Energy-Efficient Transmission Scheduling with Strict Underflow Constraints

David Shuman and Mingyan Liu
Electrical Engineering and Computer Science Department
University of Michigan
Ann Arbor, MI 48109-2122
{dishuman,mingyan}@umich.edu

Abstract

We consider a single source transmitting data to a single receiver/user over a wireless channel. The channel is time-varying, and the user has a buffer to store received packets before they are drained. At each time step, the source determines how much power to use for transmission. The source’s objective is to allocate power in a manner that minimizes an expected cost measure, while satisfying strict buffer underflow constraints and a total power constraint in each slot. The expected cost measure is composed of costs associated with power consumption from transmission and packet holding costs. The primary application motivating this problem is wireless media streaming. For this application, the buffer underflow constraints prevent the user buffer from emptying, so as to maintain playout quality. We show that a modified base-stock policy is optimal under the finite and infinite horizon discounted expected cost criteria. For a special case, we present the sequences of critical numbers that complete the characterization of the optimal control laws in each of these two problems.

Index Terms

Wireless media streaming, underflow constraints, opportunistic scheduling, resource allocation, dynamic programming, inventory theory, base-stock policy, water-filling.

Part of the work reported here was presented at the International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), April 2008, Berlin, Germany.
I. INTRODUCTION

In this paper, we examine the problem of energy-efficient transmission scheduling over a wireless channel, subject to underflow constraints. We consider a single source transmitting to a single receiver/user over a wireless channel. The user has a buffer to store received packets before they are drained at a certain rate. The available data rate of the channel varies with time, due to random fading. The transmitter’s goal is to minimize total power consumption by exploiting the temporal variation of the channel, while preventing the user’s buffer from emptying.

This problem falls into the general class of opportunistic scheduling problems. At a high level, the idea of exploiting the temporal variation of the channel via opportunistic scheduling can be explained as follows. Consider one scheduling policy that transmits data in a just-in-time fashion, without regard to the condition of the time-varying channel. Over the long run, the total power consumption will tend toward the power consumption per data packet under the average channel condition times the number of packets sent. If instead, the scheduler aims to send more data when the channel is in a “good” state (requiring less power per data packet), and less data when the channel is in a “bad” state, the total power consumption should be lower. Much of the challenge for the scheduler lies in determining how good or bad a channel condition is, and how much data to send accordingly. Moreover, as is often the case in opportunistic scheduling problems, the scheduler has a competing QoS interest - to prevent underflow at the receiver’s buffer. Thus, it may still be necessary to transmit data when the channel is in a relatively bad condition.

The primary application we have in mind to motivate this problem is wireless media streaming. For this application, the data are audio/video sequences, and the packets are drained from the receiver’s buffer in order to be decoded and played. Enforcing the underflow constraints reduces playout interruptions to the end user. In order to make the presentation concrete, we will use the above wireless media streaming terminology throughout the paper.

Transporting multimedia over wireless networks is a promising application that has seen recent advances [1]. At the same time, a number of resource allocation issues need to be addressed in order to provide high quality and efficient media over wireless. First, streaming is in general
bandwidth-demanding. Second, streaming applications tend to have stringent quality of service (QoS) requirements (e.g., they can be delay and jitter intolerant). Third, it is desirable to operate the wireless system in an energy-efficient manner. This is obvious when the source of the media streaming (the sender) is a mobile. When the media comes from a base station that is not power-constrained, it is still desirable to conserve power in order to limit potential interference to other base stations and their associated mobiles.

Of the related work in wireless media streaming, [2] has the closest setup to our model. The main differences are that [2] features a loose constraint on underflow (i.e., it is allowed, but at a cost), as opposed to our tight constraint, and the two studies adopt different wireless channel models. In the extension [3], the receiver may slow down its playout rate (at some cost) to avoid underflow. In this setting, the authors investigate the tradeoffs between power consumption and playout quality, and examine joint power/playout rate control policies. In our model, the receiver does not have the option to adjust the playout speed. Our model also bears resemblance to [4]. The first difference here is that [4] aims to minimize transmission energy subject to a constant end-to-end delay constraint on each video frame. A second difference is that the controller in [4] must assign various source coding parameters such as quantization step size and coding mode, whereas our model assumes a fixed encoding/decoding scheme. Outside of the streaming context, the most related work is [5], which considers energy-efficient transmission scheduling for sending a fixed amount of data over a wireless channel, subject to a deadline constraint. We discuss the relation between our model and [5] in Section V.

The remainder of this paper is organized as follows. In the next section, we describe the system model, formulate two optimization problems (finite and infinite horizons), and relate our model to models in inventory theory. In Section III, we analyze the structure of the optimal scheduling policy for both problems. In Section IV, we completely characterize the optimal policy for a special case. We then try to provide further intuition behind our results by comparing them to related work in inventory theory and wireless communications theory in Section V. In Section VI, we discuss the extension of our problem to the case of multiple receivers, which is ongoing work. Section VII concludes the paper.
II. PROBLEM DESCRIPTION

In this section, we present an abstraction of the transmission scheduling problem outlined in the previous section and formulate two optimization problems.

A. System Model and Assumptions

We consider a single source transmitting media sequences to a single user/receiver over a wireless channel. The sender maintains a buffer for data to be sent to the receiver, and is assumed to always have data to transmit. The receiver has a playout buffer at the receiving end, assumed to be infinite. While in reality this cannot be the case, it is nevertheless a reasonable assumption considering the decreasing cost and size of memory, and the fact that our system model penalizes holding packets in the buffer.

We consider time evolution in discrete steps, indexed backwards by \( n = N, N - 1, \ldots, 1 \), with \( n \) representing the number of slots remaining in the time horizon. \( N \) is the length of the time horizon, and slot \( n \) refers to the time interval \([n, n - 1)\).

At the beginning of each time slot, the scheduler allocates some amount of power (possibly zero) for transmission. The total power consumed in any one slot must not exceed the stationary power constraint, \( P \). Following transmission and reception in each slot, a certain number of packets are removed/purged from the receiver buffer for playing. We assume that packets transmitted in slot \( n \) arrive in time to be used for playing in slot \( n \). Holding costs may be assessed on all packets remaining in the receiver’s buffer after playout consumption in each slot. We also assume the receiver buffer is empty at the beginning of the time horizon.

We assume that the transmitter (or scheduler) knows precisely the packet requirements of the receiver in each time slot. This is justified by the assumption that the transmitter knows the encoding and decoding schemes used. We further assume that the user’s consumption of packets in each slot is constant, denoted by \( d \). This assumption is less realistic, but may be justified if the receiving buffer is drained at a constant rate at the MAC layer, before packets are decoded by the media player at the application layer. It is also worth noting that the same techniques we use in this paper to analyze the stationary demand case can be used to examine the nonstationary
demand case. We discuss the extension to the case of nonstationary demand further in Section III-A.

In general, wireless channel conditions are time-varying. Adopting a block fading model, we assume that the slot duration is within the channel coherence time such that the channel condition within a single slot is constant. The user’s channel condition in slot \( n \) is modeled as a random variable, \( C_n \). If the channel condition is in state \( c \), then the transmission of \( r \) units of data incurs a power consumption of \( c(r) \). This power-rate function \( c(\cdot) \) is commonly assumed to be convex (in the high SNR regime) or linear (in the low SNR regime). In this paper, we only consider the linear case, and \( C_n \) represents the power consumed per packet transmitted. \( C_n \) may be a discrete, continuous, or mixed random variable, and it has a sample space denoted by \( \mathcal{C} \). We assume that sending data consumes a strictly positive amount of power, and, accordingly, assume \( c_{\min} := \inf \mathcal{C} > 0 \). We also assume that even when the channel is in its worst possible condition, the maximum power constraint \( P \) is sufficient to transmit enough packets to satisfy one time slot’s playout demand; i.e., \( c \leq \frac{P}{d} \), for all \( c \in \mathcal{C} \). For the notational purpose of stating theorems more clearly, we assume there is a ”best” and a ”worst” channel condition (i.e., \( 0 < c_{\min} := \inf \mathcal{C} = \min \mathcal{C} \) and \( \frac{P}{d} \geq c_{\max} := \sup \mathcal{C} = \max \mathcal{C} \)). Thus, the sample space \( \mathcal{C} \) is contained in the interval \([c_{\min}, c_{\max}]\).

We begin by modeling the channel condition in each slot as a homogeneous Markov process. Namely, conditioned on the channel state, \( C_n \), at time \( n \), the channel states at future times \((n - 1, n - 2, \ldots)\) are independent of the channel states at past times \((n + 1, n + 2, \ldots)\). Note the somewhat unconventional notation that future times are indexed by lower epoch numbers, as \( n \) represents the number of slots remaining in the time horizon. Modeling time backwards facilitates the analysis of the infinite horizon problem, as will be seen in Section III-B. It may also be the case that the channel condition is independent and identically distributed (IID) from slot to slot. We consider this scenario in the special case discussed in Section IV. Throughout the paper, we assume the channel condition is independent of any of the transmitter’s scheduling decisions. We also assume that the transmitter learns the channel’s state through a feedback channel at the beginning of each time slot, prior to making the scheduling decision.
We are primarily concerned with two objectives in deriving a good transmission policy. One is to avoid underflow at the receiver buffer, thus avoiding disruption to the user playout. The other is to minimize power consumption. The goal of this study is to characterize the control laws that minimize the transmission power costs over a finite or infinite time horizon, subject to tight underflow constraints and a maximum power constraint in each time slot.

B. Problem Formulation

We consider two problems. The first, Problem (P1), is the finite horizon discounted expected cost problem. The second, Problem (P2), is the infinite horizon discounted expected cost problem. The two problems feature the same information state, action space, system dynamics, and cost structure, but different optimization criteria.

The information state at time $n$ is the pair $(X_n, C_n)$, where the random variable $X_n$ denotes the current receiver buffer queue length, and $C_n$ denotes the channel condition in slot $n$ (recall that $n$ is the number of steps remaining until the end of the horizon). The queue dynamics for the receiver’s buffer are governed by the simple equation $X_{n-1} = Y_n - d$ at all times $n = N, N-1, \ldots, 1$. $Y_n$ is the queue length after transmission in the $n^{th}$ slot takes place, but before playout in the $n^{th}$ slot has occurred. $Y_n$ is a controlled random variable chosen by the scheduler at each time $n$. It must satisfy the power constraint: $C_n \cdot (Y_n - X_n) \leq P$, and the underflow constraint: $Y_n \geq d$. Clearly, the scheduler cannot transmit a negative number of packets, so it must also be true that $Y_n \geq X_n$.

We now present the optimization criterion for each problem. In addition to the cost associated with power consumption from transmission, we introduce holding costs on packets stored in the user’s playout buffer at the end of a time slot. The holding costs in each slot are described by a convex, nonnegative, nondecreasing function, $h(\cdot)$, of the packets remaining in the buffer following playout. Possible holding cost models might include a linear model, $h(y) = \hat{h} \cdot y$ for some constant $\hat{h}$; a barrier-type function such as:

$$h(y) := \begin{cases} 
0, & \text{if } y \leq \mu \\
K \cdot (y - \mu), & \text{if } y > \mu \text{ (} K \text{ very large)}
\end{cases}$$
which could represent a finite receiver buffer of length \( \mu \); or \( h(\cdot) = 0 \) in the case one does not want to include holding costs. In Problem (P1), we wish to find a transmission policy \( \pi \) that minimizes \( J^\pi_N \), the expected cost under policy \( \pi \), defined as:

\[
J^\pi_N := \mathbb{E}^\pi \left\{ \sum_{t=1}^{N} \alpha^{N-t} \cdot \left\{ C_t \cdot (Y_t - X_t) + h(Y_t - d) \right\} \mid \mathcal{F}_N \right\},
\]

where \( 0 \leq \alpha \leq 1 \) is the discount factor and \( \mathcal{F}_N \) denotes all information available at the beginning of the time horizon. For Problem (P2), the cost function for minimization is defined as \( J^\pi_\infty := \lim_{N \to \infty} J^\pi_N \), and the discount factor must satisfy \( 0 \leq \alpha < 1 \). In both cases, we allow the transmission policy \( \pi \) to be chosen from the set of all randomized and deterministic control laws, \( \Pi \).

Combining the constraints and criteria, we present the optimization formulations for Problem (P1) (or (P2)):

\[
\min_{\pi \in \Pi} J^\pi_N \quad \text{or} \quad \min_{\pi \in \Pi} J^\pi_\infty
\]

s.t. \( C_n \cdot (Y_n - X_n) \leq P, \ \forall n \) and \( Y_n \geq \max(X_n, d), \ \forall n. \)

The above problem may be solved using standard dynamic programming (see, e.g., [6]). The recursive dynamic programming equations in the finite horizon case are given by:

\[
V_n(x, c) = \min_{y \in \mathcal{A}^d(x, c)} \left\{ c \cdot (y - x) + h(y - d) + \alpha \cdot \mathbb{E}\left[V_{n-1}(y - d, C_{n-1}) \mid C_n = c \right]\right\}
\]

\[
V_0(x, c) = 0, \ \forall x, \forall c
\]

(1)

where \( V(\cdot, \cdot) \) is the value function or expected cost-to-go. The relevant infinite horizon functional equation is:

\[
V_\infty(x, c) = \min_{y \in \mathcal{A}^d(x, c)} \left\{ c \cdot (y - x) + h(y - d) + \alpha \cdot \mathbb{E}\left[V_\infty(y - d, C') \mid C = c \right]\right\},
\]

(2)

where \( C' \) is the channel condition in the subsequent slot. In both functional equations (1) and
(2), the action space is defined as:

$$\mathcal{A}^d(x, c) := \left\{ y \in \mathbb{R}_+ : \max(d, x) \leq y \leq \frac{P}{c} + x \right\}.$$ 

C. Related Work in Inventory Theory

The model outlined in Section II-A corresponds closely to models used in inventory theory. Borrowing that field’s terminology, our abstraction is a multi-period, single-echelon, single-item, discrete time inventory model with random ordering prices, a budget constraint, and deterministic demands. The item corresponds to the stream of data packets, the random ordering prices to the random channel conditions, the budget constraint to the power available in each time slot, and the deterministic demands to the packet requirements for playout.

To the best of our knowledge, this particular problem has not been studied in the context of inventory theory, but similar problems have been examined. References [7] - [14] all consider single item inventory models with random ordering prices. The key result for the case of deterministic demand of a single item with no resource constraint is that the optimal policy is a base-stock policy with different target stock levels for each price. Specifically, for each possible ordering price (translates into channel condition in our context), there exists a critical number such that the optimal policy is to fill the inventory (receiver buffer) up to that critical number if the current level is lower than the critical number, and not to order (transmit) anything if the current level is above the critical number. Of the prior work, Kingsman [9], [10] is the only author to consider a resource constraint, and he imposes a maximum on the number of items that may be ordered in each slot. The resource constraint we consider is of a different nature in that we limit the amount of power available in each slot. This is equivalent to a limit on the per slot budget (regardless of the stochastic price realization), rather than a limit on the number of items that can be ordered. Of the related work on inventory models with deterministic ordering prices and stochastic demand, [15] is the most relevant; in that work, however, the resource constraint also amounts to a limit on the number of items that can be ordered in each slot, and is constant over time. References [16] and [17] consider multiple item inventory systems under resource constraints, and are more relevant for the extension we discuss in Section VI. Throughout the
paper, we discuss the above single item results further, and leverage some of their techniques and solutions in deriving our results.

III. STRUCTURE OF THE OPTIMAL POLICY

In this section, we analyze the finite and infinite horizon discounted expected problems, and describe the structure of the optimal policy for each.

A. Finite Horizon Discounted Expected Cost Problem

It may seem more natural to select the action space in our model to be the number of packets scheduled for transmission during a slot; however, choosing the action space to be the number of packets in the buffer following transmission allows us to rewrite the dynamic program (1) as:

\[
V_n(x, c) = -c \cdot x + \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \left\{ c \cdot y + h(y - d) + \alpha \cdot \mathbb{E} \left[ V_{n-1}(y - d, C_{n-1}) \mid C_n = c \right] \right\}
\]

\[
= -c \cdot x + \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \left\{ g_n(y, c) \right\},
\]  

(3)

where \( g_n(y, c) := c \cdot y + h(y - d) + \alpha \cdot \mathbb{E} \left[ V_{n-1}(y - d, C_{n-1}) \mid C_n = c \right] \). By doing so, the expected cost-to-go at time \( n \), \( V_n(x, c) \), depends on the current buffer level, \( x \), only through the fixed term \(-c \cdot x\) and the action space, \( A^d(x, c)\); i.e., the expected future costs only depend on the selected action, \( y \), and the current channel condition, \( c \). This separation allows us to leverage the inventory theory techniques of showing “single critical number” or “base-stock” policies, which date as far back as [18].

The key realization to be shown in the proof of Theorem 1 is that for all \( n \) and all \( c \), \( g_n(\cdot, c) : [d, \infty) \to \mathbb{R}_+ \) is a convex function in \( y \), with \( \lim_{y \to \infty} g_n(y, c) = \infty \). Thus, for all \( n \) and all \( c \), \( g_n(\cdot, c) \) has a global minimum, which is the ideal number of packets to have in the buffer following transmission in the \( n^{th} \) slot. This line of analysis leads to the following theorem on the structure of the optimal transmission policy.

**Theorem 1:** For every \( n \in \{1, 2, \ldots, N\} \) and \( c \in C \), define the critical number

\[
b_n(c) := \min \left\{ \hat{y} \in [d, \infty) : g_n(\hat{y}, c) = \min_{y \in [d, \infty)} g_n(y, c) \right\}.
\]
Then the optimal policy with \( n \) slots remaining is given by:

\[
y^*_n(x, c) := \begin{cases} 
  x, & \text{if } x \geq b_n(c) \\
  b_n(c), & \text{if } b_n(c) - \frac{P}{c} \leq x < b_n(c) \\
  x + \frac{P}{c}, & \text{if } x < b_n(c) - \frac{P}{c}
\end{cases}
\]  

(4)

Furthermore, for a fixed \( c \), \( b_n(c) \) is increasing in \( n \):

\[
N \cdot d \geq b_N(c) \geq b_{N-1}(c) \geq \ldots \geq b_1(c) = d.
\]  

(5)

If, in addition, the channel condition is independent and identically distributed from slot to slot, then for a fixed \( n \), \( b_n(c) \) is decreasing in \( c \); i.e., for arbitrary \( c^1, c^2 \in C \) with \( c^1 \leq c^2 \), we have:

\[
n \cdot d \geq b_n(c^1) \geq b_n(c^2) \geq b_n(c_{\max}) = d.
\]  

(6)

Details of the proof are included in the Appendix. The key idea to show (5) is to fix \( c \in C \), view \( g_n(y, c) \) as a function of \( y \) and \( n \), say \( f(y, n) \), and show that the function \( f(\cdot, \cdot) \) is submodular.

From the proof, one can also see that if we relax the stationary deterministic demand assumption to a nonstationary deterministic demand sequence, \( \{d_N, d_{N-1}, \ldots, d_1\} \) (with \( d_n \leq \frac{P}{c_{\max}} \) for all \( n \)), then the structure of the optimal policy is still as stated in (4). If, in addition to the nonstationary deterministic demand, the channel is IID, then the following statement, analogous to (6), is true for arbitrary \( c^1, c^2 \in C \) with \( c^1 \leq c^2 \):

\[
\sum_{i=1}^{n} d_i \geq b_n(c_{\min}) \geq b_n(c^1) \geq b_n(c^2) \geq b_n(c_{\max}) = d_n, \ \forall n \in \{1, 2, \ldots, N\}
\]  

(7)

However, (5), the monotonicity of critical numbers over time for a fixed channel condition, is not true in general under nonstationary deterministic demand. As one counterexample, (7) says that under an IID channel, the critical numbers for the worst possible channel condition are equal to the one period demands. Therefore, if the demand sequence is not monotonic, then the sequence of critical numbers, \( \{b_n(c_{\max})\}_{n=1,2,\ldots,N} \), is not monotonic.

The optimal transmission policy in Theorem 1 is a modified base-stock policy. At time \( n \), for each possible channel condition realization \( c \), the critical number \( b_n(c) \) describes the ideal
number of packets to have in the user’s buffer after transmission in the \( n^{th} \) slot. If that number of packets is already in the buffer, then it is optimal to not transmit any packets; if there are fewer than ideal and the available power is enough to transmit the difference, then it is optimal to do so; and if there are fewer than ideal and the available power is not enough to transmit the difference, then the sender should use the maximum power to transmit. See Fig. 1 for diagrams of the optimal policy.

![Optimal policy in slot \( n \) when the state is \((x, c)\).](image)

**B. Infinite Horizon Discounted Expected Cost Problem**

In this section, we show that the infinite horizon optimal policy is the natural extension of the finite horizon optimal policy; namely, it is a modified base-stock policy of the same form, characterized by *stationary* target buffer levels \( \{b_\infty(c)\}_{c \in \mathcal{C}} \).

**Theorem 2:**

(a) For a fixed \( x \in \mathbb{R}_+ \) and \( c \in \mathcal{C} \), \( V_n(x, c) \) is nondecreasing in \( n \). Moreover, \( \lim_{n \to \infty} V_n(x, c) \) exists and is finite, \( \forall x \in \mathbb{R}_+, \forall c \in \mathcal{C} \).

(b) Define \( V_\infty(x, c) := \lim_{n \to \infty} V_n(x, c) \). Then \( V_\infty(x, c) \) is convex in \( x \) for any fixed \( c \in \mathcal{C} \).

(c) Define \( g_\infty(y, c) := c \cdot y + h(y - d) + \alpha \cdot E [V_\infty(y - d, C') | C = c] \). Then \( g_n(y, c) \) converges monotonically to \( g_\infty(y, c) \), \( \forall y \in [d, \infty), \forall c \in \mathcal{C} \); \( g_\infty(y, c) \) is convex in \( y \) for any fixed \( c \in \mathcal{C} \);

and \( \lim_{y \to \infty} g_\infty(y, c) = \infty, \forall c \in \mathcal{C} \).
(d) Define $b_\infty(c) := \min \left\{ \hat{y} \in [d, \infty) : g_\infty(\hat{y}, c) = \min_{y \in [d, \infty)} g_\infty(y, c) \right\}$. Then $b_\infty(c) = \lim_{n \to \infty} b_n(c)$.

(e) $V_\infty(x, c)$ satisfies the infinite horizon functional equation (2), and the minimum on the right hand side of (2) is achieved by:

$$y_\infty^*(x, c) := \begin{cases} 
  x, & \text{if } x \geq b_\infty(c) \\
  b_\infty(c), & \text{if } b_\infty(c) - \frac{P}{c} \leq x < b_\infty(c) \\
  x + \frac{P}{c}, & \text{if } x < b_\infty(c) - \frac{P}{c}
\end{cases}$$

(f) The optimal stationary policy is given by $\pi_\infty^* = (y_\infty^*, y_\infty^*, \ldots)$.

The proof of Theorem 2 follows the logic conveyed in the statement of the theorem, and leverages the techniques of [15] and [19]. In part (d), we use the submodularity condition from the finite horizon proof to generalize slightly Federgruen and Zipkin’s method [15], by not requiring that the functions $g_n(\cdot, c)$ be continuously differentiable. Further details are included in the Appendix.

IV. COMPLETE CHARACTERIZATION OF THE OPTIMAL TRANSMISSION POLICY FOR A SPECIAL CASE

In this section, we consider the special case where the channel condition is independent and identically distributed from slot to slot, there are a finite number of possible channel conditions, the holding cost function is linear $\left( h(y) = h \cdot y \text{ for some } h \geq 0 \right)$, and the following technical condition is satisfied: for each possible channel condition $c$, $\frac{P}{c} = l \cdot d$ for some $l \in \mathbb{N}$; i.e., the maximum number of packets that can be transmitted in any slot covers exactly the playout requirements of some integer number of slots. Under these four assumptions, we can completely characterize the optimal transmission policy.

Using a similar approach to [9] - [14], we proceed by recursively defining a set of thresholds, and then using them to determine the critical numbers, $b_n(c)$, for each channel condition, at each time. The threshold $\gamma_{n,j}$ may be interpreted as the per packet power cost at which, with $n$ slots remaining in the horizon, the expected cost-to-go of transmitting packets to cover the user’s playout requirements for the next $j - 1$ slots is the same as the expected cost-to-go of transmitting packets to cover the user’s requirements for the next $j$ slots.
**Theorem 3:** Define the thresholds $\gamma_{n,j}$ for $n \in \{1, 2, \ldots, N\}$ and $j \in \mathbb{N}$ recursively, as follows:

(i) If $j = 1$, $\gamma_{n,j} = \infty$;

(ii) If $j > n$, $\gamma_{n,j} = 0$;

(iii) If $2 \leq j \leq n$,

\[
\gamma_{n,j} = -h + \alpha \cdot \left( \sum_{c: c \geq \gamma_{n-1,j-1}} p_c \cdot \gamma_{n-1,j-1} + \sum_{c: c < \gamma_{n-1,j-1}} p_c \cdot c \right) + \sum_{c: c < \gamma_{n-1,j} + L(c) - 1} p_c \cdot \left[ \gamma_{n-1,j} + L(c) - 1 - c \right],
\]

(8)

where $p_c$ is the probability of the channel being in state $c$ in a time slot, and $L(c) := \frac{p_c}{\alpha}$. For each $n \in \{1, 2, \ldots, N\}$ and $c \in \mathcal{C}$, if $\gamma_{n,j+1} < c \leq \gamma_{n,j}$, define $b_n(c) := j \cdot d$. The optimal control strategy is then given by $\pi^* = \{y^*_N, y^*_{N-1}, \ldots, y^*_1\}$, where

\[
y^*_n(x, c) := \begin{cases} 
x, & \text{if } x \geq b_n(c) \\
b_n(c), & \text{if } b_n(c) - \frac{p_c}{\alpha} \leq x < b_n(c) \\
x + \frac{p_c}{\alpha}, & \text{if } x < b_n(s) - \frac{p_c}{\alpha}
\end{cases}.
\]

(9)

Note that with $n$ slots remaining, $0 = \gamma_{n,n+1} \leq \gamma_{n,n} \leq \gamma_{n,n-1} \leq \ldots \leq \gamma_{n,2} \leq \gamma_{n,1} = \infty$, so $b_n(c)$ is well-defined. Compared to using the dynamic program to compute the optimal policy, the above result not only sheds more insight on the structural properties of the problem and its optimal solution, but also offers a computationally simpler method. In particular, the optimal policy is completely characterized by the thresholds $\gamma_{n,j}$. Calculating these thresholds recursively, as described in Theorem 3, is considerably simpler from a computational standpoint than solving the full dynamic program.

The proof of Theorem 3 follows a similar technique to the proof of Golabi’s Theorem 1 [14]. Specifically, we show by backwards induction that it is worse to transmit either fewer or more packets than the number suggested by the policy $\pi^*$. While not a complete repetition of [14], the proof of Theorem 3 is rather lengthy and tedious, and is therefore omitted here. The detailed proof can be found in [20], and we discuss the intuition behind the proof further in Section V-A.
The reason for the technical condition regarding the maximum number of packets that can be transmitted in any slot is as follows. The optimal action at all times (in general, without the technical condition) is either to transmit enough packets to fill the buffer up to a level satisfying the playout requirements of some number of future slots, or to transmit at maximum power. When the technical condition is satisfied, transmitting at maximum power also results in filling the buffer up to a level satisfying the playout requirements of some number of future slots. Thus, under the optimal policy, all realizations will result in the buffer level at the end of every time slot being some integer multiple of the demand, $d$. This fact makes it easier to compute the thresholds $\gamma_{n,j}$.

V. DISCUSSION

In this section, we try to provide some more intuition behind our results by comparing them to related work in both inventory theory and wireless communications theory.

A. Comparison to Related Work in Inventory Theory

Kingsman [9], [10], Magirou [11], [12], and Golabi [13], [14] all consider the single-item, discrete time inventory model with random ordering prices, deterministic demands, and no resource constraint. As mentioned in Section II-C, the optimal policy for that problem is a base-stock policy with different target stock levels for each price. The key intuition behind the calculation of the target stock levels is as follows. Let’s say that, as shown in Figure 2, the demand is two items (packets) per slot and there is enough inventory (packets in the receiver’s buffer) at time $N$ to satisfy the demand (playout requirements) for the next two slots. Thus, the first two items purchased (packets transmitted) in the current slot would be used to satisfy the demand in slot $[N-2, N-3)$. If the scheduler waits to purchase those items until time $N-2$, it must take whatever price is offered, and the per item cost of doing so is the average price (average power consumption per packet sent). Therefore, if the scheduler waits to purchase those packets until time $N-1$, it will decide if purchasing them at the offered price and holding them for one slot is less expensive (on a discounted basis) than waiting one more slot and paying the
average price at time $N - 2$. The decision, of course, will depend on the offered price in slot $[N - 1, N - 2)$. If the offered price plus the holding cost is less than the discount rate times the average price, it will purchase them at time $N - 1$; otherwise, it will not. From this set of decisions, the scheduler can calculate the expected per item cost of those items if it does not purchase them in the current slot. It can then compare that cost (discounted appropriately) to the cost of purchasing them at the current price and holding the items for one slot. If the current offered price is low enough to decide to purchase the packets to satisfy the demand in slot $[N - 2, N - 3)$, it then repeats the same process to determine whether to purchase additional packets to satisfy the demand in slot $[N - 3, N - 4)$. In this manner, the base-stock levels can be found by calculating a series of thresholds or “price breaks” [9] that determine how many periods of future demand should be satisfied, given the offered price in the current slot.

The intuition behind the computation of the base-stock levels in our restricted case is quite similar to the unrestricted case; however, the calculation of these thresholds is slightly different, as we must incorporate the transmission limitations imposed by the power constraint in each slot. Comparing the thresholds $\gamma_{n,j}$ defined in (8) to the corresponding thresholds in the unrestricted (no power constraint) single user problem [9], [14], the only difference is the third term of the right-hand side of (8):

$$\alpha \cdot \sum_{\{c: c<\gamma_{n-1,j+L(c)-1}\}} p_c \cdot \left[\gamma_{n-1,j+L(c)-1} - c\right],$$

which is absent in the unrestricted case. For all $n \in \{1, 2, \ldots, N\}$ and $j \in \mathbb{N}$, this term is
non-negative. Thus, for a fixed $n$ and $j$, the threshold in the restricted case is at least as high as the corresponding threshold in the unrestricted case. It follows that the optimal stock-up level $b_n(c)$ is also at least as high in the restricted case for all $n \in \{1, 2, \ldots, N\}$ and $c \in C$. The intuition behind this difference is that the sender should transmit more packets under the same (medium) conditions, because it is not able to take advantage of the best channel conditions to the same extent due to the power constraint.

Kingsman [9], [10] also extends the above unrestricted problem by considering a restriction on the maximum quantity that can be ordered (transmitted) at the offered price (rate) in any time slot. In Kingsman’s problem, however, the restriction is the same in every slot, regardless of the realization of the stochastic ordering price (channel condition). The same intuition of having to transmit more packets under medium conditions applies to that problem, but the calculation of the thresholds is different.

**B. Comparison to Related Work in Wireless Communications Theory**

The main idea of energy-efficient communication over a fading channel via opportunistic scheduling is to minimize power consumption by transmitting at the times when the channel is in its best condition. However, rate constraints, delay or deadline constraints, per-slot power constraints, and (in the case of multiple senders or receivers) fairness constraints may all result in higher power consumption. In this section, we try to elucidate the role of some of these constraints by examining a series of related problems. The first problem is the infinite horizon problem of minimizing average power subject to an average rate constraint and per-slot power constraints. The second problem is a finite horizon variant of the first problem. With a finite horizon, minimizing average power subject to an average rate constraint is equivalent to minimizing total power subject to a total throughput constraint; however, the restriction to a finite horizon introduces a deadline constraint, as all packets must be transmitted by the end of the horizon. Finally, we return to the problem discussed in Section IV, and show how this can be interpreted as having multiple deadline constraints. The goal of this progression is to understand better the effect of the deadline constraints on the optimal transmission policy. The basic intuition is that
the extra delay constraints prevent the transmitter from waiting as long for the best channel conditions, and, therefore, more packets need to be sent under worse channel conditions.

For ease of comparison, throughout this section, we consider an IID channel, take the discount factor, $\alpha$, to be 1, and do not include any holding costs. We start by considering a channel that can be in one of $M$ channel conditions, with probabilities $p^1, p^2, \ldots, p^M$, respectively. Associated with each channel condition $i$ is a linear power-rate function with slope $\phi^i$. We assume without loss of generality that $\phi^1 \leq \phi^2 \leq \ldots \leq \phi^M$ (i.e., $\phi^1$ is the best channel condition and $\phi^M$ is the worst). There is also a power constraint $P$ in each slot. Thus, under condition $i$, the maximum number of packets that can be transmitted is $\frac{P}{\phi^i}$. The objective is to minimize the average power consumed over an infinite horizon, subject to a minimum average rate constraint, $\bar{R}$, and the power constraints in each slot. This problem reduces to the following convex optimization problem:

$$\min_{(z^1, z^2, \ldots, z^M) \in \mathbb{R}_{+}^M} \sum_{i=1}^{M} p^i \cdot \phi^i \cdot z^i$$

s.t. \hspace{.5cm} \sum_{i=1}^{M} p^i \cdot z^i \geq \bar{R}

and \hspace{.5cm} z^i \leq \frac{P}{\phi^i}, \forall i \in \{1, 2, \ldots, M\} \hspace{.5cm} (10)$$

where $z^i$ represents the number of packets transmitted when the channel is in condition $i$. The solution to (10) is found by defining $j^* := \min \left\{ j \in \{1, 2, \ldots, M\} : \sum_{m=1}^{j} p^m \cdot \frac{P}{\phi^m} \geq \bar{R} \right\}$. Then the optimal amount of data to send under each channel condition is given by:

$$z^{m^*} := \begin{cases} 
\frac{P}{\phi^{m^*}}, & \text{if } m < j^* \\
\frac{R - \sum_{m=1}^{j^*-1} p^m \cdot \frac{P}{\phi^m}}{p^{j^*}}, & \text{if } m = j^* \\
0, & \text{if } m > j^* 
\end{cases} \hspace{.5cm} (11)$$

See Figure 3 for a diagram of this solution.

Fu et al. [5, Section III-D] consider a related problem of transmitting a given amount of data with minimum energy by a fixed deadline. They also represent the fading channel by a linear power-rate function with a different slope in each channel condition, and consider a power constraint $P$ in each slot. Let’s say the horizon to deliver $d_{\text{total}}$ data packets is $N$, and the power
consumed per packet transmitted in slot $n$ is given by the IID sequence of random variables \{\Psi_N, \Psi_{N-1}, \ldots, \Psi_1\}. If, at the beginning of the horizon, the scheduler happens to know the realizations of all future channel conditions, \{\psi_N, \psi_{N-1}, \ldots, \psi_1\}, then this problem reduces to the following convex optimization problem:

$$
\min_{(\hat{z}_N, \hat{z}_{N-1}, \ldots, \hat{z}_1)\in\mathbb{R}^N_+} \sum_{i=1}^N \psi_i \cdot \hat{z}_i \\
\text{s.t.} \quad \sum_{i=1}^N \hat{z}_i \geq d_{\text{total}} \\
\text{and} \quad \hat{z}_i \leq \frac{P}{\psi_i}, \ \forall i \in \{1, 2, \ldots, N\}, \quad (12)
$$

where $\hat{z}_i$ represents the number of packets transmitted in slot $i$. It should be clear that (12) is essentially the same problem as (10), and the solution can be found by scheduling data transmission during the slot with the best condition until all the data is sent or the power limit is reached, and then scheduling data transmission during the slot with the second best condition until all the data is sent or the power limits is reached, and so forth. See Figure 4 for a diagram of this solution. If we are focused on finding the optimal amount to transmit in the current slot, we can also aggregate the power-rate functions of all future slots, by reordering them according to the strength of the channel, as shown in Figure 5. The optimal number of packets to transmit in the current slot is then determined as follows. Define $z_0 := \min \{z : \psi_0(\tilde{z}) \geq \psi_N, \ \forall \tilde{z} > z\}$, where $\psi_0(\cdot)$ is the slope from above of the aggregate power-rate curve. Then the optimal number
of packets to transmit in slot $N$ is given by:

$$\hat{z}_N^* = \min \left\{ \frac{P}{\psi_N}, \max \left( d_{\text{total}} - z_0, 0 \right) \right\}. \quad (13)$$

This policy says that if the slope of the current curve is greater than the slope of the aggregate curve at all points up to $d_{\text{total}}$, then it is optimal to not transmit any packets in the current slot. Otherwise, the optimal number of packets to transmit in the current slot $N$ is the minimum of the maximum number of packets that can be transmitted under the current channel condition, and the number of packets that would otherwise be transmitted in worse channel conditions in future slots. Note that this policy can also be viewed as a modified base-stock policy of the form described in (4), with $x = 0$ and critical number equal to $\max \{ d_{\text{total}} - z_0, 0 \}$.

Now, as Fu et al. explain, in the more realistic case that the channel condition in slot $n$ is not learned until the beginning of the $n^{th}$ slot, a very similar aggregate method can be used as long as the number of possible channel conditions is finite, and for all realizations $\psi$ of the channel condition $\Psi$ (the power cost per packet transmitted), $\psi = \psi_{\hat{l}}^\max$ for some $\hat{l} \in \mathbb{N}$. In this situation, however, the slopes of the piecewise-linear aggregate power-rate function for future slots are not defined in terms of the actual channel conditions of future slots (which are not available), but rather by a series of thresholds that only depend on the statistics of future channel conditions. With $N$ slots remaining, the slopes of the aggregate curve can only
change at integer multiples of \( \frac{P}{\psi_{\max}} \), and the form of the optimal policy at time \( N \) is the same as (13), with \( d_{\text{total}} \) being the number of packets remaining to transmit at time \( N \), and \( z_0 \in \left\{ 0, \frac{P}{\psi_{\max}}, 2 \cdot \frac{P}{\psi_{\max}}, \ldots, (N - 1) \cdot \frac{P}{\psi_{\max}} \right\} \).

We now return to the wireless streaming model considered in Section IV of this paper, with \( d \) packets removed from the receiver’s buffer at the end of every slot. Let us once again begin by considering the unrealistic case that the scheduler knows all future channel conditions at the beginning of the horizon. The optimal solution can be found by using the same basic water-filling type principle of transmitting as much as possible in the slot with the best channel condition, and then the second best, and so forth; however, due to the additional underflow constraints, one needs to solve \( N \) sequential problems of this form. The first problem is the trivial problem of sending \( d \) packets in the first slot, \( \lfloor N, N - 1 \rfloor \). The second problem is to send \( 2d \) packets in the first two slots. If the power limit in the first slot has not been reached after allocating the initial \( d \) packets there, then the scheduler may choose to send the second batch of \( d \) packets in either the first or second slot, according to their respective channel conditions. For each sequential
problem, whatever packets have been allocated in the previous problem must be “carried over” to the subsequent problem, where there is one additional time slot available and the next \(d\) packets are allocated. The solution to the \(N^{th}\) problem represents the optimal allocation. See Figure 6 for a diagram of this solution. Comparing Figure 6 to Figure 4, we see that when \(N \cdot d = d_{\text{total}}\) and the known sequence of channel conditions is the same for both problems, the additional underflow constraints cause more data to be scheduled in earlier time slots with worse channel conditions.

Fig. 6. Pictorial representation of the solution to the wireless streaming model considered in Section IV in the somewhat unrealistic case that all future channel conditions are known at the beginning of the horizon. In the example shown, the time horizon is \(N = 6\), \(d\) packets are removed from the receiver’s buffer at the end of every slot, and the power constraint in each slot is \(P = 4d\). To satisfy the underflow constraints, 6 sequential problems are considered, with an additional \(d\) packets allocated in each problem. Packets allocated in one problem are “carried over” to all subsequent problems, and shown in solid black filling. The optimal policy, given by the solution to Problem 6, is to transmit \(d\) packets in slots 6 and 3, and \(2d\) packets in slots 5 and 2. This policy results in a cost of \(3P\) units of power.
When all future channel conditions are known ahead of time, as in Figure 6, we can also use the same aggregation technique from above to represent Problems 2 through \( N \) as comparisons between the current channel condition and the aggregate of the future channel conditions. Furthermore, when the future channel conditions are not known ahead of time, we can once again define the aggregate power-rate function for future slots in terms of a series of thresholds that only depend on the statistics of future channel conditions. Due to the underflow constraints, however, these thresholds are computed differently than those in [5, Section III-D]. The net result for this more realistic case is the same as the case when all future channel conditions are known - the additional underflow constraints make it optimal to send more data in earlier time slots with worse channel conditions.

Finally, we show that the problem in [5, Section III-D] and our finite horizon problem are equivalent when certain conditions are satisfied. There is just a single explicit underflow constraint (the deadline) in the problem in [5, Section III-D]; however, because the terminal cost is set to \( \infty \) if all the data is not transmitted by the deadline, the scheduler must transmit enough data in each slot so that it can still complete the job if the channel is in the worst possible condition in all subsequent slots. Thus, if \( d_{total} \) is the total amount of data that must be sent by the deadline and \( d_{worst} \) is the amount that can be sent in a slot under the worst channel condition, the transmitter must have sent at least \( d_{total} - d_{worst} \) packets by the beginning of the last slot, at least \( d_{total} - 2 \cdot d_{worst} \) packets by the beginning of the second to last slot, and so forth. (An unstated assumption in their formulation is that \( d_{worst} \) times the horizon length must be at least as large as \( d_{total} \).) So there are in fact implicit constraints on how much data must be transmitted by the end of slots \( N - \left\lceil \frac{d_{total}}{d_{worst}} \right\rceil + 1, \ldots, N - 2, N - 1 \). With this interpretation, we believe that our Theorem 3 is equivalent to Theorem 3 and its corollary in [5] in the special case that, in addition to the hypotheses of our Theorem 3, \( \alpha = 1, \ h = 0, \) and \( L(c_{\max}) = 1 \). For, when these conditions are met, the implicit constraints in [5] coincide exactly with the explicit underflow constraints in our problem. Of course, when these three conditions are not satisfied, the two problems are quite different, as the preceding illustrations have shown.

In summary, we first compared our problem to a problem in inventory theory that has similar
underflow constraints, but no per-slot power constraints. We then compared our problem to the wireless communication problem of transmitting data by a deadline, which has similar per-slot power constraints, but only one underflow constraint (the deadline). At a very high level, the relative effect of the extra constraints (per-slot power constraints and underflow constraints, respectively) is the same - the optimal transmission policy sends more packets under the same “medium” channel conditions, in anticipation of the need to comply with these constraints in future slots.

VI. FUTURE WORK - EXTENSION TO THE CASE OF MULTIPLE RECEIVERS

In this section, we discuss briefly the extension of the problem to the case of a single source transmitting media streams to multiple users over a shared wireless channel. The system operation is similar to the single user case discussed in this paper, with the sender maintaining separate data buffers for each user’s stream, and underflow constraints being imposed on each user’s playout buffer. The channel conditions vary asynchronously for different users as they experience different interference levels depending on location and other factors. In each time slot, the scheduler allocates some amount of power (possibly zero) for transmission to each user, subject to a per-slot total power constraint. See Fig. 7 for a diagram of the system with multiple receivers.

![System model for the case of multiple receivers.](image)

The main idea in the single user case is to reduce power consumption by exploiting the temporal variation of the channel; i.e., opportunistically scheduling transmissions for the time slots in which the channel is in its best conditions. In the case of multiple receivers, the main
idea is to reduce power consumption by exploiting both the temporal and the spatial variation of the channel. With multiple receivers, the scheduler not only gets to opportunistically choose when to transmit, but also gets to opportunistically choose which user(s) it will transmit to in each time slot. The additional benefit of doing so is commonly referred to as the multiuser diversity gain (see, e.g., [21, Ch. 6]).

Based on numerical experiments, it appears that the optimal policy for the single user case extends to the case of multiple users, in the following sense:

1) At each time and for each possible vector of the users’ channel conditions, there exists a vector of critical numbers (corresponding to \( b_n(c) \) in the single user case), with one critical number for each user.

2) Each user’s critical number depends only on its current channel condition, and is independent of its own current buffer level, other users’ current buffer levels, and other users’ current channel conditions.

3) It is optimal for the transmitter to not transmit any packets to any user whose current buffer level is greater than or equal to its critical number.

4) If it is possible for the transmitter to schedule packet transmissions to bring all other users’ buffer levels up to their respective critical numbers without exceeding the power constraint, then it is optimal for it to do so.

5) If the power constraint prevents the transmitter from doing so, then it should allocate (in some manner yet to be determined) the full power \( P \) for transmission to different receivers whose buffer levels are below their respective critical numbers. It is not optimal to transmit so as to cause any of these users’ buffer levels to exceed their critical numbers following transmission.

For a more precise problem formulation and description of our conjecture for the multiple user case, see [22]. We are currently working on the proof of this conjecture.
VII. Conclusion

In this paper, we considered the problem of transmitting data over a wireless channel in a manner that minimizes power consumption and prevents the receiver’s buffer from emptying. We showed that under both the finite and infinite horizon discounted expected cost criteria, the optimal transmission schedule is a modified base-stock policy. For the special case when holding costs are linear and the stochastic process representing the channel condition over time is IID, has a finite sample space, and satisfies an additional technical condition, we showed how to calculate the critical numbers that complete the characterization of the optimal policy.

VIII. Appendix

A. Proof of Theorem 1

Before proceeding to the proof of Theorem 1, we present a lemma due to Karush [23], which is presented in [24, pp. 237–238].

Lemma 1 (Karush, 1959): Suppose that $f : \mathbb{R} \to \mathbb{R}$ and that $f$ is convex on $\mathbb{R}$. For $v \leq w$, define $g(v, w) := \min_{z \in [v, w]} f(z)$. Then it follows that:

(a) $g$ can be expressed as $g(v, w) = F(v) + G(w)$, where $F$ is convex nondecreasing and $G$ is convex nonincreasing on $\mathbb{R}$.

(b) Suppose that $S$ is a minimizer of $f$ over $\mathbb{R}$. Then $g$ can be expressed as:

$$
g(v, w) = \begin{cases} 
    f(v), & \text{if } S \leq v \\
    f(S), & \text{if } v \leq S \leq w \\
    f(w), & \text{if } w \leq S 
\end{cases}
$$

Proof of Theorem 1: We present the proof in three parts.
**Part I - Modified Base-Stock Structure:** Recall the dynamic program (3):

\[
V_n(x, c) = -c \cdot x + \underbrace{\min_{\max(x,d) \leq y \leq x+\frac{P}{c}}} \{c \cdot y + h(y-d) + \alpha \cdot \mathbb{E}[V_{n-1}(y-d, C_{n-1}) \mid C_n = c]\}
\]

\[
V_0(x, c) = 0, \ \forall x, \forall c,
\]

where \(g_n(y, c) := c \cdot y + h(y-d) + \alpha \cdot \mathbb{E}[V_{n-1}(y-d, C_{n-1}) \mid C_n = c]\). We now show by induction on \(n\) that the following statements are true for every \(n \in \{1, 2, \ldots, N\}\) and all \(c \in C\):

(i) \(g_n(y, c)\) is convex in \(y\) on \([d, \infty)\).

(ii) \(\lim_{y \to \infty} g_n(y, c) = \infty\).

(iii) \(V_n(x, c)\) is convex in \(x\) on \(\mathbb{R}_+\).

**Base Case:** \(n = 1\)

Let \(c_1 \in C\) be arbitrary. We have \(g_1(y, c_1) = c_1 \cdot y + h(y-d)\), which clearly satisfies (i) and (ii).

\(y^*_1(x, c_1) = \max(x, d)\) and thus \(V_1(x, c_1) = c_1 \cdot (d - x)^+ + h((x - d)^+)\), which is convex in \(x\).

We conclude (i)-(iii) are true at time \(n = 1\), for all \(c \in C\).

**Induction Step:** We now assume (i)-(iii) are true for \(n = m - 1\) and all \(c \in C\), and show they hold for \(n = m\) and an arbitrary \(c_m \in C\). Let \(c_{m-1} \in C\) also be arbitrary. \(V_{m-1}(y-d, c_{m-1})\) is convex in \(y\), as it is the composition of a convex function, \(V_{m-1}(\cdot, c_{m-1})\), with an affine function, \(y-d\). \(g_m(y, c_m)\) is therefore convex in \(y\) as it is the sum of an affine function, \(c_m \cdot y\), a convex function, \(h(y-d)\), and a nonnegative weighted sum/integral of convex functions, \(\alpha \cdot \mathbb{E}[V_{m-1}(y-d, C_{m-1}) \mid C_m = c_m]\) (see, e.g., [25, Section 3.2] for the relevant results on convexity-preserving operations).

To show (ii) for \(n = m\), we have \(\lim_{y \to \infty} g_m(y, c_m) \geq \lim_{y \to \infty} c_m \cdot y = \infty\), where the inequality follows from \(V_{m-1}(x, c_{m-1}) \geq 0, \forall x \in \mathbb{R}_+, \forall c_{m-1} \in C\) and \(h(y-d) \geq 0\). Moving on to (iii), we have:

\[
V_m(x, c_m) = -c_m \cdot x + \underbrace{\min_{\max(x,d) \leq y \leq x+\frac{P}{c_m}}} \{g_m(y, c_m)\}
\]

\[
= -c_m \cdot x + F(\max(x, d)) + G(x + \frac{P}{c_m}),
\]
where, by Lemma 1, $F$ is convex nondecreasing and $G$ is convex nonincreasing. $F(\max(x,d))$ is also convex in $x$, as it is the composition of a convex increasing function with a convex function, and $V_m(x, c_m)$ is therefore convex in $x$. This concludes the induction step, and we conclude (i)-(iii) are true for all $n \in \{1, 2, \ldots, N\}$.

Next, we define the critical numbers $b_n(c)$ for all $n \in \{1, 2, \ldots, N\}$ and $c \in C$:

$$b_n(c) := \min \left\{ \hat{y} \in [d, \infty) : g_n(\hat{y}, c) = \min_{y \in [d, \infty]} g_n(y, c) \right\} .$$

Note that by properties (i) and (ii) from the above induction, the minimum of $g_n(\cdot, c)$ over $[d, \infty)$ is achieved, and the set of minimizers over $[d, \infty)$ is a closed, convex set. Thus, $b_n(c)$ is well-defined. The form of $y^*_n(x, c)$, (4), then follows from part (b) of Karush’s result, Lemma 1, with $g_n(y, c)$ playing the role of $f$, $\max(x, d)$ the role of $v$, $x + \frac{P}{c}$ the role of $w$, and $b_n(c)$ the role of $S$.

**Part II - Monotonicity of Thresholds in Time:** In this section, we prove (5). We showed above that the optimal action with one time slot remaining is $y^*_1(x, c) = \max(x, d)$, for all $c \in C$. This is precisely the policy suggested by (4) with $b_1(c) = d$, as $\frac{P}{c}$ is at least as great as $d$. Thus, we conclude the far right equality in (5) holds: $b_1(c) = d, \forall c \in C$.

In order to show the far left inequality in (5), we claim more generally that $b_n(c) \leq n \cdot d$, for all $n$ and $c$. This follows from a simple interchange argument, as all packets transmitted beyond $n \cdot d$ incur transmission costs and holding costs for the duration of the horizon; however, they do not satisfy the playout requirements in any remaining slot. Thus, a policy that transmits enough packets to fill the buffer up to $n \cdot d$ at time $n$ is strictly superior to a policy that transmits more packets.

Next, we prove:

$$b_{n+1}(c) \geq b_n(c), \forall c \in C, \forall n \in \{1, 2, \ldots, N - 1\} .$$  \hspace{1cm} (14)
To show (14), it suffices to show the following submodularity condition [26] holds for all \( c \in \mathcal{C} \), \( n \in \{1, 2, \ldots, N - 1\} \), and \( y^1, y^2 \in [d, (n + 1) \cdot d] \):

\[
g_n(y^1, c) + g_{n+1}(y^2, c) \geq g_n(y^1 \land y^2, c) + g_{n+1}(y^1 \lor y^2, c) \tag{15}
\]

Before proceeding to show (15), we show why (15) suffices to prove (14). Assume (15) holds and \( b_{\tilde{n}+1}(\tilde{c}) < b_{\tilde{n}}(\tilde{c}) \) for some \( \tilde{c} \in \mathcal{C} \) and \( \tilde{n} \in \{1, 2, \ldots, N - 1\} \). Then we have:

\[
g_{\tilde{n}}(b_{\tilde{n}}(\tilde{c}) \land b_{\tilde{n}+1}(\tilde{c}), \tilde{c}) + g_{\tilde{n}+1}(b_{\tilde{n}}(\tilde{c}) \lor b_{\tilde{n}+1}(\tilde{c}), \tilde{c})
\leq g_{\tilde{n}}(b_{\tilde{n}+1}(\tilde{c}), \tilde{c}) + g_{\tilde{n}+1}(b_{\tilde{n}}(\tilde{c}), \tilde{c})
\geq g_{\tilde{n}}(b_{\tilde{n}+1}(\tilde{c}), \tilde{c}) + g_{\tilde{n}+1}(b_{\tilde{n}+1}(\tilde{c}), \tilde{c}) \tag{16}
\]

where the last inequality follows from the facts that \( b_{\tilde{n}+1}(\tilde{c}) \) is a minimizer of \( g_{\tilde{n}+1}(\cdot, \tilde{c}) \) and \( b_{\tilde{n}}(\tilde{c}) \) is the smallest minimizer of \( g_{\tilde{n}}(\cdot, \tilde{c}) \). Since both \( b_{\tilde{n}}(\tilde{c}) \) and \( b_{\tilde{n}+1}(\tilde{c}) \) are less than or equal to \((n + 1) \cdot d\), (16) contradicts (15), and we conclude that (15) suffices to prove (14).

So we now return to showing that (15) is true. If \( y^1 \leq y^2 \), then (15) holds trivially, so we assume throughout this section that \( y^1 > y^2 \), in which case (15) reduces to:

\[
g_n(y^1, c) + g_{n+1}(y^2, c) \geq g_n(y^2, c) + g_{n+1}(y^1, c) \tag{17}
\]

We proceed to show (17) by induction on the time slot \( n \).

**Base Case:** \( n = 1 \)

Let \( c \in \mathcal{C} \) and \( y^1, y^2 \in [d, 2 \cdot d] \) be arbitrary, with \( y^1 > y^2 \). We have:

\[
g_1(y^1, c) + g_2(y^2, c)
= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) + \alpha \cdot \mathbb{E} [V_1(y^2 - d, C_1) \mid C_2 = c]
= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) + \alpha \cdot \mathbb{E} [C_1 \mid C_2 = c] \cdot (2 \cdot d - y^2) + \\
+ \alpha \cdot h((y^2 - 2 \cdot d)^+)
= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) + \alpha \cdot \mathbb{E} [C_1 \mid C_2 = c] \cdot (2 \cdot d - y^2)
\geq c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) + \alpha \cdot \mathbb{E} [C_1 \mid C_2 = c] \cdot (2 \cdot d - y^1)
\]
\[= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) + \alpha \cdot \mathbb{E} [C_1 \mid C_2 = c] \cdot (2 \cdot d - y^1) \]
\[+ \alpha \cdot h\left((y^1 - 2 \cdot d)^+\right)\]
\[= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) + \alpha \cdot \mathbb{E} [V_1 (y^1 - d, C_1) \mid C_2 = c]\]
\[= g_1 (y^2, c) + g_2 (y^1, c).\]

**Induction Step:** We assume that (17) is true for all \(n = 1, 2, \ldots, m - 1, c \in \mathcal{C},\) and \(y^1, y^2 \in [d, (n + 1) \cdot d].\) We wish to show it is true for \(n = m.\) Let \(c \in \mathcal{C}\) and \(y^1, y^2 \in [d, (m + 1) \cdot d]\) be arbitrary, with \(y^1 > y^2.\) Also, let \(\hat{c} \in \mathcal{C}\) be arbitrary. Define:

\[
\beta_1 := \min \left\{ \arg\min_{\max(y^2 - d, d) \leq \hat{y} \leq y^1 - d + \frac{P}{\hat{c}}} \{g_{m-1}(\hat{y}, \hat{c})\} \right\}
\]

and

\[
\beta_2 := \min \left\{ \arg\min_{\max(y^2 - d, d) \leq \hat{y} \leq y^2 - d + \frac{P}{\hat{c}}} \{g_m(\hat{y}, \hat{c})\} \right\}.
\]

Note that:

\[
\max\left(y^2 - d, d\right) \leq \beta_1 \land \beta_2 \leq \beta_2 \leq y^2 - d + \frac{P}{\hat{c}}, \tag{18}
\]

where the first inequality follows from \(\max(y^2 - d, d) \leq \beta_2\) and \(\max(y^2 - d, d) \leq \max(y^1 - d, d) \leq \beta_1.\) Similarly,

\[
\max\left(y^1 - d, d\right) \leq \beta_1 \lor \beta_2 \leq y^1 - d + \frac{P}{\hat{c}}, \tag{19}
\]

where the last inequality follows from \(\beta_1 \leq y^1 - d + \frac{P}{\hat{c}}\) and \(\beta_2 \leq y^2 - d + \frac{P}{\hat{c}} \leq y^1 - d + \frac{P}{\hat{c}}.\)

From these definitions, we see:

\[
\min_{\max(y^1 - d, d) \leq \hat{y} \leq y^1 - d + \frac{P}{\hat{c}}} \{g_{m-1}(\hat{y}, \hat{c})\} + \min_{\max(y^2 - d, d) \leq \hat{y} \leq y^2 - d + \frac{P}{\hat{c}}} \{g_m(\hat{y}, \hat{c})\}
\]
\[= g_{m-1}(\beta_1, \hat{c}) + g_m(\beta_2, \hat{c})
\]
\[\geq g_{m-1}(\beta_1 \land \beta_2, \hat{c}) + g_m(\beta_1 \lor \beta_2, \hat{c})
\]
\[\geq \min_{\max(y^2 - d, d) \leq \hat{y} \leq y^2 - d + \frac{P}{\hat{c}}} \{g_{m-1}(\hat{y}, \hat{c})\} + \min_{\max(y^1 - d, d) \leq \hat{y} \leq y^1 - d + \frac{P}{\hat{c}}} \{g_m(\hat{y}, \hat{c})\}, \tag{20}
\]

where the first inequality in (20) is trivial if \(\beta_1 \leq \beta_2,\) and follows from the induction hypothesis if \(\beta_2 < \beta_1.\) The second inequality in (20) follows from (18) and (19). Since \(\hat{c}\) was arbitrary, (20)
holds for all \( \hat{c} \in \mathcal{C} \). Therefore, combined with the fact that the Markov process \( \{C_n\}_{n=N,N-1,\ldots,1} \) is homogeneous, (20) implies:

\[
\mathbb{E} \left[ \min_{\max(y^1-d,d) \leq \hat{y} \leq y^1-d+\frac{\rho}{c_m-1}} \{g_{m-1}(\hat{y},C_{m-1})\} \mid C_m = c \right]
\]

\[
+ \mathbb{E} \left[ \min_{\max(y^2-d,d) \leq \hat{y} \leq y^2-d+\frac{\rho}{c_m-1}} \{g_m(\hat{y},C_m)\} \mid C_{m+1} = c \right]
\]

\[
\geq \mathbb{E} \left[ \min_{\max(y^2-d,d) \leq \hat{y} \leq y^2-d+\frac{\rho}{c_m-1}} \{g_{m-1}(\hat{y},C_{m-1})\} \mid C_m = c \right]
\]

\[
+ \mathbb{E} \left[ \min_{\max(y^1-d,d) \leq \hat{y} \leq y^1-d+\frac{\rho}{c_m-1}} \{g_m(\hat{y},C_m)\} \mid C_{m+1} = c \right].
\]

(21)

We can now use (21) to show (17) holds for \( n = m \):

\[
g_m(y^1,c) + g_{m+1}(y^2,c)
\]

\[
= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) + \alpha \cdot \mathbb{E} [V_{m-1}(y^1-d,C_{m-1}) \mid C_m = c]
\]

\[
+ \alpha \cdot \mathbb{E} [V_m(y^2-d,C_m) \mid C_{m+1} = c]
\]

\[
= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) - \alpha \cdot \mathbb{E} [C_{m-1} \mid C_m = c] \cdot (y^1 - d)
\]

\[
- \alpha \cdot \mathbb{E} [C_m \mid C_{m+1} = c] \cdot (y^2 - d) + \alpha \cdot \mathbb{E} \left[ \min_{\hat{y} \in \mathcal{A}(y^1-d,C_{m-1})} \{g_{m-1}(\hat{y},C_{m-1})\} \mid C_m = c \right]
\]

\[
+ \alpha \cdot \mathbb{E} \left[ \min_{\hat{y} \in \mathcal{A}(y^2-d,C_m)} \{g_m(\hat{y},C_m)\} \mid C_{m+1} = c \right]
\]

\[
= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) - \alpha \cdot \mathbb{E} [C_{m-1} \mid C_m = c] \cdot (y^2 - d)
\]

\[
- \alpha \cdot \mathbb{E} [C_m \mid C_{m+1} = c] \cdot (y^1 - d) + \alpha \cdot \mathbb{E} \left[ \min_{\hat{y} \in \mathcal{A}(y^1-d,C_{m-1})} \{g_{m-1}(\hat{y},C_{m-1})\} \mid C_m = c \right]
\]

\[
+ \alpha \cdot \mathbb{E} \left[ \min_{\hat{y} \in \mathcal{A}(y^2-d,C_m)} \{g_m(\hat{y},C_m)\} \mid C_{m+1} = c \right]
\]

\[
\geq c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) - \alpha \cdot \mathbb{E} [C_{m-1} \mid C_m = c] \cdot (y^2 - d)
\]

\[
- \alpha \cdot \mathbb{E} [C_m \mid C_{m+1} = c] \cdot (y^1 - d) + \alpha \cdot \mathbb{E} \left[ \min_{\hat{y} \in \mathcal{A}(y^1-d,C_{m-1})} \{g_{m-1}(\hat{y},C_{m-1})\} \mid C_m = c \right]
\]

\[
+ \alpha \cdot \mathbb{E} \left[ \min_{\hat{y} \in \mathcal{A}(y^2-d,C_m)} \{g_m(\hat{y},C_m)\} \mid C_{m+1} = c \right].
\]
\[= c \cdot (y^1 + y^2) + h(y^1 - d) + h(y^2 - d) + \alpha \cdot \mathbb{E} [V_{m-1} (y^2 - d, C_{m-1}) | C_m = c] + \alpha \cdot \mathbb{E} [V_m (y^1 - d, C_m) | C_{m+1} = c]
\]
\[= g_m (y^2, c) + g_{m+1} (y^1, c),
\]
where the third equality follow from the fact that \(\mathbb{E} [C_{m-1} | C_m = c] = \mathbb{E} [C_m | C_{m+1} = c]\), and the inequality follows from (21). This completes the induction step, and the proof of (5).

**Part III - Monotonicity of Thresholds in the Channel Condition:** Finally, we show (6), the monotonicity of the thresholds in the channel condition, when the channel condition process is IID. The far left inequality follows from the same interchange argument described above, showing \(b_m(c) \leq n \cdot d\) for all \(c\) and \(n\). We now show the far right equality of (6), \(b_n(c_{\max}) = d\). To satisfy feasibility, we must have \(b_n(c) \geq d\) for all \(n \in \{1, 2, \ldots, N\}\) and \(c \in \mathcal{C}\). To see that \(b_n(c_{\max}) \leq d\), assume the channel condition at time \(n\) is \(c_{\max}\), and consider two control policies satisfying (4), with the same critical numbers \(b_m(c)\), for all times \(m < n\). At time \(n\), the first policy, \(\pi^1\), transmits according to (4), with critical number \(b_n(c_{\max}) = d + \epsilon\) \((\epsilon > 0)\), and the second, \(\pi^2\), transmits according to (4), with critical number \(b_n(c_{\max}) = d\). These two strategies result in the same control action at time \(n\) if \(x_n \geq d + \epsilon\), and we have already shown it is not optimal to fill the buffer beyond \(n \cdot d\), so we only need to consider the case where \(x_n < d + \epsilon\) and \(\epsilon \leq (n - 1) \cdot d\).

Let \(Z^1_n, Z^1_{n-1}, \ldots, Z^1_1\) and \(Z^2_n, Z^2_{n-1}, \ldots, Z^2_1\) be random variables representing the number of packets transmitted at times \(n, n-1, \ldots, 1\) by \(\pi^1\) and \(\pi^2\), respectively. If \(d \leq x_n \leq d + \epsilon\), then \(Z^2_n = 0\) and \(Z^1_n - Z^2_n = Z^1_n = \min \left\{ \frac{P}{c_{\max}}, d + \epsilon - x_n \right\}\). If \(x_n < d\), then \(Z^2_n = d - x_n\), \(Z^1_n = \min \left\{ \frac{P}{c_{\max}}, d + \epsilon - x_n \right\}\), and \(Z^1_n - Z^2_n = \min \left\{ \frac{P}{c_{\max}} - d + x_n, \epsilon \right\}\). Thus, for all \(x_n < d + \epsilon\), we have \(Z^1_n - Z^2_n \geq 0\). If \(Z^1_n - Z^2_n = 0\), the two control policies result in the same actions for all remaining times, and therefore result in the same expected cost. So we only need to consider the case where \(\lambda := Z^1_n - Z^2_n > 0\). Because the critical numbers at times \(n-1, n-2, \ldots, 1\) are the same for both policies, for any realization, \(\omega\), of the channel condition over future times, we have \(Z^1_m(\omega) \leq Z^2_m(\omega), \forall m \in \{n-1, \ldots, 1\}\). Moreover, because the scheduler must satisfy the playout requirements for the last \(n\) slots, we have \(\sum_{m=1}^{n-1} (Z^2_m(\omega) - Z^1_m(\omega)) = \lambda\); i.e., over the remainder of the horizon, an extra \(\lambda\) packets are transmitted under the second policy. The total discounted holding costs from time \(n\) until the end of the horizon are therefore lower for \(\pi^2\).
than $\pi^1$, because the number of packets remaining after transmission in each slot is never greater under policy $\pi^2$. Furthermore, the total discounted transmission costs of the extra $\lambda$ packets are also lower for $\pi^2$ as they are transmitted at the maximum cost $c_{\text{max}}$ under $\pi^1$, and transmitted later (and therefore discounted more heavily) under $\pi^2$. Thus, the total discounted transmission plus holding costs are lower for $\pi^2$ under all realizations, and the expected discounted cost of $\pi^2$ is lower than $\pi^1$. We conclude $b_n(c_{\text{max}}) = d$.

To show $c^1 \leq c^2$ implies $b_n(c^1) \geq b_n(c^2)$, we follow Kalymon’s methodology for the proof of Theorem 1.3 in [8]. For all $y \in [d, \infty)$, we have:

$$g_n(y, c^2) = c^2 \cdot y + h(y - d) + \alpha \cdot \mathbb{E}[V_{n-1}(y - d, C_{n-1})]$$

$$= (c^2 - c^1) \cdot y + c^1 \cdot y + h(y - d) + \alpha \cdot \mathbb{E}[V_{n-1}(y - d, C_{n-1})]$$

$$= (c^2 - c^1) \cdot y + g_n(y, c^1).$$  \hspace{1cm} (22)

Assume $b_n(c^1) < b_n(c^2)$ for some $n \in \{1, 2, \ldots, N\}$ and $c^1, c^2 \in \mathcal{C}$, with $c^1 \leq c^2$. Substituting first $y = b_n(c^1)$ and then $y = b_n(c^2)$ into (22) yields:

$$(c^2 - c^1) \cdot b_n(c^1) + g_n(b_n(c^1), c^1) = g_n(b_n(c^1), c^2)$$

$$\geq g_n(b_n(c^2), c^2)$$

$$= (c^2 - c^1) \cdot b_n(c^2) + g_n(b_n(c^2), c^1).$$  \hspace{1cm} (23)

Yet, $c^1 \leq c^2$ and $b_n(c^1) < b_n(c^2)$ imply:

$$(c^2 - c^1) \cdot b_n(c^1) < (c^2 - c^1) \cdot b_n(c^2).$$  \hspace{1cm} (24)

Equations (23) and (24) imply:

$$g_n(b_n(c^1), c^1) > g_n(b_n(c^2), c^1),$$

which clearly contradicts the fact that $b_n(c^1)$ is a global minimizer of $g_n(\cdot, c^1)$. We conclude that $c^1 \leq c^2$ implies $b_n(c^1) \geq b_n(c^2)$, completing the proofs of (6) and Theorem 1.

\hspace{1cm} ■
B. Proof of Theorem 2

As mentioned in Section III-B, we leverage the techniques of [15] and [19]. We start the proof by showing part (a) of Theorem 2, \( \lim_{n \to \infty} V_n(x, s) \) exists and is finite, \( \forall x \in \mathbb{R}_+, \forall c \in C \).

Let \( x \in \mathbb{R}_+ \) and \( c \in C \) be arbitrary. First, we show inductively \( V_1(x, c) \leq V_2(x, c) \leq \ldots \leq V_n(x, c) \leq V_{n+1}(x, c) \leq \ldots \).

**Base Case:** \( n = 1 \)

\[
V_1(x, c) = -c \cdot x + \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \{c \cdot y + h(y - d)\} \\
\leq -c \cdot x + \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \{c \cdot y + h(y - d) + \alpha \cdot \mathbb{E}[V_1(y - d, C_1) | C_2 = c]\} \\
= V_2(x, c) ,
\]

where the inequality follows from \( V_1(x, c) \geq 0, \forall x, \forall c \).

**Induction Step:** Assume \( V_n(x, c) \leq V_{n+1}(x, c) \) for \( n = 1, 2, \ldots, m - 1 \). We show it is true for \( n = m \):

\[
V_m(x, c) = -c \cdot x + \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \{c \cdot y + h(y - d) + \alpha \cdot \mathbb{E}[V_{m-1}(y - d, C_{m-1}) | C_m = c]\} \\
\leq -c \cdot x + \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \{c \cdot y + h(y - d) + \alpha \cdot \mathbb{E}[V_m(y - d, C_m) | C_{m+1} = c]\} \\
= V_{m+1}(x, c) ,
\]

where the inequality follows from the induction hypothesis and the homogeneity of the Markov process representing the channel condition. So, for every \( x \in \mathbb{R}_+ \) and \( c \in C \), \( \{V_n(x, c)\}_{n=1,2,...} \) is a nondecreasing sequence. Next, consider a policy \( \pi^d \) transmitting \( d \) packets in every slot, regardless of channel condition. We have:

\[
V_n(x, c) \leq V_n^{\pi^d}(x, c) \leq (c_{\max} \cdot d + h(x)) \frac{1 - \alpha^n}{1 - \alpha} \leq (c_{\max} \cdot d + h(x)) \frac{1}{1 - \alpha} < \infty ,
\]

so \( \{V_n(x, c)\}_{n=1,2,...} \) is a bounded sequence, implying \( \lim_{n \to \infty} V_n(x, c) \) exists and is finite, \( \forall x \in \mathbb{R}_+, \forall c \in C \).

We now move on to part (b) of Theorem 2. Recall from Section VIII-A that \( V_n(x, c) \) is convex.
in $x$, for all $n$ and all $c$. Define $V_\infty(x, c) := \lim_{n \to \infty} V_n(x, c)$. Let $c \in \mathcal{C}$ be arbitrary, but fixed. $V_\infty(x, c) = \sup_{n \in \mathbb{N}} V_n(x, c)$, so $V_\infty(x, c)$ is convex in $x$ as it is the pointwise supremum of the convex functions $\{V_n(x, c)\}_{n=1,2,\ldots}$.

Define $g_\infty : [d, \infty) \times \mathcal{C} \to \mathbb{R}_+$ by

$$g_\infty(y, c) := c \cdot y + h(y - d) + \alpha \cdot \mathbb{E} [V_\infty(y - d, C') | C = c]$$

$$= c \cdot y + h(y - d) + \alpha \cdot \mathbb{E} \left[ \lim_{n \to \infty} V_n(y - d, C') | C = c \right]$$

$$= c \cdot y + h(y - d) + \alpha \cdot \lim_{n \to \infty} \mathbb{E} [V_n(y - d, C') | C = c]$$

$$= \lim_{n \to \infty} g_n(y, c),$$

where (25) follows from the homogeneity of the Markov process representing the channel condition and the Monotone Convergence Theorem. Furthermore, for each $c \in \mathcal{C}$, $g_\infty(y, c)$ is convex in $y$ as it is the sum of an affine function of $y$, a convex function of $y - d$, and a nonnegative weighted integral of the convex functions $V_\infty(y - d, c')$. Additionally, $\lim_{y \to \infty} g_\infty(y, c) \geq \lim_{y \to \infty} c \cdot y = \infty$. Thus, for every $c$, at least one finite number achieves the global minimum of $g_\infty(y, c)$. Define:

$$B_\infty(c) := \left\{ \hat{y} \in [d, \infty) : g_\infty(\hat{y}, c) = \min_{y \in [d, \infty)} \{g_\infty(y, c)\} \right\}$$

and $b_\infty(c) := \min B_\infty(c)$.

We want to show $\hat{b}_\infty(c) := \lim_{n \to \infty} b_n(c) = b_\infty(c)$, $\forall c \in \mathcal{C}$. Let $c \in \mathcal{C}$ be arbitrary. Recall from (5) that the sequence $\{b_n(c)\}_{n=1,2,\ldots}$ is nondecreasing. Thus, $\hat{b}_\infty(c) = \lim_{n \to \infty} b_n(c) = \sup_{n \in \mathbb{N}} b_n(c)$. First, assume $\hat{b}_\infty(c) > b_\infty(c)$. Then we can choose $\hat{N} \in \mathbb{N}$ such that $\hat{b}_\infty(c) \geq b_{\hat{N}}(c) > b_\infty(c)$, which implies:

$$0 < g_{\hat{N}}(b_\infty(c), c) - g_{\hat{N}}(b_{\hat{N}}(c), c).$$
By repeated application of the submodularity condition (17), (26) becomes:

\[
0 < g_N(b_\infty(c), c) - g_N(b_N(c), c) \\
\leq g_{N+1}(b_\infty(c), c) - g_{N+1}(b_N(c), c) \\
\leq g_{N+2}(b_\infty(c), c) - g_{N+2}(b_N(c), c) \\
\leq \ldots 
\]

which implies:

\[
0 < g_N(b_\infty(c), c) - g_N(b_N(c), c) \leq \lim_{m \to \infty} \left[ g_m(b_\infty(c), c) - g_m(b_N(c), c) \right] \\
= g_\infty(b_\infty(c), c) - g_\infty(b_N(c), c),
\]  
(27)

where the equality in (27) follows from \( \lim_{m \to \infty} g_m(y, c) = g_\infty(y, c) \). Yet, \( 0 < g_\infty(b_\infty(c), c) - g_\infty(b_N(c), c) \) is a contradiction, as \( b_\infty(c) \) is a global minimizer of \( g_\infty(\cdot, c) \). We conclude \( \hat{b}_\infty(c) \leq b_\infty(c) \). Next, assume \( \hat{b}_\infty(c) < b_\infty(c) \), which implies:

\[
g_\infty(\hat{b}_\infty(c), c) - g_\infty(b_\infty(c), c) > 0.
\]  
(28)

Then for all \( n \in \mathbb{N} \), we have \( b_n(c) \leq \hat{b}_\infty(c) < b_\infty(c) \), which, in combination with \( g_n(y, c) \) convex in \( y \) and minimized by \( b_n(c) \), implies:

\[
g_n(\hat{b}_\infty(c), c) \leq g_n(b_\infty(c), c).
\]  
(29)

Furthermore, by the monotone convergence of \( g_n(y, c) \) to \( g_\infty(y, c) \) in part (b) of this proof, we have:

\[
g_n(b_\infty(c), c) \leq g_\infty(b_\infty(c), c).
\]  
(30)

Equations (28), (29) and (30) imply:

\[
0 < g_\infty(\hat{b}_\infty(c), c) - g_\infty(b_\infty(c), c) \leq g_\infty(\hat{b}_\infty(c), c) - g_n(\hat{b}_\infty(c), c), \ \forall n \in \mathbb{N},
\]  
(31)
which contradicts \( \lim_{n \to \infty} g_n(y, c) = g_\infty(y, c) \) for all \( y \). Thus, we conclude \( \hat{b}_\infty(c) = b_\infty(c) \).

We are now ready to prove parts (e) and (f) of Theorem 2. Define

\[
y^*_\infty(x, c) := \begin{cases} 
x, & \text{if } x \geq b_\infty(c) \\
b_\infty(c), & \text{if } b_\infty(c) - \frac{P}{c} \leq x < b_\infty(c) \\
x + \frac{P}{c}, & \text{if } x < b_\infty(c) - \frac{P}{c}
\end{cases}
\]

Clearly, \( \lim_{n \to \infty} b_n(c) = b_\infty(c) \) implies \( \lim_{n \to \infty} y^*_n(x, c) = y^*_\infty(x, c) \), \( \forall x \in \mathbb{R}^+ \), \( \forall c \in \mathcal{C} \). Furthermore, \( g_n(y, c) \to g_\infty(y, c) \) and \( y^*_n(x, c) \to y^*_\infty(x, c) \) as \( n \to \infty \) imply \( \lim_{n \to \infty} g_n(y^*_n(x, c)) = g_\infty(y^*_\infty(x, c)) \), \( \forall x \in \mathbb{R}^+ \), \( \forall c \in \mathcal{C} \). So for all \( x \in \mathbb{R}^+ \) and \( c \in \mathcal{C} \), we have:

\[
V_\infty(x, c) = \lim_{n \to \infty} V_n(x, c) \\
= -c \cdot x + \lim_{n \to \infty} \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \{ g_n(y, c) \} \\
= -c \cdot x + \lim_{n \to \infty} g_n(y^*_n(x, c), c) \\
= -c \cdot x + g_\infty(y^*_\infty(x, c), c) \\
= -c \cdot x + \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \{ g_\infty(y, c) \} \\
= -c \cdot x + \min_{\max(x, d) \leq y \leq x + \frac{P}{c}} \left\{ c \cdot y + h(y - d) + \alpha \cdot \mathbb{E} [V_\infty(y - d, C') | C = c] \right\}.
\]

Equation (32) follows from Theorem 1, and (33) follows from Karush’s result, Lemma 1, presented in Section VIII-A. Thus, \( V_\infty(\cdot, \cdot) \), the limit of the finite horizon value functions and the infinite horizon discounted expected cost-to-go resulting from the stationary policy \( \pi_\infty^* := (y^*_\infty, y^*_\infty, \ldots) \), satisfies the infinite horizon functional equation (2). We conclude \( \pi_\infty^* \), the natural extension of the finite horizon optimal policy, is optimal for the infinite horizon problem (see, for example, [6, Propositions 9.12 and 9.16]).

REFERENCES

[1] B. Girod, M. Kalman, Y. J. Liang, and R. Zhang, “Advances in channel-adaptive video streaming,” in Proceedings of the International Conference on Image Processing, vol. 1, (Rochester, NY), pp. 9–12, September 2002.
[2] Y. Li and N. Bambos, “Power-controlled wireless links for media streaming applications,” in Proceedings of the Wireless Telecommunications Symposium, (Pomona, CA), pp. 102–111, May 2004.

[3] Y. Li, A. Markopoulou, N. Bambos, and J. Apostolopoulos, “Joint power-playout control for media streaming over wireless links,” IEEE Transactions on Multimedia, vol. 8, pp. 830–843, August 2006.

[4] C. E. Luna, Y. Eisenberg, R. Berry, T. N. Pappas, and A. K. Katsaggelos, “Joint source coding and data rate adaptation for energy efficient wireless video streaming,” IEEE Journal on Selected Areas in Communications, vol. 21, pp. 1710–1720, December 2003.

[5] A. Fu, E. Modiano, and J. N. Tsitsiklis, “Optimal transmission scheduling over a fading channel with energy and deadline constraints,” IEEE Transactions on Wireless Communications, vol. 5, pp. 630–641, March 2006.

[6] D. Bertsekas and S. E. Shreve, Stochastic Optimal Control: The Discrete-Time Case. Athena Scientific, 1996.

[7] T. Fabian, J. L. Fisher, M. W. Sasieli, and A. Yardeni, “Purchasing raw material on a fluctuating market,” Operations Research, vol. 7, pp. 107–122, January-February 1959.

[8] B. Kalymon, “Stochastic prices in a single-item inventory purchasing model,” Operations Research, vol. 19, pp. 1434–1458, October 1971.

[9] B. G. Kingsman, “Commodity purchasing,” Operational Research Quarterly, vol. 20, pp. 59–80, 1969.

[10] B. G. Kingsman, Commodity Purchasing in Uncertain Fluctuating Price Markets. PhD thesis, University of Lancaster, 1969.

[11] V. F. Magirou, “Stockpiling under price uncertainty and storage capacity constraints,” European Journal of Operational Research, vol. 11, pp. 233–246, 1982.

[12] V. F. Magirou, “Comments on ‘On Optimal Inventory Policies when Ordering Prices are Random’ by Kamal Golabi,” Operations Research, vol. 35, pp. 930–931, November-December.

[13] K. Golabi, “A single-item inventory model with stochastic prices,” in Proceedings of the Second International Symposium on Inventories, (Budapest, Hungary), pp. 687–697, 1982.

[14] K. Golabi, “Optimal inventory policies when ordering prices are random,” Operations Research, vol. 33, pp. 575–588, May-June 1985.

[15] A. Federgruen and P. Zipkin, “An inventory model with limited production capacity and uncertain demands II. The discounted-cost criterion,” Mathematics of Operations Research, vol. 11, pp. 208–215, May 1986.

[16] R. Evans, “Inventory control of a multiproduct system with a limited production resource,” Naval Research Logistics Quarterly, vol. 14, no. 2, pp. 173–184, 1967.

[17] G. A. DeCroix and A. Arreola-Risa, “Optimal production and inventory policy for multiple products under resource constraints,” Management Science, vol. 44, pp. 950–961, July 1998.

[18] R. Bellman, I. Glicksberg, and O. Gross, “On the optimal inventory equation,” Management Science, vol. 2, pp. 83–104, October 1955.

[19] D. L. Iglehart, “Optimality of (s,S) policies in the infinite horizon dynamic inventory problem,” Management Science, vol. 9, pp. 259–267, January 1963.

[20] D. Shuman and M. Liu, “Energy-efficient transmission scheduling with strict underflow constraints,” EECS Technical Report CGR 08-09, University of Michigan, Ann Arbor, 2008.
[21] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[22] D. Shuman and M. Liu, “Energy-efficient transmission scheduling for wireless media streaming with strict underflow constraints,” in *Proceedings of the International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, (Berlin, Germany), pp. 354–359, April 2008.

[23] W. Karush, “A theorem in convex programming,” *Naval Research Logistics Quarterly*, vol. 6, pp. 245–260, September 1959.

[24] E. L. Porteus, *Foundations of Stochastic Inventory Theory*. Stanford University Press, 2002.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[26] D. M. Topkis, *Supermodularity and Complementarity*. Princeton University Press, 1998.