Medium Access using Queues

Devavrat Shah
Department of EECS
MIT
Cambridge, USA
Email: devavrat@mit.edu

Jinwoo Shin
School of CS
Georgia Tech
Atlanta, USA
Email: jshin72@cc.gatech.edu

Prasad Tetali*
School of Math and School of CS
Georgia Tech
Atlanta, USA
Email: tetali@math.gatech.edu

Abstract— Consider a wireless network of n nodes represented by a (undirected) graph G where an edge (i, j) models the fact that transmissions of i and j *interfere* with each other, i.e. simultaneous transmissions of i and j become unsuccessful. Hence it is required that at each time instance a set of non-interfering nodes (corresponding to an independent set in G) access the wireless medium. To utilize wireless resources efficiently, it is required to arbitrate the access of medium among interfering nodes properly. Moreover, to be of practical use, such a mechanism is required to be totally distributed as well as simple.

As the main result of this paper, we provide such a medium access algorithm. It is randomized, totally distributed and simple: each node attempts to access medium at each time with probability that is a function of its local information. We establish efficiency of the algorithm by showing that the corresponding network Markov chain is positive recurrent as long as the demand imposed on the network can be supported by the wireless network (using *any* algorithm). In that sense, the proposed algorithm is optimal in terms of utilizing wireless resources. The algorithm is oblivious to the network graph structure, in contrast with the so-called *polynomial back-off* algorithm by Hastad-Leighton-Rogoff (STOC '87, SICOMP '96) that is established to be optimal for the complete graph and bipartite graphs (by Goldberg-MacKenzie (SODA '96, JCSS '99)).

Keywords-Wireless Medium Access, Stability, Mixing Time

1. Introduction

We consider a *single-hop* wireless network of n nodes or queues represented by $V = \{1, \dots, n\}$. Time is discrete indexed by $\tau \in \{0, 1, \dots\}$. Unit-size packets arrive at queue i as per an exogenous process. Let $A_i(\tau)$ denote the number of packets arriving at queue i in the time slot $[\tau, \tau + 1)$. For simplicity, we shall assume $A_i(\cdot)$ as an independent Bernoulli process with rate λ_i, i.e. $\lambda_i = \mathbb{P}(A_i(\tau) = 1)$ and $A_i(\tau) \in \{0, 1\}$ for all $i, \tau \geq 0$.[1] Let $Q_i(\tau) \in \mathbb{N}$ be the number of packets in queue i at the beginning of time slot $[\tau, \tau + 1)$.

The work from queues is served at the unit rate subject to *interference* constraints. Specifically, let $G = (V, E)$ denote the inference graph : $(i, j) \in E$ implies that queues i and j can not transmit simultaneously since their transmissions *interfere* with each other. Formally, let $\sigma_i(\tau) \in \{0, 1\}$ denote whether the queue i is (successfully) transmitting at time τ, and let $\boldsymbol{\sigma}(\tau) = [\sigma_i(\tau)]$. Then, for $\tau \geq 0$,

$$\boldsymbol{\sigma}(\tau) \in \mathcal{I}(G) \triangleq \{\boldsymbol{\rho} = [\rho_i] \in \{0,1\}^n : \rho_i + \rho_j \leq 1 \text{ for all } (i,j) \in E\},$$

i.e. $\mathcal{I}(G)$ is the set of independent sets of G. The resulting queueing dynamics can be summarized as

$$Q_i(\tau + 1) = Q_i(\tau) - \sigma_i(\tau)\mathbf{I}_{\{Q_i(\tau)>0\}} + A_i(\tau),$$

for $\tau \geq 0$ and $1 \leq i \leq n$. Here $\mathbf{I}_{\{x\}} = 1$ if $x =$'true' and 0 if $x =$'false'.

Now an algorithm, which we shall call *medium access*, is required to choose $\boldsymbol{\sigma}(\tau) \in \mathcal{I}(G)$ at the beginning of each time τ. A good medium access algorithm should choose $\boldsymbol{\sigma}(\tau)$ so as to utilize the wireless medium as efficiently as possible. Putting it another way, it should be able to keep queues finite for as large a set of arrival rates $\boldsymbol{\lambda} = [\lambda_i]$ as possible. Towards this, define the capacity region

$$\boldsymbol{\Lambda} \triangleq \Big\{\boldsymbol{y} \in \mathbb{R}^n_+ : \boldsymbol{y} < \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \alpha_{\boldsymbol{\sigma}} \boldsymbol{\sigma} \text{ with } \alpha_{\boldsymbol{\sigma}} \geq 0, \sum_{\boldsymbol{\sigma} \in \mathcal{I}(G)} \alpha_{\boldsymbol{\sigma}} \leq 1\Big\}.$$

Since $\boldsymbol{\sigma}(\tau) \in \mathcal{I}(G)$, the effective 'service' rate induced by any algorithm over time is essentially in the closure of $\boldsymbol{\Lambda}$. Therefore, a medium access algorithm can be considered optimal, if it can keep queues *finite, for any* $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$. Formally, if the state of the queueing system including the algorithm's decisions and queue-sizes can be described as a Markov chain, then the existence of a stationary distribution for this Markov chain and its ergodicity effectively implies that the queues remain finite. A sufficient condition for this is *aperiodicity* and *positive recurrence* of the corresponding Markov chain. This motivates the following definition.

Definition 1 (Optimal) A medium access algorithm is called *optimal* if for any $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$ the (appropriately defined) underlying network Markov chain is *positive recurrent* and *aperiodic.*

*Work of D. Shah and J. Shin supported in parts by the AFOSR complex networks grant and DARPA ITMANET grant. Research of P. Tetali supported in part by NSF grant CCF-0910584.

[1] The result in this paper extends easily even for (non-Bernoulli) adversarial arrival processes satisfying $\sum_{\tau=s}^{t-1} A_i(\tau) \leq \lambda_i(t-s) + w$, with (fixed) $w < \infty$, for all $0 \leq s < t$.

To be of practical use, medium access algorithm ought to be simple and totally distributed, i.e. should use only local information like queue-size, and past collision history. In such an algorithm, each node makes the decision to transmit or not on its own, at the beginning of each time slot. At the end of the time slot, it learns whether attempted transmission was successful or not (due to a simultaneous attempt of transmission by a neighbor). If a node does not transmit, then it learns whether any of its neighbors attempted transmission. And, ideally such an algorithm should be optimal.

1.1. Prior Work

Design of an efficient and distributed medium access algorithm has been of interest since the ALOHA algorithm for the radio network [1] and Local Area Networks [16] in the 1970s. Subsequently a variety of the so-called back-off algorithms or protocols have been extensively studied. Various negative and positive properties of back-off protocols were established in various works [13], [18], [14], [25], [2], [15].

Specifically, Hastad, Leighton and Rogoff [9] studied a medium access algorithm in which each node or queue attempts transmission at each time with probability that is inversely proportional to a polynomial function of the number of successive failures in the most recent past. They established it to be optimal when the interference graph G is complete, or equivalently all nodes are competing for one resource (as in the classical Ethernet/LAN). The optimality of this polynomial back-off algorithm was further established for G when it is induced by matching constraints in a bipartite graph by Goldberg and MacKenzie [7]. However, the optimality of polynomial back-off or any other totally distributed medium access algorithm remained open for general graphs. The interested reader may find a good summary of results until 2001, on medium access on a webpage maintained by Goldberg [8].

In the past year or so, significant progress has been made towards this open question. Specifically, Rajagopalan, Shah and Shin (RSS) [19], [20] and Jiang and Walrand (JW) [11] proposed two different medium access algorithms that operate in continuous time assuming immediate collision detections or the so called *perfect carrier sense* information. The RSS algorithm is optimal but requires a bit of information exchange between each pair of neighboring nodes per unit time. The JW algorithm was established to have a weaker form of optimality, called 'rate stability', by Jiang, Shah, Shin and Walrand [10]. The perfect carrier sense information, utilized in JW and RSS algorithms, is not available in practice in the context of wireless networks.

In summary, both algorithms stop short of being totally distributed and optimal. Further, both of them operate in continuous time (with immediate collision detections or perfect carrier sense) and thus effectively avoiding the issue of loss in performance due to contention present in discrete time (with delayed collision detections) considered in this paper. We take a note of a recent work by Jiang and Walrand that extends JW to the setting of this work [12].

1.2. Our Contribution

The main result of this paper is a totally distributed medium access algorithm that is optimal for any interference graph G. It resolves an important question in distributed computation that has been of great practical interest. The proposed medium access algorithm builds on the RSS algorithm and overcomes its two key limitations by adapting it to the discrete time and removing the need for any information exchange between neighboring nodes. In what follows, we explain in detail how we overcome such limitations.

In the proposed medium access algorithm, each node attempts transmission in each time slot based on: (a) whether it managed to successfully transmit in the previous time slot, or whether any of its neighbors attempted to transmit in the previous time slot; (b) local queue-size and estimation of the "weight" of the neighbors. Given this information, each node in the beginning of each time slot attempts transmission with probability depending upon (a) and (b). Specifically, if the node was successful in the previous time, it does not transmit in this time with probability that is inversely proportional to its own weight that depends on (b). Else if no other neighbor attempted transmission in the previous time then a node attempts transmission with probability $\frac{1}{2}$. Otherwise, with probability 1, a node does not transmit.

In such an algorithm, the only seeming non-local information is the estimation of the neighbors' weight in (b). An important contribution of this work is the design of a non-trivial learning mechanism, based only on information of type (a), that estimates the neighbors' weight without any explicit information exchange. We note that, in contrast, the RSS algorithm had required explicit information exchange for estimating the neighbors' weight.

To establish optimality of the proposed algorithm, we show that, in essence, the value of $\sum_i \sigma_i(\tau) \log\log Q_i(\tau)$ is *close* to $\max_{\boldsymbol{\rho} \in \mathcal{I}(G)} \sum_i \rho_i \log\log Q_i(\tau)$, on average for all large enough τ. That is, effectively the distributed medium access chooses $\sigma(\tau)$ that is (close to) maximum weight independent set of G when node weights are equal to $\log\log$ of the queue-sizes. Such a property is known (cf. [22], [24]) to imply that $\sum_i F(Q_i(\tau))$ (where $F(x) = \int_0^x \log\log y \; dy$) is a potential (or Lyapunov, energy) function for the system state so that the function is expected to decrease by at least a fixed amount as long as $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$. This subsequently establishes that the network as a Markov chain is positive recurrent (implying the optimality of the algorithm).

We establish the near optimality of $\sum_i \sigma_i(\tau) \log\log Q_i(\tau)$ under the medium access algorithm in two steps. To begin with, we observe that the evolution of $\boldsymbol{\sigma}(\tau)$ under the algorithm is a Markov chain on the space of independent sets $\mathcal{I}(G)$ with time-varying transition probabilities. For this Markov chain, at any particular time τ, let $\boldsymbol{\pi}(\tau)$ be the stationary distribution at time τ (given transition probabilities at time τ).

In the first step, we study this (time-varying, stationary) distribution $\boldsymbol{\pi}(\tau)$ and show that it is approximately 'product-form'. To obtain such an approximate characterization, we show that the transition probabilities of the Markov chain are well approximated by those of a reversible Markov chain which has a product-form stationary distribution. A novel comparison relation between stationary distributions of two Markov chains in terms of the relation between their transition probabilities leads to the approximate product-form characterization of $\boldsymbol{\pi}(\tau)$. We note that the RSS algorithm (and similarly, the JW algorithm) had used the continuous time setting to make sure that the corresponding Markov chain was reversible and hence had a product-form distribution to start with; such reversibility is lost in general in the discrete time setting of this paper. Using this approximate product-form characterization of $\boldsymbol{\pi}(\tau)$ in addition to the Gibbs' maximal principle (cf. [6]), we prove that $\boldsymbol{\pi}(\tau)$ has the desired property; namely, that $\sum_i \sigma_i \log\log Q_i(\tau)$ is close to $\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i \rho_i \log\log Q_i(\tau)$ if $\boldsymbol{\sigma} = [\sigma_i]$ is given by the distribution $\boldsymbol{\pi}(\tau)$. We call this the maximum-weight property at stationarity.

In the second step, we show that the Markov chain, despite it being time-varying, is always near stationarity for large enough τ by carefully estimating the effective mixing time of the time-varying Markov chain. In other words, the distribution of $\boldsymbol{\sigma}(\tau)$ is close to $\boldsymbol{\pi}(\tau)$ for large enough τ. Therefore, the maximum-weight property at stationarity (established in the first step) implies that $\sum_i \sigma_i(\tau) \log\log Q_i(\tau)$ is close to $\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i \rho_i \log\log Q_i(\tau)$. To guarantee the near stationarity property as a consequence of such a mixing analysis, it is required that a design of 'weight' maintained by each node in the medium access algorithm utilizes the neighbor's weight information. As mentioned earlier, we resolve this by developing a learning mechanism that estimates the neighbor's weight based on the information whether it transmitted or not thus far. The success in this second step is primarily due to our novel design of the learning mechanism incorporated well with the mixing time analysis of the time-varying Markov chain.

1.3. Organization

Remainder of the paper is organized as follows. Section 2 presents formally the medium access algorithm and a statement of the main result. Section 3 presents necessary technical preliminaries that are useful for establishing the results. Section 4 provides a high-level summary of the proof of the main result. The detailed proof is omitted from this extended abstract due to space constraints. An interested reader can find them in the full version of this paper [21]. Section 5 presents a generic result that compares stationary distributions of two Markov chains based on the relation between their transition probabilities. This result, utilized crucially in the proof of the main result, could be of broad interest in its own right. For example, it naturally suggests a notion of approximate product-form distributions. In Section 6, we discuss about high-level contributions of this paper.

2. Algorithm and Its Optimality

The medium access algorithm is randomized, distributed, simple and runs in discrete time with time indexed by $\tau \geq 0$. Each node $i \in V$ maintains weight $W_i(\tau) \in \mathbb{R}_+$ over $\tau \geq 0$. In the beginning of each time slot $\tau \geq 0$, each node $i \in V$ decides to attempt transmission or not as follows:

1. If the transmission of node i was successful at $\tau - 1$, then

   it attempts to transmit with probability $1 - \frac{1}{W_i(\tau)}$.

2. Else if no neighbor of i attempted transmission at $\tau - 1$, then

   it attempts to transmit with probability $\frac{1}{2}$.

3. Otherwise, it does not attempt to transmit with probability 1.

Now we describe how each node i maintains weight $W_i(\tau)$:

$$W_i(\tau) \ = \ \max\left\{\log Q_i(\tau), \max_{j\in\mathcal{N}(i)} \exp\left(\sqrt{\log g(A_j^i(\tau))}\right)\right\}, \tag{1}$$

where by $\log$ and $\log\log$ we mean $[\log]_+$ and $[\log\log]_+$ respectively; $g : \mathbb{R}_+ \to \mathbb{R}_+$ is defined as $g(x) = \exp(\log\log^4 x)$; by $\log\log^4 x$ we mean $(\log\log x)^4$, log represents $\log_e$; and $\mathcal{N}(i) = \{j \in V : (i,j) \in E\}$ represents neighbors of node i. Note that $W_i(\tau) \geq 1$ for all τ by definition. In above, $A_j^i(\cdot)$ is a 'counter' maintained by node i as a 'long term' estimate of weight $W_j(\cdot)$. This is maintained along with another 'counter' $B_j^i(\cdot)$ by node i as a 'short term' estimate of $W_j(\cdot)$. Initially, $A_j^i(0) = B_j^i(0) = 0$ for all $j \in \mathcal{N}(i)$ and $i \in V$. For each $j \in \mathcal{N}(i)$, $A_j^i(\cdot)$ and $B_j^i(\cdot)$ are updated by node i at τ as follows:

1. If $j \in \mathcal{N}(i)$ attempted transmission at $\tau - 1$, then

   $$A_j^i(\tau) = A_j^i(\tau-1) \ \text{ and } \ B_j^i(\tau) = B_j^i(\tau-1) + 1.$$

2. Else if $B_j^i(\tau - 1) \geq 2$, then

   $$A_j^i(\tau) = \begin{cases} A_j^i(\tau-1) + 1 & \text{if } \ B_j^i(\tau-1) \geq g(A_j^i(\tau-1)) \\ A_j^i(\tau-1) - 1 & \text{otherwise} \end{cases}$$

   and reset $B_j^i(\tau) = 0$.

3. Otherwise, $A_j^i(\tau) = A_j^i(\tau-1)$ and $B_j^i(\tau) = 0$.

Observe that $B_j^i(\cdot)$ counts how long neighbor j keeps attempting transmission consecutively. When j's transmissions are successful, the random period of consecutive transmissions is essentially distributed as per the geometric distribution with mean $W_j(\cdot)$ due to the nature of our algorithm. Thus $B^i_j(\cdot)$ provides a short-term (or instantaneous) estimation of $W_j(\cdot)$. To extract a robust estimation of $W_j(\cdot)$ from such short-term estimates, the long-term estimation $A^i_j(\cdot)$ is maintained: it changes by ± 1 using $B^i_j(\cdot)$ at most once per unit time. Specifically, as per the above updates, $g(A^i_j(\cdot))$ is acting as an estimation of $W_j(\cdot)$. Now it is important to note that the choice of g (defined above) plays a crucial role in the quality of estimate of $W_j(\cdot)$. The change in estimation $g(A^i_j(\cdot))$, when $A^i_j(\cdot)$ is updated by ± 1, is roughly $g'(A^i_j(\cdot))$. Since $W_j(\cdot)$ is changing over time, it is important to have $g'(\cdot)$ not too small. On the other hand, if it is too large then it is too sensitive and could be noisy just like $B^i_j(\cdot)$. A priori it is not clear if there exists a choice of any function g that allows for keeping $A^i_j(\cdot)$ as a good enough estimation of $W^i_j(\cdot)$, which subsequently leads to positive-recurrence of the network Markov chain. Somewhat surprisingly (at least to us), we find that indeed such a g exists and is as defined above: $g(x) = \exp(\log\log^4 x)$. As per our proof technique, $g(x) = \exp(\log\log^\alpha x)$ works for any $\alpha > 2$; however we shall stick to the choice of $\alpha = 4$ in the paper. Section 4 provides the reasons on why such a choice of function g is necessary and sufficient. Now we state the main result of this paper.

Theorem 1 *The medium access algorithm as described above is optimal for any interference graph.*

3. Technical Preliminaries: A Markov chain and Criteria for positive recurrence

This section provides useful technical prerliminaries for establishing Theorem 1. We start by describing an associated finite state Markov chain, characterization of its stationary distribution, bounds for mixing time and a criteria for establishing positive recurrence of a countable state-space Markov chain.

3.1. A Markov Chain (MC) of Interest

We describe a Markov chain over finite state space, whose time-varying version will describe the evolution of the medium access algorithm described in Section 2. As we described in Section 4, our strategy for proving Theorem 1 crucially relies on understanding the stationary distribution and mixing time of the (finite state) Markov chain.

Description. The Markov chain evolves on state space $\mathcal{I}(G) \times \{0,1\}^n$ and uses node weights $\boldsymbol{W} = [W_i] \in \mathbb{R}^n_+$ with $\boldsymbol{W}_{\min} \geq 1$. Given $(\boldsymbol{\sigma}, \boldsymbol{a}) \in \mathcal{I}(G) \times \{0,1\}^n$, the next (random) state $(\boldsymbol{\sigma}', \boldsymbol{a}') \in \mathcal{I}(G) \times \{0,1\}^n$ is obtained as follows:

1. Each node i chooses $r_i \in \{0,1\}$ uniformly at random, i.e. $r_i = 1$ with probability $1/2$ and 0 otherwise. Temporarily set
$$a'_i = \begin{cases} r_i & \text{if } a_j = 0 \text{ for all } j \in \mathcal{N}(i) \\ 0 & \text{otherwise} \end{cases}.$$
2. Each node i sets σ'_i (and possibly resets a'_i) as follows:
   - If $\sigma_i = 1$, then set
   $$(\sigma'_i, a'_i) = \begin{cases} (0,0) & \text{with probability } \frac{1}{W_i}, \\ (1,1) & \text{otherwise.} \end{cases}$$
   - Else if $a_j = 0$ for all $j \in \mathcal{N}(i)$, then set
   $$\sigma'_i = \begin{cases} 1 & \text{if } a'_i = 1 \text{ and } a'_j = 0 \text{ for all } j \in \mathcal{N}(i) \\ 0 & \text{otherwise} \end{cases}.$$
   - Otherwise, set $(\sigma'_i, a'_i) = (0,0)$.

Stationary distribution. Let $\Omega = \mathcal{I}(G) \times \{0,1\}^n$. Then Ω is the state space of the above described Markov chain; let $P_{\mathbf{x}\mathbf{x}'}$ denote its transition probability for $\mathbf{x} = (\boldsymbol{\sigma}, \boldsymbol{a}),\ \mathbf{x}' = (\boldsymbol{\sigma}', \boldsymbol{a}') \in \Omega$. We characterize the stationary distribution of this Markov chain as follows.

Lemma 2 *Staring from initial state $(\mathbf{0}, \mathbf{0})$, the Markov chain P is recurrent and aperiodic; let its recurrence class be denoted by $\Omega' \subset \Omega$; $(\boldsymbol{\sigma}, \mathbf{0}) \in \Omega'$ for all $\boldsymbol{\sigma} \in \mathcal{I}(G)$. Therefore, the Markov chain P has a unique stationary distribution $\boldsymbol{\pi}$ on Ω' such that for any $(\boldsymbol{\sigma}, \boldsymbol{a}) \in \Omega'$*
$$\pi_{(\boldsymbol{\sigma}, \boldsymbol{a})} \propto \exp\big(\boldsymbol{\sigma} \cdot \log \boldsymbol{W} + U(\boldsymbol{\sigma}, \boldsymbol{a})\big), \tag{2}$$
where $U : \Omega' \to \mathbb{R}_+$ is such that $|U(\boldsymbol{\sigma}, \boldsymbol{a})| \leq n4^n \log 2$ for all $(\boldsymbol{\sigma}, \boldsymbol{a}) \in \Omega'$.

To achieve the form (2), we use the classical Markov chain tree theorem [3]. Our proof strategy can be of broad interest to characterize such a form for non-reversible Markov chains via comparing reversible Markov chains. The proof of Lemma 2 is presented in the full version of this paper [21]. See Section 5 for a generic version of this comparison result.

Mixing time. Now we establish a bound on the 'mixing time' of P – the time to reach near stationary distribution starting from any initial distribution. We shall use the total-variation distance: given distributions $\boldsymbol{\nu}, \boldsymbol{\mu}$ on a finite state space Ω', define $\|\boldsymbol{\nu} - \boldsymbol{\mu}\|_{TV} = \sum_{\mathbf{x} \in \Omega'} |\nu_{\mathbf{x}} - \mu_{\mathbf{x}}|$.

Lemma 3 *Given $\varepsilon \in (0, 0.5)$ with $n \geq 2$, for any distribution $\boldsymbol{\mu}$ on Ω',*
$$\|\boldsymbol{\mu} P^\tau - \boldsymbol{\pi}\|_{TV} < \varepsilon,$$
for all $\tau \geq T_{mix}(\varepsilon, n, \boldsymbol{W}_{\max})$, where
$$\begin{aligned} T_{mix} &\equiv T_{mix}(\varepsilon, n, \boldsymbol{W}_{\max}) \\ &= 4^{n4^{n+1}+1} \boldsymbol{W}^{6n}_{\max} \ \log\Big(\frac{4^{n4^n} \boldsymbol{W}^n_{\max}}{2\varepsilon}\Big). \end{aligned} \tag{3}$$

We use the Cheeger's inequality [4], [23] to achieve the mixing bound (3). The proof of Lemma 3 is presented in the full version of this paper [21].

3.2. *Ergodicity, Positive recurrence and Lyapunov-Foster*

To establish optimality of the medium access algorithm, we need to show that the underlying network Markov chain, which has countably infinite state space, is ergodic, i.e. that it has the unique stationary distribution to which it converges. We briefly recall known methods from literature that will be helpful in doing so.

Consider a discrete time Markov chain $X(\cdot)$ on countably infinite state space X. State $\mathsf{x} \in \mathsf{X}$ is said to be *recurrent* if $\mathbb{P}(T_{\mathsf{x}} = \infty) = 0$, where $T_{\mathsf{x}} = \inf\{\tau \geq 1 : X(\tau) = \mathsf{x} : X(0) = \mathsf{x}\}$. Specifically, a recurrent state x is called *positive recurrent* $\mathbb{E}[T_{\mathsf{x}}] < \infty$, or else if $\mathbb{E}[T_{\mathsf{x}}] = \infty$ then it is called *null recurrent*. For an irreducible Markov chain, if one of its state is positive recurrent, the so are all; we call such a Markov chain positive recurrent. An irreducible, aperiodic and positive recurrent Markov chain is known to be ergodic: it has unique stationary and starting from any initial distribution, it converges (in distribution) to stationary distribution. Therefore, it is sufficient to establish positive recurrence property for establishing ergodicity of the Markov chain in addition to verifying irreducibility and aperiodicity properties. We shall recall a sufficient condition for establishing positive recurrence, known as the Lyapunov and Foster's criteria.

Lyapunov and Foster's criteria. This criteria utilizes existence of a "Lyapunov", "Potential" or "Energy" function of the state under evolution of the Markov chain. Specifically, consider a non-negative valued function $L : \mathsf{X} \to \mathbb{R}_+$ such that $\sup_{\mathsf{x}\in\mathsf{X}} L(\mathsf{x}) = \infty$. Let $h : \mathsf{X} \to \mathbb{Z}_+$ be another function that is to be interpreted as a state dependent "stopping time". The 'drift' in Lyapunov function L in h-steps starting from $\mathsf{x} \in \mathsf{X}$ is defined as

$$\mathbb{E}[L(X(h(\mathsf{x}))) - L(X(0)) \mid X(0) = \mathsf{x}].$$

Following is the criteria (see [5], for example):

Theorem 4 *For any $\kappa > 0$, let $B_\kappa = \{\mathsf{x} : L(\mathsf{x}) \leq \kappa\}$. Suppose there exist functions $h, k : \mathsf{X} \to \mathbb{Z}_+$ such that for any $\mathsf{x} \in \mathsf{X}$,*

$$\mathbb{E}\left[L(X(h(\mathsf{x}))) - L(X(0)) \mid X(0) = \mathsf{x}\right] \leq -k(\mathsf{x}),$$

that satisfy the following conditions:

(L1) $\inf_{\mathsf{x}\in\mathsf{X}} k(\mathsf{x}) > -\infty$.
(L2) $\liminf_{L(\mathsf{x})\to\infty} k(\mathsf{x}) > 0$.
(L3) $\sup_{L(\mathsf{x})\leq\gamma} h(\mathsf{x}) < \infty$ *for all* $\gamma > 0$.
(L4) $\limsup_{L(\mathsf{x})\to\infty} h(\mathsf{x})/k(\mathsf{x}) < \infty$.

Then, there exists constant $\kappa_0 > 0$ so that for all $\kappa_0 < \kappa$, the following holds:

$$\begin{aligned} \mathbb{E}\left[T_{B_\kappa} \mid X(0) = \mathsf{x}\right] &< \infty, \qquad \textit{for any } \mathsf{x} \in \mathsf{X} \\ \sup_{\mathsf{x}\in B_\kappa} \mathbb{E}\left[T_{B_\kappa} \mid X(0) = \mathsf{x}\right] &< \infty, \end{aligned}$$

where $T_{B_\kappa} := \inf\{\tau \geq 1 : X(\tau) \in B_\kappa\}$ i.e. the first return time to B_κ. In other words, B_κ is positive recurrent.

Theorem 4 implies that if *(L1) - (L4)* are satisfied and B_κ is a finite set, the Markov chain is positive recurrent.

4. Proof of Theorem 1: An Overview

This section provides an overview of the proof of Theorem 1 to explain the key challenges involved in establishing it as well as intuition behind the particular choice of function g. The goal in this section is not to provide precise arguments but only provide intuition so as to assist a reader in understanding the structure of the proof. The complete proof with all details is stated in the full version of this paper [21].

Theorem 1 requires establishing positive recurrence of an appropriate Markov chain that describes the evolution of the network state under the medium access algorithm described (as long $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$). To that end, define

$$X(\tau) = (\mathbf{Q}(\tau), \boldsymbol{\sigma}(\tau), \boldsymbol{a}(\tau), \boldsymbol{A}(\tau), \mathbf{B}(\tau))$$

where $\mathbf{Q}(\tau)$ represents vector of queue-sizes; $\boldsymbol{a}(\tau) \in \{0,1\}^n$ denotes the vector of transmission attempts by nodes at time τ; $\boldsymbol{\sigma}(\tau) \in \mathcal{I}(G)$ denotes the vector of resulting successful transmissions in time τ (clearly, $\boldsymbol{\sigma}(\tau) \leq \boldsymbol{a}(\tau)$); and $\boldsymbol{A}(\tau),\ \mathbf{B}(\tau) \in \mathbb{Z}_+^{2|E|}$ denote the vector of long-term and short-term estimations maintained at nodes as explained in Section 2. Then it follows that under medium access algorithm $X(\cdot)$ is a Markov chain. It can be easily checked that under this Markov chain, state $\mathbf{0}$ in which all components are 0, has positive probability of transiting to itself. Further, starting from any state, $X(\cdot)$ has positive probability of reaching state $\mathbf{0}$. Therefore, $X(\cdot)$ is always restricted to the recurrence class containing state $\mathbf{0}$; and over this class it is aperiodic. Therefore, it is sufficient to establish positive recurrence of $X(\cdot)$ over this recurrence class to imply that $X(\cdot)$ is ergodic.

Now as discussed in Section 3, a generic method to establish positive-recurrence of a Markov chain involves establishing that certain real-valued function over the state-space of the Markov chain is Lyapunov or Potential function for the Markov chain. Roughly speaking, this involves establishing that on average the value of this function decreases under the dynamics of the Markov chain if its value is high enough; Theorem 4 states the precise conditions that need to be verified. With this eventual goal, we consider the following function that maps state $\mathsf{x} = (\mathbf{Q}, \boldsymbol{\sigma}, \boldsymbol{a}, \boldsymbol{A}, \mathbf{B})$ to non-negative real values as

$$L(\mathsf{x}) = \sum_i F(Q_i) + \sum_{i;j\in\mathcal{N}(i)} \Big((A_j^i)^2 + g^{(-1)}(B_j^i)\Big), \quad (4)$$

where $F(x) = \int_0^x \log\log y \; dy$ with $\log\log y = [\log\log y]_+$; the inverse function of $g(x) = \exp(\log\log^4 x)$ is $g^{(-1)}(x) = \exp(\exp(\log^{1/4} x))$. With abuse of notation, we shall use $L(\tau)$ to denote $L(X(\tau))$. Now

$$L(\tau) \quad = \quad L^Q(\tau) + L^{A,B}(\tau),$$

where

$$L^Q(\tau) = \sum_i F(Q_i(\tau))$$
$$L^{A,B}(\tau) = \sum_{i;j\in\mathcal{N}(i)} \Big((A^i_j(\tau))^2 + g^{(-1)}(B^i_j(\tau)) \Big).$$

The proof is devoted to establish the *negative-drift* property of $L(\cdot)$, i.e. if $X(\tau) = \mathsf{x}$ is such that $L(\tau)$ is *large enough* (i.e. larger than some finite constant), then value of $L(\cdot)$ decreases enough on average. This property is established by considering two separate cases.

Case One. When $L(\tau)$ is large due to the component $L^{A,B}(\tau)$ being very large. Formally, when

$$\max_{i,j} \big(g(A^i_j(\tau)), B^i_j(\tau) \big) \;\geq\; \boldsymbol{W}^3_{\max}(\tau),$$

where $\boldsymbol{W}_{\max}(\tau) = \max_i W_i(\tau)$.

Case Two. When $L(\tau)$ is large due to the component $L^Q(\tau)$ being very large. Formally, when

$$\max_{i,j} \big(g(A^i_j(\tau)), B^i_j(\tau) \big) \;<\; \boldsymbol{W}^3_{\max}(\tau),$$

where $\boldsymbol{W}_{\max}(\tau) = \max_i W_i(\tau)$.

The above claim is formalized next. Recall that node weights $\boldsymbol{W}$ are determined by $\mathbf{Q}$ and $\boldsymbol{A}$ as per (1). Therefore, given state $\mathsf{x} = (\mathbf{Q}, \boldsymbol{\sigma}, \boldsymbol{a}, \boldsymbol{A}, \mathbf{B})$, the weight vector $\boldsymbol{W}$ is determined. With this in mind, let

$$C(\mathsf{x}) = \max \big\{ g(\boldsymbol{A}_{\max}),\ \mathbf{B}_{\max} \big\}, \tag{5}$$

where $\boldsymbol{A}_{\max} = \max_{i,j} A^i_j$ and $\mathbf{B}_{\max} = \max_{i,j} B^i_j$. Then h and k are defined (for two different cases) as

$$h(\mathsf{x}) = \begin{cases} C(\mathsf{x})^n & \text{if } C(\mathsf{x}) \geq \boldsymbol{W}^3_{\max}, \\ \frac{1}{2}\exp\big(\exp(\sqrt{\log \boldsymbol{W}_{\max}})\big) & \text{otherwise.} \end{cases} \tag{6}$$

$$k(\mathsf{x}) = \begin{cases} C(\mathsf{x})^{2n} & \text{if } C(\mathsf{x}) \geq \boldsymbol{W}^3_{\max}, \\ h(\mathsf{x})\sqrt{\log \boldsymbol{W}_{\max}} & \text{otherwise.} \end{cases} \tag{7}$$

With these definitions, we shall establish the following.

Lemma 5 *Let $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$. Then for any x with $L(\mathsf{x})$ large enough,*

$$\mathbb{E}\big[L(h(\mathsf{x})) - L(0) \,|\, X(0) = \mathsf{x}\big] \leq -k(\mathsf{x}). \tag{8}$$

It can be easily checked that L, h and k along with Lemma 5 satisfy conditions of Theorem 4. Now $L(\mathsf{x}) \to \infty$ as $|\mathsf{x}| \to \infty$ where $|\mathsf{x}| = |\mathbf{Q}| + |\boldsymbol{\sigma}| + |\boldsymbol{a}| + |\boldsymbol{A}| + |\mathbf{B}|$ with $|\boldsymbol{\sigma}|$, $|\boldsymbol{a}|$ being equal to the ordering of them and $|\mathbf{Q}|, |\boldsymbol{A}|$ and $|\mathbf{B}|$ are standard 1-norm. Therefore, $B_\kappa = \{\mathsf{x} : L(\mathsf{x}) \leq \kappa\}$ is a finite set. Therefore, it follows that the Markov chain $X(\cdot)$ is positive recurrent; it is aperiodic and irreducible on the recurrence class containing $\mathbf{0}$ as discussed before. Therefore, it follows that it is ergodic. That is medium access algorithm of interest is optimal establishing Theorem 1. In the remainder this section, we shall provide intuitive explanation of how Lemma 5 can be established. The preicse details can be found in the full version of this paper [21]. As mentioned above, the proof is argued for two separate cases: in the first csae, x with $C(\mathsf{x}) \geq \boldsymbol{W}^3_{\max}$, and in the second case, $C(\mathsf{x}) < \boldsymbol{W}^3_{\max}$.

Case One: $C(\mathsf{x}) \geq \boldsymbol{W}^3_{\max}$ In this case, there exists $i \in V$ and $j \in \mathcal{N}(i)$ so that $g(A^i_j(\tau))$ or $B^i_j(\tau)$ is larger than $\boldsymbol{W}^3_{\max}(\tau)$. Using the property of the estimation procedure (which updates $A^i_j(\cdot)$), we show that the $L^{A,B}(\cdot)$ decreases on average by a large amount; it is large enough so that it dominates the possible increase in any other components of $L(\cdot)$. Such a strong property holds because as per the algorithm, $g(A^i_j(\tau))$ and $B^i_j(\tau)$ continually try to estimate $W_j(\tau)$ and hence if either of them is larger than $\boldsymbol{W}^3_{\max}(\tau)$, they ought to decrease by a large amount in a short time. Indeed, to translate this property into sufficient decrease of $L(\cdot)$, the careful choice of $L^{A,B}(\cdot)$ is made.

Case Two: $C(\mathsf{x}) < \boldsymbol{W}^3_{\max}$ In this case, for each $i \in V$ and $j \in \mathcal{N}(i)$, $g(A^i_j(\tau))$ and $B^i_j(\tau)$ are smaller than $\boldsymbol{W}^3_{\max}(\tau)$. To establish the decrease in $L(\cdot)$, we show that in this case $L^Q(\cdot)$ decreases by large enough amount; large enough so that it dominates the possible increase in $L^{A,B}(\cdot)$. This case crucially utilizes the property of the medium access algorithm, the choice of the weights $W_i(\cdot)$ for $i \in V$ and the form of function g. The precise details explaining how these play roles in establishing this decrease in $L^Q(\cdot)$ is explained in the full version of this paper [21]. In this extended abstract, we shall provide key ideas behind these somewhat involved arguments.

The property that $L^Q(\cdot)$ decreases by large enough amount follows if we establish that the set of transmitting nodes $\boldsymbol{\sigma}(\tau)$ is such that

$$\sum_i \sigma_i(\tau) \log\log Q_i(\tau) \approx \max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i \rho_i \log\log Q_i(\tau). \tag{9}$$

To establish (9), using the condition of the second case $g(A^i_j(\tau)) < \boldsymbol{W}^3_{\max}(\tau)$ for all $i \in V$ and $j \in \mathcal{N}(i)$, we essentially show that

$$g(A^i_j(\tau)) \approx W_j(\tau), \quad \text{for all } i \in V,\ j \in \mathcal{N}(i), \tag{10}$$

$$\sum_i \sigma_i(\tau) \log\log Q_i(\tau) \approx \max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i \rho_i \log W_i(\tau). \tag{11}$$

To see why (10) and (11) are sufficient to imply (9), note that

$$\begin{aligned} \big|\log W_i(\tau) - \log\log Q_i(\tau)\big| &\leq \max_{j\in\mathcal{N}(i)} \sqrt{\log g(A^i_j(\tau))} \\ &\approx \max_{j\in\mathcal{N}(i)} \sqrt{\log W_j(\tau)} \\ &\ll \max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i \rho_i \log W_i(\tau), \end{aligned}$$

when $\boldsymbol{W}_{\max}(\tau)$ (or $\mathbf{Q}_{\max}(\tau)$) is very large. Therefore,

$$\max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i \rho_i \log\log Q_i(\tau) \approx \max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i \rho_i \log W_i(\tau).$$

In summary, to establish desired decrease in $L^Q(\cdot)$, it boils down to establishing (10) and (11).

To establish (10), it is essential for $g(\cdot)$ to be *growing fast enough* so that if $g(A_j^i(\tau))$ is very different (in this case, smaller) compared to $W_j(\tau)$, then under the execution of the algorithm, it quickly converges (close) to $W_j(\cdot)$. For this, it is important that $g(A_j^i(\cdot))$ should change at a faster rate compared to the rate at which $W_j(\cdot)$ changes. Towards that, note that if $A_j^i(\cdot)$ is updated (by unit amount) then $g(A_j^i(\cdot))$ roughly changes by amount $g'(A_j^i(\tau))$, which is at least

$$g'(A_j^i(\tau)) > g'(g^{(-1)}(\boldsymbol{W}_{\max}(\tau)^3)).$$

Here we have used the fact that g' is a decreasing function and $g(A_j^i(\tau))$ is at most $\boldsymbol{W}_{\max}(\tau)^3$. Using properties of function g, we establish that $W_j(\tau)$ changes per unit time by at most

$$\frac{W_j(\tau)}{g^{(-1)}\left(\exp\left(\log^2 W_j(\tau)\right)\right)}.$$

For the purpose of developing an intuition regarding the choice of g, consider $j \in \arg\max_i W_i(\tau)$, i.e. $W_j(\tau) = \boldsymbol{W}_{\max}(\tau)$. Then, such a $W_j(\tau)$ changes as

$$\frac{W_j(\tau)}{g^{(-1)}\left(\exp\left(\log^2 W_j(\tau)\right)\right)} = \frac{\boldsymbol{W}_{\max}(\tau)}{g^{(-1)}\left(\exp\left(\log^2 \boldsymbol{W}_{\max}(\tau)\right)\right)}.$$

Therefore, to have g such that the change in $W_j(\cdot)$ is slower than that in $g(A_j^i(\cdot))$, we must have

$$g'(g^{(-1)}(\boldsymbol{W}_{\max}(\tau)^3)) \;>\; \frac{\boldsymbol{W}_{\max}(\tau)}{g^{(-1)}\left(\exp\left(\log^2 \boldsymbol{W}_{\max}(\tau)\right)\right)}.$$

Our interest will be having properties holding when $\boldsymbol{W}_{\max}(\tau)$ (or $\mathbf{Q}_{\max}(\tau)$) is large enough. This leads to the condition that

$$\lim_{x\to\infty} g'(g^{(-1)}(x^3))\frac{g^{(-1)}\left(\exp\left(\log^2 x\right)\right)}{x} > \; 1.$$

It can be checked that the above condition is satisfied if $g(x)$ does not grow slower than $\exp(\log\log^\alpha x)$ for some constant $\alpha > 2$.[2] That is, we need g to be growing roughly at least as fast as the choice made in the description of our algorithm in Section 2.

Next, discussion on how we establish (11), which will require another condition on $g(\cdot)$ to be *growing slow enough*, in contrast to the fast enough growing condition for (10). Effectively, we need to establish that $\boldsymbol{\mu}(\tau)$, the distribution of $\boldsymbol{\sigma}(\tau)$ under the algorithm, is concentrated around the subset of schedules with high-weight, i.e. roughly speaking the subset

$$\Big\{\tilde{\boldsymbol{\rho}} = [\tilde{\rho}_i] \in \mathcal{I}(G) : \sum_i \tilde{\rho}_i \log W_i(\tau) \approx \max_{\boldsymbol{\rho}\in\mathcal{I}(G)} \sum_i \rho_i \log W_i(\tau)\Big\}. \quad (12)$$

To that end, consider the evolution of schedule $\boldsymbol{\sigma}(\tau) = [\sigma_i(\tau)]$ and weight $\boldsymbol{W}(\tau) = [W_i(\tau)]$ under the algorithm. Now the distribution of $\boldsymbol{\sigma}(\tau)$ depends on the schedule $\boldsymbol{\sigma}(\tau-1)$ and weight $\boldsymbol{W}(\tau-1)$. More specifically, the evolution of $\boldsymbol{\sigma}(\tau)$ can be thought of as a time-varying Markov chain with its transition matrix $P(\tau)$ being function of the time-varying $\boldsymbol{W}(\tau)$. That is, for $\Delta \geq 1$

$$\boldsymbol{\mu}(\tau) = \; \boldsymbol{\mu}(\tau-\Delta)P(\tau-\Delta)\cdots P(\tau-1).$$

In above, we assume that the distribution $\boldsymbol{\mu}(\cdot)$ represents an $|\mathcal{I}(G)|$ dimensional row vector, $P(\cdot)$ represents an $|\mathcal{I}(G)| \times |\mathcal{I}(G)|$ probability transition matrix, and their product on the right hand side should be treated as the usual vector-matrix multiplication. The first step towards establishing concentration of $\boldsymbol{\mu}(\tau)$ around the subset of $\mathcal{I}(G)$ with high-weight (cf. (12)) is establishing the existence of an appropriate $\Delta \geq 1$:

(a) Δ is small enough so that

$$P(\tau-\Delta)\cdots P(\tau-1) \approx P(\tau)^\Delta.$$

(b) Δ is large enough so that

$$\boldsymbol{\mu}(\tau-\Delta)P(\tau)^\Delta \approx \boldsymbol{\pi}(\tau),$$

where $\boldsymbol{\pi}(\tau)$ is the stationary distribution of $P(\tau)$, i.e. $\boldsymbol{\pi}(\tau) = \boldsymbol{\pi}(\tau)P(\tau)$.

By finding such Δ, it essentially follows that $\boldsymbol{\mu}(\tau) \approx \boldsymbol{\pi}(\tau)$. The second step towards establishing concentration of $\boldsymbol{\mu}(\tau)$ around the high-weight set involves establishing that $\boldsymbol{\pi}(\tau)$ is approximately product-form with respect to the weights $\boldsymbol{W}(\tau)$ (cf. Lemma 2). Therefore, as a consequence of Gibb's maximal principle for product-form distributions, it follows that $\boldsymbol{\pi}(\tau)$ is concentrated around the subset of $\mathcal{I}(G)$ with high-weight (cf. (12)). Subsequently, this establishes that $\boldsymbol{\mu}(\tau)$ is concentrated around the subset of $\mathcal{I}(G)$ with high-weight (cf. (12))

Now we discuss the remaining task of showing the existence of Δ so that (a) and (b) are satisfied. This is where we shall discover another sets of conditions on g that it must be of the form $\exp(\log\log^\alpha x)$ with $\alpha > 2$. Now for (b) to hold, it is required that Δ is larger than the mixing time of $P(\tau)$. Using Cheeger's inequality [4], [23], we prove that it is sufficient to have

$$\Delta > f_1(\boldsymbol{W}_{\max}(\tau)) \quad \text{with} \quad f_1(x) = \Theta(x^{6n+1}). \quad (13)$$

[2] We say g does not grow slower and faster than f if $\lim_{x\to\infty}\frac{g(x)}{f(x)} > 0$ and $\lim_{x\to\infty}\frac{g(x)}{f(x)} < \infty$, respectively.

The precise definition of $f_1(\cdot)$ is presented in Lemma 3.[3] Next, for Δ to satisfy (a), observe that

$$\begin{aligned}
&\|P(\tau-\Delta)\cdots P(\tau-1)-P(\tau)^{\Delta}\| \\
&\leq \sum_{s=1}^{\Delta}\|P(\tau-\Delta)\cdots P(\tau-s-1) \\
&\qquad\qquad (P(\tau-s)-P(\tau))P(\tau)^{s-1}\| \\
&\leq \sum_{s=1}^{\Delta}\|P(\tau-s)-P(\tau)\|,
\end{aligned}$$

where we use the triangle inequality with an appropriately defined norm $\|\cdot\|$. Further, by exploring algebraic properties of $P(\cdot)$ and $\boldsymbol{W}(\cdot)$, we show that

$$\|P(\tau-s)-P(\tau)\| \;\leq\; f_2(\boldsymbol{W}_{\min}(\tau))\cdot s,$$

where $\boldsymbol{W}_{\min}(\tau)=\min_i W_i(\tau)$ and

$$f_2(x)=\Theta\left(\frac{x}{g^{(-1)}\left(\exp\left(\log^2 x\right)\right)}\right).$$

Thus, it follows that

$$\|P(\tau-\Delta)\cdots P(\tau-1)-P(\tau)^{\Delta}\| \;\leq\; f_2(\boldsymbol{W}_{\min}(\tau))\cdot\Delta^2.$$

Therefore, (a) follows if Δ satisfies

$$\Delta<\frac{\varepsilon}{\sqrt{f_2(\boldsymbol{W}_{\min}(\tau))}} \quad \text{for small enough } \varepsilon>0. \tag{14}$$

From (13) and (14), it follows that a $\Delta\geq 1$ satisfying (a) and (b) exists if

$$f_1(\boldsymbol{W}_{\max}(\tau))<\frac{\varepsilon}{\sqrt{f_2(\boldsymbol{W}_{\min}(\tau))}}, \tag{15}$$

for large enough $\mathbf{Q}_{\max}(\tau)$. From (1), it follows that for any $i\in V$,

$$\begin{aligned}
W_i(\tau) \;&\geq\; \max_{j\in\mathcal{N}(i)}\exp\left(\sqrt{\log g(A_j^i(\tau))}\right) \\
&\approx\; \max_{j\in\mathcal{N}(i)}\exp\left(\sqrt{\log W_j(\tau)}\right) \\
&\geq\; \exp\left(\sqrt{\log W_j(\tau)}\right),
\end{aligned} \tag{16}$$

for any $j\in\mathcal{N}(i)$; here we have assumed $g(A_j^i(\tau))\approx W_j(\tau)$. Now let $j_*\in\arg\min_j W_j(\tau)$ and $j^*\in\arg\max_j W_j(\tau)$. Since G is connected, there exists a path connecting j_* and j^* of length at most D where $D\leq n-1$ is the diameter of G. Then by a repeated application of (16) along this path joining j_* and j^* starting with j_*, we obtain that

$$\boldsymbol{W}_{\min}(\tau)\geq\;\exp\left(\log^{1/2^D}\boldsymbol{W}_{\max}(\tau)\right). \tag{17}$$

[3]We use the asymptotic notation Θ with respect to scaling in $\boldsymbol{W}_{\max}(\cdot)$ instead of n.

Therefore, the desired inequality (15) is satisfied for large $\boldsymbol{W}_{\max}(\tau)$ if

$$f_1(\boldsymbol{W}_{\max}(\tau))<\frac{\varepsilon}{\sqrt{f_2\left(\exp\left(\log^{1/2^D}\boldsymbol{W}_{\max}(\tau)\right)\right)}}.$$

This holds if

$$\limsup_{x\to\infty} f_1(x)\sqrt{f_2\left(\exp\left(\log^{1/2^D}x\right)\right)} \;=\; 0.$$

The above can be checked to hold if g does not grow faster than $\exp(\log\log^{\alpha}x)$ for some constant $\alpha<\infty$.

5. A Comparison Lemma

Consider two Markov chains defined on a given finite state space, say Ω of N states. For simplicity, let $\Omega=\{1,\dots,N\}$. Let the transition probability matrices of these two Markov chains be represented as $P=[P_{ij}]\in[0,1]^{N\times N}$ and $Q=[Q_{ij}]\in[0,1]^{N\times N}$, respectively. Let P and Q be irreducible and aperiodic. That is, both of them have unique stationary distributions, which are denoted by $\boldsymbol{\mu}=[\mu_i]_{i\in\Omega}$ and $\boldsymbol{\nu}=[\nu_i]_{i\in\Omega}$, respectively. Define

$$C_{ij}=\max\left\{\frac{P_{ij}}{Q_{ij}},\frac{Q_{ij}}{P_{ij}}\right\}.$$

Also define

$$C^*\equiv C^*(P,Q)=\max_{1\leq i,j\leq N}C_{ij}.$$

By definition, $C^*\geq 1$: $C^*=1$ iff $P=Q$; $C_{ij}=\infty$ iff there exists (i,j) such that $P_{ij}+Q_{ij}>0$ and $P_{ij}Q_{ij}=0$. For a meaningful result, we shall restrict to scenario where $C_{ij}<\infty$. Recall that the relative entropy or Kullback-Liebler distance between distribution $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, denoted by $D_{\mathrm{KL}}(\boldsymbol{\mu},\boldsymbol{\nu})$, is defined as

$$D_{\mathrm{KL}}(\boldsymbol{\mu},\boldsymbol{\nu})=\sum_{i\in\Omega}\mu_i\log\left(\frac{\mu_i}{\nu_i}\right).$$

The following provides bound on $D_{\mathrm{KL}}(\boldsymbol{\mu},\boldsymbol{\nu})$ in terms of C^* and N.

Lemma 6 *For $C^*<\infty$,*

$$\max\left\{D_{KL}(\boldsymbol{\mu},\boldsymbol{\nu}),\; D_{KL}(\boldsymbol{\nu},\boldsymbol{\mu})\right\}\leq 2(N-1)\log C^*. \tag{18}$$

Proof: To establish this, we shall use the characterization of stationary distributions for any irreducible, aperiodic finite state Markov chain given through what is known as the ‘Markov chain tree theorem’ (cf. see [3]). To this end, define a directed graph $\mathcal{G}=(\mathcal{I}(G),\mathcal{E})$ with Ω as vertices and directed edge $(i,j)\in\mathcal{E}$ if and only if $P_{ij}>0$ (equivalently $Q_{ij}>0$). Let $\mathcal{T}_i$ be the space of all directed spanning trees of $\mathcal{G}$ rooted at $i\in\Omega$. Define the weight of a tree $T\in\mathcal{T}_i$ with respect to transition matrix P, denoted as $w(T,P)$, as

$$w(T,P)=\prod_{(i,j)\in T}P_{ij}.$$

Similarly, define the weight of $T \in \mathcal{T}_{\boldsymbol{\sigma}}$ with respect to Q, denoted as $w(T, Q)$, as

$$w(T,Q) = \prod_{(i,j)\in T} Q_{ij}.$$

Then, the Markov chain tree theorem states that for any $i \in \Omega$,

$$\mu_i \propto \sum_{T\in\mathcal{T}_i} w(T,P). \tag{19}$$

And, similarly for $i \in \Omega$,

$$\nu_i \propto \sum_{T\in\mathcal{T}_i} w(T,Q). \tag{20}$$

Since the number of edges in each spanning tree is $N-1$ and $P_{ij}/Q_{ij} \in [1/C^*, C^*]$ for all i,j, it follows that for a given $T \in \mathcal{T}_i$,

$$(C^*)^{-(N-1)} \le \frac{w(T,P)}{w(T,Q)} \le (C^*)^{(N-1)}.$$

Therefore, it follows that

$$(C^*)^{-2(N-1)} \le \frac{\mu_i}{\nu_i} \le (C^*)^{2(N-1)}. \tag{21}$$

Therefore, it can be checked that

$$\begin{aligned} D_{\mathrm{KL}}(\boldsymbol{\mu},\boldsymbol{\nu}) &= \sum_i \mu_i \log(\mu_i/\nu_i) \\ &\le \sum_i \mu_i \Big(2(N-1)\log C^*\Big) \\ &= 2(N-1)\log C^*. \end{aligned} \tag{22}$$

Similar bound applies to $D_{\mathrm{KL}}(\boldsymbol{\nu},\boldsymbol{\mu})$. This completes the proof of Lemma 6. ∎

The above Lemma is essentially tight. To see this, consider two Markov chains on a line of N nodes. The Markov chain P is such that it has the uniform stationary distribution $\boldsymbol{\mu} = [1/N]$ while Markov chain Q is such that its stationary distribution $\boldsymbol{\nu} = [(1-\rho)\rho^{i-1}/(1-\rho^N)]$ with $\rho \in (0,1)$. Examples of such P and Q are as follows: in P, nodes 1 and N have self-loop of probability $3/4$ while all other nodes have self-loop probability $1/2$ and transition between neighboring nodes (on the line) happen with probability $1/2$; in Q the self-loop probability at nodes 1 and N are of $1-\rho/4$, the self-loop probabilities at all other nodes is $1-\rho/4-1/4$, transition probability from node i to $i+1$ is $\rho/4$ while transition probability from node $i+1$ to i is $1/4$ for $1 \le i \le N-1$. For such P and Q, it can be checked that $C^* = 1/\rho$ for $\rho < 1/2$. Further,

$$\begin{aligned} D_{\mathrm{KL}}(\boldsymbol{\mu},\boldsymbol{\nu}) &= \sum_{i=1}^N \frac{1}{N}\log\Big(\frac{(1-\rho^N)}{N(1-\rho)\rho^{i-1}}\Big) \\ &= \frac{1}{N}\Big(\sum_{i=1}^N \log\frac{(1-\rho^N)}{N(1-\rho)} + (i-1)\log\frac{1}{\rho}\Big) \\ &= \log\Big(\frac{(1-\rho^N)}{N(1-\rho)}\Big) + \frac{1}{N}\log\Big(\frac{1}{\rho}\Big)\Big(\sum_{i=1}^N i-1\Big) \\ &= \log\Big(\frac{(1-\rho^N)}{N(1-\rho)}\Big) + \frac{N-1}{2}\log\Big(\frac{1}{\rho}\Big). \end{aligned} \tag{23}$$

Now note that the first term in (23) scales as $-\log N$, while the second term scales as $\frac{N}{2}\log(1/\rho)$. Therefore, effectively $D_{\mathrm{KL}}(\boldsymbol{\mu},\boldsymbol{\nu}) \approx \frac{N}{2}\log(1/\rho)$. Note that Lemma 6 implies that $D_{\mathrm{KL}}(\boldsymbol{\mu},\boldsymbol{\nu}) \le 2(N-1)\log(1/\rho)$. Thus upto a constant factor (roughly 4), this result is tight.

6. Discussion

As the main result of this paper, we presented a new medium-access algorithm for an arbitrary wireless network where simultaneously transmitting nodes must form an independent set of the network graph. The algorithm is optimal in the sense that network Markov chain is positive-recurrent as long as the imposed traffic demand can be satisfied by some scheduling algorithm. The algorithm is entirely distributed: the only information it utilizes is its own queue-size and the history of collision or successful transmission. In a sense, this work settles an important question that has been of interest in distributed computation, communication, probability and learning.

The algorithm we presented builds upon the work of [19] where the algorithm required a bit of information exchange between neighbors per unit time. Specifically, the key technical contribution of our work is to get rid of this requirement by means of designing a novel learning mechanism that essentially estimates the rate of a Bernoulli process with time varying rates. This learning mechanism could be of much broader interest.

References

[1] N. Abramson and F. Kuo (Editors). The aloha system. *Computer-Communication Networks*, 1973.

[2] D. J. Aldous. Ultimate instability of exponential back-off protocol for acknowledgement-based transmission control of random access communication channels. *IEEE Transactions on Information Theory*, 33(2):219–223, 1987.

[3] V. Anantharam and P. Tsoucas. A proof of the markov chain tree theorem. *Statistics & Probability Letters*, 8(2):189192, June 1989.

[4] M. Dyer, A. Frieze, and R. Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. ACM*, 38(1):1–17, 1991.

[5] S. Foss and T. Konstantopoulos. An overview of some stochastic stability methods. *Journal of Operations Research, Society of Japan*, 47(4), 2004.

[6] H. O. Georgii. *Gibbs measures and phase transitions*. Walter de Gruyter, 1988.

[7] L. A. Goldberg and P. D. MacKenzie. Analysis of practical backoff protocols for contention resolution with multiple servers. *Journal of Computer and System Sciences*, 58(1):232 – 258, 1999. Preliminary version in SODA '96: *Proceedings of the seventh annual ACM-SIAM Symposium on Discrete Algorithms*.

[8] L. A. Goldberg. Design and analysis of contention-resolution protocols, epsrc research grant gr/l60982. http://www.csc.liv.ac.uk/∼leslie/contention.html, Last updated, Oct. 2002.

[9] J. Hastad, T. Leighton, and B. Rogoff. Analysis of backoff protocols for multiple access channels. *SIAM Journal on Computing*, 25(4), 1996. Preliminary version in STOC '87: *Proceedings of the nineteenth annual ACM Symposium on Theory of Computing*.

[10] L. Jiang, D. Shah, J. Shin, and J. Walrand. Distributed random access algorithm: Scheduling and congesion control. *IEEE Transaction on Information Theory*, December 2010.

[11] L. Jiang and J. Walrand. A distributed csma algorithm for throughput and utility maximization in wireless networks. In *Proceedings of 46th Allerton Conference on Communication, Control, and Computing, Urbana-Champaign, IL*, 2008.

[12] L. Jiang and J. Walrand. Approaching Throughput-Optimality in Distributed CSMA Scheduling Algorithms with Collisions Accepted to appear in *IEEE Transactions on Networking*, 2011.

[13] F. P. Kelly. Stochastic models of computer communication systems. *J. R. Statist. Soc B*, 47(3):379–395, 1985.

[14] F. P. Kelly and I. M. MacPhee. The number of packets transmitted by collision detect random access schemes. *The Annals of Probability*, 15(4):1557–1568, 1987.

[15] I. M. MacPhee. On optimal strategies in stochastic decision processes, d. phil. thesis, university of cambridge, 1989.

[16] R. Metcalfe and D. Boggs. Distributed packet switching for local computer networks. *Comm. ACM*, 19:395–404, 1976.

[17] R. Montenegro and P. Tetali. Mathematical aspects of mixing times in markov chains. *Found. Trends Theor. Comput. Sci.*, 1(3):237–354, 2006.

[18] J. Mosely and P. A. Humblet. A class of efficient contention resolution algorithms for multiple access channels. *IEEE Transactions on Communications*, 33(2):145–151, 1985.

[19] S. Rajagopalan, D. Shah, and J. Shin. Network adiabatic theorem: an efficient randomized protocol for contention resolution. In *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*, pages 133–144. ACM New York, NY, USA, 2009.

[20] D. Shah and J. Shin. Randomized scheduling algorithm for queueing networks. *To appear in Annals of Applied Probability*, 2011.

[21] D. Shah, J. Shin, and P. Tetali. Efficient Distributed Medium Access. http://http://arxiv.org/abs/1104.2380.

[22] D. Shah and D. J. Wischik. Optimal scheduling algorithm for input queued switch. In *Proceeding of IEEE INFOCOM*, 2006.

[23] A. Sinclair. *Algorithms for Random Generation and Counting: A Markov Chain Approach*. Birkhäuser, Boston, 1993.

[24] A. L. Stolyar. MaxWeight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic. *Annals of Applied Probability*, 14(1):1–53, 2004.

[25] B. S. Tsybakov and N. B. Likhanov. Upper bound on the capacity of a random multiple-access system. *Problemy Peredachi Informatsii*, 23(3):64–78, 1987.