Adapted from Shuman, D. I, Liu, M., and Wu, O. Q., "Energy-efficient transmission scheduling with strict underflow constraints," IEEE Trans. Inf. Theory, Mar. 2011, pp. 1344-1367

Energy-Efficient Wireless Transmission Scheduling as an Inventory Control Problem with Stochastic Ordering Costs

David I Shuman
Institute of Electrical Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, david.shuman@epfl.ch

Mingyan Liu
Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI, mingyan@eecs.umich.edu

Owen Q. Wu
Ross School of Business, University of Michigan, Ann Arbor, MI, owenwu@bus.umich.edu

We make a novel connection between an energy-efficient wireless transmission scheduling problem and a multi-period, multi-item inventory control problem with stochastic convex ordering costs, deterministic demands, and a joint budget constraint. In the special case of a single item, we show that a state-dependent modified base-stock policy is optimal when the stochastic ordering costs are linear, and a state-dependent finite generalized base-stock policy is optimal when the stochastic ordering costs are piecewise-linear convex. We also present an efficient method to compute the target stock-up levels characterizing these policies when certain technical conditions are satisfied. For the case of two items with a joint budget constraint, we state and prove the structure of the optimal policy, which is somewhat surprisingly not a base-stock policy. We show that due to the stochastic nature of the ordering costs, structural phenomena may arise that are not possible in a two-item inventory model with a joint budget constraint and the more classical configuration of stochastic demands and deterministic ordering costs.

Key words: Wireless media streaming, underflow constraints, opportunistic scheduling, energy-delay tradeoff, resource allocation, dynamic programming, stochastic inventory theory, base-stock policy.

1. Introduction

Transporting multimedia over wireless networks is a promising application that has seen recent advances. At the same time, certain resource allocation issues need to be addressed in order to provide high quality and efficient media over wireless. First, streaming applications tend to have stringent quality of service (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.

When the sender is a battery-powered mobile device, energy-efficiency is crucial to extending its lifetime. Miao et al. (2009) explain succinctly why this problem is unlikely to disappear anytime soon: "Although silicon technology is progressing exponentially, doubling every 2 years, processor power consumption is also increasing by 150% every 2 years. In contrast, the improvement in battery technology is much slower, increasing a modest 10% every 2 years, leading to an exponentially increasing gap between the demand for energy and the battery capacity offered."

When the media comes from a base station, 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. Energy costs represent up to half of a mobile operator's operating expenses (Ericsson 2008), and Kremling (2008) estimates these energy costs will increase six-fold between 2002 and 2012. The trend is similar in mobile video, where the number of subscribers is expected to grow 37% annually from 2008 to 2014 (Pyramid Research 2009). Aleaf (2007) estimates mobile video in the U.S. will consume 17,850 gigawatt-hours in 2010, which is equivalent to the power output of two typical nuclear reactors. Most of this power consumption takes place at the base stations.

In this paper, we examine the problem of energy-efficient transmission scheduling over a wireless channel. 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*). In the context of media streaming, enforcing these strict underflow constraints reduces playout interruptions to the end users.

1.1. 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 wireless channel. The scheduler can exploit the *temporal diversity* of the channel by sending more data in the time slots when a user's channel is in a "good" state (requiring less power per data packet), and less data when the channel is in a "bad" state. 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*.

Sending more data when the channel is in a good state can increase system throughput and/or reduce total energy consumption; however, an important competing QoS consideration 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." Tassiulas and Ephremides (1993), Neely et al. (2003), and Andrews et al. (2004) examine stability in different settings. A second notion of delay is the *average* end-to-end delay (i.e., the time between a packet's arrival at the sender's buffer and its decoding by the receiver) of all packets over a long horizon. Collins and Cruz (1999) and Berry and Gallager (2002) are two of the many works that consider average delay, either as a constraint or by incorporating it directly into the objective function. However, the average delay criterion allows for the possibility of long delays; thus, *strict* end-to-end delay constraints, such as those considered by Luna et al. (2003) and Chen et al. (2009), are more appropriate for delay-sensitive applications.

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 received. Fu et al. (2006) and Lee and Jindal (2009a,b) 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. Fu et al. (2006) 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. Lee and Jindal (2009a,b) model the power-rate curve under each channel condition as convex, first in 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.[1] Uysal-Biyikoglu and Gamal (2004) and Tarello et al. (2008) consider opportunistic scheduling problems with multiple receivers and a single deadline constraint at the end of a finite horizon. The model of Tarello et al. (2008) 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 Tarello et al. (2008); and (ii) the transmitter can transmit to at most one receiver in each time slot in Tarello et al. (2008).

[1] In our notation of Section 2.1, 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 ζ is a fixed parameter.

In our model, the strict 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 3 and 4 of this paper generalize the works of Fu et al. (2006) and Lee and Jindal (2009a,b), respectively, by considering multiple deadlines in the point-to-point communication problem, rather than a single deadline at the end of the horizon.

1.2. 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. 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* (Chow and Tsitsiklis 1989, Rust 1997). 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 make a novel connection between the wireless transmission scheduling problem described above and a multi-period, multi-item inventory control problem with stochastic ordering costs, deterministic demands, and a joint budget constraint. Because the inventory models corresponding to our wireless communications models have not been examined previously, our results also represent a contribution to the inventory theory literature. Specifically, we (i) establish that in the case of a single receiver under linear power-rate curves, the optimal policy is an easily-implementable *modified base-stock* policy featuring target levels that depend on the current channel condition;[2] (ii) show that the optimal policy for a single receiver with piecewise-linear convex power-rate curves is a state-dependent *finite generalized base-stock* policy; (iii) provide an efficient method to recursively calculate the critical numbers for both of these single-receiver models, when certain technical conditions are satisfied; and (iv) state and prove the structure of the optimal policy for the case of a single sender transmitting to two receivers over a shared wireless channel, and show how the peak power constraint in each slot couples the optimal scheduling of the two receivers' packet streams. To our knowledge, contribution (iii) represents the first exact computation of the critical numbers

[2] We use the terms target level and critical number interchangeably throughout the paper.

for an optimal finite generalized base-stock policy, and (iv) represents the first structural result on any multi-item inventory model with a joint resource constraint and stochastic ordering costs.

Certain features of the wireless communications model result in technical challenges from an inventory theory perspective, but others afford a means to develop new results. The power-rate curves in the transmission scheduling problem are often convex, corresponding to convex ordering costs. Structural results exist for classical inventory problems with stochastic demands and deterministic, convex ordering costs; however, we are not aware of any methods to exactly compute the optimal decisions. The deterministic demands in our model, which are quite reasonable in the streaming context, allow us to exactly compute the optimal decisions when the channel condition is independently and identically distributed, the holding costs are linear, and the slopes of the power-rate curves change at integer multiples of the demand. This result is also important for computing good suboptimal policies when the power-rate curves are convex, but not piecewise-linear.

The time-varying stochastic channel conditions (which correspond to time-varying stochastic ordering costs) also provide a new technical challenge for multiple-item inventory theory. Namely, the value functions in our dynamic programming equations lack functional properties that lead to additional structure in the optimal policy for inventory models with deterministic ordering costs. For the case of two items, the functional property of interest is μ-difference monotonicity, which we connect to submodularity with respect to the direct value partial orders.[3] To highlight the importance of these absent functional properties, we provide an example of the two-item inventory problem with stochastic ordering costs whose optimal policy exhibits counterintuitive structural phenomena that cannot appear in the two-item inventory problem with stochastic demands and deterministic ordering costs. Namely, when both inventories start below their target levels, the unique optimal decision calls for bringing the inventory of one item past its target level while keeping the inventory of the other item below its target level. This example therefore shows that while a base-stock policy (in the sense of Porteus (1990))[4] is optimal for the two-item inventory model with stochastic demands, a joint budget constraint, and deterministic ordering costs, it may be strictly suboptimal for our two-item inventory model with deterministic demands, a joint budget constraint, and stochastic ordering costs. Moreover, the absence of submodularity with respect to the direct value orders necessitates new proof techniques for the property of supermodularity (which both problems share) and the structure of the optimal policy. We supply these techniques in the proofs of our two-item structural results of Section 5.

[3] See Appendix D for definitions and origins of these properties and partial orders.

[4] (Porteus 1990, p. 626) writes: "We formally say that a policy is a *base stock policy* if it is not possible to get at least as close to the base stock levels for every product and strictly closer for at least one item."

1.3. Organization of the Paper

In the next section, we describe the system model, formulate finite and infinite horizon MDPs, and relate our model to models in inventory theory. In Sections 3 and 4, we consider a single receiver under linear power-rate curves and piecewise-linear convex power-rate curves, respectively. We analyze the structure of the optimal policy when there are two receivers with linear power-rate curves in Section 5. Section 6 compares our model to a two-item inventory model with deterministic ordering costs, Section 7 discusses extensions, and Section 8 concludes the paper.

2. 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 more general multi-user (multi-receiver) case, so that we can discuss the most general case in Section 7.3.

2.1. 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.[5] 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 Figure 1 for a diagram of the system.

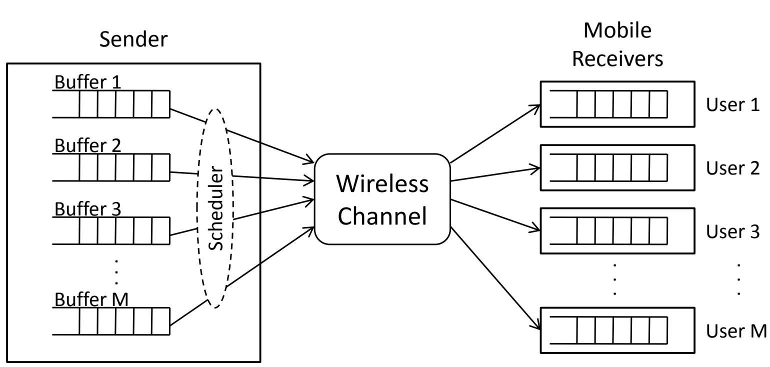

Figure 1 System model.

[5] This assumption is commonly referred to as the infinite backlog assumption.

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

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, $\{S_n^m\}_{n=N,N-1,\ldots,1}$ with state space $\mathcal{S}^m$.[6] Conditioned on the channel state, S_n^m, at time n, user m's channel states at *future* times $(n-1, n-2, \ldots)$ are independent of the channel states at *past* times $(n+1, n+2, \ldots)$. 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 3.2 and 4.2.

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.

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 $\mathbf{d} = (d^1, d^2, \ldots, d^M)$. This latter assumption is less realistic, but may be justified if the receiving buffers are drained at a constant rate by the media player. It is

[6] 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.

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 3.1. 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 7.2.

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.

2.2. Notation

Before proceeding, we introduce some notation. We define $I\!R_+ := [0, \infty)$ and $I\!N := \{1, 2, \ldots\}$. A single dot, as in $a \cdot b$, represents scalar multiplication. We use bold font to denote column vectors, such as $\mathbf{w} = (w^1, w^2, \ldots, w^M)$. We include a transpose superscript whenever a vector is meant to be a row vector, such as $\mathbf{w}^{\mathrm{T}}$. 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:

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

2.3. 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 $\mathbf{X}_n = (X_n^1, X_n^2, \cdots, X_n^M)$ denotes the receiver buffer queue lengths at time n, and $\mathbf{S}_n = (S_n^1, S_n^2, \cdots, S_n^M)$ 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 $n = N, N-1, \ldots, 1$, 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:

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

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 $\lim_{x\to\infty} h^m(x) = \infty$. 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:

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

which could represent a finite receiver buffer of length μ.[7]

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

$$J_{N,\alpha}^{\pi} := I\!\!E^{\pi}\left\{\sum_{t=1}^{N}\sum_{m=1}^{M} \alpha^{N-t} \cdot \left\{c^m\left(Z_t^m, S_t^m\right) + h^m\left(X_t^m + Z_t^m - d^m\right)\right\} \mid \mathcal{F}_N\right\} , \tag{1}$$

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 $J_{\infty,\alpha}^{\pi} := \lim_{N\to\infty} J_{N,\alpha}^{\pi}$. For Problem (**P3**), the average expected cost function for minimization is defined as $J_{\infty,1}^{\pi} := \limsup_{N\to\infty} \frac{1}{N} J_{N,1}^{\pi}$. In all three cases, we allow the transmission policy $\boldsymbol{\pi}$ to be chosen from the set of all history-dependent randomized and deterministic control laws, $\boldsymbol{\Pi}$ (e.g. Hernández-Lerma and Lasserre 1996, Definition 2.2.3, p. 15).

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

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

[7] Taking μ 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**).

Problem (**P1**) may be solved using standard dynamic programming (e.g. Bertsekas and Shreve 1996, Hernández-Lerma and Lasserre 1996). The recursive dynamic programming equations are given by:[8]

$$
\begin{aligned}
V_n(\mathbf{x},\mathbf{s}) &= \min_{\mathbf{z}\in\mathcal{A}^{\mathbf{d}}(\mathbf{x},\mathbf{s})} \left\{ \begin{array}{l} \sum_{m=1}^{M} \{c^m(z^m,s^m) + h^m(x^m+z^m-d^m)\} \\ +\alpha\cdot I\!\!E\big[V_{n-1}(\mathbf{x}+\mathbf{z}-\mathbf{d},\mathbf{S}_{n-1}) \,\big|\, \mathbf{S}_n=\mathbf{s}\big] \end{array} \right\}, \\
&\qquad\qquad\qquad\qquad n = N, N-1, \ldots, 1 \qquad (3) \\
V_0(\mathbf{x},\mathbf{s}) &= 0, \quad \forall \mathbf{x}\in I\!\!R_+^M, \forall \mathbf{s}\in\mathcal{S} := \mathcal{S}^1\times\mathcal{S}^2\times\ldots\times\mathcal{S}^M,
\end{aligned}
$$

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

$$
\mathcal{A}^{\mathbf{d}}(\mathbf{x},\mathbf{s}) := \left\{ \mathbf{z}\in I\!\!R_+^M : \mathbf{z}\succeq \max\{\mathbf{0},\mathbf{d}-\mathbf{x}\} \text{ and } \sum_{m=1}^{M} c^m(z^m,s^m) \le P \right\}, \quad \forall \mathbf{x}\in I\!\!R_+^M, \forall \mathbf{s}\in\mathcal{S}. \qquad (4)
$$

The maximum in (4) is taken element-by-element (i.e., $z^m \ge \max\{0, d^m - z^m\}\ \forall m$). 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., $\sum_{m=1}^{M} c^m(d^m,s^m) \le P,\ \forall \mathbf{s}\in\mathcal{S}$) ensures that the action space $\mathcal{A}^{\mathbf{d}}(\mathbf{x},\mathbf{s})$ is always non-empty.

2.4. Relation to Inventory Theory

The model outlined in Section 2.1 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 joint 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 playout.

This particular problem has not been studied in the context of inventory theory, but similar problems have been examined, and some of the techniques from the inventory theory literature are useful in analyzing our model. Fabian et al. (1959), Kingsman (1969a,b), Kalymon (1971), Magirou (1982), Golabi (1982, 1985), Gavirneni (2004), and Berling and Martínez de Albéniz (2011) 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

[8] As shown in Appendix C of Shuman (2010), our model satisfies the measurable selection condition 3.3.3 of Hernández-Lerma and Lasserre 1996, p. 28, justifying the use of min rather than inf in the dynamic programming equations.

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 (1969a,b) is the only one 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 inventory models with deterministic linear ordering costs and stochastic demand, Federgruen and Zipkin (1986) and Tayur (1993) 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. Sobel (1970), Bensoussan et al. (1983), and Zahrn (2009) 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., 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 (1990) 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 Karlin (1958). 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. Evans (1967), DeCroix and Arreola-Risa (1998), Chen (2004), and Janakiraman et al. (2009) consider multi-item inventory systems under deterministic ordering costs, stochastic demand, and joint resource constraints. We discuss related results from these studies in more detail in Section 6.

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.

3. 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 Figure 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.

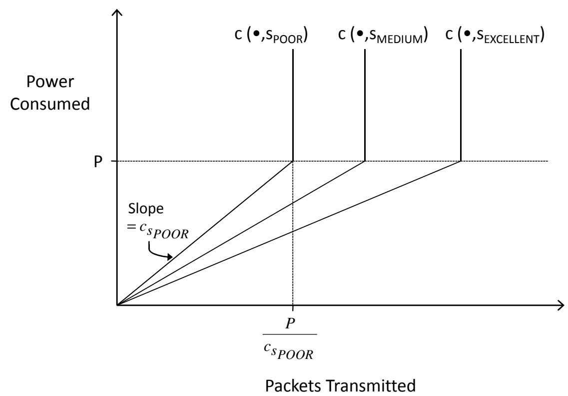

Figure 2 **A family of linear power-rate functions. Due to the power constraint, the effective power-rate function, shown above for each of the three channel conditions, is a two-segment piecewise-linear convex function. When the channel condition is s, the slope of the first segment is c_s.**

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

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

With these notations in place, the dynamic program (3) for Problem (**P1**) becomes:

$$V_n(x,s) = \min_{\max\{0,d-x\}\leq z\leq \frac{P}{c_s}} \left\{ \begin{array}{l} c_s \cdot z + h(x+z-d) \\ +\alpha \cdot I\!E\big[V_{n-1}(x+z-d,S_{n-1}) \,\big|\, S_n = s\big] \end{array} \right\} \tag{5}$$

$$= \min_{\max\{x,d\}\leq y\leq x+\frac{P}{c_s}} \left\{ \begin{array}{l} c_s \cdot (y-x) + h(y-d) \\ +\alpha \cdot I\!E\big[V_{n-1}(y-d,S_{n-1}) \,\big|\, S_n = s\big] \end{array} \right\} \tag{6}$$

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

$$V_0(x,s) = 0, \ \forall x \in I\!R_+, \forall s \in \mathcal{S},$$

where $g_n(y,s) := c_s \cdot y + h(y-d) + \alpha \cdot I\!E\big[V_{n-1}(y-d,S_{n-1}) \mid S_n = s\big]$. Here, the transition from (5) to (6) is done by the standard inventory theory 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* playout takes place (i.e., before d packets are removed from the buffer).

3.1. Structure of the Optimal Policy

With the above change of variable in the action space, the function g_n does not depend on the current buffer level, x. Without an upper bound on the action space (i.e., without the power constraint), a state-dependent base-stock policy would be optimal. Instead, a state-dependent modified base-stock policy is optimal, as stated in the following theorem.

THEOREM 1. *For every $n \in \{1,2,\ldots,N\}$ 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:*

$$y_n^*(x,s) := \begin{cases} x, & \text{if } x \geq b_n(s) \\ b_n(s), & \text{if } b_n(s) - \frac{P}{c_s} \leq x < b_n(s) \\ x + \frac{P}{c_s}, & \text{if } x < b_n(s) - \frac{P}{c_s} \end{cases} , \tag{7}$$

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

$$z_n^*(x,s) := \begin{cases} 0, & \text{if } x \geq b_n(s) \\ b_n(s) - x, & \text{if } b_n(s) - \frac{P}{c_s} \leq x < b_n(s) \\ \frac{P}{c_s}, & \text{if } x < b_n(s) - \frac{P}{c_s} \end{cases} . \tag{8}$$

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

$$N \cdot d \geq b_N(s) \geq b_{N-1}(s) \geq \ldots \geq b_1(s) = d . \tag{9}$$

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:

$$n \cdot d \geq b_n(s_{\text{best}}) \geq b_n(s^1) \geq b_n(s^2) \geq b_n(s_{\text{worst}}) = d . \tag{10}$$

The optimal transmission policy in Theorem 1 can be interpreted as follows. 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 Figure 3 for diagrams of the optimal policy.

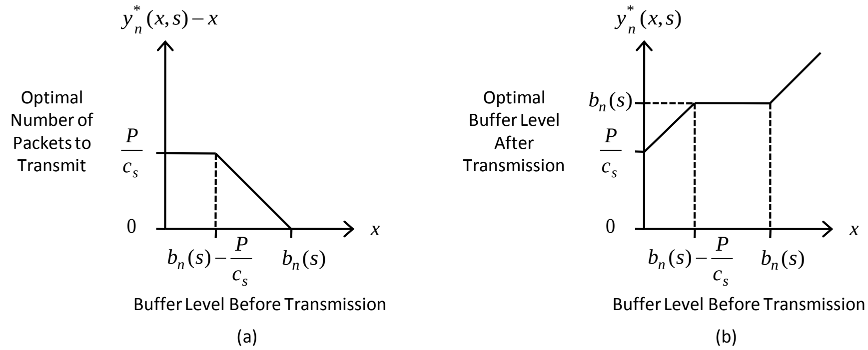

Figure 3 **Optimal policy in slot n when the state is (x,s). (a) depicts the optimal transmission quantity, and (b) depicts the resulting number of packets available for playout in slot n.**

Details of the proof of Theorem 1 are included in Appendix A of Shuman (2010). The key realization is that for all n and all s, $g_n(\cdot,s):[d,\infty)\to I\!R_+$ is a convex function in y, with $\lim_{y\to\infty} g_n(y,s)=\infty$. 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 (9), 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, $\{d_N, d_{N-1},\ldots,d_1\}$ (with $d_n\le\frac{P}{c_{\max}}$ for all n), then the structure of the optimal policy is still as stated in (7). If the channel is IID, then the following statement, analogous to (10), is true for arbitrary $s^1,s^2\in\mathcal{S}$ with $c_{s^1}\le c_{s^2}$:

$$\sum_{i=1}^{n} d_i \ge b_n(s_{\text{best}}) \ge b_n(c_{s^1}) \ge b_n(c_{s^2}) \ge b_n(s_{\text{worst}}) = d_n\ ,\ \forall n\in\{1,2,\ldots,N\}\ . \tag{11}$$

However, (9), the monotonicity of critical numbers over time for a fixed channel condition, is not true in general under nonstationary deterministic demand. As one counterexample, (11) 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, $\{b_n(s_{\text{worst}})\}_{n=1,2,\ldots,N}$, is not monotonic.

3.2. 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 $h\ge 0$), and the following technical condition is satisfied: for each possible channel condition s, $\frac{P}{c_s}=l\cdot d$ for some $l\in I\!N$; 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 $n \in \{1,2,\ldots,N\}$ and $j \in I\!N$ recursively, as follows:*

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

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

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

$$\gamma_{n,j} = -h + \alpha \cdot \left(\begin{array}{c} \displaystyle\sum_{s:\ c_s \geq \gamma_{n-1,j-1}} p(s) \cdot \gamma_{n-1,j-1} + \sum_{s:\ c_s < \gamma_{n-1,j-1}} p(s) \cdot c_s \\ \displaystyle + \sum_{s:\ c_s < \gamma_{n-1,j-1+L(s)}} p(s) \cdot \left[\gamma_{n-1,j-1+L(s)} - c_s \right] \end{array} \right), \qquad (12)$$

*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 $n \in \{1,2,\ldots,N\}$ 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 $\boldsymbol{\pi}^* = \left\{ y_N^*, y_{N-1}^*, \ldots, y_1^* \right\}$, where $y_n^*(x,s)$ is defined in* (7).[9]

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 $\{\gamma_{n,j}\}_{n \in \{1,2,\ldots,N\},\ j \in I\!N}$. 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 (see Rust 1997 and Chow and Tsitsiklis 1989). An intuitive discussion of the recursion (12) is included in Appendix A, and Theorem 2 is a special case of Theorem 4, the detailed proof of which can be found in Appendix A of Shuman (2010).

3.3. 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. 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.

[9] With n slots remaining, $0 = \gamma_{n,n+1} \leq \gamma_{n,n} \leq \gamma_{n,n-1} \leq \ \ldots\ \leq \gamma_{n,2} \leq \gamma_{n,1} = \infty$, so $b_n(s)$ is well-defined.

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 Shuman and Liu (2010).

4. 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 3.

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:

$$\begin{aligned}
&c(z,s) = z \cdot \tilde{c}_0(s) + \sum_{k=0}^{K-1} \Big\{ \big(\tilde{c}_{k+1}(s) - \tilde{c}_k(s)\big) \cdot \max\big\{z - \tilde{z}_k(s), 0\big\} \Big\} \text{ , where} \\
&0 < \tilde{c}_0(s) \leq \tilde{c}_1(s) \leq \cdots \leq \tilde{c}_K(s) \text{ , and} \\
&0 = \tilde{z}_{-1}(s) < \tilde{z}_0(s) < \tilde{z}_1(s) < \cdots < \tilde{z}_{K-1}(s) < \tilde{z}_K(s) = \infty \ .
\end{aligned} \tag{13}$$

The terms $\{\tilde{c}_k(s)\}_{k\in\{0,1,\ldots,K\}}$ represent the slopes of the segments of $c(\cdot,s)$, and the terms $\{\tilde{z}_k(s)\}_{k\in\{0,1,\ldots,K-1\}}$ 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 Figure 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 $\tilde{z}_{\max}(s) > \tilde{z}_{K-1}(s)$, $\forall s \in \mathcal{S}$.

In this case, the dynamic program (3) for Problem (**P1**) becomes:

$$\begin{aligned}
V_n(x,s) &= \min_{\big\{\max\{0,d-x\} \leq z \leq \tilde{z}_{\max}(s)\big\}} \Big\{ c(z,s) + \tilde{g}_n(x+z,s) \Big\}, \quad n = N, N-1, \ldots, 1 \\
V_0(x,s) &= 0, \ \forall x \in I\!\!R_+, \forall s \in \mathcal{S} \ ,
\end{aligned} \tag{14}$$

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

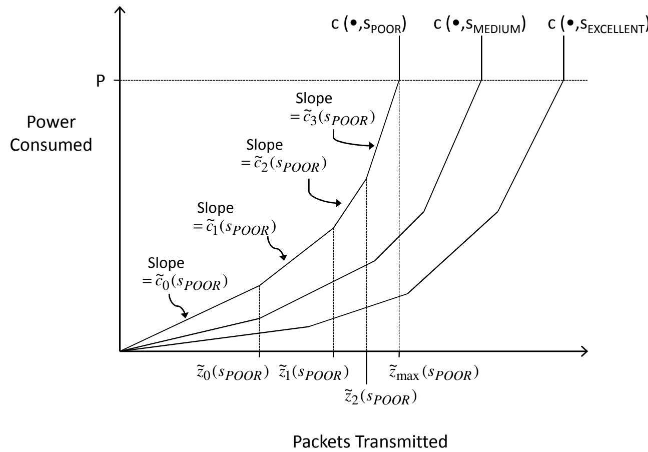

Figure 4 **A family of piecewise-linear convex power-rate functions. Like Figure 2, we incorporate the power constraint into each curve to show the effective power-rate curve. As an example, the power-rate function $c(\cdot, s_{POOR})$ is completely characterized by the sequence of slopes $\{\tilde{c}_k(s_{POOR})\}_{k\in\{0,1,2,3\}}$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

4.1. Structure of the Optimal Policy

With piecewise-linear power-rate curves, 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 Figure 3, but a more general piecewise-linear nondecreasing (respectively, nonincreasing) function.

THEOREM 3. *In Problem (**P1**) with a single receiver under piecewise-linear convex power-rate curves, for every $n \in \{1,2,\ldots,N\}$ and $s \in \mathcal{S}$, there exists a nonincreasing sequence of critical numbers $\{b_{n,k}(s)\}_{k\in\{0,1,\ldots,K\}}$ such that the optimal number of packets to transmit with n slots remaining is given by:*

$$z_n^*(x,s) := \begin{cases} \tilde{z}_{k-1}(s), & \text{if } \; b_{n,k}(s) - \tilde{z}_{k-1}(s) < x \le b_{n,k-1}(s) - \tilde{z}_{k-1}(s)\;, \\ & \qquad\qquad\qquad\qquad k \in \{0,1,\ldots,K\} \\ b_{n,k}(s) - x, & \text{if } \; b_{n,k}(s) - \tilde{z}_k(s) < x \le b_{n,k}(s) - \tilde{z}_{k-1}(s)\;, \\ & \qquad\qquad\qquad\qquad k \in \{0,1,\ldots,K-1\} \\ b_{n,K}(s) - x, & \text{if } \; b_{n,K}(s) - \tilde{z}_{\max}(s) < x \le b_{n,K}(s) - \tilde{z}_{K-1}(s) \\ \tilde{z}_{\max}(s), & \text{if } \; 0 \le x \le b_{n,K}(s) - \tilde{z}_{\max}(s) \end{cases}, \tag{15}$$

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 Figure 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 Figure 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.

4.2. 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 3.2 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

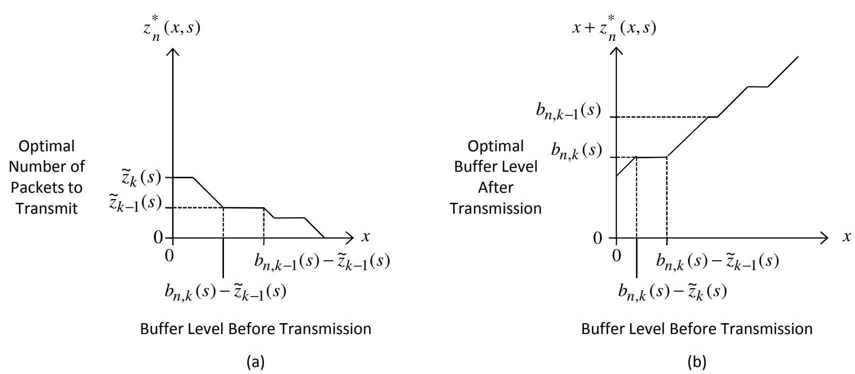

Figure 5 **Optimal transmission policy in slot n when the state is (x,s). (a) depicts the optimal transmission quantity, and (b) depicts the resulting number of packets available for playout in slot n.**

$s \in \mathcal{S}$: $\tilde{z}_{\max}(s) = \tilde{l}_{\max} \cdot d$ for some $\tilde{l}_{\max} \in I\!N$, and for every $k \in \{0, 1, \ldots, K-1\}$, $\tilde{z}_k(s) = \tilde{l}_k \cdot d$ for some $\tilde{l}_k \in I\!N$; i.e., the slopes of the effective power-rate functions only change at integer multiples of the drainage rate d.

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

THEOREM 4. *Define the thresholds $\tilde{\gamma}_{n,j}$ for $n \in \{1, 2, \ldots, N\}$ and $j \in I\!N$ recursively, as follows:*

(i) If $j = 1$, $\tilde{\gamma}_{n,j} = \infty$;

(ii) If $j > n$, $\tilde{\gamma}_{n,j} = 0$;

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

$$\tilde{\gamma}_{n,j} = -h + \alpha \cdot \left(\begin{array}{l} \displaystyle\sum_{s:\ \tilde{c}_0(s) \geq \tilde{\gamma}_{n-1,j-1}} p(s) \cdot \tilde{\gamma}_{n-1,j-1} \\ \quad + \displaystyle\sum_{k=0}^{K-1} \left\{ \begin{array}{l} \displaystyle\sum_{s:\ \tilde{\gamma}_{n-1,j-1+\tilde{L}_k(s)} \leq \tilde{c}_k(s) < \tilde{\gamma}_{n-1,j-1+\tilde{L}_{k-1}(s)}} p(s) \cdot \tilde{c}_k(s) \\ \quad + \displaystyle\sum_{s:\ \tilde{c}_k(s) < \tilde{\gamma}_{n-1,j-1+\tilde{L}_k(s)} \leq \tilde{c}_{k+1}(s)} p(s) \cdot \tilde{\gamma}_{n-1,j-1+\tilde{L}_k(s)} \end{array} \right\} \\ \quad + \displaystyle\sum_{s:\ \tilde{\gamma}_{n-1,j-1+\tilde{L}_{\max}(s)} \leq \tilde{c}_K(s) < \tilde{\gamma}_{n-1,j-1+\tilde{L}_{K-1}(s)}} p(s) \cdot \tilde{c}_K(s) \\ \quad + \displaystyle\sum_{s:\ \tilde{c}_K(s) < \tilde{\gamma}_{n-1,j-1+\tilde{L}_{\max}(s)}} p(s) \cdot \tilde{\gamma}_{n-1,j-1+\tilde{L}_{\max}(s)} \end{array} \right), \quad (16)$$

*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 $k \in \{0, 1, \ldots, K-1\}$, and $\tilde{L}_{\max}(s) := \frac{\tilde{z}_{\max}(s)}{d}$ for all $s \in \mathcal{S}$. For each $n \in \{1, 2, \ldots, N\}$ and $s \in \mathcal{S}$, define $b_{n,-1}(s) := \infty$ and for all $k \in \{0, 1, \ldots, K\}$, 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 $\boldsymbol{\pi}^* = \left\{z_N^*, z_{N-1}^*, \ldots, z_1^*\right\}$, where for all $n \in \{N, N-1, \ldots, 1\}$, $z_n^*(x, s)$ is given by* (15).

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 (16) is similar to the detailed explanation of (12) given in Appendix A, and a detailed proof of Theorem 4 is included in Appendix A of Shuman (2010).

4.3. General Convex Power-Rate Curves

As mentioned in Section 2.1, 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.

5. 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, $\{S_n^m\}_{n=N,N-1,\ldots,1}$. As discussed in Sections 1 and 2, 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 3, 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 $\mathbf{c}_\mathbf{s}^{\mathrm{T}}$, so that the total power consumption in slot n is given by $\sum_{m=1}^{2} c^m(Z_n^m, S_n^m) = \mathbf{c}_\mathbf{s}^{\mathrm{T}}\mathbf{Z}_n$. 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 (3) for Problem (**P1**) becomes:

$$
\begin{aligned}
V_n(\mathbf{x},\mathbf{s}) &= \min_{\mathbf{z}\in\mathcal{A}^{\mathbf{d}}(\mathbf{x},\mathbf{s})} \left\{ \begin{array}{l} \mathbf{c}_\mathbf{s}^{\mathrm{T}}\mathbf{z} + h(\mathbf{x}+\mathbf{z}-\mathbf{d}) \\ +\alpha\cdot I\!\!E\left[V_{n-1}(\mathbf{x}+\mathbf{z}-\mathbf{d},\mathbf{S}_{n-1}) \,\middle|\, \mathbf{S}_n=\mathbf{s}\right] \end{array} \right\} && (17) \\
&= \min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\mathbf{x},\mathbf{s})} \left\{ \begin{array}{l} \mathbf{c}_\mathbf{s}^{\mathrm{T}}[\mathbf{y}-\mathbf{x}] + h(\mathbf{y}-\mathbf{d}) \\ +\alpha\cdot I\!\!E\left[V_{n-1}(\mathbf{y}-\mathbf{d},\mathbf{S}_{n-1}) \,\middle|\, \mathbf{S}_n=\mathbf{s}\right] \end{array} \right\} && (18) \\
&= -\mathbf{c}_\mathbf{s}^{\mathrm{T}}\mathbf{x} + \min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\mathbf{x},\mathbf{s})} \left\{ G_n(\mathbf{y},\mathbf{s}) \right\} \qquad n=N,N-1,\ldots,1\ , \\
V_0(\mathbf{x},\mathbf{s}) &= 0, \quad \forall \mathbf{x}\in I\!\!R_+^2, \forall \mathbf{s}\in\mathcal{S}:=\mathcal{S}^1\times\mathcal{S}^2,
\end{aligned}
$$

where

$$
\begin{aligned}
&G_n(\mathbf{y},\mathbf{s}) := \mathbf{c}_\mathbf{s}^{\mathrm{T}}\mathbf{y} + h(\mathbf{y}-\mathbf{d}) + \alpha\cdot I\!\!E\left[V_{n-1}(\mathbf{y}-\mathbf{d},\mathbf{S}_{n-1}) \,\middle|\, \mathbf{S}_n=\mathbf{s}\right],\ \forall \mathbf{y}\in[d^1,\infty)\times[d^2,\infty), \forall \mathbf{s}\in\mathcal{S},\ \text{and} \\
&\tilde{\mathcal{A}}^{\mathbf{d}}(\mathbf{x},\mathbf{s}) := \left\{ \mathbf{y}\in I\!\!R_+^2 : \mathbf{y}\succeq \mathbf{d}\vee\mathbf{x} \text{ and } \mathbf{c}_\mathbf{s}^{\mathrm{T}}[\mathbf{y}-\mathbf{x}] \le P \right\},\quad \forall \mathbf{x}\in I\!\!R_+^2, \forall \mathbf{s}\in\mathcal{S}.
\end{aligned}
$$

The transition from (17) to (18) follows again from a change of variable in the action space from $\mathbf{Z}_n$ to $\mathbf{Y}_n$, where $\mathbf{Y}_n = \mathbf{X}_n + \mathbf{Z}_n$.

Without the per slot peak power constraint, this two-dimensional problem could be separated into two instances of the one-dimensional problem of Section 3; however, the joint power constraint couples the queues.[10] 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.

5.1. Structure of the Optimal Policy

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

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

(i) $V_{n-1}(\boldsymbol{x}, \boldsymbol{s})$ is convex in $\boldsymbol{x}$.

(ii) $V_{n-1}(\boldsymbol{x}, \boldsymbol{s})$ is supermodular in $\boldsymbol{x}$; i.e., for all $\bar{\boldsymbol{x}}, \tilde{\boldsymbol{x}} \in I\!R_+^2$,

$$V_{n-1}(\bar{\boldsymbol{x}}, \boldsymbol{s}) + V_{n-1}(\tilde{\boldsymbol{x}}, \boldsymbol{s}) \leq V_{n-1}(\bar{\boldsymbol{x}} \wedge \tilde{\boldsymbol{x}}, \boldsymbol{s}) + V_{n-1}(\bar{\boldsymbol{x}} \vee \tilde{\boldsymbol{x}}, \boldsymbol{s}) \ .$$

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

[10] This problem therefore falls into the class of *weakly coupled stochastic dynamic programs*, discussed in Hawkins (2003) and Adelman and Mersereau (2008).

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

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

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

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

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

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

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* $n \in \{1,2,\ldots,N\}$ *and* $\boldsymbol{s}\in\mathcal{S}^1\times\mathcal{S}^2$*, define the nonempty set of global minimizers of* $G_n(\cdot,\boldsymbol{s})$*:*

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

Define also

$$\begin{aligned} b_n^1(\boldsymbol{s}) &:= \min\left\{y^1\in[d^1,\infty):(y^1,y^2)\in\mathcal{B}_n(\boldsymbol{s}) \text{ for some } y^2\in[d^2,\infty)\right\} \;, \text{ and} \\ b_n^2(\boldsymbol{s}) &:= \min\left\{y^2\in[d^2,\infty):\left(b_n^1(\boldsymbol{s}),y^2\right)\in\mathcal{B}_n(\boldsymbol{s})\right\} \;. \end{aligned}$$

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

$$\begin{aligned} f_n^1(x^2,\boldsymbol{s}) &:= \inf\left\{\underset{y^1\in[d^1,\infty)}{\arg\min}\left\{G_n\left(y^1,x^2,s^1,s^2\right)\right\}\right\}, \text{ for } x^2\in[d^2,\infty), \text{ and} \\ f_n^2(x^1,\boldsymbol{s}) &:= \inf\left\{\underset{y^2\in[d^2,\infty)}{\arg\min}\left\{G_n\left(x^1,y^2,s^1,s^2\right)\right\}\right\}, \text{ for } x^1\in[d^1,\infty). \end{aligned}$$

Note that by construction, $f_n^1\left(b_n^2(\boldsymbol{s}),\boldsymbol{s}\right) = b_n^1(\boldsymbol{s})$ *and* $f_n^2\left(b_n^1(\boldsymbol{s}),\boldsymbol{s}\right) = b_n^2(\boldsymbol{s})$*. Partition* $I\!R_+^2$ *into the following seven regions:*

$$\begin{aligned} \mathcal{R}_I(n,\boldsymbol{s}) &:= \left\{\boldsymbol{x}\in I\!R_+^2 : \boldsymbol{x}\succeq\left(f_n^1(x^2,\boldsymbol{s}),f_n^2(x^1,\boldsymbol{s})\right) \text{ and } \boldsymbol{x}\neq\boldsymbol{b}_n(\boldsymbol{s})\right\} \\ \mathcal{R}_{II}(n,\boldsymbol{s}) &:= \left\{\boldsymbol{x}\in I\!R_+^2 : \boldsymbol{x}\preceq\boldsymbol{b}_n(\boldsymbol{s}) \text{ and } \boldsymbol{c}_{\boldsymbol{s}}^{\mathrm{T}}\left[\boldsymbol{b}_n(\boldsymbol{s})-\boldsymbol{x}\right]\leq P\right\} \\ \mathcal{R}_{III-A}(n,\boldsymbol{s}) &:= \left\{\boldsymbol{x}\in I\!R_+^2 : x^2 > b_n^2(\boldsymbol{s}) \text{ and } f_n^1(x^2,\boldsymbol{s})-P/c_{s^1}\leq x^1 < f_n^1(x^2,\boldsymbol{s})\right\} \\ \mathcal{R}_{III-B}(n,\boldsymbol{s}) &:= \left\{\boldsymbol{x}\in I\!R_+^2 : x^1 > b_n^1(\boldsymbol{s}) \text{ and } f_n^2(x^1,\boldsymbol{s})-P/c_{s^2}\leq x^2 < f_n^2(x^1,\boldsymbol{s})\right\} \\ \mathcal{R}_{IV-A}(n,\boldsymbol{s}) &:= \left\{\boldsymbol{x}\in I\!R_+^2 : x^2 > b_n^2(\boldsymbol{s}) \text{ and } x^1 < f_n^1(x^2,\boldsymbol{s})-P/c_{s^1}\right\} \\ \mathcal{R}_{IV-B}(n,\boldsymbol{s}) &:= \left\{\boldsymbol{x}\in I\!R_+^2 : \boldsymbol{x}\preceq\boldsymbol{b}_n(\boldsymbol{s}) \text{ and } \boldsymbol{c}_{\boldsymbol{s}}^{\mathrm{T}}\left[\boldsymbol{b}_n(\boldsymbol{s})-\boldsymbol{x}\right]> P\right\} \\ \mathcal{R}_{IV-C}(n,\boldsymbol{s}) &:= \left\{\boldsymbol{x}\in I\!R_+^2 : x^1 > b_n^1(\boldsymbol{s}) \text{ and } x^2 < f_n^2(x^1,\boldsymbol{s})-P/c_{s^2}\right\} \;, \end{aligned}$$

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

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

$$\boldsymbol{y}_n^*(\boldsymbol{x}, \boldsymbol{s}) := \begin{cases} \boldsymbol{x}, & \text{if } \boldsymbol{x} \in \mathcal{R}_I(n, \boldsymbol{s}) \\ \boldsymbol{b}_n(\boldsymbol{s}), & \text{if } \boldsymbol{x} \in \mathcal{R}_{II}(n, \boldsymbol{s}) \\ \left(f_n^1(x^2, \boldsymbol{s}), x^2\right), & \text{if } \boldsymbol{x} \in \mathcal{R}_{III-A}(n, \boldsymbol{s}) \\ \left(x^1, f_n^2(x^1, \boldsymbol{s})\right), & \text{if } \boldsymbol{x} \in \mathcal{R}_{III-B}(n, \boldsymbol{s}) \end{cases} . \tag{19}$$

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

$$\boldsymbol{c}_s^{\mathrm{T}} \left[\boldsymbol{y}_n^*(\boldsymbol{x}, \boldsymbol{s}) - \boldsymbol{x}\right] = P \ . \tag{20}$$

A detailed proof is included in Appendix C. Equation (20) 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 $\mathcal{R}_{IV}(n, \mathbf{s})$. Figure 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 $\left(b_n^m(\mathbf{s})\right)$ such that it is optimal to bring both receivers' buffer levels up to those critical numbers if it is possible to do so $\left(\text{region } \mathcal{R}_{II}(n, \mathbf{s})\right)$, and it is optimal to not transmit any packets if both receivers' buffer levels start beyond their critical numbers $\left(\text{region } \mathcal{R}_I(n, \mathbf{s})\right)$. 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 Figure 6 by the ★ and ⧫, 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.

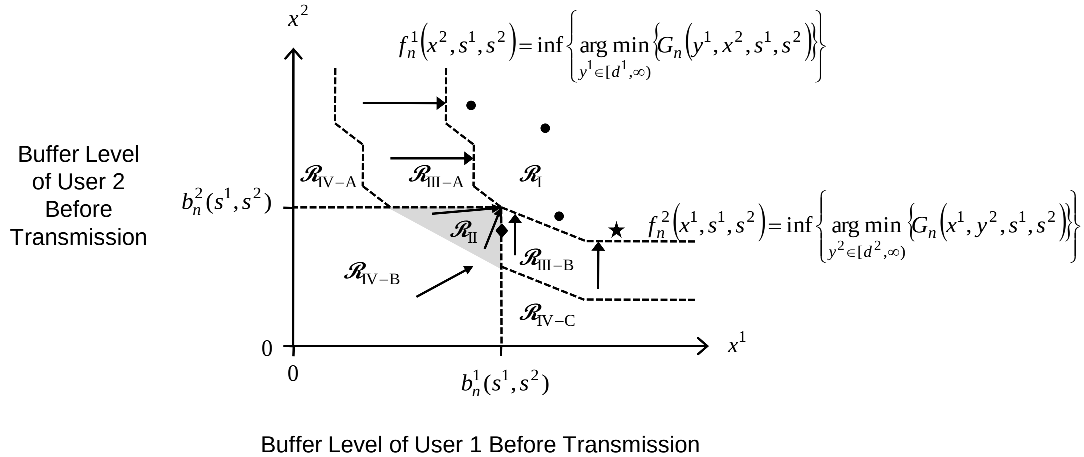

Figure 6 **Optimal transmission policy for the two receiver case in slot n when the state is $(\mathbf{x}, \mathbf{s})$. The seven regions described in Theorem 6 are labeled. The tails of the arrows represent the vectors of the receiver buffer levels at the beginning of slot n, and the heads of the arrows 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 ★ and ⧫ represent possible starting buffer levels for Scenarios 1 and 2, respectively, in Example 1.**

6. Comparison with Deterministic Ordering Costs

In this section, we compare the results of the previous section to the two-item resource-constrained inventory model with the more classical assumptions of deterministic, time-invariant prices and stochastic demands. Variants of this model are considered in Evans (1967), Chen (2004), DeCroix and Arreola-Risa (1998), and Janakiraman et al. (2009). Our purpose in discussing this model is to compare the qualitative properties of the optimal policy to those of Problem (**P1**). Specifically, the question at hand is whether models with stochastic prices deserve their own analysis or if the qualitative behavior follows in a straightforward manner from analysis of models with deterministic prices. The main thesis of the section is that inventory models with stochastic prices do indeed merit their own line of analysis as structural phenomena that cannot appear in the corresponding models with deterministic prices are liable to appear in the stochastic price inventory models.

6.1. Problem Formulation with Stochastic Demands and Deterministic Ordering Costs

We consider a two-item inventory model where the total ordering cost in each period cannot exceed a joint budget, P. The ordering costs for each item are linear, with the deterministic, time-invariant vector of ordering prices given by $\mathbf{c}$. The vector of inventories in period n is given by $\mathbf{X}_n$, and the vector of controlled order quantities is denoted by $\mathbf{Z}_n$. The demands for each item are stochastic, and represented by the random vector $\mathbf{D}_n$ in period n. We assume the vector of demands is IID across time. Unmet demands are completely backlogged until future slots (i.e., $\mathbf{X}$ can take on negative values). The total shortage and holding costs at the end of each period are given by $l(\mathbf{x}) := l^1(x^1) + l^2(x^2)$, where $l^j(0) = 0$ and $l^j(\cdot)$ is convex, nondecreasing above 0, and nonincreasing below 0, for item $j = 1, 2$. We consider the following finite horizon discounted expected cost problem, which we refer to as Problem (**D1**):

$$\begin{aligned} &\inf_{\pi \in \Pi} \quad I\!E^{\pi}\left\{\sum_{n=1}^{N} \alpha^{N-n} \cdot \left\{\mathbf{c}^{\mathrm{T}}\mathbf{Z}_n + l(\mathbf{X}_n + \mathbf{Z}_n - \mathbf{D}_n)\right\} \mid \mathcal{F}_N\right\} \\ \text{s.t.} \quad &\mathbf{c}^{\mathrm{T}}\mathbf{Z}_n \leq P, \ w.p.1, \ \forall n \\ \text{and} \quad &\mathbf{Z}_n \succeq \mathbf{0}, \ w.p.1, \ \forall n \ . \end{aligned}$$

Using the normal change of variable $\mathbf{Y}_n = \mathbf{X}_n + \mathbf{Z}_n$, the dynamic program for Problem (**D1**) is given by:

$$\begin{aligned} V_n(\mathbf{x}) &= -\mathbf{c}^{\mathrm{T}}\mathbf{x} + \min_{\mathbf{y} \in \hat{\mathcal{A}}(\mathbf{x})} \left\{\hat{G}_n(\mathbf{y})\right\}, \ n = N, N-1, \ldots, 1, \\ V_0(\mathbf{x}) &= 0, \quad \forall \mathbf{x} \in I\!R^2, \end{aligned}$$

where $\hat{G}_n(\mathbf{y}) := \mathbf{c}^{\mathrm{T}}\mathbf{y} + I\!E\big[l(\mathbf{y} - \mathbf{D})\big] + \alpha \cdot I\!E\big[V_{n-1}(\mathbf{y} - \mathbf{D})\big]$, and the action space is defined as:

$$\hat{\mathcal{A}}(\mathbf{x}) := \left\{\mathbf{y} \in I\!R^2 : \mathbf{x} \preceq \mathbf{y} \text{ and } \mathbf{c}^{\mathrm{T}}(\mathbf{y} - \mathbf{x}) \leq P\right\} \ .$$

6.2. Structure of the Optimal Policy

Problem (**D1**) is essentially the same problem as the one considered in Chen (2004). However, in order to maintain comparability with the corresponding statements about Problem (**P1**), we present the structure of the optimal policy for Problem (**D1**) in our notation.[11] We define $\hat{\mathcal{B}}_n$, the global minimizers of $\hat{G}_n(\cdot, \mathbf{s})$; the vector $\hat{\mathbf{b}}_n$; the functions $\hat{f}_n^1$ and $\hat{f}_n^2$; and the seven regions $\hat{\mathcal{R}}_I(n)$ through $\hat{\mathcal{R}}_{IV-c}(n)$ in an analogous manner to Theorem 6 with the exception that none depend on the current ordering cost condition s, which is time-invariant in Problem (**D1**).

THEOREM 7. *For Problem (**D1**), for all $\boldsymbol{x} \notin \hat{\mathcal{R}}_{IV-B}(n)$, an optimal control action with n slots remaining is given by:*

$$\boldsymbol{y}_n^*(\boldsymbol{x}) := \begin{cases} \boldsymbol{x}, & \text{if} \ \ \boldsymbol{x} \in \hat{\mathcal{R}}_I(n) \\ \hat{\boldsymbol{b}}_n, & \text{if} \ \ \boldsymbol{x} \in \hat{\mathcal{R}}_{II}(n) \\ \left(\hat{f}_n^1(x^2), x^2\right), & \text{if} \ \ \boldsymbol{x} \in \hat{\mathcal{R}}_{III-A}(n) \\ \left(x^1, \hat{f}_n^2(x^1)\right), & \text{if} \ \ \boldsymbol{x} \in \hat{\mathcal{R}}_{III-B}(n) \\ \left(x^1 + \frac{P}{c^1}, x^2\right), & \text{if} \ \ \boldsymbol{x} \in \hat{\mathcal{R}}_{IV-A}(n) \\ \left(x^1, x^2 + \frac{P}{c^2}\right), & \text{if} \ \ \boldsymbol{x} \in \hat{\mathcal{R}}_{IV-B}(n) \end{cases} . \tag{21}$$

[11] We derived this result independently before learning of Chen (2004), and our proof of Theorem 7 is different from that of (Chen 2004, Theorem 2), as we decouple the proof of the supermodularity from the submodularity with respect to the direct value orders.

For all $\boldsymbol{x} \in \hat{\mathcal{R}}_{IV-B}(n)$*, there exists an optimal control action with* n *slots remaining,* $\boldsymbol{y}_n^*(\boldsymbol{x})$*, which satisfies:*[12]

$$\boldsymbol{c}^{\mathrm{T}}\left[\boldsymbol{y}_n^*(\boldsymbol{x}) - \boldsymbol{x}\right] = P \ \textit{and}\ \boldsymbol{y}_n^*(\boldsymbol{x}) \preceq \hat{\boldsymbol{b}}_n \ . \tag{22}$$

6.3. Comparison of Problems (P1) and (D1)

As first glance, the structures of the optimal policies for Problems (**P1**) and (**D1**), described in Theorems 6 and 7, respectively, may seem extremely similar. However, there are two fundamental differences that distinguish these two problems.

First, in addition to convexity and supermodularity, the function $\hat{G}_n(\cdot)$ in Problem (**D1**) has the properties of submodularity with respect to the direct $(\mathbf{c}, 1)$ and $(\mathbf{c}, 2)$ value orders. These partial orders, introduced in Antoniadou (1996, 2007), are defined and briefly discussed in Appendix D.

The extra functional properties of $\hat{G}_n(\cdot)$ lead to two additional structural results on the optimal control action: (i) when the initial vector of inventories (corresponds to the vector of receivers' buffer levels in Problem (**P1**)) is in region $\hat{\mathcal{R}}_{IV-B}(n)$, there exists an optimal control action such that $\mathbf{y}_n^*(\mathbf{x}) \preceq \mathbf{b}_n$; and (ii) when the initial vector of inventories is in region $\hat{\mathcal{R}}_{IV-A}(n)$ (respectively, $\hat{\mathcal{R}}_{IV-C}(n)$), there exists an optimal control action that includes not ordering any of item 2 (respectively, item 1), corresponding to not transmitting any packets to user 2 (respectively, user 1) in Problem (**P1**). Due to the time-varying ordering prices (channel conditions), this property does not hold for the function $G_n(\cdot, \mathbf{s})$ in Problem (**P1**), and these two additional statements on the structure of the optimal policy are not true in general for Problem (**P1**), as shown by the following example.

Example 2. Consider the following instance of Problem (**P1**). The two items are statistically identical, the ordering costs are linear, and the possible ordering prices are 1.750, 2.000, 2.001, and 2.100. The associated probabilities of these ordering prices are 0.4, 0.4, 0.1, and 0.1, respectively. The total budget constraint in each slot is $P = 4.2$, and the deterministic demand is $\mathbf{d} = (1, 1)$. We consider a finite horizon problem with the discount rate $\alpha = 1$, and no holding costs. We are interested in the optimal control action with $T = 3$ time slots remaining, and the current prices are 2.000 and 2.001 for users 1 and 2, respectively.

Exactly solving the dynamic program by hand shows that the unique global minimizer of the function $G_3(\cdot, \cdot, \mathbf{s}_3)$ is the vector $(\frac{101}{75}, \frac{101}{75})$. However, if the vector of starting receiver inventory levels at time $T = 3$ is $\mathbf{x}_3 = (0.2, 0.2)$, the unique optimal ordering decision in the slot is to order 0.8 units

[12] Chen (2004) elaborates on the optimal allocation of the budget between the two items in Region $\hat{\mathcal{R}}_{IV-B}(n)$ by defining a curve splitting the region into the area where item 1 should be ordered and the area where item 2 should be ordered. Chen refers to this policy as a *hedging point policy.*

of item 2, and use the remaining budget on item 1, which results in the purchase of 1.2996 units of item 1. A diagram of this optimal control action is shown in Figure 7. The interesting thing to note here is that despite being budget-constrained (the vector of starting inventory levels is in Region $\mathcal{R}_{IV-B}$), the unique optimal decision calls for filling item 1's inventory beyond its critical number $b_3^1(\mathbf{s}_3) = \frac{101}{75}$. That is, the optimal decision brings the vector of inventory levels from Region $\mathcal{R}_{IV-B}$ to Region $\mathcal{R}_{III-B}$ rather than Region $\mathcal{R}_{II}$. Thus, the optimal policy is somewhat surprisingly not a base-stock policy in the sense of Porteus (1990).

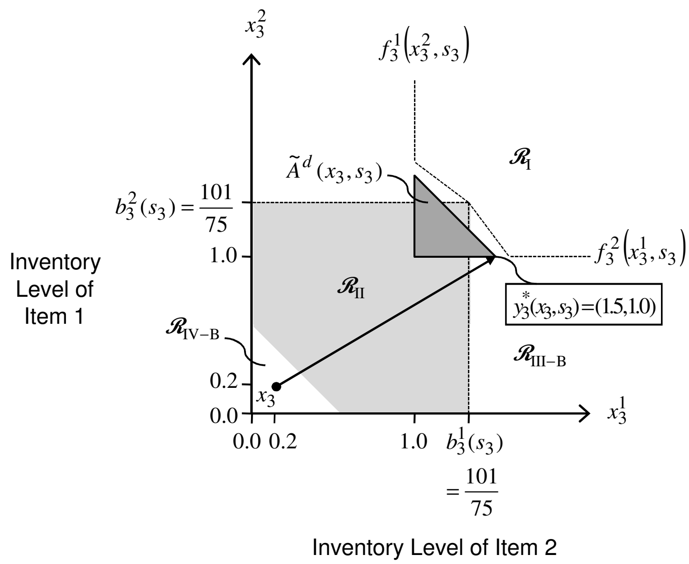

Figure 7 **Optimal scheduling decision with 3 slots remaining in Example 2. The action space is represented by the triangle $\tilde{\mathcal{A}}^{\mathbf{d}}(\mathbf{x}_3, \mathbf{s}_3)$. The critical vector $\mathbf{b}_3(\mathbf{s}_3)$ is not reachable from the starting inventory levels $\mathbf{x}_3 = (0.2, 0.2)$. The unique optimal control action is to choose $\mathbf{y}_3(\mathbf{x}_3, \mathbf{s}_3)$ to be (1.5, 1.0). The interesting feature of the example is that even though $\mathbf{x}_3 \preceq \mathbf{b}_3(\mathbf{s}_3)$, we have $\mathbf{y}_3^*(\mathbf{x}_3, \mathbf{s}_3) \not\preceq \mathbf{b}_3(\mathbf{s}_3)$.**

The second fundamental difference is also a consequence of the time-varying ordering prices in Problem (**P1**). In the infinite horizon version of Problem (**D1**), the critical vector $\hat{\mathbf{b}}$ is time-invariant. Combined with the above property that it is optimal to not order inventory so as to move out of regions $\hat{\mathcal{R}}_{II}$ and $\hat{\mathcal{R}}_{IV-B}$, the time-invariant critical vector means that the region $\hat{\mathcal{R}}_{II} \cup \hat{\mathcal{R}}_{IV-B}$ (i.e., the lower-left square below the critical vector) is a "stability" region.[13] Eventually, the vector of inventories enters this region under the optimal ordering policy, and once it does, it never leaves. This behavior both simplifies the analysis and opens the door for new mathematical techniques. For instance, Tayur (1993) and Janakiraman et al. (2009) analyze *shortfall* to compute the critical numbers and determine the optimal allocation between items in the budget-constrained region. In

[13] A key assumption needed to ensure the stability region is that $\mathbf{c}^{\mathrm{T}} \mathbb{E}[\mathbf{D}] < P$; that is, the budget in a period suffices to purchase inventory to fulfill the aggregate average demand. Without this assumption, the infinite horizon average cost is infinite as the shortage costs accumulate (Janakiraman et al. 2009).

the infinite horizon Problems (**P2**) and (**P3**) (which are discussed further in Section 7.1), even though the boundaries of the seven regions for each possible channel condition are time-invariant, no such stability region exists, because the critical numbers vary over time due to the time-varying channel conditions. Therefore, the same vector of inventories may be in region $\mathcal{R}_{II}(\mathbf{s})$ in one time slot and say $\mathcal{R}_{III-A}(\mathbf{s}')$ in the next time slot. This makes it significantly more difficult to determine optimal and near-optimal policies.

7. Extensions

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

7.1. 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 Figure 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 Figure 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* (e.g. Hernández-Lerma and Lasserre 1996). 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 5.5, 5.6, 5.10, and 5.11 of Shuman (2010).

7.2. 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.[14]

7.3. The Most General Case of M Receivers

Our ongoing work includes 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. One simple lower bound to the value function can be found by relaxing the per slot joint power constraint of P, and allowing up to P units of power to be allocated to *each* receiver in a single slot (for a total of up to $M \cdot P$). The advantage of this technique is that it is easy to compute the lower bound, as the M-dimensional problem separates into M instances of the 1-dimensional problem we know how to solve from Section 3. However, the resulting bound is likely to be loose. A second lower bounding method we are investigating is the information relaxation method of Brown et al. (2010). The main idea is to assume the scheduler has access to future channel conditions, but penalize the scheduler for using this information. A clever choice of the penalty function often leads to tight lower bounds on the value function. A third method is the Lagrangian relaxation method discussed in Hawkins (2003) and Adelman and Mersereau (2008). For our problem, this method is equivalent to relaxing the per slot peak power constraint to an average power constraint (i.e., the scheduler may allocate more than P units of power in some slots, but the average power consumed per slot over the duration of

[14] 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. Li and Bambos (2004) consider such a model, which corresponds to the *lost sales* assumption in inventory theory.

the horizon cannot exceed P). In the case of M statistically identical receivers, the resulting relaxed problem under this method separates into M instances of a 1-dimensional problem, this time with an average power constraint of $\frac{P}{M}$ instead of a strict power constraint of P for each receiver. A fourth lower bounding method is the linear programming approach to approximate dynamic programming discussed in Schweitzer and Seidmann (1985), de Farias and Van Roy (2003), and Adelman and Mersereau (2008). The main idea is to formulate the dynamic program as a linear program, and approximate the value functions as linear combinations of a set of basis functions. For a more in-depth comparison of the Lagrangian relaxation and approximate linear programming approaches, see Adelman and Mersereau (2008). Once lower bounds to the value function are determined from any of these methods, feasible policies can be generated based on our structural results or via one-step greedy optimization with the lower bounds substituted into the right-hand side of the dynamic programming equation.

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 $\mathcal{R}_{IV}(n, \mathbf{s})$.

8. Conclusion

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 an efficient method to compute the critical numbers that fully characterize the optimal modified base-stock and finite generalized base-stock policies. For the case of two receivers, the structure of the optimal policy is in some sense an extension of the modified base-stock policy; however, the peak power constraint couples the optimal scheduling of the two data streams.

The literature on inventory models with stochastic ordering costs is relatively thin compared to the more classical inventory models with deterministic ordering costs and stochastic demands. While some of our results follow in an expected manner from the more classical setup (e.g., the optimality of a modified base-stock policy), the time-varying channel conditions may result in counterintuitive optimal scheduling decisions that are not possible in the analogous inventory theory problems with deterministic ordering costs. The class of multi-item inventory models with stochastic ordering costs and a joint resource constraint therefore merits its own line of analysis. Specifically, the M-item inventory problem with stochastic prices and a joint resource constraint remains open, and does not lend itself to known techniques such as shortfall analysis. Many of the numerical approaches to this most general case of M items proposed in Section 7.3 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 items may also indirectly improve the quality of approximate numerical solutions to related higher dimensional problems.

Appendix A: Intuitive Explanation of the Recursion (12)

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:

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

which is equivalent to:

$$\begin{aligned}
&\gamma_{n,j} \\
&= -h + \frac{\alpha}{d}\cdot \mathbb{E}\left[V_{n-1}\Big((j-2)\cdot d, S_{n-1}\Big) - V_{n-1}\Big((j-1)\cdot d, S_{n-1}\Big)\right] && (23)\\
&= -h + \frac{\alpha}{d}\cdot \sum_{s\in\mathcal{S}} p(s)\cdot \left[V_{n-1}\Big((j-2)\cdot d, s\Big) - V_{n-1}\Big((j-1)\cdot d, s\Big)\right] \\
&= -h + \frac{\alpha}{d}\cdot \left[\begin{array}{l}
\displaystyle\sum_{s:\ b_{n-1}(s)\le (j-2)\cdot d} p(s)\cdot \left\{-h\cdot d + \alpha\mathbb{E}\left[V_{n-2}\Big((j-3)\cdot d, S_{n-2}\Big) - V_{n-2}\Big((j-2)\cdot d, S_{n-2}\Big)\right]\right\} \\
\displaystyle + \sum_{s:\ (j-2)\cdot d < b_{n-1}(s) \le \big(j-2+L(s)\big)\cdot d} p(s)\cdot c_s \cdot d \\
\displaystyle + \sum_{s:\ b_{n-1}(s) > \big(j-2+L(s)\big)\cdot d} p(s)\cdot \left\{-h\cdot d + \alpha\mathbb{E}\left[\begin{array}{l} V_{n-2}\Big(\big(j-3+L(s)\big)\cdot d, S_{n-2}\Big) \\ -V_{n-2}\Big(\big(j-2+L(s)\big)\cdot d, S_{n-2}\Big)\end{array}\right]\right\}
\end{array}\right] && (24)\\
&= -h + \alpha\cdot\left\{ \sum_{s:\ b_{n-1}(s)\le (j-2)\cdot d} p(s)\cdot \gamma_{n-1,j-1} + \sum_{s:\ (j-2)\cdot d < b_{n-1}(s) \le \big(j-2+L(s)\big)\cdot d} p(s)\cdot c_s + \sum_{s:\ b_{n-1}(s) > \big(j-2+L(s)\big)\cdot d} p(s)\cdot \gamma_{n-1,j-1+L(s)} \right\} && (25)\\
&= -h + \alpha\cdot\left\{ \sum_{s:\ c_s \ge \gamma_{n-1,j-1}} p(s)\cdot \gamma_{n-1,j-1} + \sum_{s:\ \gamma_{n-1,j-1+L(s)} \le c_s < \gamma_{n-1,j-1}} p(s)\cdot c_s + \sum_{s:\ c_s < \gamma_{n-1,j-1+L(s)}} p(s)\cdot \gamma_{n-1,j-1+L(s)} \right\}. && (26)
\end{aligned}$$

Here, (24) follows from the structure of the optimal control action (7). If the channel condition s in the $(n-1)^{st}$ slot is such that $b_{n-1}(s) \le (j-2)\cdot d$, 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 $(j-2)\cdot d < b_{n-1}(s) \le \big(j-2+L(s)\big)\cdot d$, 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 $b_{n-1}(s) > \big(j-2+L(s)\big)\cdot d$, 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 (25) follows from (23), with $n-1, j-1$ substituted for n, j, and (26) follows from the definition that $b_n(s) = j\cdot d$ if $\gamma_{n,j+1} \le c_s < \gamma_{n,j}$.

Comparing the threshold $\gamma_{n,j}$ defined in (12) to the corresponding threshold in the uncapacitated (no power constraint) single user problem studied in Kingsman (1969a) and Golabi (1985), the only difference is the third term of the right-hand side of (12):

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

which is absent in the uncapacitated case. For all $n \in \{1,2,\ldots,N\}$ 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 as high in the capacitated case for all $n \in \{1,2,\ldots,N\}$ 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.

Appendix B: 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. $G_1(\mathbf{y}_1,\mathbf{s}_1) = \mathbf{c}_{\mathbf{s}_1}^{\mathrm{T}}\mathbf{y}_1 + h(\mathbf{y}_1 - \mathbf{d})$, which is convex and supermodular. Thus, (iii) and (iv) are true. Additionally,

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

so $\inf \left\{ \underset{y_1^2 \in [d^2,\infty)}{\arg\min} \left\{ G_1\left(y_1^1, y_1^2, s_1^1, s_1^2\right) \right\} \right\}$ is independent of y_1^1, and vice versa.

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

(i) Consider two arbitrary points, $\bar{\mathbf{x}}, \tilde{\mathbf{x}} \in I\!R_+^2$. 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:

$$\begin{aligned}
\lambda \cdot V_{l-1}(\bar{\mathbf{x}},\mathbf{s}) + (1-\lambda)\cdot V_{l-1}(\tilde{\mathbf{x}},\mathbf{s}) &= -\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\hat{\mathbf{x}} + \lambda \cdot G_{l-1}\big(\mathbf{y}^*(\bar{\mathbf{x}},\mathbf{s}),\mathbf{s}\big) + (1-\lambda)\cdot G_{l-1}\big(\mathbf{y}^*(\tilde{\mathbf{x}},\mathbf{s}),\mathbf{s}\big) \\
&\geq -\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\hat{\mathbf{x}} + G_{l-1}\big(\lambda\mathbf{y}^*(\bar{\mathbf{x}},\mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}},\mathbf{s}),\mathbf{s}\big) \\
&\geq -\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\hat{\mathbf{x}} + \min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\hat{\mathbf{x}},\mathbf{s})} \{G_{l-1}(\mathbf{y},\mathbf{s})\} \\
&= V_{l-1}(\hat{\mathbf{x}},\mathbf{s}) = V_{l-1}(\lambda\bar{\mathbf{x}} + (1-\lambda)\tilde{\mathbf{x}},\mathbf{s})\ ,
\end{aligned} \tag{27}$$

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

$$\mathbf{y}^*(\bar{\mathbf{x}},\mathbf{s}) \succeq \mathbf{d} \vee \bar{\mathbf{x}} \quad \text{and} \quad \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\left[\mathbf{y}^*(\bar{\mathbf{x}},\mathbf{s}) - \bar{\mathbf{x}}\right] \leq P\ . \tag{28}$$

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

$$\mathbf{y}^*(\tilde{\mathbf{x}},\mathbf{s}) \succeq \mathbf{d} \vee \tilde{\mathbf{x}} \quad \text{and} \quad \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\left[\mathbf{y}^*(\tilde{\mathbf{x}},\mathbf{s}) - \tilde{\mathbf{x}}\right] \leq P\ . \tag{29}$$

Multiplying the equations in (28) by λ and the equations in (29) by $1-\lambda$, and summing, we have:

$$\lambda\mathbf{y}^*(\bar{\mathbf{x}},\mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}},\mathbf{s}) \succeq \lambda(\mathbf{d}\vee\bar{\mathbf{x}}) + (1-\lambda)(\mathbf{d}\vee\tilde{\mathbf{x}}) \succeq \mathbf{d}\vee\hat{\mathbf{x}}, \tag{30}$$

and

$$\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\left[\lambda\mathbf{y}^*(\bar{\mathbf{x}},\mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}},\mathbf{s}) - \hat{\mathbf{x}}\right] = \lambda\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\left[\mathbf{y}^*(\bar{\mathbf{x}},\mathbf{s}) - \bar{\mathbf{x}}\right] + (1-\lambda)\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\left[\mathbf{y}^*(\tilde{\mathbf{x}},\mathbf{s}) - \tilde{\mathbf{x}}\right] \leq P\ . \tag{31}$$

From (30) and (31), we conclude $\lambda\mathbf{y}^*(\bar{\mathbf{x}},\mathbf{s}) + (1-\lambda)\mathbf{y}^*(\tilde{\mathbf{x}},\mathbf{s}) \in \tilde{\mathcal{A}}^{\mathbf{d}}(\hat{\mathbf{x}},\mathbf{s})$. Thus, the value of $G_{l-1}(\cdot,\mathbf{s})$ at this point is greater than or equal to the minimum of $G_l(\cdot,\mathbf{s})$ over the region $\tilde{\mathcal{A}}^{\mathbf{d}}(\hat{\mathbf{x}},\mathbf{s})$. From (27), we conclude $V_{l-1}(\cdot,\mathbf{s})$ is convex. This argument is similar to the one used to show convexity in Evans (1967).

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

$$\min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\} + \min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\tilde{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\} \leq \min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\} + \min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\}. \tag{32}$$

If $\bar{\mathbf{x}}$ and $\tilde{\mathbf{x}}$ are comparable (i.e., $\tilde{x}^1 \geq \bar{x}^1$ and $\tilde{x}^2 \geq \bar{x}^2$ or $\tilde{x}^1 \leq \bar{x}^1$ and $\tilde{x}^2 \leq \bar{x}^2$), then (32) is trivial. So we assume they are not comparable, and also assume without loss of generality that $\bar{x}^1 < \tilde{x}^1$ and $\tilde{x}^2 < \bar{x}^2$. 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 $\bar{\mathbf{y}} \in \tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}},\mathbf{s})$ and $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^{\mathbf{d}}(\tilde{\mathbf{x}},\mathbf{s})$ such that:

$$\begin{aligned}
G_{l-1}(\bar{\mathbf{y}},\mathbf{s}) + G_{l-1}(\tilde{\mathbf{y}},\mathbf{s}) &\leq G_{l-1}\Big(\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}),\mathbf{s}\Big) + G_{l-1}\Big(\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s}),\mathbf{s}\Big) \\
&= \min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\} + \min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\}.
\end{aligned} \tag{33}$$

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

$$\min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\}+\min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\tilde{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\}\leq G_{l-1}(\bar{\mathbf{y}},\mathbf{s})+G_{l-1}(\tilde{\mathbf{y}},\mathbf{s}). \tag{34}$$

Combining equations (33) and (34) yields the desired result, (32). 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 (33) for two exhaustive cases.

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

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

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

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:

$$\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})\succ\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})\succeq\mathbf{d}\vee(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}})\ . \tag{35}$$

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

$$\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}[\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})-\bar{\mathbf{x}}\vee\tilde{\mathbf{x}}]\leq\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}[\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})-\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}}]\leq P\ . \tag{36}$$

Equations (35) and (36) imply $\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})$, and thus:

$$G_{l-1}\big(\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s}),\mathbf{s}\big)=\min_{\mathbf{y}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})}\{G_{l-1}(\mathbf{y},\mathbf{s})\}\leq G_{l-1}\big(\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}),\mathbf{s}\big). \tag{37}$$

However, we also have:

$$\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})\succeq\mathbf{d}\vee(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}})\succeq\mathbf{d}\vee(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}})\ , \tag{38}$$

and

$$\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}[\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})-\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}}]\leq\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}[\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})-\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}}]\leq P\ . \tag{39}$$

Equations (38) and (39) imply $\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})$, which, in combination with (37), 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 $\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})$. □

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 $\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})\not\succ\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})$. We show (32) by considering two exhaustive cases.

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

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

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

Then

$$f(\boldsymbol{z})+f(\hat{\boldsymbol{z}})\geq f(\boldsymbol{z}^{\sigma,\beta})+f(\boldsymbol{z}^{1-\sigma,1-\beta})\ . \tag{40}$$

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:

$$f(\mathbf{z})+f(\hat{\mathbf{z}})\geq f(\mathbf{z}^{\sigma,\sigma})+f(\mathbf{z}^{1-\sigma,1-\sigma})\ , \tag{41}$$

and

$$f(\mathbf{z}^{1-\sigma,1-\sigma})+f(\mathbf{z}^{1-\sigma,\sigma})\geq f(\mathbf{z}^{1-\sigma,\beta})+f(\mathbf{z}^{1-\sigma,1-\beta})\ . \tag{42}$$

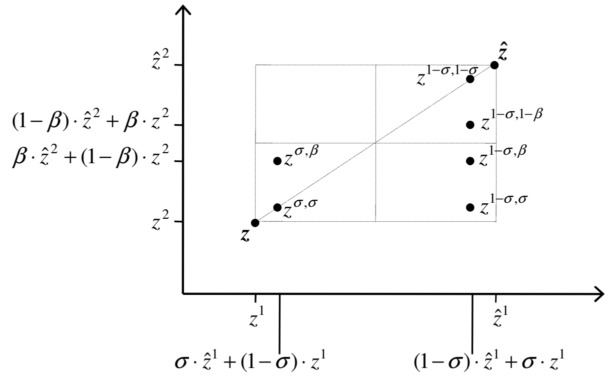

Figure 8 Diagram of the points referred to in Step 1 of the proof of Lemma 2.

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

$$f(\mathbf{z}^{1-\sigma,\beta})+f(\mathbf{z}^{\sigma,\sigma})\geq f(\mathbf{z}^{\sigma,\beta})+f(\mathbf{z}^{1-\sigma,\sigma})\ . \tag{43}$$

Figure 8 shows these relationships. Combining (41)-(43), we have:

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

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:

$$f(\mathbf{z})+f(\hat{\mathbf{z}})\geq f(\mathbf{z}^{\hat{\sigma},\hat{\beta}})+f(\mathbf{z}^{1-\hat{\sigma},1-\hat{\beta}})\ . \tag{44}$$

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:

$$f(\mathbf{z}^{\hat{\sigma},\hat{\beta}})+f(\mathbf{z}^{1-\hat{\sigma},1-\hat{\beta}})\geq f(\mathbf{z}^{\sigma,\beta})+f(\mathbf{z}^{1-\sigma,1-\beta})\ . \tag{45}$$

Combining (44) and (45) yields the desired result, (40). $\square$

Next, define the following points, shown in the left-hand side of Figure 9:

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

and

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

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:

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

By a similar argument, $\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}(\tilde{\mathbf{y}}-\tilde{\mathbf{x}})\leq P$, and thus $\bar{\mathbf{y}}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}},\mathbf{s})$, and $\tilde{\mathbf{y}}\in\tilde{\mathcal{A}}^{\mathbf{d}}(\tilde{\mathbf{x}},\mathbf{s})$. So (34) is true.
Now define[15]:

$$\sigma:=\frac{y^{*^1}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})-\tilde{y}^1}{y^{*1}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})-y^{*1}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})}\ \text{and}\ \beta:=\frac{y^{*^2}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})-\tilde{y}^2}{y^{*2}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})-y^{*2}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s})}\ .$$

[15] 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 σ 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 β be arbitrary in $[0,1]$.

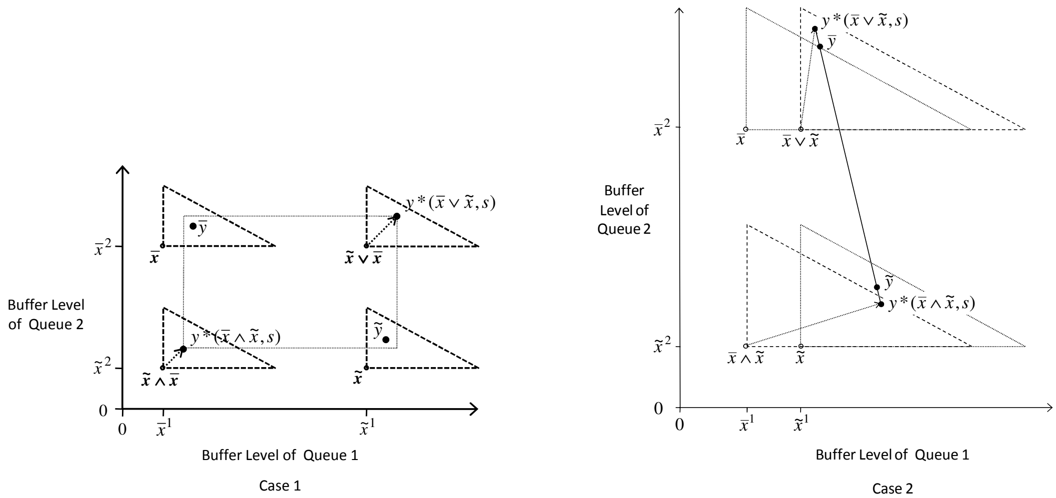

Figure 9 **Construction of feasible points $\bar{\mathbf{y}}$ and $\tilde{\mathbf{y}}$ in Cases 1 and 2 of the proof of supermodularity of $V_{l-1}(\cdot,\mathbf{s})$.**

Rearranging the definitions of σ and β yields:

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

It is also straightforward to check that:

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

We also have:

$$\begin{aligned} y^{*^1}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) &= \min\Big\{y^{*^1}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}), y^{*^1}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) + (\tilde{x}^1-\tilde{x}^2)\Big\} \\ &\leq \min\Big\{y^{*^1}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s}), y^{*^1}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) + (\tilde{x}^1-\tilde{x}^2)\Big\} = \tilde{y}^1 \leq y^{*^1}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s}) \ , \end{aligned}$$

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 (33), which, in combination with (34), implies (32).

Case 2: $\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s}) \ngeq \mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) \ngeq \mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s})$

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 (32) under the first possibility, and a symmetric argument can be used to show (32) under the second possibility. We have:

$$y^{*^1}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) > y^{*^1}(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s}) \geq \max\left\{(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}})^1, d^1\right\} = \max\left\{\tilde{x}^1, d^1\right\}, \tag{46}$$

$$y^{*^2}(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) \geq \max\left\{(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}})^2, d^2\right\} = \max\left\{\tilde{x}^2, d^2\right\}, \tag{47}$$

and

$$\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\Big[\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) - \tilde{\mathbf{x}}\Big] \leq \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\Big[\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) - (\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}})\Big] \leq P \ . \tag{48}$$

Equations (46), (47), and (48) imply $\mathbf{y}^*(\bar{\mathbf{x}}\wedge\tilde{\mathbf{x}},\mathbf{s}) \in \tilde{\mathcal{A}}^{\mathbf{d}}(\tilde{\mathbf{x}},\mathbf{s})$. If it also happens that $\mathbf{y}^*(\bar{\mathbf{x}}\vee\tilde{\mathbf{x}},\mathbf{s}) \in \tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}},\mathbf{s})$, then we have:

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

Otherwise, define:

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

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

$$\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\mathbf{y}^*(\bar{\mathbf{x}} \vee \tilde{\mathbf{x}}, \mathbf{s}) > \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\bar{\mathbf{x}} + P \geq \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}) + P \geq \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\mathbf{y}^*(\bar{\mathbf{x}} \wedge \tilde{\mathbf{x}}, \mathbf{s}) \ . \tag{49}$$

It is clear from (49) that the numerator and denominator of γ are positive, and $\gamma \in [0,1]$. Now define:

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

It is somewhat tedious but straightforward to show that $\bar{\mathbf{y}} \in \tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}, \mathbf{s})$ and $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^{\mathbf{d}}(\tilde{\mathbf{x}}, \mathbf{s})$, implying (34). In the right-hand side of Figure 9, $\bar{\mathbf{y}}$ 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 $\tilde{\mathcal{A}}^{\mathbf{d}}(\bar{\mathbf{x}}, \mathbf{s})$, 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 (33) follows from the convexity of $G_{l-1}(\cdot, \mathbf{s})$ along this line segment. Combining (33) and (34) again yields the desired result, (32).

(iii) $G_l(\mathbf{y}, \mathbf{s}) = \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\mathbf{y} + h(\mathbf{y} - \mathbf{d}) + \alpha \cdot I\!\!E\left[V_{l-1}(\mathbf{y} - \mathbf{d}, \mathbf{S})\right]$. 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. $I\!\!E\left[V_{l-1}(\mathbf{y} - \mathbf{d}, \mathbf{S})\right]$ 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 $\mathbf{y}$.

(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.

(v) This step follows from Topkis 1998, Theorem 2.8.1, p. 76.

Appendix C: Proof of Theorem 6

Let $n \in \{1, 2, \ldots, N\}$ and $\mathbf{s} \in \mathcal{S}$ be arbitrary. We start by proving (19). First, let $\mathbf{x} \in \mathcal{R}_I(n, \mathbf{s})$ and $\hat{\mathbf{y}} \in \tilde{\mathcal{A}}^{\mathbf{d}}(\mathbf{x}, \mathbf{s})$ 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 $[d^1, \infty) \times [d^2, \infty)$ (e.g. Rockafellar 1970, 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:

$$G_n(\hat{\mathbf{y}}, \mathbf{s}) \geq G_n\big((\hat{y}^1, x^2), \mathbf{s}\big) \geq G_n\Big(\big(\hat{y}^1, f_n^2(\hat{y}^1, \mathbf{s})\big), \mathbf{s}\Big) \ . \tag{50}$$

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:

$$G_n\big((\hat{y}^1, x^2), \mathbf{s}\big) \geq G_n(\mathbf{x}, \mathbf{s}) \geq G_n\Big(\big(f_n^1(x^2, \mathbf{s}), x^2\big), \mathbf{s}\Big) \ . \tag{51}$$

Combining (50) and (51) yields:

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

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

Second, let $\mathbf{x} \in \mathcal{R}_{II}(n, \mathbf{s})$ be arbitrary. Then $\mathbf{b}_n(\mathbf{s}) \in \tilde{\mathcal{A}}^{\mathbf{d}}(\mathbf{x}, \mathbf{s})$ 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 $\tilde{\mathbf{y}} \in \tilde{\mathcal{A}}^{\mathbf{d}}(\mathbf{x}, \mathbf{s})$ be arbitrary. By definition of $f_n^1(\cdot, \mathbf{s})$, we have:

$$G_n(\tilde{\mathbf{y}}, \mathbf{s}) \geq G_n\Big(\big(f_n^1(\tilde{y}^2, \mathbf{s}), \tilde{y}^2\big), \mathbf{s}\Big) \ . \tag{52}$$

Furthermore, the function $\min_{y^1 \in [d^1, \infty)} \Big\{G_n\Big((y^1, y^2), \mathbf{s}\Big)\Big\}$ is convex in y^2 since $[d^1, \infty)$ is a convex set (e.g. Boyd and Vandenberghe 2004, pp. 101-102). Thus, $\tilde{y}^2 \geq x^2 \geq b_n^2(\mathbf{s})$ implies:

$$G_n\Big(\big(f_n^1(\tilde{y}^2, \mathbf{s}), \tilde{y}^2\big), \mathbf{s}\Big) \geq G_n\Big(\big(f_n^1(x^2, \mathbf{s}), x^2\big), \mathbf{s}\Big) \geq G_n\Big(\big(f_n^1(b_n^2(\mathbf{s}), \mathbf{s}), b_n^2(\mathbf{s})\big), \mathbf{s}\Big) = G_n\Big(\mathbf{b}_n(\mathbf{s}), \mathbf{s}\Big) \ . \tag{53}$$

Combining (52) and (53) yields:

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

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

The optimality of $\mathbf{y}_n^*(\mathbf{x},\mathbf{s}) = \Big(x^1, f_n^2(x^1,\mathbf{s})\Big)$ 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 $\Big(x^1, f_n^2(x^1,\mathbf{s})\Big)$.

Finally, we prove (20). Define:

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

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

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

Note that $\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}[\hat{\mathbf{y}}-\mathbf{x}] < P$ and $\mathbf{c}_{\mathbf{s}}^{\mathrm{T}}[\mathbf{b}_n(\mathbf{s})-\mathbf{x}] > P$ 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:

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

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

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

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

<u>Case 1</u>: $\hat{y}^2 < f_n^2(\hat{y}^1,\mathbf{s})$ and $\bar{\mathbf{y}} := \Big(\hat{y}^1, f_n^2(\hat{y}^1,\mathbf{s})\Big) \notin \tilde{\mathcal{A}}^{\mathbf{d}}(\mathbf{x},\mathbf{s})$

Let $\tilde{\mathbf{y}} := \Big(\hat{y}^1, x^2 + \frac{P - c_{s^1}\cdot[\hat{y}^1-x^1]}{c_{s^2}}\Big)$. 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:

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

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

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

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

$$G_n(\hat{\mathbf{y}},\mathbf{s}) \geq G_n\Big(\big(\hat{y}^1, f_n^2(\hat{y}^1,\mathbf{s})\big),\mathbf{s}\Big) \ . \tag{54}$$

Define:

$$\tilde{y}^1 := \sup\Big\{y^1 \in \big[x^1,\hat{y}^1\big) : \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\Big(y^1, f_n^2(y^1,\mathbf{s})\Big) \geq \mathbf{c}_{\mathbf{s}}^{\mathrm{T}}\mathbf{x} + P\Big\} \text{ and } \tilde{y}^2 := \frac{P - c_{s^1}\cdot[\tilde{y}^1 - x^1]}{c_{s^2}} \ .$$

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

$$G_n\Big(\big(\hat{y}^1, f_n^2(\hat{y}^1,\mathbf{s})\big),\mathbf{s}\Big) \geq G_n\Big(\big(\tilde{y}^1, f_n^2(\tilde{y}^1,\mathbf{s})\big),\mathbf{s}\Big) \ . \tag{55}$$

Furthermore, we have:

$$G_n\Big(\big(\tilde{y}^1, f_n^2(\tilde{y}^1,\mathbf{s})\big),\mathbf{s}\Big) = G_n\Big(\big(\tilde{y}^1, \tilde{y}^2\big),\mathbf{s}\Big) = G_n(\tilde{\mathbf{y}},\mathbf{s}) \ . \tag{56}$$

If $\tilde{y}^2 = f_n^2(\tilde{y}^1,\mathbf{s})$, (56) is trivial. Otherwise, there is a discontinuity in $f_n^2(\cdot,\mathbf{s})$ at $\tilde{y}^1$, and we have:

$$\lim_{y^1 \nearrow \tilde{y}^1} f_n^2(y^1,\mathbf{s}) \geq \tilde{y}^2 \geq \lim_{y^1 \searrow \tilde{y}^1} f_n^2(y^1,\mathbf{s}) \ , \tag{57}$$

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

$$G_n\Big(\big(\tilde{y}^1, \lim_{y^1 \nearrow \tilde{y}^1} f_n^2\left(y^1,\mathbf{s}\right)\big),\mathbf{s}\Big) = G_n\Big(\big(\tilde{y}^1, \lim_{y^1 \searrow \tilde{y}^1} f_n^2\left(y^1,\mathbf{s}\right)\big),\mathbf{s}\Big) = G_n\Big(\big(\tilde{y}^1, f_n^2\left(\tilde{y}^1,\mathbf{s}\right)\big),\mathbf{s}\Big). \tag{58}$$

The convexity of $G_n(\cdot,\mathbf{s})$ along the line $y^1 = \tilde{y}^1$ and (58) imply:

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

which, in combination with (57), implies (56). Combining (54)-(56) yields the desired result: $G_n\left(\tilde{\mathbf{y}},\mathbf{s}\right) \leq G_n\left(\hat{\mathbf{y}},\mathbf{s}\right)$ for a $\tilde{\mathbf{y}} \in \mathcal{H}^{\mathbf{d}}(\mathbf{x},\mathbf{s})$. The validity of (20) for $\mathbf{x} \in \mathcal{R}_{IV-A}(n,\mathbf{s})$ follows from a symmetric argument, completing the proof of (20) and Theorem 6. $\square$

Appendix D: The Direct Value Order

Definition 1 (Antoniadou, 1996). Let $\mathbf{c} \in I\!R_{++}$ and let $\bar{\mathbf{x}}, \tilde{\mathbf{x}} \in I\!R^2$. Then the direct $(\mathbf{c},i)$ value order, $\leq_{dv(\mathbf{c},i)}$ for $i = 1,2$, is defined by:

$$\bar{\mathbf{x}} \leq_{dv(\mathbf{c},i)} \tilde{\mathbf{x}} \quad \text{if and only if} \quad \mathbf{c}^{\mathrm{T}}\bar{\mathbf{x}} \leq \mathbf{c}^{\mathrm{T}}\tilde{\mathbf{x}} \text{ and } \bar{x}^i \leq \tilde{x}^i\ .$$

The meet and join of two points with respect to the usual Euclidean, $dv(\mathbf{c},1)$, and $dv(\mathbf{c},2)$ partial orders are shown in Figure 10.

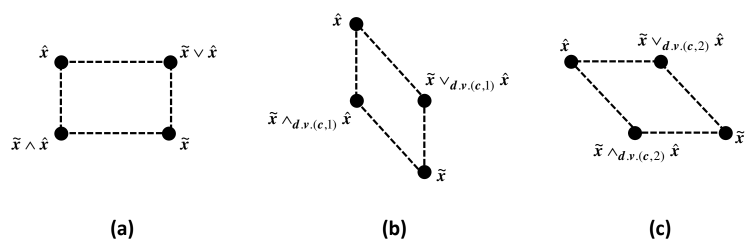

Figure 10 The direct value order. (a) shows the standard Euclidean order; (b) shows the $dv(\mathbf{c},1)$ order; and (c) shows the $dv(\mathbf{c},2)$ order.

The following proposition relates the direct $(\mathbf{c},i)$ value orders to the conditions Evans (1967) refers to as “dominance of the second partials over the mixed partials.”

Proposition 1. *If $f: I\!R^2 \to I\!R$ is twice continuously differentiable on $I\!R^2$, then f is supermodular (submodular) with respect to the direct $(\boldsymbol{c},1)$ value order if and only if:*

$$\frac{\partial^2 f}{\partial x^2 \partial x^2} \leq (\geq)\ \ \frac{c^2}{c^1} \cdot \frac{\partial^2 f}{\partial x^1 \partial x^2}\ , \tag{59}$$

and is supermodular (submodular) with respect to the direct $(\boldsymbol{c},2)$ value order if and only if:

$$\frac{\partial^2 f}{\partial x^1 \partial x^1} \leq (\geq)\ \ \frac{c^1}{c^2} \cdot \frac{\partial^2 f}{\partial x^1 \partial x^2}\ . \tag{60}$$

Definition 2 (Chen, 2004). A convex function $f: I\!R^2 \to I\!R$ is ***μ**-difference monotone* if for any $t > 0$,

(*i*) $f(x^1+t, x^2) - f(x^1,x^2)$ is nondecreasing in x^1 and nondecreasing in x^2;
(*ii*) $f(x^1, x^2+t) - f(x^1,x^2)$ is nondecreasing in x^1 and nondecreasing in x^2;
(*iii*) $f(x^1+\mu^1\cdot t, x^2) - f(x^1, x^2+\mu^2\cdot t)$ is nondecreasing in x^1 and nonincreasing in x^2.

Proposition 2. *A function $f: I\!R^2 \to I\!R$ satisfies statement (iii) of Definition 2 if and only if it is submodular with respect to both the $dv(\boldsymbol{c},1)$ and $dv(\boldsymbol{c},2)$ partial orders, where $c^i = \frac{1}{\mu^i}$ for $i = 1,2$.*

References

Adelman, D., A. J. Mersereau. 2008. Relaxations of weakly coupled stochastic dynamic programs. *Oper. Res.* **56**(3) 712–727.

Aleaf, A. 2007. Optimizing power consumption in the mobile multimedia delivery chain. http://www.eetimes.com/design/power-management-design/4012197/Optimizing-power-consumption-in-the-mobile-multimedia-delivery-chain–Part-I.

Andrews, M., K. Kumaran, K. Ramanan, A. Stolyar, R. Vijayakumar, P. Whiting. 2004. Scheduling in a queueing system with asynchronously varying service rates. *Probab. Eng. Inform. Sc.* **18** 191–217.

Antoniadou, E. 1996. Lattice programming and economic optimization. Ph.D. thesis, Stanford University.

Antoniadou, E. 2007. Comparative statics for the consumer problem. *Econ. Theory* **31**(1) 189–203.

Bensoussan, A., M. Crouhy, J.-M. Proth. 1983. *Mathematical Theory of Production Planning*. Elsevier Science.

Berling, P., V. Martínez de Albéniz. 2011. Optimal inventory policies when purchase price and demand are stochastic. *Oper. Res.* **59**(1) 109–124.

Berry, R. A., R. G. Gallager. 2002. Communication over fading channels with delay constraints. *IEEE T. Inform. Theory* **48**(5) 1135–1149.

Bertsekas, D., S. E. Shreve. 1996. *Stochastic Optimal Control: The Discrete-Time Case*. Athena Scientific.

Boyd, S., L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press.

Brown, D. B., J. E. Smith, P. Sun. 2010. Information relaxations and duality in stochastic dynamic programs. *Oper. Res.* **58**(4) 785–801.

Chen, S. X. 2004. The optimality of hedging point policies for stochastic two-product flexible manufacturing systems. *Oper. Res.* **52**(2) 312–322.

Chen, W., U. Mitra, M. J. Neely. 2009. Energy-efficient scheduling with individual packet delay constraints over a fading channel. *Wireless Networks* **15**(5) 601–618.

Chow, C.-S., J. N. Tsitsiklis. 1989. The complexity of dynamic programming. *J. Complexity* **5**(4) 466–488.

Collins, B. E., R. L. Cruz. 1999. Transmission policies for time varying channels with average delay constraints. *Proc. Allerton Conf. Comm., Control, and Computing*. Monticello, IL.

de Farias, D. P., B. Van Roy. 2003. The linear programming approach to approximate dynamic programming. *Oper. Res.* **51**(6) 850–865.

DeCroix, G. A., A. Arreola-Risa. 1998. Optimal production and inventory policy for multiple products under resource constraints. *Management Sci.* **44**(7) 950–961.

Ericsson. 2008. Energy-saving solutions helping mobile operators meet commercial and sustainability goals worldwide. www.ericsson.com/ericsson/press/facts_figures/doc/energy_efficiency.pdf.

Evans, R. 1967. Inventory control of a multiproduct system with a limited production resource. *Naval Res. Logist. Quart.* **14**(2) 173–184.

Fabian, T., J. L. Fisher, M. W. Sasieni, A. Yardeni. 1959. Purchasing raw material on a fluctuating market. *Oper. Res.* **7**(1) 107–122.

Federgruen, A., P. Zipkin. 1986. An inventory model with limited production capacity and uncertain demands II. The discounted-cost criterion. *Math. Oper. Res.* **11**(2) 208–215.

Fu, A., E. Modiano, J. N. Tsitsiklis. 2006. Optimal transmission scheduling over a fading channel with energy and deadline constraints. *IEEE T. Wireless Comm.* **5**(3) 630–641.

Gavirneni, S. 2004. Periodic review inventory control with fluctuating purchasing costs. *Oper. Res. Lett.* **32** 374–379.

Golabi, K. 1982. A single-item inventory model with stochastic prices. *Proc. Second Internat. Symp. Inventories*. Budapest, Hungary, 687–697.

Golabi, K. 1985. Optimal inventory policies when ordering prices are random. *Oper. Res.* **33**(3) 575–588.

Hawkins, J. T. 2003. A Lagrangian decomposition approach to weakly coupled dynamic optimization problems and its applications. Ph.D. thesis, Massachusetts Institute of Technology.

Hernández-Lerma, O., J. B. Lasserre. 1996. *Discrete-Time Markov Control Processes*. Springer-Verlag.

Janakiraman, G., M. Nagarajan, S. Veeraraghavan. 2009. Simple policies for managing flexible capacity. *Management Sci.* (Under review).

Kalymon, B. 1971. Stochastic prices in a single-item inventory purchasing model. *Oper. Res.* **19**(6) 1434–1458.

Karlin, S. 1958. Optimal inventory policy for the Arrow-Harris-Marschak dyanmic model. K. J. Arrow, S. Karlin, H. Scarf, eds., *Studies in the Mathematical Theory of Inventory and Production*. Stanford University Press, 135–154.

Kingsman, B. G. 1969a. Commodity purchasing. *Oper. Res. Quart.* **20** 59–80.

Kingsman, B. G. 1969b. Commodity purchasing in uncertain fluctuating price markets. Ph.D. thesis, University of Lancaster.

Kremling, H. 2008. Making mobile broadband a success - operator requirements. *Next Generation Mobile Networks Conf. at CeBIT*. Hannover, Germany.

Lee, J., N. Jindal. 2009a. Delay constrained scheduling over fading channels: Optimal policies for monomial energy-cost functions. *Proc. IEEE Internat. Conf. Comm.*. Dresden, Germany.

Lee, J., N. Jindal. 2009b. Energy-efficient scheduling of delay constrained traffic over fading channels. *IEEE T. Wirless Comm.* **8**(4) 1866–1875.

Li, Y., N. Bambos. 2004. Power-controlled wireless links for media streaming applications. *Proc. Wireless Telecomm. Symp.*. Pomona, CA, 102–111.

Luna, C. E., Y. Eisenberg, R. Berry, T. N. Pappas, A. K. Katsaggelos. 2003. Joint source coding and data rate adaptation for energy efficient wireless video streaming. *IEEE J. Sel. Area. Comm.* **21**(10) 1710–1720.

Magirou, V. F. 1982. Stockpiling under price uncertainty and storage capacity constraints. *Eur. J. Oper. Res.* **11** 233–246.

Miao, G., M. Himayat, Y. Li, A. Swami. 2009. Cross-layer optimization for energy-efficient wireless communications: a survey. *Wirel. Commun. Mob. Comput.* **9**(4) 529–542.

Neely, M. J., E. Modiano, C. E. Rohrs. 2003. Dynamic power allocation and routing for time varying wireless networks. *Proc. IEEE INFOCOM*, vol. 1. San Francisco, CA, 745–755.

Porteus, E. L. 1990. Stochastic inventory theory. D. P. Heyman, M. J. Sobel, eds., *Stochastic Models*. Elsevier Science, 605–652.

Pyramid Research. 2009. Mobile video service subscriptions to grow five-fold by 2014. http://www.pyramidresearch.com/pr_prlist/PR060409_VIDEO.htm.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press.

Rust, J. 1997. Using randomization to break the curse of dimensionality. *Econometrica* **65**(3) 487–516.

Schweitzer, P. J., A. Seidmann. 1985. Generalized polynomial approximations in Markovian decision processes. *J. Math. Anal. Appl.* **110**(2) 568–582.

Shuman, D. I. 2010. From sleeping to stockpiling: Energy conservation via stochastic scheduling in wireless networks. Ph.D. thesis, University of Michigan, Ann Arbor.

Shuman, D. I, M. Liu. 2010. Opportunistic scheduling with deadline constraints in wireless networks. N. Gulpinar, P. Harrison, B. Rustem, eds., *Performance Models and Risk Management in Communication Systems*. Springer, 127–155.

Sobel, M. J. 1970. Making short-run changes in production when the employment level is fixed. *Oper. Res.* **18**(1) 35–51.

Tarello, A., J. Sun, M. Zafer, E. Modiano. 2008. Minimum energy transmission scheduling subject to deadline constraints. *ACM Wireless Networks* **14**(5) 633–645.

Tassiulas, L., A. Ephremides. 1993. Dynamic server allocation to parallel queues with randomly varying connectivity. *IEEE T. Inform. Theory* **39**(2) 466–478.

Tayur, S. 1993. Computing the Optimal Policy for Capacitated Inventory Models. *Comm. in Statist. Stochastic Models* **9**(4) 585–598.

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

Uysal-Biyikoglu, E., A. El Gamal. 2004. On adaptive transmission for energy efficiency in wireless data networks. *IEEE T. Inform. Theory* **50**(12) 3081–3094.

Zahrn, F. C. 2009. Studies of inventory control and capacity planning with multiple sources. Ph.D. thesis, Georgia Institute of Technology.