Age of Information Optimization in RIS-Assisted Wireless Networks

Ali Muhammad*, Mohamed Elhattab*, Mohamed Amine Arfaoui*, Ahmed Al-Hilo, and Chadi Assi*, Fellow, IEEE

Abstract—In this paper, we consider a wireless network consisting of a base station that is serving multiple real-time traffic streams forwarding information updates to their destinations in order to sustain the freshness of information for time-critical applications. Since the wireless channels may be unreliable due to the impurities of the propagation environments, such as deep fading, blockages, etc., we integrate a reconfigurable intelligent surface to the wireless system in order to mitigate the propagation-induced impairments, enhance the quality of the wireless links, and ensure that the required freshness of information is achieved for these real-time applications. For this network set-up, we investigate the joint optimization of the traffic streams scheduling and the reconfigurable intelligent surface phase-shift matrix with the goal of minimizing the long-term average Age of Information. The formulated optimization problem is a mixed integer non-convex optimization problem, which is difficult to solve. To circumvent the high-coupled optimization variables, and with the aid of bi-level optimization, we decompose the original problem into an outer traffic stream scheduling problem and an inner reconfigurable intelligent surface phase-shift matrix problem. For the outer problem, owing to its complexity and stochastic nature of packet arrivals, we resort to deep reinforcement learning solution where the traffic stream scheduling is modeled as a Markov Decision Process, and Proximal Policy Optimization is invoked to solve it. Whereas, the inner problem that determines the reconfigurable intelligent surface configuration is solved through semi-definite relaxation. Finally, we show through extensive simulations that our approach evaluates the combined impact of scheduling policy and reconfigurable intelligent surface configuration on the long-term average Age of Information, where we demonstrate its superiority against other baseline schemes.

Index Terms—Information freshness, deep reinforcement learning, passive beamforming, scheduling, 6G.

I. INTRODUCTION

A. Motivation

The Sixth-Generation (6G) of mobile networks is poised to transform the landscape of communication systems by offering a massive connectivity, ultra reliable and low-latency communications, and soaring broadband speeds. Such transformation will give rise to a wide range of low-latency cyber-physical system applications. For example, camera images from vehicles are being utilized in intelligent transportation systems to describe the surroundings, speed and position of the vehicles. Video streams from security cameras augmented with informative labels are used to enhance the safety procedures in traffic monitoring and control systems. Another application pertains to remote surgery systems, which needs to update the position of robotically controlled surgical tools in tactile Internet environment. All these applications share a common description, i.e., a source generates multiple time-stamped update messages that are transmitted via communication network to one or more destinations. Thus, the awareness of the state of the system must be as timely as possible at any time to make critical decisions. For example, changing lanes in autonomous driving systems may lead to a catastrophe if the state of the vehicles is not timely updated. Precisely, if the information delivered is not “fresh”, i.e., if the information updates are not timely delivered, there may be severe consequences impacting not only the performance of these intelligent and critical systems, but also the safety and wellness of people. Thus, reliability and timeliness in delivering status updates are of primordial importance for these real-time applications.

Recently, information freshness has been investigated through defining a new performance metric that is Age of Information (AoI). AoI quantifies the freshness of status updates from the destination perspective [2]. AoI is defined as the elapsed time since the most recent delivered status message was generated [3]. AoI has brought a sheer novelty in specifying the information freshness against other metrics such as delay and latency for time-critical applications. These traditional performance metrics may not be sufficient to quantify the timeliness of these status update systems based on the following two reasons: firstly, timely updating is not same as maximizing the generation of updates as it may cause the backlogging in the communication systems and will lead to delayed updates at the receiver, secondly, throttling the update rate will also result in having outdated
In this paper, we investigate the AoI minimization problem in RIS-aided time-sensitive applications. Specifically, we aim at minimizing the expected sum AoI by optimizing the user scheduling decisions and the phase shifts of RIS elements. The main contributions of this work are summarized as follows:

- We formulate a joint user scheduling and phase-shift matrix (passive beamforming) optimization problem with the objective of minimizing the expected sum AoI of multiple traffic streams.
- Owing to the stochastic nature of arrival of packets, the combinatorial nature of the user scheduling task, and the non-convexity of the different system constraints, it is extremely challenging to solve the formulated problem. Alternatively, with the aid of bi-level optimization, the original problem is reformulated into an outer user scheduling problem and an inner phase-shift matrix optimization problem. Due to the uncertainties in the environment (e.g., random arrival of packets), it is often hard and impractical to solve the outer user scheduling problem using standard optimization techniques even for small number of instances. Hence, our problem is reformulated as Markov Decision Process (MDP) and in the absence of the transition probabilities of the MDP, a model-free RL based on Actor-Critic is exploited to find the best policy for the transmission scheduling. The policy-based DRL algorithms show significant performance improvements compared to state-of-the-art algorithms. For the inner problem, an efficient algorithm based on semi-definite relaxation (SDR) is proposed.
- The performance of the proposed approach is assessed through extensive simulations, where different baseline methods were considered for comparison purposes. We demonstrate that our proposed scheme achieves the minimum expected sum AoI in contrast with the other considered methods.

In the simulation results, we show how the integration of RIS can significantly reduce the AoI of time-critical applications as compared to the case where there is no RIS. In addition, for comparison purposes, three baseline schemes, namely, greedy scheduling with SDR, Round-Robin scheduling with SDR and DRL with a random RIS configuration, were adopted and we demonstrated the superiority of the proposed scheme. Finally, we evaluate the performance of the proposed scheme and baseline approaches with respect to different system parameters, including the size and the location of RIS and the arrival rate of the updates’ packets.

C. Outline and Notations

The remainder of this paper is organized as follows. The related works are presented in Section II. The system model is presented in Section III. The definition of age of information accompanied by an illustrative example is presented in Section IV. The problem formulation, the proposed scheduling and phase shift optimization algorithms are explained in Section V. Section VI demonstrates the performance evaluations on the proposed algorithm with different parameter settings. Finally, Section VII concludes the paper. The notations used throughout the paper are summarized in Table I.
II. RELATED WORKS

The aim of this work is to investigate the AoI improvement that can be brought by RIS. Based on this, the two main components of this study are AoI and RIS. Here, we present the relevant works related to AoI and RIS that are reported in the literature.

1) AoI Based Data Transmissions: AoI metric has received a considerable interest from the research community, accentuating its benefits especially for time-sensitive systems. The AoI minimization problem has been investigated in various domains, such as, vehicular networks [11], [12], softwarized networks [13], UAV-assisted communications [14], [15], edge caching [16], [17], and mobile edge computing assisted networks [18], [19]. More relevant to this research are the works that have investigated the AoI minimization problem with stochastic arrivals. In [20], a lower bound on the average AoI performance is derived for networks with stochastic packet arrivals under three different queuing scenarios, namely, no queue, single queue, and first-in-first-out queue. The authors of [21] investigated the AoI in a carrier-sense multiple access based system employing the stochastic hybrid system tools where $N$ links contend for a channel. They aimed to optimize the average AoI by adjusting the back-off time of each link.

In [22], the authors proposed a near-optimal solution to address the optimization of AoI in wireless communication networks wherein Whittle’s index was used to capture the transmission urgency of terminals. The authors of [23] considered various sampling periods and sample sizes for each source node and proposed a low-complexity scheduling algorithm that achieves near-optimal performance when there is no synchronization among the nodes during the sampling process. The authors of [24] investigated the AoI minimization problem in the context of cellular Internet of UAVs and formulated a framework that jointly optimize the sensing and transmission time, the UAV trajectory and the task scheduling (i.e., the selection of the sensing tasks). The formulated problem, which is NP-hard, was decoupled into two sub-problems and that were solved using an iterative algorithm and a dynamic programming approach, respectively.

Different from the above background, this work leverages RIS in a wireless network to enhance information freshness at the end users by minimizing the AoI. In the following part, we will discuss the recent research contributions on the integration of RIS in wireless cellular networks.

2) RIS Aided Wireless Networks: In [25], the authors addressed the minimization problem of total transmit power at the transmitter by jointly optimizing the transmit beamforming through the active antenna array of the transmitter and the passive beamforming through the phase-shift elements of the RIS. The authors of [26] developed different free-space path-loss models for RIS-assisted wireless communications, with the goal of enhancing the network coverage in a cost-effective and energy-efficient way through optimizing the phase-shifts of the RIS elements. Considering the potential challenges pertaining to spectrum and energy usage in Device-to-Device (D2D) communication, the authors in [9] focused on an RIS-assisted uplink D2D-enabled cellular networks and investigated the joint power allocation and RIS phase-shift optimization problem with an objective to maximize the sum rate. The authors of [27] investigated the resource allocation problem for multi users communication leveraging the RIS. More specifically, the total transmit power is minimized through an optimal design of the transmit power at the base station (BS) and the passive beamforming at the RIS. In [28], the authors proposed a two-way communication model assisted by an RIS, where the objective was to maximize the minimum received signal-to-interference-plus-noise ratio (SINR) at the cellular users by optimizing the RIS configuration. The paramount security performance of multi-input and multi-output wireless communication systems is probed by invoking the RIS in [10], where the aim was maximizing the secrecy rates through a proper design of the RIS configuration and the transmit power. The authors of [29] proposed the concept of holographic multiple-input and multiple-output (MIMO) surfaces (HMIMOS), and discussed both active and passive RISs, explaining the hardware architectures, operation modes, and applications in communications. The authors of [8] presented a hybrid beamforming architecture for multi-hop RIS-assisted MIMO systems with an objective to improve the coverage range in the terahertz frequency band based on DRL approach. The substantial performance of DDPG to find the RIS phase configurations, as well as the BS precoding vectors and power allocation strategies was demonstrated. The authors of [30] presented a joint design of transmit beamforming and phase shifts in a RIS-assisted MIMO system by leveraging DRL. Different from the conventional alternating optimization techniques to determine the RIS phase shift and transmit beamforming, this work utilizes deep neural network to obtain a joint design simultaneously. Despite that aforementioned works presented RIS based solutions, none of these works studied the effect of the RIS on improving the AoI.

Recently, the authors of [31] explored the integration of RIS to minimize AoI. However, the authors considered a deterministic transmission policy where new update packets are instantaneously generated by all devices as soon as the last
sent packets are received at the BS. The authors of [32] investigated the AoI minimization problem in an RIS-assisted SWIPT network. An SCA based alternating optimization algorithm was presented to solve the scheduling problem with joint active and passive beamforming design. However, the long-term AoI evolution was not considered in the design. The authors of [33] addressed the AoI minimization problem in UAV-assisted RIS networks. The work considered scheduling a single user within a given time-slot and ignored the direct channels from the BS to the users. To the best of our knowledge, the integration of RIS in time-sensitive applications, where the freshness of information is of critical importance, is still far from being mature. Motivated by this, in this work, as opposed to [33], we consider the problem of scheduling a finite number of streams in order to transmit their information update messages in a more general setting by including the direct channels from the BS to the users. We summarize the main differences between our paper and other closely related works in Table II.

III. SYSTEM MODEL

A. Network Model

We consider a downlink wireless network consisting of one BS equipped with a single antenna, that is serving \( I \) traffic streams to forward their status-update messages to \( I \) destinations as depicted in Fig. 1.\(^1\) We assume that the BS is equipped with \( I \) virtual queues, within which the BS only stores the most recent packet of each stream. The time dimension is slotted into time-slots, where each is represented by a time-slot index \( t \in [1, \infty) \). Let \( T \) denote the time horizon of this discrete-time system. In addition, let \( T = \{1, 2, \ldots, T\} \) denote the set of time slot indices within the time interval \([1, T]\) and let \( \mathcal{I} = \{1, 2, \ldots, I\} \) denote the set of the traffic streams. In this setting, at the beginning of every time-slot \( t \in T \), a packet from stream \( i \in \mathcal{I} \) arrives to the system with a probability \( \lambda_i \in (0, 1] \). Accordingly, for all \( t \in T \) and \( i \in \mathcal{I} \), let \( u_i(t) \) be the binary variable that indicates whether a packet from the \( i \)th traffic stream arrives to the BS at the \( t \)th time-slot or not. Based on its definition, for all \( i \in \mathcal{I} \), the arrival process \( u_i(t) \) is a Bernoulli arrival process that is i.i.d over time, where \( P(u_i(t) = 1) = \lambda_i \). Moreover, the arrival processes \((u_i(t))_{1 \leq i \leq I}\) are independent across the different streams.

Due to the obstacles of the wireless propagation environment, the existence of a strong direct line-of-sight (LoS) communication link between the BS and each destination is difficult to obtain. For this purpose, an RIS equipped with \( F \) reflecting elements is assumed to be deployed within the considered wireless network to assist the transmission from the BS by passively relaying the status update information to the destinations. The BS continuously controls the phase-shift of the reflecting elements in order to maintain the quality of service (QoS) required by the destinations. In this context, for all \( t \in \{1, 2, \ldots, T\} \), let \( \Phi(t) = \text{diag}(\exp[\theta(t)]) \in \mathbb{C}^{F \times F} \) denotes the \( F \times F \) phase-shift matrix of the RIS, where \( \theta(t) = [\theta_1(t), \theta_2(t), \ldots, \theta_F(t)]^T \) is the \( F \times 1 \) vector that contains the phase-shifts of the RIS, such that, for all \( f \in \mathcal{F} \triangleq \{1, 2, \ldots, F\} \), \( \theta_f(t) \in [0, 2\pi) \) is the phase-shift of the \( f \)th reflecting element of the RIS.

The total bandwidth available at the BS is divided into \( N \) channels, where each channel consists of one spectrum resource. The channel diversity exists between different channels and BS can schedule different traffic streams to at most \( N \) channels where each traffic stream is assumed to be allocated to only one channel [23]. Let \( \mathcal{N} = \{1, 2, \ldots, N\} \) denote the set of the \( N \) channels. Moreover, for all \( t \in T \), \( i \in \mathcal{I} \) and \( n \in \mathcal{N} \), let \( x_{i,n}(t) \) be the indicator whether the \( i \)th stream has been scheduled by the BS on \( n \)th channel in the \( t \)th time-slot or not.\(^2\) This is to note that the term scheduling is collectively used for selecting a traffic stream and allocating it a channel for transmission. On the other hand, scheduling a traffic stream without allocating a channel resource and vice versa have no meaning at all. \( x_{i,n}(t) \) is defined as follows:

\[
x_{i,n}(t) = \begin{cases} 
1 & \text{if traffic stream } i \text{ is scheduled on channel } n \text{ in time-slot } t, \\
0 & \text{otherwise},
\end{cases}
\]

Based on this, the transmission scheduling constraints are given as follows.

\[
\sum_{i=1}^{I} \sum_{n=1}^{N} x_{i,n}(t) \leq N, \quad \forall t \in T. 
\]

\[
\sum_{n=1}^{N} x_{i,n}(t) \leq 1, \quad \forall t \in T, \ i \in \mathcal{I}. 
\]

B. Channel Model and SNR Analysis

For all \( t \in T \), \( i \in \mathcal{I} \), and \( n \in \mathcal{N} \), the channel coefficients between the BS and the RIS, between the RIS and the \( i \)th destination, and between the BS and the \( i \)th destination on the \( n \)th spectrum resource are denoted, respectively, by \( h_{b\rightarrow R,n}(t) \in \mathbb{C}^{F \times 1} \), \( h_{R\rightarrow i,n}(t) \in \mathbb{C}^{F \times 1} \) and \( h_{b\rightarrow i,n}(t) \in \mathbb{C}^{F \times 1} \). All channel coefficients consist of both the small-scale fading and the

\(^1\)In this work, we focus on studying the fundamentals and presenting a proof of concept for RIS-enabled single-input-single-output wireless networks that generate real-time information updates, where our main target is characterizing the corresponding performance, in terms of AoI, in order to provide succinct insights. The use of multiple antennas at the BS can significantly boost the performance of the proposed model. This will be considered in future works, where the proposed techniques in the current work can be exploited.

\(^2\)Note that the transmission from BS to each of the destinations through each direct link (BS-destination) and indirect link (BS-RIS-destination) takes only one time-slot. It is worth mentioning that the RIS is a full-duplex technology with interference-free transmission.
large-scale fading. In fact, for all $t \in T$, $i \in \mathcal{I}$, and $n \in \mathcal{N}$, the channel coefficients $h_{b \rightarrow R, n}(t)$, $h_{R \rightarrow i, n}(t)$ and $h_{b \rightarrow i, n}(t)$ are expressed, respectively, as

$$h_{b \rightarrow R, n}(t) = \hat{h}_{b \rightarrow R, n}(t) \Delta_{b \rightarrow R}$$  \hspace{1cm} (4)
$$h_{R \rightarrow i, n}(t) = \hat{h}_{R \rightarrow i, n}(t) \Delta_{R \rightarrow i}$$  \hspace{1cm} (5)
$$h_{b \rightarrow i, n}(t) = \bar{h}_{b \rightarrow i, n}(t) \Delta_{b \rightarrow i}$$ \hspace{1cm} (6)

where $\hat{h}_{b \rightarrow R, n}(t)$, $\hat{h}_{R \rightarrow i, n}(t)$ and $\hat{h}_{b \rightarrow i, n}(t)$ represent the small-scale fading coefficients between the BS and the RIS, between the RIS and the $i$th destination, and between the BS and the $i$th destination on the $n$th frequency resource, respectively, whereas $\Delta_{b \rightarrow R}$, $\Delta_{R \rightarrow i}$ and $\Delta_{b \rightarrow i}$ represent the large-scale fading coefficients between the BS and RIS, between the RIS and the $i$th destination, and between the BS and the $i$th destination respectively. Additionally, for all $i \in \mathcal{I}$, and $n \in \mathcal{N}$, the large-scale fading coefficients can be modeled as

$$\Delta_{b \rightarrow R} = \sqrt{\gamma_0 d_{b \rightarrow R}^{-\eta_b}}$$ \hspace{1cm} (7)
$$\Delta_{R \rightarrow i} = \sqrt{\gamma_0 d_{R \rightarrow i}^{-\eta_R}}$$ \hspace{1cm} (8)
$$\Delta_{b \rightarrow i} = \sqrt{\gamma_0 d_{b \rightarrow i}^{-\eta_b}}$$ \hspace{1cm} (9)

where $\gamma_0$ is the path-loss average channel power gain at a reference distance $d_0 = 1$m, $\eta_b$ is the path-loss exponent for the wireless link $k \in \{br, Ri, bi\}$, $d_{R \rightarrow i}$ represents the distance between the RIS and $i$th destination, $d_{b \rightarrow i}$ represents the distance between the BS and $i$th destination, and $d_{b \rightarrow R}$ represents the distance between the BS and RIS. The small scale fading of the direct links between the BS and the destinations is modelled as a Rayleigh fading channel with zero mean and unit variance [35]. Meanwhile, the communication links between the BS and the RIS and between the RIS and the destinations are considered to have LoS components. These links experience small-scale fading that is modelled as Rician fading [35]. Accordingly, for all $t \in T$ and $n \in \mathcal{N}$, the small-scale fading $\bar{h}_{b \rightarrow R, n}(t)$ between the BS and the RIS on the $n$th frequency resource is defined as:

$$\bar{h}_{b \rightarrow R, n}(t) = \frac{K_1}{K_1 + 1} h_{b \rightarrow R, n}(t) + \frac{1}{K_1 + 1} h_{b \rightarrow R, n}(t), \hspace{1cm} (10)$$

where $K_1$ is the Rician factor, and $h_{b \rightarrow R, n}(t)$ and $h_{b \rightarrow R, n}(t)$ are the deterministic LoS and Rayleigh fading coefficients respectively. Similarly, for all $t \in T$, $i \in \mathcal{I}$, and $n \in \mathcal{N}$, the small-scale fading $h_{R \rightarrow i, n}(t)$ between the RIS and the $i$th destination on the $n$th frequency resource is given as:

$$h_{R \rightarrow i, n}(t) = \frac{K_2}{K_2 + 1} h_{R \rightarrow i, n}(t) + \frac{1}{K_2 + 1} h_{R \rightarrow i, n}(t), \hspace{1cm} (11)$$

where $K_2$ is the Rician factor and $h_{R \rightarrow i, n}(t)$ and $h_{R \rightarrow i, n}(t)$ are the deterministic LoS and Rayleigh fading coefficients respectively. Additionally, similar to other works in literature [36], [37], we assume that the channel state information (CSI) of the considered wireless links is perfectly estimated at the BS. Although, obtaining the perfect CSI is quite challenging, recent studies [38], [39] have provided means to obtain efficient channel estimation techniques for RIS-enabled networks that can be embraced with our system model to obtain accurate CSI.

Based on the above discussion, and for all $t \in T$, $i \in \mathcal{I}$, and $n \in \mathcal{N}$, the signal-to-noise ratio (SNR) at the $i$th destination

| Table II |
| Summary of Closely Related Works |
| --- |
| **Ref.** | **UL/DL** | **Position of RIS** | **Random Packets** | **Objective function** | **Machine learning based solution** | **Others** |
| [32] | DL | No | Yes | Minimize sum AoI | No | SCA based algorithm, SWIPT network |
| [14] | UL | No RIS | Yes | Minimize weighted sum AoI | DQN | UAVs trajectory constraint |
| [25] | DL | Yes | No | Minimize transmit power | No | Multi-antenna AP direct transmission BS and users is considered |
| [24] | UL | No RIS | Yes | Minimize sum AoI | No | UAVs trajectory constraint |
| [34] | DL | No | Yes | Minimize weighted sum AoI | No | SCA based algorithm, Aerial RISs are considered |
| [30] | DL | No | No | Maximize sum rate | DDPG | direct transmission between users and BS not considered |
| [8] | DL | No | No | Maximize sum rate | DDPG | considered multi-hop transmission |
| [33] | UL | No | No | Minimize expected sum AoI | PPO algorithm | single user scheduled per timeslot, direct transmission between users and BS not considered |
| [1] | DL | No | Yes | Minimize sum AoI | No | multiple-user scheduling per timeslot, direct transmission BS and users is considered |
| Our paper | DL | Yes | Yes | Minimize expected sum AoI | PPO algorithm |
at the $t$ time-slot and for the $n$th channel can be expressed as
\[
\gamma_{i,n}(t) = \frac{P|h_{i\to R,i,n}(t)\Phi(t)h_{R\to i,n}(t) + h_{b\to i,n}(t)|^2}{\sigma^2}, \quad (12)
\]
where $\sigma^2$ is the noise power at each destination and $P$ is the transmit power of the BS. So far, we have discussed the main components related to the SNR at each destination. Next, we will discuss the main elements for the AoI problem.

IV. AOI EVOLUTION AND ILLUSTRATIVE EXAMPLE

The AoI illustrates how old the information is from a destination’s perspective and is defined as the time elapsed since the most recent successful transmission of the valid information update [3]. For all $t \in T$ and $i \in I$, let $y_i(t)$ denote the AoI for a destination $i$ in time-slot $t$. In addition, it is important to mention that a successful delivery of a packet at the destination in a given time slot $t$, for all $t \in T$, is conditioned on two realizations:

1) The stream selected by the BS for scheduling in time-slot $t$ has a packet available in its queue.

2) The SNR of the channel between the BS and the destination including the impact of both the direct and indirect links is above a given threshold.

Precisely, for all $i \in I$ and $t \in T$, if a packet of the $i$th traffic stream is scheduled by the BS and it is successfully delivered at the $t$th time-slot, then the corresponding AoI in the subsequent time-slot will be given by $y_i(t + 1) = z_i(t) + 1$, where $z_i(t)$ represents the system time of the packet in queue $i$ at the beginning of slot $t$. Conversely, if the transmission remained unsuccessful, then the AoI in the subsequent time-slot will be given by $y_i(t + 1) = y_i(t) + 1$. Hence, for all $i \in I$, the evolution of AoI of destination $i$ [20] is given as
\[
y_i(t + 1) = \begin{cases} 
z_i(t) + 1 & \text{if } x_{i,n}(t) = 1, \beta_i(t) = 1, \text{and} \\ y_i(t) + 1 & \text{otherwise,}
\end{cases} \quad (13)
\]
where $y_i(0) = 0$ and $\beta_i(t)$ is a binary variable that indicates whether the $i$th stream has an available packet for transmission at the beginning of time-slot $t$ or not. It is worth mentioning that, for all $i \in I$, the value of $z_i$ is reset to 0 when a new packet of the $i$th stream arrives in its queue. However, if no new packet is available at the $i$th queue, then the value of $z_i$ is linearly increased by 1 in the subsequent time-slot. Based on this, for all $i \in I$, the evolution of $z_i$ [20] is given as
\[
z_i(t + 1) = \begin{cases} 0 & \text{if } u_i(t + 1) = 1, \forall i, t, \\
z_i(t) + 1 & \text{otherwise.}
\end{cases} \quad (14)
\]
In addition, it is important to mention that, for all $i \in I$, the value of $\beta_i(t)$ changes to 0 only when the packet of stream $i$ is scheduled and successfully delivered and there is no new arrival in the same queue, i.e., $u_i(t) = 0$. Based on this, for

\[\text{all } i \in I, \text{ the evolution of } \beta_i(t) \text{ [41] can be written as:}
\]
\[
\beta_i(t + 1) = \begin{cases} 1 & \text{if } u_i(t + 1) = 1, \\
0 & \text{if } \beta_i(t)x_{i,n}(t) = 1 \land \\
\gamma_{i,n}(\Phi(t)) \geq \gamma_{th}, & \beta_i(t), \text{otherwise.}
\end{cases} \quad (15)
\]
which can be rewritten as
\[
\beta_i(t + 1) = u_i(t + 1) + \beta_i(t)(1 - x_{i,n}(t))(1 - u_i(t + 1)). \quad (16)
\]
For the sake of tractability, the AoI can be explained by the following [41]:
\[
y_i(t + 1) = 1 + x_{i,n}(t) \beta_i(t) z_i(t) + (1 - x_{i,n}(t) \beta_i(t)) y_i(t) \quad (17)
\]
\[
\gamma_{i,n}(\Phi(t)) \geq x_{i,n}(t) \beta_i(t) \gamma_{th}. \quad (18)
\]
To better understand the definition of AoI and to determine its calculation in the studied system model, an example is provided by Fig. 2, which illustrates the evolution of AoI associated with a single traffic stream and the horizontal axis represents the slot with no activity (i.e., no packet was arrived, scheduled or received at the destination).

3In this work, we considered that the transmission of each packet occupies one time-slot from BS to each destination [40].
at $t = 3$, a fresh packet $j_2$ arrives. The arrival of $j_2$ causes the $j_1$ to get discarded and resets the $z_i(t)$ to 0. However, $y_i(t)$ still increases linearly. At $t = 4$, the $j_2$ is scheduled but the delivery remained unsuccessful probably due to the channel conditions. However, another scheduling of $j_2$ at $t = 6$ resulted in a successful delivery at $t = 7$, which causes the age to drop. Afterwards, at $t = 8, t = 10$ and $t = 14$, the packets $j_2, j_4$ and $j_5$ are arrived back to back and were scheduled and delivered to the $i$th destination such that the SNR was above the threshold. Thus, the delivery of packets without errors causes the AoI to get reduced. Fig. 2 demonstrates that was above the threshold. Thus, the delivery of packets without the delivery remained unsuccessful probably due to the channel impairments in the system. Fortunately, the RIS will consider the RIS scheduling of AoI, the received signals strengths can be improved at the destination, which increases the chances of the successful delivery of the packets and ultimately helps to reduce the AoI.

V. PROBLEM FORMULATION

In this section, we leverage the communication model and the AoI definition presented in the previous sections to formulate a joint optimization of packet scheduling and RIS configuration to minimize the AoI of the system.

A. Problem Formulation

To ensure the freshness of the received information at each destination, we aim to minimize the expected sum AoI for the $I$ streams over the time horizon $T$. Let $X$ and $R$ denote the sets of the scheduling policies and the AoI configurations over the time horizon $T$, which are defined, respectively, as

$$X = \{x_{i,n}(t) | \forall t \in T, i \in I, n \in N\},$$

$$R = \{\Phi(t) | \forall t \in T\}.$$

Hence, the optimization problem can be formulated as:

$$\min_{X,R} \frac{1}{T} \mathbb{E}\left[\sum_{t=1}^{T} \sum_{i=1}^{I} y_i(t)|y_i(0) = 0\right],$$

s.t. (2), (3), (16)–(18),

$$\theta_f(t) \in [0, 2\pi), \quad \forall t \in T, f \in F,$$

$$x_{i,n}(t) \in \{0, 1\}, \quad \forall t \in T, i \in I, n \in N,$$

$$(21a)$$

$$(21b)$$

$$(21c)$$

In problem $\mathcal{OP}$, the objective function in (21a) seeks to minimize the expected sum AoI. On the other hand, constraint (2) ensures that no more than $N$ traffic streams are scheduled for transmission in a given time-slot and constraint (3) guarantees that each traffic stream is scheduled on at most one frequency channel. Moreover, constraint (16) shows the current status of the queue of each information stream at each time slot whether it is empty or has a packet available for transmission. In addition, constraints (17) and (18) ensure the correct evolution of AoI over the time horizon $T$ considering that the received SNR is above a certain threshold at each time slot.

Furthermore, constraint (21b) restrains the range of the phase shift at each RIS element. Finally, constraint (21c) ensures the binarity of the traffic streams scheduling variables over the available frequency channels at each time slot. Given the uncertainties in the arrival of packets from each traffic stream at a given time-slot, $\mathcal{OP}$ is a stochastic optimization problem over the time horizon $T$. We further observe that $\mathcal{OP}$ is a mixed-integer non-convex optimization problem which is difficult to be solved. This is due to the existence of both binary decision variables for packet scheduling and the RIS phase shift optimization. Therefore, we solve the $\mathcal{OP}$ by using the concept of bi-level optimization [42].

B. Solution Approach

In this section, we introduce our road-map to solve the joint traffic streams scheduling and RIS phase shift optimization problem towards the objective of minimizing the expected sum AoI. Leveraging the concept of bi-level optimization, the above problem is decomposed into an outer traffic stream scheduling problem and an inner phase shift matrix optimization problem. The stochastic arrival of the traffic into each stream makes the outer problem quite challenging. Hence, we resort to DRL to observe the environment and train an agent that performs scheduling. While, the inner problem of phase shift matrix optimization is solved using SDR technique. The algorithm framework that is used in this paper is depicted in Fig. 3, where the incorporated functional blocks are well illustrated to explain the whole solution approach. At the high level, the action taken by the Proximal Policy Optimization (PPO) agent gets into the scheduling decision and accordingly the phase shift matrix optimization takes place. The reward obtained at each iteration is fed back to the agent to update its policy and improve decision towards achieving the minimization objective. Next, we provide the details of these two main problems:

1) Traffic Streams Scheduling Problem: The outer problem aims to obtain the traffic stream scheduling having the RIS phase shift matrix obtained from the $\mathcal{OP}_{inner}$ problem as modelled as an MDP. A DRL based on PPO algorithm is hereby proposed to determine the policy that governs the scheduling of traffic streams. The $\mathcal{OP}_{outer}$ can be written as:

$$\min_{\mathcal{X}} \frac{1}{T} \mathbb{E}\left[\sum_{t=1}^{T} \sum_{i=1}^{I} y_i(t)|y_i(0) = 0\right],$$

s.t. (2), (3), (16), (17), (21c)

$$R = \mathcal{OP}_{inner},$$

An MDP is generally defined as a 4-tuple $(S, A, R, P)$, where: $S$ is a finite set of all possible states $s(t)$ at any time-slot $t$, where $s(t) \in S$; $A$ is a set of all feasible actions $a(t)$ at any time-slot $t$, where $a(t) \in A$; $R$ is the reward distribution, given by a measurable function $P(r(t)|s(t), a(t))$, which grants immediate reward $r(t) \in R$ after an action $a(t) \in A$ has been chosen in a state $s(t) \in S$ at time-slot $t$; $P$ is a Markovian transition model, where $P(s(t+1)|s(t), a(t)), s(t), s(t+1) \in S, a(t) \in A$ represents the probability of going from state $s(t)$ to state $s(t+1)$.
with action $a(t)$. We will next elaborate the state, action and reward functions under the MDP framework as under:

- **State $S$**: The system state at time $t$ is defined as $s(t) = (y(t), \beta(t), Z(t))$, where $s(t) \in S$. The $y(t) = (y_1(t), y_2(t), \ldots, y_I(t))$ is a vector of size $I$ containing the AoI of all traffic streams at time-slot $t$, $\beta(t) = (\beta_1(t), \beta_2(t), \ldots, \beta_I(t))$ is a vector of size $I$ containing the indicator that traffic streams have packets available for transmission and $Z(t) = (Z_1(t), Z_2(t), \ldots, Z_I(t))$ is the system time related to the $I$ streams at time slot $t$.

- **Action $A$**: An action $a(t)$ is executed at each time-slot $t$ denoted by $a(t) \in A$ consists of channel allocation decisions. The $a(t)$ is a vector of size $\alpha$, where $\alpha$ represents the number of channels to be assigned to users.

- **Reward $R$**: The immediate reward $r(t)$ at time slot $t$ is the negative summation of AoI, $r(t) = -\sum_{i=1}^{I} y_i(t)$, where $r(t) \in R$. Considering the objective of minimizing the expected sum AoI, the RL-agent aims to optimize the scheduling decision that leads to minimize the AoI.

Next, we summarize the steps of our algorithm, given as Algorithm 1, and provide a detailed explanation of how our solution framework works.

- The agent first initializes a random sampling policy and a value function for neural networks as given by line 3 and line 4 in Algorithm 1.

- At each episode, the agent observes the environment which is composed of current AoI of all the destinations and the current system time in each queue up to $t$ slot.

- At each time-slot, the agent selects an action which is a vector carrying the channels in a specific order to be mapped with the traffic streams that have a packet available for transmission.

- The action thus results to invoke the SDR (Algorithm 2) in order to configure the RIS phases shift matrix for the selected traffic streams such that the channel gain is maximized.

- The time step reward is then calculated which is the negative sum of age of information of all the streams.

- Once the set of samples have been gathered and rewards have been computed, the agent determines the advantage estimate (line 15 of Algorithm 1) which is the resultant of the difference of the expected value function from the actual reward. The advantage estimate helps the system to analyze how good it is performing based on its normal estimate function value.

The total computational complexity of DRL frameworks such as PPO algorithm can be expressed as the number of multiplications: $O(\sum_{p=1}^{P-1} n_p n_{p-1})$, based on [41], where $n_p$ is the number of neural units in the $p$-th hidden layer.

![Algorithm 1: Proposed Solution Approach for Minimizing the Expected Sum AoI](image)

1. **Input**: Number of users ($I$), Number of time-slots ($T$), Learning Rate, Episodes $K$, threshold ($\gamma_{th}$).
2. **Output**: User scheduling, Resource allocation and Phase shift matrix.
3. **Initialize policy $\pi$ with random parameter $\theta$**
4. **Initial value function $V$ with random parameters $\phi$**
5. for $k \leftarrow 1:K$ do
   6. for $t \leftarrow 1:T$ do
      7. Get $(y(t), \beta(t), Z(t))$ from the environment.
      8. sample action $a(t) \sim \pi_{\theta,old}$.
      9. Take action $a(t)$ that specifies the channels (in a specific order).
      10. Obtain the resource allocation by mapping the top $N$ traffic streams that have a packet available for transmission.
      11. Configure $\Phi(t)$ that maximizes the SNR of the mapped users to the respective channels using SDR approach using Algorithm 2.
      12. Perform the feasibility check to determine if SNR threshold constraint is satisfied.
      13. Get relevant reward $r(t)$ and $s(t+1)$.
      14. Store $(s(t), a(t), r(t), s(t+1))$ as one transition in the experience replay.
    15. Compute advantage estimate $\hat{A}$ for all epochs.
    16. Optimize surrogate loss function using Adam optimizer.
    17. Update current policy $\pi_{\theta,old} \leftarrow \pi_{\theta}$.

Fig. 3. An illustration of the proposed solution.
Algorithm 2: Design of Phase Shift Matrix via SDR

1. Input: Number of users, Number of RIS elements
2. Output: Phase shift matrix, i.e., $\Phi$
3. Initialize the maximum generation of candidate random vector as $\xi$
4. Solve the relaxed SDR problem (23a).
5. if rank($V$) = $\xi$ then
6. With the obtained $V$, calculate the eigenvalue $\omega$ and
7. eigen vector $u$ according to $Vu = \omega u$.
8. Update the value of the phase-shift matrix
   $\Phi^* := \text{diag}(\sqrt{\omega}u)$.
9. else
10. obtain the eigenvalue decomposition using Eq. (28)
11. for $x \leftarrow 1 : \xi$ do
12. Generate a Gaussian random vector $r_x$, i.e.,
   $r_x \sim \mathcal{CN}(0; I_{F+1})$
13. Obtain a candidate solution $\Theta_x$ using Eq. (29) and
   Eq. (30).
14. Find the optimal $\Theta^* := \Theta_x$ that maximizes the combined
   channel gain for all users.

2) SDR for RIS Phase Shift Coefficients: Referring to the definition of AoI given in Section IV, if no successful status update is delivered, the age for a destination will increase linearly with the time axis. Therefore, if the updated packets of a stream are scheduled by the BS but the corresponding channels do not satisfy the SNR constraints, the total AoI in $T$ time-slots will increase. Therefore, the phase shifts of the reflective elements should be configured to maximize the SNR of the channels corresponding to the selected streams. The SDR technique is applied to obtain $\Theta$ that can maximize the overall channel gain.

$$\text{OP}_{\text{inner}} : \max_\theta |h_{b\rightarrow R,n}(t)\Phi(t)h_{R\rightarrow i,n}(t) + h_{b\rightarrow i,n}(t)|^2 \quad \text{s.t.} \quad 0 \leq \theta_f(t) \leq 2\pi, \forall f \in [1, F]$$

Let us define, $v = [v_1, v_2, \ldots, v_F]^H$, where $v_f = e^{j\theta_f}$, $\forall f \in F$. By applying the change of variables, $h_{b\rightarrow R,n}(t)\Phi(t)h_{R\rightarrow i,n}(t)$ can be represented as $v^H W(t)$, where $W(t) = \text{diag}(h_{b\rightarrow R,n}(t))h_{R\rightarrow i,n}(t)$. Thus, we have

$$|h_{b\rightarrow R,n}(t)\Phi(t)h_{R\rightarrow i,n}(t) + h_{b\rightarrow i,n}(t)|^2$$

$$= |v^H W(t) + h_{b\rightarrow i,n}(t)|^2$$

An expression of overall channel gain denoted by $Z$ can be given as:

$$Z = |v^H W(t) + h_{b\rightarrow i,n}(t)|^2, \quad = v^H W(t)v + h_{b\rightarrow i,n}(t)W(t)v + v^H W(t)h_{b\rightarrow i,n}(t) + |h_{b\rightarrow i,n}(t)|^2, \quad \text{(25)}$$

The above equation can be written as follows

$$Z = \bar{v}^H \Phi \bar{v} + |h_{b\rightarrow i,n}(t)|^2, \quad \text{(26)}$$

where

$$\Phi = \begin{bmatrix} W(t)W^H(t) & W(t)h_{b\rightarrow i,n}(t) \\ h_{b\rightarrow i,n}(t)W^H(t) & 0 \end{bmatrix}$$

$$\bar{v} = \begin{bmatrix} v \\ 1 \end{bmatrix}$$

Note that $\bar{v}^H \Phi \bar{v} = \text{tr}(\Phi \bar{v} \bar{v}^H)$. Additionally, we define $V = v\bar{v}^H$, which needs to satisfy rank($V$) = 1 and $V \succeq 0$. This rank constraint (rank($v$) = 1) is non-convex [25]. By dropping this constraint, the problem OP$_{\text{inner}}$ can be rewritten as:

$$P1 : \max_\Phi Z(\Phi) \quad \text{s.t.} \quad V \succeq 0, \quad [V]_{F,F} = 1.$$ 

After the proposed transformation, the above problem can be solved by any convex optimization solver such as CVX [25]. Generally, the optimal $V$ obtained by solving problem P1 does not satisfy the rank one constraint. This implies that the optimal solution of the P1 only serves as an upper bound for the problem OP$_{\text{inner}}$. Therefore, other steps are needed to construct a rank one solution. The rank one solution is hence achieved by applying the Gaussian randomization scheme. We now describe it in detail. Firstly, we obtain the eigenvalue decomposition of $V$ as

$$V = U\Sigma U^H,$$ 

where $U = [u_1, u_2, \ldots, u_{F+1}]$ is a unitary matrix and $\Sigma = \text{diag}(\omega_1, \omega_2, \ldots, \omega_{F+1})$ is a diagonal matrix, respectively. Next, a random vector is generated as follows

$$\bar{v} = U\Sigma^{1/2}r,$$ 

where $r$ is a random vector that follows a circularly symmetric complex Gaussian (CSCG) distribution with a zero mean and a co-variance matrix equal to the identity matrix of order $F + 1$, denoted by $I_{F+1}$, i.e., $r \sim \mathcal{CN}(0; I_{F+1})$. Furthermore, we generate the scalar $v$

$$v = \exp\left[j \arg\left(\bar{v}_{1:F}^{1:F+1} \right)\right],$$

where $\bar{v}_{1:F}$ represents the vector with first $F$ elements in $v$. It is significant to highlight that the SDR approach followed by a large number of Gauss randomization can guarantee a minimum accuracy of $\pi/4$ of the optimal objective value [25]. The core details of the phase shift matrix optimization is given by Algorithm 2. Regarding the complexity of Algorithm 2, obtaining the phase-shift matrix is a semi-definite programming (SDP) problem which can be solved by the interior point method and its order of computational complexity with $m$ SDP constraints that contain an $n \times n$ positive semi-definite matrix is given as $O(\sqrt{m}\log(1/\epsilon)(mn^3 + m^2n^2 + m^3))$, where $\epsilon > 0$ is the solution accuracy [35]. The approximate computational complexity to solve SDP can be written as $O(\log(1/\epsilon)(F^{4.5} + wT_{GR}))$ with $F = F$ and $n = F + 1$. Meanwhile, let $w$ be the maximal number of generated Gaussian random vectors and $T_{GR}$ is the complexity of performing one Gaussian random iteration. Hence, the approximate complexity of obtaining phase shift matrix can be written as $O(\log(1/\epsilon)(F^{4.5} + wT_{GR}))$.
VI. SIMULATION AND NUMERICAL ANALYSIS

In this section, we present a series of simulations to evaluate the performance of the proposed algorithm. The simulation parameters are first presented, followed by the adopted benchmark schemes and then the results and discussions.

A. Simulations Setup

We consider a 3-D area where a BS is communicating with a set of spatially dispersed destinations through an RIS. We assume that the global coordinate system \((X, Y, Z)\) is Cartesian. As shown in Fig. 4, the BS is located at \((0, 0, H_B)\) and the RIS is located at \((x_i, 0, H_i)\), where \(x_i = d_{B,R}\) is the distance from the BS to the RIS, and \(H_B\) and \(H_i\) are the heights of the transmit antenna of the BS and of the RIS, respectively. In addition, multiple destinations are randomly distributed at the ground level within a given area in the network, where for all \(i \in \mathcal{I}\), the locations of destinations are \((x_i, y_i, 0)\). Precisely, based on Fig. 4, the coordinates of the \(ith\) destination, for all \(i \in \mathcal{I}\), are given by

\[
\begin{align*}
    x_i &= d_{U} \cos(\theta_i) + d_{B,U}, \\
    y_i &= d_{U} \sin(\theta_i) + d_{B,U},
\end{align*}
\]

where \(d_{U}\) is the radius of the area where the destinations are located, \(d_{B,U}\) is the distance from the BS to the center of this area and, \(\theta_i \in [\pi, 2\pi]\) is a polar angle. Unless otherwise indicated, all the simulation parameters are given by Table III, which are derived from [33], [35].

B. Benchmark Schemes

To the best of our knowledge, there is no existing approach that aims to solve the problem of minimizing the age of information in RIS-assisted wireless networks by optimizing the scheduling of traffic streams and the design of the RIS configuration considering the impacts of the stochastic arrivals of the packets and the multi-user scheduling. Thus, for the sake of comparison, we develop three other baseline schemes in order to assess the performance of the proposed scheme.

1) Greedy Scheduling With SDR (GS-SDR): In this scheme, the scheduling problem is solved using a greedy approach, whereas the RIS configuration problem is solved using the SDR approach. The greedy scheduling approach is explained as follows. At each time-slot \(t \in \mathcal{T}\), the traffic streams are first ranked based on their current AoI. The top \(N\) streams are selected to get scheduled and the RIS phase shift matrix optimization is performed to maximize the SNR of these selected streams. If the obtained SNR satisfies the given threshold, the selected streams are assumed to be scheduled and the corresponding age is calculated accordingly. However, the scheduling decisions are taken irrespective of the knowledge that the queue of the selected streams are empty or have packets to deliver. In case, if there is no status update packet in the selected stream’s queue, a time-slot is lost.

2) Round-Robin Scheduling With SDR (RRS-SDR): This algorithm is based on round-robin scheme, where at each time-slot, the BS alternately selects an input stream \(i \in \mathcal{I}\), starting from the first stream, to upload its status update packet to the destination node. The RIS configuration optimization is performed to maximize the channel gain of the scheduled streams. However, and similar to the GS-SDR baseline, the scheduling decisions are taken irrespective of the knowledge that the queue of the selected streams are empty or have packets to deliver.

3) DRL With Random Phase-Shift Matrix (DRL-RPM): In this approach, the proposed DRL algorithm is used to obtain the scheduling of the traffic streams. However, the RIS configuration is not optimally designed. Instead, a random RIS phase shift matrix is employed.
C. Results and Discussions

To get the better understanding of our proposed approach, we first investigate the behavior of the DRL agent and verify the convergence of the proposed algorithm as shown in Fig. 5. It can be seen that the cumulative reward, which is the negative of the minimum average sum AoI, will converge with the increase of number of iterations, or episodes. It can be noticed that the proposed PPO algorithm starts to converge after 3000 iterations. Similarly, we investigate the impact of varying the size of the RIS (number of elements) on the performance of different schemes as shown in Fig. 6. The impact of RIS elements is simulated by varying the number of RIS elements from 10 to 50 with a step size of 10. It can be observed from Fig. 6(a) that the integration of the RIS has a significant impact on the AoI as compared to the case when RIS is not utilized, i.e., when the direct links from the BS to destinations are solely relied on to transmit time-sensitive information. Indeed, this highlights that the RIS can significantly improve the channel quality of the scheduled users, which subsequently results in a high success rate of packets delivery. We can see that the curves of the expected sum age of information for all schemes decrease as the number of RIS elements augments. Specifically, the channel quality of the potential scheduled users can be greatly enhanced by adding more RIS elements as it enhances the chances of successful delivery at the destination and eventually ends up decreasing the AoI. From Fig. 6(b), one can remark that the proposed algorithm outperforms other benchmark approaches, e.g., when the RIS elements are 50, the expected sum AoI obtained by the proposed algorithm is around 22% lower than the one obtained by the RRS-SDR approach. This is due to the fact that the proposed PPO-based approach leverages the learning of the packet arrivals of the traffic streams and then adjusts the RIS configuration accordingly. However, the other approaches do not consider this important factor which eventually results in their worse age values.

Although the expected sum AoI of the proposed algorithm is decreased by around 35% when the number of RIS elements are increased from 10 to 50 elements, one can further note from Fig. 6 that the decrease in the AoI is not linear with the number of RIS elements, where the decrease in the AoI is not sharp when the number of elements are increased from 40 to 50, which is 3% in this case. This can be explained as increasing the number of RIS elements helps to improve the channel gains which eventually leads to satisfy the SNR threshold constraint. However, once it is satisfied, increasing number of RIS elements may not further bring the AoI down. We also observe that the GS-SDR scheme performs better than all other approaches except the proposed approach. To sum up, the greedy approach opts to schedule the streams with the worst AoI by ranking the streams against their AoI values. However, we should also point out that, the scheduling decisions are taken irrespective of whether the scheduled stream has a packet ready for transmission or not, a waste of resources occurs, which eventually lowers the efficiency of the method. Contrary to that, our proposed approach learns the arrival of packets for scheduling and thus attains better performance leveraging the informed scheduling decisions.

We next analyze the impact of a variable network load on the AoI, which is depicted in Fig. 7, where the impact of increasing the network load on the AoI is investigated. The impact of increasing the load is simulated by varying the arrival rate of the packets from 0.1 to 0.5, with a step size of 0.1. The results are plotted for the expected sum AoI versus the arrival rate. The time-horizon used for this experiment is $T = 100$ time slots. As learnt from the theory of AoI, frequent information updates along with their successful delivery results in keeping the information fresh at a destination. Precisely, a low arrival rate leads to an increase in the expected...
rate increases. These facts are validated by Fig. 7, where we pack-
tect the AoI decreases when the packets arrival properly prop-
arrive to the system and replace the old ones. Hence, under
AoI. However, as the arrival rate increases, more fresh packets
increase as the arrival rate increases. On top of this, our
proposed approach outperforms all the other
approaches.

It can be seen that as the RIS is placed neither close to
BS, nor close to destinations, the resultant AoI values start to
increase for all approaches. Once, the RIS is installed close to
the destinations, a significant improvement in terms of decreasing
AoI against the No-RIS case can be provided. Again, the
proposed approach outperforms the other baseline approaches.

As explained earlier, the proposed approach takes advantage
of the learning of the packets arrivals and also uses the SDR
to efficiently configure the phase shift matrix of the RIS in
order to maximize the SNR of the scheduled streams, which
eventually results in reducing the AoI. To conclude, a well
reasoned placement of the RIS can for sure lead to improving
the overall system performance. Finally, we also should point
out that the obtained results are in accordance with [35] which
confirms that the best location for the RIS is either besides the
BS or the users of interest.

Finally, to understand the impact of different scheduling
and phase shift optimization techniques on the AoI evolution
over time, the AoI evolution is presented in Fig. 9 for
all the algorithms. For a fair comparison, we have simulated a
system where $f = 5$ traffic streams are competing to forward
their information update packets to the destinations and 3 traffic
streams are selected to determine their AoI evolution over
time. Fig. 9 depicts that the AoI evolution is substantially
different for the different methods. It can be observed that with
the proposed approach, the AoI of all the streams is consid-
erably smaller than those of the baseline methods. This is due
to the fact that, as previously explained, the DRL agent learns
how to schedule the traffic streams with packets to transmit
such that the SNR on the selected channel is high enough to
make the transmissions successful, which eventually reduces
the AoI. However, the baseline approaches may undergo trans-
mission failures due to inefficient scheduling, which results
in packets’ loss and re-transmission by the BS that increases
the age. On the one hand, as delineated by Fig. 9(b)-(d), it
can be seen that baseline approaches significantly decrease the
AoI for some streams. Furthermore, the AoI gets significantly
increased to the maximum for other streams. This is because
(i) the RRS-SDR schedules the traffic streams in a round robin
fashion irrespective of looking at the current AoI or the arrival
time of the packets in the queue, (ii) the DRL-RPM utilizes
the RIS agent to learn to do scheduling but without a proper

![Fig. 7. Impact of arrival rate on the AoI.](image)

![Fig. 8. Impact of the position of RIS on the AoI.](image)

AoI. However, as the arrival rate increases, more fresh packets
arrive to the system and replace the old ones. Hence, under
proper propagation environment through the RIS and a proper
packets scheduling, the AoI decreases when the packets arrival
rate increases. These facts are validated by Fig. 7, where we
observe that the curves of the expected AoI for all the schemes
decrease as the arrival rate increases. On top of this, our
proposed method achieves the lowest AoI as compared to the
other methods. For example when the arrival rate is increased
from 0.1 to 0.5, the expected sum age is decreased by 70%
for the proposed method. This observation is aligned to the
principle of AoI that was discussed in previous sections, that
frequent arrival of packets directly impacts towards enhancing
the information freshness. We also notice that the GS-SDR
scheme has higher performance than the RR-SDR and the
DRL-RPM schemes even when the arrival rate is low. This
is because of its scheduling policy and RIS phase shift matrix
optimization approach. Precisely, since the GS-SDR scheme
aims to schedule the streams that give the largest decrease in
the sum AoI, it ends up getting the lower age than the other
baseline approaches.

Fig. 8 illustrates the impact of the RIS location from the
perspective of the BS and the destinations. As delineated by
Fig. 8, the distance from the BS to the RIS, $d_{B,R}$ is varied
starting by placing it next to the BS (at 1m distance) then
increasing the distance up to 200m with an increment of 50m.

Some interesting observations can be collected here. First,
since the destinations are at least 200m apart from the BS,
the direct link from the BS to each destination is expected to
undergo severe fading which will result in very high AoI val-
ues for the case where the RIS is not used. The same has been
experienced through simulations. Next, with the integration of
the RIS, the quality of the transmitted signals can be greatly
improved, which will eventually result in decreasing the AoI.
It can be seen that, when the RIS is placed in close proximity to
the BS while the destinations were at least 200m away from
the BS, the resulting AoI values were considerably low. This
is because that the direct link was not sufficient to perform
successful transmissions, and hence, RIS played its perfect
role with a well designed phase shift matrix that resulted in
lower AoI values for all the methods that had utilized RIS. On
top of this, our proposed approach outperformed all the other
approaches.

Finally, to understand the impact of different scheduling
and phase shift optimization techniques on the AoI evolution
over time, the AoI evolution is presented in Fig. 9 for
all the algorithms. For a fair comparison, we have simulated a
system where $f = 5$ traffic streams are competing to forward
their information update packets to the destinations and 3 traffic
streams are selected to determine their AoI evolution over
time. Fig. 9 depicts that the AoI evolution is substantially
different for the different methods. It can be observed that with
the proposed approach, the AoI of all the streams is consid-
erably smaller than those of the baseline methods. This is due
to the fact that, as previously explained, the DRL agent learns
how to schedule the traffic streams with packets to transmit
such that the SNR on the selected channel is high enough to
make the transmissions successful, which eventually reduces
the AoI. However, the baseline approaches may undergo trans-
mission failures due to inefficient scheduling, which results
in packets’ loss and re-transmission by the BS that increases
the age. On the one hand, as delineated by Fig. 9(b)-(d), it
can be seen that baseline approaches significantly decrease the
AoI for some streams. Furthermore, the AoI gets significantly
increased to the maximum for other streams. This is because
(i) the RRS-SDR schedules the traffic streams in a round robin
fashion irrespective of looking at the current AoI or the arrival
time of the packets in the queue, (ii) the DRL-RPM utilizes
the RIS agent to learn to do scheduling but without a proper
RIS configuration, which may not achieve the required SNR for the selected streams and results in poor performance, and (iii) despite trying to schedule the streams with the worst AoI in each time-slot and properly configuring the RIS for the selected streams, the GS-SDR is limited due to the fact that it does scheduling attempt without having any knowledge about the arrival of packets. That’s why, optimizing the RIS phase shifts alone may not guarantee that the scheduled stream would also have a packet to transmit and would increase the AoI. At the end, Fig. 9 (e) delineates the AoI evolution under the No-RIS case. It can be observed that for all the three traffic streams, the AoI increases linearly over time and no drop in AoI can be visualized. The reason being the substantial path loss and the poor quality of wireless environment.

To conclude, the proposed DRL-based approach gets highest decrease in AoI as compared to all the baseline methods. The rationale behind the improved performance of the proposed approach is its efficient scheduling policy that comprises learning of packet arrivals and eventual adjustment of RIS configuration for the traffic streams that have packets available for transmission. Imagine, if the scheduling decisions are made irrespective of the knowledge that the traffic streams have packets available for transmission or not, a waste of resource may undeniably occur by scheduling the streams that have no packets as is the case for GR-SDR approach. On the same note, if the RIS configuration is not optimally designed and instead, a random RIS phase shift matrix is employed, it may not be guaranteed that the SNR gets above the required threshold as is the case with employed RPM approach. In short, an efficient scheduling policy as well as an optimal RIS configuration design are both needed to achieve the lower AoI values, which has been confirmed through a series of simulations.

VII. CONCLUSION

In this paper, a joint design of traffic stream scheduling and RIS phase-shift matrix based on the advances in DRL technology was proposed, which strives to formulate a framework that incorporates the DRL technique into the designs for reflecting RIS-assisted wireless networks to investigate AoI optimization problem. The formulated optimization problem, which is a mixed integer non-convex optimization problem, aims to find the efficient scheduling policy that minimizes the expected sum AoI by evaluating the combined impact of stochastic packet arrivals, scheduling policy and RIS phase shift. To circumvent the high-coupled optimization variables, we decompose the original problem into an outer traffic stream scheduling problem and an inner RIS phase-shift matrix problem. For the outer problem, owing to its complexity and stochastic nature of packet arrivals, we resort to deep reinforcement learning solution where the traffic stream scheduling is modeled as an MDP, and PPO is invoked to solve it. On the other hand, the inner problem to determine the RIS configuration is solved through SDR. Numerical results demonstrate that the proposed DRL based algorithm is able to learn from the environment by observing the achieved rewards in each step and by further improving its decisions to schedule traffic streams such that the SNR on the selected channel is high enough to make the transmissions successful. It is also observed that a well reasoned placement of the RIS can definitely lead to improving the overall system performance. For future work, we can extend the current framework to consider the impact of other objective functions on age of information evolution, such as optimizing Min Max AoI, where the proposed techniques in the current work can be exploited.
