

# A Palm Calculus Approach to the Distribution of the Age of Information

Amr Rizk, *Senior Member, IEEE*, and Jean-Yves Le Boudec, *Fellow, IEEE*

**Abstract**—A key metric to express the timeliness of status updates in latency-sensitive networked systems is the age of information (AoI), i.e., the time elapsed since the generation of the last received informative status message. This metric allows studying a number of applications including updates of sensory and control information in cyber-physical systems and vehicular networks as well as, job and resource allocation in cloud clusters. State-of-the-art approaches to analyzing the AoI rely on queueing models that are composed of one or many queueing systems endowed with service order, e.g., FIFO, LIFO, or last-generated-first-out order. A major difficulty arising in these analysis methods is capturing the AoI under message reordering when the delivery is non-preemptive and non-FIFO, i.e., when messages can overtake each other and the reception of informative messages may obsolete some messages that are underway. In this paper, we derive an exact formulation for the distribution of AoI in non-preemptive, non-FIFO systems where the main ingredients of our analysis are Palm calculus and time inversion. Owing to the rationality of the Laplace-Stieltjes transforms that are used in our approach, we obtain computable exact expressions for the distribution of AoI.

## I. INTRODUCTION

Cyber-physical systems (CPS) constitute a type of hybrid system that combines physical processes and computation [1]. Often, the considered physical process such as those arising in chemical plants or platoons of automated vehicles are controlled via feedback loops. Due to this sensor-computation-actuator feedback loop, CPS are characterized by the mutual interaction of the physical process, the software computation, and essentially, the network.

While CPS encompass diverse key interactions worth accurate modeling such as control correctness and concurrency we are concerned in the following with the effort of characterizing the timeliness of sensor data when received at the controller. This is the first step to ensure that the actions taken by the controller and hence executed by the actuator are based on fresh information. A key metric to express this timeliness of sensor data at the controller is the Age of Information (AoI) [2], which has recently been a vivid object of study [3, 4]. AoI is a semi-continuous function that denotes the age of the sensor (sender) status at the controller (receiver). The status age is hence best described by a *jump-and-drift* process that grows linearly with time and jumps downwards at the time points when informative messages arrive at the receiver. An informative message is defined as a message containing an update that was generated after the generation time point

of the last received update at the controller. Now, the time points at which messages arrive at the receiver, as well as the timestamps contained in these messages, are random and essentially dependent on the generation and transmission of messages at the sender and at every network node on the path from the sender to the receiver. Figure 2 shows a sketch of this scenario where a newer message<sup>1</sup> ( $m_2$ ) overtakes an older message ( $m_1$ ). Note that the age at the receiver does not jump downwards at the reception of the outdated message  $m_1$ .

One research direction to optimize the timeliness of sensor information in CPS is through advancing the state-of-the-art physical layer techniques such as deterministically reserved transmission time slots over all available frequencies as in low-latency 5G network slices known as URLLC [5]. While this eliminates contention on the wireless link in 5G, data packet interactions and sporadic network congestion still occur on the end-to-end path between the sensors and the controller.

Research on the topic of AoI has been characterized by the analysis of mathematical models that capture the stochastic process of the age at the receiver given a combination of ingredients, i.e., (i) the process of data generation and transmission at the source, (ii) a model of the network interactions such as traffic scheduling and the variability of the link transmission rate. The prevalent approach in many works on AoI is to capture these ingredients in form of a queueing system (or a series thereof) that naturally capture the former and models the latter through the service process. In many works the arrival process is often considered as a Poisson process for tractability [6, 7] or as a periodic process to capture simple sensor device implementations [8]. The variety of queueing models ranges from simple M/M/1 queues with FIFO service to preemptive Last-Generated, First-Served (LGFS) systems. A remarkable difficulty of some AoI models is due to the lack of FIFO service. Allowing messages to overtake each other leads to considerable complexity as shown in the basic example in Figure 2. A direct approach to model AoI systems with non-FIFO service is presented in [9] using the Stochastic-Hybrid-System (SHS) technique. This technique essentially depends on the Fokker-Planck partial differential equation (PDE) satisfied by the time-dependent probability density of the AoI as shown in [3] and quickly becomes intractable. For a comprehensive overview we point the reader to [10].

The **key differences** of this work to the works in [7, 10, 11] are: (i) Instead of considering the time variant PDE of the density of the age we reduce the problem to a simple model where we are interested in the stationary distribution of the age in a system where message overtaking is allowed. (ii)

A. Rizk is with the University of Duisburg-Essen, Essen, Germany. (e-mail: amr.rizk@uni-due.de).

Jean-Yves Le Boudec is with the École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland (e-mail: jean-yves.leboudec@epfl.ch).

<sup>1</sup>in terms of the generation time point

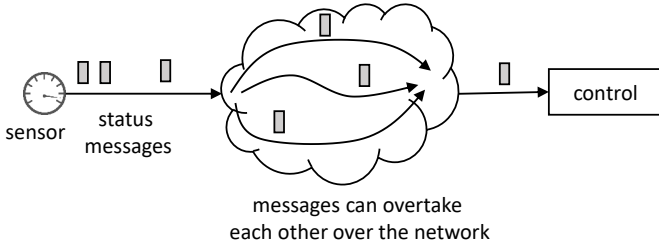


Fig. 1: The sensory information is generated and immediately transmitted in form of messages. These can overtake each other on the network. Informative messages keep the total message order at the receiver and reduce the age of the status information at the receiver to their respective one way delay.

We obtain the distribution of the age using Palm calculus and time inversion where we essentially require to know the joint distribution of the age when a message arrives and the time until the next message arrives. Note that the obtained inversion formula applies to all types of arrivals in this model but we apply it to the informative messages to obtain the age density at the receiver. This model naturally captures the distribution of functions of the age. (iii) Due to the used mathematical tools our results only require *stationarity* of the underlying queuing model, which is a Markov process on a discrete state space and can thus be analyzed with elementary techniques. (iv) The model considered in this paper is different from [7, 11] as we consider a window flow controlled sender that injects at most a fixed number of non-obsolete messages into the network channel. We denote this model as  $M/M/I_{\max}/I_{\max}^*$ .

Our contributions in this paper are summarized as follows:

- We use Palm calculus and time inversion to derive the probability distribution of the age of information in a stationary  $M/M/I_{\max}/I_{\max}^*$  system.
- We calculate the Laplace-Stieltjes transform of the distribution of the age at the arrival time points of informative messages as well as at any point in time.

## II. PROBLEM STATEMENT AND SYSTEM MODEL

We consider cyber-physical systems as depicted in Fig. 1 where sensors transmit status updates to a central control and data acquisition function. We assume that timestamped messages are transmitted at the sender according to some parameterized stochastic process. When a message is generated, it obsoletes any previous message. However, every message is subject to a one way delay and *messages can overtake each other*. We say that a message is “informative” if its timestamp was generated after the generation time of all messages received so far. When a non informative message arrives, it is of no use and is discarded.

We are interested in the age of information at the receiver,  $X_t$ , which is formally defined as follows. Timestamped messages are generated at times  $\{\tau_i\}$  and received at times  $\{\tau'_i\}$  respectively (with  $\tau_i \leq \tau'_i$ ). Then

$$X_t = t - \max_{i: \tau'_i \leq t} \tau_i \quad (1)$$

The dynamic evolution of  $X_t$  is such that the age  $X_t$  increases at rate 1 between arrival events; furthermore, when message  $i$

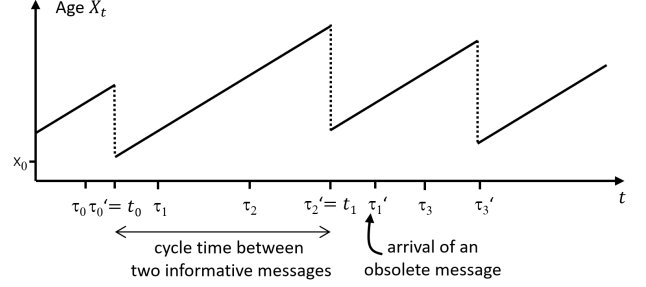


Fig. 2: The age process  $X_t$ . Message  $m_i$  is emitted at time  $\tau_i$  and received at time  $\tau'_i$ . Observe that message  $m_1$  is overtaken by message  $m_2$  hence the age process at the receiver does not change when  $m_1$  arrives.

arrives, the value of  $X_t$  just after the arrival, namely  $X_{\tau'_i+}$ , is set to  $\min(\tau'_i - \tau_i, X_{\tau'_i-})$  as seen in Figure 2.

We assume that messages are generated according to a Poisson process of rate  $\lambda$ . The channel is modelled as a number of independent parallel servers each serving at most one message at a time with exponentially distributed service times, i.e. the random variables  $\tau'_i - \tau_i$  are independent of each other and of the arrival process  $\{\tau_i\}$  and they are exponentially distributed with same parameter  $\mu$ . Furthermore, in order to not overwhelm the channel the sender is window-flow-controlled and allows only a fixed number of outstanding informative messages  $I_{\max}$  in the channel: arriving messages are dropped if the number of outstanding informative messages is equal to  $I_{\max}$ . We assume that the sender knows the number of informative messages in the channel (presumably via some instantaneous reverse channel). We use the notation  $M/M/I_{\max}/I_{\max}^*$  for this queueing system, where the  $*$  here means that the departure of a message flushes all older messages out of the system.

In this paper we are interested in the stationary distribution of the age  $X_t$  given the process parameters  $\lambda, \mu$  and  $I_{\max}$ .

The global notation used in the paper is recalled in Table I.

## III. A PALM CALCULUS APPROACH TO THE AOI

### A. The Underlying Queuing Model

First we consider a continuous time Markov jump process  $\{Z_t\}_{t \geq 0}$  that models the  $M/M/I_{\max}/I_{\max}^*$  queue described in the previous section. Let  $Z_t$  represent the number of messages underway from the sender to the receiver, for  $t \in \mathbb{R}^+$ , with  $Z_t \in E = \{0, 1, \dots, I_{\max}\}$ . Recall that, by our modelling assumption, this counts only informative messages. At any time  $t$  such that  $Z_t = n > 0$ , and for  $i \in \{1, \dots, n\}$  we call  $i$ th message, the message with the  $i$ th smallest timestamp among all messages present in the channel.

When the sender generates a new message at time  $t$  (which occurs at constant rate  $\lambda$ ), if  $Z_{t-} < I_{\max}$  then the message is accepted in the channel and  $Z_t$  is incremented by 1, i.e.  $Z_{t+} = Z_{t-} + 1$ ; else, i.e. if  $Z_{t-} = I_{\max}$ , the message is discarded and  $Z_t$  is unchanged.

Consider now message departures from the channel. Whenever  $Z_t = n > 0$  all  $n$  messages in the channel can leave the channel with same rate  $\mu$ , thus the rate of message departure

TABLE I: Notation List

$\tilde{d}_n, d'_n$	$\tilde{d}_n = \sum_{n'} Q_{n,n'}, d'_n = \sum_{n'} Q'_{n,n'}$
$f(x)$	PDF of age of information at received at an arbitrary point in time;
$f^\circ(x_0, t_1)$	Joint PDF of age of information $x_0$ and time to wait until next delivery of informative message, sampled when an informative message arrives at receiver;
$f_A^\circ(x_0)$	PDF of age of information sampled when an informative message arrives at receiver;
$f_{n',n}$	Laplace-Stieltjes Transform of $x_0 \mapsto g^\circ(x_0 n', n)$
$\tilde{f}_n$	Laplace-Stieltjes Transform of $t_1 \mapsto h(t_1 n)$
$g^\circ(x_0 n', n)$	PDF of the age $x_0$ just after an informative message arrival given that the state of the Markov chain is $n'$ just before the arrival of the informative message and $n$ just after the arrival;
$h(t_1 n)$	PDF of the time that will elapse from time $t$ until the next informative message arrives, given that $Z_t = n$ ;
$I_{\max}$	Maximum number of messages in transit; messages generated when $Z_t = I_{\max}$ are discarded;
$\lambda$	Rate of generation of messages;
$\mu$	Message transit time is exponential with rate $\mu$ ;
$\bar{N}$	$= \sum_{i=1}^{I_{\max}} i p_i$
$p_n$	Stationary probability of $Z_t$
$p'_{n',n}$	Probability that an arbitrary informative message arrival happens at a transition ( $n' \rightarrow n$ ) of the Markov chain $Z_t$
$Q_{i,j}, Q'_{i,j}$	Rate of transition of $Z_t$ [resp. $Z_t^r$ ] from state $i$ to state $j$
$X_t$	Age of information at receiver at time $t$ ;
$Z_t$	Number of messages in transit at time $t$ ;
$Z_t^r$	Time-reversed process derived from $Z_t$

is  $n\mu$  and all messages are equally likely to leave the channel. Assume that a departure occurs at time  $t$  and  $Z_{t-} = n > 0$ . For  $i \in \{1 \dots n\}$ , the probability that the departing message is the  $i$ th message is  $\frac{1}{n}$ . In this case,  $Z_t$  is decremented by  $i$ , i.e.  $Z_{t+} = Z_{t-} - i$ ; in other words, the transition  $n \rightarrow n - i$  occurs at rate  $\mu$  for every  $i \in \{1 \dots n\}$ .

Thus  $Z_t$  is a continuous-time Markov chain with finite state space  $E$  and with transition rates (Fig. 3):

$$\begin{aligned} Q_{i,i+1} &= \lambda, \quad i = 0 \dots I_{\max} - 1 \\ Q_{i,j} &= \mu, \quad i = 1 \dots I_{\max}, 0 \leq j \leq i - 1 \\ Q_{i,j} &= 0, \quad \text{otherwise.} \end{aligned} \quad (2)$$

Observe that  $Z_t$  can be regarded as the number of messages in a FIFO queue with Poisson arrivals of rate  $\lambda$  and drained using a batch service process. It is ergodic as the state space is finite and fully connected.

Using the balance equation, the steady state probabilities  $p_n$  can be computed and are given by:

$$p_n = \begin{cases} \frac{(n+1)\lambda^n \mu}{\prod_{j=1}^{n+1} (\lambda + j\mu)} & \text{for } 0 \leq n < I_{\max} \\ \frac{\lambda^n}{\prod_{j=1}^n (\lambda + j\mu)} & \text{for } n = I_{\max} \end{cases} \quad (3)$$

The derivation of (3) is given in appendix IX-A.

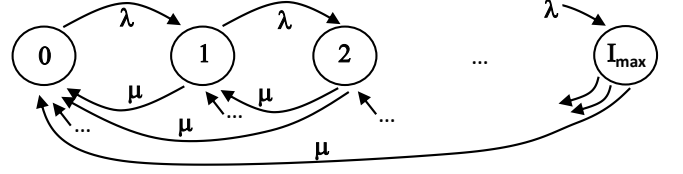


Fig. 3: State transition diagram of the Markov chain  $Z_t$  representing the number of non-obsolete messages underway.

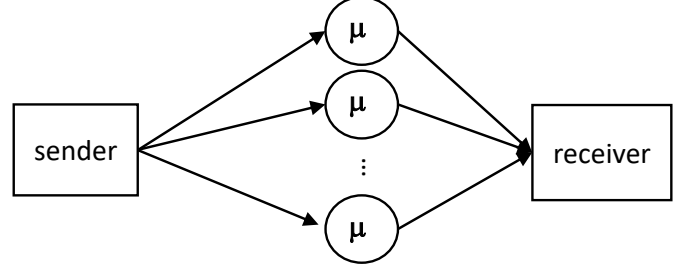


Fig. 4: Our system model assumes that messages are transmitted upon generation and take random iid one way delay to reach the receiver. Hence the system model assumes for every message an independent channel each with exponentially distributed service time with identical parameter  $\mu$ . We assume a network channel (as sketched in Fig. 1) that is constrained by a finite number of informative messages under way denoted by  $I_{\max}$ . This assumption corresponds to a window flow constrained sender with a maximum number of outstanding informative messages  $I_{\max}$  given a perfect reverse channel.

### B. A Palm Calculus Approach

In the following we use Palm Calculus to compute the stationary distribution of age. To this end, we assume that the continuous time Markov chain  $Z_t$  is in its unique stationary regime, which, since it is ergodic, occurs in practice if the system has been operating for a long time. With Palm calculus, we are able to relate the stationary distribution of the age to quantities that are computed for the Markov chain  $Z_t$ .

Palm calculus [12, 13, 14] applies to a stationary point process  $T_n$  ( $n \in \mathbb{Z}$ ) and an observable (random) process  $X_t$  ( $t \in \mathbb{R}$ ) that are jointly stationary. Here we take for  $T_n$  the sequence of times at which a departure occurs from the  $M/M/I_{\max}/I_{\max}^*$  queue  $Z_t$  (i.e. when  $Z_t$  is decremented, which also corresponds to arrivals of informative messages at the receiver). Since we assume  $Z_t$  is in its stationary regime, this point process is also stationary. In the context of Palm calculus, it is customary to assume that the numbering convention is such that  $T_0 \leq 0 < T_1$ . The observable  $X_t$  is the age of information at the receiver, as defined earlier. Note that  $X_t$  can be computed in a deterministic way from the trajectory  $Z_{(-\infty, t]}$  and is invariant with respect to change of time origins, therefore it is jointly stationary with  $Z_t$ , hence with  $T_n$  [14, Section 7.2.1]. Also note that Palm calculus does not require the point process to be Poisson (the arrival process is Poisson by definition, but it can be seen that the departure process is not).

We can now apply Palm's inversion formula [14, Theorem 7.1], which states that, for any bounded, measurable test

function  $\varphi$  we have

$$\mathbb{E}[\varphi(X_t)] = \hat{\lambda} \mathbb{E}^\circ \left[ \int_{T_0=0}^{T_1} \varphi(X_s) ds \right] \quad (4)$$

In the above,  $\mathbb{E}^\circ$  stands for the Palm expectation, which is the conditional expectation given that the point process has a point at time 0 (i.e. given that there is a departure from the  $M/M/I_{\max}/I_{\max}^*$  queue at time 0)<sup>2</sup>. Also, under this conditional expectation,  $T_0 = 0$  and  $T_1$  is the following departure instant. Last,  $\hat{\lambda}$  is the intensity of the point process of departures, which can be calculated from the Markov chain as  $\hat{\lambda} = \sum_i ip_i \mu = \mu \bar{N}$  with the stationary expectation of  $Z_t$  denoted as  $\bar{N} := \sum_{i=1}^{I_{\max}} ip_i$ .

Observe that obtaining  $\mathbb{E}[\varphi(X_t)]$  for arbitrary  $\varphi$  is equivalent to finding the stationary distribution of the age of information at an arbitrary point in time. Applying these ideas to the AoI gives the following theorem:

**Theorem III.1.** *The stationary PDF of the age of information at an arbitrary point in time,  $f(x)$ , is given by*

$$f(x) = \hat{\lambda} \int_{x-x_0}^{\infty} \int_0^x f^\circ(x_0, t_1) dx_0 dt_1 \quad (5)$$

where  $f^\circ(x_0, t_1)$  denotes the joint PDF of the age  $x_0$  just after an informative message arrival and of the time  $t_1$  that will elapse until the next informative message arrives.

Note that  $f^\circ(x_0, t_1)$  is a Palm PDF, i.e. it corresponds to observations made upon the arrival of an informative message. Following the conventions in [12], we use a  $^\circ$  superscript to denote a Palm PDF.

*Proof:* We apply Palm's inversion formula (4). Next, note that for  $0 \leq s \leq T_1$  we have  $X_s = s + X_{0+}$ , therefore

$$\mathbb{E}^\circ \left[ \int_0^{T_1} \varphi(X_s) ds \right] = \mathbb{E}^\circ \left[ \int_0^{T_1} \varphi(s + X_{0+}) ds \right] \quad (6)$$

By definition of  $f^\circ(x_0, t_1)$ , it follows that

$$\begin{aligned} & \mathbb{E}^\circ \left[ \int_0^{T_1} \varphi(X_s) ds \right] \\ &= \int_0^\infty \int_0^\infty \int_0^{t_1} \varphi(x_0 + s) ds f^\circ(x_0, t_1) dx_0 dt_1 \\ &= \int_0^\infty \varphi(u) \int_0^u \int_{u-x_0}^\infty f^\circ(x_0, t_1) dt_1 dx_0 du, \end{aligned} \quad (7)$$

where, in the last line we substituted  $u = x + s$  with  $x \leq u \leq x + t$ .

Now we find the PDF of the age at any arbitrary point in time  $f(x)$  from comparing (7) with

$$\mathbb{E}[\varphi(X_t)] = \int_0^\infty \varphi(x) f(x) dx \quad (8)$$

where we find

$$f(x) = \hat{\lambda} \int_0^x \int_{x-x_0}^\infty f^\circ(x_0, t_1) dt_1 dx_0 \quad (9)$$

<sup>2</sup>Note that the definition of the conditional expectation can be given a meaning even though the probability of the point process having a point at time 0 exactly is 0 [14, Section 7.2.2].

In the following, we calculate the Palm distribution  $f^\circ(x_0, t_1)$ . This is tractable because it involves  $Z_t$  only, which is a Markov process on a finite state space. ■

#### IV. COMPUTING THE PALM DISTRIBUTIONS

##### A. Decomposition into Forward and Backward Components

In order to compute the Palm PDF  $f^\circ(x_0, t_1)$  we observe that the part on  $x_0$  (the age) involves the past of  $Z_t$  whereas the part on  $t_1$  (time until a new arrival) involves the future of  $Z_t$ . This is captured by the following theorem.

**Theorem IV.1.** *The joint PDF of both the age  $x_0$  just after the informative message arrival and the length of the cycle  $t_1$  until the arrival of the next informative message is given by*

$$f^\circ(x_0, t_1) = \sum_{(n', n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n', n}^\circ g^\circ(x_0 | n', n) h(t_1 | n) \quad (10)$$

where

- $p_{n', n}^\circ$  is the probability that an arbitrary message arrival happens at a transition ( $n' \rightarrow n$ ) of the Markov chain  $Z_t$  and is given by

$$p_{n', n}^\circ = \frac{p_{n'}}{\bar{N}} \mathbf{1}_{1 \leq n+1 \leq n' \leq I_{\max}} \quad (11)$$

in the above,  $\bar{N}$  is the stationary expectation of  $Z_t$  and  $p_i$  is given in (3);

- $g^\circ(x_0 | n', n)$  is the PDF of the Palm distribution of the age  $x_0$  just after the informative message arrival given that the state of the Markov chain is  $n'$  just before the arrival of the informative message and  $n$  just after the arrival, where  $n' \geq n + 1$ ;
- $h(t_1 | n)$  is the stationary PDF of the time that will elapse from time  $t$  until the next informative message arrives, given that  $Z_t = n$ .

The proof exploits the Markov property of  $Z_t$ , which expresses that the future depends on the past only through the present state.

*Proof:* Define  $f^\circ(x_0, t_1 | n', n)$  as the joint PDF of the Palm distribution of both the age  $x_0$  just after the informative message arrival and the length of the cycle  $t_1$  until the arrival of the next informative message, given that the state of the Markov chain is  $n'$  just before the arrival of the informative message and  $n$  just after the arrival, where  $n' \geq n + 1$ . It follows that the required PDF  $f^\circ(x_0, t_1)$  is given by

$$f^\circ(x_0, t_1) = \sum_{(n', n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n', n}^\circ f^\circ(x_0, t_1 | n', n) \quad (12)$$

where  $p_{n', n}^\circ$  is the probability that an arbitrary message arrival happens at a transition ( $n' \rightarrow n$ ) of the Markov chain  $Z_t$ . By [15, Thm 7.1.2], such a probability is given by

$$p_{n', n}^\circ = \eta p_{n'} Q_{n', n}, \quad \mathbf{1}_{1 \leq n+1 \leq n' \leq I_{\max}} \quad (13)$$

where  $\mathbf{1}_{\{\cdot\}}$  is the indicator function, equal to 1 when the condition is true and 0 otherwise,  $p_{n'}$  is the stationary probability

given in (3),  $Q_{n',n}$  is the transition rate in (2) and  $\eta$  is a normalizing constant. Observe that  $Q_{n',n} = \mu$ , which gives  $\eta^{-1} = \sum_{i=1}^{I_{\max}} ip_i = \bar{N}$  where  $\bar{N}$  is the stationary expectation of  $Z_t$ . It finally comes

$$p_{n',n}^\circ = \frac{p_{n'}}{\bar{N}} \mathbf{1}_{1 \leq n+1 \leq n' \leq I_{\max}} \quad (14)$$

Let  $g^\circ(x_0|n',n)$  denote the PDF of the Palm distribution of the age  $x_0$  just after the informative message arrival given that the state of the Markov chain is  $n'$  just before the arrival of the informative message and  $n$  just after the arrival, where  $n' \geq n+1$ .

Recall that  $h(t_1|n)$  denotes the stationary PDF of the time that will elapse from time  $t$  until the next informative message arrives, given that  $Z_t = n$ . By the Markov property, this is also the PDF of the Palm distribution of the time until the next informative message arrives given that the state of the Markov chain is  $n'$  just before the arrival of the informative message and  $n$  just after the arrival. Again by the Markov property,  $f(x_0, t_1|n',n) = g^\circ(x_0|n',n)h(t_1|n)$ , which proves (10). ■

We next compute  $h(t_1|n)$ , which we call the forward component of (10). The computation of the backward component  $g^\circ(x_0|n',n)$  will involve a similar method plus a time-reversal argument.

### B. Computation of the Forward Component

First, we will introduce the following lemma to calculate the Laplace-Stieltjes Transform (LST) of the time until the occurrence of the *next transition of interest* in a continuous-time Markov chain  $\{\tilde{Z}(t)\}_{t \in \mathbb{R}_+}$  conditioned on  $Z_t = n$ . The transitions of interest are defined by some subset  $\tilde{\mathcal{F}}$  of  $E \times E$ , where  $E \subseteq \mathbb{N}$  is the state-space of the Markov chain.

**Lemma IV.2.** *Consider a time-homogeneous, continuous-time Markov chain  $(\tilde{Z}_t)_{t \in \mathbb{R}_+}$  with state space  $E \subseteq \mathbb{N}$  and with transition rates  $\tilde{Q}_{n,n'}$ ; let  $\tilde{d}_n = \sum_{n' \in E} \tilde{Q}_{n,n'}$  denote the sum of all outgoing rates from state  $n$  and assume that  $\tilde{d}_n > 0$  for all  $n \in E$ . Let  $\tilde{\mathcal{F}} \subseteq E$  such that  $\tilde{Q}_{n,n'} > 0$  for all  $(n, n') \in \tilde{\mathcal{F}}$ .*

*Call  $\tilde{Y}_t$  the time that will elapse from  $t$  until the next jump in  $\tilde{\mathcal{F}}$  of the Markov chain, i.e.  $\tilde{Y}_t = \inf\{s > 0, (Z_{(t+s)^-}, Z_{(t+s)^+}) \in \tilde{\mathcal{F}}\}$ . The conditional LST of  $\tilde{Y}_t$  given that  $\tilde{Z}_t = n$ , denoted as  $\tilde{f}_n(\nu)$ , satisfies*

$$\begin{aligned} \tilde{f}_n(\nu) &:= \mathbb{E} \left[ e^{-\nu \tilde{Y}_t} | \tilde{Z}_t = n \right] \\ &= \frac{1}{\tilde{d}_n + \nu} \left( \sum_{\substack{n', \\ (n,n') \notin \tilde{\mathcal{F}}}} \tilde{f}_{n'}(\nu) \tilde{Q}_{n,n'} + \sum_{\substack{n', \\ (n,n') \in \tilde{\mathcal{F}}}} \tilde{Q}_{n,n'} \right). \end{aligned} \quad (15)$$

*Proof:* Fix some arbitrary time  $t$  and define  $\tilde{S}_t$  as the time until the next transition (of interest or not) out of state  $\tilde{Z}_t$  and let  $N'_t := Z_{t+\tilde{S}_t}$  denote the next state. It is known [16] that, conditional to  $\tilde{Z}_t = n$ ,  $N'_t$  and  $\tilde{S}_t$  are independent, the distribution of  $\tilde{S}_t$  is exponential with rate  $\tilde{d}_n$  and the distribution of  $N'_t$  is given by  $\mathbb{P}(N'_t = n' | \tilde{Z}_t = n) = \frac{\tilde{Q}_{n,n'}}{\tilde{d}_n}$ .

It follows that

$$\mathbb{P} \left[ N'_t = n' | \tilde{Z}_t = n, \tilde{S}_t = s \right] = \frac{\tilde{Q}_{n,n'}}{\tilde{d}_n} \quad (16)$$

and

$$\mathbb{E} \left[ e^{-\nu \tilde{S}_t} \right] = \frac{\tilde{d}_n}{\tilde{d}_n + \nu} \quad (17)$$

Also let  $\tilde{R}_t$  denote the residual time from the next transition until the next transition of interest, i.e.  $\tilde{R}_t = 0$  whenever  $(\tilde{Z}_t, N'_t) \in \tilde{\mathcal{F}}$  and otherwise  $\tilde{R}_t = \tilde{Y}_{t+\tilde{S}_t}$ . Hence

$$\tilde{Y}_t = \tilde{S}_t + \tilde{R}_t \quad (18)$$

By conditioning on  $\tilde{S}_t = s$  we can write

$$\begin{aligned} \mathbb{E} \left[ e^{-\nu \tilde{Y}_t} | \tilde{Z}_t = n, \tilde{S}_t = s \right] \\ = e^{-\nu s} \mathbb{E} \left[ e^{-\nu \tilde{R}_t} | \tilde{Z}_t = n, \tilde{S}_t = s \right] \end{aligned} \quad (19)$$

By conditioning with respect to  $N'_t$  in the latter term and applying (16) we obtain

$$\begin{aligned} \mathbb{E} \left[ e^{-\nu \tilde{R}_t} | \tilde{Z}_t = n, \tilde{S}_t = s \right] \\ = \sum_{n' \in E} \left( \mathbb{E} \left[ e^{-\nu \tilde{R}_t} | \tilde{Z}_t = n, \tilde{S}_t = s, N'_t = n' \right] \times \right. \\ \left. \mathbb{P} \left[ N'_t = n' | \tilde{Z}_t = n, \tilde{S}_t = s \right] \right) \\ = \sum_{n' \in E} \left( \mathbb{E} \left[ e^{-\nu \tilde{R}_t} | \tilde{Z}_t = n, \tilde{S}_t = s, N'_t = n' \right] \frac{\tilde{Q}_{n,n'}}{\tilde{d}_n} \right) \end{aligned} \quad (20)$$

Now if  $(n, n') \in \tilde{\mathcal{F}}$  then  $\tilde{R}_t = 0$  hence

$$\mathbb{E} \left[ e^{-\nu \tilde{R}_t} | \tilde{Z}_t = n, \tilde{S}_t = s, N'_t = n' \right] = 1 \text{ if } (n, n') \in \tilde{\mathcal{F}} \quad (21)$$

Else, i.e. if  $(n, n')$  is not in  $\tilde{\mathcal{F}}$ ,  $\tilde{R}_t = \tilde{Y}_{t+\tilde{S}_t}$  is the time that remains to elapse until the next transition of interest; by the Markov property, the future of the Markov chain depends on the history only via the current state, i.e.

$$\begin{aligned} \mathbb{E} \left[ e^{-\nu \tilde{Y}_{t+\tilde{S}_t}} | \tilde{Z}_t = n, \tilde{S}_t = s, N'_t = n' \right] \\ = \mathbb{E} \left[ e^{-\nu \tilde{Y}_{t+s}} | N'_t = n' \right] = \mathbb{E} \left[ e^{-\nu \tilde{Y}_{t+s}} | Z_{t+s} = n' \right] \\ = \tilde{f}_{n'}(\nu) \end{aligned} \quad (22)$$

where the last equality is because the Markov chain is time-homogeneous.

Combining (19) with (20)-(22) gives

$$\begin{aligned} \mathbb{E} \left[ e^{-\nu \tilde{Y}_t} | \tilde{Z}_t = n, \tilde{S}_t = s \right] \\ = e^{-\nu s} \left( \sum_{\substack{n', \\ (n,n') \notin \tilde{\mathcal{F}}}} \tilde{f}_{n'}(\nu) \frac{\tilde{Q}_{n,n'}}{\tilde{d}_n} + \sum_{\substack{n', \\ (n,n') \in \tilde{\mathcal{F}}}} \frac{\tilde{Q}_{n,n'}}{\tilde{d}_n} \right) \end{aligned} \quad (23)$$

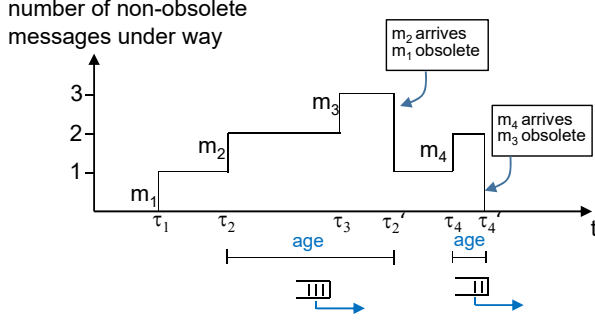


Fig. 5: A sample path of the number of non-obsolete messages under way given the system model from Sect. II. Looking at the forward process: The downward jumps mark the arrival of informative messages at the receiver which make previous messages obsolete. As messages depart in batches of random sizes in this model the waiting time of the freshest message of the batch corresponds to the age value set upon the arrival of that message at the receiver.

By the law of total expectation we can now write

$$\begin{aligned} \tilde{f}_n(\nu) &= \mathbb{E} \left[ e^{-\nu \tilde{Y}_t} | \tilde{Z}_t = n \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ e^{-\nu \tilde{Y}_t} | \tilde{Z}_t = n, \tilde{S}_t \right] \right] \\ &= \mathbb{E} \left[ e^{-\nu \tilde{S}_t} \right] \left( \sum_{\substack{n', \\ (n, n') \notin \tilde{\mathcal{F}}}} \tilde{f}_{n'}(\nu) \frac{\tilde{Q}_{n, n'}}{\tilde{d}_n} + \sum_{(n, n') \in \tilde{\mathcal{F}}} \frac{\tilde{Q}_{n, n'}}{\tilde{d}_n} \right) \end{aligned} \quad (24)$$

Using (17) completes the proof.  $\blacksquare$

Now we can use Lem. IV.2 to calculate the stationary PDF  $h(t_1|n)$  of the time that will elapse from a fixed time  $t$  until the next informative message arrives conditioned on  $Z_t = n$ . The set of transitions of interest is  $\tilde{\mathcal{F}} := \{(i, j)\}_{i > j}$ , i.e. the transitions associated with the arrival of informative messages. The transition rates  $Q_{i,j}$  are given in (2) and

$$\tilde{d}_n = \lambda \mathbf{1}_{\{n < I_{\max}\}} + \mu n \quad (25)$$

where  $\mathbf{1}_{\{\cdot\}}$  is the indicator function, equal to 1 when the condition is true and 0 otherwise.

The Laplace-Stieltjes Transform of  $h(t_1|n)$  continues to be denoted by  $\tilde{f}_n(\nu)$ ; the application of Lem. IV.2 gives:

$$\tilde{f}_n(\nu) = \frac{1}{\tilde{d}_n + \nu} \left[ n\mu + \lambda \tilde{f}_{n+1}(\nu) \right]. \quad (26)$$

This recursive relation can be rewritten using matrix notation as

$$(\nu \mathbf{I} + \tilde{\mathbf{D}}) \tilde{\mathbf{f}} = \tilde{\boldsymbol{\mu}} + \tilde{\mathbf{A}} \tilde{\mathbf{f}} \quad (27)$$

with the identity matrix  $\mathbf{I}$ , the vectors  $\tilde{\mathbf{f}} = [\tilde{f}_0(\nu), \dots, \tilde{f}_{I_{\max}}(\nu)]^T$ , and  $\tilde{\boldsymbol{\mu}} = [0, \mu, 2\mu, \dots, I_{\max}\mu]^T$ ,

and the matrices  $\tilde{\mathbf{D}} := \text{diag}(\tilde{d}_0, \tilde{d}_1, \dots, \tilde{d}_{I_{\max}})$ , and

$$\tilde{\mathbf{A}} := \begin{bmatrix} 0 & \lambda & \dots & 0 \\ \vdots & \ddots & \lambda & \vdots \\ \vdots & \ddots & \ddots & \lambda \\ 0 & \dots & 0 & 0 \end{bmatrix}. \quad (28)$$

Now we can directly solve for the conditional LSTs as

$$\tilde{\mathbf{f}} = (\nu \mathbf{I} + \tilde{\mathbf{D}} - \tilde{\mathbf{A}})^{-1} \tilde{\boldsymbol{\mu}} \quad (29)$$

### C. Computation of the Backward Component

Recall that  $g^\circ(x_0|n', n)$  denotes the PDF of the Palm distribution of the age  $x_0$  just after the informative message arrival given that the state of the Markov chain is  $n'$  just before the arrival of the informative message and  $n$  just after the arrival, where  $n' \geq n + 1$ . since the arrival of the freshest message in the served batch. For the computation of  $g^\circ(x_0|n', n)$  we resort to time-reversal as this allows to use a similar method as for the forward component.

The time-reversed process  $Z_t^r$  is defined by  $Z_t^r = Z_{-t}$ . In a nutshell, time reversal allows us to change the underlying queueing model from Sect. III-A into a FIFO queue where arrivals occur in message batches of random size while the server removes exactly *one* message on each visit. To illustrate this, consider the sample path shown in Fig. 5 in the reverse time direction.

It is shown in [17] [Section 1.7] that if  $Z_t$  is endowed with its stationary probability, then the time-reversed process is also a time-homogeneous continuous time Markov chain with same state space and same stationary probability, but with different transition rates. Specifically, by [17] [Theorem 1.12] the transition rates  $Q'_{i,j}$  for  $Z_t^r$  depend on the transition rates of the original Markov process (2) and its stationary distribution (3). We obtain  $Q'_{i,i-1} = \lambda'_i$  for  $i = 1 \dots I_{\max}$ ,  $Q'_{i,j} = \mu'_{ij}$  for  $i = 0 \dots I_{\max} - 1, i < j \leq I_{\max}$  and  $Q'_{i,j} = 0$  otherwise, with

$$\lambda'_i = \begin{cases} \frac{i}{i+1} (\lambda + (i+1)\mu) & \text{for } 1 \leq i < I_{\max} \\ I_{\max}\mu & \text{for } i = I_{\max} \end{cases} \quad (30)$$

and

$$\mu'_{ij} = \begin{cases} \frac{(j+1)\mu\lambda^{j-i}}{(i+1) \prod_{k=i+2}^{j+1} (\lambda + k\mu)} & \text{for } j \neq I_{\max} \\ \frac{\lambda^{I_{\max}-i}}{(i+1) \prod_{k=i+2}^{I_{\max}} (\lambda + k\mu)} & \text{for } j = I_{\max} \end{cases} \quad (31)$$

for  $0 \leq i < I_{\max}$ . The derivation of (30), (31) is given in the appendix.

In  $Z_t$ , upon serving a batch of messages, the sojourn time of the freshest message of that batch is the age of that particular informative message. In  $Z_t^r$ , this is given by the sojourn time of the  $(n+1)$ st message of an arriving batch of size  $n' - n$ . In  $Z_t^r$ , the arrival of a batch corresponds to a transition  $n \rightarrow n'$  with  $n' \geq n + 1$  and the size of the arriving batch is  $n' - n$ . It

follows that, for  $n' \geq n+1$ ,  $g^\circ(x_0|n', n)$  can be re-interpreted as the PDF of the time from now until the  $(n+1)$ st departure of  $Z_t^r$ , given that  $Z_t^r$  is doing a transition  $n \rightarrow n'$  now. Since  $Z_t^r$  is also Markov, we can apply the Markov property and obtain that this is simply the PDF of the time from now until the  $(n+1)$ st departure of  $Z_t^r$  given that  $Z_t^r = n'$ . For  $n+1 = 1$  this is the conditional PDF of the time until a next departure, which is exactly the problem that is solved in Lemma IV.2, and which we now extend as follows (the proof is similar and is not given).

**Lemma IV.3.** *Consider a time-homogeneous, continuous-time Markov chain  $(\tilde{Z}_t)_{t \in \mathbb{R}_+}$  with state space  $E \subseteq \mathbb{N}$  and with transition rates  $\tilde{Q}_{n,n'}$ ; let  $\tilde{d}_n = \sum_{n' \in E} \tilde{Q}_{n,n'}$  denote the sum of all outgoing rates from state  $n$  and assume that  $\tilde{d}_n > 0$  for all  $n \in E$ . Let  $\tilde{\mathcal{F}} \subseteq E$  such that  $\tilde{Q}_{n,n'} > 0$  for all  $(n, n') \in \tilde{\mathcal{F}}$ .*

*For  $k \geq 1$ , call  $\tilde{Y}_t^k$  the time that will elapse from  $t$  until the  $k$ th jump in  $\tilde{\mathcal{F}}$  of the Markov chain, i.e.  $\tilde{Y}_t^1 = \inf\{s \geq 0, (\tilde{Z}_{(t+s)^-}, \tilde{Z}_{(t+s)^+}) \in \tilde{\mathcal{F}}\}$  and for  $k \geq 2$ ,  $\tilde{Y}_t^k = \inf\{s > \tilde{Y}_t^{k-1}, (\tilde{Z}_{(t+s)^-}, \tilde{Z}_{(t+s)^+}) \in \tilde{\mathcal{F}}\}$ . The conditional LST of  $\tilde{Y}_t^k$  given that  $\tilde{Z}_t = n$ , denoted as  $\tilde{f}_{n,k}(\theta)$ , satisfies, for  $k \geq 1$ :*

$$\tilde{f}_{n,k}(\theta) = \frac{1}{\tilde{d}_n + \theta} \times \left( \sum_{\substack{n', \\ (n,n') \notin \tilde{\mathcal{F}}}} \tilde{f}_{n',k}(\theta) \tilde{Q}_{n,n'} + \sum_{\substack{n', \\ (n,n') \in \tilde{\mathcal{F}}}} \tilde{f}_{n',k-1}(\theta) \tilde{Q}_{n,n'} \right). \quad (32)$$

where  $\tilde{f}_{n,0}(\theta) = 1$  by convention.

Let  $f_{n',k}(\theta)$  be the LST of the time from now until the  $k$ th departure of  $Z_t^r$  given that  $Z_t^r = n'$ , so that the LST of  $g^\circ(x_0|n', n)$  is  $f_{n',n+1}$ . To compute  $f_{n',k}(\theta)$ , we now apply Lemma IV.3 to the Markov chain  $Z_t^r$ , with transition rate matrix  $Q'$ , and obtain:

$$f_{n',1}(\theta) [d'_{n'} + \theta] = \lambda'_{n'} \mathbf{1}_{n' > 0} + \sum_{j > n'} \mu'_{n',j} f_{j,1}(\theta) \quad (33)$$

$$f_{n',k}(\theta) [d'_{n'} + \theta] = \lambda'_{n'} \mathbf{1}_{n' > 0} f_{n',k-1}(\theta) + \sum_{j > n'} \mu'_{n',j} f_{j,k}(\theta) \quad (34)$$

for  $0 \leq n' \leq I_{\max}$  and  $k \geq 2$ . In the above,  $\lambda'$  and  $\mu'$  are given in (30) and (31) and

$$d'_{n'} = \sum_{n=0}^{I_{\max}} Q'_{n',n} \quad (35)$$

We use the following matrix notation:  $\mathbf{D} := \text{diag}(d'_0, d'_1, \dots, d'_{I_{\max}})$ ,  $\mathbf{f}_{\cdot,k} = [f_{0,k}(\theta), \dots, f_{I_{\max},k}(\theta)]^\top$ ,  $\bar{\lambda}' = [0, \lambda'_1, \dots, \lambda'_{I_{\max}}]^\top$ .  $\mathbf{M}$  is the upper triangular matrix

$$\mathbf{M} = \begin{cases} \mu'_{ij} & \text{for } i < j \\ 0 & \text{for } i \geq j \end{cases}$$

and  $\Lambda$  is the matrix with  $\lambda'_n$  on the subdiagonal defined by

$$\Lambda := \begin{bmatrix} 0 & \dots & \dots & 0 \\ \lambda'_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \lambda'_{I_{\max}} & 0 \end{bmatrix}. \quad (36)$$

for  $i, j \in \{0, 1, \dots, I_{\max}\}$ . We can rewrite the recursive relation of the conditional LST in (33) as

$$(\theta \mathbf{I} + \mathbf{D}) \mathbf{f}_{\cdot,1} = \bar{\lambda}' + \mathbf{M} \mathbf{f}_{\cdot,1} \quad (37)$$

The previous equation can be solved and we obtain:

$$\mathbf{f}_{\cdot,1} = (\theta \mathbf{I} + \mathbf{D} - \mathbf{M})^{-1} \bar{\lambda}' \quad (38)$$

Similarly, we can re-write (34) as

$$(\theta \mathbf{I} + \mathbf{D}) \mathbf{f}_{\cdot,k} = \Lambda \mathbf{f}_{\cdot,k-1} + \mathbf{M} \mathbf{f}_{\cdot,k} \quad (39)$$

for  $k \geq 2$ . Now we can construct a block matrix form that takes (37) as well as (39) to follow the form

$$\begin{bmatrix} \theta \mathbf{I} + \mathbf{D} & 0 \\ 0 & \theta \mathbf{I} + \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{f}_{\cdot,1} \\ \mathbf{f}_{\cdot,I_{\max}} \end{bmatrix} = \begin{bmatrix} \bar{\lambda}' \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{M} & 0 \\ 0 & \Lambda \end{bmatrix} \begin{bmatrix} \mathbf{f}_{\cdot,1} \\ \mathbf{f}_{\cdot,I_{\max}} \end{bmatrix} \quad (40)$$

We can directly find the vector of conditional LST as

$$\begin{bmatrix} \mathbf{f}_{\cdot,1} \\ \vdots \\ \mathbf{f}_{\cdot,I_{\max}} \end{bmatrix} = \begin{bmatrix} \theta \mathbf{I} + \mathbf{D} - \mathbf{M} & 0 \\ -\Lambda & \ddots \\ 0 & -\Lambda & \theta \mathbf{I} + \mathbf{D} - \mathbf{M} \end{bmatrix}^{-1} \begin{bmatrix} \bar{\lambda}' \\ \vdots \\ 0 \end{bmatrix} \quad (41)$$

Since the computation of (41) requires the inversion of a matrix of the order of  $I_{\max}^2 \times I_{\max}^2$  we show in the following how to calculate the conditional LST recursively from (40). We observe that  $\mathbf{M} - \mathbf{D} = \mathbf{Q}' - \Lambda$  where  $\mathbf{Q}'$  denotes the transition rate matrix of the continuous Markov chain associated with the reversed process  $Z_t^r$ . We define  $\Phi := \theta \mathbf{I} + \mathbf{D} - \mathbf{M}$  and obtain the following recursion in block matrix form

$$\begin{bmatrix} \Phi & & & \\ -\Lambda & \ddots & 0 & \\ 0 & \ddots & \ddots & \\ 0 & 0 & -\Lambda & \Phi \end{bmatrix} \begin{bmatrix} \mathbf{f}_{\cdot,1} \\ \vdots \\ \vdots \\ \mathbf{f}_{\cdot,I_{\max}} \end{bmatrix} = \begin{bmatrix} \bar{\lambda}' \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$$

Now we obtain the conditional LST  $\mathbf{f}_{\cdot,n}$  recursively with the initial condition

$$\mathbf{f}_{\cdot,1} = \Phi^{-1} \bar{\lambda}' \quad (42)$$

and for  $k \geq 2$

$$\mathbf{f}_{\cdot,k} = \Psi^{k-1} \Phi^{-1} \bar{\lambda}' \quad (43)$$

where we used the shorthand notation  $\Psi := \Phi^{-1} \Lambda$ .

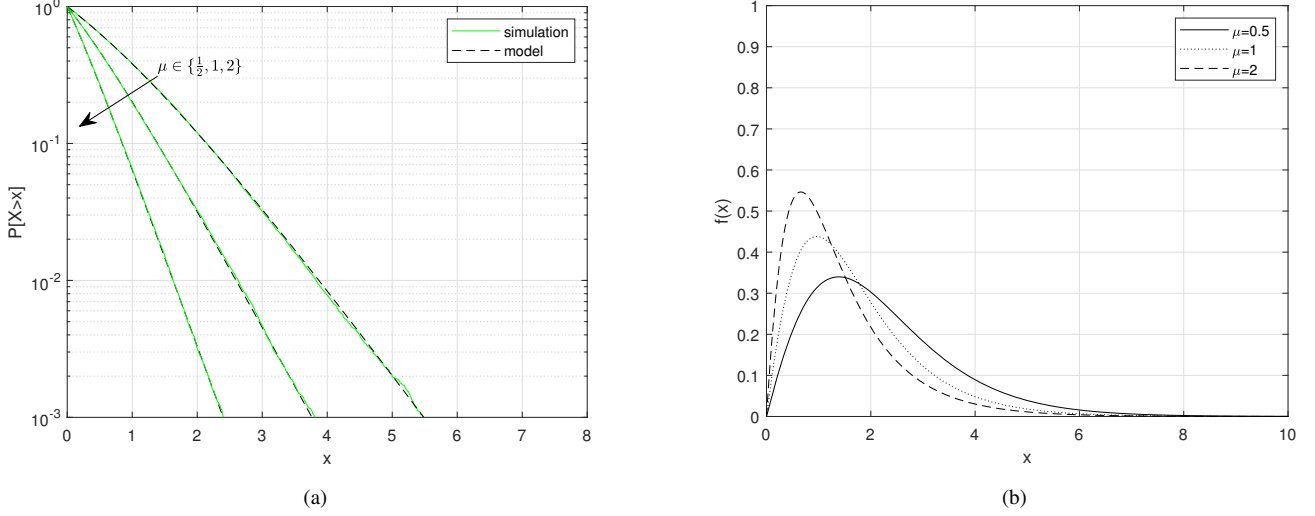


Fig. 6: (a) CCDF of the age at the arrival times of informative messages for arrival rate  $\lambda = 1$  and varying OWD parameter  $\mu$ .  $I_{\max} = 20$ . (b) Probability density of the age  $f(x)$  at any point in time obtained from (9) for  $\lambda = 1$  and varying OWD parameter  $\mu$ .  $I_{\max} = 20$ .

#### D. Computing $f^\circ(x_0, t_1)$

We can now put together the forward and backward elements. Let  $\hat{f}(\nu, \theta)$  denote the LST of  $f^\circ(x_0, t_1)$ , specifically

$$\hat{f}(\nu, \theta) := \int_0^\infty \int_0^\infty f^\circ(x_0, t_1) e^{-\nu t_1} e^{-\theta x_0} dt_1 dx_0$$

From (10) this becomes

$$\hat{f}(\nu, \theta) = \sum_{(n', n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n', n}^\circ f_{n', n+1}(\theta) \tilde{f}_n(\nu) \quad (44)$$

where  $p_{n', n}^\circ$  is in (11),  $\tilde{f}_n(\nu)$  is the  $n$ th component of (29) and  $f_{n', n+1}(\theta)$  is obtained by setting  $k = n + 1$  in (43).

As all the Laplace-Stieltjes transforms encountered here are rational fractions in  $\theta$  [resp.  $\nu$ ], the distributions associated with them are matrix-exponential and can be computed in closed form given  $\lambda, \mu$ . Specifically, we obtain

$$g^\circ(x_0 | n', n) = \sum_{i=0}^{I_{\max}} \pi^{i, n', n}(x_0) e^{-x_0 d'_i} \quad (45)$$

where  $d'_i$  is given in (35) and  $\pi^{i, n', n}$  is a polynomial in  $x_0$ , the coefficients of which are computed numerically for every  $(\lambda, \mu)$ . Similarly, we obtain

$$h(t_1 | n) = \sum_{j=0}^{I_{\max}} \tilde{\pi}^{j, n}(t_1) e^{-t_1 \tilde{d}_j} \quad (46)$$

where  $\tilde{d}_j$  is given in (25) and  $\tilde{\pi}^{j, n}$  is a polynomial in  $t_1$ , the coefficients of which are computed numerically for every  $(\lambda, \mu)$ . Putting things together we obtain

$$f^\circ(x_0, t_1) = \sum_{i=0}^{I_{\max}} \sum_{j=0}^{I_{\max}} e^{-x_0 d'_i - t_1 \tilde{d}_j} \sum_{(n', n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n', n}^\circ \pi^{i, n', n}(x_0) \tilde{\pi}^{j, n}(t_1) \quad (47)$$

#### V. COMPUTING AGE PERFORMANCE METRICS

##### A. Age Distribution at Arbitrary Points in Time

To obtain the PDF of the age at any point in time we insert the formulation of the PDF (47) into (9). To calculate this expression, we first show the calculation of a generic term that represents the core of this expression. We compute

$$\begin{aligned} & \int_0^x \int_{x-x_0}^{+\infty} x_0^k t_1^\ell e^{-d'x_0 - \tilde{d}t_1} dt_1 dx_0 \\ &= \int_0^x x_0^k e^{-d'x_0} \left[ -e^{-\tilde{d}t_1} \sum_{i=0}^{\ell} \frac{\ell!}{i! \tilde{d}^{\ell-i+1}} t_1^i \right]_{x-x_0}^{\infty} dx_0 \\ &= \int_0^x x_0^k e^{-d'x_0} \left( e^{-\tilde{d}(x-x_0)} \sum_{i=0}^{\ell} \frac{\ell!}{i! \tilde{d}^{\ell-i+1}} (x-x_0)^i \right) dx_0. \end{aligned} \quad (48)$$

The expression in (48) stems from the fact that a primitive of  $e^{-\tilde{d}t_1} P(t_1)$ , with polynomial  $P(t_1) = t_1^\ell$ , is  $-e^{-\tilde{d}t_1} \sum_{i=0}^{\deg(P)} \frac{P^{(i)}(t_1)}{\tilde{d}^{i+1}}$  where  $P^{(i)}$  is the  $i$ th derivative of  $P$  and  $P^{(0)} = P$ . This sum can be written in a compact form as  $\sum_{i=0}^{\ell} \frac{\ell!}{i! \tilde{d}^{\ell-i+1}} t_1^i$ . For the evaluation of the integral we used that  $\lim_{t_1 \rightarrow \infty} t_1^\ell e^{-\tilde{d}t_1} = 0$  for any fixed  $\ell$  and positive  $\tilde{d}$ .

Using the binomial theorem to expand the term  $(x-x_0)^i$  in the expression above we can rewrite (48) as

$$\begin{aligned} & \frac{\ell! e^{-\tilde{d}x}}{\tilde{d}^{\ell+1}} \int_0^x x_0^k e^{-(d'-\tilde{d})x_0} \sum_{i=0}^{\ell} \frac{\tilde{d}^i}{i!} (x-x_0)^i dx_0 \\ &= \frac{\ell! e^{-\tilde{d}x}}{\tilde{d}^{\ell+1}} \int_0^x x_0^k e^{-(d'-\tilde{d})x_0} \sum_{i=0}^{\ell} c_{i, \ell}(x) x_0^i dx_0. \end{aligned} \quad (49)$$

where we expanded  $(x-x_0)^i = \sum_{k=0}^i (-1)^k \binom{i}{k} x^{i-k} x_0^k$ . Then we rearrange the sum terms in the first line to express it as a polynomial in  $x_0$  with coefficients

$$c_{i, \ell}(x) = \sum_{j=i}^{\ell} (-1)^j \binom{j}{i} x^{j-i} \frac{\ell!}{j!} \quad (50)$$



with  $j \geq i$ . Here, we explicitly express the dependency of the coefficients on  $x$  through  $c_{i,\ell}(x)$ .

In a last step to compute (49) we calculate for  $d' \neq \tilde{d}$

$$\begin{aligned} & \frac{\ell! e^{-\tilde{d}x}}{\tilde{d}^{\ell+1}} \sum_{i=0}^{\ell} c_{i,\ell}(x) \int_0^x x_0^{k+i} e^{-(d'-\tilde{d})x_0} dx_0 \\ &= \frac{\ell! e^{-\tilde{d}x}}{\tilde{d}^{\ell+1}} \sum_{i=0}^{\ell} c_{i,\ell}(x) \left[ -e^{-(d'-\tilde{d})x_0} \sum_{j=0}^{k+i} \frac{(k+i)! x_0^j}{j! (d'-\tilde{d})^{k+i-j+1}} \right]_0^x \\ &= \frac{\ell! e^{-\tilde{d}x}}{\tilde{d}^{\ell+1}} \sum_{i=0}^{\ell} c_{i,\ell}(x) \left[ \frac{(k+i)!}{(d'-\tilde{d})^{k+i+1}} \right. \\ & \quad \left. - e^{-(d'-\tilde{d})x} \sum_{j=0}^{k+i} \frac{(k+i)! x^j}{j! (d'-\tilde{d})^{k+i-j+1}} \right]. \end{aligned} \quad (51)$$

For the case when  $d' = \tilde{d}$  we obtain as a solution of (49)

$$\frac{\ell! e^{-\tilde{d}x}}{\tilde{d}^{\ell+1}} \sum_{i=0}^{\ell} c_{i,\ell}(x) \frac{x^{k+i+1}}{k+i+1} \quad (52)$$

Now, using the steps from above we insert the formulation of the PDF (47) into (9) and obtain the PDF of the age at any point in time in the following theorem.

**Theorem V.1.** *In a stationary  $M/M/I_{\max}/I_{\max}^*$  system, the PDF of the age of information at an arbitrary point in time,  $f(x)$ , is given by*

$$\begin{aligned} f(x) &= \hat{\lambda} \sum_{i=0}^{I_{\max}} \sum_{j=0}^{I_{\max}} \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} e^{-x\tilde{d}_j} \times \\ & \sum_{l=0}^{\tilde{\xi}_{j,n} + \xi_{i,n'},n} \tilde{c}_l(x) \left( \frac{l!}{(d'_i - \tilde{d}_j)^{l+1}} - e^{-x(d'_i - \tilde{d}_j)} \sum_{v=0}^l \frac{l!}{v! (d'_i - \tilde{d}_j)^{l-v+1}} x^v \right) \end{aligned} \quad (53)$$

where

- $\lambda$  is the message generation rate at the sender,  $1/\mu$  is the mean message transit time and  $I_{\max}$  is the maximum number of messages in transit;
- $p_{n',n}^{\circ}$  from (11),  $d'_i$  from (35), and  $\tilde{d}_j$  from (25);
- $\xi_{i,n'},n, \tilde{\xi}_{j,n}$  are the degrees of the polynomials  $\pi^{i,n'},n(x_0)$  and  $\tilde{\pi}^{j,n}(t_1)$  from (45) and (46). Specifically, these are given as

$$\pi^{i,n'},n(x_0) := \sum_{k=0}^{\xi_{i,n'},n} a_k^{i,n'},n x_0^k$$

and

$$\tilde{\pi}^{j,n}(t_1) := \sum_{k=0}^{\tilde{\xi}_{j,n}} \tilde{a}_k^{j,n} t_1^k.$$

- $\tilde{c}_l(x)$  are the polynomial coefficients obtained through the convolution

$$\tilde{c}_l(x) = \sum_{v=0}^l \tilde{z}_v^{j,n}(x) a_{l-v}^{i,n'},n \quad (54)$$

with  $\tilde{z}_k^{j,n}(x) = \sum_{i=k}^{\tilde{\xi}_{j,n}} \tilde{a}_i^{j,n} c_{k,i}(x)$ , where  $c_{k,i}(x)$  is given in (50).

*Proof:* By inserting the formulation of the PDF (47) into (9) we obtain the PDF of the age at any point in time as

$$\begin{aligned} f(x) &= \hat{\lambda} \int_0^x \int_{x-x_0}^{\infty} f^{\circ}(x_0, t_1) dt_1 dx_0 \\ &= \hat{\lambda} \sum_{i=0}^{I_{\max}} \sum_{j=0}^{I_{\max}} \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} \times \\ & \int_0^x e^{-x_0 d'_i} \pi^{i,n'},n(x_0) e^{-(x-x_0)\tilde{d}_j} \times \\ & \int_{x-x_0}^{\infty} \tilde{\pi}^{j,n}(t_1) e^{-t_1 \tilde{d}_j} dt_1 dx_0 \\ &= \hat{\lambda} \sum_{i=0}^{I_{\max}} \sum_{j=0}^{I_{\max}} \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} \times \\ & \int_0^x e^{-x_0 d'_i} \sum_{k=0}^{\xi_{i,n'},n} a_k^{i,n'},n x_0^k e^{-(x-x_0)\tilde{d}_j} \times \\ & \int_{x-x_0}^{\infty} \sum_{k=0}^{\tilde{\xi}_{j,n}} \tilde{a}_k^{j,n} t_1^k e^{-t_1 \tilde{d}_j} dt_1 dx_0 \end{aligned} \quad (55)$$

Looking closely at (55) after rearranging terms and swapping the sum in  $\tilde{\pi}^{j,n}(t_1)$  and the integral over  $t_1$  we observe that at the core of the problem we need to compute the expression in (48). The additional complexity compared to the result in (51) arises due to the sums in  $\tilde{\pi}^{j,n}(t_1)$  and  $\pi^{i,n'},n(x_0)$ . In the next step, we evaluate the second integral in (55) using the same method as for (48) to obtain

$$\begin{aligned} & \hat{\lambda} \sum_{i=0}^{I_{\max}} \sum_{j=0}^{I_{\max}} \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} \times \\ & \int_0^x e^{-x_0 d'_i} \sum_{k=0}^{\xi_{i,n'},n} a_k^{i,n'},n x_0^k \sum_{k=0}^{\tilde{\xi}_{j,n}} \tilde{a}_k^{j,n} \sum_{v=0}^k \frac{k!}{v! \tilde{d}_j^{k-v+1}} (x-x_0)^v dx_0 \\ &= \hat{\lambda} \sum_{i=0}^{I_{\max}} \sum_{j=0}^{I_{\max}} \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} \times \\ & \int_0^x e^{-x_0 d'_i} \sum_{k=0}^{\xi_{i,n'},n} a_k^{i,n'},n x_0^k \sum_{k=0}^{\tilde{\xi}_{j,n}} \tilde{z}_k^{j,n}(x) x_0^k dx_0. \end{aligned} \quad (56)$$

Here, in the second line we expanded  $(x-x_0)^v$  in the same way as in (49). The difference to (49) stems from the additional sum leading to the intermediate form  $\sum_{k=0}^{\tilde{\xi}_{j,n}} \tilde{a}_k^{j,n} \sum_{i=0}^k c_{i,k}(x) x_0^i$  after using the expansion in (49). Now, after collecting the terms we can rewrite this sum as  $\sum_{k=0}^{\tilde{\xi}_{j,n}} \tilde{z}_k^{j,n}(x) x_0^k$  with  $\tilde{z}_k^{j,n}(x) = \sum_{i=k}^{\tilde{\xi}_{j,n}} \tilde{a}_i^{j,n} c_{k,i}(x)$ , where  $c_{k,i}(x)$  is given in (50).

Inspecting (56), we see that the product of the two given polynomials in  $x_0$  can be rewritten as one polynomial  $\sum_{l=0}^{\tilde{\xi}_{j,n} + \xi_{i,n'},n} \tilde{c}_l(x) x_0^l$  with  $\tilde{c}_l(x) = \sum_{v=0}^l \tilde{z}_v^{j,n}(x) a_{l-v}^{i,n'},n$ , i.e., the convolution of the coefficients of the two polynomials. Equipped with the integral evaluation in (51) we can now

compute

$$\begin{aligned}
& \hat{\lambda} \sum_{i=0}^{I_{\max}} \sum_{j=0}^{I_{\max}} \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} \times \\
& \int_0^x e^{-x_0 d'_i} \sum_{l=0}^{\xi_{i,n',n} + \tilde{\xi}_{j,n}} \tilde{c}_l(x) x_0^l dx_0 \\
& = \hat{\lambda} \sum_{i=0}^{I_{\max}} \sum_{j=0}^{I_{\max}} \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} e^{-x \tilde{d}_j} \times \\
& \sum_{l=0}^{\tilde{\xi}_{j,n} + \xi_{i,n',n}} \tilde{c}_l(x) \left( \frac{l!}{(d'_i - \tilde{d}_j)^{l+1}} - e^{-x(d'_i - \tilde{d}_j)} \sum_{v=0}^l \frac{l!}{v!(d'_i - \tilde{d}_j)^{l-v+1}} x^v \right)
\end{aligned} \tag{57}$$

### B. Age Distribution at the Arrival of Informative Messages

In addition to calculating the age distribution at any point in time we can easily calculate the distribution of the age at the *arrival instants of informative messages*. The corresponding PDF  $f_A^{\circ}(x_0)$  is readily obtained as

$$f_A^{\circ}(x_0) = \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} g^{\circ}(x_0 | n', n)$$

Using (45) we obtain

$$f_A^{\circ}(x_0) = \sum_{i=0}^{I_{\max}} e^{-x_0 d'_i} \sum_{(n',n) \text{ s.t. } 1 \leq n+1 \leq n' \leq I_{\max}} p_{n',n}^{\circ} \pi^{i,n',n}(x_0) \tag{58}$$

## VI. EVALUATION

In this section, we compare the obtained expressions for the age distributions to results from empirical discrete event simulations. We consider the system as described in Sect. III with messages arriving as a Poisson process with rate  $\lambda$  where each message observes a service time sampled from an exponential distribution with parameter  $\mu$ . The system simulation results are obtained from simulation runs over  $10^5$  messages and we set the number of non-obsolete messages under way to  $I_{\max} = 20$ .

Figure 6a shows the age distribution at the arrival time points of informative messages. The dashed curve is obtained from the model (8) (with test function  $\varphi$  being the identity function). The figure also shows the empirical age distribution obtained from simulations. We observe that these two distributions match very well and the impact of the average service rate of one message  $\mu$  on the tail of the age distribution. Fig. 6b shows the probability density of the age at any point in time that is obtained from (9) using the Laplace inverse of (44). Observe the skewness of the density function. This shows that approximations based on the first few (two) moments, e.g. obtained based on work that calculate the moments of the age distribution [3], will be inaccurate.

Figure 7a shows a comparison of the expected age at any point in time as a function of the message arrival rate  $\lambda$ . The expected age that is obtained from the model is computed using the density in (9) in closed form.

To empirically obtain the average age from the event based simulation we utilize the Palm inversion formula (4) with  $\varphi$  set as identity function. Hence we can write

$$\begin{aligned}
\mathbb{E}[X_t] &= \mu \bar{N} \mathbb{E}^{\circ} \left[ \int_{T_0}^{T_1} X_s ds \right] = \mu \bar{N} \mathbb{E}^{\circ} \left[ \int_0^{T_1} (A_0 + s) ds \right] \\
&= \mu \bar{N} \mathbb{E}^{\circ} \left[ A_0(T_1 - T_0) + \frac{1}{2}(T_1 - T_0)^2 \right]
\end{aligned} \tag{59}$$

where  $A_0$  is the age of the informative message received at time  $T_0$ . The estimate of  $\mathbb{E}[X_t]$  obtained from one simulation run is

$$\mu \hat{N} \sum_{n=1}^{n_{\text{tot}}-1} \left( A_n(T_{n+1} - T_n) + \frac{1}{2}(T_{n+1} - T_n)^2 \right) \tag{60}$$

where  $A_n$  [resp.  $T_n$ ] is the age upon delivery at the receiver [resp. delivery time] of the  $n$ th informative message,  $n_{\text{tot}}$  is the total number of informative messages delivered in the simulation run, and  $\hat{N}$  is the time-average number of messages in the channel.

Here too, the comparison with the empirically obtained average age from (59) shows a close match. Note that the empirically obtained average age still requires invoking the Palm inversion formula (4) as given in (59). We observe in Fig. 7a that the expected age decreases monotonically with the message arrival rate  $\lambda$ . This stand in line with similar age models with finite message capacity assuming, however, FIFO message delivery, such as in [18]. In Fig. 7b we show the quantiles of the age at any point in time based on the age probability density in (9). These quantiles can be utilized for system dimensioning by providing operating points, in terms of setting the service rate  $\mu$  or throttling the message generation rate  $\lambda$ , to retain a corresponding age quantile  $x_{\varepsilon}$  that is only violated with probability  $P[X > x_{\varepsilon}] = \varepsilon$ .

## VII. RELATED WORK

The problem of status updating to combat data staleness in distributed systems that use a shared and unreliable network was first discussed in the context of real-time database systems in [19]. Essentially, a recent reincarnation of this problem in the context of IoT that is known as Age of information (AoI) considers transmission scheduling strategies to update the status at some receiver in a way that optimizes the freshness of that information [2, 20, 21]. This problem has been in particular of interest in the context of vehicular networks [22, 23] and sensory information transmission over wireless networks [24, 25] as the freshness of information such as the captured environment model that is exchanged between vehicles is safety critical. For a comprehensive review see [10].

Given a single sender and a system modeled as an M/M/1 queue the work in [2] derives the *sample path average age* at the receiver as  $\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \Delta(t) dt = \lambda (E[XS] + E[X^2]/2)$  with  $\Delta(t)$  denoting the age of information at the receiver at time  $t$ ,  $\lambda$  being the message arrival

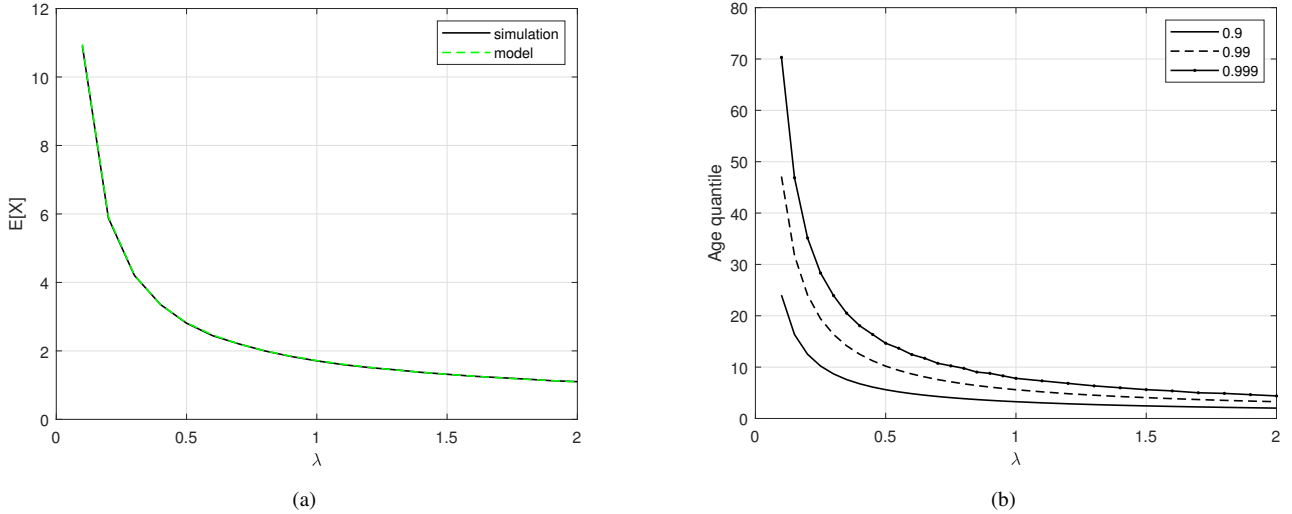


Fig. 7: (a) Expected age  $E[X]$  obtained from simulations compared to the model in (8) with  $\varphi$  being the identity function for  $I_{\max} = 20$  and  $\mu = 1$ . (b) Quantiles of the age  $P[X > x_\varepsilon] = \varepsilon$  obtained from integrating the age density in (9) for  $I_{\max} = 20$  and  $\mu = 1$ .

rate and the random variables  $X, S$  that denote the message inter-arrival time and message response time, respectively. The same seminal work provides expressions for the sample path average age in a D/M/1 system using a transcendental function. Given the forms derived above the work [2] also provides the parametrization that minimizes the average age.

Similarly, works such as [6, 7] consider the age at the receiver given a D/G/1 and M/M/ $\infty$  systems where messages may arrive out of order. The work in [6] considers the distribution of the age process for a deterministic transmission schedule and a single server with general service time distribution under the FIFO assumption. The authors derive in [7] an expression for the *average age* using a similar reasoning as [2] as a function of the average message arrival rate and service time distribution. Note that the provided expression there is not directly computable as it contains multiple infinite sums and infinite products. Applications of the methods above to the special case of G/G/1/1 queueing systems under message blocking and message preemption is found in [26]. In contrast to this work, however, our approach here provides the distribution of the age using a computable closed form that is constructed as the combination of Laplace-Stieltjes transforms of elementary functions

Going beyond elementary queues, the work in [27] considers the AoI for a path consisting of a concatenation of multiple links where the random delay at each of the links is only due random access. The authors show that given this delay model and a graph model of the network the problem of finding a transmission strategy for minimizing the AoI at  $N$  sender-receiver pairs can be decomposed into a simpler equivalent optimization problem. We believe that the main reason for this lies in the delay due to random access model that does not incorporate queueing and scheduling effects. A similar work considering multihop networks, i.e., [4], that considers, however, a multihop queueing network, shows that a preemptive Last Generated First Served (LGFS) policy at all nodes minimizes any non-decreasing functional of the age

in stochastic dominance sense. This result is obtained under the assumption that all message transmission times are iid exponentially distributed at all nodes.

The works in [3, 9] provide a method to calculate the MGF and the moments of the AoI at a network monitor for networks of *preemptive finite buffer servers* based on the so called stochastic hybrid system (SHS) framework [28] by leveraging a notion of a hybrid state  $\{q(t)\}_{t \geq 0}, \mathbf{x}(t)\}$  where  $\{q(t)\}_{t \geq 0}$  describes a continuous time Markov chain over finite state and  $\mathbf{x}(t) \geq 0 \in \mathbb{R}^{1 \times n}$  describes a vector of positive real values of the age. By attaching deterministic matrices  $\mathbf{A}_l \in \{0, 1\}^{n \times n}$  to the transitions  $l \in L$  of the Markov chain  $\{q(t)\}_{t \geq 0}$ , where  $L$  denotes the set transitions, one is able to track the jumps of the age vector as  $\mathbf{x}' = \mathbf{x}\mathbf{A}_l$ . The key to finding a formulation for the expected age and for the Moment generating function (MGF) of the age is based on a set of first order differential equations that assume  $\{q(t)\}_{t \geq 0}$  is ergodic and utilize its steady state stationary probability distribution [3, 9]. Note that there exists a direct relation between the presented SHS framework and utilizing the Master equation  $\frac{d}{dt} E[\varphi(q_t, X_t)] = E[G\varphi(q_t, X_t)]$  with  $\varphi$  being a test function,  $G$  a generator and  $q_t$  and  $X_t$  describing the queue state, as well as, the age at time  $t$  respectively. This relation is explored in [3, 9] to simplify the SHS formulation. We note that, in general, applying the master equation to the infinitesimal generator that describes the jump and drift evolution of the age results in the Fokker-Plank equation describing the time evolution of the age density. Analytical closed-form results to solve this formulation, even for the stationary age distribution, are yet to be shown. Computable solutions to the presented SHS system of equations in are provided for the examples of a single M/M/1/1 queue, a line network of M/M/1/1 queues in [3]. Note that the SHS framework was applied in [11] to obtain a close form for the expected age for a system of parallel servers where a new message arrival preempts the oldest message under way.

Concerning message reordering, a seminal work on packet

reordering is [29] that provides a recursive expression for the total delay distribution of a system in which messages that arrive in order are delivered through a disordering system, hence the out-of-order arrivals require resequencing. The author provides an analytical solution for the case of an  $M/G/\infty$  disordering system given in terms of a Laplace-Stieltjes transform.

A min-plus approach to the Age of Information is given in [8] showing that the virtual delay at a FIFO system, i.e., the horizontal deviation of cumulative arrivals and departures at a min-plus system, is an upper bound on the age. Equipped with lower bounds on the cumulative arrival traffic, the work in [8] shows deterministic and statistical upper bound on the age of information for different combinations of arrivals and systems with deterministic or probabilistic description. In contrast to [8], we consider here non-FIFO systems with possible message reordering.

### VIII. CONCLUSION

In this paper, we considered the problem of computing the distribution of the Age of Information (AoI) at any point in time for non-preemptive, non-FIFO systems. Our key observation is that this networked system can be modeled as a batch queueing system where the served batch size is random and the sojourn time of the freshest message in the batch corresponds to the age of an arriving informative message. The batch (except for the its freshest message) essentially models the set of messages that are generated before the freshest message and hence are rendered obsolete by its arrival. This captures message reordering due to the non-FIFO system property.

Equipped with this queueing model we use Palm calculus together with time inversion to decompose the elements that form the joint distribution of the age and the time between the arrival of informative messages at the receiver. Then, Palm inversion allows us to compute the distribution of the age at any point in time given this joint distribution. We find recursions for the corresponding Laplace-Stieltjes transforms of the conditional age and informative message inter-arrival time distributions owing to the Markovian nature of the underlying model. As these transforms turn to be rational we obtain a computable expression for the AoI distribution composed of matrix-exponential terms. This main result allows further to deduce formulations for the expected age, as well as, the age distribution at the arrival time points of informative messages. We validate the exact model using discrete-event simulations and show the skewness of the PDFs of the age. Further, we show the impact of the arrival and service rates on the age CDF and its quantiles. We leave the extension of this model to multi-stage queueing networks to future work.

### IX. APPENDIX

#### A. Calculation of the steady state probabilities in (3)

We consider the queueing model from Sect. III-A with the corresponding Markov chain as sketched in Fig. 3. In the following we prove the formulation of the steady state probabilities  $p_n$  of the given Markov chain through induction. We first show that the formulation holds for  $p_0$  and  $p_1$ . Then

we show that given that the formulation holds for all  $p_k$  for  $k < n$  it also holds for  $p_n$ .

From the balance equations we can write  $p_0\lambda = \sum_{i=1}^{I_{\max}} p_i\mu$ . From the normalization condition  $\sum_{i=0}^{I_{\max}} p_i = 1$  we obtain  $p = \frac{\mu}{\lambda+\mu}$  as  $\sum_{i=1}^{I_{\max}} p_i = 1 - p_0$ . For  $p_1$  we can write  $p_1(\lambda + \mu) = p_0\lambda + \sum_{i=2}^{I_{\max}} p_i\mu$ . Using  $p_0$  and the normalization condition this reduces to  $p_1\lambda = p_0\lambda + \mu(1 - p_0 - p_1) - p_1\mu$  which leads to  $p_1 = \frac{2\lambda\mu}{(\lambda+\mu)(\lambda+2\mu)}$ .

Now, considering (3) we directly see that it holds for  $n = 0$  and  $n = 1$ . For a state  $k$  of the given Markov chain we can write using the balance equations

$$\begin{aligned} p_k(\lambda + k\mu) &= p_{k-1}\lambda + \mu \sum_{i=k+1}^{I_{\max}} p_i \\ &= p_{k-1}\lambda + \mu(1 - \sum_{i=0}^k p_i), \end{aligned} \quad (61)$$

which we can rewrite as

$$\begin{aligned} p_k(\lambda + (k+1)\mu) &= p_{k-1}(\lambda - \mu) + \mu - \mu \sum_{i=0}^{k-2} p_i \\ &= \frac{k\lambda^{k-1}\mu(\lambda - \mu)}{\prod_{j=1}^k (\lambda + j\mu)} + \mu - \mu \sum_{i=0}^{k-2} \frac{(i+1)\lambda^i\mu}{\prod_{j=1}^{i+1} (\lambda + j\mu)} \\ &= \frac{k\lambda^{k-1}\mu(\lambda - \mu)\Gamma\left(\frac{\lambda+\mu}{\mu}\right)}{\mu^k\Gamma\left(\frac{\lambda+(k+1)\mu}{\mu}\right)} + \mu \\ &\quad - \mu\Gamma\left(\frac{\lambda+\mu}{\mu}\right) \left( \frac{\lambda+\mu}{\mu\Gamma\left(\frac{\lambda+2\mu}{\mu}\right)} - \frac{\lambda^{k-1}(k\mu+\lambda)}{\mu^k\Gamma\left(\frac{\lambda+(k+1)\mu}{\mu}\right)} \right), \end{aligned} \quad (62)$$

where we used the identity  $\prod_{j=1}^k (\lambda + j\mu) = \frac{\mu^k\Gamma\left(\frac{\lambda+(k+1)\mu}{\mu}\right)}{\Gamma\left(\frac{\lambda+\mu}{\mu}\right)}$ . Now, we can further simplify (62) using an instance of this the identity  $\Gamma\left(\frac{\lambda+2\mu}{\mu}\right) = \frac{\lambda+\mu}{\mu}\Gamma\left(\frac{\lambda+\mu}{\mu}\right)$ . Finally through rearranging terms we obtain

$$p_k = \frac{(k+1)\lambda^k\mu\Gamma\left(\frac{\lambda+\mu}{\mu}\right)}{\mu^k\Gamma\left(\frac{\lambda+(k+1)\mu}{\mu}\right)}, \quad (63)$$

which completes the proof. Calculating  $p_k$  for  $k = I_{\max}$  follows along using the balance equation as shown above.

#### B. Calculation of the transition rates of the reversed process (30), (31)

The transition rates for the reversed Markov process are obtained directly from [Theorem 1.12] from [17] as  $Q'(i, j) = \frac{p_j Q(j, i)}{p_i}$ ,  $i, j \in E$  with the same steady state distribution  $p_n$ ,  $n \in E$ . Now given the forward Markov process with transition rates in (2) (as sketched in Fig. 3) we obtain the following transition rates for the reverse process  $Q'_{i, i-1} := \lambda'_i$  for  $i = 1 \dots I_{\max}$ ,  $Q'_{i, j} := \mu'_{ij}$  for  $i = 0 \dots I_{\max} - 1, i < j \leq I_{\max}$  and  $Q'_{i, j} := 0$

otherwise. Hence, we obtain from (2) and (3) for  $i < I_{\max}$

$$\begin{aligned}\lambda'_i &= \lambda \frac{p_{i-1}}{p_i} = \lambda \frac{i\lambda^{i-1}\mu}{\prod_{j=1}^i (\lambda + j\mu)} \frac{\prod_{j=1}^{i+1} (\lambda + j\mu)}{(i+1)\lambda^i\mu} \\ &= \frac{i}{i+1} (\lambda + (i+1)\mu).\end{aligned}\quad (64)$$

For  $i = I_{\max}$  the derivation goes accordingly to find  $\lambda'_i = I_{\max}\mu$ . Similarly, for  $i < j < I_{\max}$  we obtain

$$\begin{aligned}\mu'_{ij} &= \mu \frac{p_j}{p_i} = \mu \frac{(j+1)\lambda^j\mu}{\prod_{k=1}^{j+1} (\lambda + k\mu)} \frac{\prod_{k=1}^{i+1} (\lambda + k\mu)}{(i+1)\lambda^i\mu} \\ &= \frac{(j+1)\mu\lambda^{j-i}}{(i+1) \prod_{k=i+2}^{j+1} (\lambda + k\mu)}.\end{aligned}\quad (65)$$

Again, for  $j = I_{\max}$  the derivation goes similarly.

### C. On the numerical calculation of the LSTs in (44)

For completeness, we show in the following an alternative method to the direct calculation of the conditional LST in (44) that we used in this paper. We underline that the following numerical computation may be beneficial in speeding up computations especially for evaluation purposes.

The direct computation of (44) by computing the conditional LST vectors (43) through recursion and matrix inversions, as well as, the LST vector in (29) and the following insertion of the vector components  $f_{n,k}$  and  $\tilde{f}_n$  into (44) becomes computationally intensive when  $I_{\max}$  is large. The reason for this is the computation of the matrix inverse  $\Phi^{-1}$  in (42) as well its exponentiation in form of  $\Psi^{k-1}$  in (43). Next we discuss an alternative numerical method to compute the quantities in (42) - (43).

First, we recognize that  $\Phi^{-1} = (\theta\mathbf{I} + \mathbf{D} - \mathbf{M})^{-1}$  used in (42) is a fraction by the adjugate matrix formula, i.e.,

$$\Phi^{-1} = \frac{\mathbf{R}(\theta)}{\rho(\theta)}, \quad (66)$$

where  $\mathbf{R}(\theta)$  is a polynomial with matrix coefficients given by

$$\mathbf{R}(\theta) = \sum_{k=0}^{I_{\max}} \theta^k \mathbf{P}_k, \quad (67)$$

and the denominator is the characteristic polynomial of  $\mathbf{M} - \mathbf{D}$ , i.e.

$$\rho(\theta) := \sum_{k=0}^{I_{\max}+1} r_k \theta^k = \det(\Phi) \quad (68)$$

Also observe that  $r_0 = \det(\mathbf{D} - \mathbf{M})$  and that  $r_{I_{\max}+1} = 1$ . Note that the matrices  $\mathbf{P}_k$  are of dimensions  $(I_{\max} + 1) \times (I_{\max} + 1)$  large. Next, we obtain  $\mathbf{P}_k$  iteratively using LeVerrier's method. In a nutshell, we plug (66) into the identity  $\Phi^{-1}\Phi = \mathbf{I}$  and rearrange the terms to obtain

$$\sum_{k=1}^{I_{\max}+1} \theta^k \mathbf{P}_{k-1} + \sum_{k=0}^{I_{\max}} \theta^k \mathbf{P}_k (\mathbf{D} - \mathbf{M}) = \left( \sum_{k=0}^{I_{\max}+1} r_k \theta^k \right) \mathbf{I}. \quad (69)$$

Now, we can compare the coefficients of  $\theta^k$  on both sides of (69) and obtain the recursive form for the matrices  $\mathbf{P}_k$  as

$$\mathbf{P}_k = r_{k+1} \mathbf{I} - \mathbf{P}_{k+1} (\mathbf{D} - \mathbf{M}), \quad (70)$$

for  $k \in \{I_{\max}-1, \dots, 1\}$ . From the comparison of the coefficients in (69) we know that  $\mathbf{P}_0 = (\det(\mathbf{D} - \mathbf{M})) (\mathbf{D} - \mathbf{M})^{-1}$  and  $\mathbf{P}_{I_{\max}} = \mathbf{I}$  such that we can iteratively find the matrices  $\mathbf{P}_k$  using (70), hence, calculate the coefficients of  $\mathbf{R}(\theta)$ .

Now, given that we calculate  $\Phi^{-1}$  using the method above we can use this result to simplify the matrix multiplication in  $\Psi^k$  as  $\Psi = \Phi^{-1}\Lambda$ . Hence, we can write

$$\Psi^k = \frac{\tilde{\mathbf{R}}(\theta)^k}{\rho(\theta)^k}, \quad (71)$$

where we used the polynomial  $\tilde{\mathbf{R}}(\theta)$  that is defined as

$$\tilde{\mathbf{R}}(\theta) = \sum_{k=0}^{I_{\max}} \theta^k \tilde{\mathbf{P}}_k, \quad (72)$$

with  $\tilde{\mathbf{P}}_k = \mathbf{P}_k \Lambda$ . Now calculating the denominator of (71) is simple as  $\Phi$  is triangular and its determinant is obtained in closed form as

$$\det(\Phi) = \prod_{i=0}^{I_{\max}} \left( \theta + \sum_j Q'_{i,j} \right). \quad (73)$$

As  $\tilde{\mathbf{R}}(\theta)^k$  is a product of polynomials with matrix coefficients, we calculate the numerator in (71) using an iterative convolution operation of the coefficients  $\tilde{\mathbf{P}}_k$ . Now, given the calculation method above we can numerically obtain  $\Psi^k$  for insertion in (43). Note that the same procedure can be used to obtain the elements  $\tilde{f}_n(\nu)$  in (44) by numerically calculating the inversion in (29).

Finally, calculating the inverse Laplace transform of the Palm joint density  $f^\circ(t_1, x_0)$  entails taking the inverse Laplace transform of the right hand side (RHS) of (44). Given the factorization of  $\rho(\theta)$  and the matrix coefficient form of the polynomial  $\mathbf{R}(\theta) = \sum_{k=0}^{I_{\max}} \theta^k \mathbf{P}_k$  we observe that  $f_{n',n+1}(\theta)$  on the RHS of (44) has the form  $\sum_i \frac{\alpha_i}{(\theta + d'_i)^{\kappa_i}}$  with constants  $\alpha_i$  and  $\kappa_i \leq n$  due to the partial fraction decomposition of (71). To obtain the Palm joint density  $f^\circ(t_1, x_0)$  we calculate the inverse Laplace-Stieltjes transform of the RHS of (44). Given the observation that  $f_{n',n+1}(\theta)$  can be rewritten as  $\sum_i \frac{\alpha_i}{(\theta + d'_i)^{\kappa_i}}$  we know that the inverse LST of  $f_{n',n+1}(\theta)$  has the form  $\sum_i c_i x_0^{\kappa_i} e^{-d'_i x_0}$  with constants  $c_i$  and  $\kappa_i \leq n$ . The same observation holds for  $\tilde{f}_n(\nu)$  in (44), i.e., by calculating the inversion of (29) using the method above we finally obtain a partial fraction decomposition and subsequent inverse LST that has the form  $\sum_j h_j t_1^{\zeta_j} e^{-\tilde{d}_j t_1}$ .

## REFERENCES

- [1] E. A. Lee, "Cyber physical systems: Design challenges," in *11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC)*, 2008, pp. 363–369.
- [2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proceedings IEEE INFOCOM*, 2012, pp. 2731–2735.
- [3] R. D. Yates, "The age of information in networks: Moments, distributions, and sampling," *IEEE Transactions on Information Theory*, vol. 66, no. 9, pp. 5712–5728, 2020.

- [4] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Age-optimal information updates in multihop networks," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 576–580.
- [5] P. Popovski, C. Stefanović, J. J. Nielsen, E. de Carvalho, M. Angelichinoski, K. F. Trillingsgaard, and A.-S. Bana, "Wireless access in ultra-reliable low-latency communication (urllc)," *IEEE Transactions on Communications*, vol. 67, no. 8, pp. 5783–5801, 2019.
- [6] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *2013 IEEE International Symposium on Information Theory*, 2013, pp. 66–70.
- [7] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, "Effect of message transmission path diversity on status age," *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1360–1374, 2016.
- [8] M. Noroozi and M. Fidler, "A min-plus model of age-of-information with worst-case and statistical bounds," *CoRR*, vol. abs/2112.11934, 2021.
- [9] 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, 2019.
- [10] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. H. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, 2021.
- [11] R. D. Yates, "Status updates through networks of parallel servers," in *2018 IEEE International Symposium on Information Theory, ISIT 2018, Vail, CO, USA, June 17-22, 2018*. IEEE, 2018, pp. 2281–2285.
- [12] F. Baccelli and P. Brémaud, *Palm probabilities and stationary queues*. Springer Science & Business Media, 2012, vol. 41.
- [13] R. Serfozo, *Basics of applied stochastic processes*. Springer Science & Business Media, 2009.
- [14] J.-Y. Le Boudec, *Performance evaluation of computer and communication systems*. Epfl Press Lausanne, 2010, vol. 2.
- [15] J. Le Boudec, *Performance Evaluation of Computer and Communication Systems*, ser. Computer and communication sciences. CRC Press, 2011.
- [16] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *Journal of computational physics*, vol. 22, no. 4, pp. 403–434, 1976.
- [17] F. P. Kelly, *Reversibility and Stochastic Networks*. USA: Cambridge University Press, 2011.
- [18] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, "Effect of message transmission path diversity on status age," *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1360–1374, 2016.
- [19] X. Song and J. Liu, "Performance of multiversion concurrency control algorithms in maintaining temporal consistency," in *Proceedings., Fourteenth Annual International Computer Software and Applications Conference*, 1990, pp. 132–139.
- [20] Y. Sun, E. Uysal-Biyikoglu, and S. Kompella, "Age-optimal updates of multiple information flows," in *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2018, pp. 136–141.
- [21] S. Ioannidis, A. Chaintreau, and L. Massoulie, "Optimal and scalable distribution of content updates over a mobile social network," in *IEEE INFOCOM 2009*, 2009, pp. 1422–1430.
- [22] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 2011, pp. 350–358.
- [23] S. Kaul, R. Yates, and M. Gruteser, "On piggybacking in vehicular networks," in *IEEE Global Telecommunications Conference - GLOBECOM*, 2011, pp. 1–5.
- [24] Q. He, G. Dan, and V. Fodor, "Minimizing age of correlated information for wireless camera networks," in *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2018, pp. 547–552.
- [25] J. Hribar, M. Costa, N. Kaminski, and L. A. DaSilva, "Updating strategies in the internet of things by taking advantage of correlated sources," in *IEEE Global Communications Conference*, 2017, pp. 1–6.
- [26] A. Soysal and S. Ulukus, "Age of information in g/g/1/1 systems: Age expressions, bounds, special cases, and optimization," *IEEE Transactions on Information Theory*, vol. 67, no. 11, pp. 7477–7489, 2021.
- [27] R. Talak, S. Karaman, and E. H. Modiano, "Minimizing age-of-information in multi-hop wireless networks," *55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 486–493, 2017.
- [28] J. Hu, J. Lygeros, and S. Sastry, "Towards a theory of stochastic hybrid systems," in *Hybrid Systems: Computation and Control*. Springer Berlin Heidelberg, 2000, pp. 160–173.
- [29] F. Baccelli, E. Gelenbe, and B. Plateau, "An end-to-end approach to the resequencing problem," *J. ACM*, vol. 31, no. 3, pp. 474–485, 1984.