A Learning-Based Approach to Energy Efficiency Maximization in Wireless Networks

Salvatore D'Oro*, Alessio Zappone†, Sergio Palazzo*, Marco Lops†
*: University of Catania, Catania, Italy,
†: University of Cassino and Southern Lazio, Cassino, Italy

Abstract—**This work develops a learning-based framework for energy-efficient power control in multi-carrier wireless networks. The problem is formulated as the maximization of the network global energy efficiency, defined as the ratio between the network sum-rate and the total consumed power, and is tackled by a novel approach which merges tools from learning, non-cooperative game theory, and fractional programming theory. The proposed algorithm is provably convergent, enjoys near-optimal performance, while requiring a much lower complexity than previous alternatives.**

I. Introduction

The exponential increase of the number of wireless devices, which is forecast to reach 50 billions by 2020 [1], poses serious sustainable growth concerns because: 1) an unprecedented amount of energy will be required to serve this massive amount of devices; 2) the computational complexity and overheads to manage such large networks will be enormous. It is common opinion, that future wireless networks will have to increase the bit-per-Joule energy efficiency, defined as the system data-rate over the consumed power, by a factor 2000 [2]–[4].

Traditionally, energy efficiency maximization in the context of multi-carrier networks has been performed assuming exclusive subcarrier assignment. This approach has the advantage of nulling out multi-user interference, leading to a problem that can be globally solved by fractional programming theory [5]. However, exclusive subcarrier assignment does not appear practical for future networks, as it is difficult to implement in multi-cell and heterogeneous scenarios, and because of the exponentially increasing spectrum demands [6]. Nevertheless, allowing subcarrier reuse inevitably leads to multi-user interference, which makes the problem of energy efficiency maximization more complex, and in general NP-hard. This calls for low-complexity solutions for practical applications.

One widely used approach to reduce the complexity of resource allocation problems is non-cooperative game theory [7], [8]. In this context, the network nodes are modeled as utility-driven, rational agents, and the problem of maximizing the system-wide energy efficiency is decomposed into a set of simpler, user-dependent, resource allocation problems. A seminal work in this field is [9], where a multi-carrier interference network is considered, and the energy efficiency of each communication link is defined as the ratio between the link achievable rate and consumed power. Heterogeneous networks are considered in [10], where the tool of quasi-variational inequalities is used to analyze the non-cooperative power control problem. The non-cooperative energy-efficient power control for relay-assisted networks is investigated in [11], while [12] extends the results of [11] to the problem of joint energy-efficient power and receiver design. Similarly, the non-cooperative energy-efficient power control problem is considered in [13], [14] where, instead, a minimum rate level for each communication link is considered. Nevertheless, all previous works show that purely game-theoretic approaches in general perform quite far from global optimality.

More recently, a more performing approach tackles the problem of network energy efficiency maximization by merging fractional programming and sequential optimization theory [14], [15]. Again, the energy efficiency maximization problem is decomposed into a sequence of simpler resource allocation sub-problems, but unlike game-theoretic approaches, each sub-problem is not a local, user-dependent problem, but rather a network-wide problem still aiming at network energy efficiency maximization. This approach enjoys near-optimal performance, even though a higher computational complexity is required, which could become problematic in large networks with many users and subcarriers.

Motivated by this scenario, this works aims at developing a novel energy efficiency optimization framework, which exhibits a comparable or lower complexity than game-theoretic approaches, but near-optimal performance as the sequential method. This is hereby achieved complementing the tools of fractional programming and game theory with the tool of learning theory [16], [17]. In particular, it should be mentioned that radio resource allocation by learning theory has been mostly considered to develop online algorithms for long-term resource allocation in fast-fading scenarios [18]–[21]. However, the very recent study [22] demonstrates how learning theory can also be used to reduce the computational complexity of resource allocation in slow-fading environments.

This work extends the approach in [22] in at least two major directions:

1) Energy efficiency maximization is considered, while [22] only focuses on rate optimization.
2) Generic interference networks are considered, which includes multi-cell scenarios as a special case, whereas

S. D'Oro is now with the Northeastern University, Boston, MA, USA.

This paper has been accepted for publication on *IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain, April 15-18, 2018*. This is a preprint version of the accepted paper. The final paper will be available in short.

[22] only focuses on single-cell systems.

More in detail, assuming a generic interference network model, the problem is formulated as the maximization of the global energy efficiency (GEE), defined as the ratio between the system sum-rate and consumed power. Next, merging learning theory, game theory, and fractional programming theory, a low-complexity, but near-optimal algorithm is devised. The algorithm is also provably convergent, and its performance is numerically evaluated both against the near-optimal, but more complex method from [15], and against purely game-theoretic approaches.

II. System model and problem statement

Let us consider a synchronous, multi-carrier, interference network, with K transmitters, J receivers, and N subcarriers. Denoting by $p_{k,n}$ and $h_{k,j,n}$ the k-th user's transmit power and complex channel gain to receiver j, over subcarrier n, the k-th user's signal to interference plus noise ratio (SINR) at its intended receiver $a(k)$, over subcarrier n is:

$$\gamma_{k,n} = \frac{p_{k,n}|h_{k,a(k),n}|^2}{\sigma^2_{k,n} + \sum_{\ell \neq k} p_{\ell,n}|h_{\ell,a(k),n}|^2} \tag{1}$$

wherein $\sigma^2_{k,n}$ is the thermal noise power at receiver $a(k)$ over subcarrier n, modeled as a zero-mean, circularly symmetric Gaussian process. In this context, the system bit-per-Joule GEE is defined as the network benefit-cost ratio, in terms of amount of data reliably transmitted and consumed power, namely:

$$\text{GEE} = \frac{W \sum_{k=1}^{K} \sum_{n=1}^{N} \log_2(1+\gamma_{k,n})}{P_c + \sum_{k=1}^{K}\sum_{n=1}^{N} \mu_{k,n} p_{k,n}} , \tag{2}$$

wherein W is the subcarrier bandwidth, $\mu_{k,n}$ is the inefficiency of the k-th user's power amplifier over subcarrier n, while P_c is the total static hardware power dissipated in all system hardware blocks other than the transmit amplifiers (e.g. ADC/DAC, cooling equipments, analog filters).

In this context, the problem of power allocation for GEE maximization is mathematically stated as

$$\max_{\boldsymbol{p}_k} \frac{\sum_{k=1}^{K} \sum_{n=1}^{N} \log_2(1+\gamma_{k,n})}{P_c + \sum_{k=1}^{K}\sum_{n=1}^{N} \mu_{k,n} p_{k,n}} , \tag{3a}$$

$$\text{s.t. } 0 \leq \sum_{n=1}^{N} p_{k,n} \leq P_{max,k} \; , \; \forall \, k = 1, \ldots, K \; . \tag{3b}$$

Being a single-ratio fractional problem, the most widely used approach to tackle Problem (3) is based on the use of fractional programming [5], [23]. In particular, available approaches resort to the popular Dinkelbach's algorithm, that is stated in Algorithm 1 for the case of (3), and that is based on the following result from fractional programming theory.

Proposition 1. *Consider Problem* (3)*, with $\mathcal{S}$ its feasible set. Define the auxiliary function $F : \lambda \in \mathbb{R} \to F(\lambda)$ as*

$$F(\lambda) = \max_{\boldsymbol{p}} \sum_{k=1}^{K}\sum_{n=1}^{N} \log_2(1+\gamma_{k,n}) - \lambda\left(P_c + \sum_{k=1}^{K}\sum_{n=1}^{N}\mu_{k,n}p_{k,n}\right). \tag{4}$$

Then, $\boldsymbol{p}^\star$ is a global solution of (3) *if and only if $F(\lambda^\star) = 0$, with $\lambda^\star$ being the unique zero of $F(\lambda)$.*

Algorithm 1 Dinkelbach's algorithm

```
Set j = 0; λ_j = 0; ε > 0;
while F(λ_j) ≥ ε do
    p* = arg max_{p∈S} Σ_{k=1}^{K} Σ_{n=1}^{N} log2(1 + γ_{k,n}) − λ_j (P_c + Σ_{k=1}^{K} Σ_{n=1}^{N} μ_{k,n} p_{k,n});
    F(λ_j) = Σ_{k=1}^{K} Σ_{n=1}^{N} log2(1 + γ_{k,n}(p*_k)) − λ_j (P_c + Σ_{k=1}^{K} Σ_{n=1}^{N} μ_{k,n} p*_{k,n});
    λ_{j+1} = [Σ_{k=1}^{K} Σ_{n=1}^{N} log2(1 + γ_{k,n}(p*))] / [P_c + Σ_{k=1}^{K} Σ_{n=1}^{N} μ_{k,n} p*_{k,n}];   j = j + 1;
end while
```

While Algorithm 1 is guaranteed to globally solve (3), it requires to *globally* solve (4) in each iteration in order to compute the true value of $F(\lambda)$ [23], [24]. If this requirement is not fulfilled, neither the convergence nor the optimality of Algorithm 1 can be guaranteed based on available results in the open literature. Unfortunately, for the case at hand, globally solving (4) appears computationally challenging, since (4) is not a convex Problem, due to the fact that its objective function is not concave. In [15] it is shown that implementing Algorithm 1 in interference-limited scenarios like the one considered here has exponential complexity in general. A more practical, yet not provably optimal, method is also proposed in [15], by merging sequential optimization and fractional programming. This approach has been numerically shown to achieve global optimality in several relevant problem instances, but it still requires to numerically solve a sequence of convex problems, each one having KN optimization variables. In large networks with many users and/or subcarriers, the resulting complexity might still be too high. Thus, the goal of this work is to develop a novel approach to tackle (3), which enjoys near-optimal performance, like the method from [15], but a much lower computational complexity. This will be achieved by merging fractional programming, game theory, and learning theory.

III. Proposed approach

Consider Algorithm 1, and define the objective of (4) as the function:

$$V(\boldsymbol{p}) = \sum_{i=1}^{K}\sum_{n=1}^{N} \log_2(1+\gamma_{i,n}) - \lambda_j \left(P_c + \sum_{i=1}^{K}\sum_{n=1}^{N} \mu_{i,n}p_{i,n}\right). \tag{5}$$

As already mentioned, finding the global maximum of (5) is in general computationally prohibitive. For this reason, the approach will be to provide a computationally-efficient method to derive (possibly suboptimal) solutions of (4). To this end, let us introduce the non-cooperative game in normal form

$$\mathcal{G} = \{\mathcal{K}, \{\mathcal{S}_k\}_{k=1}^{K}, \{u_k\}_{k=1}^{K}\} , \tag{6}$$

wherein $\mathcal{K}$ is the players' set,

$$\mathcal{S}_k = \left\{ \boldsymbol{p}_k = [p_{k,1}, \ldots, p_{k,N}]^T : p_{k,n} \geq 0, \; \forall n = 1, \ldots, N \right.$$
$$\left. \sum_{n=1}^{N} p_{k,n} \leq P_{max,k} \right\} , \tag{7}$$

is the k-th player's strategy set, and u_k is the k-th player's utility function defined as

$$u_k(\boldsymbol{p}_k, \boldsymbol{p}_{-k}) = V(\boldsymbol{p}_k, \boldsymbol{p}_{-k}) = V(\boldsymbol{p}) \tag{8}$$

with $\boldsymbol{p}_{-k} = \{\boldsymbol{p}_j\}_{j \in \mathcal{K}, j \neq k}$. In particular, it can be seen that every player has the same utility function $V(\boldsymbol{p})$, which makes $\mathcal{G}$ a potential game [25], and specifically a so-called *identical interest* game [26]. Potential games enjoy several pleasant properties, among which the existence of at least one pure-strategy Nash equilibrium (NE), under the mild assumptions that the potential function is continuos and the strategy sets are compact [25]. Moreover, pure-strategy NE can be reached by implementing the game best-response dynamics, i.e. letting each player k iteratively maximize the common utility function $V(\boldsymbol{p})$ with respect to their own strategy $\boldsymbol{p}_k$, assuming all other players' power vectors $\{\boldsymbol{p}_\ell\}_{\ell \neq k}$ are fixed. Unfortunately, implementing the best-response dynamics of $\mathcal{G}$ is no easy task, since the common utility $V(\boldsymbol{p})$ is not concave even with respect to only a single power vector $\boldsymbol{p}_k$. To circumvent this problem, in the following we replace the concept of best-response, with the milder notion of better response, formally defined next.

Definition 1 (Better-Response Strategy)**.** *A strategy $\boldsymbol{p}_k^*$ is a better-response strategy for user $k \in \mathcal{K}$ to $(\boldsymbol{p}_k, \boldsymbol{p}_{-k})$, if $u_k(\boldsymbol{p}_k^*, \boldsymbol{p}_{-k}) \geq u_k(\boldsymbol{p}_k, \boldsymbol{p}_{-k})$. That is, $\boldsymbol{p}_k^*$ dominates $\boldsymbol{p}_k$ when other players choose $\boldsymbol{p}_{-k}$.*

Otherwise stated, each player k does not aim at computing the strategy which maximizes its utility function u_k given the strategies of the other players. Instead, the goal is to find a strategy $\boldsymbol{p}_k^*$ which improves the value of u_k as compared to the present strategy $\boldsymbol{p}_k$, and given the strategies of the other players $\boldsymbol{p}_{-k}$. Following Definition 1, we can introduce the notion of better-response dynamics:

Definition 2 (Better-Response Dynamics (BRD))**.** *A Better-Response Dynamics is a sequence of strategies $\{p_k^*\}_k$, where p_k^* is a better response of player k, for all k.*

Definition 2 immediately implies the following result.

Proposition 2. *Let $\{p_k^*\}_k$ be a better-response dynamics for the game $\mathcal{G}$. Then, the sequence $\{V(p_k^*)\}_k$ converges.*

Proof: According to the definition of better-response strategy, any unilateral strategy update performed by user $k \in \mathcal{K}$ improves the utility function, and thus the potential function $V(\boldsymbol{p})$. Moreover, $V(\boldsymbol{p})$ admits a finite maximizer over $\mathcal{S} = \mathcal{S}_1 \cap \mathcal{S}_2 \ldots \cap \mathcal{S}_K$, because $V(\boldsymbol{p})$ is continuous and $\mathcal{S}_k$ is compact for each k. Hence, $\{V(p_k^*)\}_k$ must converge. ■

We hasten to stress that Proposition 2 holds without requiring any concavity/convexity property for the potential $V(\boldsymbol{p})$ and for the strategy sets $\mathcal{S}_k$. Also, it is worth remarking that while Proposition 2 ensures convergence towards an efficient power allocation policy, convergence towards a NE cannot be guaranteed by Proposition 2. In fact, the definition of NE is built upon the concept of best response functions, whose computation in our case is made unfeasible by the non-convexity of the problem.

At this point, having ensured this convergence result, it remains to provide a low complexity method to obtain a better response for the generic player k. This is accomplished by employing the tool of learning with exponential mappings [22]. In particular, let us consider the power update scheme:

$$\begin{cases} y_{k,n}(t+1) = y_{k,n}(t) + \delta_t \hat{v}_{k,n}(p_{k,n}(t), \boldsymbol{p}_{-k}, \boldsymbol{p}_k^*) \\ p_{k,n}(t+1) = P_{max,k} \frac{e^{y_{k,n}(t)}}{1+\sum_{m=1}^{N} e^{y_{k,m}(t)}} \end{cases} \tag{9}$$

wherein $\boldsymbol{p}_k^*$ is the previous better response of user k, t is the iteration index of the learning procedure, δ_t is the step-size, and $\hat{v}_{k,n}(p_{k,n}(t), \boldsymbol{p}_{-k}, \boldsymbol{p}_k^*)$ is defined as

$$\hat{v}_{k,n}(p_{k,n}(t), \boldsymbol{p}_{-k}) = \frac{|h_{k,a(k),n}|^2/\log(2)}{p_{k,n}(t)|h_{k,a(k),n}|^2 + z_n(\boldsymbol{p}_{-k})} - \beta_{k,n} \ ,$$

with $z_n(\boldsymbol{p}_{-k}) = \sigma_n^2 + \sum_{\ell \neq k} p_{\ell,n} |h_{\ell,a(k),n}|^2$, $\beta_{k,n} = \frac{\lambda_j \mu_{k,n} - \phi_{k,n}}{K}$, and

$$\phi_{k,n} = -\sum_{i \neq k} \frac{|h_{i,a(k),n}|^2/\log(2)}{\sigma_n^2 + p_{k,n}^* |h_{i,a(k),n}|^2 + \sum_{\ell \neq i, \ell \neq k} p_{\ell,n} |h_{\ell,a(k),n}|^2}.$$

In (9), the *scores* $y_{k,n}$ are iteratively updated according to the marginal utility $\hat{v}_{k,n}$, thus providing an effective reinforcement learning mechanism. Instead, the exponential mapping is used to enforce the feasibility constraints in (3b).

The following result can be shown.

Proposition 3. *If $\sum_{t=1}^{+\infty} \delta_t^2 < \sum_{t=1}^{+\infty} \delta_t = +\infty$, the learning procedure* (9) *converges to a better response for player k.*

Proof: The proof is omitted due to space constraints. ■

Thus, any user k can determine a better response by implementing (9), and the overall energy-efficient power control algorithm can be formally stated as in Algorithm 2.

Algorithm 2 Dinkelbach's algorithm for GEE

Set $j = 0$; $\lambda_j = 0$; $\varepsilon > 0$;
while $\bar{F}(\lambda_j) \geq \varepsilon \ || \ |\lambda_j - \lambda_{j-1}| \geq \varepsilon$ **do**
  while Convergence has not been reached **do**
    for each $k = 1, \ldots, K$ **do**
      $\boldsymbol{p}_k^* \leftarrow$ The convergence point of (9);
      $\boldsymbol{p} = (\boldsymbol{p}_k^*, \boldsymbol{p}_{-k})$;
    end for
  end while

$$\bar{F}(\lambda_j) = \sum_{k=1}^{K} \sum_{n=1}^{N} \log_2(1 + \gamma_{k,n}(\boldsymbol{p})) - \lambda_j \left(\sum_{k=1}^{K} P_{c,k} + \sum_{n=1}^{N} \mu_{k,n} p_{k,n} \right) \tag{10}$$

$$\lambda_{j+1} = \frac{\sum_{k=1}^{K} \sum_{n=1}^{N} \log_2(1 + \gamma_{k,n}(\boldsymbol{p}))}{\sum_{k=1}^{K} P_{c,k} + \sum_{n=1}^{N} \mu_{k,n} p_{k,n}}; \quad j = j + 1;$$

end while

Let us note that Algorithm 2 is an implementation of Algorithm 1 where the equilibrium point of the better-response dynamics is used in place of the optimal solution of the auxiliary Problem (4). Thus, recalling the discussion in Section II, the convergence and optimality of Algorithm 2 do not follow from the known properties of Dinkelbach's algorithm. Indeed, in Algorithm 2 the notation $\bar{F}(\lambda_j)$ is used to stress the difference between $F(\lambda)$, i.e. the maximum value of the objective of (4), and (10) which is in general lower than $F(\lambda)$. Nevertheless, while Algorithm 2 is not guaranteed to be globally optimal, its convergence can be theoretically guaranteed.

Proposition 4. *Algorithm 2 converges in a finite number of iterations. Moreover, one of the following two cases occurs:*

1) *Algorithm 2 monotonically increases the value of* (3a) *after each iteration and converges. In addition, if upon convergence it holds $F(\lambda) = 0$, then global optimality is attained.*
2) *Let $\bar{j}$ be the index of the first iteration for which $\bar{F}(\lambda_{\bar{j}}) < 0$. In this case, Algorithm 2 stops at iteration $\bar{j}$, after having monotonically increased the value of* (3a) *for all $0 \leq j \leq \bar{j}$.*

Proof: The proof is omitted due to space constraints. ■

Remark 1. *Algorithm 2 can be implemented either in a centralized fashion, by a network controller with global CSI, or in a distributed way, letting each user compute his better response. In the latter scenario, each user k needs to know its own channels $\{h_{k,\alpha(k),n}\}_n$, plus the quantities $z_n(\boldsymbol{p}_{-k})$ and $\beta_{k,n}$ for all $n = 1, \ldots, N$. The first two quantities are typically locally available since each user knows its own channels $\{h_{k,\alpha k,n}\}_n$, and is aware of the suffered interference term z_n over the N subcarriers. However, the same is not true for the coefficients $\{\beta_{k,n}\}_n$, which need to be explicitly sent to user k, by his associated base station.*

A. Computational complexity

The computational complexity of Algorithm 2 depends on the complexity of the inner loop to compute the convergence point of (9), and on the number of outer iterations occurring before convergence is reached in the two outer loops.

In each iteration of the inner loop to compute the convergence point of (9), the variables $p_{k,n}$ and $y_{k,n}$ are simultaneously updated for each of the N subcarriers, which has complexity $\mathcal{O}(N)$, whereas no closed-form result is available for the number I_L of iterations before (9) converges. Nevertheless, the numerical analysis presented in Section IV shows that a few tens of iterations are required in practical scenarios.

As for the number of outer iterations required for the convergence of Algorithm 2, say I_D, again no closed-form result is available in general. However, as for I_D, as observed in Section II, the global maximum of the GEE corresponds to the zero of the auxiliary function $F(\lambda)$. By using a bisection approach, such a zero could be found within a tolerance ε by performing a number of bisection steps equal to $\log_2\left(\left\lceil \frac{U-L}{\varepsilon} \right\rceil\right)$, with U and L the initial upper and lower bounds of the maximum of the GEE [27]. Instead, the update rule for the parameter λ is typically faster than the bisection method. Indeed, Dinkelbach's algorithm is known to be equivalent to Newton's method applied to find the zero of the differentiable function $F(\lambda)$, and thus converges with superlinear rate [24], which is typically faster than the bisection method. This argument is corroborated by our numerical results, which shows how I_D is of the order of a few units.

Finally, the number of better responses to compute is equal to the number of users K, times the number of iterations until the better response dynamics converges, say I_B, which is of the order of a few units in game-theoretic power control algorithms [7], [14], [28]. This behavior is again confirmed by our numerical analysis.

Finally, we can obtain the overall complexity of Algorithm 2 as $\mathcal{O}(NKI_DI_BI_L)$. It is important to stress how the complexity is linear in the number of users K, and above all in the number of subcarriers N. On the other hand, other methods which do not exploit learning theory have a complexity which is either polynomial (typically cubic) in both N and K, or linear in K but polynomial in N. The former is the case with sequential optimization methods, which requires solving a sequence of pseudo-concave fractional sub-problems with respect to all of the NK transmit powers. The latter is the case when purely game-theoretic methods are employed, which requires solving a sequence of user-dependent sub-problems, in which each user k has to determine its optimal transmit powers over the N available subcarriers.

IV. Numerical Results

This section assesses the achievable performance of the proposed solution through numerical simulation. We assume $N = 4$ subcarriers are available, and users are uniformly distributed over a square area of edge $L = 200$ m. The noise power spectral density is $N_0 = -174\,\text{dBm/Hz}$ for all $n \in \mathcal{N}$, the subcarrier bandwidth is set to $W = 10.93$kHz, and $P_{c,k} = -20$ dB for all $k \in \mathcal{K}$. Unless stated otherwise, the step-size of the learning mechanism in (9) is $\delta_t = 1/t^{0.51}$ [1]. Finally, the inefficiency of the power amplifier is set to 2% for all users and over all subcarriers, i.e., $\mu = 1.02$. The presented results are averaged over 1000 simulation runs.

Let GEE^{OPT} be the optimal solution of the GEE maximization problem calculated through the centralized approach from [15]. Furthermore, let GEE^{P} be the GEE obtained through the proposed approach described in Section III. Similarly, let GEE^{C} denote the GEE achieved by the system under a purely competitive game-theoretic approach in which a best-response dynamics is run until convergence, with each user's best-response being the maximization of his own EE.

Figs. 1 and 2 compare the GEE of the proposed approach with the GEE achieved by both the optimal and competitive solutions as a function of P_{max} for $K = 6$ and $K = 12$. In Fig. 1 it is shown that the GEE achieved by both the proposed and the optimal solutions increases as P_{max} increases, eventually saturating for large P_{max}. The saturation effect is expected, since the GEE function is not increasing in the transmit powers, but rather admits finite optimal power levels. Thus, when P_{max} is large enough, the optimal powers levels are attained, and further increasing P_{max} does not change the power allocation policy anymore. On the contrary, the GEE of the competitive solution is a decreasing function of P_{max}. Furthermore, it is worth noting that the GEE of the system when a small number of users access the network (i.e., dashed lines) is higher than that achieved when the number K of users to be scheduled is larger (i.e., solid lines).

[1] It is worth noting that $\delta_t = 1/t^{\beta}$ satisfies Proposition 3 for any $\beta \in (0.5, 1]$. Also, it has been shown that fast convergence rates can be acheived when $\beta \to 0.5$ [22].

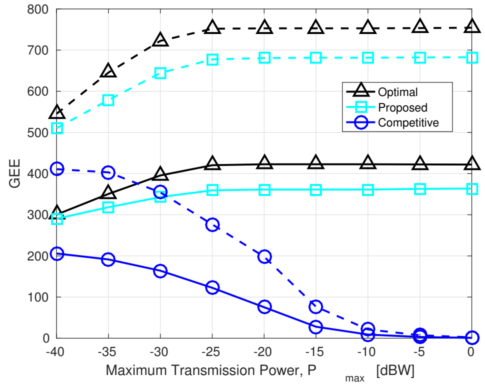

Fig. 1. GEE of different schemes a function of P_{max} for different number K of users ($K = 12$: solid lines; $K = 6$: dashed lines).

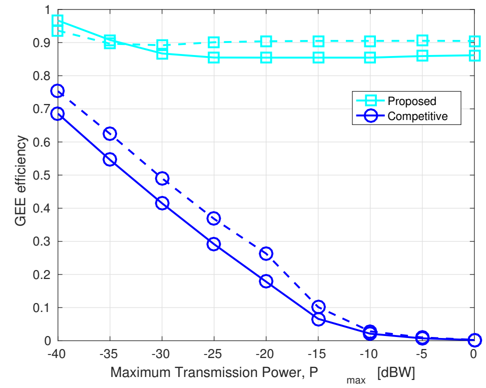

Fig. 2. Comparison of different schemes for the GEE maximization as a function of P_{max} for different number K of users ($K = 12$: solid lines; $K = 6$: dashed lines).

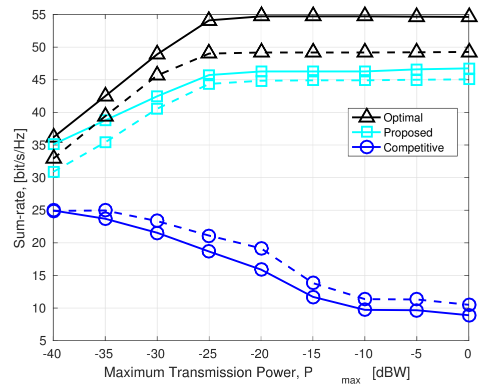

Fig. 3. Sum-rate of different schemes a function of P_{max} for different number K of users ($K = 12$: solid lines; $K = 6$: dashed lines).

In Fig. 2 we show the two ratios $\text{GEE}^{P}/\text{GEE}^{\text{OPT}}$ and $\text{GEE}^{C}/\text{GEE}^{\text{OPT}}$ which represent the efficiency of the proposed and competitive solutions. The higher the value of the ratio, the higher the efficiency. It is shown that the competitive approach achieves good performance when very small transmission power levels are considered. However, in general this approach is very inefficient, and its efficiency decreases as the maximum transmission power level increases. This is because it is designed to perform individual EE maximization, and if a large P_{max} is available, then each user can create a significant interference to other users, in order to selfishly increase its own EE. Instead, the approach proposed in Section III is almost optimal and $\text{GEE}^{C}/\text{GEE}^{\text{OPT}} \approx 0.9$ in all of the considered cases. However, it is worth noting that such a high efficiency comes at the expense of higher overhead with respect to the competitive solution. In fact, while the competitive approach can be implemented in a fully-distributed fashion and only requires local information, the identical interest game-based approach requires additional feedback as discussed in Remark 1.

The system sum-rate achieved by the GEE-maximizing power allocation is shown in Fig. 3 as a function of P_{max} for $K = 6$ and $K = 12$ users. It is shown that the sum-rate of the system when the proposed solution is considered increases as the value of P_{max} increases as well. However, let us note that the sum-rate of the system asymptotically converges towards a saturation point. Indeed, as shown in Fig. 1, the GEE eventually saturates when P_{max} is large enough to attain the peak of the GEE. At this point, the optimal power allocation remains constant with P_{max}, since further increasing the transmit power would only decrease the GEE value. On the contrary, the sum-rate of the system under the competitive approach is generally decreasing in P_{max}.

In Fig. 4 we investigate the show the average number of iterations needed to converge by each nested loop in Algorithm 2 as a function of P_{max} when $K = 12$. In the inner loop of Algorithm 2 we execute the learning mechanism (9), the middle loop computes the solution of the better-response dynamics, and the outer loop is used to iteratively update the λ parameter of the Dinkelbach's algorithm. It is shown that the number of iterations of Dinkelbach's algorithm is almost constant and lower than 5. The average number of better-response updates for each user at each iteration of the Dinkelbach's algorithm decreases as the value of P_{max} increases. However, such a number is low, i.e., the equilibrium is attained after few iterations of the BRD. On the contrary, the per-user average number of iterations of the learning mechanism increases as the maximum transmission power P_{max} increases. This is caused by the larger action state space, which increases the number of power updates of the learning mechanism. However, it is worth noting that the iterates of the learning mechanism have linear complexity $\mathcal{O}(N)$. Therefore, though the number of iterations to find a better-response increases, it does not considerably impact the overall convergence speed of the proposed approach.

To further support this point, in Table I we show the ratio between the asymptotic complexity of the approach from[2] [15] and the proposed approach, and between the game-theoretic competitive method and the proposed approach. The asymptotic complexity of the proposed method has been evaluated as derived in Section III-A. The asymptotic complexity of the method from [15], can be lower-bounded as $\mathcal{O}(I_{S_b} I_{D_b} K^3 N^3)$,

[2]This method merges Dinkelbach's algorithm and sequential optimization, solving a sequence of approximate fractional problems, refining the approximation after each iteration.

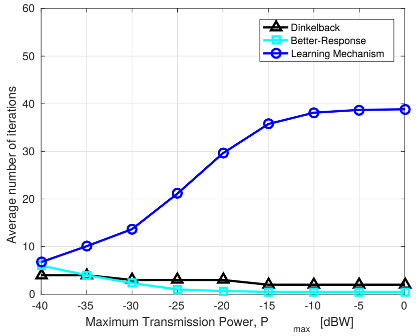

Fig. 4. Average number of iterations as a function of P_{max}.

TABLE I
ASYMPTOTIC COMPLEXITY RATIO OF A CENTRALIZED SOLUTION W.R.T. THE PROPOSED AND COMPETITIVE APPROACHES.

	-40dBW	-30dBW	-20dBW	-10dBW	0dBW
Proposed	$1.36*10^4$	$6.75*10^3$	$3.11*10^3$	$2.42*10^3$	$2.37*10^3$
Competitive	$1.20*10^2$	$4.11*10^2$	$1.44*10^3$	$2.88*10^3$	$2.88*10^3$

with I_{S_b} and I_{D_b} the number of approximate fractional problems to solve and the number of Dinkelbach's iterations for each approximate fractional problem. The choice of the cubic exponent is motivated by the fact that a cubic complexity is what is required to solve a KKT system in the case of linear problems, which is a simpler scenario as the one at hand here[3]. The complexity of the competitive method can be evaluated as $\mathcal{O}(I_{B_c} I_{D_c} K N^3)$, with I_{B_c} being the number of per-user best-responses until convergence, I_{D_c} the number of Dinkelbach's iterations to solve each best-response problem, while the term N^3 stems from the fact that in each Dinkelbach's iteration a convex problem with N variables must be solved. The results clearly show that the proposed method has a much lower complexity than the method from [15] and even of the game-theoretic competitive solution. This result is particularly relevant when taken together with the results in Fig. 2.

V. CONCLUSIONS

This work has developed a learning-based approach to power control for GEE maximization in multi-carrier interference networks. The proposed algorithm is provably convergent and has been numerically shown to enjoy near-optimal performance, despite requiring a much lower complexity than competing alternatives.

REFERENCES

[1] Ericsson White Paper, "More than 50 billion connected devices," Ericsson, Tech. Rep. 284 23-3149 Uen, Feb. 2011.

[2] "NGMN alliance 5G white paper," *https://www.ngmn.org/5g-white-paper/5g-white-paper.html*, 2015.

[3] S. Buzzi, C.-L. I, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5g networks and challenges ahead," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 5, 2016.

[3] We also observe that an upper-bound for the asymptotic complexity of a generic convex problem is $\mathcal{O}(I_{S_b} I_{D_b} K^4 N^4)$ [29]

[4] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5G: how to empower SON with big data for enabling 5G," *IEEE Network*, vol. 28, no. 6, pp. 27–33, November 2014.

[5] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Foundations and Trends in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185–396, 2015.

[6] "Reducing the net energy consumption in communications networks by up to 90% by 2020," GreenTouch Green Meter Research Study, Tech. Rep., June 2013.

[7] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks: Fundamentals and Applications*, ser. Academic Press. Elsevier, 2011.

[8] Z. Han, D. Niyato, W. Saad, T. Baar, and A. Hjrungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*, 1st ed. New York, NY, USA: Cambridge University Press, 2012.

[9] G. Miao, N. Himayat, G. Y. Li, and S. Talwar, "Distributed interference-aware energy-efficient power optimization," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1323–1333, April 2011.

[10] I. Stupia, L. Sanguinetti, G. Bacci, and L. Vandendorpe, "Power control in networks with heterogeneous users: A quasi-variational inequality approach," *IEEE Trans. Signal Process.*, vol. vol. 63, no. 21, pp. 5691–5705, November 2015.

[11] A. Zappone, Z. Chong, E. A. Jorswieck, and S. Buzzi, "Energy-aware competitive power control in relay-assisted interference wireless networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1860–1871, April 2013.

[12] A. Zappone, E. A. Jorswieck, and S. Buzzi, "Energy efficiency and interference neutralization in two-hop MIMO interference channels," *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6481–6495, December 2014.

[13] G. Bacci, E. Belmega, P. Mertikopoulos, and L. Sanguinetti, "Energy-aware competitive power allocation for heterogeneous networks under QoS constraints," *IEEE Trans. Wireless Communications*, vol. 14, no. 9, pp. 4728 – 4742, September 2015.

[14] A. Zappone, L. Sanguinetti, G. Bacci, E. A. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1668–1683, 2016.

[15] A. Zappone, E. Björnson, L. Sanguinetti, and E. Jorswieck, "Globally optimal energy-efficient power control and receiver design in wireless networks," *IEEE Transactions on Signal Processing*, vol. 65, no. 11, pp. 2844–2859, June 2017.

[16] R. Dilts and T. A. Epstein, *Dynamic Learning*. Capitola, CA, USA, 1995.

[17] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 1998.

[18] P. Mertikopoulos and E. V. Belmega, "Transmit without regrets: Online optimization in MIMO-OFDM cognitive radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 11, pp. 1987–1999, November 2014.

[19] F. Shams, G. Bacci, and M. Luise, "Energy-efficient power control for multiple-relay cooperative networks using Q-learning," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1567–1580, March 2015.

[20] P. Mertikopoulos and E. V. Belmega, "Learning to be green: Robust energy efficiency maximization in dynamic MIMO–OFDM systems," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 743–757, April 2016.

[21] Z. Gao, B. Wen, L. Huang, C. Chen, and Z. Su, "Q-learning-based power control for LTE enterprise femtocell networks," *IEEE Systems Journal*, vol. PP, no. 99, 2016.

[22] S. D'Oro, P. Mertikopoulos, A. L. Moustakas, and S. Palazzo, "Interference-based pricing for opportunistic multi-carrier cognitive radio systems," *IEEE Trans. Wireless Communications*, vol. 14, no. 12, pp. 6536–6549, July 2015.

[23] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, Mar. 1967.

[24] J.-P. Crouzeix and J. A. Ferland, "Algorithms for generalized fractional programming," *Mathematical Programming*, vol. 52, no. 1-3, pp. 191–207, 1991.

[25] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[26] ——, "Fictitious play property for games with identical interests," *Journal of economic theory*, vol. 68, no. 1, pp. 258–265, 1996.

[27] A. Zappone, E. A. Jorswieck, and A. Leshem, “Distributed resource allocation for energy efficiency in MIMO OFDMA wireless networks,” *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3451–3465, December 2016.

[28] S. Buzzi and H. V. Poor, “Joint receiver and transmitter optimization for energy-efficient CDMA communications,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 3, pp. 459–472, April 2008.

[29] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.