Analyzing Age Performance of Hybrid-ARQ: A Unified Explicit Result

Aimin Li, Shaohua Wu, Yajing Deng, Jian Jiao, Ning Zhang, and Qinyu Zhang.

GLOBECOM 2022 - 2022 IEEE Global Communications Conference | 978-1-6654-3540-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/GLOBECOM48099.2022.10001126

***Abstract*—In this paper, we offer an explicit, unified result that can generally depict the age performance of error-correcting techniques at the physical layer. We first propose a more realistic code-based status update system, wherein different types of delay elements, e.g., the coding delay, transmission delay, propagation delay, decoding delay and feedback delay are comprehensively considered. Under this system, we derive closed-form average Age of Information (AoI) expressions for reactive HARQ and proactive HARQ, respectively. On the basis of these explicit expressions, and utilizing the existing results for finite-length codes, we formulate an AoI minimization problem to investigate the age-optimal codeblock assignment strategy in the finite blocklength (FBL) regime. Through case studies and analytical results, we provide comparative insights between reactive HARQ and proactive HARQ from the perspective of *freshness* of information. The numerical results and optimization solutions reveal that proactive HARQ draws its strength from both superior age performance and system robustness, thus enabling the potential to provide new system advancement for a freshness-critical status update system. The full paper version of this work is available on the arXiv at https://arxiv.org/abs/2204.01257.**

***Index Terms*—5G NR, proactive HARQ, reactive HARQ, age of information, finite blocklength regime.**

I. INTRODUCTION

A. Background

In recent ten years since Kaul *et al.* proposed a framework to quantify the timeliness of information in 2012 [1], one of the most popular ideas in timely update system design has been how to keep information as fresh as possible and ensure timely information delivery. For timely update systems such as vehicle networks [2], environmental sensor systems [3], and wireless communication networks [4], achieving timely delivery can freshen the monitor's awareness of the sources and thus assist correct and efficient decision making.

This has aroused new interest in analyzing *Age of Information (AoI)* performance metrics. AoI has been broadly used to capture the *freshness* of a monitor's knowledge of an entity or process. Different from conventional performance metrics such as delay and throughput, AoI comprehensively measures the effects of update rate, latency, and system utilization. Initial works on this issue were mainly based on queue analysis, which originated from single-source single-server queues [1], [5]–[7], and subsequently developed to multiple-sources single-server queues [8]–[11] and wireless queuing networks [12]–[17]. These works are based on an ideal assumption that the status update is transmitted through a perfect channel without packet errors and losses. In practice, however, packet errors and losses are inevitable due to ubiquitous noises, signal interference, and channel fading. As the incorrectly decoded message does not bring about *fresh* awareness, the packet errors and losses will result in *staleness* of information, leading to uncontrollable residual errors, system instability, and wrong decisions. Therefore, it is imperative to analyze the AoI over unreliable channels.

B. Related Works

Some recent works have noticed this limitation and have extended the AoI analyses to the unreliable physical layer (PHY-layer). One pioneering work concerning this issue was accomplished by Chen, *et al.* in 2016 [18], in which the update is delivered over an erasure channel and the Peak Age of Information (PAoI) are studied. This work has aroused extensive research interest in understanding the effect of transmission reliability on AoI. From then on, including but not limited to the follow-up works that also analyzed AoI over the erasure channel [19]–[21], various transmission protocols, ranging from conventional protocols like non-ARQ, classical ARQ and truncated ARQ (TARQ) protocols, to state-of-the-art protocols such as HARQ with Chase Combing (HARQ-CC) and HARQ with Incremental Redundancy (HARQ-IR) protocols, have been investigated under different types of noisy channels [22]–[26].

We notice that the above AoI analyses focus on the transmission delay only, while neglecting other types of delay elements such as coding delay, propagation delay, decoding delay and feedback delay. An exception work is [27], which considers non-trivial propagation delay and studies the AoI of HARQ-IR with a fixed number of retransmitted packets $m = 2$ under Satellite-IoT Systems, but also assumes negligible coding delay and decoding delay. Nevertheless, in practical communication systems, especially the short-packet communication, the coding delay and decoding delay are also nontrivial compared to the transmission delay, resulting in the *staleness* of information by nature. In light of this, it is imperative to extend the existing PHY-layer AoI-oriented works to a more realistic (or general) setting, wherein different

---

This work has been supported in part by the National Key Research and Development Program of China under Grant no. 2020YFB1806403, and in part by the National Natural Science Foundation of China under Grant nos. 61871147, 62071141, and in part by the Shenzhen Science and Technology Program under Grant nos. GXWD20201230155427003-20200730122528002, ZDSYS20210623091808025, and in part by the Major Key Project of PCL under Grant no. PCL2021A03-1. *(Corresponding author: Shaohua Wu.)*

A. Li, S. Wu, Y. Deng, J. Jiao and Q. Zhang are with the Department of Electronic Engineering, Harbin Institute of Technology (Shenzhen), Guangdong, China. N. Zhang is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, N9B 3P4, Canada (e-mail: liaimin@stu.hit.edu.cn; hitwush@hit.edu.cn; hitdengyj@163.com; jiaojian@hit.edu.cn; zqy@hit.edu.cn; ning.zhang@uwindsor.ca).

978-1-6654-3540-6/22/$31.00 ©2022 IEEE

types of delay elements naturally exist and the number of retransmitted packets is not fixed to $m = 2$.

Moreover, it is worth noting that the majority of works respectively analyze the age performance of PHY-layer error-correcting techniques on a protocol-by-protocol basis. For instance, [22] derived explicit average AoI for non-ARQ fixed-rate codes, ARQ protocol, and TARQ protocol, respectively. [21] analyzed and optimized the average AoI for HARQ-IR protocol. Capitalizing on the obtained explicit results of different protocols, the comparative insights of AoI among various PHY-layer protocols have been demonstrated through numerical results. In this work, we try to further unify the available 'protocol-distinguished' results into a single closed-form explicit expression. This unified explicit form enables potential to reveal the comparative insights among different PHY-layer protocols by a singular explicit expression.

Up to this point, we have only introduced AoI research based on conventional *reactive* HARQ (also known as stop-and-wait HARQ), which allows for retransmissions only upon the reception of a Negative ACKnowledgment (NACK). As such, the retransmission process is not truly automatic. In the Release-16 5G NR specifications by the 3^{rd} Generation Partnership Project (3GPP), a new HARQ protocol named *proactive* HARQ is designated for the Up-Link Grant-Free communication to meet the stringent requirements for URLLC [28]. Some recent works have shown the superiority of *proactive* HARQ in terms of latency and throughput [29]–[31]. Inspiringly, these available studies also witness the potential for *proactive* HARQ to be applied in the freshness-critical status update system.

C. Contributions

Motivated by the above, this work achieves several key contributions, and we summarize them as follows:

- We derive a unified closed-form average AoI expression for *reactive* HARQ, which fulfills three-fold extensions to existing research: i) To the best of our knowledge, we are the first to comprehensively consider various kinds of delay elements that naturally exit in a realistic status update system to analyze the PHY-layer AoI performance; ii) in our framework, the number of retransmitted packets is not fixed as $m = 2$, but is relaxed to a variable parameter; iii) based on the above two points, we further unify different types of PHY-layer protocols into a single explicit expression.
- We applied our generalized framework to analyze the AoI of *proactive* HARQ as well and obtained a unified explicit result, which is the first time that *proactive* HARQ is analyzed from the AoI perspective. Theoretical and numerical comparisons are given to show the superiority of *proactive* HARQ in enabling timely information delivery.
- We further formulate an AoI minimization problem for both *reactive* HARQ and *proactive* HARQ. The problem explores the optimal number of retransmissions and their incremental lengths, which yields an age-optimal block assignment strategy. The results show that for *proactive* HARQ the finest grained symbol-by-symbol transmission strategy is always the optimal, whereas the optimal strategy for *reactive* HARQ is highly dependent on the propagation delay and SNR.

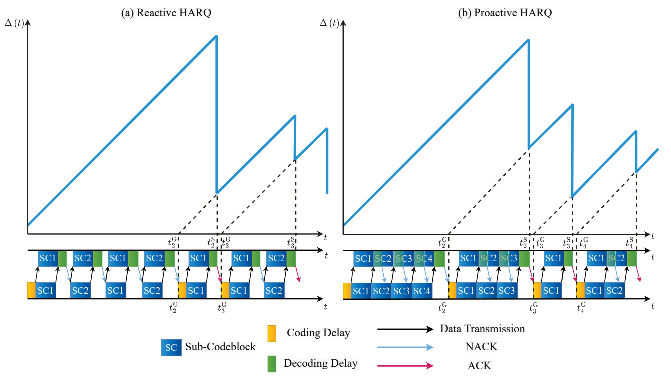

Fig. 1. Instantaneous age evolutions of reactive HARQ and proactive HARQ. Here the maximum retransmissions (the number of sub-codeblocks) is set as $m = 4$.

II. SYSTEM MODEL

We consider an end-to-end (E2E) code-based timely status update system. At the transmitter end, the update generator (source) generates a k-bit short-packet update and encodes it to a *parent codeword* with length $\sum_{i=1}^{m} \ell_i$ channel uses, which is then divided into m sub-codeblocks with length $\ell_i, i = 1, 2, \cdots, m$ and stored in a buffer waiting to be transmitted. The coding process above takes up τ_c channel uses, and we call it as coding delay. Next, the stored sub-codeblocks are transmitted over a power-limited AWGN channel sub-codeblock by sub-codeblock, with each transmission taking a transmission delay $\ell_i, i = 1, 2, \cdots, m$ channel uses. At the receiver end, we assume that the decoding process can be conducted once receiving any complete sub-codeblock. The decoding delay in each transmission round is assumed to be the same and is denoted by τ_d. If the update is decoded correctly, an ACKnowledgment (ACK) will be fed back to the transmitter; otherwise, a NACK will be sent back [1]. The feedback, similar to the forwarding information propagation, generally takes time and results in delay by nature, and we denote the delay as τ_f channel uses.

The *generate-at-will* model is adopted in the considered E2E status update system. That is, when the transmitter receives an ACK, the process of sensing and sampling will be performed, and a new update will be generated. In such a case, we mainly focus on two types of HARQ schemes: *reactive* HARQ and *proactive* HARQ. The detailed processes are shown in Fig. 1 (a) and Fig. 1 (b), respectively.

Reactive HARQ: The *reactive* HARQ scheme with maximum number of sub-codeblocks $m = 4$ is shown in Fig. 1 (a). The so-called *reactive* scheme implies that the transmitter allows for retransmissions only upon reception of a NACK. As such, the transmitter should always wait for feedback to

---

[1] In this article, we assume that the feedback is error-free. The research with erroneous feedback can be extended following this work.

decide whether to generate a new update or retransmit the old update's sub-codeblocks.

Proactive HARQ: The *proactive* HARQ scheme with maximum sub-codeblocks $m = 4$ is shown in Fig. 1 (b). As its name indicates, the retransmission process is completely spontaneous and consecutive, which is interrupted only when an ACK is received. By *proactive* retransmitting, the latency caused by waiting for feedback is reduced, and thus the issue of long HARQ round trip time (RTT) is resolved.

Definition 1. ***(AoI)*** *Denote* t_i^{G} *as the generation time instant of the* i^{th} *status update packet that can be correctly decoded, and denote* t_i^{S} *as the time instant at which this packet is correctly decoded. At a time instant* τ, *the index of the most recently generated update can be given by* $N(\tau) = \max\left\{i | t_i^{\mathrm{S}} \leq \tau\right\}$ *and the time stamp is* $U(\tau) = t_{N(\tau)}^{\mathrm{G}}$. *Then, the instantaneous AoI is defined as* $\Delta\left(t\right) \triangleq t - U\left(t\right)$.

We focus on AoI analysis and optimization in this paper. Fig. 1 shows the instantaneous age evolutions of reactive HARQ and proactive HARQ. We consider the discrete symbol-level analysis, where the time is divided into some time slots in units of channel use. In this manner, the average AoI is defined by $\Delta \triangleq \lim\limits_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \Delta\left(t\right)$.

III. ANALYTICAL RESULTS

In this section, we study the symbol-level AoI of reactive HARQ and proactive HARQ. We first give the closed-form expressions for the AoI in **Proposition 1** and **Proposition 2**, and then conduct a theoretical AoI comparison between the two considered transmission protocols in **Corollary** 1. The AoI expressions, given in (1) and (2), are functions of the block assignment vector $\mathbf{n}$ and its dependent error probability vector $\mathbf{e}$, where the element n_i in vector $\mathbf{n}$ denotes the number of cumulative transmitted symbols up to the i^{th} transmission round with $n_i = \sum_{j=1}^{i} \ell_j$, and the element ϵ_i in vector $\mathbf{e}$ denotes the probability that the i^{th} re-transmitted message remains incorrectly decoded.

By flexible design of the vector $\mathbf{n}$ and the vector $\mathbf{e}$, we also demonstrate that the derived expression for reactive transmission protocol also unifies the available AoI analyses conducted in [22]. Moreover, by using the result of the achievable rate of finite-length codes, we obtain the AoI closed-form expression under the FBL regime.

A. Reactive Scheme

1) Average AoI

Proposition 1. ***(The Generalized Closed-form Average AoI Expression for Reactive HARQ)*** *For reactive HARQ with maximum retransmissions* m, *block assignment vector* $\mathbf{n} = (n_1, n_2, \cdots, n_m)$ *and error probability vector* $\mathbf{e} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_m)$, *the average AoI can be calculated by*

$$\Delta_{\mathrm{Reactive}} = -\tfrac{1}{2} - \tau_{\mathrm{f}} + \frac{\tau_{\mathrm{c}}+n_1+\mathcal{T}+\sum\limits_{i=1}^{m-1}(n_{i+1}-n_i+\mathcal{T})\epsilon_i}{1-\epsilon_m} + \frac{(\tau_{\mathrm{c}}+n_1+\mathcal{T})^2+\sum\limits_{i=1}^{m-1}(n_{i+1}-n_i+\mathcal{T})(2\tau_{\mathrm{c}}+n_{i+1}+n_i+(2i+1)\mathcal{T})\epsilon_i}{2\left(\tau_{\mathrm{c}}+n_1+\mathcal{T}+\sum\limits_{i=1}^{m-1}(n_{i+1}-n_i+\mathcal{T})\epsilon_i\right)}, \tag{1}$$

where $\mathcal{T} = \tau_{\mathrm{f}} + \tau_{\mathrm{p}} + \tau_{\mathrm{d}}$ *with* τ_{c}, τ_{p}, τ_{d} *and* τ_{f} *denoting the coding delay, propagation delay, decoding delay and feedback delay, respectively.*

Proof. Please see Section III. A in [32] for detailed derivations. ☐

2) Case Study: A Unified Result

With Proposition 1 in hand, we can conduct some case studies by flexibly considering the choices of the block assignment vector $\mathbf{n}$ and the error probability vector $\mathbf{e}$. By this means, we theoretically show that the closed-form AoI expressions given in this paper is a unified result. Though the given examples are not exhaustive in this paper, we can observe from these case studies that the unified expression given in (1) enables potential for exploring the intrinsic relationship and comparative insights among different types of transmission protocols.

Case 1. ***(Average AoI for non-ARQ Fixed-Rate Codes)*** *We show that the available average AoI expression for fixed-rate codes in [22] is a specific case of our unified result in (1). For fixed-rate codes without ARQ, the maximum retransmissions turns to* $m = 1$. *Substitute* $m = 1$ *into (1) and remove the effect of delay elements such that* $\tau_{\mathrm{c}} = \tau_{\mathrm{p}} = \tau_{\mathrm{d}} = \tau_{\mathrm{f}} = 0$, *we can obtain the average AoI as the Proposition 1 in [22]:*

$$\bar{\Delta}_{\mathrm{Non-ARQ}} = -\frac{1}{2} + \frac{n_1}{1-\epsilon_1} + \frac{n_1}{2},$$

where n_1 *is the code length and* ϵ_1 *is the error probability of the fixed-rate codes.*

Case 2. ***(Average AoI for TARQ)*** *We demonstrate that the average AoI expression for TARQ is also a specific case of our unified result in (1). For truncated ARQ, the transmitter retransmits the same packet till the allowable maximum retransmissions* m *is reached or this packet is successfully received. Since the retransmitted packet is the same as the first packet, the cumulative transmitted message length is* $n_i = in_1$ *and the corresponding error probability is* $\epsilon_i = {\epsilon_1}^i$. *Then, by substituting them back into (1) and similarly remove the effect of delay elements, we can obtain the average AoI as the Proposition 3 in [22]:*

$$\bar{\Delta}_{\mathrm{TARQ}} = -\frac{1}{2} + n_1\left(\frac{2}{1-\epsilon_1} - \frac{1}{2} - \frac{m\epsilon_1^m}{1-\epsilon_1^m}\right).$$

Case 3. ***(Average AoI for Classical ARQ)*** *We also find that the average AoI expression for TARQ is a specific case of our unified result in (1). For classical ARQ, the transmitter re-transmits the same packet till the packet is successfully* *received, while the maximum retransmissions is not limited. The classical ARQ is a special case of TARQ where $m \to \infty$. Then, by calculating the limit $\lim_{m\to\infty} \Delta_{\mathrm{TARQ}}$, we can obtain the average AoI as the Proposition 2 in [22]:*

$$\bar{\Delta}_{\mathrm{Classical-ARQ}} = -\frac{1}{2} + n_1\left(\frac{2}{1-\epsilon_1} - \frac{1}{2}\right).$$

Case 4. ***(Average AoI for HARQ-IR)*** *By similarly removing the effect of delay elements, we find that the result in (1) turns to a variant of that in [21], given as:*

$$\bar{\Delta}_{\mathrm{HARQ-IR}} = -\tfrac{1}{2} + \frac{n_1+\sum\limits_{i=1}^{m-1}(n_{i+1}-n_i)\epsilon_i}{1-\epsilon_m} + \frac{n_1^2+\sum\limits_{i=1}^{m-1}\left(n_{i+1}^2-n_i^2\right)\epsilon_i}{2\left(n_1+\sum\limits_{i=1}^{m-1}(n_{i+1}-n_i)\epsilon_i\right)}.$$

B. Proactive HARQ

By leveraging our methodology from *reactive* HARQ, we analyze the AoI of *proactive* HARQ as well and obtained a unified explicit result, as shown in the following.

Proposition 2. ***(The Generalized Closed-form Average AoI Expression for Proactive HARQ)*** *For proactive HARQ with maximum retransmissions m, block assignment vector $\mathbf{n} = (n_1, n_2, \cdots, n_m)$ and error probability vector $\mathbf{e} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_m)$, the average AoI can be calculated by*

$$\bar{\Delta}_{\mathrm{Proactive}} = -\tfrac{1}{2} - \tau_{\mathrm{f}} + \frac{\tau_{\mathrm{c}}+n_1+\mathcal{T}+\sum\limits_{i=1}^{m-1}(n_{i+1}-n_i)\epsilon_i}{1-\epsilon_m} + \frac{(\tau_{\mathrm{c}}+n_1+\mathcal{T})^2+\sum\limits_{i=1}^{m-1}(n_{i+1}-n_i)(2\tau_{\mathrm{c}}+2\mathcal{T}+n_{i+1}+n_i)\epsilon_i}{2\left(\tau_{\mathrm{c}}+n_1+\mathcal{T}+\sum\limits_{i=1}^{m-1}(n_{i+1}-n_i)\epsilon_i\right)}. \tag{2}$$

Proof. Please see Section III.B in [32] for detailed derivations. ☐

C. Reactive HARQ vs. Proactive HARQ

Corollary 1. ***(Reactive HARQ vs. Proactive HARQ)*** *The average age performance of reactive HARQ does not exceed that of proactive HARQ given the same block assignment vector $\mathbf{n}$ and the same error probability vector $\mathbf{e}$. The necessary and sufficient condition for their equivalence is $m = 1$ or $\mathcal{T} = 0$.*

Proof. Please see Appendix D in [32] for detailed derivations. ☐

D. Average Age in the FBL Regime

In addition, the given generalized expressions also allow us to adopt the FBL results in [33] to evaluate the AoI performance of the considered HARQ protocols. Over the power-limited AWGN channel with SNR γ, and error probability ϵ_i can be approximated by the Theorem 54 of [33] as:[2]

$$\epsilon_i \approx Q\left(\frac{C(\gamma) - k/n_i - \frac{\log_2 n_i}{2n_i}}{\sqrt{V(\gamma)/n_i}}\right). \tag{3}$$

where C is the channel capacity with $C = \log_2(1+\gamma)$ and V is the channel dispersion with $V = (1 - \frac{1}{(1+\gamma)^2})\log_2^2 e$.

Substitute Eq. (3) into Eq. (1) and Eq. (2), and we obtain the AoI closed-form expressions in the FBL regime.

IV. AGE-OPTIMAL BLOCK ASSIGNMENT

A. Problem Formulation

We establish an average AoI minimization problem here to further explore the age-optimal transmission mechanism in the FBL regime with non-trivial delay:

1) Objective function: To minimize the average age Δ.

2) Decision variable: The block assignment vector $\mathbf{n} = (n_1, n_2, \cdots, n_m)$.

Problem 1. *Age-optimal block assignment for reactive HARQ (or proactive HARQ)*

$$\begin{aligned} \min_{\mathbf{n}} \quad & \bar{\Delta}_{\mathrm{Reactive}} \quad \text{or} \quad \bar{\Delta}_{\mathrm{Proactive}} \\ \textit{s.t.} \quad & c_1 : n_{\min} \le n_1 < n_2 < \ldots < n_m \le n_{\max}, \\ & c_2 : 1 \le m \le n_{\max} - n_{\min} + 1, \\ & c_3 : \epsilon_i = Q\left(\frac{C(\gamma) - k/n_i - \frac{\log_2 n_i}{2n_i}}{\sqrt{V(\gamma)/n_i}}\right), \\ & c_4 : m, n_i \in \mathbb{Z}^+, i = 1, \ldots, m. \end{aligned} \tag{4}$$

Note that the decision variable $\mathbf{n}$ is a variable-length vector with infinite solution space, we introduce $n_{\min}$ and $n_{\max}$ as constraints of the solution space, which denotes the lower bound and the upper bound of the range of block length, respectively.

B. Solutions and Discussions

Problem 1 is a nonlinear integer problem. To solve the optimal solution of **Problem 1**, an auxiliary vector $\mathbf{p} \in \mathcal{S} \triangleq \{0,1\}^{n_{\max}-n_{\min}+1}$ can be introduced here.

Lemma 1. *There exists a one-to-one mapping between vectors $\mathbf{n}$ and $\mathbf{p}$.*

Proof. Please see Appendix E of [32] for the detailed proof. ☐

Lemma 1 illustrates that the auxiliary vector $\mathbf{p}$ can be regarded as an index of the solution space of Problem 1, which can help us traverse the entire solution space efficiently and find the optimal solution. The detailed algorithm process is provided in Algorithm 1.

Lemma 1 helps us to efficiently traverse the entire solution space and find the optimal solution. The detailed algorithm process is provided in Algorithm 1.

Fig. 2. gives some detailed examples of the solved optimal block assignment vector $\mathbf{n}_{\mathrm{optimal}}$ under different protocols, SNRs, and propagation delays. For example, under SNR$= 0.7$ dB and $\tau_{\mathrm{p}} = 0$, the optimal block assignment vector for reactive HARQ is $\mathbf{n}_{\mathrm{optimal}} = (105, 120)$; under SNR$= 1.9$ dB and $\tau_{\mathrm{p}} = 20$, the optimal block assignment vector for reactive HARQ is $\mathbf{n}_{\mathrm{optimal}} = (100, 112, 120)$.

Fig. 2. leads to the following conclusions:

- For proactive HARQ, the finest grained symbol-by-symbol strategy minimizes the average AoI.

---

[2] Here we focus on FBL analysis as a case study. Note that ϵ_i can also be characterized by other specific error-correcting techniques.

Algorithm 1: The algorithm for solving Problem 1.

```
Input: The signal-to-noise ratio $\gamma$; The message length $k$; The lower bound of the range of block length $n_{\min}$; The upper bound of the range of block length $n_{\max}$; The system delay $\tau_c$, $\tau_p$, $\tau_d$ and $\tau_f$;
Output: The optimal block assignment vector $\mathbf{n}_{\text{optimal}}$; The minimum average age $\Delta_{\min}$;
Initialization: $\Delta_{\min} = \infty$; $\mathcal{S} = \{0,1\}^{n_{\max}-n_{\min}+1}$;
for $\mathbf{p}$ in $\mathcal{S}$ do
    Map vector $\mathbf{p}$ to the block assignment vector $\mathbf{n}$;
    According to the obtained $\mathbf{n}$, calculate the average age $\Delta$ by using Eq. (1) or Eq. (2) ;
    if $\Delta < \Delta_{\min}$ then
        Update $\Delta_{\min} = \Delta$;
        Update $\mathbf{n}_{\text{optimal}} = \mathbf{n}$;
return $\mathbf{n}_{\text{optimal}}$ and $\Delta_{\min}$
```

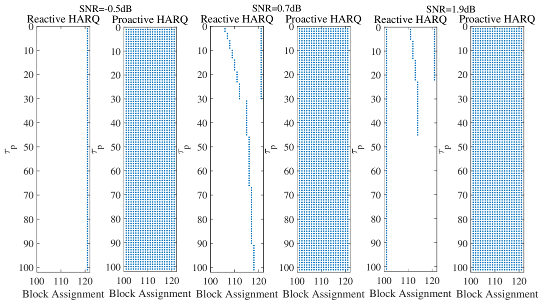

Fig. 2. Age-optimal block assignment of reactive HARQ and proactive HARQ. Here the length of message is $k = 100$, the minimum code length $n_{\min} = 100$, the maximum code length $n_{\max} = 120$, the coding delay is $\tau_c = 20$, the decoding delay is $\tau_d = 30$.

- For reactive HARQ, the age-optimal block assignment varies among different SNR and τ_p. As the propagation delay increases, the number of retransmissions will monotonically decrease and finally converge to $m = 1$, and the transmission scheme turns out to be non-ARQ. This indicates that there exists a threshold of the propagation delay, only within which retransmission is beneficial to AoI.
- From a perspective of channel coding, we can see that the trade-off between reliability and effectiveness can be well evaluated by the new metric, AoI. It is well known that a longer code length can improve the reliability while sacrificing the effectiveness; however, what is not fully explored is that an appropriate choice of code length can minimize the AoI.

V. NUMERICAL RESULTS

Fig. 3 demonstrates the average AoI comparison between reactive HARQ and proactive HARQ. The comparisons are conducted among different settings of m and τ_p. It is shown that, intuitively, the proactive HARQ surface remains below the reactive HARQ surface. Also, they intersect with each other at $m = 1$. These numerical results are consistent with **Corollary 1**. In addition, Fig. 3 also illustrates the impact that τ_p and m exert on average AoI. On the one hand, the average AoI is monotonically increasing with respect to the propagation delay τ_p. On the other hand, the impact of m on average AoI could be complex: i) for proactive HARQ, retransmitting redundancy remains beneficial for AoI performance metric; ii) for reactive HARQ, however, retransmitting redundancy naturally brings about RTT and thus results in staleness of information when the SNR is high enough to achieve reliable communication; in contrast, if the channel condition is poor, retransmitting redundancy is essential for reliable delivery, and thus may even compensate the AoI losses due to RTT.

Fig. 4 holds up a comparison among conventional reactive HARQ, optimal reactive HARQ, and proactive HARQ. It is shown that the optimized reactive HARQ approaches proactive HARQ in average AoI performance. The reason is that the age-optimal m for reactive HARQ tends to remain small (see Fig. 2), which eliminates the need for extra RTT and thus achieve age performance close to that of finest grained proactive HARQ. However, we notice that this gain lies in

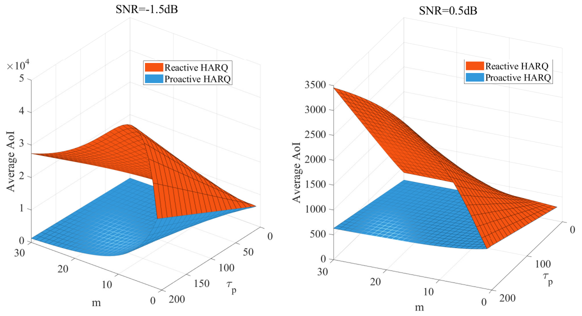

Fig. 3. Reactive HARQ vs proactive HARQ. Here the message length is $k = 100$, the encoding delay is $\tau_c = 20$, the decoding delay is $\tau_c = 30$, the propagation delay range is $\tau_p = 0 : 200$ and the feedback delay is calculated by $\tau_f = \tau_p + 1$. For fairness, here we consider the finest-granularity block assignment vectors for both reactive HARQ and proactive HARQ, with $n_1 = 100, n_i = n_{i-1} + 1$.

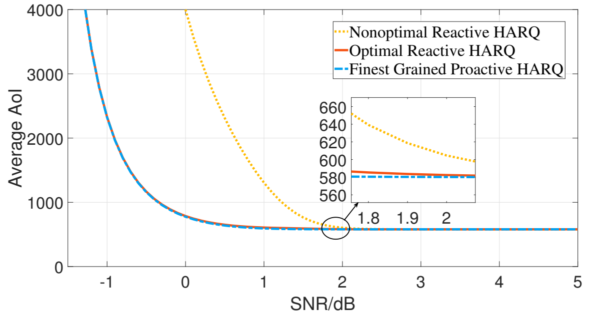

Fig. 4. Average age comparison among conventional reactive HARQ, optimal reactive HARQ and finest grained proactive HARQ. Here the message length is $k = 100$, the encoding delay is $\tau_c = 20$, the decoding delay is $\tau_c = 200$, the propagation delay is $\tau_p = 50$ and the feedback delay is $\tau_f = 51$.

an adaptive block assignment strategy which requires accurate channel status information. In this regard, we find that adopting proactive HARQ for freshness-critical status update systems would be a robust and timeliness-efficient approach.

VI. CONCLUSIONS

In this paper, we comprehensively consider various types of nontrivial system delay and derive generalized closed-form average AoI expressions for reactive HARQ and proactive HARQ. Through comparisons and optimizations, the finest-granularity proactive HARQ remains superior to conventional reactive HARQ in terms of both age performance and system robustness, indicating the potential of proactive HARQ to be applied in the status update system.

The research in this paper also leaves some open issues for future research. First, it would be interesting to conduct AoI analyses and comparisons for some specific state-of-the-art channel coding techniques, such as polar codes, LDPC codes, Turbo codes, and rateless Raptor codes, etc. As such, from the AoI perspective, the trade-off among coding complexity, decoding complexity, codelength, the number of retransmitted packets and error probability can be explored. Second, since this work is based on an ideal assumption of perfect feedback, the analysis considering lossy feedback can be further conducted. Third, notwithstanding the AoI superiority of proactive HARQ compared to reactive HARQ, proactive HARQ may consume more energy due to the consecutive retransmissions. To this end, to further investigate the trade-off of proactive HARQ between timeliness and energy efficiency would be an interesting topic.

REFERENCES

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

[2] P. Papadimitratos, A. D. La Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, “Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation,” *IEEE Communications Magazine*, vol. 47, no. 11, pp. 84–95, 2009.

[3] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valencia, and D. Moore, “Environmental wireless sensor networks,” *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1903–1917, 2010.

[4] A. A. Reddy, S. Banerjee, A. Gopalan, S. Shakkottai, and L. Ying, “On distributed scheduling with heterogeneously delayed network-state information,” *Queueing Systems*, vol. 72, no. 3, pp. 193–218, 2012.

[5] M. Costa, M. Codreanu, and A. Ephremides, “On the Age of Information in Status Update Systems With Packet Management,” *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1897–1910, 2016.

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

[7] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, “Update or Wait: How to Keep Your Data Fresh,” *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[8] R. D. Yates and S. K. Kaul, “The Age of Information: Real-Time Status Updating by Multiple Sources,” *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1807–1827, 2019.

[9] L. Huang and E. Modiano, “Optimizing age-of-information in a multi-class queueing system,” in *2015 IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 1681–1685.

[10] E. Najm and E. Telatar, “Status updates in a multi-stream M/G/1/1 preemptive queue,” in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2018, pp. 124–129.

[11] M. Moltafet, M. Leinonen, and M. Codreanu, “On the Age of Information in Multi-Source Queueing Models,” *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 5003–5017, 2020.

[12] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, “Effect of Message Transmission Path Diversity on Status Age,” *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1360–1374, 2016.

[13] A. M. Bedewy, Y. Sun, and N. B. Shroff, “Minimizing the Age of Information Through Queues,” *IEEE Transactions on Information Theory*, vol. 65, no. 8, pp. 5215–5232, 2019.

[14] ——, “The Age of Information in Multihop Networks,” *IEEE/ACM Transactions on Networking*, vol. 27, no. 3, pp. 1248–1257, 2019.

[15] E. Najm, R. Nasser, and E. Telatar, “Content Based Status Updates,” *IEEE Transactions on Information Theory*, vol. 66, no. 6, pp. 3846–3863, 2020.

[16] S. K. Kaul and R. D. Yates, “Age of Information: Updates with Priority,” in *2018 IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 2644–2648.

[17] R. D. Yates, “Status Updates through Networks of Parallel Servers,” in *2018 IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 2281–2285.

[18] K. Chen and L. Huang, “Age-of-information in the presence of error,” in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 2579–2583.

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

[20] E. Najm, E. Telatar, and R. Nasser, “Optimal Age over Erasure Channels,” in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 335–339.

[21] S. C. Bobbili, P. Parag, and J.-F. Chamberland, “Real-Time Status Updates With Perfect Feedback Over Erasure Channels,” *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5363–5374, 2020.

[22] M. Xie, Q. Wang, J. Gong, and X. Ma, “Age and Energy Analysis for LDPC Coded Status Update With and Without ARQ,” *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 10 388–10 400, 2020.

[23] J. You, S. Wu, Y. Deng, J. Jiao, and Q. Zhang, “An Age Optimized Hybrid ARQ Scheme for Polar Codes via Gaussian Approximation,” *IEEE Wireless Communications Letters*, vol. 10, no. 10, pp. 2235–2239, 2021.

[24] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor, “On Timely Channel Coding with Hybrid ARQ,” in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[25] B. Yu, Y. Cai, D. Wu, and Z. Xiang, “Average Age of Information in Short Packet Based Machine Type Communication,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 10 306–10 319, 2020.

[26] R. Devassy, G. Durisi, G. C. Ferrante, O. Simeone, and E. Uysal-Biyikoglu, “Delay and Peak-Age Violation Probability in Short-Packet Transmissions,” in *2018 IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 2471–2475.

[27] D. Li, S. Wu, Y. Wang, J. Jiao, and Q. Zhang, “Age-Optimal HARQ Design for Freshness-Critical Satellite-IoT Systems,” *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 2066–2076, 2020.

[28] *Discussion on Explicit HARQ-ACK Feedback for Configured Grant Transmission*, document R1-1903079, 3GPP TSG RAN WG1 96, 2019.

[29] R. Abreu, T. Jacobsen, G. Berardinelli, K. Pedersen, I. Z. Kovács, and P. Mogensen, “Power control optimization for uplink grant-free URLLC,” in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1–6.

[30] N. H. Mahmood, R. Abreu, R. Böhnke, M. Schubert, G. Berardinelli, and T. H. Jacobsen, “Uplink Grant-Free Access Solutions for URLLC services in 5G New Radio,” in *2019 16th International Symposium on Wireless Communication Systems (ISWCS)*, 2019, pp. 607–612.

[31] Y. Liu, Y. Deng, M. Elkashlan, A. Nallanathan, and G. K. Karagiannidis, “Analyzing Grant-Free Access for URLLC Service,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 741–755, 2021.

[32] A. Li, S. Wu, J. Jiao, N. Zhang, and Q. Zhang, “Age of information with hybrid-arq: A unified explicit result,” *arXiv preprint arXiv:2204.01257*, 2022.

[33] Y. Polyanskiy, H. V. Poor, and S. Verdu, “Channel Coding Rate in the Finite Blocklength Regime,” *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.