A Theory of Second-Order Wireless Network Optimization and Its Application on AoI

Daojing Guo, Khaled Nakhleh, I-Hong Hou
Dept. of ECE, Texas A&M University
Email: {daojing_guo, khaled.jamal, ihou}@tamu.edu

Sastry Kompella, Clement Kam
Naval Research Laboratory
Email: {sk, ckk}@ieee.org

Abstract—This paper introduces a new theoretical framework for optimizing second-order behaviors of wireless networks. Unlike existing techniques for network utility maximization, which only considers first-order statistics, this framework models every random process by its mean and temporal variance. The inclusion of temporal variance makes this framework well-suited for modeling stateful fading wireless channels and emerging network performance metrics such as age-of-information (AoI). Using this framework, we sharply characterize the second-order capacity region of wireless access networks. We also propose a simple scheduling policy and prove that it can achieve every interior point in the second-order capacity region. To demonstrate the utility of this framework, we apply it for an important open problem: the optimization of AoI over Gilbert-Elliott channels. We show that this framework provides a very accurate characterization of AoI. Moreover, it leads to a tractable scheduling policy that outperforms other existing works.

I. INTRODUCTION

There are two seemingly contradictory trends happening in the field of wireless network optimization. On one hand, the study of network utility maximization (NUM) has witnessed tremendous success in the past two decades. Techniques based on dual decomposition, Lyapunov function, etc., have been shown to produce tractable and optimal solutions in complex networks for a wide range of objectives, including maximizing spectrum efficiency, minimizing power consumption, enforcing fairness among clients, and the combination of these objectives. Recent studies have also established iterative algorithms that not only converge to the optimum, but also have provably fast convergence rate [1]–[5]. On the other hand, there have been growing interests in new performance metrics for emerging network applications, such as quality-of-experience (QoE) for the application of video streaming and age-of-information (AoI) for the application of real-time state estimation. Surprisingly, except for a few special cases, the problem of optimizing these new performance metrics remains largely open. This raises the question: Why do existing NUM techniques fail to solve the optimization problem for these new performance metrics?

The fundamental reason is that current NUM techniques are only applicable to first-order performance metrics, while emerging new performance metrics involve higher-order behaviors. Existing NUM problems typically define the utility of a flow as $U_n(x_n)$, where $x_n$ is an asymptotic first-order performance metric, such as throughput (long-term average number of packet deliveries per unit time), power consumption (long-term average amount of energy consumption per unit time), and channel utilization (long-term average number of transmissions per unit time). However, emerging performance metrics like QoE and AoI require the characterization of short-term network behaviors, and hence cannot be fully captured by asymptotic first-order statistics.

To bridge the gap between NUM techniques and emerging performance metrics, we present a new framework of second-order wireless optimization. This framework consists of the second-order models, that is, the means and the temporal variances, of all random processes, including the channel qualities and packet deliveries of wireless clients. The incorporation of temporal variances enables this framework to better characterize stateful fading wireless channels, such as Gilbert-Elliott channels, and emerging performance metrics.

Using this framework, we sharply characterize the second-order capacity region of wireless networks, which entails the set of means and temporal variances of packet deliveries that are feasible under the constraints of the second-order models of channel qualities. As a result, the problem of optimizing emerging performance metrics is reduced to one that finds the optimal means and temporal variances of packet deliveries within the second-order capacity region. We also propose a simple scheduling policy and show that it can achieve every interior point of the second-order capacity region.

To demonstrate the utility of our framework, we apply it for an important open problem: Finding the optimal scheduling policy to minimize system-wide AoI over Gilbert-Elliott channels. We theoretically derive the closed-form expressions of the second-order models for Gilbert-Elliott channels. We also show that the AoI of each wireless client can be well-approximated by the mean and the temporal variance of its packet delivery process. We compare the system-wide AoI of our scheduling policy against other policies from recent studies on AoI minimization. Simulation results show that our policy achieves a smaller system-wide AoI. These results are especially significant when one considers that our policy is a generic second-order optimization policy, while the other policies are tailor-made to minimize the system-wide AoI.

The rest of the paper is organized as follows: Section II formally defines the second-order models of channel qualities and packet deliveries and the problem of second-order opti-
Section III uses second-order models to formulate the problem of minimizing system-wide AoI over Gilbert-Elliott channels. Section IV derives an outer bound of the second-order capacity region. Section V proposes a simple scheduling policy and shows that it achieves every interior point of the second-order capacity region. Section VI presents our simulation results. Section VII surveys some related studies. Finally, Section VIII concludes the paper.

II. System Model for Second-Order Wireless Network Optimization

We begin by describing a generic network optimization problem. Consider a wireless system where one AP serves $N$ clients, numbered as $\{1, 2, \ldots, N\}$. Time is slotted and denoted by $t = 1, 2, 3, \ldots$. We consider the ON-OFF channel model where the AP can schedule a client for transmission if and only if the channel for the client is ON. Let $X_i(t)$ be the indicator function that client $i$ is ON at time $t$. We assume that the sequence $\{X_i(1), X_i(2), \ldots\}$ is governed by a stochastic positive-recurrent Markov process with finite states. In each time slot, if there is at least one client having an ON channel, then the AP selects a client with an ON channel and transmits a packet to it. Let $Z_i(t)$ be the indicator function that client $i$ receives a packet at time $t$. The empirical performance of client $i$ is modeled as a function of the entire sequence $\{Z_i(1), Z_i(2), \ldots\}$. We note that the performance model is very general and covers virtually all existing network performance metrics, including both traditional ones like throughput and emerging ones like AoI. The network optimization problem is to find a scheduling policy that maximizes the total performance of the network.

Solving this generic network optimization problem is difficult because it requires solving an $N$-dimensional Markov decision process. As a result, except for a few special cases, there remains no tractable optimal solutions for many emerging network performance metrics like AoI. To circumvent this challenge, we propose capturing each random process by its second-order model, namely, its mean and temporal variance.

We first define the second-order model for channels. With a slight abuse of notation, let $X_S(t) := \max \{X_i(t) | i \in S\}$ be the indicator function that at least one client in $S$ has an ON channel at time $t$. Since all channels are governed by stochastic positive-recurrent Markov processes, the strong law of large numbers for Markov chains states that $\frac{\sum_{t=1}^{T} X_S(t)}{T}$ converges to a constant almost surely as $T \to \infty$. Hence, we can define the mean of $X_S$ as

$$m_S := \lim_{T \to \infty} \frac{\sum_{t=1}^{T} X_S(t)}{T}. \quad (1)$$

The Markov central limit theorem further states that $\frac{\sum_{t=1}^{T} X_S(t)-Tm_S}{\sqrt{\sum_{t=1}^{T} X_S(t)}}$ converges in distribution to a Gaussian random variable as $T \to \infty$. Hence, we define the temporal variance of $X_S$ as

$$v_S^2 := E[(\lim_{T \to \infty} \frac{\sum_{t=1}^{T} X_S(t) - Tm_S}{\sqrt{T}})^2]. \quad (2)$$

The second-order channel model is then expressed as the collection of the means and temporal variances of all $X_S$, namely, $\{(m_S, v_S^2) | S \subseteq \{1, 2, \ldots, N\}\}$.

The second-order model for packet deliveries is defined similarly. Assuming that the AP’s scheduling policy is ergodic, we can define the mean and the temporal variance of $Z_i$ as

$$\mu_i := \lim_{T \to \infty} \frac{\sum_{t=1}^{T} Z_i(t)}{T}, \quad \sigma_i^2 := E[(\lim_{T \to \infty} \frac{\sum_{t=1}^{T} Z_i(t) - T\mu_i}{\sqrt{T}})^2]. \quad (3)$$

The second-order delivery model is $\{ (\mu_i, \sigma_i^2) | 1 \leq i \leq N \}$. The performance client $i$ is modeled as a function of $(\mu_i, \sigma_i^2)$, which we denote by $F_i(\mu_i, \sigma_i^2)$.

Since clients want to have large means and small variances for their delivery processes, we define the second-order capacity region of a network as follows:

**Definition 1** (Second-order capacity region). Given a second-order channel model $\{(m_S, v_S^2) | S \subseteq \{1, 2, \ldots, N\}\}$, the second-order capacity region is the set of all $\{(\mu_i, \sigma_i^2) | 1 \leq i \leq N\}$ such that there exists a scheduling policy under which $\lim_{T \to \infty} \frac{\sum_{t=1}^{T} Z_i(t)}{T} = \mu_i$ and $E[(\lim_{T \to \infty} \frac{\sum_{t=1}^{T} Z_i(t) - T\mu_i}{\sqrt{T}})^2] \leq \sigma_i^2, \forall i$. \(\square\)

The second-order network optimization problem entails finding the scheduling policy that maximizes $\sum_{i=1}^{N} F_i(\mu_i, \sigma_i^2)$.

III. The Second-Order Model for AoI Optimization over Gilbert-Elliott Channels

To demonstrate the utility of our second-order models, we derive the second-order models for an important, but unsolved, problem: the optimization of AoI over Gilbert-Elliott channels.

A. The Second-Order Model of Gilbert-Elliott Channels

In Gilbert-Elliott channels [6], [7], the channel for each client $i$ is modeled as a two-state Markov process, as shown in Fig. 1. The channel is ON if it is in the good (G) state, and is OFF if it is in the bad (B) state. The transition probabilities from G to B and from B to G are $p_i$ and $q_i$, respectively. The channels are independent from each other.

We now show the second-order model of Gilbert-Elliott channels.

**Theorem 1.** Under the Gilbert-Elliott channels, for all $S$,

$$m_S = 1 - \prod_{i \in S} \frac{p_i}{p_i + q_i}. \quad (4)$$

$$v_S^2 = 2 \sum_{k=1}^{\infty} \left( \prod_{i \in S} G_i(k+1) - \prod_{i \in S} \frac{p_i}{p_i + q_i} \right) \prod_{i \in S} \frac{p_i}{p_i + q_i} + \prod_{i \in S} \frac{p_i}{p_i + q_i} - \left( \prod_{i \in S} \frac{p_i}{p_i + q_i} \right)^2. \quad (5)$$
where \( G_i(k) = \frac{p_k}{p_i + q_i} + \frac{q_k}{p_i + q_i} (1 - p_i - q_i)^{k-1} \).

**Proof.** Let \( Y_i(t) := 1 - X_i(t) \) be the indicator function that client \( i \) has an OFF channel at time \( t \). Let \( Y_S(t) := 1 - X_S(t) \) be the indicator function that all clients in the subset \( S \) have OFF channels at time \( t \). Hence, we have \( Y_S(t) = \prod_{i \in S} Y_i(t) \).

Suppose the Markov process of each channel is in the steady-state at time \( t \), then we have \( \text{Prob}(Y_i(t) = 1) = \frac{p_i}{p_i + q_i} \). Hence, \( E[Y_S(t)] = \sum_{i \in S} \frac{p_i}{p_i + q_i} \) and \( E[X_S(t)] = 1 - E[Y_S(t)] = 1 - \prod_{i \in S} \frac{p_i}{p_i + q_i} \). This establishes (4).

Next, we establish (5). We have

\[
E[Y_i(t)] = \prod_{k=1}^{\infty} Cov(Y_S(1), Y_S(1+k)).
\]

Since \( Y_S(1) \) is a Bernoulli random variable with mean \( \prod_{i \in S} \frac{p_i}{p_i + q_i} \), we have

\[
\text{Var}(Y_S(1)) = \sum_{i \in S} \frac{p_i}{p_i + q_i} - (\prod_{i \in S} \frac{p_i}{p_i + q_i})^2.
\]

Let \( G_i(k) = \text{Prob}(Y_i(k) = 1|Y_i(1) = 1) \). Then,

\[
E[Y_S(1)Y_S(1+k)] = \text{Prob}(Y_S(1+k) = 1|Y_S(1) = 1) \times \text{Prob}(Y_S(1) = 1)
\]

\[
= \text{Prob}(Y_i(1+k) = 1, \forall i \in S|Y_i(1) = 1, \forall i \in S) \prod_{i \in S} \frac{p_i}{p_i + q_i}
\]

\[
= \prod_{i \in S} G_i(k+1) \prod_{i \in S} \frac{p_i}{p_i + q_i},
\]

and

\[
\text{Cov}(Y_S(1), Y_S(1+k)) = E[Y_S(1)Y_S(1+k)] - E[Y_S(1)]E[Y_S(1+k)]
\]

\[
= \left( \prod_{i \in S} G_i(k+1) - \prod_{i \in S} \frac{p_i}{p_i + q_i} \right) \prod_{i \in S} \frac{p_i}{p_i + q_i}
\]

Combining (7) and (9) establishes (5).

It remains to find the closed-form expression of \( G_i(k) \). We have

\[
G_i(k) = \text{Prob}(Y_i(k) = 1|Y_i(1) = 1) = G_i(k-1)(1 - q_i) + (1 - G_i(k-1))p_i = p_i + (1 - p_i - q_i)G_i(k-1),
\]

if \( k > 1 \), and \( G_i(k) = 1 \), if \( k = 1 \). Solving this recursive equation yields

\[
G_i(k) = \frac{p_k}{p_i + q_i} + \frac{q_k}{p_i + q_i} (1 - p_i - q_i)^{k-1}.
\]

This completes the proof.

When \( p_i + q_i = 1 \), the Gilbert-Elliott channel reduces to the i.i.d. channel model where \( X_i(t) = 1 \) with probability \( q_i \), independent from any prior events. By replacing \( p_i = 1 - q_i \), we obtain the second-order model of i.i.d. channels as below:

**Corollary 1.** Under the i.i.d. channels with \( \text{Prob}(X_i(t) = 1) = q_i \),

\[
m_S = 1 - \sum_{i \in S} (1 - q_i), v_S^2 = \prod_{i \in S} (1 - q_i) - \prod_{i \in S} (1 - q_i)^2,
\]

for all \( S \). \( \square \)

**B. The Second-Order Model of AoI Optimization**

Age-of-Information (AoI) has been proposed to model the performance of real-time remote sensing applications, where a controller is obtaining status updates from a number of sensors. In a nutshell, the AoI corresponding to a sensor at a given time is defined as the age of the newest information update that it has ever delivered to the controller. In terms of our network model, the AP is the controller and each client is a sensor.

Similar to the case studied in \([8]\), we consider that each sensor \( i \) generates new updates by a Bernoulli random process. In each time slot \( t \), sensor \( i \) generates a new update with probability \( \lambda_i \), independent from any prior events. To minimize AoI, each sensor only keeps the most recent update in its memory, and it transmits the most recent update whenever it is scheduled for transmission. In other words, a sensor discards all its prior updates every time it generates a new update. The prior work \([8]\) considers that the controller knows when each sensor generates a new update. In practice, however, the controller cannot know whether a sensor has generated a new update until it schedules the sensor for transmission. In this paper, we further address the issue that the controller only knows \( \lambda_i \) but not the exact times at which sensors generate new updates. Hence, we assume that the scheduling decision is independent from update generations.

Let \( A_i(n) := \min\{\tau|\sum_{\tau=1}^n Z_i(t) = n\} \) be the time of the \( n \)-th delivery for client \( i \), and let \( B_i(n) := A_i(n+1) - A_i(n) \) be the time between the \( n \)-th and the \((n+1)\)-th deliveries. Since scheduling decisions are independent from update generations, we have the following:

**Lemma 1.** If \( \{B_i(0), B_i(1), \ldots\} \) is independent from the update generation processes of sensor \( i \), then the long-term average AoI of sensor \( i \) is

\[
\bar{\text{AoI}}_i = \frac{E[B_i^2]}{2E[B_i]} + \frac{1}{\lambda_i} - 1 - \frac{1}{2},
\]

where \( E[B_i^2] := \lim_{m \to \infty} \sum_{n=1}^m B_i(n)^2/m \) and \( E[B_i] := \lim_{m \to \infty} \sum_{n=1}^m B_i(n)/m \).

**Proof.** This lemma can be established by combining techniques in the proof of Proposition 2 in \([8]\) and the fact that \( B_i(n) \) is independent from update generations. The complete proof is omitted due to space limitation. \( \square \)

We aim to express \( \bar{\text{AoI}}_i \) as a function of the second-order delivery model of client \( i \), \( (\mu_i, \sigma_i^2) \). Since there can be multiple sequences of \( \{Z_i(1), Z_i(2), \ldots\} \) with the same \( (\mu_i, \sigma_i^2) \), we will derive \( \bar{\text{AoI}}_i \) with respect to a second-order reference delivery process as defined below.
Let $BM_{\mu_i, \sigma_i^2}(t)$ be a Brownian motion random process with mean $\mu_i$ and variance $\sigma_i^2$. An important property of the Brownian motion random process is that for any $t_1 < t_2$, $BM_{\mu_i, \sigma_i^2}(t_1) - BM_{\mu_i, \sigma_i^2}(t_2)$ is a Gaussian random variable with mean $(t_2 - t_1)\mu_i$ and variance $(t_2 - t_1)\sigma_i^2$. Our goal is to define a sequence $\{Z'_i(1), Z'_i(2), \ldots\}$ such that $\sum_{\tau=1}^{m} Z'_i(\tau) \approx BM_{\mu_i, \sigma_i^2}(t)$.

**Definition 2.** Given $(\mu_i, \sigma_i^2)$, the second-order reference delivery process, denoted by $\{Z'_i(1), Z'_i(2), \ldots\}$ is defined to be

$$Z'_i(t) = \begin{cases} 1 & \text{if } BM_{\mu_i, \sigma_i^2}(t) - BM_{\mu_i, \sigma_i^2}(t^-) \geq 1, \\ 0 & \text{else,} \end{cases}$$

where $t^- := \max\{\tau|\tau < t, Z'_i(\tau) = 1\}$. □

We now derive $\overline{AoI}$ with respect to the sequence $\{Z'_i(1), Z'_i(2), \ldots\}$. Consider the time between the $n$-th and the $(n + 1)$-th deliveries, which is denoted by $B_i(n)$, under the sequence $\{Z'_i(1), Z'_i(2), \ldots\}$. From [13], $B_i(n)$ can be approximated by the amount of time needed for the Brownian motion random process to increase by 1, which is equivalent to the first-hitting time for a fixed level 1 and we denote it by $H_i$. It has been shown that the first-hitting time for a fixed level 1 follows the inverse Gaussian distribution $\text{IG}(\frac{1}{\mu_i}, \frac{1}{\sigma_i^2})$ [9, 10]. Hence, we have $E[H_i] = \frac{1}{\mu_i}$ and $E[H_i^2] = \frac{\sigma_i^2}{\mu_i^3} + \frac{1}{\mu_i}$.

We now have

$$\overline{AoI} = \frac{E[B_i^2]}{2E[B_i]} + \frac{1}{\lambda_i} - \frac{1}{2} \approx \frac{E[H_i^2]}{2E[H_i]} + \frac{1}{\lambda_i} - \frac{1}{2} = \left(\frac{\sigma_i^2}{\mu_i^3} + \frac{1}{\mu_i}\right) + \frac{1}{\lambda_i} - \frac{1}{2}. \quad (14)$$

**C. Model Validation**

We now verify whether the second-order model provides a good approximation of $\text{AoI}$ over Gilbert-Elliott channels. We consider a system with only one client (sensor). The AP (controller) schedules the client for transmission whenever the client has an ON channel. Hence, we have $p_i = m(1)$ and $\sigma_i^2 = \frac{1}{(1)}$. Given, $p_1, q_1$, and $\lambda_1$, we can combine (1), (5), and (14) to obtain a theoretical approximation of the $\text{AoI}$. We note that (5) involves a summation of infinite terms $\sum_{k=1}^{\infty} G_1(k) \frac{p_1}{p_1 + q_1}$. Since $G_1(k)$ converges to $\frac{p_1}{p_1 + q_1}$ exponentially fast, we replace this term with $\sum_{k=1}^{100} G_1(k) \frac{p_1}{p_1 + q_1}$ when calculating $v_1^2$.

We evaluate the accuracy of the theoretical $\text{AoI}$ over a wide range of $(p_1, q_1, \lambda_1)$. For each $(p_1, q_1, \lambda_1)$, we obtain the empirical $\text{AoI}$ by simulation the system for 1000 runs, where each run contains 50,000 time slots. The results are shown in Fig. 2. It can be observed that the theoretical $\text{AoI}$ is almost identical to the empirical $\text{AoI}$ under all settings. The largest difference between theoretical and empirical $\text{AoI}$ among all evaluated case is only 0.00558.

**IV. AN OUTER BOUND OF THE SECOND-ORDER CAPACITY REGION**

In this section, we derive a necessary condition for the second-order delivery model $\{(\mu_i, \sigma_i^2)|1 \leq i \leq N\}$ to be in the second-order capacity region.

**Theorem 2.** Given a second-order channel model $\{(m_S, \sigma_S^2)|S \subseteq \{1, 2, \ldots, N\}\}$, a second-order delivery model $\{(\mu_i, \sigma_i^2)|1 \leq i \leq N\}$ can be in the second-order capacity region only if

$$\sum_{i \in S} \mu_i \leq m_S, \forall S \subseteq \{1, 2, \ldots, N\}, \quad (15)$$

$$\sum_{i=1}^{N} \mu_i = m\{1, 2, \ldots, N\}, \quad (16)$$

$$\sum_{i=1}^{N} \sqrt{\sigma_i^2} \geq \sqrt{\sigma_S^2}\{1, 2, \ldots, N\}, \quad (17)$$

$$\mu_i \geq 0, \forall i. \quad (18)$$

**Proof.** We first establish (15). The AP can transmit a packet to a client $i$ at time $t$ only if the client has an ON channel, that is, $X_i(t) = 1$. Moreover, the AP can transmit to at most one client in each time slot. Hence, we have $\sum_{i \in S} X_i(t) \leq X_S(t)$ under any scheduling policy. This gives us

$$\sum_{i \in S} \mu_i = \lim_{T \to \infty} \frac{1}{T} \sum_{i \in S} \sum_{t=1}^{T} X_i(t) \leq \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} X_S(t) = m_S, \forall S \subseteq \{1, 2, \ldots, N\}. \quad (19)$$

We can similarly establish (16) by noting that $\sum_{i=1}^{N} X_i(t) = X\{1, 2, \ldots, N\}(t)$, since the AP always transmits one packet as long as at least one client has an ON channel.

Finally, we establish (17). Let $\hat{X}_S$ be the random variable $\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} X_S(t) - TM_S$ and $Z_i$ be the random variable
\[ \lim_{T \to \infty} \sum_{i=1}^{N} \frac{Z_i(t) - T \mu_i}{\sqrt{T}} = X_{(1,2,\ldots,N)}(t) \] and (16), we have \( \sum_{i=1}^{N} Z_i(t) = \hat{X}_{(1,2,\ldots,N)} \). We then have
\[ \left( \sum_{i=1}^{N} \sqrt{\sigma_i^2} \right)^2 = \sum_{i=1}^{N} E[\hat{Z}_i^2] = \sum_{i \neq j} \left( \sum_{i=1}^{N} E[\hat{Z}_i^2] E[\hat{Z}_j^2] \right) \geq \sum_{i=1}^{N} E[\hat{Z}_i^2] + 2 \sum_{i \neq j} E[\hat{Z}_i \hat{Z}_j] \] (Cauchy-Schwarz inequality)
\[ = E\left[ \sum_{i=1}^{N} \hat{Z}_i^2 \right] = E[\hat{X}_{(1,2,\ldots,N)}^2] = \sigma_i^2 \] \( \text{for all } i \).

This completes the proof. \( \square \)

V. SCHEDULING POLICY WITH TIGHT INNER BOUND

In this section, we derive a sufficient condition for the second-order delivery model \( \{(\mu_i, \sigma_i^2)|1 \leq i \leq N\} \) to be in the second-order capacity region. We also propose a simple scheduling policy that delivers the desirable second-order delivery models as long as they satisfy the sufficient condition.

We state the sufficient condition as follows:

**Theorem 3.** Given a second-order channel model \( \{(m_S, \sigma_S^2)|S \subseteq \{1,2,\ldots,N\}\} \), a second-order delivery model \( \{(\mu_i, \sigma_i^2)|1 \leq i \leq N\} \) is in the second-order capacity region if

\[ \sum_{i \in S} \mu_i < m_S, \forall S \subseteq \{1,2,\ldots,N\}, \]
\[ \sum_{i=1}^{N} \mu_i = m_{\{1,2,\ldots,N\}}, \]
\[ \sum_{i=1}^{N} \sqrt{\sigma_i^2} \geq \sqrt{\sigma_S^2}, \]
\[ \mu_i \geq 0, \sigma_i^2 > 0 \forall i. \]

Before proving Theorem 3, we first discuss its implications. Comparing the conditions in Theorems 2 and 3, we note that the only difference is that the sufficient condition requires strict inequality for (15) for all proper subsets. Hence, the sufficient condition describes an inner bound that is almost tight except on some boundaries.

We prove Theorem 3 by proposing a scheduling that achieves every point in the inner bound. Given \( \{(\mu_i, \sigma_i^2)|1 \leq i \leq N\} \), define the **deficit** of a client \( i \) at time \( t \) as \( d_i(t) = t \mu_i - \sum_{\tau=1}^{t} Z_i(\tau) \). In each time slot \( t \), the AP chooses the client with the largest \( d_i(t-1)/\sqrt{\sigma_i^2} \) among those with ON channels and transmits a packet to the chosen client. We call this scheduling policy the **variance-weighted-deficit (VWD)** policy.

We now analyze the performance of the VWD policy. Let \( D(t) := \sum_{i=1}^{N} d_i(t)/\sqrt{\sum_{i=1}^{N} \sigma_i^2} \). We then have
\[ \Delta d_i(t) := d_i(t) - d_i(t-1) = \mu_i - Z_i(t), \]
\[ \Delta D(t) := D(t) - D(t-1) = \sum_{i=1}^{N} \mu_i - \sum_{i=1}^{N} Z_i(t) = \frac{m_{\{1,2,\ldots,N\}} - X_{(1,2,\ldots,N)}(t)}{\sqrt{\sum_{i=1}^{N} \sigma_i^2}}. \]

Consider the Lyapunov function \( L(t) := \frac{1}{2} \sum_{i=1}^{N} \frac{d_i(t)}{\sqrt{\sigma_i^2}} - D(t) \). Let \( H^t \) be the system history up to time \( t \). We can derive the expected one-step Lyapunov drift as
\[ \Delta(L(t)) := E[L(t) - L(t-1)|H^{t-1}] \]
\[ = E\left[ \frac{1}{2} \sum_{i=1}^{N} \frac{d_i(t)}{\sqrt{\sigma_i^2}} - D(t) \right]^2 \]
\[ - \frac{1}{2} \sum_{i=1}^{N} \frac{d_i(t-1)}{\sqrt{\sigma_i^2}} - D(t-1) \right]^2 |H^{t-1}] \]
\[ = E\left[ \frac{1}{2} \sum_{i=1}^{N} \frac{d_i(t-1)}{\sqrt{\sigma_i^2}} - D(t-1) \right]^2 |H^{t-1}] \]
\[ \leq B + E\left[ \sum_{i=1}^{N} \frac{d_i(t-1)}{\sqrt{\sigma_i^2}} - D(t-1) \right]\Delta d_i(t) \]
\[ - \sum_{i=1}^{N} \frac{d_i(t-1)}{\sqrt{\sigma_i^2}} - D(t-1) \right]\Delta D(t)|H^{t-1}] \]
\[ = B + E\left[ \sum_{i=1}^{N} \frac{d_i(t-1)}{\sqrt{\sigma_i^2}} - D(t-1) \right]\Delta d_i(t)|H^{t-1}], \]
where \( B \) is a bounded constant. The last two steps follow because \( \Delta d_i(t) \) and \( \Delta D(t) \) are bounded and because \( \sum_{i=1}^{N} d_i(t-1) = \sum_{i=1}^{N} \sqrt{\sigma_i^2} D(t-1) \).

The VWD policy schedules the client with the largest \( d_i(t-1)/\sqrt{\sigma_i^2} \), which is also the client with the largest \( d_i(t-1)/\sqrt{\sigma_i^2} - D(t-1) \), among those with ON channels. Hence, under the VWD policy, the system can be modeled as a Markov process whose state consists of the channel states and \( d_i(t-1)/\sqrt{\sigma_i^2} - D(t-1) \) of all clients. Further, the VWD policy is the policy that minimizes \( E[\sum_{i=1}^{N} \frac{d_i(t-1)}{\sqrt{\sigma_i^2}} - D(t-1) \right]\Delta d_i(t)|H^{t-1}] \) for all \( t \). We first show that the Markov process is positive-recurrent.

**Lemma 2.** Assume that (21) - (24) are satisfied. Then, under the VWD policy, the system-wide Markov process, whose state consists of the channel states and \( d_i(t-1)/\sqrt{\sigma_i^2} - D(t-1) \) of all clients, is positive-recurrent.

**Proof.** Due to (21), we can define
\[ \delta := \min\{m_S - \sum_{i \in S} \mu_i | S \subseteq \{1,2,\ldots,N\} \} > 0. \]
Further, since the channel of each client follows a positive-recurrent Markov process with finite states, there exists a finite number $T$ such that

$$Tm_{S} - \frac{\delta}{2} \leq E\left[\sum_{t=r+1}^{\tau+T} X_{S}(t)|H^{r}\right] \leq Tm_{S} + \frac{\delta}{2}, \quad (29)$$

for any $H^{r}$.

Let $L^{V}(t)$ and $\Delta d^{V}(t)$ be the values of $L(t)$ and $d_{i}(t)$ under the VWD policy. From (27), we can bound the $T$-step Lyapunov drift by

$$E[L^{V}(\tau + T) - L^{V}(\tau)|H^{r}] \leq BT + E\left[\sum_{t=r+1}^{\tau+T} \sum_{i=1}^{N} \left(\frac{d_{i}(t-1)}{\sqrt{\sigma_{i}^{2}}} - \frac{d(t-1)}{\sqrt{\sigma_{i}^{2}}}ight)\Delta d_{i}^{V}(t)|H^{r}\right] \leq BT + E[\sum_{t=r+1}^{\tau+T} \sum_{i=1}^{N} \left(\frac{d_{i}(t-1)}{\sqrt{\sigma_{i}^{2}}} - \frac{d(t-1)}{\sqrt{\sigma_{i}^{2}}}ight)\Delta d_{i}^{V}(t)|H^{r}] \leq A + E[\sum_{i=1}^{N} \left(\frac{d_{i}(\tau)}{\sqrt{\sigma_{i}^{2}}} - \frac{d(t)}{\sqrt{\sigma_{i}^{2}}}\right)(\sum_{t=r+1}^{\tau+T} \Delta d_{i}^{V}(t))|H^{r}], \quad (30)$$

for any other scheduling policy $\eta$, where $d_{i}^{V}(t)$ is the value of $d_{i}(t)$ under $\eta$ and $A$ is a bounded constant. The last inequality follows because $T$, $|d_{i}(t) - d_{i}(\tau)|$, and $\Delta d_{i}(t)$ are all bounded for all $t \in [\tau + 1, \tau + T]$.

We can consider the scheduling policy $\eta$ that schedules the flow with the largest $d_{i}(t)/\sqrt{\sigma_{i}^{2}}$ among those with ON channels in all time slots $t \in [\tau + 1, \tau + T]$.

Without loss of generality, we assume that $d_{1}(\tau)/\sqrt{\sigma_{1}^{2}} \geq d_{2}(\tau)/\sqrt{\sigma_{2}^{2}} \geq \ldots$. Under $\eta$, a client $i$ will be scheduled in time slot $t$ if it has an ON channel and all clients in $\{1, 2, \ldots, i-1\}$ have OFF channels, that is, $X_{1,2,\ldots,i-1}(t) = 1$ and $X_{1,2,\ldots,i-1}(t) = 0$. Hence have $\sum_{t=r+1}^{\tau+T} X_{1,2,\ldots,i}(t) = \sum_{t=r+1}^{\tau+T} X_{1,2,\ldots,i-1}(t) = \sum_{t=r+1}^{\tau+T} X_{1,2,\ldots,i-2}(t) = \sum_{t=r+1}^{\tau+T} X_{1,2,\ldots,i-3}(t) = \ldots$. Therefore,

$$E\left[\sum_{i=1}^{N} \left(\frac{d_{i}(\tau)}{\sqrt{\sigma_{i}^{2}}} - \frac{d(t)}{\sqrt{\sigma_{i}^{2}}}\right)(\sum_{t=r+1}^{\tau+T} \Delta d_{i}^{V}(t))|H^{r}\right] \leq A + E\left[\sum_{i=1}^{N} \left(\frac{d_{i}(\tau)}{\sqrt{\sigma_{i}^{2}}} - \frac{d(t)}{\sqrt{\sigma_{i}^{2}}}\right)(\sum_{t=r+1}^{\tau+T} \Delta d_{i}^{V}(t))|H^{r}\right] \leq \frac{d_{i}(\tau)}{\sqrt{\sigma_{i}^{2}}} - \frac{d(t)}{\sqrt{\sigma_{i}^{2}}},$$

where the inequality holds due to (22), (28), and (29).

Combining (30) and (31), and we have

$$E[L^{V}(\tau + T) - L^{V}(\tau)|H^{r}] \leq -\delta, \quad (32)$$

if $\max_{i} \left(\frac{d_{i}(\tau)}{\sqrt{\sigma_{i}^{2}}} - D(\tau)\right) > 2(A/\delta + 1)$, and

$$E[L^{V}(\tau + T) - L^{V}(\tau)|H^{r}] \leq A, \quad (33)$$

if $\max_{i} \left(\frac{d_{i}(\tau)}{\sqrt{\sigma_{i}^{2}}} - D(\tau)\right) \leq 2(A/\delta + 1)$. Recall that $\sum_{i} \left(\frac{d_{i}(\tau-1)}{\sqrt{\sigma_{i}^{2}}} - D(\tau-1)\right) = 0$ and the channel of each client follows a Markov process with finite states. Hence, all states of the system with $\max_{i} \left(\frac{d_{i}(\tau)}{\sqrt{\sigma_{i}^{2}}} - D(\tau)\right) \leq 2(A/\delta + 1)$ belong to a finite set of states. By the Foster-Lyapunov Theorem, the system-wide Markov process is positive-recurrent.

We now show that the VWD policy delivers all desirable second-order delivery models that satisfy the sufficient conditions (21) – (24), and thereby establishing Theorem 3.

**Theorem 4.** Assume that (27) – (24) are satisfied. Then, under the VWD policy, $\lim_{T \to \infty} \sum_{i=1}^{N} Z_{i}(t)/\sqrt{T} = \mu_{i}$ and $E[\lim_{T \to \infty} \sum_{i=1}^{N} Z_{i}(t)/\sqrt{T}]^{2} \leq \sigma_{i}^{2}, \forall i$.

**Proof.** Since the system-wide Markov process is positive recurrent under the VWD policy, we have:

$$\lim_{T \to \infty} \frac{d_{i}(T)/\sqrt{\sigma_{i}^{2}} - D(T)}{T} \to 0, \forall i, \quad (34)$$

$$\lim_{T \to \infty} \frac{d(T)/\sqrt{T} - D(T)}{\sqrt{T}} \to 0, \forall i. \quad (35)$$

First, we show that $\lim_{T \to \infty} \sum_{i=1}^{N} Z_{i}(t)/\sqrt{T} = \mu_{i}, \forall i$. Recall that $d_{i}(t) = t\mu_{i} - \sum_{t=1}^{T} Z_{i}(t)$ and $D(t) = \sum_{i=1}^{N} d_{i}(t)/\sum_{i=1}^{N} \mu_{i}$. By (22), we have:

$$\lim_{T \to \infty} \frac{D(T)/\sqrt{T} - D(T)}{T} = \lim_{T \to \infty} \frac{\sum_{i=1}^{N} T\mu_{i} - \sum_{i=1}^{N} \sum_{t=1}^{T} Z_{i}(t)}{T \sum_{i=1}^{N} \mu_{i}} = \lim_{T \to \infty} \frac{Tm_{1,2,\ldots,N} - \sum_{i=1}^{N} X_{1,2,\ldots,N}(t)}{T \sum_{i=1}^{N} \mu_{i}} = 0. \quad (36)$$

Hence, by (34), we have $\lim_{T \to \infty} \frac{d_{i}(T)}{T} = \mu_{i} - \lim_{T \to \infty} \frac{\sum_{i=1}^{N} Z_{i}(t)}{T} = 0, \forall i$.

Next, we show that $E[\lim_{T \to \infty} \sum_{i=1}^{N} Z_{i}(t)/\sqrt{T}]^{2} \leq \sigma_{i}^{2}, \forall i$. We have, by (23),

$$E[\lim_{T \to \infty} \frac{D(T)/\sqrt{T} - D(T)}{\sqrt{T}}^{2}] = \frac{\sqrt[2]{N}}{\sqrt{T}} (\sum_{i=1}^{N} \mu_{i})^{2} \leq 1, \quad (37)$$

and, hence,

$$E[\lim_{T \to \infty} \frac{\sum_{i=1}^{N} Z_{i}(t) - T\mu_{i}}{\sqrt{T}}^{2}] = E[\lim_{T \to \infty} \frac{d_{i}(T)}{\sqrt{T}}^{2}] \leq \sigma_{i}^{2}. \quad (38)$$

We conclude this section by discussing how to leverage Theorems 3 and 4 to solve the second-order network optimization problem. Recall that the performance of a client $i$ is modeled by $F_{i}(\mu_{i}, \sigma_{i}^{2})$. For example, when the goal is to minimize total...
Aöi, we can define $F_i(\mu_i, \sigma_i^2) = -\frac{1}{2}(\frac{\sigma_i^2}{\mu_i^2} + \frac{1}{\mu_i}) - \frac{1}{2}$. Hence, the second-order optimization problem can be written as the following:

$$\max \sum_{i=1}^{N} F_i(\mu_i, \sigma_i^2)$$

s.t. $\sum_{i \in S} \mu_i \leq m_S - \delta$, where $\delta$ is a small positive number. After the change, the optimization problem can be directly solved by standard solvers to find the optimal $\{\mu_i, \sigma_i^2| 1 \leq i \leq N\}$. After finding the optimal $\{\mu_i, \sigma_i^2| 1 \leq i \leq N\}$, one can use the VWD policy to attain the optimal network performance.

VI. Simulation Results

In this section, we present the simulation results for the proposed scheduler VWD. The objective is to minimize the total weighted Aöi, $\sum_{i} \alpha_i \text{Aöi}_i$, where $\alpha_i$ is the weight of client $i$. The system model is the one discussed in Section III. Each Gilbert-Elliott channel with transition probabilities $p_i$ and $q_i$. Each time slot, each client $i$ generates a new packet with probability $\lambda_i$. VWD is evaluated against three recent scheduling policies on this problem. We provide a description of each policy, along with modifications needed to fit the testing setting.

- **Whittle index policy**: This policy is based on the Whittle index policy in [11]. Under our setting, the policy calculates an index on ON clients based on their Aöi as $W_i(t) = \frac{A\text{AoI}_i(t)}{2} - \frac{A\text{AoI}_i(t)}{2} + \frac{\lambda_i}{1 - a_i}$. and then schedules the ON client with the largest index. [11] has shown that $W_i(t)$ is indeed the Whittle index of a client when the channel is i.i.d., i.e., $p_i + q_i = 1$, and $\lambda_i = 1$.

- **Stationary randomized policy**: This policy calculates a weight $\mu_i$ for each client. In each time slot, it randomly picks an ON client, with the probability of picking $i$ being proportional to $\mu_i$. In the setting of [8], it has been shown that, when $\mu_i$ is properly chosen, this policy achieves an approximation ratio of four in terms of total weighted Aöi. In our setting, we choose $\mu_i$ to be the optimal $\mu_i$ from solving [39].

- **Max weight policy [8]**: This policy schedules the ON client with the largest $(A\text{AoI}_i(t) - z_i(t))/\mu_i$. In the setting of [8], $z_i(t)$ is the time since client $i$ generates the latest packet. It has been shown that the total weighted AoI under this policy is no larger than that under the stationary randomized policy, and therefore this policy also achieves an approximation ratio of four. In our setting, the AP does not know when each client generates a new packet. Hence, we choose $z_i(t)$ to be $\frac{1}{\mu_i}$, which is the expected time since client $i$ generates the latest packet.

We consider three different systems, each with 5 clients, 10 clients, and 20 clients, respectively. For each system, $p_i$ and $q_i$ are randomly chosen from the range $[0.05, 0.95]$, and $\{\lambda_i\}$ is randomly chosen from $[\frac{0.1}{N}, \frac{1}{N}]$. After determining the values of $p_i$, $q_i$, and $\lambda_i$, we generate 1000 independent traces of channels and packet arrivals. The performance of each policy is the average over these 1000 independent traces. We consider both the unweighted case, i.e., $\alpha_i \equiv 1, \forall i$, and the weighted case. In addition to the evaluated policies, we also include the numerical solutions from solving the problem [39], which is referred to as the Theoretical Aöi.

A. Empirical Aöi Performance With Equal Aöi Weights

Fig. [3] shows the average total Aöi for different network sizes $N = \{5, 10, 20\}$ when $\alpha_i \equiv 1$. It can be observed that VWD achieves the smallest total Aöi in all systems, with max weight performing virtually the same as VWD when $N = 5$. VWD’s superiority becomes more significant as $N$ increases. It can also be observed that the empirical Aöi under VWD is very close to the theoretical Aöi based on the solution to [39], and the difference decreases as $N$ increases. The differences between the empirical Aöi under VWD and the theoretical one are 10.7%, 7.8%, and 6.1% for $N = 5, 10, 20$, respectively.

To understand why VWD performs much better than the other three policies, we evaluate the total empirical variance under each policy. Specifically, let $d_i(t)$ be the total number of packet deliveries for client $i$ from time 1 to time $t$. The
empirical variance of a client $i$ at time $t$ is defined as the variance of $\frac{d_i(t)}{\sqrt{t}}$ across all 1000 independent runs. The total empirical variance is then the sum of the empirical variances of all clients. Fig. 4 shows that VWD has much smaller variances than the other three policies. The ability to properly control variance enables VWD to achieve small AoIs.

We also evaluate the convergence time of VWD. For each system, we randomly select two clients and plot their empirical means, i.e., the average of $\frac{d_i(t)}{t}$ across all independent runs, and empirical variances. Since the objective is to minimize the unweighted sum of AoIs, the optimal solution to (39) has $\mu_i = \mu_j$ and $\sigma_i^2 = \sigma_j^2$ for all $i \neq j$. We call the optimal $\mu_i$ and $\sigma_i^2$ obtained from solving (39) the theoretical mean and the theoretical variance, respectively. The results are shown in Figs. 5 and 6. It can be observed that both the empirical means and the empirical variances of clients indeed converge to their
respective theoretical values. The empirical means converges to the theoretical ones very fast. On the other hand, it takes up to 355 slots for the empirical variances to be within 0.001 from the theoretical variances. This convergence time may be the reason why the empirical AoI is larger than the theoretical one.

B. Weighed Total AoI Evaluation

We now present the results for the weighted AoI. The weights $\alpha_1, \alpha_2, \ldots$ are randomly chosen from the range $(1, 5)$ and independently from each other. All other parameters are the same as in the unweighted case. Fig. 7 shows results for network sizes $N = \{5, 10, 20\}$. VWD still outperforms other policies for all tested systems. Similar to the unweighted case, it can be observed that the superiority of VWD becomes more significant, and the gap between VWD and theoretical AoI becomes smaller, with more clients in the system.

VII. RELATED WORKS

There have been many works on scheduling in wireless networks for minimizing AoI. In [12], the Tripathi and Moharir schedule over multiple orthogonal channels and propose Max-Age Matching and Iterative Max-Age Scheduling, which they show to be asymptotically optimal. Hsu, Modiano and Duan [13] studied the problem of scheduling updates for multiple clients where the updates arrive i.i.d. Bernoulli, and formulate the Markov decision process (MDP) and prove structural results and finite-state approximations. In [11], Hsu follows up this work by showing that a Whittle index policy can achieve near optimal performance with much lower complexity. Sun et al. [14] studied scheduling for multiple flows over multiple servers, and show that maximum age first (MAF)-type policies are nearly optimal for i.i.d. servers. In [15], Talak, Karaman and Modiano study scheduling a set of links in a wireless network under general interference constraints. The optimization of AoI and timely-throughput were studied in [16], [17]. All of these works assume i.i.d. channels.

There have been a limited number of works on Markov channel and source models related to AoI. In the recent work [18], Pan et al. study scheduling a single source and choosing between a Gilbert-Elliott channel and a deterministic lower rate channel. Buyukates and Ulukus [19] study the age-optimal policy for a system where the server is a Gilbert-Elliott model and one where the sampler follows a Gilbert-Elliott model. In [20], Nguyen et al. analyze the Peak Age of Information (PAoI) of a two-state Markov channel with differing cases of channel state information (CSI) knowledge. Kam et al. [21] study the remote estimation of a Markov source, and they propose effective age metrics that capture the estimation error. Our work differs in that we focus on scheduling for multiple clients from a single AP over parallel non-i.i.d. channels.

There have been some recent efforts on studying short-term performance through Brownian motion approximation [22]– [25], but each of them is limited to a specific channel model and a specific application.

VIII. CONCLUSION

In this paper, we presented a theoretical second-order framework for wireless network optimization. This framework captures the behaviors of all random processes by their second-order models, namely, their means and temporal variances. We analytically established a simple expression of the second-order capacity region of wireless networks. A new scheduling policy, VWD, was proposed and proved to achieve every interior point of the second-order capacity region. The framework utility is demonstrated by applying it to the problem of AoI optimization over Gilbert-Elliott channels. We derived closed-form expressions of second-order models for both Gilbert-Elliott channels and AoIs, and formulated the problem of minimizing weighted total AoI as an optimization problem over the means and temporal variances of delivery processes. The solution of this optimization problem can then be used as parameters for VWD. Simulation results show that VWD achieves much smaller weighted total AoI than other policies.
REFERENCES

[1] L. Huang, X. Liu, and X. Hao, “The power of online learning in stochastic network optimization,” in ACM SIGMETRICS Performance Evaluation Review, vol. 42, no. 1. ACM, 2014, pp. 153–165.

[2] L. Huang and M. J. Neely, "Delay reduction via lagrange multipliers in stochastic network optimization," in 2009 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks. IEEE, 2009, pp. 1–10.

[3] J. Liu, A. Eryilmaz, N. B. Shroff, and E. S. Bentley, “Heavy-ball: A new approach to tame delay and convergence in wireless network optimization,” in IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications. IEEE, 2016, pp. 1–9.

[4] T. Chen, Q. Ling, and G. B. Giannakis, “Learn-and-adapt stochastic dual gradients for network resource allocation,” IEEE Transactions on Control of Network Systems, vol. 5, no. 4, pp. 1941–1951, 2017.

[5] J. Liu, “Achieving low-delay and fast-convergence in stochastic network optimization: A nesterovian approach,” in ACM SIGMETRICS Performance Evaluation Review, vol. 44, no. 1. ACM, 2016, pp. 221–234.

[6] E. N. Gilbert, “Capacity of a burst-noise channel,” The Bell System Technical Journal, vol. 39, no. 5, pp. 1253–1265, 1960.

[7] E. O. Elliott, “Estimates of error rates for codes on burst-noise channels,” The Bell System Technical Journal, vol. 42, no. 5, pp. 1977–1997, 1963.

[8] I. Kadota and E. Modiano, “Minimizing the age of information in wireless networks with stochastic arrivals,” IEEE Transactions on Mobile Computing, 2019.

[9] E. Schrödinger, “Zur Theorie der Fall- und Steigversuche an Teilchen mit Brownscher Bewegung,” Physikalische Zeitschrift, vol. 16, pp. 289–295, 1915.

[10] J. L. Folks and R. S. Chhikara, “The inverse gaussian distribution and its statistical application—a review,” Journal of the Royal Statistical Society: Series B (Methodological), vol. 40, no. 3, pp. 263–275, 1978.

[11] Y.-P. Hsu, “Age of information: Whittle index for scheduling stochastic arrivals,” in 2018 IEEE International Symposium on Information Theory (ISIT), 2018, pp. 2634–2638.

[12] V. Tripathi and S. Moharir, “Age of information in multi-source systems,” in GLOBECOM 2017 - 2017 IEEE Global Communications Conference, 2017, pp. 1–6.

[13] Y.-P. Hsu, E. Modiano, and L. Duan, “Age of information: Design and analysis of optimal scheduling algorithms,” in 2017 IEEE International Symposium on Information Theory (ISIT), 2017, pp. 561–565.

[14] Y. Sun, E. Uysal-Biyikoglu, and S. Kompella, “Age-optimal updates of multiple information flows,” in IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2018, pp. 136–141.

[15] R. Talak, S. Karaman, and E. Modiano, “Optimizing information freshness in wireless networks under general interference constraints,” IEEE/ACM Transactions on Networking, vol. 28, no. 1, pp. 15–28, 2020.

[16] I. Kadota, A. Sinha, and E. Modiano, “Optimizing age of information in wireless networks with throughput constraints,” in IEEE INFOCOM 2018 - IEEE Conference on Computer Communications, 2018, pp. 1844–1852.

[17] N. Lu, B. Ji, and B. Li, “Age-based scheduling: Improving data freshness for wireless real-time traffic,” in Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing, New York, NY, USA: Association for Computing Machinery, 2018, p. 191–200.

[18] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, “Minimizing age of information via scheduling over heterogeneous channels,” arXiv preprint arXiv:2012.09403, 2020.

[19] B. Buyukates and S. Ulukus, “Age of information with gilbert-elliot servers and samplers,” arXiv preprint arXiv:2002.05711, 2020.

[20] G. D. Nguyen, S. Kompella, C. Kam, and J. E. Wieselthier, “Information freshness over a markov channel: The effect of channel state information,” Ad Hoc Networks, vol. 86, pp. 63–71, 2019.

[21] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, “Towards an effective age of information: Remote estimation of a markov source,” in IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2018, pp. 367–372.

[22] P.-C. Hsieh and I.-H. Hou, “Heavy-traffic analysis of qoe optimality for on-demand video streams over fading channels,” in IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications, 2016, pp. 1–9.

[23] I.-H. Hou and P.-C. Hsieh, “The capacity of qoe for wireless networks with unreliable transmissions;” Queueing Systems, vol. 87, no. 1, pp. 131–159, 2017.

[24] P.-C. Hsieh, X. Liu, and I.-H. Hou, “Fresher content or smoother playback? a brownian-approximation framework for scheduling real-time wireless video streams,” in Proceedings of the Twenty-First International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing. New York, NY, USA: Association for Computing Machinery, 2020, p. 41–50.

[25] D. Guo, P.-C. Hsieh, and I.-H. Hou, “Optimal wireless scheduling for remote sensing through brownian approximation,” in IEEE INFOCOM 2021 - IEEE Conference on Computer Communications, 2021, to appear.