How to Price Fresh Data

Meng Zhang, Ahmed Arafa, Jianwei Huang, and H. Vincent Poor

***Abstract*—We introduce the concept of a *fresh data market*, in which a destination user requests, and pays for, fresh data updates from a source provider. Data freshness is captured by the *age of information* (AoI) metric, defined as the time elapsed since the latest update has reached the destination. The source incurs an operational cost, modeled as an increasing convex function of the number of updates. The destination incurs an age-related cost, modeled as an increasing convex function of the AoI. The source charges the destination for each update and designs a pricing mechanism to maximize its profit; the destination on the other hand chooses a data update schedule to minimize the summation of its payments to the source and its age-related cost. The interaction among the source and destination is hence game-theoretic. Motivated by the existing pricing literature, we first study a *time-dependent* pricing scheme, in which the price for each update depends on when it is requested. We show in this case that the game equilibrium leads to only one data update, which does not yield the maximum profit to the source. This motivates us to consider a *quantity-based* pricing scheme, in which the price of each update depends on how many updates have been previously requested. We show that among all pricing schemes in which the price of an update may vary according to both time and quantity, the quantity-based pricing scheme performs best: it maximizes the source's profit and minimizes the social cost of the system, defined as the aggregate source's operational cost and the destination's age-related cost. Numerical results show that the optimal quantity-based pricing can be 27% more profitable for the source and incurs 54% less social cost, compared with the optimal time-dependent pricing.**

I. INTRODUCTION

A. Motivation

Information usually has the greatest value when it is fresh [1, p. 56]. Data freshness is becoming increasingly significant due to the fast growth of the number of mobile devices and the dramatic increase of real-time applications: news updates, traffic alerts, stock quotes, and social media updates. In addition, timely information updates are also critical in real-time monitoring, data analytics, and control systems. For instance, real-time knowledge of traffic information and the speed of motor vehicles is crucial in autonomous driving and unmanned aerial vehicles. Another instance is for phasor data updates in power grid stabilization systems and application program interface (API) monitoring [2]. Examples of real-time datasets include real-time map data and traffic data, such as the Google Maps Platform [3]. A suitable candidate metric to measure the freshness of data is the *age of information* (AoI) metric, introduced in [4], [5], which measures the amount of time elapsed since the most recent data update.

M. Zhang and J. Huang are with Department of Information Engineering, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, E-mail: {zm015, jwhuang}@ie.cuhk.edu.hk; J. Huang is also with School of Engineering and Science, The Chinese University of Hong Kong, Shenzhen; A. Arafa and H. V. Poor are with the Department of Electrical Engineering, Princeton University, NJ, USA, E-mail: {aarafa, poor}@princeton.edu. This work was supported in part by the Global Scholarship Programme for Research Excellence from CUHK, in part by the Overseas Research Attachment Programme from School of Engineering at CUHK, in part by the General Research Fund CUHK 14219016 from Hong Kong UGC, in part by the Presidential Fund from the Chinese University of Hong Kong, Shenzhen, and in part by the U.S. National Science Foundation under Grants CCF-0939370 and CCF-1513915.

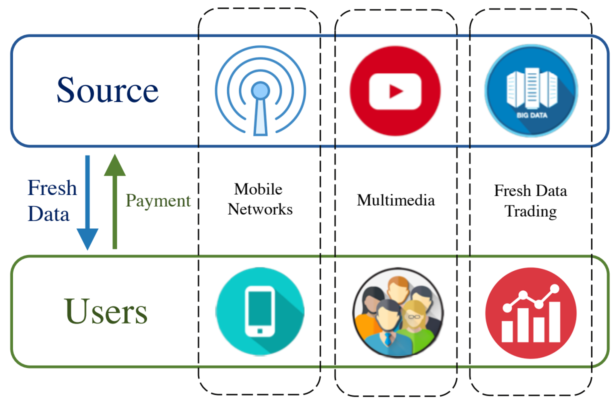

Fig. 1. Examples of potential fresh data markets.

However, the availability of fresh data relies on frequent data generation, processing, and transmission, which can lead to significant operational costs for the data provider. Such operational costs make the pricing design of an essential role in the data market, as pricing provides an incentive for the data provider to update the data and prohibits the data users (receivers) from requesting data updates unnecessarily often. This is quite different from the traditional study of pricing in networks, which often aims at maximizing a network operator's revenue and control the network congestion level. The pricing for fresh data is under-explored, as all existing pricing schemes for communication systems assume that a consumer's satisfaction with the service depends mainly on the quantity/quality of the service received without considering its timeliness. Such an interaction between data providers and users requesting fresh data leads to the *fresh data market*, examples of which are shown in Fig. 1. This paper tries to partially fill in the gap by considering a single source-destination pair, and addressing the following key question:

Question 1. *How should the source choose the pricing scheme to maximize its profit in a fresh data market?*

B. Solution Approach and Contributions

As the first step toward studying the pricing mechanism design for fresh data, we consider two types of pricing schemes. The first one is a *time-dependent* pricing scheme, in which the source of fresh data prices each data update based on the time at which the update is requested. Due to the nature of the AoI, the destination's desire for updates increases as time (since the most recent update) goes by, which makes it potentially profitable to explore this time sensitivity. This pricing scheme is also motivated by many existing time-dependent pricing schemes (in which users are not age-sensitive) of mobile networks, e.g., [6]–[14].

The second pricing scheme that we consider is a *quantity-based* pricing scheme, in which the price for each update depends on the number of updates requested so far (but does not depend on the timing of the updates). Such a pricing scheme is also known as *second-degree price discrimination* or *volume discount* [15], and is motivated by practical pricing schemes (e.g., for mobile data plans and data analytics [2]).

The challenge of designing a proper pricing scheme for fresh data is two-fold. First, different from the classical pricing setting, e.g., [6]–[14], the demands for fresh data over time are interdependent due to the nature of AoI. That is, the desire for an update at each time instance depends on the time elapsed since the latest update. Hence, the source's pricing scheme choice needs to take such interdependence overall the entire period into consideration. Second, in the case of the time-dependent pricing scheme design, one needs to optimize a continuous-time pricing function, i.e., solve an infinite dimensional optimization problem. The above discussion motivates our consideration of the following question:

Question 2. *How profitable is it for the source to exploit the time sensitivity in designing the pricing scheme for fresh data?*

We summarize our approaches and contributions as follows:

- *Fresh Data Market Modeling.* To the best of our knowledge, this paper presents the first model of a fresh data market, in which an age-sensitive destination interacts with a source data provider.
- *Time-Dependent Pricing Scheme.* We study a time-dependent pricing scheme for the fresh data market, aiming at exploiting the time sensitivity. We show that at the optimal (equilibrium) time-dependent pricing scheme, the source sends only one update, and hence exploiting time sensitivity may not enhance profitability.
- *Quantity-Based Pricing Scheme.* We propose a quantity-based pricing scheme, and show that it is more profitable than the time-dependent pricing scheme. We further prove that it maximizes the profit among all classes of time-and-quantity dependent pricing schemes, and that it minimizes the social cost of the system: the sum of the source's operational cost and the destination's age-related cost.
- *Simulation Results.* The numerical results show that the optimal quantity-based pricing scheme can be 27% more profitable and incurs 54% less social cost, compared with the optimal time-dependent pricing scheme on average.

We organize the rest of this paper as follows. In Section II, we discuss some related work. In Section III, we describe the system model and the game-theoretic problem formulation. In Sections IV and V, we develop the time-dependent and the quantity-based pricing schemes, respectively. We then relate the two schemes and mention some relevant properties in Section VI. We provide some numerical results in Section VII to evaluate the performance of the two pricing schemes, and conclude the paper in Section VIII.

II. Related Work

The concept age-of-information was first proposed as a metric of data freshness in the studies of databases [4], [5] in the 1990s. In recent years, there have been many excellent works focusing on the optimization of scheduling policies in terms of minimizing the AoI in various system settings, see, e.g., [16]–[33]. In [16], Kaul *et al.* recognized the importance of real-time status updates in networks. In [17], [18], He *et al.* investigated the NP-hardness of minimizing the AoI in scheduling general wireless networks. In [19], Kadota *et al.* studied the scheduling problem in a wireless network with a single base station and multiple destinations. In [20], Kam *et al.* investigated the AoI for a status updating system through a network cloud. In [21], Sun *et al.* studied the optimal management of the fresh information updates. References [22] and [23] studied the optimal wireless network scheduling with the interference constraint and the throughput constraint, respectively. The AoI consideration has recently gained some attention in energy harvesting communication systems, e.g., [24]–[28] and Internet of Things systems, e.g., [29], [30]. Several existing studies focused on game-theoretic interactions in interference channels, without considering the interactions in a fresh data market or the pricing scheme design [31]–[33].

There exists a rich literature on the pricing mechanism design and revenue management in communication networks (please refer to [6]–[14], surveys in [34], [35], and references therein). Specifically, time-dependent pricing has also been extensively studied, e.g., [6]–[11] while a few works focused on the quantity-based pricing and other forms of price differentiation for the Internet service providers, e.g., [12]–[14]. These works assumed that destinations are only interested in the throughput/rate received instead of the data freshness.

References [36], [37] are the most closely-related works to ours. In [36], a repeated game is studied between two AoI-aware platforms, yet without studying pricing schemes. While in [37], the authors considered a system in which the destination designs a dynamic pricing scheme to incentivize sensors to provide fresh updates, with random data arrivals. Different from [37], our considered pricing schemes are designed by the source, which is motivated by most practical communication/data systems in which sources are price designers while the destinations are price takers.

III. System Model

A. System Overview

1) Single-Source Single-destination System: We consider an information update system, in which one source node generates data packets and sends them to one destination through a channel.[1]

[1]We note that the single-source single-destination model has been widely considered in the AoI literature (e.g., [20], [21], [24], [26], [28]). In addition, the insights derived from this model allow us to potentially extend the results to the multi-destination scenarios.

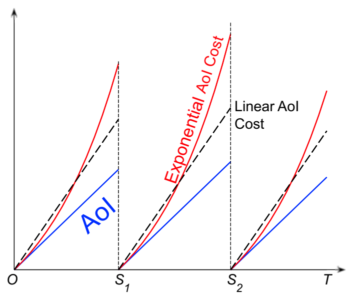

Fig. 2. Illustrations of AoI Δ_t and two types of AoI costs $f(\Delta_t)$. There are two updates at S_1 and S_2.

2) Data Updates and Age of Information: We consider a fixed time period of $\mathcal{T} = [0, T]$, during which the source sends its updates to the destination. We consider a generate-at-will model, as in, e.g., [24]–[28], in which the source is able to generate and send a new update when requested by the destination. Updates reach the destination instantly, with negligible transmission time, as in, e.g., [25]–[27].

We denote by $S_k \in \mathcal{T}$ the transmission time of the k-th update. The set of all update time instances is $\mathcal{S} \triangleq \{S_k\}$. Let K denote the number of total updates, i.e., $|\mathcal{S}| = K$, where $|\cdot|$ denotes the cardinality of a set. The set $\mathcal{S}$ (and hence the value of K) is a decision variable of the destination.

To measure the freshness of data, let us define the AoI $\Delta_t(\mathcal{S})$ at time t as [4], [16]

$$\Delta_t(\mathcal{S}) = t - U_t, \tag{1}$$

where U_t is the time stamp of the most recently received update before time t, i.e., $U_t = \max_{S_k \leq t}\{S_k\}$.

3) Source's Operational Cost and Pricing: We denote the source's operational cost by $C(K)$, which is modeled as an *increasing convex* function in the number of updates K, with $C(0) = 0$.[2] This can represent transmission costs in case the source is a network operator, or the cost of generating, processing and transmission commission in case the source is a content/data provider.

The source designs the pricing scheme for sending the data updates. We consider a general scheme in which the price for a particular update may depend on the time of the update request and the number of previously requested updates. We denote by $p(t,k) : \mathcal{T} \times \mathbb{N} \to \mathbb{R}_+$ the pricing function, with $p(t', k')$ being the price of the k'th update if requested at time t'. Note that we denote by $p(t,k)$ the function itself, while we denote by $p(t', k')$, i.e., using any argument other than (t,k) specifically, the value of the function. As mentioned, such a pricing scheme is motivated by (i) the time-sensitive demand for an update due to the nature of AoI, and (ii) the wide consideration of both time-dependent and quantity-based pricing schemes in practice. Under a pricing scheme $p(t,k)$, the destination's total payment to the source over the entire period is $P(\mathcal{S}) = \sum_{k=1}^{K} p(S_k, k)$.

4) Destination's AoI Cost: Besides the payment $P(\mathcal{S})$, the destination also experiences an AoI cost $f(\Delta_t)$ related to the destination's desire for the new data update.[3] We assume that $f(\Delta_t)$ is *increasing* and *convex* in Δ_t.[4] Let $\Gamma(\mathcal{S})$ denote the aggregate AoI cost over the entire period $\mathcal{T}$, defined as

$$\Gamma(\mathcal{S}) \triangleq \int_0^T f(\Delta_t(\mathcal{S}))dt. \tag{2}$$

Fig. 2 illustrates the AoI, an exponential AoI cost, and a linear AoI cost.

Stage I
The source determines the pricing scheme $p(t,k)$.

$$\Downarrow$$

Stage II
The destination determines its update policy $\mathcal{S}$.

Fig. 3. Two-stage Stackelberg game.

B. Stackelberg Game

We model the interaction between the source and the destination as a two-stage Stackelberg game as shown in Fig. 3. Specifically, in Stage I, the source determines the pricing scheme function $p(t,k)$ at the beginning of the period, in order to maximize its profit, given by the payment it receives minus its operational cost, as follows:

$$\textbf{Source}: \max_{p(t,k)} \; P(\mathcal{S}^*(p(t,k))) - C(|\mathcal{S}^*(p(t,k))|), \tag{3a}$$

$$\text{s.t.} \quad p(t', k') \geq 0, \; \forall t' \in \mathcal{T}, k' \in \mathbb{N}, \tag{3b}$$

where $\mathcal{S}^*(p(t,k))$ is the destination's optimal update policy, in response to the pricing scheme chosen by the source, which is defined below.

In Stage II, the destination decides its update policy to minimize its overall cost (aggregate AoI cost plus payment):

$$\textbf{Destination}: \mathcal{S}^*(p(t,k)) = \arg\min_{\mathcal{S} \in \Phi} \; \Gamma(\mathcal{S}) + P(\mathcal{S}), \tag{4}$$

where Φ is the set of all feasible $\mathcal{S}$, given by $\Phi = \cup_{K' \in \mathbb{N}} \Phi^{K'}$ and $\Phi^{K'}$ is the set of all transmission times $\mathcal{S} = \{S_1, S_2, ..., S_{K'}\}$, with $S_k' \in \mathcal{T}$ and $S_k \geq S_{k-1}$ for all $1 \leq k \leq K'$.

In the following two sections, we will separately consider two special cases of $p(t,k)$: $p(t)$ and $p_q(k)$. We note that analyzing the simplified pricing schemes is still challenging. First, the optimal pure time-dependent pricing scheme involves solving an infinite-dimensional optimization problem. Second, the pricing scheme needs to take the optimal decisions $\mathcal{S}^*(\cdot)$ over the whole period into consideration.

IV. Time-Dependent Pricing Scheme

In this section, we consider a (pure) *time-dependent* pricing scheme, in which the price function only depends on the time at which the update is requested and does not depend on the number of updates.

[2] Formally, $C(\cdot)$ is defined on the set of non-negative real numbers $\mathbb{R}_+$, and then evaluated on the set of natural numbers $\mathbb{N}$.

[3] As the first work considering the pricing scheme design for fresh data, we assume the benefit of receiving the data is constant, i.e. independent of the total number of updates.

[4] An example of this AoI cost model exists in the online learning in real-time applications such as online advertisement placement and online Web ranking, in which fresh data is critical [38]–[40].

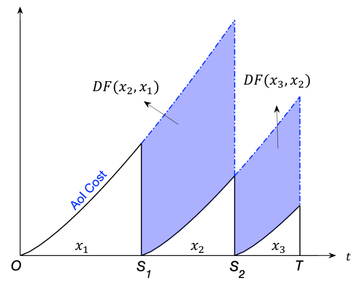

Fig. 4. An illustrative example of the differential aggregate AoI cost function and Lemma 1.

We derive the *(Stackelberg)* equilibrium price-update profile $(p^*(t), \mathcal{S}^*(p^*(t)))$ using backward induction. First, given any pricing scheme $p(t)$ in Stage I, we characterize the destination's update policy $\mathcal{S}^*(p(t))$ that minimizes its overall cost in Stage II. Then in Stage I, by characterizing the equilibrium pricing structure, we convert the continuous pricing function optimization into a vector one, based on which we characterize the source's optimal pricing scheme $p^*(t)$.

A. Destination's Update Policy in Stage II

Recall that K is the total number of updates. Let x_k denote the kth interarrival time, which is the time elapsed between the generation of $(k-1)$-th update and k-th update, i.e., x_k is

$$x_k \triangleq S_k - S_{k-1},\ \forall k \in \mathcal{K}(K+1), \tag{5}$$

where $\mathcal{K}(K+1) = \{1, 2, ..., K+1\}$, $S_0 = 0$, and $S_{K+1} = T$.

To analyze the aggregate AoI cost function $\Gamma(\mathcal{S})$ in (2), we define

$$F(x) \triangleq \int_0^x f(t)dt, \tag{6}$$

based on which we have $\Gamma(\mathcal{S}) = \sum_{k\in\mathcal{K}(K+1)} F(x_k)$.

Given the pricing scheme $p(t)$, the destination's problem in (4) is equivalent to

$$\min_{K\in\mathbb{N}\cup\{0\}, \boldsymbol{x}\in\mathbb{R}_{++}^{K+1}} \quad \sum_{k=1}^{K+1} F(x_k) + \sum_{k=1}^{K} p\left(\sum_{j\le k} x_j\right), \tag{7a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K+1} x_k = T, \tag{7b}$$

where $\boldsymbol{x} = \{x_k\}_{k\in\mathcal{K}(K+1)}$ and $\mathbb{R}_{++}^{K+1}$ is the space of $(K+1)$-dimensional positive vectors (i.e., the value of every entry is positive).

To understand when the destination would choose to update, we define the differential aggregate AoI cost function as

$$DF(x, y) \triangleq \int_0^x [f(t+y) - f(t)]dt. \tag{8}$$

As illustrated in Fig. 4, for each update k, $DF(x_{k+1}, x_k)$ is the aggregate AoI cost increase if the destination changes its update policy from $\mathcal{S}$ to $\mathcal{S}\backslash\{S_k\}$ (removes the update at S_k). We now introduce the following lemma:

Lemma 1. *Any equilibrium price-update tuple $(p^*(t), K^*, \boldsymbol{x}^*)$ should satisfy*

$$p^*\left(\sum_{j=1}^{k} x_j^*\right) = DF(x_{k+1}^*, x_k^*), \quad \forall k \in \mathcal{K}(K^*+1). \tag{9}$$

Proof Sketch: For each k-th update, the differential aggregate AoI cost equals the destination's maximal willingness to pay. Hence, if $p^*\left(\sum_{j=1}^{k} x_j^*\right) > DF(x_{k+1}^*, x_k^*)$, then the destination would prefer not to update at S_k, contradicting to the fact that $(p^*(t), K^*, \boldsymbol{x}^*)$ is an equilibrium. In addition, if $p^*\left(\sum_{j=1}^{k} x_j^*\right) < DF(x_{k+1}^*, x_k^*)$, we can show that the source can always properly increase $p^*\left(\sum_{j=1}^{k} x_j^*\right)$. The increase in the price does not change the destination's optimal solution $(K^*, \boldsymbol{x}^*)$, and hence increases the source's profit. This contradicts the fact that $(p^*(t), K^*, \boldsymbol{x}^*)$ is an equilibrium. $\square$

Note that given that the optimal pricing scheme satisfies (9), there might exist multiple optimal update policies as the solutions of problem (4). This may lead to a multi-valued source's profit and thus an ill-defined problem (3). To ensure the uniqueness of the received profit for the source, one can impose infinitely large prices to ensure that the destination does not update at any time instance other than $\sum_{j=1}^{k} x_j^*$ for all $k \in \mathcal{K}(K^*+1)$.

B. Source's Time-Dependent Pricing Design in Stage I

Based on Lemma 1, we can reformulate the time-dependent pricing scheme as follows (the proof is omitted due to space limits).

Theorem 1. *The time-dependent pricing problem in* (3) *is equivalent to the following problem:*

$$\max_{K\in\mathbb{N}\cup\{0\}, \boldsymbol{x}\in\mathbb{R}_{++}^{K+1}} \quad \sum_{k=1}^{K} DF(x_{k+1}, x_k) - C(K), \tag{10a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K+1} x_k = T. \tag{10b}$$

The decision variables in problem (10) correspond to the interarrival time interval vector $\boldsymbol{x}$ instead of the continuous-time pricing function $p(t)$. By converting a continuous function optimization problem into a vector optimization problem, we significantly simplify the problem. We are now ready to present the following result:

Theorem 2. *There will be only one update (i.e., $K^* = 1$) under any equilibrium time-dependent pricing scheme.*

One can prove Theorem 2 by induction, showing that for an arbitrary time-dependent pricing scheme yielding more than $K > 1$ updates (K-update pricing), there always exists a pricing scheme leading to a single-update equilibrium that is more profitable. The following example illustrates this with a linear AoI cost function:

Example 1. *Consider a linear AoI cost $f(\Delta_t) = \Delta_t$ and an arbitrary update policy $(K, \boldsymbol{x})$, as shown in Fig. 5.*

- ***Base case:*** *When there are $K = 2$ updates, as shown in Fig. 5, the source's profit (the objective value in* (10a)*) is $x_1x_2 + x_2x_3 - C(2)$. Consider another update policy*

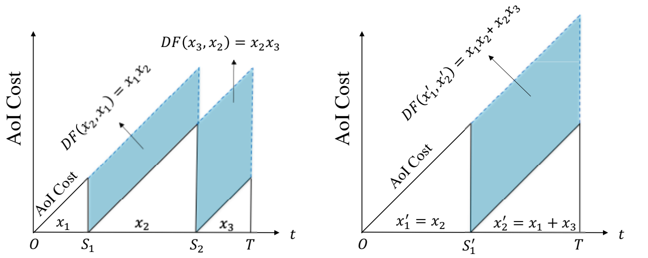

Fig. 5. Illustrations of Example 1 with a linear cost function. Combining the first interval into the third interval maintains the payment.

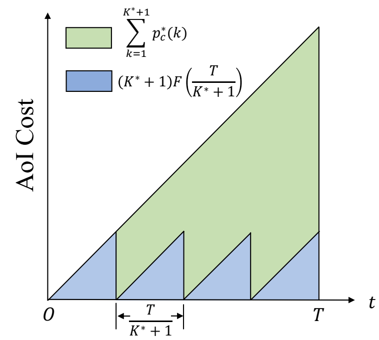

Fig. 6. An illustrative example of Proposition 2 with an linear AoI cost function $f(\Delta_t) = \Delta_t$ and $K^* = 3$.

$(K' = 1, x'_1, x'_2)$ where $x'_1 = x_2$ and $x'_2 = x_3 + x_1$. The objective value in (10a) *becomes $x_2(x_1 + x_3) - C(1)$. Comparing these two values, we see that (K', x'_1, x'_2) is strictly more profitable than $(K, \boldsymbol{x})$.*

- ***Induction step:*** *Let $K \geq n$ and suppose the statement that, for an arbitrary K-update pricing, there exists a more profitable $(K-1)$-update pricing is true for $K = n$. The objective value in* (10a) *is $\sum_{k=1}^{K} x_k x_{k+1} - C(K)$. Consider another update policy $(K' = K-1, \boldsymbol{x}')$ where $x'_1 = x_2$, $x'_2 = x_3 + x_1$, and $x'_k = x_{k+1}$ for all other k. The objective value in* (10a) *becomes $(x_1 + x_3)(x_2 + x_4) + \sum_{k=4}^{K} x_k x_{k+1} - C(K-1)$, which is strictly larger than P. It is then readily verified that $(K', \boldsymbol{x}')$ is strictly more profitable than $(K, \boldsymbol{x})$. Based on induction, we can show that we can find a $(K'-1)$-update policy would be more profitable than the $(K', \boldsymbol{x}')$ policy. This eventually leads to the conclusion that a single update policy is the most profitable.*

Based on the above technique, we can show that the above argument works for any increasing convex AoI cost function. The complete proof is omitted due to space limits.

To rule out trivial cases in which there is no update at the equilibrium, we adopt the following assumption:

Assumption 1. *The source's operational cost function $C(K)$ satisfies $C(1) \leq DF(T/2, T/2)$.*

Assumption 1 ensures that the operational cost for one update $C(1)$ is not larger than the maximal revenue for one update $DF(T/2, T/2)$, as shown in the following optimal time-dependent pricing scheme:

Proposition 1. *There exists an optimal time-dependent pricing scheme such that*[5]

$$p^*(t) = DF\left(T/2, T/2\right), \ \forall t \in \mathcal{T}, \tag{11}$$

where the equilibrium update takes place at $S_1^* = T/2$.

Proposition 1 suggests the existence of an optimal time-dependent pricing scheme that is in fact *time-invariant*. That is, although our original intention is to exploit the time sensitivity/flexibility of the destination through the time-dependent pricing, it turns out not to be very effective. This motivates us to consider a quantity-based pricing scheme next.

[5]There exist multiple optimal pricing schemes; the only difference among all optimal pricing schemes are the prices for time instances other than $T/2$, which can be arbitrarily larger than $DF(T/2, T/2)$.

V. Quantity-Based Pricing Scheme

In this section, we focus on a (pure) *quantity-based* pricing scheme, in which the price for each update depends on the number of updates that the destination has requested so far.

The source determines the quantity-based pricing scheme $p_q(k)$ in Stage I, in which $p_q(k)$ represents the price for the kth update. The payment from the destination will be $P_q(K) = \sum_{k=1}^{K} p_q(k)$. Based on $\boldsymbol{p}_q = \{p_q(k)\}_{k \in \mathbb{N}}$, the destination in Stage II chooses its update policy $(K, \boldsymbol{x})$.

We derive the *(Stackelberg)* price-update equilibrium using the bilevel optimization framework [42]. Specifically, the bilevel optimization problem embeds the optimality condition of the low-level problem (the destination's problem (12)) into the upper-level problem (the source's problem (3)). We first characterize the conditions of the destination's update policy $(K^*(\boldsymbol{p}_q), \boldsymbol{x}^*(\boldsymbol{p}_q))$ that minimize its overall cost in Stage II. We then substitute such conditions into the constraint set of the source's pricing problem in Stage I in order to characterize the source's optimal pricing $\boldsymbol{p}_q^*$ accordingly. We use $(K^*, \boldsymbol{x}^*)$ to denote the equilibrium update policy, i.e., $(K^*, \boldsymbol{x}^*) = (K^*(\boldsymbol{p}_q^*), \boldsymbol{x}^*(\boldsymbol{p}_q^*))$.

A. Destination's Update Policy in Stage II

Given the quantity-based pricing scheme $\boldsymbol{p}_q$, the destination solves the following overall cost minimization problem:

$$\min_{K \in \mathbb{N} \cup \{0\}, \boldsymbol{x} \in \mathbb{R}_{++}^{K+1}} \ \sum_{k=1}^{K+1} F(x_k) + \sum_{k=1}^{K} p_q(k), \tag{12a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K+1} x_k = T. \tag{12b}$$

If we fix the value of K in (12), then problem (12) is convex with respect to $\boldsymbol{x}$. The convexity allows us to exploit the Karush-Kuhn-Tucker (KKT) conditions on $\boldsymbol{x}$ to derive the following lemma (the proof is omitted due to space limits):

Lemma 2. *Under any given quantity-based pricing scheme $\boldsymbol{p}_q$ in Stage I, the destination's optimal update policy $(K^*(\boldsymbol{p}_q), \boldsymbol{x}^*(\boldsymbol{p}_q))$ satisfies*

$$x_k^*(\boldsymbol{p}_q) = \frac{T}{K^*(\boldsymbol{p}_q) + 1}, \quad \forall k \in \mathcal{K}(K^*(\boldsymbol{p}_q) + 1). \tag{13}$$

Lemma 2 indicates that the optimal update policy for the destination equalizes the inter-update time intervals. Hence, once the optimal interarrival time intervals is set according to (13), the destination would search for the optimal $K^*(\boldsymbol{p}_q)$ to minimize the objective in (12):

$$K^*(\boldsymbol{p}_q) \in \arg \min_{K' \in \mathbb{N}\cup\{0\}} \Upsilon(K', \boldsymbol{p}_q), \tag{14}$$

where $\Upsilon(K', \boldsymbol{p}_q)$ is the overall cost given the equalized interarrival time intervals:

$$\Upsilon(K', \boldsymbol{p}_q) \triangleq (K'+1)F\left(\frac{T}{K'+1}\right) + \sum_{k=1}^{K'} p_q(k). \tag{15}$$

B. Source's Quantity-Based Pricing in Stage I

Instead of solving both $K^*(\boldsymbol{p}_q)$ and $\boldsymbol{x}^*(\boldsymbol{p}_q)$ explicitly in Stage II, we apply the bilevel optimization to solving the optimal quantity-based pricing $\boldsymbol{p}_q^*$ in Stage I. Doing so would lead to the price-update equilibrium of our entire two-stage game [42]. By substituting the solutions (13)-(14) into the source's pricing in (3), we obtain the following bilevel problem:

$$\textbf{Bilevel}: \max_{\boldsymbol{p}_q, K, \boldsymbol{x}} \sum_{k=1}^{K} p_q(k) - C(K), \tag{16a}$$

$$\text{s.t. } x_k = \frac{T}{K+1}, \ \forall k \in \mathcal{K}(K+1), \tag{16b}$$

$$K \in \arg \min_{K' \in \mathbb{N}\cup\{0\}} \Upsilon(K', \boldsymbol{p}_q). \tag{16c}$$

In problem (16), we treat $(K, \boldsymbol{x})$ as variables with the destination's behavior being part of the source's constraints. The optimal solution to the bilevel optimization problem (16) is exactly the equilibrium $(\boldsymbol{p}_q^*, K^*, \boldsymbol{x}^*)$ [42].

The bilevel optimization in (16) leads to the following result:

Proposition 2. *The equilibrium update count K^* and the optimal quantity-based pricing scheme $\boldsymbol{p}_q^*$ satisfy*

$$\sum_{k=1}^{K^*} p_q^*(k) = F(T) - (K^*+1)F\left(\frac{T}{K^*+1}\right), \tag{17}$$

$$\sum_{k=1}^{K'} p_q^*(k) \geq F(T) - (K'+1)F\left(\frac{T}{K'+1}\right), \forall K' \in \mathbb{N}\backslash\{K^*\}. \tag{18}$$

Proof Sketch: Fig. 6 provides an illustrative example to understand Proposition 2. The area of the blue region is the aggregate AoI cost of the optimal updates, $(K^*+1)F(T/(K^*+1))$; the area of the blue region plus the green region is the aggregate AoI cost of a no-update scheme $F(T)$. The area of the green region $F(T) - (K^*+1)F\left(T/(K^*+1)\right)$ is the aggregate AoI cost difference between these two schemes.

We prove that inequality (18) together with (17) will ensure that constraint (16c) holds. Specifically, if (18) is not satisfied or if $\sum_{k=1}^{K^*} p_q^*(k) > F(T) - (K^*+1)F\left(T/(K^*+1)\right)$, then K^* would violate constraint (16c). If $\sum_{k=1}^{K^*} p_q^*(k) < F(T) - (K^*+1)F\left(T/(K^*+1)\right)$, then the source can always properly increase $p_q^*(1)$ until (17) is satisfied. Such an increase does not violate constraint (16c) but improves the source's profit, contradicting to the optimality of $\boldsymbol{p}_q^*$. $\square$

Substituting the pricing structure in (17) into (16), we can obtain K^* through solving the following problem:[6]

$$\max_{K \in \mathbb{N}\cup\{0\}} \ -(K+1)F\left(\frac{T}{K+1}\right) - C(K). \tag{19}$$

To solve problem (19), we first relax the constraint $K \in \mathbb{N}\cup\{0\}$ into $K \in \mathbb{R}_+$, hence transforming the integer programing problem (19) into a continuous optimization problem as follows:

$$\max_{K \in \mathbb{R}_+} \ -(K+1)F\left(\frac{T}{K+1}\right) - C(K), \tag{20}$$

which is a convex problem.[7] We take the derivative of objective in (20) and obtain

$$\underbrace{f\left(\frac{T}{K+1}\right)\frac{T}{K+1} - F\left(\frac{T}{K+1}\right)}_{\text{Marginal Revenue}} - \underbrace{C'(K)}_{\text{Marginal Cost}}. \tag{21}$$

We can interpret the first term as the source's *marginal revenue* in K and the second term as the source's *marginal cost* in K.

Based on the marginal revenue and the marginal cost in (21), we define a threshold update count $\hat{K}$ satisfying

$$f\left(\frac{T}{\hat{K}+1}\right)\frac{T}{\hat{K}+1} - F\left(\frac{T}{\hat{K}+1}\right) \geq C'(\hat{K}), \tag{22a}$$

$$f\left(\frac{T}{\hat{K}+2}\right)\frac{T}{\hat{K}+2} - F\left(\frac{T}{\hat{K}+2}\right) < C'(\hat{K}+1). \tag{22b}$$

The threshold count $\hat{K}$ serves as one of the candidates for the optimal update count to problem (19) as shown next.[8]

Proposition 3. *The optimal update count K^* to problem in* (16) *satisfies*

$$K^* = \arg \min_{K \in \{\hat{K}, \hat{K}+1\}} (K+1)F\left(\frac{T}{K+1}\right) + C(K). \tag{23}$$

Proof: Let K^* be the optimal solution to problem (20). By the definition of $\hat{K}$ in (22), we have $\hat{K} \leq K^* \leq \hat{K}+1$. The convexity of the objective in (20) implies that the objective of (20) (which is also the objective of problem (19)) is non-decreasing in K for all $K \leq K^*$ and is non-increasing in K for all $K \geq K^*$. This implies an optimal solution to problem (19) is either $\hat{K}$ or $\hat{K}+1$. $\blacksquare$

After obtaining K^*, we can construct an optimal pricing scheme based on Proposition 2 as follows:

Proposition 4. *An optimal quantity-based pricing $\boldsymbol{p}_q^*$ is*

$$p_q^*(k) = \begin{cases} 0, & \text{if } k = 0, \\ F(T) - (k+1)F(\frac{T}{k+1}) - \sum_{j=1}^{k-1} p_q^*(j) + \epsilon, & \text{if } 1 < k < K^*, \\ F(T) - (K^*+1)F(\frac{T}{K^*+1}) - \sum_{j=1}^{K^*-1} p_q^*(j), & \text{if } k \geq K^*, \end{cases} \tag{24}$$

[6] Note that $F(T)$ in (17) is a constant and hence is not considered in (19).

[7] To see the convexity of $(K+1)F\left(T/(K+1)\right)$, note that $(K+1)F\left(T/(K+1)\right)$ is the *perspective* of function $F(T)$. The perspective of $F(T)$ is convex since $F(T)$ is convex [38].

[8] Assumption 1 leads to the existence of a unique $\hat{K}$ satisfying (22).

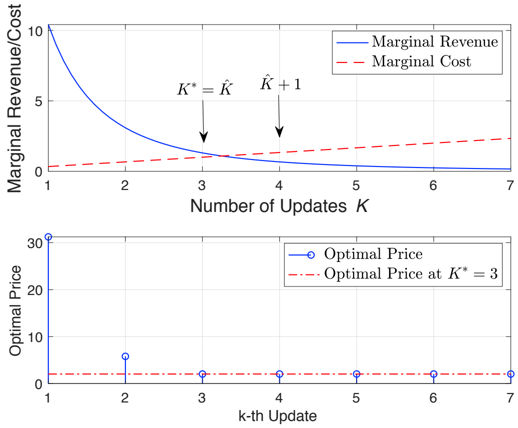

Fig. 7. An illustrative example of the optimal quantity-based pricing scheme in Proposition 4. The destination's AoI cost is $f(\Delta_t) = \Delta_t^2$ and the source's operational cost is $C(K) = 1/6K^2$.

where $\epsilon > 0$ *is an infinitesimal value to ensure* (18).

Fig. 7 presents an illustrative example of (24). In Fig. 7 (up), the marginal revenue intersects with the marginal cost in (21) at around $K = 3.3$. Hence, the threshold update count is $\hat{K} = 3$ based on (22), and we can further verify based on (23) that the optimal update count is $K^* = 3$. In Fig. 7 (down), we present the optimal quantity-based pricing scheme described in (24). As we can see, the optimal price drops until the third update. The relatively high prices value of the first two update prices are to ensure (18) holds for $K' = \{1, 2\}$ while the relatively lower price starting from the third update is to ensure (17) holds.

VI. Properties

In this section, we study several properties of the pure quantity-based pricing and the pure time-dependent pricing. Let Π_q denote the achievable profit of the pure quantity-based pricing and Π_t denote that of the pure time-dependent pricing. We first compare Π_q with Π_t in the following Proposition:

Proposition 5. *The achievable profit of the optimal quantity-based pricing* Π_q *and that of the optimal time-dependent pricing* Π_t *satisfy*

$$\Pi_t \overset{(a)}{\leq} \Pi_q \overset{(b)}{<} 2\Pi_t. \tag{25}$$

Proof Sketch: We can show that the time-dependent pricing scheme is a special case of the quantity-based pricing in Proposition 4 by fixing $K^* = 1$, which proves (a). To prove (b), the destination's payment under the optimal time-dependent pricing scheme in Proposition 1 is $DF(T/2, T/2)$, which we can show to be at least $F(T)/2$; while the destination's payment under the optimal quantity-based pricing is at most $F(T)$. □

We are now ready to introduce the key result of this paper:

Theorem 3 (Profit Maximizing Structure)**.** *The optimal quantity-based pricing achieves the maximum source profit among all possible time-and-quantity dependent pricing schemes in the form of* $p(t, k)$.

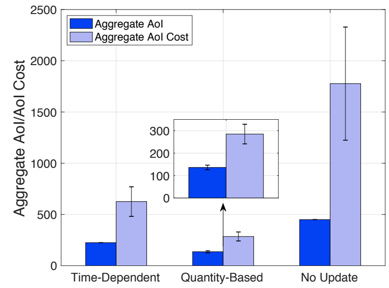

Fig. 8. Performance comparison in terms of the aggregate AoI and the aggregate AoI cost.

Proof Sketch: We first prove that, regardless of the pricing choice, the destination's payment is always upper-bounded by the AoI cost reduction as we discussed in Proposition 3. Meanwhile, the optimal quantity-based pricing in (18) attains such a bound. We then prove it is profit-maximizing. □

Theorem 3 implies that the relatively-simple quantity-based pricing scheme is already optimal. Hence, even without exploiting the time flexibility explicitly, it is still possible to obtain the optimal pricing structure, which again implies that utilizing time flexibility may not be necessary.

Next we introduce the social cost minimization problem, which minimizes the sum of the destination's aggregate AoI cost and the source's operational cost:

$$\min_{K \in \mathbb{N} \cup \{0\}, \boldsymbol{x} \in \mathbb{R}_{++}^{K+1}} \sum_{k=1}^{K+1} F(x_k) + C(K), \tag{26a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K+1} x_k = T. \tag{26b}$$

Proposition 6 (Social Cost Minimization)**.** *The optimal quantity-based pricing scheme in* (24) *leads to the optimal solution of* (26)*, i.e., minimizing the social cost.*

Proof Sketch: We first prove that the social cost minimizing update policy equalizes the time intervals, due to the similar reason as in Lemma 2. In this case, problem (26) becomes equivalent to problem (19). Hence, the solution in (19) yields the minimal social cost. □

VII. Numerical Results

In this section, we perform simulations to numerically compare both proposed pricing schemes regarding the aggregate AoI, the source's profit, and the social cost.

We consider a time interval of $T = 30$ (days).[9] The destination's AoI cost function is $f(\Delta_t) = \Delta_t^\kappa$, where the exponent κ is the *destination's age sensitivity*. Hence, the function $F(t)$ is $F(t) = t^{\kappa+1}/(\kappa+1)$. The source has a cubic operational cost function, i.e., $C(K) = c \cdot K^3$, where c is the source's *operational cost coefficient*. Let κ follow a normal distribution $\mathcal{N}(1.5, 0.2)$ truncated into the interval $[1, 2]$; let c follow a normal distribution $\mathcal{N}(6, 1.5)$ truncated into the interval $[2, 10]$.

[9]Examples of such a period of interest include the API monitoring platform [2] and Google Map platform [3].

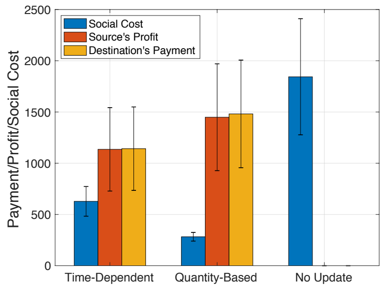

Fig. 9. Performance comparison in terms of the social cost, the source's profit, and the destination's payment.

We compare the performance of three schemes: the optimal time-dependent pricing, the optimal quantity-based pricing, and a no-update benchmark. In Fig. 8, we first compare the three schemes in terms of the aggregate AoI and the aggregate AoI cost. The no-update scheme incurs a much larger aggregate AoI than both proposed pricing schemes. Moreover, the optimal quantity-based pricing scheme incurs an aggregate AoI which is only 59% of that incurred by the optimal time-dependent pricing. In terms of the aggregate AoI cost, we observe a similar trend.

In Fig. 9, we compare the three schemes in terms of the social cost, profit (of the source), and payment (of the destination). First, we observe that the optimal quantity-based pricing is 27% more profitable than the optimal time-dependent pricing. Such an improvement is consistent with the analytical bounds in Proposition 5. Finally, the optimal time-dependent pricing only incurs 34% of the social cost of the no-update scheme, while the optimal quantity-based pricing further reduces the social cost and incurs only 46% of that of the optimal time-dependent pricing. Note that the large standard deviations for both profits and payments of the proposed pricing schemes are mainly due to the large standard deviation of the aggregate AoI cost for the no-update scheme as shown in Fig. 8.

VIII. Conclusions

We have presented the first pricing scheme design for a fresh data market and proposed two pricing schemes. Our results have revealed that (i) the optimal time-dependent pricing scheme yields a single-update equilibrium, which does not effectively exploit the time flexibility, and (ii) the optimal quantity-based pricing scheme achieves the maximum profit for the source among all time-and-quantity dependent pricing schemes, and leads to the minimal social cost. Future work includes the extension to multi-destination scenarios, and studying incomplete user information settings.

References

[1] C. Shapiro and H. Varian, "Information rules: A strategic guide to the network economy." Harvard Business Press, 1999.
[2] https://cloud.google.com/apigee-api-management/
[3] https://cloud.google.com/maps-platform/
[4] X. Song and J. W. S. Liu, "Performance of multiversion concurrency control algorithms in maintaining temporal consistency, in *Proc. COMPSAC*, 1990.
[5] A. Segev and W. Fang, "Optimal update policies for distributed materialized views, in *Manage. Sci.*, 1991.
[6] C. Joe-Wong, S. Ha, and M. Chiang, "Time-dependent broadband pricing: Feasibility and benefits," in *Proc. ICDCS*, 2011.
[7] P. Hande, M. Chiang, R. Calderbank, and J. Zhang, "Pricing under constraints in access networks: Revenue maximization and congestion management," in *Proc. IEEE INFOCOM*, 2010.
[8] L. Zhang, W. Wu, and D. Wang, "Time dependent pricing in wireless data networks: Flat-rate vs. usage-based schemes," in *Proc. IEEE INFOCOM*, 2014.
[9] L. Jiang, S. Parekh, and J. Walrand, "Time-dependent network pricing and bandwidth trading," in *Proc. IEEE NOMS*, 2008.
[10] S. Ha, S. Sen, C. Joe-Wong, Y. Im, and M. Chiang, "TUBE: Time-dependent pricing for mobile data." *SIGCOMM Comput. Commun. Rev*, 2012.
[11] Q. Ma, Y.-F. Liu, and J. Huang,"Time and location aware mobile data pricing." *IEEE Trans. Mobile Comput.*, 2016.
[12] A. Sundararajan, "Nonlinear pricing of information goods," *Manage. Sci* 2004.
[13] H. Shen and T. Basar, "Optimal nonlinear pricing for a monopolistic network service provider with complete and incomplete information," *IEEE J. Sel. Areas Commun.*, 2007.
[14] S. Li and J. Huang, "Price differentiation for communication networks," *IEEE/ACM Trans. Netw.*, 2014.
[15] R. L. Phillips, "Pricing and revenue optimization." *Stanford University Press,* 2005.
[16] S. Kaul, R. D. Yates, and M. Gruteser, "Real-time status: How often should one update? in *Proc. IEEE INFOCOM*, 2012.
[17] Q. He, D. Yuan, and A. Ephremides, "Optimal link scheduling for age minimization in wireless systems". *IEEE Trans. Inf. Theory*, 2018.
[18] Q. He, D. Yuan, and A. Ephremides, "On optimal link scheduling with min-max peak age of information in wireless systems, in *Proc. IEEE ICC*, 2016.
[19] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks, in *Proc. IEEE Allerton*, 2016.
[20] C. Kam, S. Kompella, and A. Ephremides, Age of information under random updates, in *Proc. IEEE ISIT*, 2013.
[21] 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 Trans. Inf. Theory*, 2017.
[22] R. Talak, S. Karaman, and E. Modiano. "Optimizing information freshness in wireless networks under general interference constraints," in *Proc. ACM Mobihoc*, 2018.
[23] I. Kadota, A. Sinha, and E. Modiano. "Optimizing age of information in wireless networks with throughput constraints," in *Proc. IEEE INFOCOM*, 2018.
[24] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc. IEEE ISIT*, 2015.
[25] X. Wu, J. Yang, and J. Wu, "Optimal status update for age of information minimization with an energy harvesting source," *IEEE Trans. on Green Commun. Netw.*, 2(1):193204, March 2018.
[26] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies", available online: arXiv:1806.07271.
[27] B. T. Bacinoglu, Y. Sun, E. Uysal-Biyikoglu, and V. Mutlu. "Achieving the age-energy tradeoff with a finite-battery energy harvesting source," in *Proc. IEEE ISIT*, 2018.
[28] A. Arafa and S. Ulukus, "Timely updates in energy harvesting two-hop networks: Offline and online policies," available online: arXiv:1812.01005.
[29] B. Zhou and W. Saad, "Optimal sampling and updating for minimizing age of information in the Internet of things", in *Proc. of IEEE GLOBECOM*, 2018.
[30] M. A. Abd-Elmagid, N. Pappas, and H. S. Dhillon, "On the role of age-of-information in Internet of things", available online: arXiv:1812.08286.
[31] G. D. Nguyen, S. Kompella, C. Kam, J. E. Wieselthier, and A. Ephremides, "Impact of hostile interference on information freshness: A game approach," in *Proc. WiOpt*, 2017.
[32] G. D. Nguyen, S. Kompella, C. Kam, J. E. Wieselthier, and A. Ephremides, "Information freshness over an interference channel: A game theoretic view," in *Proc. IEEE INFOCOM*, 2018.
[33] Y. Xiao and Y. Sun, "A dynamic jamming game for real-time status updates" in *Proc. IEEE INFOCOM Age of Information Workshop*, 2018.
[34] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang, "A survey of smart data pricing: Past proposals, current plans, and future trends." *ACM Computing Surveys*, 2013.
[35] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang, "Incentivizing time-shifting of data: A survey of time-dependent pricing for internet access," *IEEE Commun. Magazine*, 2012.
[36] S. Hao and L. Duan, "Economics of age of information management under network externalities," in *Proc. ACM MobiHoc*, 2019.
[37] X. Wang and L. Duan. "Dynamic pricing for controlling age of information," in *Proc. IEEE ISIT*, 2019.
[38] A. Even and G. Shankaranarayanan, "Utility-driven assessment of data quality" *SIGMIS Database*, 2007.
[39] S. Shalev-Shwartz, "Online learning and online convex optimization, *Found. Trends Mach. Learn.,* vol. 4, no. 2, pp. 107194, 2012.
[40] X. He *et al.*, "Practical lessons from predicting clicks on ads at Facebook," in *Proc. 8th Int. Workshop Data Mining Online Advertising*, 2014, pp. 19.
[41] A. Mas-Colell, M. D. Whinston, and J. R. Green, "Microeconomic theory," *Oxford university press*, 1995.
[42] B. Colson, P. Marcotte, and G. Savard, "An overview of bilevel optimization, *Ann. Operations Res.*, 2007.