

# Energy-Efficient Transmission Scheduling With Strict Underflow Constraints

David I Shuman, *Member, IEEE*, Mingyan Liu, *Member, IEEE*, and Owen Q. Wu

**Abstract**—This paper considers a single source transmitting data to one or more receivers/users over a shared wireless channel. Due to random fading, the wireless channel conditions vary with time and from user to user. Each 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 to each user. The source’s objective is to dynamically allocate power in a manner that minimizes total power consumption and packet holding costs, while satisfying strict buffer underflow constraints and a joint power constraint in each slot. The primary application motivating this problem is wireless media streaming. For this application, the buffer underflow constraints prevent the user buffers from emptying, so as to maintain playout quality. In the case of a single user, a state-dependent modified base-stock policy is shown to be optimal with linear power-rate curves, and a state-dependent finite generalized base-stock policy is shown to be optimal with piecewise-linear convex power-rate curves. When certain technical conditions are satisfied, efficient methods to compute the critical numbers that complete the characterizations of the optimal control laws in each of these cases are presented. The structure of the optimal policy for the case of two users is then analyzed.

**Index Terms**—Base-stock policy, dynamic programming, energy-delay tradeoff, opportunistic scheduling, resource allocation, stochastic inventory theory, underflow constraints, wireless media streaming.

## I. INTRODUCTION

**I**N this paper, we examine the problem of energy-efficient transmission scheduling over a wireless channel, subject to strict underflow constraints. We consider a single source wirelessly transmitting data to one or more receivers/users. Each user has a buffer to store received packets before they are drained at a known rate. Due to random fading, the condition of the wireless channel, which determines how much power

is required to reliably transmit a given amount of data, varies with time and from user to user. The transmitter’s goal is to minimize total power consumption, while preventing any user’s buffer from emptying (we refer to the latter as *strict underflow constraints*).

### A. Opportunistic Scheduling and Related Work

This problem falls into the general class of opportunistic scheduling problems, where the common theme is to exploit the temporal and spatial variation of the channel.<sup>1</sup> At a high level, the idea of exploiting the *temporal diversity* of the channel via opportunistic scheduling can be explained as follows. Consider the case of a single sender transmitting to a single receiver with different linear power-rate curves for each possible channel condition. 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 tends 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.

Similarly, in the case of multiple receivers, the scheduler can exploit the *spatial diversity* of the channel by transmitting only to those receivers who have the best channel conditions in each time slot. The benefit of increasing system throughput and reducing total power consumption through such a joint resource allocation policy is commonly referred to as the *multiuser diversity gain* [2]. It was introduced in the context of the analogous *uplink* problem, where multiple sources transmit to a single destination (e.g., the base station) [3]. Since, there has been a wide range of literature on opportunistic scheduling problems in wireless networks.

Sending more data when the channel is in a good state can increase system throughput and/or reduce total energy consumption; however, in opportunistic scheduling problems, it is often the case that the transmission scheduler has competing quality of service (QoS) interests. For instance, one QoS interest commonly considered is *fairness*. If, when a single source is transmitting to multiple receivers, the scheduler only considers total throughput and energy consumption across all users, it may end up transmitting to a single user or the same small group of users

<sup>1</sup>Opportunistic scheduling problems are also referred to as multi-user variable channel scheduling problems [1].

Manuscript received December 24, 2008; revised February 15, 2010; accepted August 04, 2010. Date of current version February 18, 2011. This work was supported in part by Grants NNX06AD47G and NNX09AE91G from the National Aeronautics and Space Administration (NASA) Earth Science Technology Office (ESTO), the Advanced Information Systems Technologies program, and the National Science Foundation (NSF) under Career Award ANI-0238035. This work was presented (in part) at the International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), Berlin, Germany, April 2008.

D. I Shuman is with the Institute of Electrical Engineering, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: david.shuman@epfl.ch).

M. Liu is with the Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI 49109 USA (e-mail: mingyan@eecs.umich.edu).

O. Q. Wu is with the Ross School of Business, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: owenwu@bus.umich.edu).

Communicated by A. Nosratinia, Associate Editor for Communication Networks.

Digital Object Identifier 10.1109/TIT.2011.2105002

in every slot. This can happen, for instance, if a base station requires less power to send data to a nearby receiver, even when the nearby receiver's channel is in its worst possible condition and a farther away receiver's channel is in its best possible condition. Thus, fairness constraints are often imposed to ensure that the transmitter sends packets to all receivers.

A number of different fairness conditions have been examined in the literature. For example, [4] and [5] consider *temporal fairness*, where the scheduler must transmit to each receiver for some minimum fraction of the time over the long run. Under the *proportional fairness* considered by [2] and [6], the scheduler considers the current channel conditions relative to the average channel condition of each receiver. Reference [5] considers a more general *utilitarian fairness*, where the focus is on system performance from the receiver's perspective, rather than on resources consumed by each user. The authors of [7] incorporate fairness directly into the objective function by setting relative throughput target values for each receiver and maximizing the minimum relative long-run average throughput.

Another QoS consideration that is important in many applications is delay. Different notions of delay have been incorporated into opportunistic scheduling problems. One proxy for delay is the *stability* of all of the sender's queues for arriving packets awaiting transmission. The motivation for this criterion is that if none of these queues blows up, then the delay is not "too bad." With stability as an objective, it is common to restrict attention to *throughput optimal* policies, which are scheduling policies that ensure the sender's queues are stable, as long as this is possible for the given arrival process and channel model. References [8]–[11] present such throughput optimal scheduling algorithms, and examine conditions guaranteeing stabilizability in different settings.

When an arriving packet model is used for the data, one can also define *end-to-end delay* as the time between a packet's arrival at the sender's buffer and its decoding by the receiver. A number of opportunistic scheduling studies have considered the *average end-to-end delay* of all packets over a long horizon. For instance, [12]–[23] all consider average delay, either as a constraint or by incorporating it directly into the objective function to be minimized. However, the average delay criterion allows for the possibility of long delays (albeit with small probability); thus, for many delay-sensitive applications, *strict end-to-end delay* is often a more appropriate consideration for studies with arriving packet models. In [24] and [25], Chen *et al.* place strict constraints on the end-to-end delay of each packet in a point-to-point system, examine the optimal scheduling policy assuming all future channel conditions are known, and suggest heuristics based on this optimal *offline* scheduling policy for the more realistic *online* case where the scheduler only learns the channel conditions in a causal fashion. Rajan, Sabharwal, and Aazhang also consider strict constraints on the end-to-end delay in an arriving packet model in [16, Section IV].

A strict constraint on the end-to-end delay of each packet is one particular form of a *deadline* constraint, as each packet has a deadline by which it must be transmitted. This notion can be generalized to impose individual deadlines on each packet, whether the packets are arriving over time or are all in the sender's buffer from the beginning. References [26]–[31]

consider point-to-point communication when a fixed amount of data is in the sender's buffer at the start of the time horizon and the individual deadlines coincide, so that all packets must be transmitted and received by a common deadline, the end of the time horizon under consideration. In [26, Section III-D] and [27, Section III-D], Fu *et al.* specify the optimal transmission policy when the power-rate curves under each channel condition are linear and the transmitter is subject to a per slot peak power constraint. In [28]–[31], Lee and Jindal model the power-rate curve under each channel condition as convex, first of the form of the so-called Shannon cost function based on the capacity of the additive white Gaussian noise channel, and then as a convex monomial function.<sup>2</sup>

References [32] and [33] consider opportunistic scheduling problems with multiple receivers and a single deadline constraint at the end of a finite horizon. Packets arrive over time and the emphasis is on offline scheduling policies in [32], whereas [33] considers a fixed amount of data destined for each receiver, and assumes the data is already in the sender's buffers at the beginning of the horizon. The model of [33] is perhaps the closest to our general model for  $M$  receivers; however, two key differences are (i) the transmitter is not subject to a power constraint in [33] and (ii) the transmitter can transmit to at most one receiver in each time slot in [33].

In our model, the strict underflow constraints serve as a notion of both fairness and delay. The notion of fairness is that none of the receivers' buffers are allowed to empty, guaranteeing the required level of service to all users. The underflow constraints also serve as a notion of delay, and can be seen as multiple deadline constraints—certain packets must arrive by the end of the first slot, another group by the end of the second slot, and so forth. Therefore, Sections III and IV of this paper generalize the works of [26], [27] and [28]–[31], respectively, by considering multiple deadlines in the point-to-point communication problem, rather than a single deadline at the end of the horizon. In addition to better representing some delay-sensitive applications, this extension of the model also allows us to consider infinite horizon problems. We compare related work in opportunistic scheduling problems with deadline constraints further in [34]. For more complete surveys of opportunistic scheduling studies in wireless networks, see [35] and [36].

## B. Wireless Media Streaming and Related Work

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 receivers' buffers in order to be decoded and played. Enforcing the underflow constraints reduces playout interruptions to the end users. In order to make the presentation concrete, we use the above wireless media streaming terminology throughout the paper.

Transporting multimedia over wireless networks is a promising application that has seen recent advances [37]. At the same time, certain resource allocation issues need to

<sup>2</sup>In our notation of Section II-A, these two cases correspond to power-rate curves of the form  $c(z, s) = \frac{2^z - 1}{g_1(s)}$  and  $c(z, s) = \frac{-z^\zeta}{g_2(s)}$ , respectively, where  $c(z, s)$  is the power required to transmit  $z$  bits under channel condition  $s$ ,  $g_1(\cdot)$  and  $g_2(\cdot)$  are known functions, and  $\zeta$  is a fixed parameter.

be addressed in order to provide high quality and efficient media over wireless. First, streaming applications tend to have stringent QoS requirements (e.g., they can be delay and jitter intolerant). Second, because streaming is bandwidth-demanding, it is especially 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: (i) limit potential interference to other base stations and their associated mobiles; (ii) maximize the number of receivers the sender can support; and (iii) reduce the energy costs incurred by operators.

Of the related work in wireless media streaming, [38] has the closest setup to our model. The main differences are that [38] 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 [39], 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 speeds. Our model also bears resemblance to [40]. The first difference here is that [40] 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 [40] must assign various source coding parameters such as quantization step size and coding mode, whereas our model assumes a fixed encoding/decoding scheme.

### C. Summary of Contribution

We formulate the task of energy-efficient transmission scheduling subject to strict underflow constraints as three different Markov decision problems (MDPs), with the finite horizon discounted expected cost, infinite horizon discounted expected cost, and infinite horizon average expected cost criteria, respectively. These three MDPs feature a continuous component of the state space and a continuous action space at each state. Therefore, unlike *finite* MDPs, they cannot in general be solved exactly via dynamic programming, and suffer from the well-known *curse of dimensionality* [41], [42]. Our aim is to analyze the dynamic programming equations in order to (i) determine if there are circumstances under which we can analytically derive optimal solutions to the three problems; and (ii) leverage our mathematical analysis and results on the structures of the optimal scheduling policies to improve our intuitive understanding of the problems.

We begin by showing that in the case of a single receiver under linear power-rate curves, the optimal policy is an easily-implementable *modified base-stock* policy. In each time slot, it is optimal for the sender to transmit so as to bring the number of packets in the receiver's buffer level after transmission as close as possible to a target level or critical number.<sup>3</sup> The target level depends on the current channel condition, with a better channel condition corresponding to a higher target level. We

also show that the strict underflow constraints may cause the scheduler to be less opportunistic than it otherwise would be, and transmit more packets under "medium" channel conditions in anticipation of deadline constraints in future time slots.

We then generalize this result in two different directions. First, we relax the assumption that the power-rate curves under each channel condition are linear, and model them as piecewise-linear convex to better approximate more realistic convex power-rate curves. Under piecewise-linear power-rate curves, we show the optimal policy is a *finite generalized base-stock* policy, and provide an intuitive explanation of this structure in terms of multiple target levels in each time slot. In addition to the structural results on the optimal policy for the case of a single receiver under either linear or piecewise-linear convex power-rate curves, we provide an efficient method to calculate the critical numbers that complete the characterization of the optimal policy when certain technical conditions are satisfied.

The second generalization of the single receiver model under linear power-rate curves is to a single user transmitting to two receivers over a shared wireless channel. In this case, we state and prove the structure of the optimal policy, and show how the peak power constraint in each slot couples the optimal scheduling of the two receivers' packet streams.

In all three setups, the structure of the optimal policy in the finite horizon discounted expected cost problem extends to the infinite horizon discounted and average expected cost problems.

Throughout the analysis, we make a novel connection with inventory models that may prove useful in other wireless transmission scheduling problems. Because the inventory models corresponding to our wireless communication models have not been previously examined, our results also represent a contribution to the inventory control literature.

The remainder of this paper is organized as follows. In the next section, we describe the system model, formulate finite and infinite horizon MDPs, and relate our model to inventory control models. In Section III, we consider the case of a single receiver under linear power-rate curves. While this case can be considered a special case of the models of Sections IV and V, we present it first in order to: (i) state additional structural properties of the optimal transmission policy to a single user under linear power-rate curves that are not true in general for the cases discussed in Sections IV and V; (ii) highlight some intuitive takeaways that carry over to the generalized models, but are more transparent in the simpler model; and (iii) compare it to related problems in the wireless communications literature. We analyze the structure of the optimal scheduling policy for the finite horizon problem, provide a method to compute the critical numbers that complete the characterization of the optimal policy when some additional technical conditions are met, and provide sufficient conditions for this problem to be equivalent to a previously-studied single deadline problem. Section IV generalizes the analysis of Section III to the case of a single receiver under piecewise-linear convex power-rate curves. In Section V, we analyze the structure of the optimal policy when there are two receivers with linear power-rate curves. We discuss the infinite horizon problems, relaxation of the strict underflow constraints, and the extension to the general case of  $M$  receivers in Section VI. Section VII concludes the paper.

<sup>3</sup>We use the terms target level and critical number interchangeably throughout the paper.

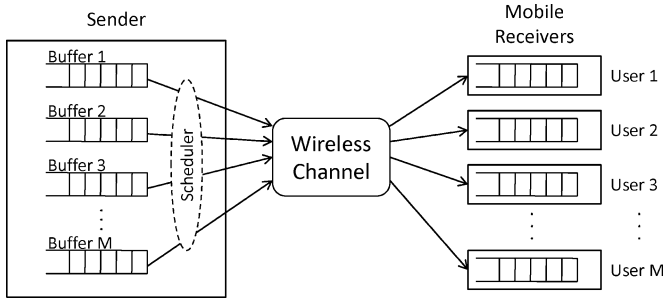


Fig. 1. System model.

## II. PROBLEM DESCRIPTION

In this section, we present an abstraction of the transmission scheduling problem outlined in the previous section and formulate three optimization problems. While most of this paper focuses on the cases of one and two users, the formulation in this section is for the general multi-user (multi-receiver) case, so that we can discuss this most general case in Section VI-C.

### A. System Model and Assumptions

We consider a single source transmitting media sequences to  $M$  users/receivers over a shared wireless channel. The sender maintains a separate buffer for each receiver, and is assumed to always have data to transmit to each receiver.<sup>4</sup> We consider a *fluid* packet model that allows packets to be split, with the receiver reassembling fractional packets. Each 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 allows holding costs to be assessed on packets in the receiver buffers. See Fig. 1 for a diagram of the system.

We consider time evolution in discrete steps, indexed backwards by  $n = N, N-1, \dots, 1$ , 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  $[n, n-1]$ .

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 conditions within a single time slot are constant. User  $m$ 's channel condition in slot  $n$  is modeled as a random variable,  $S_n^m$ . We assume that the evolution of a given user's channel condition is independent of all other users' channel conditions and the transmitter's scheduling decisions.

We begin by modeling the evolution of each user's channel condition as a finite-state ergodic homogeneous Markov process,  $\{S_n^m\}_{n=N, N-1, \dots, 1}$  with state space  $\mathcal{S}^m$ .<sup>5</sup> Conditioned on the channel state,  $S_n^m$ , at time  $n$ , user  $m$ 's channel states at future times ( $n-1, n-2, \dots$ ) are independent of the channel states at past times ( $n+1, n+2, \dots$ ). 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

<sup>4</sup>This assumption is commonly referred to as the infinite backlog assumption.

<sup>5</sup>Theorems 1, 3, 5, 6, and their proofs remain valid as stated when each user's channel condition is given by a more general homogeneous Markov process that is not necessarily finite-state and ergodic.

analysis of the infinite horizon problems. We also consider the case that each user's channel condition is independent and identically distributed (IID) over time. In this case, we can say more about the optimal transmission policy, as will be seen in Sections III-B and IV-B.

Associated with each channel condition for a given user is a power-rate function. If user  $m$ 's channel is in condition  $s^m$ , then the transmission of  $r$  units of data to user  $m$  incurs a power consumption cost of  $c^m(r, s^m)$ . This power-rate function  $c^m(\cdot, s^m)$  is commonly assumed to be linear (in the low SNR regime) or convex (in the high SNR regime). In this paper, we consider power-rate functions that are linear or piecewise-linear convex, the latter of which can be used to approximate more general convex power-rate functions. We assume that sending data consumes a strictly positive amount of power, and therefore take the power-rate functions to be strictly increasing under all channel conditions.

At the beginning of each time slot, the scheduler learns all the channel states through a feedback channel. It then allocates some amount of power (possibly zero) for transmission to each user, which is equivalent to deciding how many packets to send to each user. The total power consumed in any one slot must not exceed the fixed power constraint,  $P$ . Following transmission and reception in each slot, a certain number of packets are removed/purged from each receiver buffer for playing. The transmitter (or scheduler) knows precisely the packet requirements of each receiver (i.e., the number of packets removed from the buffer) in each time slot. This is justified by the fact that the transmitter knows the encoding and decoding schemes used. We assume that packets transmitted in slot  $n$  arrive in time to be used for playing in slot  $n$ , and that the users' consumption of packets in each slot is constant, denoted by  $\mathbf{d} = (d^1, d^2, \dots, d^M)$ . This latter assumption is less realistic, but may be justified if the receiving buffers are drained at a constant rate at the media access control (MAC) layer, before packets are decoded at the application layer by the media player. It is also worth noting that the same techniques we use to analyze the constant drainage rate case can be used to examine the case of time-varying drainage rates, which we discuss further in Section III-A. We assume the receiver buffers are empty at the beginning of the time horizon, and that even when the channels are in their worst possible condition, the maximum power constraint  $P$  is sufficient to transmit enough packets to satisfy one time slot's packet requirements for every user. We discuss the relaxation of this assumption in Section VI-B.

The goal of this study is to characterize the control laws that minimize the transmission power and packet holding costs over a finite or infinite time horizon, subject to strict underflow constraints and a maximum power constraint in each time slot.

### B. Notation

Before proceeding, we introduce some notation. We define  $\mathbb{R}_+ := [0, \infty)$  and  $\mathbb{N} := \{1, 2, \dots\}$ . A single dot, as in  $a \cdot b$ , represents scalar multiplication. We use bold font to denote column vectors, such as  $\mathbf{w} = (w^1, w^2, \dots, w^M)$ . We include a transpose superscript whenever a vector is meant to be a row vector, such as  $\mathbf{w}^T$ . The notations  $\mathbf{w} \preceq \tilde{\mathbf{w}}$  and  $\mathbf{w} \succeq \tilde{\mathbf{w}}$  denote component-wise inequalities; i.e.,  $w^m \leq$  (respectively,  $\geq$ )  $\tilde{w}^m$ ,  $\forall m$ .

Finally, we use the standard definitions of the meet and join of two vectors:

$$\begin{aligned} \mathbf{w} \wedge \tilde{\mathbf{w}} &:= (\min\{w^1, \tilde{w}^1\}, \min\{w^2, \tilde{w}^2\}, \dots, \\ &\quad \min\{w^M, \tilde{w}^M\}), \\ \text{and } \mathbf{w} \vee \tilde{\mathbf{w}} &:= (\max\{w^1, \tilde{w}^1\}, \max\{w^2, \tilde{w}^2\}, \dots, \\ &\quad \max\{w^M, \tilde{w}^M\}). \end{aligned}$$

### C. Problem Formulation

We consider three MDPs. Problem **(P1)** is the finite horizon discounted expected cost problem; Problem **(P2)** is the infinite horizon discounted expected cost problem; and Problem **(P3)** is the infinite horizon average expected cost problem. The three 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  $(\mathbf{X}_n, \mathbf{S}_n)$ , where the random vector  $\mathbf{X}_n = (X_n^1, X_n^2, \dots, X_n^M)$  denotes the receiver buffer queue lengths at time  $n$ , and  $\mathbf{S}_n = (S_n^1, S_n^2, \dots, S_n^M)$  denotes the channel conditions in slot  $n$  (recall that  $n$  is the number of steps remaining until the end of the horizon). The dynamics for the receivers' queues are governed by the simple equation  $\mathbf{X}_{n-1} = \mathbf{X}_n + \mathbf{Z}_n - \mathbf{d}$  at all times  $n = N, N-1, \dots, 1$ , where  $\mathbf{Z}_n$  is a controlled random vector chosen by the scheduler at each time  $n$  that represents the number of packets transmitted to each user in the  $n$ th slot. At each time  $n$ ,  $\mathbf{Z}_n$  must be chosen to meet the peak power constraint

$$\sum_{m=1}^M c^m(Z_n^m, S_n^m) \leq P$$

and the underflow constraints

$$X_n^m + Z_n^m \geq d^m, \quad \forall m \in \{1, 2, \dots, M\}.$$

Clearly, the scheduler cannot transmit a negative number of packets to any user, so it must also be true that  $Z_n^m \geq 0$  for all  $m$ .

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 each user's playout buffer at the end of a time slot. The holding costs associated with user  $m$  in each slot are described by a convex, nonnegative, nondecreasing function,  $h^m(\cdot)$ , of the packets remaining in user  $m$ 's buffer following playout, with  $\lim_{x \rightarrow \infty} h^m(x) = \infty$ . We assume without loss of generality that  $h^m(0) = 0$ . Possible holding cost models include a linear model,  $h^m(x) = \hat{h}^m \cdot x$  for some positive constant  $\hat{h}^m$ , or a barrier-type function such as:

$$h^m(x) := \begin{cases} 0, & \text{if } x \leq \mu \\ \kappa \cdot (x - \mu), & \text{if } x > \mu \text{ } (\kappa \text{ very large}) \end{cases}$$

which could represent a finite receiver buffer of length  $\mu$ .<sup>6</sup>

<sup>6</sup>Taking  $\mu$  to be greater than the time horizon  $N$  in the finite horizon expected cost problem is equivalent to not assessing any holding costs in Problem **(P1)**.

In Problem **(P1)**, we wish to find a transmission policy  $\pi$  that minimizes  $J_{N,\alpha}^\pi$ , the finite horizon discounted expected cost under policy  $\pi$ , defined as

$$J_{N,\alpha}^\pi := \mathbb{E}^\pi \left\{ \sum_{n=1}^N \sum_{m=1}^M \alpha^{N-n} \cdot \{c^m(Z_n^m, S_n^m) + h^m(X_n^m + Z_n^m - d^m)\} | \mathcal{F}_N \right\}$$

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 discount factor must satisfy  $0 \leq \alpha < 1$ , and the infinite horizon discounted expected cost function for minimization is defined as

$$J_{\infty,\alpha}^\pi := \lim_{N \rightarrow \infty} J_{N,\alpha}^\pi.$$

For Problem **(P3)**, the average expected cost function for minimization is defined as

$$J_{\infty,1}^\pi := \limsup_{N \rightarrow \infty} \frac{1}{N} J_{N,1}^\pi.$$

In all three cases, we allow the transmission policy  $\pi$  to be chosen from the set of all history-dependent randomized and deterministic control laws,  $\Pi$  (see, e.g., [43, Definition 2.2.3, p. 15]).

Combining the constraints and criteria, we present the optimization formulations for Problem **(P1)** (or **(P2)** or **(P3)**)

$$\begin{aligned} &\inf_{\pi \in \Pi} J_{N,\alpha}^\pi \left( \text{or } \inf_{\pi \in \Pi} J_{\infty,\alpha}^\pi \text{ or } \inf_{\pi \in \Pi} J_{\infty,1}^\pi \right) \\ \text{s.t. } &\sum_{m=1}^M c^m(Z_n^m, S_n^m) \leq P, \text{ w.p. } 1, \forall n \\ \text{and } &Z_n^m \geq \max\{0, d^m - X_n^m\}, \text{ w.p. } 1, \\ &\forall n, \forall m \in \{1, 2, \dots, M\}. \end{aligned}$$

Problem **(P1)** may be solved using standard dynamic programming (see, e.g., [43], [44]). The recursive dynamic programming equations are given by<sup>7</sup>

$$\begin{aligned} V_n(\mathbf{x}, \mathbf{s}) &= \min_{\mathbf{z} \in \mathcal{A}^d(\mathbf{x}, \mathbf{s})} \left\{ \sum_{m=1}^M \{c^m(z^m, s^m) + h^m(x^m + z^m - d^m)\} \right. \\ &\quad \left. + \alpha \cdot \mathbb{E}[V_{n-1}(\mathbf{x} + \mathbf{z} - \mathbf{d}, \mathbf{S}_{n-1}) | \mathbf{S}_n = \mathbf{s}] \right\}, \\ &\quad n = N, N-1, \dots, 1 \\ V_0(\mathbf{x}, \mathbf{s}) &= 0, \forall \mathbf{x} \in \mathbb{R}_+^M, \forall \mathbf{s} \in \mathcal{S} := \mathcal{S}^1 \times \mathcal{S}^2 \times \dots \times \mathcal{S}^M \quad (1) \end{aligned}$$

<sup>7</sup>As shown in Appendix C of [45] and [46], our model satisfies the measurable selection condition 3.3.3 of [43, p. 28], justifying the use of min rather than inf in the dynamic programming equations.

where  $V(\cdot, \cdot)$  is the value function or expected cost-to-go, and the action space is defined as

$$\mathcal{A}^d(\mathbf{x}, \mathbf{s}) := \left\{ \mathbf{z} \in \mathbb{R}_+^M : \begin{array}{l} \mathbf{z} \succeq \max\{\mathbf{0}, \mathbf{d} - \mathbf{x}\} \text{ and} \\ \sum_{m=1}^M c^m(z^m, s^m) \leq P \\ \forall \mathbf{x} \in \mathbb{R}_+^M, \forall \mathbf{s} \in \mathcal{S} \end{array} \right\}, \quad (2)$$

The maximum in (2) is taken element-by-element (i.e.,  $z^m \geq \max\{0, d^m - z^m\}, \forall m \in \{1, 2, \dots, M\}$ ). Note that our assumption that the maximum power constraint  $P$  is always sufficient to transmit enough packets to satisfy one time slot's packet requirements for every user (i.e.,  $\sum_{m=1}^M c^m(d^m, s^m) \leq P, \forall \mathbf{s} \in \mathcal{S}$ ) ensures that the action space  $\mathcal{A}^d(\mathbf{x}, \mathbf{s})$  is always nonempty.

#### D. Relation to 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, multi-item, discrete-time inventory model with random (linear or piecewise-linear convex) ordering costs, a budget constraint, and deterministic demands. The items correspond to the streams of data packets, the random ordering costs 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 payout.

This particular problem has not been studied in the context of inventory management, but similar problems have been examined, and some of the techniques from the inventory control literature are useful in analyzing our model. References [47]–[54] all consider single-item inventory models with linear ordering costs and random prices. The key result for the case 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 works, Kingsman [49], [50] 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 single-item discrete-time inventory models with deterministic linear ordering costs and stochastic demand, [55] and [56] are the most relevant; in those studies, 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 [57]–[59] consider single-item inventory models with deterministic piecewise-linear convex ordering costs and stochastic demand. The key result in this setup is that the optimal inventory level after ordering is a piecewise-linear nondecreasing function of the current inventory level (i.e.,

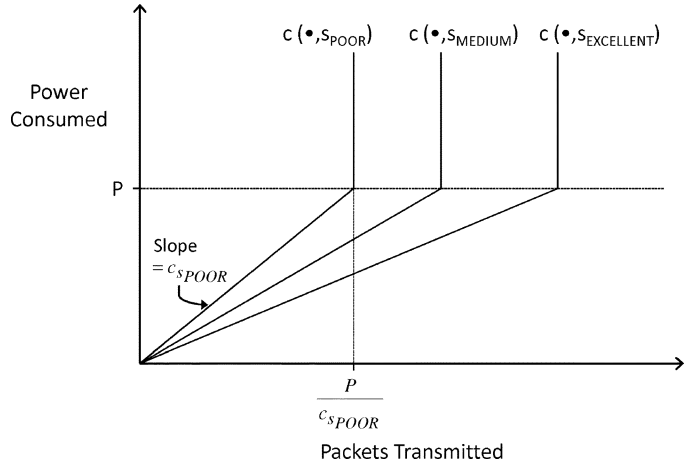


Fig. 2. Family of linear power-rate functions. Due to the power constraint, the effective power-rate function, shown above for each of the three channel conditions, is a two-segment piecewise-linear convex function. When the channel condition is  $s$ , the slope of the first segment is  $c_s$ .

there are a finite number of target stock levels), and the optimal ordering quantity is a piecewise-linear nonincreasing function of the current inventory level. Porteus [60] refers to policies of this form as *finite generalized base-stock policies*, to distinguish them from the superclass of *generalized base-stock policies*, which are optimal when the deterministic ordering costs are convex (but not necessarily piecewise-linear), as first studied in [61]. Under a generalized base-stock policy, the optimal inventory level after ordering is a nondecreasing function of the current inventory level, and the optimal ordering quantity is a nonincreasing function of the current inventory level. References [62]–[65] consider multi-item inventory systems under deterministic ordering costs, stochastic demand, and joint resource constraints.

We are not aware of any prior work on: (i) single-item inventory models with random piecewise-linear convex ordering costs; (ii) exact computation of the critical numbers in any sort of finite generalized base-stock policy; or (iii) multi-item inventory models with random ordering costs and joint resource constraints. Therefore, not only is this connection between wireless transmission scheduling problems and inventory models novel, but the results we present in this paper also represent a contribution to inventory theory.

### III. SINGLE RECEIVER WITH LINEAR POWER-RATE CURVES

In this section, we analyze the finite horizon discounted expected cost problem when there is only a single receiver ( $M = 1$ ), and the power-rate functions under different channel conditions are linear. One such family of power-rate functions is shown in Fig. 2, where there are three possible channel conditions, and a different linear power-rate function associated with each channel condition. Note that due to the power constraint  $P$  in each slot, the effective power-rate function is a two-segment piecewise-linear convex function under all channel conditions. We subsequently simplify our notation and use  $c_s$  to denote the power consumption per unit of data transmitted when the channel condition is in state  $s$ . Because there is just a single receiver, we also drop the dependence of the functions and random variables on  $m$ .

We denote the “best” and “worst” channel conditions by  $s_{\text{best}}$  and  $s_{\text{worst}}$ , respectively, and denote the slopes of the power-rate functions under these respective conditions by  $c_{\text{min}}$  and  $c_{\text{max}}$ . That is

$$\begin{aligned} 0 < c_{s_{\text{best}}} = c_{\text{min}} &:= \min_{s \in \mathcal{S}} \{c_s\} \leq \max_{s \in \mathcal{S}} \{c_s\} \\ &=: c_{\text{max}} = c_{s_{\text{worst}}} \leq \frac{P}{d}. \end{aligned}$$

With these notations in place, the dynamic program (1) for Problem (P1) becomes as shown in (3)–(4), found at the bottom of the page, where  $g_n(y, s) := c_s \cdot y + h(y-d) + \alpha \cdot \mathbb{E}[V_{n-1}(y-d, S_{n-1}) | S_n = s]$ . Here, the transition from (3) to (4) is done by a change of variable in the action space from  $Z_n$  to  $Y_n$ , where  $Y_n = X_n + Z_n$ . The controlled random variable  $Y_n$  represents the queue length of the receiver buffer *after* transmission takes place in the  $n$ th slot, but *before* ployout takes place (i.e., before  $d$  packets are removed from the buffer). The restrictions on the action space,  $\max(x, d) \leq y \leq x + \frac{P}{c_s}$ , ensure: (i) a nonnegative number of packets is transmitted; (ii) there are at least  $d$  packets in the receiver buffer following transmission, in order to satisfy the underflow constraint; and (iii) the power constraint is satisfied.

#### A. Structure of the Optimal Policy

With the above change of variable in the action space, the expected cost-to-go at time  $n$ ,  $V_n(x, s)$ , depends on the current buffer level,  $x$ , only through the fixed term  $-c_s \cdot x$  and the action space; i.e., the function  $g_n$  does not depend on  $x$ . This separation allows us to leverage the inventory management techniques of showing “single critical number” or “base-stock” policies, which date as far back as [66]. The following theorem gives the structure of the optimal transmission policy for the finite horizon discounted expected cost problem.

*Theorem 1:* For every  $n \in \{1, 2, \dots, N\}$  and  $s \in \mathcal{S}$ , define the critical number

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

Then, for Problem (P1) in the case of a single receiver with linear power-rate curves, the optimal buffer level after transmission with  $n$  slots remaining is given by

$$y_n^*(x, s) := \begin{cases} x, & \text{if } x \geq b_n(s) \\ b_n(s), & \text{if } b_n(s) - \frac{P}{c_s} \leq x < b_n(s) \\ x + \frac{P}{c_s}, & \text{if } x < b_n(s) - \frac{P}{c_s} \end{cases} \quad (5)$$

or, equivalently, the optimal number of packets to transmit in slot  $n$  is given by

$$z_n^*(x, s) := \begin{cases} 0, & \text{if } x \geq b_n(s) \\ b_n(s) - x, & \text{if } b_n(s) - \frac{P}{c_s} \leq x < b_n(s) \\ \frac{P}{c_s}, & \text{if } x < b_n(s) - \frac{P}{c_s} \end{cases}.$$

Furthermore, for a fixed  $s$ ,  $b_n(s)$  is nondecreasing in  $n$ :

$$N \cdot d \geq b_N(s) \geq b_{N-1}(s) \geq \dots \geq b_1(s) = d. \quad (6)$$

If, in addition, the channel condition is independent and identically distributed from slot to slot, then for a fixed  $n$ ,  $b_n(s)$  is nonincreasing in  $c_s$ ; i.e., for arbitrary  $s^1, s^2 \in \mathcal{S}$  with  $c_{s^1} \leq c_{s^2}$ , we have:

$$n \cdot d \geq b_n(s_{\text{best}}) \geq b_n(s^1) \geq b_n(s^2) \geq b_n(s_{\text{worst}}) = d. \quad (7)$$

The optimal transmission policy in Theorem 1 is a state-dependent modified base-stock policy. At time  $n$ , for each possible channel condition realization  $s$ , the critical number  $b_n(s)$  describes the target 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, it is optimal to not transmit any packets; if there are fewer than the target and the available power is enough to transmit the difference, it is optimal to do so; and if there are fewer than the target and the available power is not enough to transmit the difference, the sender should use the maximum power to transmit. See Fig. 3 for diagrams of the optimal policy.

Details of the proof of Theorem 1 are included in Appendix A of [45] and [46]. The key realization is that for all  $n$  and all  $s$ ,  $g_n(\cdot, s) : [d, \infty) \rightarrow \mathbb{R}_+$  is a convex function in  $y$ , with  $\lim_{y \rightarrow \infty} g_n(y, s) = \infty$ . Thus, for all  $n$  and all  $s$ ,  $g_n(\cdot, s)$  has a global minimum  $b_n(s)$ , the target number of packets to have in the buffer following transmission in the  $n$ th slot. To prove (6) we fix  $s \in \mathcal{S}$ , view  $g_n(y, s)$  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 (time-invariant) deterministic demand assumption to a nonstationary (time-varying) deterministic demand sequence,  $\{d_N, d_{N-1}, \dots, d_1\}$  (with  $d_n \leq \frac{P}{c_{\text{max}}}$  for all  $n$ ), then the structure of the optimal policy is still as stated in (5). If the channel is IID, then the following statement, analogous to (7), is true for arbitrary  $s^1, s^2 \in \mathcal{S}$  with  $c_{s^1} \leq c_{s^2}$ :

$$\sum_{i=1}^n d_i \geq b_n(s_{\text{best}}) \geq b_n(c_{s^1}) \geq b_n(c_{s^2}) \geq b_n(s_{\text{worst}}) = d_n, \quad \forall n \in \{1, 2, \dots, N\}. \quad (8)$$

$$V_n(x, s) = \min_{\max\{0, d-x\} \leq z \leq \frac{P}{c_s}} \left\{ c_s \cdot z + h(x+z-d) + \alpha \cdot \mathbb{E}[V_{n-1}(x+z-d, S_{n-1}) | S_n = s] \right\} \quad (3)$$

$$= \min_{\max\{x, d\} \leq y \leq x + \frac{P}{c_s}} \left\{ c_s \cdot (y-x) + h(y-d) + \alpha \cdot \mathbb{E}[V_{n-1}(y-d, S_{n-1}) | S_n = s] \right\} \quad (4)$$

$$= -c_s \cdot x + \min_{\max\{x, d\} \leq y \leq x + \frac{P}{c_s}} \{g_n(y, s)\}, \quad n = N, N-1, \dots, 1,$$

$$V_0(x, s) = 0, \quad \forall x \in \mathbb{R}_+, \quad \forall s \in \mathcal{S}$$

However, (6), the monotonicity of critical numbers over time for a fixed channel condition, is not true in general under nonstationary deterministic demand. As one counterexample, (8) says that under an IID channel, the critical numbers for the worst possible channel condition are equal to the single period demands. Therefore, if the demand sequence is not monotonic, the sequence of critical numbers,  $\{b_n(s_{\text{worst}})\}_{n=1,2,\dots,N}$ , is not monotonic.

### B. Computation of the Critical Numbers

In this section, we consider the special case where the channel condition is independent and identically distributed from slot to slot, the holding cost function is linear (i.e.,  $h(x) = h \cdot x$  for some  $h \geq 0$ ), and the following technical condition is satisfied: for each possible channel condition  $s$ ,  $\frac{P}{c_s} = l \cdot d$  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 three assumptions, we can completely characterize the optimal transmission policy.

*Theorem 2:* Define the threshold  $\gamma_{n,j}$  for  $n \in \{1, 2, \dots, N\}$  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  $2 \leq j \leq n$ , let  $\gamma_{n,j}$  be as shown in (9), found at the bottom of the page,

where  $p(s)$  is the probability of the channel being in state  $s$  in a time slot, and  $L(s) := \frac{P}{d \cdot c_s}$ . For each  $n \in \{1, 2, \dots, N\}$  and  $s \in \mathcal{S}$ , if  $\gamma_{n,j+1} \leq c_s < \gamma_{n,j}$ , define  $b_n(s) := j \cdot d$ . The optimal control strategy for Problem (P1) is then given by  $\pi^* = \{y_N^*, y_{N-1}^*, \dots, y_1^*\}$ , where  $y_n^*(x, s)$  is defined in (5).<sup>8</sup>

Compared to using standard numerical techniques to *approximately* solve the dynamic program and find a near-optimal policy, the above result not only sheds more insight on the structural properties of the problem and its *exactly*-optimal solution, but also offers a computationally simpler method. In particular, the optimal policy is completely characterized by the thresholds  $\{\gamma_{n,j}\}_{n \in \{1,2,\dots,N\}, j \in \mathbb{N}}$ . Calculating these thresholds recursively, as described in Theorem 2, requires  $O(N^2|\mathcal{S}|)$  operations, which is considerably simpler from a computational standpoint than approximately solving the dynamic program [41], [42].

To prove Theorem 2, we show by backwards induction that it is worse to transmit either fewer or more packets than the number suggested by the policy  $\pi^*$ . Theorem 2 is a special case of Theorem 2, the detailed proof of which can be found in [45]

<sup>8</sup>With  $n$  slots remaining,  $0 = \gamma_{n,n+1} \leq \gamma_{n,n} \leq \gamma_{n,n-1} \leq \dots \leq \gamma_{n,2} \leq \gamma_{n,1} = \infty$ , so  $b_n(s)$  is well defined.

and [46]. We discuss some intuition behind the proof and the recursion (9) in the Appendix of this paper.

### C. Sufficient Conditions for Equivalence With the Single Deadline Problem

In [27, Section III-D], Fu *et al.* consider the related single user 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. There is just a single explicit underflow constraint (the deadline) in their problem; 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  $d_{\text{total}}$  is the total amount of data that must be sent by the deadline and  $d_{\text{worst}}$  is the amount that can be sent in a slot under the worst channel condition, the transmitter must have sent at least  $d_{\text{total}} - d_{\text{worst}}$  packets by the beginning of the last slot, at least  $d_{\text{total}} - 2 \cdot d_{\text{worst}}$  packets by the beginning of the second to last slot, and so forth.<sup>9</sup> So there are in fact implicit constraints on how much data must be transmitted by the end of slots  $N - \lceil \frac{d_{\text{total}}}{d_{\text{worst}}} \rceil + 1$ ,  $N - \lceil \frac{d_{\text{total}}}{d_{\text{worst}}} \rceil + 2, \dots, N - 2, N - 1$ . With this interpretation, we believe that our Theorem 2 is equivalent to Theorem 3 and its corollary in [27] in the special case that, in addition to the hypotheses of our Theorem 2,  $\alpha = 1$ ,  $h = 0$ , and  $L(s_{\text{worst}}) = 1$ . For, when these conditions are met, the implicit constraints in [27] 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. For a more detailed comparison of these two problems, see [34].

### D. Intuitive Takeaways on the Role of the Strict Underflow Constraints

As mentioned earlier, the main idea of energy-efficient communication over a fading channel via opportunistic scheduling is to minimize power consumption by transmitting more data when the channel is in a “good” state, and less data when the channel is in a “bad” state. However, in order to comply with the underflow or deadline constraints, the transmitter may be forced to send data under poor channel conditions.

One intuitive takeaway from the analysis is that it is better to anticipate the need to comply with these constraints in future slots by sending more packets (than one would without the deadlines) under “medium” channel conditions in earlier slots.

<sup>9</sup>An unstated assumption in the formulation in [27, Section III-D] is that  $d_{\text{worst}}$  times the horizon length must be at least as large as  $d_{\text{total}}$ .

$$\gamma_{n,j} = -h + \alpha \cdot \left( \begin{aligned} & \sum_{s:c_s \geq \gamma_{n-1,j-1}} p(s) \cdot \gamma_{n-1,j-1} + \sum_{s:c_s < \gamma_{n-1,j-1}} p(s) \cdot c_s \\ & + \sum_{s:c_s < \gamma_{n-1,j-1+L(s)}} p(s) \cdot [\gamma_{n-1,j-1+L(s)} - c_s] \end{aligned} \right) \quad (9)$$



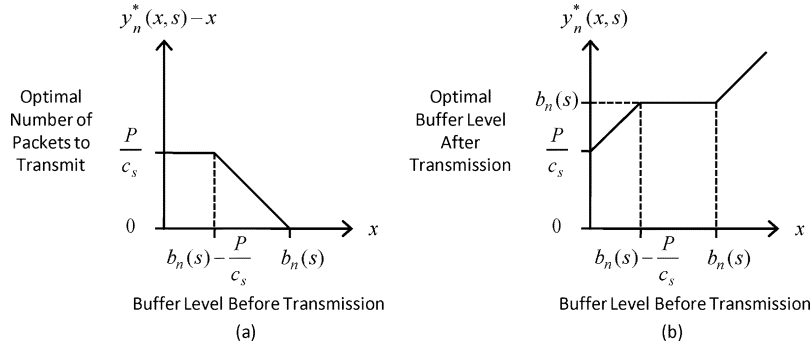


Fig. 3. Structure of the optimal policy for Problem (P1) in the case of a single receiver with linear power-rate curves. When the state is  $(x, s)$  in slot  $n$ , (a) depicts the optimal transmission quantity, and (b) depicts the resulting number of packets available for payout in slot  $n$ .

Doing so is a way to manage the risk of being stuck sending a large amount of data over a poor channel to meet an imminent deadline constraint. Another intuitive takeaway is that the closer the deadlines and the more deadlines it faces, the less “opportunistic” the scheduler can afford to be. In summary, both the underflow constraints and the power constraints shift the definition of what constitutes a “good” channel, and how much data to send accordingly. For more detailed comparisons of single-receiver opportunistic scheduling problems highlighting the role of the deadline constraints, see [34].

#### IV. SINGLE RECEIVER WITH PIECEWISE-LINEAR CONVEX POWER-RATE CURVES

In this section, we analyze Problem (P1) when there is only a single receiver ( $M = 1$ ), and the power-rate functions under different channel conditions are piecewise-linear convex. Note that this is a generalization of the case considered in Section III.

We assume without loss of generality that under each channel condition  $s$ , the power-rate function has  $K + 1$  segments, and thus the power consumed in transmitting  $z$  packets under channel condition  $s$  can be represented as follows:

$$c(z, s) = z \cdot \tilde{c}_0(s) + \sum_{k=0}^{K-1} \{(\tilde{c}_{k+1}(s) - \tilde{c}_k(s)) \max\{z - \tilde{z}_k(s), 0\}\}$$

where  $0 < \tilde{c}_0(s) \leq \tilde{c}_1(s) \leq \dots \leq \tilde{c}_K(s)$ , and

$$0 = \tilde{z}_{-1}(s) < \tilde{z}_0(s) < \tilde{z}_1(s) < \dots < \tilde{z}_{K-1}(s) < \tilde{z}_K(s) = \infty.$$

The terms  $\{\tilde{c}_k(s)\}_{k \in \{0, 1, \dots, K\}}$  represent the slopes of the segments of  $c(\cdot, s)$ , and the terms  $\{\tilde{z}_k(s)\}_{k \in \{0, 1, \dots, K-1\}}$  represent the points at which the slopes of  $c(\cdot, s)$  change. An example of a family of such power-rate functions is shown in Fig. 4. For each channel condition  $s \in \mathcal{S}$ , we define the maximum number of packets that can be transmitted without exceeding the per slot power constraint  $P$  as

$$\tilde{z}_{\max}(s) := \{z : c(z, s) = P\}.$$

Note that  $\tilde{z}_{\max}(s)$  is well defined due to the strictly increasing nature of  $c(\cdot, s)$ . Recall that we assume  $\tilde{z}_{\max}(s) \geq d, \forall s \in \mathcal{S}$ . We also assume without loss of generality that  $\tilde{z}_{\max}(s) > \tilde{z}_{K-1}(s), \forall s \in \mathcal{S}$ .

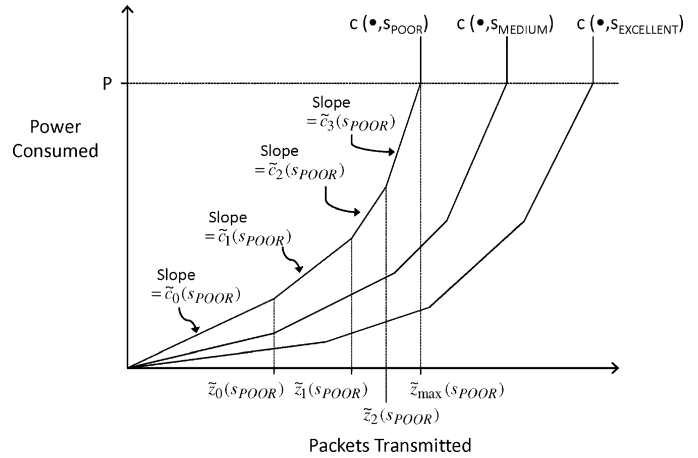


Fig. 4. Family of piecewise-linear convex power-rate functions. Like Fig. 2, we incorporate the power constraint into each curve to show the effective power-rate curve. As an example, the power-rate function  $c(\cdot, s_{\text{POOR}})$  is completely characterized by the sequence of slopes  $\{\tilde{c}_k(s_{\text{POOR}})\}_{k \in \{0, 1, 2, 3\}}$  and the sequence of points where the slopes change  $\{\tilde{z}_k(s_{\text{POOR}})\}_{k \in \{0, 1, 2\}}$ . The maximum number of packets that can be transmitted in a slot when the channel condition is  $s_{\text{POOR}}$  is  $\tilde{z}_{\max}(s_{\text{POOR}})$ .

In this case, the dynamic program (1) for Problem (P1) becomes

$$V_n(x, s) = \min_{\max\{0, d-x\} \leq z \leq \tilde{z}_{\max}(s)} \{c(z, s) + \tilde{g}_n(x+z, s)\},$$

$$n = N, N-1, \dots, 1$$

$$V_0(x, s) = 0, \quad \forall x \in \mathbb{R}_+, \forall s \in \mathcal{S}$$

where  $\tilde{g}_n(y, s) := h(y-d) + \alpha \cdot \mathbb{E}[V_{n-1}(y-d, S_{n-1}) | S_n = s]$ .

##### A. Structure of the Optimal Policy

We showed in Theorem 1 that the optimal transmission policy to a single receiver in the case of linear power-rate curves is a state-dependent modified base-stock policy characterized by a single critical level for each channel condition. In this section, we generalize this result to the case of piecewise-linear power-rate curves, and show that the optimal receiver buffer level after transmission (respectively, optimal number of packets to transmit) is no longer a three-segment piecewise-linear nondecreasing (respectively, nonincreasing) function of the starting buffer level as in Fig. 3, but a more general piecewise-linear nondecreasing (respectively, nonincreasing) function.

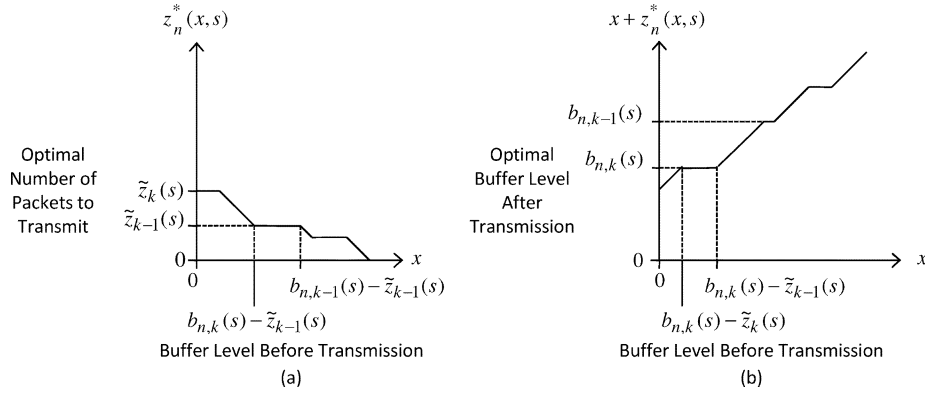


Fig. 5. Structure of the optimal transmission policy for Problem (P1) in the case of a single receiver with piecewise-linear convex power-rate curves. When the state is  $(x, s)$  in slot  $n$ , (a) depicts the optimal transmission quantity, and (b) depicts the resulting number of packets available for payout in slot  $n$ .

**Theorem 3:** In Problem (P1) with a single receiver under piecewise-linear convex power-rate curves, for every  $n \in \{1, 2, \dots, N\}$  and  $s \in \mathcal{S}$ , there exists a nonincreasing sequence of critical numbers  $\{b_{n,k}(s)\}_{k \in \{0,1,\dots,K\}}$  such that the optimal number of packets to transmit with  $n$  slots remaining is given by (10), as shown at the bottom of the page, where  $b_{n,-1}(s) := \infty, \forall s \in \mathcal{S}$ . The optimal receiver buffer level after transmission is given by  $y_n^*(x, s) = x + z_n^*(x, s)$ .

The optimal transmission policy in Theorem 3, which is shown in Fig. 5, is a finite generalized base-stock policy. It can be interpreted as follows. Under each channel condition  $s$ , there is a target level or critical number associated with each segment of the associated piecewise-linear convex power-rate curve shown in Fig. 4. If the starting buffer level is below the critical number associated with the first segment,  $b_{n,0}(s)$ , the scheduler should try to bring the buffer level as close as possible to the target,  $b_{n,0}(s)$ . If the maximum number of packets sent at this per packet power cost,  $\tilde{z}_0(s)$ , does not suffice to reach the critical number  $b_{n,0}(s)$ , then those  $\tilde{z}_0(s)$  packets are scheduled, and the next segment of the power-rate curve is considered. This second segment has a slope of  $\tilde{c}_1(s)$  and an associated critical number  $b_{n,1}(s)$ , which is no higher than  $b_{n,0}(s)$ , the first critical number. If the starting buffer level plus the  $\tilde{z}_0(s)$  already-scheduled packets brings the buffer level above  $b_{n,1}(s)$ , then no more packets are scheduled for transmission. Otherwise, it is optimal to transmit so as to bring the buffer level as close as possible to  $b_{n,1}(s)$ , by transmitting up to  $\tilde{z}_1(s) - \tilde{z}_0(s)$  additional packets at a cost of  $\tilde{c}_1(s)$  power units per packet. This process continues with the sequential consideration of each segment of the power-rate curve. At each successive iteration, the target level is lower and the starting buffer level, updated to include already-scheduled packets, is higher. Eventually, the buffer level reaches or exceeds a

critical number, or the full power  $P$  is consumed. Note that this sequential consideration is not actually done online, but only meant to provide an intuitive explanation of the optimal policy.

### B. Computation of Critical Numbers

While finite generalized base-stock policies have been considered in the inventory literature for almost three decades, we are not aware of any previous studies that explicitly compute the critical numbers for any model where such a policy is optimal. In this section, we compute the critical numbers under each channel condition when technical conditions similar to those of Section III-B are satisfied. We consider the special case when the channel condition is independent and identically distributed from slot to slot; the holding cost function is linear (i.e.,  $h(x) = h \cdot x$ ); and the following technical condition on the power-rate functions is satisfied for each possible channel condition  $s \in \mathcal{S}$ :  $\tilde{z}_{\max}(s) = \tilde{l}_{\max} \cdot d$  for some  $\tilde{l}_{\max} \in \mathbb{N}$ , and for every  $k \in \{0, 1, \dots, K-1\}$ ,  $\tilde{z}_k(s) = \tilde{l}_k \cdot d$  for some  $\tilde{l}_k \in \mathbb{N}$ ; i.e., the slopes of the effective power-rate functions only change at integer multiples of the drainage rate  $d$ . Under these conditions, we can completely characterize the optimal transmission policy.

As in Theorem 2, we recursively define a set of thresholds, and use them to determine the critical numbers,  $\{b_{n,k}(s)\}_{k \in \{-1,0,\dots,K\}}$ , for each channel condition, at each time.

**Theorem 4:** Define the thresholds  $\tilde{\gamma}_{n,j}$  for  $n \in \{1, 2, \dots, N\}$  and  $j \in \mathbb{N}$  recursively, as follows:

- i) If  $j = 1$ ,  $\tilde{\gamma}_{n,j} = \infty$ ;
- ii) If  $j > n$ ,  $\tilde{\gamma}_{n,j} = 0$ ;

$$z_n^*(x, s) := \begin{cases} \tilde{z}_{k-1}(s), & \text{if } b_{n,k}(s) - \tilde{z}_{k-1}(s) < x \leq b_{n,k-1}(s) - \tilde{z}_{k-1}(s), \quad k \in \{0, 1, \dots, K\} \\ b_{n,k}(s) - x, & \text{if } b_{n,k}(s) - \tilde{z}_k(s) < x \leq b_{n,k}(s) - \tilde{z}_{k-1}(s), \quad k \in \{0, 1, \dots, K-1\} \\ b_{n,K}(s) - x, & \text{if } b_{n,K}(s) - \tilde{z}_{\max}(s) < x \leq b_{n,K}(s) - \tilde{z}_{K-1}(s) \\ \tilde{z}_{\max}(s), & \text{if } 0 \leq x \leq b_{n,K}(s) - \tilde{z}_{\max}(s) \end{cases} \quad (10)$$

iii) If  $2 \leq j \leq n$ , let  $\tilde{\gamma}_{n,j}$  be as shown in (11), at the bottom of the page,

where  $p(s)$  is the probability of the channel being in state  $s$  in a time slot,  $\tilde{L}_k(s) := \frac{\tilde{z}_k(s)}{d}$  for all  $s \in \mathcal{S}$  and  $k \in \{0, 1, \dots, K-1\}$ , and  $\tilde{L}_{\max}(s) := \frac{\tilde{z}_{\max}(s)}{d}$  for all  $s \in \mathcal{S}$ . For each  $n \in \{1, 2, \dots, N\}$  and  $s \in \mathcal{S}$ , define  $b_{n,-1}(s) := \infty$  and for all  $k \in \{0, 1, \dots, K\}$ , if  $\tilde{\gamma}_{n,j+1} \leq \tilde{c}_k(s) < \tilde{\gamma}_{n,j}$ , define  $b_{n,k}(s) := j \cdot d$ . The optimal control strategy for Problem (P1) is then given by  $\pi^* = \{z_N^*, z_{N-1}^*, \dots, z_1^*\}$ , where for all  $n \in \{N, N-1, \dots, 1\}$ ,  $z_n^*(x, s)$  is given by (10).

It is straightforward to check that Theorem 4 is in fact a generalization of Theorem 2. To see this, set  $K = 0$  so that the summation from  $k = 0$  to  $k = K-1$  on the right-hand side of (11) drops out. Then  $\tilde{\gamma}_{n,j}$  in (11) is the same as  $\gamma_{n,j}$  in (9),  $\tilde{c}_0(s)$  corresponds to  $c_s$  in (9),  $b_{n,0}(s)$  corresponds to  $b_n(s)$ ,  $\tilde{z}_{\max}(s)$  corresponds to  $\frac{P}{c_s}$ ,  $\tilde{L}_{\max}(s)$  corresponds to  $L(s)$ , and  $\tilde{L}_{K-1}(s) = 0$ . The resulting optimal transmission policies are also the same.

In Theorem 4, the threshold  $\tilde{\gamma}_{n,j}$  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. The intuition behind the recursion (11) is similar to the detailed explanation of (9) given in Section A of the Appendix. A detailed proof of Theorem 4 is included in [45, Appendix A] and [46].

### C. General Convex Power-Rate Curves

As mentioned in Section II-A, in general, the power-rate curve under each possible channel condition is convex. It can be shown that under convex power-rate curves at each time, the optimal number of packets to send is a nonincreasing function of the starting buffer level. However, without any further structure on the power-rate curves, it is not computationally tractable to compute such optimal policies, known as generalized base-stock policies (a superclass of the *finite* generalized base-stock policies discussed above). This is why we have chosen to analyze piecewise-linear convex power-rate curves,

which can be used to approximate general convex power-rate curves. More specifically, our analysis suggests approximating the general convex power-rate curves by piecewise-linear convex power-rate curves where the slopes change at integer multiples of the demand  $d$ , in order to be able to apply Theorem 4 to compute the critical numbers in an extremely efficient manner. Doing so represents an approximation at the modeling stage followed by an exact solution, as compared to modeling the power-rate curves as more general convex functions and having to approximate the solution.

### V. TWO RECEIVERS WITH LINEAR POWER-RATE CURVES

In this section, we analyze the finite horizon discounted expected cost problem when there are two receivers ( $M = 2$ ), and the power-rate functions under different channel conditions are linear for each user. Each user  $m$ 's channel condition evolves as a homogeneous Markov process,  $\{S_n^m\}_{n=N, N-1, \dots, 1}$ . As discussed in Sections I and II, the time-varying channel conditions of the two users are independent of each other, and the transmission scheduler can exploit this spatial diversity. Like Section III, we denote the power consumption per unit of data transmitted to receiver  $m$  under channel condition  $s^m$  by  $c_s^m$ . The row vector of these per unit power consumptions is given by  $\mathbf{c}_s^T$ , so that the total power consumption in slot  $n$  is given by  $\sum_{m=1}^2 c^m(Z_n^m, S_n^m) = \mathbf{c}_s^T \mathbf{Z}_n$ . We denote the total holding costs  $\sum_{m=1}^2 h^m(X_n^m + Z_n^m - d^m)$  by  $h(\mathbf{X}_n + \mathbf{Z}_n - \mathbf{d})$ .

With these notations, the dynamic program (1) for Problem (P1) becomes

$$\begin{aligned} V_n(\mathbf{x}, \mathbf{s}) &= \min_{\mathbf{z} \in \mathcal{A}^d(\mathbf{x}, \mathbf{s})} \left\{ \mathbf{c}_s^T \mathbf{z} + h(\mathbf{x} + \mathbf{z} - \mathbf{d}) \right. \\ &\quad \left. + \alpha \cdot \mathbb{E}[V_{n-1}(\mathbf{x} + \mathbf{z} - \mathbf{d}, \mathbf{S}_{n-1}) | \mathbf{S}_n = \mathbf{s}] \right\} \quad (12) \\ &= \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})} \left\{ \mathbf{c}_s^T [\mathbf{y} - \mathbf{x}] + h(\mathbf{y} - \mathbf{d}) \right. \\ &\quad \left. + \alpha \cdot \mathbb{E}[V_{n-1}(\mathbf{y} - \mathbf{d}, \mathbf{S}_{n-1}) | \mathbf{S}_n = \mathbf{s}] \right\} \quad (13) \\ &= -\mathbf{c}_s^T \mathbf{x} + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})} \{G_n(\mathbf{y}, \mathbf{s})\}, \quad n = N, N-1, \dots, 1, \\ V_0(\mathbf{x}, \mathbf{s}) &= 0, \quad \forall \mathbf{x} \in \mathbb{R}_+^2, \quad \forall \mathbf{s} \in \mathcal{S} := \mathcal{S}^1 \times \mathcal{S}^2 \end{aligned}$$

$$\tilde{\gamma}_{n,j} = -h + \alpha \cdot \left( \begin{aligned} &\sum_{s: \tilde{c}_0(s) \geq \tilde{\gamma}_{n-1, j-1}} p(s) \cdot \tilde{\gamma}_{n-1, j-1} \\ &+ \sum_{k=0}^{K-1} \left\{ \begin{aligned} &\sum_{s: \tilde{\gamma}_{n-1, j-1} + \tilde{L}_k(s) \leq \tilde{c}_k(s) < \tilde{\gamma}_{n-1, j-1} + \tilde{L}_{k-1}(s)} p(s) \cdot \tilde{c}_k(s) \\ &+ \sum_{s: \tilde{c}_k(s) < \tilde{\gamma}_{n-1, j-1} + \tilde{L}_k(s) \leq \tilde{c}_{k+1}(s)} p(s) \cdot \tilde{\gamma}_{n-1, j-1} + \tilde{L}_k(s) \end{aligned} \right\} \\ &+ \sum_{s: \tilde{\gamma}_{n-1, j-1} + \tilde{L}_{\max}(s) \leq \tilde{c}_K(s) < \tilde{\gamma}_{n-1, j-1} + \tilde{L}_{K-1}(s)} p(s) \cdot \tilde{c}_K(s) \\ &+ \sum_{s: \tilde{c}_K(s) < \tilde{\gamma}_{n-1, j-1} + \tilde{L}_{\max}(s)} p(s) \cdot \tilde{\gamma}_{n-1, j-1} + \tilde{L}_{\max}(s) \end{aligned} \right) \quad (11)$$

where

$$G_n(\mathbf{y}, \mathbf{s}) := \mathbf{c}_s^T \mathbf{y} + h(\mathbf{y} - \mathbf{d}) \\ + \alpha \cdot \mathbb{E}[V_{n-1}(\mathbf{y} - \mathbf{d}, \mathbf{S}_{n-1}) | \mathbf{S}_n = \mathbf{s}], \\ \forall \mathbf{y} \in [d^1, \infty) \times [d^2, \infty), \forall \mathbf{s} \in \mathcal{S}, \quad \text{and} \\ \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s}) := \{\mathbf{y} \in \mathbb{R}_+^2 : \mathbf{y} \succeq \mathbf{d} \vee \mathbf{x} \text{ and } \mathbf{c}_s^T [\mathbf{y} - \mathbf{x}] \leq P\}, \\ \forall \mathbf{x} \in \mathbb{R}_+^2, \forall \mathbf{s} \in \mathcal{S}.$$

The transition from (12) to (13) follows again from a change of variable in the action space from  $\mathbf{Z}_n$  to  $\mathbf{Y}_n$ , where  $\mathbf{Y}_n = \mathbf{X}_n + \mathbf{Z}_n$ . The controlled random vector  $\mathbf{Y}_n$  represents the queue lengths of the receiver buffers *after* transmission takes place in the  $n$ th slot, but *before* playout takes place (i.e., before  $d^m$  packets are removed from user  $m$ 's buffer).

Without the per slot peak power constraint, this two-dimensional problem could be separated into two instances of the one-dimensional problem of Section III; however, the joint power constraint couples the queues.<sup>10</sup> As a result, the optimal transmission quantity to one receiver depends on the other receivers' queue length, as the following example shows.

*Example 1:* Assume receiver 1's channel is currently in a "poor" condition, receiver 2's channel is currently in a "medium" condition, and receiver 2's buffer contains enough packets to satisfy the demand for the next few slots. We consider two different scenarios for receiver 1's buffer level to show how the optimal transmission quantity to receiver 2 depends on receiver 1's buffer level. In Scenario 1, receiver 1's buffer already contains many packets. In this scenario, it may be beneficial for the scheduler to wait for receiver 2 to have a better channel condition, because it will be able to take full advantage of an "excellent" condition when it comes. In Scenario 2, receiver 1's queue only contains enough packets to satisfy the demand in the current slot. It may be optimal to transmit some packets to receiver 2 in the current slot in this scenario. To see this, note that even if receiver 2 experiences the best possible channel condition in the next slot, the scheduler will need to allocate some power to receiver 1 in order to prevent receiver 1's buffer from emptying. Therefore, the scheduler anticipates not being able to take full advantage of receiver 2's "excellent" condition in the next slot, and may compensate by sending some packets in the current slot under the "medium" condition.

#### A. Structure of the Optimal Policy

Before proceeding to the structure of the optimal transmission policy, we state some key properties of the value functions in the following theorem, a detailed proof of which is included in the Appendix.

*Theorem 5:* With two receivers and linear power-rate curves, the following statements are true for  $n = 1, 2, \dots, N$ , and for all  $\mathbf{s} \in \mathcal{S}$ :

- i)  $V_{n-1}(\mathbf{x}, \mathbf{s})$  is convex in  $\mathbf{x}$ .
- ii)  $V_{n-1}(\mathbf{x}, \mathbf{s})$  is supermodular in  $\mathbf{x}$ ; i.e., for all  $\bar{\mathbf{x}}, \tilde{\mathbf{x}} \in \mathbb{R}_+^2$ ,  $V_{n-1}(\bar{\mathbf{x}}, \mathbf{s}) + V_{n-1}(\tilde{\mathbf{x}}, \mathbf{s}) \leq V_{n-1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) + V_{n-1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ .

<sup>10</sup>This problem therefore falls into the class of *weakly coupled stochastic dynamic programs* [67], [68].

iii)  $G_n(\mathbf{y}, \mathbf{s})$  is convex in  $\mathbf{y}$ .

iv)  $G_n(\mathbf{y}, \mathbf{s})$  is supermodular in  $\mathbf{y}$ ; i.e., for all  $\bar{\mathbf{y}}, \tilde{\mathbf{y}} \in [d^1, \infty) \times [d^2, \infty)$ ,

$$G_n(\bar{\mathbf{y}}, \mathbf{s}) + G_n(\tilde{\mathbf{y}}, \mathbf{s}) \leq G_n(\bar{\mathbf{y}} \wedge \tilde{\mathbf{y}}, \mathbf{s}) + G_n(\bar{\mathbf{y}} \vee \tilde{\mathbf{y}}, \mathbf{s}).$$

v)  $y_n^1 < \hat{y}_n^1$  implies

$$\inf \left\{ \operatorname{argmin}_{y_n^1 \in [d^1, \infty)} \{G_n(y_n^1, y_n^2, s^1, s^2)\} \right\} \\ \geq \inf \left\{ \operatorname{argmin}_{y_n^1 \in [d^1, \infty)} \{G_n(\hat{y}_n^1, y_n^2, s^1, s^2)\} \right\}$$

and  $y_n^2 < \hat{y}_n^2$  implies

$$\inf \left\{ \operatorname{argmin}_{y_n^1 \in [d^1, \infty)} \{G_n(y_n^1, y_n^2, s^1, s^2)\} \right\} \\ \geq \inf \left\{ \operatorname{argmin}_{y_n^1 \in [d^1, \infty)} \{G_n(y_n^1, \hat{y}_n^2, s^1, s^2)\} \right\}.$$

The following theorem on the structure of the optimal transmission policy for the finite horizon discounted expected cost problem leverages the functional properties of Theorem 5.

*Theorem 6:* For every  $n \in \{1, 2, \dots, N\}$  and  $\mathbf{s} \in \mathcal{S}^1 \times \mathcal{S}^2$ , define the nonempty set of global minimizers of  $G_n(\cdot, \mathbf{s})$ :

$$\mathcal{B}_n(\mathbf{s}) := \left\{ \hat{\mathbf{y}} \in [d^1, \infty) \times [d^2, \infty) : \right.$$

$$\left. G_n(\hat{\mathbf{y}}, \mathbf{s}) = \min_{\mathbf{y} \in [d^1, \infty) \times [d^2, \infty)} G_n(\mathbf{y}, \mathbf{s}) \right\}.$$

Define also

$$b_n^1(\mathbf{s}) := \min\{y^1 \in [d^1, \infty) : c(y^1, y^2) \in \mathcal{B}_n(\mathbf{s}) \\ \text{for some } y^2 \in [d^2, \infty)\}$$

$$b_n^2(\mathbf{s}) := \min\{y^2 \in [d^2, \infty) : (b_n^1(\mathbf{s}), y^2) \in \mathcal{B}_n(\mathbf{s})\}.$$

Then the vector  $\mathbf{b}_n(\mathbf{s}) = (b_n^1(\mathbf{s}), b_n^2(\mathbf{s})) \in \mathcal{B}_n(\mathbf{s})$  is a global minimizer of  $G_n(\cdot, \mathbf{s})$ . Define also the functions

$$f_n^1(x^2, \mathbf{s}) := \inf \left\{ \operatorname{argmin}_{y^1 \in [d^1, \infty)} \{G_n(y^1, x^2, s^1, s^2)\} \right\}, \\ \text{for } x^2 \in [d^2, \infty), \text{ and}$$

$$f_n^2(x^1, \mathbf{s}) := \inf \left\{ \operatorname{argmin}_{y^2 \in [d^2, \infty)} \{G_n(x^1, y^2, s^1, s^2)\} \right\}, \\ \text{for } x^1 \in [d^1, \infty).$$

Note that by construction,  $f_n^1(b_n^2(\mathbf{s}), \mathbf{s}) = b_n^1(\mathbf{s})$  and  $f_n^2(b_n^1(\mathbf{s}), \mathbf{s}) = b_n^2(\mathbf{s})$ . Partition  $\mathbb{R}_+^2$  into the following seven regions:

$$\mathcal{R}_I(n, \mathbf{s}) := \{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x} \neq \mathbf{b}_n(\mathbf{s}) \text{ and} \\ \mathbf{x} \succeq (f_n^1(x^2, \mathbf{s}), f_n^2(x^1, \mathbf{s}))\}$$

$$\mathcal{R}_{II}(n, \mathbf{s}) := \{\mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x} \preceq \mathbf{b}_n(\mathbf{s}) \text{ and} \\ \mathbf{c}_s^T [\mathbf{b}_n(\mathbf{s}) - \mathbf{x}] \leq P\}$$

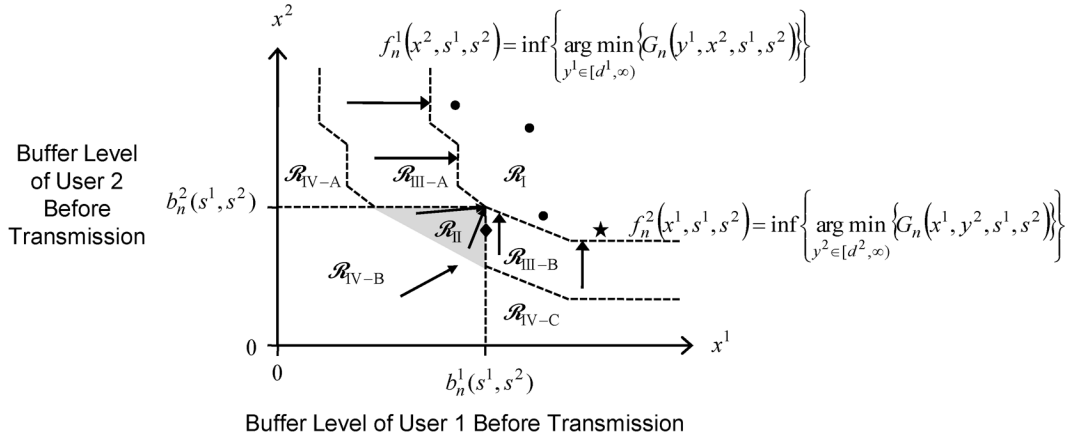


Fig. 6. Structure of the optimal transmission policy for the case of two receivers with linear power-rate curves. The state in slot  $n$  is  $(\mathbf{x}, \mathbf{s})$ . The seven regions described in Theorem 6 are labeled. The tails of the arrows represent the vectors of the receiver buffer levels at the beginning of slot  $n$ , and the heads of the arrows represent the vectors of the receiver buffer levels after transmission but before playout in slot  $n$  under the optimal transmission policy. In region  $\mathcal{R}_I(n, \mathbf{s})$ , a single dot represents that it is optimal to not transmit any packets to either user. The  $\star$  and  $\blacklozenge$  represent possible starting buffer levels for Scenarios 1 and 2, respectively, in Example 1.

$$\mathcal{R}_{III-A}(n, \mathbf{s}) := \left\{ \mathbf{x} \in \mathbb{R}_+^2 : x^2 > b_n^2(\mathbf{s}) \text{ and } f_n^1(x^2, \mathbf{s}) - \frac{P}{c_{s^1}} \leq x^1 < f_n^1(x^2, \mathbf{s}) \right\}$$

$$\mathcal{R}_{III-B}(n, \mathbf{s}) := \left\{ \mathbf{x} \in \mathbb{R}_+^2 : x^1 > b_n^1(\mathbf{s}) \text{ and } f_n^2(x^1, \mathbf{s}) - \frac{P}{c_{s^2}} \leq x^2 < f_n^2(x^1, \mathbf{s}) \right\}$$

$$\mathcal{R}_{IV-A}(n, \mathbf{s}) := \left\{ \mathbf{x} \in \mathbb{R}_+^2 : x^2 > b_n^2(\mathbf{s}) \text{ and } x^1 < f_n^1(x^2, \mathbf{s}) - \frac{P}{c_{s^1}} \right\}$$

$$\mathcal{R}_{IV-B}(n, \mathbf{s}) := \left\{ \mathbf{x} \in \mathbb{R}_+^2 : \mathbf{x} \preceq \mathbf{b}_n(\mathbf{s}) \text{ and } \mathbf{c}_s^T[\mathbf{b}_n(\mathbf{s}) - \mathbf{x}] > P \right\}$$

$$\mathcal{R}_{IV-C}(n, \mathbf{s}) := \left\{ \mathbf{x} \in \mathbb{R}_+^2 : x^1 > b_n^1(\mathbf{s}) \text{ and } x^2 < f_n^2(x^1, \mathbf{s}) - \frac{P}{c_{s^2}} \right\}$$

and define  $\mathcal{R}_{IV}(n, \mathbf{s}) := \mathcal{R}_{IV-A}(n, \mathbf{s}) \cup \mathcal{R}_{IV-B}(n, \mathbf{s}) \cup \mathcal{R}_{IV-C}(n, \mathbf{s})$ .

Then for Problem (P1) in the case of two receivers with linear power-rate curves, for all  $\mathbf{x} \notin \mathcal{R}_{IV}(n, \mathbf{s})$ , an optimal control action with  $n$  slots remaining is given by

$$\mathbf{y}_n^*(\mathbf{x}, \mathbf{s}) := \begin{cases} \mathbf{x}, & \text{if } \mathbf{x} \in \mathcal{R}_I(n, \mathbf{s}) \\ \mathbf{b}_n(\mathbf{s}), & \text{if } \mathbf{x} \in \mathcal{R}_{II}(n, \mathbf{s}) \\ (f_n^1(x^2, \mathbf{s}), x^2), & \text{if } \mathbf{x} \in \mathcal{R}_{III-A}(n, \mathbf{s}) \\ (x^1, f_n^2(x^1, \mathbf{s})), & \text{if } \mathbf{x} \in \mathcal{R}_{III-B}(n, \mathbf{s}) \end{cases} \quad (14)$$

Furthermore, for all  $\mathbf{x} \in \mathcal{R}_{IV}(n, \mathbf{s})$ , there exists an optimal control action with  $n$  slots remaining,  $\mathbf{y}_n^*(\mathbf{x}, \mathbf{s})$ , which satisfies

$$\mathbf{c}_s^T[\mathbf{y}_n^*(\mathbf{x}, \mathbf{s}) - \mathbf{x}] = P. \quad (15)$$

A detailed proof is included in the Appendix. Equation (15) says that it is optimal for the transmitter to allocate the full power budget for transmission when the vector of receiver buffer levels at the beginning of slot  $n$  falls in region  $\mathcal{R}_{IV}(n, \mathbf{s})$ . We cannot say anything in general about the optimal allocation (split) of the full power budget between the two receivers when the starting buffer levels lie in region  $\mathcal{R}_{IV}(n, \mathbf{s})$ . Fig. 6 shows the partition of  $\mathbb{R}_+^2$  into the seven regions, and a diagram of the structure of the optimal transmission policy. Note that the figure shows the seven regions of the optimal policy for a *fixed realization* of the pair of channel conditions. Under different pairs of channel realizations, the seven regions have the same general form, but the targets  $\mathbf{b}_n(\mathbf{s})$  are shifted and the boundary functions  $f_n^1(x^2, \mathbf{s})$  and  $f_n^2(x^1, \mathbf{s})$  are different.

In some sense, the structure of the optimal policy outlined in Theorem 6 can be interpreted as an extension of the modified base-stock policy for the case of a single receiver outlined in Theorem 1. Namely, under each channel condition at each time, there is a critical number for each receiver ( $b_n^m(\mathbf{s})$ ) such that it is optimal to bring both receivers' buffer levels up to those critical numbers if it is possible to do so (region  $\mathcal{R}_{II}(n, \mathbf{s})$ ), and it is optimal to not transmit any packets if both receivers' buffer levels start beyond their critical numbers (region  $\mathcal{R}_I(n, \mathbf{s})$ ). However, this extended notion of the modified base-stock policy only captures the optimal behavior in two of the seven regions, and does not account for the coupling behavior between users that arises through the joint power constraint. For instance, possible starting buffer levels for Scenario 1 and Scenario 2 in Example 1 are illustrated in Fig. 6 by the  $\star$  and  $\blacklozenge$ , respectively. Even though the buffer level of receiver 2 before transmission is the same under both scenarios, the optimal transmission quantity to receiver 2 is different under the two scenarios due to the different starting buffer levels of receiver 1.

## VI. EXTENSIONS

In this section, we discuss the infinite horizon problems, the relaxation of the strict underflow constraints, and the extension to the general case of  $M$  receivers.

### A. The Infinite Horizon Problems

The structure of the optimal *stationary* (or *time-invariant*) policy for the infinite horizon discounted expected cost problem, Problem (P2), is the same as the structure of the optimal policy for the finite horizon discounted expected cost problem. Namely, for the case of a single receiver under linear power-rate curves, it is a modified base-stock policy; for the case of a single receiver under piecewise-linear convex power-rate curves, it is a finite generalized base-stock policy; and for the case of two receivers under linear power-rate curves, it is of the seven-region form shown in Fig. 6. Moreover, with a single receiver, the time-invariant sequences of critical numbers that complete the characterizations of the modified base-stock and finite generalized base-stock policies are equal to the limits of the sequences of critical numbers that characterize the finite horizon optimal policies as the time horizon  $N$  goes to infinity. Similarly, with two receivers, the boundaries of the seven regions of the finite horizon optimal policy shown in Fig. 6 converge to the boundaries of the seven regions of the infinite horizon discounted expected cost optimal policy as the time horizon  $N$  goes to infinity.

For all three cases, optimal policies for the infinite horizon average expected cost problem, Problem (P3), exist and can be represented as the limit of optimal policies for Problem (P2) as the discount factor increases to one. This technique is called the *vanishing discount approach* (see, e.g., [43]). Thus, the modified base-stock, finite generalized base-stock, and seven-region structures are also average cost optimal for the three cases, respectively. For precise statements and proofs of the structures of the optimal policies for Problems (P2) and (P3), see Theorems 5, 6, 9, and 10 of [45] or Theorems 5.5, 5.6, 5.10, and 5.11 of [46].

### B. Relaxation of the Strict Underflow Constraints

In some applications, it may not be the case that the peak power per slot is always sufficient to transmit one slot's worth of packets to each receiver, even under the worst channel conditions. In this case, a more appropriate model is to relax the strict underflow constraints, and allow underflow at a cost. One way to model this situation is to allow the receivers' queues to be negative, with a negative buffer level representing the number of packets that the playout process is behind. Then, in addition to the holding costs assessed on positive buffer levels, shortage costs are assessed on negative buffer levels. With some minor alterations to the proofs, it is straightforward to show that as long as the shortage cost function is a convex function of the negative buffer level, the structural results of Theorems 1, 3, and 6 are essentially unchanged by the relaxation of the strict underflow constraints to loose underflow constraints with penalties on underflow. This is not too surprising as the strict underflow constraint case we consider can be thought of as the limiting case as the penalties on underflow go to infinity.<sup>11</sup>

<sup>11</sup>Tracking the number of packets that the playout process is behind in this manner corresponds to the *complete backlogging* assumption in inventory theory. An alternate model is to say that a packet is of no use once it misses its deadline, penalize missed packets, and keep the receiver queue length at zero. This model corresponds to the *lost sales* assumption in inventory theory.

### C. Extension to the General Case of $M$ Receivers

Our ongoing work includes examining the extension to the most general case of  $M$  receivers. It is unlikely that the structure of the optimal policy in this case has a simple, intuitive, and implementable form. Therefore, our approach is to find lower bounds on the value function and a feasible policy whose expected cost is as close as possible to these bounds. See [45, Section VI-B] or [46, Section 8.1.3] for a discussion of lower bounding methods, some of which call for approximating the  $M$ -dimensional problem by lower dimensional subproblems. These same numerical techniques are most likely also the best way to approximate the boundaries of the seven regions of the two receiver optimal policy, and determine a near-optimal split of the power  $P$  between the two receivers when the vector of starting receiver buffer levels is in the power-constrained region  $\mathcal{R}_{IV}(n, \mathbf{s})$ .

The results we have presented in this paper are useful not only in terms of the intuition they provide, but also in generating feasible policies for the most general case of  $M$  receivers and solving the subproblems that are used to find lower bounds for the  $M$ -dimensional problem.

## VII. CONCLUSION

In this paper, we considered the problem of transmitting data to one or more receivers over a shared wireless channel in a manner that minimizes power consumption and prevents the receivers' buffers from emptying. We presented a novel connection between this wireless communications model and an inventory model with stochastic ordering costs. We showed that the optimal transmission policy to a single receiver has an easily implementable modified base-stock structure when the power-rate curves are linear, and a finite generalized base-stock structure when they are piecewise-linear convex. When additional technical conditions are satisfied, we presented efficient methods to compute the critical numbers that fully characterize the optimal modified base-stock and finite generalized base-stock policies.

We also analyzed the structure of the optimal transmission policy for the case of two receivers. In some sense, the structure of the optimal policy is an extension of the modified base-stock policy; however, the peak power constraint couples the optimal scheduling of the two data streams, and the time-varying channel conditions may result in counterintuitive optimal scheduling decisions.

The extension to the most general case of  $M$  receivers is quite complex, and it is likely that numerical approximation techniques need to be used to develop further insights on the nature of the optimal policy. Many of the numerical approaches to the  $M$  receiver case relax the higher dimensional problem, so as to decouple it into multiple instances of a lower dimensional subproblem. Therefore, the results presented in this paper for the cases of one and two receivers may also indirectly improve the quality of approximate numerical solutions to related higher dimensional problems.

## APPENDIX

## A. Discussion of Theorem 2

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 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  $\{\gamma_{n,j}\}_{n \in \{1,2,\dots,N\}, j \in \mathbb{N}}$ .

An intuitive explanation of the recursion (9) is as follows. 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. That is,  $\gamma_{n,j}$  should satisfy

$$\alpha \cdot \mathbb{E}[V_{n-1}((j-1) \cdot d, S_{n-1})] + \gamma_{n,j} \cdot d + h \cdot d \\ = \alpha \cdot \mathbb{E}[V_{n-1}((j-2) \cdot d, S_{n-1})]$$

which is equivalent to (16)–(19), as shown at the bottom of the page. Here, (17) follows from the structure of the optimal control action (5). If the channel condition  $s$  in the  $(n-1)$ st slot is

such that  $b_{n-1}(s) \leq (j-2) \cdot d$ , then no packets are transmitted when the starting buffer level is either  $(j-2) \cdot d$  or  $(j-1) \cdot d$ , and the respective buffer levels at the beginning of slot  $n-2$  are  $(j-3) \cdot d$  and  $(j-2) \cdot d$ . The instantaneous costs resulting from the two starting buffer levels differ by  $-h \cdot d$ . When  $(j-2) \cdot d < b_{n-1}(s) \leq (j-2+L(s)) \cdot d$ , the power constraint is not tight starting from  $(j-1) \cdot d$ , so the buffer level after transmission is the same starting from  $(j-2) \cdot d$  or  $(j-1) \cdot d$ . The instantaneous costs resulting from the two starting buffer levels differ by  $c_s \cdot d$ , as an extra  $d$  packets are transmitted if the starting buffer is  $(j-2) \cdot d$ . Finally, when  $b_{n-1}(s) > (j-2+L(s)) \cdot d$ , the power constraint is tight starting from both  $(j-2) \cdot d$  and  $(j-1) \cdot d$ . Therefore, the instantaneous cost difference is  $-h \cdot d$ , and the respective buffer levels at the beginning of slot  $n-2$  are  $(j-3+L(s)) \cdot d$  and  $(j-2+L(s)) \cdot d$ . Equation (18) follows from (16), with  $n-1, j-1$  substituted for  $n, j$ , and (19) follows from the definition that  $b_n(s) = j \cdot d$  if  $\gamma_{n,j+1} \leq c_s < \gamma_{n,j}$ .

Comparing the threshold  $\gamma_{n,j}$  defined in (9) to the corresponding threshold in the uncapacitated (no power constraint) single user problem [49], [54], the only difference is the third term of the right-hand side of (9)

$$\alpha \cdot \sum_{\{s: c_s < \gamma_{n-1, j-1+L(s)}\}} p(s) \cdot [\gamma_{n-1, j-1+L(s)} - c_s]$$

which is absent in the uncapacitated case. For all  $n \in \{1, 2, \dots, N\}$  and  $j \in \mathbb{N}$ , this term is nonnegative. Thus, for a fixed  $n$  and  $j$ , the threshold in the capacitated case is at least as high as the corresponding threshold in the uncapacitated case. It follows that the optimal stock-up level  $b_n(s)$  is also at least

$$\gamma_{n,j} = -h + \frac{\alpha}{d} \cdot \mathbb{E}[V_{n-1}((j-2) \cdot d, S_{n-1}) - V_{n-1}((j-1) \cdot d, S_{n-1})] \\ = -h + \frac{\alpha}{d} \cdot \sum_{s \in \mathcal{S}} p(s) \cdot [V_{n-1}((j-2) \cdot d, s) - V_{n-1}((j-1) \cdot d, s)] \quad (16)$$

$$= -h + \frac{\alpha}{d} \cdot \left\{ \begin{aligned} & \sum_{s: b_{n-1}(s) \leq (j-2) \cdot d} p(s) \cdot \left\{ -h \cdot d + \alpha \cdot \mathbb{E} \left[ \begin{array}{l} V_{n-2}((j-3) \cdot d, S_{n-2}) \\ -V_{n-2}((j-2) \cdot d, S_{n-2}) \end{array} \right] \right\} \\ & + \sum_{s: (j-2) \cdot d < b_{n-1}(s) \leq (j-2+L(s)) \cdot d} p(s) \cdot c_s \cdot d \\ & + \sum_{s: b_{n-1}(s) > (j-2+L(s)) \cdot d} p(s) \cdot \left\{ -h \cdot d + \alpha \cdot \mathbb{E} \left[ \begin{array}{l} V_{n-2}((j-3+L(s)) \cdot d, S_{n-2}) \\ -V_{n-2}((j-2+L(s)) \cdot d, S_{n-2}) \end{array} \right] \right\} \end{aligned} \right\} \quad (17)$$

$$= -h + \alpha \cdot \left\{ \begin{aligned} & \sum_{s: b_{n-1}(s) \leq (j-2) \cdot d} p(s) \cdot \gamma_{n-1, j-1} \\ & + \sum_{s: (j-2) \cdot d < b_{n-1}(s) \leq (j-2+L(s)) \cdot d} p(s) \cdot c_s \\ & + \sum_{s: b_{n-1}(s) > (j-2+L(s)) \cdot d} p(s) \cdot \gamma_{n-1, j-1+L(s)} \end{aligned} \right\} \quad (18)$$

$$= -h + \alpha \cdot \left\{ \begin{aligned} & \sum_{s: c_s \geq \gamma_{n-1, j-1}} p(s) \cdot \gamma_{n-1, j-1} \\ & + \sum_{s: \gamma_{n-1, j-1+L(s)} \leq c_s < \gamma_{n-1, j-1}} p(s) \cdot c_s \\ & + \sum_{s: c_s < \gamma_{n-1, j-1+L(s)}} p(s) \cdot \gamma_{n-1, j-1+L(s)} \end{aligned} \right\} \quad (19)$$

as high in the capacitated case for all  $n \in \{1, 2, \dots, N\}$  and  $s \in \mathcal{S}$ . 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.

### B. Proof of Theorem 3

While the proof is similar in spirit to the proof of a finite generalized base-stock policy in [58, pp. 324–334], some key differences include the introduction of: (i) stochastic channel conditions (ordering costs); (ii) the underflow constraint  $x + z \geq d$ ; and (iii) the power constraint  $z \leq \tilde{z}_{\max}(s)$ .

We show by induction on  $n$  that the following two statements are true for every  $n \in \{1, 2, \dots, N\}$  and  $s \in \mathcal{S}$ :

- i)  $V_n(x, s)$  is convex in  $x$  on  $\mathbb{R}_+$ .
- ii) There exists a nonincreasing sequence of critical numbers  $\{b_{n,k}(s)\}_{k \in \{-1, 0, 1, \dots, K\}}$  such that the optimal control action with  $n$  slots remaining is given by (20), as shown below.

**Base Case:**  $n = 1$

$$\begin{aligned} V_1(x, s) &= \min_{\max\{0, d-x\} \leq z \leq \tilde{z}_{\max}(s)} \{c(z, s) + h(x + z - d)\} \\ &= c(\max\{0, d-x\}, s) + h(\max\{0, x-d\}) \end{aligned} \quad (21)$$

which is convex because  $c(\cdot, s)$  and  $h(\cdot)$  are both convex and nondecreasing functions, and  $\max\{0, d-x\}$  and  $\max\{0, x-d\}$  are both convex functions (see, e.g., [69, Section 3.2] for the relevant results on convexity-preserving operations). Further, let  $b_{1,-1}(s) = \infty$  and  $b_{1,k}(s) = d$  for all  $k \in \{0, 1, \dots, K\}$ . Then

(20) is equivalent to  $z_1^*(x, s) = \max\{0, d-x\}$ , which clearly achieves the minimum in (21).

**Induction Step:** We now assume (i)–(ii) are true for  $n = m-1$  and all  $s \in \mathcal{S}$ , and show they hold for  $n = m$  and an arbitrary  $s \in \mathcal{S}$ . Let  $\check{x}, \hat{x} \in \mathbb{R}_+$  and  $\theta \in [0, 1]$  be arbitrary, and define  $\bar{x} := \theta \cdot \check{x} + (1-\theta) \cdot \hat{x}$ . Then we have the series of equations (22)–(23) shown at the bottom of the page, where (23) follows from the convexity of  $c(\cdot, s)$ ,  $h(\cdot)$ , and  $\mathbb{E}[V_{m-1}(\cdot, S_{m-1}) | S_m = s]$ , the last of which follows from the induction hypothesis. Equation (22) follows from the fact that for every  $\max\{0, d-\check{x}\} \leq \check{z} \leq \tilde{z}_{\max}(s)$  and  $\max\{0, d-\hat{x}\} \leq \hat{z} \leq \tilde{z}_{\max}(s)$ , there exists a  $\max\{0, d-\bar{x}\} \leq \bar{z} \leq \tilde{z}_{\max}(s)$  (namely,  $\bar{z} := \theta \cdot \check{z} + (1-\theta) \cdot \hat{z}$ ) such that

$$\begin{aligned} &c(\bar{z}, s) + h(\bar{x} + \bar{z} - d) \\ &+ \alpha \cdot \mathbb{E}[V_{m-1}(\bar{x} + \bar{z} - d, S_{m-1}) | S_m = s] \\ &= c(\theta \cdot \check{z} + (1-\theta) \cdot \hat{z}, s) \\ &+ h(\bar{x} + \theta \cdot \check{z} + (1-\theta) \cdot \hat{z} - d) \\ &+ \alpha \cdot \mathbb{E}[V_{m-1}(\bar{x} + \theta \cdot \check{z} + (1-\theta) \cdot \hat{z} - d, S_{m-1}) | S_m = s]. \end{aligned}$$

This concludes the induction step for (i) and we now proceed to (ii).

Note first that  $\check{g}_m(y, s) = h(y-d) + \alpha \cdot \mathbb{E}[V_{m-1}(y-d, S_{m-1}) | S_m = s]$  is convex in  $y$ , as  $h(\cdot)$  is convex, and  $V_{m-1}(x, s)$  is convex in  $x$  for every  $s \in \mathcal{S}$  by the induction hypothesis. Let  $b_{m,-1}(s) := \infty$  and

$$b_{m,k}(s) := \max \{d, \inf \{b | \check{g}_m^+(b, s) \geq -\check{c}_k(s)\}\}, \quad \forall k \in \{0, 1, \dots, K\}$$

$$z_n^*(x, s) := \begin{cases} \tilde{z}_{k-1}(s), & \text{if } b_{n,k}(s) - \tilde{z}_{k-1}(s) \leq x < b_{n,k-1}(s) - \tilde{z}_{k-1}(s), \quad k \in \{0, 1, \dots, K\} \\ b_{n,k}(s) - x, & \text{if } b_{n,k}(s) - \tilde{z}_k(s) \leq x < b_{n,k}(s) - \tilde{z}_{k-1}(s), \quad k \in \{0, 1, \dots, K-1\} \\ b_{n,K}(s) - x, & \text{if } b_{n,K}(s) - \tilde{z}_{\max}(s) \leq x < b_{n,K}(s) - \tilde{z}_{K-1}(s) \\ \tilde{z}_{\max}(s), & \text{if } 0 \leq x < b_{n,K}(s) - \tilde{z}_{\max}(s) \end{cases} \quad (20)$$

$$\begin{aligned} &V_m(\theta \cdot \check{x} + (1-\theta) \cdot \hat{x}, s) \\ &= V_m(\bar{x}, s) \\ &= \min_{\max\{0, d-\bar{x}\} \leq z \leq \tilde{z}_{\max}(s)} \{c(z, s) + h(\bar{x} + z - d) + \alpha \cdot \mathbb{E}[V_{m-1}(\bar{x} + z - d, S_{m-1}) | S_m = s]\} \\ &\leq \min_{\substack{\max\{0, d-\check{x}\} \leq \check{z} \leq \tilde{z}_{\max}(s) \\ \max\{0, d-\hat{x}\} \leq \hat{z} \leq \tilde{z}_{\max}(s)}} \left\{ \begin{aligned} &c(\theta \cdot \check{z} + (1-\theta) \cdot \hat{z}, s) + h(\bar{x} + \theta \cdot \check{z} + (1-\theta) \cdot \hat{z} - d) \\ &+ \alpha \cdot \mathbb{E}[V_{m-1}(\bar{x} + \theta \cdot \check{z} + (1-\theta) \cdot \hat{z} - d, S_{m-1}) | S_m = s] \end{aligned} \right\} \end{aligned} \quad (22)$$

$$\begin{aligned} &\leq \min_{\substack{\max\{0, d-\check{x}\} \leq \check{z} \leq \tilde{z}_{\max}(s) \\ \max\{0, d-\hat{x}\} \leq \hat{z} \leq \tilde{z}_{\max}(s)}} \left\{ \begin{aligned} &\theta \cdot c(\check{z}, s) + (1-\theta) \cdot c(\hat{z}, s) \\ &+ \theta \cdot h(\check{x} + \check{z} - d) + (1-\theta) \cdot h(\hat{x} + \hat{z} - d) \\ &+ \alpha \cdot \theta \cdot \mathbb{E}[V_{m-1}(\check{x} + \check{z} - d, S_{m-1}) | S_m = s] \\ &+ \alpha \cdot (1-\theta) \cdot \mathbb{E}[V_{m-1}(\hat{x} + \hat{z} - d, S_{m-1}) | S_m = s] \end{aligned} \right\} \end{aligned} \quad (23)$$

$$\begin{aligned} &= \theta \cdot \min_{\max\{0, d-\check{x}\} \leq \check{z} \leq \tilde{z}_{\max}(s)} \left\{ \begin{aligned} &c(\check{z}, s) + h(\check{x} + \check{z} - d) \\ &+ \alpha \cdot \mathbb{E}[V_{m-1}(\check{x} + \check{z} - d, S_{m-1}) | S_m = s] \end{aligned} \right\} \\ &+ (1-\theta) \cdot \min_{\max\{0, d-\hat{x}\} \leq \hat{z} \leq \tilde{z}_{\max}(s)} \left\{ \begin{aligned} &c(\hat{z}, s) + h(\hat{x} + \hat{z} - d) \\ &+ \alpha \cdot \mathbb{E}[V_{m-1}(\hat{x} + \hat{z} - d, S_{m-1}) | S_m = s] \end{aligned} \right\} \\ &= \theta \cdot V_m(\check{x}, s) + (1-\theta) \cdot V_m(\hat{x}, s) \end{aligned}$$



where  $\tilde{g}_m^+(b, s)$  represents the right derivative:

$$\tilde{g}_m^+(b, s) := \lim_{y \downarrow b} \frac{\tilde{g}_m(y, s) - \tilde{g}_m(b, s)}{y - b}$$

which is nondecreasing and continuous from the right, by the convexity of  $\tilde{g}_m(\cdot, s)$  [70, Section 24]. Note that  $\{b_{m,k}(s)\}_{k \in \{-1, 0, 1, \dots, K\}}$  is a nonincreasing sequence, because the sequence  $\{\tilde{c}_k(s)\}_{k \in \{0, 1, \dots, K\}}$  is nondecreasing. We show the optimal control action  $z_m^*(x, s)$  is then given by (20), by considering four exhaustive cases.

**Case 1:**  $b_{m,k}(s) - \tilde{z}_{k-1}(s) \leq x < b_{m,k-1}(s) - \tilde{z}_{k-1}(s), k \in \{0, 1, \dots, K\}$

In order to show  $z_m^*(x, s)$  is given by (20), it suffices to show

$$c^+(z, s) + \tilde{g}_m^+(x + z, s) < 0 \quad \text{for } \max\{0, d - x\} \leq z < \tilde{z}_{k-1}(s) \quad (24)$$

and

$$c^+(z, s) + \tilde{g}_m^+(x + z, s) \geq 0 \quad \text{for } \tilde{z}_{k-1}(s) \leq z \leq \tilde{z}_{\max}(s). \quad (25)$$

First, let  $z \in [\max\{0, d - x\}, \tilde{z}_{k-1}(s)]$  be arbitrary, and let  $j \in \{0, 1, \dots, k - 1\}$  be such that  $z \in [\tilde{z}_{j-1}(s), \tilde{z}_j(s))$ . If  $b_{m,k-1}(s) = d$ , then  $b_{m,k}(s) = d$ , as  $d \leq b_{m,k}(s) \leq b_{m,k-1}(s) = d$ . Yet,  $b_{m,k}(s) = b_{m,k-1}(s) = d$  implies  $d - \tilde{z}_{k-1}(s) \leq x < d - \tilde{z}_{k-1}(s)$ , which is vacuous. Therefore, we need only consider  $b_{m,k-1}(s) = \inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_{k-1}(s)\}$ . By the construction of the piecewise-linear function  $c(\cdot, s)$ ,  $z < \tilde{z}_{k-1}(s)$  implies

$$c^+(z, s) \leq \tilde{c}_{k-1}(s). \quad (26)$$

We also have

$$x + z < x + \tilde{z}_{k-1}(s) < b_{m,k-1}(s) = \inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_{k-1}(s)\}$$

which implies

$$\tilde{g}_m^+(x + z, s) < -\tilde{c}_{k-1}(s). \quad (27)$$

Summing (26) and (27) yields (24).

Next, let  $z \in [\tilde{z}_{k-1}(s), \tilde{z}_{\max}(s)]$  be arbitrary, so that by construction of  $c(\cdot, s)$

$$c^+(z, s) \geq \tilde{c}_k(s). \quad (28)$$

We also have

$$x + z \geq x + \tilde{z}_{k-1}(s) \geq b_{m,k}(s) \geq \inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_k(s)\}$$

which, in combination with the nondecreasing nature of  $\tilde{g}_m^+(\cdot, s)$ , implies

$$\tilde{g}_m^+(x + z, s) \geq \tilde{g}_m^+(\inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_k(s)\}, s). \quad (29)$$

Because  $\tilde{g}_m^+(\cdot, s)$  is continuous from the right

$$\tilde{g}_m^+(\inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_k(s)\}, s) \geq -\tilde{c}_k(s). \quad (30)$$

Combining (29) and (30), and summing with (28) yields (25).

**Case 2:**  $b_{m,k}(s) - \tilde{z}_k(s) \leq x < b_{m,k}(s) - \tilde{z}_{k-1}(s), k \in \{0, 1, \dots, K - 1\}$

In order to show  $z_m^*(x, s)$  is given by (20), it suffices to show:

$$c^+(z, s) + \tilde{g}_m^+(x + z, s) < 0 \quad \text{for } \max\{0, d - x\} \leq z < b_{m,k}(s) - x \quad (31)$$

and

$$c^+(z, s) + \tilde{g}_m^+(x + z, s) \geq 0 \quad \text{for } b_{m,k}(s) - x \leq z \leq \tilde{z}_{\max}(s). \quad (32)$$

First, let  $z \in [\max\{0, d - x\}, b_{m,k}(s) - x]$  be arbitrary. This case is vacuous if  $b_{m,k}(s) = d$ , so  $b_{m,k}(s) = \inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_k(s)\}$ . Thus, we have

$$x + z < b_{m,k}(s) = \inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_k(s)\}$$

which implies

$$\tilde{g}_m^+(x + z, s) < -\tilde{c}_k(s). \quad (33)$$

Furthermore, from  $z < b_{m,k}(s) - x \leq \tilde{z}_k(s)$  and the construction of the piecewise-linear function  $c(\cdot, s)$

$$c^+(z, s) \leq \tilde{c}_k(s). \quad (34)$$

Summing (33) and (34) yields (31).

Next, let  $z \in [b_{m,k}(s) - x, \tilde{z}_{\max}(s)]$  be arbitrary, so that  $z \geq b_{m,k}(s) - x > \tilde{z}_{k-1}(s)$ , which by the construction of the piecewise-linear function  $c(\cdot, s)$  implies:

$$c^+(z, s) \geq \tilde{c}_k(s). \quad (35)$$

We also have  $x + z \geq b_{m,k}(s) \geq \inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_k(s)\}$ . Therefore, because  $\tilde{g}_m^+(\cdot, s)$  is nondecreasing and continuous from the right

$$\tilde{g}_m^+(x + z, s) \geq \tilde{g}_m^+(\inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_k(s)\}, s) \geq -\tilde{c}_k(s). \quad (36)$$

Summing (35) and (36) yields (32).

**Case 3:**  $b_{m,K}(s) - \tilde{z}_{\max}(s) \leq x < b_{m,K}(s) - \tilde{z}_{K-1}(s)$

This case is the same as Case 2, with  $K$  in place of  $k$ , and  $\tilde{z}_{\max}(s)$  in place of  $\tilde{z}_k(s)$ .

**Case 4:**  $0 \leq x < b_{m,K}(s) - \tilde{z}_{\max}(s)$   
Let  $z \in [\max\{0, d - x\}, \tilde{z}_{\max}(s)]$  be arbitrary.  $\tilde{z}_{\max}(s) \geq d$  by assumption, so this case is vacuous if  $b_{m,K} = d$ . Thus, we have  $b_{m,K}(s) = \inf\{b \mid \tilde{g}_m^+(b, s) \geq -\tilde{c}_K(s)\}$ , which, in combination with  $x + z < x + \tilde{z}_{\max} < b_{m,K}(s)$ , implies

$$\tilde{g}_m^+(x + z, s) < -\tilde{c}_K(s). \quad (37)$$

Additionally,  $z < \tilde{z}_{\max}(s)$  implies

$$c^+(z, s) \leq \tilde{c}_K(s). \quad (38)$$

Summing (37) and (38) yields  $c^+(z, s) + \tilde{g}_m^+(x + z, s) < 0$  for all  $z \in [\max\{0, d - x\}, \tilde{z}_{\max}(s)]$ , which implies  $z_m^*(x, s) = \tilde{z}_{\max}(s)$ .  $\square$

### C. Proof of Theorem 5

We prove statements (i)–(v) by joint induction on the time remaining,  $n$ .

**Base Case:**  $n = 1$

$V_0(\mathbf{x}, \mathbf{s}_0) = 0$ , for all  $\mathbf{s}_0$ , so (i) and (ii) hold trivially. Let  $\mathbf{s}_1 \in \mathcal{S}$  be arbitrary.  $G_1(\mathbf{y}_1, \mathbf{s}_1) = \mathbf{c}_{\mathbf{s}_1}^T \mathbf{y}_1 + h(\mathbf{y}_1 - \mathbf{d})$ , which is convex and supermodular. Thus, (iii) and (iv) are true. Additionally

$$G_1(\mathbf{y}_1, \mathbf{s}_1) = \sum_{m=1}^2 \{c_s^m \cdot y_1^m + h^m (y_1^m - d^m)\}$$

so  $\inf\{\operatorname{argmin}_{y_1^2 \in [d^2, \infty)} \{G_1(y_1^1, y_1^2, s_1^1, s_1^2)\}\}$  is independent of  $y_1^1$ , and vice versa.

**Induction Step:**

Assume statements (i)–(v) are true for  $n = 2, 3, \dots, l-1$ . We want to show they are true for  $n = l$ . We let  $\mathbf{s} \in \mathcal{S}$  be arbitrary, and proceed in order.

Statement (i): Consider two arbitrary points,  $\bar{\mathbf{x}}, \tilde{\mathbf{x}} \in \mathbb{R}_+^2$ . Let  $\lambda \in [0, 1]$  be arbitrary, and define  $\hat{\mathbf{x}} := \lambda \bar{\mathbf{x}} + (1-\lambda)\tilde{\mathbf{x}}$ . Let  $\mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s})$ ,  $\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s})$ , and  $\mathbf{y}^*(\hat{\mathbf{x}}, \mathbf{s})$  be optimal buffer levels after transmission in slot  $l-1$ , for each of the respective starting points. We have

$$\begin{aligned} & \lambda \cdot V_{l-1}(\bar{\mathbf{x}}, \mathbf{s}) + (1-\lambda) \cdot V_{l-1}(\tilde{\mathbf{x}}, \mathbf{s}) \\ &= -\mathbf{c}_s^T \hat{\mathbf{x}} + \lambda \cdot G_{l-1}(\mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}), \mathbf{s}) \\ & \quad + (1-\lambda) \cdot G_{l-1}(\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}) \\ & \geq -\mathbf{c}_s^T \hat{\mathbf{x}} + G_{l-1}(\lambda \mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}) \\ & \geq -\mathbf{c}_s^T \hat{\mathbf{x}} + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\hat{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} \\ & = V_{l-1}(\hat{\mathbf{x}}, \mathbf{s}) = V_{l-1}(\lambda \bar{\mathbf{x}} + (1-\lambda)\tilde{\mathbf{x}}, \mathbf{s}) \end{aligned} \quad (39)$$

where the first inequality follows from the convexity of  $G_{l-1}(\cdot, \mathbf{s})$  from the induction hypothesis. The second inequality follows from the following argument.  $\mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})$  implies:

$$\mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) \succeq \mathbf{d} \vee \bar{\mathbf{x}} \quad \text{and} \quad \mathbf{c}_s^T [\mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) - \bar{\mathbf{x}}] \leq P. \quad (40)$$

Similarly,  $\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})$  implies

$$\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}) \succeq \mathbf{d} \vee \tilde{\mathbf{x}} \quad \text{and} \quad \mathbf{c}_s^T [\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}) - \tilde{\mathbf{x}}] \leq P. \quad (41)$$

Multiplying the equations in (40) by  $\lambda$  and the equations in (41) by  $1-\lambda$ , and summing, we have

$$\begin{aligned} & \lambda \mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}) \\ & \succeq \lambda(\mathbf{d} \vee \bar{\mathbf{x}}) + (1-\lambda)(\mathbf{d} \vee \tilde{\mathbf{x}}) \succeq \mathbf{d} \vee \hat{\mathbf{x}} \end{aligned} \quad (42)$$

and

$$\begin{aligned} & \mathbf{c}_s^T [\lambda \mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}) - \hat{\mathbf{x}}] \\ & = \lambda \mathbf{c}_s^T [\mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) - \bar{\mathbf{x}}] \\ & \quad + (1-\lambda) \mathbf{c}_s^T [\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}) - \tilde{\mathbf{x}}] \leq P. \end{aligned} \quad (43)$$

From (42) and (43), we conclude  $\lambda \mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\hat{\mathbf{x}}, \mathbf{s})$ , as shown in Fig. 7. Thus, the value of  $G_{l-1}(\cdot, \mathbf{s})$  at this point is greater than or equal to the minimum of  $G_{l-1}(\cdot, \mathbf{s})$  over the region  $\tilde{\mathcal{A}}^d(\hat{\mathbf{x}}, \mathbf{s})$ . From (39), we conclude  $V_{l-1}(\cdot, \mathbf{s})$  is convex. This argument is similar to the one used by Evans to show convexity in [62].

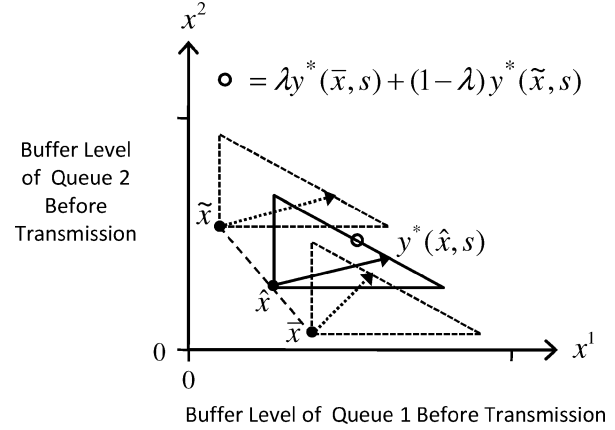


Fig. 7. Diagram showing  $\lambda \mathbf{y}^*(\bar{\mathbf{x}}, \mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\hat{\mathbf{x}}, \mathbf{s})$  in the proof of the convexity of  $V_{l-1}(\cdot, \mathbf{s})$ .

Statement (ii): Recall that  $V_{l-1}(\mathbf{x}, \mathbf{s}) = -\mathbf{c}_s^T \mathbf{x} + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\}$ . The first term,  $-\mathbf{c}_s^T \mathbf{x}$ , is clearly supermodular in  $\mathbf{x}$ , so it suffices to show that the second term,  $\min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\}$ , is also supermodular in  $\mathbf{x}$ . Let  $\bar{\mathbf{x}}, \tilde{\mathbf{x}} \in \mathbb{R}^2$  be arbitrary. We want to show

$$\begin{aligned} & \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} \\ & \leq \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\}. \end{aligned} \quad (44)$$

If  $\bar{\mathbf{x}}$  and  $\tilde{\mathbf{x}}$  are comparable (i.e.,  $\hat{x}^1 \geq \bar{x}^1$  and  $\hat{x}^2 \geq \bar{x}^2$  or  $\hat{x}^1 \leq \bar{x}^1$  and  $\hat{x}^2 \leq \bar{x}^2$ ), then (44) is trivial. So we assume they are not comparable, and also assume without loss of generality that  $\bar{x}^1 < \tilde{x}^1$  and  $\hat{x}^2 < \bar{x}^2$ . The main idea going forward is to cleverly construct—depending on the relative locations of  $\bar{\mathbf{x}}, \tilde{\mathbf{x}}, \mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$ , and  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ —points  $\bar{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})$  and  $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})$  such that

$$\begin{aligned} & G_{l-1}(\bar{\mathbf{y}}, \mathbf{s}) + G_{l-1}(\tilde{\mathbf{y}}, \mathbf{s}) \\ & \leq G_{l-1}(\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}) \\ & \quad + G_{l-1}(\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}) \\ & = \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} \\ & \quad + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\}. \end{aligned} \quad (45)$$

Then,  $\bar{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})$  and  $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})$  imply

$$\begin{aligned} & \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} \\ & \leq G_{l-1}(\bar{\mathbf{y}}, \mathbf{s}) + G_{l-1}(\tilde{\mathbf{y}}, \mathbf{s}). \end{aligned} \quad (46)$$

Combining (45) and (46) yields the desired result, (44). We proceed with a lemma on the relative locations of the points  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$  and  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ , and then construct  $\bar{\mathbf{y}}$  and  $\tilde{\mathbf{y}}$  to satisfy (45) for two exhaustive cases.

**Lemma 1:** There exist optimal buffer levels after transmission in slot  $l-1$ ,  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$  and  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ , such that  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \not\preceq \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ ; i.e., such that  $y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \leq y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$  or  $y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \leq y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ .

*Proof of Lemma 1:* Fix a choice of  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$  such that

$$G_{l-1}(\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}) = \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\}.$$

Assume that for all optimal choices of  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$ , we have  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \succ \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ , where we define  $\mathbf{a} \succ \mathbf{b}$  to mean  $a^1 > b^1$  and  $a^2 > b^2$ . Fix one such choice of  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$ , and we have:

$$\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \succ \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \succeq \mathbf{d} \vee (\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}). \quad (47)$$

Further,  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$  implies  $\mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}] \leq P$ , and thus

$$\begin{aligned} \mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \bar{\mathbf{x}} \vee \tilde{\mathbf{x}}] \\ \leq \mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}] \leq P. \end{aligned} \quad (48)$$

Equations (47) and (48) imply  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ , and thus

$$\begin{aligned} G_{l-1}(\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}) &= \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} \\ &\leq G_{l-1}(\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}). \end{aligned} \quad (49)$$

However, we also have

$$\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \succeq \mathbf{d} \vee (\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}) \succeq \mathbf{d} \vee (\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}), \quad (50)$$

and

$$\begin{aligned} \mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}] \\ \leq \mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}] \leq P. \end{aligned} \quad (51)$$

Equations (50) and (51) imply  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$ , which, in combination with (49), implies it is optimal to move from  $\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}$  to  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ , contradicting the assumption that  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \succ \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$  for all possible choices of  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$ .  $\square$

Now let  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$  and  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$  be arbitrary optimal actions such that  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \not\succeq \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ . We show (44) by considering two exhaustive cases.

**Case 1:**  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \succeq \mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$

We start with another lemma.

**Lemma 2:** Let  $f: [d^1, \infty) \times [d^2, \infty) \rightarrow \mathbb{R}$  be convex and supermodular, let  $\sigma, \beta \in [0, 1]$  be arbitrary, and let  $\mathbf{z} = (z_1, z_2) \preceq (\hat{z}_1, \hat{z}_2) = \hat{\mathbf{z}}$ . Define

$$\mathbf{z}^{\lambda_1, \lambda_2} := (\lambda_1 \hat{z}_1 + (1 - \lambda_1) z_1, \lambda_2 \hat{z}_2 + (1 - \lambda_2) z_2).$$

Then

$$f(\mathbf{z}) + f(\hat{\mathbf{z}}) \geq f(\mathbf{z}^{\sigma, \beta}) + f(\mathbf{z}^{1-\sigma, 1-\beta}). \quad (52)$$

*Proof of Lemma 2:* **Step 1:** Assume  $\sigma, \beta \leq \frac{1}{2}$ . Assume without loss of generality that  $\sigma \leq \beta$ . By the convexity of  $f(\cdot)$ , we have

$$f(\mathbf{z}) + f(\hat{\mathbf{z}}) \geq f(\mathbf{z}^{\sigma, \sigma}) + f(\mathbf{z}^{1-\sigma, 1-\sigma}) \quad (53)$$

and

$$f(\mathbf{z}^{1-\sigma, 1-\sigma}) + f(\mathbf{z}^{1-\sigma, \sigma}) \geq f(\mathbf{z}^{1-\sigma, \beta}) + f(\mathbf{z}^{1-\sigma, 1-\beta}). \quad (54)$$

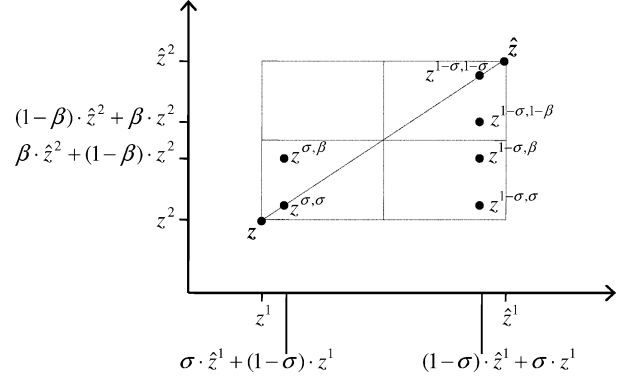


Fig. 8. Diagram of the points referred to in Step 1 of the proof of Lemma 2.

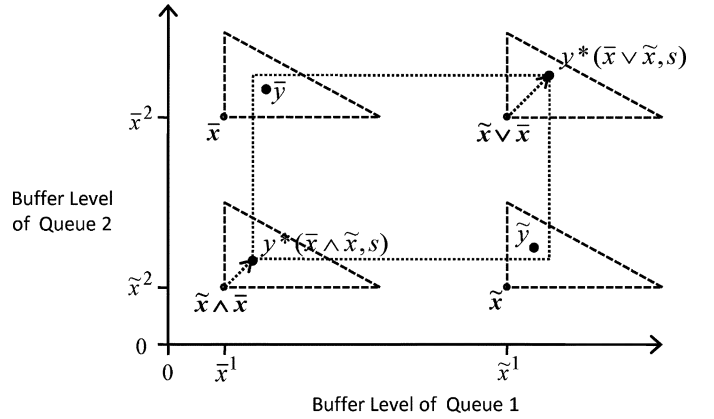


Fig. 9. Construction of feasible points  $\tilde{\mathbf{y}}$  and  $\bar{\mathbf{y}}$  in Case 1 of the proof of supermodularity of  $V_{l-1}(\cdot, \mathbf{s})$ .

By the supermodularity of  $f(\cdot)$ , we have

$$f(\mathbf{z}^{1-\sigma, \beta}) + f(\mathbf{z}^{\sigma, \sigma}) \geq f(\mathbf{z}^{\sigma, \beta}) + f(\mathbf{z}^{1-\sigma, \sigma}). \quad (55)$$

Fig. 8 shows these relationships. Combining (53)–(55), we have

$$\begin{aligned} f(\mathbf{z}) + f(\hat{\mathbf{z}}) &\geq f(\mathbf{z}^{\sigma, \sigma}) + f(\mathbf{z}^{1-\sigma, 1-\sigma}) \\ &\geq f(\mathbf{z}^{\sigma, \sigma}) + f(\mathbf{z}^{1-\sigma, \beta}) - f(\mathbf{z}^{1-\sigma, \sigma}) + f(\mathbf{z}^{1-\sigma, 1-\beta}) \\ &\geq f(\mathbf{z}^{\sigma, \beta}) + f(\mathbf{z}^{1-\sigma, 1-\beta}). \end{aligned}$$

**Step 2:** Now let  $\sigma, \beta \in [0, 1]$ , and define  $\hat{\sigma} := \min\{\sigma, 1 - \sigma\}$  and  $\hat{\beta} := \min\{\beta, 1 - \beta\}$ . Then  $\hat{\sigma}, \hat{\beta} \leq \frac{1}{2}$ , so by Step 1, we have

$$f(\mathbf{z}) + f(\hat{\mathbf{z}}) \geq f(\mathbf{z}^{\hat{\sigma}, \hat{\beta}}) + f(\mathbf{z}^{1-\hat{\sigma}, 1-\hat{\beta}}). \quad (56)$$

Note that  $\mathbf{z}^{\sigma, \beta} \wedge \mathbf{z}^{1-\sigma, 1-\beta} = \mathbf{z}^{\hat{\sigma}, \hat{\beta}}$ , and  $\mathbf{z}^{\sigma, \beta} \vee \mathbf{z}^{1-\sigma, 1-\beta} = \mathbf{z}^{1-\hat{\sigma}, 1-\hat{\beta}}$ , so by the supermodularity of  $f(\cdot)$ , we have

$$f(\mathbf{z}^{\hat{\sigma}, \hat{\beta}}) + f(\mathbf{z}^{1-\hat{\sigma}, 1-\hat{\beta}}) \geq f(\mathbf{z}^{\sigma, \beta}) + f(\mathbf{z}^{1-\sigma, 1-\beta}). \quad (57)$$

Combining (56) and (57) yields the desired result, (52).

Next, define the following points, shown in Fig. 9:

$$\begin{aligned} \bar{\mathbf{y}} &:= \left( \bar{x}^1 + \max\{y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \bar{x}^1, y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \bar{x}^1\} \right) \\ &\quad \left( \bar{x}^2 + \min\{y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \bar{x}^2, y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \bar{x}^2\} \right) \\ \tilde{\mathbf{y}} &:= \left( \tilde{x}^1 + \min\{y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \tilde{x}^1, y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \tilde{x}^1\} \right) \\ &\quad \left( \tilde{x}^2 + \max\{y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \tilde{x}^2, y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \tilde{x}^2\} \right). \end{aligned}$$

Note that  $\bar{\mathbf{y}} \succeq \mathbf{d} \vee \bar{\mathbf{x}}$  and  $\tilde{\mathbf{y}} \succeq \mathbf{d} \vee \tilde{\mathbf{x}}$ . Furthermore, we have the equations shown at the bottom of the page. By a similar argument,  $\mathbf{c}_s^T(\tilde{\mathbf{y}} - \tilde{\mathbf{x}}) \leq P$ , and thus  $\bar{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})$ , and  $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})$ . So (46) is true.

Now define<sup>12</sup>

$$\sigma := \frac{y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \tilde{y}^1}{y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})} \quad \text{and}$$

$$\beta := \frac{y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \tilde{y}^2}{y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})}.$$

Rearranging the definitions of  $\sigma$  and  $\beta$  yields

$$\tilde{\mathbf{y}} = \begin{pmatrix} (1 - \sigma) \cdot y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) + \sigma \cdot y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \\ (1 - \beta) \cdot y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) + \beta \cdot y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \end{pmatrix}.$$

It is also straightforward to check that

$$\bar{\mathbf{y}} = \begin{pmatrix} \sigma \cdot y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) + (1 - \sigma) \cdot y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \\ \beta \cdot y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) + (1 - \beta) \cdot y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \end{pmatrix}.$$

We also have

$$\begin{aligned} y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) &= \min \left\{ y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}), y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) + (\tilde{x}^1 - \hat{x}^1) \right\} \\ &\leq \min \left\{ y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}), y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) + (\tilde{x}^1 - \hat{x}^1) \right\} \\ &= \tilde{y}^1 \leq y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \end{aligned}$$

and thus,  $\sigma \in [0, 1]$ . Similarly,  $y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \leq \tilde{y}^2 \leq y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ , and thus,  $\beta \in [0, 1]$ . Since  $G_{l-1}(\cdot, \mathbf{s})$  is convex and supermodular, we can now apply Lemma 2, with  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$  playing the role of  $\mathbf{z}$ ;  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$  the role of  $\hat{\mathbf{z}}$ ;  $\bar{\mathbf{y}}$  the role of  $\mathbf{z}^{\sigma, \beta}$ ; and  $\tilde{\mathbf{y}}$  the role of  $\mathbf{z}^{1-\sigma, 1-\beta}$ , to get (45), which, in combination with (46), implies (44).

**Case 2:**  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \not\leq \mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \not\leq \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ .

There are two possibilities for this case. The first possibility is that  $y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) > y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$  and  $y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \leq y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ . The second possibility is that  $y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \leq y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$  and  $y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) > y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ . We show (44) under the first possibility, and a symmetric argument can be used to show (44) under the second possibility. We have

$$\begin{aligned} y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) &> y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \\ &\geq \max\{(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}})^1, d^1\} = \max\{\hat{x}^1, d^1\} \end{aligned} \quad (58)$$

<sup>12</sup>If  $y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) = 0$ , let  $\sigma$  be arbitrary in  $[0, 1]$ . Similarly, if  $y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) = 0$ , let  $\beta$  be arbitrary in  $[0, 1]$ .

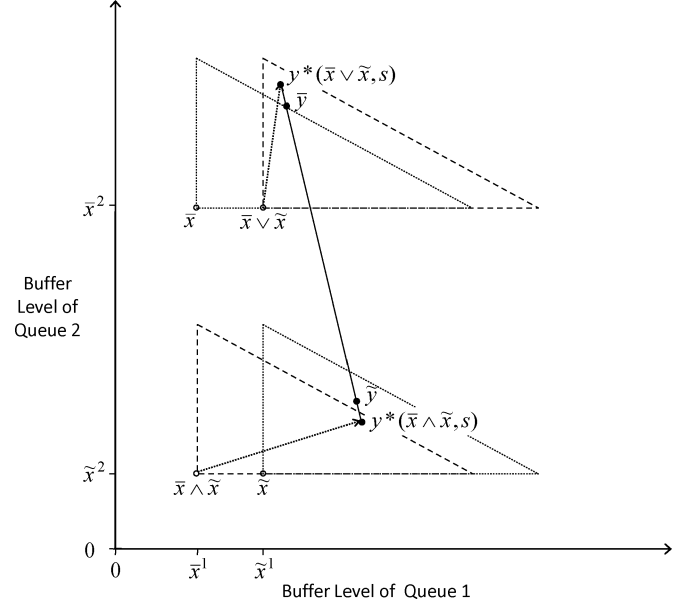


Fig. 10. Construction of feasible points  $\bar{\mathbf{y}}$  and  $\tilde{\mathbf{y}}$  in Case 2 of the proof of supermodularity of  $V_{l-1}(\cdot, \mathbf{s})$ .

$$\begin{aligned} y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) &\geq \max\{(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}})^2, d^2\} = \max\{\tilde{x}^2, d^2\} \end{aligned} \quad (59)$$

and

$$\begin{aligned} \mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \tilde{\mathbf{x}}] &\leq \mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - (\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}})] \leq P. \end{aligned} \quad (60)$$

Equations (58), (59), and (60) imply  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})$ . If it also happens that  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})$ , then we have

$$\begin{aligned} &\min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} \\ &\leq G_{l-1}(\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}) + G_{l-1}(\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}), \mathbf{s}) \\ &= \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\} + \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})} \{G_{l-1}(\mathbf{y}, \mathbf{s})\}. \end{aligned}$$

Otherwise, define

$$\gamma := \frac{\mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \bar{\mathbf{x}}] - P}{\mathbf{c}_s^T[\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})]}.$$

From  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \notin \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})$  and  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$ , we know

$$\begin{aligned} \mathbf{c}_s^T \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) &> \mathbf{c}_s^T \bar{\mathbf{x}} + P \geq \mathbf{c}_s^T(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}) + P \\ &\geq \mathbf{c}_s^T \mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}). \end{aligned} \quad (61)$$

$$\begin{aligned} \mathbf{c}_s^T(\bar{\mathbf{y}} - \bar{\mathbf{x}}) &= \mathbf{c}_s^T \begin{pmatrix} \max\{y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \bar{x}^1, y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \hat{x}^1\}, \\ \min\{y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \hat{x}^2, y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \bar{x}^2\} \end{pmatrix} \\ &\leq \max \left\{ \begin{array}{l} \mathbf{c}_s^T(y^{*1}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \bar{x}^1, y^{*2}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - \hat{x}^2), \\ \mathbf{c}_s^T(y^{*1}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \hat{x}^1, y^{*2}(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - \bar{x}^2) \end{array} \right\} \\ &= \max \left\{ \mathbf{c}_s^T(\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) - (\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}})), \mathbf{c}_s^T(\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) - (\bar{\mathbf{x}} \vee \tilde{\mathbf{x}})) \right\} \leq P \end{aligned}$$

It is clear from (61) that the numerator and denominator of  $\gamma$  are positive, and  $\gamma \in [0, 1]$ . Now define

$$\begin{aligned}\bar{\mathbf{y}} &:= \gamma \mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) + (1 - \gamma) \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) \quad \text{and} \\ \tilde{\mathbf{y}} &:= (1 - \gamma) \mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) + \gamma \mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}).\end{aligned}$$

It is somewhat tedious but straightforward to show that  $\bar{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})$  and  $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\tilde{\mathbf{x}}, \mathbf{s})$ , implying (46). In Fig. 10,  $\bar{\mathbf{y}}$  is the point where the line segment connecting  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$  and  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$  intersects the budget constraint (hypotenuse) of  $\tilde{\mathcal{A}}^d(\bar{\mathbf{x}}, \mathbf{s})$ , and  $\tilde{\mathbf{y}}$  is a point along this line segment the same distance from  $\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s})$  as  $\bar{\mathbf{y}}$  is from  $\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s})$ . Equation (45) follows from the convexity of  $G_{l-1}(\cdot, \mathbf{s})$  along this line segment. Combining (45) and (46) again yields the desired result, (44).

Statement (iii):  $G_l(\mathbf{y}, \mathbf{s}) = \mathbf{c}_s^T \mathbf{y} + h(\mathbf{y} - \mathbf{d}) + \alpha \cdot \mathbb{E}[V_{l-1}(\mathbf{y} - \mathbf{d}, \mathbf{S})]$ . By (i), for all  $\mathbf{s}$ ,  $V_{l-1}(\mathbf{x}, \mathbf{s})$  is convex in  $\mathbf{x}$ ; thus,  $V_{l-1}(\mathbf{y} - \mathbf{d}, \mathbf{s})$  is convex in  $\mathbf{y}$  as it is the composition of a convex function with an affine function.  $\mathbb{E}[V_{l-1}(\mathbf{y} - \mathbf{d}, \mathbf{S})]$  is also convex as it is the nonnegative weighted sum/integral of convex functions. It follows that  $G_l(\mathbf{y}, \mathbf{s})$ , the sum of convex functions, is convex in  $\mathbf{y}$ .

Statement (iv): Supermodularity of  $G_l(\mathbf{y}, \mathbf{s})$  follows from the same series of arguments as (iii), because, like convexity, supermodularity is preserved under addition and scalar multiplication.<sup>13</sup>

Statement (v): This step basically follows from Topkis' Theorem 2.8.1 [72, p. 76], but, for the reader's benefit, we reproduce the proof here with our notation. Let  $y^2, \hat{y}^2 \in [d^2, \infty)$  be arbitrary with  $y^2 < \hat{y}^2$ . Let  $\bar{y}^1 \in \operatorname{argmin}_{y^1 \in [d^1, \infty)} \{G_l(y^1, y^2, \mathbf{s})\}$  and  $\tilde{y}^1 \in \operatorname{argmin}_{y^1 \in [d^1, \infty)} \{G_l(y^1, \hat{y}^2, \mathbf{s})\}$  be arbitrary. We want to show

$$\bar{y}^1 \wedge \tilde{y}^1 \in \operatorname{argmin}_{y^1 \in [d^1, \infty)} \{G_l(y^1, \hat{y}^2, \mathbf{s})\}.$$

If  $\tilde{y}^1 \leq \bar{y}^1$ , this is trivial, so we check that it is true for  $\tilde{y}^1 > \bar{y}^1$ . Since  $\bar{y}^1$  is a minimizer of  $G_l(\cdot, y^2, \mathbf{s})$ , we have

$$G_l(\bar{y}^1, y^2, \mathbf{s}) \leq G_l(\tilde{y}^1, y^2, \mathbf{s}) \quad (62)$$

and since  $\tilde{y}^1$  is a minimizer of  $G_l(\cdot, \hat{y}^2, \mathbf{s})$ , we have

$$G_l(\tilde{y}^1, \hat{y}^2, \mathbf{s}) \leq G_l(\bar{y}^1, \hat{y}^2, \mathbf{s}). \quad (63)$$

By the supermodularity of  $G_l(\cdot, \mathbf{s})$ , we have

$$\begin{aligned}G_l(\tilde{y}^1, y^2, \mathbf{s}) + G_l(\bar{y}^1, \hat{y}^2, \mathbf{s}) \\ \leq G_l(\tilde{y}^1 \wedge \bar{y}^1, y^2 \wedge \hat{y}^2, \mathbf{s}) \\ + G_l(\tilde{y}^1 \vee \bar{y}^1, y^2 \vee \hat{y}^2, \mathbf{s}) \\ = G_l(\bar{y}^1, y^2, \mathbf{s}) + G_l(\tilde{y}^1, \hat{y}^2, \mathbf{s})\end{aligned}$$

or, rearranging terms

$$\begin{aligned}G_l(\tilde{y}^1, y^2, \mathbf{s}) - G_l(\bar{y}^1, y^2, \mathbf{s}) \\ \leq G_l(\tilde{y}^1, \hat{y}^2, \mathbf{s}) - G_l(\bar{y}^1, \hat{y}^2, \mathbf{s}). \quad (64)\end{aligned}$$

Combining (62), (63), and (64) yields

$$\begin{aligned}0 \leq G_l(\tilde{y}^1, y^2, \mathbf{s}) - G_l(\bar{y}^1, y^2, \mathbf{s}) \\ \leq G_l(\tilde{y}^1, \hat{y}^2, \mathbf{s}) - G_l(\bar{y}^1, \hat{y}^2, \mathbf{s}) \leq 0. \quad (65)\end{aligned}$$

<sup>13</sup> Smith and McCardle [71] refer to these as *closed convex cone properties*.

So (65) holds with equality throughout, implying  $G_l(\tilde{y}^1, \hat{y}^2, \mathbf{s}) = G_l(\bar{y}^1, \hat{y}^2, \mathbf{s})$ , and we conclude

$$\tilde{y}^1 \wedge \bar{y}^1 = \bar{y}^1 \in \operatorname{argmin}_{y^1 \in [d^1, \infty)} \{G_l(y^1, \hat{y}^2, \mathbf{s})\}.$$

Since  $\bar{y}^1$  and  $\tilde{y}^1$  were chosen arbitrarily, we have

$$\begin{aligned}\inf \left\{ \operatorname{argmin}_{y_n^1 \in [d^1, \infty)} \{G_n(y_n^1, y_n^2, s^1, s^2)\} \right\} \\ \geq \inf \left\{ \operatorname{argmin}_{y_n^1 \in [d^1, \infty)} \{G_n(y_n^1, \hat{y}_n^2, s^1, s^2)\} \right\}.\end{aligned}$$

The first implication in (v) follows from a symmetric argument.  $\square$

#### D. Proof of Theorem 6

Let  $n \in \{1, 2, \dots, N\}$  and  $\mathbf{s} \in \mathcal{S}$  be arbitrary. We start by proving (14). First, let  $\mathbf{x} \in \mathcal{R}_I(n, \mathbf{s})$  and  $\hat{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$  be arbitrary. We know from Theorem 5 that  $G_n(\cdot, \mathbf{s})$  is convex on  $[d^1, \infty) \times [d^2, \infty)$ , which implies that  $G_n(\cdot, \mathbf{s})$  is also convex on any line segment in  $[d^1, \infty) \times [d^2, \infty)$  (see, e.g., [70, Theorem 4.1]). Specifically, by the convexity of  $G_n(\cdot, \mathbf{s})$  along the line  $y^1 = \hat{y}^1$  and the fact that  $\hat{y}^2 \geq x^2 \geq f_n^2(\hat{y}^1, \mathbf{s})$ , we have

$$G_n(\hat{\mathbf{y}}, \mathbf{s}) \geq G_n((\hat{y}^1, x^2), \mathbf{s}) \geq G_n((\hat{y}^1, f_n^2(\hat{y}^1, \mathbf{s})), \mathbf{s}). \quad (66)$$

Similarly, by the convexity of  $G_n(\cdot, \mathbf{s})$  along the line  $y^2 = x^2$  and the fact that  $\hat{y}^1 \geq x^1 \geq f_n^1(x^2, \mathbf{s})$ , we have:

$$G_n((\hat{y}^1, x^2), \mathbf{s}) \geq G_n(\mathbf{x}, \mathbf{s}) \geq G_n((f_n^1(x^2, \mathbf{s}), x^2), \mathbf{s}). \quad (67)$$

Combining (66) and (67) yields

$$G_n(\hat{\mathbf{y}}, \mathbf{s}) \geq G_n((\hat{y}^1, x^2), \mathbf{s}) \geq G_n(\mathbf{x}, \mathbf{s})$$

and we conclude  $G_n(\mathbf{x}, \mathbf{s}) = \min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})} \{G_n(\mathbf{y}, \mathbf{s})\}$ .

Second, let  $\mathbf{x} \in \mathcal{R}_{II}(n, \mathbf{s})$  be arbitrary. Then  $\mathbf{b}_n(\mathbf{s}) \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$  and  $\mathbf{b}_n(\mathbf{s})$  is a global minimizer of  $G_n(\cdot, \mathbf{s})$ , so it is clearly optimal to transmit to bring the receivers' buffer levels up to  $\mathbf{b}_n(\mathbf{s})$ .

Next, let  $\mathbf{x} \in \mathcal{R}_{III-A}(n, \mathbf{s})$  and  $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$  be arbitrary. By definition of  $f_n^1(\cdot, \mathbf{s})$ , we have

$$G_n(\tilde{\mathbf{y}}, \mathbf{s}) \geq G_n((f_n^1(\tilde{y}^2, \mathbf{s}), \tilde{y}^2), \mathbf{s}). \quad (68)$$

Furthermore, the function  $\min_{y^1 \in [d^1, \infty)} \{G_n((y^1, y^2), \mathbf{s})\}$  is convex in  $y^2$  since  $[d^1, \infty)$  is a convex set (see, e.g., [69, pp. 101–102]). Thus,  $\tilde{y}^2 \geq x^2 \geq b_n^2(\mathbf{s})$  implies

$$\begin{aligned}G_n((f_n^1(\tilde{y}^2, \mathbf{s}), \tilde{y}^2), \mathbf{s}) \\ \geq G_n((f_n^1(x^2, \mathbf{s}), x^2), \mathbf{s}) \\ \geq G_n((f_n^1(b_n^2(\mathbf{s}), \mathbf{s}), b_n^2(\mathbf{s})), \mathbf{s}) \\ = G_n(\mathbf{b}_n(\mathbf{s}), \mathbf{s}). \quad (69)\end{aligned}$$

Combining (68) and (69) yields

$$G_n(\tilde{\mathbf{y}}, \mathbf{s}) \geq G_n((f_n^1(x^2, \mathbf{s}), x^2), \mathbf{s})$$

and  $\mathbf{x} \in \mathcal{R}_{III-A}(n, \mathbf{s})$  implies  $(f_n^1(x^2, \mathbf{s}), x^2) \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$ . Since  $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$  was arbitrary, we conclude  $\mathbf{y}_n^*(\mathbf{x}, \mathbf{s}) = (f_n^1(x^2, \mathbf{s}), x^2)$  is optimal.

The optimality of  $\mathbf{y}_n^*(\mathbf{x}, \mathbf{s}) = (x^1, f_n^2(x^1, \mathbf{s}))$  for  $\mathbf{x} \in \mathcal{R}_{III-B}(n, \mathbf{s})$  follows from a symmetric argument, using the convexity of  $G_n(\cdot, \mathbf{s})$  along the curve  $(x^1, f_n^2(x^1, \mathbf{s}))$ .

Finally, we prove (15). Define

$$\mathcal{H}^d(\mathbf{x}, \mathbf{s}) := \{ \mathbf{y} \in [d^1, \infty) \times [d^2, \infty) : \mathbf{y} \succeq \mathbf{x} \text{ and } \mathbf{c}_s^T[\mathbf{y} - \mathbf{x}] = P \} \subset \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s}).$$

First, let  $\mathbf{x} \in \mathcal{R}_{IV-B}(n, \mathbf{s})$  and  $\hat{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$  be arbitrary such that  $\mathbf{c}_s^T[\hat{\mathbf{y}} - \mathbf{x}] < P$ . Define

$$\lambda_0 := \frac{\mathbf{c}_s^T \mathbf{b}_n(\mathbf{s}) - \mathbf{c}_s^T \mathbf{x} - P}{\mathbf{c}_s^T \mathbf{b}_n(\mathbf{s}) - \mathbf{c}_s^T \hat{\mathbf{y}}}.$$

Note that  $\mathbf{c}_s^T[\hat{\mathbf{y}} - \mathbf{x}] < P$  and  $\mathbf{c}_s^T[\mathbf{b}_n(\mathbf{s}) - \mathbf{x}] > P$  imply  $\lambda_0 \in (0, 1)$ . Then define

$$\tilde{\mathbf{y}} := \lambda_0 \hat{\mathbf{y}} + (1 - \lambda_0) \mathbf{b}_n(\mathbf{s}).$$

By the convexity of  $G_n(\cdot, \mathbf{s})$  along the line segment from  $\hat{\mathbf{y}}$  to  $\mathbf{b}_n(\mathbf{s})$ , we have

$$G_n(\hat{\mathbf{y}}, \mathbf{s}) \geq G_n(\tilde{\mathbf{y}}, \mathbf{s}) \geq G_n(\mathbf{b}_n(\mathbf{s}), \mathbf{s}).$$

Since  $\hat{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$  was arbitrary, we conclude

$$\min_{\mathbf{y} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})} \{G_n(\mathbf{y}, \mathbf{s})\} = \min_{\mathbf{y} \in \mathcal{H}^d(\mathbf{x}, \mathbf{s})} \{G_n(\mathbf{y}, \mathbf{s})\}.$$

Next, let  $\mathbf{x} \in \mathcal{R}_{IV-C}(n, \mathbf{s})$  and  $\hat{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$  be arbitrary such that  $\mathbf{c}_s^T[\hat{\mathbf{y}} - \mathbf{x}] < P$ . We consider two exhaustive cases, and for each case, we construct a  $\tilde{\mathbf{y}} \in \mathcal{H}^d(\mathbf{x}, \mathbf{s})$  such that  $G_n(\tilde{\mathbf{y}}, \mathbf{s}) \leq G_n(\hat{\mathbf{y}}, \mathbf{s})$ .

**Case 1:**  $\hat{y}^2 < f_n^2(\hat{y}^1, \mathbf{s})$  and  $\bar{\mathbf{y}} := (\hat{y}^1, f_n^2(\hat{y}^1, \mathbf{s})) \notin \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$ . Let  $\tilde{\mathbf{y}} := (\hat{y}^1, x^2 + \frac{P - c_{s^1} \cdot [\hat{y}^1 - x^1]}{c_{s^2}})$ . Then, by the convexity of  $G_n(\cdot, \mathbf{s})$  along  $y^1 = \hat{y}^1$ , the definition of  $f_n^2(\hat{y}^1, \mathbf{s})$ , and  $\hat{y}^2 \leq \tilde{y}^2 \leq f_n^2(\hat{y}^1, \mathbf{s})$ , we have

$$G_n(\bar{\mathbf{y}}, \mathbf{s}) = G_n((\hat{y}^1, f_n^2(\hat{y}^1, \mathbf{s})), \mathbf{s}) \leq G_n(\tilde{\mathbf{y}}, \mathbf{s}) \leq G_n(\hat{\mathbf{y}}, \mathbf{s}).$$

It is also straightforward to check that  $\tilde{\mathbf{y}} \in \mathcal{H}^d(\mathbf{x}, \mathbf{s})$ , as desired.

**Case 2:** All other  $\hat{\mathbf{y}} \in \tilde{\mathcal{A}}^d(\mathbf{x}, \mathbf{s})$  such that  $\mathbf{c}_s^T[\hat{\mathbf{y}} - \mathbf{x}] < P$

By the definition of  $f_n^2(\hat{y}^1, \mathbf{s})$ , we have

$$G_n(\hat{\mathbf{y}}, \mathbf{s}) \geq G_n((\hat{y}^1, f_n^2(\hat{y}^1, \mathbf{s})), \mathbf{s}). \quad (70)$$

Define

$$\begin{aligned} \tilde{y}^1 &:= \sup \{ y^1 \in [x^1, \hat{y}^1) : \mathbf{c}_s^T(y^1, f_n^2(y^1, \mathbf{s})) \geq \mathbf{c}_s^T \mathbf{x} + P \} \\ \text{and} \\ \tilde{y}^2 &:= \frac{P - c_{s^1} \cdot [\tilde{y}^1 - x^1]}{c_{s^2}}. \end{aligned}$$

By the convexity of  $G_n(\cdot, \mathbf{s})$  along  $(y^1, f_n^2(y^1, \mathbf{s}))$ , we have

$$G_n((\hat{y}^1, f_n^2(\hat{y}^1, \mathbf{s})), \mathbf{s}) \geq G_n((\tilde{y}^1, f_n^2(\tilde{y}^1, \mathbf{s})), \mathbf{s}). \quad (71)$$

Furthermore, we have

$$G_n((\tilde{y}^1, f_n^2(\tilde{y}^1, \mathbf{s})), \mathbf{s}) = G_n((\tilde{y}^1, \tilde{y}^2), \mathbf{s}) = G_n(\tilde{\mathbf{y}}, \mathbf{s}). \quad (72)$$

If  $\tilde{y}^2 = f_n^2(\tilde{y}^1, \mathbf{s})$ , (72) is trivial. Otherwise, there is a discontinuity in  $f_n^2(\cdot, \mathbf{s})$  at  $\tilde{y}^1$ , and we have

$$\lim_{y^1 \nearrow \tilde{y}^1} f_n^2(y^1, \mathbf{s}) \geq \tilde{y}^2 \geq \lim_{y^1 \searrow \tilde{y}^1} f_n^2(y^1, \mathbf{s}) \quad (73)$$

with at least one of the inequalities being strict. Nonetheless,  $G_n((y^1, f_n^2(y^1, \mathbf{s})), \mathbf{s})$  is a continuous function of  $y^1$ , and therefore

$$\begin{aligned} &G_n\left(\left(\tilde{y}^1, \lim_{y^1 \nearrow \tilde{y}^1} f_n^2(y^1, \mathbf{s})\right), \mathbf{s}\right) \\ &= G_n\left(\left(\tilde{y}^1, \lim_{y^1 \searrow \tilde{y}^1} f_n^2(y^1, \mathbf{s})\right), \mathbf{s}\right) \\ &= G_n((\tilde{y}^1, f_n^2(\tilde{y}^1, \mathbf{s})), \mathbf{s}). \end{aligned} \quad (74)$$

The convexity of  $G_n(\cdot, \mathbf{s})$  along the line  $y^1 = \tilde{y}^1$  and (74) imply

$$\begin{aligned} G_n((\tilde{y}^1, y^2), \mathbf{s}) &= G_n((\tilde{y}^1, f_n^2(\tilde{y}^1, \mathbf{s})), \mathbf{s}), \\ \forall y^2 \in &\left[ \lim_{y^1 \searrow \tilde{y}^1} f_n^2(y^1, \mathbf{s}), \lim_{y^1 \nearrow \tilde{y}^1} f_n^2(y^1, \mathbf{s}) \right] \end{aligned}$$

which, in combination with (73), implies (72). Combining (70)–(72) yields the desired result:  $G_n(\tilde{\mathbf{y}}, \mathbf{s}) \leq G_n(\hat{\mathbf{y}}, \mathbf{s})$  for a  $\tilde{\mathbf{y}} \in \mathcal{H}^d(\mathbf{x}, \mathbf{s})$ .

The validity of (15) for  $\mathbf{x} \in \mathcal{R}_{IV-A}(n, \mathbf{s})$  follows from a symmetric argument, completing the proof of (15) and Theorem 6.

#### ACKNOWLEDGMENT

The authors would like to thank Aditya Mahajan, Ashutosh Nayyar, Demosthenis Teneketzis, and the anonymous reviewers for helpful suggestions.

#### REFERENCES

- [1] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Comm. Mag.*, vol. 39, pp. 150–154, Feb. 2001.
- [2] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.
- [3] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. Int. Conf. Comm.*, Seattle, WA, Jun. 1995, vol. 1, pp. 331–335.
- [4] F. Berggren and R. Jäntti, "Asymptotically fair transmission scheduling over fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 326–336, Jan. 2004.
- [5] X. Liu, E. K. P. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Comput. Netw.*, vol. 41, pp. 451–474, Mar. 2003.
- [6] J. M. Holtzman, "CDMA forward link waterfilling power control," in *Proc. IEEE Veh. Technol. Conf.*, Tokyo, Japan, May 2000, vol. 3, pp. 1663–1667.
- [7] S. Borst and P. Whiting, "Dynamic channel-sensitive scheduling algorithms for wireless data throughput optimization," *IEEE Trans. Veh. Technol.*, vol. 52, no. 5, pp. 569–586, May 2003.
- [8] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, and P. Whiting, "Scheduling in a queueing system with asynchronously varying service rates," *Prob. Eng. Inform. Sci.*, vol. 18, pp. 191–217, 2004.
- [9] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queues with randomly varying connectivity," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 466–478, Mar. 1993.
- [10] S. Shakkottai, R. Srikant, and A. Stolyar, "Pathwise optimality of the exponential scheduling rule for wireless channels," *Adv. Appl. Prob.*, vol. 36, no. 4, pp. 1021–1045, Dec. 2004.
- [11] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time varying wireless networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, Apr. 2003, vol. 1, pp. 745–755.

- [12] B. E. Collins and R. L. Cruz, "Transmission policies for time varying channels with average delay constraints," in *Proc. Allerton Conf. Comm., Control, and Computing*, Monticello, IL, 1999.
- [13] R. A. Berry and R. G. Gallager, "Communication over fading channels with delay constraints," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1135–1149, May 2002.
- [14] M. Goyal, A. Kumar, and V. Sharma, "Power constrained and delay optimal policies for scheduling transmission over a fading channel," in *Proc. IEEE INFOCOM*, San Francisco, CA, Mar.–Apr. 2003, pp. 311–320.
- [15] H. Wang, "Opportunistic Transmission of Wireless Data Over Fading Channels Under Energy and Delay Constraints," Ph.D. thesis, Rutgers University, New Brunswick, NJ, 2003.
- [16] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 125–144, Jan. 2004.
- [17] R. A. Berry and E. M. Yeh, "Cross-layer wireless resource allocation," *IEEE Signal Process. Mag.*, vol. 21, no. 9, pp. 59–69, Sep. 2004.
- [18] D. V. Djonin and V. Krishnamurthy, "Structural results on the optimal transmission scheduling policies and costs for correlated sources and channels," in *Proc. IEEE Conf. Dec. and Contr.*, Seville, Spain, Dec. 2005, pp. 3231–3236.
- [19] A. Borkar, A. Karandikar, and V. S. Borkar, "Power optimal opportunistic scheduling," in *Proc. IEEE Glob. Telecommun. Conf.*, San Francisco, CA, Nov. 2006.
- [20] S. Kittipiyakul and T. Javidi, "Resource allocation in OFDMA with time-varying channel and bursty arrivals," *IEEE Commun. Lett.*, vol. 11, pp. 708–710, Sep. 2007.
- [21] M. Agarwal, V. S. Borkar, and A. Karandikar, "Structural properties of optimal transmission policies over a randomly varying channel," *IEEE Trans. Autom. Contr.*, vol. 53, pp. 1476–1491, Jul. 2008.
- [22] M. Goyal, A. Kumar, and V. Sharma, "Optimal cross-layer scheduling of transmissions over a fading multiaccess channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3518–3537, Aug. 2008.
- [23] S. Kittipiyakul and T. Javidi, "Delay-optimal server allocation in multi-queue multi-server systems with time-varying connectivities," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2319–2333, May 2009.
- [24] W. Chen, U. Mitra, and M. J. Neely, "Energy-efficient scheduling with individual delay constraints over a fading channel," in *Proc. Int. Symp. Model. and Optim. in Mobile, Ad Hoc, and Wirel. Netw.*, Limassol, Cyprus, Apr. 2007.
- [25] W. Chen, U. Mitra, and M. J. Neely, "Energy-efficient scheduling with individual packet delay constraints over a fading channel," *Wireless Netw.*, vol. 15, pp. 601–618, July 2009.
- [26] A. Fu, E. Modiano, and J. Tsitsiklis, "Optimal energy allocation for delay-constrained data transmission over a time-varying channel," in *Proc. IEEE INFOCOM*, San Francisco, CA, Apr. 2003, vol. 2, pp. 1095–1105.
- [27] A. Fu, E. Modiano, and J. N. Tsitsiklis, "Optimal transmission scheduling over a fading channel with energy and deadline constraints," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 630–641, Mar. 2006.
- [28] J. Lee and N. Jindal, "Energy-efficient scheduling of delay constrained traffic over fading channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008.
- [29] J. Lee and N. Jindal, "Energy-efficient scheduling of delay constrained traffic over fading channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1866–1875, Apr. 2009.
- [30] J. Lee and N. Jindal, "Delay constrained scheduling over fading channels: Optimal policies for monomial energy-cost functions," in *Proc. IEEE Int. Conf. Comm.*, Dresden, Germany, Jun. 2009.
- [31] J. Lee and N. Jindal, "Asymptotically optimal policies for hard-deadline scheduling over fading channels," *IEEE Trans. Inf. Theory*, 2009, submitted for publication.
- [32] E. Uysal-Biyikoglu and A. E. Gamal, "On adaptive transmission for energy efficiency in wireless data networks," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3081–3094, Dec. 2004.
- [33] A. Tarello, J. Sun, M. Zafer, and E. Modiano, "Minimum energy transmission scheduling subject to deadline constraints," *Wireless Netw.*, vol. 14, pp. 633–645, Oct. 2008.
- [34] D. I. Shuman and M. Liu, "Opportunistic scheduling with deadline constraints in wireless networks," in *Performance Models and Risk Management in Communication Systems*, N. Gulpinar, P. Harrison, and B. Rustem, Eds. New York: Springer, 2010.
- [35] X. Liu, E. K. P. Chong, and N. B. Shroff, "Optimal opportunistic scheduling in wireless networks," in *Proc. IEEE Veh. Technol. Conf.*, Orlando, FL, Oct. 2003, vol. 3, pp. 1417–1421.
- [36] X. Liu, N. B. Shroff, and E. K. P. Chong, "Opportunistic scheduling: An illustration of cross-layer design," *Telecommun. Rev.*, vol. 14, pp. 947–959, Dec. 2004.
- [37] B. Girod, M. Kalman, Y. J. Liang, and R. Zhang, "Advances in channel-adaptive video streaming," in *Proc. Int. Conf. Image Process.*, Rochester, NY, Sep. 2002, vol. 1, pp. 9–12.
- [38] Y. Li and N. Bambos, "Power-controlled wireless links for media streaming applications," in *Proc. Wireless Telecommun. Symp.*, Pomona, CA, May 2004, pp. 102–111.
- [39] Y. Li, A. Markopoulou, N. Bambos, and J. Apostolopoulos, "Joint power-playout control for media streaming over wireless links," *IEEE Trans. Multimedia*, vol. 8, no. 8, pp. 830–843, Aug. 2006.
- [40] 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 J. Select Areas Commun.*, vol. 21, no. 12, pp. 1710–1720, Dec. 2003.
- [41] J. Rust, "Using randomization to break the curse of dimensionality," *Econometrica*, vol. 65, pp. 487–516, May 1997.
- [42] C.-S. Chow and J. N. Tsitsiklis, "The complexity of dynamic programming," *J. Complexity*, vol. 5, pp. 466–488, Dec. 1989.
- [43] O. Hernández-Lerma and J. B. Lasserre, *Discrete-Time Markov Control Processes*. New York: Springer-Verlag, 1996.
- [44] D. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*. New York: Athena Scientific, 1996.
- [45] D. I. Shuman, M. Liu, and O. Q. Wu, "Energy-Efficient Transmission Scheduling With Strict Underflow Constraints," Univ. Michigan, Ann Arbor, EECs Tech. Rep. CGR 10-01, 2010 [Online]. Available: <http://arxiv.org/abs/0908.1774>
- [46] D. I. Shuman, "From Sleeping to Stockpiling: Energy Conservation via Stochastic Scheduling in Wireless Networks," Ph.D. dissertation, Univ. Michigan, Ann Arbor, 2010.
- [47] T. Fabian, J. L. Fisher, M. W. Sasieni, and A. Yardeni, "Purchasing raw material on a fluctuating market," *Oper. Res.*, vol. 7, pp. 107–122, Jan.–Feb. 1959.
- [48] B. Kalyon, "Stochastic prices in a single-item inventory purchasing model," *Oper. Res.*, vol. 19, pp. 1434–1458, Oct. 1971.
- [49] B. G. Kingsman, "Commodity purchasing," *Oper. Res. Quart.*, vol. 20, pp. 59–79, Mar. 1969.
- [50] B. G. Kingsman, "Commodity Purchasing in Uncertain Fluctuating Price Markets," Ph.D. dissertation, University of Lancaster, Lancaster, U.K., 1969.
- [51] V. F. Magirou, "Stockpiling under price uncertainty and storage capacity constraints," *Eur. J. Oper. Res.*, vol. 11, no. 3, pp. 233–246, Nov. 1982.
- [52] V. F. Magirou, "Comments on 'on optimal inventory policies when ordering prices are random' by Kamal Golabi," *Oper. Res.*, vol. 35, pp. 930–931, Nov.–Dec. 1987.
- [53] K. Golabi, "A single-item inventory model with stochastic prices," in *Proc. 2nd Int. Symp. Inventories*, Budapest, Hungary, 1982, pp. 687–697.
- [54] K. Golabi, "Optimal inventory policies when ordering prices are random," *Oper. Res.*, vol. 33, pp. 575–588, May–Jun. 1985.
- [55] A. Federgruen and P. Zipkin, "An inventory model with limited production capacity and uncertain demands II. The discounted-cost criterion," *Math. Oper. Res.*, vol. 11, pp. 208–215, May 1986.
- [56] S. Tayur, "Computing the optimal policy for capacitated inventory models," *Comm. Statist. Stochast. Models*, vol. 9, no. 4, pp. 585–598, 1993.
- [57] M. J. Sobel, "Making short-run changes in production when the employment level is fixed," *Oper. Res.*, vol. 18, pp. 35–51, Jan.–Feb. 1970.
- [58] A. Bensoussan, M. Crouhy, and J.-M. Proth, *Mathematical Theory of Production Planning*. New York: Elsevier Science, 1983.
- [59] F. C. Zahn, "Studies of Inventory Control and Capacity Planning With Multiple Sources," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, Aug. 2009.
- [60] E. L. Porteus, "Stochastic inventory theory," in *Stochastic Models*, D. P. Heyman and M. J. Sobel, Eds. New York: Elsevier Science, 1990, pp. 605–652.
- [61] S. Karlin, "Optimal inventory policy for the Arrow-Harris-Marschak dynamic model," in *Studies in the Mathematical Theory of Inventory and Production*, K. J. Arrow, S. Karlin, and H. Scarf, Eds. Stanford, CA: Stanford Univ. Press, 1958, pp. 135–154.
- [62] R. Evans, "Inventory control of a multiproduct system with a limited production resource," *Naval Res. Logist. Quart.*, vol. 14, no. 2, pp. 173–184, 1967.
- [63] G. A. DeCroix and A. Arreola-Risa, "Optimal production and inventory policy for multiple products under resource constraints," *Manag. Sci.*, vol. 44, pp. 950–961, Jul. 1998.
- [64] S. X. Chen, "The optimality of hedging point policies for stochastic two-product flexible manufacturing systems," *Oper. Res.*, vol. 52, pp. 312–322, Mar.–Apr. 2004.

- [65] G. Janakiraman, M. Nagarajan, and S. Veeraraghavan, "Simple policies for managing flexible capacity," *Management Sci.*, submitted for publication.
- [66] R. Bellman, I. Glicksberg, and O. Gross, "On the optimal inventory equation," *Manag. Sci.*, vol. 2, pp. 83–104, Oct. 1955.
- [67] J. T. Hawkins, "A Lagrangian Decomposition Approach to Weakly Coupled Dynamic Optimization Problems and its Applications," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2003.
- [68] D. Adelman and A. J. Mersereau, "Relaxations of weakly coupled stochastic dynamic programs," *Oper. Res.*, vol. 56, pp. 712–727, May–Jun. 2008.
- [69] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, MA: Cambridge Univ. Press, 2004.
- [70] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton Univ. Press, 1970.
- [71] J. E. Smith and K. F. McCardle, "Structural properties of stochastic dynamic programs," *Oper. Res.*, vol. 50, no. 5, pp. 796–809, 2002.
- [72] D. M. Topkis, *Supermodularity and Complementarity*. Princeton, NJ: Princeton Univ. Press, 1998.

**David I Shuman** (M'10) received the B.A. degree in economics and the M.S. degree in engineering-economic systems and operations research from Stanford University, Stanford, CA, in 2001 and the M.S. degree in electrical engineering: systems, the M.S. degree in applied mathematics, and the Ph.D. degree in electrical engineering: systems from the University of Michigan, Ann Arbor, in 2006, 2009, and 2010, respectively.

Currently, he is a Postdoctoral Researcher at the Institute of Electrical Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. His research interests include stochastic control, stochastic scheduling and resource allocation problems, energy-efficient design of wireless communication networks, and inventory theory.

**Mingyan Liu** (M'00) received the B.Sc. degree in electrical engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1995 and the M.Sc. degree in systems engineering and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 1997 and 2000, respectively.

She joined the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, in September 2000, where she is currently an Associate Professor. Her research interests are in optimal resource allocation, performance modeling and analysis, and energy efficient design of wireless, mobile *ad hoc*, and sensor networks. Dr. Liu is the recipient of the 2002 National Science Foundation (NSF) CAREER Award, and the University of Michigan Elizabeth C. Crosby Research Award in 2003. She serves on the editorial board of the IEEE/ACM TRANSACTIONS ON NETWORKING and the IEEE TRANSACTIONS ON MOBILE COMPUTING.

**Owen Q. Wu** received the M.S. degree in industrial engineering and engineering management from the Hong Kong University of Science and Technology in 2001 and the Ph.D. degree in business administration (management science) from the University of British Columbia, Vancouver, BC, Canada, in 2006.

He joined the Ross School of Business at the University of Michigan, Ann Arbor, as an Assistant Professor in 2006. His research interests include operations management of energy resources, interfaces of operations and energy markets, and game theory applications in energy supply contracts.