An Age of Information Characterization of Frameless ALOHA

Andrea Munari, Francisco Lázaro, Giuseppe Durisi, Gianluigi Liva

Abstract—**We provide a characterization of the peak age of information (AoI) achievable in a random-access system operating according to the frameless ALOHA protocol. Differently from previous studies, our analysis accounts for the fact that the number of terminals contending the channel may vary over time, as a function of the duration of the previous contention period. The exact characterization of the AoI provided in this paper, which is based on a Markovian analysis, reveals the impact of some key protocol parameters such as the maximum length of the contention period, on the average peak AoI. Specifically, we show that setting this parameter so as to maximize the throughput may result in an AoI degradation.**

I. Introduction

Internet of things (IoT) systems often involve a large number of terminals that sense a physical process and report time-stamped status updates to a common receiver. This scenario is relevant in, e.g., environmental monitoring and asset tracking, where a primary objective is to maintain an up-to-date record of the status of an observed source. A number of performance metrics related to the notion of information freshness have recently been proposed to quantify the ability of a system to reach this goal [1], [2]. Among them, a prominent role is played by the *age of information (AoI)* [3], which quantifies the amount of time elapsed since the newest update available at the receiver was generated at the source, and has been shown to effectively capture fundamental trends in a number of relevant scenarios [4], [5].

Accurate AoI characterizations are available for traditional grant-based link-layer medium-access policies. Unfortunately, such policies are often highly suboptimal or non-viable in IoT networks, because of the need to share a common channel among a possibly massive number of battery-powered, low-complexity devices that generate traffic in a sporadic fashion. Accordingly, random access strategies based on variations of ALOHA [6] are the de-facto choice in a number of commercial systems [7], [8]. Preliminary insights on the information-freshness trade-offs that emerge in random-access systems were derived in [9], [10]. Specifically, these contributions illustrate that throughput and AoI can be jointly optimized under ALOHA policies by properly tuning the channel access probability. Further improvements in the presence of feedback were discussed in [11], [12].

In parallel to this line of research, a family of advanced grant-free protocols for IoT, often dubbed *modern random access* [13]–[15], has recently been proposed. Such protocols allow terminals to transmit multiple copies of their packets over time, and employ successive interference cancellation

at the receiver to resolve collisions. This leads to significant throughput improvements, which makes these solutions excellent candidates for next-generation IoT networks. Unfortunately, little is known about the behavior of modern random-access protocols in terms of information freshness. The first results in this direction were presented in [16], where the focus was on irregular repetition slotted ALOHA [14]. There, non-trivial trade-offs between spectral efficiency and average AoI were revealed.

Contributions: To further tackle this open research question, we concentrate in the present work on frameless ALOHA. This protocol, originally proposed in [15], operates according to the same principle of rateless codes, and has emerged as a particularly promising approach. Specifically, frameless ALOHA allows terminals to transmit copies of their packets over a contention period whose duration is dynamically tuned by the receiver. In contrast to previous works, which assume a fixed number of contending users, we focus on a more general and realistic setup in which the number of users accessing the channel may vary over time, driven by the duration of previous contention periods. We track the rich dynamic evolution of the system by means of a Markovian analysis, and derive its stationary throughput. Moreover, by obtaining an exact formula for the attainable average peak AoI, we provide the first study of the information freshness achievable by this protocol. Our analysis highlights the critical role played by some key protocol parameters, such as the maximum length of the contention period, and shows that operating the system to maximize throughput comes at the expense of an AoI degradation—a trade-off that is fundamentally different from what previously noted for traditional ALOHA strategies.

II. System Model and Preliminaries

We focus on a system in which U users share a wireless channel to communicate with a common receiver (sink). Time is divided in slots of fixed duration, equal to the length of a packet, and all terminals are slot-synchronous. The medium is shared among all users according to a grant-free approach, and a collision channel model is assumed. Specifically, the transmission of two or more packets over a slot leads to a destructive collision, which prevents immediate retrieval of all colliding packets at the sink. On the contrary, packets sent over *singleton slots* are always decoded correctly.

Channel access is regulated by the frameless ALOHA protocol [15], which operates in successive contention periods (CPs) of not necessarily equal length. The receiver initiates a new CP by broadcasting a beacon, whose duration is considered negligible throughout our analysis. At this point, every user with data attempts transmission of its packet over each subsequent slot with probability q, potentially sending multiple copies of the same packet over the CP. Conversely,

users that do not have a packet to send at the time of beacon reception refrain from accessing the channel for the whole duration of the CP. The procedure continues until a new beacon sent by the sink notifies the end of the current CP and the start of the next one.

At the receiver side, decoding of a packet over a singleton slots triggers successive interference cancellation (SIC). Specifically, the interference contribution of all the copies of the retrieved packet is removed, possibly leading to new singleton slots and thus to the decoding of previously collided packets.[1] The receiver proceeds with this operation mode on a slot-by-slot basis, and terminates the CP when either all transmitting users have been decoded or a maximum number $d_{\max}$ of slots has been reached.[2] An example of the frameless-ALOHA operations is discussed in Fig. 1.

As to traffic generation, we assume every user to independently generate a new packet over each slot with probability γ. This packet is stored in a one-packet-sized buffer for later delivery. A pre-emption policy with replacement in waiting is implemented, so that, at any given time instant, a user either has one packet to send (the last generated one) or has an empty buffer. Accordingly, a user will attempt transmission over a CP only if it has generated at least one packet over the previous CP. Assume that the previous CP lasted for d slots. Then, an arbitrary user has a packet to transmit with probability

$$\gamma_d := 1 - (1 - \gamma)^d.$$

All copies of the packet sent by each user during a CP are marked with a common time stamp, set to the start time of the CP. Since all users are assumed to generate traffic independently, the number U of users that become active at the end of a CP of d slots is binomial distributed with parameters (U, γ_d). In the remainder of the paper, we shall denote the corresponding probability mass function (PMF) as

$$P_{U|D}(u|d) := \binom{\mathsf{U}}{u} \gamma_d{}^u (1 - \gamma_d)^{\mathsf{U}-u}. \tag{1}$$

Finally, no retransmissions are considered: if a packet is not decoded during the CP it is sent over, it is simply discarded.

We are interested in evaluating the ability of the system to maintain an up-to-date record of the state of each user at the sink. To this aim, we consider the AoI $\Delta(t)$ of a generic user:

$$\Delta(t) := t - \sigma(t).$$

Here, $\sigma(t)$ is the time stamp of the last update received by the sink from the user of interest as of time t. This metric grows linearly over time, and drops each time the receiver successfully decodes a packet from the user under observation. For simplicity, we will assume throughout that these refreshes take place at the end of the CP over which the status update

was received, i.e., we do not track the exact slot in which the corresponding packet was decoded.[3] This yields the saw-tooth profile exemplified in Fig. 2.

We will focus on the *average peak AoI* Ω^*, defined as the mean value of $\Delta(t)$ when sampled right before an update from the user of interest is decoded. This metric, introduced in [17], characterizes the maximum value reached on average by the AoI of a user.

A. Operational Details

We next describe some operational details of the protocol that will be relevant for the subsequent analyses. At the end of each slot, the sink attempts to decode as many users as possible, canceling also their interference. When no more users can be decoded, i.e., when the contention contains no more singleton slots, the receiver decides whether to terminate the CP or not. Specifically, the CP is concluded only if all active users have been decoded, or alternatively if a maximum number of slots has elapsed since the beginning of the contention. Note, however, that, without further assumptions, it is in general not possible for the sink to determine whether all active users have been decoded, since the sink cannot discriminate between inactive users, who do not have a packet to transmit, and active users who do have a packet to transmit, but have not (yet) transmitted their packet since the beginning of the CP.

To allow the sink to determine whether all active users have been decoded, we set the slot access probability to 1 in the first slot of every contention period. This implies that all active users will transmit their packet in the first slot. Furthermore, we make the reasonable assumption that the receiver can distinguish among empty slots, singleton slots containing exactly one packet, and *collided slots* containing two or more packets. Under this assumption, the sink can use the first slot of every CP to determine whether all active users have been decoded or not. In particular, after canceling the interference from a decoded user, the sink can check whenever the first slot becomes empty to infer whether there are no more undecoded active users and the CP can be terminated. This strategy allows the receiver also to detect empty CPs. Indeed, these CPs are characterized by an empty initial slot. Note that the minimum CP duration in our setting is one slot, reached when either no users or a single user have data to transmit.

We emphasize that more realistic and sophisticated methods may be devised to estimate the number of active users in the CP, as discussed for instance in [18]. For the purpose of the analysis provided in this paper, the proposed technique suffices, in the sense that it provides a simple model for the cost (i.e., overhead) required for the estimation of the number of active users.

B. Notation

In the remainder of the paper, we denote a discrete r.v. and its realization as X and x, respectively, whereas the corresponding PMF is indicated by $P_X(x)$. A conditional PMF

---

[1]Note that, in order to implement this procedure, the sink needs to know of the position of all the replicas of a packet. This can be achieved, for instance, by using a hash function of the payload as seed of a pseudo-random generator, used by the transmitter to determine the slots of the CP over which to transmit. Upon decoding the payload, the sink becomes thus aware of all the slots occupied by the user, effectively allowing the removal of the interference of that user throughout the CP.

[2]Details on how the sink can determine whether all users have been decoded will be presented in Sec. II-A.

[3]As will be clarified in Sec. V, this assumption does not change the fundamental trade-offs of interest, and the analysis can be easily adapted to account for this additional factor.
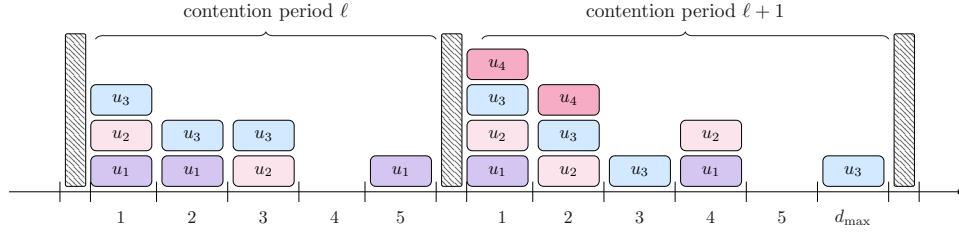
Fig. 1. Example of operations for frameless ALOHA over two successive CPs. We assume $\mathsf{U} = 4$ users in the system and a maximum contention duration of $d_{\max} = 6$ slots. Within the first CP, only three users are active. The receiver decodes the first packet in slot 5, retrieving the status update of user u_1. By removing its interference contribution from slot 2, the sink can then decode the packet of user u_3. Finally, after removing the interference caused by user 3, the sink can also obtain the packet of user 2. Having decoded all users, the sink sends a new beacon at the end of slot 5, initiating the next CP. All four users attempt transmission. The first decoding occurs at slot 3, leading to the retrieval of u_3. Removal of such packet from slot 2, however, does not resolve completely the existing collision, and SIC stops. The situation does not change after slot 4 (collision not involving u_3), slot 5 (idle), or slot 6, which contains the transmission of a resolved user, and the receiver terminates the CP as the maximum number of slots has been reached, even if some users (i.e., u_1, u_2 and u_4) have not been decoded. Note that the first slot of each CP is used by all active users to send a packet, allowing the sink to infer when complete decoding has occurred (see Sec. II-A).
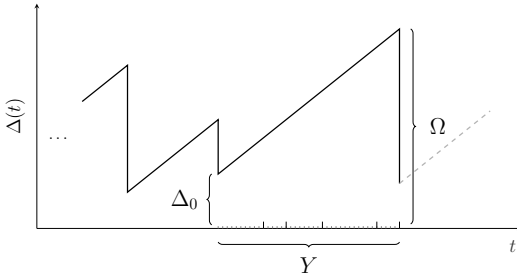

Fig. 2. Evolution over time of the AoI $\Delta(t)$ for a user. In the plot, Y denotes the duration (in slots) of an inter-update period, possibly composed by multiple CPs. The value at which AoI was reset upon reception of the last update from the node is denoted by Δ_0, which in our case corresponds to the duration of the CP in which the previous update was received. Accordingly, Ω indicates the value reached by $\Delta(t)$ at the end of the CP over which the next update from the node is decoded. An in-depth discussion of these quantities will be presented in Sec. V.

is denoted as $P_{X|Y}(x|y)$. We further write the state of a discrete-time Markov chain at time ℓ as $X^{(\ell)}$, and express its one-step transition probability from state i to state j as

$$p_X(i,j) := \mathbb{P}\{X^{(\ell+1)} = j \mid X^{(\ell)} = i\}.$$

In the case of bi-dimensional Markov chains, we maintain the same notation, but denote the state by means of a two-element vector, e.g., $j = (j_1, j_2)$.

III. FRAMELESS ALOHA ANALYSIS

Following [19], we model the iterative SIC process at the sink using a finite-state machine. A state is identified by the triplet (w, c, r), where w denotes the number of unresolved users, c denotes the number of collided slots (ignoring the initial slot), and r is the number of singleton slots. We denote by Pre_d the pre-decoding state, i.e., the state right after the sink observes the d-th slot within a CP and before it tries to decode any new packets, whereas Pos_d denotes the post-decoding state, i.e., the state after SIC decoding. To describe the decoding process, we next provide a characterization of the conditional probability of Pos_d given Pre_d and of the conditional probability of Pre_d given Pos_{d-1}.

A. State Initialization

Assume that u users are active. The state is initialized as

$$\mathrm{Pre}_1 = \begin{cases} (0,0,0) & \text{if } u = 0 \\ (1,0,1) & \text{if } u = 1 \\ (u,0,0) & \text{if } u \geq 1. \end{cases}$$

B. Conditional Probability of Pos_d Given Pre_d

We next derive the conditional probability of the post-decoding state Pos_d given the pre-decoding state $\mathrm{Pre}_d = (w, c, r)$. Two cases need to be distinguished: $r = 0$ and $r = 1$. Indeed, in the pre-decoding state, we always have $r \in \{0, 1\}$, since the reception of a new slot yields at most one new singleton slot. If $r = 0$, the state remains unchanged since no users can be resolved. Hence, we have

$$\mathbb{P}\{\mathrm{Pos}_d = (w', c', r')|\mathrm{Pre}_d = (w, c, 0)\}$$
$$= \begin{cases} 1 & \text{if } w' = w, \ c' = c, \ r' = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Let us now focus on the case $r = 1$. It is convenient to describe SIC decoding as an iterative process in which one user is resolved at a each iteration, potentially resulting in new singleton slots. The iterative process is terminated when no singleton slots are available. This implies that the post-decoding state must have $r = 0$. To characterize the state evolution at each SIC iteration, we use [19, Theorem 1]. This theorem, when specialized to the scenario considered here, implies that, if the state is (w, c, r) with $w \geq 1$ and $r \geq 1$, after resolving exactly one user, the state becomes $(w-1, c-j, r-i+j+a)$ with probability

$$I_w(a)\binom{c}{j}h_w^j(1-h_w)^{c-j}\binom{r-1}{i-1}\left(\frac{1}{w}\right)^{i-1}\left(1-\frac{1}{w}\right)^{r-i}$$

for $0 \leq a \leq 1$, $0 \leq j \leq c$, $1 \leq i \leq r-1$ and $i - j - a \leq r$, where

$$I_w(a) = \begin{cases} 1 & \text{if } w \neq 2, \ a = 0 \\ 1 & \text{if } w = 2, \ a = 1 \\ 0 & \text{otherwise} \end{cases}$$

and with

$$h_w = \frac{\sum\limits_{k=2}^{u-w+2} \Lambda_k k(k-1)\frac{1}{u}\frac{w-1}{u-1}\frac{\binom{u-w}{k-2}}{\binom{u-2}{k-2}}}{1 - \sum\limits_{k=1}^{u-w+1} \Lambda_k w\frac{\binom{u-w}{k-1}}{\binom{u}{k}} - \sum\limits_{k=0}^{u-w} \Lambda_k \frac{\binom{u-w}{k}}{\binom{u}{k}}}$$

where $\Lambda_k = \binom{u}{k}q^k(1-q)^{u-k}$.

To derive the desired conditional probability $\mathbb{P}\{\mathrm{Pos}_d = (w',c',0)|\mathrm{Pre}_d = (w,c,1)\}$ for all values of w', and c', we apply the result just stated iteratively, stopping when we reach a state with no singleton slots.

C. Contention Termination

The CP is terminated after $d < d_{\max}$ slots only if all u active users are resolved, i.e., only if the post-decoding state is $\mathrm{Pos}_d = (0,0,0)$. However, when $d = d_{\max}$, the CP is terminated, no matter what the value of $\mathrm{Pos}_{d_{\max}}$ is.

D. Conditional Probability of Pre_d Given Pos_{d-1}

We now analyze how the state changes when one slot is added to the CP. To do so, we derive the conditional probability of the pre-decoding state Pre_d given the post-decoding state $\mathrm{Pos}_{d-1} = (w,c,0)$, for $d \geq 2$. Three different cases must be considered. In the first case, the extra slot contains no packet from any of the w unresolved users. This event, which occurs with probability $(1-q)^w$, yields a pre-decoding state $\mathrm{Pre}_d = (w,c,0)$. Hence, we have

$$\mathbb{P}\{\mathrm{Pre}_d = (w,c,0)|\mathrm{Pos}_{d-1} = (w,c,0)\} = (1-q)^w.$$

In the second case, the extra slot contains the packet of exactly one of the w unresolved users. It can then be verified that

$$\mathbb{P}\{\mathrm{Pre}_d = (w,c,1)|\mathrm{Pos}_{d-1} = (w,c,0)\} = wq(1-q)^{w-1}.$$

Finally, in the third case, the extra slot contains the transmission of two or more unresolved users, which yields

$$\mathbb{P}\{\mathrm{Pre}_d = (w,c+1,0)|\mathrm{Pos}_{d-1} = (w,c,0)\} = 1 - (1-q)^w$$

$$- wq(1-q)^{w-1}.$$

E. Derivation of Some Useful Quantities

We will next use the state-transition probabilities just introduced to derive three main quantities that will turn out important for the characterization of the average peak AoI.

The first quantity is the probability that the contention period is terminated after exactly d slots, given that the number of active users is u. We denote this quantity by $P_{D|U}(d|u)$. To characterize it, we need to consider three different cases. The first one is $d = 1$. In this case, we have

$$P_{D|U}(1|u) = \begin{cases} 1 & \text{if } u \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

The second case covers $1 < d < d_{\max}$. Recall that, in this case, the CP is terminated only if all u active users are resolved. Hence,

$$P_{D|U}(d|u) = \mathbb{P}\{\mathrm{Pos}_d = (0,0,0)\}. \tag{2}$$

The probability of the remaining case, $d = d_{\max}$, can be easily obtained as

$$P_{D|U}(d_{\max}|u) = 1 - \sum_{d=1}^{d_{\max}-1} P_{D|U}(d|u). \tag{3}$$

The second quantity we are interested in is the conditional probability that exactly m users were decoded given that u users were active. We denote this quantity by $P_{M|U}(m|u)$. To characterize it, we must distinguish two cases: $m < u$ and $m = u$. When $m < u$, since not all users were resolved, the CP was terminated after $d_{\max}$ slots. Hence, we have

$$P_{M|U}(m|u) = \sum_c \mathbb{P}\{\mathrm{Pos}_{d_{\max}} = (u-m,c,0)\}.$$

Assume now $m = u$. To obtain $P_{M|U}(u|u)$, we need to add the probabilities of all post-decoding states in which all users are decoded:

$$P_{M|U}(u|u) = \sum_{d=1}^{d_{\max}} \mathbb{P}\{\mathrm{Pos}_d = (0,0,0)\}$$

$$= 1 - \sum_{m=0}^{u-1} P_{M|U}(m|u).$$

The third quantity of interest, which we denote by $\beta(m,u)$, is the conditional probability that m users are resolved, given that u users accessed the CP and that the CP ran until its maximum duration $d_{\max}$. We can obtain $\beta(m,u)$ by summing the probabilities of all post-decoding states $\mathrm{Pos}_{d_{\max}}$ in which exactly $u-m$ active users are unresolved, and then normalizing by the sum of the probabilities of all states $\mathrm{Pos}_{d_{\max}}$:

$$\beta(m,u) = \frac{\sum_c \mathbb{P}\{\mathrm{Pos}_{d_{\max}} = (u-m,c,0)\}}{\sum_w \sum_c \mathbb{P}\{\mathrm{Pos}_{d_{\max}} = (w,c,0)\}}.$$

IV. THROUGHPUT PERFORMANCE

We provide in this section an analysis of the stationary throughput achievable with the frameless ALOHA protocol. This analysis will turn out useful for the characterization of the AoI. Previous works, e.g., [15], [18]–[20], have studied the protocol behavior either over a single CP, or under the assumption that the number of contending terminals is fixed. For this scenario, the number of packets that can be decoded under an optimized access probability has been characterized. The setting under consideration in this paper, however, is characterized by a richer dynamic, since the number of users with packets to transmit, and thus the level of contention, may vary over time. To appreciate this aspect, observe how, for instance, a long CP increases the probability for more users to generate at least one packet over its duration. This leads to a harsher contention over the successive period, which, in turn, is likely to last longer. Similarly, contentions resolved in few slots will instead drive the system on average towards shorter and less loaded CPs.

To capture the impact on throughput of this non-trivial evolution, we focus on the homogeneous Markov processes $D^{(\ell)}$ and $U^{(\ell)}$, tracking the duration of the ℓ-th CP and the number of users contending over it, respectively. Let us first consider the former, which takes values in the set

$\{1, \ldots, d_{\max}\}$. Recalling that the duration of the $(\ell + 1)$-th CP is driven by the number of users contending over it, we compute the transition probabilities for the chain as

$$
\begin{aligned}
p_D(i, j) &= \sum_{u=0}^{\mathsf{U}} \mathbb{P}\{D^{(\ell+1)} = j \,|\, U^{(\ell+1)} = u\} \\
&\qquad \times \mathbb{P}\{U^{(\ell+1)} = u \,|\, D^{(\ell)} = i\} \\
&\overset{(a)}{=} \sum_{u=0}^{\mathsf{U}} P_{D|U}(j|u)\, P_{U|D}(u|i)
\end{aligned}
$$

where (a) follows from (1), (2), and (3). Similarly, the transitions for the finite-state chain $U^{(\ell)}$ are

$$
\begin{aligned}
p_U(i, j) &= \sum_{d=1}^{d_{\max}} \mathbb{P}\{U^{(\ell+1)} = j \,|\, D^{(\ell)} = d\} \\
&\qquad \times \mathbb{P}\{D^{(\ell)} = d \,|\, U^{(\ell)} = i\} \\
&= \sum_{d=1}^{d_{\max}} P_{U|D}(j|d)\, P_{D|U}(d|i).
\end{aligned}
$$

In both cases, it is easy to verify that the finite-state chains are irreducible and aperiodic, and thus ergodic. In the remainder of the discussion, we shall indicate their stationary distributions, derived by solving the corresponding balance equations, as $\pi_D(d)$ and $\pi_U(u)$, respectively.

Let us now denote by $M^{(\ell)}$ the number of successfully decoded users over the ℓ-th CP. The system throughput S, defined as the *average* number of decoded packets per slot, is

$$
\mathsf{S} := \lim_{t \to \infty} \frac{\frac{1}{t} \sum_{\ell=1}^{t} M^{(\ell)}}{\frac{1}{t} \sum_{\ell=1}^{t} D^{(\ell)}}. \tag{4}
$$

Observing now that

$$
\mathbb{P}\{M^{(\ell)} = m\} = \sum_{u=0}^{\mathsf{U}} P_{M|U}(m|u)\, \mathbb{P}\{U^{(\ell)} = u\}
$$

we conclude that the statistics of $M^{(\ell)}$ can be directly derived from that of the number of contending users over the corresponding CP. Hence, this process has also a stationary distribution. Accordingly, both numerator and denominator in (4) admit finite limits for $t \to \infty$ by virtue of the ergodicity of the involved chains. This allows us to to compute S as the ratio of the expected values of the processes in stationary conditions:

$$
\mathsf{S} = \frac{\sum_{m=0}^{\mathsf{U}} \sum_{u=0}^{\mathsf{U}} m\, P_{M|U}(m|u)\, \pi_U(u)}{\sum_{d=1}^{d_{\max}} d\, \pi_D(d)}.
$$

Leaning on this result, we provide a first characterization of the behavior of the system in Fig. 3. The plot shows how the stationary throughput changes as a function of the transmission probability q, considering a population of $\mathsf{U} = 100$ users, and a maximum duration for a CP of $d_{\max} = 100$ slots. The reported results were obtained by setting the activation probability γ such that the average number of users generating a new packet over each slot, $\gamma\mathsf{U}$, equals 0.6. The exhibited trend confirms the existence of an optimal medium access probability for throughput performance. Indeed, too low values of q tend to result in successful yet unnecessarily long CPs, where many
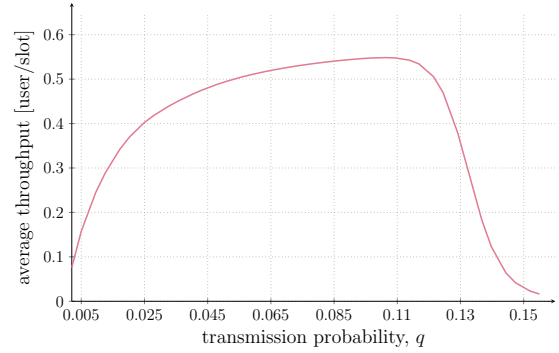

Fig. 3. System throughput S vs transmission probability q. A population of $\mathsf{U} = 100$ users and a maximum CP duration of $d_{\max} = 100$ slots are considered. The packet generation probability is set so that $\gamma\mathsf{U} = 0.6$.

slots may remain unused. Conversely, when users become too aggressive, collisions become predominant, leading to the sharp decrease in throughput which is typically observed in grant-free schemes that resort to SIC.

Such a behavior is confirmed in Fig. 4, which shows the stationary distribution of the CP duration, $\pi_D(d)$, and of the number of decoded packets per CP, denoted as $\pi_M(m)$, for low ($q = 0.01$), intermediate ($q = 0.1$) and high ($q = 0.15$) values of the transmission probability. Consider first the case $q = 0.01$. In this situation, an active node will send no copy of its packet even over a CP of maximum duration with probability $(1 - q)^{d_{\max}} \simeq 0.37$. As a result, the sink tends to operate with long CPs, awaiting for packets of users that have not yet transmitted. In terms of throughput, while an average of around 45 users will participate in the contention for a CP of $d_{\max} = 100$ slots, only a significantly lower fraction of them is decoded, as reported by the stationary distribution $\pi_M(m)$. Notably, the system operates at CPs of maximum duration also for high values of transmission probability ($q = 0.15$). In this case, though, hardly any packet is decoded, due to the excessive level of congestion, which leads to unresolvable collisions. Both settings result in poor throughput performance, as illustrated in Fig. 3. An efficient utilization of the channel is instead achieved for $q = 0.1$, where a proper balance between frame duration and number of active users is reached. A first relevant trade-off driven by the dynamics of the system thus emerges, calling for a proper balance of the operating parameters.

V. AVERAGE PEAK AGE OF INFORMATION

We now analyze the performance of frameless ALOHA in terms of information freshness. We start with some preliminary results that will facilitate the characterization of the average peak AoI Ω^*.

A. Preliminaries

Fix a generic user for which the AoI is tracked, and denote by $\nu(u, d)$ the conditional probability that the user delivers a status update over the current CP, given that u users contend, and that the CP is terminated after d slots. Recall that whenever the CP is terminated prior to its maximum duration,
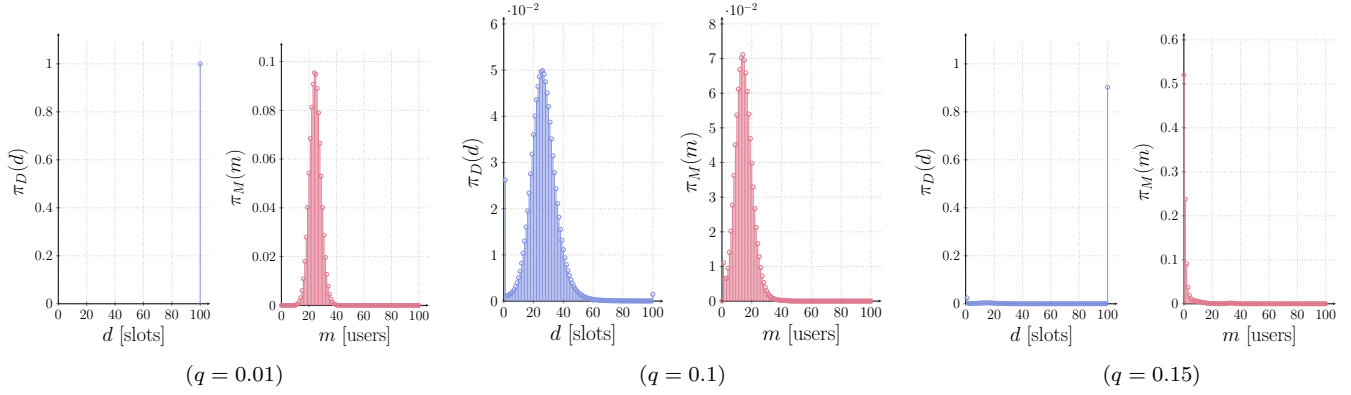
Fig. 4. Stationary distribution of the CP duration, $\pi_D(d)$, and of the number of decoded users per CP, $\pi_M(m)$, for three different values of transmission probability q. In all cases, results were generated considering $\mathsf{U} = 100$, $d_{\max} = 100$, and $\gamma\mathsf{U} = 0.6$. The *average* number of active users per CP in stationary conditions in the different cases is 45.22 for $q = 0.01$, 14.49 for $q = 0.1$, and 41.68 for $q = 0.15$.

all contending users are successfully decoded. The user of interest belongs to this pool with probability u/U. Conversely, if the CP runs for $d_{\max}$ slots, the conditional probability for the user to deliver a packet given that m users are successfully decoded is $(m/\mathsf{U})\beta(m,u)$. Combining these two results we then have:

$$\nu(u,d) = \begin{cases} \dfrac{u}{\mathsf{U}} & d < d_{\max} \\ \displaystyle\sum_{m=0}^{\mathsf{U}} \dfrac{m}{\mathsf{U}}\,\beta(m,u) & d = d_{\max}. \end{cases} \quad (5)$$

Next, we introduce a simple ancillary Markov chain, whose state is defined as $Z^{(\ell)} = (D^{(\ell)}, S^{(\ell)})$. The first component, which we have already discussed, characterizes the duration of the ℓ-th CP, whereas $S^{(\ell)}$ is a binary r.v. taking value 1 if an update from the user of interest has been successfully received over the ℓ-th CP, and 0 otherwise. Consider now the probability for the chain to transition from state (j, s) to state $(d, 1)$. By definition, this event occurs when the current CP has duration d slots, and the user delivers an update. Observing that the user's success does not depend on its outcome over the previous CP, we can simplify the transition probability to

$$p_Z((j,s),(d,1)) := \mathbb{P}\{Z^{(\ell+1)} = (d,1)\,|\,Z^{(\ell)} = (j,s)\}$$
$$= \mathbb{P}\{S^{(\ell+1)} = 1, D^{(\ell+1)} = d\,|\,D^{(\ell)} = j\}.$$

Conditioning now on the number of users contending over the CP, we further have

$$\mathbb{P}\{S^{(\ell+1)} = 1, D^{(\ell+1)} = d\,|\,D^{(\ell)} = j\}$$
$$= \sum_{u=0}^{\mathsf{U}} \mathbb{P}\{S^{(\ell+1)} = 1\,|\,U^{(\ell+1)} = u, D^{(\ell+1)} = d\}$$
$$\times \mathbb{P}\{U^{(\ell+1)} = u, D^{(\ell+1)} = d\,|\,D^{(\ell)} = j\}$$
$$= \sum_{u=0}^{\mathsf{U}} \mathbb{P}\{S^{(\ell+1)} = 1\,|\,U^{(\ell+1)} = u, D^{(\ell+1)} = d\}$$
$$\times \mathbb{P}\{D^{(\ell+1)} = d\,|\,U^{(\ell+1)} = u\}\,\mathbb{P}\{U^{(\ell+1)} = u\,|\,D^{(\ell)} = j\}.$$

Finally, using (5), (1), (2), and (3), we can write $p_Z((j,s),(d,1))$ compactly as

$$p_Z((j,s),(d,1)) = \sum_{u=0}^{\mathsf{U}} \nu(u,d)\,P_{D|U}(d|u)\,P_{U|D}(u|j). \quad (6)$$

Following similar steps, we can express the transition probabilities from a generic state (j, s) to a state $(d, 0)$ in which the user does not deliver an update as

$$p_Z((j,s),(d,0)) = \sum_{u=0}^{\mathsf{U}} (1 - \nu(u,d))\,P_{D|U}(d|u)\,P_{U|D}(u|j). \quad (7)$$

It is immediate to verify that the finite-state Markov chain $Z^{(\ell)}$ is irreducible and aperiodic, and admits thus a stationary distribution, which we denote as $\pi_Z(d,s)$.

B. Average Peak AoI

Let us now focus on the calculation of the average peak AoI achieved by the frameless ALOHA policy. To this aim, denote by Ω the r.v. describing the value of the AoI in stationary conditions at the end of a CP over which the user of interest delivers an update. As exemplified in Fig. 2, this quantity can be conveniently expressed as $\Omega = \Delta_0 + Y$. Here, Δ_0 captures the value at which the AoI $\Delta(t)$ was lastly reset, and corresponds to the duration of the CP in which the previous update was received. The r.v. Y accounts for the duration of the inter-update time, i.e., the number of slots that have elapsed between the last and the current successful reception of a status update from the user. The average peak AoI can then be computed as

$$\Omega^* = \mathbb{E}[\Delta_0] + \mathbb{E}[Y]. \quad (8)$$

Consider first Δ_0. Its PMF $P_{\Delta_0}(\delta_0)$ can be readily computed from the stationary distribution of the Markov chain $Z^{(\ell)}$. We have indeed

$$P_{\Delta_0}(\delta_0) = \frac{\pi_Z(\delta_0, 1)}{\sum_{\delta=1}^{d_{\max}} \pi_Z(\delta, 1)}$$

where the numerator denotes the probability for the system to be in a CP of duration δ_0 slots in which the tracked user

is decoded, and the denominator is a normalization factor, capturing that we are interested only in CPs with successful updates from that user. The first addend in (8) can be evaluated simply by applying the definition of expected value.

We analyze now $\mathbb{E}[Y]$. We start by noting that the statistics of Y depends on the r.v. Δ_0. In fact, the duration of the CP over which the last update was received does influence the number of users contending on the subsequent one, impacting both the probability for the user of interest to transmit and be decoded as well as the duration of the subsequent CPs. It is then convenient to express $\mathbb{E}[Y]$ as

$$\mathbb{E}[Y] = \sum_{\delta_0=1}^{d_{\max}} \mathbb{E}[Y \mid \Delta_0 = \delta_0] \cdot P_{\Delta_0}(\delta_0). \qquad (9)$$

Without loss of generality, let us denote by $\ell = 1$ the index of the first CP that contributed to the inter-update time being tracked. Accordingly, we reformulate the conditional expectation in (9) considering the value of $Z^{(1)}$ as

$$\mathbb{E}[Y \mid \Delta_0 = \delta_0]$$
$$= \sum_z \mathbb{E}[Y \mid Z^{(1)} = z, \Delta_0 = \delta_0]\, \mathbb{P}\{Z^{(1)} = z \mid \Delta_0 = \delta_0\}$$
$$\overset{(a)}{=} \sum_z \mathbb{E}[Y \mid Z^{(1)} = z]\, \mathbb{P}\{Z^{(1)} = z \mid \Delta_0 = \delta_0\} \qquad (10)$$

where the summation is taken over all the possible states $z = (d, s)$, $d \in \{1, \ldots, d_{\max}\}$, $s \in \{0, 1\}$, and (a) follows from the Markov property of the involved processes. Note that the factors $\mathbb{P}\{Z^{(1)} = z \mid \Delta_0 = \delta_0\}$ on the right-hand side of (10) can be computed using (6) and (7).

The conditional expectation of Y given the outcome of the first CP of the inter-update period, $\mathbb{E}[Y \mid Z^{(1)} = z, \Delta_0 = \delta_0]$, can be derived resorting to a first step analysis [21]. To this aim, let us denote by $\mathcal{A}$ the set of states for the Markov chain $Z^{(\ell)}$ corresponding to an update delivery for the user of interest:

$$\mathcal{A} := \{z = (d, s) \mid d \in \{1, \ldots, d_{\max}\}, s = 1\}.$$

We refer to $\mathcal{A}$ as the set of *absorbing* states, and define the chain absorption time as

$$\ell^* := \min\{\ell \geq 1 \mid Z^{(\ell)} \in \mathcal{A}\}.$$

Furthermore, let us assign to each state $z = (d, s)$ a cost

$$g((d, s)) = d.$$

Then, the overall duration (in slots) of the inter-update period Y is simply given by the sum of the costs undergone for all states traversed by the process $Z^{(\ell)}$ from time $\ell = 1$ until absorption, i.e.,

$$\mathbb{E}[Y \mid Z^{(1)} = z] = \mathbb{E}\left[\left. \sum_{\ell=1}^{\ell^*} g\big(Z^{(\ell)}\big) \right| Z^{(1)} = z \right].$$

To compute this quantity, consider first the situation in which the packet from the user of interested is decoded already in the initial CP. In this case, the chain is immediately absorbed, and Y coincides with the length of the initial CP:

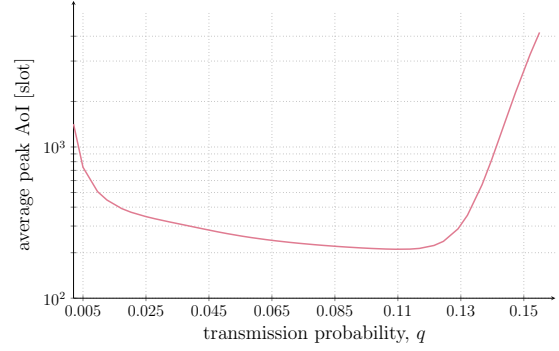$$\mathbb{E}[Y \mid Z^{(1)} = (d, 1)] = d. \qquad (11)$$

Fig. 5. Average peak AoI Ω^* vs. transmission probability q. Results obtained with $\mathsf{U} = 100$ users, $d_{\max} = 100$ slots. The packet generation probability has been set such that $\gamma\mathsf{U} = 0.6$.

When $Z^{(1)} = (d, 0) \notin \mathcal{A}$, instead, the average cost prior to absorption can be computed by conditioning on the outcome of the first transition. Specifically,

$$\mathbb{E}[Y \mid Z^{(1)} = (d, 0)] = d + \sum_z \mathbb{E}[Y \mid Z^{(1)} = z] \cdot p_Z((d, 0), z) \qquad (12)$$

where the Markov property ensures that the average cost once the transition to state z is taken is equal to the one that we would have by starting from such state. Combining (11) and (12), we get a full-rank system of $d_{\max}$ equations in the $d_{\max}$ unkowns $\mathbb{E}[Y \mid Z^{(1)} = (d, 0)]$. We obtain $\mathbb{E}[Y]$ by solving such system, and by insert the results into (10) and (9). Finally, the sought average peak AoI Ω^* follows as per (8).

In Fig. 5, we plot the average peak AoI against the channel access probability q for the same setting considered in Fig. 3, i.e., $\mathsf{U} = 100$, $d_{\max} = 100$, $\gamma\mathsf{U} = 0.6$. As illustrated in the plot, both too low and too high values of q result in poor performance in terms of AoI. In the former case, an excessively conservative behavior is likely to result in a user missing opportunities to deliver a status update, refraining from transmission for the whole duration of a CP even when a packet is available. Conversely, collisions dominate when users become too aggressive, hindering the capability of the receiver to decode transmitted updates prior to reaching the maximum contention duration.

It is also important to note that, for a given traffic profile (U, γ) and a given maximum frame duration, the optimal operating points in terms of throughput and average peak AoI coincide. In other words, there exists a value q^* of the transmission probability that jointly maximizes S and minimizes Ω^*. This outcome is common to other random access solutions under symmetric traffic conditions, as epitomized by the inverse proportionality of AoI and throughput exhibited by slotted ALOHA [9], [16]. From this standpoint, indeed, any choice of $q \neq q^*$ reducing the probability to deliver a status update would also be harmful in terms of information freshness.

In contrast, frameless ALOHA exhibits a more complex behavior when performance are analyzed versus $d_{\max}$. In Fig. 6, we report the optimal throughput and peak AoI pairs that can be achieved by tuning the maximum CP duration. Specifically, we explore different values of $d_{\max}$ between 10
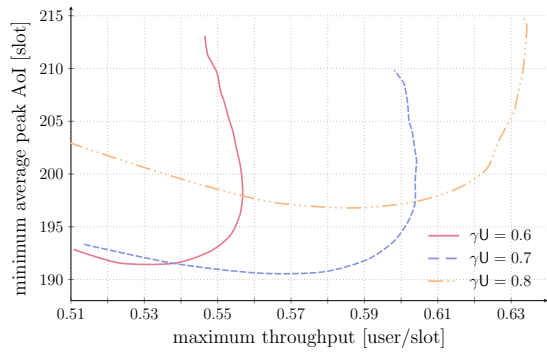
Fig. 6. Maximum throughput and minimum average peak AoI obtained when varying the maximum CP duration $d_{\max}$ in the range $\{10, 150\}$ slots. In all cases, $\mathsf{U} = 100$ users were considered.

and 150 slots, and pick, for each setting, the optimal access probability q^*, plotting the corresponding values obtained for S and Ω^*. Distinct curves in the figure refer to different packet generation probabilities. In all cases, $\mathsf{U} = 100$.

Consider first the case $\gamma\mathsf{U} = 0.6$, represented by the solid curve in Fig. 6, and focus on throughput performance. For low values of $d_{\max}$, the system operates in the lower-left corner of the plot. Too short CPs hinder packet decoding, not allowing enough slots for SIC to be fully efficient. By increasing the maximum duration of the CP, S improves, approaching the elbow exhibited by the reported trend. After a certain point, though, a further increase of $d_{\max}$ enables to decode only a limited additional number of users, and such diminishing-return behavior leads to a decrease in throughput.

Notably, while a similar trend emerges also for the average peak AoI, the impact of operating over excessively long CPs is far more pronounced. The rationale behind this lies in the dependency of Ω^* on the inter-update time, i.e., the number of CPs between two updates as well as their duration in slots. From this standpoint, higher values of $d_{\max}$ may reduce the former (increasing throughput), yet entail a larger average cost in terms of elapsed slots over a CP. While initially the first factor prevails, and Ω^* improves together with S, the impact of longer CPs quickly emerges and yields a reduction of Ω^*.

This trend is confirmed when analyzing the results obtained by increasing the channel load $\gamma\mathsf{U}$. As shown in Fig. 6, frameless ALOHA can reach larger values of throughput by supporting higher levels of channel congestion. Yet operating in this region is detrimental in terms of average peak AoI.

VI. CONCLUSIONS

We provided a characterization of the average throughput and the average peak AoI achievable in a random access system operating according to the frameless ALOHA protocol. One novel aspect of our analysis, compared to results available in the literature, is that we are able to capture the rich system dynamics that result from letting the number of users contending on a given CP depend on the duration of the previous CP. We show that this modeling assumption is critical to capture practically relevant trade-offs in the design of the system. Specifically, choosing the maximum length of the CP

so as to maximize the system throughput may result in a significant deterioration of the average peak AoI. We hasten to add that the performance of frameless ALOHA protocol considered in the paper can be further improved by making some of the protocol parameters dependent on the duration of the previous CP. Such a generalization will be addressed in a future work.

REFERENCES

[1] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.

[2] E. Uysal, O. Kaya, A. Ephremides, J. Gross, M. Codreanu, P. Popovski, M. Assaad, G. Liva, A. Munari, T. Soleymani, B. Soret, and K. Johansson, "Semantic communications in networked systems," 2021. [Online]. Available: http://arxiv.org/abs/2103.05391

[3] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minizing age of information in vehicular networks," in *Proc. IEEE SECON*, June 2011.

[4] O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, "Age-of-information vs. value-of-information scheduling for cellular networked control systems," in *Proc. ACM/IEEE ICCPS*, Apr. 2019.

[5] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the Wiener process for remote estimation over a channel with random delay," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 1118–1135, Feb. 2020.

[6] N. Abramson, "The ALOHA System - Another Alternative for Computer Communications," in *Proc. 1970 Fall Joint Computer Conference*. AFIPS Press, Nov. 1970.

[7] LoRa Alliance, "The LoRa Alliance Wide Area Networks for Internet of Things," www.lora-alliance.org.

[8] Sigfox, "SIGFOX: The Global Communications Service Provider for the Internet of Things," www.sigfox.com.

[9] R. Yates and S. Kaul, "Status updates over unreliable multiaccess channels," in *Proc. IEEE ISIT*, June 2017.

[10] ——, "Age of information in uncoordinated unslotted updating," in *Proc. IEEE ISIT*, Jun. 2020.

[11] O. Yavascan and E. Uysal, "Analysis of slotted ALOHA with an age threshold," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1456–1470, May 2021.

[12] X. Chen, K. Gatsis, H. Hassani, and S. Bidokhti, "Age of information in random access channels," in *Proc. IEEE ISIT*, Jun. 2020.

[13] M. Berioli, G. Cocco, G. Liva, and A. Munari, "Modern random access protocols," *Foundations and Trends® in Networking*, vol. 10, no. 4, pp. 317–446, 2016.

[14] E. Paolini, G. Liva, and M. Chiani, "Coded Slotted ALOHA: A Graph-Based Method for Uncoordinated Multiple Access," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6815–6832, Dec. 2015.

[15] C.Stefanović, P. Popovski, and D. Vukobratovic, "Frameless ALOHA Protocol for Wireless Networks," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 2087–2090, Dec. 2012.

[16] A. Munari, "Modern random access: an age of information perspective on irregular repetition slotted ALOHA," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3572–3585, Jun. 2021.

[17] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *Proc. IEEE ISIT*, Jun. 2014.

[18] C. Stefanović and P. Popovski, "ALOHA Random Access that Operates as a Rateless Code," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4653–4662, Nov. 2013.

[19] F. Lázaro, C. Stefanović, and P. Popovski, "Reliability-latency performance of frameless ALOHA with and without feedback," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6302–6316, Oct. 2020.

[20] C. Stefanović, K. F. Trilingsgaard, N. K. Pratas, and P. Popovski, "Joint estimation and contention-resolution protocol for wireless random access," in *Proc. IEEE ICC*, Jun. 2013.

[21] H. Taylor and S. Karlin, *An Introduction to Stochastic Modeling*, 3rd ed. London: Academic Press, 1998.