $\boldsymbol{x}'$ with $\|\boldsymbol{x}' - \boldsymbol{x}\| \leq \delta$ is implementable. Moreover, the implementable load region $\mathcal{L}$ is open.

Proof: Let $\widetilde{\boldsymbol{p}} = \beta \boldsymbol{p}$ with $\beta > 1$, and let the corresponding load satisfying the NLCE with rate $\boldsymbol{r}$ be $\widetilde{\boldsymbol{x}}$. Note that $\widetilde{\boldsymbol{x}}$ exists, because the existence of load does not depend on power (cf. Lemma 1). By applying the IAL in (7) to obtain $\widetilde{\boldsymbol{x}}$ (using power $\widetilde{\boldsymbol{p}}$) with the initial load set as $\boldsymbol{x}^0 = \boldsymbol{x}$, it can be easily checked that the load vector decreases in every iteration. Since $\widetilde{\boldsymbol{x}} > \boldsymbol{0}$, the iterations must converge to $\widetilde{\boldsymbol{x}} = \lim_{L \to \infty} \boldsymbol{x}^L < \boldsymbol{x}$. By Theorem 3, any $\boldsymbol{x}' \geq \widetilde{\boldsymbol{x}}$ is implementable. As $\widetilde{\boldsymbol{x}} < \boldsymbol{x}$, there is an implementable neighbourhood of $\boldsymbol{x}$. That is, there exists $\delta > 0$, for which any load vector $\boldsymbol{x}'$ satisfying $\|\boldsymbol{x}' - \boldsymbol{x}\| \leq \delta$ is implementable. Since the result holds for any $\boldsymbol{x}$ in $\mathcal{L}$, it follows that $\mathcal{L}$ is open. ■

D. Algorithm for Optimal Power Vector

By Theorem 2, we can use the IAP to compute the optimal power $\boldsymbol{p}^\star$ for (8) in Theorem 1 for implementable load $\boldsymbol{x}^\star$. To obtain the output of the function $\boldsymbol{h}(\cdot)$ in each step of the IAP, bisection search is able to determine the power p_i such that $\eta_i(p_i) = 1$ (see Remark 4). Putting together the theoretical insights results in the following formal algorithmic description (Algorithm 1) for computing $\boldsymbol{p}^\star$.

Given:
- target load vector $\boldsymbol{x}^\star = [x_1^\star, x_2^\star, \cdots, x_n^\star]^T$
- rate vector $\boldsymbol{r}$ such that $\rho(\boldsymbol{\Lambda}(\boldsymbol{r})) < 1$
- arbitrary initial power vector $\boldsymbol{p}$
- tolerance $\epsilon > 0$

Output: $\boldsymbol{p}^\star$ with $\boldsymbol{x}^\star = \boldsymbol{f}(\boldsymbol{x}^\star; \boldsymbol{r}, \boldsymbol{p}^\star)$

Initialize $\boldsymbol{x} \leftarrow \boldsymbol{f}(\boldsymbol{x}^\star; \boldsymbol{r}, \boldsymbol{p})$.
while $\|\boldsymbol{x} - \boldsymbol{x}^\star\| > \epsilon$ **do**
  for $i = 1 : n$ **do**
    $p_i^{\text{left}} \leftarrow \xi$ for any ξ such that $\eta_i(\xi) > 1$
    $p_i^{\text{right}} \leftarrow \psi$ for any ψ such that $\eta_i(\psi) < 1$
    while $|\eta_i(p_i) - 1| > \epsilon$ **do**
      if $\eta_i(p_i) \leq 1$ **then**
        $p_i^{\text{right}} \leftarrow p_i$
      else if $\eta_i(p_i) > 1$ **then**
        $p_i^{\text{left}} \leftarrow p_i$
      end if
      $p_i \leftarrow (p_i^{\text{left}} + p_i^{\text{right}})/2$
    end while
  end for
  $\boldsymbol{x} \leftarrow \boldsymbol{f}(\boldsymbol{x}^\star; \boldsymbol{r}, \boldsymbol{p})$
end while
$\boldsymbol{p}^\star \leftarrow \boldsymbol{p}$, return $\boldsymbol{p}^\star$

Algorithm 1: IAP algorithm for computing optimal power.

TABLE I
NETWORK AND SIMULATION PARAMETERS

Parameter	Value
Service area size	7500 × 7500 m^2
Pixel resolution	50 × 50 m^2
Number of sites	50
Number of cells	148
Number of pixels	22,500
Number of users	1480
Thermal noise spectral density	-145.1 dBm/Hz
Total bandwidth per cell	4.5 MHz
Bandwidth per resorce unit	180 KHz
Tolerance ϵ in IAP	10^{-6}
Initial power vector $\boldsymbol{p}$ in IAP	$\mathbf{1}$ W

Algorithm 1 solves the NPCE for given $\boldsymbol{x}^{\star}$, by iteratively updating the power vector and re-evaluating the resulting load $\boldsymbol{f}(\boldsymbol{x}^{\star};\boldsymbol{r},\boldsymbol{p})$. The bulk of the algorithm starts at Line 2. The outer loop terminates if the load vector $\boldsymbol{x}$ has converged to $\boldsymbol{x}^{\star}$. For each outer iteration, the inner loop is run starting at Line 3, for which the power vector for each cell i is updated. In each update, the power range is first initialized to $[\xi,\psi]$, where $\xi<\psi$, such that $\eta_i(\xi)>1$ and $\eta_i(\psi)<1$. Since the function $\eta_i(\cdot)$ is a strictly decreasing function, the bisection search from Lines 7-12 ensures convergence to the unique solution for $\eta_i(p_i)=1$, or equivalently, the value of $h_i(\bar{\boldsymbol{p}}_i;\boldsymbol{x},\boldsymbol{r})$. Load re-evaluation is then carried out in Line 15.

VI. NUMERICAL EVALUATION

A. Simulation Setup

In this section, we provide numerical results to illustrate the theoretical findings. The simulations have been performed for a real-life based cellular network scenario, with publicly available data provided by the European MOMENTUM project [25]. In our simulations, the data we have used are derived from real measurements for the network of a sub-area of Alexanderplatz in the city of Berlin. The scenario is illustrated in Fig. 2. The scenario has 50 base station sites, sectorized into 148 cells. In Fig. 2, the red dots indicate base station sites and the green dots represent the location of users. Most of the sites have three sectors (cells) equipped with directional antennas. The blue short lines represent the antenna directions of the cells. The entire service area of the Berlin network scenario is divided into $22,500$ pixels as shown Fig. 2. That is, each pixel represents a small square area, with resolution 50 × 50 m^2, for which signal propagation is considered uniform. Users located in the same pixel are assumed to have the same channel gains. The cell-pixel gain values originate from real measurements. In our simulations, each cell serves up to ten randomly distributed users in its serving area as defined in the MOMENTUM data set. The total bandwidth MB of each cell is 4.5 MHz. Following the LTE standards, we use one resource block to represent a resource unit with 180 KHz bandwidth each in the simulation. Network and simulation parameters are summarized in Table I.

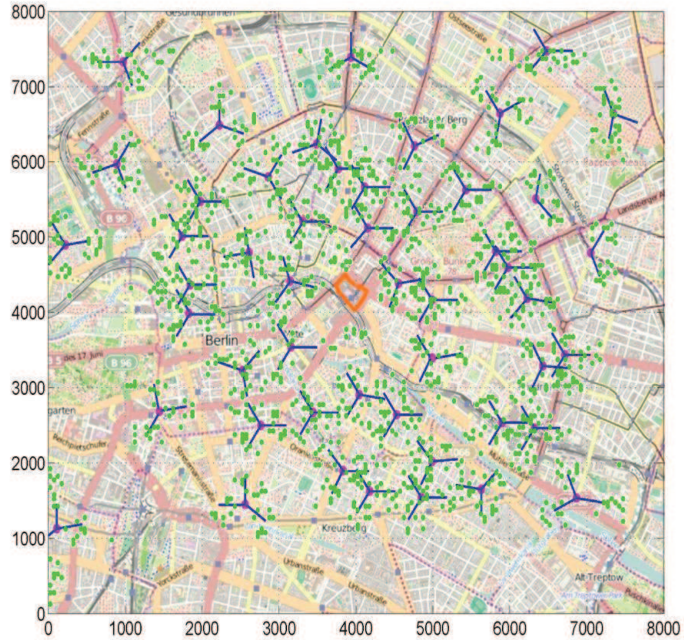


Fig. 2. Network layout and user distribution in an area of Alexanderplatz, Berlin. The units of the axes are in meters. Digital Map: © OpenStreetMap contributors, the map data is available under the Open Database License.

B. Results

Our objective is to numerically illustrate the relationship among the load, power, and sum transmission energy. First, we consider the use of uniform load with $\boldsymbol{x}=\phi\mathbf{1}$ for various $0<\phi\leq 1$, with $\phi=1$ being the case of full load. Given the load vector $\boldsymbol{x}$, the optimal power solution $\boldsymbol{p}$ is then obtained by using the IAP described by Algorithm 1. Next, for benchmarking, we consider the conventional scheme that employs uniform power allocation $\boldsymbol{p}=\beta\mathbf{1},\beta>0$. We choose β that results in the minimum sum energy subject to the constraint that the corresponding load satisfies $\mathbf{0}\leq\boldsymbol{x}\leq\mathbf{1}$, as follows. From the proof of Lemma 5, the energy (given by the product of load and power) for each cell strictly decreases as the power strictly decreases. Thus, to minimize the sum energy, we choose the smallest β such that $\boldsymbol{x}\leq\mathbf{1}$; this can be obtained by a bisection search starting with sufficiently small and large values of β. For any β under consideration, the IAL is used to obtain the load corresponding to the power $\boldsymbol{p}=\beta\mathbf{1}$.

In the first numerical experiment, we consider the sum energy for rate demand $\boldsymbol{r}=\xi\mathbf{1}$ with ξ being successively increased, while keeping $\boldsymbol{r}$ satisfiable. Fig. 3 compares the sum energy for various uniform load levels, including full load, and that obtained uniform power allocation. From Fig. 3, the sum energy for all cases appear to grow exponentially fast as the rate demand increases, and will approach infinity as the rate demand approaches the respective dotted vertical lines. The lines correspond to the boundary when the rate demand is not satisfiable, i.e., $\rho(\boldsymbol{\Lambda}(\boldsymbol{r}))=1$. This behaviour is consistent with Lemma 1. Deploying full load achieves the smallest sum energy, in accordance with Theorem 5. The reduction in sum energy is particularly evident in comparison to the scheme

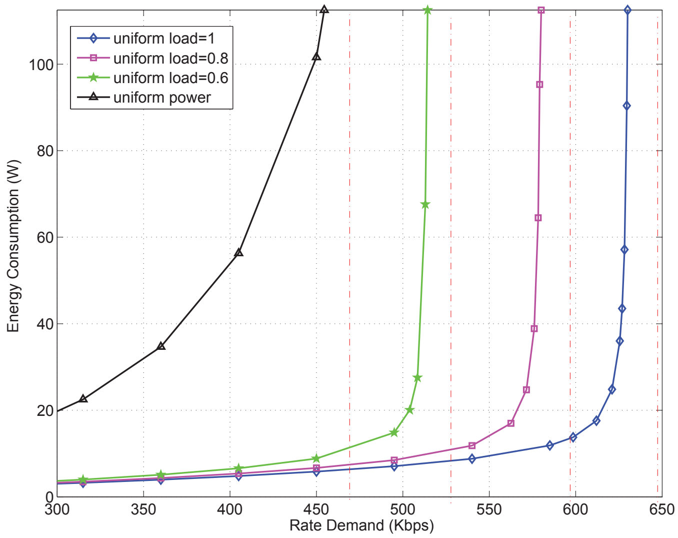


Fig. 3. Sum Transmission energy with respect to user's rate demand.

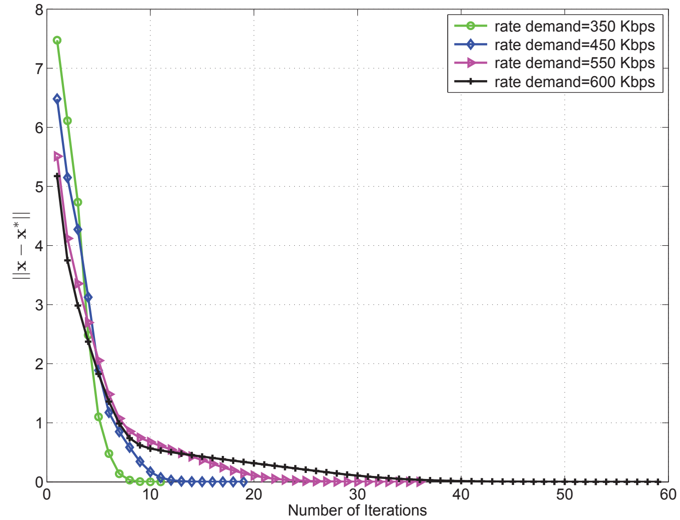


Fig. 5. The evolution of the Euclidean distance between the iterate $\boldsymbol{x}$ and the target $\boldsymbol{x}^\star$, given by $\|\boldsymbol{x}-\boldsymbol{x}^\star\|$, over iterations with tolerance $\epsilon = 10^{-6}$ and $\boldsymbol{x}^\star = \mathbf{1}$.

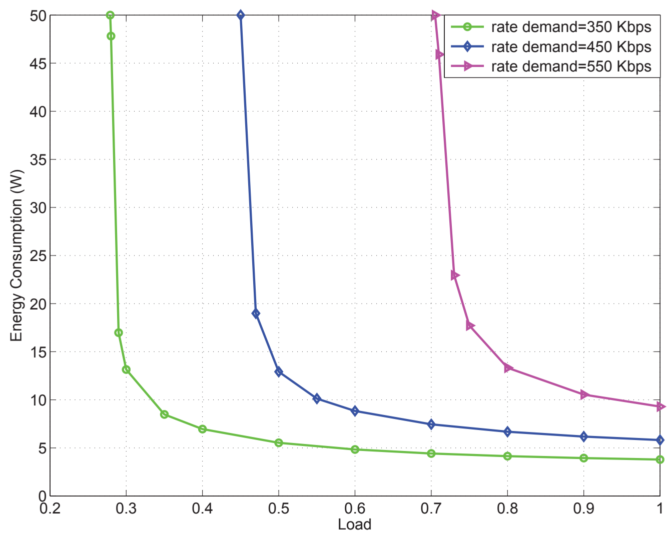


Fig. 4. Sum transmission energy with respect to cell load.

of uniform power – the relative saving is 90% or higher for the rate demand shown in Fig. 3. Conversely, for a fixed amount of sum energy, deploying full load and optimizing the corresponding power allows for maximizing the rate demand that can be served.

Next, we examine the energy consumption by progressively increasing the uniform load for three rate demand levels $\boldsymbol{r} = \xi\mathbf{1}$ with ξ taking the values of 350 Kbps, 450 Kbps and 550 Kbps. The results are shown in Fig. 4. We observe that the sum energy decreases monotonically by increasing the load. The reduction of sum energy appears to be exponentially fast in the low-load regime, but is much slower in the high-load regime. In addition, the numerical results also reinforce the fact that some load vectors are not implementable. In particular, it is not always possible to obtain a power vector $\boldsymbol{p}$ for a load vector $\boldsymbol{x} = \phi\mathbf{1}$ with very small $\phi > 0$. From Fig. 4, the sum energy surges to infinity when the load approaches some fixed (small) value, which suggests that, for any $\boldsymbol{r} > 0$, the load cannot become arbitrarily small, irrespective of power.

In the last part of experiments, we investigate the convergence behavior of the IAP. In Fig. 5, we set the target load vector $\boldsymbol{x}^\star = \mathbf{1}$ and the initial power vector $\boldsymbol{p} = \mathbf{1}$ Watt, with rate demand $\boldsymbol{r} = \xi\mathbf{1}$ where $\xi \in \{350, 450, 550, 600\}$ Kpbs. The Euclidean distance between the iterate $\boldsymbol{x}$ and target $\boldsymbol{x}^\star$ is given by $\|\boldsymbol{x}-\boldsymbol{x}^\star\|$. The evolution of $\|\boldsymbol{x}-\boldsymbol{x}^\star\|$ for the four different rate demand cases is illustrated in Fig. 5. We consider the algorithm as converged if $\|\boldsymbol{x}-\boldsymbol{x}^\star\| \leq \epsilon$, with $\epsilon = 10^{-6}$. For the four rate demand cases, convergence is reached after 11, 19, 36 and 59 iterations, respectively. Given the size of the network (148 cells), the values are moderate. Also, we notice that when the rate demand increases, more iterations are required for convergence with a longer tail-off. This is mainly because of a high rate demand which means that, in general, the NPCE is operating in the high SINR regime. The amount of progress in load in an IAP iteration is mainly dependent on the denominator in (3). For high SINR regime, the relative change in load is lesser due to the logarithm operator, thus slowing down the progress. Moreover, the number of iterations depends on the initial power point. In general, fewer iterations are required if the starting power point is closer to the optimum. Note that no matter what the initialization is, the convergence of the IAP is guaranteed by Theorem 2.

In case of the presence of some time constraints in a practical application, the IAP may be terminated before full convergence is reached. Thus, the capability of delivering a load-feasible and close-to-convergent solution within few iterations is of significance. It can be seen in Fig. 5 that a majority of the iterations is due to the tailing-off effect – the load vector is in fact close to the target value within about half of the iterations. For all the rate demand levels, convergence is in effect achieved in less than 20 iterations; this is promising for the practical relevance of the proposed IAP scheme. Finally, to ensure that the load is strictly less than full load for practical implementation, we may set $\boldsymbol{x}^\star = (1-\epsilon')\mathbf{1}$ with $\epsilon' > \epsilon$.

VII. Conclusion

We have obtained some fundamental properties for the cellular network modeled by a non-linear load coupling equation (NLCE), from the perspective of minimizing the energy consumption of all the base stations. To obtain analytical results on the optimality of full load, and the computation and existence of the power allocation, we have investigated a dual to the NLCE, given by a non-linear power coupling equation (NPCE). Interestingly, although the NPCE cannot be stated in closed-form, we have obtained useful properties that is instrumental in proving the analytical results. Our results suggest that in practice, communication systems that are active should focus on using all available resources to satisfy the users' rate demand, so as to minimize the total energy consumption.

Appendix

Lemma 5: For Problem $P0$, the optimal solution is such that the load vector satisfies $\boldsymbol{x}^\star = \mathbf{1}$.

Proof: Note that the the load satisfies $x_k > 1$ for all cell k to satisfy non-trivial rate demands. Assume that at optimality, we have $\mathbf{0} < \boldsymbol{x}^\star \leq \mathbf{1}$ where there exists at least one cell $i \in \mathcal{N}$ with load $0 < x_i^\star < 1$ and power $p_i^\star$. With all other power $p_k^\star$ and load $x_k^\star$ fixed, $k \neq i$, we reduce the power $p_i^\star$ to $p' = p_i^\star - \epsilon, \epsilon > 0$. Using (2), the corresponding load $x_i^\star$ strictly increases to $x' = x_i^\star + \epsilon', \epsilon' > 0$. We choose $\epsilon > 0$ such that $x' \leq 1$. With this new power-load pair (p', x') for cell i, we claim that: (i) the objective function is reduced, and (ii) the corresponding rate vector $\boldsymbol{r}'$ is such that $\boldsymbol{r}' \geq \boldsymbol{r}^\star$, i.e., the NLCE constraint is satisfied since $\boldsymbol{r}^\star \geq \boldsymbol{d}_{\min}$. The two claims together imply that $\boldsymbol{x}^\star$ with $0 < x_i^\star < 1$ is not optimal, independent of the actual cell i. By contradiction, $x_i^\star = 1$ for all i, i.e, $\boldsymbol{x}^\star = \mathbf{1}$.

We now prove the first claim. Denote the energy used in cell i, as a function of its power p_i, as $e_i = x_i p_i = \sum_{j\in\mathcal{J}_i} \frac{r_{ij} p_i}{\log(1+c_{ij}p_i)}$ where $c_{ij} \triangleq g_{ij}/(\sum_{k\in\mathcal{N}\setminus\{i\}} p_k g_{kj} x_k + \sigma^2)$ does not depends on p_i nor x_i. Then

$$\frac{\partial e_i}{\partial p_i} = \sum_{j\in\mathcal{J}_i} r_{ij} \frac{(1+c_{ij}p_i)\log(1+c_{ij}p_i) - c_{ij}p_i}{\log^2(1+c_{ij}p_i)(1+c_{ij}p_i)}. \quad (18)$$

It can be verified by calculus that the numerator of each summand is strictly increasing for $p_i > 0$, i.e., $\frac{\partial e_i}{\partial p_i} > 0$. Hence, when the power for cell i is decreased, the energy e_i decreases. Thus, the objective function also decreases.

To prove the second claim, we first note that for cell i, we have constrained the new power-load pair (p', x') to satisfy (2). Thus, the new rate for cell i, denoted by $r'_{ij}, j \in \mathcal{J}_i$, is the same as the optimal rate $r_{ij}^\star$ corresponding to the power-load pair $(p_i^\star, x_i^\star)$. Next, we observe that the product $x'p'$ is strictly smaller as compared to $x_i^\star p_i^\star$, according to the first claim. Thus, for user $j \in \mathcal{J}_k$ in cell $k \neq i$, $\mathsf{SINR}_{kj}(\boldsymbol{x})$ strictly increases. It follows that the NLCE for cell k is satisfied with the same load x_k but with a larger rate r'_{kj} as compared to the optimal rate $r_{kj}^\star$. In summary, we thus have $\boldsymbol{r}' \geq \boldsymbol{r}^\star$. ■

Lemma 6 (Theorem 2, [23]): Consider the NLCE (3) with power $\boldsymbol{p}$ fixed. Given the rate vectors $\boldsymbol{r}'$ and $\boldsymbol{r}$ with $\boldsymbol{r}' \geq \boldsymbol{r}$ and $\boldsymbol{r}' \neq \boldsymbol{r}$, the corresponding load vectors $\boldsymbol{x}'$ and $\boldsymbol{x}$ satisfy $\boldsymbol{x}' > \boldsymbol{x}$.

We omit the proof of Lemma 6, which is given in [23].

Lemma 7: For Problem $P0$, the optimal rate vector satisfies $\boldsymbol{r}^\star = \boldsymbol{d}_{\min}$.

Proof: Suppose that at optimality, there exists at least one rate element $r_i^\star$ that is strictly greater than its corresponding (minimum) rate demand $d_{i,\min}$. Taking the power to be fixed as $\boldsymbol{p}^\star$, if we decrease $r_i^\star$ to $d_{i,\min}$, then the load will strictly decrease while satisfying the constraint (4b) by Lemma 6. Thus, the objective function value decreases. This contradicts the optimality of $r_i^\star$. Thus $\boldsymbol{r}^\star = \boldsymbol{d}_{\min}$. ■

References

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2011-2016," Feb. 2012, White Paper. [Online]. Available: http://tinyurl.com/b9berc

[2] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. V. der Perre, "Challenges and enabling technologies for energy aware mobile radio networks," *IEEE Commun. Mag.*, vol. 48, pp. 66–72, Nov. 2010.

[3] G. Fettweis and E. Zimmermann, "ICT energy consumption-trends and challenges," in *Proc. of the 11th Int. Symp. on Wireless Personal Multimedia Commun.*, vol. 2, no. 4, Sep. 2008, p. 6.

[4] W. Vereecken, W. V. Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester, "Power consumption in telecommunication networks: overview and reduction strategies," *IEEE Commun. Mag.*, vol. 49, pp. 62–69, Jun. 2011.

[5] H. Bogucka and A. Conti, "Degree of freedom for energy savings in practical adaptive wireless systems," *IEEE Commun. Mag.*, vol. 49, pp. 38–45, Jun. 2011.

[6] M. Gruber, O. Blume, D. Ferling, D. Zeller, and M. A. Imran, "EARTH–energy aware radio and network technologies," in *Proc. Personal, Indoor and Mobile Radio Commun. Symp.*, Tokyo, Japan, Sep. 2009.

[7] J. Joung, C. K. Ho, and S. Sun, "Power amplifier switching (PAS) for energy efficient systems," *IEEE Wireless Commun. Lett.*, vol. 2, no. 1, pp. 14–17, Feb. 2013.

[8] ——, "Spectral efficiency and energy efficiency of OFDM systems: impact of power amplifiers and countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 2, pp. 208–220, Feb. 2014.

[9] K. Adachi, J. Joung, S. Sun, and P. H. Tan, "Adaptive coordinated napping (CoNap) for wireless networks energy saving," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5656–5667, Nov. 2013.

[10] C. K. Ho, P. H. Tan, and S. Sun, "Energy-efficient relaying over multiple slots with causal CSI," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1494–1505, Aug. 2013.

[11] J. Joung, C. K. Ho, P. H. Tan, and S. Sun, "Energy minimization in OFDMA downlink systems: A sequential linear assignment algorithm for resource allocation," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 300–303, Aug. 2012.

[12] J. Joung, D. Yuan, C. K. Ho, and S. Sun, "Energy efficient network-flow-based algorithm with flow size adaptation for multiuser multicarrier systems," *IET Networks*, vol. 1, no. 2, pp. 66–73, Aug. 2012.

[13] R. Kwan, "Distributed energy-saving mechanism for self-organizing Femto LTE networks," in *Proc. IEEE VTC Fall Conf.*, Sep. 2012.

[14] L. Lei, D. Yuan, C. K. Ho, and S. Sun, "A unified graph labeling algorithm for consecutive-block channel allocation in SC-FDMA," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5767–5779, Nov. 2013.

[15] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 28–35, 2011.

[16] Z. Hasan, H. Boostanimehr, and V. Bhargava, "Green cellular networks: A survey, some research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 4, pp. 524–540, Fourth Quar. 2011.

[17] I. Siomina, A. Furuskar, and G. Fodor, “A mathematical framework for statistical QoS and capacity studies in OFDM networks,” in *Proc. IEEE Personal, Indoor and Mobile Radio Commun. Symp.*, Sep. 2009, pp. 2772 –2776.

[18] K. Majewski and M. Koonert, “Conservative cell load approximation for radio networks with Shannon channels and its application to LTE network planning,” in *Proc. Advanced Int. Conf. on Telecomm. (AICT)*, May 2010, pp. 219–225.

[19] I. Siomina and D. Yuan, “Analysis of cell load coupling for LTE network planning and optimization,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2287–2297, Jun. 2012.

[20] A. J. Fehske and G. P. Fettweis, “Aggregation of variables in load models for interference-coupled cellular data networks,” in *Proc. IEEE Int. Conf. on Commun.*, Jun. 2012, pp. 5102–5107.

[21] ——, “On flow level modeling of multi-cell wireless networks,” in *Proc. Int. Symp. on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Japan, 2013, pp. 572–579.

[22] C. K. Ho, D. Yuan, and S. Sun, “Data offloading in load coupled networks: Solution characterization and convexity,” in *Proc. IEEE Int. Conf. on Commun.*, Jun. 2013, pp. 5102–5107.

[23] ——, “Data offloading in load coupled networks: A utility maximization framework,” *IEEE Trans. Wireless Commun.*, accepted, Jan. 2014. [Online]. Available: http://arxiv.org/abs/1312.6918

[24] R. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[25] “Momentum project IST-2000-28088,” 2003 (updated in 2005). [Online]. Available: http://momentum.zib.de