Age of Information in Random Access Channels

Xingran Chen, Konstantinos Gatsis, Hamed Hassani and Shirin Saeedi Bidokhti
Department of Electrical and Systems Engineering
University of Pennsylvania
{xingranc, kgatsis, hassani, saeedi}@seas.upenn.edu

***Abstract*—In applications of remote sensing, estimation, and control, timely communication is not always ensured by high-rate communication. Oftentimes, it is observed that as the capacity of a system is approached, delay increases significantly and so does age of information – a metric recently proposed to capture freshness and timeliness of information. This work proposes distributed age-efficient transmission policies for random access channels with M transmitters. In the first part of this work, we analyze the age performance of stationary randomized policies by relating the problem of finding age to the absorption time of a related Markov chain. In the second part of this work, we propose the notion of *age-gain* of a packet to quantify how much the packet will reduce the instantaneous age of information at the receiver side upon successful delivery. We then utilize this notion to propose a transmission policy in which transmitters act in a distributed manner based on the age-gain of their available packets. In particular, each transmitter sends its latest packet only if its corresponding age-gain is beyond a certain threshold which could be computed adaptively using the collision feedback or found as a fixed value analytically in advance. Both methods improve age of information significantly compared to the state of the art. In the limit of large M, we prove that when the arrival rate is small (below $\frac{1}{eM}$), slotted ALOHA-type algorithms are asymptotically optimal. As the arrival rate increases beyond $\frac{1}{eM}$, while age increases under slotted ALOHA, it decreases significantly under the proposed age-based policies. For arrival rates θ, $\theta = \frac{1}{o(M)}$, the proposed algorithms provide a multiplicative factor of at least two compared to the minimum age under slotted ALOHA (minimum over all arrival rates). We conclude that, as opposed to the common practice, it is beneficial to increase the sampling rate (and hence the arrival rate) and transmit packets selectively based on their age-gain.**

***Index Terms*—Age of Information, Random Access, Collision Channel, Distributed Algorithms, Stochastic Arrival, Slotted ALOHA, Carrier Sensing Multiple Access.**

I. Introduction

Communication networks have witnessed rapid growth in the past few decades and they have laid a path to the integration of intelligence into cyber-physical systems, the Internet of Things, smart cities, as well as healthcare systems. Today, state-of-the-art network communication strategies are considered reliable and high speed; nevertheless, they often do not perform satisfactorily for time-sensitive applications. For example, in applications of remote sensing, estimation, and control, high-rate communication does not ensure timely communication of data. As a matter of fact, it is often observed that as the capacity of a system is approached, the delay increases significantly and hence so does the age of information.

Age of information (AoI), introduced in [1], [2], measures the freshness of information at the receiver side. AoI is a function of both how often packets are transmitted and how much delay packets experience in the system. When the rate of communication is low, the receiver's AoI will increase (implying that the receiver's information is stale) because the transmitter is not sending packets frequently enough. But even when the transmitter is sending packets frequently, if the system design imposes a large delay for the packets, the information at the receiver will still be stale. The metric of AoI is of great importance in the Internet of Things applications where timeliness of information is crucial (e.g. in monitoring the status of a system). Another interesting application domain of AoI is in communication for estimation and control where estimation error increases (exponentially) by time before new packets (samples) are received at the destination. It is believed that minimizing AoI may be a good proxy for minimizing estimation error.

Assuming a first come first serve (FCFS) policy, the work in [3], [4] show in queue theoretic setups that AoI is minimized at an optimal update rate. Relaxing the restriction of FCFS policies, [4], [5] propose packet management policies that discard old packets and improve AoI in wide regimes of operation. This already points to the fact that, under the metric of AoI, rate and reliability have little relevance in the design of communication schemes. This is because AoI implicitly assumes that the information content of the packets form a Markov process and hence fresh packets render older packets obsolete. In the past few years, various extensions and new dimensions have also been studied in the paradigm of timely communication: source and channel coding were studied in [6]–[9], multi-hop networks were studied in [10]–[12], and scheduling algorithms were studied in [13]–[19].

This paper considers the problem of minimizing age of information over a random access channel. This setup is particularly relevant in remote estimation and control of processes that are observed from distributed sensors in wireless networks. For these applications, coordination and scheduling policies that are proposed in prior work such as [**?**] are not practical. Towards designing distributed algorithms for minimizing age of information, [20], [21] analyze stationary randomized policies under the assumption that sources generate packets in every time slot (i.e., all sources are active at all times). Considering the more realistic scenario where packets are generated at random times, [22] analyzes round-robin scheduling techniques with and without packet management and also presents partial results for stationary randomized policies. Round-robin policies are proved to be age-optimal in [23] when the number of transmitters are large and the arrival rate is constant. The followup work [24] additionally assumes that nodes are provided with carrier sensing capabilities and proposes distributed schemes that have good performance in simulations; Nevertheless, [24] does not address how the parameters of the proposed algorithms should be designed theoretically.

In the first part of this paper, we provide a full analysis of age for stationary randomized policies with stochastic arrivals, contributing to the body of work in [20]–[22]. In the second part of this work, we design distributed age-based transmission policies and provide analytical results on achievable AoI in interesting regimes of operation. The major part of this paper deals with random access technologies such as slotted ALOHA that do not assume carrier sensing capabilities. The underlying reason is threefold: (i) Status packets are generally very short (as opposed to traditional settings such as streaming where packets are long) and so CSMA is not efficient, (ii) Transmitters have low power capabilities. As such, it is not vert efficient (in terms of energy and cost) to perform carrier sensing when the rate is large and CSMA is not useful when the rate is small. More importantly, since transmission power is low, the hidden node problem will be a major issue under CSMA-type protocols, (iii) Our analytical results are more clear without the additional complexity of CSMA, but we describe how our findings generalize and apply for CSMA as well.

The contributions of this paper are as follows. In presenting our results below, we assume large symmetric networks in which we have M transmitters and each transmitter has arrival rate θ.

- We derive general lower bound on AoI for any transmission policy and show the asymptotic tightness of the bounds in two regimes of operation, namely when $\theta \leq \frac{1}{eM}$ and when $\theta = \frac{1}{o(M)}$.
- We provide an analytic expression for AoI under stationary randomized policies (in which collision feedback and local AoI information are not utilized). In particular, we relate the problem of finding age to the absorption time of a particular Markov chain. This framework can account for the stochastic nature of arrivals which was disregarded in previous work such as [20], [21] and was only partially addressed in [22] without explicit characterization.
- We prove that when the sum arrival rate $M\theta$ is below the infamous critical point $\frac{1}{e}$, the normalized age performance of a (stabilized) slotted ALOHA algorithm, properly defined later, is approximately $\frac{1}{M\theta}$ in the limit of large M and is optimal. We further show numerically that this is close to the performance of centralized max-weight policies that schedule based on *age-gain*, a notion defined formally in Section IV.
- The maximum (sum) throughput that slotted ALOHA can support is provably $\frac{1}{e}$. Hence, as the sum arrival rate increases beyond this critical point, the age of slotted ALOHA and its unit-buffer-size variants such as [20], [22] increase. We propose two age-based thinning methods (adaptive and stationary thresholding) in which transmitters disregard packets in order to mimic an effective (sum) arrival rate equal to $\frac{1}{e}$. In particular, we develop a threshold policy that can be implemented in a distributed manner at the transmitters and in which packets that offer large age-gains are transmitted and those that offer small age-gains are disregarded. Using the stationary thinning method, we then prove asymptotically ($M \to \infty$) that for any θ that is not too small ($\theta = \frac{1}{o(M)}$), the normalized age is approximately $\frac{e}{2}$. Furthermore, numerical results show that as θ approaches 1, the normalized age approaches 1 using the adaptive thresholding method.
- Finally, we demonstrate how our proposed thinning mechanism is useful for other random access technologies (e.g. CSMA). In particular, we prove that given a technology that can achieve the throughput C, age-based thinning methods attain the normalized age of $\frac{1}{2C}$. In particular, using CSMA, we approach optimality in the limit of large M.

The rest of the paper is organized as follows. Section II introduces the system model and notations. Section III provides lower bounds on NEWSAoI and Section IV proposes centralized Max-Weight scheduling policies to avoid collisions and ensure small NEWSAoI. We analyzes the age performance of stationary randomized policies in Section V. Section VI introduces novel distributed age-based policies and provides asymptotic analysis of their corresponding NEWSAoI (as $M \to \infty$). In Section VII, we numerically compare the achievable age of the proposed distributed transmission policies with centralized policies as well as the derived lower bounds and demonstrate that our asymptotic results hold approximately for moderate values of M as well. We finally conclude in Section VIII and discuss future research directions.

II. System Model and Notation

We consider a wireless symmetric architecture where a controller monitors the status of M identical source nodes over a shared wireless medium. Let time be slotted. At the beginning of every slot k, $k = 1, 2, \ldots$, the source node i, $i = 1, \ldots, M$, generates a new packet encoding information about its current status with probability θ and this packet becomes available at the transmitter immediately. We denote this generation/arrival process at the transmitter by $A_i(k)$, where $A_i(k) = 1$ indicates that a new packet is generated at time slot k and $A_i(k) = 0$ corresponds to the event where there is no new update. New packets are assumed to replace undelivered older packets at the source (i.e., older packets are discarded), relying on the fact that the underlying processes that are monitored in physical systems are oftentimes Markovian[1].

The communication media is modeled by a collision channel: If two or more source nodes transmit at the beginning of the same slot, then the packets interfere with each other (collide) and do not get delivered at the receiver. We use the binary variable $d_i(k)$ to indicate whether a packet is transmitted from source i and received at the destination in

[1]We show in Appendix A that this assumption can be made without loss of generality when the performance measure is Age of Information.

time slot k. Specifically, $d_i(k) = 0$ if source i does not transmit at the beginning of time slot k or if collision occurs; $d_i(k) = 1$ otherwise.

We assume a delay of one time unit in delivery of packets, meaning that packets are transmitted at the beginning of time slots and, if there is no collision, they are delivered at the end of the same time slot. We assume that all transmitters are provided with channel collision feedback at the end of each time slot. Specifically, at the end of time slot k, $c(k) = 1$ if collision happened and $c(k) = 0$ otherwise. In the event that collision occurs, the involved transmitters can keep the undelivered packets and retransmit them according to their transmission policy (until the packets are successfully delivered or replaced by new packets).

Our objective is to design *distributed* transmission mechanisms to minimize time-average age of information per source node. A distributed transmission policy is one in which the decision of transmitter i at time k is dependent only on its own history of actions, the packets arrived so far, $\{A_i(j)\}_{j=1}^{k}$, as well as the collision feedback received so far, $\{c(j)\}_{j=1}^{k-1}$.

The measure of performance in this work is Age of Information (AoI). Originally defined in [1], [2], AoI captures the timeliness of information at the receiver side. We extend the definition a bit further, formally defined below, to also account for the age of information at the source side. Aging at the source/transmitter is caused by the *stochastic nature of arrivals*.

Definition 1. *Consider a source-destination pair. Let $\{k_\ell\}_{\ell\geq 1}$ be the sequence of generation times of packets and $\{k'_\ell\}_{\ell\geq 1}$ be the sequence of times at which those packets are received at the destination. At any time τ, denote the index of the last generated packet by $n_s(\tau) = \max\{\ell | k_\ell \leq \tau\}$ and the index of the last received packet by $n_d(\tau) = \max\{\ell | k'_\ell \leq \tau\}$. The* source's age of information *is defined by $w(k) = k - k_{n_s(k)}$ and the* destination's age of information *is defined by by $h(k) = k - k_{n_d(k)}$.*

It is clear from the above definition that once there is a new packet available at the transmitter, the older packet(s) cannot contribute to reducing the age of the system. We hence assume without loss of generality that buffers at transmitters are of size 1 and new packets replace old packets upon arrival. We formalize and prove this claim in Appendix A.

Following Definition 1, let $h_i(k)$ denote the destination's AoI at time slot k with respect to source i. The age $h_i(k)$ increases linearly as a function of k when there is no packet delivery from source i and it drops with every delivery to a value that represents how old the received packet is; within our framework, this would be the corresponding source's AoI (in previous time slot) plus 1. Without loss of generality, we assume $w_i(1) = 0$ and $h_i(1) \geq 0$, and write the recursion of AoI as follows:

$$h_i(k) = \begin{cases} w_i(k-1) + 1 & d_i(k-1) = 1 \\ h_i(k-1) + 1 & d_i(k-1) = 0 \end{cases} \qquad (1)$$

and

$$w_i(k) = \begin{cases} 0 & A_i(k) = 1 \\ w_i(k-1) + 1 & A_i(k) = 0. \end{cases} \qquad (2)$$

Note that at the beginning of each time slot k, given the collision feedback $\{c(j)\}_{j\leq k-1}$ and local information about $\{A_i(j)\}_{j\leq k}$, transmitter i can compute its corresponding source's AoI $\{w_i(j)\}_{j\leq k}$ and destination's AoI $\{h_i(j)\}_{j\leq k}$.

We define the Normalized Expected Weighted Sum AoI (NEWSAoI) as our performance metric of choice[2]:

$$\lim_{K\to\infty} \mathbb{E}[J_K^\pi],\ J_K^\pi = \frac{1}{MK}\sum_{i=1}^{M}\sum_{k=1}^{K}\left(\alpha_i h_i^\pi(k)\right) \qquad (3)$$

where $(\alpha_1, \ldots, \alpha_M)$ belongs to the probability simplex and π refers to the underlying transmission policy. In particular for symmetric networks, we set $\alpha_i = \frac{1}{M}$ for all i.

We consider three classes of policies: *centralized policies*, *stationary randomized policies*, and *distributed age-based policies*. Centralized policies need a central scheduler who receives information about all arrival processes and previous transmission actions, and can coordinate all the transmitters. When the number of transmitters M gets large, facilitating such scales of coordination is not feasible and we are hence interested in distributed mechanisms. In the class of stationary randomized policies, each transmitter sends its packet with some fixed probability that can be optimized ahead of time as a function of M and θ. Randomized policies are easy to implement in a distributed manner. Previous works [20], [21] fall into this class and they have the weakness of not utilizing local collision feedback at the transmitters. In the third class of policies, we utilize collision feedback to make age-based decisions at the transmitters in a distributed manner.

A. Notation

We use the notations $\mathbb{E}[\cdot]$ and $\Pr(\cdot)$ for expectation and probability, respectively. We denote scalars with lower case letters, e.g. s; vectors with underlined lowercase letters, e.g. $\underline{s}$, and matrices with boldface capital letters, e.g. $\mathbf{S}$. Notation $[\underline{s}]_i$ represents the i^{th} element of $\underline{s}$ and $[\mathbf{S}]_{ij}$ denotes the element in the i^{th} row and j^{th} column. Random variables are denoted by capital letters, e.g. S. We use M to denote the number of transmitters, K to denote the time horizon, and C to denote the capacity of a channel (under a given technology). The operator $(s)^+$ returns 0 if $s < 0$ and it returns s if $s \geq 0$. $\lfloor s \rfloor$ represents the largest integer j such that $j \leq s$. $O(\cdot)$ and $o(\cdot)$ represent the Big O and little o notations according to Bachmann-Landau notation, respectively.

[2]For any distributed transmission scheme, it is clear that the expected weighted sum of AoI increases with the number of source node M for any fixed arrival rate θ. Note that our problem setup allows M to become very large, so to offset the effect introduced by the number of source nodes, we consider the proposed normalized expected weighted sum AoI.

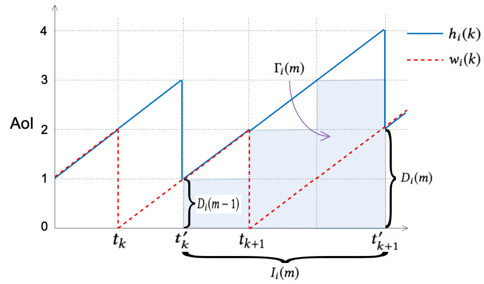


Fig. 1: an example of $D_i(m)$, $I_i(m)$, and $\Gamma_i(m)$

III. Lower Bound

We start by deriving two lower bounds on the achievable age performance. The first lower bound is derived by assuming that there is always a fresh packet to be transmitted (and hence delivered packets are assumed to experience unit-time delays). The second lower bound is derived by assuming that all packets are delivered instantaneously upon their arrivals (with unit-time delays, but without experiencing collisions). The former is tight as θ approaches 1 and the latter is tight when θ is small (when the inter-arrival time is the dominant term of the inter-delivery time).

Fix a large time horizon K and look at the packets of source i. Let $N_i(K)$ denote the number of delivered packets (from source i) up to and including time slot K. Now consider the m^{th} and $(m+1)^{th}$ *deliveries* at the receiver and denote the delivery time of them at the receiver by $T_i(m)$ and $T_i(m+1)$, respectively. The inter-delivery time

$$I_i(m) = T_i(m+1) - T_i(m)$$

is the time between these two consecutive deliveries. Upon arrival of the m^{th} delivered packet at the receiver, the age of information at the receiver drops to the value $D_i(m)$ which represents how much delay the packet has experienced in the system. Fig. 1 illustrates the introduced notation. Let L_i be the number of remaining time slots after the last packet delivery in source i. Now define $\Gamma_i(m)$ as the sum of age functions $h_i(k)$, where k is in the interval $[T_i(m), T_i(m+1))$:

$$\Gamma_i(m) = \sum_{k=T_i(m)}^{T_i(m)+I_i(m)-1} h_i(k) \tag{4}$$

$$= \frac{1}{2}I_i^2(m) - \frac{1}{2}I_i(m) + D_i(m-1)I_i(m). \tag{5}$$

It follows that in the limit of large K, we have

$$\lim_{K\to\infty} \mathbb{E}[J_K^\pi] = \lim_{K\to\infty} \mathbb{E}\left[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{m=1}^{N_i(K)}\Gamma_i(m)\right].$$

Using this formulation, we next lower bound NEWSAoI. Let C_{RA} denote the sum-capacity of the underlying random access channel. Note that in the limit of large K, $\frac{N_i(K)}{K}$ is the throughput of transmitter i and

$$\lim_{K\to\infty}\sum_{i=1}^{M}\frac{N_i(K)}{K} \leq C_{RA}. \tag{6}$$

Then, we have the following propositions.

Proposition 1. *For any transmission policy π,*

$$\lim_{K\to\infty} \mathbb{E}[J_K^\pi] \geq \frac{1}{2C_{RA}} + \frac{1}{2M}.$$

Proof. The proof is given in Appendix B. □

Proposition 2. *For any transmission policy π,*

$$\lim_{K\to\infty} \mathbb{E}[J_K^\pi] \geq \frac{1}{M\theta}.$$

Proof. The proof is given in Appendix C. □

Note that C_{RA} is not known in general. Nevertheless, any upper bound on C_{RA} gives a lower bound on the normalized age. Note that $C_{RA} \leq 1$, Proposition 1 and Proposition 2 thus lead to the following lower bound on NEWSAoI.

Theorem 1. *For any transmission policy π, we have*

$$\lim_{K\to\infty} \mathbb{E}[J_K^\pi] \geq \begin{cases} \dfrac{1}{M\theta} & \theta \leq \dfrac{2}{M+1} \\ \dfrac{1}{2} + \dfrac{1}{2M} & \theta > \dfrac{2}{M+1}. \end{cases}$$

IV. Centralized Scheduling

The first class of schemes that we consider are centralized schemes that avoid collision by scheduling transmitters one by one. In particular, Max-Weight policies are shown to perform close to optimal in various works such as [13], [14], [25]. Although such schemes are not practical (due to the scale of required coordination), it turns out that they provide useful intuitions and they also serve as a benchmark for comparison in Section V. We assume a central scheduler that can observe all arrival processes (at different source nodes) and coordinate/control all senders' actions in order to avoid collision.

Denote by $\lambda_i(k) = 1$ the event that transmitter i sends a packet and recall that $d_i(k)$ indicates delivery of packets. Note that if $\lambda_j(k) = 1$ for another source $j \neq i$, then the packets collide and no packets will be delivered. One can thus write

$$d_i(k) = \lambda_i(k)\prod_{j\neq i}\big(1-\lambda_j(k)\big). \tag{7}$$

The goal of a central scheduler is to select one source for transmission at each time. Denote $\underline{h}(k) = (h_1(k), h_2(k), \cdots, h_M(k))$. Following the work in [26], an age-based max-weight policy can be designed by considering the following Lyapunov function:

$$\mathcal{L}(\underline{h}(k)) = \sum_{i=1}^{M} h_i(k) \tag{8}$$

and minimizing its corresponding one-step Lyapunov Drift:

$$\Delta(\underline{h}(k)) = \mathcal{L}(\underline{h}(k+1)) - \mathcal{L}(\underline{h}(k)). \tag{9}$$

It turns out that an optimal max-weight policy selects, in each time slot k, the transmitter that offers the highest *age-gain* $\delta_i(k)$, defined below:

$$\delta_i(k) := h_i(k) - w_i(k). \tag{10}$$

$\delta_i(k)$ quantifies how much the instantaneous receiver's age of information reduces upon successful delivery from transmitter i. Proposition 3 states the above max-weight policy more formally.

Proposition 3. *For every time slot k, define*

$$\ell(k) = \arg\max_i \delta_i(k). \tag{11}$$

An optimal policy to minimize the one-step drift in (9) is to choose $\lambda_{\ell(k)}(k) = 1$ and $\lambda_j(k) = 0$ for all $j \neq \ell(k)$.

Remark 1. *We will show in Section VI how the notion of age-gain plays a central role also in the design of distributed age-based policies.*

V. Stationary Randomized Policies

One of the simplest distributed transmission policies that senders can adopt is perhaps the stationary randomized policy. In this class of policies, decisions are based only on availability of new packets. In particular, if transmitter i's queue is non-empty, then it sends its (latest) packet with probability $\lambda \in [0,1]$ independent across time slots $k = 1, 2, \ldots$.

Using the notations introduced in Section III, for stationary randomized policies, $\{(D_i(m-1), I_i(m), \Gamma_i(m))\}_m$ is a renewal-reward process and has stationary and ergodic distribution. Let (D_i, I_i, Γ_i) have the same (joint) distribution as $\big(D_i(m-1), I_i(m), \Gamma_i(m)\big)$ for $m \to \infty$ (i.e., when the system has reached stationarity). Also, note that D_i and I_i are independent. Following a similar analysis as in [27], we obtain

$$\begin{aligned}\lim_{K\to\infty} \mathbb{E}[J_K^\pi] &= \frac{1}{M^2}\sum_{i=1}^{M}\frac{\mathbb{E}[\Gamma_i]}{\mathbb{E}[I_i]} \\ &= \frac{1}{M^2}\sum_{i}^{M}\left(\mathbb{E}[D_i] - \frac{1}{2} + \frac{\mathbb{E}[I_i^2]}{2\mathbb{E}[I_i]}\right).\end{aligned} \tag{12}$$

Thus, (12) is reduced to

$$\lim_{K\to\infty} \mathbb{E}[J_K^\pi] = \frac{1}{M}\left(\mathbb{E}[D_1] - \frac{1}{2} + \frac{\mathbb{E}[I_1^2]}{2\mathbb{E}[I_1]}\right). \tag{13}$$

Under these assumptions, we have λ to be optimized. In order to characterize $NEWSAoI$, we need to find $\mathbb{E}[I_1]$, $\mathbb{E}[I_1^2]$, and $\mathbb{E}[D_1]$.

Note that the system has 2^M states depending on the queue state of each transmitter. Let $\underline{s}(k)$ denote the state at the beginning of time slot k before the arrival of new packets. State $\underline{s}(k) = \underline{s}$, $\underline{s} \in \{0,1\}^M$, corresponds to the event that transmitter i, $i = 1, 2, \ldots, M$, has $[\underline{s}]_i$ packets in its queue. For simplicity of presentation, we order the states lexicographically and rename them as $1, 2, \ldots, 2^M$ when needed (we

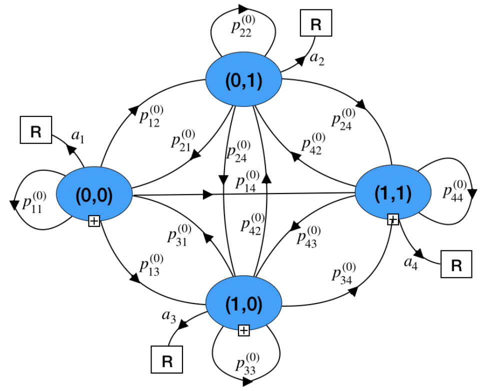


Fig. 2: Evolution of queues

use the two definitions of state interchangeably and make a distinction, when necessary, by using underlines for the vector notation). For $u, v \in \{1, \ldots, 2^M\}$, denote

$$p_{uv}^{(0)} = \Pr(s(k) = v \;\&\; d_1(k-1) = 0 | s(k-1) = u) \tag{14}$$

$$p_{uv}^{(1)} = \Pr(s(k) = v \;\&\; d_1(k-1) = 1 | s(k-1) = u) \tag{15}$$

$$p_{uv} = p_{uv}^{(0)} + p_{uv}^{(1)}. \tag{16}$$

We give the expressions of $p_{uv}^{(0)}$ and $p_{uv}^{(1)}$ for completeness in Appendix E. Note that the system's state evolves according to the transition probability matrix $\mathbf{P} = [p_{uv}]$ and has a unique stationary distribution $\underline{\pi}$. In particular, $\underline{\pi}$ is the unique solution of $\underline{\pi} = \underline{\pi}\mathbf{P}$ and we write

$$\pi_s := [\underline{\pi}]_s = \lim_{k\to\infty} \Pr(s(k) = s). \tag{17}$$

Now define a (virtual) absorbing state R to capture the event that a packet is delivered from source 1. Starting from a random state $\underline{s}$, we count the number of steps before the first delivery (i.e., before getting absorbed in state R). An example with two source nodes is given in Figure 2. The transition probabilities of this absorbing Markov chain is given by

$$\begin{bmatrix}\mathbf{Q} & \underline{a}' \\ \underline{0} & 1\end{bmatrix}$$

where

$$[\mathbf{Q}]_{uv} = p_{uv}^{(0)} \tag{18}$$

$$\begin{aligned}[\underline{a}]_u &= \Pr(d_1(k-1) = 1 | s(k-1) = u) \\ &= \sum_{v=1}^{2^M} p_{uv}^{(1)}.\end{aligned} \tag{19}$$

We can now find the first and second moments of the absorption time using the identity in [28, Chapter 3]. In particular, define

$$\underline{\mu} = (\mathbf{I} - \mathbf{Q})^{-1}\underline{1} \tag{20}$$

$$\underline{\nu} = (\mathbf{I} - \mathbf{Q})^{-1}(\mathbf{I} + \mathbf{Q})(\mathbf{I} - \mathbf{Q})^{-1}\underline{1} \tag{21}$$

where $\mathbf{I}$ is an identity matrix of size $2^M \times 2^M$, $\mathbf{Q}$ is the matrix of transition probabilities among transient states, and $\underline{1}$ denotes a tall all-one vector of size 2^M. Condition on starting in the transient state s, $s \in \{1, \dots, 2^M\}$, [28, Chapter 3] states that the expected number of steps before getting absorbed in state R is $\left[\underline{\mu}\right]_s$. Similarly, the second moment of absorption time from state s is $[\underline{\nu}]_s$.

Recall that in finding the inter-delivery time, we shall assume the initial state to be a feasible states after a successful delivery. We thus need to find the conditional probability of being in state s, condition on successful delivery in the previous time slot:

$$\tilde{\pi}_s := \lim_{k\to\infty} \Pr(s(k) = s | d_1(k-1) = 1) \tag{22}$$

$$= \lim_{k\to\infty} \frac{\sum_u \Pr(s(k) = s, s(k-1) = u, d_1(k-1) = 1)}{\Pr(d_1(k-1) = 1)} \tag{23}$$

$$= \frac{\sum_u \pi_u p_{us}^{(1)}}{\sum_v \pi_v \sum_{w=1}^{2^M} p_{vw}^{(1)}}. \tag{24}$$

Given the distribution on the starting transient states, we now arrive at the following expressions for the first two moments of the inter-delivery time I_1:

$$\mathbb{E}[I_1] = \sum_{s=1}^{2^M} \tilde{\pi}_s \cdot [\underline{\mu}]_s \tag{25}$$

$$\mathbb{E}[I_1^2] = \sum_{s=1}^{2^M} \tilde{\pi}_s \cdot [\underline{\nu}]_s \tag{26}$$

where $\tilde{\pi}_s$, $s = 1, 2, \cdots, 2^M$, is obtained from (24).

It remains to find the expected delay $\mathbb{E}[D_1]$. Recall that D_1 represents how much delay a packet has experienced in the system and note that $D_1 \geq 1$. We first prove that D_1 has a geometric distribution (see Appendix F) and then find its corresponding parameter q. Suppose a delivery occurs at the end of time slot k and suppose this corresponds to the m_k^{th} delivered packet. Let k (and hence m_k) be very large. Condition on the delivery of the mth packet, we have

$$q = \lim_{k\to\infty} \Pr(D(m_k) = 1 | d_1(k) = 1)$$
$$= \lim_{k\to\infty} \frac{\Pr(D(m_k) = 1, d_1(k) = 1)}{Pr(d_1(k) = 1)}$$

Since D is geometric with parameter q, we have

$$\mathbb{E}[D_1] = \frac{1}{q}. \tag{27}$$

Calculating q is straightforward and leads to the following expression for $\mathbb{E}[D_1]$ (see Appendix G):

$$\mathbb{E}[D_1] = \frac{\sum_v \pi_v \sum_{w=1}^{2^M} p_{vw}^{(1)}}{\sum_{i=2^{M-1}+1}^{2^M} p_{i,i-2^{M-1}}^{(1)} (\pi_i + \pi_{i-2^{M-1}})}. \tag{28}$$

Finally, substituting $\mathbb{E}[I_1]$, $\mathbb{E}[I_1^2]$ and $\mathbb{E}[D_1]$ into (13), we get the NEWSAoI under stationary randomized policies. Moreover, letting the NEWSAoI be the objective function and $0 \leq \lambda \leq 1$ be the constraint, we can also obtain the optimum stationary randomized policies by minimizing the NEWSAoI over all λ. We remark that the objective function is in the form of a fraction of polynomials and hence, for moderate values of M, the optimization can be done efficiently using methods such as [29], [30] that build on the sum of squares approximation framework [31].

Remark 2. *In a related work, [22] studies the age performance of stationary randomized policies in random access channels. The proposed framework assumes that the probability of collision is given and finds age in terms of this probability. Note, however, that the probability of collision depends on the arrival and departure rates of queues. Our work develops an alternative approach and provides an explicit characterization of the age performance.*

VI. Distributed Age-Based Policies

While stationary randomized policies allow simple distributed decision making at the transmitters, they are not designed to *prioritize* transmissions for the purpose of minimizing age of information. We propose a new class of distributed policies in which transmitter i decides whether or not to send its packet at time slot k depending on its local AoI, and in particular, based on $\delta_i(k)$ $\big($defined in (10)$\big)$.

To develop a deeper understanding of our proposed algorithm, let us focus on two regimes of operation assuming large M:

- The regime of infrequent arrivals, where $\theta \leq \frac{1}{eM}$,
- The regime of frequent arrivals, where $\theta > \frac{1}{eM}$.

The choice of these two regimes is made based on the well-established performance of slotted ALOHA with respect to rate (throughput) [32, Chapter 4]. As explained earlier in Section I, we will develop our framework particularly for the slotted-ALOHA random access technology, but we will expand on how it generalizes for other random access technologies in Section VI-E.

The basic idea of slotted ALOHA is as follows: At every time slot k, transmitters send their packets immediately upon arrival unless they are "backlogged" after a collision in which case they transmit with a backoff probability. In this section, we focus on Rivest's stabilized slotted ALOHA. In this algorithm, all arrivals are regarded as backlogged nodes that transmit with the backoff probability $p_b(k)$. Let $c(k) = 1$ denote the event that collision occurred at time k and $c(k) = 0$ denote the complementary event. Then, the backoff probability is calculated through a pseudo-Bayesian algorithm based on an estimate of the number of backlogged nodes $n(k)$ [32, Chapter 4.2.3]:

$$p_b(k) = \min\{1, \frac{1}{n(k)}\}$$
$$n(k) = \begin{cases} \min\{n(k-1) + M\theta + (e-2)^{-1}, M\} & \text{if } c(k) = 1 \\ \min\{\max\{M\theta, n(k-1) + M\theta - 1\}, M\} & \text{if } c(k) = 0 \end{cases}. \tag{29}$$

Remark 3. *Since we assume the buffer size is 1, then the number of backlogged transmitters is at most M.*

Performing slotted ALOHA, transmitters can reliably send packets with a sum-rate up to $\frac{1}{e}$ in a distributed manner

[32, Chapter 4.2.3]. Asymptotically, when $M \to \infty$, the probability of delivering a packet in each time slot is $1/e$, the probability of collisions is $1-2/e$, and the probability of having an idle channel is $1/e$ (see Appendix H). Note that since $M\theta \le \frac{1}{e}$, the expected total number of delivered packets in every time slot is $M\theta$.

We find the asymptotic NEWSAoI (in the limit of large M) in Theorem 2 below.

Theorem 2. *In the limit of large M, and given θ, if*

$$\lim_{M\to\infty} \theta M = \eta \le \frac{1}{e},$$

then (any) stabilized slotted ALOHA scheme achieves

$$\lim_{M\to\infty} \mathbb{E}[J^\pi] = \frac{1}{M\theta}.$$

Hence, slotted ALOHA is optimal for $\theta \le \frac{1}{eM}$.

Proof. The proof is presented in Appendix I. The idea is to divide the sources into two groups in every time slot k: sources with $\delta_i(k) = 0$ and sources with $\delta_i(k) > 0$. We show that (i) the contribution of the first group of sources to $NEWSAoI$ is equal to $\frac{1}{M\theta}$, and (ii) the second group constitutes only a vanishing fraction of the nodes and therefore, even though the sources in this group have larger $\delta_i(k)$'s, their total contribution vanishes as $M \to \infty$. □

A. Age-Based Thinning

When the arrival rate θ increases beyond $\frac{1}{eM}$, slotted ALOHA is not optimal anymore, leading to a large delay and age. Noting that the maximum rate that ALOHA can support is $\frac{1}{eM}$ per source, a natural question rises: What should the transmitters do in order to ensure small age of information at the destination when $\theta > \frac{1}{eM}$? A naive solution to the above question would be to have each transmitter randomly drop packets and perform at the effective rate $\frac{1}{eM}$. But as we argued in Theorem 2, this only leads to $NEWSAoI \approx e$ which means that we will not be able to benefit from the frequency of fresh packets to reduce age.

To benefit from the availability of fresh packets, we devise an age-based transmission policy in which transmitters prioritize packets that have larger age-gains. In particular, in each time slot k, transmitters find a proper threshold $\mathtt{T}(k)$ in order to distinguish and keep packets that offer high age-gains. The core idea is to still use the channel at its capacity (depending on the available technology) but to carefully select, in a distributed manner, what packets to send to minimize age.

Note that no matter how the transmission policy is designed, since it is distributed, it may happen that multiple transmitters try to access the channel at the same time, leading to collision. For simplicity and clarity of ideas, we will restrict attention to slotted ALOHA techniques to resolve such collisions[3].

[3]Other technologies such as CSMA achieve sum rates equal to 1, but for the purposes of this paper, we restrict attention to the class of ALOHA-type algorithms, We describe in Section VI-E how our findings extend to other contention resolution technologies.

Recall that $\delta_i(k)$ denotes the age-gain of scheduling transmitter i. At time k, we propose to discard a fresh packet at transmitter i if $0 \le \delta_i(k) < \mathtt{T}(k)$ and to keep it otherwise. We refer to this process as thinning and note that this is done based on the local AoI at the source/destination. The main underlying challenge is in the design of $\mathtt{T}(k)$. We propose two algorithms: an adaptive method of calculating $\mathtt{T}(k)$ for each time slot based on the local collision feedback and a fixed threshold value $\mathtt{T}^*$ that is found in advance and remains fixed for all time slots k.

In the remainder of this section, we assume that M is large, and $\theta > \frac{1}{eM}$. The following definition comes in handy in presenting our results.

Definition 2. *Consider transmitter i at time slot k. If $\delta_i(k) = m$, we say that transmitter i is an m-order node. Now let $\ell_m(k)$ be the fraction of m-order nodes in time slot k. We define $\{\ell_m(k)\}_{m=0}^\infty$ as the node distribution (of the age-gain) at time k.*

B. Adaptive Threshold

Let $\mathtt{T}(k)$ denote the threshold for decision making in slot k. We design $\mathtt{T}(k)$ in three steps:

(i) Compute an estimate of the node distribution of the age-gain;
(ii) Find $\mathtt{T}(k)$ based on the estimated distribution;
(iii) Update the estimate of the node distribution based on the chosen $\mathtt{T}(k)$ and the collision feedback.

Suppose the (estimated) node distribution $\{\ell_m(k-1)\}_{m=0}^\infty$ is known at (the end of) time slot $k-1$. We now describe how threshold $\mathtt{T}(k)$ is designed and how $\{\ell_m(k)\}_{m=0}^\infty$ is updated. For clarity of ideas, let us view time slot k in three stages: The first stage corresponds to the beginning of the time slot when new packets may arrive and replace the old packets. We denote the time just before the arrival of new packets by k^- and the time just after the arrival of packets by k^+. After the arrival of new packets, at time k^+, the source's AoI changes from $w_i(k^-)$ to $w_i(k^+)$ and the destination's AoI $h_i(k^+)$ remains the same as $h_i(k^-)$. So the age-gain values and their node distributions change. We denote the resulting node distribution in this stage by $\{\ell_m(k^+)\}_{m=0}^\infty$. In the second stage, transmitters determine the threshold $\mathtt{T}(k)$ based on $\{\ell_m(k^+)\}_{m=0}^\infty$. Transmissions happen according to the designed threshold $\mathtt{T}(k)$. In the third phase, at the end of time slot k when collision feedback is also available, the node distribution is once again estimated. We slightly abuse notation and denote the final estimate of the node distribution at the end of time slot k with $\{\ell_m(k)\}_{m=0}^\infty$. The aforementioned three stages of calculating $\mathtt{T}(k)$ is described next.

Stage 1: Suppose the node distribution $\{\ell_m(k-1)\}_m$ is known at the beginning of slot k before the arrival of new packets (we will describe later how $\{\ell_m(k)\}_m$ is estimated recursively at the end of each time slot in Stage 3). The expected fraction of $m-$order nodes, $m \ge 0$, that receive new packets is $\theta\ell_m(k-1)$. The order of these nodes increase and this changes the expected node distribution to $\{\ell_m(k^+)\}_m$ as a function of

$\{\ell_m(k-1)\}_m$. Let $a_m(k)$ denote the expected fraction of nodes that have just become $m-$order nodes at time k^+.

Lemma 1. *The expected fraction of nodes that have just become $m-$order nodes at time k^+ is*

$$a_m(k) = \theta^2 \sum_{j=0}^{m-1} \ell_j(k-1)(1-\theta)^{m-j-1} \tag{30}$$

and the expected node distribution of age-gain at time k^+ is

$$\ell_m(k^+) = \begin{cases} (1-\theta)\ell_0(k-1) & m=0 \\ (1-\theta)\ell_m(k-1) + a_m(k) & m \geq 1. \end{cases} \tag{31}$$

Proof. The proof is straightforward and delegated to Appendix J. □

Stage 2: The threshold $\mathtt{T}(k)$ is determined based on $\{\ell_m(k^+)\}_m$. We design $\mathtt{T}(k)$ such that the *effective arrival rate* of packets that have an age-gain above $\mathtt{T}(k)$ is close to $\frac{1}{e}$. In other words, we *thin* the arrival process using local age information. The critical point $\frac{1}{e}$ is the maximum sum arrival rate that ALOHA can support. So if the effective sum arrival rate falls below $\frac{1}{e}$, we do not use the full channel capacity[4] and if we operate above $\frac{1}{e}$, then we incur additional collisions and delay.

Recall that $a_m(k)$ is the expected fraction of nodes that have just become m-order nodes at time k^+ (coming from lower order nodes). So the total fraction of nodes whose age-gain would, for the first time, pass the threshold $\mathtt{T}(k)$ is

$$\sum_{m \geq \mathtt{T}(k)} a_m(k).$$

We propose to choose $\mathtt{T}(k)$ according to the following rule:

$$\mathtt{T}(k) = \max \left\{ t | \sum_{m \geq t} a_m(k) \geq \frac{1}{eM} \right\}. \tag{32}$$

Remark 4. *We chose $\mathtt{T}(k)$ to be the maximum threshold value that does not bring effective sum arrival rate below $\frac{1}{e}$. This is due to the integer nature of age and hence K. One can also time share between $\mathtt{T}(k)-1$ and $\mathtt{T}(k)$ to operate at an effective sum arrival rate equal to $\frac{1}{e}$.*

Remark 5. *The threshold $\mathtt{T}(k)$ is computed based on $\{\ell_m(k-1)\}$. Note that $\{\ell_m(k-1)\}$ is an estimate of the expected node distribution of age-gain. As we describe in stage 3, this estimate is found recursively by computing the expected node distribution condition on the estimated distribution in the previous time slot.*

Stage 3: Once the threshold $\mathtt{T}(k)$ is determined, each transmitter verifies locally if its age-gain is above the specified threshold. If so, it transmits its packet with probability $p_b(k)$ defined in (29) mimicking slotted ALOHA . If collision happens or if all nodes abstain from transmitting, then AoI at the destination increases by 1 for all sources. If only one node transmits, then its packet will be delivered successfully and the corresponding age at the destination drops to the source's AoI.

[4]Here, capacity refers to the maximum achievable sum rate under ALOHA

C. Estimating the node distribution

It remains to estimate $\ell_m(k)$ at the end of time slot k, which will serve in computing $\mathtt{T}(k+1)$ in the next time slot. We assume that at the end of time slot k, all transmitters are provided with collision feedback from the channel and we hence consider two cases separately: $c(k)=0$ and $c(k)=1$.

If collision has occurred, i.e., $c(k)=1$, then the order of nodes will not change:

$$\ell_m(k) = \ell_m(k^+), \quad m \geq 0. \tag{33}$$

Moreover, we estimate $n(k)$ from $n(k-1)$, as done in (29), by accounting for the expected number of active nodes in the next time slot plus the number of nodes who will be backlogged because of the collision that just occurred. In using (29), note that we have to replace $M\theta$ by the effective arrival rate $\frac{1}{e}$.

If there was no collision, i.e., $c(k)=0$, then either a packet was delivered or no packet was delivered. Since we design $\mathtt{T}(k)$ to impose (in the limit of large M) an effective sum arrival rate almost equal to $\frac{1}{e}$, then we conclude from Lemma 4 that the two events are almost equiprobable:

$$\lim_{k\to\infty} \Pr\big(\sum_{i=1}^{M} d_i(k) = 1, c(k)=0\big) \approx \frac{1}{e}$$
$$\lim_{k\to\infty} \Pr\big(\sum_{i=1}^{M} d_i(k) = 0, c(k)=0\big) \approx \frac{1}{e}.$$

Thus, condition on $c(k)=0$, a packet is delivered with probability $1/2$, i.e., the expected number of delivered packet is $1/2$ and by the inherent symmetry of the system, each active node has the same chance to deliver a new packet. For any $m \geq \mathtt{T}(k)$, a packet is delivered by m-order nodes with probability

$$r_m(k) = \frac{\ell_m(k^+)}{\sum_{t\geq \mathtt{T}(k)} \ell_t(k^+)}.$$

The expected number of $m-$ order nodes is $M\ell_m(k^+)$ and the expected number of delivered packets by m-order nodes (condition on $c(k)=0$) is $\frac{r_m(k)}{2}$. Therefore, the expected fraction of $m-$ order nodes with a successful delivery is:

$$\frac{1}{M}\min\big(\frac{r_m(k)}{2}, M\ell_m(k^+)\big).$$

Consequently, the update rule of the node distribution of age, $\{\ell_m(k)\}_m$, is given as follows:

$$\begin{aligned} \ell_0(k) =& \ell_0(k^+) + \sum_{m=\mathtt{T}(k)}^{\infty} \min\big(\frac{r_m(k)}{2M}, \ell_m(k^+)\big) \\ \ell_m(k) =& \ell_m(k^+), \qquad\qquad 1 \leq m \leq \mathtt{T}(k)-1 \\ \ell_m(k) =& \left(\ell_m(k^+) - \frac{r_m(k)}{2M}\right)^+, \qquad m \geq \mathtt{T}(k). \end{aligned} \tag{34}$$

Finally, in this case, the probability of transmitting a new packet is given by (29), where $M\theta$ is replace by the effective arrival rate $\frac{1}{e}$.

Algorithm 1 below describes the proposed distributed age-based transmission policy.

Algorithm 1 Adaptive Age-based Thinning

Set a large integer N and the time horizon K.
Set initial points: $h_i(0)=1$, $w_i(0)=0$ for $i=1,2,\cdots,M$; $c(0)=0$; $\mathtt{T}(0)=1$; $p_b(0)=1$; $n(0)=0$; $k=1$.
repeat
Step 1: Calculate $\{\ell_m(k^+)\}_{m=1}^N$ by (31).
Step 2: Calculate $\mathtt{T}(k)$ by (32).
Step 3: For transmitter i, $i=1,\ldots,M$: compute $\delta_i(k^+)=h_i(k^+)-w_i(k^+)$; if $\delta_i(k^+)<\mathtt{T}(k)$, then it does not transmit packets; if $\delta_i(k^+)\geq \mathtt{T}(k)$, then it transmits a packet with probability $p_b(k)$ (if it is not empty).
Step 4: If $c(k)=0$, calculate $\{\ell_m(k)\}_{m=1}^N$ by (33), and if $c(k)=1$, calculate $\{\ell_m(k)\}_{m=1}^N$ by (34). Calculate $p_b(k+1)$ by (29) $M\theta$ is replace by the effective arrival rate $\frac{1}{e}$.
until $k=K$
Calculate

$$NEWSAoI=\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=0}^{K}h_i(k).$$

D. Fixed Threshold

A simple variant of the age-based thinning method is found when the threshold $\mathtt{T}(k)=\mathtt{T}^*$ is fixed throughout the transmission phase. In particular, we design $\mathtt{T}^*$ ahead of time based on the node distribution in the stationary regime. By doing so, we cannot benefit from the collision feedback to adaptively choose $\mathtt{T}(k)$. However, this framework is preferable for deriving analytical results.

We use the framework and derivation we developed for adaptive thinning in order to find the optimal $\mathtt{T}^*$ as the limit behavior of $\mathtt{T}(k)$. The major difference here is in the update rules (33)-(34) because $c(k)$ is not known when $\mathtt{T}^*$ is designed. In particular, the update rule (33)-(34) is replaced by an average rule that weighs $c(k)=1$ with probability $1-\frac{2}{e}$ and $c(k)=0$ with probability $\frac{2}{e}$ (see Lemma 4).

Let $k\to\infty$. Abusing the notation slightly, define $\{\ell_m^*\}_{m=0}^\infty$ and $\{\ell_m^{+*}\}_{m=0}^\infty$ as the limit of $\{\ell_m(k)\}_{m=0}^\infty$ and $\{\ell_m(k^+)\}_{m=0}^\infty$, respectively. From (31), the update rule of Stage 1 implies

$$\begin{aligned}&\ell_0^{+*}=(1-\theta)\ell_0^*\\&\ell_m^{+*}=(1-\theta)\ell_m^*+a_m^* \qquad m\geq 1\end{aligned}\tag{35}$$

where

$$a_m^*=\theta^2\sum_{j=0}^{m-1}\ell_j^*(1-\theta)^{m-j-1}\qquad m\geq 1.\tag{36}$$

The threshold proposed in Stage 2 is then

$$\mathtt{T}^*=\max\Big\{t|\sum_{m\geq t}a_m^*\geq\frac{1}{eM}\Big\}.\tag{37}$$

Next, consider Stage 3. In contrast to Section VI-C, we don't have collision feedback available and hence estimating the fraction of $m-$order nodes at the end of time slot k will account for $c(k)=1$ with probability $1-\frac{2}{e}$ and $c(k)=0$ with probability $\frac{2}{e}$ (see Lemma 4). We hence obtain

$$\begin{aligned}&\ell_0^*=\ell_0^{+*}+\frac{1}{eM}\\&\ell_m^*=\ell_m^{+*}, && 1\leq m\leq \mathtt{T}^*-1\\&\ell_m^*=\ell_m^{+*}-\frac{r_m^*}{eM}, && m\geq\mathtt{T}^*\end{aligned}\tag{38}$$

where

$$r_m^*=\ell_m^{+*}/\sum_{i=\mathtt{T}}^{\infty}\ell_i^{+*}.$$

Putting together (35) - (38), we obtain

$$\begin{aligned}&\ell_0^*=(1-\theta)\ell_0^*+\frac{1}{eM}\\&\ell_m^*=(1-\theta)\ell_m^*+a_m^* && 1\leq m\leq\mathtt{T}^*-1\\&\ell_m^*=(1-\theta)\ell_m^*+a_m^*-\frac{r_m^*}{eM} && m\geq\mathtt{T}^*\end{aligned}\tag{39}$$

and conclude the following lemma (see Appendix K for the proof).

Lemma 2. *As $k\to\infty$, the stationary distributions $\{\ell_m^*\}_{m=0}^{\mathtt{T}^*-1}$, $\{\ell_m^{+*}\}_{m=1}^{\mathtt{T}^*}$ and $\{a_m^*\}_{m=1}^{\mathtt{T}^*}$ are given by*

$$\ell_m^*=\begin{cases}\frac{1}{eM\theta} & m=0\\ \frac{1}{eM} & 1\leq m\leq\mathtt{T}^*-1\end{cases}\tag{40}$$

$$\ell_m^{+*}=\frac{1}{eM}\qquad 1\leq m\leq\mathtt{T}^*-1\tag{41}$$

$$a_m^*=\frac{\theta}{eM}\qquad 1\leq m\leq\mathtt{T}^*.\tag{42}$$

The closed form expression of the fixed threshold $\mathtt{T}^*$ is given below (see Appendix L for the proof) and Algorithm 2 describes the age-based transmission policy.

Theorem 3. *The fixed threshold $\mathtt{T}^*$ in (37) has the following closed form expression:*

$$\mathtt{T}^*=\max\big(1,\lfloor eM-\frac{1}{\theta}+1\rfloor\big).$$

We finally prove asymptotically (as $M\to\infty$) that the distributed age-based policy proposed in Algorithm 2 significantly reduces age when $1/\theta=o(M)$. Recall that at $\theta=\frac{1}{eM}$, we have $\lim_{M\to\infty}NEWSAoI=e$. For larger arrival rates θ where $1/\theta=o(M)$, we prove that Algorithm 2 sharply reduces AoI from e to $\frac{e}{2}$.

Theorem 4. *For any $\theta=\frac{1}{o(M)}$,*

$$\lim_{M\to\infty}\mathbb{E}[J^\pi]=\frac{e}{2}.$$

Proof. The proof is given in Appendix M. □

Roadmap of the proof: In every time slot k, the sources can be divided into two groups: 1) source i with $\delta_i(k)<\mathtt{T}^*$ 2) source i with $\delta_i(k)\geq\mathtt{T}^*$. The first group of sources have main contribution to the $NEWSAoI$ (equals to $\frac{e}{2}$) in the limit $M\to\infty$, the contribution of the second group of sources to the $NEWSAoI$ vanishes when $M\to\infty$.

Algorithm 2 Stationary Age-based Thinning

Set the time horizon K.
Set initial points: $h_i(0) = 1, w_i(0) = 0$ for $i = 1, 2, \cdots, M$; $c(0) = 0$; $T(0) = 1$; $p_b(0) = 1$; $n(0) = 0$; $k = 1$.
Calculate $\mathtt{T}^* = \max\big(1, \lfloor eM - \frac{1}{\theta} + 1 \rfloor\big)$.
repeat
Step 1: For i^{th} source node, compute $\delta_i(k) = h_i(k) - w_i(k)$, if $\delta_i(k) < \mathtt{T}^*$, then it does not transmit packets; if $\delta_i(k) \geq \mathtt{T}^*$, then it transmits a packet with probability $p_b(k)$ (if it is not empty).
Step 2: Calculate $p_b(k)$ by (29) in which $M\theta$ is replaced by $\min(M\theta, e^{-1})$.
until $k = K$
Calculate

$$NEWSAoI = \frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=0}^{K} h_i(k).$$

E. Extensions to Other Random Access Technologies

So far, we restricted attention to slotted ALOHA as the main random access technology. However, in the past decade, novel technologies such as Carrier Sensing Multiple Access (CSMA) technologies have proved significant improvement in terms of throughput and it is hence of interest to know how they perform with regard to age, especially since they are known to have large delays. In this section, we outline how the age-based thinning method described in Section VI-D (with a fixed threshold) can be applied to any given random access technology).

Suppose we are given a random access technology with maximum throughput C which we will treat as a black box. Consider the age-based thinning process in two steps in every time slot: (i) the threshold $\mathtt{T}^*$ is calculated, (ii) all nodes with age-gains larger than or equal to $\mathtt{T}^*$ become active and transmit using the prescribed random access technology. This process can be combined with any random access technology.

Since the throughput of the proposed random access policy is C, the expected number of delivered packets per time slot is $\min(M\theta, C)$. Therefore, (35) remains the same and (37) will take the following form:

$$\mathtt{T}^* = \max\{t | \sum_{m\geq t} a^*_m \geq \frac{C}{M}\}. \tag{43}$$

Following a similar argument as in Section VI-D, the equations in (38) can be written more generally as follows:

$$\begin{aligned} \ell^*_m &= \ell^{+*}_0 + \min(\theta, \frac{C}{M}) \\ \ell^*_m &= \ell^{+*}_m \qquad\qquad 1 \leq m \leq \mathtt{T}^* - 1 \\ \ell^*_m &= \ell^{+*}_m - r^*_m \min(\theta, \frac{C}{M}) \quad m \geq \mathtt{T}^* \end{aligned} \tag{44}$$

where

$$r^*_m = \ell^{+*}_m / \sum_{i=\mathtt{T}^*}^{\infty} \ell^{+*}_i.$$

Combining (35), (43), (44), we thus find

$$\ell^*_0 = \begin{cases} \min(1, \frac{C}{M\theta}) & m = 0 \\ \min(\theta, \frac{C}{M}) & 1 \leq m \leq \mathtt{T}^* - 1 \end{cases} \tag{45}$$

$$\ell^{+*}_m = \min(\theta, \frac{C}{M}) \qquad 1 \leq m \leq \mathtt{T}^* - 1 \tag{46}$$

$$a^*_m = \min(\theta^2, \frac{\theta C}{M}) \qquad 1 \leq m \leq \mathtt{T}^*. \tag{47}$$

Moreover, the threshold T^* takes a simple closed-form expression as stated below (and proved in Appendix N).

Theorem 5. *The fixed threshold* $\mathtt{T}^*$ *in* (43) *has the following closed form expression:*

$$\mathtt{T}^* = \max\big(1, \lfloor \frac{M}{C} - \frac{1}{\theta} + 1 \rfloor\big).$$

Using this result, Algorithm 3 proposes a distributed age-based thinning method for any given random access technology with capacity C.

Algorithm 3 General Stationary Age-based Thinning

Set the time horizon K.
Set initial points: $h_i(0) = 1, w_i(0) = 0$ for $i = 1, 2, \cdots, M$; $c(0) = 0$; $T(0) = 1$; $p_b(0) = 1$; $n(0) = 0$; $k = 1$.
Calculate the threshold $\mathtt{T}(C) = \max\big(1, \lfloor \frac{M}{C} - \frac{1}{\theta} + 1 \rfloor\big)$.
repeat
For the source node i, compute $\delta_i(k) = h_i(k) - w_i(k)$. If $\delta_i(k) < \mathtt{T}(C)$ re main silent; If $\delta_i(k) \geq \mathtt{T}(C)$, transmits according to the policy prescribed by the given random access technology.
until $k = K$
Calculate

$$NEWSAoI = \frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=0}^{K} h_i(k).$$

Finally, we prove an analogue to Theorem 4, showing that age-based thinning reduces age to $\frac{1}{2C}$ as θ increases.

Theorem 6. *For any* $\theta = \frac{1}{o(M)}$, $\lim_{M\to\infty} \mathbb{E}[J^\pi] = \frac{1}{2C}$.

Proof. The proof of Theorem 6 is given in Appendix O. □

Remark 6. *The results in this section are stronger than [24] in three aspects: (i) we gave a simple and explicit expression for the threshold T^*, while the threshold has to be computed numerically in [24]; (ii) we found the asymptotical $NEWSAoI$ analytically; (iii) the threshold in this section can be applied not only to CSMA, but also to any other transmission policy.*

Roadmap of the proof: In every time slot k, the sources can be divided into two groups: 1) source i with $\delta_i(k) < ttT^*$ 2) source i with $\delta_i(k) > 0$. The first group of sources have main contribution to the $NEWSAoI$ (equals to $\frac{1}{2C}$) in the limit $M \to \infty$, the contribution of the second group of sources to the $NEWSAoI$ vanishes when $M \to \infty$.

VII. NUMERICAL RESULTS

In this section, we verify our findings through simulations.

Figure 3a shows the normalized age under adaptive and stationary age-based transmission policies for $M = 50, 100, 500$. For stationary age-based policies, the normalized age converges to $\frac{e}{2}$ when M is large, validating our findings in Theorem 4. The performance of the adaptive policy is better than that of the stationary age-based policy for $\theta > \frac{1}{M}$ and the efficacy (the gap between the two curves) increases with θ. Since the maximum sum throughput of slotted ALOHA is $\frac{1}{e}$, one may ask if this contradicts the lower bound of Proposition 2. To answer this question, we remark that the adaptive age-based transmission policy is *not* a slotted ALOHA scheme and therefore the maximum throughput of slotted ALOHA would not apply. As a matter of fact, Fig. 3b shows that the throughput of the scheme increases beyond $\frac{1}{e}$ with θ, supporting Proposition 1. One can also observe that adaptive policy performs worse than the stationary policy for $\theta \leq \frac{1}{M}$. We believe this is because the estimation of the node distribution $\{\ell_m\}_{m\geq 0}$ is imprecise for small θ.

Finally the age-performance of our proposed distributed age-based policies are compared with the lower bounds of Section III, state-of-the-art distributed schemes such as [20], as well as centralized Max-Weight policies such as [26]. For clarity, we consider two regimes of θ: $\theta \in (\frac{1}{M}, 1]$ (see Fig. 3c), and $\theta \in (0, \frac{1}{M}]$ (see Fig. 3d). Fig. 3d, in particular, shows that when $\theta \leq \frac{1}{eM}$, the normalized age of slotted ALOHA coincides with centralized Max-Weight policies and the lower bound of Proposition 2. When θ increases beyond $\frac{1}{eM}$, our proposed age-based thinning methods provide significant gains compared to randomized stationary and slotted ALOHA schemes. Finally, we numerically observe that the normalized age of the centralized Max-Weight policy is approximately attained by stationary age-based thinning in CSMA.

VIII. FUTURE RESEARCH

Future research includes generalization to accommodate 1) dynamic channels, i.e., the number of nodes M, or the arrival rates θ are time-variant 2) asymmetric channels, i.e., the arrival rates θ_i is different. In the first case, the method we proposed above can be applied directly. Suppose the expressions of the number of nodes, $M(k)$ and the arrival rates, $\theta(k)$ are known, we replace M and θ by $M(k)$ and $\theta(k)$, respectively, in every time slot. Sequently, the fixed threshold hold $\mathtt{T}^*$ is also a time-variant variable, $\mathtt{T}^*(k)$. In the second case, the method we proposed above can not be applied directly. This is because we use the profile of all sources as an estimate on any individual source. A more general estimation should be proposed in the second case.

REFERENCES

[1] V. R. S.Kaul, M.Gruteser and J. Kenny, "Minimizing age of information in vehicular networks," in *2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, June 2011, pp. 350 – 358.

[2] R. Y. S, Kaul and M. Gruteser, "On piggybacking in vehicular networks," in *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, June 2011, pp. 1 – 5.

[3] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *2012 Proceedings IEEE INFOCOM*, March 2012, pp. 2731–2735.

[4] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1807–1827, March 2019.

[5] M. Costa, M. Codreanu, and A. Ephremides, "On the age of information in status update systems with packet management," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1897–1910, April 2016.

[6] R. D. Yates, E. Najm, E. Soljanin, and J. Zhong, "Timely updates over an erasure channel," in *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 316–320.

[7] P. Mayekar, P. Parag, and H. Tyagi, "Optimal lossless source codes for timely updates," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 1246–1250.

[8] R. Devassy, G. Durisi, G. C. Ferrante, O. Simeone, and E. Uysal-Biyikoglu, "Delay and peak-age violation probability in short-packet transmissions," *CoRR*, vol. abs/1805.03271, 2018. [Online]. Available: http://arxiv.org/abs/1805.03271

[9] E. Najm, E. Telatar, and R. Nasser, "Optimal age over erasure channels," *CoRR*, vol. abs/1901.01573, 2019. [Online]. Available: http://arxiv.org/abs/1901.01573

[10] B. Buyukates, A. Soysal, and S. Ulukus, "Age of information in multihop multicast networks," *CoRR*, vol. abs/1812.10455, 2018. [Online]. Available: http://arxiv.org/abs/1812.10455

[11] S. Farazi and A. G. Klein and J. A. McNeill and D. Richard Brown, "On the Age of Information in Multi-Source Multi-Hop Wireless Status Update Networks,," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC),*, 2018, pp. 1 – 5.

[12] A. M. Bedewy, Y. Sun, and N. B. Shroff, "The age of information in multihop networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 1248–1257, Jun. 2019. [Online]. Available: https://doi.org/10.1109/TNET.2019.2915521

[13] E. U.-B. R. S. I. Kadota, A. Sinha and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2637 – 2650, Dec 2018.

[14] A. S. I. Kadota and E. Modiano, "Scheduling algorithms for optimizing age of information in wireless networks with throughput constraints," *IEEE/ACM Transactions on Networking*, vol. 27, no. 4, pp. 1359 – 1372, 2019.

[15] Y. S. A. M. Bedewy and N. B. Shroff, "Minimizing the age of information through queues," *IEEE Transactions on Information Theory*, vol. 65, no. 8, pp. 5215 – 5232, Aug. 2019.

[16] D. Y. Q. He and A. Ephremides, "Optimal link scheduling for age minimization in wireless systems," *IEEE Transactions on Information Theory*, vol. 64, no. 7, pp. 5381 – 5394, July 2018.

[17] Y. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018.

[18] Y. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Transactions on Mobile Computing*, 2019.

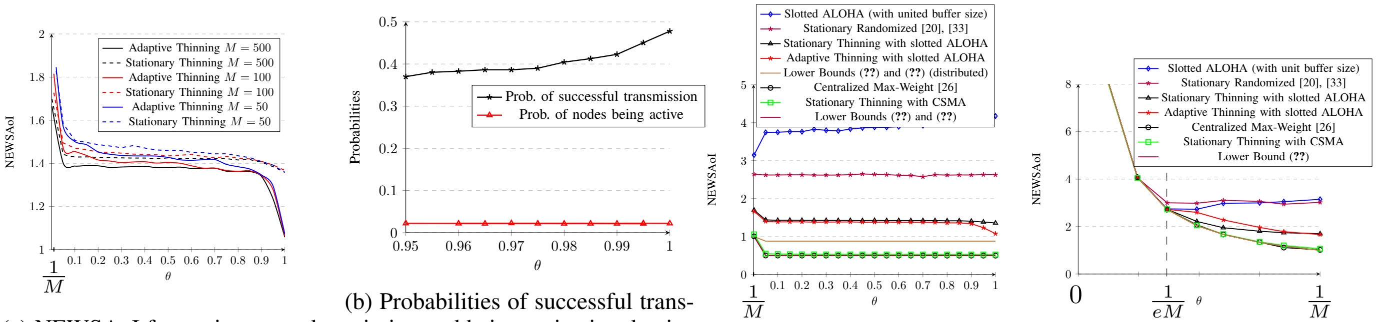
[19] C. Joo and A. Eryilmaz, "Wireless Scheduling for Information Freshness and Synchrony: Drift-Based Design and Heavy-Traffic Analysis," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2556 – 2568, Dec 2018.

[20] S. K. Kaul and R. D. Yates, "Status updates over unreliable multiaccess channels," arXiv:1705.02521v1, May 2017.

[21] R. Talak, S. Karaman, and E. Modiano, "Distributed scheduling algorithms for optimizing information freshness in wireless networks," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2018, pp. 1–5.

[22] A. E. A. Kosta, N. Pappas and V. Angelakis, "Age of information performance of multiaccess strategies with packet management," arXiv:1812.09201v2, Jul 2019.

[23] X. Z.-S. Z. Z. Jiang, B. Krishnamachari and Z. Niu, "Timely status update in massive iot systems: Decentralized scheduling for wireless uplinks," arXiv:1801.03975v1, Jan 2018.



(a) NEWSAoI for stationary and adaptive age-based policies.

(b) Probabilities of successful transmission and being active in adaptive thinning policies.

(c) $M = 500$, $\theta \in (\frac{1}{M}, 1]$

(d) $M = 500$, $\theta \in (0, \frac{1}{M}]$

Fig. 3: Normalized age $NEWSAoI$ and success transmission probabilities

[24] S. Z.-Z. N. Z. Jiang, B. Krishnamachari, "Can decentralized status update achieve universally near-optimal age-of-information in wireless multiaccess channels?" arXiv:1803.08189v1, Mar 2018.

[25] Igor Kadota and Eytan Modiano, "Minimizing the age of information in wireless networks with stochastic arrivals," *CoRR*, vol. abs/1905.07020, 2019. [Online]. Available: http://arxiv.org/abs/1905.07020

[26] I. Kadota and E. Modiano, "Minimizing the age of information in wireless networks with stochastic arrivals," arXiv: 1905.07020v1, 2019.

[27] X. Chen and S. S. Bidokhti, "Benefits of coding on age of information in broadcast networks," arXiv:1904.10077, 2018.

[28] J. G. Kemeny and J. L. Snell, *Finite Markov Chains (Second ed.)*. New York Berlin Heidelberg Tokyo: Springer-Verlag, 1976.

[29] Z. Y. E. L. Kaltofen, B. Li and L. Zhi, "Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients," *Journal of Symbolic Computation*, vol. 47, pp. 1–15, Jan 2012.

[30] A. Z. A. Pizzo and L. Sanguinetti, "Solving fractional polynomial problems by polynomial optimization theory," arXiv: 1806.07220v1, Jun 2018.

[31] A. Ahmadi and G. Hall, "Sum of squares basis pursuit with linear and second order cone programming," *Algebraic and Geometric Methods in Discrete Mathematics, Contemp. Math*, pp. 27–53, 2015.

[32] Bertsekas, Dimitri and Gallager, Robert, *Data Networks (2Nd Ed.)*. Prentice-Hall, Inc., 1992.

[33] H. H. Xingran Chen, Konstantinos Gatsis and S. S. Bidokhti, "Age of information in random access channels," arXiv:1912.01473, 2019.

[34] E. Modiano, *Packet multiple access and the Aloha protocol*, Massachusetts Institute of Technology Department of Aeronautics and Astronautics, 2009.

Appendix A
Sufficiency of Unit Buffer Size

Consider two types of policies: policies with buffer size 1, denoted by π_1, and policies with larger buffer sizes, denoted by π_2. To differentiate the two policies and their corresponding queues, we label the packets inside the queues by *new* and *old*. A new packet in a queue refers to the latest arrival. A packet in a queue is considered old if there is a newer packet in the same queue or if the packet (or a fresher packet) from that source is already delivered at the receiver. In the following, we refer to the freshest old packet as the old packet. At a given time slot, denote the new packet and the old packet of source i by $p_{new}^{(i)}$ and $p_{old}^{(i)}$, respectively. Denote the arrival times of the new and old packets as $t_i^{(\mathrm{n})}$ and $t_i^{(\mathrm{o})}$. It is clear that $t_i^{(\mathrm{n})} > t_i^{(\mathrm{o})}$. We will show that no matter what policy π_2 does, there is always a policy of type π_1 whose resulting age is at least as low as π_2 with respect to every source node.

At time slot t', suppose policy π_2 chooses certain action, then we design policy π_1 to follow the same action with the new packet. In this time slot, under π_2 a subset of sources transmit packets. Denote the index of these sources by $\mathcal{I}$. For the sources which do not transmit packets, the AoI under both policies will increase by 1. For the sources in $\mathcal{I}$, we have the following two cases:

Case 1. Suppose collision happens in time slot t'. Then, no packet is delivered, and the AoI of these sources under both policies will increase by 1.

Case 2. If a packet is delivered, which implies the cardinal of $\mathcal{I}$, $|\mathcal{I}| = 1$. Denote the index of this source by i. Then at the next time slot, the AoI under π_1 drops to $h_i^{\pi_1}(t'+1) = t' - t_i^{(\mathrm{n})} + 1$, and the AoI under π_2 drops to $h_i^{\pi_2}(t'+1) = t' - t_i^{(\mathrm{o})} + 1 > h_i^{\pi_1}(t'+1)$. This means that from t' onward $h_i^{\pi_2}(t)$ will be point-wise larger or equal to $h_i^{\pi_1}(t)$, $t > t'$.

Appendix B
Proof of Proposition 1.

Consider any transmission policy and a large time-horizon K. The $NEWSAoI$ defined in (3) can be re-written in terms of $\Gamma_i(m)$:

$$\begin{aligned} J_K^{\pi} =& \frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}h_i(k) \\ =& \frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\Big(\sum_{m=1}^{N_i(K)}\Gamma_i(m)+\frac{1}{2}L_i^2+D_i(N_i(K))L_i-\frac{1}{2}L_i\Big). \end{aligned}$$

Since $D_i(m) \geq 1$ for all $1 \leq m \leq N_i(K)$, we can lower bound (3) by substituting $D_i(m-1) = 1$. Using similar steps as [13, Eqns. (9) - (14)], we find

$$J^{\pi} \geq \lim_{K\to\infty}\frac{1}{2M^2}\sum_{i=1}^{M}\frac{K}{N_i(K)}+\frac{1}{2M}. \tag{48}$$

Recall that $N_i(K)$ is the total number of packets delivered by source i. In the limit of $K \to \infty$, $\frac{N_i(K)}{K}$ is the throughput of source i. By the model assumption, in every time slot, at most one packet is delivered in the system. Therefore,

$$\sum_{i=1}^{M}\lim_{K\to\infty}\frac{N_i(K)}{K} \leq C_{RA}.$$

Now note that by the Cauchy-Schwarz inequality, we have

$$\sum_{i=1}^{M}\lim_{K\to\infty}\frac{N_i(K)}{K}\sum_{i=1}^{M}\lim_{K\to\infty}\frac{K}{N_i(K)}\geq M^2,$$

then

$$\sum_{i=1}^{M}\lim_{K\to\infty}\frac{K}{N_i(K)}\geq\frac{M^2}{C_{RA}}.$$

Inserting this back into (48), we obtain

$$J^\pi\geq\frac{1}{2C_{RA}}+\frac{1}{2M}.$$

Appendix C
Proof of Proposition 2.

Suppose all packets are delivered instantaneously with one time-unit delay and without experiencing collisions. A lower bound to NEWSAoI in this scenario constitutes a lower bound to NEWSAoI in our setup. Let $X_i(m)$ denote the inter arrival time between the mth and $m+1$th packets. $\{X_i(m)\}_m$ is a geometric iid sequence. Under the assumption of instantaneous delivery, $I_i(m)=X_i(m)$. It hence follows from (4) that

$$\Gamma_i(m)=\sum_{k=T_i(m)}^{T_i(m)+X_i(m)-1}h_i(k)=\frac{1}{2}X_i(m)^2+\frac{1}{2}X_i(m).$$

Thus, similar with [13], the time-average AoI of source i is

$$\mathbb{E}[h_i]=\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}h_i(k)=\frac{\mathbb{E}[X^2]}{2\mathbb{E}[X]}+\frac{1}{2}.$$

Since X has a geometric distribution with parameter θ, we find

$$\mathbb{E}[h_i]=\frac{2-\theta}{2\theta}+\frac{1}{2}$$

and conclude

$$\mathbb{E}[J^\pi]\geq\frac{1}{M\theta}.$$

Appendix D
Proof of Proposition 3.

First consider a source node i whose queue is empty. This means that no new packet has arrived at that transmitter since the last delivery (from that source node) at the receiver; i.e., $h_i(k)=w_i(k)$ and hence $\delta_i(k)=0$. Such nodes i are thus irrelevant because $\delta_j(k)\geq 0$ for all source nodes j. Now consider nodes with non-empty queues. Among these nodes, $d_i(k)$ is non-zero if and only if $(\lambda_1,\dots,\lambda_M)$ is a vector consisting of 0's except for $\lambda_i=1$. Hence at most one $d_i(k)$ can be equal to 1. Call the corresponding source node $\ell(k)$. Expression (9) is minimized when $d_{\ell(k)}(k)$ picks the largest $\delta_j(k)$.

Appendix E
Expressions of $p_{uv}^{(0)}$ and $p_{uv}^{(1)}$

Lemma 3. *Consider $p_{uv}^{(1)}$ and $p_{uv}^{(0)}$ defined in (14) and (15). From $\underline{u}$ to $\underline{v}$, suppose m_1 0s remain the same, m_2 1s remain the same, m_3 0s become 1s, and m_4 1s become 0s. Denote $\underline{m}_{uv}=(m_4,[\underline{u}]_1,[\underline{v}_1])$, $\tilde{\theta}=\theta^{m_3}\bar{\theta}^{m_1-1}$ and $\tilde{\lambda}=\lambda\bar{\lambda}^{m_2+m_3}$[5],*

$$p_{uv}^{(1)}=\begin{cases}\theta\tilde{\theta}\tilde{\lambda} & \underline{m}_{uv}=(0,0,0)\\ \bar{\theta}\tilde{\lambda}\tilde{\theta} & \underline{m}_{uv}=(1,1,0)\\ 0 & \textit{otherwise}\end{cases}\tag{49}$$

and

$$p_{uv}^{(0)}=\begin{cases}\bar{\theta}\tilde{\theta}+\tilde{\theta}\tilde{\lambda}\big((m_1-1)\theta-\frac{\bar{\theta}(m_3+m_2)}{\bar{\lambda}}\big),\underline{m}_{uv}=(0,0,0)\\ \bar{\theta}\tilde{\theta}+\tilde{\theta}\tilde{\lambda}\big(m_1\theta-\frac{\bar{\theta}(m_3+m_2)}{\bar{\lambda}}\big),\underline{m}_{uv}=(0,*,1)\\ \theta\tilde{\theta}\tilde{\lambda},\underline{m}_{uv}=(1,0,*),(1,1,1)\\ 0,\textit{otherwise}\end{cases}\tag{50}$$

where $*$ means the corresponding element is 0 or 1.

Proof. Denote $\underline{m}_{uv}=(m_4,[\underline{u}]_1,[\underline{v}]_1)$. Recall that $\underline{u}$ and $\underline{v}$ represent two positive integers between $[1,2^M]$ and also represent the corresponding two 2^M dimensional vectors (in which every element is 0 or 1), then $[\underline{u}]_1$ and $[\underline{v}]_1$ is 0 or 1. Collisions must happen when two or more packets are transmitted, then m_4 is 0 or 1. So each element of $\underline{m}_{uv}$ is 0 or 1, we have 8 cases. First, we consider $p_{uv}^{(1)}$, i.e., a packet is delivered from 1^{st} source. Suppose $\underline{m}_{uv}=(0,0,0)$, note that $d_1(k-1)=1$, then a packet has arrived at source 1 at the beginning of time slot $k-1$, and it is delivered at the end of the same slot, so

$$p_{uv}^{(1)}=\theta\lambda\cdot\bar{\theta}^{m_1-1}\theta^{m_3}\bar{\lambda}^{m_2+m_3}=\theta\tilde{\theta}\tilde{\lambda}.$$

If $\underline{m}_{uv}=(0,0,1)$, then a new packet has arrived at source 1 and it remains at source 1. This is contradicted to $d_1(k-1)=1$, so $p_{uv}^{(1)}=0$. By similar analysis, we get (49).

Similarly, we consider $p_{uv}^{(0)}$. Suppose $\underline{m}_{uv}=(0,0,0)$, note that $d_1(k-1)=0$, then we have two cases: 1) no successful delivery occurs, 2) a new packet has arrived at an empty source (except source 1), and then it is delivered. Then,

$$\begin{aligned}p_{uv}^{(0)}=&\theta^{m_3}\bar{\theta}^{m_1}\big(1-(m_2+m_3)\lambda\bar{\lambda}^{m_2+m_3-1}\big)\\&+\theta^{m_3+1}\bar{\theta}^{m_1-1}(m_1-1)\lambda\bar{\lambda}^{m_3+m_2}\\=&\bar{\theta}\tilde{\theta}+\tilde{\theta}\tilde{\lambda}\big((m_1-1)\theta-\frac{\bar{\theta}(m_3+m_2)}{\bar{\lambda}}\big).\end{aligned}$$

By similar analysis, we get (50). □

Appendix F
D_1 has a geometric distribution

First consider the special case of $M=1$ in which there exists only one source. The case of arbitrary M is similar and discussed afterwards.

Suppose source 1 is the only source in the network. In every time slot, a packet arrives at the source with probability

[5]Here, $\tilde{\theta}$ and $\tilde{\lambda}$ are functions of m_1,m_2,m_3,m_4. For notational clearness, we slightly abuse the notation $\tilde{\theta}$ and $\tilde{\lambda}$.

θ, and delivered with probability λ. If there is a packet at the source (the queue is non-empty), it is replaced by a new packet with probability θ, and is delivered with probability λ. Since these two events are independent, this packet remains at the source for one time slot with probability $\bar{\lambda}\bar{\theta}$.

Note that every packet is either delivered or replaced. Suppose the l^{th} packet stays in the channel for $v(l)$ time slots, then $\{v(l)\}_{l=1}^{\infty}$ are i.i.d. Recall that D_1 is how much delay a packet has experienced in the system, so suppose the l^{th} packet arrived at the system at the beginning of time slot t_0 and is delivered at the end of time slot t_1, where t_0, t_1 are both large. Then we have

$$\begin{aligned} &\Pr\{D_1 = t_1 - t_0 + 1\} \\ =& \Pr\{v(l) = t_1 - t_0 + 1 | d_1(t_1) = 1\} \\ =& \frac{\Pr\{v(l) = t_1 - t_0 + 1, d_1(t_1) = 1\}}{\Pr\{d_1(t_1) = 1\}}. \end{aligned} \tag{51}$$

It is straightforward to write:

$$\Pr\{v(l) = t_1 - t_0 + 1, d_1(t_1) = 1\} = \theta \big(\bar{\lambda}\bar{\theta}\big)^{t_1 - t_0} \lambda.$$

Moreover, in the limit of large t_1, the term $\Pr\{d_1(t_1) = 1\}$ approaches its corresponding stationary distribution and is independent of t_1 as long as $\theta > 0, \lambda > 0$. Therefore, we have

$$\Pr\{D_1 = t_1 - t_0 + 1\} = \big(\bar{\lambda}\bar{\theta}\big)^{t_1 - t_0}(1 - \bar{\lambda}\bar{\theta}),$$

which is the desired geometric distribution.

Recall that in the above argument we assumed one source in order to disregard collisions. In reality we have $M > 1$ sources. In this case, if two or more than two sources transmit packet at the same time slot, then collisions occurs. Consider

$$\Pr(d_1(t){=}1 | \{c(k)\}_{k<t}, \{A_1(k)\}_{k\le t}, \text{Queue 1 non-empty}). \tag{52}$$

Since the stationary randomized policy of Section V does not utilize collision feedback and is independent of the packet arrivals condition of the queue being non-empty, we have

$$\Pr(d_1(t) = 1 | \{c(k)\}_{k<t}, \{A_1(k)\}_{k\le t}, \text{Queue 1 non-empty}) \tag{53}$$

$$= \Pr(d_1(t) = 1 | \text{Queue 1 non-empty}). \tag{54}$$

It is clear that the right hand side approaches its stationary distribution λ_c for large t. So the derivation above for $M = 1$ applies with λ replaced by λ_c.

Appendix G
Proof of (28).

First, from (24),

$$\lim_{k\to\infty} \Pr(d_1(k) = 1) = \sum_v \pi_v \sum_{w=1}^{2^M} p_{vw}^{(1)}.$$

Now we calculate

$$\lim_{k\to\infty} \Pr(D_1(m_k) = 1, d_1(k) = 1).$$

Note that

$$\{D(m_k) = 1, d_1(k) = 1\} \Leftrightarrow \{A_1(k) = 1, d_1(k) = 1\}.$$

Denote the state of channel at the beginning of time slot k by $s(k)$, then

$$\begin{aligned} &\lim_{k\to\infty} \Pr\big(A_1(k) = 1, d_1(k) = 1\big) \\ =& \lim_{k\to\infty} \sum_{i=1}^{2^M} \Pr\big(A_1(k) = 1, d_1(k) = 1 | s(k) = i\big) \Pr(s(k) = i). \end{aligned}$$

We focus on the event $\{A_1(k) = 1, d_1(k) = 1\}$ condition on $\{s(k) = i\}$. Note that the state of the first transmitter at the beginning of time slot k is irrelevant to the resulting state after the occurrence of $\{A_1(k) = 1, d_1(k) = 1\}$. So for any $\underline{s}' \in \{0,1\}^{2^M - 1}$, define $\underline{s} = (*, \underline{s}')$ as the system's state at the beginning of time slot k before the arrival of new packets. Here, $*$ stands for 0 or 1. Note, furthermore, that the resulting state is $(0, \underline{s}')$ because $d(k) = 1$. Ordering the states lexicographically, we can thus write

$$\lim_{k\to\infty} \Pr\big(A_1(k) = 1, d_1(k) = 1\big) \tag{55}$$

$$= \sum_{i=1}^{2^M} \lim_{k\to\infty} \Pr\big(A_1(k) = 1, d_1(k) = 1 | s(k) = i\big) \tag{56}$$

$$\cdot \Pr(s(k) = i) \tag{57}$$

$$= \sum_{i=2^{M-1}+1}^{2^M} p_{i,i-2^{M-1}}^{(1)} \big(\pi_i + \pi_{i-2^{M-1}}\big). \tag{58}$$

where π_s is from (17) and $p_{uv}^{(1)}$ is from (15). Therefore from (27), we have

$$\mathbb{E}[D_1] = \frac{\sum_v \pi_v \sum_{w=1}^{2^M} p_{vw}^{(1)}}{\sum_{i=2^{M-1}+1}^{2^M} p_{i,i-2^{M-1}}^{(1)} (\pi_i + \pi_{i-2^{M-1}})}.$$

Appendix H
Probabilities of idle, deliveries and collisions under Slotted ALOHA

Lemma 4. *Consider any stabilized slotted ALOHA scheme. Define G as the expected number of attempted transmissions in a slot. Then, for M large, the probability of delivering a packet is (asymptotically) Ge^{-G}, the probability of idle system is (asymptotically) e^{-G}, and the probability of collisions is (asymptotically) $1 - e^{-G} - Ge^{-G}$. In particular, when $G = 1$, the maximum probability of delivery is $1/e$, the corresponding probabilities of collisions and idle system are $1 - 2/e$ and $1/e$, respectively.*

Proof. The proof is similar to [32, Chapter 4]. Define the nodes which are not backlogged as *fresh* nodes. Each fresh node transmits a packet directly in a slot if it is not empty, and it generates/receives a packet with probability θ, thus a fresh node transmits a packet with probability θ. Let $P_a\big(i, n(k)\big)$ be the probability that i fresh nodes transmit a packet in a time slot and let $P_s\big(j,n(k)\big)$ be the probability that j backlogged nodes transmit. We have:

$$P_a\big(i,n(k)\big)=\binom{M-n(k)}{i}(1-\theta)^{M-n(k)-i}\theta^i \tag{59}$$

$$P_s\big(j,n(k)\big)=\binom{n(k)}{j}(1-p_b(k))^{n(k)-i}p_b(k)^i. \tag{60}$$

Thus, in slot k, when a packet is delivered, i.e., $\sum_{i=1}^{M} d_i(k) = 1$, the probability is

$$\begin{aligned}&\Pr(\sum_{i=1}^{M} d_i(k)=1)\\ =&P_a\big(1,n(k)\big)P_s\big(0,n(k)\big)+P_a\big(0,n(k)\big)P_s\big(1,n(k)\big).\end{aligned} \tag{61}$$

If the channel does not transmit a packet in a slot, i.e., we have an idle channel, $\sum_{i=1}^{M} d_i(k) = 0$, $c(k) = 0$. The probability of idle system in slot k is

$$\Pr(\sum_{i=1}^{M} d_i(k)=0,c(k)=0)=P_a\big(0,n(k)\big)P_s\big(0,n(k)\big). \tag{62}$$

Define the attempt rate $G = (M-n(k))\theta + n(k)p_b(k)$ as the expected number of attempted transmissions in a slot. From (59) and (60), the probability of delivery is

$$\begin{aligned}&\Pr(\sum_{i=1}^{M} d_i(k)=1)\\ =&\big(M-n(k)\big)(1-\theta)^{M-n(k)-1}\theta(1-p_b)^{n(k)}\\ &+(1-\theta)^{M-n(k)}n(k)(1-p_b)^{n(k)-1}p_b\end{aligned}$$

and the probability of an idle channel is

$$\Pr(\sum_{i=1}^{M} d_i(k)=0,c(k)=0)=(1-\theta)^{M-n(k)}(1-p_b)^{n(k)}.$$

Note that the valid regime of θ is $\theta M < \frac{1}{e}$, so θ, p_b are small, using the approximation $(1-x)^{-y} \approx \exp(-xy)$ for small x, then

$$\begin{aligned}\Pr(\sum_{i=1}^{M} d_i(k)=1)&\approx Ge^{-G}\\ \Pr(\sum_{i=1}^{M} d_i(k)=0,c(k)=0)&\approx e^{-G}.\end{aligned}$$

Then,

$$\Pr(c(k)=1)\approx 1-Ge^{-G}-e^{-G}.$$

Taking the first derivative of function Ge^{-G}, we can find the maximum point is 1 for $0 < G \leq 1$. So the maximum probability of delivery is

$$\Pr(\sum_{i=1}^{M} d_i(k)=1)\approx 1/e,$$

and the corresponding

$$\begin{aligned}&\Pr(\sum_{i=1}^{M} d_i(k)=0,c(k)=0)\approx 1/e,\\ &\Pr(c(k)=1)=1-2/e.\end{aligned}$$

□

APPENDIX I
PROOF OF THEOREM 2.

The proof is organized in three parts:

Part 1: Preliminaries. In time slot k, denote the time just before the arrival of new packets by k^- and the time just after the arrival of new packets by k^+. We hence write $\delta_i(k^-) = h_i(k^-) - w_i(k^-)$ and $\delta_i(k^+) = h_i(k^+) - w_i(k^+)$. Suppose a packet is delivered from the i^{th} source at the end of time slot $k-1$. We then have $\delta_i\big(k^-\big) = 0$. From (29), since all nodes have the same arrival rate and transmission policy, the sequences $\{h_i(k^-)\}_{k=1}^{\infty}$, $\{h_i(k^+)\}_{k=1}^{\infty}$, $\{w_i(k^-)\}_{k=1}^{\infty}$, $\{w_i(k^+)\}_{k=1}^{\infty}$, $\{\delta_i(k^-)\}_{k=1}^{\infty}$, $\{\delta_i(k^+)\}_{k=1}^{\infty}$ are identical random variables across $i = 1, 2, \cdots, M$, respectively. Recall that source nodes with $\delta_i(k^-) = 0$ are 0-order nodes and define $N_0(k^-)$ as the number of 0-order nodes at time k^-.

In the beginning of time slot k, on average, θM new packets arrive at the sources, and $\theta N_0(k^-)$ 0-order nodes receive new packets. Suppose source i is a 0-order node and $h_i(k^-) - w_i(k^-) = 0$. If source i receives new packets, then the source's AoI changes from $w_i(k^-)$ to $w_i(k^+) = 0$ and the destination's AoI $h_i(k^-)$ remains the same as $h_i(k^+)$. Thus,

$$\begin{aligned}\delta_i(k^+)=&h_i(k^+)-w_i(k^+)\\ =&h_i(h^+)\\ =&h_i(k^-)\\ >&h_i(k^-)-w_i(k^-)\\ =&\delta_i(k^-)\\ =&0,\end{aligned}$$

which implies that if a 0-order source receives a new packet, then it is not a 0-order source at k^+.

Fix any large M and denote the maximum throughput of Slotted ALOHA with $C_{SA}(M)$. We know that

$$\lim_{M\to\infty} C_{SA}(M)=e^{-1}.$$

The recursion of the expected number of $0-$order nodes is:

$$\mathbb{E}[N_0\big((k+1)^-\big)]=(1-\theta)\mathbb{E}[N_0(k^-)]+\min(M\theta,C_{SA}(M))$$

where the second term on the right hand side is the average number of delivered packets per time slot. Since we consider a stabilized slotted ALOHA, then $\lim_{k\to\infty}\mathbb{E}[N_0(k^-)]$ exists. Denote

$$N_0^*=\lim_{k\to\infty}\mathbb{E}[N_0(k^-)].$$

Note that

$$\lim_{M\to\infty}\min(M\theta,C_{SA}(M))=\lim_{M\to\infty}M\theta=\eta,$$

and hence

$$\lim_{M\to\infty}\frac{N_0^*}{M}=1. \tag{63}$$

Part 2: Find the expression of $NEWSAoI$. From (3), and letting $\alpha_i=\frac{1}{M}$ for $i=1,2,\cdots,M$, we find

$$\begin{aligned}\mathbb{E}[J^\pi]=&\lim_{K\to\infty}\mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}h_i(k^-)]\\ \triangleq& NEWSAoI_1+NEWSAoI_2\end{aligned}$$

where

$$NEWSAoI_1=\lim_{K\to\infty}\mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}w_i(k^-)]$$
$$NEWSAoI_2=\lim_{K\to\infty}\mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}\delta_i(k^-)].$$

Part 3: Find the limit of $NEWSAoI$. First, we consider $NEWSAoI_1$. $w_i(k^-)$ has a geometric distribution starting from 1 with parameter θ for all i. Employing the law of large number, we find

$$NEWSAoI_1=\frac{1}{M\theta}.$$

Next, we consider $NEWSAoI_2$ and prove that its limit in large M approaches zero. Note from (10) that $\delta_i(k)=0$ if source i is empty in time slot k and $\delta_i(k)>0$ if a packet remains in source i in time slot k. We first note that $\delta_i(k)$ is upper bounded by $h_i(k)$. Let us consider a worse case in which buffer sizes are infinite. In this case, assuming stationarity[6], denote the inter-arrival time and delay of packets with respect to source i by X_i and D_i. Since the Bernoulli arrival process has parameter θ, then $\mathbb{E}[X_i]=\frac{1}{\theta}=\frac{M}{\eta}$. Moreover $\mathbb{E}[D_i]$ is approximately bounded by some constant independent of the number of sources M [34]. Now we observe that for each packet delivery the expected peak age at the destination is upper bounded by $\mathbb{E}[X_i]+\mathbb{E}[D_i]$. We can hence write

$$\mathbb{E}[\delta_i^*|\delta_i^*>0]\le\mathbb{E}[X_i]+\mathbb{E}[D_i] \tag{64}$$

which implies that $\mathbb{E}[\delta_i^*]$ is $O(M)$.

[6]This assumption approximately holds for infinite time horizon T

Now expand $NEWSAoI_2$:

$$\begin{aligned}NEWSAoI_2=&\lim_{K\to\infty}\mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}\delta_i(k^-)]\\ =&\lim_{K\to\infty}\mathbb{E}[\frac{1}{M^2}\frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{M}\delta_i(k^-)1_{\delta_i(k^-)>0}]\\ =&\lim_{K\to\infty}\frac{1}{M^2}\frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{M}\mathbb{E}[\delta_i(k^-)1_{\delta_i(k^-)>0}]\\ \le&\limsup_{k\to\infty}\frac{1}{M^2}\sum_{i=1}^{M}\mathbb{E}[\delta_i(k^-)1_{\delta_i(k^-)>0}]\\ =&\limsup_{k\to\infty}\frac{1}{M^2}\sum_{i=1}^{M}\Big(\Pr(\delta_i(k^-)>0)\\ &\times\mathbb{E}[\delta_i(k^-)|\delta_i(k^-)>0]\Big)\end{aligned}$$

Since for k large enough the conditional expectation $\mathbb{E}[\delta_i(k^-)|\delta_i(k^-)>0]$ is $O(M)$, it remains to prove that in the limit of large M, $\lim_{k\to\infty}\frac{1}{M}\sum_{i=1}^{M}\Pr(\delta_i(k^-)>0)$ vanishes. But this holds because we can write

$$\begin{aligned}\lim_{k\to\infty}\frac{1}{M}\sum_{i=1}^{M}\Pr(\delta_i(k^-)>0)=&\lim_{k\to\infty}\mathbb{E}[\frac{1}{M}\sum_{i=1}^{M}1_{\delta_i(k^-)>0}]\\ =&\lim_{k\to\infty}\mathbb{E}[\frac{1}{M}(M-N_0(k^-))]\\ =&\frac{M-N_0^*}{M}\end{aligned} \tag{65}$$

and (65) goes to zero by (63).

Finally, we prove that for any scheme, NEWSAoI is lower bounded by $1/\eta$. From Theorem 1, let $M\to\infty$, we have

$$\lim_{M\to\infty}NEWSAoI\ge\lim_{M\to\infty}\frac{1}{M}(\frac{2-2\theta}{\theta}+\frac{1}{2})=\frac{1}{\eta}.$$

Therefore, slotted ALOHA can reach the lower bound when $\theta\in(0,\frac{1}{eM}]$ and is hence optimal.

Appendix J
Proof of Lemma 1.

Before presenting the proof, we state the following straightforward lemma (whose proof is omitted).

Lemma 5. *At the beginning of time slot k, before new packets arrive at source i,* $w_i(k^-)>0$ *and its probability distribution is*

$$\Pr\left(w_i(k^-)=j\right)=\theta(1-\theta)^{j-1},\ j=1,2,3,\cdots. \tag{66}$$

First consider $m=0$ and suppose source i is a 0-order node. From Lemma 5, we know that $w_i(k^-)>0$. Moreover, since $\delta_i(k^-)=0$, we conclude $h_i(k^-)=w_i(k^-)>0$. Once the 0-order node has a new arrival, $w_i(k^+)=0$ and $h_i(k^+)=h_i(k^-)$, resulting in $\delta_i(k^+)=h_i(k^+)>0$; i.e., the order of the node increases. In other words, the order of a 0-order node increases once it receives a new packet. In total, the fraction of 0-order nodes that become of higher order is on average $\theta\ell_0(k-1)$. Thus,

$$\ell_0(k^+) = (1-\theta)\ell_0(k-1).$$

Similarly, we consider $m \geq 1$. The fraction of m-order nodes that have new arrivals is $\theta\ell_m(k-1)$. These nodes will have larger orders. Suppose source i is of order m, $m \geq 1$, i.e., $\delta_i(k^-) = h_i(k^-) - w_i(k^-) = m$, once a new packet arrives, then $w_i(k^+) = 0$, $h_i(k^+) = h_i(k^-)$, and $\delta_i(k^+) = h_i(k^+) = m + w_i(k^-)$. From Lemma 5, $w_i(k^-) > 0$, then $\delta_i(k^+) > \delta_i(k^-) = m$. The order of a m-order node increases once it receives a new packet. In total, the fraction of m-order nodes have larger orders is $\theta\ell_m(k-1)$.

More precisely, consider a j-order node, $j < m$. This node becomes an m-order node if it receives a new packet and $w_i(k^-) = m - j$. Using Lemma 5, we cam write

$$\begin{aligned}\ell_m(k^+) =&(1-\theta)\ell_m(k-1)\\ &+\sum_{j=0}^{m-1}\theta\ell_j(k-1)\Pr\Big(w_i(k^-)=m-j\Big)\\ =&(1-\theta)\ell_m(k-1)+\sum_{j=0}^{m-1}\theta\ell_j(k-1)\theta(1-\theta)^{m-j-1}\\ =&(1-\theta)\ell_m(k-1)+\theta^2\sum_{j=0}^{m-1}\ell_j(k-1)(1-\theta)^{m-j-1}.\end{aligned}$$

where the second term on the left hand side is the average fraction of nodes that have just become of order m. Denoting it by a_m, we have

$$a_m(k) = \theta^2\sum_{j=0}^{m-1}\ell_j(k-1)(1-\theta)^{m-j-1}.$$

APPENDIX K
PROOF OF LEMMA 2.

From the expression of ℓ_m^* in (39), $0 \leq m \leq T^*-1$, we obtain

$$\ell_0^* = \frac{1}{eM\theta} \tag{67}$$

$$\ell_m^* = \frac{a_m^*}{\theta} \qquad 0 \leq m \leq T^*-1. \tag{68}$$

From (36), a_m^* depends on $\{\ell_j^*\}_{j\leq m-1}$ and from (68), ℓ_m^* depends on a_m^* for $1 \leq m \leq T^*-1$. So they can be recursively found and in particular, it is not difficult to prove for all $1 \leq m \leq T^*-1$:

$$a_m^* = \frac{\theta}{eM} \tag{69}$$

$$\ell_m^* = \frac{1}{eM}. \tag{70}$$

We prove this by mathematical induction on $T^* \geq 2$. For $T^*-1 = 1$, the statement holds because

$$a_1^* = \theta^2\ell_0^* = \frac{\theta}{eM}$$

$$\ell_1^* = \frac{a_1^*}{\theta} = \frac{1}{eM}.$$

Now suppose the statements (69)-(70) hold for $m \leq \mathtt{T}^*-1 = k$. We prove the statement for $\mathtt{T}^*-1 = k+1$ and in particular we find a_{k+1}^* and ℓ_{k+1}^* below:

$$\begin{aligned}a_{k+1}^* &= \theta^2\sum_{j=0}^{k}\ell_j^*(1-\theta)^{k-j}\\ &=\theta^2\frac{1}{eM}\sum_{j=1}^{k}(1-\theta)^{k-j}+\theta^2(1-\theta)^k\frac{1}{eM\theta}\\ &=\theta^2\frac{1}{eM}\frac{1-(1-\theta)^k}{\theta}+\theta(1-\theta)^k\frac{1}{eM}\\ &=\frac{\theta}{eM}.\end{aligned} \tag{71}$$

Next, using (68), we find

$$\ell_{k+1} = \frac{1}{eM}.$$

Moreover, using the derivation in (71), we also find

$$a_{T^*}^* = \frac{\theta}{eM}.$$

Finally, from (35), we obtain

$$\ell_m^{+*} = \frac{1}{eM} \qquad 1 \leq m \leq \mathtt{T}^*-1.$$

APPENDIX L
PROOF OF THEOREM 3.

Summing (39) on both sides, we have

$$\sum_{m\geq 1} a_m^* = \theta.$$

Moreover, $\mathtt{T}^*$ satisfies

$$\mathtt{T}^* = \max\{t|\sum_{m\geq t} a_m^* \geq \frac{1}{eM}\} \tag{72}$$

by its definition in (32). The term $\sum_{m\geq t} a_m^*$ can be re-written as follows:

$$\begin{aligned}\sum_{m\geq t} a_m^* &= \sum_{m\geq 1} a_m^* - \sum_{m<\mathtt{T}^*} a_m^*\\ &\overset{(a)}{=}\theta - (\mathtt{T}^*-1)\frac{\theta}{eM}\\ &\geq\frac{1}{eM}\end{aligned}$$

where (a) follows by Lemma 2 (3). Since $\mathtt{T}^*$ is an integer, we find

$$\mathtt{T}^* = \lfloor eM - \frac{1}{\theta} + 1\rfloor.$$

APPENDIX M
PROOF OF THEOREM 4.

The proof is organized in three parts:

Part 1: Preliminaries. Since $\theta = \frac{1}{o(M)}$ and in particular $\theta > \frac{1}{eM}$, from Lemma 2, $\ell_m^{+*} = \frac{1}{eM}$ for $m = 1, 2, \cdots, \mathtt{T}^*$ and

$\ell_0^{+*} = \frac{o(M)}{eM}$. From Theorem 3, $\mathtt{T}^* = \lfloor eM - 1/\theta + 1 \rfloor = \lfloor eM - o(M) + 1 \rfloor$. Denote $s_{\mathtt{T}^*} = \sum_{m=0}^{\mathtt{T}^*-1} \ell_m^{+*}$, then

$$\lim_{M\to\infty} s_{\mathtt{T}^*} = \lim_{M\to\infty} \frac{o(M) + \lfloor eM - o(M) + 1 \rfloor - 1}{eM} = 1.$$

The expected number of inactive nodes is $Ms_{\mathtt{T}^*}$ and the expected number of active nodes is $M(1 - s_{\mathtt{T}^*})$.

Denote the time just before arrival of new packets by k^- and the time just after arrival of new packets by k^+. Since we have assumed that all nodes are identical, then the sequence $\{h_i(k^+)\}_{k=1}^{\infty}$ is identical (but not independent) across all $i = 1, 2, \cdots, M$. From (2), $\{w_i(k^+)\}_{k=1}^{\infty}$ are i.i.d with respect to i. Then, the sequence $\{\delta_i(k^+)\}_{k=1}^{\infty}$ is identical but not independent for all $i = 1, 2, \cdots, M$.

Part 2: Find the expression of $NEWSAoI$. Let $\alpha_i = \frac{1}{M}$ for $i = 1, 2, \cdots, M$ in (3):

$$\begin{aligned}\mathbb{E}[J^\pi] = &\lim_{K\to\infty} \mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K} h_i(k^+)]\\ \triangleq &NEWSAoI_1 + NEWSAoI_2\end{aligned}$$

where

$$NEWSAoI_1 = \lim_{K\to\infty} \mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K} w_i(k^+)]$$
$$NEWSAoI_2 = \lim_{K\to\infty} \mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K} \delta_i(k^+)].$$

In addition,

$$NEWSAoI_2 = NEWSAoI_{21} + NEWSAoI_{22},$$

where

$$NEWSAoI_{21} = \lim_{K\to\infty} \mathbb{E}[\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i:\delta_i(k^+)<\mathtt{T}^*} \delta_i(k^+)]$$
$$NEWSAoI_{22} = \lim_{K\to\infty} \mathbb{E}[\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i:\delta_i(k^+)\geq\mathtt{T}^*} \delta_i(k^+)].$$

Part 3: Find the limit of $NEWSAoI$. First, we consider $NEWSAoI_1$. From (2), $w_i(k^+)$ has a geometric distribution with parameter θ for all i. Let w have the same distribution as $w_i(k^+)$. We thus have

$$NEWSAoI_1 = \frac{1}{M}\mathbb{E}[w] = \frac{1}{M\theta}.$$

Next, we consider $NEWSAoI_{21}$:

$$\begin{aligned}&\lim_{M\to\infty} NEWSAoI_{21}\\ =&\lim_{M\to\infty}\lim_{K\to\infty} \mathbb{E}[\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i:\delta_i(k^+)<\mathtt{T}^*}\delta_i(k^+)]\\ =&\lim_{M\to\infty}\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}\mathbb{E}[\delta_i(k^+)1_{(\delta_i(k^+)<\mathtt{T}^*)}]\\ =&\lim_{M\to\infty}\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{\mathtt{T}^*-1} j\Pr(\delta_i(k^+)=j)]\\ \overset{(a)}{=}&\lim_{M\to\infty}\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M}\sum_{j=1}^{\mathtt{T}^*-1} j\mathbb{E}[\ell_j(k^+)]\\ \overset{(b)}{=}&\lim_{M\to\infty}\frac{1}{M}\sum_{j=1}^{\mathtt{T}^*-1} j\frac{1}{eM}\\ =&\lim_{M\to\infty}\frac{1}{M}\frac{\mathtt{T}^*(\mathtt{T}^*-1)}{2}\frac{1}{eM}\\ =&\frac{e}{2}.\end{aligned}$$

where (a) follows because

$$\begin{aligned}&\frac{1}{M}\sum_{i=1}^{M}\Pr(\delta_i(k^+)=j)\\ &=\frac{1}{M}\sum_{i=1}^{M}\mathbb{E}[1_{\delta_i(k^+)=j}]\\ &=\mathbb{E}[\frac{1}{M}\sum_{i=1}^{M}1_{\delta_i(k^+)=j}]\\ &=\ell_j(k^+)\end{aligned}$$

and (b) follows because $\ell_j(k^+)$ converges to ℓ_j^*+ as $k\to\infty$.

Finally, we consider $NEWSAoI_{22}$:

$$\begin{aligned}
&\lim_{M\to\infty} NEWSAoI_{22}\\
=&\lim_{M\to\infty}\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}\mathbb{E}[\delta_i(k^+)1_{\delta_i(k^+)>\mathtt{T}^*}]\\
=&\lim_{M\to\infty}\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}\Big(\mathbb{E}[\delta_i(k^+)|\delta_i(k^+)>\mathtt{T}^*]\\
&\qquad\qquad\qquad\qquad\qquad\times\Pr(\delta_i(k^+)>\mathtt{T}^*)\Big)\\
\overset{(a)}{\le}&\lim_{M\to\infty}\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}cM\Pr(\delta_i(k^+)>\mathtt{T}^*)\\
=&\lim_{M\to\infty}\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M}cM\left(\sum_{j=T^*}^{M}\ell_j(k^+)\right)\\
=&\lim_{M\to\infty}\frac{1}{M}cM\left(\sum_{j=T^*}^{M}\ell_j^*\right)\\
=&\lim_{M\to\infty}\frac{1}{M}cM(1-s_T^*)\\
=&0.
\end{aligned}$$

In the above chain of inequalities, step (a) holds because $\mathbb{E}[\delta_i(k^+)|\delta_i(k^+)>\mathtt{T}^*]=O(M)$. To show this, we first observe that $\delta_i(k)$ is increasing in k until a delivery occurs. Now, note that $\delta_i(k^+)$ is upper bounded by $\mathtt{T}^*$ plus the peak age at the first delivery after time slot k. The peak age is bounded by X_i (the inter arrival time), which is $o(M)$ on average, plus delay D_i, which is constant on average (similar to (64)). The threshold $\mathtt{T}^*$ is also $O(M)$. So overall, we have

$$\mathbb{E}[\delta_i(k^+)|\delta_i(k^+)>\mathtt{T}^*]\le cM$$

for some constant c.

Finally, summing $NEWSAoI_1$, $NEWSAoI_{21}$ and $NEWSAoI_{22}$, we have

$$\lim_{M\to\infty}\mathbb{E}[J^\pi]=\frac{e}{2}.$$

Appendix N
Proof of Theorem 5.

Summing (39) on both sides, we have

$$\sum_{m\ge1}a_m^*=\theta.$$

From the definition of the threshold in (32), $\mathtt{T}^*$ satisfies

$$\mathtt{T}^*=\max\{t|\sum_{m\ge t}a_m^*\ge\min\big(\theta,\frac{C}{M}\big)\},\tag{73}$$

If $\theta\le\frac{C}{M}$, then from (73), we have $\mathtt{T}^*=1$. If $\theta>\frac{C}{M}$, then from Lemma 2 (3), $a_m^*=\min(\theta^2,\frac{\theta C}{M})=\frac{\theta C}{M}$ for $1\le m\le\mathtt{T}^*-1$, then

$$\sum_{m\ge\mathtt{T}^*}a_m^*=\theta-(\mathtt{T}^*-1)\frac{\theta C}{M}\ge\frac{C}{M},$$

note that $\mathtt{T}^*$ is an integer, thus

$$\mathtt{T}^*=\lfloor\frac{M}{C}-\frac{1}{\theta}+1\rfloor.$$

Appendix O
Proof of Theorem 5.

The proof is organized in three parts:

Part 1: Preliminaries. Since $\theta=\frac{1}{o(M)}$ and in particular $\theta>\frac{1}{eM}$, from (45) - (47), $\ell_m^{+*}=\frac{C}{M}$ for $m=1,2,\cdots,\mathtt{T}^*$ and $\ell_0^{+*}=\frac{o(M)}{M/C}$. From Theorem 5, $\mathtt{T}^*=\lfloor\frac{M}{C}-1/\theta+1\rfloor=\lfloor\frac{M}{C}-o(M)+1\rfloor$. Denote $s_{\mathtt{T}^*}=\sum_{m=0}^{\mathtt{T}^*-1}\ell_m^{+*}$, then

$$\lim_{M\to\infty}s_{\mathtt{T}^*}=\lim_{M\to\infty}\frac{o(M)+\lfloor\frac{M}{C}-o(M)+1\rfloor-1}{\frac{M}{C}}=1.$$

The expected number of inactive nodes is $Ms_{\mathtt{T}^*}$ and the expected number of active nodes is $M(1-s_{\mathtt{T}^*})$.

Denote the time just before arrival of new packets by k^- and the time just after arrival of new packets by k^+. Since we have assumed that all nodes are identical, then the sequence $\{h_i(k^+)\}_{k=1}^\infty$ is identical (but not independent) across all $i=1,2,\cdots,M$. From (2), $\{w_i(k^+)\}_{k=1}^\infty$ are i.i.d with respect to i. Then, the sequence $\{\delta_i(k^+)\}_{k=1}^\infty$ is identical but not independent for all $i=1,2,\cdots,M$.

Part 2: Find the expression of $NEWSAoI$. Let $\alpha_i=\frac{1}{M}$ for $i=1,2,\cdots,M$ in (3):

$$\begin{aligned}\mathbb{E}[J^\pi]&=\lim_{K\to\infty}\mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}h_i(k^+)]\\&\triangleq NEWSAoI_1+NEWSAoI_2\end{aligned}$$

where

$$NEWSAoI_1=\lim_{K\to\infty}\mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}w_i(k^+)]$$
$$NEWSAoI_2=\lim_{K\to\infty}\mathbb{E}[\frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}\delta_i(k^+)].$$

In addition,

$$NEWSAoI_2=NEWSAoI_{21}+NEWSAoI_{22},$$

where

$$NEWSAoI_{21}=\lim_{K\to\infty}\mathbb{E}[\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i:\delta_i(k^+)<\mathtt{T}^*}\delta_i(k^+)]$$
$$NEWSAoI_{22}=\lim_{K\to\infty}\mathbb{E}[\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i:\delta_i(k^+)\ge\mathtt{T}^*}\delta_i(k^+)].$$

Part 3: Find the limit of $NEWSAoI$. First, we consider $NEWSAoI_1$. From (2), $w_i(k^+)$ has a geometric distribution with parameter θ for all i. Let w have the same distribution as $w_i(k^+)$. We thus have

$$NEWSAoI_1=\frac{1}{M}\mathbb{E}[w]=\frac{1}{M\theta}.$$

Next, we consider $NEWSAoI_{21}$:

$$\begin{aligned}
&\lim_{M\to\infty} NEWSAoI_{21} \\
&= \lim_{M\to\infty}\lim_{K\to\infty} \mathbb{E}[\frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i:\delta_i(k^+)<\mathtt{T}^*}\delta_i(k^+)] \\
&= \lim_{M\to\infty}\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}\mathbb{E}[\delta_i(k^+)1_{(\delta_i(k^+)<\mathtt{T}^*)}] \\
&= \lim_{M\to\infty}\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{\mathtt{T}^*-1} j\Pr(\delta_i(k^+)=j)] \\
&\overset{(a)}{=} \lim_{M\to\infty}\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K}\frac{1}{M}\sum_{j=1}^{\mathtt{T}^*-1} j\mathbb{E}[\ell_j(k^+)] \\
&\overset{(b)}{=} \lim_{M\to\infty}\frac{1}{M}\sum_{j=1}^{\mathtt{T}^*-1} j\frac{C}{M} \\
&= \lim_{M\to\infty}\frac{1}{M}\frac{\mathtt{T}^*(\mathtt{T}^*-1)}{2}\frac{C}{M} \\
&= \frac{1}{2C}.
\end{aligned}$$

where (a) follows because

$$\begin{aligned}
&\frac{1}{M}\sum_{i=1}^{M}\Pr(\delta_i(k^+)=j) \\
&\quad= \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}[1_{\delta_i(k^+)=j}] \\
&\quad= \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M}1_{\delta_i(k^+)=j}] \\
&\quad= \ell_j(k^+)
\end{aligned}$$

and (b) follows because $\ell_j(k^+)$ converges to ℓ_j^*+ as $k\to\infty$.

Finally, we consider $NEWSAoI_{22}$:

$$\begin{aligned}
&\lim_{M\to\infty} NEWSAoI_{22} \\
&= \lim_{M\to\infty}\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}\mathbb{E}[\delta_i(k^+)1_{\delta_i(k^+)>\mathtt{T}^*}] \\
&= \lim_{M\to\infty}\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M}\Big(\mathbb{E}[\delta_i(k^+)|\delta_i(k^+)>\mathtt{T}^*] \\
&\qquad\qquad \times\Pr(\delta_i(k^+)>\mathtt{T}^*)\Big) \\
&\overset{(a)}{\le} \lim_{M\to\infty}\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K}\frac{1}{M^2}\sum_{i=1}^{M} cM\Pr(\delta_i(k^+)>\mathtt{T}^*) \\
&= \lim_{M\to\infty}\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K}\frac{1}{M}cM\left(\sum_{j=T^*}^{M}\ell_j(k^+)\right) \\
&= \lim_{M\to\infty}\frac{1}{M}cM\left(\sum_{j=T^*}^{M}\ell_j^*\right) \\
&= \lim_{M\to\infty}\frac{1}{M}cM(1-s_T^*) \\
&=0.
\end{aligned}$$

In the above chain of inequalities, step (a) holds because $\mathbb{E}[\delta_i(k^+)|\delta_i(k^+) > \mathtt{T}^*] = O(M)$. To show this, we first observe that $\delta_i(k)$ is increasing in k until a delivery occurs. Now, note that $\delta_i(k^+)$ is upper bounded by $\mathtt{T}^*$ plus the peak age at the first delivery after time slot k. The peak age is bounded by X_i (the inter arrival time), which is $o(M)$ on average, plus delay D_i, whose expectation is upper bounded by a constant times M as formulated in Lemma 6 below. Note that the threshold $\mathtt{T}^*$ is also $O(M)$. So overall, we have

$$\mathbb{E}[\delta_i(k^+)|\delta_i(k^+)>\mathtt{T}^*] \le cM$$

for some constant c.

Lemma 6. *When M is large, the expectation of delay, $\mathbb{E}[D_i] \le c'M$, where c' is a constant depends on a particular transmission policy.*

Proof. Since we assume the rate of the transmission policy is C, then for any large M, denote the maximum throughput is $C(M)$. We know that

$$\lim_{M\to\infty} C(M) = C.$$

Then, denote the inter-delivery time for source i is I_i, thus the expected number of received packets from source i from 0 up to K is $\frac{K}{\mathbb{E}[I_i]}$. From the definition of maximum throughput,

$$C(M) = \lim_{K\to\infty}\frac{\sum_{i=1}^{M}\frac{K}{\mathbb{E}[I_i]}}{K}.$$

Since we assume all nodes are identical, then

$$C(M) = \frac{M}{\mathbb{E}[I_i]},$$

so

$$\mathbb{E}[I_i] = \frac{M}{C(M)}.$$

Note that

$$\mathbb{E}[D_i] \leq \mathbb{E}[I_i],$$

and for any $\epsilon > 0$, there exists a $N_0 > 0$, for all $M \geq N_0$, $C(M) \geq C - \epsilon$.

$$\mathbb{E}[D_i] \leq \frac{M}{C - \epsilon} \triangleq c' M.$$

$\square$

Finally, summing $NEWSAoI_1$, $NEWSAoI_{21}$ and $NEWSAoI_{22}$, we have

$$\lim_{M \to \infty} \mathbb{E}[J^\pi] = \frac{1}{2C}.$$