Optimal SINR-based Random Access

Amir-Hamed Mohsenian-Rad, Vincent W.S. Wong, and Robert Schober
Department of Electrical and Computer Engineering
University of British Columbia, Vancouver, Canada
e-mails: {hamed, vincentw, rschober}@ece.ubc.ca

Abstract— Random access protocols, such as Aloha, are commonly modeled in wireless ad-hoc networks by using the *protocol model*. However, it is well-known that the protocol model is *not* accurate and particularly it cannot account for *aggregate interference* from multiple interference sources. In this paper, we use the more accurate *physical model*, which is based on the *signal-to-interference-plus-noise-ratio* (SINR), to study *optimization-based* design in wireless random access systems, where the optimization variables are the transmission probabilities of the users. We focus on *throughput maximization*, *fair resource allocation*, and *network utility maximization*, and show that they entail non-convex optimization problems if the physical model is adopted. We propose two schemes to solve these problems. The first design is centralized and leads to the global optimal solution using a *sum-of-squares* technique. However, due to its complexity, this approach is only applicable to small-scale networks. The second design is distributed and leads to a close-to-optimal solution using the *coordinate ascent* method. This approach is applicable to medium-size and large-scale networks. Based on various simulations, we show that it is highly preferable to use the physical model for optimization-based random access design. In this regard, even a *sub-optimal* design based on the physical model can achieve a significantly better performance than an *optimal* design based on the inaccurate protocol model.

I. Introduction

In random access protocols, such as Aloha, each transmitter node *randomly* attempts to access a shared wireless channel with a certain *transmission probability*. Random access protocols are scalable and inherently flexible. They are widely used in practice in wireless local area and ad-hoc networking.

The wireless interference in random access systems is commonly modeled using the *protocol model* [1]. In the protocol model, the interference is *pair-wise*. That is, the model determines whether a pair of simultaneous transmissions from two transmitter nodes can be received successfully by the respective receivers, *assuming* that no other nodes are also transmitting. As a result, the protocol model *cannot* incorporate the impact of *aggregate interference* from *multiple* interfering sources. Examples where the protocol model is used for random access design include [2]–[9]. In [2], an optimization problem is formulated to select transmission probabilities in order to achieve proportional fairness among the nodes in terms of accessing the shared channel. The framework in [2] is extended in [3]–[5] to incorporate various other optimization objectives such as network throughput maximization, harmonic-mean fairness, and max-min fairness. Joint optimal random access and congestion control design is studied in [6], [7]. Most optimization problems in [2]–[7] are either convex or can be transformed into convex problems. Another thread of research focuses on using the protocol model to study non-cooperative random access using game theory [8], [9].

As an alternative, random access systems can also be modeled using the *physical model* [1]. In the physical model, the impact of the wireless interference is evaluated based on the *signal-to-interference-plus-noise-ratio* (SINR) at each receiver node. In this regard, a transmission is deemed successful if and only if the SINR at the receiver node is above a certain *threshold*, which depends on the modulation and coding schemes being used. Similarly, a *collision* occurs if the SINR drops below the threshold. The latter case can be caused by either a single interferer or *multiple* interferers. Therefore, the physical model incorporates the effect of *aggregate interference*. This can improve the accuracy of the model and better predict how random access protocols (e.g., the IEEE 802.11 distributed coordination function [10]) perform in practice (cf. [11], [12]).

In this paper, we formulate various optimization-based random access design problems using the physical model. Unlike the case when the protocol model is used to formulate the same problems, the optimization problems here are non-convex and cannot be transformed into convex problems. The main contributions of this paper are summarized as follows.

- We adopt the physical model to formulate *throughput maximization*, *fair resource allocation*, and *network utility maximization* problems in random access systems in order to find the optimal transmission probabilities.
- We show that the performance loss caused by the inaccuracy of the protocol model in solving these problems can be arbitrarily large in certain scenarios.
- We propose two design schemes to solve the formulated optimization problems. The first design is centralized. It uses the *sum-of-squares* technique to find the exact global optimal solution. Due to its complexity, this scheme is only applicable to small networks. It serves as a benchmark to assess the performance of our second design.
- Our second design is distributed. It uses the *coordinate ascent* method to find a close-to-optimal solution. It is applicable to medium-size and large-scale networks.
- We compare the network performance achievable with the physical and protocol models via extensive simulations.

This paper differs from the existing related work in the literature in several aspects. Unlike the studies in [11], [12], which focus on using the physical model to *analyze* the performance of existing random access protocols, here we adopt the physical model for developing *new* designs with optimal performance. The optimization problems that we consider in this paper have also been addressed in [2]–[5].

However, here we use the physical model, while the previous work in [2]–[5] has used the protocol model. It is also worth mentioning that the SINR is already widely considered for analyzing and designing random access systems with *multiple packet reception*, based on an information-theoretic approach, e.g., in [13], [14]. However, our work here addresses the more practical *single packet reception* setting by using an optimization-based approach to find the optimal transmission probabilities. Last but not least, some of the optimization techniques that we use in this paper, such as the sum-of-squares method, have also been used to tackle non-convexity in other communications and networking problems, e.g., for congestion control in [15], [16].

The rest of this paper is organized as follows. The system model is introduced in Section II. Our first design and the proof for its optimality are presented in Section III. Our second design is discussed in Section IV. Simulation results are provided in Section V. The paper is concluded in Section VI.

II. RANDOM ACCESS: PHYSICAL VS. PROTOCOL MODEL

Consider a wireless ad-hoc network with the set of one-hop transmitter/receiver pairs denoted by $\mathcal{N}$, with size $N = |\mathcal{N}|$. We refer to each transmitter/receiver pair as a *user*. For any user $n \in \mathcal{N}$, the transmitter node is denoted by t_n and the receiver node is denoted by d_n. Let q_n denote the transmission power of node t_n. Throughout this paper, we assume that transmission powers are *fixed* for all users. Let G_{nm} denote the *channel gain* from the transmitter node of user m (i.e., node t_m) to the receiver node of user n (i.e., node d_n). We notice that G_{nn} denotes the channel gain from node t_n to node d_n. We also notice that $q_n G_{nn}$ denotes the *received signal power* at receiver node d_n. Furthermore, if node t_n transmits a packet to node d_n and at the *same* time, node t_m transmits a packet to node d_m (with $m \neq n$), then the *interference power* caused by user m to user n is given by $q_m G_{nm}$.

We assume that time is divided into equal-length time slots. As an example, the length of a time slot can be equal to the time it takes to transmit one packet. Let $p_n \in [0, 1]$ denote the *transmission probability* of user $n \in \mathcal{N}$. That is, in each time slot, node t_n transmits a packet to node d_n with probability p_n. A user $n \in \mathcal{N}$ is called *active* in a time slot, if transmitter node t_n transmits a packet to receiver node d_n in that particular time slot. In each time slot, if user n is active, then we denote γ_n as its SINR. For each user $n \in \mathcal{N}$, we have

$$\gamma_n = \frac{q_n G_{nn}}{i_n + \sigma_n^2}, \qquad (1)$$

where i_n denotes the interference power in the time slot of interest, and σ_n^2 denotes the noise power. Clearly, the value of i_n (and consequently the value of γ_n) depends on how many other users $m \in \mathcal{N}\setminus\{n\}$ are active in the same time slot. For a transmission from node t_n to be *successful*, the SINR γ_n should be greater than or equal to a *threshold* value, denoted by γ_n^{th}. Therefore, in any time slot, receiver node d_n, $n \in \mathcal{N}$, successfully receives a packet from node t_n with probability

$$p_n \, \text{Prob}\left\{\gamma_n \geq \gamma_n^{th}\right\} = p_n \, \text{Prob}\left\{i_n \leq \frac{q_n G_{nn}}{\gamma_n^{th}} - \sigma_n^2\right\}. \quad (2)$$

We notice that since the transmission power, channel gain, and noise power are fixed, the term $(q_n G_{nn})/\gamma_n^{th} - \sigma_n^2$ has a fixed value. Therefore, $\text{Prob}\left\{i_n \leq (q_n G_{nn})/\gamma_n^{th} - \sigma_n^2\right\}$ represents the probability that the *aggregate* interference power at receiver node d_n, caused by *all* active interfering transmissions in the neighborhood, does *not* exceed a fixed threshold.

A. Physical Model

Given the modulation and coding scheme and the targeted SINR, let μ_n denote the *peak* data rate for user $n \in \mathcal{N}$. As in [2]–[7], we assume that the users always have data to transmit. In that case, the *average* data rate for user n is given by

$$r_n(\boldsymbol{p}) = \mu_n \, p_n \, \text{Prob}\left\{i_n \leq \frac{q_n G_{nn}}{\gamma_n^{th}} - \sigma_n^2\right\}, \qquad (3)$$

where $\boldsymbol{p} = (p_m, \, \forall m \in \mathcal{N})$ denotes the vector of transmission probabilities of *all* users. Recall that the term i_n in (3) is due to the interference powers caused by the transmissions of the active users in set $\mathcal{N}\setminus\{n\}$ when user n is also active. Let $\mathbb{P}_n$ denote the *power set* of set $\mathcal{N}\setminus\{n\}$, i.e., the set of all its *subsets*. For example, if $\mathcal{N} = \{n, m, a\}$, then $\mathbb{P}_n = \{\{\}, \{m\}, \{a\}, \{m, a\}\}$. Assuming that a subset $\mathcal{M} \in \mathbb{P}_n$ of *other* users in the network is active at the same time that user n is active, the transmission of user n is *successful* if and only if

$$\sum_{m \in \mathcal{M}} q_m G_{nm} \leq \frac{q_n G_{nn}}{\gamma_n^{th}} - \sigma_n^2. \qquad (4)$$

On the other hand, the probability of having subset $\mathcal{M} \subseteq \mathbb{P}_n$ of other users active at a time slot is given by

$$\left(\prod_{m \in \mathcal{M}} p_m\right)\left(\prod_{k \in \mathcal{N}\setminus\mathcal{M}, k \neq n} (1 - p_k)\right). \qquad (5)$$

The term $\prod_{m \in \mathcal{M}} p_m$ is the probability that all users in set $\mathcal{M}$ are active. The term $\prod_{k \in \mathcal{N}\setminus\mathcal{M}, k \neq n}(1 - p_k)$ is the probability that all users, other than user n and those users in set $\mathcal{M}$, are *not* active. Next, we define set $\mathbb{M}_n$ for each user $n \in \mathcal{N}$ as

$$\mathbb{M}_n = \left\{\mathcal{M} \in \mathbb{P}_n : \sum_{m \in \mathcal{M}} q_m G_{nm} \leq \frac{q_n G_{nn}}{\gamma_n^{th}} - \sigma_n^2\right\}. \quad (6)$$

That is, set $\mathbb{M}_n$ includes all subsets $\mathcal{M} \in \mathbb{P}_n$ of other users for which the inequality in (4) holds. Therefore, we have

$$\text{Prob}\left\{i_n \leq \frac{q_n G_{nn}}{\gamma_n^{th}} - \sigma_n^2\right\}$$
$$= \sum_{\mathcal{M} \in \mathbb{M}_n}\left(\prod_{m \in \mathcal{M}} p_m\right)\left(\prod_{k \in \mathcal{N}\setminus\mathcal{M}, k \neq n} (1 - p_k)\right), \qquad (7)$$

Replacing (7) in (3), for each user $n \in \mathcal{N}$, we have

$$r_n(\boldsymbol{p}) = \mu_n \, p_n$$
$$\times \left[\sum_{\mathcal{M} \in \mathbb{M}_n}\left(\prod_{m \in \mathcal{M}} p_m\right)\left(\prod_{k \in \mathcal{N}\setminus\mathcal{M}, k \neq n} (1 - p_k)\right)\right]. \qquad (8)$$

We will further explain the SINR-based data rate model in (8) in an example in Section II-C. In Sections III and IV, we will use (8) for the proposed optimization-based designs.

B. Protocol Model

The physical model is usually *not* used in random access design. Instead, the protocol model is used [2]–[7]. The protocol model takes into account only *pair-wise* interference. For each user $n \in \mathcal{N}$, let $\mathcal{I}_n$ denote the set of interfering users:

$$\mathcal{I}_n = \left\{ m \in \mathcal{N} \backslash \{n\} : q_m G_{nm} > \frac{q_n G_{nn}}{\gamma_n^{th}} - \sigma_n^2 \right\}. \quad (9)$$

From (9), if any user $m \in \mathcal{I}_n$ is active while user n is active, then the transmission of user n will *not* be successful. This is because the individual interference power from user m is already high enough to bring the SINR at the receiver node of user n below threshold γ_n^{th}. According to the protocol model, the *approximate* average data rate for user n is obtained as

$$\tilde{r}_n(\boldsymbol{p}) = \mu_n \, p_n \left(\prod_{m \in \mathcal{I}_n} (1 - p_m) \right). \quad (10)$$

Comparing the approximate data rate in (10) with the accurate data rate in (8), we can show the following.

Theorem 1: We always have

$$\tilde{r}_n(\boldsymbol{p}) \geq r_n(\boldsymbol{p}), \quad \forall n \in \mathcal{N}. \quad (11)$$

That is, the protocol model always leads to an *over-estimation* of the average data rates in wireless random access systems.

The proof of Theorem 1 is given in Appendix A.

C. Example

In this section, we present an example to compare the physical model and the protocol model in random access systems. Consider a network with $N = 4$ users. We have

$$\boldsymbol{q} = \begin{bmatrix} 5 \\ 4 \\ 4 \\ 3 \end{bmatrix}, \; \boldsymbol{\sigma}^2 = \begin{bmatrix} 0.5 \\ 0.3 \\ 0.6 \\ 0.3 \end{bmatrix}, \; \mathbf{G} = \begin{bmatrix} 1 & 0.8 & 0.4 & 0.5 \\ 0.8 & 1 & 0.5 & 0.3 \\ 0.4 & 0.5 & 1 & 0.2 \\ 0.5 & 0.3 & 0.7 & 1 \end{bmatrix},$$

where $\boldsymbol{q} = (q_n, \, \forall \, n \in \mathcal{N})$ denotes the transmission power vector, $\mathbf{G} = (G_{nm}, \, \forall \, n, m \in \mathcal{N})$ denotes the channel gain matrix, and $\boldsymbol{\sigma}^2 = (\sigma_n^2, \, \forall \, n \in \mathcal{N})$ denotes the noise power vector. We assume that $\gamma_1^{th} = \ldots = \gamma_4^{th} = 1$. Considering user $n = 1$, set $\mathcal{N} \backslash \{1\} = \{2, 3, 4\}$ and the power set $\mathbb{P}_1 = \{\{\}, \{2\}, \{3\}, \{4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}, \{2, 3, 4\}\}$. We can verify that $\mathbb{M}_1 = \{\{\}, \{2\}, \{3\}, \{4\}, \{3, 4\}\}$. Thus,

$$\begin{aligned}
\text{Prob} &\left\{ i_1 \leq \frac{q_1 G_{11}}{\gamma_1^{th}} - \sigma_1^2 \right\} \\
&= (1 - p_2)(1 - p_3)(1 - p_4) \\
&\quad + p_2(1 - p_3)(1 - p_4) + p_3(1 - p_2)(1 - p_4) \\
&\quad + p_4(1 - p_2)(1 - p_3) + p_3 p_4 (1 - p_2) \\
&= 1 - p_2(p_3 + p_4 - p_3 p_4),
\end{aligned} \quad (12)$$

and we have

$$r_1(\boldsymbol{p}) = \mu_1 p_1 - \mu_1 p_1 p_2 (p_3 + p_4 - p_3 p_4). \quad (13)$$

Similarly, and after reordering the terms, we can show that

$$\begin{aligned}
r_2(\boldsymbol{p}) &= \mu_2 p_2 (1 - p_1), & (14) \\
r_3(\boldsymbol{p}) &= \mu_3 p_3 (1 - p_1 p_2), & (15) \\
r_4(\boldsymbol{p}) &= \mu_4 p_4 (1 - p_3)(1 - p_1 p_2). & (16)
\end{aligned}$$

On the other hand, if we use the protocol model, we have

$$\begin{aligned}
\tilde{r}_1(\boldsymbol{p}) &= \mu_1 p_1, & (17) \\
\tilde{r}_2(\boldsymbol{p}) &= \mu_2 p_2 (1 - p_1), & (18) \\
\tilde{r}_3(\boldsymbol{p}) &= \mu_3 p_3, & (19) \\
\tilde{r}_4(\boldsymbol{p}) &= \mu_4 p_4 (1 - p_3). & (20)
\end{aligned}$$

A comparison of (13)-(16) with (17)-(20), confirms Theorem 1. We notice that, except for the second user, estimations of data rates for all other users are *inaccurate*. Next, we address optimization-based random access using the physical model.

III. DESIGN I: CENTRALIZED AND OPTIMAL

It is usually desired to select the transmission probabilities $\boldsymbol{p}$ such that a certain objective is achieved across the network. In this section, we formulate two optimization problems, namely *network throughput maximization* and *fair resource allocation*, in random access systems when the data rates are modeled using the physical model as in Section II-A. We then propose a centralized scheme to find the exact global optimal solution of each optimization problem accordingly. The optimal design scheme in this section can be used as a benchmark to assess the optimality of various other design schemes including the close-to-optimal design presented in Section IV.

A. Network Throughput Maximization

Consider the following optimization problem

$$\begin{aligned}
\max_{\boldsymbol{p}} \quad & \sum_{n \in \mathcal{N}} r_n(\boldsymbol{p}) \\
\text{s.t.} \quad & P_n^{\min} \leq p_n \leq P_n^{\max}, \qquad \forall n \in \mathcal{N},
\end{aligned} \quad (21)$$

where for each user $n \in \mathcal{N}$, the average data rate $r_n(\boldsymbol{p})$ is as in (8) and $P_n^{\min}$ and $P_n^{\max}$ are constant parameters such that $0 \leq P_n^{\min} \leq P_n^{\max} \leq 1$. Clearly, the objective function in (21) is the network aggregate throughput. Such an objective function is a *polynomial* function in transmission probabilities $\boldsymbol{p}$:

$$\sum_{n \in \mathcal{N}} \mu_n \, p_n \left[\sum_{\mathcal{M} \in \mathbb{M}_n} \left(\prod_{m \in \mathcal{M}} p_m \right) \left(\prod_{k \in \mathcal{N} \backslash \mathcal{M}, k \neq n} (1 - p_k) \right) \right].$$

Thus, problem (21) is *non-convex*. In fact, problem (21) is a *polynomial programming* problem which is usually difficult to solve. It is worth mentioning that a special class of polynomial programming problems, called *geometric programs*, are tractable [17]. A geometric program involves *minimizing* a *posynomial*, i.e., a polynomial which has only *positive* coefficients. However, problem (21) is usually not a geometric program. For example, the sum $r_1(\boldsymbol{p}) + r_2(\boldsymbol{p}) + r_3(\boldsymbol{p}) + r_4(\boldsymbol{p})$ of (13)-(16) has *both* positive and negative coefficients. Next, we explain how we can solve problem (21) in a centralized fashion using the *sum-of-squares* technique [16], [18]–[21].

We start by writing the *dual* problem corresponding to the *primal* problem in (21) as

$$\begin{aligned}
\min_{\boldsymbol{\lambda}, \boldsymbol{\delta}} \quad & \max_{\boldsymbol{p}} \; \sum_{n \in \mathcal{N}} r_n(\boldsymbol{p}) + \sum_{n \in \mathcal{N}} \lambda_n (p_n - P_n^{\min}) \\
& \qquad + \sum_{n \in \mathcal{N}} \delta_n (P_n^{\max} - p_n) \\
\text{s.t.} \quad & \lambda_n \geq 0, \; \delta_n \geq 0, \qquad \forall n \in \mathcal{N},
\end{aligned} \quad (22)$$

where for each $n \in \mathcal{N}$, λ_n denotes the Lagrange multiplier corresponding to constraint $p_n \geq P_n^{\min}$ and δ_n denotes the Lagrange multiplier corresponding to constraint $p_n \leq P_n^{\max}$. From *weak duality* [22], the optimum of dual problem (22) provides an *upper bound* for the optimum of primal problem (21). However, this upper bound may *not* be tight as problem (21) is *not* convex. In fact, this bound can be quite loose.

Next, we replace scalar λ_n with *polynomial* $\lambda_n(\boldsymbol{p})$ and scalar δ_n with *polynomial* $\delta_n(\boldsymbol{p})$ for each $n \in \mathcal{N}$. In other words, we replace *linear* Lagrange multipliers in dual problem (22) with *nonlinear* Lagrange multipliers. Furthermore, for each $n \in \mathcal{N}$, we replace constraint $\lambda_n \geq 0$ with

$$\lambda_n(\boldsymbol{p}) \geq 0, \quad \forall \boldsymbol{p}. \tag{23}$$

The above inequality indicates that for any choice of vector $\boldsymbol{p}$, the polynomial $\lambda_n(\boldsymbol{p})$ is non-negative for any $n \in \mathcal{N}$. Similarly, we replace constraint $\delta_n \geq 0$ with

$$\delta_n(\boldsymbol{p}) \geq 0, \quad \forall \boldsymbol{p}. \tag{24}$$

In that regard, a *generalized* dual problem for primal problem (21) is obtained as

$$\min_{\boldsymbol{\lambda}(\boldsymbol{p}),\boldsymbol{\delta}(\boldsymbol{p})} \quad \max_{\boldsymbol{p}} \sum_{n \in \mathcal{N}} r_n(\boldsymbol{p}) + \sum_{n \in \mathcal{N}} \lambda_n(\boldsymbol{p})(p_n - P_n^{\min})$$
$$+ \sum_{n \in \mathcal{N}} \delta_n(\boldsymbol{p})(P_n^{\max} - p_n) \tag{25}$$
$$\text{s.t.} \quad \lambda_n(\boldsymbol{p}) \geq 0, \ \delta_n(\boldsymbol{p}) \geq 0, \quad \forall n \in \mathcal{N}, \ \forall \boldsymbol{p}.$$

The optimization variables in (25) are the *coefficients* of polynomials $\boldsymbol{\lambda}(\boldsymbol{p}) = (\lambda_n(\boldsymbol{p}), \forall n \in \mathcal{N})$ and $\boldsymbol{\delta}(\boldsymbol{p}) = (\delta_n(\boldsymbol{p}), \forall n \in \mathcal{N})$. If the degrees of polynomials $\boldsymbol{\lambda}(\boldsymbol{p})$ and $\boldsymbol{\delta}(\boldsymbol{p})$ are zero, then problem (25) reduces to the original dual problem in (22).

Theorem 2: The generalized dual problem (25) provides an upper bound for the primal problem (21), which is *no worse* than the upper bound obtained from the dual problem (22).

The proof of Theorem 2 is given in Appendix B. Next, we will show that by replacing the scalar Lagrange multipliers $\lambda_1, \ldots, \lambda_N$ and $\delta_1, \ldots, \delta_N$ with the polynomial Lagrange multipliers $\lambda_1(\boldsymbol{p}), \ldots, \lambda_N(\boldsymbol{p})$ and $\delta_1(\boldsymbol{p}), \ldots, \delta_N(\boldsymbol{p})$, we can actually improve the upper bound eventually obtained by the dual problem and obtain the exact maximum. But first, let us introduce an auxiliary scalar variable t and rewrite problem (25) in the following *equivalent* form:

$$\min_{\boldsymbol{\lambda}(\boldsymbol{p}), \boldsymbol{\delta}(\boldsymbol{p}), t} \quad t$$
$$\text{s.t.} \quad t \geq \sum_{n \in \mathcal{N}} r_n(\boldsymbol{p}) + \sum_{n \in \mathcal{N}} \lambda_n(\boldsymbol{p})(p_n - P_n^{\min})$$
$$+ \sum_{n \in \mathcal{N}} \delta_n(\boldsymbol{p})(P_n^{\max} - p_n), \quad \forall \boldsymbol{p},$$
$$\lambda_n(\boldsymbol{p}) \geq 0, \ \delta_n(\boldsymbol{p}) \geq 0, \quad \forall n \in \mathcal{N}, \quad \forall \boldsymbol{p}. \tag{26}$$

In problem (26), the variables are scalar t and the *coefficients* of the polynomials $\boldsymbol{\lambda}(\boldsymbol{p})$ and $\boldsymbol{\delta}(\boldsymbol{p})$. Notice that the vector of transmission probabilities $\boldsymbol{p}$ is *not* a variable here as the constraints in (26) should hold for *any* choice of $\boldsymbol{p}$.

A polynomial $a(\boldsymbol{p})$ is a *sum-of-squares* (SOS) if there exist polynomials $a_1(\boldsymbol{p}), \ldots, a_\eta(\boldsymbol{p})$, for *some* $\eta \geq 1$, such that

$$a(\boldsymbol{p}) = \sum_{j=1}^{\eta} a_j^2(\boldsymbol{p}). \tag{27}$$

It is clear that $a(\boldsymbol{p})$ being an SOS naturally implies $a(\boldsymbol{p}) \geq 0$ for *any* real-valued $\boldsymbol{p}$. That is,

$$a(\boldsymbol{p}) \text{ is SOS} \quad \Rightarrow \quad a(\boldsymbol{p}) \geq 0, \quad \forall \boldsymbol{p}. \tag{28}$$

Furthermore, we notice that there exist efficient algorithms that can check if a polynomial is SOS, by using *semi-definite relaxation* techniques (cf. [19]). From this, together with the relation in (28), we replace all the inequality constraints in (26) with corresponding SOS-form constraints and rewrite problem (26) as the following *sum-of-squares programming* problem

$$\min_{\substack{\boldsymbol{\lambda}(\boldsymbol{p}), t, \\ \boldsymbol{\delta}(\boldsymbol{p})}} \quad t$$
$$\text{s.t.} \quad \left(t - \sum_{n \in \mathcal{N}} r_n(\boldsymbol{p}) - \sum_{n \in \mathcal{N}} \lambda_n(\boldsymbol{p})(p_n - P_n^{\min}) \right.$$
$$\left. - \sum_{n \in \mathcal{N}} \delta_n(\boldsymbol{p})(P_n^{\max} - p_n) \right) \text{ is SOS,}$$
$$\lambda_n(\boldsymbol{p}) \text{ is SOS,} \quad \forall n \in \mathcal{N},$$
$$\delta_n(\boldsymbol{p}) \text{ is SOS,} \quad \forall n \in \mathcal{N}. \tag{29}$$

We are now ready to state the following key results.

Theorem 3: For any choice of system parameters, we have
(a) Problem (29) is *equivalent* to problem (26).
(b) Problem (29) is a tractable optimization problem. That is, it can be solved in polynomial time.
(c) There exists an *even* and *finite* degree $d \geq 2$ for polynomials $\boldsymbol{\lambda}(\boldsymbol{p})$ and $\boldsymbol{\delta}(\boldsymbol{p})$ such that the minimum objective value of the SOS programming problem (29), denoted by $t^\star$, becomes *equal* to the maximum objective value of problem (21).

The proof of Theorem 3 is given in Appendix C. Theorem 3 summarizes some of the key results on optimization over *compact semi-algebraic sets* (cf. [16], [18]–[21]). From Theorem 3, when the physical model is used, the optimal network throughput can be obtained by solving the SOS programming problem (29), e.g., by for a *large enough* and *even* degree d for polynomials $\boldsymbol{\lambda}(\boldsymbol{p})$ and $\boldsymbol{\delta}(\boldsymbol{p})$. Problem (29) can be solved using efficient techniques such as the *interior-point method* [22]. In particular, problem (29) can be solved by the automated software SOSTOOLS [23]. SOSTOOLS uses a *sufficiency test* (cf. [24]) for the right selection of degree d which guarantees finding the optimal solution of the SOS program (29).

B. Max-Min Fair Resource Allocation

In this section, we consider the following *max-min fair* resource allocation problem in a random access system:

$$\max_{\boldsymbol{p}} \quad \min_{n \in \mathcal{N}} \quad r_n(\boldsymbol{p})$$
$$\text{s.t} \quad P_n^{\min} \leq p_n \leq P_n^{\max}, \quad \forall n \in \mathcal{N}. \tag{30}$$

Notice that in problem (30), our goal is to *balance* the performance among all users by maximizing the minimum

average data rate a user may achieve. This is a common design objective in random access literature (e.g., see [5]). Next, we explain how we can find the exact global optimal solution of problem (30) using similar techniques as in Section III-A.

By introducing a new *auxiliary* variable β, we can rewrite problem (30) as the following *equivalent* problem:

$$
\begin{aligned}
\max_{\boldsymbol{p}} \quad & \beta \\
\text{s.t} \quad & \beta \le r_n(\boldsymbol{p}), && \forall n \in \mathcal{N}, \quad (31) \\
& P_n^{\min} \le p_n \le P_n^{\max}, && \forall n \in \mathcal{N},
\end{aligned}
$$

Problem (31) is a *polynomial programming* problem. Thus, following similar steps as in Section III-A, we can formulate the *generalized* dual problem for primal problem (31) as the following *sum-of-squares programming* problem:

$$
\begin{aligned}
\min_{\substack{\boldsymbol{\zeta}(\boldsymbol{p}), \boldsymbol{\lambda}(\boldsymbol{p}), \\ \boldsymbol{\delta}(\boldsymbol{p}), t, \beta}} \quad & t \\
\text{s.t} \quad & \left(t - \sum_{n \in \mathcal{N}} \zeta_n(\boldsymbol{p})(r_n(\boldsymbol{p}) - \beta) - \sum_{n \in \mathcal{N}} \lambda_n(\boldsymbol{p})(p_n - P_n^{\min}) \right. \\
& \left. \quad - \sum_{n \in \mathcal{N}} \delta_n(\boldsymbol{p})(P_n^{\max} - p_n) \right) \text{ is SOS}, \\
& \zeta_n(\boldsymbol{p}) \text{ is SOS}, \quad \forall n \in \mathcal{N}, \\
& \lambda_n(\boldsymbol{p}) \text{ is SOS}, \quad \forall n \in \mathcal{N}, \\
& \delta_n(\boldsymbol{p}) \text{ is SOS}, \quad \forall n \in \mathcal{N}.
\end{aligned}
$$

$$(32)$$

Here, $\boldsymbol{\lambda}(\boldsymbol{p})$ and $\boldsymbol{\delta}(\boldsymbol{p})$ are as in Section III-A. Furthermore, we have $\boldsymbol{\zeta}(\boldsymbol{p}) = (\zeta_n(\boldsymbol{p}), \ \forall n \in \mathcal{N})$, where $\zeta_n(\boldsymbol{p})$ is a *nonlinear* Lagrange multiplier corresponding to constraint $\beta \le r_n(\boldsymbol{p})$ for each $n \in \mathcal{N}$. In problem (32), the variables are scalars t and β and the *coefficients* of polynomials $\boldsymbol{\zeta}(\boldsymbol{p})$, $\boldsymbol{\lambda}(\boldsymbol{p})$, and $\boldsymbol{\delta}(\boldsymbol{p})$.

Similar to Theorem 2, we can show that there exists an *even* and *finite* degree $d \ge 2$ for the choice of polynomials $\boldsymbol{\zeta}(\boldsymbol{p})$, $\boldsymbol{\lambda}(\boldsymbol{p})$, and $\boldsymbol{\delta}(\boldsymbol{p})$ such that the minimum objective value of the SOS programming problem (32), denoted by $t^\star$, becomes *exactly equal* to the maximum objective value of problem (30), i.e., the maximized minimum average data rate among all users. Thus, the max-min fair resource allocation problem in (30) can be solved easily by using SOSTOOLS [23].

C. Example

Consider the example in Section II-C. Assume that the peak data rates are $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 1$. Solving problem (32) with SOSTOOLS with degree $d = 6$, the optimal transmission probabilities are obtained as $p_1^\star = 0.4474$, $p_2^\star = 0.4800$, $p_3^\star = 0.3378$, and $p_4^\star = 0.7704$, and the corresponding data rates are

$$r_1(\boldsymbol{p}^\star) = r_2(\boldsymbol{p}^\star) = r_3(\boldsymbol{p}^\star) = r_4(\boldsymbol{p}^\star) = 0.2652. \quad (33)$$

In this case, the minimum data rate among users, i.e., the max-min fairness objective function becomes

$$\min_{n \in \mathcal{N}} r_n(\boldsymbol{p}^\star) = 0.2652. \quad (34)$$

On the other hand, if the protocol model is used, i.e., the data rates are assumed to be as in (17)-(20), then the optimal trans-

mission probabilities become $\tilde{p}_1^\star = \frac{1}{2}$, $\tilde{p}_2^\star = 1$, $\tilde{p}_3^\star = \frac{1}{2}$, and $\tilde{p}_4^\star = 1$. If we apply these probabilities to (17)-(20), we obtain

$$\tilde{r}_1(\tilde{\boldsymbol{p}}^\star) = \tilde{r}_2(\tilde{\boldsymbol{p}}^\star) = \tilde{r}_3(\tilde{\boldsymbol{p}}^\star) = \tilde{r}_4(\tilde{\boldsymbol{p}}^\star) = 0.5. \quad (35)$$

The results in (35) *misleadingly* indicate that all users achieve a data rate of 0.5. However, to find the actual data rates of the users in this case, we need to apply the above probability values to (13)-(16). In that case, we have

$$r_1(\tilde{\boldsymbol{p}}^\star) = 0, \ \ r_2(\tilde{\boldsymbol{p}}^\star) = 0.5, \ \ r_3(\tilde{\boldsymbol{p}}^\star) = 0.25, \ \ r_4(\tilde{\boldsymbol{p}}^\star) = 0.25. \quad (36)$$

Thus, the minimum data rate among the users at $\tilde{\boldsymbol{p}}^\star$ becomes

$$\min_{n \in \mathcal{N}} r_n(\tilde{\boldsymbol{p}}^\star) = \min\{0, 0.5, 0.25, 0.25\} = 0. \quad (37)$$

By comparing (34) and (37), we conclude that the *performance loss* due to the inaccuracy of the protocol model is *unbounded*: $\frac{0.2652 - 0}{0}$. This underlines the importance of using the physical model in an optimization-based random access design.

IV. DESIGN II: DISTRIBUTED AND NEAR-OPTIMAL

Both the throughput maximization and the max-min fair resource allocation problems in Section III are polynomial programming programs. However, this property may not hold for many other optimization problems in random access systems, making the sum-of-squares technique an inappropriate tool for solving them. Furthermore, we are usually interested in selecting transmission probabilities in a *distributed* fashion rather than a centralized fashion. Finally, the complexity of the sum-of-squares programming makes our design in Section III most appropriate for small-scale networks.

We will address all the above performance bottlenecks in this section. In particular, we consider a more general *network utility maximization* problem in the following form:

$$
\begin{aligned}
\max_{\boldsymbol{p}} \quad & \sum_{n \in \mathcal{N}} U_n(r_n(\boldsymbol{p})) \\
\text{s.t} \quad & P_n^{\min} \le p_n \le P_n^{\max}, \quad \forall n \in \mathcal{N},
\end{aligned}
\quad (38)
$$

where for each user $n \in \mathcal{N}$, the average data rate $r_n(\boldsymbol{p})$ is as in (3) and the utility function $U_n(r_n(\boldsymbol{p}))$ is an *increasing*, *concave*, and *differentiable* function in $r_n(\boldsymbol{p})$. In general, the utility function for each user indicates the user's level of satisfaction upon its achieved average data rate.

An important class of utility functions is α-fair utilities [25], which are formulated with parameter $\alpha \ge 0$ as

$$
U_n(r_n(\boldsymbol{p})) = \begin{cases} (1 - \alpha)^{-1} \, r_n(\boldsymbol{p})^{1-\alpha}, & \text{if } \alpha \neq 1, \\ \log r_n(\boldsymbol{p}), & \text{if } \alpha = 1. \end{cases} \quad (39)
$$

Using (39), a wide range of efficient and fair resource allocations can be modeled. In particular, problem (38) reduces to throughput maximization with $\alpha = 0$, to proportional fair allocation with $\alpha = 1$, to harmonic-mean fair allocation with $\alpha = 2$, and to max-min fairness with $\alpha \to \infty$ [3], [25]. In fact, the two optimization problems studied in Sections III-A and III-B are special cases[1] of the network utility maximization problem in (38) with α-fair utility functions as in (39).

---

[1]The optimal solution of problem (38), when $\alpha \to \infty$, is also an optimal solution for problem (30) [25, Lemma 3]. However, the converse is not true.

A. Algorithm

For each $n \in \mathcal{N}$, let $\boldsymbol{p}_{-n} = (p_1, \ldots, p_{n-1}, p_{n+1}, \ldots p_N)$ denote the $(N-1)$-dimensional vector of transmission probabilities of all users *other than* user n. Consider the following *local* and *myopic* optimization problem for user $n \in \mathcal{N}$:

$$\max_{p_n} \quad U_n\big(p_n\, o_n(\boldsymbol{p}_{-n})\big)$$
$$+ \sum_{m \neq n} U_m\big(p_n\, v_{nm}(\boldsymbol{p}_{-n}) + (1-p_n)\, w_{nm}(\boldsymbol{p}_{-n})\big) \quad (40)$$
$$\text{s.t} \quad P_n^{\min} \leq p_n \leq P_n^{\max}.$$

where

$$o_n(\boldsymbol{p}_{-n}) = \mu_n \left[\sum_{\mathcal{M} \in \mathbb{M}_n} \left(\prod_{m \in \mathcal{M}} p_m \right) \left(\prod_{k \in \mathcal{N} \setminus \mathcal{M}, k \neq n} (1 - p_k) \right) \right],$$

and for each user $m \in \mathcal{N} \setminus \{n\}$, we have

$$v_{nm}(\boldsymbol{p}_{-n}) = \mu_m\, p_m$$
$$\times \left[\sum_{\mathcal{M} \in \mathbb{M}_m,\, n \in \mathcal{M}} \left(\prod_{k \in \mathcal{M} \setminus \{n\}} p_k \right) \left(\prod_{l \in \mathcal{N} \setminus \mathcal{M}, l \neq m} (1 - p_l) \right) \right],$$

and

$$w_{nm}(\boldsymbol{p}_{-n}) = \mu_m\, p_m$$
$$\times \left[\sum_{\mathcal{M} \in \mathbb{M}_m,\, n \notin \mathcal{M}} \left(\prod_{k \in \mathcal{M}} p_k \right) \left(\prod_{l \in \mathcal{N} \setminus \mathcal{M}, l \neq m, n} (1 - p_l) \right) \right].$$

The objective functions in problems (38) and (40) are the *same*. Thus, by solving problem (40), user n can select its own transmission probability p_n to maximize the objective value in network utility maximization problem (38), assuming that $\boldsymbol{p}_{-n}$ is *fixed*, i.e., *no* other user changes its transmission probability.

Theorem 4: For each user $n \in \mathcal{N}$, the *local* optimization problem (40) is a convex optimization problem in p_n.

The proof of Theorem 4 is evident and is omitted for brevity. From Theorem 4, each user $n \in \mathcal{N}$ can easily solve problem (40) by using the *interior-point method* [22] via *local* iterations. This motivates us to introduce the proposed distributed algorithm, Algorithm 1, for solving problem (38). Algorithm 1 works based on the *coordinate ascent* method [26, p. 207], where we fix all of the components of vector $\boldsymbol{p}$, except for the nth component, and maximize the network aggregate utility only with respect to p_n as in (40). This procedure is repeated, leading to an iterative algorithm. Let $\mathcal{T}_n$ denote the set of time slots at which user $n \in \mathcal{N}$ solves problem (40) using the interior-point method. We assume that: (a) For any users $n \neq m$, we have $\mathcal{T}_n \cap \mathcal{T}_m = \{\}$. That is, the iterative local maximizations are carried out *successively* as in the *Gauss-Seidel* mapping [26, p. 21]. (b) There is a constant $T^{\max}$ such that for each user $n \in \mathcal{N}$, there exist $t_1, t_2 \in \mathcal{T}_n$ such that $|t_1 - t_2| \leq T^{\max}$. In other words, all users update their transmission probabilities at least once every $T^{\max}$ time slots.

In Algorithm 1, lines 6 to 14 are executed repeatedly until user n stops operation, switches off, or leaves the network. In line 6, user n may attempt transmitting a packet according to its transmission probability p_n. In lines 8 to 10, user n updates p_n and then broadcasts a control message including the new

---

Algorithm 1 : Executed by each user $n \in \mathcal{N}$.

1: Allocate memory for $\boldsymbol{p}_n$ and $\boldsymbol{p}_{-n}$.
2: Randomly initialize $p_n \in [P_n^{\min}, P_n^{\max}]$.
3: Randomly initialize $\boldsymbol{p}_{-n}$ such that $p_m \in [P_m^{\min}, P_m^{\max}]$ for $m \neq n$.
4: Set clock timer t.
5: **repeat**
6:     Transmit with probability p_n.
7:     **if** $t \in \mathcal{T}_n$ **then**
8:       Solve problem (40) using the interior-point method [22].
9:       Update p_n according to the solution.
10:      Broadcast a control message to announce p_n to other users.
11:     **end if**
12:     **if** a control message is received **then**
13:       Update $\boldsymbol{p}_{-n}$ accordingly.
14:     **end if**
15: **until** user n stops operation or leaves the network.

---

value of p_n, whenever the current time slot is in set $\mathcal{T}_n$. Upon reception of a control message from another user $m \neq n$, user n updates its *local memory* $\boldsymbol{p}_{-n}$ accordingly in line 13.

It is worth mentioning that if we use the protocol model (instead of the physical model) to formulate problem (38), then Algorithm 1 reduces to [4, Algorithm 2].

B. Convergence and Optimality

In this section, we investigate the convergence and optimality of Algorithm 1. Let $\Theta(t)$ denote the current network aggregate utility, i.e., the objective value of problem (38), at time slot $t \geq 0$. We can show the following.

Theorem 5: For any choice of system parameters and starting from any initial point, there exists some Θ^* such that

$$\Theta^* = \lim_{t \to \infty} \Theta(t). \quad (41)$$

That is, Algorithm 1 converges.

Theorem 5 directly results from [26, Proposition 2.5, p. 208] on convergence analysis of the coordinate ascent method with Gauss-Seidel updates. It is shown in [27] that the rate of convergence for the coordinate ascent method in *Gauss-Seidel* mappings is at least linear. We can further show that:

Theorem 6: Any fixed point of Algorithm 1 is a stationary point of problem (38). That is, it is at least a *local* optimal solution for the non-convex optimization problem (38).

The proof of Theorem 6 is given in Appendix D. From Theorems 5 and 6, convergence and local optimality of Algorithm 1 are guaranteed. Clearly, the obtained transmission probabilities may *not* be globally optimal. However, simulation results in Section V show that Algorithm 1 usually results in close-to-optimal performance, making it a practical algorithm for large networks when using SOSTOOLS is not feasible.

V. SIMULATION RESULTS

In this section, we assess the performance of the proposed design schemes via simulations. Our main focus is on investigating how the physical model can improve the performance of optimization-based random access compared to the protocol model. In this regard, we also study the impact of various system parameters, such as the SINR threshold, the number of users in the network, and the choice of utility functions.

Unless we state otherwise, the common simulation setting in all experiments is as follows. The peak data rates $\mu_1, \ldots, \mu_N$ are selected randomly between 1 to 11 Mbps as in the IEEE 802.11b standard [10]. For any pair of users $n, m \in \mathcal{N}$, the channel gain G_{nm} is calculated as the inverse of the square of the distance between nodes t_m and d_n according to the Friis free space model [28]. For each user $n \in \mathcal{N}$, the transmission power q_n and the noise power σ_n^2 are selected randomly to achieve a *signal-to-noise-ratio* which is 0 to 3 dB *higher* than the threshold $\gamma_n^{th} = 1$ in the absence of any interference. We also set the bound parameters $P_n^{\min} = 0.01$ and $P_n^{\max} = 0.99$ for each user $n \in \mathcal{N}$. Finally, all the results for the *protocol model* are obtained by using [4, Algorithm 2].

A. Optimality

Recall that Design I in Section III leads to the global optimal solutions of the throughput maximization and max-min fair resource allocation problems, while Design II in Section IV only guarantees finding a local optimal solution. In this section, we use Design I as a benchmark to evaluate the performance of Design II, and also compare the network performance achievable with the physical model and the protocol model, respectively. Here, we limit our attention to only small networks such that Design I is applicable. Larger network topologies will be studied in Sections V-B, V-C, and V-D.

Simulation results for 20 random scenarios are shown in Fig. 1. In each scenario, the network topology includes $N = 4$ users which are randomly located in a 50 m $\times$ 50 m field. For each user $n \in \mathcal{N}$, the distance between node t_n and node d_n is selected randomly between 5 to 25 m. For the results in Fig. 1, the network is designed to solve the max-min fair resource allocation problem in (30). We can see that both Design I and Design II significantly outperform the case when the protocol model is used. The average performance gain[2] across all 20 scenarios is 4.22 for Design I and 4.17 for Design II. In fact, as we discussed in Section III-C, it is important for the max-min fair resource allocation to be based on the physical model. Another key observation in Fig. 1 is that Design II achieves *close-to-optimal* performance for all considered scenarios. On average, Design II achieves 98.8% optimality, compared to Design II. Similar results can be obtained for the network throughput maximization problem (21).

For the rest of the simulations, we focus only on Design II which is applicable to large-scale networks and leads to a close-to-optimal performance.

B. Impact of the SINR Threshold

Next, we investigate the impact of changing the SINR threshold $\gamma_1^{th}, \ldots, \gamma_N^{th}$ on the performance gain of using the physical model versus the protocol model. We consider 100 scenarios, where in each scenario, $N = 10$ users are randomly located in a 100 m $\times$ 100 m square field. The simulation results when the SINR threshold varies from 0 to 5 are shown in Fig. 2. For the results in this figure, the network is designed

[2]Here, the *performance gain* is defined as the *ratio* of the minimum data rate achieved among the users when Design I (or Design II) is used to the minimum data rate among the users when the protocol model is used.
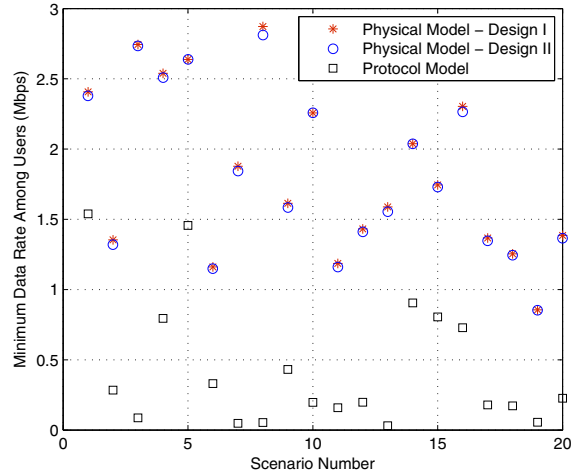

Fig. 1. Comparison of the achievable network performance with the physical and protocol models for 20 random scenarios. The performance is measured in terms of solving the max-min fair resource allocation problem (30). By using the physical model, we can significantly improve the performance in all scenarios. Furthermore, Design II leads to a *close-to-optimal* performance compared to the optimal performance achieved by Design I.
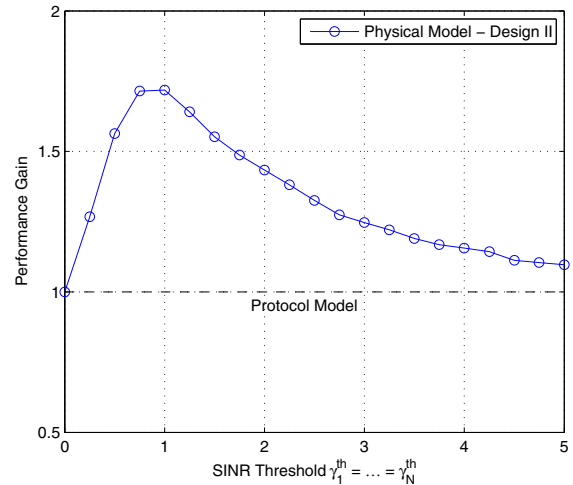

Fig. 2. The impact of changing the SINR thresholds on the performance gain achieved due to using the physical model. Here, the design objective is to solve the network utility maximization problem (38) for logarithmic utility functions in order to achieve proportional fairness among users. Each point in this figure indicates the average results for 100 simulated scenarios.

to solve the network utility maximization problem (38) when the utility functions are *logarithmic* in order to achieve *proportional fairness* among the users. Each point represents the average results across all 100 simulated scenarios. We can see that if the SINR threshold becomes too small (i.e., approaches zero) or too large, the performance gain tends to 1. The first case corresponds to a scenario where *all* users can be active at the same time. The second case implies a scenario where only *one* user can be active at a time. In either case, the data rates obtained from the physical model and the protocol model will be almost the same. For other cases, the difference between the physical model and the protocol model (and consequently the performance gain achievable with the physical model) is significant. In particular, when $\gamma_1^{th} = \ldots = \gamma_N^{th} = 1$, the average performance gain is as high as 1.72.
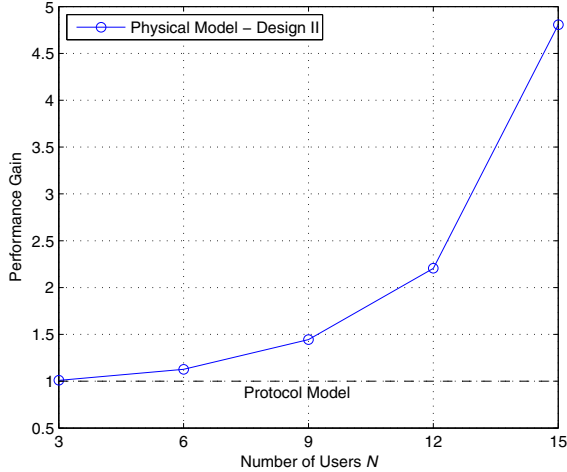
Fig. 3. The impact of changing the number of users in the network on the performance gain achievable by adopting the physical model. Each point in this figure indicates the average results for 100 simulated scenarios.

Fig. 4. The impact of changing the utility parameter α on network efficiency and fairness. Each point indicates the average results for 100 simulated scenarios with $N = 10$ users. (a) Aggregate network throughput, (b) Jain's fairness index. We can see that for Design II, parameter α acts as a *knob* to control the tradeoff between efficiency and fairness.

C. Impact of the Number of Users

Next, we change the number of users N from 3 to 15 and set $\gamma_n^{th} = 1$ for all $n \in \mathcal{N}$. The rest of the parameters are as in Section V-B. The results are shown in Fig. 3. As N increases, the network becomes more *dense* and the *aggregate interference* from multiple interfering sources can be a major cause for packet loss, making the protocol model very inaccurate. The performance improvement due to using the physical model when the number of users is $N = 15$ is as high as 4.82.

D. Impact of the Utility Parameter α

It is known that if the utility functions are selected to be α-fair as in (39), then parameter α can act as a *knob* to control the tradeoff between network efficiency and fairness in network utility maximization problem (38) [3], [4], [7]. In this section, we investigate the impact of parameter α on the network performance achievable with the physical and protocol models.

The simulation results when the utility parameter α varies from 0.5 to 5 are shown in Fig. 4. Each point represents the average results for 100 random scenarios where the simulation setting is as in Section V-B. We measure the network efficiency in terms of the achieved network throughput. We also use Jain's fairness index [29] to measure fairness: $(\sum_{n \in \mathcal{N}} r_n(\boldsymbol{p}))^2 / (N \sum_{n \in \mathcal{N}} r_n(\boldsymbol{p})^2)$. We can see that if the physical model is used, then we can make the system fairer but less efficient (and vice versa) by changing parameter α. However, this trend is *not* seen when the protocol model is used. In fact, in this case, the performance is more or less independent of the choice of α, particularly when it comes to the achievable throughput. Furthermore, we can see that the performance gain due to using the physical model is significant for *both* throughput and fairness index.

VI. CONCLUSION

In this paper, we studied optimization-based design in wireless random access systems. Unlike most of the previous work along this line that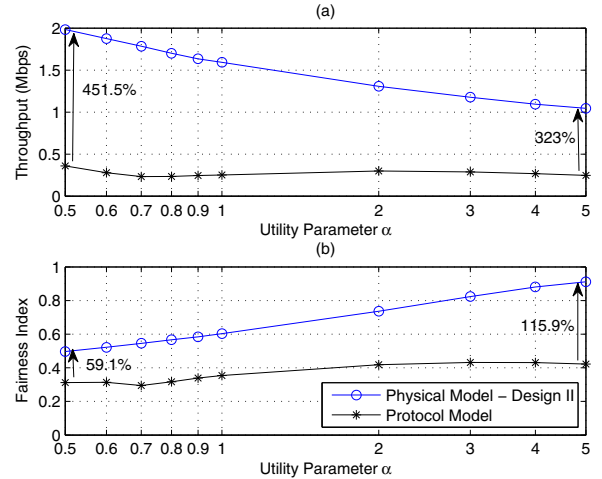 has used the *protocol model*, here we adopted the *physical model* which is based on the *signal-to-interference-plus-noise-ratio* (SINR), in order to formulate three problems: *throughput maximization*, *max-min fair resource allocation*, and *network utility maximization*. The new formulations entail difficult non-convex optimization problems. We proposed two schemes to solve these problems. The first design is centralized and leads to the optimal solutions. However, due to its complexity, it is only applicable to small networks. The second design is distributed, leads to a close-to-optimal performance, and is applicable to large networks. We performed various simulations to compare the network performance achievable with the physical and protocol models. We showed that it is highly preferable to use the physical model for optimization-based random access design, particularly when the network is *dense*. In this regard, even a sub-optimal design based on the (accurate) physical model can achieve a significant performance gain compared to an optimal design based on the (inaccurate) protocol model.

APPENDIX

A. Proof of Theorem 1

From (6) and (9), for each user $n \in \mathcal{N}$, we have

$$\mathcal{I}_n \subseteq \mathcal{N} \backslash \mathcal{M}, \quad \forall \mathcal{M} \in \mathbb{M}_n. \tag{42}$$

Thus, $(\prod_{m \in \mathcal{I}_n} (1 - p_m))$ is a common factor in *all* terms in the summation in (7) and we further have

$$
\text{Prob} \left\{ i_n \leq \frac{q_n G_{nn}}{\gamma_n^{th}} - \sigma_n^2 \right\} = \left(\prod_{m \in \mathcal{I}_n} (1 - p_m) \right)
$$

$$
\times \left[\sum_{\mathcal{M} \in \mathbb{M}_n} \left(\prod_{m \in \mathcal{M}} p_m \right) \left(\prod_{k \in \mathcal{N} \backslash (\mathcal{M} \cup \mathcal{I}_n), \, k \neq n} (1 - p_k) \right) \right] \tag{43}
$$

$$
\leq \left(\prod_{m \in \mathcal{I}_n} (1 - p_m) \right).
$$

From (43), (3), and (10), the inequality in (11) results. ∎

B. Proof of Theorem 2

(a) Suppose $\tilde{\boldsymbol{p}}$ is a feasible point for the primal problem in (21); i.e., $0 \leq \tilde{p}_n \leq 1$ for all $n \in \mathcal{N}$. Then, for any choice of polynomials $\boldsymbol{\lambda}(\boldsymbol{p})$ and $\boldsymbol{\delta}(\boldsymbol{p})$ which satisfy (23) and (24),

$$\sum_{n\in\mathcal{N}} \lambda_n(\tilde{\boldsymbol{p}})(\tilde{p}_n - P_n^{\min}) + \sum_{n\in\mathcal{N}} \delta_n(\tilde{\boldsymbol{p}})(P_n^{\max} - \tilde{p}_n) \geq 0. \quad (44)$$

For notational simplicity, we define $f(\boldsymbol{p}) = \sum_{n\in\mathcal{N}} r_n(\boldsymbol{p})$. By adding $f(\tilde{\boldsymbol{p}})$ to both sides of (44), we have

$$\begin{aligned}
\max_{\boldsymbol{p}} \;\; & f(\boldsymbol{p}) + \sum_{n\in\mathcal{N}} \lambda_n(\boldsymbol{p})(p_n - P_n^{\min}) + \sum_{n\in\mathcal{N}} \delta_n(\boldsymbol{p})(P_n^{\max} - p_n) \\
\geq & f(\tilde{\boldsymbol{p}}) + \sum_{n\in\mathcal{N}} \lambda_n(\tilde{\boldsymbol{p}})(\tilde{p}_n - P_n^{\min}) + \sum_{n\in\mathcal{N}} \delta_n(\tilde{\boldsymbol{p}})(P_n^{\max} - \tilde{p}_n) \\
\geq & f(\tilde{\boldsymbol{p}}).
\end{aligned} \quad (45)$$

Since (45) holds for *any* feasible $\tilde{\boldsymbol{p}}$, the minimization in (25) forms an *upper bound* for the primal problem in (21).

(b) This part results from the fact that the dual problem (22) is a special case of the generalized dual problem (25) when the polynomials $\boldsymbol{\lambda}(\boldsymbol{p})$ and $\boldsymbol{\delta}(\boldsymbol{p})$ have degree zero. ∎

C. Proof of Theorem 3

(a) This part directly results from the *Positivstellensatz* theorem in [21], which was based on the earlier results in [20]. A short proof is provided in [18, Proof of Theorem 3].

(b) This part is based on the results in [19, Section 3].

(c) This part is based on the results in [18, Theorem 12 and Corollary 13]. Sufficient upper bounds on the choice of degree d are also available in [30] for a few special cases. ∎

D. Proof of Theorem 6

Let $\boldsymbol{p}^*$ denote any fixed point of Algorithm 1. That is, for any $n \in \mathcal{N}$, given $\boldsymbol{p}_{-n} = \boldsymbol{p}_{-n}^*$, $p_n = p_n^*$ is optimal for (40). Since (40) is *convex*, the fixed point $\boldsymbol{p}^*$ satisfies the necessary and sufficient Karush-Kuhn-Tucker (KKT) optimality conditions [22] corresponding to (40) for all $n \in \mathcal{N}$. By definition, each stationary point [31] of *non-convex* problem (38) also satisfies the KKT conditions for (38). Since the objectives in (38) and (40) are the same and the set of constraints in (38) is the union of those in (40) for all $n \in \mathcal{N}$, the KKT conditions for (38) are equal to the *union* of those for (40) for all $n \in \mathcal{N}$. Thus, since $\boldsymbol{p}^*$ satisfies the KKT conditions of (40) for *all* users, it also satisfies the KKT conditions for (38), i.e., each fixed point $\boldsymbol{p}^*$ is a local optimal solution for problem (38). ∎

REFERENCES

[1] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. on Information Theory*, vol. 46, pp. 388–404, Mar. 2000.

[2] K. Kar, S. Sarkar, and L. Tassiulas, "Achieving proportionally fair rates using local information in Aloha networks," *IEEE Trans. on Automatic Control*, vol. 49, no. 10, pp. 1858–1862, Oct. 2004.

[3] J. Lee, M. Chiang, and R. Calderbank, "Utility-optimal random-access control," *IEEE Trans. on Wireless Communications*, vol. 25, pp. 1135–1147, Aug. 2007.

[4] A. H. Mohsenian-Rad, J. Huang, M. Chiang, and V. W. S. Wong, "Utility-optimal random access: Reduced complexity, fast convergence, and robust performance," *IEEE Trans. on Wireless Communications*, vol. 8, pp. 898–911, Feb. 2009.

[5] X. Wang and K. Kar, "Distributed algorithms for max-min fair rate allocation in Aloha networks," in *Proc. of 42nd Annual Allerton Conference*, Urbana-Champaign, IL, Oct. 2004.

[6] ——, "Cross-layer rate control for end-to-end proportional fairness in wireless networks with random access," *IEEE J. on Selected Areas in Communications*, vol. 24, pp. 1548–1559, Aug. 2006.

[7] J. Lee, M. Chiang, and R. Calderbank, "Jointly optimal congestion and contention control in wireless ad hoc networks," *IEEE Communications Letters*, vol. 10, no. 3, pp. 216–218, Mar. 2006.

[8] T. Cui, L. Chen, and S. Low, "A game-theoretic framework for medium access control," *IEEE J. on Selected Areas in Comm.*, vol. 26, pp. 1116–1127, Sept. 2008.

[9] H. Inaltekin and S. B. Wicker, "The analysis of Nash equilibria of the one-shot random-access game for wireless networks and the behavior of selfish nodes," *IEEE/ACM Trans. on Networking*, vol. 16, no. 5, pp. 1094–1107, Oct. 2008.

[10] "IEEE 802.11b," http://standards.ieee.org/getieee802/download/802.11b-1999.pdf.

[11] M. M. Carvalho and J. J. Garcia-Luna-Aceves, "A scalable model for channel access protocols in multihop ad-hoc networks," in *Proc. of ACM MobiCom'04*, Philadelphia, PA, Oct. 2004.

[12] H. Chang, V. Misra, and D. Rubenstein, "A general model and analysis of physical layer capture in 802.11 networks," in *Proc. of IEEE INFO-COM*, Barcelona, Spain, April 2006.

[13] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted Aloha with multipacket reception," *IEEE Trans. on Information Theory*, vol. 51, pp. 2636–2656, July 2005.

[14] R. H. Gau, "Performance analysis of slotted Aloha in interference-dominating wireless ad-hoc networks," *IEEE Communications Letter*, vol. 10, pp. 402–404, May 2006.

[15] M. Fazel and M. Chiang, "Network utility maximization with nonconcave utilities using sum-of-squares method," in *Proc. of IEEE Conference on Decision and Control*, Seville, Spain, Dec. 2005.

[16] M. Chiang, "Nonconvex optimization of communication systems," *Advances in Mechanics and Mathematics*, vol. 3, pp. 137–196, 2008.

[17] ——, "Geometric programming for communication systems," *Short monograph in Foundations and Trends in Communications and Information Theory*, vol. 2, pp. 1–154, Aug. 2005.

[18] M. Schweighofer, "Optimization of polynomials on compact semi-algebraic sets," *SIAM J. on Optimization*, vol. 15, pp. 805–825, 2005.

[19] P. A. Parrilo, "Semidefinite programming relaxations for semi-algebraic problems," *Mathematical Programming*, vol. 96, pp. 293–320, 2003.

[20] K. Schmudgen, "The k-moment problem for compact semi-algebraic sets," *Mathematische Annalen*, vol. 289, pp. 203–206, 1991.

[21] M. Putinar, "Positive polynomials on compact semi-algebraic sets," *Indiana University Mathematics Journal*, vol. 42, pp. 969–984, 1993.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[23] S. Pranja, A. Papachristodoulou, P. Seiler, and P. A. Parrilo, "SOS-TOOLS: Sum of squares optimization toolbox for MATLAB - user's guide," http://www.mit.edu/~parrilo/sostools/sostools.pdf, June 2004.

[24] D. Henrion and J. B. Lasserre, "Detecting global optimality and extracting solutions in GloptiPoly," *Positive Polynomials in Control (Lecture Notes in Control and Info. Sci.)*, vol. 312, pp. 293–310, Jan. 2005.

[25] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. on Networking*, vol. 8, pp. 556–567, Oct. 2000.

[26] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Prentice Hall, 1989.

[27] Z. Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *Journal of Optimization Theory and Applications*, vol. 72, pp. 7–35, Jan. 1992.

[28] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw-Hill, 2007.

[29] R. Jain, W. Hawe, and D. Chiu, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Tech. Rep. DEC-TR-301, Sept. 1984.

[30] M. Schweighofer, "An algorithmic approach to Schmudgen's Positivstellensatz," *J. of Pure and Applied Algebra*, vol. 166, pp. 307–319, 2002.

[31] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Sci., 2004.