Probability in the Engineering and Informational Sciences, 6, 1992, 29-61. Printed in the U.S.A.

OPTIMAL SCHEDULING OF A FINITE CAPACITY SHUTTLE UNDER DELAYED INFORMATION

MARK P. VAN OYEN AND DEMOSTHENIS TENEKETZIS
Department of Electrical Engineering and Computer Science
University of Michigan
1301 Beal Avenue
Ann Arbor, Michigan 48109-2122

We consider the optimal scheduling of a finite capacity shuttle in a two-node network with imperfect information. When shuttle trips do not depend on the number of passengers carried, we prove optimality and monotonicity of threshold policies. We derive conditions for dispatching that reduce the computational effort required to compute an optimal threshold policy. We prove a counterexample to the optimality of threshold policies for finite horizon problems where trip lengths increase stochastically in the number of passengers carried.

1. INTRODUCTION

The significant role of transportation, communication, and manufacturing networks in today's society motivates the need to develop further insight into the fundamental issues of control and optimization associated with these networks. An open issue for networks is control in the absence of complete state observations. Transportation and communication networks are often characterized by nodes that act individually, each possessing only a local knowledge of its immediate environment. Even when information can be exchanged among nodes, there are propagation and processing delays; also, faults and transmission er-

This work was supported in part by ONR Grant No. N00014-87-K-0540 and by a Regents Fellowship from the University of Michigan.

rors may render the data inaccurate. Thus, an understanding of the effects of partial information on optimal control policies will be useful for effectively designing and controlling networks in which incomplete (imperfect) information is a realistic consideration. As a modest step in this direction, we focus on the effect that delayed observations have on an optimal shuttle scheduling policy in a simple two-node transportation network.

In discrete time, we examine a two-terminal network with a single, finite capacity shuttle providing transportation between the terminals. Passengers arrive at either terminal and must be transported to the other terminal, whereupon they exit the system. At a given terminal at any time t, the controller's (shuttle dispatcher's) decisions are based on the following information: (i) the history of the arrival process to that terminal through time t and (ii) the history of the arrival process of the other terminal through time $t - I$. By imposing a holding cost per passenger per unit time held at either terminal, we provide an incentive for prompt service. On the other hand, a dispatching cost is incurred by each shuttle trip, thus discouraging frequent dispatching.

The objective is to characterize a shuttle dispatching policy that is a function of the aforementioned information and minimizes an expected discounted cost due to passenger waiting and the dispatching of the shuttle.

Results for this type of network have implications for many existing systems, mass transit and shipment of goods being obvious examples. The queueing network considered here captures fundamental features of transportation networks because (i) it models service occurring in batches of up to Q customers and (ii) it incorporates a switching cost that reflects the cost of initiating service.

The preceding transportation problem was introduced by Ignall and Kolesar [4,5], who investigated various dispatching schemes for the case where the shuttle carries at most one passenger and the case of infinite shuttle capacity. For the sake of practical application, Ignall and Kolesar devoted considerable attention to ad hoc schemes based solely on the number of customers at the terminal where the shuttle waits. Deb [2] was the first to solve the optimal dispatching problem for a two-node network under perfect information. For a continuous time version of the problem with complete state observations and with equal linear holding and dispatching costs at both terminals, Deb characterized the nature of an optimal dispatching policy as being of threshold type: Dispatch the shuttle if, and only if, the number of customers at the present terminal exceeds a threshold depending on the number of customers at the other terminal. Moreover, Deb discovered that the threshold is a monotone nonincreasing function of the queue length at the terminal opposite the shuttle and takes values in only a finite set. Dror [3] treated the problem of Deb for the case where the shuttle can carry at most one passenger. Interestingly, Dror used an idea proposed by Ignall and Kolesar [4] to analyze the network as a modified M/G/1 system and thereby determined the monotonicity property for the threshold function.

Our contribution is the analysis of the shuttle dispatching problem with delayed information as stated at the beginning of this section. (i) We prove that an optimal shuttle dispatching policy at any terminal at any time is characterized by a threshold that depends on the probability distribution of the number of customers at the other terminal. In addition, we prove that the thresholds are monotone functions of the most recent delayed observation. (ii) We expose further qualitative properties of an optimal policy that reduce the computational effort required to determine optimal threshold functions; these results are new even in the special case of perfect information ($I = 0$). First, for the case of linear holding and dispatching costs, which are not equal at the two terminals, we derive one necessary and several sufficient conditions for optimally dispatching the shuttle from a given terminal. Second, for the case of linear and symmetrical costs in the network, we prove that the thresholds characterizing an optimal dispatching policy take values in a finite set. This feature simplifies the search for an optimal policy.

We formulate the dispatching problem with imperfect information in Section 2. In Section 3, we determine the structure of an optimal policy for the general problem of Section 2. Section 4 presents further qualitative properties of an optimal policy for the case of linear holding costs. In Section 5, we discuss problems in which the shuttle trip length distribution depends on the load carried. Conclusions are presented in Section 6.

2. PROBLEM FORMULATION

Consider a single shuttle that provides transportation between two passenger terminals labeled one and two. Let $\delta \in \{1,2\}$ represent the terminal number and denote by $\mathbb{N}$ (resp. $\mathbb{Z}^+$) the positive (resp. nonnegative) integers. In discrete time, customers arrive to each terminal according to an arbitrary, prespecified, independent batch arrival process. The arrival processes at the two terminals are independent of each other and all else. All arrivals to one terminal desire passage to the other and exit the system upon reaching this destination. The shuttle may carry at most $Q \in \mathbb{N}$ passengers per trip. Interterminal trips made from a given terminal require durations that are i.i.d. integer random variables that are independent of the load carried and all else. Denote by $(\Omega,\mathcal{A},\mathrm{P})$ the probability space underlying the previously defined random variables. Let the probability mass function (p.m.f.) governing the number of arrivals to terminal δ at time t be $a_t^\delta \triangleq (a_t^\delta(0), a_t^\delta(1), \ldots, a_t^\delta(M))$ for some $M \in \mathbb{Z}^+$ and the p.m.f. governing trip durations from node δ be $b^\delta \triangleq (b^\delta(D_1), b^\delta(D_1+1), \ldots, b^\delta(D_2))$ where $0 < D_1 \leq D_2 \leq \infty$.

During shuttle trips, no control actions are possible. However, if at time t the shuttle has either just arrived to or is waiting at one of the terminals, one of two control actions, $U_t \in \{0,1\}$, must be taken: dispatch ($U_t = 1$) or wait ($U_t = 0$). We assume the shuttle controller exercises decisions immediately following the instant at which potential arrivals enter the system. Control $U_t = 0$

causes the shuttle to be held at the present node until time $t+1$, whereas $U_t = 1$ dispatches the shuttle (and no further control actions are possible until reaching the destination). Upon dispatching the shuttle, as many passengers as possible are boarded, but never more than Q (the shuttle capacity).

The control decision at time t is based on the information available to the shuttle controller up to and including time t. The controller has perfect memory of its observations and control actions. When at a given terminal δ at time t, the controller knows (i) the initial ($t = 0$) queue length of both terminals, (ii) the history of the arrival process for terminal δ up to and including time t, and (iii) the history of the arrival process up to and including time $(t-I)^+$ for the other terminal, where $I \in \mathbb{Z}^+$ and $(t-I)^+ \triangleq \max(t-I, 0)$. Clearly this information combined with prior control decisions yields the queue length of terminal δ at t and that of the other terminal at $(t-I)^+$. Thus, at any time t the controller has *imperfect information* about the state of the system (the queue lengths at both nodes). We make the important assumption that $I \leq D_1$: The information delay does not exceed the lower bound on interterminal transit time. As a result of the preceding information pattern, the controller's information state at time t can be represented by the triplet $(x_t, y_{(t-I)^+}, 1)$ (resp. $(x_{(t-I)^+}, y_t, 2)$), which has the following interpretation: (i) the last component, 1 (resp. 2), indicates the terminal at which the shuttle is at time t; (ii) x_t (resp. y_t) indicates the number of customers seen by the controller when the shuttle is at terminal 1 (resp. 2) at time t; (iii) $y_{(t-I)^+}$ (resp. $x_{(t-I)^+}$) is the queue length of terminal 2 (resp. 1) at time $(t-I)^+$. The p.m.f. for the queue length of terminal 2 (resp. 1) at time t can be easily computed from $y_{(t-I)^+}$ (resp. $x_{(t-I)^+}$) and the arrival process $\{a_s^2 : (t-I)^+ + 1 \leq s \leq t\}$ (resp. $\{a_s^1 : (t-I)^+ + 1 \leq s \leq t\}$).

We consider real valued nondecreasing convex instantaneous holding cost rate functions $c_1(x)$ and $c_2(y)$ for the customers in terminals one and two, respectively. We assume that

$$c_\delta(0) = 0. \tag{2.1}$$

We also include an affine dispatching cost, thus introducing a tradeoff between the incentive to provide prompt service and the competing incentive to minimize the number of trips. That is, a dispatching (switching) cost of $K_\delta + R_\delta z$ units is incurred at each instant the shuttle is dispatched carrying z customers from terminal δ. We assume that

$$0 \leq R_\delta \leq c_\delta(1), \tag{2.2}$$

and that passengers incur no further holding or carrying costs after boarding the shuttle.

The objective is to determine a nonanticipative policy g^* that minimizes, over an infinite horizon, the total expected β-discounted cost ($0 < \beta < 1$) due to the waiting as well as the dispatching of customers at the two terminals under the information pattern described earlier. If $n_\delta^h(t)$ is the number of customers held at terminal δ from time t until $t+1$, and $n_\delta^d(t)$ is the number dispatched from terminal δ at time t, then an optimal policy g^* is one that minimizes

$$\mathrm{E}^{g^*}\left\{\sum_{t=1}^{\infty}\beta^t\sum_{\delta=1}^{2}[c_\delta(n_\delta^h(t)) + \mathbf{1}(U_t = 1, \delta_t = \delta)(K_\delta + R_\delta n_\delta^d(t))]\right\}, \tag{2.3}$$

where δ_t denotes the terminal at which the shuttle resides at time t, and $\mathbf{1}(s)$ is the indicator function of event s. Without loss of optimality (see Chapter 6 of Kumar and Varaiya [6]), we restrict attention to the class of policies that are functions of the information state.

We proceed as follows. In Section 3, we consider the problem formulated above with nondecreasing convex holding costs and derive qualitative properties of the optimal policy g^* using stochastic dynamic programming arguments. In Section 4, we examine the case where the holding costs are linear. We derive, via coupling arguments, necessary and sufficient conditions for dispatching from a given terminal. These conditions simplify the search for an optimal threshold policy.

3. OPTIMALITY OF THRESHOLD POLICIES

We adopt the stochastic dynamic programming approach to the problem formulated in the previous section. We start with the finite horizon problem and then extend the results to the infinite horizon by limiting arguments.

We note that for the finite horizon problem our results hold when $\beta = 1$ and the arrival process is time-varying (provided the independence assumption is retained). We believe that the finite horizon undiscounted problem with time-varying independent batch arrival processes is not unrealistic for shuttle dispatching systems such as people movers (see Barnett [1]). For this reason, our finite horizon solution treats the case of time-varying arrivals and the possibility of $\beta = 1$.

3.1. The Finite Horizon Problem

Consider a finite horizon T and $0 < \beta \leq 1$. Let $V_t^\beta(x,y,\delta)$ be the minimal expected β-discounted cost-to-go from time t through T conditioned on the information state (x,y,δ). Let $x, y \in \mathbb{Z}^+$, $\delta \in \{1,2\}$, and $t \in \{1,2,\ldots,T\}$. Because β is fixed, we let $V_t \triangleq V_t^\beta$. The optimality equation is

$$V_t(x,y,\delta) = \min(h_t(x,y,\delta), d_t(x,y,\delta)), \tag{3.1}$$

where $h_t(x,y,\delta)(d_t(x,y,\delta))$ is the minimal expected β-discounted cost-to-go from time t through T conditioned on state (x,y,δ) and the decision to hold (dispatch) at t. The functions h_t and d_t are given by

$$h_{T+1}(\cdot,\cdot,\cdot) \triangleq 0, \tag{3.2}$$

$$h_t(x,y,1) = c_1(x) + E\{c_2(\bar{Y}_t) + \beta V_{t+1}(X_{t+1}, Y_{(t+1-l)^+}, 1) \mid X_t = x, Y_{(t-l)^+} = y, U_t = 0\} \tag{3.3}$$

$$h_t(x,y,2) = c_2(y) + E\{c_1(\bar{X}_t) + \beta V_{t+1}(X_{(t+1-I)^+}, Y_{t+1}, 2) \mid X_{(t-I)^+} = x,\ Y_t = y,\ U_t = 0\} \tag{3.4}$$

$$d_{T+1}(\cdot,\cdot,\cdot) \triangleq 0, \tag{3.5}$$

$$d_t(x,y,1) = K_1 + R_1(x \wedge Q) + E\Bigg\{ \sum_{j=t}^{(t+\tau_1-1)\wedge T} \beta^{j-t}(c_1(\bar{X}_j) + c_2(\bar{Y}_j)) + \mathbf{1}(t+\tau_1 \le T)\beta^{\tau_1} V_{t+\tau_1}(X_{t+\tau_1-I}, Y_{t+\tau_1}, 2) \,\Bigg|\, X_t = x,\ Y_{(t-I)^+} = y,\ U_t = 1 \Bigg\} \tag{3.6}$$

$$d_t(x,y,2) = K_2 + R_2(y \wedge Q) + E\Bigg\{ \sum_{j=t}^{(t+\tau_2-1)\wedge T} \beta^{j-t}(c_1(\bar{X}_j) + c_2(\bar{Y}_j)) + \mathbf{1}(t+\tau_2 \le T)\beta^{\tau_2} V_{t+\tau_2}(X_{t+\tau_2}, Y_{t+\tau_2-I}, 1) \,\Bigg|\, X_{(t-I)^+} = x,\ Y_t = y,\ U_t = 1 \Bigg\} \tag{3.7}$$

where the following notation has been used: $a \vee b$ and $a \wedge b$ denote the maximum and minimum, respectively, of a and b and $(a)^+ = a \vee 0$; X_t and Y_t are random variables describing the queue lengths of terminals one and two, respectively, at time t just prior to the application of U_t; $\bar{X}_t$ and $\bar{Y}_t$ are random variables denoting the queue lengths of nodes one and two immediately following the application of U_t; and the expectations are with respect to the arrival processes and trip durations. Let $\tau_\delta, \delta \in \{1,2\}$ be the random variable defining the duration of the shuttle trip from node δ to the other node. According to Section 2, the p.m.f. of τ_δ is $b^\delta \triangleq (b^\delta(D_1), \ldots, b^\delta(D_2))$. Define

$$A_t^\delta(\tau, m) \triangleq P\left(\sum_{j=t-\tau+1}^{t} \alpha_j^\delta = m \right), \tag{3.8}$$

where α_j^δ is the number of customers that arrive to terminal δ at time j. For $t \ge 1$, $A_t^\delta(\tau, m)$ is the mth element of the τ-fold convolution $a_{t-\tau+1}^\delta * \cdots * a_t^\delta$. Define $A_t^\delta(0,m) = \mathbf{1}(m=0)$.

Based on the preceding terminology, we obtain the following alternative form for Eqs. (3.3), (3.4), (3.6), and (3.7):

$$h_t(x,y,1) = c_1(x) + \sum_{\ell=0}^{(t\wedge I)M} A_t^2(t \wedge I, \ell) c_2(y+\ell) + \beta \sum_{k,n=0}^{M} a_{t+1}^1(k) a_{t+1-I}^2(n) V_{t+1}(x+k, y+n, 1), \tag{3.9}$$

$$h_t(x,y,2) = c_2(y) + \sum_{\ell=0}^{(t\wedge I)M} A_t^1(t\wedge I,\ell)c_1(x+\ell)$$

$$+\beta \sum_{k,n=0}^{M} a_{t+1}^2(k)a_{t+1-I}^1(n)V_{t+1}(x+n,\ y+k,2), \tag{3.10}$$

$$d_t(x,y,1) = K_1 + R_1(x\wedge Q)$$

$$+\sum_{\tau=D_1}^{D_2} b^1(\tau)\left\{\sum_{j=t}^{(t+\tau-1)\wedge T}\beta^{j-t}\left[\sum_{k=0}^{(j-t)M} A_j^1(j-t,k)c_1((x-Q)^+ + k)\right.\right.$$

$$\left.+\sum_{\ell=0}^{(j-(t-I)^+)M} A_j^2(j-(t-I)^+,\ell)c_2(y+\ell)\right]$$

$$+\mathbf{1}(t+\tau\le T)\beta^\tau \sum_{m=0}^{(\tau-I)M}\ \sum_{n=0}^{(t+\tau-(t-I)^+)M}$$

$$\times A_{t+\tau-I}^1(\tau-I,m)A_{t+\tau}^2(t+\tau-(t-I)^+,n)$$

$$\left.\times V_{t+\tau}((x-Q)^+ + m,\ y+n,2)\right\}, \tag{3.11}$$

$$d_t(x,y,2) = K_2 + R_2(y\wedge Q)$$

$$+\sum_{\tau=D_1}^{D_2} b^2(\tau)\left\{\sum_{j=t}^{(t+\tau-1)\wedge T}\beta^{j-t}\left[\sum_{k=0}^{(j-t)M} A_j^2(j-t,k)c_2((y-Q)^+ + k)\right.\right.$$

$$\left.+\sum_{\ell=0}^{(j-(t-I)^+)M} A_j^1(j-(t-I)^+,\ell)c_1(x+\ell)\right]$$

$$+\mathbf{1}(t+\tau\le T)\beta^\tau \sum_{m=0}^{(\tau-I)M}\ \sum_{n=0}^{(t+\tau-(t-I)^+)M}$$

$$\times A_{t+\tau-I}^2(\tau-I,m)A_{t+\tau}^1(t+\tau-(t-I)^+,n)$$

$$\left.\times V_{t+\tau}(x+n,(y-Q)^+ + m,1)\right\}. \tag{3.12}$$

We point out that in Eqs. (3.9) and (3.10) we have $a_{t+1-I}^{\delta}(n) = \mathbf{1}(n=0)$ (i.e., no arrivals) for $t < I$ because delayed observations are received only at times $I+1$, $I+2, \ldots$. The expression for $d_t(x,y,1)$ explicitly uses the fact that once the shuttle is dispatched, no further control actions are possible until time $t+\tau_1$, when the shuttle arrives at node two. Upon arrival at two, the shuttle controller observes $X_{t+\tau_1-I}$ because $\tau_1 \ge D_1 \ge I$. Similar comments apply to $d_t(x,y,2)$.

We note that the symmetry of the problem with respect to the two terminals, which can be seen in the dynamic programming equations, implies that any property proved for one terminal will hold analogously for the other as well. In the proofs, we frequently refer to this symmetry.

Our investigation of the qualitative properties of the optimal dispatching policy is based on the study of the properties of expected incremental β-discounted cost-to-go functions, which we define below. Let

$$\Delta_1 W_t(x,y,\delta) \triangleq W_t(x+1,y,\delta) - W_t(x,y,\delta) \tag{3.13}$$

$$\Delta_2 W_t(x,y,\delta) \triangleq W_t(x,y+1,\delta) - W_t(x,y,\delta), \tag{3.14}$$

where W may be replaced by the symbols h, d, and V. Thus for state (x,y,δ) at time t, $\Delta_\delta V_t(x,y,\delta)$ is the expected incremental (β-discounted) cost-to-go function resulting from the placement of an additional customer in node δ at t. If $i \neq \delta$, then $\Delta_i V_t(x,y,\delta)$ is the expected incremental (β-discounted) cost-to-go induced by an extra customer in node i at $(t-I)^+$. Similar interpretations can be given to $\Delta_1 d_t(x,y,\delta)$, $\Delta_2 d_t(x,y,\delta)$, $\Delta_1 h_t(x,y,\delta)$, and $\Delta_2 h_t(x,y,\delta)$.

Using Eqs. (3.2), (3.5), and (3.9)–(3.12), we find that

$$\Delta_i h_{T+1}(\cdot,\cdot,\cdot) = 0 \qquad \text{for } i = 1,2, \tag{3.15}$$

$$\begin{aligned}\Delta_1 h_t(x,y,1) &= c_1(x+1) - c_1(x)\\ &\quad + \beta \sum_{k,n=0}^{M} a^1_{t+1}(k) a^2_{t+1-I}(n) \Delta_1 V_{t+1}(x+k, y+n, 1),\end{aligned} \tag{3.16}$$

$$\begin{aligned}\Delta_2 h_t(x,y,1) &= \sum_{\ell=0}^{(t\wedge I)M} A^2_t(t\wedge I, \ell)(c_2(y+1+\ell) - c_2(y+\ell))\\ &\quad + \beta \sum_{k,n=0}^{M} a^1_{t+1}(k) a^2_{t+1-I}(n) \Delta_2 V_{t+1}(x+k, y+n, 1),\end{aligned} \tag{3.17}$$

$$\Delta_i d_{T+1}(\cdot,\cdot,\cdot) = 0 \qquad \text{for } i = 1,2, \tag{3.18}$$

$$\begin{aligned}\Delta_1 d_t(x,y,1) &= \mathbf{1}(x<Q)R_1\\ &\quad + \sum_{\tau=D_1}^{D_2} b^1(\tau)\Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \beta^{j-t} \sum_{k=0}^{(j-t)M} A^1_j(j-t,k)\\ &\quad \times (c_1((x+1-Q)^+ + k) - c_1((x-Q)^+ + k))\\ &\quad + \mathbf{1}(t+\tau \le T)\beta^\tau \sum_{m=0}^{(\tau-I)M}\\ &\quad \times \sum_{n=0}^{(t+\tau-(t-I)^+)M} A^1_{t+\tau-I}(\tau-I,m) A^2_{t+\tau}(t+\tau-(t-I)^+, n)\\ &\quad \times \big(V_{t+\tau}((x+1-Q)^+ + m, y+n, 2)\\ &\quad - V_{t+\tau}((x-Q)^+ + m, y+n, 2)\big)\Bigg\},\end{aligned} \tag{3.19}$$

$$\Delta_2 d_t(x,y,1) = \sum_{\tau=D_1}^{D_2} b^1(\tau)\Bigg\{\sum_{j=t}^{(t+\tau-1)\wedge T} \beta^{j-t}$$
$$\times \sum_{\ell=0}^{(j-(t-l)^+)M} A_j^2(j-(t-l)^+,\ell)(c_2(y+1+\ell)-c_2(y+\ell))$$
$$+ \mathbf{1}(t+\tau\le T)\beta^\tau \sum_{m=0}^{(\tau-l)M}\sum_{n=0}^{(t+\tau-(t-l)^+)M} A^1_{t+\tau-l}(\tau-l,m)$$
$$\times A^2_{t+\tau}(t+\tau-(t-l)^+,n)\Delta_2 V_{t+\tau}((x-Q)^+ + m, y+n,2)\Bigg\}. \tag{3.20}$$

Having defined the incremental expected cost-to-go functions, we now proceed to investigate the qualitative properties of an optimal dispatching policy. We proceed via several lemmas. Lemma 3.1 states that the incremental expected cost-to-go induced by an additional customer is nonnegative.

LEMMA 3.1: *For any* $x, y \in \mathbb{Z}^+$, $1 \le t \le T$, *and* $i,\delta \in \{1,2\}$;

$$\Delta_i V_t(x,y,\delta) \ge 0.$$

PROOF: Consider an arbitrary information state (u,v,δ) at time t. Let g^* denote an optimal policy. Along each realization $\omega \in \Omega$ of the arrival process and trip durations, let $\sigma(\omega)$ be the vector of times at which the shuttle is dispatched under g^* from time t upon given the information state (u,v,δ) at t.

Subtract a customer from either node at t, thus yielding $(u-1,v,\delta)$ or $(u,v-1,\delta)$. Consider this new information state at t, and a policy g that, along each $\omega \in \Omega$, dispatches the shuttle at the times indicated by $\sigma(\omega)$. Such a policy is feasible. Denote by $V_t^g(u-1,v,\delta)$ (resp. $V_t^g(u,v-1,\delta)$) the expected β-discounted cost-to-go corresponding to g when the information state at t is $(u-1,v,\delta)$ (resp. $(u,v-1,\delta)$). Because of the definition of g,

$$V_t(u,v,\delta) \ge V_t^g(u-1,v,\delta), \tag{3.21}$$
$$V_t(u,v,\delta) \ge V_t^g(u,v-1,\delta). \tag{3.22}$$

Moreover, because g may not be optimal,

$$V_t^g(u-1,v,\delta) \ge V_t(u-1,v,\delta), \tag{3.23}$$
$$V_t^g(u,v-1,\delta) \ge V_t(u,v-1,\delta). \tag{3.24}$$

Combination of Eqs. (3.21)–(3.24) yields the result. ■

The following two lemmas are presented in an abstract light to emphasize their fundamental nature. Lemma 3.2 can be interpreted as follows. If the incremental expected cost-to-go from t given that the shuttle is held at t is greater than the incremental expected cost-to-go given the shuttle is dispatched at t, then the incremental expected optimal cost-to-go lies between them and can be no smaller than that given when the shuttle is dispatched at t.

LEMMA 3.2: *Let* $h(x), h(y), d(x), d(y) \in \mathbb{R}$, *and* $V(\cdot) \triangleq \min(h(\cdot), d(\cdot))$. *If* $h(x) - h(y) \geq d(x) - d(y)$, *then* $h(x) - h(y) \geq V(x) - V(y) \geq d(x) - d(y)$.

Discussion. This property is most easily seen graphically and is merely due to the nature of the minimum function. The result is similar to that in Deb [2] but the proof differs. We can interpret $h(x) - h(y)$ and $d(x) - d(y)$ as incremental cost-to-go functions for state y and time t given that at t the shuttle must be held and dispatched, respectively.

PROOF: We consider three cases.

Case I: Let $h(y) \geq d(y)$. Then beginning with the hypothesis, we find $0 \leq h(x) - h(y) + d(y) - d(x) \leq h(x) - d(x)$. Hence, $d(x) \leq h(x)$ and consequently $V(x) - V(y) = d(x) - d(y) \leq h(x) - h(y)$.

Case II: Let $h(y) < d(y)$ and $h(x) < d(x)$. Then $V(x) - V(y) = h(x) - h(y) \geq d(x) - d(y)$.

Case III: Let $h(y) < d(y)$ and $h(x) \geq d(x)$. Because $V(x) - V(y) = d(x) - h(y)$, we use the Case III hypothesis directly to conclude $h(x) - h(y) \geq V(x) - V(y) \geq d(x) - d(y)$. ■

Remark: Note that Lemma 3.2 would still hold for $V(\cdot) \triangleq \max(h(\cdot), d(\cdot))$.

Lemma 3.3 gives a sufficient condition for the minimum of two submodular functions to be submodular (see Chapter 1, Section 4 of Ross [7]). This result will be used in Lemma 3.4 to prove the submodularity of the value function. Before proceeding with Lemma 3.3, consider functions h, d, and V, where $h(a,b): \mathbb{Z}^+ \times \mathbb{Z}^+ \to \mathbb{R}$, $d(a,b): \mathbb{Z}^+ \times \mathbb{Z}^+ \to \mathbb{R}$, and define $V(a,b) = \min(h(a,b), d(a,b))$.

LEMMA 3.3: *If the following relations hold for all* $a, b \in \mathbb{Z}^+$:

(i) $h(a+1,b) - h(a,b) \geq h(a+1,b+1) - h(a,b+1)$,

(ii) $d(a+1,b) - d(a,b) \geq d(a+1,b+1) - d(a,b+1)$,

(iii) $h(a+1,b) - h(a,b) \geq d(a+1,b) - d(a,b)$,

(iv) $h(a,b+1) - h(a,b) \geq d(a,b+1) - d(a,b)$;

then $V(a+1,b) - V(a,b) \geq V(a+1,b+1) - V(a,b+1)$ *for all* $a, b, \in \mathbb{Z}^+$.

Discussion. Assumptions (i) and (ii) require h and d to be submodular functions. In (iii) and (iv), we require h to increase at least as quickly as d with respect to either argument.

PROOF: The proof makes use of four cases.

Case I: Let $h(a,b) > d(a,b)$. Then

$$\begin{aligned} V(a+1,b) - V(a,b) &\geq d(a+1,b) - d(a,b) \quad \text{by Lemma 3.2 applied to (iii)} \\ &\geq d(a+1,b+1) - d(a,b+1) \quad \text{by (ii)} \\ &= d(a+1,b+1) - V(a,b+1) \quad \text{because the} \\ &\quad \text{hypothesis and (iv) imply } h(a,b+1) > d(a,b+1) \\ &\geq V(a+1,b+1) - V(a,b+1). \end{aligned}$$

Case II: Let $h(a,b) \leq d(a,b)$ and $h(a+1,b) \leq d(a+1,b)$. Then

$$\begin{aligned} V(a+1,b) - V(a,b) &= h(a+1,b) - h(a,b) \\ &\geq h(a+1,b+1) - h(a,b+1) \quad \text{by (i)} \\ &\geq V(a+1,b+1) - V(a,b+1) \\ &\qquad \text{by Lemma 3.2 applied to (iii).} \end{aligned}$$

Case III: Let $h(a,b) \leq d(a,b)$, $h(a+1,b) > d(a+1,b)$, and $h(a,b+1) > d(a,b+1)$. Then

$$\begin{aligned} V(a+1,b) - V(a,b) &\geq d(a+1,b) - d(a,b) \quad \text{by hypothesis} \\ &\geq d(a+1,b+1) - d(a,b+1) \quad \text{by (ii)} \\ &= d(a+1,b+1) - V(a,b+1) \quad \text{by hypothesis} \\ &\geq V(a+1,b+1) - V(a,b+1). \end{aligned}$$

Case IV: Let $h(a,b) \leq d(a,b)$, $h(a+1,b) > d(a+1,b)$, and $h(a,b+1) \leq d(a,b+1)$. Then

$$\begin{aligned} &V(a+1,b) - V(a,b) - [V(a+1,b+1) - V(a,b+1)] \\ &\quad = d(a+1,b) - h(a,b) - V(a+1,b+1) + h(a,b+1) \\ &\quad \geq d(a+1,b) - h(a,b) - d(a+1,b+1) + h(a,b+1) \\ &\quad \geq d(a,b+1) - d(a,b) + d(a+1,b) - d(a+1,b+1) \quad \text{by (iv)} \\ &\quad \geq d(a,b+1) - d(a,b+1) + d(a+1,b+1) - d(a+1,b+1) \quad \text{by (ii)} \\ &\quad = 0. \end{aligned}$$

■

We now present Lemma 3.4, the backbone of our analysis. The result of Lemma 3.4 is a sufficient condition that guarantees the threshold property of an optimal dispatching policy as well as the monotonicity of the optimal threshold functions.

LEMMA 3.4: *For any* $x, y \in \mathbb{Z}^+$, $1 \le t \le T + 1$, *and* $i, j \in \{1,2\}$ *such that* $i \neq j$;

(i) $\Delta_i d_t(x, y, i) \le \Delta_i h_t(x, y, i)$,

(ii) $\Delta_j d_t(x, y, j) \le \Delta_j h_t(x, y, j)$,

and the following relations hold for the value function:

(iiia) $\Delta_1 V_t(x, y, 1) \ge \Delta_1 V_t(x, y + 1, 1)$,

(iiib) $\Delta_2 V_t(x, y, 2) \ge \Delta_2 V_t(x + 1, y, 2)$,

(iva) $\Delta_2 V_t(x, y, 1) \ge \Delta_2 V_t(x + 1, y, 1)$,

(ivb) $\Delta_1 V_t(x, y, 2) \ge \Delta_1 V_t(x, y + 1, 2)$.

Discussion. Part (i) is the key statement of the lemma as it is sufficient to guarantee the threshold property of the optimal policy. Its meaning becomes more clear when rewritten as follows (let $i = 1$ for convenience; the interpretation is similar for $i = 2$):

$$h_t(x, y, 1) - d_t(x, y, 1) \le h_t(x + 1, y, 1) - d_t(x + 1, y, 1).$$

The difference $h_t(x, y, 1) - d_t(x, y, 1)$ provides the incentive to dispatch the shuttle at time t when the information state is $(x, y, 1)$. For $i = 1$, part (i) states that the incentive to dispatch from terminal one increases as the number of customers waiting at that terminal increases and the controller's perception of the number of customers at terminal two (expressed by y) remains fixed.

Part (ii) of the lemma establishes the monotonicity of the threshold with respect to the delayed observation. Again considering $i = 1$ for convenience, rewriting (ii) as

$$h_t(x, y, 1) - d_t(x, y, 1) \le h_t(x, y + 1, 1) - d_t(x, y + 1, 1),$$

we see that the incentive to dispatch the shuttle from terminal one at time t increases as the number of customers observed I time units ago at terminal two increases and the queue length at terminal one remains fixed.

Combining (i) and (ii), we see that if it is optimal to dispatch at t while at state $(x, y, 1)$, then it is also optimal to dispatch under both states $(x + 1, y, 1)$ and $(x, y + 1, 1)$ at time t.

Parts (i) and (ii) imply the main results of the paper. The remaining parts, (iiia), (iiib), (iva), and (ivb), express the submodularity of the value function and are used to support the proof of (i) and (ii).

PROOF: The proof proceeds by induction. Because $h_{T+1}(\cdot,\cdot,\cdot) = 0$ and $d_{T+1}(\cdot,\cdot,\cdot) = 0$, all the statements of the lemma are true for time $T + 1$. Assume the lemma holds at $t + 1, t + 2, \ldots, T + 1$.

PROOF OF (i): We prove part (i) for $i = 1$ by considering three cases. In each case we prove the result by a series of inequalities and explain how these inequalities are obtained. The result holds for $i = 2$ by symmetry.

Case I: Suppose $x < Q$. Then, using Eq. (3.16) we obtain

$$\begin{aligned}\Delta_1 h_t(x, y, 1) &= c_1(x+1) - c_1(x) \\ &\quad + \beta \sum_{k,n} a^1_{t+1}(k) a^2_{t+1-l}(n) \Delta_1 V_{t+1}(x+k, y+n, 1) \\ &\geq c_1(x+1) - c_1(x) \\ &\geq R_1 \\ &= \Delta_1 d_t(x, y, 1).\end{aligned}$$

The first inequality follows by Lemma 3.1 and the second by Eqs. (2.1) and (2.2) and the convexity of $c_1(\cdot)$. The last step follows from Eq. (3.19) because under the hypothesis, $(x+1-Q)^+ = (x-Q)^+ = 0$.

Case II: Suppose $x \geq Q$ and $t + D_1 \leq T$. Then

$$\begin{aligned}&\Delta_1 h_t(x, y, 1) \\ &\quad = c_1(x+1) - c_1(x) + \beta \sum_{k,n} a^1_{t+1}(k) a^2_{t+1-l}(n) \Delta_1 V_{t+1}(x+k, y+n, 1) \\ &\quad \geq c_1(x+1-Q) - c_1(x-Q) + \beta \sum_{k,n} a^1_{t+1}(k) a^2_{t+1-l}(n) \Delta_1 d_{t+1}(x+k, y+n, 1)\end{aligned} \tag{3.25}$$

$$\begin{aligned}&= c_1(x+1-Q) - c_1(x-Q) + \beta \sum_{k,n} a^1_{t+1}(k) a^2_{t+1-l}(n) \Bigg\{ \sum_{\tau=D_1}^{D_2} b^1(\tau) \\ &\quad \times \Bigg[\sum_{j=t+1}^{(t+\tau)\wedge T} \beta^{j-t-1} \sum_{\ell} A^1_j(j-t-1, \ell)(c_1(x-Q+1+k+\ell) \\ &\quad - c_1(x-Q+k+\ell)) + \mathbf{1}(t+1+\tau \leq T)\beta^\tau \sum_{p,q} A^1_{t+1+\tau-l}(\tau - I, p) \\ &\quad \times A^2_{t+1+\tau}(t+1+\tau-(t+1-I)^+, q) \\ &\quad \times \Delta_1 V_{t+1+\tau}(x-Q+k+p, y+n+q, 2) \Bigg] \Bigg\}\end{aligned} \tag{3.26}$$

$$\begin{aligned}&= \sum_{\tau=D_1}^{D_2} b^1(\tau) \Bigg\{ \sum_{j=t}^{(t+\tau)\wedge T} \sum_{L} A^1_j(j-t, L)\beta^{j-t}(c_1(x-Q+1+L) - c_1(x-Q+L)) \\ &\quad + \mathbf{1}(t+1+\tau \leq T)\beta^{\tau+1} \sum_{m,n} A^1_{t+1+\tau-l}(\tau - I + 1, m) \\ &\quad \times A^2_{t+1+\tau}(t+1+\tau-(t-I)^+, n) \Delta_1 V_{t+1+\tau}(x-Q+m, y+n, 2) \Bigg\}\end{aligned} \tag{3.27}$$

$$= \sum_{\tau=D_1}^{D_2} b^1(\tau)\Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \sum_L A_j^1(j-t,L)\beta^{j-t}$$
$$\times (c_1(x-Q+1+L) - c_1(x-Q+L))$$
$$+ \mathbf{1}(t+\tau \le T) \sum_{\ell} A_{t+\tau}^1(\tau,\ell)\beta^{\tau}(c_1(x-Q+1+\ell) - c_1(x-Q+\ell))$$
$$+ \mathbf{1}(t+\tau \le T-1)\beta^{\tau+1} \sum_{m,n} A_{t+1+\tau-I}^1(\tau-I+1,m)$$
$$\times A_{t+1+\tau}^2(t+1+\tau-(t-I)^+,n)\Delta_1 V_{t+1+\tau}(x-Q+m,y+n,2)\Bigg\} \tag{3.28}$$

$$= \sum_{\tau=D_1}^{D_2} b^1(\tau)\Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \sum_L A_j^1(j-t,L)\beta^{j-t}$$
$$\times (c_1(x+1-Q+L) - c_1(x-Q+L))$$
$$+ \mathbf{1}(t+\tau \le T)\beta^{\tau} \sum_{k,i} A_{t+\tau-I}^1(\tau-I,k)A_{t+\tau}^2(t+\tau-(t-I)^+,i)$$
$$\times \Bigg[\sum_m A_{t+\tau}^1(I,m)(c_1(x+1-Q+k+m) - c_1(x-Q+k+m))$$
$$+ \beta \sum_{p,q} a_{t+1+\tau-I}^1(p)a_{t+1+\tau}^2(q)\Delta_1 V_{t+1+\tau}(x-Q+k+p,y+i+q,2)\Bigg]\Bigg\} \tag{3.29}$$

$$= \sum_{\tau=D_1}^{D_2} b^1(\tau)\Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \sum_L A_j^1(j-t,L)\beta^{j-t}$$
$$\times (c_1(x+1-Q+L) - c_1(x-Q+L))$$
$$+ \mathbf{1}(t+\tau \le T)\beta^{\tau} \sum_{k,i} A_{t+\tau-I}^1(\tau-I,k)A_{t+\tau}^2(t+\tau-(t-I)^+,i)$$
$$\times \Delta_1 h_{t+\tau}(x-Q+k,y+i,2)\Bigg\} \tag{3.30}$$

$$\ge \sum_{\tau=D_1}^{D_2} b^1(\tau)\Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \sum_L A_j^1(j-t,L)\beta^{j-t}$$
$$\times (c_1(x+1-Q+L) - c_1(x-Q+L))$$
$$+ \mathbf{1}(t+\tau \le T)\beta^{\tau} \sum_{k,i} A_{t+\tau-I}^1(\tau-I,k)A_{t+\tau}^2(t+\tau-(t-I)^+,i)$$
$$\times \Delta_1 V_{t+\tau}(x-Q+k,y+i,2)\Bigg\} \tag{3.31}$$

$$= \Delta_1 d_t(x,y,1).$$

The first step merely restates Eq. (3.16). In Eq. (3.25), we cite the convexity of $c_1(\cdot)$ and use (i) of the induction hypothesis at time $t+1$ to apply Lemma 3.2. Equation (3.26) is immediate from Eq. (3.19). We group terms and condense the convolutions defining the expectation with respect to the arrival processes to conclude Eq. (3.27). One must be careful in determining $A^2_{t+1+\tau}(t+1+\tau-(t-l)^+, n)$ of Eq. (3.27) to realize that if $t+1 \leq l$, then $a^2_{t+1-l}(0) = 1$. Equations (3.28) and (3.29) present rearrangements of the terms appearing in Eq. (3.27). The intent of this rearrangement is to group terms appropriately so that we relate $\Delta_1 h_t(x,y,1)$ with $\Delta_1 h_{t+\tau}(\cdot,\cdot,2)$. After using the fact that $\Delta_1 V_{T+1}(\cdot,\cdot,\cdot) = 0$ in Eq. (3.29), such a relationship is achieved in Eq. (3.30). By the induction hypothesis and Lemma 3.2, Eq. (3.30) leads to Eq. (3.31), and finally we conclude the result by Eq. (3.19).

Case III: Suppose $x \geq Q$ and $t + D_1 > T$. Then

$$\Delta_1 h_t(x,y,1)$$

$$\geq c_1(x+1-Q) - c_1(x-Q) + \beta \sum_{k,n} a^1_{t+1}(k) a^2_{t+1-l}(n) \Delta_1 d_{t+1}(x+k, y+n, 1) \tag{3.32}$$

$$= c_1(x+1-Q) - c_1(x-Q) + \beta \sum_k a^1_{t+1}(k) \left\{ \sum_{j=t+1}^{T} \beta^{j-t-1} \times \sum_{\ell} A^1_j(j-t-1,\ell)(c_1(x+1-Q+k+\ell) - c_1(x-Q+k+\ell)) \right\} \tag{3.33}$$

$$= c_1(x+1-Q) - c_1(x-Q) + \sum_{j=t+1}^{T} \beta^{j-t} \times \sum_m A^1_j(j-t,m)(c_1(x+1-Q+m) - c_1(x-Q+m)) \tag{3.34}$$

$$= \Delta_1 d_t(x,y,1).$$

Equation (3.32) merely abbreviates the first two steps in the proof of Case II and Eq. (3.33) uses Eq. (3.19) where $x \geq Q$ and $t + \tau \geq t + D_1 > T$. Equation (3.34) merely condenses the expectation with respect to arrivals and the last step follows directly from Eq. (3.19).

Proof of (ii): The proof of (ii) for $i = 1$ follows in three cases. The first case is the most involved and requires (iiib) of the induction hypothesis. The result holds for $i = 2$ by symmetry.

Case I: Suppose $x < Q$ and $t + D_1 \le T$. Then

$$\Delta_2 h_t(x, y, 1)$$

$$= \sum_{\ell} A_t^2(t \wedge I, \ell)(c_2(y + 1 + \ell) - c_2(y + \ell)$$

$$+ \beta \sum_{k,n} a_{t+1}^1(k) a_{t+1-I}^2(n) \Delta_2 V_{t+1}(x + k, y + n, 1)$$

$$\ge \sum_{\ell} A_t^2(t \wedge I, \ell)(c_2(y + 1 + \ell) - c_2(y + \ell))$$

$$+ \beta \sum_{k,n} a_{t+1}^1(k) a_{t+1-I}^2(n) \Delta_2 d_{t+1}(x + k, y + n, 1) \tag{3.35}$$

$$= \sum_{\ell} A_t^2(t \wedge I, \ell)(c_2(y + 1 + \ell) - c_2(y + \ell))$$

$$+ \beta \sum_{k,n} a_{t+1}^1(k) a_{t+1-I}^2(n) \Bigg\{ \sum_{\tau=D_1}^{D_2} b^1(\tau) \Bigg[\sum_{j=t+1}^{(t+\tau)\wedge T} \beta^{j-t-1}$$

$$\times \sum_{i} A_j^2(j - (t + 1 - I)^+, i)(c_2(y + n + 1 + i) - c_2(y + n + i))$$

$$+ \mathbf{1}(t + 1 + \tau \le T)\beta^{\tau} \sum_{p,q} A_{t+1+\tau-I}^1(\tau - I, p)$$

$$\times A_{t+1+\tau}^2(t + 1 + \tau - (t + 1 - I)^+, q)$$

$$\times \Delta_2 V_{t+1+\tau}((x + k - Q)^+ + p, y + n + q, 2) \Bigg] \Bigg\} \tag{3.36}$$

$$\ge \sum_{\ell} A_t^2(t \wedge I, \ell)(c_2(y + 1 + \ell) - c_2(y + \ell))$$

$$+ \sum_{k,n} a_{t+1}^1(k) a_{t+1-I}^2(n) \Bigg\{ \sum_{\tau=D_1}^{D_2} b^1(\tau) \Bigg[\sum_{j=t+1}^{(t+\tau)\wedge T} \beta^{j-t}$$

$$\times \sum_{i} A_j^2(j - (t + 1 - I)^+, i)(c_2(y + n + 1 + i) - c_2(y + n + i))$$

$$+ \mathbf{1}(t + 1 + \tau \le T)\beta^{\tau+1} \sum_{p,q} A_{t+1+\tau-I}^1(\tau - I, p)$$

$$\times A_{t+1+\tau}^2(t + 1 + \tau - (t + 1 - I)^+, q)$$

$$\times \Delta_2 V_{t+1+\tau}(k + p, y + n + q, 2 \Bigg] \Bigg\} \tag{3.37}$$

$$= \sum_{\ell} A_t^2(t \wedge I, \ell)(c_2(y+1+\ell) - c_2(y+\ell)) + \sum_{\tau=D_1}^{D_2} b^1(\tau) \Bigg\{ \sum_{j=t+1}^{(t+\tau)\wedge T} \beta^{j-t}$$

$$\times \sum_L A_j^2(j-(t-I)^+, L)(c_2(y+1+L) - c_2(y+L))$$

$$+ \mathbf{1}(t+1+\tau \le T)\beta^{\tau+1} \sum_{m,k} A^1_{t+1+\tau-I}(\tau - I + 1, m)$$

$$\times A^2_{t+1+\tau}(t+1+\tau-(t-I)^+, k)\Delta_2 V_{t+1+\tau}(m, y+k, 2)\Bigg\} \quad \textbf{(3.38)}$$

$$= \sum_{\tau=D_1}^{D_2} b^1(\tau) \Bigg\{ \sum_{i=t}^{(t+\tau)\wedge T} \beta^{i-t} \sum_{\ell} A_i^2(i-(t-I)^+, \ell)$$

$$\times (c_2(y+1+\ell) - c_2(y+\ell))$$

$$+ \mathbf{1}(t+1+\tau \le T)\beta^{\tau+1} \sum_{m,k} A^1_{t+1+\tau-I}(\tau - I + 1, m)$$

$$\times A^2_{t+1+\tau}(t+1+\tau-(t-I)^+, k)$$

$$\times \Delta_2 V_{t+1+\tau}(m, y+k, 2)\Bigg\} \quad \textbf{(3.39)}$$

$$= \sum_{\tau=D_1}^{D_2} b^1(\tau) \Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \beta^{j-t} \sum_{\ell} A_j^2(j-(t-I)^+, \ell)$$

$$\times (c_2(y+1+\ell) - c_2(y+\ell))$$

$$+ \mathbf{1}(t+\tau \le T)\Bigg[\beta^{\tau} \sum_L A^2_{t+\tau}(t+\tau-(t-I)^+, L)$$

$$\times (c_2(y+1+L) - c_2(y+L))$$

$$+ \beta^{\tau+1} \sum_{m,k} A^1_{t+1+\tau-I}(\tau - I + 1, m)\, A^2_{t+1+\tau}(t+1+\tau-(t-I)^+, k)$$

$$\times \Delta_2 V_{t+1+\tau}(m, y+k, 2)\Bigg]\Bigg\} \quad \textbf{(3.40)}$$

$$= \sum_{\tau=D_1}^{D_2} b^1(\tau) \Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \beta^{j-t} \sum_{\ell} A_j^2(j-(t-I)^+, \ell)$$

$$\times (c_2(y+1+\ell) - c_2(y+\ell))$$

$$+ \mathbf{1}(t+\tau \le T)\beta^{\tau} \sum_{i,L} A^1_{t+\tau-I}(\tau - I, i) A^2_{t+\tau}(t+\tau-(t-I)^+, L)$$

$$\times \Bigg[(c_2(y+1+L) - c_2(y+L))$$

$$+ \beta \sum_{p,q} a^1_{t+\tau-I+1}(p) a^2_{t+\tau+1}(q) \Delta_2 V_{t+\tau+1}(i+p, y+L+q, 2)\Bigg]\Bigg\} \quad \textbf{(3.41)}$$

$$= \sum_{\tau=D_1}^{D_2} b^1(\tau)\Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \beta^{j-t} \sum_{\ell} A_j^2(j-(t-I)^+,\ell)$$

$$\times\, (c_2(y+1+\ell) - c_2(y+\ell))$$

$$+\, \mathbf{1}(t+\tau \le T)\beta^{\tau} \sum_{i,L} A^1_{t+\tau-I}(\tau - I, i)$$

$$\times\, A^2_{t+\tau}(t+\tau-(t-I)^+, L)\Delta_2 h_{t+\tau}(i, y+L, 2)\Bigg\} \tag{3.42}$$

$$\ge \sum_{\tau=D_1}^{D_2} b^1(\tau)\Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \beta^{j-t} \sum_{\ell} A_j^2(j-(t-I)^+,\ell)$$

$$\times\, (c_2(y+1+\ell) - c_2(y+\ell))$$

$$+\, \mathbf{1}(t+\tau \le T)\beta^{\tau} \sum_{i,L} A^1_{t+\tau-I}(\tau - I, i)$$

$$\times\, A^2_{t+\tau}(t+\tau-(t-I)^+, L)\Delta_2 V_{t+\tau}(i, y+L, 2)\Bigg\} \tag{3.43}$$

$$= \Delta_2 d_t(x, y, 1).$$

We begin in the first step by merely restating Eq. (3.17) and invoke (ii) of the induction hypothesis together with Lemma 3.2 to yield Eq. (3.35). Equation (3.36) invokes Eq. (3.20). We apply (iiib) of the induction hypothesis precisely $k - (x + k - Q)^+$ times (because $x < Q$ in Case I) for each k and each time $t + \tau + 1$ to obtain Eq. (3.37). The expectation convolutions are condensed in Eq. (3.38) by recalling that if $t + 1 \le I$, $a^2_{t+1-I}(0) = 1$ and so $A_j^2(j - (t-I)^+, L) = A_j^2(j, L)$. Equation (3.39) further simplifies Eq. (3.38) by noting that for $i = t$, $A_i^2(i - (t-I)^+, \ell) = A_i^2(t \wedge I, \ell)$. We use the fact that $V_{t+1+\tau}(\cdot,\cdot,\cdot) = 0$ if $t + \tau = T$ to justify Eq. (3.40) and expanding the expectation convolutions yields Eq. (3.41). Equation (3.42) is immediate from the definition of $\Delta_2 h_{t+\tau}(i, y + L, 2)$ (analogous to Eq. (3.16)). Part (i) of the induction hypothesis and Lemma 3.2 yield the inequality in Eq. (3.43). The proof of Case I concludes by applying Eq. (3.20).

Case II: Suppose $x \ge Q$ and $t + D_1 \le T$. The proof of this case is the same as for Case I with one exception. The step of Eq. (3.37) and its appeal to (iiib) of the induction hypothesis is no longer necessary because for $X \ge Q$, $(x + k - Q)^+ = x - Q + k$; i.e., none of the k arrivals at time $t + 1$ will be dispatched (see Eq. (3.35)) at $t + 1$. The remainder of the argument proceeds as before.

Case III: Let $t + D_1 > T$. Then

$$\begin{aligned}\Delta_2 h_t(x, y, 1) &\geq \sum_{\ell} A_t^2(t \wedge I, \ell)(c_2(y + 1 + \ell) - c_2(y + \ell)) \\ &\quad + \beta \sum_{k,n} a_{t+1}^1(k) a_{t+1-I}^2(n) \Delta_2 d_{t+1}(x + k, y + n, 1)\end{aligned} \tag{3.44}$$

$$\begin{aligned}&= \sum_{\ell} A_t^2(t \wedge I, \ell)(c_2(y + 1 + \ell) - c_2(y + \ell)) \\ &\quad + \beta \sum_{k,n} a_{t+1}^2(k) a_{t+1-I}^2(n) \Bigg\{ \sum_{j=t+1}^{T} \beta^{j-t-1} \\ &\quad \times \sum_{L} A_j^2(j + (t + 1 - I)^+, L)(c_2(y + 1 + n + L) - c_2(y + n + L)) \Bigg\}\end{aligned} \tag{3.45}$$

$$\begin{aligned}&= \sum_{\ell} A_t^2(t \wedge I, \ell)(c_2(y + 1 + \ell) - c_2(y + \ell)) \\ &\quad + \sum_{j=t+1}^{T} \beta^{j-t} \sum_{k} A_j^2(j - (t - I)^+, k)(c_2(y + 1 + k) - c_2(y + k))\end{aligned} \tag{3.46}$$

$$= \Delta_2 d_t(x, y, 1).$$

We begin in Eq. (3.44) by using the first two steps of the proof in Case I. Equation (3.45) follows from Eq. (3.20) and the fact that $t + \tau \geq t + D_1 > T$. Recalling that $a_{t+1-I}^2(0) = 1$ for $t + 1 - I \leq 0$ yields Eq. (3.46). We conclude the result by using Eq. (3.20) and noting that $j - (t - I)^+ = t \wedge I$ for $j = t$.

Proof of (iii): Statement (iii) was used in the proof of (ii); the proof of (iii) (as well as (iv)) is based on Lemma 3.3. Proving (iii) for the functions $\Delta h(\cdot)$ and $\Delta d(\cdot)$ allows us to apply Lemma 3.3 and conclude the property for the value function.

We begin by verifying that $\Delta_1 h_t(x, y, 1) \geq \Delta_1 h_t(x, y + 1, 1)$. By Eq. (3.16) and (iiia) of the induction hypothesis at time $t + 1$ we obtain

$$\begin{aligned}\Delta_1 h_t(x, y, 1) &= c_1(x + 1) - c_1(x) + \beta \sum_{k,n} a_{t+1}^1(k) a_{t+1-I}^2(n) \Delta_1 V_{t+1}(x + k, y + n, 1) \\ &\geq c_1(x + 1) - c_1(x) + \beta \sum_{k,n} a_{t+1}^1(k) a_{t+1-I}^2(n) \Delta_1 V_{t+1}(x + k, y + 1 + n, 1) \\ &= \Delta_1 h_t(x, y + 1, 1).\end{aligned}$$

We prove a similar property for $\Delta d_t(\cdot)$ by considering two cases, $x < Q$ and $x \geq Q$. If $x < Q$, we simply note that $\Delta_1 d_t(x, y, 1) = R_1 = \Delta_1 d_t(x, y + 1, 1)$. On the other hand, if $x \geq Q$, then by Eq. (3.19) and the induction hypothesis (ivb) at time $t + \tau$ for each τ we get

$$\begin{aligned}
&\Delta_1 d_t(x, y, 1) \\
&= \sum_{\tau=D_1}^{D_2} b_1(\tau) \Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \beta^{j-t} \sum_k A_j^1(j-t, k) \\
&\quad \times (c_1(x + 1 - Q + k) - c_1(x - Q + k)) \\
&\quad + \mathbf{1}(t + \tau \leq T)\beta^\tau \sum_{m,n} A_{t+\tau-I}^1(\tau - I, m) A_{t+\tau}^2(t + \tau - (t - I)^+, n) \\
&\quad \times \Delta_1 V_{t+\tau}(x - Q + m, y + n, 2) \Bigg\} \\
&\geq \sum_{\tau=D_1}^{D_2} b_1(\tau) \Bigg\{ \sum_{j=t}^{(t+\tau-1)\wedge T} \beta^{j-t} \sum_k A_j^1(j-t, k) \\
&\quad \times (c_1(x + 1 - Q + k) - c_1(x - Q + k)) \\
&\quad + \mathbf{1}(t + \tau \leq T)\beta^\tau \sum_{m,n} A_{t+\tau-I}^1(\tau - I, m) A_{t+\tau}^2(t + \tau - (t - I)^+, n) \\
&\quad \times \Delta_1 V_{t+\tau}(x - Q + m, y + 1 + n, 2) \Bigg\} \\
&= \Delta_1 d_t(x, y + 1, 1).
\end{aligned}$$

Hypotheses (i) and (ii) of Lemma 3.3 correspond to $\Delta_1 h_t(x, y, 1) \geq \Delta_1 h_t(x, y + 1, 1)$ and $\Delta_1 d_t(x, y, 1) \geq \Delta_1 d_t(x, y + 1, 1)$, respectively. The remaining hypotheses of Lemma 3.3 correspond to $\Delta_1 h_t(x, y, 1) \geq \Delta_1 d_t(x, y, 1)$ and $\Delta_2 h_t(x, y, 1) \geq \Delta_2 d_t(x, y, 1)$, which hold by (i) and (ii) of Lemma 3.4. Since all the hypotheses of Lemma 3.3 are satisfied, it follows that

$$\Delta_1 V_t(x, y, 1) \geq \Delta_1 V_t(x, y + 1, 1).$$

Because of symmetry, (iiib) holds as well.

PROOF OF (iv): Property (iv) is complementary to (iii) and the structure of its proof is the same as that of (iii). We omit the details. ■

The structural properties presented in Lemma 3.4 are sufficient to prove that there exists an optimal dispatching strategy that is of a threshold type and to prove a monotonicity property for the threshold. These statements are formally proved in Theorem 3.1. Let $\theta_t^\delta(w)$ denote the threshold function for terminal $\delta \in \{1,2\}$ at time t given that w customers are known to have waited at the other terminal at time $(t - I)^+$. Then let $\theta_t^\delta : \mathbb{Z}^+ \to \mathbb{Z}^+ \cup \{+\infty\}$ such that

$$\theta_t^1(w) \triangleq \inf\{z \in \mathbb{Z}^+ : d_t(z, w, 1) < h_t(z, w, 1)\}, \tag{3.47}$$

$$\theta_t^2(w) \triangleq \inf\{z \in \mathbb{Z}^+ : d_t(w, z, 2) < h_t(w, z, 2)\}; \tag{3.48}$$

where $\inf(\emptyset) = +\infty$.

THEOREM 3.1: *Let u be the number of customers observed at terminal δ at time t and v be the delayed observation of the queue length of the other terminal at $(t - I)^+$. For all $u, v \in \mathbb{Z}^+$, $\delta \in \{1,2\}$ and $t \in \{1,2,\ldots, T\}$, there exist threshold functions $\theta_t^\delta(\cdot) : \mathbb{Z}^+ \to \mathbb{Z}^+ \cup \{+\infty\}$ defined by Eqs.* (3.47) *and* (3.48) *such that the following shuttle control policy is optimal:*

$$\textit{dispatch the shuttle from terminal } \delta \textit{ at time } t \textit{ if and only if } u \geq \theta_t^\delta(v). \tag{3.49}$$

Moreover,

$$\theta_t^\delta(v) \geq \theta_t^\delta(v + 1). \tag{3.50}$$

Discussion. The threshold property of the optimal policy is stated in Eq. (3.49) and expresses the following relationship. If at time t it is optimal to dispatch the shuttle when in state $(x, y, 1)$ (resp. $(x, y, 2)$), then it is optimal to dispatch at t when in state $(x + 1, y, 1)$ (resp. $(x, y + 1, 2)$). The threshold function for node δ at time t depends on the probability distribution of the queue length at t in the other node, which in turn is determined by the most recent delayed observation, v.

Equation (3.50) states that the threshold function is monotone nonincreasing in the delayed observation. That is, the threshold at t cannot increase (the shuttle is dispatched more readily) if the number of customers at time $(t - I)^+$ in the terminal not occupied by the shuttle is increased.

PROOF: We begin with Eq. (3.49) and present only the case where $\delta = 1$ because of symmetry.

Sufficiency: Applying (i) of Lemma 3.4 $u - \theta_t^1(v)$ times in succession gives

$$\begin{aligned} h_t(u, v, 1) - d_t(u, v, 1) &\geq h_t(\theta_t^1(v), v, 1) - d_t(\theta_t^1(v), v, 1) \\ &> 0 \qquad \text{by Eq. (3.47)}. \end{aligned}$$

Necessity: By Eq. (3.47), $V_t(u, v, 1) = d_t(u, v, 1)$ implies $u \geq \theta_t^\delta(v)$.

Finally, we prove Eq. (3.50). Rewriting (ii) of Lemma 3.4,

$$\begin{aligned} h_t(\theta_t^1(v), v + 1, 1) - d_t(\theta_t^1(v), v + 1, 1) &\geq h_t(\theta_t^1(v), v, 1) - d_t(\theta_t^1(v), v, 1) \\ &> 0 \qquad \text{by Eq. (3.47)}. \end{aligned}$$

Using Eq. (3.47) again, we conclude that $\theta_t^1(v) \geq \theta_t^1(v + 1)$. ∎

3.2. The Infinite Horizon Problem

For the problem with the infinite horizon expected β-discounted ($\beta < 1$) cost criterion, we restrict attention to i.i.d. batch arrival processes and holding cost rate functions that are of polynomial order; i.e., there exists $c \in \mathbb{N}$ such that

$$c_\delta(z) \leq cz^c \qquad \forall z \in \mathbb{Z}^+. \tag{3.51}$$

Because times $1,2,\ldots,I$ represent a transient period in the evolution of the information state, we restrict attention to characterizing the structure of a stationary Markov policy that is optimal at time $t \geq I + 1$.

In the analysis that follows, the superscript on V, d, and h indicates the horizon of optimization. Recall that the dependence of V, d, and h on β is implicit. We begin by observing that for all $t, T \in \mathbb{N}$, $x, y \in \mathbb{Z}^+$, and $\delta \in \{1,2\}$;

$$V_t^{T+1}(x,y,\delta) \geq V_t^T(x,y,\delta) \tag{3.52}$$

and because of Eqs. (3.9)–(3.12)

$$h_t^{T+1}(x,y,\delta) \geq h_t^T(x,y,\delta), \tag{3.53}$$

$$d_t^{T+1}(x,y,\delta) \geq d_t^T(x,y,\delta). \tag{3.54}$$

Because of Eqs. (3.52) and (3.9)–(3.12), we obtain by taking the limit as $T \to \infty$ and applying the monotone convergence theorem

$$V^\infty(x,y,\delta) = \min(h^\infty(x,y,\delta), d^\infty(x,y,\delta)) \tag{3.55}$$

where we define

$$V^\infty(x,y,\delta) = \lim_{T\to\infty} V_t^T(x,y,\delta),$$

$$h^\infty(x,y,\delta) = \lim_{T\to\infty} h_t^T(x,y,\delta),$$

$$d^\infty(x,y,\delta) = \lim_{T\to\infty} d_t^T(x,y,\delta).$$

Denote by Π the space of all policies that are functions of the information state specified in Section 2. let $V^\pi(x,y,\delta)$ be the infinite horizon expected β-discounted cost for $\pi \in \Pi$ with (x,y,δ) as the initial information state at some time $t \geq I + 1$. Define

$$V(x,y,\delta) \triangleq \inf_{\pi\in\Pi} V^\pi(x,y,\delta) \tag{3.56}$$

We show that $V(x,y,\delta)$ is finite. To bound $V(x,y,\delta)$, we compute $V^g(x,y,\delta)$, where policy g never dispatches the shuttle. Using Eq. (3.51) we find

$$V^g(x,y,1) = E^g\left\{\sum_{j=0}^{\infty}\beta^j(c_1(X_{t+j}) + c_2(Y_{t+j}))\,\middle|\,X_t = x,\ Y_{t-I} = y, \delta_t = 1\right\}$$

$$< \sum_{j=0}^{\infty}\beta^j(c(x+jM)^c + c(y + M(j+I))^c)$$

$$< \infty.$$

Similarly, $V^g(x,y,2) < \infty$.

Because $V(x,y,\delta)$ is finite, Propositions 1–3 of Sennot [8] yield $V(x,y,\delta) = \lim_{T\to\infty} V_t^T(x,y,\delta) = V^\infty(x,y,\delta)$ and the stationary policy π determined by the right-hand side of Eq. (3.55) is β-discounted optimal. Therefore, to determine the properties of the infinite horizon β-discounted optimal policy, we study Eq. (3.55).

We begin by considering Lemma 3.4 as the horizon of optimization goes to infinity. This yields

$$\Delta_i d^\infty(x,y,i) \le \Delta_i h^\infty(x,y,i), \tag{3.57}$$

$$\Delta_i d^\infty(x,y,j) \le \Delta_i h^\infty(x,y,j). \tag{3.58}$$

for all $x, y \in \mathbb{Z}^+$ and $i, j \in \{1,2\}$ such that $j \neq i$. The threshold property is implied by Eq. (3.57); the monotonicity of the threshold is implied by Eq. (3.58). The following result follows from Eqs. (3.57) and (3.58) by the arguments of Theorem 3.1.

THEOREM 3.2: *Let u be the number of customers observed at terminal δ at time t, where $t \ge I + 1$, and let v be the delayed observation of the queue length of the other terminal at $(t - I)^+$. For all $u, v \in \mathbb{Z}^+$ and $\delta \in \{1,2\}$, there exist threshold functions $\theta^\delta(\cdot): \mathbb{Z}^+ \to \mathbb{Z}^+ \cup \{+\infty\}$ (defined similar to Eqs.* (3.47) *and* (3.48)) *such that the following stationary shuttle control policy is optimal:*

$$\textit{dispatch the shuttle from terminal } \delta \textit{ at time } t \textit{ if and only if } u \ge \theta^\delta(v). \tag{3.59}$$

Moreover,

$$\theta^\delta(v) \ge \theta^\delta(v+1). \tag{3.60}$$

4. LINEAR HOLDING COSTS: COMPUTATIONAL REDUCTION

In Section 3, we showed that a threshold type policy is optimal. This reduces the computational effort required to determine an optimal dispatching policy by limiting the search to the class of policies possessing the threshold property. Moreover, the determination of the threshold functions that define an optimal policy is further simplified by the monotonic dependence of the threshold on the delayed observation, as expressed by Eq. (3.60) of Theorem 3.2. However, the determination of the threshold functions defining an optimal policy still remains a difficult computational problem. Thus, further characterization of the optimal threshold functions is of great interest. In this section, we consider the problem of Section 2 with linear holding costs, that is, for $\delta \in \{1,2\}$ and $z \in \mathbb{Z}^+$,

$$c_\delta(z) = c_\delta z. \tag{4.1}$$

We present one necessary and several sufficient conditions for dispatching that reduce the computational effort required to determine optimal threshold functions. Because the ideas are similar under either the finite or infinite horizon cost criterion, we develop the results in the infinite horizon context (assuming $0 < \beta < 1$) and merely state them for the finite horizon case (assuming a horizon T and $0 < \beta \leq 1$). As before, the infinite horizon results assume i.i.d. arrival processes and focus on the characteristics of a policy that is optimal at times $I + 1$ and beyond.

LEMMA 4.1: *Let $u \geq Q$, $\delta \in \{1,2\}$, and*

$$\gamma^\delta \triangleq (1-\beta)^{-1} c_\delta Q - K_\delta - R_\delta Q.$$

Suppose $\gamma^m \geq \gamma^k$ where $m, k \in \{1,2\}$ and $m \neq k$; then the following actions are optimal:

(i) *when the shuttle is at terminal m with u customers in terminal m, dispatch if and only if $\gamma^m > 0$,*

(ii) *when the shuttle is at terminal k with u customers in terminal k, hold the shuttle if $\gamma^m \leq 0$,*

(iii) *in* (ii), *dispatch if $\gamma^k > 0$.*

PROOF OF SUFFICIENCY IN (i): For simplicity, assume $m = 1$ and $\gamma^1 > 0$. Consider at time t ($t \geq I + 1$) the information state $(u, v, 1)$, where $u \geq Q$. For a policy, π, define $J_t(\omega, \pi, (u, v, 1))$ as the cost-to-go from information state $(u, v, 1)$ at t along the sample path $\omega \in \Omega$ (Ω is the underlying sample space defined in Section 2). Suppose that the optimal policy, g, holds the shuttle at t for state $(u, v, 1)$. We construct an alternative policy, $\tilde{g}$, which dispatches the shuttle at t for state $(u, v, 1)$ and achieves for all $\omega \in \Omega$

$$J_t(\omega, \tilde{g}, (u, v, 1)) < J_t(\omega, g, (u, v, 1)). \tag{4.2}$$

Therefore, g does not achieve the minimum expected cost-to-go and is not optimal. Thus it is optimal to dispatch the shuttle at t, i.e., $d(u, v, 1) < h(u, v, 1)$.

We consider the arrival and service process realizations along $\omega \in \Omega$ and present a coupling argument to prove Eq. (4.2). We begin with the case where g never dispatches the shuttle along ω. We construct policy $\tilde{g}$ so as to dispatch at time t and never to dispatch thereafter. This construction is feasible because after dispatching at t, $\tilde{g}$ uses the delayed observations of the arrival history at node one and its knowledge of the arrival history at node two to determine that g never dispatches. Then $J_t(\omega, g, (u, v, 1)) - J_t(\omega, \tilde{g}, (u, v, 1)) = \gamma^1 > 0$.

To complete the proof, we consider the case where g dispatches the shuttle along ω at time $\sigma(\omega)$, where $t < \sigma(\omega)$. Let $\tau_1(\omega), D_1 \leq \tau_1(\omega)$, denote the trip duration along ω. Thus, the shuttle arrives at node two at $\sigma(\omega) + \tau_1(\omega)$. Along ω, the duration of the trip begun at t under $\tilde{g}$ is equal to $\tau_1(\omega)$ (because the durations of trips made from node one are i.i.d.). Upon arriving to node two, $\tilde{g}$ holds the shuttle there until $\sigma(\omega) + \tau_1(\omega)$. Because $I \leq D_1$, along ω the same information state exists at $\sigma(\omega) + \tau_1(\omega)$ under policies g and $\tilde{g}$ (see Fig. 1). From time $\sigma(\omega) + \tau_1(\omega)$ on, we make $\tilde{g}$ identical to g. We pause to justify the construction of $\tilde{g}$ by examining two distinct scenarios. Suppose $\sigma(\omega) \leq t + \tau_1(\omega) - I$. Then upon arrival to node two at $t + \tau_1(\omega)$, $\tilde{g}$ knows the history of the arrival process of node one through time $\sigma(\omega)$ and thereby determines that policy g dispatched the shuttle at $\sigma(\omega)$. Having determined $\sigma(\omega)$, $\tilde{g}$ uses its own trip length realization in determining to follow g from time $\sigma(\omega) + \tau_1(\omega)$ on. On the other hand, if $\sigma(\omega) > t + \tau_1(\omega) - I$, then $\tilde{g}$ cannot determine $\sigma(\omega)$ upon arrival to node two. Instead, $\tilde{g}$ waits at node two until $\sigma(\omega) + I$, at which time $\tilde{g}$ determines $\sigma(\omega)$. Because $I \leq D_1$, $\tilde{g}$ always determines $\sigma(\omega)$ at or before time $\sigma(\omega) + \tau_1(\omega)$.

Comparing policy g with $\tilde{g}$, we find that along ω, g holds Q additional customers in node one at times t through $\sigma(\omega) - 1$ and matches $\tilde{g}$ thereafter. No difference exists at node two. The policies differ with respect to dispatching only in that g dispatches at $\sigma(\omega)$, whereas $\tilde{g}$ dispatches at t. Thus,

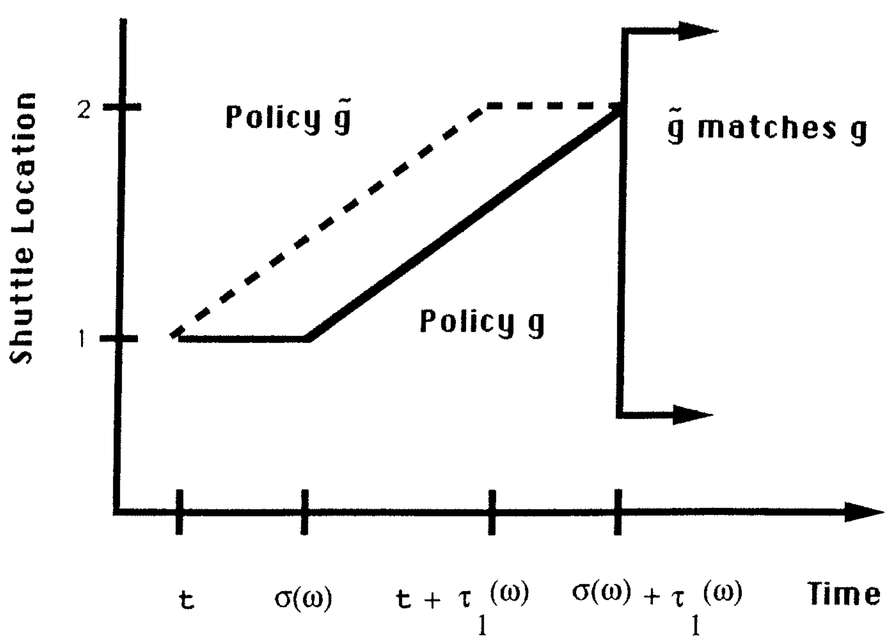

FIGURE 1. Illustration of policies $\tilde{g}$ and g.

$$
\begin{aligned}
J_t(\omega, g, (u, v, 1)) - J_t(\omega, \tilde{g}, (u, v, 1)) &= \sum_{j=t}^{\sigma(\omega)-1} \beta^{j-t} c_1 Q + (\beta^{\sigma(\omega)-t} - 1)(K_1 + R_1 Q) \\
&= (1 - \beta^{\sigma(\omega)-t})(c_1 Q(1-\beta)^{-1} \\
&\quad - (K_1 + R_1 Q)) \\
&= (1 - \beta^{\sigma(\omega)-t})\gamma^1 \\
&> 0.
\end{aligned}
$$

Proof of (ii) and Necessity in (i): Because the argument is the same for both statements regardless of the value of m, simply assume $\gamma^1 \le 0, \gamma^2 \le 0$ and consider the state $(u, v, 1)$ at any time t, $t \ge I + 1$. Suppose policy g is optimal and dispatches at time t. We present a coupling argument to show that there is a policy, $\tilde{g}$, that never dispatches the shuttle and performs at least as well as g. Thus, $h_t(u, v, 1) \le d_t(u, v, 1)$.

Consider the arrival and service process realizations along $\omega \in \Omega$. Let $\sigma_1(\omega)$, $\sigma_2(\omega), \ldots$ denote the sequence of dispatching epochs at times t ($\sigma_1(\omega) = t$) and beyond made under g along ω. Let $\tilde{g}$ be a policy that never dispatches the shuttle. Denote by $\delta_{\sigma_j(\omega)}$ the node from which the shuttle is dispatched under g at time $\sigma_j(\omega)$. Then, because $c_1 \ge R_1$, $c_2 \ge R_2$, $\gamma^1 \le 0$, $\gamma^2 \le 0$, and each dispatch clears at most Q customers from the system, the cost advantage of policy $\tilde{g}$ vs. g for any dispatch, say the one at time $\sigma_k(\omega)$, is bounded below by $\beta^{\sigma_k(\omega)}(-\gamma^{\delta_{\sigma_k(\omega)}})$, which is positive. Hence, along ω, the cost advantage of policy $\tilde{g}$ vs. g is

$$
\begin{aligned}
J_t(\omega, g, (u, v, 1)) - J_t(\omega, \tilde{g}, (u, v, 1)) &\ge \sum_j \beta^{\sigma_j(\omega)-t}(-\gamma^{\delta_{\sigma_j(\omega)}}) \\
&\ge 0.
\end{aligned}
$$

Consequently, it is optimal to hold at t.

Proof of (iii): The result follows by the argument made for the sufficiency statement of (i). ■

For $\gamma^m \ge \gamma^k$, Theorem 3.2 and Lemma 4.1 yield $\theta^1 = \theta^2 = +\infty$ provided $\gamma^m \le 0$; otherwise, they reduce the range of θ^m to a finite set, thus yielding $\theta^m : \mathbb{Z}^+ \to \{0, 1, 2, \ldots, Q\}$. Moreover, a sufficient condition is given for this reduction to apply to θ^k. This is the essence of the following theorem.

Theorem 4.1: *Suppose* $\gamma^m \ge \gamma^k$ *where* $m, k \in \{1,2\}$, $m \ne k$, $v \in \mathbb{Z}^+$;

(i) *if* $\gamma^m > 0$, *then* $\theta^m(v) \le Q$;

(ii) *if* $\gamma^k > 0$, *then* $\theta^1(v) \le Q$ *and* $\theta^2(v) \le Q$;

(iii) *if* $\gamma^m \le 0$, *then* $\theta^1(v) = \theta^2(v) = +\infty$.

Proof of (i): Without loss of generality, suppose $m = 1$. If $\gamma^1 > 0$, Lemma 4.1 yields $h(Q, v, 1) > d(Q, v, 1)$. In light of Eq. (3.57), the result holds by the definition of γ^1.

PROOF OF (ii): This case follows by the argument of (i) and the fact that $\gamma^m \geq \gamma^k$.

PROOF OF (iii): By Lemma 4.1, $\gamma^m \leq 0$ implies $h(u, v, 1) \leq d(u, v, 1)$ and $h(v, u, 2) \leq d(v, u, 2)$ for all $u \geq Q$. This in turn implies by Eq. (3.57) that it is optimal to hold the shuttle for $u < Q$. To be consistent with Theorem 3.2, we set $\theta^1(v) = \theta^2(v) = +\infty$. ■

The following corollary is an immediate consequence of Theorem 4.1 for the case of cost symmetry in the network.

COROLLARY 4.1: *Let $c_1 = c_2$, $K_1 = K_2$, and $R_1 = R_2$. Then for all $v \in \mathbb{Z}^+$; $\gamma^1 = \gamma^2$ and*

(i) *if $\gamma^1 > 0$, then $\theta^1(v) \leq Q$ and $\theta^2(v) \leq Q$,*
(ii) *if $\gamma^1 \leq 0$, then $\theta^1(v) = \theta^2(v) = +\infty$.*

It is possible to derive a finite horizon result analogous to that of Theorem 4.1. This is the following:

THEOREM 4.2: *For $t \in \{1,2, \ldots, T\}$, let*

$$\gamma_t^\delta \triangleq c_\delta Q \sum_{j=0}^{T-t} \beta^j - K_\delta - R_\delta Q$$

and suppose $\gamma_t^m \geq \gamma_t^k$ where $m, k \in \{1,2\}$ and $m \neq k$. Then

(i) *if $\gamma_t^m > 0$, then $\theta_t^m(v) \leq Q$;*
(ii) *if $\gamma_t^k > 0$, then $\theta_t^1(v) \leq Q$ and $\theta_t^2(v) \leq Q$;*
(iii) *if $\gamma_t^m \leq 0$, then $\theta_t^1(v) = \theta_t^2(v) = +\infty$.*

In some cases, it may be possible to determine that the maximum value of an optimal threshold lies below Q. Lemma 4.2 and Theorem 4.3 present a sufficient condition for this to be the case.

LEMMA 4.2: *Suppose that the shuttle is at terminal δ with u, $u < Q$, passengers waiting in that terminal. It is optimal to dispatch the shuttle if*

$$c_\delta u - \beta(1-\beta)^{-1} c_\delta (Q - u) - (1 - \beta)(K_\delta + R_\delta u) > 0. \quad \textbf{(4.3)}$$

PROOF: We begin by noting that Eq. (4.3) implies $\gamma^\delta > 0$, because $c_1 \geq R_1$, $c_2 \geq R_2$, and $u < Q$. Let g denote an optimal policy. Suppose g holds the shuttle at node δ at time t when u passengers wait there. We construct a modified policy, $\tilde{g}$, that dispatches the shuttle at t and achieves a smaller cost-to-go from time t on for every $\omega \in \Omega$.

Fix $\omega \in \Omega$. We begin with the case where g never dispatches the shuttle along ω. We construct policy $\tilde{g}$ so as to dispatch the shuttle at time t and never dispatch thereafter. Along ω, the cost advantage of $\tilde{g}$ over g is $\gamma^\delta > 0$. On the other hand, g may dispatch the shuttle at time $\sigma(\omega)$, where $\sigma(\omega) \geq t + 1$. As in the proof of the sufficiency statement of (i) in Lemma 4.1, we construct a policy unloading times for larger loads. Thus, the idea is to model the shuttle trip length as a random variable $\tau(u)$, which increases stochastically in u, the number of passengers carried.

Under this new assumption, the issue we investigate is whether or not the threshold property continues to characterize an optimal policy. We provide a counterexample, which demonstrates that for a finite horizon problem, even with perfect information, threshold policies are not always optimal. The infinite horizon problem remains open.

Counterexample: Using the same notation as before, we consider a deterministic two-stage optimization problem ($T = 2$) with complete observations ($I = 0$). Assume the shuttle's capacity is at least two ($Q \geq 2$). If u is the number of customers carried, the length of a trip made from either node depends on u according to $\tau(u) = u \vee 1$. We study three initial conditions for (X_1, Y_1, δ) at time $t = 1$: (0,3,1); (1,3,1); and (2,3,1). Because any arrivals at $t = 1$ are included in the initial condition, we need only specify the arrival process at $t = 2$. We consider the case of deterministic arrivals at $t = 2$: 3 at node one and 1 at node two. Let $\beta = 1$, $c_1 = 1$, $c_2 = \frac{3}{4}$, $K_1 = 1$, $K_2 = \frac{1}{2}$, and $R_1 = R_2 = 0$. We derive the optimal control action at $t = 1$ for each of the initial conditions.

State (0,3,1): If the shuttle is held at $t = 1$, the state at $t = 2$ is (3,4,1), and it is optimal to dispatch at that time. Thus, $h_1(0,3,1) = 6.25$. Dispatching at $t = 1$ yields state (3,4,2) at $t = 2$, and thus it is optimal to dispatch a second time. We find $d_1(0,3,1) = 6.75$.

State (1,3,1): If held at $t = 1$, the shuttle must be dispatched from state (4,4,1) at $t = 2$; thus, $h_1(1,3,1) = 7.25$. On the other hand, $d_1(1,3,1) = d_1(0,3,1) = 6.75$.

State (2,3,1): We find $h_1(2,3,1) = 8.25$. If dispatched at $t = 1$, the trip length of two units precludes a second dispatch: $d_1(2,3,1) = 9.25$.

To conclude, the threshold property is violated at $t = 1$ because the shuttle must be held for states (0,3,1) and (2,3,1) but must be dispatched for state (1,3,1).

6. CONCLUSIONS

We have analyzed a simple two-node shuttle system under imperfect state observations. We have shown that under the conditions specified in Section 2 an optimal dispatching policy is of the threshold type; furthermore, we proved that the optimal thresholds are monotone functions of the most recent delayed observation. Knowledge of these properties of an optimal policy can guide its computation by reducing the search to functions described by a threshold. The results of Section 4, which present necessary and sufficient conditions for optimally dispatching the shuttle from a given terminal, further reduce the computational effort required to determine the optimal threshold functions, because, in certain cases, they limit the range of the thresholds to a finite set. The counterexample of Section 5 demonstrates that the threshold property does not hold in general for finite horizon problems in which the trip length increases stochastically with increasing shuttle load.

With an analysis in hand for the problem with trip durations longer than the information delay, we are interested in the cases where (with positive probability) the information delay exceeds the trip duration. An important example is the case where the shuttle controller observes only those customers waiting at the terminal at which the shuttle currently waits. Thus, upon leaving a terminal, no new information about that terminal is gained until returning there (this corresponds to an infinite delay in our formulation). Such problems remain open. Although we still believe that a threshold-type scheduling rule is optimal, we have not been able to extend either the dynamic programming approach of Section 3 or the coupling arguments of Section 4 to more general information patterns.

The insight gained from this work can guide the design of dispatching policies in more realistic transportation networks and can support future analyses of more than two nodes, multiple vehicles, and decentralized information.

References

1. Barnett, A. (1973). On operating a shuttle service. *Networks* 3: 305–313.
2. Deb, R.K. (1978). Optimal dispatching of a finite capacity shuttle. *Management Science* 24: 1362–1372.
3. Dror, I.M. (1978). Shuttle systems with one carrier and two passenger queues. Ph.D. thesis, Columbia University, New York.
4. Ignall, E. & Kolesar, P. (1972). Operating characteristics of a simple shuttle under local dispatching rules. *Operations Research* 20: 1077–1088.
5. Ignall, E. & Kolesar, P. (1974). Optimal dispatching of an infinite-capacity shuttle: Control at a single terminal. *Operations Research* 22: 1008–1024.
6. Kumar, P.R. & Varaiya, P. (1986). *Stochastic systems: Estimation, identification, and adaptive control.* Englewood Cliffs, NJ: Prentice-Hall.
7. Ross, S. (1983). *Introduction to stochastic dynamic programming.* New York: Academic Press.
8. Sennot, L.I. (1989). Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Operations Research* 37: 626–633.