Age-Optimal Sampling and Routing under Intermittent Links and Energy Constraints

A. Utku Atasayar, Aimin Li, *Member, IEEE*, Çağrı Arı, and Elif Uysal, *Fellow, IEEE*

***Abstract*—Links in practical systems, such as satellite–terrestrial integrated networks, exhibit distinct delay distributions, intermittent availability, and heterogeneous energy costs. These characteristics pose significant challenges to maintaining timely and energy-efficient status updates. While link availability restricts feasible transmission routes, routing decisions determine the actual delay and energy expenditure. This paper tackles these challenges by jointly optimizing sampling and routing decisions to minimize monotonic, non-linear Age of Information (AoI). The proposed formulation incorporates key system features, including multiple routes with correlated random delays, stochastic link availability, and route-dependent energy consumption. We model the problem as an infinite-horizon Constrained Semi-Markov Decision Process (CSMDP) with a hybrid state–action space and develop an efficient nested algorithm, termed Bisec-ReaVI, to solve this problem. We reveal a well-defined jointly optimal policy structure: (i) The optimal routing policy is a monotonic handover policy that adapts to the availability of routes and their mean delays; and (ii) The optimal sampling policy is a piecewise linear waiting policy, with at most $\binom{N}{2}+N$ breakpoints given N routes. Numerical experiments in a *satellite–terrestrial* integrated routing scenario demonstrate that the proposed scheme efficiently balances energy usage and information freshness, and reveal a counter-intuitive insight: *even routes with higher average delay, higher delay variance or lower availability can still play a critical role in minimizing monotonic functions of AoI*.**

I. Introduction

A. Background

In an increasingly connected world where systems rely on remotely sampled data to make real-time decisions, the freshness of data samples has become a key driver of application performance. Hence, information freshness is emerging as a Key Performance Indicator (KPI) across a wide range of applications, supported by next-generation communication networks spanning wired, wireless, and non-terrestrial links. For instance, in remote-sensing-based emergency response systems, access to *fresh* data regarding environmental variables supports real-time risk assessment and enhances response efficiency. Similarly, in Vehicle to Everything (V2X) scenarios, vehicles rely on continuously updated information collected through multi-sensor data fusion to navigate safely and adapt to rapidly changing environmental conditions. Moreover, in Industrial Internet of Things (IIoT) applications, the staleness of sensor data negatively impacts production efficiency, equipment maintenance timing, and overall operational safety.

This growing emphasis on information freshness has led to the development of the Age of Information (AoI) metric that quantifies it [2]. Distinct from traditional latency, AoI provides a receiver-centric, flow-level measure of information timeliness. Formally, at any time t, the AoI is defined as $\Delta(t) \triangleq t - U(t)$, where $U(t)$ denotes the generation time of the latest received sample. Maintaining a low AoI requires both sufficiently frequent updates and low-latency delivery, thus coupling throughput and delay in a novel performance metric. Hence, AoI combines the conventional performance metrics of latency and throughput in a novel way. Over the past few years, AoI minimization has been studied under various constraints and network models, including energy-limited update policies [3]–[5], multi-hop and multi-source networks [6]–[10], and broadcast settings [11]. Furthermore, in scenarios where the significance of data extends beyond temporal freshness, more sophisticated evaluation frameworks have been developed [12]–[18]. Some of these frameworks utilize AoI as an intermediate metric to capture task-specific relevance through the freshness of data samples [17].

B. Motivation

In this paper, we focus on the optimization of AoI. Our goal is to extend the formulation of Age-optimal sampling first proposed in [19] to a case where there are multiple routing options between the source and the destination. The new formulation proposed in this paper was inspired and motivated by the growing interest in Satellite IoT and integrated TN-NTN in 5G and 6G, where data transmission decisions are sometimes faced with a choice between routing through non-terrestrial links versus terrestrial connections:

(i) Routing through **terrestrial links** typically offers low-latency and energy-efficient transmission, owing to the relatively short propagation distance and mature ground-based infrastructure such as optical fiber and cellular networks. These links are generally stable under normal operating conditions and can support high-throughput, delay-sensitive services. However, their performance and reliability are susceptible to *network congestion*, *coverage holes*, and *infrastructure failures.*

(ii) Routing through **non-terrestrial links** often exhibits

An earlier version of this work was presented in part at the ACM International Symposium on Mobile Ad Hoc Networking and Computing (ACM MobiHoc) 2025 [1].

This work is supported by the European Union through ERC Advanced Grant 101122990-GO SPACE-ERC-2023-A. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. *(Corresponding author: Elif Uysal).*

A. Utku Atasayar, Aimin Li, Çağrı Arı, and Elif Uysal are with Communication Networks Research Group (CNG), EE Dept, Middle East Technology University (METU), Ankara, 06800, Turkiye (e-mail: atasayar.utku@metu.edu.tr; liaimingogo@gmail.com; ari.cagri@metu.edu.tr; uelif@metu.edu.tr).

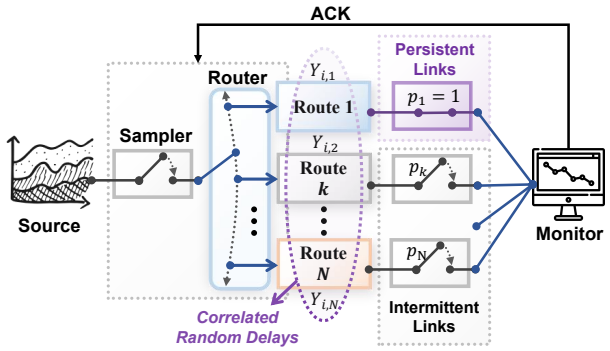

Fig. 1. A remote monitoring system, where status updates are transmitted through N heterogeneous routes.

intermittent availability because the space–atmosphere channel and satellite geometry change over time. In addition to orbital dynamics and visibility windows, beam/footprint handovers, gateway reassociation, and ISL route reconfiguration introduce short disruptions that fragment service into available/unavailable epochs. Propagation conditions vary with rain attenuation (Ku/Ka), cloud ice/water content, ionospheric irregularities, and geomagnetic storms, which can drive rapid SNR fluctuations and temporary outages. Even during physical available periods, random-access collisions, beam scheduling, and backhaul bottlenecks can make access bursty.

Moreover, delay statistics across routes may be *correlated* due to shared infrastructures, spectrum usage, or satellite visibility patterns, which further couples the sampling and routing decisions. These practical considerations result in route-dependent delay statistics, stochastic link availability, and heterogeneous energy costs, which motivates a unified framework that jointly optimizes both the sampling time and the routing path. Our formulation addresses the fundamental question of how route selection impacts information freshness in hybrid terrestrial/non-terrestrial environments, and provides a theoretical foundation for age-aware joint sampling and routing in next-generation communication networks.

C. Related Works and Contributions

• **System Model**: We formulate a joint sampling and routing problem in which a transmitter optimizes both the *sampling interval* and the *transmission route* to minimize the *long-term average AoI* at the destination, subject to an average routing energy constraint and route availiability constraint. Our formulation is a direct extension of the problem in [19], and in contrast to this and other prior work where the sampling problem is attacked under given delay statistics [14]–[16], [20], our system model takes a *proactive* approach by actively selecting and switching routes to control the delay experienced by status updates. Meanwhile, unlike existing multi-channel scheduling problems that either focus on homogeneous channels with uniform one-slot delays [21] or heterogeneous channels where each channel experiences an on-off constant discrete delay [22], [23], our work considers distinct *continuous delay distributions* across different routes and jointly optimizes sampling and route selection strategies. This works generalizes our prior work in [1] by incorporating the following practical aspects: (i) *Correlated route delays*, motivated by shared infrastructures or congestion effects that couple delay statistics across different routes; (ii) *Stochastic route availability*, modeled as independent random on/off links with fixed availability probabilities, capturing real-world uncertainties such as satellite visibility, atmospheric attenuation, and space-weather-induced disruptions and (iii) *Route-dependent energy consumption* under an average energy constraint, which is essential for energy-limited platforms such as satellites or IoT devices.

• **Solution Methodology**: We show the problem can be formulated by a Constrained Semi-Markov Decision Process (CSMDP) with *uncountable hybrid* state and action spaces. The state captures discrete link availability and continuous delay, while the action includes both routing choices and sampling intervals. Such CSMDPs are known to be challenging to solve due to the size of the state and action spaces, and previous research has addressed this complexity by: (i) discretizing the *uncountable* state space and the action space [24], [25], which introduces *quantization error* or; (ii) focusing on a special case of SMDP where the state transitions are independent of actions[1] [14], [19], [20], which, however, does not hold in our scenario. In this paper, we develop a new nested algorithm namely Bisection ***R**elative **E**xpected **A**ction **V**alue **I**teration* (BISEC-REAVI) that directly solves this class of SMDPs without discretizing the space. *To the best of our knowledge, this is the first algorithm that efficiently solves hybrid-state CSMDPs while preserving structural optimality and avoiding discretization error.*

• **Structural Results**: We prove that each of the jointly optimal sampling and routing policies exhibit a graceful *threshold structure*: (i) The routing policy is a thresholds-based handover policy, where a specific route is selected when the current AoI at the receiver falls within certain range, precisely determined by multiple thresholds; (ii) A new sample is taken and transmitted when the AoI at the receiver reaches a threshold that is a function of the selected transmission route. These structural properties deem the policies suitable for practical implementation. We designed an efficient algorithm to determine these thresholds.

• **Counter-Intuitive Insights:** We test our algorithms on the model of an *integrated satellite-terrestrial* communication network scenario. Our simulation results reveal an intriguing insight: routes with higher mean delay, greater variance, or lower availability can still contribute to minimizing AoI. This finding challenges conventional wisdom that may prioritize routes with minimal mean delay or delay variance characteristics. It demonstrates that the strategic utilization of diverse routing options in complex network environments can lead to superior information freshness.

[1]This allows the Markov decision process to be reduced to a renewal reward process, thus simplifying the analysis.

II. System Model

We consider a remote monitoring system, as illustrated in Fig. 1, consisting of a source, a sampler, a router, and a monitor. Status updates are timely generated, and each is transmitted through one of the N heterogeneous routes, with the objective of maintaining the freshest possible information on the monitor at all times.

A. Persistent and Intermittent Links

In this work, we consider a heterogeneous network comprising multiple communication routes, which we categorize into two distinct types based on their physical characteristics and long-term availability: $\mathcal{R}_\mathrm{s}$ for persistently available links, and $\mathcal{R}_\mathrm{i}$ for intermittent links with stochastic availability.

- **Persistent Routes** ($\mathcal{R}_\mathrm{s}$): The routes in set $\mathcal{R}_\mathrm{s}$ are continuously accessible over time and typically correspond to terrestrial links such as fiber-optic or cellular infrastructure. Due to their stable physical environment and minimal susceptibility to external disruptions, these links exhibit deterministic availability, and are modeled with an availability probability of $p_k = 1, k \in \mathcal{R}_s$. At least one route in the network belongs to this category, ensuring that the system can always transmit data, albeit potentially at higher latency or lower throughput. Without loss of generality, we denote this persistent route by index $k = 1$, as shown in Fig. 1.
- **Intermittent Routes** ($\mathcal{R}_\mathrm{i}$): whose availability fluctuates over time due to stochastic factors. These routes represent satellite links or other opportunistic channels, whose availability is inherently stochastic due to physical factors such as satellite orbital motion, line-of-sight (LOS) constraints, or environmental interference (e.g., weather). Each intermittent route $k \in \mathcal{R}_\mathrm{i}$ is characterized by a stationary availability probability $p_k \in (0, 1)$, which denotes the long-term fraction of time the route is usable.

B. Correlated Random Delays

The transmission delays across different routes at a given transmission instance may be *statistically dependent*. That is, the random variables $\{_k\}_{k\in\mathcal{N}}$ representing the route-wise delays are jointly distributed according to a stationary multivariate distribution Q, which captures the variability and potential correlation between different routes. This setting reflects realistic phenomena such as correlated queuing delays, weather-induced slowdowns, or congestion affecting multiple routes simultaneously. The delay vector is denoted by $\mathbf{T} \triangleq (T_1, T_2, \ldots, T_N) \sim Q$, with joint cumulative distribution function (CDF) $F_\mathbf{T}(t_1, \ldots, t_N)$.

We denote the marginal delay distribution of route k by Q_k, i.e., $T_k \sim Q_k$, and assume that each Q_k has finite first and second moments. Specifically, the mean and variance of the delay over route k are given by $\mu_k \triangleq \mathbb{E}_{Q_k}[T_k] < \infty$ and $\sigma_k^2 \triangleq \mathbb{E}_{Q_k}[(T_k - \mu_k)^2] < \infty$, for all $k \in \mathcal{N}$.

C. Heterogeneous Energy Costs

The system is subject to a long-term average energy constraint denoted by E_max, which limits the energy consumption over time. This constraint is particularly relevant in *energy-constrained systems* such as remote sensing applications or satellite-terrestrial networks, where power sources (e.g., battery-powered ground terminals or solar-powered satellite relays) are limited.

Each update cycle incurs two types of energy costs:

- **Sampling Energy Cost:** Every time a new status update is generated, a fixed amount of energy $C_s > 0$ is incurred. This cost accounts for sensing, computation, and other acquisition overheads required to produce a fresh update.
- **Transmission Energy Cost:** Upon sampling, the generated update is transmitted through a selected route k. Each route incurs a *per-unit-time transmission energy cost*, denoted by a function $g_t(k)$. This cost reflects route-specific factors such as propagation loss, transmission power requirements, protocol configurations, and hardware-level energy consumption. The energy consumed for transmitting an update over route k is:

$$E_\mathrm{tx}(k, T_k) = g_t(k) \cdot T_k,$$

where T_k is the stochastic delay experienced on route k.

This energy model introduces a trade-off between *timeliness* (i.e., age of information) and *energy efficiency*. For example, lower-delay routes may be intermittently available but may also incur higher energy per unit time (e.g., high-bandwidth satellite links), whereas persistent links might offer lower energy efficiency due to higher latency.

We also assume a *non-preemptive* system, where a new transmission can begin only after the previous one has been completed. Upon receiving each data sample, the monitor sends an ideal acknowledgment (ACK) to the transmitter, indicating that the system is ready to initiate the next transmission.

We adopt the *generate-at-will* model [4], [19], in which the sampler can become active at any time, provided that a new transmission is allowed. We next introduce some notation. After receiving the ACK corresponding to the previous transmission, the $(i+1)$-th data sample is generated and submitted to route R_i, selected from the pool of available routes $\mathcal{R}_i$, at time instant S_{i+1}. It is subsequently delivered to the monitor at time instant $D_{i+1} = S_{i+1} + Y_{i+1}$, where Y_{i+1} denotes the random transmission delay experienced by the $(i+1)$-th sample. The overall energy cost for transmitting the i-th update is:

$$E_i = C_s + Y_i \; g_t(R_i).$$

Since only one route is selected per transmission epoch, and the delays of the remaining routes are not observed, we model Y_{i+1} as a random variable drawn from the marginal distribution Q_{R_i}. This modeling choice remains valid even in the presence of correlation among routes, as the scheduler observes and utilizes only the delay associated with the chosen route.

The initial conditions of the system are set as follows: a sample is submitted to route k, arbitrarily selected from the set of available routes A_0, at time instant $S_0 = 0$. Consequently,

the corresponding delivery time is $D_0 = Y_0$, where $Y_0 = T_k \sim Q_k$.

D. Age of Information

The Age of Information (AoI) is the metric of our interest to measure the *freshness* of information. This metric is defined as the time elapsed since the generation of the most recently received data sample [2]. Specifically, the AoI $\Delta(t)$ at time t is defined by[2]

$$\Delta(t) \triangleq t - S_i, \quad \text{if } D_i \le t < D_{i+1}. \tag{1}$$

The AoI $\Delta(t)$ is a stochastic process that increases linearly over time and experiences downward jumps to the most recent delay value Y_i upon the delivery of the i-th data sample at time D_i, as illustrated in Fig. 2. The value of $\Delta(t)$ between the time instants $S_0 = 0$ and D_0 is assumed to increase linearly, starting from an arbitrary finite initial real value $\Delta(0) = \Delta_0 < \infty$.

III. Problem Formulation

We aim at minimizing the long-term average AoI by designing a joint sampling and routing policy $\pi \triangleq (R_0, Z_0, R_1, Z_1 \dots)$. This policy consists of two distinct sequences: (i) a sequence of routing decisions $(R_0, R_1, R_2, \dots)$, where R_i specifies the route selected for transmitting the $(i+1)$-th packet, and (ii) a sequence of finite waiting times $(Z_0, Z_1, Z_2, \dots)$, where $Z_i < \infty$ determines the $(i+1)$-th sampling (or submission) time as $S_{i+1} = D_i + Z_i$. Let Π denote the set of all causal and stationary policies π. The corresponding optimization problem is then formulated as follows:

Problem 1 (Average Age Minimization with Energy Constraint)**.**

$$\begin{aligned} \lambda^\star = \min_{\pi\in\Pi} \ & \limsup_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\int_0^T \Delta(t)\,dt\right] \\ \text{s.t.} \quad & \limsup_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\left(\int_0^T g_t(R(t))\,dt + C_s\,N_s(T)\right)\right] \\ & \le E_{\max}, \end{aligned} \tag{2}$$

where $R(t)$ denotes the transmission route at time t (with $g_t(R(t)) = 0$ if idle), $N_s(T)$ is the number of sampling actions up to time T, and $\lambda^\star$ is the optimal long-term average AoI.

A. Lagrangian Techniques

The AoI process $\Delta(t)$ is inherently a piecewise linear function as defined by equation (1). Hence, it is natural to rewrite Problem 1 as an average sum over the renewal intervals $[D_i, D_{i+1})$ corresponding to consecutive successful updates.

[2]The framework in this work can be readily extended to general AoI penalty functions $g(\Delta(t))$ as alternative metrics of interest, where $g(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ represents any monotonically non-decreasing function. However, due to page limits, we focus on the linear case in this paper.

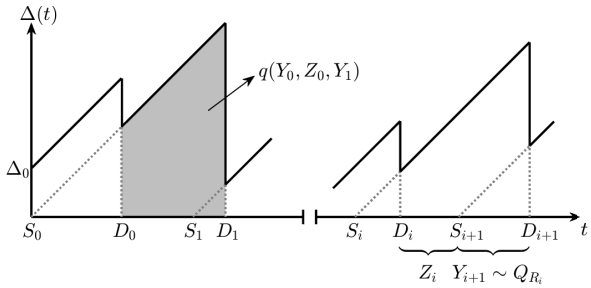

Fig. 2. Sample evolution of the AoI process $\Delta(t)$.

The long-term time average in (2) can be expressed as

$$\begin{aligned} &\limsup_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\int_0^T \Delta(t)\,dt\right] = \\ &\limsup_{n\to\infty} \frac{\mathbb{E}\left[\sum_{i=0}^{n-1} q(Y_i, Y_{i+1}, Z_i)\right]}{\mathbb{E}\left[\sum_{i=0}^{n-1}(Z_i + Y_{i+1})\right]}, \end{aligned} \tag{3}$$

where $q(Y_i, Y_{i+1}, Z_i)$ represents the accumulated AoI (the area under $\Delta(t)$) during the i-th cycle.

Similarly, the average energy consumption in the constraint of Problem 1 can be rewritten in terms of discrete renewal cycles. During each cycle, the transmitter may expend energy both through transmission and sampling operations. Specifically, the expected total energy consumed up to the n-th update can be written as

$$\mathbb{E}\left[\sum_{i=0}^{n-1}\left(g_t(R_i)Y_{i+1} + C_s\right)\right]. \tag{4}$$

Normalizing both the objective and the constraint by the expected cycle duration $\mathbb{E}\left[\sum_{i=0}^{n-1}(Z_i + Y_{i+1})\right]$, the continuous-time Problem 1 can thus be reformulated in discrete form as

$$\begin{aligned} \lambda^\star \triangleq \min_{\pi} \ & \limsup_{n\to\infty} \frac{\mathbb{E}\left[\sum_{i=0}^{n-1} q(Y_i, Y_{i+1}, Z_i)\right]}{\mathbb{E}\left[\sum_{i=0}^{n-1}(Z_i + Y_{i+1})\right]} \\ \text{s.t.} \quad & \limsup_{n\to\infty} \frac{\mathbb{E}\left[\sum_{i=0}^{n-1}\left(g_t(R_i)Y_{i+1} + C_s\right)\right]}{\mathbb{E}\left[\sum_{i=0}^{n-1}(Z_i + Y_{i+1})\right]} \le E_{\max}. \end{aligned} \tag{5}$$

The problem in equation (5) is an infinite-horizon Constrained Semi-Markov Decision Process (CSMDP) with a hybrid state space and a state-dependent hybrid action space, making it particularly challenging to solve due to the fractional structure of its objective and the additional long-term energy constraint. To address this, we adopt a three-step approach.

Since the denominator in (5) is strictly positive, the constraint in (5) can be equivalently written as

$$\begin{aligned} \limsup_{n\to\infty} \Big(&\mathbb{E}\big[\textstyle\sum_{i=0}^{n-1}\left(g_t(R_i)Y_{i+1} + C_s\right)\big] \\ &- E_{\max}\,\mathbb{E}\big[\textstyle\sum_{i=0}^{n-1}(Z_i + Y_{i+1})\big]\Big) \le 0. \end{aligned} \tag{6}$$

B. Dinkelbach's Reformulation

The objective in Problem (5) is a nonlinear fractional function of the policy π. To eliminate this fractional structure, we apply *Dinkelbach's method* for nonlinear fractional programming. For a given $\lambda \geq 0$, we define the auxiliary function

Problem 2 (Dinkelbach's Reformulation)**.**

$$h(\lambda) \triangleq \min_{\pi\in\Pi} \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\Big[q(Y_i, Y_{i+1}, Z_i) - \lambda(Z_i + Y_{i+1})\Big]. \tag{7}$$

The function $h(\lambda)$ measures the difference between the expected accumulated AoI cost and the weighted accumulated epoch duration. As established by Dinkelbach's theorem, there exists a unique $\lambda^\star$ such that

$$h(\lambda^\star) = 0,$$

and the policy attaining this zero point yields the optimal long-term average AoI. Thus, the unconstrained problem can be solved iteratively by updating λ until $h(\lambda) = 0$.

C. Lagrangian Formulation for the Energy-Constrained Case

When the average energy constraint in (6) is imposed, the corresponding Dinkelbach function becomes

$$h(\lambda, c) \triangleq \min_{\pi\in\Pi} \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\Big[q(Y_i, Y_{i+1}, Z_i) - \lambda(Z_i + Y_{i+1}) - c\big(E_{\max}(Z_i + Y_{i+1}) - g_t(R_i)Y_{i+1} - C_s\big)\Big], \tag{8}$$

where the nonnegative Lagrange multiplier $c \geq 0$ penalizes violations of the long-term average energy constraint.

The above formulation can be interpreted as the Lagrangian relaxation of the constrained problem:

$$L(c, \lambda) = \min_{\pi} \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\Big[q(Y_i, Y_{i+1}, Z_i) - \lambda(Z_i + Y_{i+1}) - c\big(E_{\max}(Z_i + Y_{i+1}) - g_t(R_i)Y_{i+1} - C_s\big)\Big]. \tag{9}$$

For a fixed c, solving $h(\lambda, c) = 0$ yields the conditionally optimal long-term average AoI $\lambda_c^\star$. Let $E(\lambda_c^\star, c)$ denote the corresponding average energy consumption. Since $E(\lambda_c^\star, c)$ is monotonically decreasing in c by complementary slackness [26], the optimal Lagrange multiplier is obtained as

$$c^\star = \inf\{\, c > 0 : E(\lambda_c^\star, c) \leq E_{\max} \,\}. \tag{10}$$

At $(\lambda^\star, c^\star)$, both the average AoI and energy constraint are simultaneously optimal, satisfying $h(\lambda^\star, c^\star) = 0$.

The following lemma establishes the relationship between $h(\lambda, c)$ and the conditionally optimal long-term average AoI $\lambda^\star(c)$ in the energy-constrained case.

Lemma 1. *For any fixed $c \geq 0$, the following assertions hold:*

1) *$\lambda^\star(c) \gtreqless \lambda$ if and only if $h(\lambda, c) \gtreqless 0$.*
2) *If $h(\lambda, c) = 0$, the solutions to the constrained problem (5) and Problem 2 coincide for the corresponding multiplier c.*

Proof. See Appendix A. ∎

According to Lemma 1, the solution to the energy-constrained problem in equation (5) can be obtained by identifying the value of λ for which $h(\lambda, c) = 0$ for a given multiplier c, and then optimizing over c. The root of the function $h(\lambda, c)$ thus corresponds to the optimal long-term average age $\lambda^\star$ under the average energy constraint, and the pair $(\lambda^\star, c^\star)$ jointly characterizes the optimal policy $\pi^\star$.

D. Average-Cost MDP for a Given λ

In the third step, we show that Problem 2 can be formulated as an average-cost Markov Decision Process (MDP) for a fixed value of λ, described by the quadruple

$$\mathscr{M}(\lambda) \triangleq (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{C}).$$

Each component of this MDP is defined as follows.

- *State Space* $\mathcal{S} = [0, \infty) \times \{0, 1\}^N$: At each decision epoch i, the system occupies a state
$$(Y_i, \mathbf{L}_i) \in \mathcal{S},$$
where the continuous component $Y_i = y < \infty$ denotes the delay observed in the previous transmission, and
$$\mathbf{L}_i = (l_{1,i}, l_{2,i}, \ldots, l_{N,i}) \in \{0, 1\}^N$$
encodes the availability of the N routes at epoch i. Specifically,
$$l_{k,i} = \begin{cases} 0, & \text{if route } k \text{ is available at time } i, \\ 1, & \text{otherwise.} \end{cases}$$
The set $\mathcal{S}$ itself is time-invariant (homogeneous), although the realization $\mathbf{L}_i$ evolves over time according to fixed underlying statistics.
- *Action Space* $\mathcal{A}(\mathbf{L}_i)$: Given the current availability vector $\mathbf{L}_i$, the decision maker selects an action (R_i, Z_i), where
  - R_i denotes the chosen route for transmission, and
  - $Z_i \geq 0$ represents the waiting time before generating the next update.

  The set of admissible actions depends on $\mathbf{L}_i$:
$$\mathcal{A}(\mathbf{L}_i) = \{(r, z) : r \in \{k \mid l_{k,i} = 0\},\ z \in \mathbb{R}^+\}.$$
The global action space is the union of all state-dependent sets:
$$\mathcal{A} = \bigcup_{\mathbf{L}\in\{0,1\}^N} \mathcal{A}(\mathbf{L}).$$
We assume $\mathcal{S}$ and $\{\mathcal{A}(\mathbf{L})\}$ are time-homogeneous, meaning they do not explicitly depend on the decision epoch i; only the realized state $\mathbf{L}_i$ varies stochastically over time.

- *State Transition Probability* $\mathcal{P}: \mathcal{S} \times \mathcal{A} \times \mathscr{B}(\mathcal{S}) \to [0,1]$: Let $\mathscr{B}(\mathcal{S})$ denote the Borel σ-algebra generated by the measurable subsets of $\mathcal{S} = [0,\infty) \times \{0,1\}^N$. For any current state $(y, \mathbf{l}) \in \mathcal{S}$, admissible action $(r,z) \in \mathcal{A}(\mathbf{l})$, and measurable set $C \in \mathscr{B}(\mathcal{S})$, the transition kernel is defined as $\mathcal{P}\big(C \mid y, \mathbf{l}, r, z\big) = P\big((Y_{i+1}, \mathbf{L}_{i+1}) \in C \mid Y_i = y, \mathbf{L}_i = \mathbf{l}, R_i = r, Z_i = z\big)$. We assume that, conditioned on the current state and action, the next delay Y_{i+1} and the next availability vector $\mathbf{L}_{i+1}$ are *conditionally independent*. That is, $P\big(Y_{i+1}, \mathbf{L}_{i+1} \mid Y_i = y, \mathbf{L}_i = \mathbf{l}, R_i = r, Z_i = z\big) = P(Y_{i+1} \mid R_i = r) P(\mathbf{L}_{i+1})$, meaning that the delay distribution depends only on the chosen route r, while the availability process evolves independently of the previous state or action. Hence, for any measurable set $C \subseteq \mathcal{S}$,

$$\mathcal{P}\big(C \mid y, \mathbf{l}, r, z\big) = \sum_{\mathbf{l}' \in \{0,1\}^N} P(\mathbf{L}_{i+1} = \mathbf{l}') \int_{C_Y(\mathbf{l}')} Q_r(y') dy',$$

where $C_Y(\mathbf{l}') := y' \geq 0 : (y', \mathbf{l}') \in C$. The availability process $\mathbf{L}_i$ is independent across epochs and identically distributed with stationary distribution

$$P(\mathbf{L}_{i+1} = \mathbf{l}') = \prod_{k=1}^{N} p_k^{1-l'_k} (1-p_k)^{l'_k}.$$

- *Cost Function* $\mathcal{C}: \mathcal{S} \times \mathcal{A} \to \mathbb{R}$: The one-step cost incurred when the system is in state $(y, \mathbf{l})$ and action (r,z) is taken is denoted by $g(y, \mathbf{l}, z, r; \lambda)$, defined as

$$g(y, \mathbf{l}, z, r; \lambda) = \mathbb{E}_{Q_r}\left[\frac{(2y + Y_{i+1} + z)(Y_{i+1} + z)}{2}\right] - \lambda z - \lambda \, \mathbb{E}_{Q_r}[Y_{i+1}].$$

By substituting $\mu_r = \mathbb{E}_{Q_r}[Y_{i+1}]$ and $\sigma_r^2 = \mathrm{Var}_{Q_r}[Y_{i+1}]$, this simplifies to

$$g(y, \mathbf{l}, z, r; \lambda) = \frac{z^2}{2} + (y + \mu_r - \lambda)\, z + (y - \lambda)\mu_r + \frac{\mu_r^2 + \sigma_r^2}{2}. \tag{11}$$

Using Lemma 1 and the constructed MDP $\mathscr{M}(\lambda)$ for a fixed λ, we can design a *nested* three-layer optimization algorithm (e.g., [9], [15]) to solve the CMDP in Problem (1). The details of this numerical solution are presented in Section V.

IV. Main Results

Before presenting the numerical solution to Problem 1, we first establish several structural results regarding the jointly optimal sampling and routing policies.

A. Structural Results of Optimal Policies

The following Theorem 1 establishes the piecewise-threshold structure of the jointly optimal sampling and routing policies.

Theorem 1. *For an N-route problem where the mean delay of each route satisfies $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_N$ and the delay distribution of each route has infinite support, the jointly optimal sampling and routing policies exhibit the following threshold structure:*

1) *Optimal Routing: The optimal routing action at the i-th epoch, denoted by $R_i^\star$, is a monotonic non-decreasing step function of the observed delay Y_i, and can be determined by $K \leq |\mathcal{R}_i| - 1$ positive thresholds $0 < \tau_1(\mathbf{l}_i) < \tau_2(\mathbf{l}_i) \cdots < \tau_K(\mathbf{l}_i)$ and $K+1$ monotonic increasing index values $a_1(\mathbf{l}_i) < a_2(\mathbf{l}_i) < \cdots < a_{K+1}(\mathbf{l}_i) \in \mathcal{R}_i$:*

$$R_i^\star = \sum_{k=1}^{K+1} \big(a_k(\mathbf{l}_i) - a_{k-1}(\mathbf{l}_i)\big) u\big(Y_i - \tau_{k-1}(\mathbf{l}_i)\big), \tag{12}$$

where $\tau_0(\mathbf{l}_i) \triangleq 0$, $a_0(\mathbf{l}_i) \triangleq 0$, and $u(t)$ is the unit step function:

$$u(t) \triangleq \begin{cases} 0, & t < 0 \\ 1, & t \geq 0. \end{cases} \tag{13}$$

Furthermore, the number of unique thresholds is upper bounded by $\frac{N(N-1)}{2}$.

2) *Optimal Sampling: The optimal waiting time at the i-th epoch, denoted by $Z_i^\star$, follows a water-filling structure and can be determined by $K+1$ thresholds $\beta_1^\star(\mathbf{l}_i) < \beta_2^\star(\mathbf{l}_i) < \cdots < \beta_{K+1}^\star(\mathbf{l}_i)$ with $\beta_k^\star(\mathbf{l}_i) = \lambda^\star - \mu_{a_k(\mathbf{l}_i)}$,*

$$Z_i^\star = \begin{cases} \big(\beta_1^\star(\mathbf{l}_i) - Y_i\big)^+, & 0 \leq Y_i < \tau_1(\mathbf{l}_i) \\ \vdots & \vdots \\ \big(\beta_K^\star(\mathbf{l}_i) - Y_i\big)^+, & \tau_{K-1}(\mathbf{l}_i) \leq Y_i < \tau_K(\mathbf{l}_i) \\ \big(\beta_{K+1}^\star(\mathbf{l}_i) - Y_i\big)^+, & \tau_K(\mathbf{l}_i) \leq Y_i \end{cases}, \tag{14}$$

or equivalently,

$$Z_i^\star = (\lambda^\star + c^\star E_{\max} - \mu_{R_i^\star} - Y_i)^+. \tag{15}$$

where $\lambda^\star$ is the optimal average AoI defined in Problem 1, and $(\cdot)^+$ is defined as $(\cdot)^+ \triangleq \max\{0, \cdot\}$.

Proof Sketch. With the MDP $\mathscr{M}(\lambda)$ (11), we can establish the *Average-Cost Optimality Equation* (ACOE) [27, Eq. 4.1]:

$$V^*(y, \mathbf{l}; \lambda, c) + h(\lambda, c) = \min_{z,\, r \in \mathcal{R}} \Big\{ g(y, \mathbf{l}, z, r; \lambda, c) + \mathbb{E}_{Q_r, \boldsymbol{p}}\big[V^*(Y_{i+1}, \mathcal{R}_{i+1}; \lambda, c)\big] \Big\}, \tag{16}$$

where $\mathcal{R} = \{\, k \mid l_k = 0 \,\}$, $V^*(y, \mathbf{l}; \lambda, c)$ is the relative value function, and $h(\lambda, c)$ is the optimal value of the reformulated MDP in Problem 2. Given any λ, c and route $r \in \mathcal{N}$, we first prove that the optimal waiting time that solves the right hand-side of (16) follows a *water-filling* structure, given by:

$$z^\star(y, \mathbf{l}; r, \lambda, c) = (\lambda + cE_{\max} - \mu_r - y)^+. \tag{17}$$

As $h(\lambda_c^\star, c) = 0$ for the conditionally optimal $\lambda_c^\star$ for any c, applying $\lambda = \lambda_c^\star$ in (16) and (17) yields:

$$V^*(y, \mathbf{l}; \lambda_c^\star, c) = \min_{r \in \mathcal{R}} \{ g(y, \mathbf{l}, z^\star(y, \mathbf{l}; r, \lambda_c^\star, c), r; \lambda_c^\star, c) + \mathbb{E}_{Q_r, \boldsymbol{p}}\big[V^*(Y_{i+1}, \mathcal{R}_{i+1}; \lambda_c^\star, c)\big] \}. \tag{18}$$

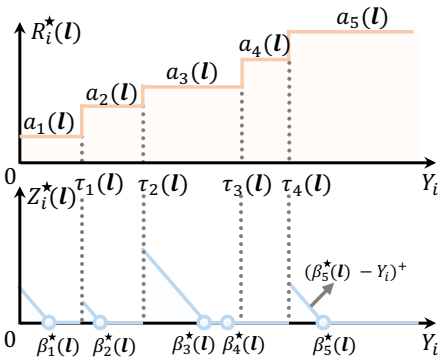

Fig. 3. Visualization of the jointly optimal policies.

For short-hand notations, we define the action-value function as:

$$Q(y,\mathbf{l},r) \triangleq g\big(y,\mathbf{l},z^\star(y;r,\lambda_c^\star,c),r;\lambda_c^\star,c\big) + \mathbb{E}_{Q_r,\boldsymbol{p}}\big[V^*(Y_{i+1},\mathcal{R}_{i+1};\lambda_c^\star,c)\big], \tag{19}$$

and the optimal routing policy $r^\star(y,\mathbf{l})$ turns to

$$r^\star(y,\mathbf{l}) = \arg\min_{r\in\mathcal{R}}\{Q(y,\mathbf{l},r)\}. \tag{20}$$

Then, we analyze a series of properties of the function $Q(y,\mathbf{l},r)$ and prove that $r^\star(y,\mathbf{l})$ is a non-decreasing step function, thus accomplishing the proof. See Section VI for the detailed proof. ■

We have established that the AoI-optimal routing policy follows a threshold-based structure. However, one can argue that these thresholds never actually exist and that the optimal policy always uses a single route. To counter this, we show that there exist system configurations where these thresholds must exist.

The following Lemma 2 demonstrates an important relationship between the sampling threshold $\beta_k^\star(\mathbf{l}_i)$ and the routing threshold $\tau_k(\mathbf{l}_i)$.

Lemma 2. *The following assertion holds true:*

$$\beta_k^\star(\mathbf{l}_i) < \tau_k(\mathbf{l}_i) \quad k \le K. \tag{21}$$

Proof. See Appendix A. ■

Consequently, for any interval $Y_i \in [\tau_{k-1}(\mathbf{l}_i),\tau_k(\mathbf{l}_i)$ associated with a constant optimal routing option $a_k(\mathbf{l}_i)$, there exists a corresponding sub-interval $\big[\beta_k^\star(\mathbf{l}_i),\tau_k(\mathbf{l}_i)\big)$ in which a zero-waiting policy, defined by $Z_i^\star = \big(\beta_k^\star(\mathbf{l}_i) - Y_i\big)^+ = 0$, is optimal. An example sketch for the structure of the jointly optimal sampling and routing policy is depicted in Fig. 3. The threshold-based structure derived in this subsection enables highly efficient deployment in complex networks. Terminals can maximize information *freshness* simply by storing and applying the derived thresholds. In Section V, we present a series of algorithms to compute these thresholds efficiently.

B. Bounds on the Optimal Average AoI

In this subsection, we establish the upper and lower bounds on the optimal average age $\lambda^\star$. These bounds will later serve as initialization points for the bisection search described in Section V.

Lemma 3. *$\lambda^\star$ is upper and lower bounded by:*

$$0 \le \lambda^\star \le \frac{3\mu_1}{2} + \frac{\sigma_1^2}{2\mu_1}. \tag{22}$$

Proof. See Appendix A. ■

V. Numerical Solutions

In this section, we develop numerical algorithms to solve the energy-constrained Average Age Minimization problem in (1) and determine the thresholds introduced in Theorem 1. By leveraging the Lagrangian relaxation and Dinkelbach's method introduced in Section III, the problem can be reformulated as a two-layer nested structure.

In the *inner layer*, for a fixed pair (λ,c), we approximate the auxiliary function $h(\lambda,c)$ defined in (7) through the Average Cost Optimality Equation (ACOE) (16). In the *outer layer*, we update (λ,c) iteratively until the Dinkelbach equilibrium $h(\lambda^\star,c^\star)=0$ and the energy constraint are simultaneously satisfied.

A. Challenges in Approximating $h(\lambda,c)$

1) Challenge 1: Hybrid Action Space

The first challenge in computing $h(\lambda,c)$ arises from the *hybrid* nature of the action space, where sampling actions $z\in\mathbb{R}^+$ are continuous, while routing actions $r\in\mathcal{R}$ are discrete and depend on the current route availability. By leveraging (17), which gives the optimal sampling policy

$$z^\star(y,\mathbf{l},r;\lambda,c) = (\lambda + cE_{\max} - \mu_r - y)^+,$$

we substitute $z^\star(\cdot)$ into the ACOE (16). This transformation yields a simplified SMDP with a countable routing action space:

$$V^*(y,\mathbf{l};\lambda,c) + h(\lambda,c) = \min_{r\in\mathcal{R}}\big\{g(y,\mathbf{l},z^\star(y;\lambda,c),r;\lambda,c) + \mathbb{E}_{Q_r,\boldsymbol{p}}[\,V^*(Y_{i+1},\mathcal{R}_{i+1};\lambda,c)\,]\big\}. \tag{23}$$

This effectively decouples the continuous and discrete components of the hybrid action space, reducing the dimensionality of the optimization.

2) Challenge 2: Uncountable State Space

The second challenge stems from the uncountable nature of the continuous state variable $y\in\mathbb{R}^+$. Evaluating (23) over the entire state space is computationally intractable. Traditional discretization methods approximate $\mathcal{S}$ by a finite grid of M points $\{y_1,\ldots,y_M\}$, introducing a quantization error ϵ_M that vanishes asymptotically but increases computational cost quadratically with M. The overall complexity of the Relative Value Iteration (RVI) approach scales as $\mathcal{O}(NM^2)$, where $N=|\mathcal{R}|$.

B. Proposed REAVI Algorithm with Energy Constraint

To overcome the trade-off between accuracy and computational cost, we extend the Relative Expected Action Value Iteration (REAVI) algorithm to handle the Lagrangian form of the constrained problem. Define the relative value function

$$W^*(y,\mathbf{l};\lambda,c) \triangleq V^*(y,\mathbf{l};\lambda,c) - V^*(0,\mathbf{l};\lambda,c), \tag{24}$$

where $W^*(0,A;\lambda,c)=0$. Substituting (24) into (23) gives

$$\begin{aligned} &W^*(y,\mathbf{l};\lambda,c) + h(\lambda,c) = \\ &\min_{r\in\mathcal{R}} \Big\{ g(y,\mathbf{l},z^\star(y;r,\lambda,c),r;\lambda,c) + \\ &\mathbb{E}_{Q_r,\boldsymbol{p}}[\, W^*(Y_{i+1},\mathcal{R}_{i+1};\lambda,c)\,] \Big\}, \end{aligned} \tag{25}$$

where $h(\lambda,c)$ is obtained by evaluating (25) at $y=0$.

We define the **Relative Expected Action Value (REAV)** function as

$$G(r;\lambda,c) \triangleq \mathbb{E}_{Q_r,\boldsymbol{p}}\big[W^*(Y_{i+1},\mathcal{R}_{i+1};\lambda,c)\big], \quad r\in\mathcal{R}. \tag{26}$$

This converts the uncountable state space $\mathbb{R}^+$ into a fixed-point problem over the finite route set $\mathcal{R}$.

Taking the expectation over the random variables $(Y,\mathcal{R})$ yields the *Relative Expected Action Value Optimality Equation (REAVOE)*:

$$\begin{aligned} &G(q;\lambda,c) + h(\lambda,c) = \mathbb{E}_{Y\sim Q_q}\mathbb{E}_{\boldsymbol{p}} \\ &\left[\min_{r\in\mathcal{R}} \big\{ g(Y,\mathbf{L},z^\star(Y;r,\lambda,c),r;\lambda,c) + G(r;\lambda,c) \big\}\right], \end{aligned} \tag{27}$$

with

$$\begin{aligned} &h(\lambda,c) \\ &= \mathbb{E}_{\boldsymbol{p}}\left[\min_{r\in\mathcal{R}} \big\{ g(0,\mathbf{L},z^\star(0;r,\lambda,c),r;\lambda,c) + G(r;\lambda,c) \big\}\right]. \end{aligned} \tag{28}$$

Algorithmic Solution: The resulting REAVI algorithm iteratively updates $\{G(r;\lambda,c)\}_{r\in\mathcal{R}}$ and $h(\lambda,c)$ until convergence. In the first outer loop, the Dinkelbach update for λ is done via bisection search for a fixed c. In the outermost loop a bisection search for $c^\star$ is done to enforce the energy constraint.

C. Policy and Energy Evaluation

Given (λ,c) and a solution $\{G(r;\lambda,c)\}$, the one-step route selection is

$$\begin{aligned} r^\star(y,\mathbf{l};\lambda,c) \in \arg\min_{r\in\mathcal{R}} \big\{ &g(y,\mathbf{l},z^\star(y;r,\lambda,c),r;\lambda,c) \\ &+ G(r;\lambda,c) \big\}. \end{aligned} \tag{29}$$

The corresponding long-run average energy of this stationary policy is

$$E(\lambda,c) = \limsup_{n\to\infty} \frac{\mathbb{E}\left[\sum_{i=0}^{n-1} \big(g_t(R_i)\, Y_{i+1} + C_s\big)\right]}{\mathbb{E}\left[\sum_{i=0}^{n-1} \big(z^\star(Y_{i+1};R_i,\lambda,c) + Y_{i+1}\big)\right]}. \tag{30}$$

Algorithm 1: *Energy-Constrained ReaVI with Nested Bisection on (λ,c) and Embedded Threshold Extraction*

Input: Bounds $c^-=0$, $c^+>0$; tolerances $\epsilon_\lambda,\epsilon_c,\epsilon_{\text{fp}}>0$

while $c^+ - c^- > \epsilon_c$ **do**

  Outer bisection on c: $c \leftarrow \dfrac{c^- + c^+}{2}$

  Inner bisection on λ (Dinkelbach root): choose $\lambda^- < \lambda^+$;

  while $\lambda^+ - \lambda^- > \epsilon_\lambda$ **do**

    $\lambda \leftarrow \dfrac{\lambda^- + \lambda^+}{2}$

    REAV fixed-point for given (λ,c): initialize $G(r)\leftarrow 0$ for all r and

$$h \leftarrow \mathbb{E}_{\boldsymbol{p}}\Big[\min_{r\in\mathcal{R}}\{g(0,\mathbf{l},z^\star(0;r,\lambda,c),r;\lambda,c)+G(r)\}\Big].$$

    repeat

      $h_{\text{old}} \leftarrow h$

      for $q\in\mathcal{R}$ **do**

$$G(q) \leftarrow -h_{\text{old}} + \mathbb{E}_{Y\sim Q_q}\mathbb{E}_{\boldsymbol{p}} \Big[\min_{r\in\mathcal{R}}\{g(Y,\mathbf{L},z^\star(Y;r,\lambda,c),r;\lambda,c) + G(r)\}\Big].$$

$$h \leftarrow \mathbb{E}_{\boldsymbol{p}}\Big[\min_{r\in\mathcal{R}}\{g(0,\mathbf{l},z^\star(0;r,\lambda,c),r;\lambda,c) + G(r)\}\Big].$$

    until $\big|h - h_{old}\big| < \epsilon_{fp}$;

    Evaluate the Dinkelbach sign: use $h(\lambda,c)=h$;

    if $h(\lambda,c)>0$ **then** $\lambda^-\leftarrow\lambda$;

    else $\lambda^+\leftarrow\lambda$;

  Fix λ as mid-point root: $\lambda \leftarrow \dfrac{\lambda^- + \lambda^+}{2}$.

  Define policy: $r^\star(y,A)$ via (29).

  Energy check: compute $E(\lambda,c)$ by (30).

  if $E(\lambda,c) \ge E_{\max}$ **then** $c^-\leftarrow c$;

  else $c^+\leftarrow c$;

Mixing to hit the boundary: with $c^- < c^+$, compute

$$q \leftarrow \frac{E_{\max} - \bar{E}(\lambda,c^-)}{\bar{E}(\lambda,c^+) - \bar{E}(\lambda,c^-)} \in [0,1],$$

and randomize between the two stationary policies.

Output: $(\lambda^\star,c^\star)$, $h(\lambda^\star,c^\star)=0$, and $\{G(r;\lambda^\star,c^\star)\}_{r\in\mathcal{R}}$.

VI. Proof of Optimal Threshold Structures

For short-hand notations, we define $Q(y,\mathbf{l},z,r;\lambda,c)$ as the *state-action function* in the right-hand side of (18):

$$\begin{aligned} Q(y,\mathbf{l},z,r;\lambda,c) \triangleq & g(y,\mathbf{l},z,r;\lambda,c) \\ & + \mathbb{E}_{Q_r,\boldsymbol{p}}[V^*(Y_{i+1},\mathcal{R}_{i+1};\lambda,c)]. \end{aligned} \tag{31}$$

Given a specific route r for $Q(y,\mathbf{l},z,r;\lambda,c)$, we first solve the conditionally optimal $z^\star(y,\mathbf{l};r,\lambda,c)$.

- Case 1: If $\lambda + cE_{\max} - \mu_r - y \le 0$, we have that

$$\frac{\partial Q(y,\mathbf{l},z,r;\lambda,c)}{\partial z} = z + y + \mu_r - \lambda - cE_{\max} \ge 0. \tag{32}$$

  In this case, $Q(y,\mathbf{l},z,r;\lambda,c)$ is monotonically increasing with z given a specific r and y, which indicates that $z^\star(y,\mathbf{l};r,\lambda,c) = 0$.
- Case 2: If $\lambda + cE_{\max} - \mu_r - y > 0$, from (32) we can establish that if $z \in (0, \lambda + cE_{\max} - \mu_r - y)$, $Q(y,\mathbf{l},z,r;\lambda,c)$ is monotonically decreasing with z; if $z \in [\lambda + cE_{\max} - \mu_r - y, \infty)$, $Q(y,\mathbf{l},z,r;\lambda,c)$ is monotonically increasing with z. As a result, $z^\star(y,\mathbf{l};r,\lambda,c) = z^\star(y;r,\lambda,c) = \lambda + cE_{\max} - \mu_r - y$.

Combining the aforementioned two cases yields:

$$z^\star(y;r,\lambda,c) = (\lambda + cE_{\max} - \mu_r - y)^+. \tag{33}$$

Substituting (33) into (31) and setting $\lambda = \lambda^\star$ yields a compact form of $Q(y,\mathbf{l},r)$, whose definition has been given in (19):

$$\begin{aligned} Q(y,\mathbf{l},r) = & -\frac{\left((\lambda^\star + cE_{\max} - \mu_r - y)^+\right)^2}{2} \\ & + (y + g_t(r)c - cE_{\max} - \lambda^\star)\mu_r + cC_s \\ & + \frac{\sigma_r^2 + \mu_r^2}{2} + \mathbb{E}_{Q_r,\boldsymbol{p}}\left[V^*(Y_{i+1},\mathcal{R}_{i+1};\lambda^\star)\right]. \end{aligned} \tag{34}$$

With the notation $Q(y,\mathbf{l},r)$, the ACOE turns to:

$$V^*(y,\mathbf{l};\lambda^\star,c) = \min_{r\in\mathcal{R}}\{Q(y,\mathbf{l},r)\}, y \in \mathbb{R}^+. \tag{35}$$

Meanwhile, the optimal routing policy is given by:

$$r^\star(y,\mathbf{l}) = \arg\min_{r\in\mathcal{R}}\{Q(y,\mathbf{l},r)\}. \tag{36}$$

To analyze the threshold structure of $r^\star(y,\mathbf{l})$, the following lemmas discuss some important properties of $Q(y,\mathbf{l},r)$ and $V^*(y,\mathbf{l};\lambda^\star,c)$.

Lemma 4. *The action-value function $Q(y,\mathbf{l},r)$ is independent of the state $\mathbf{l}$, i.e*

$$Q(y,\mathbf{l},r) = Q(y,r), \quad \forall\mathcal{R}. \tag{37}$$

Hence,

$$V^*(y,\mathbf{l};\lambda^\star,c) = \min_{r\in\mathcal{R}}\{Q(y,r)\}. \tag{38}$$

Proof. Even though $V^*(y,\mathbf{l};\lambda^\star)$ depends on $\mathbf{l}$, once route r has been selected, the term $\mathbb{E}_{Q_r,\boldsymbol{p}}\left[V^*(Y_{i+1},\mathcal{R}_{i+1};\lambda^\star,c)\right]$ in the right-hand side of (34) is independent of $\mathbf{l}$. This is because $\mathbf{l}$ and $\mathbf{l}_{i+1}$ are independent. This completes the proof. ■

Lemma 5. *The following assertions hold true:*

1) *$\forall r \in \mathcal{N}$, $Q(y,r)$ is monotonically increasing with y.*
2) *For a given $\mathbf{l}$, $V^*(y,\mathbf{l};\lambda^\star,c)$ is monotonically increasing with y.*
3) *For any routes j,k such that $\mu_j > \mu_k$, we have*

$$\frac{\partial Q(y,j)}{\partial y} \ge \frac{\partial Q(y,k)}{\partial y}, \forall y \in \mathbb{R}^+. \tag{39}$$

Proof. See Appendix A. ■

With (1) and (3) of Lemma 5 in hand, we can then establish the following lemma, which indicates that the optimal routing policy $r^\star(y,\mathbf{l})$ is monotonically non-decreasing with y:

Lemma 6. *Consider N routes with their mean delays satisfying $\mu_1 \ge \mu_2 \cdots \ge \mu_N$, if route j is optimal at $y = y^\star, \mathcal{R}$, we have that*

$$\begin{aligned} r^\star(y,\mathbf{l}) &= \arg\min_{r\in\mathcal{R}}\{Q(y,r)\} \ge j, \text{if } y > y^\star, \\ r^\star(y,\mathbf{l}) &= \arg\min_{r\in\mathcal{R}}\{Q(y,r)\} \le j, \text{if } y < y^\star. \end{aligned} \tag{40}$$

Proof. See Appendix A ■

As Lemma 6 holds for $\forall y, \forall\mathbf{l}$ and $y^\star$, $r^\star(y,\mathbf{l})$ is a monotonically non-decreasing function with respect to y. As $r^\star(y,\mathbf{l})$ belongs to a discrete set $\mathcal{R}$, it forms a non-decreasing step function as shown in (12). Substituting the step function $r^\star(y,\mathbf{l})$ into (33) yields:

$$z^\star(y) = z^\star(y;r^\star(y),\lambda^\star,c) = (\lambda^\star + cE_{\max} - \mu_{r^\star(y,\mathbf{l})} - y)^+. \tag{41}$$

For a given constant-value interval $\left[\tau_{k-1}(\mathbf{l}),\tau_k(\mathbf{l})\right)$ where $r^\star(y,\mathbf{l}) = a_k(\mathbf{l})$, the optimal sampling policy is defined as:

$$z^\star(y) = (\lambda^\star + cE_{\max} - \mu_{a_k(\mathbf{l})} - y)^+. \tag{42}$$

Defining $\beta_k^\star(\mathbf{l}) \triangleq \lambda^\star - \mu_{a_k(\mathbf{l})}$, we next prove that the *water-filling levels* $\beta_k^\star(\mathbf{l})$ are strictly increasing with the index k. First, we can show that for any $i < j$, it follows that $\mu_{a_i(\mathbf{l})} \le \mu_{a_j(\mathbf{l})}$, which leads to

$$\begin{aligned} \beta_i^\star(\mathbf{l}) &= (\lambda^\star + cE_{\max} - \mu_{a_i(\mathbf{l})})^+ \le (\lambda^\star + cE_{\max} - \mu_{a_j(\mathbf{l})})^+ \\ &= \beta_j^\star(\mathbf{l}). \end{aligned} \tag{43}$$

Next, we prove that $\beta_i^\star(\mathbf{l}) \ne \beta_j^\star(\mathbf{l})$ for $i \ne j$. This is achieved by the following lemma, which indicates that $\mu_{a_i(\mathbf{l})} \ne \mu_{a_j(\mathbf{l})}$ for $i \ne j$.

Lemma 7. *Let $\mathcal{R}^\star = \left\{a_1(\mathbf{l}),\ldots,a_{K+1}(\mathbf{l})\right\}$ denote the set of routes used by the age-optimal policy for a given $\mathbf{l}$ and let $\mathcal{G}_\mu$ be defined as*

$$\mathcal{G}_\mu \triangleq \{r \in \mathcal{R} : \mu_r = \mu\}. \tag{44}$$

Then, at most one route from $\mathcal{G}_\mu$ can belong to the optimal set $\mathcal{R}^\star$:

$$|\mathcal{R}^\star \cap \mathcal{G}_\mu| \le 1, \forall\mu \in \mathbb{R}^+. \tag{45}$$

Proof. See Appendix A. ■

With lemma 7 and (43), we establish that $\beta_1(\mathbf{l})^\star < \cdots < \beta_{K+1}^\star(\mathbf{l})$.

Finally, we prove that there are at most $\frac{N(N-1)}{2}$ unique routing thresholds. For $a < b$, we define $\tau_{a,b}$ as where the equality $Q(\tau_{a,b}, a) = Q(\tau_{a,b}, b)$ is satisfied. By Lemma 6, $\tau_{a,b}$ exists if and only if $Q(0,b) \leq Q(0,a)$. Then, if the inequality

$$Q(0,1) \leq Q(0,2) \leq \cdots \leq Q(0,N) \tag{46}$$

is satisfied, $\tau_{a,b}$ exists $\forall a, b \in \mathcal{N},\ a \neq b$. Hence, there are at most $\binom{N}{2} = \frac{N(N-1)}{2}$ unique routing thresholds. Note that the existence of $\tau_{a,b}$ does not necessitate its presence in the optimal solution.

VII. Special Case: $\boldsymbol{p} = \mathbf{1}$, $E_{\max} = \infty$

This section focuses on the case where all routes are always available, i.e $\boldsymbol{p} = \mathbf{1}$, and there is no energy constraint upon the system. Since all routes are available at each interval, availability is no longer a part of the state space. At each decision instance a route is picked from the set $\mathcal{N}$ instead of $\mathcal{R}_i$ since $\mathcal{R}_i = \mathcal{N}, \forall i$. A detailed analysis of this problem is given in Here, we present the differences of the optimal solution from problem 1.

Since $\mathcal{R}_i = \mathcal{N},\ \forall i$, the maximum number of routing thresholds is reduced to $N-1$. However, due to the independence of the waiting time to the availability probabilities $\boldsymbol{p}$, the maximum number of waiting thresholds is again given by N.

The optimal routing action at the i-th epoch is given by

$$R_i^\star = \sum_{k=1}^{K+1} (a_k - a_{k-1}) u(Y_i - \tau_{k-1}), \tag{47}$$

where $K \leq N-1$, $\tau_0 \triangleq 0$, $a_0 \triangleq 0$, and $u(t)$ is the unit step function.

The optimal waiting time at the i-th epoch has a modified expression where the energy constraint is removed from(15) and is given by

$$Z_i^\star = \left(\lambda^\star - \mu_{R_i^\star} - Y_i\right)^+. \tag{48}$$

Lemma 2 turns into

$$\beta_k^\star < \tau_k. \tag{49}$$

The upper bound to $\lambda^\star$ can be updated as $\min_i \left\{\frac{3\mu_i}{2} + \frac{\sigma_i^2}{2\mu_i}\right\}$ since all routes are available at all times.

VIII. Simulation Results

This section presents simulation results for practical scenarios to validate the analytical findings and evaluate the performance of our proposed algorithm.

A. Comparing Benchmarks

In this subsection, we refer to our designed jointly optimal sampling and routing policy as the "optimal policy" and evaluate its performance against the following benchmark policies:

• *Minimum Average Delay Routing with AoI-Optimal Sampling (**MAD-Optimal**)*: This policy always selects the route with the minimum average delay over the set of available routes at each instance. Given this selection, a modified version of the *ReAVI* with the minimization over the routing options removed is implemented to find the AoI-optimal waiting strategy and minimize the long-term average AoI.

• *Minimum Average Delay Routing with Zero-Wait Sampling (**MAD-Zero Wait**)*: This policy always selects the route with the minimum average delay over the set of available routes at each instance. It is combined with a zero-wait strategy, where a new packet is sampled and transmitted immediately upon the delivery of the previous packet[3]. The long term average AoI achieved by this policy can be analytically calculated.

Lemma 8. *For a system with N routes satisfying $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_N$, MAD-ZW policy given by:*

$$\pi^{\text{MAD-ZW}}(y, \mathbf{1}) \triangleq (r = \max A, z = 0), \tag{50}$$

achieves a long-term average AoI given by:

$$\begin{aligned}\lambda^{\text{MAD-ZW}} = &\sum_{i=1}^{N} p_i \mu_i \prod_{k=i+1}^{N} (1-p_k) \\ &+ \frac{\sum_{i=1}^{N} p_i \mu_i \left(\frac{\mu_i}{2} + \frac{\sigma_i^2}{2\mu_i}\right) \prod_{k=i+1}^{N}(1-p_k)}{\sum_{i=1}^{N} p_i \mu_i \prod_{k=i+1}^{N}(1-p_k)}.\end{aligned} \tag{51}$$

Proof. See Appendix A. ∎

This policy can be undesirable over simpler policies like route k-Zero Wait. Following is an analysis where $N = 3$ and $\boldsymbol{p} = [1, p, p]$.

a) MAD-Zero Wait vs. route 1-Zero Wait

The analytical expression for λ^{MAD-ZW} can be obtained from (51) by setting $N = 3$ and $\boldsymbol{p} = [1, p, p]$. Then,

$$\lambda^{MAD-ZW} = A(p) + \frac{B(p)}{A(p)} \tag{52}$$

where

$$\begin{aligned}A(p) &= \left(1 - p^2\right)\mu_1 + \left(p - p^2\right)\mu_2 + p\mu_3 \\ B(p) &= \left(\frac{\mu_1}{2} + \frac{\sigma_1^2}{2\mu_1}\right)\left(1 - p^2\right)\mu_1 \\ &\quad + \left(\frac{\mu_2}{2} + \frac{\sigma_2^2}{2\mu_2}\right)\left(p - p^2\right)\mu_2 \\ &\quad + \left(\frac{\mu_3}{2} + \frac{\sigma_3^2}{2\mu_3}\right)p\mu_3.\end{aligned} \tag{53}$$

As a result, if the first derivative of λ^{MAD-ZW} with respect to p, given by:

$$\frac{d}{dp}\lambda^{MAD-ZW} = A'(p) + \frac{B'(p)\,A(p) - B(p)\,A'(p)}{A(p)^2}, \tag{54}$$

satisfies $\frac{d}{dp}\lambda^{MAD-ZW} > 0$ for $p \in [0,1]$, then the MAD-Zero Wait age will increase with p. In such cases, route 1-Zero Wait policy will outperform the MAD-Zero Wait.

• *Minimum Delay Variance Routing with Zero-Wait Sampling (**MDV-Zero Wait**)*: This policy consistently selects the route with the lowest delay variance. It is combined with a zero-wait strategy.

[3]Zero-wait policy [19] is work-conserving, hence, it achieves maximum throughput on any given route.

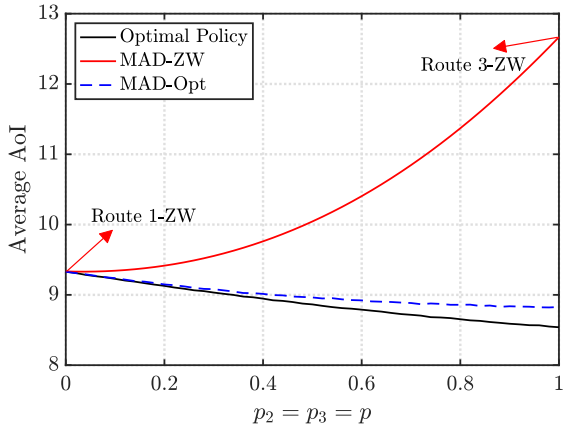

Fig. 4. Simulation results of systems with $N = 3$ and competitive route 1.

• *Minimum Delay Variance Routing with AoI-Optimal Sampling (**MDV-Optimal**)*: This policy always selects the route with the minimum delay variance and follows the AoI-optimal waiting strategy as outlined in [19, Theorem 4].

B. Satellite-Terrestrial Integrated Routes

We consider two distinct classes of routes, denoted by $\mathcal{N}_{\text{Sat}}$ and $\mathcal{N}_{\text{Ter}}$. Here, $\mathcal{N}_{\text{Sat}}$ represents the set of Low Earth Orbit (LEO) Satellite routes with stochastic delays, while $\mathcal{N}_{\text{Ter}}$ represents the set of terrestrial routes with stochastic delays.

1) LEO Satellite Routes with Stochastic Delays

For $l \in \mathcal{N}_{\text{Sat}}$, the delay is modeled by a *log-normal distribution*, characterized by the following probability density function [28]:

$$P_{Y\sim Q_l}(y) = \frac{1}{y\beta_l\sqrt{2\pi}} \exp\left(-\frac{(\ln y - \alpha_l)^2}{2\beta_l^2}\right), l \in \mathcal{N}_{\text{Sat}}, \tag{55}$$

where α_l and β_l correspond to the mean and standard deviation of the underlying normal distribution.

The mean μ_l and the variance σ_l^2 of $Y \sim Q_l$ are given by:

$$\mu_l = \exp(\alpha_l + \frac{\beta_l^2}{2}), l \in \mathcal{N}_{\text{Sat}} \tag{56a}$$

$$\sigma_l^2 = (\exp(\beta_l^2) - 1)\exp(2\alpha_l + \beta_l^2), l \in \mathcal{N}_{\text{Sat}}. \tag{56b}$$

2) Terrestrial Routes with Stochastic Delays

If $l \in \mathcal{N}_{\text{Ter}}$, we leverage the *gamma distribution* to simulate the statistics of delay y, where the probability density function is given by [28]:

$$P_{Y\sim Q_l}(y) = \frac{1}{\Gamma(\theta_l)\gamma_l{}^{\theta_l}} y^{\theta_l - 1} e^{-y/\gamma_l}, l \in \mathcal{N}_{\text{Ter}}. \tag{57}$$

The mean μ_l and the variance σ_l^2 of $Y \sim Q_l$ are given by:

$$\mu_l = \theta_l\gamma_l, \text{ and } \sigma_l^2 = \theta_l\gamma_l^2, l \in \mathcal{N}_{\text{Ter}}. \tag{58}$$

C. Parameter Settings

We first consider a scenario where there are 2 available routes that are always available with the varied energy constraint $E_{\max} \in [1, 13]$. The routes has the delay statistics: Log-normal distribution, $\mu_1 = 5, \sigma_1 = 1.4$; Gamma distribution, $\mu_1 = 5, \sigma_1 = 4$. Route dependent transmission costs are given by: $g_t(1) = 9$, $g_t(2) = 13$, and the sampling cost is $C_s = 2$.

We then consider two different scenarios with no energy constraint. First of which is a scenario with three available routes where $\mathcal{N} = \{1, 2, 3\}$, and two different availability settings. The parameter setting for the simulations where all

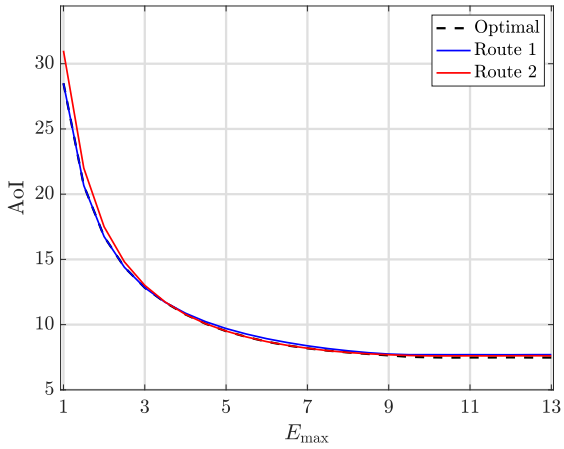

Fig. 5. AoI vs. $E_{\max}$

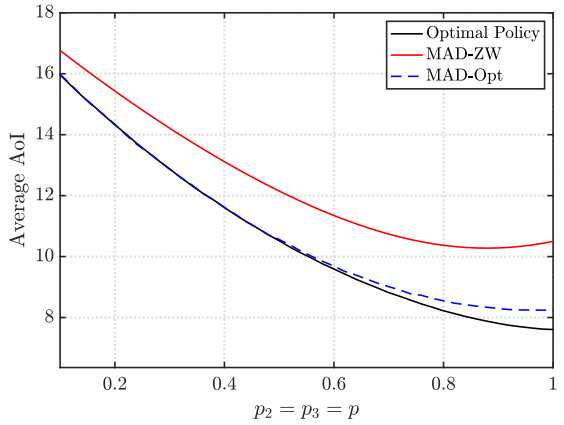

Fig. 6. Simulation results of systems with $N = 3$ and uncompetitive route 1.

TABLE I
SIMULATION PARAMETERS $\boldsymbol{p} = \mathbf{1}, E_{\max} = \infty$

Route	**Route** 1	**Route** 2	**Route** 3
Distribution Parameters	Log-normal (μ_1, σ_1)	Gamma (μ_2, σ_2)	Gamma (μ_3, σ_3)
Fig. 7 (a)	$(3.4, [0, 3])$	$(0.7, 5)$	—
Fig. 7 (b)	$([0.7, 5.2], 2)$	$(0.7, 5)$	—
Fig. 7 (c)	$(2.4, [0, 3])$	$(1.2, 3)$	$(0.7, 3.4)$
Fig. 7 (d)	$(2.4, 0.7)$	$(1.2, 3)$	$([0.7, 1.2], 3.4)$
Fig. 8	$(2.4, 0.7)$	$(1.2, 3)$	$(0.7, 3.4)$

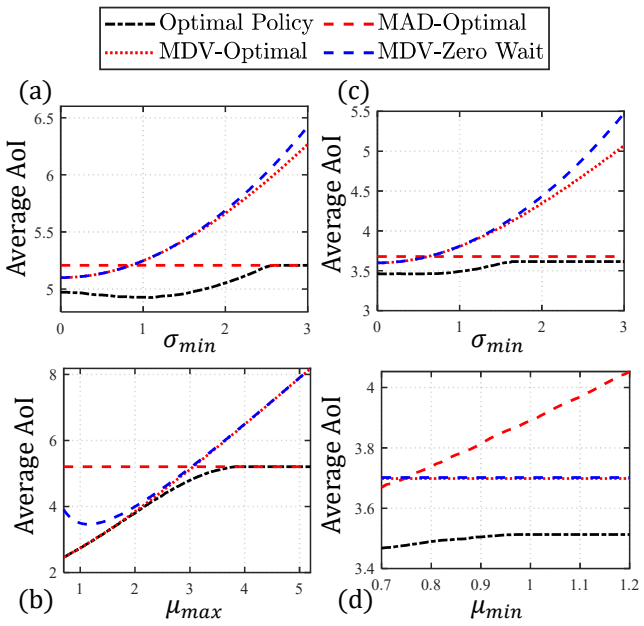

Fig. 7. Simulation results of systems with $N = 2$ and $N = 3$.

TABLE II
SIMULATION PARAMETERS $E_{\max} = \infty$

Route	**Route** 1	**Route** 2	**Route** 3
Parameters	$(Q_1, \mu_1, \sigma_1, p_1)$	$(Q_2, \mu_2, \sigma_2, p_2)$	$(Q_3, \mu_3, \sigma_3, p_3)$
Fig. 4, $E_{\max} = \infty$	(Gamma, 6, 2, 1)	(Log-normal, 5, 4, p)	(Gamma, 3, 7, p)
Fig. 6, $E_{\max} = \infty$	(Gamma, 10, 8, 1)	(Log-normal, 4, 4, p)	(Log-normal, 3, 6, p)

routes are always ON ($\boldsymbol{p} = \mathbf{1}$) is presented in Table I. In this table, parameters specified as intervals indicate the values that are varied along the horizontal axis of the corresponding simulation figure. For notational convenience, we define:

$$\sigma_{\min} = \min_{r\in\mathcal{N}} \sigma_r, \quad \mu_{\min} = \min_{r\in\mathcal{N}} \mu_r, \quad \mu_{\max} = \max_{r\in\mathcal{N}} \mu_r. \quad (59)$$

The parameter setting for the simulations where a single route (route 1) is always ON is presented in Table II where $p_2 = p_3 = p$ is the varying parameter.

D. Discussions

Fig. 5 demonstrates that the long-term average AoI decreases as the energy thresholds is loosened. Furthermore, when E_{max} exceeds the energy output of the unconstrained system, the AoI converges to a set value as expected.

Fig. 4 demonstrates that MAD-Zero Wait can underperform against a simpler policy. When $p = 0$, MAD-Zero Wait performs similar to the optimal policy since route 1's delay has a small variance in this setting. As p increases, the MAD-Zero Wait policy uses the other routes, which are not suitable to zero wait policies, increasingly often. When $p = 1$, MAD-Zero Wait policy is equivalent to the route 3-Zero Wait policy. An analysis showing when we can expect MAD-Zero Wait to exhibit the behavior in Fig. 4 follows:

Figures 4 and 6 demonstrate that the advantage of using the optimal policy over the MAD-Optimal policy dwindles

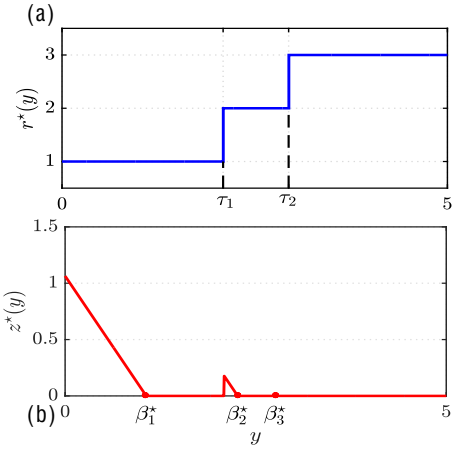

Fig. 8. Visualization of simulated optimal policies.

in the low availability region (small p). It can also be seen that the separation between the policies happens for a larger p when route 1 has worse delay characteristics. As a result, it may be beneficial to apply the MAD-Optimal policy (smaller complexity) in cases where route 1 is uncompetitive and routes $2, \ldots, N$ have low joint availability (i.e, when $\prod_2^N 1 - p_i$ is large).

Fig. 7(a) highlights a surprising finding: *a higher delay variance may counterintuitively improve the average AoI performance*. In a single-route setting, reducing the delay variance typically leads to more regular update arrivals and thus lower AoI. However, this intuitive conclusion breaks down when an additional route is available, as evidenced by Fig. 7(a). Moreover, the figure shows that when $\sigma_{\min} = \sigma_1$ is below a certain threshold (approximately 2.5), the optimal policy actively utilizes both routes. Once this threshold is crossed, route 1 is no longer selected, and the route 2 (MAD-Optimal) with minimum delay is always used.

Fig. 7(b) demonstrates that route 1 (MDV-Optimal) provides better AoI performance than route 2 (MAD-Optimal) when $\mu_1 = \mu_{\max}$ is relatively small. While route 2 is used in the optimal policy due to its shorter average delay, its role is marginal. Notably, the benefit of joint routing peaks when $\mu_{\max}$ is just above 3. As $\mu_{\max}$ continues to increase and exceeds approximately 4, the route 2 (MAD-Optimal) policy becomes age-optimal.

Fig. 7(c) presents the long-term average AoI values in a three-route scenario, where $\sigma_1 = \sigma_{\min}$ is varied from 0 to 3. The optimal policy utilizes all three routes until $\sigma_{\min}$ exceeds a threshold of approximately 1.5, beyond which route 1 is no longer selected, and the policy relies solely on routes 2 and 3.Fig. 7(d) shows the long-term average AoI values for another three-route scenario, where $\mu_3 = \mu_{\min}$ is varied from 0.7 to 1.2. In this case, all three routes consistently appear in the optimal policy, as route 3 remains the minimum delay route. However, its contribution to the overall performance becomes negligible as $\mu_3 = \mu_{\min}$ approaches $\mu_2 = 1.2$.

Fig. 8(a) and Fig. 8(b) show threshold structure of the op-

timal routing decision $R_i^\star = r^\star(y)$ and waiting time decision $Z_i^\star = z^\star(y)$ when $Y_i = y$. The parameter configuration is specified in Table I. These results verify Theorem 1.

Overall, the proposed joint sampling and routing policy demonstrates robust improvements in AoI performance under diverse parameter settings. In particular, our simulations show that even in a basic three-route example, average AoI can be reduced by as much as 10%. This finding challenges conventional intuition and reveals a critical insight: *routes that appear suboptimal in isolation—due to higher mean delays or variances—can meaningfully contribute to AoI minimization under a well-designed optimized handover policy*.

IX. Conclusion

In this work, we investigated a multi-route status update system and proved that a threshold-based joint sampling and routing policy can minimize the long-term average AoI. We introduced an efficient algorithm namely Bisec-ReaVI to compute this optimal policy. Our simulations consistently show improvements in AoI, revealing that higher variance or mean delays in certain routes can still help minimize AoI when jointly optimized. This challenges the common intuition that lower delay variance always leads to better AoI performance and provides insights into routing design for future TN-NTN networks.

References

[1] A. U. Atasayar, A. Li, Arı, and E. Uysal, "Fresh data delivery: Joint sampling and routing for minimizing the age of information," in *Proc. ACM Mobihoc*, 2025.

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

[3] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *IEEE ITA*, 2015, pp. 25–31.

[4] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *IEEE ISIT*, 2015, pp. 3008–3012.

[5] B. T. Bacinoglu, Y. Sun, E. Uysal, and V. Mutlu, "Optimal status updating with a finite-battery energy harvesting source," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 280–294, 2019.

[6] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *55th Annual Allerton Conference on Communication, Control, and Computing*. IEEE, 2017, pp. 486–493.

[7] H. B. Beytur and E. Uysal-Biyikoglu, "Minimizing age of information for multiple flows," in *IEEE BlackSeaCom*. IEEE, 2018, pp. 1–5.

[8] I. Kadota, A. Sinha, and E. Modiano, "Scheduling algorithms for optimizing age of information in wireless networks with throughput constraints," *IEEE/ACM Transactions on Networking*, vol. 27, no. 4, pp. 1359–1372, 2019.

[9] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Optimal sampling and scheduling for timely status updates in multi-source networks," *IEEE Transactions on Information Theory*, vol. 67, no. 6, pp. 4019–4034, 2021.

[10] E. T. Ceran, D. Gündüz, and A. György, "A reinforcement learning approach to age of information in multi-user networks with harq," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1412–1426, 2021.

[11] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2637–2650, 2018.

[12] G. Chen, S. C. Liew, and Y. Shao, "Uncertainty-of-information scheduling: A restless multiarmed bandit framework," *IEEE Transactions on Information Theory*, vol. 68, no. 9, pp. 6151–6173, 2022.

[13] X. Chen, A. Li, and S. Wu, "Optimal sampling for uncertainty-of-information minimization in a remote monitoring system," in *Profc. IEEE Information Theory Workshop (ITW)*. IEEE, 2024, pp. 115–120.

[14] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the wiener process for remote estimation over a channel with random delay," *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 1118–1135, 2019.

[15] A. Li, S. Wu, G. C. Lee, X. Chen, and S. Sun, "Sampling to achieve the goal: An age-aware remote markov decision process," in *2024 IEEE Information Theory Workshop (ITW)*, 2024, pp. 121–126.

[16] X. Chen, A. Li, and S. Wu, "Optimal sampling for uncertainty-of-information minimization in a remote monitoring system," in *2024 IEEE Information Theory Workshop (ITW)*, 2024, pp. 115–120.

[17] M. K. C. Shisher, Y. Sun, and I.-H. Hou, "Timely communications for remote inference," *IEEE/ACM Transactions on Networking*, 2024.

[18] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 204–219, 2019.

[19] 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.

[20] H. Tang, Y. Chen, J. Wang, P. Yang, and L. Tassiulas, "Age optimal sampling under unknown delay statistics," *IEEE Transactions on Information Theory*, vol. 69, no. 2, pp. 1295–1314, 2023.

[21] Y. Sun and S. Kompella, "Age-optimal multi-flow status updating with errors: A sample-path approach," *Journal of Communications and Networks*, vol. 25, no. 5, pp. 570–584, 2023.

[22] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, "Minimizing age of information via scheduling over heterogeneous channels," in *ACM MobiHoc*, 2021, pp. 111–120.

[23] ——, "Age-optimal scheduling over hybrid channels," *IEEE Transactions on Mobile Computing*, vol. 22, no. 12, pp. 7027–7043, 2023.

[24] S. Leng and A. Yener, "Age of information minimization for an energy harvesting cognitive radio," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 2, pp. 427–439, 2019.

[25] M. Zhou, M. Zhang, H. H. Yang, and R. D. Yates, "Age-minimal cpu scheduling," in *IEEE INFOCOM 2024-IEEE Conference on Computer Communications*. IEEE, 2024, pp. 401–410.

[26] E. Altman, *Constrained Markov decision processes*. Routledge, 2021.

[27] R. A. Howard, *Dynamic programming and markov processes*. Cambridge, MA; MIT Press, 1960.

[28] P. Z. Peebles Jr, *Probability, random variables, and random signal principles*. McGraw-Hill, 2001.

[29] S. M. Ross, *Stochastic processes*. John Wiley & Sons, 1995.

Appendix

Part 1. We first prove that

$$\lambda^\star \le \lambda \iff h(\lambda, c) \le 0. \tag{60}$$

If $\lambda^\star \le \lambda$,

$$\exists\ \pi,\ \limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}_\pi[q(Y_i,\, Z_i,\, Y_{i+1})]}{\sum_{i=0}^{n-1} \mathbb{E}_\pi[Z_i + Y_{i+1}]} \le \lambda. \tag{61}$$

Moving λ to the left-hand side yields: $\exists\ \pi$

$$\limsup_{n\to\infty} \frac{\frac{1}{n}\sum_{i=0}^{n-1} \Big(\mathbb{E}_\pi\big[q(Y_i,\, Z_i,\, Y_{i+1})\big] - \lambda\, \mathbb{E}_\pi[Z_i + Y_{i+1}] \Big)}{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}_\pi[Z_i + Y_{i+1}]} \le 0. \tag{62}$$

Since Y_i's over the same route are independent, the inter-sampling times $T_i^{\text{int}} = Y_i + Z_i$ are regenerative. Since there are N routes, the expected period of the most frequently used route satisfies $\mathbb{E}[n_{k+1} - n_k] \le N$, where n_k denotes the k-th epoch a particular route is used. Because T_i^{int}'s are regenerative and we have $0 < \mathbb{E}[D_{n_{k+1}} - D_{n_k}] < \infty$, for all k, the renewal theory [29] tells us that $\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]$ exists and is positive. Thus, there exists a policy π such that the numerator of the left-hand side (62) is less than zero, which

indicates that the infimum of the numerator in (62) is less than zero, indicating that $h(\lambda) \leq 0$.

Conversely, if $h(\lambda, c) \leq 0$, as $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]$ exists and is positive, we can derive (61) and (62), which indicates that $\lambda^\star \leq \lambda$. The corollary $\lambda^\star > \lambda \iff h(\lambda, c) > 0$ can be derived directly from (60) by leveraging *Modus Tollens*.

Part 2: $\lambda^\star = \lambda \iff h(\lambda, c) = 0$. If $h(\lambda, c) = 0$, from part 1 of the proof, we can first establish that $\lambda^\star \leq \lambda$. We then show that the policy π such that $h(\lambda, c) = 0$ can lead to

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}_\pi[q(Y_i,\, Z_i,\, Y_{i+1})]}{\sum_{i=0}^{n-1} \mathbb{E}_\pi[Z_i + Y_{i+1}]} = \lambda, \tag{63}$$

which indicates that $\lambda \geq \lambda^\star$. Combining these together, we can obtain $\lambda = \lambda^\star$. Conversely, if $\lambda = \lambda^\star$, we can establish from part 1 that $h(\lambda, c) \leq 0$; Meanwhile, the definition of $\lambda^\star$ in (5) leads to

$$\forall\ \pi,\ \limsup_{n\to\infty} \sum_{i=0}^{n-1} \mathbb{E}_\pi[q(Y_i,\, Z_i,\, Y_{i+1})] - \lambda \mathbb{E}_\pi[Z_i + Y_{i+1}] \geq 0, \tag{64}$$

which indicates that $h(\lambda, c) \geq 0$. Combining these together, we establish that $h(\lambda, c) = 0$.

Differentiating the action-value function given in (34) with respect to y, we obtain:

$$\frac{\partial Q(y, r)}{\partial y} = \begin{cases} \lambda^\star - y, & \text{if } y < \lambda^\star - \mu_r \\ \mu_r, & \text{if } y \geq \lambda^\star - \mu_r \end{cases} \tag{65}$$

For all y, the derivative is positive. Hence, $\forall r \in \mathcal{N}$, $Q(y, r)$ is monotonically increasing with y. As a result, for any $y_2 \geq y_1$, we can establish that

$$\begin{aligned} V^*(y_2, \mathcal{R}; \lambda^\star) - V^*(y_1, \mathcal{R}; \lambda^\star) &= \min_{r\in\mathcal{R}} Q(y_2, r) - \min_{r\in\mathcal{R}} Q(y_1, r) \\ &\geq \min_{r\in\mathcal{R}} \{Q(y_2, r) - Q(y_1, r)\} \geq 0, \end{aligned} \tag{66}$$

which indicates that $V^*(y, \mathbf{l}; \lambda^\star) = \min_{r\in\mathcal{R}} Q(y, r)$ is monotonically increasing with y.

Since $\mu_j > \mu_k$, it follows that $\lambda^\star - \mu_k > \lambda^\star - \mu_j$. Then, using (65), we compute the difference:

$$\frac{\partial Q(y, j)}{\partial y} - \frac{\partial Q(y, k)}{\partial y} = \begin{cases} 0, & \text{if } y < \lambda^\star - \mu_j \\ \mu_j + y - \lambda^\star, & \text{if } \lambda^\star - \mu_k > y \geq \lambda^\star - \mu_j \\ \mu_j - \mu_k, & \text{if } y \geq \lambda^\star - \mu_k. \end{cases} \tag{67}$$

In all cases, the difference is non-negative, thus

$$\frac{\partial Q(y, j)}{\partial y} - \frac{\partial Q(y, k)}{\partial y} \geq 0, \tag{68}$$

which completes the proof. Since route j is optimal at $y = y^\star, A$, we have

$$Q(y^\star, j) \leq Q(y^\star, i), \tag{69}$$

for any $i \in A$. Now, for $i < j$ we know $\mu_i \geq \mu_j$. Then, combining (39) with (69) we obtain

$$Q(y, j) \leq Q(y, i),\ y \geq y^\star, A, \tag{70}$$

which proves that no route $i < j,\ i \in A$ can be optimal for $y > y^\star, A$. The proof for the converse statement follows the same logic. As given by (65), the sole dependence of $\partial Q(y, i)/\partial y$ on i is μ_i. Thus, $\forall j, k \in \mathcal{G}_\mu$, we have

$$\frac{\partial Q(y, j)}{\partial y} = \frac{\partial Q(y, k)}{\partial y}, \forall y \in \mathbb{R}^+, \forall j, k \in \mathcal{G}_\mu. \tag{71}$$

Thus, if route i is optimal at $y = 0$, it is also optimal for every $y \in \mathbb{R}^+$:

$$Q(y, i) = \min_{r\in\mathcal{G}_\mu} \{Q(y, r\}), \forall y \in \mathbb{R}^*, \tag{72}$$

which indicates that this route *dominates* the space $\mathcal{G}_\mu$. As a result, only one route from $\mathcal{G}_\mu$ will be included in $\mathcal{R}^\star$.

We know from Lemma 6 that the optimal route j at $y = \tau_k(\mathbf{l}_i)$ satisfies $\mu_j < \mu_{a_k(\mathbf{l}_i)}$. Then, we have

$$\frac{\partial Q(y, j)}{\partial y} = \frac{\partial Q\big(y, \mathbf{l}_k(\mathbf{l}_i)\big)}{\partial y} = \lambda^\star - y, \quad y < \beta_k^\star(\mathbf{l}_i). \tag{73}$$

Therefore, $\tau_k(\mathbf{l}_i)$ must be greater than $\beta_k^\star(\mathbf{l}_i)$.

Consider a policy $\pi = (1, 0, 1, 0, \ldots)$ that selects a single route 1 (always available) and generates a new sample immediately after the previous update packet is delivered (i.e., $Z_i = 0,\ \forall i$). We denote the long-term average age under this zero-wait route 1 policy as λ_1^{zw}. This age simplifies to $\frac{3\mu_1}{2} + \frac{\sigma_1^2}{2\mu_1}$ since Y_i's over a single route are i.i.d.

Let λ^π denote the average age achieved by a policy $\pi \in \Pi$. By definition of optimality, we have

$$\lambda^\star \leq \lambda^\pi, \quad \forall \pi \in \Pi. \tag{74}$$

In particular, this implies: $\lambda^\star \leq \lambda_1^{zw} = \frac{3\mu_1}{2} + \frac{\sigma_1^2}{2\mu_1}$. Since the AoI $\Delta(t)$ is non-negative, $\lambda^\star \geq 0$. This completes the proof.

Let us derive the long-term average age achieved by using route j under the MAD-ZW policy. The expected age over such intervals can be given by:

$$\lambda_j^{av} = \frac{E\left[YY' + \frac{(Y')^2}{2}\right]}{E[Y']} = E[Y] + \frac{E[Y']}{2} + \frac{var(Y')}{2E[Y']}, \tag{75}$$

where Y denotes the randomized previous delay and Y' denotes the delay over route j. Since the availability of routes are i.i.d over intervals, so is the route selection process of the MAD-ZW policy. The probability that route k is used under π^{MAD-ZW} is the probability of the event that route k is ON and no route $l > k, l \leq N$ is ON, i.e $k = \max \mathcal{R}$. Therefore, we can express the $E[Y]$ term in (75) as:

$$E[Y] = \sum_{i=1}^{N} p_i \mu_i \prod_{k=i+1}^{N} (1 - p_k). \tag{76}$$

Then, combining (75) and (76), the average age attained over route j can be expressed as:

$$\lambda_j^{av} = \sum_{i=1}^{N} p_i \mu_i \prod_{k=i+1}^{N} (1 - p_k) + \left(\frac{\mu_j}{2} + \frac{\sigma_j^2}{2\mu_j}\right). \tag{77}$$

Finally, since we can express the average age attained over

any route, we take the time average over all j:

$$\lambda^{MAD-ZW} = \sum_{j=1}^{N} \frac{p_j \mu_j \lambda_j^{av} \prod_{k=j+1}^{N} (1-p_k)}{\sum_{i=1}^{N} p_i \mu_i \prod_{k=i+1}^{N} (1-p_k)}, \qquad (78)$$

which gives (51).