Simple Contention Resolution via Multiplicative Weight Updates

Yi-Jun Chang
University of Michigan
cyijun@umich.edu

Wenyu Jin
University of Michigan
wyjin@umich.edu

Seth Pettie[1]
University of Michigan
pettie@umich.edu

Abstract

We consider the classic *contention resolution* problem, in which devices conspire to share some common resource, for which they each need temporary and exclusive access. To ground the discussion, suppose (identical) devices wake up at various times, and must send a single *packet* over a shared *multiple-access channel*. In each time step they may attempt to send their packet; they receive ternary feedback $\{0, 1, 2^+\}$ from the channel, 0 indicating *silence* (no one attempted transmission), 1 indicating *success* (one device successfully transmitted), and 2^+ indicating *noise*. We prove that a simple strategy suffices to achieve a channel utilization rate of $1/e - O(\epsilon)$, for any $\epsilon > 0$. In each step, device i attempts to send its packet with probability p_i, then applies a rudimentary multiplicative weight-type update to p_i.

$$p_i \leftarrow \begin{cases} p_i \cdot e^{\epsilon} & \text{upon hearing silence } (0) \\ p_i & \text{upon hearing success } (1) \\ p_i \cdot e^{-\epsilon/(e-2)} & \text{upon hearing noise } (2^+) \end{cases}$$

This scheme works well even if the introduction of devices/packets is adversarial, and even if the adversary can *jam* time slots (make noise) at will. We prove that if the adversary jams J time slots, then this scheme will achieve channel utilization $1/e - \epsilon$, excluding $O(J)$ wasted slots. Results similar to these (Bender, Fineman, Gilbert, Young, SODA 2016) were already achieved, but with a lower constant efficiency (less than 0.05) and a more complex algorithm.

2012 ACM Subject Classification Networks → Network protocols, Mathematics of computing → Probabilistic algorithms

Keywords and phrases Contention resolution, multiplicative weight update method

Digital Object Identifier 10.4230/OASIcs.SOSA.2019.16

1 Introduction

Suppose n identical devices have packets that they wish to transmit over a shared multiple access channel. For simplicity we assume that time is divided into discrete time slots and that the devices are synchronized. In each time slot they decide whether to attempt to transmit their packet or remain idle. In order to succeed the devices must monopolize the

[1] Supported by NSF grants CCF-1514383, CCF-1637546, and CCF-1815316.

2nd Symposium on Simplicity in Algorithms (SOSA 2019).
Editors: Jeremy Fineman and Michael Mitzenmacher; Article No. 16; pp. 16:1–16:16
OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

channel for one time slot: if two or more devices transmit there is *noise* and if zero devices transmit there is *silence*. We assume that after each time step, all devices receive ternary feedback $\{0, 1, 2^+\}$ from the channel indicating how many devices attempted to transmit their packets. The reader should remember that the problem we are considering here is *abstract* contention resolution. The terms *packet*, *channel*, *noise*, etc. are merely meant to keep an easily visualized instance of the problem in mind.

The traditional way to solve this contention resolution problem is via *exponential backoff* [22]. Each device i holds a parameter p_i, initialized to some constant, say 1/2. In each time step it executes the following protocol.[2]

Binary Exponential Backoff:

$$\text{Device } i \begin{cases} \text{remains silent with probability } 1 - p_i \\ \text{transmits with probability } p_i\text{; if unsuccessful, set } p_i \leftarrow p_i/2. \end{cases}$$

Although binary exponential backoff is empirically useful in many applications [22, 20, 18, 27, 19, 24], it has numerous shortcomings. Even if packets are injected into the system according to a Poisson distribution with some low expectation $\lambda > 0$ (i.e., a plausible and *non*-adversarial input distribution), binary exponential backoff will *eventually* become deadlocked and no more packets will ever be successfully sent [1, 4]. When all n packets are injected simultaneously, binary exponential backoff requires $n \log n$ steps to transmit all of them, and each device attempts to transmit its packet $\Theta(\log n)$ times [4], whereas $O(n)$ and $O(1)$ are optimal in this situation.

Recent research has tried to fix all the deficiencies of exponential backoff, and along many metrics this research has been quite successful. Bender et al. [4] studied the behavior of backoff-type protocols when all n packets arrive simultaneously, and proved that $O(n \log\log n / \log\log\log n)$ time is necessary and sufficient for *monotone* protocols (p_i decreases over time) whereas $O(n)$ time is possible with a non-monotone protocol. In the case that a *jammer* can jam slots at will, it is possible to achieve a (small) constant throughput on the unjammed slots [6, 3], even when the adversary controls the injection rate of new packets. In the [6] protocol each device makes $O(\log^2(n + J))$ transmissions, on average, where J is the number of jammed slots.

Bender, Kopelowitz, Pettie, and Young [7] considered a model motivated by battery-powered devices in which *both* transmitting and listening to the channel cost one unit of energy. They proved that constant channel utilization could be achieved with $O(\log(\log^* n))$ energy per device when an adversary controls the packet insertions (but cannot jam). The bound $O(\log(\log^* n))$ was later proved to be optimal [10].

Unfortunately, these recent advances come nowhere close to the minimalism and elegance of binary exponential backoff. In this work we design a contention resolution protocol that matches and substantially improves the main result of [6], while at the same time achieving something close to the simplicity of binary exponential backoff. Like backoff, our algorithm keeps a single numerical parameter (p_i) and is otherwise *stateless*: it keeps no information on its previous actions or the history of the channel.

[2] Exponential backoff comes in several more-or-less equivalent varieties. In the *windowed* version each device partitions its time in the system into consecutive *windows* $W_1, W_2, W_3, \ldots$, $|W_j| = 2^j$, and attempts to transmit at a uniformly chosen time slot in each window, until successful. In the *homogeneous* version, any device i in the system for t steps transmits with probability $p_i = 1/t$. The version of exponential backoff presented here requires the devices to keep track of less information.

Organization

In Section 2 we introduce our contention resolution protocol and analyze some parts of it. In Section 3 we design an unusual (continuous, real valued) potential function, and use it to to argue that the channel utilization of our protocol can get arbitrarily close to $1/e$. In Section 4 we give a more thorough literature survey on contention resolution and multiple access channels. We conclude in Section 5 with some observations and open problems.

2 Contention Resolution

If the devices have no distinguishing features to break symmetry, but they know what 'n' is, then a reasonable strategy is for everyone to transmit with probability $p = 1/n$, decrementing 'n' every time a packet is transmitted successfully. Observe that they succeed with probability $p_{\text{suc}} = \binom{n}{1}p(1-p)^{n-1} > e^{-1}$, and that $\lim_{n\to\infty} p_{\text{suc}} = 1/e$. More generally, the number of devices that transmit is, in the limit, a Poisson-distributed random variable:

$$\lim_{n\to\infty} \Pr(t \text{ devices transmit}) = \binom{n}{t} p^t (1-p)^{n-t} = e^{-1}/t!$$

Of course, for any finite n the distribution is merely *almost*-Poisson. In order to simplify things, we begin by considering an algorithm that creates channel feedback consistent with a number of transmitters that is Poisson-distributed. Each device i holds a variable p_i which it uses to determine its behavior according to the *Transmission Rule*.

Transmission Rule:

$$\text{Device } i \begin{cases} \text{remains silent with probability } e^{-p_i} \\ \text{transmits its packet with probability } p_i e^{-p_i} \\ \text{makes noise with probability } 1-(1+p_i)e^{-p_i} \end{cases}$$

If device i successfully transmits its packet, it halts.

We will later argue that "*making noise*" (even if the devices were capable of this) is unnecessary, and that the algorithm is improved if we simply transmit with probability $1 - e^{-p_i}$.

If the number of devices in the system is n, the probability of the three channel feedbacks (silence, success, and noise) is exactly:

$$p_{\text{sil}} = \prod_{i\in[n]} e^{-p_i} = e^{-c}$$
$$p_{\text{suc}} = \sum_{i\in[n]} p_i e^{-p_i} \cdot \prod_{j\in[n]\setminus\{i\}} e^{-p_j} = ce^{-c}$$
$$p_{\text{noi}} = 1-(1+c)e^{-c}$$

where c measures the *aggregate contention* in the system

$$c = \sum_{i\in[n]} p_i$$

The probability of success is maximized when $c = 1$. We would like to design an update rule such that c tends to move toward 1 whenever it is too small or too large. Observe that when $c = 1$, the probability of hearing silence and noise are $1/e$ and $(e-2)/e$, respectively. In order for the update rule to be unbiased at $c = 1$, we must respond to noise and silence proportionately. Assuming device i has not successfully transmitted its packet, it applies the Update Rule to change p_i.

Update Rule:

$$p_i \leftarrow \begin{cases} p_i \cdot e^{\epsilon} & \text{upon hearing silence} \\ p_i & \text{upon hearing success} \\ p_i \cdot e^{-\epsilon/(e-2)} & \text{upon hearing noise} \end{cases}$$

Here the *step size* $\epsilon > 0$ is the only parameter of the algorithm. Since probabilities are updated multiplicatively, it is natural to measure the contention c on a logarithmic scale, so we define

$$\gamma = \ln(c)$$

In the absence of packet arrivals/departures, γ evolves according to a random walk on the reals that has a certain positive *attraction* towards the origin.[3] Observe that if γ and γ' are the values before and after an update, $\gamma' \in \{\gamma - \frac{\epsilon}{e-2}, \gamma, \gamma + \epsilon\}$.[4] We define the *attraction at* γ to be the expectation of $\gamma' - \gamma$, expressed in units of the step size ϵ.

$$\mathrm{attr}(\gamma) = p_{\mathrm{sil}}(\gamma) - \tfrac{1}{e-2} \cdot p_{\mathrm{noi}}(\gamma) = e^{-e^{\gamma}} - \tfrac{1}{e-2} \cdot (1 - (1 + e^{\gamma})e^{-e^{\gamma}}).$$

In other words, $\mathrm{E}[\gamma'] = \gamma + \epsilon \cdot \mathrm{attr}(\gamma)$. Observe that because of the different step sizes, $\mathrm{attr}(\gamma)$ is asymptotic to 1 as $\gamma \to -\infty$ and asymptotic to $-1/(e-2)$ as $\gamma \to \infty$. We do not deal with the actual expression for $\mathrm{attr}(\gamma)$, but with a piecewise-linear approximation. See Figure 1.

▶ **Approximation 1.** *Define* $\widetilde{\mathrm{attr}}(\gamma)$ *as follows.*

$$\widetilde{\mathrm{attr}}(\gamma) = \begin{cases} 3/5 & \gamma < -1 \\ -(3/5)\gamma & \gamma \in [-1, 1] \\ -3/5 & \gamma > 1 \end{cases}$$

Then $\mathrm{attr}(0) = \widetilde{\mathrm{attr}}(0) = 0$ *and* $\mathrm{attr}(\gamma)/\widetilde{\mathrm{attr}}(\gamma) \geq 1$ *when* $\gamma \neq 0$.

2.1 Interlude: Homesick Random Walks

In order to build some intuition for how γ evolves, it is useful to think about what the stationary distribution of a *simplified* random walk looks like when the walk exhibits an attraction towards the origin. Consider a random walk on the integers $[-\delta^{-1}, \ldots, \delta^{-1}]$, $\delta > 0$, with the following transition probabilities. If the token is at $\pm i$ at step t, at step $t+1$ it moves toward 0 (i.e., to $\pm(i-1)$) with probability $(1 + i\delta)/2$ and away from 0 (i.e., to $\pm(i+1)$) with probability $(1 - i\delta)/2$. When it is at 0 it moves to -1 and 1 with equal probability. Let

[3] Interestingly, this attraction is qualitatively different in the positive and negative halves of the γ-axis, though it is numerically similar. Observe that when $\gamma > 0$ is large, the probability of hearing silence $p_{\mathrm{sil}} = e^{-c}$ is exponentially small in c, but when $\gamma < 0$ is small, the probability of hearing noise $p_{\mathrm{noi}} = 1 - (1 + c)e^{-c} \approx c^2$ is quadratic in c.

[4] Actually, in the event of a successful transmission, γ will be reduced by some amount after one device withdraws from the system; we do not take that effect into account when calculating *attraction*.

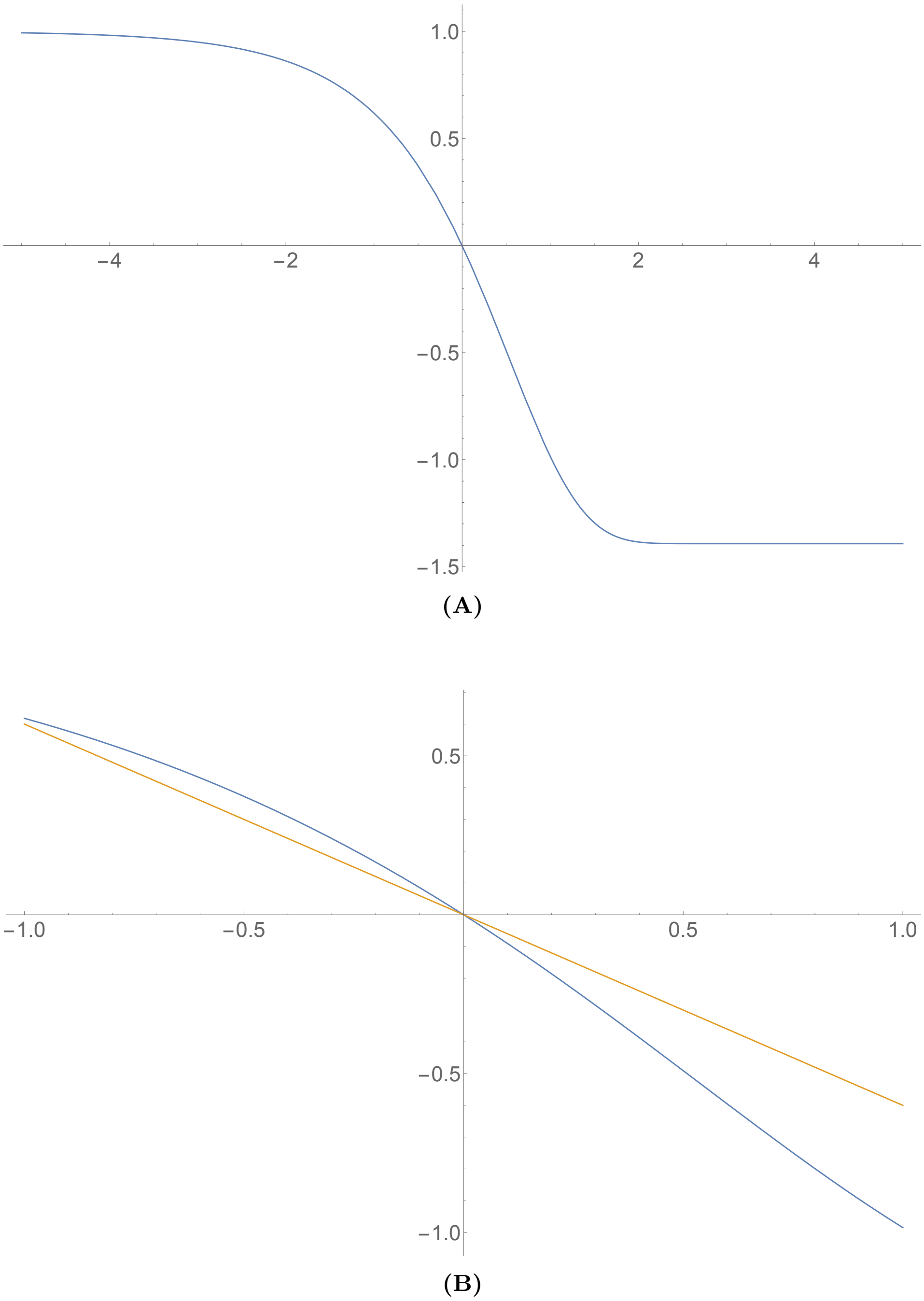

Figure 1 (A) The attraction function $\mathrm{attr}(\gamma)$ is monotone decreasing in γ. (B) In the interval $[-1, 1]$, $\widetilde{\mathrm{attr}}(\gamma) = -\frac{3}{5}\gamma$ is a conservative approximation in the sense that $\mathrm{attr}(\gamma)/\widetilde{\mathrm{attr}}(\gamma) > 1$.

$\pi(i)$ be the probability of being at either i or $-i$ under the stationary distribution. Then π satisfies the following equations.

$$\pi(0) = \pi(1) \cdot \frac{1+\delta}{2}$$
$$\pi(i-1) \cdot \frac{1-(i-1)\delta}{2} = \pi(i) \cdot \frac{1+i\delta}{2}$$

and hence

$$\pi(i) = 2\pi(0) \cdot \prod_{j=1}^{i} \frac{1-(j-1)\delta}{1+j\delta} = 2\pi(0) \cdot \prod_{j=0}^{i-1} \frac{1-j\delta}{1+(i-j)\delta} > 2\pi(0)(1-i\delta)^i \approx 2\pi(0)e^{-i^2\delta}.$$

In other words, a constant fraction of the mass of π is in the interval $[-\sqrt{\delta^{-1}}, \sqrt{\delta^{-1}}]$.

2.2 The Efficiency Curve

The back-of-the-envelope calculations above *suggest* that in its stationary distribution, the random walk generated by our contention resolution protocol puts a constant fraction of the probability mass in the real interval $[-\sqrt{\epsilon}, \sqrt{\epsilon}]$. Given that the efficiency of the algorithm is $1/e$ at $\gamma = 0$, it is natural to ask how the efficiency *degrades* as γ deviates from optimum. The overall efficiency of the algorithm will be determined by its behavior at the extremes of $\gamma \in [-\sqrt{\epsilon}, \sqrt{\epsilon}]$.

Recall that $p_{\text{suc}}(\gamma) = e^{\gamma}e^{-e^{\gamma}}$ is the probability of success as a function of $\gamma = \ln(c)$. By taking the first few terms of the Taylor expansion of p_{suc} at $\gamma = 0$, we have the following approximation. See Figure 2.

▶ **Approximation 2.**

$$p_{\text{suc}}(\gamma) = \frac{1}{e} - \frac{\gamma^2}{2e} - \frac{\gamma^3}{6e} + O(\gamma^4) > \frac{1}{e} - \gamma^2\left(\frac{1}{2e} + \frac{1}{6e}\right) > 1/e - \gamma^2/4.$$

To recap, we expect that γ will spend a constant fraction of its time in $[-\sqrt{\epsilon}, \sqrt{\epsilon}]$, and in this interval the expected channel utilization of the algorithm is at least $1/e - \epsilon/4$.

3 Amortized Analysis

Our goal is to show that the channel utilization of the algorithm is $1/e - O(\epsilon)$ by analyzing the *expected* change in a certain potential function Φ. Let Φ_t be the potential after time slot t and $n_t \geq 0$ be the number of packets inserted into the system just before time slot t begins. We intend to show that

$$\mathrm{E}[\Phi_t - \Phi_{t-1}] \leq -(1 - O(\epsilon)) + (e + O(\epsilon)) \cdot n_t.$$

In other words, each new packet carries with it $e + O(\epsilon)$ units of potential, and the combined effect of the Transmission & Update Rules reduces the potential by $1 - O(\epsilon)$ in expectation, thereby "paying for" this slot in a probabilistic sense. As a consequence, the channel utilization is $(1 - O(\epsilon))/(e + O(\epsilon)) = 1/e - O(\epsilon)$; the formal definition of channel utilization and its analysis will be presented in Section 3.4. For this analysis to work, it is important that newly injected devices initialize p_i properly.

Initialization Rule:

Upon activation, device i sets $p_i \leftarrow \epsilon^2$.

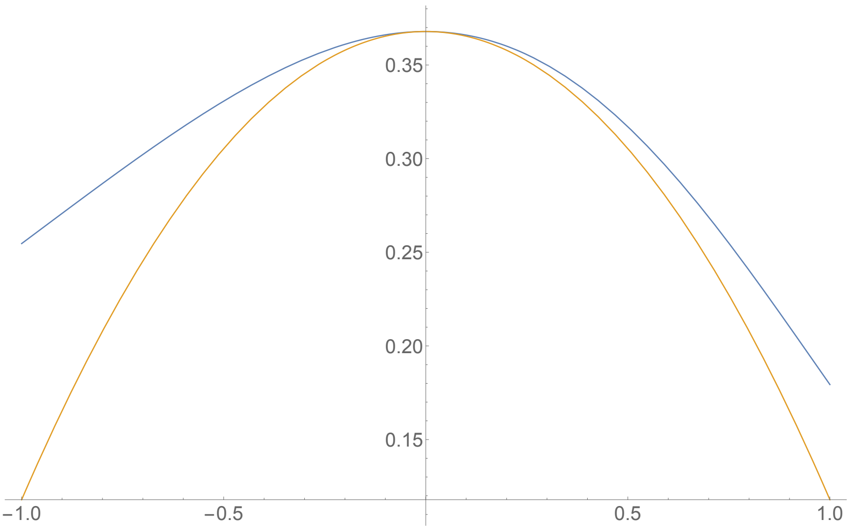

Figure 2 Top curve: the probability of successful transmission, as a function of γ, is $e^{\gamma}e^{-e^{\gamma}}$. Bottom curve: it is lower bounded by $1/e - \gamma^2/4$.

3.1 The Potential Function

The potential function Φ has three components.

$$\Phi = A(n) + B(\gamma) + C.$$

A depends only on n, the number of active devices still in the system, B depends on the contention γ, and C depends on the *relative* magnitude of the variables (p_i). The main term is

$$A(n) = en.$$

If γ is in the "efficient" range $[-\sqrt{\epsilon}, \sqrt{\epsilon}]$, then by Approximation 2 the expected change in A is $p_{\text{suc}}(\gamma) \cdot (-e) = -(1 - O(\epsilon))$, which pays for the time slot.

When $\gamma \in (-\infty, -\sqrt{\epsilon}) \cup (\sqrt{\epsilon}, \infty)$ we make up for the loss in efficiency by showing the expected contention becomes *closer* to optimum in the next time step; this is where the $B(\gamma)$ term comes into play. We define B to be the unique continuous function with $B(0) = 0$ and derivative

$$B'(\gamma) = \begin{cases} -\dfrac{5}{3\epsilon} & \text{when } \gamma < -1 \\ \dfrac{5}{3\epsilon}\gamma & \text{when } \gamma \in [-1, 1] \\ \dfrac{5}{3\epsilon} & \text{when } \gamma > 1 \end{cases}$$

In other words, when $\gamma \in [-1, 1]$, $B(\gamma) = \frac{5}{6\epsilon}\gamma^2$; see Figure 3.

Recall that $\mathrm{E}[\gamma'] = \gamma + \epsilon \cdot \mathrm{attr}(\gamma)$, and by Approximation 1, $\mathrm{attr}(\gamma)/\widetilde{\mathrm{attr}}(\gamma) \geq 1$ when $\gamma \neq 0$. Therefore, whenever γ, γ' are in the "far off" range $(-\infty, -1] \cup [1, \infty)$, the expected change in B is smaller than $B'(\gamma) \cdot \epsilon \cdot \widetilde{\mathrm{attr}}(\gamma) \leq -1$. When $\gamma \in [-1, 1]$ is close to the origin, it is also possible to show that the combined change in $A + B$ is less than $-(1 - O(\epsilon))$ in expectation. The formal analysis is in Section 3.2.

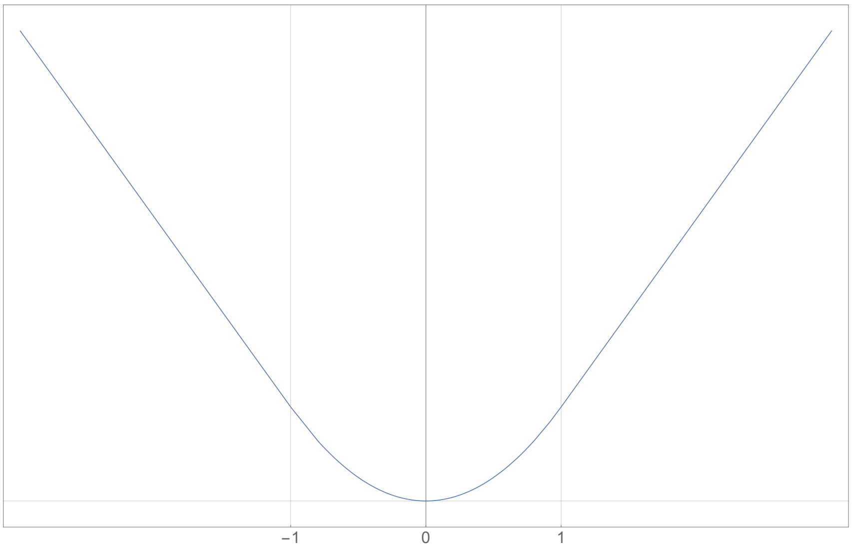

Figure 3 A schematic depiction of B. It is linear in the ranges $(-\infty, 1]$ and $[1, \infty)$, and quadratic in $[-1, 1]$.

In view of the above, under most of the circumstances, the expected change in A and B suffices to pay for each time slot. Occasionally, one packet p_i contributes the lion's share of the aggregate contention c, and when packet i is transmitted, $\gamma = \ln(c)$ drops sharply, increasing $B(\gamma)$. The third component C of Φ compensates for this effect:

$$C = \frac{5}{3\epsilon}(\gamma - \ln p_{\min}), \text{ where } p_{\min} = \min_i p_i$$

Observe that C remains unchanged by the Update Rule since γ and $\ln p_{\min}$ are increased/decreased by the same amount; but the value of C might change when the packet of the smallest transmission probability is successfully transmitted.

3.2 Analysis

We analyze the effect of one time slot on Φ by considering three actions sequentially (i) the increase in $\Phi = A + B + C$ caused by the insertion of new packets, (ii) the expected increase in $A + B$ caused by executing the Transmission & Update Rules, and (iii) in the event of a successful transmission from device i, the increase in $B + C$ caused by subtracting p_i from the aggregate contention. Part (i) is considered in Lemma 1; parts (ii) and (iii) are considered in Lemma 2.

In the analysis below, we work under the following assumption (*).

$$p_{\min} \leq \epsilon^2 \tag{*}$$

Assumption (*) often fails to hold when the number of active devices in the system is less than ϵ^{-2}. Section 3.4 justifies Assumption (*) by showing that it suffices to bound the long term channel utilization of our contention resolution protocol.

▶ **Lemma 1.** *Suppose all new packets follow the Initialization Rule and that Assumption (*) holds. Inserting m packets increases Φ by $(e + O(\epsilon))m$.*

Proof. We consider the contribution of each packet insertion individually. $A(n) = en$ clearly increases by e. If $\gamma \in [-1, \infty)$, C increases by $\frac{5}{3\epsilon}\ln(\frac{e^\gamma+\epsilon^2}{e^\gamma}) < \frac{5}{3\epsilon}(\epsilon^2/e^\gamma) = O(\epsilon)$ and B also increases by $O(\epsilon)$. (If $\gamma \in [-1, 0]$ then B is actually reduced; this only helps us.) If $\gamma \in (-\infty, -1]$, B is reduced by $\frac{5}{3\epsilon}(\ln(\frac{e^\gamma+\epsilon^2}{e^\gamma}))$ and C is increased by precisely the same amount. (Note that when $\gamma \ll -1$, the positive and negative changes to C and B can be very large.) ◀

▶ **Lemma 2.** *In each time step, the expected change in $A + B$ is $-1 + O(\epsilon)$. In the event of a successful transmission, the worst case change to $B + C$ is at most zero.*

Proof. Let γ and γ' be the values before and after applying the Update Rule in this time step. We consider three cases. Case 1 is when $B(\gamma)$ and $B(\gamma')$ are both on the *linear* parts, when $\gamma, \gamma' \in (-\infty, -1] \cup [1, \infty)$. Case 2 is when $B(\gamma)$ is on the quadratic part of B. Case 3 is when $B(\gamma)$ is on the linear parts of B but $B(\gamma')$ has a chance to be on the quadratic part of B.

Case 1: $\boldsymbol{\gamma \in (-\infty, -(1+\epsilon)] \cup [(1+\frac{\epsilon}{e-2}, \infty)}$. $B(\gamma)$ and $B(\gamma')$ are both guaranteed to be on the linear parts of B. The expected change[5] in B is therefore at most

$$\begin{aligned} B'(\gamma) \cdot (\mathrm{E}[\gamma'] - \gamma) &= B'(\gamma) \cdot (\epsilon \cdot \mathrm{attr}(\gamma)) \\ &\le \left(\mathrm{sign}(\gamma)\frac{5}{3\epsilon}\right) \cdot \left(\epsilon \cdot \widetilde{\mathrm{attr}}(\gamma)\right) \\ &\le \mathrm{sign}(\gamma) \cdot \frac{5}{3} \cdot \left(-\,\mathrm{sign}(\gamma)\frac{3}{5}\right) \le -1. \end{aligned}$$

In this range we do not count on successful transmissions; if they do occur, this reduces A even further.

Case 2: $\boldsymbol{\gamma \in [-1, 1]}$. Here $B(\gamma) = \frac{5}{6\epsilon}\gamma^2$ behaves as a quadratic function. The expected change in $A + B$ is at most

$$p_{\mathrm{suc}}(\gamma)(-e) + \frac{5}{6\epsilon}\left[p_{\mathrm{sil}}(\gamma)\left(-\gamma^2 + (\gamma+\epsilon)^2\right) + p_{\mathrm{noi}}(\gamma)\left(-\gamma^2 + \left(\gamma - \frac{\epsilon}{e-2}\right)^2\right)\right]$$

Cancelling the γ^2 terms, we have

$$= p_{\mathrm{suc}}(\gamma)(-e) + \frac{5}{6\epsilon}\left[2\gamma\epsilon\left(p_{\mathrm{sil}}(\gamma) - \frac{1}{e-2} \cdot p_{\mathrm{noi}}(\gamma)\right) + \epsilon^2\left(p_{\mathrm{sil}}(\gamma) + \frac{1}{(e-2)^2} \cdot p_{\mathrm{noi}}(\gamma)\right)\right]$$

Observe that the term following $2\gamma\epsilon$ is exactly the definition of the *attraction* at γ, i.e., $\mathrm{attr}(\gamma)$. Because $e - 2 < 1$, the term following ϵ^2 is maximized over $\gamma \in [-1, 1]$ when $p_{\mathrm{noi}}(\gamma)$ is maximized. At $\gamma = 1$, $p_{\mathrm{sil}}(\gamma) + (e-2)^{-2}p_{\mathrm{noi}}(\gamma) < 1.53 < 8/5$. Simplifying, we have

$$< p_{\mathrm{suc}}(\gamma)(-e) + \frac{5}{6\epsilon}\left[2\gamma\epsilon \cdot \mathrm{attr}(\gamma) + \frac{8}{5}\epsilon^2\right]$$

[5] Here we are only considering the effect of the Update Rule on γ; decreases in γ due to successful transmission are considered when we analyze the effect on $B + C$.

Applying Approximations 1 and 2 (which state $p_{\text{suc}}(\gamma) > \frac{1}{e} - \frac{\gamma^2}{4}$ and $\gamma \cdot \text{attr}(\gamma) \leq -\frac{3}{5}\gamma^2$) and cancelling an ϵ factor, we have

$$\begin{aligned}&< \left(\frac{1}{e} - \frac{\gamma^2}{4}\right)(-e) + \frac{5}{6}\left[-2 \cdot \frac{3}{5}\gamma^2 + \frac{8}{5}\epsilon\right] \\ &= -1 + \gamma^2\left[\frac{e}{4} - 1\right] + \frac{4}{3}\epsilon \\ &\leq -1 + \frac{4}{3}\epsilon\end{aligned}$$

In other words, in each time step we lose at least $1 - O(\epsilon)$ units of potential in expectation, independent of γ.

Case 3: The remaining case covers the transition between the linear and quadratic parts, when $\gamma \in (-(1+\epsilon), -1) \cup (1, 1 + \frac{\epsilon}{e-2})$. The case 1 analysis applies here, up to a $(1 - O(\epsilon))$-factor since the slope of B between γ and γ' is either in the narrow interval $[-\frac{5}{3\epsilon}, -\frac{5}{3\epsilon}(1-\epsilon))$ or $(\frac{5}{3\epsilon}(1 - \frac{\epsilon}{e-2}), \frac{5}{3\epsilon}]$. This $O(\epsilon)$ loss is more than compensated for by the expected change in A, which is at most

$$p_{\text{suc}}(\gamma)(-e) \leq \left(1 - \frac{(1 + \epsilon/(e-2))^2}{4}\right)(-e) = -\frac{e}{4} + O(\epsilon) \ll -\Theta(\epsilon).$$

This concludes our analysis of the change in $A + B$ caused by the Update Rule.

If device i successfully transmits its packet, the new γ is $\gamma'' = \ln(e^\gamma - p_i)$. The term C *decreases* by at least $\frac{5}{3\epsilon}(\gamma - \gamma'')$. Note that C decreases even more if the successful transmission causes $p_{\min}$ to increase. Because the derivative of B is always at least $-\frac{5}{3\epsilon}$, B *increases* by at most $\frac{5}{3\epsilon}(\gamma - \gamma'')$. Thus, the change in $B + C$ due to successful transmission is always at most zero. ◀

3.3 Variants and Extensions

Jamming

Our analysis easily extends to handle an adversarial *jammer*. In any time step, the jammer can make noise during the time slot; no packets are sent successfully and all active devices receive channel feedback 2^+ (noise). If they are following the Update Rule, then γ is reduced by $\epsilon/(e-2)$, and the increase in $B(\gamma)$ is *at most* $\frac{5}{3\epsilon} \cdot \frac{\epsilon}{e-2} < 2.33$. We charge the jammer $3.33 \cdot J$ for jamming a total of J slots: $1 \cdot J$ pays for the jammed slots and $2.33 \cdot J$ pays for the increase in potential. In other words, we expect the efficiency of our algorithm to be completely unchanged, if we ignore $3.33 \cdot J$ wasted time slots.

A Simpler Transmission Rule

Recall that in order to effect channel feedback consistent with a *precisely* Poisson distribution, the Transmission Rule allowed device i to "make noise" in a time slot (as if ≥ 2 devices were transmitting) with small probability $1 - (1 + p_i)e^{-p_i}$. Intuitively this is unwise. From a device's perspective, it is always better to attempt to transmit its packet rather than make noise. We show that from a system-wide perspective, the efficiency of Transmission Rule* is better than Transmission Rule.

Transmission Rule*:

$$\text{Device } i \begin{cases} \text{remains silent with probability } e^{-p_i} \approx 1 - p_i \\ \text{transmits its packet with probability } 1 - e^{-p_i} \approx p_i \end{cases}$$

If device i successfully transmits its packet, it halts.

▶ **Lemma 3.** *Let Φ be the current potential at the beginning of a time step. Let Φ' (resp., $\Phi^\star$) be the potential after applying Update Rule and Transmission Rule (resp., Transmission Rule*) in this time step. Then $\Phi' \geq \Phi^\star$.*

Proof. The only situation the two protocols differ in their behavior is when all devices remain silent, except for one, which chooses to make noise (Transmission Rule) or transmit its packet (Transmission Rule*). Observe that in this situation, following Transmission Rule* decreases Φ by at least e.[6] On the other hand, following Transmission Rule reduces Φ by at most $\frac{5}{3(e-2)}$ (when $\gamma > 1$), which is smaller than e. Thus, we must have $\Phi' \geq \Phi^\star$. ◀

3.4 Channel Utilization

One unfortunate aspect of our potential function is that it does not perform very well when the number of packets in the system is very small. For example, if there are a constant number of packets and γ is close to 0, then inserting a new packet with $p_i = \epsilon^2$ will likely *increase* C by $\Omega(\epsilon^{-1} \ln \epsilon^{-1})$, not $O(\epsilon)$ like we would hope. It turns out that in order to guarantee channel utilization of $1/e - O(\epsilon)$ over the long term, it is *not necessary* that the system be this efficient when number of active packets drops below a certain threshold, e.g., $O(\text{poly}(\epsilon^{-1}))$. Indeed, if the number of active packets is small, this is proof that the protocol is already functioning at the maximum possible efficiency (successful transmission rate = packet injection rate). Theorem 5 captures this intuition more formally. We first define a class of adversaries that strikes a nice balance between allowing essentially *arbitrary* adversarial behavior and adhering to some long-term average injection rate. This definition is more permissive than (λ, T)-adversaries [4].

▶ **Definition 4.** A λ-*adversary* injects packets and jams time slots indefinitely, under the constraint that $N_t < \lambda(t - \alpha J_t)$, for infinitely many values of t, where N_t and J_t are the total number of packets inserted and slots jammed by time t. Note that in our case, $\alpha = 3.33$.

In other words, if we delete αJ_t wasted slots from consideration, the adversary inserts λ packets per slot, on average, over the time period $[1, t]$. This condition is only required to hold infinitely often, which means the adversary is nearly always unconstrained.

Let us normalize the constants implicit in Lemmas 1, 2, and 3 so that whenever Assumption (*) holds, every packet insertion increases Φ by at most e, and every time step reduces Φ by at least $1 - \hat{\epsilon}$ in expectation, where $\hat{\epsilon} = \Theta(\epsilon)$ depends on ϵ.

▶ **Theorem 5.** *Suppose the packet-injection and channel jamming is controlled by a λ-adversary, with $\lambda + \epsilon < \frac{1-\hat{\epsilon}}{e}$. If the devices adhere to the Initialization, Transmission*, and Update Rules, then for infinitely many time slots, the number of active devices in the system will be less than ϵ^{-3}.*

[6] A decreases by e; B is unchanged by Update Rule, and the effect on $B + C$ caused by reducing γ is non-positive.

In other words, for infinitely many time slots, the channel utilization is optimal (up to an *additive* ϵ^{-3}).

Proof. We partition time into consecutive *epochs*, alternating between periods when there are at most ϵ^{-3} active devices and periods when there are greater than ϵ^{-3} active devices. We are not concerned with epochs of the first type. Suppose an epoch of the second type begins at time slot t_0. At this moment we evaluate the potential Φ of the system, with one minor change. In the definition of C, let

$$p_{\min} = \min\{\epsilon^2,\ \min_i p_i\}.$$

We argue that our previous analysis also applies when $p_{\min}$ is redefined in this way. We only need to consider the situation where we hear silence, which would ordinarily make $p_{\min}$ greater than ϵ^2, but it is forced to remain at ϵ^2. Since the epoch has not ended, the contention is $c \geq n\epsilon^2 \geq \epsilon^{-1}$. The probability of hearing silence is $e^{-c} \leq e^{-1/\epsilon}$ and this causes an extra increase in C-potential of $5/3$. On the other hand, the probability of seeing a successful transmission is ce^{-c}, and if this occurs, we see a reduction in potential of $e > 5/3$. The net expected effect of these two phenomena is negative. (Recall that our previous analysis did not take successful transmission into account when c was this large, so we are not double-counting this effect.)

Let Φ_0 be the initial potential endowment at time t_0.[7] Let t_1 be a time sufficiently far in the future when the adversary hits average insertion rate at most $\lambda = \frac{1-\hat{\epsilon}}{e} - \epsilon$. The number of packets inserted during the interval $[t_0, t_1]$ is at most the number of packets inserted by t_1, which is at most λt_1, and so the increase in potential due to packet insertion during the interval $[t_0, t_1]$ is always at most $e\lambda t_1$ (Lemma 1). In the interval $[t_0, t_1]$, the *expected* drop in potential is $(1-\hat{\epsilon})(t_1 - t_0 + 1)$ (Lemma 2).

We choose t_1 to be sufficiently large so that the expected net change in potential is $e\lambda t_1 - (1-\hat{\epsilon})(t_1 - t_0) < -\epsilon t_1$, and $-\epsilon t_1 + \Phi_0 < -\epsilon t_1/2$. Of course, if Φ ever reaches zero the epoch surely has ended. Seeing such a large deviation from the expectation is unlikely.

Let X_i be the potential drop at time step t_i (without taking into account the potential increase due to packet insertion), and let $X = \sum_{i=t_0}^{t_1} X_i$. The probability that the epoch has *not* ended by time t_1 is at most $\Pr[X \geq -(\Phi_0 + e\lambda t_1)]$. Note that $-(\Phi_0 + e\lambda t_1) \geq \mathrm{E}[X] + \epsilon t_1/2$ by our choice of t_1. By Azuma's inequality, this occurs with probability $\exp\left(-\Omega(\frac{(\epsilon t_1/2)^2}{t_1 - t_0 + 1})\right) = \exp(-\Omega(\epsilon^2 t_1))$.[8]

In the unlikely event that the epoch has not ended by time t_1, we can do the analysis with a sufficiently distant point $t_2 > t_1$ in the future. Thus, with probability 1 every epoch with $n > \epsilon^{-3}$ eventually ends. ◀

Theorem 5 establishes the main result of [6] but in a stronger form. In their protocol the efficiency is some constant much smaller than $1/e$. If there are n device injections, the protocol of [6] guarantees that the devices make $O(\log^2(n+J))$ transmission attempts each, on average. Our protocol also improves this aspect of [6], by showing that the number of transmission attempts is independent of n and J.

---

[7] Typically Φ_0 will be $\Theta(\epsilon^{-3})$ but we do not require this.

[8] Note that $|X_i|$ can be upper bounded by a universal Lipchitz constant. In each time step, the term A can only be decreased by at most e; the absolute change of B is at most $\frac{5}{3(e-2)}$; the extra increase in the component C in the modified potential is at most $\frac{5}{3}$.

▶ **Theorem 6.** *If the devices adhere to the Initialization, Transmission***, and Update Rules, the average number of transmission attempts per device is* $e + O(\epsilon)$*, under any adversarial strategy.*

Proof. The analysis is similar, except that the *expected* cost of a slot is now less than $c = e^\gamma$ rather than 1.[9] We redefine the potential Φ to be

$$\Phi = en + \frac{5}{3\epsilon} \cdot \max\{c, 1\}$$

An insertion increases n by 1 and c by ϵ^2, so the cost per insertion is at most $e + O(\epsilon)$. When $\gamma \in [0, \infty)$, the expected change in Φ caused by applying the Update Rule is

$$\begin{aligned} & p_{\text{suc}}(\gamma)(-e) + \frac{5}{3\epsilon}\left[p_{\text{sil}}(\gamma)(-c + ce^{\epsilon}) + p_{\text{noi}}(\gamma)(-c + ce^{-\epsilon/(e-2)})\right] \\ &= p_{\text{suc}}(\gamma)(-e) + \frac{5c}{3\epsilon}\left[p_{\text{sil}}(\gamma)(e^{\epsilon} - 1) + p_{\text{noi}}(\gamma)(e^{-\epsilon/(e-2)} - 1)\right] \end{aligned}$$

We apply the approximation $e^\epsilon \leq 1 + \epsilon + \epsilon^2$ obtained from the Taylor expansion of e^x, yielding:

$$\begin{aligned} &\leq p_{\text{suc}}(\gamma)(-e) + \frac{5c}{3\epsilon}\left[\epsilon \cdot p_{\text{sil}}(\gamma) - \frac{\epsilon}{e-2} p_{\text{noi}}(\gamma) + 2\epsilon^2\right] \\ &= p_{\text{suc}}(\gamma)(-e) + \frac{5c}{3}\left[\text{attr}(\gamma) + 2\epsilon\right] \qquad (**) \end{aligned}$$

We bound (**) depending on γ. When $\gamma \in [1, \infty)$, $\text{attr}(\gamma) + 2\epsilon < -3/5$, in which case (**) is

$$< \frac{5c}{3}\left(-\frac{3}{5}\right) = -c$$

and the slot is paid for, in a probabilistic sense. If $\gamma \in [0, 1]$, we bound (**) as

$$\begin{aligned} &< ce^c(-e) + \frac{5c}{3}\left(-(3/5)\ln c + 2\epsilon\right) \\ &= -c\left(e^{1-c} + \frac{5}{3}\ln c\right) + O(\epsilon) \\ &\leq -c + O(\epsilon) \end{aligned}$$

When $\gamma \in (-\infty, 0]$, the Update Rule (alone) has no effect on Φ; only a successful packet transmission can decrease Φ. Let s be the number of transmitters in a given time slot. When $s = 0$ the cost is zero, so we can consider what the distribution on s looks like, normalized by the event that $s \geq 1$. The event $s = 1$ is good (it costs 1 and decreases Φ by e) and the events when $s \geq 2$ are bad (they cost s and leave Φ unchanged). Within the range $(-\infty, 0]$, the worst distribution on s occurs when $\gamma = 0$, simultaneously minimizing $\Pr(s = 1 | s \geq 1)$ and maximizing $\Pr(s = r | s \geq 1)$ for all $r \geq 2$. The efficiency at $\gamma = 0$ was already handled in the case $\gamma \in [0, 1]$ above. ◀

[9] $c = e^\gamma$ is an *upper bound* on the expected number of packets that transmit in this time step.

4 Related Work

The “new” idea in this work is to create a protocol that is optimal, in a sense, in its lowest energy configuration, by taking inspiration from the *multiplicative weight update* meta-algorithm [2]. Of course, there is nothing new under the sun, and even in the area of backoff-type protocols, updating parameters in response to channel feedback is quite common. Coming from a systems perspective, researches have evaluated variants of exponential backoff that use exponential increase/exponential decrease heuristics [28], multiplicative-increase/linear-decrease [8, 16], additive-increase/multiplicative-decrease [21], and a mixture of linear or multiplicative increase/linear decrease [11].[10] In the theoretical literature, Awerbuch et al. [3] used a multiplicative-weight-type update rule to achieve a (*very* small) constant rate of efficiency, in a model in which a jammer can jam up to a $(1-\epsilon)$-fraction of the slots. To our knowledge, no prior work has analyzed MWU-type contention resolution protocols in both a rigorous and numerically precise fashion.

We have shown that our protocol is *stable* for long-term injection rates approaching $1/e$. The stability of binary exponential backoff (BEB) and its variants has been studied extensively. Aldous [1] showed that for any constant Poisson injection rate $\lambda > 0$, BEB is unstable. Improving this, Bender et al. [4] proved that BEB is unstable at rate $\Omega(\log\log n/\log n)$[11] and stable at rate $O(1/\log n)$. See [15, 17] for other results on the stability of BEB. The failure of BEB to achieve stability even under constant injection rates motivated the development of more complex stable protocols [6, 3, 7]. Unlike BEB, these protocols (like ours) require that the channel feedback differentiate silence and noise.

Although the “$1/e$” threshold of our algorithm is optimal for *stateless* algorithms,[12], it is known that $1/e$ can be beaten, assuming the arrival times of packets are *Poisson-distributed*. The most efficient algorithms of Mosely and Humblet [25] and Tsybakov and Mikhailov [29] (slightly improving [9, 12]) are stable under arrival rates up to ≈ 0.48776. The best known upper bound on contention resolution (in Poisson-distributed injections, which also applies to adversarial injections) is 0.5874 [23]. The assumption of *ternary* feedback is essential here for both the upper and lower bounds. Goldberg et al. [13] have shown that if only transmitters receive feedback from the channel, then no protocol is stable at injections rates above 0.42. Pippenger [26] showed that if the channel reports the *exact* number of transmitters, that a batch of n synchronized devices can solve contention resolution in $n + o(n)$ time slots, i.e., achieving efficiency $1 - o(1)$.

Bender et al. [5] considered variants of BEB that are efficient with heterogeneous packet sizes (as opposed to unit-size packets). Goldberg et al. [14] designed a protocol in which the expected *delay* per packet is $O(1)$, assuming Poisson-injection at rates less than $1/e$.

5 Conclusions

In this work we proved that a simple and natural contention resolution protocol achieves channel utilization arbitrarily close to $1/e$, which is also resilient to a jammer that can jam a constant fraction of the slots. The “$1/e$” threshold of our algorithm cannot be improved by a *stateless* algorithm, and so in this sense its efficiency cannot be improved without a measurable increase in algorithmic complexity. We are confident that the protocols [25, 29]

[10] In these works, ‘multiplicative’ and ‘exponential’ are used interchangeably; ‘additive’ and ‘linear’ are used interchangeably.
[11] I.e., the number of packets that arrive at slot n is Poisson-distributed with expectation $\Omega(\log\log n/\log n)$.
[12] (meaning every device executes the same algorithm in each time step)

with efficiency 0.48776 for Poisson injections can be successfully adapted to adversarial injections using the same multiplicative weight update machinery developed here.

Although our protocol is very efficient in terms of transmission attempts ($e + O(\epsilon)$ vs. the $O(\log^2(n+J))$ of [6]) it does require that the devices listen for channel feedback in every step. In [7], "energy" is defined to be the number of slots spent accessing/listening to the channel. Is it possible to simultaeneously achieve energy cost $\text{poly}(\epsilon^{-1}, \log T)$[13] and $1/e - O(\epsilon)$ channel utilization?

References

1. D. J. Aldous. Ultimate instability of exponential back-off protocol for acknowledgment-based transmission control of random access communication channels. *IEEE Trans. Information Theory*, 33(2):219–223, 1987. doi:10.1109/TIT.1987.1057295.
2. S. Arora, E. Hazan, and S. Kale. The Multiplicative Weights Update Method: a Meta-Algorithm and Applications. *Theory of Computing*, 8(1):121–164, 2012. doi:10.4086/toc.2012.v008a006.
3. B. Awerbuch, A. W. Richa, and C. Scheideler. A jamming-resistant MAC protocol for single-hop wireless networks. In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Principles of Distributed Computing (PODC)*, pages 45–54, 2008. doi:10.1145/1400751.1400759.
4. M. A. Bender, M. Farach-Colton, S. He, B. C. Kuszmaul, and C. E. Leiserson. Adversarial contention resolution for simple channels. In *Proceedings of the 17th Annual ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pages 325–332, 2005. doi:10.1145/1073970.1074023.
5. M. A. Bender, J. T. Fineman, and S. Gilbert. Contention Resolution with Heterogeneous Job Sizes. In *Proceedings 14th Annual European Symposium on Algorithms (ESA)*, pages 112–123, 2006. doi:10.1007/11841036_13.
6. M. A. Bender, J. T. Fineman, S. Gilbert, and M. Young. How to Scale Exponential Backoff: Constant Throughput, Polylog Access Attempts, and Robustness. In *Proceedings 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 636–654, 2016. doi:10.1137/1.9781611974331.ch47.
7. M. A. Bender, T. Kopelowitz, S. Pettie, and M. Young. Contention resolution with log-logstar channel accesses. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, pages 499–508, 2016. doi:10.1145/2897518.2897655.
8. V. Bharghavan, A. J. Demers, S. Shenker, and L. Zhang. MACAW: A media access protocol for wireless LAN's. In *Proceedings of the ACM SIGCOMM Conference on Communications Architectures, Protocols and Applications*, pages 212–225, 1994. doi:10.1145/190314.190334.
9. J. Capetanakis. Tree algorithms for packet broadcast channels. *IEEE Trans. Information Theory*, 25(5):505–515, 1979. doi:10.1109/TIT.1979.1056093.
10. Y.-J. Chang, T. Kopelowitz, S. Pettie, R. Wang, and W. Zhan. Exponential separations in the energy complexity of leader election. In *Proceedings 49th Annual ACM Symposium on Theory of Computing (STOC)*, pages 771–783, 2017. doi:10.1145/3055399.3055481.
11. J. Deng, P. K. Varshney, and Z. Haas. A New Backoff Algorithm for the IEEE 802.11 Distributed Coordination Function. In *In Communication Networks and Distributed Systems Modeling and Simulation*, pages 215–225, 2004.

[13] I.e., if a packet is transmitted at time slot T, it listens in at most $\text{poly}(\epsilon^{-1}, \log T)$ slots.

12 R. G. Gallager. Conflict resolution in random access broadcast networks. In *Proceedings AFOSR Workshop on Communications Theory Applications, Provincetown, MA, Sept 17–20*, pages 74–76, 1978.

13 L. A. Goldberg, M. Jerrum, S. Kannan, and M. Paterson. A bound on the capacity of backoff and acknowledgment-based protocols. *SIAM J. Comput.*, 33(2):313–331, 2004. doi:10.1137/S0097539700381851.

14 L. A. Goldberg, P. D. MacKenzie, M. Paterson, and A. Srinivasan. Contention resolution with constant expected delay. *J. ACM*, 47(6):1048–1096, 2000. doi:10.1145/355541.355567.

15 J. Goodman, A. G. Greenberg, N. Madras, and P. March. Stability of binary exponential backoff. *J. ACM*, 35(3):579–602, 1988. doi:10.1145/44483.44488.

16 Z. J. Haas and J. Deng. On optimizing the backoff interval for random access schemes. *IEEE Trans. Communications*, 51(12):2081–2090, 2003. doi:10.1109/TCOMM.2003.820754.

17 J. Håstad, F. Thomson Leighton, and B. Rogoff. Analysis of Backoff Protocols for Multiple Access Channels. *SIAM J. Comput.*, 25(4):740–774, 1996. doi:10.1137/S0097539792233828.

18 M. Herlihy and J. E. B. Moss. Transactional Memory: Architectural Support for Lock-Free Data Structures. In *Proceedings of the 20th Annual International Symposium on Computer Architecture (ISCA)*, pages 289–300, 1993. doi:10.1145/165123.165164.

19 V. Jacobson. Congestion avoidance and control. In *Proceedings of the ACM Symposium on Communications Architectures and Protocols (SIGCOMM)*, pages 314–329, 1988. doi:10.1145/52324.52356.

20 J. F. Kurose and K. W. Ross. *Computer networking: a top-down approach*, volume 4. Addison-Wesley, Boston, 2009.

21 K. Li, I. Nikolaidis, and J. J. Harms. The analysis of the additive-increase multiplicative-decrease MAC protocol. In *Proceedings 10th Annual Conference on Wireless On-demand Network Systems and Services (WONS)*, pages 122–124, 2013. doi:10.1109/WONS.2013.6578335.

22 R. M. Metcalfe and D. R. Boggs. Ethernet: Distributed packet switching for local computer networks. *Communications of the ACM*, 19(7):395–404, 1976.

23 V. A. Mikhailov and B. S. Tsybakov. Upper bound for the capacity of a random multiple access system. *Problemy Peredachi Informatsii*, 17(1):90–95, 1981.

24 A. Mondal and A. Kuzmanovic. Removing exponential backoff from TCP. *Computer Communication Review*, 38(5):17–28, 2008. doi:10.1145/1452335.1452338.

25 J. Mosely and P. A. Humblet. A Class of Efficient Contention Resolution Algorithms for Multiple Access Channels. *IEEE Trans. Communications*, 33(2):145–151, 1985. doi:10.1109/TCOM.1985.1096261.

26 N. Pippenger. Bounds on the performance of protocols for a multiple-access broadcast channel. *IEEE Trans. Information Theory*, 27(2):145–151, 1981. doi:10.1109/TIT.1981.1056332.

27 R. Rajwar and J. R. Goodman. Speculative lock elision: enabling highly concurrent multithreaded execution. In *Proceedings of the 34th Annual International Symposium on Microarchitecture (MICRO)*, pages 294–305, 2001. doi:10.1109/MICRO.2001.991127.

28 N.-O. Song, B.-J. Kwak, and L. E. Miller. Analysis of EIED backoff algorithm for the IEEE 802.11 DCF. In *Proceedings 62nd IEEE Vehicular Technology Conference (VTC)*, volume 4, pages 2182–2186, 2005.

29 B. S. Tsybakov and V. A. Mikhailov. Slotted multiaccess packet broadcasting feedback channel. *Problemy Peredachi Informatsii*, 14(4):32–59, 1978.