Age of Information Process Under Strongly Mixing Communication – Moment Bound, Mixing Rate and Strong Law

Adrian Redder, Arunselvan Ramaswamy and Holger Karl

October 3, 2022

Age of Information Process under Strongly Mixing Communication – Moment Bound, Mixing Rate and Strong Law

Adrian Redder, Arunselvan Ramaswamy and Holger Karl

***Abstract*—The decentralized nature of multi-agent systems requires continuous data exchange to achieve global objectives. In such scenarios, Age of Information (AoI) has become an important metric of the freshness of exchanged data due to the error-proneness and delays of communication systems. Communication systems usually possess dependencies: the process describing the success or failure of communication is highly correlated when these attempts are "close" in some domain (e.g. in time, frequency, space or code as in wireless communication) and is, in general, non-stationary. To study AoI in such scenarios, we consider an abstract event-based AoI process $\Delta(n)$, expressing time since the last update: If, at time n, a monitoring node receives a status update from a source node (event $A(n-1)$ occurs), then $\Delta(n)$ is reset to one; otherwise, $\Delta(n)$ grows linearly in time. This AoI process can thus be viewed as a special random walk with resets. The event process $A(n)$ may be non-stationary and we merely assume that its temporal dependencies decay sufficiently, described by α-mixing. We calculate moment bounds for the resulting AoI process as a function of the mixing rate of $A(n)$. Furthermore, we prove that the AoI process $\Delta(n)$ is itself α-mixing from which we conclude a strong law of large numbers for $\Delta(n)$. These results are new, since AoI processes have not been studied so far in this general strongly mixing setting. This opens up future work on renewal processes with non-independent interarrival times.**

***Index Terms*—Age of Information, Temporal Communication Dependencies, Moment Bounds, Strong LLN, α-Mixing**

I. Introduction

Age of Information (AoI) is a metric that measures the freshness of status updates from a source available at a monitor. AoI arises if information traverses via a network that requires a certain transmission, service or processing time until the information arrives from a source at a monitor. AoI matters in problems such as distributed optimization, learning and monitoring, where small AoI usually lets algorithms that use the aged data converge faster.

Many effects give rise to AoI. To encompass them, we consider an abstract model for the information exchange between two nodes, a source and a monitor. Information exchange is described by a sequence of events $A(n)$ that represent successful fresh status updates from the source at the monitor (Figure 1).

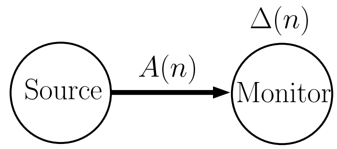

Fig. 1: A source sends status updates through a channel to a monitor. At time $n+1$ an update is successfully received and the AoI process $\Delta(n)$ is reset to one if the event $A(n)$ has occurred.

We use a discrete time model and denote time steps by $n \geq 0$. We refer to a time slot n as the time interval from time step n to $n+1$. The source node sends fresh status updates to the monitor. A fresh update sent at time step n is received either at time step $n+1$ (more precisely, at the start of time slot $n+1$) or not at all. Here, *fresh* means that the update from the source is from time step n. If a fresh update sent at time step n is received at time step $n+1$, then we say the event $A(n)$ has occurred. The event $A(n)$ is thus associated with the n-th time slot. Hence, whenever an event $A(n)$ occurs $\Delta(n+1) = 1$. The resulting AoI process $\Delta(n)$ is defined as

$$\Delta(n+1) := \begin{cases} 1, & A(n) \text{ has occurred}, \\ \Delta(n)+1, & \text{otherwise}, \end{cases} \tag{1}$$

with $\Delta(0) := 1$. This simple AoI model thus considers status updates that require one time-slot for communication. However, it also serves as a basis for more general AoI models where communication requires a certain (possibly stochastic) transmission time, as to be discussed later.

The AoI process (1) can be seen as a special random walk on the positive real line that restarts whenever an event $A(n)$ occurs. From this perspective, the AoI process can represent various hill-climbing phenomena, where a process growths over time and then resets at certain events. This has for example been used in physics to study Doppler laser cooling, where atoms raise from a ground level in the presence of a light field [1]. In the context of wireless communication, $A(n)$ may represent the joint event that information is sent and that the used communication channel is in a good state, e.g., because the signal-to-noise ratio of a received signal is above a certain threshold, thus guaranteeing successful communication [2].

In literature, communication channels have been studied under various time-correlation models. Most common are channel representations as linear-time invariant systems [3] or finite-state Markov chains [4]. However, the abstract AoI process (1) has only been studied under the assumption that the event process $A(n)$ is an independent stochastic process [1]. Independence assumptions are often particularly unrealistic for communication systems. For example, success of sequential communication attempts via wireless fading channels can be highly correlated, but the correlation usually decays over time. In general, many real-world systems possess potentially “long range” dependencies that decay as the events get more separated in time [5].

This work studies the abstract AoI process (1) in the fundamental setting where the event process $A(n)$ merely possesses a temporal dependency decay property, where the process’ dependency decays over time in a way that can be described by α-mixing. Due to this abstract view, our results can be applied to AoI processes that arise in communication, computing or as the consequence of other phenomena that delay the transport of information. Our results are fundamental in that we relate for the first time the rate of mixing of the event process $A(n)$ to properties of the AoI process $\Delta(n)$. We calculate moment bounds for the AoI process $\Delta(n)$ as a function of the α-mixing rate of $A(n)$. We show that $\Delta(n)$ is itself α-mixing, which leads to a strong law of large numbers (SLLN) for $\Delta(n)$ under sufficiently rapid α-mixing of the event process $A(n)$.

A. Prior and Related Work

Age of Information has repeatedly risen in popularity over time. Historically, processes related to aging have been studied to understand the age distribution of populations as individuals live and die [6, 7]. Most notably, age processes were studied as part of Blackwall’s renewal theory, see e.g. [8, Chapter V]; here renewals occur after i.i.d. interarrival times with a certain fixed distribution. In Markov renewal theory, see e.g. [9], the sequence of interarrivals constitute a Markov chain. Renewal theory with dependent interarrivals has recently been considered in [10] from an abstract symbolic dynamics perspective. Notably, renewal theory with temporally dependent interarrivals, where the dependency is described by a probabilistic mixing notion has not been considered so far.

Freshness of information has first been considered in [11]. More recently, AoI has become popular in the networking and information theory community as a representative metric for the freshness of data [12, 13]. Here, AoI has been studied for various queuing and scheduling models [14, 15] that require i.i.d. service and waiting times as in renewal theory. The analysis presented therein relies on the saw-tooth nature of AoI processes in combination with certain stationarity and ergodicity assumptions for the waiting and service time processes. In wireless communication settings, AoI has been considered in edge-based network models where success of communication via individual edges has been considered as i.i.d. across time [16]. In the above cases, we observed that i.i.d. assumptions are prominent for arrival and service times as well as for the success of communication. In the former case, this is due to the tractability of sums of independent random variables using laws of large numbers. In the latter case, it is due to the tractability of products of independent random variables. However, interarrival times can be highly dependent as arrivals tend to form clumps [17]. Similarly, communication success over wireless fading channels can be highly correlated [18].

The previous two paragraphs motivate our consideration of the fundamental AoI process (1) as a function of an event process $A(n)$ with dependencies described by suitably probabilistic mixing notion. Close to our work is the perspective of the AoI process (1) as a random walk with restarts, as studied in [1] and followup papers. Here, the process is called Sisyphus random walk due to its analogy with climbing and falling down a hill indefinitely. This line of work analyzes (1) as a countable-state Markov process where $A(n)$ is an independent stochastic process. The AoI process (1) has not yet been studied when $A(n)$ is a general (not necessarily stationary) stochastic process with time dependencies. We intend to close this gap.

B. Paper Summary

We first recall necessary background from probability theory in Section II. Notably, we use the notion of first-order stochastic dominance (Definition 1), which is a stochastic order for random variables (r.v.s) denoted by $X \leq_{\text{st}} \overline{X}$ for some r.v.s X and $\overline{X}$. Further, we use the notion of α-mixing (Definition 2). For every (not necessarily stationary) sequence of r.v.s $\{X_n\}_{n\geq 0}$, one can define α-mixing coefficients $\alpha(X, m)$, which provides a worst-case measure of temporal dependency of events generated by $\{X_n\}_{n=0}^{k}$ and $\{X_n\}_{n=k+m}^{\infty}$ for all $k \geq 0$. We will then discuss details of the AoI process (1) and associated assumptions in Section III.

Our main results are proven in Sections IV and V. Our first result is a set of sufficient conditions that guarantees that the average p-th moment of the AoI process (1) is finite. We will denote by $\alpha(A, n)$ the α-mixing coefficients of the indicator process $\mathbb{1}_{A(n)}$. Theorem 1 shows that sufficiently fast decay of $\alpha(A, n)$ guarantees that all AoI r.v.s $\Delta(n)$ are stochastically dominated by a single r.v. that has certain finite moments depending on the decay rate of $\alpha(A, n)$.

Theorem 1. *Suppose there are $\kappa \geq 0$ and $\varepsilon \in (0, 1)$ such that $\mathbb{P}\left(\cup_{k=n}^{n+\kappa} A(k)\right) \geq \varepsilon$ for all $n \geq 0$. If $\sum_{m\geq 0} m^{p-1}\alpha(A, m) < \infty$ for some $p > 0$, then there is a r.v. $\overline{\Delta}$ with $\Delta(n) \leq_{st} \overline{\Delta}$ for all $n \geq 0$ and $\mathbb{E}\left[\overline{\Delta}^p\right] < \infty$.*

The p-th moment of the dominating r.v. $\overline{\Delta}$ is calculated in the proof of Theorem 1 as a function of ε, κ and the mixing coefficients $\alpha(A, n)$. Notice that Theorem 1 does not require that $A(n)$ is a stationary process. Hence, $\frac{1}{N}\sum_{n=0}^{N-1} \Delta^p(n)$ might not converge. However, by stochastic dominance, we have a bound for the limiting average moments of the AoI process:

Corollary 1. *If Theorem 1 holds for some $p > 0$, then*

$$\limsup_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\sum_{n=0}^{N-1} \Delta^p(n)\right] \leq \mathbb{E}\left[\overline{\Delta}^p\right]. \tag{2}$$

Secondly, we study the mixing rate of the AoI process (1) in Section V. Here, our main result bounds the mixing rate of

(1) by a combination of the mixing coefficient $\alpha(A,n)$ and the tail decay of the dominating r.v. from Theorem 1.

Theorem 2. *Suppose the assumptions of Theorem 1 are satisfied, then the AoI process $\Delta(n)$ is α-mixing with coefficients $\alpha(\Delta,n)$, such that*

$$\alpha(\Delta,n) \leq \min_{0\leq m\leq n}\{\alpha(A,n-m)+\mathbb{P}\left(\overline{\Delta}>m\right)\} \tag{3}$$

with $\overline{\Delta}$ from Theorem 1.

We also show that the mixing rate of $\alpha(\Delta,n)$ is almost as fast as the mixing rate of $\alpha(A,n)$. Specifically, we show that

$$\sum_{n=0}^{\infty} n^{p-1}\alpha(A,n)<\infty \implies \sum_{n=0}^{\infty} n^{q-1}\alpha(\Delta,n) \quad \forall q<p. \tag{4}$$

Using (3), we conclude Section V with a SLLN for the AoI process (1) under sufficiently rapid α-mixing of the event process $A(n)$.

We close our work with conclusions, future work and open problems in Section VI. Most notably, we highlight that our analysis of the abstract AoI process (1) under merely α-mixing communication is a gateway to study more complex AoI processes. For example, renewal processes with non-independent interarrivals or real-world AoI processes where the mixing rate of the communication process has been estimated from data.

II. Probability Theory Background and Notation

We consider an underlying, sufficiently rich probability space $(\Omega,\mathcal{F},\mathbb{R})$ [19, p. 482]. All events are to be understood as elements of $\mathcal{F}$ and all r.v.s are measurable functions from Ω to another measurable space.

Definition 1. A r.v. X is said to be *stochastically dominated* by a r.v. $\overline{X}$, denoted by $X\leq_{\text{st}}\overline{X}$, if $\mathbb{P}\left(X>m\right)\leq\mathbb{P}\left(\overline{X}>m\right)$ for all $m\geq 0$.

Proposition 1 ([20])**.** *Suppose X is a non-negative integer-valued r.v., then for every $p>0$:*

$$\mathbb{E}\left[X^p\right]=\sum_{m=0}^{\infty}((m+1)^p-m^p)\mathbb{P}\left(X>m\right). \tag{5}$$

Let $\mathcal{A}$ and $\mathcal{B}$ be two sub-σ-algebras of $\mathcal{F}$. The following is a measure of dependency between $\mathcal{A}$ and $\mathcal{B}$

$$\alpha(\mathcal{A},\mathcal{B}) \coloneqq \sup_{A\in\mathcal{A},B\in\mathcal{B}}|\mathbb{P}\left(A\cap B\right)-\mathbb{P}\left(A\right)\mathbb{P}\left(B\right)|. \tag{6}$$

Consider a (not necessarily stationary) stochastic process $X=\{X_n\}_{n\geq 0}$. For $0\leq l\leq m\leq\infty$, define the sub-σ-algebra generated from X_l up to X_m by

$$\mathcal{F}_l^m \coloneqq \sigma\left(X_n \mid l\leq n\leq m\right), \tag{7}$$

Informally, the σ-algebra generated by a stochastic process from a time interval describes the information that can be extracted from the associated process realizations, cmp. [21] for details. With these σ-algebras we can now define α-*mixing*, which it a notion of asymptotic independence. We refer to [22] for a survey about α-mixing and other mixing notions.

Definition 2. The α-mixing coefficients of the process X are

$$\alpha(X,n) \coloneqq \sup_{l\geq 0}\alpha(\mathcal{F}_0^l,\mathcal{F}_{l+n}^{\infty}), \quad n\geq 0. \tag{8}$$

The process X is called α-*mixing* (or strongly mixing) if $\alpha(X,n)\to 0$ as $n\to\infty$.

A. Notation

We make frequent use of the small o and big $\mathcal{O}$ notation: Consider two real-valued sequences x^n, y^n. Then $x^n\in o(y^n)$ if $\limsup_{n\to\infty}\frac{x^n}{y^n}=0$ and $x^n\in\mathcal{O}(y^n)$ if $\limsup_{n\to\infty}\frac{x^n}{y^n}<\infty$. Further, we use the floor and ceiling function. For some $x\in\mathbb{R}$, the floor function is defined as $\lfloor x\rfloor \coloneqq \max\{n\in\mathbb{Z}:n\leq x\}$ and the ceiling function is defined as $\lceil x\rceil \coloneqq \min\{n\in\mathbb{Z}:n\geq x\}$.

III. The Age of Information Process

We study the discrete-time AoI stochastic process $\Delta(n)$ in (1) as depicted in Figure 1. We refer to a time slot n as the time interval from time step n to $n+1$. Successful status updates from the source node are received at the monitor whenever an event $A(n)$ occurs. The event $A(n)$ is associated with the n-th time slot. Hence, whenever an event $A(n)$ occurs, the monitor receives a fresh update during time slot n and thus $\Delta(n+1)=1$. We refer to $A(n)$ as the *event process* of the AoI process $\Delta(n)$. With a slight "abuse" of notation, we refer by $A(n)$ interchangeably to the event $A(n)$ as well as to its corresponding indicator function $\mathbb{1}_{A(n)}$.[1] We make the following assumptions for the event process $A(n)$:

Assumption III.1. There is some $\varepsilon\in[0,1)$, such that

$$\mathbb{P}\left(A(n)\right)\geq 1-\varepsilon, \quad \forall n\geq 0.$$

Assumption III.2. The event process $A(n)$ is α-mixing (Definition 2) with coefficients $\alpha(A,n)$.

Assumption III.1 requires that at every time step the monitor receives an update from the source with non-zero probability. A slightly weaker assumption that is also sufficient for our results is that there is some $\kappa\geq 0$ and some $\varepsilon\in[0,1)$, such that $\mathbb{P}\left(\bigcup_{k=n}^{n+\kappa}A(k)\right)\geq 1-\varepsilon, \quad \forall n\geq 0$. The weaker version thus requires that for every time interval of the form $[n,n+\kappa]$ the probability that the monitor receives an update from the source is greater than zero. This weaker assumption is in fact necessary for the existence of a r.v. $\overline{\Delta}$ such that $\Delta(n)\leq_{\text{st}}\overline{\Delta}$ for all $n\geq 0$; without it, there exists a subsequence $\{n_k\}$ such that $\mathbb{P}\left(\lim_{k\to\infty}\Delta(n_k)=\infty\right)=1$. For simplicity, we will present our results in Section IV and Section V for $\kappa=0$, though all proofs can be easily extended to $\kappa>0$.

Assumption III.2 requires that the dependency of events $A(n)$ and $A(m)$ decays as $|m-n|\to\infty$. There are many examples where $A(n)$ will be α-mixing. In general, $A(n)$ will be α-mixing if it can be written as a Borel-measurable function of another α-mixing process [22, Thm. 5.2]. Here, an important class of examples are scenarios where the

[1]The abuse of notation actually follows the de Finetti notation [23] as popularized by David Pollard.

communication events $A(n)$ are a Borel-measurable function of a geometrically ergodic Markov process [22, 24]. This class has been a common tool to approximate wireless fading channels [4, 18, 25]. In addition, it was shown in [26] that an event process that represents successful communication via a geometrically ergodic wireless fading channel is α-mixing even when online, AoI-aware medium access control protocols are used to decide when to communicate.

Another class of examples comes from renewal theory. [27, Theorem 6.1] shows that the event process resulting from a lattice renewal processes with i.i.d. interarrival times is α-mixing if the waiting time distribution has a finite moment greater than one. In [28], the α-mixing rates of linear processes and ARMA processes are determined. Finally, it is also possible to directly estimate α-mixing coefficients from observed data [29, 30]. We will further discuss this in Section VI on future work.

IV. AoI Moment Bounds under α-Mixing Communication

In this section, we analyze the moments of the AoI process (1). As the event process $A(n)$ may in general be non-stationary, we use stochastic dominance (Definition 1) to construct uniformly dominating r.v.s for the AoI process $\Delta(n)$. This immediately leads to bounds for the limiting average moments of the AoI process. Furthermore, uniformly dominating r.v.s also immediately leads to bounds for the asymptotic growth of the AoI process.

To construct uniformly dominating r.v.s, we derive an upper bound to the complementary cumulative distribution functions (CCDF) $\mathbb{P}(\Delta(n) > m)$ uniformly over all $n \geq 0$; the CCDF is also known as the violation probability in the AoI context [15]. In the next subsection, we will construct a function $u : \mathbb{N}_0 \to [0, 1]$ such that $\mathbb{P}(\Delta(n) > m) \leq u(m)$ for all $m \geq 0$ independent of $n \geq 0$ with $\lim_{m\to\infty} u(m) = 0$. We can now use this bound to define the CCDF of a new random variable. Specifically, define a non-negative integer-valued r.v. $\overline{\Delta}$ by defining

$$\mathbb{P}\left(\overline{\Delta} > m\right) := u(m), \quad \forall m \geq 0. \tag{9}$$

This uniquely defines $\overline{\Delta}$ and by construction $\overline{\Delta}$ stochastically dominates all $\Delta(n)$ for all $n \geq 0$. Moreover, if $\sum_{m=0}^{\infty}((m+1)^p - m^p)u(m) < \infty$ for some $p > 0$, then it follows from Proposition 1 that $\mathbb{E}\left[\overline{\Delta}^p\right] < \infty$.

A. Violation probability under α-mixing communication

By construction of the AoI process (1), we have that

$$\mathbb{P}(\Delta(n) > m) = \mathbb{P}\left(\bigcap_{l=n-m}^{n-1} A^c(l)\right), \tag{10}$$

where $A^c(n)$ denotes the complement of the event $A(n)$. Observe that if the events $A(n)$ where independent, then probability on the right-hand side above could directly be written as the product of the individual probabilities. However, we merely consider that $A(n)$ is α-mixing, Assumption III.2. To use this temporal dependency decay, we separate events in (10) by intervals of certain length.

Let $\{a_m\}$ and $\{b_m\}$ be two non-decreasing sequences of non-negative integers with $a_m \leq m$ and $b_m \leq m$. Now fix $n \geq m \geq 0$ and define time indices

$$n_1 := n - m + a_m, \qquad n_k := n_{k-1} + a_m + b_m \tag{11}$$

as long as $n_k \leq n$. Let $L(m)$ be the number of constructed time indices. Notice, that removing events from the intersection in (10) leads to an upper bound. Thus

$$\begin{aligned} \mathbb{P}(\Delta(n) > m) &= \mathbb{P}\left(\bigcap_{l=n-m}^{n-1} A^c(l)\right) &(12)\\ &\leq \mathbb{P}\left(\bigcap_{k=1}^{L(m)}\left(\bigcap_{l=n_k-a_m}^{n_k-1} A^c(l)\right)\right) &(13)\\ &= \mathbb{P}\left(\bigcap_{k=1}^{L(m)}\{\Delta(n_k) > a_m\}\right). &(14)\end{aligned}$$

Notice that by construction of the time indices n_k, the events $\{\Delta(n_k) > a_m\}$ in (14) are separated by b_m steps. The following lemma uses this separation to formulate a bound for the joint event in (14) using the marginals $\mathbb{P}(\Delta(n_k) > a_m)$ and the mixing coefficients $\alpha(A, n)$ of $A(n)$.

Lemma 1. *For $n \geq m \geq 0$,*

$$\begin{aligned}\mathbb{P}(\Delta(n) > m) \leq& \prod_{k=1}^{L(m)} \mathbb{P}(\Delta(n_s) > a_m) &(15)\\ &+ \alpha(A, b_m)\left(\sum_{k=1}^{L(m)-1}\left(\prod_{s=1}^{k-1}\mathbb{P}(\Delta(n_s) > a_m)\right)\right),\end{aligned}$$

where n_s is defined in (11)*.*

Proof. We expand the right-hand side of (14) using Assumption III.2. Consider the following σ-algebras:

$$\mathcal{F}_l^s := \sigma(A(n) \mid l \leq n \leq s), \quad l \geq 0, s \geq 0. \tag{16}$$

Each event $\{\Delta(n_k) > a_m\}$ in (14) is generated by the events $A(n)$ with $n \in \{n_k - a_m, \dots, n_k - 1, n_k - 1\}$. By construction of the above sub-σ-algebras, we have that

$$\{\Delta(n_k) > a_m\} \in \mathcal{F}_{n_k-a_m}^{n_k-1}. \tag{17}$$

Recall that by the construction of the time indices (11), the events $\{\Delta(n_k) > a_m\}$ are separated by b_m steps. We thus have that

$$\{\Delta(n_{L(m)}) > a_m\} \in \mathcal{F}_{n_{L(m)}-a_m}^{\infty} \tag{18}$$

and

$$\bigcap_{k=1}^{L(m)-1}\{\Delta(n_k) > l_m\} \in \mathcal{F}_0^{n_{L(m)-1}-1}, \tag{19}$$

where $L(m)$ is the number of constructed time indices. Due to the aforementioned separation, we have

$$n_{L(m)} - a_m - (n_{L(m)-1} - 1) = b_m + 1 \geq b_m. \tag{20}$$

By Assumption III.2 $\{A(n)\}_{n\geq 0}$ is α-mixing with coefficient $\alpha(A,n)$. It then follows from (18) and (19) and the construction of the time indices n_k that

$$\mathbb{P}\left(\Delta(n) > m\right) \leq \mathbb{P}\left(\{\Delta(n_{L(m)}) > a_m\}\right) \mathbb{P}\left(\bigcap_{k=1}^{L(m)-1} \{\Delta(n_k) > a_m\}\right) + \alpha(A, b_m). \tag{21}$$

The lemma then follows by applying the described procedure successively. □

With Assumption III.1, a preliminary bound for the violation probability follows:

Corollary 2. *Let $n \geq m \geq 0$, then $\mathbb{P}\left(\Delta(n) > m\right) \leq p(m,\delta)$ with*

$$p(m,\delta) := \varepsilon^{L_\delta(m)} + \alpha(A, \lceil m^\delta \rceil)\left(\sum_{k=1}^{L_\delta(m)-1} \varepsilon^{k-1}\right) \tag{22}$$

for every $\delta \in (0,1)$ with $L_\delta(m) := \left\lfloor \frac{m}{1+\lceil m^\delta\rceil}\right\rfloor$.

Proof. For $\delta \in (0,1)$ choose $a_m = 0$ and $b_m = \lceil m^\delta \rceil$ for all $m \geq 0$ in Lemma 1 and use that $\mathbb{P}\left(\Delta(n) > 0\right) \leq \varepsilon$ for all $n \geq 0$ by Assumption III.1. □

Notice that Lemma 1 and Corollary 2 hold without Assumption III.2 since the mixing coefficients are defined for all stochastic process. Under Assumption III.2 it now follows that $\alpha(A,n) \to 0$ as $n \to \infty$. Hence, the right-hand side in (22) decays to zero as $m \to \infty$, since $\lim_{m\to\infty} \lceil m^\delta \rceil = \infty$ and $\lim_{m\to\infty} L_\delta(m) = \infty$ for every $\delta \in (0,1)$. The recipe described at the beginning of this section thus immediately shows that a r.v. exists that jointly stochastically dominates $\Delta(n)$ for all $n \geq 0$. Moreover, for every $q < p$ with p from Assumption III.2, we can choose δ sufficiently close to one such that $p(m,\delta)$ decays sufficiently fast such that

$$\sum_{m=0}^{\infty} ((m+1)^q - m^q) p(m,\delta) < \infty \tag{23}$$

Thus for every $q < p$, we can find a r.v. $\overline{\Delta}$ with $\Delta(n) \leq_{\text{st}} \overline{\Delta}$ for all $n \geq 0$ and $\mathbb{E}\left[\overline{\Delta}^q\right] < \infty$.

B. Proof of Theorem 1

The previous paragraph shows Theorem 1 for all $q < p$. For $q = p$ the situation is different and the bound in Corollary 2 is insufficient to complete Theorem 1. This is because b_m in (15) has to be chosen such that $\liminf \frac{b_m}{m} > 0$, to guarantee that $\sum_{m=0}^{\infty}((m+1)^p - m^p)\alpha(A, b_m) < \infty$. In this case $L(m) < \infty$ and thus a_m has to increase to infinity to guarantee that the first term in (15) decays to zero. The central observation is that both sequences a_m and b_m, as used to construct the time indices in (11), have to been chosen to jointly drift to infinity. To do this, we use the bound from Corollary 2 in Lemma 1 to improve the bound on the violation probability.

Lemma 2. *Let $n \geq m \geq 0$,*

$$\mathbb{P}(\Delta(n) > m) \leq p(\lceil \lambda m \rceil, \delta)^{L_\lambda(m)} + \alpha(A, \lceil \lambda m \rceil)\left(\sum_{k=1}^{L_\lambda(m)-1}\left(\prod_{s=1}^{k-1} p(\lceil \lambda m\rceil, \delta)\right)\right) =: u(m,\delta,\lambda) \tag{24}$$

for every $(\delta,\lambda) \in (0,1)^2$ with $L_\lambda(m) := \lfloor \frac{m}{2\lceil \lambda m\rceil} \rfloor$ and $p(m,\delta)$ as defined in Corollary 2.

Proof. For $\lambda \in (0,1)$ choose $a_m = b_m = \lceil \lambda m \rceil$ for all $m \geq 0$ in Lemma 1. Then, (22) shows that for every $\delta \in (0,1)$ and every $n \geq 0$, we have the bound

$$\mathbb{P}\left(\Delta(n) > a_m\right) \leq p(\lceil \lambda m \rceil, \delta). \tag{25}$$

□

We are now ready to prove Theorem 1.

Proof of Theorem 1. Fix $(\delta,\lambda) \in (0,1)^2$ and observe that $u(m,\delta,\lambda)$ is by construction decreasing in m. Now define a non-negative integer-valued r.v. $\overline{\Delta}$ by describing its CCDF as follows:

$$\mathbb{P}\left(\overline{\Delta} > m\right) := u(m,\delta,\lambda), \quad m \geq 0. \tag{26}$$

By Lemma 2, $\overline{\Delta}$ stochastically dominates all $\Delta(n)$ for $n \geq 0$. To prove Theorem 1, we have to show that there exist $(\delta,\lambda) \in (0,1)^2$, such that $\sum_{m=0}^{\infty}((m+1)^p - m^p)u(m,\delta,\lambda) < \infty$.

We start by showing that

$$\sum_{m=0}^{\infty} ((m+1)^p - m^p)\alpha(A, \lceil \lambda m \rceil) < \infty \tag{27}$$

for all $\lambda \in (0,1)$. We claim that

$$\sum_{m=1}^{\infty} ((m+1)^p - m^p)\alpha(\lceil \lambda m \rceil) \leq \frac{2^p}{\lambda^p} \sum_{m=1}^{\infty} m^{p-1}\alpha(m) \tag{28}$$

This claim follows from the observations that for $x \in \mathbb{R}_{\geq 0}$ and $m \geq 1$, $((m+1)^x - m^x) \leq 2^x m^{x-1}$, and that $|\{n \geq 0 : \lceil \lambda n \rceil = m\}| \leq \frac{1}{\lambda}$. Hence, by Assumption III.2 we have that (27) holds for all $\lambda \in (0,1)$. Further, $L_\lambda(m) \leq \frac{1}{\lambda}$. To complete the proof it is therefore left to show that

$$\sum_{m=0}^{\infty} ((m+1)^p - m^p) p(\lceil \lambda m \rceil, \delta)^{L_\lambda(m)} < \infty \tag{29}$$

for some $(\delta,\lambda) \in (0,1)^2$. Since $(m+1)^p - m^p \leq 2^p m^{p-1}$ for $m \geq 1$, its enough to show that $\sum_{m=1}^{\infty} m^{p-1} p(\lceil \lambda m \rceil, \delta)^{L_\lambda(m)} < \infty$. Using the summability property of $\alpha(A,n)$ from Assumption III.2, we can show that

$$p(\lceil \lambda m \rceil, \delta)^{L_\lambda(m)} \in \mathcal{O}\left(\left(\varepsilon^{\left(m^{1-\delta}\right)} + m^{-\mu\delta}\right)^{\frac{1}{\lambda}}\right) \tag{30}$$

for $\mu := p$, if $p \leq 1$ and $\mu := p-1$, if $p > 1$.[2] Asymptotically, $m^{-\mu\delta}$ will dominate $\varepsilon^{\left(m^{1-\delta}\right)}$ for any $\delta \in (0,1)$. It is thus enough to show that $\sum_{m=1}^{\infty} m^{p-1} m^{-\mu\frac{\delta}{\lambda}} < \infty$, which holds for $\delta\mu > \lambda p$. This completes the proof. □

[2] The distinction follows, since $m^{p-1}\alpha(A,n)$ is not necessarily monotone for $p > 1$.

V. Mixing Rates of AoI Processes

Section IV shows that if $A(n)$ is α-mixing with $\sum_{n=0}^{\infty} n^{p-1}\alpha(A,n) < \infty$, then

$$\mathbb{P}\left(\Delta(n) > m\right) \leq \mathbb{P}\left(\overline{\Delta} > m\right) \in o(m^{-p}) \tag{31}$$

for a random variable $\overline{\Delta}$. This uniform tail decay of the distributions of each $\Delta(n)$ indicates that the dependency of $\Delta(n)$ on $A(m)$ decays as $|n-m| \to \infty$. This in turn indicates that the dependency of $\Delta(n)$ on $\Delta(m)$ also decays. Theorem 2 affirms this indication and that $\Delta(n)$ is itself α-mixing. Now, we present the proof of Theorem 2.

A. Proof of Theorem 2

For $0 \leq l \leq m \leq \infty$, define the σ-algebra generated from $\Delta(l)$ up to $\Delta(m)$ by

$$\mathcal{F}_l^m(\Delta) \coloneqq \sigma\left(\Delta(n) \mid l \leq n \leq m\right). \tag{32}$$

Similarly, define the σ-algebra generated from $A(l)$ up to $A(m)$ by

$$\mathcal{F}_l^m(A) \coloneqq \sigma\left(A(n) \mid l \leq n \leq m\right). \tag{33}$$

The α-mixing coefficients of $\Delta(n)$ are

$$\alpha(\Delta, n) \coloneqq \sup_{l\geq 0}\{\sup_{A,B}|\mathbb{P}\left(A \cap B\right) - \mathbb{P}\left(A\right)\mathbb{P}\left(B\right)|\}, \tag{34}$$

where the supremum is taken over $A \in \mathcal{F}_0^l(\Delta), B \in \mathcal{F}_{l+n}^{\infty}(\Delta)$.

The idea of Theorem 2, is to condition events $A \cap B$ and B in (34) on events $\{\Delta(l+n) \leq m\}$ and $\{\Delta(l+n) > m\}$ for $0 \leq m \leq n$. Then, we use the key property that an event B as above conditioned on $\{\Delta(l+n) \leq m\}$ is already an element of $\mathcal{F}_{n+l-M}^{\infty}(A)$, which thus allows us to use that $A(n)$ is α-mixing.

Lets develop some intuition as to why for $B \in \mathcal{F}_{l+n}^{\infty}(\Delta)$ it holds that

$$B \cap \{\Delta(l+n) \leq m\} \in \mathcal{F}_{l+n-m}^{\infty}(A). \tag{35}$$

Observe that information $\tilde{B} \in \sigma(\Delta(n+l)) = \mathcal{F}_{n+l}^{n+l}(\Delta)$ is of the form

$$\tilde{B} = \Delta(n+l)^{-1}(C),$$

where C is a subset of $\{1, \ldots, n+l\}$. By definition, we have

$$\{\Delta(n+l) \leq m\} = \Delta(n+1)^{-1}(\{1, \ldots, m\}). \tag{36}$$

Using the construction of the AoI process (1) it is then not difficult to see that $\tilde{B} \cap \{\Delta(n+l) \leq m\} \in \mathcal{F}_{l+n-m}^{\infty}(A)$. A monotone class argument now completes the reasoning. A formal proof of property (35) is given in Appendix A. We are now ready to proof Theorem 2.

Proof of Theorem 2. Let $l, m, n \geq 0$, $A \in \mathcal{F}_0^l(\Delta)$ and $B \in \mathcal{F}_{l+n}^{\infty}(\Delta)$. Further, let $\overline{\Delta}$ be a dominating r.v. as given by Theorem 1. By the law of total probability, we have that $|\mathbb{P}\left(A \cap B\right) - \mathbb{P}\left(A\right)\mathbb{P}\left(B\right)|$ is thus equal to

$$\begin{aligned}
&|\mathbb{P}\left(A \cap B \cap \{\Delta(l+n) \leq m\}\right) \\
&\quad + \mathbb{P}\left(A \cap B \cap \{\Delta(l+n) > m\}\right) \\
&\quad - \big(\mathbb{P}\left(A\right)\mathbb{P}\left(B \cap \{\Delta(l+n) \leq m\}\right) \\
&\quad + \mathbb{P}\left(A\right)\mathbb{P}\left(B \cap \{\Delta(l+n) > m\}\right)\big)| \qquad (37)\\
&\leq |\mathbb{P}\left(A \cap B \cap \{\Delta(l+n) \leq m\}\right) \\
&\quad - \mathbb{P}\left(A\right)\mathbb{P}\left(B \cap \{\Delta(l+n) \leq m\}\right)| \\
&\quad + \mathbb{P}\left(\Delta(l+n) > m\right)\Big(|\mathbb{P}\left(A \cap B \mid \{\Delta(l+n) > m\}\right) \\
&\quad - \mathbb{P}\left(A\right)\mathbb{P}\left(B \mid \{\Delta(l+n) > m\}\right)|\Big) \qquad (38)\\
&\leq |\mathbb{P}\left(A \cap B \cap \{\Delta(l+n) \leq m\}\right) \\
&\quad - \mathbb{P}\left(A\right)\mathbb{P}\left(B \cap \{\Delta(l+n) \leq m\}\right)| \\
&\quad + \mathbb{P}\left(\overline{\Delta} > m\right) \qquad (39)
\end{aligned}$$

The first inequality uses conditional probability and triangular inequality; the second inequality uses that $\overline{\Delta} \leq_{\text{st}} \Delta(n)$ for all $n \geq 0$ and that

$$\begin{aligned}
&|\mathbb{P}\left(A \cap B \mid \{\Delta(l+n) > m\}\right) \\
&\quad - \mathbb{P}\left(A\right)\mathbb{P}\left(B \mid \{\Delta(l+n) > m\}\right)| \leq 1.
\end{aligned} \tag{40}$$

By construction of the AoI process, we have that $\mathcal{F}_0^l(\Delta) \subset \mathcal{F}_0^l(A)$ for all $l \geq 0$. Thus $A \in \mathcal{F}_0^l(A)$. By (35), we have that $B \cap \{\Delta(l+n) \leq m\} \in \mathcal{F}_{l+n-m}^{\infty}(A)$. Since $A(n)$ is α-mixing it follows that

$$\begin{aligned}
&|\mathbb{P}\left(A \cap B \cap \{\Delta(l+n) \leq m\}\right) \\
&\quad - \mathbb{P}\left(A\right)\mathbb{P}\left(B \cap \{\Delta(l+n) \leq m\}\right)| \leq \alpha(A, n-M)
\end{aligned} \tag{41}$$

Thus for all $n \geq 0$, we found that

$$\alpha(\Delta, n) \leq \alpha(A, n-m) + \mathbb{P}\left(\overline{\Delta} > m\right) \tag{42}$$

for all $0 \leq m \leq n$. To see that $\alpha(\Delta, n) \to 0$ as $n \to \infty$ choose, e.g., $m(n) = \lceil \frac{n}{2} \rceil$. Minimizing over m in (42) yields the statement of Theorem 2. $\square$

B. Mixing rate of $\Delta(n)$

The next natural question is to analyze the rate of convergence of $\alpha(\Delta, n)$. For this, let $q < p$ and let $\delta \in (0,1)$. Then

$$\begin{aligned}
\sum_{n=0}^{\infty} n^{q-1}\alpha(\Delta, n) &\leq \sum_{n=0}^{\infty} n^{q-1}\alpha(A, \lceil n^{\delta} \rceil) \\
&\quad + \sum_{n=0}^{\infty} n^{q-1}\mathbb{P}\left(\overline{\Delta} > n - \lceil n^{\delta} \rceil\right).
\end{aligned} \tag{43}$$

By Assumption III.2, we have that $\sum_{n=0}^{\infty} n^{p-1}\alpha(A,n)$ and we can show using $q < p$ that the first summation is finite *for some* $\delta \in (0,1)$. We also claim that, the second summation is finite *for every* $\delta \in (0,1)$. This follows from (31) and the observation that $\sum_{n=0}^{\infty} n^{q-1-p} < \infty$ for $q < p$. Here, we used that $\lim_{n\to\infty} \frac{n^p}{(n-\lceil n^\delta \rceil)^p} = 1$ for every $\delta \in (0,1)$. We have therefore shown that

$$\sum_{n=0}^{\infty} n^{p-1}\alpha(A,n) < \infty \implies \sum_{n=0}^{\infty} n^{q-1}\alpha(\Delta,n) \quad \forall q < p. \tag{44}$$

Thus $\Delta(n)$ has almost the same asymptotic mixing rate as $A(n)$.

As a corollary to Theorem 2 and the mixing rate in (44), we can now formulate a SLLN for $\Delta(n)$. The SLLN is based on a SLLN for strongly mixing stochastic processes presented in [31] that we state here paraphrased to suit our purpose:

Theorem 3 ([31, Thm. 2.10]). *Suppose $\{X_n\}_{n\geq 0}$ is a zero mean α-mixing sequence of r.v. with $\sum_{n=0}^{\infty} \alpha(X,n)^{\frac{1}{q}} < \infty$ for some $q > 1$. If there is $1 < r < q$ and $\frac{r}{r-1} < s \leq \frac{2r}{r-1}$ such that*

$$\sum_{n=0}^{\infty} n^{-\frac{r-1}{r}s}\mathbb{E}\left[|X_n|^{\frac{r-1}{r}s}\right] < \infty, \tag{45}$$

then $\frac{1}{N}\sum_{n=0}^{N-1} X_n \to 0$ a.s.

Using the moment bound in Theorem 1 and the mixing rate in Theorem 2, we can carefully choose r, s in Theorem 3 to conclude with the following SLLN for $\Delta(n)$. Details are given in Appendix B.

Corollary 3. *Suppose the assumptions of Theorem 1 hold for some $p > 1$, then*

$$\frac{1}{N}\sum_{n=0}^{N-1}(\Delta(n) - \mathbb{E}[\Delta(n)]) \xrightarrow{a.s.} 0. \tag{46}$$

If in addition $A(n)$ is stationary, then

$$\frac{1}{N}\sum_{n=0}^{N-1}\Delta(n) \xrightarrow{a.s.} \lim_{n\to\infty}\mathbb{E}[\Delta(n)] \leq \mathbb{E}\left[\overline{\Delta}\right] \tag{47}$$

VI. Conclusions, Open Problems and Future Work

AoI has so far only been studied with independent or Markovian interarrivals. To study AoI, renewal theory is a suitable tool, but there is little to no literature on renewal theory with general (non-Markovian) dependent interarrival times. Our presented results offer a path to analyze AoI in such settings even when new information is received at a monitor after general dependent interarrival times. To apply our results, we seek to compute the α-mixing rate of a renewal sequence (i.e. the indicator process that takes the value one when new data arrives) when the sequence of dependent interarrival times is itself α-mixing. This is still an important open problem, since as of now the α-mixing rates of renewal sequences are only known (asymptotically) when the associated interarrival times are i.i.d. [27]. Then, we seek to apply our results to the AoI associated with queuing models with non-independent service times.

Notably, to apply our results to the aforementioned renewal and queuing theory problems, we have to take into account stochastic transmission times as mentioned in Section I. Interestingly, a completely new analysis is not necessary here since it turns out that one can represent the AoI of a renewal process with interarrival times by the concatenation of two identical AoI processes of the simpler form (1), which were analyzed herein. This representation will be discussed in an upcoming paper.

Another interesting line of work will be to identify scenarios where modeling of the process that gives rise to AoI is difficult. For such scenarios, we envision that α-mixing rates of the event process $A(n)$ can be directly estimated from data [30] and then use our results to draw conclusions about the average AoI using Theorem 2.

Appendix A

We verify that for $B \in \mathcal{F}_{l+n}^{\infty}(\Delta)$, we have that

$$B \cap \{\Delta(l+n) \leq M\} \in \mathcal{F}_{l+n-M}^{\infty}(A), \tag{48}$$

as used in Section V-A.

First consider $B \in \sigma(\Delta(n+l))$. Every AoI random variable is a measurable map $\Delta(n+l) : \Omega \to 2^{\{1,\ldots,n+l\}}$. Thus B is of the form $B = \Delta(n+l)^{-1}(C)$ for some $C \in 2^{\{1,\ldots,n+l\}}$, i.e. C is a subset of $\{1,\ldots,n+l\}$. Second, for any $0 \leq m \leq n+l$, we have

$$\{\Delta(n+l) \leq m\} = \Delta(n+1)^{-1}(\{1,\ldots,m\}). \tag{49}$$

Therefore, using properties of the preimage of intersections, we have that

$$B \cap \{\Delta(n+l) \leq m\} = \Delta(n+l)^{-1}\left(C \cap \{1,\ldots,m\}\right) \tag{50}$$

$$= \bigcup_{c\in C\cap\{1,\ldots M\}} \Delta(n+l)^{-1}(\{c\}). \tag{51}$$

By construction of the AoI process, we have that

$$\Delta(n+l)^{-1}(c) = A(n+l-c) \cap \bigcap_{k=1}^{c-1} A^c(n+l-c+k). \tag{52}$$

Since $\mathcal{F}(A)_{l+n-m}^{\infty}$ is a σ-algebra, we thus conclude from (51) and (52) that

$$B \cap \{\Delta(n+l) \leq m\} \in \mathcal{F}(A)_{l+n-m}^{\infty}. \tag{53}$$

Next we consider elements of the join σ-algebra $\mathcal{F}(\Delta)_{n+l}^{\infty}$. It can be expressed as

$$\mathcal{F}(\Delta)_{n+l}^{\infty} = \sigma\left(\bigcup_{k=n+l}^{\infty}\sigma(\Delta(k))\right). \tag{54}$$

For all $B \in \bigcup_{k=n+l}^{\infty}\sigma(\Delta(k))$, the previous paragraph shows that

$$B \cap \{\Delta(n+l) \leq m\} \in \mathcal{F}(A)_{l+n-m}^{\infty}, \tag{55}$$

using the stability of σ-algebras under countable unions. A generating-class argument now completes the proof: Let $\mathcal{G} := \{B \in \mathcal{F}(\Delta)_{n+l}^{\infty} : B \text{ satisfies (53)}\}$. Clearly, $\Omega \in \mathcal{G}$ and countable unions of elements from $\mathcal{G}$ are in $\mathcal{G}$. Finally, let $B \in G$. Then

$$B^c \cup \{\Delta(n+l) > m\} \in \mathcal{F}(A)_{l+n-m}^{\infty}, \tag{56}$$

since $B \cap \{\Delta(n+l) \leq m\} \in \mathcal{F}(A)_{l+n-m}^{\infty}$ and $\mathcal{F}(A)_{l+n-m}^{\infty}$ is a σ-algebra. Finally,

$$B^c \cap \{\Delta(n+l) \leq m\} = (B^c \cup \{\Delta(n+l) > m\}) \cap \{\Delta(n+l) \leq m\} \in \mathcal{F}(\Delta)_{n+l-m}^{\infty}, \quad (57)$$

again, since $\{\Delta(n+l) \leq m\} \in \mathcal{F}(\Delta)_{n+l-m}^{\infty}$ and $\mathcal{F}(A)_{l+n-m}^{\infty}$ is a σ-algebra. We have therefore shown that $\mathcal{G}$ is itself a σ-algebra, hence $\mathcal{G} = \mathcal{F}(\Delta)_{n+l-m}^{\infty}$.

Appendix B

Proof of Corollary 3. Suppose that $A(n)$ is α-mixing with $\sum_{n=0}^{\infty} n^{p-1}\alpha(A,n) < \infty$ for some $p > 1$. Theorem 2 shows that $\Delta(n)$ is α-mixing with $\sum_{n=0}^{\infty} n^{q-1}\alpha(\Delta,n) < \infty$ for every $q < p$.

Let $1 < q < p$. To apply Theorem 3 to $\Delta(n)$, we first have to show that

$$\sum_{n=0}^{\infty} \alpha(\Delta,n)^{\frac{1}{q}} = \sum_{n=0}^{\infty} \alpha(\Delta,n)^{\frac{1}{q}-1}\alpha(\Delta,n) < \infty \quad (58)$$

Since $\sum_{n=0}^{\infty} n^{q-1}\alpha(\Delta,n) < \infty$ and $\alpha(\Delta,n)$ is monotone, we especially have that $\alpha(\Delta,n) \in o(n^{-1})$ and hence $\alpha(\Delta,n)^{\frac{1}{q}-1} \in o(n^{1-\frac{1}{q}})$. Finally (58) follows, since $1 - \frac{1}{q} \leq q-1$ for $q \geq 0$ and $\sum_{n=0}^{\infty} m^{q-1}\alpha(\Delta,n) < \infty$ by Assumption III.2.

To complete the proof, recall that Theorem 1 showed that $\mathbb{E}[\Delta(n)^p] \leq \mathbb{E}\left[\overline{\Delta}^p\right] < \infty$. It is now easy to see that we can choose r, s in Theorem 3, such that $1 < \frac{r-1}{r}s < p$ and thus (45) holds. Theorem 3 therefore shows that $\frac{1}{N}\sum_{n=0}^{N-1}(\Delta(n) - \mathbb{E}[\Delta(n)]) \xrightarrow{\text{a.s.}} 0$. If in addition $A(n)$ is strictly stationary, then $\mathbb{E}[\Delta(n)]$ is monotonically increasing and the additional statement follows. □

References

[1] M. Montero and J. Villarroel, "Directed random walk with random restarts: The sisyphus random walk," *Physical Review E*, vol. 94, no. 3, p. 032132, 2016.

[2] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[3] W. Liu and P. Shi, "Optimal linear filtering for networked control systems with time-correlated fading channels," *Automatica*, vol. 101, pp. 345–353, 2019.

[4] G. Bianchi, "Performance analysis of the ieee 802.11 distributed coordination function," *IEEE Journal on selected areas in communications*, vol. 18, no. 3, pp. 535–547, 2000.

[5] G. Samorodnitsky *et al.*, *Stochastic processes and long range dependence*. Springer, 2016, vol. 26.

[6] W. Feller, "On the integral equation of renewal theory," *The Annals of Mathematical Statistics*, vol. 12, no. 3, pp. 243–267, 1941.

[7] J. L. Doob, "Renewal theory from the point of view of the theory of probability," *Transactions of the American Mathematical Society*, vol. 63, no. 3, pp. 422–438, 1948.

[8] S. Asmussen, *Applied probability and queues*. Springer Science & Business Media, 2008, vol. 51.

[9] G. Alsmeyer, *The Markov renewal theorem and related results*. Citeseer, 1996.

[10] S. Kombrink, "Renewal theorems for processes with dependent interarrival times," *Advances in Applied Probability*, vol. 50, no. 4, pp. 1193–1216, 2018.

[11] M. Xiong and K. Ramamritham, "Deriving deadlines and periods for real-time update transactions," *IEEE Transactions on Computers*, vol. 53, no. 5, pp. 567–583, 2004.

[12] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *2012 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2012, pp. 2666–2670.

[13] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *2012 46th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2012, pp. 1–6.

[14] Y. Sun, I. Kadota, R. Talak, and E. Modiano, "Age of information: A new metric for information freshness," *Synthesis Lectures on Communication Networks*, vol. 12, no. 2, pp. 1–224, 2019.

[15] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, 2021.

[16] S. Farazi, A. G. Klein, and D. R. Brown, "Average age of information in update systems with active sources and packet delivery errors," *IEEE Wireless Communications Letters*, vol. 9, no. 8, pp. 1164–1168, 2020.

[17] D. Daley and T. Rolski, "Finiteness of waiting-time moments in general stationary single-server queues," *The Annals of Applied Probability*, pp. 987–1008, 1992.

[18] M. Boban, X. Gong, and W. Xu, "Modeling the evolution of line-of-sight blockage for v2v channels," in *2016 IEEE 84th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2016, pp. 1–7.

[19] P. Billingsley, *Probability & measure*. Wiley&Sons, 2008.

[20] S. Chakraborti, F. Jardim, and E. Epprecht, "Higher-order moments using the survival function: The alternative expectation formula," *The American Statistician*, 2018.

[21] R. Durrett, *Probability: Theory and Examples*. Cambridge University Press, 2019.

[22] R. C. Bradley, "Basic properties of strong mixing conditions. a survey and some open questions," *Probability surveys*, vol. 2, pp. 107–144, 2005.

[23] B. De Finetti, *Teoria delle probabilità: sintesi introduttiva con appendice critica. 2 (1970)*. Einaudi, 1970.

[24] Y. A. Davydov, "Mixing conditions for markov chains," *Theory of Probability & Its Applications*, vol. 18, no. 2, pp. 312–328, 1974.

[25] H. S. Wang and N. Moayeri, "Finite-state markov channel-a useful model for radio communication channels," *IEEE transactions on vehicular technology*, vol. 44, no. 1, pp. 163–171, 1995.

[26] A. Redder, A. Ramaswamy, and H. Karl, "Practical

network conditions for the convergence of distributed optimization," in *IFAC Conference on Networked Systems*, 2022.

[27] H. Berbee, "Convergence rates in the strong law for bounded mixing sequences," *Probability theory and related fields*, vol. 74, no. 2, pp. 255–270, 1987.

[28] T. D. Pham and L. T. Tran, "Some mixing properties of time series models," *Stochastic processes and their applications*, vol. 19, no. 2, pp. 297–303, 1985.

[29] D. J. McDonald, C. R. Shalizi, and M. Schervish, "Estimating beta-mixing coefficients via histograms," *Electronic Journal of Statistics*, vol. 9, no. 2, pp. 2855–2883, 2015.

[30] A. Khaleghi and G. Lugosi, "Inferring the mixing properties of an ergodic process," *arXiv preprint arXiv:2106.07054*, 2021.

[31] D. L. McLeish, "A maximal inequality and dependent strong laws," *The Annals of probability*, vol. 3, no. 5, pp. 829–839, 1975.