This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2017.2764081, IEEE Transactions on Wireless Communications

Game Theoretic Perspective of Optimal CSMA

Hyeryung Jang, Se-Young Yun, Jinwoo Shin, and Yung Yi

***Abstract*—Game-theoretic approaches have provided valuable insights into the design of robust local control rules for the individuals in multi-agent systems, e.g., Internet congestion control, road transportation networks, etc. In this paper, we introduce a non-cooperative Medium Access Control game for wireless networks and propose new fully-distributed CSMA (Carrier Sense Multiple Access) algorithms that are provably optimal in the sense that their long-term throughputs converge to the optimal solution of a utility maximization problem over the maximum throughput region. The most significant part of our approach lies in introducing novel price functions in agents' utilities so that the proposed game admits an ordinal potential function with no Price-of-Anarchy. The game formulation naturally leads to game-based dynamics finding a Nash equilibrium, but they often require global information. Towards our goal of designing fully-distributed operations, we propose new game-inspired dynamics by utilizing a certain property of CSMA that enables links to estimate their temporary throughputs without message passing. They can be thought as stochastic approximations to the standard dynamics, which is a new feature in our work, not prevalent in other traditional game-theoretic approaches. We show that they converge to a Nash equilibrium, and numerically evaluate their performance to support our theoretical findings.**

***Index Terms*—CSMA, Distributed algorithms, Game theory, Wireless ad-hoc network, Stochastic approximation.**

I. Introduction

In many engineering systems, we often observe the trade-off between efficiency and complexity, where optimal algorithms require heavy computational challenges or extensive message passing, but light-weight approximate algorithms incur severe efficiency degradation. MAC (Medium Access Control) in wireless networks is no exception. The seminal work is done by Tassiulas and Ephremides [2], referred to as Max-Weight scheduling, which is centralized and computationally intractable (for a large-scale network). The high complexity in Max-Weight stems from the fact that an NP-hard problem (maximum weight independent set problem) has to be solved repeatedly over time. Since then, various subsequent papers based on many principles, *e.g.*, random access [3], [4], pick-and-compare [5]–[8], and maximal/greedy [9]–[11], have been published, and most of them more or less show that the trade-off between efficiency and complexity indeed exists, *e.g.*, see [12], [13] for surveys.

Manuscript received December 29, 2016; revised April 20, 2017, and August 17, 2017. The associate editor coordinating the review of this paper and approving it for publication was S. Wang.

Part of this work was presented at IEEE INFOCOM 2014, Toronto [1].

This work was partly supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No.B0717-16-0119), and the Center for Integrated Smart Sensors funded by the Ministry of Science, ICT & Future Planning as Global Frontier Project (CISS-2012M3A6A6054195).

H. Jang, J. Shin, and Y. Yi are with the Department of Electrical Engineering, KAIST, South Korea (e-mails: {hrjang357, jinwoos, yiyung}@kaist.ac.kr). S. Yun is with the Department of Industrial and Systems Engineering, KAIST, South Korea (e-mail:yunseyoung@gmail.com).

In this paper, we aim at developing fully-distributed MAC algorithms (no message passing of information between neighbors) that are utility-optimal over the maximum throughput region (as achieved by Max-Weight). To that end, we adopt CSMA (Carrier Sense Multiple Access) as a base-line MAC, where we smartly update CSMA operational parameters, so that the long-term throughput over links forms the optimal solution of a utility maximization problem. Recently, there have been theoretical works that fully-distributed MAC algorithms based on CSMA can achieve optimality in both throughput and utility, *e.g.*, [14]–[17] and [18] as a survey, as well as some studies on practicalizing such theoretically developed algorithms, *e.g.*, redesigning 802.11 DCF [19]. We refer to this family of work as *optimal CSMA*. All theory-based optimal CSMA works in literatures develop different algorithms commonly based on the framework of optimization. In this paper, we take a different angle by formulating a non-cooperative CSMA game, to study complex interactions due to interference among independent wireless links, and furthermore to present insights into the design of different types of distributed control rules.

Game theory has often been emerged as a powerful tool not only to understand the strategic behavior but also to optimize its distributed process in a competitive multi-agent systems [20], where agents just optimize their local objectives reacting to limited network information, yet their local decisions often result in a system-wide efficient performance. In applying game theory, there are some important questions, such as: (i) Is there Nash equilibrium (*i.e.*, the steady state of the game)? (ii) Is it unique? (iii) Is it also a global optimum of the system, *i.e.*, does it maximize the social welfare? Especially in (iii), it is aimed that a game is artificially designed, so that a game has a desirable solution in terms of uniqueness and social optimality as a steady state.

A game theoretic approach naturally leads to popular learning dynamics in classical game theory, *e.g.*, the best response dynamics, which are run by each player, yet require the information of other players, thereby incurring heavy message passing in general. In particular, game dynamics are interactive: each player's learning process correlates and affects what has to be learned by every other players over time. Then, one more important but challenging question in this perspective is: (iv) Does the system converge to a good equilibrium? It has been shown by Hart and Mas-Colell [21] that for a broad class of games, there is no general algorithm which allows the players' period-by-period behavior (even not fully-distributed) to converge to a Nash equilibrium (even if it exists). Also, there exists a distributed game learning dynamics in [22], which, however, provides only probabilistic convergence guarantee under highly strict conditions such as finiteness of a game. This paper inherits such philosophy of a game-theoretic approach for distributed optimization. Our main contributions are summarized in what follows:

C1. We construct a non-cooperative CSMA game with artificially-selected payoff functions, and characterize the existence, uniqueness, efficiency of Nash equilibrium (in terms of achieving the social optimality, *i.e.*, Price-of-Anarchy). In our game, each link uses its own CSMA operational parameter, backoff and holding time, as a strategy, and the payoff function is designed to reflect both (i) the long-term utility from the network, and (ii) the price measured by the harmful effects on other links. We prove that the game is an ordinal potential game and has the unique (non-trivial) Nash equilibrium which is equivalent to the socially optimal point.

C2. Inspired by the popular dynamics in game theory, we develop three fully-distributed dynamics exploiting the feature of CSMA that each link's temporary throughput is naturally locally-observable without message passing. Each link adjusts its parameter only in response to its locally realized throughput without knowledge of the game structure, and without observing the behaviors and/or throughputs of the other links. We theoretically provide provable convergence to the unique non-trivial Nash equilibrium that corresponds to the utility-maximizing solution (*i.e.*, socially optimal point) from ***C1.*** Technical challenges lie in complex inter-plays between the update of CSMA's operational parameters and the underlying dynamics of MAC in wireless networks, where non-trivial time-scale issues exist, *e.g.*, the parameter-update before the underlying Markov chain for a given parameter reaches the stationary regime. Our algorithms can be thought as *stochastic approximations* to the standard game dynamics, which is a new feature of our work, not popular in other traditional game-theoretic studies of MAC. We also provide the discussions on how our results can be extended for more practical interference model and the case when there exist collisions.

To connect the study of CSMA to the angles from machine learning or statistical physics, it can be regarded as a problem of finding the parameters of the hard-core graphical model[1] [23] in a distributed manner, leading to the marginal distribution that is the optimal solution of a utility maximization problem. The hard-core model then corresponds to the interference graph of multi-hop wireless network and the parameters are related to operational parameters in CSMA. The proposed dynamics in this paper can be also interpreted as variants of the *contrastive divergence learning* [24] in hard-core graphical models, which is of intellectually independent interest in the area of machine learning and statistical physics.

The rest of this paper is organized as follows: In Section II, we present a large array of related work with their difference from our work. In Section III, we describe the system model and the game-theoretic problem formulation with its objective, followed by the equilibrium analysis in Section IV. In Section V, we provide three optimal distributed game-inspired learning dynamics, and demonstrate their performance through numerical results in Section VI. Finally, we provide the discussions on extensions for more practical situations in Section VII and conclude in Section VIII. Appendix includes the detail of the mathematical proofs.

[1] This corresponds to a graphical model that neighboring nodes cannot be active simultaneously.

II. Related Work

There exists an extensive array of researches (i) on the design and analysis of random access based MAC protocols in wireless networks from the game-theoretic perspective, and (ii) on the fully-distributed CSMA based MAC protocols that achieve optimality in throughput and/or utility from the optimization perspective. We refer the readers to survey papers, *e.g.*, [18], [25] for an exhaustive list. We summarize a part of them, which is closely related to this paper.

D1. *Single-hop random access MAC games:* There exist game-theoretic studies on non carrier sense multiple access such as ALOHA or Slotted-ALOHA with selfish users [26]–[31]. To summarize some of those papers, the authors in [26] considered a multi packet reception model for selfish users and analyzed a Nash equilibrium and its stability region with the assumption of perfect information. The case with partial information has been studied in [27]. In [28], non-cooperative two-player ALOHA game was shown to have two different Nash equilibria, where only one was locally asymptotically stable. The authors in [30], [31] studied the impact of channel-state information. As another class of random access MAC protocol, CSMA has also been studied from the game-theoretic view, where the strategy of a game is usually contention window, transmission power, or data rate, see [25] for a survey and references therein. In [32], it has been studied how selfish users can cheat those who obey the standard CSMA/CA. The authors in [33] abstract 802.11 DCF by focusing on 802.11's average behavior and connecting its window-based access and backoff to transmission probability. Then, the stability of 802.11 has been studied when heterogeneous selfish users exist, where each user dynamically changes its contention window size based on its disutility in terms of contention degree. All these papers considered a single-hop wireless network, *e.g.*, WLAN (Wireless LAN).

D2. *Multi-hop random access MAC games:* The authors in [34] reverse-engineered exponential backoff based contention resolution mechanism in ad-hoc networks which can be modeled by a non-cooperative game with a player's strategy being channel access probability (*i.e.*, ALOHA-like MAC). They also showed that the resulting Nash equilibrium (NE) is not generally socially optimal. This motivates the work [35] which forward-engineers utility-optimal contention resolution algorithms using a standard optimization decomposition approach. The authors in [36], [37] take a similar medium access model based on access probability and study how cost function in each player's payoff function should be designed to achieve a good NE and propose a dynamic access probability update rule converging to the NE of the network with multiple contention measure signals. The authors in [38], [39] proposed distributed ALOHA-type MAC algorithms which have provable convergence, optimality, and robustness under a wider range of utility functions with single message passing for each node in [38] for general topologies, and without message passing for fully interfered topologies in [39].

D3. *Optimal CSMA:* As mentioned in Section I, CSMA has recently been studied from an optimization based framework to achieve optimality in throughput and/or fairness, *e.g.*, see [14]–[17], [40], [41] and [18] for a survey. The main intuition

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2017.2764081, IEEE Transactions on Wireless Communications

underlying these results is that links dynamically adjust their CSMA parameters, *backoff* and *channel holding* times, using local information such as queue-length so that they solve a certain network-wide optimization problem. This research has been regarded as an exciting progress to achieve both simplicity and optimality in the area of wireless cross-layer design. Toward practical scenarios, more practical situations, *e.g.*, CSMA under the SINR interference model [42], or CSMA networks with collisions [43], [44], have been considered for optimal CSMA.

Major difference from prior work. Out work enhances the research in ***D1*** and ***D2*** in the sense that (i) we propose a non-cooperative game whose NE is utility optimal over the maximum throughput region (*i.e.*, throughput region achieved by Max-Weight), whereas the studies in ***D1, D2*** are utility-optimal over a throughput region from ALOHA or Slotted-ALOHA (a much smaller region than the maximum throughput region), and (ii) the dynamic update rules in ***D1, D2*** require message passing (*i.e.*, partially-distributed algorithms). CSMA's utility-optimality has been studied by the researches in ***D3*** mainly from an optimization perspective, but our work starts from a non-cooperative game, followed by the resulting NE's efficiency (*i.e.*, asymptotically no Price-of-Anarchy), and proposes three optimal learning dynamics not revealed by the work in ***D3***. In particular, applying game theory to CSMA-based wireless networks presents powerful insights of developing new *fully*-distributed CSMA algorithms.

III. Model and Problem Description

A. System Model

Network, interference, and traffic. In a wireless network, links share the common wireless medium where they may interfere in their transmissions. As a popular model that captures a certain degree of interference relationships among wireless links in a static and binary ones, called *protocol model*, a network topology can be represented as an undirected graph $G = (V, E)$, called *interference graph,* where n links correspond to nodes V, *i.e.*, $|V| = n$, and undirected edges E are generated among interfering links. In other words, we assume that the interference is symmetric, captured by undirected edges in the graph, *i.e.*, $(i, j) \in E$ if and only if the transmission over link i cannot be successful if a transmission over link j occurs simultaneously. However, the results presented here also can be readily extended to more practical situations where the transmission success is decided by the aggregated interference level among links, called signal-to-interference-plus-noise ratio (SINR) model, see Section VII for this extension. The network is assumed to handle single-hop link-level traffic, and be saturated, *i.e.*, each link has infinite backlog to transmit. However, the results presented here can be readily extended to *multi-hop* flows if a classical combination of back-pressure routing and source congestion control [13] are inserted to.

Schedule and throughput region. We consider a continuous time framework, where our primary interest is to track which links transmit over time. Let $\sigma_i(\tau) \in \{0, 1\}$ denote whether link i is transmitting at time τ or not, where $\sigma_i(\tau) = 1$ means that the transmission at link i is active at time τ and $\sigma_i(\tau) = 0$ otherwise. We also denote by $\boldsymbol{\sigma}(\tau) = [\sigma_i(\tau)]_{i \in V}$ a *schedule* for links at time τ. A scheduling algorithm is regarded as a policy that chooses a sequence of schedules $\{\boldsymbol{\sigma}(\tau)\}_{\tau=0}^{\infty}$ over time. Since interfering links cannot successfully transmit packets simultaneously, a schedule $\boldsymbol{\sigma}(\tau)$ is said to be *feasible* (*i.e.*, no collisions) unless there exists $(i, j) \in E$ such that both $\sigma_i(\tau)$ and $\sigma_j(\tau)$ are 1. Then, the set of all feasible schedules $\mathcal{I}(G)$ is defined as:

$$\mathcal{I}(G) \triangleq \{\boldsymbol{\sigma} \in \{0, 1\}^n : \sigma_i + \sigma_j \leq 1, \forall (i, j) \in E\}.$$

We now define the maximum throughput region (or simply throughput region) Λ of a given network, which is the convex hull of $\mathcal{I}(G)$, *i.e.*,

$$\Lambda = \Lambda(G) \triangleq \Bigg\{ \sum_{\boldsymbol{\rho} \in \mathcal{I}(G)} \alpha_{\boldsymbol{\rho}} \boldsymbol{\rho} : \sum_{\boldsymbol{\rho} \in \mathcal{I}(G)} \alpha_{\boldsymbol{\rho}} = 1, \quad \text{where } \alpha_{\boldsymbol{\rho}} \geq 0 \text{ for all } \boldsymbol{\rho} \in \mathcal{I}(G) \Bigg\}.$$

The intuition behind this notion of throughput region comes from the fact that any scheduling algorithm has to choose a schedule from $\mathcal{I}(G)$ at each time, where $\alpha_{\boldsymbol{\rho}}$ denotes the fraction of time selecting schedule $\boldsymbol{\rho}$. Hence, the time average of the 'service rate' in each link induced by any scheduling algorithm must belong to Λ.

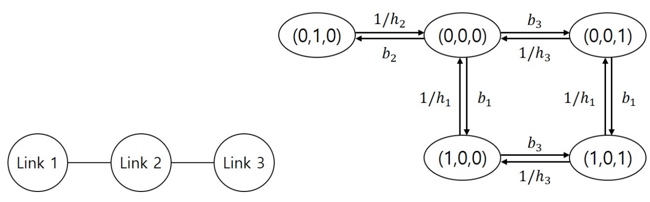
CSMA (Carrier Sense Multiple Access) and its Markov chain. As mentioned in Section I, our interest lies in a simple, fully-distributed CSMA scheduling algorithm to avoid interferences efficiently in wireless networks. Under a CSMA algorithm, prior to trying to transmit a packet, links first check whether the medium is busy or idle, and then transmit a packet only when the medium is sensed idle, *i.e.*, no interfering link is transmitting. To control the aggressiveness of such medium access, each link maintains a backoff timer, which is reset to a random value when it expires. The timer runs only when the medium is idle, and stops otherwise. With the backoff timer, links try to avoid collisions by the following procedure:

- Each link does not start transmission immediately when the medium is sensed idle, but keeps silent until its *backoff time* expires.
- After a link grabs the channel (or medium), the link holds the channel for a random amount of time, called the *holding time*.

For simplicity, we assume backoff and holding times of link i follow exponential distributions with means $1/b_i$ and h_i, respectively, for some positive real numbers b_i and h_i. Then, the probability for two interfering links to start transmission simultaneously is zero (*i.e.*, no collisions)[2], and a sequence of schedules $\{\boldsymbol{\sigma}(\tau)\}_{\tau=0}^{\infty}$ of CSMA constructs a continuous-time reversible Markov chain[3] [12], [16], [45], [46], so-called *CSMA Markov chain*. To illustrate, consider a simple three-link interference graph and its resulting CSMA Markov chain in Fig. 1.

[2]This CSMA algorithm is called *idealized CSMA*, meaning that the sensing is instantaneous so that collisions do not occur, which comes from the assumption of continuous distributions of backoff and holding times.

[3]We refer the readers to previous works [12], [16], [45], [46] to follow the theoretical details of CSMA Markov chain, *e.g.*, reversibility from detailed balance equations, ergodicity due to finite state space, and a form of stationary distribution (1).



(a) (b)

Fig. 1. (a) 3-link interference graph, where links 1 and 3 do not interfere with each other, but interfere only with link 2. (b) The resulting CSMA Markov chain for fixed parameters (b_1, b_2, b_3) and (h_1, h_2, h_3), where for example the state $(1, 0, 1)$ means that links 1 and 3 are active and link 2 is inactive.

For fixed CSMA parameters $\boldsymbol{b} = [b_i]_{i\in V}$ and $\boldsymbol{h} = [h_i]_{i\in V}$, using the reversibility of Markov process $\{\boldsymbol{\sigma}(\tau)\}_{\tau=0}^{\infty}$, its stationary distribution $\pi^{\boldsymbol{b},\boldsymbol{h}} = [\pi^{\boldsymbol{b},\boldsymbol{h}}_{\boldsymbol{\sigma}}]_{\boldsymbol{\sigma}\in\mathcal{I}(G)}$ can be characterized as follows:

$$\pi^{\boldsymbol{b},\boldsymbol{h}}_{\boldsymbol{\sigma}} = \frac{\prod_{i\in V}(b_i h_i)^{\sigma_i}}{\sum_{\boldsymbol{\sigma'}\in\mathcal{I}(G)}\prod_{i\in V}(b_i h_i)^{\sigma'_i}}. \quad (1)$$

For a simple presentation, we let $r_i = \log(b_i h_i)$ and call r_i the *transmission intensity* (or simply *intensity*) of link i, intuitively meaning the transmission aggressiveness of the link. Hence, for a given $\boldsymbol{r} = [r_i]_{i\in V}$, we also use $\pi^{\boldsymbol{r}}$ instead of $\pi^{\boldsymbol{b},\boldsymbol{h}}$, *i.e.*,

$$\pi^{\boldsymbol{r}}_{\boldsymbol{\sigma}} = \frac{\exp(\sum_{i\in V}\sigma_i r_i)}{\sum_{\boldsymbol{\sigma'}\in\mathcal{I}(G)}\exp(\sum_{i\in V}\sigma'_i r_i)}.$$

It now follows from the reversibility and ergodicity of CSMA Markov chain that the marginal probability $s_i(\boldsymbol{r})$ that link i is scheduled under the stationary distribution $\pi^{\boldsymbol{r}}$ becomes the link i's long-term (average) throughput or service rate, *i.e.*, $\lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau \sigma_i(s)\,\mathrm{d}s$, which is given by:

$$\begin{aligned} s_i(\boldsymbol{r}) = \mathbb{E}_{\pi^{\boldsymbol{r}}}[\sigma_i] &= \sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1}\pi^{\boldsymbol{r}}_{\boldsymbol{\sigma}} \\ &= \frac{\sum_{\boldsymbol{\sigma}\in\mathcal{I}(G):\sigma_i=1}\exp(\sum_{i\in V}\sigma_i r_i)}{\sum_{\boldsymbol{\sigma'}\in\mathcal{I}(G)}\exp(\sum_{i\in V}\sigma'_i r_i)}. \quad (2) \end{aligned}$$

This CSMA Markov chain models the behavior of links in a wireless LAN, thus it can be used to assess the performance of large-scale wireless LAN with appropriate selection of intensity.

B. Problem Description: Utility Maximization

We aim at developing a CSMA algorithm that controls the intensity of each link so as to make the long-term service rate close to some fairness point of the boundary of Λ. Specifically, each link i finds its intensity r_i (*i.e.*, CSMA's backoff and holding time parameters b_i and h_i) in a distributed manner, by adaptively changing r_i over time, so that the resulting long-term service rates over links form a solution of the following utility maximization problem:

$$\textbf{(OPT)} \qquad \max_{\boldsymbol{\lambda}\in\Lambda}\ \sum_{i\in V} U_i(\lambda_i), \quad (3)$$

where we denote by $\boldsymbol{\lambda}^\star$ the optimal solution. In the above, we assume that each link i has a concave, increasing, and (twice continuously) differentiable utility function, $U_i : [0,1] \to \mathbb{R}$, where its value represents the utility when rate $\lambda_i \in [0,1]$ is allocated at link i. As is well-known, the various forms of utility function enforce different concepts of fairness, *e.g.*, α-fairness [47].

To achieve the desired utility maximization using a CSMA algorithm, the core question is how each link i chooses its intensity r_i so that $s_i(\boldsymbol{r})$ is the solution of (3). To this end, we take a game-theoretic approach, where a smart design of payoff functions for links is necessary to have the desired property such that the solution of (3) is attained at the (Nash) equilibrium of the game. Under such a game design, we will consider various game learning dynamics that provably converge to the equilibrium, where major technical challenges for proving the convergence lie in handling a non-trivial coupling between CSMA Markov chain and its parameter updates.

IV. Game Design and Equilibrium Analysis

A. Optimal CSMA Game: oCSMA(β)

We first design a non-cooperative game, denoted by oCSMA(β), which is parameterized by a constant $\beta > 0$, then explain the rationale behind our game design.

---

oCSMA(β)

(i) ***Players.*** Each link $i \in V$ (*i.e.*, a node in the interference graph G) acts as a player.

(ii) ***Strategy.*** Each player i chooses an intensity $r_i \in (-\infty, \infty)$ as its own strategy, which determines how aggressively i accesses the medium. We use the conventional notation that the strategy vector for all players except i is $r_{-i} := (r_1, r_2, \cdots, r_{i-1}, r_{i+1}, \cdots, r_n)$ and write a strategy profile $\boldsymbol{r} = (r_i, r_{-i})$.

(iii) ***Payoff.*** The payoff function $\Phi_i(\boldsymbol{r})$ of player i is designed to be utility U_i subtracted by an incurring price C_i (scaled by $1/\beta$) as follows:

$$\Phi_i(\boldsymbol{r}) = U_i(s_i(\boldsymbol{r})) - \frac{1}{\beta}C_i(\boldsymbol{r}),$$
$$\text{where} \quad C_i(\boldsymbol{r}) = \int_{-\infty}^{r_i} x s'_i(x, r_{-i})\mathrm{d}x. \quad (4)$$

Note that the payoff function of player i is determined by how aggressively other links access the medium (*i.e.*, r_{-i}) as well as how itself does (*i.e.*, r_i), and it consists of two major terms: (i) individual utility $U_i(s_i(\boldsymbol{r}))$ and (ii) incurring price $C_i(\boldsymbol{r})$. The parameter $\beta > 0$ quantifies 'price level' in the players' payoffs, and we realize that it balances the trade-off between the quality of equilibria in the game (*i.e.*, Price-of-Anarchy) and the convergence speed to the equilibria under the learning dynamics (see Section V). Before to analyze the game oCSMA(β), we provide popular notions of non-cooperative game: (i) Nash equilibrium and (ii) Price-of-Anarchy, whose definitions are presented as follows:

Definition 1. *(i) A strategy profile* $\boldsymbol{r}^{NE}$ *is a* Nash equilibrium *(NE) if*

$$\Phi_i(r_i^{NE}, r_{-i}^{NE}) \geq \Phi_i(r_i, r_{-i}^{NE}), \quad \forall r_i \in \mathbb{R}, \forall i \in V.$$

(ii) A Price-of-Anarchy *(PoA) is*

$$\textit{PoA} = \frac{\max_{\boldsymbol{r}}\sum_{i\in V}U_i(s_i(\boldsymbol{r}))}{\min_{\boldsymbol{r}^{NE}}\sum_{i\in V}U_i(s_i(\boldsymbol{r}^{NE}))}.$$

Intuitively, we see that NE is a strategy profile for which no player has an incentive to deviate unilaterally. Furthermore, we say that a NE $\boldsymbol{r}^{\text{NE}}$ (if exists) is *non-trivial*, if each player i's service rate at equilibrium $s_i(\boldsymbol{r}^{\text{NE}})$ is positive for all players $i \in V$, and *trivial* otherwise. The PoA indicates the ratio between the social optimum and the worst equilibrium of the game, and we say *no PoA* if PoA $= 1$.

B. Rationale of Price Function

To have nice properties of our game, *e.g.*, good equilibria or provable transfer to fully-distributed dynamics (converging to a good equilibrium), the choice of price function is of critical importance. This subsection is devoted to explaining how such nice properties can be obtained from our price function (4) which has two following design features ***P1*** and ***P2***.

P1. Appropriate measure of link's contention: We choose a price function so that it appropriately measures each link's contention impact on other links' throughput. This choice differs from other price function choices, *e.g.*, one in ALOHA systems, which is the key to provably have almost no Price-of-Anarchy (see Section IV-C). Specifically, a simple algebra gives us the following expression of our price function: $C_i(\boldsymbol{r}) = s_i(r_i, r_{-i}) \cdot \int_{-\infty}^{r_i} x \frac{s_i'(x, r_{-i})}{s_i(r_i, r_{-i})} \mathrm{d}x$. Let $R_i \in [-\infty, r_i]$ be a continuous random variable with the density $f_{R_i}(\cdot)$: $f_{R_i}(x) = \frac{1}{s_i(r_i, r_{-i})} \frac{\partial s_i(x, r_{-i})}{\partial x}$. Then, our price function can be regarded as the product of the service rate and R_i's expectation: $C_i(\boldsymbol{r}) = s_i(r_i, r_{-i}) \cdot \mathbb{E}[R_i]$.

We remark that a popular choice of price function is $C_i(\boldsymbol{r}) = r_i \cdot s_i(r_i, r_{-i})$. The intuition behind such a choice is that the price to be paid by a link's access intensity is proportional to the current aggressiveness in media usage r_i multiplied by its achieved gain (*i.e.*, throughput $s_i(\boldsymbol{r})$). This type of choice has already been made in other works, *e.g.*, [37] which studies ALOHA-based MAC. However, it is unclear that this price function provides a provable framework for no Price-of-Anarchy. On the other hand, our design of price function considers the expected aggressiveness $\mathbb{E}[R_i]$ which depends on the *relative increasing speed of the service rate* in the interval $(-\infty, r_i)$. As shown later in Section IV-C, this way of measuring link's contention impact leads to system-wide efficient behaviors from each link's local decision.

P2. Indirect coupling of players' strategies: Our price function is a function of self-strategy and its marginal distribution of the given strategy profile, not the individual strategy values of others. This feature enables us to develop fully-distributed dynamics that work based only on throughput measurements (see Section V). We first re-express the price function to better understand how it is structured in terms of the local strategy and other players' strategies, as follows:

$$C_i(r_i, r_{-i}) = \Big[x s_i(x, r_{-i}) \Big]_{-\infty}^{r_i} - \int_{-\infty}^{r_i} s_i(x, r_{-i}) \mathrm{d}x$$
$$= r_i s_i(\boldsymbol{r}) + \ln\big(1 - s_i(\boldsymbol{r})\big), \tag{5}$$

where for the second term we use: first,

$$s_i(x, r_{-i}) = \frac{\sum_{\boldsymbol{\sigma} \in \mathcal{I}(G):\sigma_i = 1} \exp\big(\sum_{j \in V} r_j \sigma_j\big)}{\sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \exp\big(\sum_{j \in V} r_j \sigma_j\big)}$$
$$= \frac{\sum_{\boldsymbol{\sigma} \in \mathcal{I}(G):\sigma_i = 1} \exp\big(\sum_{j \in V \setminus \{i\}} r_j \sigma_j\big) \exp(x)}{\sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \exp\big(\sum_{j \in V \setminus \{i\}} r_j \sigma_j\big) \exp(x \sigma_i)}$$
$$= \frac{B \exp(x)}{A + B \exp(x)},$$

where $A \equiv \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G) | \sigma_i = 0} \exp(\sum_{j \in V \setminus \{i\}} r_j \sigma_j)$ and $B \equiv \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G) | \sigma_i = 1} \exp(\sum_{j \in V \setminus \{i\}} r_j \sigma_j)$, then

$$\int_{-\infty}^{r_i} s_i(x, r_{-i}) \mathrm{d}x = \int_{-\infty}^{r_i} \frac{B \exp(x)}{A + B \exp(x)} \mathrm{d}x$$
$$= \ln \frac{A + B \exp(r_i)}{A} = -\ln\Big(1 - s_i(\boldsymbol{r})\Big).$$

From (5), the payoff function becomes:

$$\Phi_i(\boldsymbol{r}) \quad = \quad U_i(s_i(\boldsymbol{r})) - \frac{1}{\beta}\Big(r_i s_i(\boldsymbol{r}) + \ln\big(1 - s_i(\boldsymbol{r})\big)\Big).$$

It is important to see that the payoff function depends only on the local strategy r_i and local service rate $s_i(r_i, r_{-i})$, not directly on strategies or service rates of other players. This *indirect coupling*, which is a unique feature in our game, is highly convenient to develop a fully-distributed dynamics, because $s_i(\cdot)$ can be measured in the midst of playing a player's own strategy via CSMA without message passing, even though the exact computation of local service rate requires heavy computation.

C. Equilibrium Analysis: Existence, Uniqueness and Price-of-Anarchy

In this subsection, we provide the equilibrium analysis of our game. Three main questions of our interests are: existence, uniqueness, and Price-of-Anarchy of the equilibrium. We now present our main results on the equilibrium analysis in the following theorem.

Theorem 1 (Uniqueness and no PoA)**.** *In the oCSMA(β), for any $\beta > 0$,*

(i) Existence and uniqueness. *There exists a unique non-trivial NE $\boldsymbol{r}^{NE}$.*

(ii) Price-of-Anarchy. *Furthermore, at the non-trivial NE $\boldsymbol{r}^{NE}$,*

$$\sum_{i \in V} U_i\Big(s_i(\boldsymbol{r}^{NE})\Big) \geq \sum_{i \in V} U_i\Big(s_i(\boldsymbol{r}^{\star})\Big) - \frac{\log |\mathcal{I}(G)|}{\beta}, \tag{6}$$

where $\boldsymbol{r}^\star$ represents a strategy profile such that the service rate vector $[s_i(\boldsymbol{r}^\star)]_{i \in V}$ is the solution of the optimization problem **OPT** *in (3),* i.e., $[s_i(\boldsymbol{r}^\star)]_{i \in V} = \boldsymbol{\lambda}^\star$.

Theorem 1, whose proof is presented in Section IV-D, implies that there is almost no PoA (Price-of-Anarchy) in our game, *i.e.*, the aggregate utility at the unique non-trivial NE can be arbitrarily close to the social optimum by choosing β sufficiently large. Namely, PoA can become arbitrarily small: $\lim_{\beta \to \infty} \text{PoA} = 1$.

D. Proof of Theorem 1

Proof. *(i) Existence and uniqueness:* We first prove the existence and uniqueness of non-trivial NE using a potential game approach. Consider the following function $P(\boldsymbol{r})$ on the space $\mathcal{R}^+ = \{\boldsymbol{r} | \boldsymbol{s}(\boldsymbol{r}) > 0\}$ (the set of strategies producing "non-trivial" service rates), defined by:

$$P(\boldsymbol{r}) \quad \triangleq \quad - \sup_{\boldsymbol{\lambda} \in [0,1]^n, \boldsymbol{\mu} \in \mathcal{P}} L(\boldsymbol{\lambda}, \boldsymbol{\mu}; \frac{\boldsymbol{r}}{\beta}),$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2017.2764081, IEEE Transactions on Wireless Communications

where $\mathcal{P}$ is the set of all probability measures over the set of all feasible schedules $\mathcal{I}(G)$, and

$$L(\boldsymbol{\lambda}, \boldsymbol{\mu}; \frac{\boldsymbol{r}}{\beta}) \triangleq \sum_{i \in V} U_i(\lambda_i) - \frac{1}{\beta} \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \mu_{\boldsymbol{\sigma}} \log \mu_{\boldsymbol{\sigma}} + \sum_{i \in V} \frac{r_i}{\beta} \left(\sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \mu_{\boldsymbol{\sigma}} \sigma_i - \lambda_i \right).$$

It is easy to check that $P(\boldsymbol{r})$ is strictly concave in $\boldsymbol{r}$, since $P(\cdot)$ is the infimum of $-L(\cdot)$ which is a family of affine functions in $\boldsymbol{r}$. We now show that oCSMA(β) is an *ordinal potential game* [48] with the potential function $P(\boldsymbol{r})$, *i.e.*, $\mathrm{sgn}\frac{\partial \Phi_i(\boldsymbol{r})}{\partial r_i} = \mathrm{sgn}\frac{\partial P(\boldsymbol{r})}{\partial r_i}$, for all $i \in V$. We first have:

$$\begin{aligned} \frac{\partial \Phi_i(\boldsymbol{r})}{\partial r_i} &= \frac{\partial}{\partial r_i}\left(U_i(s_i(\boldsymbol{r})) - \frac{1}{\beta}\int_{-\infty}^{r_i} x s_i'(x, r_{-i}) \mathrm{d}x \right) \\ &= \frac{\partial s_i(\boldsymbol{r})}{\partial r_i}\left(U_i'(s_i(\boldsymbol{r})) - \frac{r_i}{\beta} \right) \\ &= s_i(\boldsymbol{r})\Big(1 - s_i(\boldsymbol{r})\Big)\left(U_i'(s_i(\boldsymbol{r})) - \frac{r_i}{\beta} \right), \end{aligned} \tag{7}$$

where the last equality comes from a simple algebra:

$$\frac{\partial s_i(r_i, r_{-i})}{\partial r_i} = s_i(r_i, r_{-i})\Big(1 - s_i(r_i, r_{-i})\Big), \tag{8}$$

and second we have:

$$\frac{\partial P(\boldsymbol{r})}{\partial r_i} = \frac{1}{\beta}\Big(U_i'^{-1}(r_i/\beta) - s_i(\boldsymbol{r}) \Big). \tag{9}$$

Thus on the space $\{\boldsymbol{r} | \boldsymbol{s}(\boldsymbol{r}) > 0\}$, $\mathrm{sgn}\frac{\partial \Phi_i(\boldsymbol{r})}{\partial r_i} = \mathrm{sgn}\frac{\partial P(\boldsymbol{r})}{\partial r_i}$. From the standard results in potential games and strict concavity of $P(\cdot)$, the solution that maximizes $P(\cdot)$ is a NE $\boldsymbol{r}^{\mathrm{NE}}$, where each player's strategy is a *best response* to the others' strategies at NE, and is non-trivial and unique.

(ii) Price-of-Anarchy: Consider an approximated problem **A-OPT** of **OPT**, given by:

(A-OPT)

$$\begin{aligned} \max_{\boldsymbol{\lambda} \in [0,1]^n, \boldsymbol{\mu} \in \mathcal{P}} \quad & \sum_{i \in V} U_i(\lambda_i) - \frac{1}{\beta} \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \mu_{\boldsymbol{\sigma}} \log \mu_{\boldsymbol{\sigma}} \\ \text{subject to} \quad & \lambda_i \le \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \mu_{\boldsymbol{\sigma}} \sigma_i, \quad \forall i \in V. \end{aligned} \tag{10}$$

Since the objective function is concave and the entropy follows $-\sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \mu_{\boldsymbol{\sigma}} \log \mu_{\boldsymbol{\sigma}} \le \log |\mathcal{I}(G)|$, **A-OPT** problem has a unique solution $(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)$ and the solution $\boldsymbol{\lambda}^o$ satisfies that

$$\sum_{i \in V} U_i(\lambda_i^o) \ge \max_{\boldsymbol{\lambda} \in \Lambda} \sum_{i \in V} U_i(\lambda_i) - \frac{\log |\mathcal{I}(G)|}{\beta}. \tag{11}$$

We now consider the Lagrangian $\mathcal{L}$ of **A-OPT** with dual variables $\boldsymbol{k} = [k_i]_{i \in V}$:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\mu}; \boldsymbol{k}) &= \sum_{i \in V} U_i(\lambda_i) - \frac{1}{\beta} \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \mu_{\boldsymbol{\sigma}} \log \mu_{\boldsymbol{\sigma}} \\ &\quad + \sum_{i \in V} k_i (\sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \mu_{\boldsymbol{\sigma}} \sigma_i - \lambda_i) \\ &= \frac{1}{\beta}\left(\sum_{i \in V} \beta k_i \cdot \mathbb{E}_{\boldsymbol{\mu}}[\sigma_i] - \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \mu_{\boldsymbol{\sigma}} \log \mu_{\boldsymbol{\sigma}} \right) \\ &\quad + \sum_{i \in V} (U_i(\lambda_i) - k_i \lambda_i), \end{aligned}$$

where $\mathbb{E}_{\boldsymbol{\mu}}[\cdot]$ denotes the expectation for distribution $\boldsymbol{\mu}$. The solution of **A-OPT** is the minimum point of the dual function, which is given by

$$D(\boldsymbol{k}) = \sup_{\boldsymbol{\lambda} \in [0,1]^n, \boldsymbol{\mu}} \mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\mu}; \boldsymbol{k}),$$

where $\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{\mu}; \cdot)$ is maximized at $(\boldsymbol{\lambda}^o, \boldsymbol{\mu}^o)$ when $\mu_{\boldsymbol{\sigma}}^o = \pi_{\boldsymbol{\sigma}}^{\boldsymbol{r}}$ with $\boldsymbol{r} = \beta \boldsymbol{k}$, and $\lambda_i^o = U_i'^{-1}(k_i)$. Let $\boldsymbol{r}^o$ be the strategy such that $\boldsymbol{r}^o = \beta \boldsymbol{k}^o$, where $\boldsymbol{k}^o$ is the solution of $\min D(\boldsymbol{k})$, thus

$$s_i(\boldsymbol{r}^o) = U_i'^{-1}(r_i^o/\beta) = \lambda_i^o \quad \text{for all } i \in V. \tag{12}$$

Note that the strategy vector $\boldsymbol{r}^{\mathrm{NE}}$ maximizing the potential function $P(\boldsymbol{r})$ satisfies that $s_i(\boldsymbol{r}^{\mathrm{NE}}) = U_i'^{-1}(r_i^{\mathrm{NE}}/\beta)$, from (9). This completes the proof, because $\boldsymbol{r}^{\mathrm{NE}}$ coincides with the strategy vector $\boldsymbol{r}^o = \beta \boldsymbol{k}^o$ that minimizes $D(\boldsymbol{k})$, and thus (6) holds from (11) and (12). □

V. Game-inspired Dynamics: Distributedness, Convergence, and Optimality

In Section IV, we have established desirable equilibrium properties in oCSMA(β), such as uniqueness and no Price-of-Anarchy (thus asymptotic utility optimality). In this section, we consider *game (learning) dynamics* to study how players' strategy (*i.e.*, intensity in oCSMA(β)) evolves over time and converges (if it does). We aim at developing dynamics that *(i)* operate in a fully-distributed manner, and *(ii)* converge to the unique non-trivial equilibrium in Section IV. By "fully-distributed", we mean that links update their strategies without any message passing among them, relying only on pure local information and observations. In other words, the strategy update of a link does not require the payoffs and the strategies of other links. To that end, we first consider three popular dynamics in non-cooperative game theory (best response dynamics, Jacobi dynamics, and gradient dynamics), and discuss the major challenges in terms of fully distributed operations, when applying such standard dynamics to oCSMA(β). Then, we develop new fully-distributed dynamics, investigate their convergence and optimality on the strength of stochastic approximation theory, and finally discuss the rationale behind the dynamics compared to the existing *optimal CSMA* algorithms.

A. Preliminaries: Three Standard Game Dynamics

Best response dynamics. The most popular dynamics in non-cooperative games is the *best response (BR) dynamics*. In the BR dynamics, each player i chooses its best strategy, given the strategy vector (at the previous frame) for all other players except i, which gives a maximum payoff to the player, *i.e.*, at t-th frame

$$r_i(t+1) = \mathrm{BR}_i(r_{-i}(t)) := \arg\max_{r_i \in \mathbb{R}} \Phi_i\Big(r_i, r_{-i}(t)\Big),$$

which leads to a fixed point of following function in oCSMA(β): for all $i \in V$,

$$r_i(t+1) = \beta U_i'\Big(s_i\Big(r_i(t+1), r_{-i}(t) \Big) \Big). \tag{13}$$

Jacobi dynamics. The second dynamics is *Jacobi dynamics*. Its basic idea is to adjust each player's strategy gradually towards its best response strategy, *i.e.*, at t-th frame

$$r_i(t+1) = r_i(t) + \alpha\Big(\mathrm{BR}_i(r_{-i}(t)) - r_i(t)\Big), \qquad (14)$$

where $\alpha \in (0,1]$ is a smoothing parameter[4]. The smoothing parameter α captures how aggressively the dynamics follows the BR dynamics, where $\alpha = 1$ corresponds to the BR dynamics.

Gradient dynamics. Finally, we investigate the *gradient dynamics* [50], which can be viewed as a 'better reply' dynamics[5]. At each frame, each player i first determines the gradient of its payoff (4), $\nabla \Phi_i(\boldsymbol{r})$, then updates its strategy in the gradient direction, *i.e.*, at t-th frame,

$$r_i(t+1) = r_i(t) + \alpha \cdot \frac{\partial \Phi_i(r_i, r_{-i}(t))}{\partial r_i},$$

where $\alpha > 0$ is the step-size. In oCSMA(β), the gradient dynamics becomes: for all $i \in V$,

$$\begin{aligned} & r_i(t+1) \\ & = r_i(t) + \alpha \cdot \frac{\partial s_i(r_i, r_{-i}(t))}{\partial r_i}\left(U_i'\left(s_i(\boldsymbol{r}(t))\right) - \frac{r_i(t)}{\beta}\right), \end{aligned} \qquad (15)$$

since the gradient of player i from (4) is given by:

$$\begin{aligned} \nabla \Phi_i(\boldsymbol{r}) = \frac{\partial \Phi_i(\boldsymbol{r})}{\partial r_i} &= \frac{\partial U_i(s_i(\boldsymbol{r}))}{\partial r_i} - \frac{1}{\beta}\frac{\partial C_i(\boldsymbol{r})}{\partial r_i} \\ &= \frac{\partial s_i(\boldsymbol{r})}{\partial r_i}\left(U_i'\Big(s_i(\boldsymbol{r})\Big) - \frac{r_i}{\beta}\right). \end{aligned}$$

The interpretation of the gradient dynamics from an economic perspective is that if the marginal utility exceeds the marginal price, *i.e.*, $\nabla \Phi_i(\boldsymbol{r}) > 0$, link i's intensity is increased to achieve more utility, and if the marginal price exceeds the marginal utility, *i.e.*, $\nabla \Phi_i(\boldsymbol{r}) < 0$, link i's intensity is decreased to reduce the transmission price.

In developing the learning dynamics for oCSMA(β), two primary goals are *fully-distributed* operation and *convergence* to the unique NE (which is also socially optimal). To that end, an immediate and obvious attempt is to apply the aforementioned three standard dynamics to our case. However, we have the following two challenges in achieving our goals.

(i) Hardness of convergence to NE under fully-distributed dynamics. It is known that that for a broad class of games, there exists no generalized uncoupled dynamics which operates even in a "partially"-distributed manner (*i.e.*, operating based on the observation of other players' payoff, and thus with message passing), converging to NE [21]. However, in a special class of games, *e.g.*, finite ordinal potential games, it has been shown that many adaptive learning dynamics are guaranteed to converge to a pure NE [48], *e.g.*, best response dynamics, better reply dynamics, fictitious play, and regret matching. Therefore, by applying three dynamics in Section V-A to oCSMA(β), convergence to the unique NE is guaranteed, since oCSMA(β) is an ordinal potential game. As described in Section V-A, however, in each of three standard dynamics (13), (14), and (15), player i requires message passing with other players to know their current strategies to determine the next strategy, in particular for computing the service rate $s_i(\boldsymbol{r})$ or its derivative $\nabla s_i(\boldsymbol{r})$, even though they do not directly need r_{-i} or $s_{-i}(\boldsymbol{r})$. We overcome this challenge by smartly exploiting the *locally-observed service rate* rather than computing the exact marginal distribution at every frame.

(ii) Long convergence time for classical dynamics. In updating strategies, locally-observed service rate is not the actual marginal distribution, because after a strategy is played, it takes long time to reach the stationary regime. In other words, the observed service rates may be far from the 'stationary' service rates. This time-scale issue incurs additional challenges of extremely long convergence times, because a certain amount of time (formally called mixing time) to reach the stationary regime is required for each strategy update, and for convergence, long strategy update iterations are necessary. This challenge prevents us from ensuring the convergence to NE under the three dynamics in Section V-A when exploiting locally-observed service rates. We tackle this challenge by adopting special learning dynamics, called *stochastically-approximated dynamics* that utilize the *time-aggregated service rates* in the strategy updates.

B. Fully-Distributed Stochastically-Approximated Dynamics

We now propose three learning dynamics, all of which provably converge to the unique non-trivial NE and operate in a fully-distributed manner. They are theoretically supported by the theory of stochastic approximation, and therefore, we call them *(i)* ***SA-BRD*** (SA-Best Response Dynamics), *(ii)* ***SA-JD*** (SA-Jacobi Dynamics), and *(iii)* ***SA-GD*** (SA-Gradient Dynamics).

Algorithm description. In all three dynamics, time is divided into discrete frames $t = 0, 1, \ldots,$ where the frame duration is fixed by, say the time to transmit a MAC packet of a fixed size. We first let $\hat{s}_i(t)$ and $\bar{s}_i(t)$ be the *instantaneous* and *aggregate* service rate of player i at and until frame t, respectively. $\hat{s}_i(t)$ denotes the number of transmitted packets at link i over frame t, and

$$\bar{s}_i(t) \quad = \quad \frac{1}{t}\sum_{n=0}^{t-1} \hat{s}_i(n). \qquad (16)$$

We now describe three algorithms.

At t-th frame, each link $i \in V$ updates $r_i(t)$ as follows:

(i) SA-BRD (Stochastically-Approximated Best Response Dynamics)[6]

$$r_i(t+1) = \Big[\beta U_i'\Big(\bar{s}_i(t)\Big)\Big]_{r^{\min}}^{r^{\max}}, \qquad (17)$$

(ii) SA-JD (Stochastically-Approximated Jacobi Dynamics)

$$r_i(t+1) = \left[r_i(t) + \alpha\left(\beta U_i'\Big(\bar{s}_i(t)\Big) - r_i(t)\right)\right]_{r^{\min}}^{r^{\max}}, \qquad (18)$$

where $\alpha \in (0,1]$ is a constant.

[4]Jacobi dynamics generally makes a smoother move than the best response dynamics, where a small smoothing parameter plays the role of compensating for the instability of the best response dynamics, see [49].

[5]Sometimes, it is called better response dynamics.

[6]$[\cdot]_a^b \equiv \max(b, \min(a, \cdot))$.

(iii) SA-GD (Stochastically-Approximated Gradient Dynamics)

$$r_i(t+1) = \left[r_i(t) + \alpha \frac{\partial s_i(\boldsymbol{r}(t))}{\partial r_i(t)} \left(U_i'\Big(\bar{s}_i(t)\Big) - \frac{r_i(t)}{\beta} \right) \right]_{r^{\min}}^{r^{\max}}, \quad (19)$$

where $\alpha \in (0,1]$ is a constant.

The above three dynamics are seemingly simple variants of the classical dynamics in Section V-A. Recall that computing the service rate directly produces message passing among players, and measuring the 'stationary' service rate $s_i(r_i(t), r_{-i}(t-1))$ under a CSMA algorithm might incur the long convergence issue (*i.e.*, it takes the mixing time of the underlying CSMA Markov chain). Note that $r_i(t), \bar{s}_i(t), U_i(\cdot), \nabla s_i(\boldsymbol{r}(t))$ are locally obtained due to the feature of CSMA in (8). The key idea of our new fully-distributed dynamics to overcome such challenges is to use the locally maintained aggregate service rate $\bar{s}_i(t)$ instead of $s_i(\boldsymbol{r}(t))$ when computing the best response or gradient direction.

Convergence and optimality. For provable convergence analysis, we make the following assumption **(A1)**, which means that we choose $r^{\min}$ and $r^{\max}$ such that the interval $[r^{\min}, r^{\max}]$ is large enough to include the optimal solution of **A-OPT**. The explicit values of $r^{\min}$ and $r^{\max}$ can be also computable [41].

(A1) If $\boldsymbol{r}^0 \in \mathbb{R}^n$ solves for all $i \in V$, $r_i^0 = \beta U_i'\Big(\mathbb{E}_{p_{\boldsymbol{r}^0}}[\sigma_i]\Big)$, then $r^{\min} \leq r_i^0 \leq r^{\max}$ for all $i \in V$. Note that, for example, if the utility function $U_i(\cdot)$ is such that $U_i'(0) < \infty$, then **(A1)** for r_i^0 is satisfied when $r^{\min} \leq \beta U_i'(1)$ and $r^{\max} \geq \beta U_i'(0)$. One can easily verify that **(A1)** provides a guarantee that the convergent point of our three dynamics actually belongs to a bounded region.

We now present the main theorem, which states that all of three dynamics converge to the unique non-trivial NE, which is in turn asymptotically equals to the socially optimum, as described in Section IV-C.

Theorem 2 (Convergence and Optimality)**.** *Under* **(A1)***, for arbitrary initial condition* $\boldsymbol{r}(0)$*,* ***SA-BRD****,* ***SA-JD****, and* ***SA-GD*** *converge to the unique non-trivial NE* $\boldsymbol{r}^{NE}$*, where* $\boldsymbol{r}^{NE}$ *is defined in Theorem 1, in the sense that*

$$\lim_{t\to\infty} \boldsymbol{r}(t) = \boldsymbol{r}^{NE}, \quad \textit{component-wise, almost surely.}$$

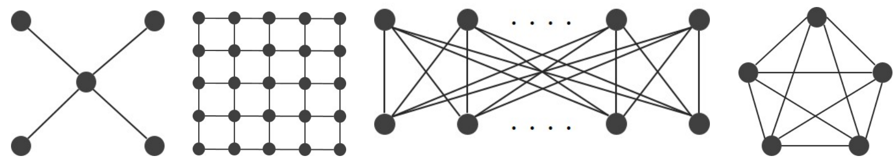
The proof of Theorem 2 is presented in Appendix, but we briefly provide the proof sketch for readers' convenience. Each of ***SA-BRD***, ***SA-JD*** and ***SA-GD*** is interpreted as a stochastic approximation procedure with the controlled Markov noise, and the main technical challenge lies in handling a non-trivial coupling between the underlying CSMA Markov chain and CSMA parameter $\boldsymbol{r}$ updates. The use of aggregate service rates $\bar{\boldsymbol{s}}(\cdot)$ provides a provable convergence, in the way that we intuitively expect that by averaging empirical service rates which has an effect of $1/t$ step-size (see the relation (16)), the speed of variations of the intensity $\boldsymbol{r}$ tends to zero after sufficiently long time, and its limiting behavior is understood by ordinary differential equations (ODE), see Appendix for the mathematical detail. The additional technical challenge dealing with ***SA-JD, SA-GD*** (not existing for ***SA-BRD***) is that they have higher-order temporal dependencies in their updating rules, *i.e.*, use the current strategy $r_i(t)$ for obtaining the next strategy $r_i(t+1)$. To handle the issue, we define a 'virtual' process (see $\nu_i(t)$ and $\upsilon_i(t)$ in Appendix) and argue its convergence under the relation to that of the original process $\{r_i(t)\}_{t \in \mathbb{Z}_{\geq 0}}$.

Comparison with existing algorithms. The dynamics update rules in optimal CSMA have been studied in other papers, *e.g.*, [41], [15], and [40]. We briefly summarize their underlying rationale, and compare them with three algorithms in this paper. In [41] and [15], the authors develop utility optimal CSMA algorithms, which we refer to as ***JW*** and ***EJW***, respectively, and they show that the (asymptotically) optimal solution of **A-OPT** can be achieved by the following algorithm:

$$r_i(t+1) = r_i(t) + a_i(t) \left(U_i'^{-1}\left(\frac{r_i(t)}{\beta} \right) - \hat{s}_i(\boldsymbol{r}(t)) \right), \quad (20)$$

where $a_i(t)$ is a positive step-size function of link i at frame t. The key idea of this algorithm is that the transmission intensity is updated by quantifying the supply-demand differential, and the new intensity is locally applied to the system with more moderate updates with the belief that the system approaches to what is desired. Technically, both algorithms iteratively update each link's intensity based on the gradient of the dual problem of **A-OPT**, where the transmission intensity plays a role of Lagrangian multiplier and step-size function is set to decrease to sufficiently small positive values for convergence guarantees. In updating the intensities as per (20), the empirical service rate $\hat{s}_i(\boldsymbol{r}(t))$ has been used instead of computing the exact service rate $s_i(\boldsymbol{r}(t))$ towards fully-distributed operation. The authors in [41] take an exponentially increasing length of the update intervals in ***JW***, *i.e.*, $\exp(\sqrt{t})$ for t-th frame, with $a_i(t) = 1/t$, so that $\hat{s}_i(\cdot)$ becomes close to $s_i(\boldsymbol{r}(t))$, while the authors in [15] use a fixed duration of update intervals with a decreasing step-size function $a_i(\cdot)$ with the condition of $\sum_t a_i(t) = \infty, \sum_t a_i(t)^2 < \infty$ in ***EJW***.

Instead, all of our proposed dynamics ***SA-BRD***, ***SA-JD***, and ***SA-GD*** are motivated by game-inspired learning dynamics by designing the *rational behavior* of wireless links under CSMA in a non-cooperative game framework. Nevertheless, our algorithms may be interpreted from an optimization perspective. First, ***SA-BRD*** and ***SA-JD*** correspond to approximations (in the sense that they exploit locally-observed aggregate service rates) of the perfect steepest ascent algorithm of **OPT** using the smoothing parameter $\alpha \in (0,1]$ to smooth out the effect of random behavior of wireless links. In particular, we found that the algorithm in [40] becomes equivalent to ***SA-BRD***, where the transmission intensity of each link i is set to be proportional to the first derivative of the utility function $U_i(\cdot)$ at the empirical rate until the corresponding instant. This algorithm is inspired by a steepest ascent method of **OPT**, *i.e.*, $r_i(t+1) = \beta \cdot U_i'\Big(\bar{s}_i(t)\Big)$. This implies that our work reverse-engineers that in [40] in the game-theoretic framework. However, other two algorithms ***SA-JD*** and ***SA-GD*** are the new algorithms which can be developed by our game-theoretic framework, which shows the value of investigating optimal CSMA from a different angle. To interpret ***SA-GD*** from an optimization context, it is a variant (*i.e.*, by using the locally maintained aggregate service rate) of the steepest ascent method of the optimization problem: for $i \in V$, $\max_{r_i} \Phi_i(\boldsymbol{r})$,



(a) star (b) grid (c) bipa (d) comp

Fig. 2. Interference graph topologies: basic topologies

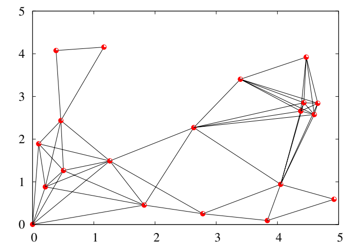


Fig. 3. rand topology of 20 nodes, 28 links

with a constant step-size α. Despite various forms of fully-distributed algorithms from optimization or game theoretic perspective, their convergence and optimality is from smartly applying the stochastic approximation theory to the problem.

VI. Numerical Results

We provide numerical results that show our analytical findings of game dynamics, which we simply call ***SA***-dynamics, by considering various interference network topologies.

Setup. For numerical experiments, we consider proportional fairness across users, *i.e.*, $U_i(\cdot) = \log(\cdot)$ for all users i, and simply fix the unit duration for all dynamics. Since network utility has negative value in our framework of $\log(\cdot)$, to get more intuitive values, we use GAT (Geometric Average of user Throughput) instead, which is defined as $(\Pi_{i\in V} s_i(\boldsymbol{r}(t)))^{1/n}$. Note that under the proportional fairness, $\max$ GAT equals to $\max$ the aggregate $\log$ utilities.

In this paper, we consider "basic" topologies to show that our game dynamics converge to the accurate non-trivial NE, and a random topology that is regarded as a collection of such basic topologies, for more general results. The interference network topologies $G = (V, E)$ under which our results are presented here are star, grid, bipartite, complete, and random graphs, as classified into the following 5 topologies and depicted in Fig. 2 and Fig. 3:

- star: Star interference graph of 5 nodes, 4 links.
- grid: Grid interference graph with of 25 nodes, 40 links.
- bipa: Bipartite interference graph of 20 nodes, 100 links.
- comp: Complete interference graph of 5 nodes, 10 links.
- rand: Random interference graph of 20 nodes, 28 links.

(i) SA-dynamics converge to the unique non-trivial NE: We first demonstrate the result of Theorem 2 for the star topology in Fig. 4(a) by showing the convergence of intensity and GAT to the unique non-trivial NE under ***SA-BRD***, ***SA-JD***, and ***SA-GD***, where we use $\beta = 1.0$ and $\alpha = 0.5$. We consider this simple case to rigorously support that our dynamics find the "accurate" solution (*i.e.*, the unique NE), where the exact solution can be easily characterized. For the star topology, the accurate solution with $\beta = 1.0$ is attained at $r_1^\star = 5.35$ for the hub node, and $r_2^\star = 1.5$ for the other spoke nodes, thus the optimal GAT value, say $U^\star$, becomes 0.516. The intensity updates of the hub node and the corresponding GAT under our ***SA***-dynamics are shown to converge to the unique non-trivial NE $r_1^\star$ and $U^\star$, in Fig. 4(a). The convergence speeds of all algorithms do not show much difference in this simple setup. Figs. 5(a), 6(a), 7(a), and 8(a) also illustrate the convergence of our game-inspired dynamics to the unique non-trivial NE for grid, bipa, comp and rand graph, respectively.

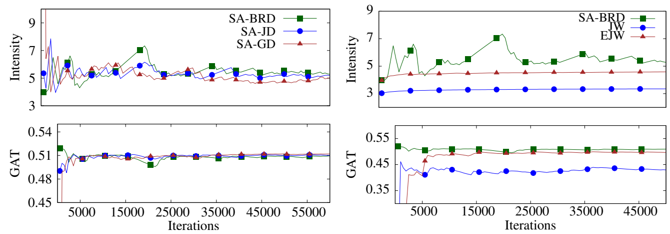
(ii) SA-dynamics converge faster than optimization-based algorithms: Second, we compare the convergence speed of the proposed game dynamics and other utility optimal CSMA algorithms ***JW*** and ***EJW*** described in Section V-B. For both algorithms, we run the simulation under the same setup as in ***SA***-dynamics. Fig. 4(b), Fig. 5(b), Fig. 6(b), Fig. 7(b), and Fig. 8(b) show the traces of transmission intensities and GATs of ***SA-BRD, JW*** and ***EJW*** for star, grid, bipa, comp, and rand topology, where we plot ***SA-BRD*** instead of plotting all three ***SA***-dynamics since the traces are similar as we see in **(i)**. In particular, regarding the results for the star topology in Fig. 4(b), we observe that ***SA-BRD*** converges faster to the $r_1^\star = 5.35$ and $U^\star = 0.516$ within 50000 iterations, while ***JW*** and ***EJW*** are still slowly moving towards the optimal solution.

(iii) The convergence speed is relatively slower under higher-connected interference graphs: As the figures demonstrate, the convergence speed of the dynamics is dependent on the topology characteristics. Based on numerical results, bipa and comp topologies have relatively slower convergence speed than other topologies. As the interference graph has complex interfering relation among links, *e.g.*, comp and bipa, the learning dynamics require more iterations to converge, since the CSMA Markov chain is likely to stay at one state and thus it has much longer mixing time. We run simulations under various interference topology, and find the slower convergence speed under bipa and comp graph than other cases, see Figs. 6 and 7.

(iv) Trade-off between efficiency and convergence speed: Finally, we present the results that show the trade-off between Price-of-Anarchy of ***SA***-dynamics and their convergence speeds, to support the findings stated in Theorem 1, *i.e.*, PoA of ***SA***-dynamics is asymptotically $1/\beta$. To that end, we vary β and plot the GAT at the converged non-trivial NE. For the relation between convergence speed and β, we also measure the convergence time to reach the NE. Fig. 9 shows that, as β grows, ***SA***-dynamics require exponentially long time to reach the equilibrium point, and the corresponding point becomes closer to the socially optimal point. According to the numerical experiments, the GAT with $\beta = 3.0$ is 0.4986 and converges after almost 5×10^8 iterations, while that with $\beta = 1.0$ is 0.4342 and converges after 4×10^6 iterations.

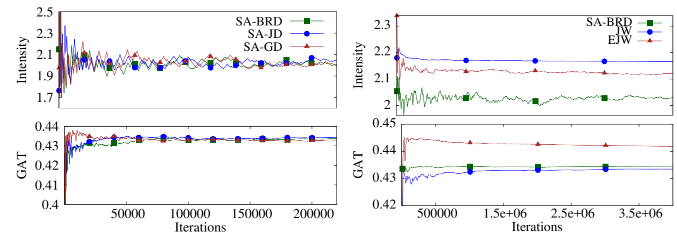
VII. Extension: SINR-based Interference and Collisions

We have so far considered a simple interference model that are captured just by the topological condition and an idealized scenario, *i.e.*, no collision. However, in practical situations, (i) collisions could not be avoided completely, and/or (ii)



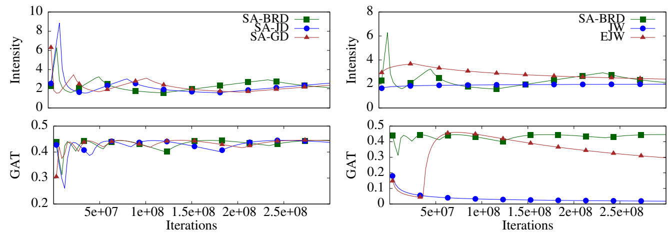
(a) ***SA***-dynamics (b) ***SA-BRD*** vs. ***JW, EJW***

Fig. 4. Convergence of intensity and GAT: `star` topology of 5 nodes.



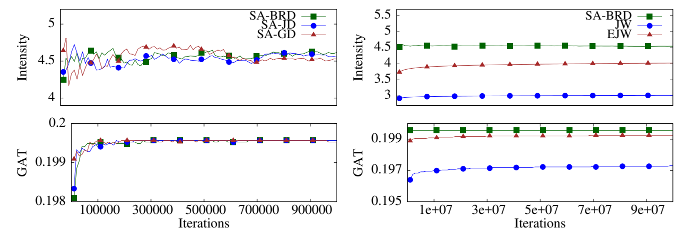
(a) ***SA***-dynamics (b) ***SA-BRD*** vs. ***JW, EJW***

Fig. 5. Convergence of intensity and GAT: `grid` topology of 25 nodes.



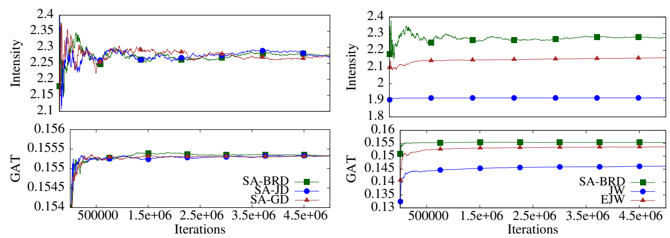
(a) ***SA***-dynamics (b) ***SA-BRD*** vs. ***JW, EJW***

Fig. 6. Convergence of intensity and GAT: `bipa` topology of 20 nodes.



(a) ***SA***-dynamics (b) ***SA-BRD*** vs. ***JW, EJW***

Fig. 7. Convergence of intensity and GAT: `comp` topology of 5 nodes.



(a) ***SA***-dynamics (b) ***SA-BRD*** vs. ***JW, EJW***

Fig. 8. Convergence of intensity and GAT: `rand` topology of 20 nodes, 28 links.

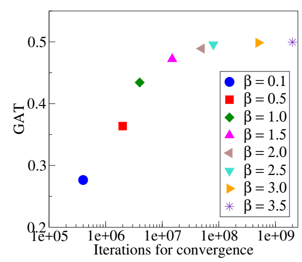


Fig. 9. Trade-off: efficiency(β) and convergence speed.

successful transmission is indeed determined by the aggregated interference level among links. This section is devoted to how our game-inspired approach can be also applied to those practical scenarios. The key step in these extensions is the construction of a CSMA Markov chain with a product-form of stationary distribution. To this end, we can use the techniques developed for (i) and (ii) by [42] and [43], respectively.

A. CSMA with no collisions under the SINR model

Model. Let d_{ji} denote the distance between the links j and i, $\zeta \geq 2$ denote the path loss exponent, and P denote the transmission power at each link. We assume white Gaussian thermal noise with variance ω^2 at all links. Although all the active links in the network can potentially contribute to the interference, the aggregate interference from the links beyond the certain distance, referred to as the *close-in* radius, can be safely neglected to be less than amount of ω_{ci}^2 [51]. Denote $\mathcal{N}(i)$ as the set of links who are within the close-in radius of link i. Now, for a fixed schedule $\boldsymbol{\sigma}$, the total interference power at link i is given by $I_i(\boldsymbol{\sigma}) = \sum_{j\in\mathcal{N}(i):\sigma_j=1} P d_{ji}^{-\zeta}$, and thus the SINR at link i under the schedule $\boldsymbol{\sigma}$ is given by: $\eta_i(\boldsymbol{\sigma}) = \frac{P d_{ii}^{-\zeta}}{I_i(\boldsymbol{\sigma})+\omega_{ci}^2+\omega^2}$. Under this SINR model, we say that link i can successfully transmit if its SINR exceeds a pre-determined SINR requirement κ_i, *i.e.*, $\eta_i(\boldsymbol{\sigma}) \geq \kappa_i$. In contrast to the protocol model where two interfering links cannot transmit simultaneously, active links can coexist under the SINR model if they can make successful transmissions at the same time. Thus, schedule $\boldsymbol{\sigma}$ is said to be *feasible* if the set of active links satisfy the SINR requirements. Then, the set of all feasible schedules under the SINR model, denoted by $\mathcal{S}$, is defined as: $\mathcal{S} \triangleq \{\boldsymbol{\sigma} \in \{0,1\}^n \mid \eta_i(\boldsymbol{\sigma}) \geq \kappa_i, \forall i : \sigma_i = 1\}$, and the throughput region under the SINR model $\Lambda^{\mathcal{S}}$ is the convex hull of $\mathcal{S}$.

SINR-aware feasibility probing mechanism. One significant challenge under the SINR model is that multiple links can transmit successfully at the same time, and such a coexistence relationship is dynamic and complicated, while the feasibility of a schedule under the protocol model can be easily verified only by the topological structure, see Section III-A. To tackle this challenge and construct a (continuous-time) Markov chain structure of a CSMA network under the SINR model as we did under the protocol model, where each state in the Markov chain corresponds to a feasible schedule, we need to ensure that the network always stays in a feasible schedule. To this end, we employ a *SINR-aware feasibility probing* mechanism, motivated by the approach in [42]. The role of this mechanism is to enable each link to judge its coexistence feasibility with the existing active links and avoid possible violations to the SINR requirements, by utilizing carrier sensing and control messages (*i.e.*, interference tolerance level) exchange.

In this probing mechanism, we use the *interference tolerance level*, defined as the interference power that the link can further tolerate without violating the SINR requirement, as follows: for link i at time τ,

$$T_i(\tau) = \frac{Pd_{ii}^{-\zeta}}{\kappa_i} - I_i(\tau) - \omega_{ci}^2 - \omega^2, \tag{21}$$

where $I_i(\tau) = I_i(\boldsymbol{\sigma}(\tau))$ is link i's aggregated interference power from active links at time τ. When a link is activated, it should tolerate the aggregated interference from other active links, and meanwhile, its incurring interference would not violate the SINR requirements of other ongoing transmissions. At time τ, each active link j keeps broadcasting its interference tolerance level $T_j(\tau)$ in the control message (using the separate frequency from data signal) to its nearby links. Then, when an inactive link i is about to be active, link i can estimate its coexistence feasibility, *i.e.*, (i) $T_i(\tau) > 0$ and (ii) for any active link $j \in \mathcal{N}(i)$, the interference from link i to j is no greater than $T_j(\tau)$. Along the same line as in conventional idealized CSMA in Section III-A, each link i contends for transmission using the random backoff and holding time (following exponential distributions with means $1/b_i$ and h_i), except that the backoff timer is suspended when it would violate any existing transmission of nearby links (using SINR-aware feasibility probing mechanism).

CSMA and its Markov chain under the SINR model. By the aid of the probing mechanism, a sequence of schedules $\{\boldsymbol{\sigma}(\tau)\}_{\tau=0}^{\infty}$ of CSMA under the SINR model constructs a continuous-time Markov chain [42]. The Markov chain is shown to be ergodic and time-reversible with the stationary distribution of the following form (see Section IV in [42] for the details): for $\boldsymbol{\sigma} \in \mathcal{S}$,

$$\pi_{\boldsymbol{\sigma}}^{\boldsymbol{r}} = \frac{\exp\left(\sum_{i\in V} \sigma_i r_i\right)}{\sum_{\boldsymbol{\sigma}'\in\mathcal{S}} \exp\left(\sum_{i\in V} \sigma_i' r_i\right)}, \tag{22}$$

where $\boldsymbol{r}$ is transmission intensity of links, as defined in Section III-A. This stationary distribution has similar product form to that of CSMA Markov chain under the protocol model, see (1) and (22), except the set of feasible states. Now, the marginal probability $s_i(\boldsymbol{r})$ under the stationary distribution (22) becomes the link i's long-term throughput as in (2), and the utility maximization problem in this model can be expressed as: $\max_{\boldsymbol{\lambda}\in\Lambda^{\mathcal{S}}} \sum_{i\in V} U_i(\lambda_i)$. All the remaining analysis including (i) game design, (ii) NE analysis, and (iii) game-inspired distributed dynamics can be applied without significant changes.

B. CSMA with collisions under the protocol model

The assumption of continuous distributions of backoff and holding times easily removes the need to consider collisions under the protocol model. However, a real system is discrete, thus collisions naturally occur when two interfering links contend at a same slot. In this discrete system, link throughputs are characterized in more complicated way by considering the transmission loss due to collisions. Following the approach in [43], we will present a *discrete-time Markov chain* for a CSMA with collisions under the protocol model, which also enables us to extend our game-inspired CSMA dynamics achieving utility optimality to the case with collisions.

Model and CSMA algorithm with collisions. We describe a basic CSMA/CA algorithm with fixed transmission probabilities and fixed duration of slot. In each slot, if link i is inactive and if the medium is idle, link i starts transmission with probability p_i. If at a certain slot, link i did not choose to transmit, but an interfering link (captured by an interference graph G) starts transmitting, then link i keeps silent until that transmission finishes. A collision occurs when interfering links try transmitting at the same slot.

To deal with issues not presented in CSMA network with no collisions, let each link transmit a short *probe packet* with length ℓ before the data is transmitted, which enables to avoid collisions of long data packets. Assume that a successful transmission of link i lasts μ_i, which includes a constant overhead $\bar{\mu}$ (due to RTS, CTS, ACK, DIFS, etc.) and the data payload μ_i^{d}, which is a random variable with mean $M_i = \mathbb{E}[\mu_i]$. For a fixed schedule $\boldsymbol{\sigma}$, let $\mathcal{T}(\boldsymbol{\sigma})$ denote the set of links having a successful transmission (*i.e.*, no collisions with interfering links at the slot), $\mathcal{C}(\boldsymbol{\sigma})$ denote the set of links that are experiencing collisions, and $h(\boldsymbol{\sigma})$ denote the *collision number*, *i.e.*, the size of $\mathcal{C}(\boldsymbol{\sigma})$. Clearly, any active link of a schedule $\boldsymbol{\sigma}$ is in either $\mathcal{T}(\boldsymbol{\sigma})$ or $\mathcal{C}(\boldsymbol{\sigma})$.

Markov chain of CSMA with collisions. Now, we can construct the underlying discrete-time Markov chain, which evolves slot by slot. The state of the Markov chain in a slot t is $\boldsymbol{w}(t) \triangleq \{\boldsymbol{\sigma}(t), [(l_i(t), e_i(t)) \mid \forall i : \sigma_i(t) = 1]\}$, where $l_i(t)$ is the total length of the current packet that link i is transmitting, and $e_i(t)$ is the remaining time before the transmission of link i ends. As detailed in Section II and Appendix of [43], a sequence of $\{\boldsymbol{w}(t)\}_{t\in\mathbb{N}}$ is shown to be ergodic and almost time-reversible discrete-time Markov chain. Moreover, for a given transmission probability $\boldsymbol{p} = [p_i]_{i\in V}$, its stationary distribution has a simple product-form, from which the probability of a schedule $\boldsymbol{\sigma}$ can be obtained as follows (see Theorem 1 in [43] for details):

$$\pi_{\boldsymbol{\sigma}} = \frac{1}{Z}\left(\ell^{h(\boldsymbol{\sigma})} \cdot \prod_{i\in\mathcal{T}(\boldsymbol{\sigma})} M_i\right) \cdot \prod_{i\in V} p_i^{\sigma_i}(1-p_i)^{1-\sigma_i}. \tag{23}$$

To see how we can apply our game-based approach for idealized CSMA to the CSMA with collisions, we first re-write the distribution (23) with parameter $\boldsymbol{\varrho} = [\varrho_i]_{i\in V}$, defined as $M_i = \bar{\mu} + M_0 \cdot \exp(\varrho_i)$. In particular, $\boldsymbol{\varrho}$ is a vector representing *transmission length* of links, *i.e.*, $M_0 \cdot \exp(\varrho_i)$ is the mean length of the data payload where M_0 is a constant reference payload length. Then, for a given $\boldsymbol{\varrho}$, the distribution is re-written as:

$$\pi_{\boldsymbol{\sigma}}^{\boldsymbol{\varrho}} = \frac{1}{Z(\boldsymbol{\varrho})} g(\boldsymbol{\sigma}) \cdot \prod_{i\in\mathcal{T}(\boldsymbol{\sigma})} \left(\bar{\mu} + M_0 \cdot \exp(\varrho_i)\right),$$

where $g(\boldsymbol{\sigma}) = \ell^{h(\boldsymbol{\sigma})} \cdot \prod_{i\in V} p_i^{\sigma_i}(1-p_i)^{1-\sigma_i}$ is not related to $\boldsymbol{\varrho}$, and $Z(\boldsymbol{\varrho})$ is a normalizing term. Then, a throughput (or service rate) of link i in terms of data payload, *i.e.*, the stationary probability that link i transmitting a data payload, is given by: $s_i(\boldsymbol{\varrho}) = \frac{M_0 \cdot \exp(\varrho_i)}{\bar{\mu} + M_0 \cdot \exp(\varrho_i)} \sum_{\boldsymbol{\sigma}: i\in\mathcal{T}(\boldsymbol{\sigma})} \pi_{\boldsymbol{\sigma}}^{\boldsymbol{\varrho}}$.

If link i is transmitting successfully at a schedule $\boldsymbol{\sigma}$, *i.e.*, $i \in \mathcal{T}(\boldsymbol{\sigma})$, it is either transmitting the overhead or the data payload. Now, this can be captured by the *detailed state* $(\boldsymbol{\sigma}, \boldsymbol{z})$, where $z_i = 1$ if $i \in \mathcal{T}(\boldsymbol{\sigma})$ and it transmits the payload, and $z_i = 0$ if $i \in \mathcal{T}(\boldsymbol{\sigma})$ and it transmits the overhead. Denote the set of all possible detailed states by $\mathcal{D}(G)$, and define the throughput region $\Lambda^{\mathcal{D}}$ of a CSMA network with collisions as the convex hull of $\mathcal{D}(G)$. Then, we have the following product-form stationary distribution of the detailed state $(\boldsymbol{\sigma}, \boldsymbol{z}) \in \mathcal{D}(G)$ [43]:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2017.2764081, IEEE Transactions on Wireless Communications

$$\pi^{\boldsymbol{\varrho}}_{(\boldsymbol{\sigma},\boldsymbol{z})} = \frac{1}{E(\boldsymbol{\varrho})} g(\boldsymbol{\sigma},\boldsymbol{z}) \cdot \exp\left(\sum_{i\in V} z_i \varrho_i\right),$$

where $g(\boldsymbol{\sigma},\boldsymbol{z}) = g(\boldsymbol{\sigma}) \cdot (\bar{\mu})^{|\mathcal{T}(\boldsymbol{\sigma})|-\mathbf{1}'\boldsymbol{z}} M_0^{\mathbf{1}'\boldsymbol{z}}$, and $E(\boldsymbol{\varrho})$ is a normalizing term. Clearly, this provides alternative expression of the service rate $s_i(\boldsymbol{\varrho}) = \sum_{(\boldsymbol{\sigma},\boldsymbol{z})\in\mathcal{D}(G):z_i=1} \pi^{\boldsymbol{\varrho}}_{(\boldsymbol{\sigma},\boldsymbol{z})}$. Now, the utility maximization problem in this model can be expressed as: $\max_{\boldsymbol{\lambda}\in\Lambda^{\mathcal{D}}} \sum_{i\in V} U_i(\lambda_i)$. Note that the stationary distribution and service rate parameterized by $\boldsymbol{\varrho}$ have a similar form to those of CSMA Markov chain in (1) and (2). The only difference is that this discrete-time Markov chain is controlled by a parameter $\boldsymbol{\varrho}$ with a set of feasible states $\mathcal{D}(G)$. Therefore, all the remaining analysis including (i) game design, (ii) NE analysis, and (iii) game-inspired distributed dynamics can be extended to this case (*i.e.*, CSMA with collisions), by applying the results with respect to transmission length $\boldsymbol{\varrho}$ instead of transmission intensity $\boldsymbol{r}$.

VIII. Conclusion

Despite a large array of game-theoretic studies on wireless MAC, to the best of knowledge, this is the first game-theoretic work that is utility optimal over the maximum throughput region. We start by designing a non-cooperative CSMA game whose equilibrium properties (*i.e.*, uniqueness and Price-of-Anarchy) are first analyzed, and propose three game-inspired dynamics based on the idea of stochastic approximation theory. Our theoretical findings exploit the features of CSMA, where the price function is smartly designed so that the NE is unique and asymptotically close to the social optimum as well as fully-distributed dynamics are feasible. Our main contribution lies in developing a different style of fully-distributed optimal CSMA algorithms inspired from a game-theoretic approach, which we believe, provides useful insights and interests with the support from several efforts in literature ensuring practical values of our theoretical results.

Appendix: Proof of Theorem 2

A. Proof of Theorem 2: *SA-BRD*

To prove that ***SA-BRD*** converges to the non-trivial NE, we show that in **Step 1**, the evolutions of strategies in ***SA-BRD*** asymptotically approach deterministic trajectory. In next step, we prove that the resulting deterministic trajectory converges to the non-trivial NE. To do this, we use a similar approach as that used in Theorem 1 of [52] and Corollary 8 of [53][pp.74].

Step 1. The first step is to approximate the dynamics of strategies when t is large by dynamics of a continuous-time ordinary differential equation (ODE) system, by introducing continuous-time interpolation of strategies. We start with the definition of $\bar{s}_i(t)$,

$$\bar{s}_i(t) = \frac{1}{t}\sum_{n=0}^{t} \hat{s}_i(n) = \bar{s}_i(t-1) - \frac{1}{t}(\bar{s}_i(t-1) - \hat{s}_i(t)). \quad (24)$$

SA-BRD algorithm (17) then becomes approximately as follows when t grows large:

$$\begin{aligned} & r_i(t+1) \\ & \overset{(a)}{=} \beta\Big(U_i'(\bar{s}_i(t-1)) - \frac{1}{t}(\bar{s}_i(t-1) - \hat{s}_i(t))U_i''(\bar{s}_i(t-1))\Big) \\ & = \beta U_i'(\bar{s}_i(t-1)) + \frac{1}{t}\Big(-\beta U_i''(\bar{s}_i(t-1))\Big)(\bar{s}_i(t-1) - \hat{s}_i(t)) \\ & \overset{(b)}{=} r_i(t) + \frac{1}{t} g(r_i(t))\left(U_i'^{-1}\left(\frac{r_i(t)}{\beta}\right) - \hat{s}_i(t)\right), \end{aligned} \quad (25)$$

where

$$g(x) = -\beta U_i''(U_i'^{-1}(\frac{x}{\beta})) > 0, \quad (26)$$

for concave, increasing utility function. The equality (a) holds when t is sufficiently large, and the equality (b) comes from the ***SA-BRD*** rule at time t:

$$r_i(t) = \beta U_i'(\bar{s}_i(t-1)), \text{ and thus } \bar{s}_i(t-1) = U_i'^{-1}\left(\frac{r_i(t)}{\beta}\right).$$

Now, we define $\kappa(t) := \sum_{j=1}^{t} \frac{1}{j}$ where $\kappa(0) = 0$. We take continuous-time interpolation from the discrete-time sequence $\{\boldsymbol{r}(t)\}_{t\in\mathbb{Z}_{\geq 0}}$ in the following way: define $\{\bar{\boldsymbol{r}}(\tau)\}_{\tau\in\mathbb{R}_+}$ as: $\forall t \in \mathbb{Z}_{\geq 0}$, for all $\tau \in [\kappa(t), \kappa(t+1))$, $\bar{r}_i(\tau) = r_i(t) + (r_i(t+1) - r_i(t)) \times \frac{\tau - \kappa(t)}{\kappa(t+1)-\kappa(t)}$. Also define continuous-time instantaneous service rate as $\hat{s}_i(\tau) = \hat{s}_i(t) \cdot \mathbf{1}_{\kappa(t)\leq\tau<\kappa(t+1)}$[7]. It should be clear that ***SA-BRD*** (25) is a stochastic approximation algorithm with controlled Markov noise as defined in [15], [52], [53]. We can now easily check follows:

i. We set the strategy domain to the compact set $[r^{\min}, r^{\max}]$.
ii. The transition kernel of $[\hat{s}_i(t)]_{i\in V}$ is continuous in $\boldsymbol{r}(t)$ and corresponding Markov chain is ergodic with stationary distribution $\boldsymbol{\pi}^{\boldsymbol{r}}$ for fixed $\boldsymbol{r}$.
iii. The function $g(r_i)\Big(U_i'^{-1}\Big(\frac{r_i}{\beta}\Big) - \hat{s}_i\Big)$ is continuous and Lipschitz in r_i, uniformly in $\hat{s}_i$ due to given properties of utility function and boundedness of strategy set.
iv. $\hat{\boldsymbol{s}}(t)$ is a stochastic process with values in a finite space.

Then applying Theorem 1 of [52] to ***SA-BRD***, we get followings: Let $T > 0$ and fix $w > 0$. Denote by $\tilde{\boldsymbol{r}}^w(\cdot)$ the solution on $[w, w+T]$ of the following ODE system: $\forall i \in V$,

$$\dot{r}_i(\tau) = g(r_i(\tau))\left[U_i'^{-1}\left(\frac{r_i(\tau)}{\beta}\right) - \sum_{\boldsymbol{\sigma}} \pi^{\boldsymbol{r}(\tau)}_{\boldsymbol{\sigma}} \sigma_i\right], \quad (27)$$

with $\tilde{\boldsymbol{r}}^w(w) = \bar{\boldsymbol{r}}(w)$. Then, we have almost surely, $\lim_{w\to\infty} \sup_{\tau\in[w,w+T]} \|\bar{\boldsymbol{r}}(\tau) - \tilde{\boldsymbol{r}}^w(\tau)\| = 0$. Note that if the ODE system (27) has a unique fixed point $\boldsymbol{r}_\star$, then we would have $\lim_{\tau\to\infty} \bar{\boldsymbol{r}}(\tau) = \boldsymbol{r}_\star$. As a consequence, we would also have $\lim_{t\to\infty} \boldsymbol{r}(t) = \boldsymbol{r}_\star$.

Step 2. The second step of the proof consists in showing that (27) is interpreted as the subgradient algorithm of the dual of **A-OPT**, using a similar technique as in [14]. The Karush-Kuhn-Tucker(KKT) condition for dual variable $\boldsymbol{k}$ is given by:

$$U_i'^{-1}(k_i) - \sum_{\boldsymbol{\sigma}} \pi^{\boldsymbol{r}}_{\boldsymbol{\sigma}} \sigma_i = 0, \quad \forall i \in V, \quad (28)$$

with $\boldsymbol{r} = \beta\boldsymbol{k}$. Now the subgradient algorithm corresponding to (28) is: for all $i \in V$,

$$\dot{r}_i = \left(U_i'^{-1}\left(\frac{r_i}{\beta}\right) - \sum_{\boldsymbol{\sigma}} \pi^{\boldsymbol{r}}_{\boldsymbol{\sigma}} \sigma_i\right). \quad (29)$$

Note that (29) is equivalent to (27) since $g(\cdot) > 0$ for concave, increasing utility function. Finally, **A-OPT** is a strictly convex

[7] $\mathbf{1}_E$ is the indicator function for the event E.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2017.2764081, IEEE Transactions on Wireless Communications

optimization problem with unique solution, and hence this subgradient algorithm converges to its unique solution $\boldsymbol{r}^o$. Using $\boldsymbol{r}^o = \boldsymbol{r}^{\text{NE}}$ and **Step 1**, we complete the proof of ***SA-BRD***'s convergence to the unique non-trivial $\boldsymbol{r}^{\text{NE}}$ of oCSMA(β). □

B. Proof of Theorem 2: *SA-JD*

Step 1. We first define a sequence $\{\boldsymbol{\nu}(t)\}_{t\in\mathbb{Z}_{\ge 0}}$ derived from $\{\boldsymbol{r}(t)\}_{t\in\mathbb{Z}_{\ge 0}}$ as:

$$\begin{aligned} \nu_i(t) &= r_i(t) - (1-\alpha)\cdot r_i(t-1), \\ r_i(t) &= \alpha \cdot \frac{\nu_i(t)}{\alpha} + (1-\alpha)\cdot r_i(t-1). \end{aligned} \tag{30}$$

Recall ***SA-JD***'s update rule (18), then under **(A1)**, it is represented as

$$\begin{aligned} &\nu_i(t+1) = r_i(t+1) - (1-\alpha) r_i(t) = \alpha\beta U_i'(\bar{s}_i(t)) \\ &\overset{(a)}{=} \alpha\beta U_i'\bigg(\bar{s}_i(t-1) - \frac{1}{t}(\bar{s}_i(t-1) - \hat{s}_i(t))\bigg) \\ &\overset{(b)}{=} \alpha\beta\bigg(U_i'(\bar{s}_i(t-1)) - \frac{1}{t}(\bar{s}_i(t-1) - \hat{s}_i(t))U_i''(\bar{s}_i(t-1))\bigg) \\ &= \alpha\beta U_i'(\bar{s}_i(t-1)) + \frac{1}{t}\alpha g\bigg(\frac{\nu_i(t)}{\alpha}\bigg)(\bar{s}_i(t-1) - \hat{s}_i(t)) \\ &\overset{(c)}{=} \nu_i(t) + \frac{1}{t}\alpha g\bigg(\frac{\nu_i(t)}{\alpha}\bigg)\bigg(U_i'^{-1}\bigg(\frac{\nu_i(t)}{\alpha\beta}\bigg) - \hat{s}_i(t)\bigg), \end{aligned} \tag{31}$$

where $g(\cdot)$ is defined in (26). The equality (a) holds due to (24), the equality (b) holds when t is sufficiently large, and the equality (c) comes from the ***SA-JD*** rule at time t:

$$\nu_i(t) = \alpha\beta U_i'(\bar{s}_i(t-1)), \text{ and thus } \bar{s}_i(t-1) = U_i'^{-1}\bigg(\frac{\nu_i(t)}{\alpha\beta}\bigg).$$

As we did in the analysis of ***SA-BRD***, we now define a continuous-time interpolation $\{\bar{\boldsymbol{\nu}}(\tau)\}_{\tau\in\mathbb{R}_+}$ from the discrete-time sequence $\{\boldsymbol{\nu}(t)\}_{t\in\mathbb{Z}_{\ge 0}}$ as follows: $\forall t \in \mathbb{Z}_{\ge 0}$, $\forall \tau \in [\kappa(t), \kappa(t+1))$, $\bar{\nu}_i(\tau) = \nu_i(t) + (\nu_i(t+1) - \nu_i(t)) \times \frac{\tau - \kappa(t)}{\kappa(t+1)-\kappa(t)}$. It should be clear that **SA-JD** (31) is also a stochastic approximation algorithm with controlled Markov noise, and we can verify the conditions for convergence (i)-(iv) of **Step 1** in proof of ***SA-BRD***. The only difference is that (31) is a stochastic process of $\nu_i(t)$ instead of $r_i(t)$, and it has $\alpha\beta$ instead of α. Having applying Theorem 1 of [52] in the framework of ***SA-JD***, we have followings: Let $T > 0$ and fix $w > 0$. Denote by $\tilde{\boldsymbol{\nu}}^w(\cdot)$ the solution on $[w, w+T]$ of the following ODE system: for all $i \in V$,

$$\dot{\nu}_i(\tau) = \alpha g\bigg(\frac{\nu_i(\tau)}{\alpha}\bigg)\bigg[U_i'^{-1}\bigg(\frac{\nu_i(\tau)}{\alpha\beta}\bigg) - \sum_{\boldsymbol{\sigma}} \pi_{\boldsymbol{\sigma}}^{\boldsymbol{\nu}(\tau)/\alpha}\sigma_i\bigg], \tag{32}$$

with $\tilde{\boldsymbol{\nu}}^w(w) = \bar{\boldsymbol{\nu}}(w)$. Then, we have almost surely, $\lim_{w\to\infty}\sup_{\tau\in[w,w+T]}\|\bar{\boldsymbol{\nu}}(\tau) - \tilde{\boldsymbol{\nu}}^w(\tau)\| = 0$. As a consequence, if the above ODE system (32) has a unique fixed point $\boldsymbol{\nu}_\star$, then we would have that almost surely, $\lim_{t\to\infty}\boldsymbol{\nu}(t) = \boldsymbol{\nu}_\star$.

Step 2. This step proves the equivalence of the convergence of virtual process $\{\boldsymbol{\nu}(t)\}_{t\in\mathbb{Z}_{\ge 0}}$ and that of $\{\boldsymbol{r}(t)\}_{t\in\mathbb{Z}_{\ge 0}}$. Since r_i lies in compact region, from (30), there exist L and M such that for all t,

$$|\nu_i(t)| < L \text{ and } \bigg|g\bigg(\frac{\nu_i(t)}{\alpha}\bigg)\bigg(U_i'^{-1}\bigg(\frac{\nu_i(t)}{\alpha\beta}\bigg) - \hat{s}_i(t)\bigg)\bigg| < M.$$

For $\varepsilon > 0$, let $T(\varepsilon) := \frac{4\log(\frac{\varepsilon\alpha M}{4L})}{\varepsilon\log(1-\alpha)}$. Then, for all $t \ge T(\varepsilon)$, $\left|\frac{\nu_i(t)}{\alpha} - r_i(t)\right| \le \frac{5}{4}\varepsilon M$, because [8]

$$\begin{aligned} \bigg|\frac{\nu_i(t)}{\alpha} - r_i(t)\bigg| &\overset{(a)}{=} \bigg|\frac{\nu_i(t)}{\alpha} - \sum_{j=0}^{t-1}\frac{\nu_i(t-j)}{\alpha}\alpha(1-\alpha)^j\bigg| \\ &\le \sum_{j=0}^{t-1}\bigg|\frac{\nu_i(t)}{\alpha} - \frac{\nu_i(t-j)}{\alpha}\bigg|\alpha(1-\alpha)^j + \bigg|\frac{\nu_i(t)}{\alpha}\cdot(1-\alpha)^t\bigg| \\ &\overset{(b)}{\le} \frac{\varepsilon t/4}{t - \varepsilon t/4}M + \frac{2L}{\alpha}\sum_{j=\varepsilon t/4}^{t-1}\alpha(1-\alpha)^j + \frac{L}{\alpha}(1-\alpha)^t \\ &\overset{(c)}{\le} \frac{\varepsilon}{2}M + \frac{2L}{\alpha}(1-\alpha)^{\varepsilon t/4} + \frac{\varepsilon}{4}M \;\le\; \frac{5}{4}\varepsilon M, \end{aligned} \tag{33}$$

where (a) comes from (30) by applying recursion as:

$$\begin{aligned} r_i(t) &= \alpha\cdot\frac{\nu_i(t)}{\alpha} + (1-\alpha)\cdot r_i(t-1) \\ &= \alpha\frac{\nu_i(t)}{\alpha} + (1-\alpha)\bigg(\alpha\frac{\nu_i(t-1)}{\alpha} + (1-\alpha)r_i(t-2)\bigg) \\ &= \cdots = \sum_{j=0}^{t-1}\frac{\nu_i(t-j)}{\alpha}\cdot\alpha(1-\alpha)^j, \end{aligned}$$

and where (b) comes from the followings:

$$\begin{aligned} &\sum_{j=0}^{\varepsilon t/4-1}\bigg|\frac{\nu_i(t)}{\alpha} - \frac{\nu_i(t-j)}{\alpha}\bigg|\alpha(1-\alpha)^j \\ &\le \sum_{j=1}^{\varepsilon t/4-1}\sum_{k=1}^{j}\bigg|\frac{\nu_i(t-k+1)}{\alpha} - \frac{\nu_i(t-k)}{\alpha}\bigg|\alpha(1-\alpha)^j \\ &\le \sum_{j=1}^{\varepsilon t/4-1}\frac{j\cdot M}{t-j}\alpha(1-\alpha)^j \;\le\; \frac{\varepsilon t/4}{t-\varepsilon t/4}M, \end{aligned}$$

and where (c) holds for $t \ge T(\varepsilon)$ and $\varepsilon \le 2$. Therefore,

$$\lim_{t\to\infty}\frac{\nu_i(t)}{\alpha} - r_i(t) = 0.$$

Step 3. From **Step 1** and **Step 2**, the ODE system (32) is equivalent ODE system to (27), thus it converges to a fixed point $\boldsymbol{\nu}^\dagger$ such that $\boldsymbol{r}^\dagger = \frac{\boldsymbol{\nu}^\dagger}{\alpha}$ and

$$\alpha\beta U_i'(s_i(\boldsymbol{r}^\dagger)) = \alpha\beta U_i'\bigg(s_i\bigg(\frac{\boldsymbol{\nu}^\dagger}{\alpha}\bigg)\bigg) = \nu_i^\dagger = \alpha r_i^\dagger.$$

Note that $\boldsymbol{r}^\dagger$ satisfies $r_i^\dagger = \beta U_i'(s_i(\boldsymbol{r}^\dagger))$, and thus it is clear that $\boldsymbol{r}^\dagger = \boldsymbol{r}^{\text{NE}}$. Now, we can conclude that $\boldsymbol{r}(t)$ of ***SA-JD*** converges to the unique non-trivial $\boldsymbol{r}^{\text{NE}}$ of oCSMA(β). □

C. Proof of Theorem 2: *SA-GD*

In case of ***SA-GD***, similar proof strategy is applied as in ***SA-JD***, thus we provide brief proof due to the space limit.

Step 1. We start with defining following alternative discrete sequence $\{\boldsymbol{\upsilon}(t)\}_{t\in\mathbb{Z}_{\ge 0}}$ as follows:

$$\upsilon_i(t) \;=\; \frac{r_i(t)}{\nabla s_i(\boldsymbol{r}(t))} - \bigg(\frac{1}{\nabla s_i(\boldsymbol{r}(t))} - \frac{\alpha}{\beta}\bigg)r_i(t-1),$$

[8]Here, we use just $\varepsilon t/4$ instead of $\lceil \varepsilon t/4 \rceil$ for notional simplicity.

and thus we have following property:

$$r_i(t) = \upsilon_i(t) \cdot \nabla s_i(\boldsymbol{r}(t)) + \left(1 - \frac{\alpha}{\beta}\nabla s_i(\boldsymbol{r}(t))\right) \cdot r_i(t-1).$$

then, ***SA-GD*** is understood as the update rule of $\{\boldsymbol{\upsilon}(t)\}_{t\in\mathbb{Z}_{\geq 0}}$ with following representation:

$$\upsilon_i(t+1) = \alpha U_i'\Big(\bar{s}_i(t)\Big), \text{ and thus } \bar{s}_i(t-1) = U_i'^{-1}\left(\frac{\upsilon_i(t)}{\alpha}\right).$$

Now, under **(A1)** and from the property of service rate function in (8), we have that $s_i'(\boldsymbol{r})$ is positive and in a range of $(0, 1/4]$. Then, for sufficiently large t, ***SA-GD***'s update rule is represented as follows: where $g(\cdot)$ is defined in (26),

$$\begin{aligned}\upsilon_i(t+1) &= \alpha U_i'(\bar{s}_i(t)) \\ &= \upsilon_i(t) + \frac{1}{t}\frac{\alpha}{\beta}g\left(\frac{\beta\upsilon_i(t)}{\alpha}\right)\left(U_i'^{-1}\left(\frac{\upsilon_i(t)}{\alpha}\right) - \hat{s}_i(t)\right).\end{aligned}$$

Now, similar arguments as in ***SA-JD*** can be applied. First, we define a continuous-time interpolation $\{\bar{\boldsymbol{\upsilon}}(\tau)\}_{\tau\in\mathbb{R}_+}$ of the alternative process $\{\boldsymbol{\upsilon}(t)\}_{t\in\mathbb{Z}_{\geq 0}}$, and verify the conditions for convergence to the solution of corresponding ODE system:

$$\dot{\upsilon}_i(\tau) = \frac{\alpha}{\beta}g_i\left(\frac{\beta\upsilon_i(\tau)}{\alpha}\right)\left[U_i'^{-1}\left(\frac{\upsilon_i(\tau)}{\alpha}\right) - \sum_{\boldsymbol{\sigma}}\pi_{\boldsymbol{\sigma}}^{\boldsymbol{\upsilon}(\tau)/(\alpha/\beta)}\sigma_i\right].$$

Then, the asymptotic closeness of $\{\bar{\boldsymbol{\upsilon}}(\tau)\}_{\tau\in\mathbb{R}_+}$ to the solution trajectory of the above ODE system can be claimed as in **Step 1** of ***SA-JD***.

<u>**Step 2.**</u> Now, the equivalence of convergence of $\{\boldsymbol{\upsilon}(t)\}_{t\in\mathbb{Z}_{\geq 0}}$ and that of $\{\boldsymbol{r}(t)\}_{t\in\mathbb{Z}_{\geq 0}}$ in (19) is shown as follows: First, for notational simplicity we use $\gamma = \alpha/\beta$, and we denote by $\gamma_{\min}, \gamma_{\max}$ the minimum and maximum value of the sequence $\{\gamma_t := \gamma\nabla s_i(\boldsymbol{r}(t))\}_{t\in\mathbb{Z}_{\geq 0}}$, respectively. Note that $\gamma_{\min}, \gamma_{\max} < \gamma$ since $s_i'(\boldsymbol{r}) \in (0, 1/4]$. Then, for $\varepsilon > 0$, let $S(\varepsilon) := \frac{4\log(\frac{\varepsilon\gamma M}{4L}\cdot\frac{\gamma_{\max}}{\gamma_{\min}})}{\varepsilon\log(1-\gamma_{\min})}$. Since r_i lies in compact region, from (34), there exist constants L and M such that for all t,

$$|\upsilon_i(t)| < L\ ,\ \left|g\left(\frac{\beta\upsilon_i(t)}{\alpha}\right)\left(U_i'^{-1}\left(\frac{\upsilon_i(t)}{\alpha}\right) - \hat{s}_i(t)\right)\right| < M,$$

$$\text{and } \gamma > \left(1 - \frac{\gamma_{\min}}{\gamma_{\max}}\right)\cdot\frac{4L}{\varepsilon M}.$$

Then, $\forall t \geq S(\varepsilon)$, we can verify that $|\frac{\upsilon_i(t)}{\gamma} - r_i(t)| \leq \frac{5}{4}\varepsilon M$, following similar arguments in (33):

$$\begin{aligned}\left|\frac{\upsilon_i(t)}{\gamma} - r_i(t)\right| &\overset{(a)}{=} \left|\frac{\upsilon_i(t)}{\gamma} - \sum_{j=0}^{t-1}\frac{\upsilon_i(t-j)}{\gamma}\prod_{k=1}^{j}(1-\gamma_k)\gamma_{j+1}\right| \\ &\leq \sum_{j=0}^{t-1}\left|\frac{\upsilon_i(t)}{\gamma} - \frac{\upsilon_i(t-j)}{\gamma}\right|\prod_{k=1}^{j}(1-\gamma_k)\gamma_{j+1} \\ &\quad + \left|\frac{\upsilon_i(t)}{\gamma}\cdot\left(1 - \sum_{j=1}^{t-1}\prod_{k=1}^{j}(1-\gamma_k)\gamma_{j+1}\right)\right| \\ &\overset{(b)}{\leq} \frac{\varepsilon t/4}{t - \varepsilon t/4}M + \frac{2L}{\gamma}\sum_{j=\varepsilon t/4}^{t-1}\gamma_{\max}(1-\gamma_{\min})^j \\ &\quad + \frac{L}{\gamma}\left(1 - \sum_{j=1}^{t-1}\gamma_{\min}(1-\gamma_{\max})^j\right) \\ &\overset{(c)}{\leq} \frac{\varepsilon}{2}M + \frac{2L}{\gamma}\cdot\frac{\gamma_{\max}}{\gamma_{\min}}\cdot(1-\gamma_{\min})^{\varepsilon t/4} + \frac{\varepsilon}{4}M \ \leq\ \frac{5}{4}\varepsilon M,\end{aligned}$$

where (a) comes from (34) by applying recursion, and (b) is straightforward as in ***SA-JD***, and (c) holds for $t \geq S(\varepsilon)$ and $\varepsilon \leq 2$. Now, we have following consequence:

$$\lim_{t\to\infty}\frac{\upsilon_i(t)}{\alpha/\beta} - r_i(t) = 0.$$

<u>**Step 3.**</u> We can finally conclude that $\boldsymbol{r}(t)$ of ***SA-GD*** converges to $\boldsymbol{r}^{\ddagger}$ such that $r_i^{\ddagger} = \beta U_i'(s_i(\boldsymbol{r}^{\ddagger}))$, and it is clear that $\boldsymbol{r}^{\ddagger} = \boldsymbol{r}^{\mathrm{NE}}$, which completes the proof. □

References

[1] H. Jang, S. Yun, J. Shin, and Y. Yi, "Distributed learning for utility maximization over CSMA-based wireless multihop networks," in *Proc. of Infocom*, 2014.

[2] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1949, December 1992.

[3] C. Joo and N. B. Shroff, "Performance of random access scheduling schemes in multi-hop wireless networks," in *Proc. of Infocom*, 2007.

[4] X. Lin and S. Rasool, "Constant-time distributed scheduling policies for ad hoc wireless networks," in *Proc. of CDC*, 2006.

[5] L. Tassiulas, "Linear complexity algorithms for maximum throughput in radio networks and input queued switches," in *Proc. of Infocom*, 1998.

[6] E. Modiano, D. Shah, and G. Zussman, "Maximizing throughput in wireless networks via gossiping," in *Proc. of Sigmetrics*, New York, NY, USA, 2006.

[7] A. Eryilmaz, A. Ozdaglar, and E. Modiano, "Polynomial complexity algorithms for full utilization of multi-hop wireless networks," in *Proc. of Infocom*, 2007.

[8] S. Sanghavi, L. Bui, and R. Srikant, "Distributed link scheduling with constant overhead," in *Proc. of Sigmetrics*, 2007.

[9] P. Chaporkar, K. Kar, and S. Sarkar, "Throughput guarantees through maximal scheduling in wireless networks," in *Proc. of Allerton*, 2005.

[10] X. Lin and N. B. Shroff, "The impact of imperfect scheduling on cross-layer rate control in wireless networks," in *Proc. of Infocom*, 2005.

[11] X. Wu and R. Srikant, "Bounds on the capacity region of multi-hop wireless networks under distributed greedy scheduling," in *Proc. of Infocom*, 2006.

[12] L. Jiang and J. Walrand, *Scheduling and Congestion Control for Wireless and Processing Networks*. Morgan-Claypool, 2010.

[13] Y. Yi and M. Chiang, "Stochastic network utility maximization: A tribute to kelly's paper published in this journal a decade ago," *European Transactions on Telecommunications*, vol. 19, no. 4, pp. 421–442, 2008.

[14] L. Jiang and J. Walrand, "A distributed CSMA algorithm for throughput and utility maximization in wireless networks," *IEEE/ACM Transactions on Networking*, vol. 18, no. 3, pp. 960 –972, June 2010.

[15] J. Liu, Y. Yi, A. Proutiere, M. Chiang, and H. V. Poor, "Towards utility-optimal random access without message passing," *Wiley Journal of Wireless Communications and Mobile Computing*, vol. 10, no. 1, pp. 115–128, Jan. 2010.

[16] S. Rajagopalan, D. Shah, and J. Shin, "Network adiabatic theorem: An efficient randomized protocol for contention resolution," in *Proc. of Sigmetrics*, 2009.

[17] J. Ni, B. Tan, and R. Srikant, "Q-csma: Queue-length based CSMA/CA algorithms for achieving maximum throughput and low delay in wireless networks," in *Proc. of Infocom*, 2010.

[18] S. Y. Yun, Y. Yi, J. Shin, and D. Y. Eun, "Optimal csma: A survey," in *Proc. of ICCS*, 2012.

[19] J. Lee, H. Lee, Y. Yi, S. Chong, E. W. Knightly, and M. Chiang, "Making 802.11 DCF near-optimal: Design, implementation, and evaluation," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1745 –1758, 2016.

[20] N. Li and J. R. Marden, "Designing games for distributed optimization," *J. Sel. Topics Signal Processing*, vol. 7, no. 2, pp. 230–242, 2013.

[21] S. Hart and A. M. Colell, "Uncoupled dynamics do not lead to nash equilibrium," *American Economic Review*, vol. 93, no. 5, pp. 1830–1836, 2003.

[22] P. H. Young, "Learning by trial and error," *Games and Economic Behavior*, vol. 65, no. 2, pp. 626–643, March 2009.

[23] Z. Ghahramani, "Graphical models: Parameter learning," in *Handbook of Brain Theory and Neural Networks*, 2003.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2017.2764081, IEEE Transactions on Wireless Communications

[24] G. E. Hinton, “Training products of experts by minimizing contrastive divergence,” *Neural Computation.*, vol. 14, no. 8, pp. 1771–1800, Aug. 2002.

[25] M. Ghazvini, N. Movahedinia, K. Jamshidi, and N. Moghim, “Game theory applications in csma methods,” *IEEE Communications Surveys and Tutorials*, vol. PP, no. 99, pp. 1167–1178, Nov. 2012.

[26] A. MacKenzie and S. Wicker, “Stability of multipacket slotted aloha with selfish users and perfect information,” in *Proc. of Infocom*, 2003.

[27] E. Altman, R. El-Azouzi, and T. Jimenezy, “Slotted aloha as a game with partial information,” *Computer Networks*, vol. 45, no. 6, pp. 701–713, 2004.

[28] Y. Jin and G. Kesidis, “Equilibria of a noncooperative game for heterogeneous users of an aloha network,” *IEEE Communications Letters*, vol. 6, no. 7, pp. 282–284, 2002.

[29] G. Kesidis, Y. Jin, A. P. Azad, and E. Altman, “Stable nash equilibria of aloha medium access games under symmetric, socially altruistic behavior,” in *Proc. of CDC*, 2010.

[30] I. Menache and N. Shimkin, “Fixed-rate equilibrium in wireless collision channels,” *Network Control and Optimization (NET-COOP)*, pp. 23–32, 2007.

[31] Y. Jin, G. Kesidis, and J. W. Jang, “A channel aware mac protocol in an aloha network with selfish users,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 128–137, 2012.

[32] M. Cagalj, S. Ganeriwal, I. Aad, and J. P. Hubaux, “On smart cheating in csma/ca networks,” in *Proc. of Infocom*, 2005.

[33] Y. Jin and G. Kesidis, “Distributed contention window control for selfish users in ieee 802.11 wireless lans,” *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 6, pp. 1113–1123, 2007.

[34] J. W. Lee, A. Tang, J. Huang, M. Chiang, and A. Calderbank, “Reverse-engineering mac: a non-cooperative game model,” *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 1135–1147, 2007.

[35] J. W. Lee, M. Chiang, and R. A. Calderbank, “Utility-optimal medium access control: reverse and forward engineering,” in *Proc. of Infocom*, 2006.

[36] T. Cui, L. Chen, and S. Low, “A game-theoretic framework for medium access control,” *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 1116–1127, 2008.

[37] L. Chen, S. Low, and J. Doyle, “Random access game and medium access control design,” *IEEE/ACM Trans. Netw.*, vol. 18, no. 4, pp. 1303–1316, Aug. 2010.

[38] A. H. M. Rad, J. Huang, M. Chiang, and V. Wong, “Utility-optimal random access: Reduced complexity, fast convergence, and robust performance,” *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 898–911, 2009.

[39] A. H. M. Rad, J. Huang, M. Chiang, and V. W. S. Wong, “Utility-optimal random access without message passing,” *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1073–1079, 2009.

[40] N. Hegde and A. Proutiere, “Simulation-based optimization algorithms with applications to dynamic spectrum access,” in *Proc. of CISS*, 2012.

[41] L. Jiang, D. Shah, J. Shin, and J. Walrand, “Distributed random access algorithm: Scheduling and congestion control,” *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6182 –6207, dec. 2010.

[42] D. Qian, D. Zheng, J. Zhang, N. B. Shroff, and C. Joo, “Distributed csma algorithms for link scheduling in multihop mimo networks under sinr model,” *IEEE/ACM Transactions on Networking*, vol. 21, no. 3, pp. 746–759, 2013.

[43] L. Jiang and J. Walrand, “Approaching throughput-optimality in distributed csma scheduling algorithms with collisions,” *IEEE/ACM Trnasactions on Networking*, vol. 19, no. 3, pp. 816–829, 2011.

[44] T. H. Kim, J. Ni, R. Srikant, and N. H. Vaidya, “Throughput-optimal csma with imperfect carrier sensing,” *IEEE/ACM Trnasactions on Networking*, vol. 21, no. 5, pp. 1636–1650, 2013.

[45] X. Wang and K. Kar, “Throughput modelling and fairness issues in csma/ca based ad-hoc networks,” in *Proc. of Infocom*, 2005.

[46] S. C. Liew, C. H. Kai, H. C. Leung, and P. Wong, “Back-of-the-envelope computation of throughput distributions in csma wireless networks,” *IEEE Transactions on Mobile Computing*, vol. 9, no. 9, pp. 1319–1331, 2010.

[47] J. Mo and J. Walrand, “Fair end-to-end window-based congestion control,” *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 556–567, 2000.

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

[49] L. Chen, T. Cui, S. H. Low, and J. C. Doyle, “A game-theoretic model for medium access control,” in *Proc. of WICON*, 2007.

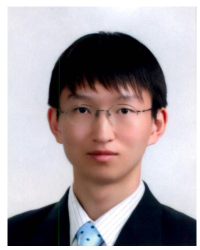
[50] S. D. Flam, “Equilibrium, evolutionary stability and gradient dynamics,” *International game Theory Review*, vol. 4, no. 4, pp. 357–370, 2002.

[51] O. Goussevskaia, V. A. Oswald, and R. Wattenhofer, “Complexity in geometric sinr,” in *Proc. of MobiHoc*, 2007.

[52] A. Proutiere, Y. Yi, T. Lan, and M. Chiang, “Resource allocation over network dynamics without timescale separation,” in *Proc. of Infocom*, 2010.

[53] V. S. Borkar, *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, 2008.

Hyeryung Jang Hyeryung Jang is currently a postdoctoral research at Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, from March 2017. She received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from KAIST in 2010, 2012, and 2017, respectively. Her research interests include network economics, game theory, analysis of communication systems, and machine learning, especially applying learning of graphical models to communication systems.



Se-Young Yun Se-Young Yun is an Assistant Professor in the Department of Industrial and Systems Engineering at KAIST, Republic of Korea. He was a visiting researcher at Microsoft Research, Cambridge, UK (hosted by Prof. Milan Vojnovic) and a postdoctoral researcher at Los Alamos National Lab, NM, USA (hosted by Dr. Michael Chertkov), Microsoft Research-INRIA Joint center, Paris, France (hosted by Dr. Marc Lelarge), and KTH, Stockholm, Sweden (hosted by Prof. Alexandre Proutiere). He received the B.S. and Ph.D in electrical engineering from the KAIST, Daejeon, Republic of Korea, in 2006 and 2012 under the supervision of Prof. Dong-Ho Cho. From 2012 to 2013, he was a postdoc researcher at the same department (hosted by Prof. Yung Yi). He received the best paper award in ACM MOBIHOC 2013 and NIPS outstanding reviewer award 2016.

Jinswoo Shin Jinwoo Shin is currently an Assistant Professor at the Department of Electrical Engineering at KAIST, South Korea. His major research interest is to develop advanced algorithms for large-scale data and network analytics. He obtained the Ph.D. degree from Massachusetts Institute of Technology in 2010 and B.S. degrees (in Math and CS) from Seoul National University in 2001. After spending two years (20102012) at Algorithms & Randomness Center, Georgia Institute of Technology, one year (20122013) at Business Analytics and Mathematical Sciences Department, IBM T. J. Watson Research, he started teaching at KAIST in Fall 2013. He received the Kennneth C. Sevcik Award at SIGMETRICS 2009, George M. Sprowls Award 2010, Best Paper Award at MobiHoc 2013, Best Publication Award from INFORMS Applied Probability Society 2013, Bloomberg Scientific Research Award 2015 and ACM SIGMETRICS Rising Star Award 2015.

Yung Yi Yung Yi received the B.S. and M.S. degrees from the School of Computer Science and Engineering, Seoul National University, South Korea, in 1997 and 1999, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Texas at Austin, in 2006. From 2006 to 2008, he was a Post-Doctoral Research Associate with the Department of Electrical Engineering, Princeton University. He is currently an Associate Professor with the Department of Electrical Engineering, KAIST, South Korea. His current research interests include the design and analysis of computer networking and wireless communication systems, especially congestion control, scheduling, and interference management, with applications in wireless ad hoc networks, broadband access networks, economic aspects of communication networks, and green networking systems. He received the best paper awards at the IEEE SECON 2013, the ACM MOBIHOC 2013, and the IEEE William R. Bennett Award 2016.