# State-dependent Control of a Single Stage Hybrid System with Poisson Arrivals

**Kagan Gokbayrak**

**Abstract** We consider a single-stage hybrid manufacturing system where jobs arrive according to a Poisson process. These jobs undergo a deterministic process which is controllable. We define a stochastic hybrid optimal control problem and decompose it hierarchically to a lower-level and a higher-level problem. The lower-level problem is a deterministic optimal control problem solved by means of calculus of variations. We concentrate on the stochastic discrete-event control problem at the higher level, where the objective is to determine the service times of jobs. Employing a cost structure composed of process costs that are decreasing and strictly convex in service times, and system-time costs that are linear in system times, we show that receding horizon controllers are state-dependent controllers, where state is defined as the system size. In order to improve upon receding horizon controllers, we search for better state-dependent control policies and present two methods to obtain them. These stochastic-approximation-type methods utilize gradient estimators based on Infinitesimal Perturbation Analysis or Imbedded Markov Chain techniques. A numerical example demonstrates the performance improvements due to the proposed methods.

## 1 Introduction

Systems that are described both with time-driven and event-driven dynamics are referred to as hybrid systems. The former are represented by differential (or difference) equations, while the latter may be described through various frameworks

K. Gokbayrak (✉)
Department of Industrial Engineering, Bilkent University, Ankara, 06800, Turkey
e-mail: kgokbayr@bilkent.edu.tr

used for discrete event systems, such as timed automata, max-plus equations, or Petri nets (see in Cassandras and Lafortune 2007).

A hybrid system has different modes of operation at which its physical state evolves according to mode-specific time-driven dynamics. Discrete events (either controlled or uncontrolled) cause transitions from one mode to another not only changing the time-driven dynamics but also the physical state of the system. The timings and the types of events are determined by the underlying discrete event dynamics and the applied control.

We consider a hybrid manufacturing system where the time-driven dynamics for each mode is deterministic while one of the events causing mode transitions, the arrival event, occurs according to a Poisson process. Therefore, our system can be viewed as a stochastic hybrid system (SHS), similar to the ones in Cassandras and Lygeros (2006). In order to tackle an optimal control problem for our SHS, we first decompose it hierarchically to a deterministic optimal control problem for the time-driven component and a stochastic discrete-event control problem for the event-driven component, which we call lower-level and higher-level problems, respectively (see in Gokbayrak and Cassandras 2000a and b). The lower-level problem is a classical optimal control problem that can be solved via calculus of variations (see in Bryson and Ho 1975). In order to tackle the higher-level problem of determining the optimal service times of jobs, we start by obtaining receding horizon (RH) controller (Cassandras and Mookherjee 2003a, b, c) results. For the system we consider in this paper, the RH controller selects service times based only on the system size at the moment of service initiation. Therefore, defining the system size as the state of the discrete event system, we can view the RH controller result as a state-dependent control policy. Hence, our discrete event system is modeled as an M/D(n)/1 system, where D(n) denotes a deterministic service-time selection based on the state.

In this paper, motivated by the RH controller results, we search for the optimal state-dependent control policy for the service time selection. For this purpose, starting at RH controller results, we apply Stochastic Approximation (SA) algorithms (e.g., see in Kushner and Yin 2003) that are driven by gradient estimates obtained from two alternative techniques: Infinitesimal Perturbation Analysis (IPA) and Imbedded Markov Chains (IMC).

Perturbation Analysis (PA) of discrete event dynamic systems allows gradient estimation from a single sample path with minimal effort (see in Glasserman 1991 and Ho and Cao 1991). IPA, a flavor of PA, yields unbiased and strongly consistent gradient estimates (e.g. see in Suri and Zazanis 1988 and L'Ecuyer and Glynn 1994) and has been used along with SA for optimization of various single-stage queueing systems (e.g., in Chong and Ramadge 1992 and Fu 1990). The IPA gradient estimator in this paper is a modified version of the one in Suri and Zazanis (1988) (and in Zazanis and Suri 1994), which is for the system time with respect to service time parameters of GI/GI/1 queues. This estimator has been shown to be strongly consistent for M/G/1 systems in Suri and Zazanis (1988).

The Imbedded Markov Chain (IMC) is a method first used in Kendall (1951) to determine the system-size distribution of an M/G/1 system at the steady state. Since the service times are not necessarily exponentially distributed in M/G/1 systems, a discrete-time Markov chain, which is called as the IMC, is formed at the departure instants. Utilizing the PASTA (Poisson arrivals see time averages) property (see in Wolff 1982), the steady-state distribution at the departure instants can be shown to

be equal to the steady-state distribution at all points in time. Foster presented the ergodicity condition for the IMC of M/G/1 in Foster (1953), and Harris, in Harris (1967), applied this condition to prove the ergodicity of the IMC of M/M(n)/1. A similar analysis is performed in this paper for the ergodicity of the IMC of M/D(n)/1 systems. Once the existence of the steady-state distribution is established, it is employed in the gradient estimation via matrix truncation and finite difference method.

The rest of the paper is organized as follows: In Section 2, we formulate the optimal control problem and decompose it into a classical optimal control problem and a stochastic discrete-event control problem. The RH controller's state-dependent policy, the motivation for this work, is also presented in this section. Section 3 presents the stochastic approximation algorithm with projections along with two gradient estimators. A numerical example in Section 4 demonstrates the cost performance improvements due to the proposed methods. Finally, in Section 5, we present our conclusions and possible directions for future work.

## 2 Problem formulation

We consider a single stage make-to-order manufacturing system. The Palm-Khintchine Theorem (see in Heyman and Sobel 2003) states that the superimposition of a large number of properly normalized independent renewal processes approaches to a Poisson process. Therefore, modeling the order arrival as a Poisson process is a widely accepted assumption. Each order is associated with a job. These jobs undergo a deterministic process to change their physical characteristics, such as temperature, chemical composition, etc., according to the following time-driven dynamics described by the first-order differential equation

$$\dot{z}_i(t) = f(z_i(t), u_i(t)), \ \tau_i \leq t \leq \tau_i + s_i \tag{1}$$

with the boundary condition

$$z_i(\tau_i) = \zeta^0 \tag{2}$$

for the $i$th job $C_i$. In Eq. 1, $z_i(t)$ denotes the physical state of job $C_i$, $u_i(t)$ is the admissible control input applied, the temporal variable $\tau_i$ is the time job $C_i$ enters process, and the temporal variable $s_i$ is the duration of service, which is a decision variable in our analysis. All jobs start out at the same initial state $\zeta^0$, and are processed to reach a desired state $\zeta^d$. The physical state of the system assumes the physical state of the job in process.

Assuming that jobs are processed one-by-one in the order they arrived, in order to capture the dynamics on the temporal variables, we employ an M/G/1 queueing model operating under first-in-first-out (FIFO) policy. Let $\{a_i\}_{i \in \mathbb{N}}$ be the arrival times of the jobs, where $a_i \leq a_{i+1}$ for $i \in \mathbb{N}$. According to FIFO policy, the departure time of job $C_i$ from the system, denoted by $d_i$, has the event-driven dynamics given by

$$d_i = \max(a_i, d_{i-1}) + s_i \tag{3}$$

for $i \in \mathbb{N}$, where $d_0 = -\infty$. Note that the service starting time for job $C_i$ is given by $\tau_i = \max(a_i, d_{i-1})$.

The stochastic hybrid control problem that we are interested in is the infinite horizon average cost minimization

$$
\min_{s_i \geq 0, u_i(t), i \in \mathbb{N}} \lim_{N \to \infty} \frac{1}{N} E \left\{ \sum_{i=1}^{N} \left[ \phi_i(d_i) + \varphi(z_i(d_i)) + \int_{\tau_i}^{d_i} L(u_i(t)) dt \right] \right\}
$$

subject to Eqs. 1–3, where $\phi_i$ is the linear system time cost of departing job $C_i$ at $d_i$ given by

$$
\phi_i(d_i) = \alpha(d_i - a_i) \tag{4}
$$

$\varphi$ is the cost penalizing the deviation of $z_i(d_i)$ from the desired state $\zeta^d$, and $L$ is the instantaneous cost of control. The expectation is taken over a probability space $(\Omega, \mathcal{F}, P)$ defined for the Poisson arrival process, where $\Omega$ is the sample space, $\mathcal{F}$ is the set of events, and $P$ is the probability measure. Note that each outcome $\omega$ in $\Omega$ corresponds to a sequence of arrival times $\{a_i\}_{i \in \mathbb{N}}$.

Following the arguments in Gokbayrak and Cassandras (2000a and b), we can rewrite the optimal control problem as

$$
\min_{s_i \geq 0, i \in \mathbb{N}} \lim_{N \to \infty} \frac{1}{N} E \left\{ \sum_{i=1}^{N} \left[ \phi_i(d_i) + \min_{\bar{u}_i(t, s_i)} \left\{ \varphi(z_i(d_i)) + \int_{\tau_i}^{d_i} L(\bar{u}_i(t, s_i)) dt \right\} \right] \right\} \tag{5}
$$

subject to Eqs. 1–3. Note that the inner minimization is performed for given $s_i$ and subject to only Eqs. 1 and 2. Hence, we can hierarchically decompose our problem to a lower-level deterministic optimal control problem of obtaining the optimal process cost function $\theta(s_i)$ and the corresponding optimal control function $\bar{u}_i^*(t, s_i)$ where

$$
\theta(s_i) = \min_{\bar{u}_i(t, s_i)} \left\{ \varphi(z_i(d_i)) + \int_{\tau_i}^{d_i} L(\bar{u}_i(t, s_i)) dt \right\} \tag{6}
$$

subject to Eqs. 1–2, and a higher-level stochastic problem of determining the optimal service durations $s_i^*$ for each job

$$
\min_{s_i \geq 0, i \in \mathbb{N}} J = \lim_{N \to \infty} \frac{1}{N} E \left\{ \sum_{i=1}^{N} \left[ \phi_i(d_i) + \theta(s_i) \right] \right\} \tag{7}
$$

subject to Eq. 3. Note that once the optimal service durations are determined, the control input $\bar{u}_i^*(t, s_i^*)$ is applied on job $C_i$.

The deterministic lower level problem is a classical optimal control problem whose solution methodologies can be found in e.g. Bryson and Ho (1975). Instead, we concentrate on solving the higher-level stochastic discrete-event control problem.

In general, the process cost $\theta(.)$ can take any form. The analysis in this paper is applicable, however, only if the following assumption on the process cost function is satisfied:

**Assumption 1** Process cost $\theta(.)$ is a decreasing, differentiable, and strictly convex function.

This is a realistic assumption which can be observed in e.g. some linear time-invariant systems with quadratic cost functions $\varphi$ and $L$. (Section 4 presents such an example) A trade-off is observed in this assumption: Longer service times will decrease the process costs, while increasing the departure times and hence the system time costs. Our aim in this paper is to determine the optimal service times.

A deterministic version of this problem, where the arrival times are given, was first considered by Pepyne and Cassandras (1998) in which the objective function was designed to complete jobs as fast as possible with the least amount of control effort. In Pepyne and Cassandras (2000), they extended this line of work to a system with completion deadlines penalizing both earliness and tardiness. The uniqueness of the optimal solution was shown in Cassandras et al. (2001), and the algorithms to determine this solution were given in Wardi et al. (2001) and Cho et al. (2001), which was later improved in Zhang and Cassandras (2002). A parallel line of work in Gokbayrak and Cassandras (2000a and b) showed that, under some conditions, it is possible to decompose the hybrid optimal control problem in Eq. 5 to classical optimal control and discrete-event optimal control problems, as we have performed above.

In this paper, we assume that the times of future arrivals are unavailable. According to the RH control method presented in Cassandras and Mokherjee (2003a, b and c), this setting corresponds to a zero-length information window: At each service time decision instant $t$, we have the arrival times of jobs that are already in the system and the next arrival is assumed to occur after an idle period, hence decoupling the optimization problem (see in Cho et al. 2001). If jobs $\{C_k, ..., C_n\}$ are currently in the system at the service initiation time $\tau_k$, in order to determine the service time $s_k$ for job $C_k$, the RH controller solves the following deterministic optimization problem

$$R(k, n): \min_{\substack{s_i \geq 0 \\ i=k,...,n}} \frac{1}{n-k+1} \sum_{i=k}^{n} \left( \theta(s_i) + \phi_i \left( \tau_k + \sum_{j=k}^{i} s_j \right) \right) \tag{8}$$

The optimal service times obtained from this problem are denoted by $\{s_i^*(k, n)\}_{i=k}^{n}$. For the linear system time cost in Eq. 4, the RH controller determines a service time $s_k^*(k, n) \geq 0$ for job $C_k$ that satisfies

$$\frac{d\theta(s_k^*(k, n))}{ds} + (n-k+1)\alpha = 0 \tag{9}$$

If $\frac{d\theta(0)}{ds} > -(n-k+1)\alpha$, then, by Assumption 1, the equality in Eq. 9 can not be satisfied. In that case $s_k^*(k, n)$ is selected to be zero.

Note that the service time decision for job $C_k$ is based on the number of jobs in system, which is equal to $(n-k+1)$, but not on their arrival times. Therefore, defining the state as the system size, the RH controller implements a state-dependent control policy. Note also that this solution is optimal only if an idle period is observed after job $C_n$. In other words, it is possible to improve the policy to account for additional jobs that may be served in the same busy period.

Motivated by the RH controller, we focus on the state-dependent control policies. We propose two alternative methods to improve the RH controller solution employing the given interarrival time distribution information.

## 3 Policy improvements

We assume that the arrival rate $\lambda$ is available and propose a stochastic approximation algorithm to improve the RH controller solution. The control policy we propose is to select a service duration of $S_i$ if there are $i$ jobs in the system at the service initiation of a job. We assume that once the service is started, the service time decision is not altered due to new arrivals.

According to Eq. 9, the RH controller selects $S_i$ values that satisfy

$$\frac{d\theta(S_i)}{ds} + i\alpha = 0$$

for all $i \in \mathbb{N}^+$ such that $i \leq -\frac{1}{\alpha}\frac{d\theta(0)}{ds}$. For system sizes larger then $-\frac{1}{\alpha}\frac{d\theta(0)}{ds}$, zero service time is selected.

Starting at the RH controller policy, we utilize the iterative stochastic approximation algorithm to determine the optimal service times for each state. Let us represent the policy at iteration $n$ by $S^{(n)} = [S_1^{(n)}, S_2^{(n)}, ...]$. Initializing $S^{(0)}$ with the RH controller result, at each iteration $n = 1, 2, ...$, the policy is updated via

$$S^{(n)} = \Pi_H \left( S^{(n-1)} - \epsilon_n g \left( S^{(n-1)} \right) \right) \tag{10}$$

where $g$ is an estimate of $\nabla J$ evaluated at the last policy $S^{(n-1)}$, $\Pi_H$ is the projection onto the constraint set $H = [0, 1/\lambda)^\infty$, and the step-size sequence $\{\epsilon_n\}$ is properly selected to satisfy the standard conditions for convergence

$$\epsilon_n > 0, \epsilon_n \longrightarrow 0, \sum_{n=0}^{\infty} \epsilon_n = \infty, \sum_{n=0}^{\infty} \epsilon_n^2 < \infty,$$

as in Robbins and Monro (1951).

The necessary gradient information $g$ is supplied by either an IPA estimator or a finite forward difference method applied on the costs obtained from the IMC. The constraint set $H$ guarantees that the busy periods have finite mean length as required by the IPA estimator and that long-run distributions exist for the IMC from which the costs are calculated.

### 3.1 IPA gradient estimation

In this method, we observe a sample path of the M/D(n)/1 queueing system over a long time interval, and seek to construct an estimate of $\nabla J$.

Let us define the sample path cost for a given policy $S$ as $L(S, \omega)$ so that

$$J(S) = E_\omega[L(S, \omega)]$$

where $\omega$ is an element of the sample space $\Omega$ corresponding to the arrival times of this particular sample path, and the expectation is taken over the probability space $(\Omega, \mathcal{F}, P)$ for the Poisson arrival process. We will determine the sample path gradient $\nabla L$ by applying IPA techniques and employ it as the gradient estimate $g$ in Eq. 10.

A part of the IPA gradient estimator below is a modified version of the one in Suri and Zazanis (1988) and Zazanis and Suri (1994), which is for the mean system time of a GI/GI/1 queue with respect to a service time parameter. This estimator has been shown to be strongly consistent for M/G/1 systems in Suri and Zazanis (1988).

Employing the interarrival time distribution information, let us generate offline the arrival times of $N$ jobs. Due to the projection mapping $\Pi_H$ in Eq. 10, the maximum service time of the current state-dependent control policy, $S_{\max} = \max_i S_i$ is less than the mean interarrival time $1/\lambda$, so that the busy periods have finite mean length. Let us separate the sample path cost into process and system-time costs

$$L = L_p + L_s$$

where

$$L_p = \frac{1}{N} \sum_{i=1}^{N} \theta(s_i)$$

and

$$L_s = \frac{1}{N} \sum_{i=1}^{N} \alpha(d_i - a_i)$$

Defining the indicator function as

$$I\{s_n = S_i\} = \begin{cases} 1 & s_n = S_i \\ 0 & s_n \neq S_i \end{cases}$$

the gradient for the process cost $L_p$ can simply be given as

$$\frac{\partial L_p}{\partial S_i} = \frac{1}{N} \frac{d\theta(S_i)}{ds} \sum_{n=1}^{N} I\{s_n = S_i\} \tag{11}$$

In order to estimate the gradient for the system time cost $L_s$, we modify the unbiased IPA estimator for system times developed in Zazanis and Suri (1994) to our setting: Let us decompose the sample path into busy periods, and analyze the effects of perturbations in the service times. We denote the sample path with perturbations as the perturbed sample path, while the nominal sample path is the one without perturbations. Take, for example, the busy period in Fig. 1 and analyze the effect of an infinitesimal increment $\Delta_2 > 0$ in $S_2$ on the job departure times. This increment is assumed to be small enough so that the system sizes at the service initializations of all jobs in both the nominal and the perturbed sample paths are the same, i.e., the busy period structures stay the same.
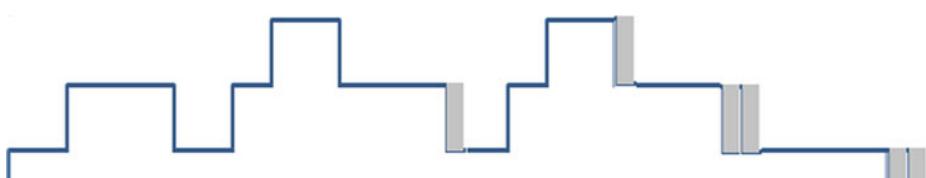


**Fig. 1** Perturbation propagation and accumulation

Let us denote the nominal and the perturbed departure times of job $C_i$ by $d_i$ and $\bar{d}_i$, respectively. For the sample path in Fig. 1, since the first two jobs are applied $S_1$, their departure times are not altered in the perturbed sample path.

$$d_1 = \bar{d}_1, d_2 = \bar{d}_2$$

The third job, however, is applied a service duration of $S_2$, therefore, its perturbed departure time is

$$\bar{d}_3 = d_3 + \Delta_2$$

Since the service starting time of the fourth job is delayed by $\Delta_2$, the perturbation is propagated so that

$$\bar{d}_4 = d_4 + \Delta_2$$

The fifth job is also applied a service duration of $S_2$, and there is a propagated delay of $\Delta_2$ from the previous job. Therefore, we have

$$\bar{d}_5 = d_5 + 2\Delta_2$$

i.e., the perturbation is accumulated. The departure of the sixth job is also delayed with $2\Delta_2$ due to the perturbation propagation.

$$\bar{d}_6 = d_6 + 2\Delta_2$$

Note that the departure perturbation of each job is equal to the number of $S_2$ services observed within the same busy period prior to (and including) the job. Since the arrival times of these jobs are the same in the perturbed sample path, the change in $L_s$ from this busy period only is given by

$$\frac{\alpha}{N} \sum_{i=1}^{6} \left( \bar{d}_i - d_i \right) = \frac{6\Delta_2}{N}$$

for a sample path formed of $N$ jobs.

In order to generalize, let us denote the index of the last job of the $b$th busy period by $n_b$, where $n_0 = 0$. Then, the amount of delay due to an infinitesimal increment $\Delta_i$ in $S_i$ on the departure time of job $C_k$ residing in the $b$th busy period of the perturbed sample path can be given as

$$\bar{\delta}_i^k(\Delta_i) = \sum_{j=n_{b-1}+1}^{k} I\{s_j = S_i\}\Delta_i$$

where $k \leq n_b$, which accounts for both the propagation and the accumulation of perturbation. If a total of $B$ busy periods are observed over the sample path, then the change in $L_s$ due to the infinitesimal increment in $S_i$ is

$$L_s^p - L_s = \frac{\alpha}{N} \sum_{b=1}^{B} \sum_{k=n_{b-1}+1}^{n_b} \bar{\delta}_i^k(\Delta_i) = \frac{\alpha}{N} \sum_{k=1}^{N} \bar{\delta}_i^k(\Delta_i) \qquad (12)$$

where $n_B = N$, and $L_s$ and $L_s^p$ are the nominal and the perturbed sample path system-time costs, respectively. Consequently, we get

$$\frac{\partial L_s}{\partial S_i} = \frac{\alpha}{N} \sum_{k=1}^{N} \delta_i^k \tag{13}$$

where $\delta_i^k$ is the number of prior jobs within the same busy period as job $C_k$ (and including job $C_k$) that the service duration $S_i$ is applied

$$\delta_i^k = \sum_{j=n_{b-1}+1}^{k} I\{s_j = S_i\} \tag{14}$$

for $k \leq n_b$.

From Eqs. 11 and 13, we obtain the gradient estimates $g = [\frac{\partial L}{\partial S_1}, \frac{\partial L}{\partial S_2}, ...]$ where

$$\frac{\partial L}{\partial S_i} = \frac{1}{N} \sum_{n=1}^{N} \left[ I\{s_n = S_i\} \frac{d\theta(S_i)}{ds} + \alpha \delta_i^n \right] \tag{15}$$

Note that the components in this estimate due to any job $C_i$ depends only on already observed service times; therefore, analyzing the sample path in a forward fashion, we can obtain the gradient estimate. An algorithm for this purpose can be given as:

Step 1   Initialize the gradients $\left(\frac{\partial L}{\partial S_i}\right)^{(0)} = 0$ for all $i = 1, 2, ...$
         For each job $C_n, n = 1, ...N$

Step 2   Check the system size $X_n'$ at the service starting time of job $C_n$

   a.   For all $i$, assign $\delta_i^n = \begin{cases} \delta_i^{n-1} + 1 & \text{if } X_n' = i \text{ and } a_n < d_{n-1} \\ \delta_i^{n-1} & \text{if } X_n' \neq i \text{ and } a_n < d_{n-1} \\ 1 & \text{if } X_n' = i \text{ and } a_n \geq d_{n-1} \\ 0 & \text{if } X_n' \neq i \text{ and } a_n \geq d_{n-1} \end{cases}$

   b.   For all $i$, assign $\left(\frac{\partial L}{\partial S_i}\right)^{(n)} = \begin{cases} \left(\frac{\partial L}{\partial S_i}\right)^{(n-1)} + \frac{1}{N}\left(\alpha\delta_i^n + \frac{d\theta(S_i)}{ds}\right) & \text{if } X_n' = i \\ \left(\frac{\partial L}{\partial S_i}\right)^{(n-1)} + \frac{\alpha}{N}\delta_i^n & \text{otherwise} \end{cases}$

Step 3   Set $\frac{\partial L}{\partial S_i} = \left(\frac{\partial L}{\partial S_i}\right)^{(N)}$ for all $i = 1, 2, ...$

Note that this IPA gradient estimation method does not require the arrival process to be Poisson. The next method, however, requires a Poisson arrival process.

### 3.2 Imbedded Markov chain based gradient estimation

Following the development in Gross et al. (2008) for M/G/1 systems, let us form an imbedded discrete time Markov chain $\{X_i : i \in \mathbb{N}\}$ by defining the state $X_i$ as the

number of jobs left in the system immediately after the departure of job $C_i$. If $A_i$ is the number of jobs arrived during the service of job $C_i$, then the state of the chain evolves according to

$$X_i = \begin{cases} X_{i-1} - 1 + A_i & X_{i-1} \geq 1 \\ A_i & X_{i-1} = 0 \end{cases}$$

where $X_0$ is defined to be zero.

The state-dependent control policy dictates the service times

$$s_i = \begin{cases} S_{X_{i-1}} & X_{i-1} \geq 1 \\ S_1 & X_{i-1} = 0 \end{cases}$$

for all jobs $C_i$. Note that the first job of a busy period starts service when the system size is one; therefore $S_1$ duration is picked in the case of $X_{i-1} = 0$.

If the Poisson arrival rate is $\lambda$, then the state transition matrix for our IMC can be given as

$$\bar{P} = \begin{bmatrix} k_0^1 & k_1^1 & k_2^1 & k_3^1 & ... \\ k_0^1 & k_1^1 & k_2^1 & k_3^1 & ... \\ 0 & k_0^2 & k_1^2 & k_2^2 & ... \\ 0 & 0 & k_0^3 & k_1^3 & ... \\ ... & ... & ... & ... & ... \end{bmatrix} \qquad (16)$$

where $k_n^m$ denotes the probability of $n$ arrivals during the service time $S_m$

$$k_n^m = \frac{(\lambda S_m)^n}{n!} e^{-\lambda S_m}$$

Since the maximum service time of the policy satisfies $S_{\max} = \max_i S_i < 1/\lambda$ as guaranteed by the $\Pi_H$ projection mapping in Eq. 10, the following theorem enables us to determine the steady-state average costs.

**Theorem 1** *Since $\lambda S_{\max} < 1$ due to the projection mapping $\Pi_H$, the IMC is ergodic and it has a steady-state probability distribution.*

*Proof* As it can be observed in the transition matrix (Eq. 16), the chain is both aperiodic and irreducible, independent of the $\lambda S_{\max}$ value. In order to guarantee a steady-steady probability distribution, it remains to show that the chain is also positive recurrent when $\lambda S_{\max} < 1$.

According to Theorem 2 in Foster (1953), this IMC is positive recurrent if there exists a nonnegative solution $\{y_i\}_{i \in \mathbb{N}}$ of the inequalities

$$\sum_{j=0}^{\infty} p_{ij} y_j \leq y_i - 1 \text{ for all } i > 0$$

such that

$$\sum_{j=0}^{\infty} p_{0j} y_j < \infty$$

Note that from Eq. 16, the transition probabilities are given by

$$p_{ij} = \begin{cases} k_j^1 & i = 0 \\ k_{j-i+1}^i & j \geq i-1, i \geq 1 \\ 0 & j < i-1, i \geq 1 \end{cases}$$

Let us suggest the solution

$$y_j = \frac{j}{1 - \lambda S_{\max}}$$

so that

$$\sum_{j=0}^{\infty} p_{ij} y_j = \sum_{j=i-1}^{\infty} k_{j-i+1}^i \frac{j}{1 - \lambda S_{\max}} \tag{17}$$

$$= \sum_{l=0}^{\infty} k_l^i \frac{i-1+l}{1 - \lambda S_{\max}}$$

$$= \frac{(i-1)\sum_{l=0}^{\infty} k_l^i}{1 - \lambda S_{\max}} + \frac{\sum_{l=0}^{\infty} l k_l^i}{1 - \lambda S_{\max}}$$

$$= \frac{i-1+\lambda S_i}{1 - \lambda S_{\max}} \tag{18}$$

$$\leq \frac{i + \lambda S_{\max} - 1}{1 - \lambda S_{\max}} = y_i - 1 \tag{19}$$

because the entries in the $i$th row of a transition matrix should add up to one, and the expected number of arrivals during the service time $S_i$ is $\lambda S_i$. Note that for this solution

$$\sum_{j=0}^{\infty} p_{0j} y_j = \frac{\sum_{j=0}^{\infty} j k_j^1}{1 - \lambda S_{\max}} = \frac{\lambda S_1}{1 - \lambda S_{\max}} < \infty$$

is satisfied. Therefore, the IMC is also positive recurrent, completing the proof. □

This theorem proves the existence of a steady-state probability distribution $\{\pi_i\}_{i \in \mathbb{N}}$ for our IMC. Employing the Little's Law in Little (1961), the steady-state average cost can be evaluated as

$$J = \sum_{i=0}^{\infty} \left( \theta(S_i)\pi_i + \alpha \frac{i\pi_i}{\lambda} \right) \tag{20}$$

for a given state-dependent control policy.

In order to get the gradient of $J$ numerically, we truncate the single-step transition matrix $\bar{P}$, i.e., we consider the transition matrix for an M/D(n)/1/K system for a large enough $K$ to obtain

$$
\bar{P}_t = \begin{bmatrix}
k_0^1 & k_1^1 & k_2^1 & k_3^1 & \ldots & 1 - \sum_{n=0}^{K-2} k_n^1 \\
k_0^1 & k_1^1 & k_2^1 & k_3^1 & \ldots & 1 - \sum_{n=0}^{K-2} k_n^1 \\
0 & k_0^2 & k_1^2 & k_2^2 & \ldots & 1 - \sum_{n=0}^{K-3} k_n^2 \\
0 & 0 & k_0^3 & k_1^3 & \ldots & 1 - \sum_{n=0}^{K-4} k_n^3 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & \ldots & \ldots & 1 - k_0^{K-1}
\end{bmatrix}
$$

Then, we apply the finite forward difference method on the policy to approximate the derivatives. Note that the estimation error is due to both the matrix truncation and the finite difference method.

In the next section, we demonstrate the performances of the proposed methods via a numerical example.

## 4 A numerical example

In order to illustrate the methods proposed above, we solve the following hybrid control problem of a single-stage manufacturing system: Determine the optimal service times $\{s_i^*\}_{i\in\mathbb{N}}$ and the optimal control inputs $\{u_i^*(t)\}_{i\in\mathbb{N}}$ by solving

$$
\min_{s_i \geq 0, u_i(t), i\in\mathbb{N}} \lim_{N\longrightarrow\infty} \frac{1}{N} E\left\{ \sum_{i=1}^{N} \left[ \alpha(d_i - a_i) + \frac{1}{2} h \left( z_i(d_i) - \zeta^d \right)^2 + \int_{\tau_i}^{d_i} \frac{1}{2} r u_i^2(t) dt \right] \right\}
$$

subject to the time-driven dynamics

$$
\dot{z}_i(t) = b u_i(t), \quad \tau_i \leq t \leq \tau_i + s_i \tag{21}
$$

$$
z_i(\tau_i) = \zeta^0 \tag{22}
$$

and the event-driven dynamics

$$
d_i = \tau_i + s_i = \max(d_{i-1}, a_i) + s_i \tag{23}
$$

for given initial and desired states, $\zeta^0$ and $\zeta^d$, and parameters $\alpha, h, r, b$ that are all positive.

Applying calculus of variations to the lower-level optimization problem

$$
\theta(s_i) = \min_{u_i(t)} \left\{ \frac{1}{2} h \left( z_i(d_i) - \zeta^d \right)^2 + \int_{\tau_i}^{d_i} \frac{1}{2} r u_i^2(t) dt \right\}
$$

subject to Eqs. 21 and 22, we obtain the optimal control input function $\bar{u}_i^*(t, s_i)$ and the corresponding process cost function $\theta(s_i)$ as

$$
\bar{u}_i^*(t, s_i) = \frac{\zeta^d - \zeta^0}{\frac{r}{bh} + b s_i}, \quad \tau_i \leq t \leq \tau_i + s_i \tag{24}
$$

and

$$\theta(s_i) = \frac{r\left(\zeta^d - \zeta^0\right)^2}{2b^2\left(\frac{r}{b^2h} + s_i\right)} = \frac{\beta}{\sigma + s_i}$$

where

$$\beta = \frac{r\left(\zeta^d - \zeta^0\right)^2}{2b^2}$$

and

$$\sigma = \frac{r}{b^2h}$$

Note that $\theta(s_i)$ satisfies Assumption 1, as it is decreasing, differentiable, and strictly convex.

The stochastic optimal control problem we undertake is to choose the optimal service times $\{s_i^*\}_{i\in\mathbb{N}}$ that minimize the infinite horizon average cost

$$\min_{s_i \geq 0, i\in\mathbb{N}} \lim_{N\longrightarrow\infty} \frac{1}{N} E\left[\sum_{i=1}^{N}\left[\frac{\beta}{\sigma + s_i} + \alpha(d_i - a_i)\right]\right] \tag{25}$$

subject to Eq. 23. Once the optimal service times are determined, the optimal control inputs are simply determined as

$$u_i^*(t) = \bar{u}_i^*(t, s_i^*), \quad \tau_i^* \leq t \leq \tau_i^* + s_i^*$$

where $\tau_i^*$ is the service initiation time for job $C_i$ on the optimal sample path obtained by applying the optimal service times.

From Eq. 9, the RH controller for this system implements the state-dependent control policy with

$$S_n = \begin{cases} \sqrt{\frac{\beta}{n\alpha}} - \sigma & n \leq \frac{\beta}{\alpha\sigma^2} \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

for all system sizes $n$, which we need to improve. Note that the arrival rate $\lambda$ is not considered in this calculation, so it is possible that some of these service times are larger than the average interarrival time. Therefore, drastic performance improvements can be expected for high arrival rates.

*Example 1* Consider the cost function in Eq. 25 with $\alpha = 2$, $\sigma = 1$, and $\beta = 15$. For Poisson arrivals with arrival rates from the set $\{0.25, 0.5, 1.0, 2.0\}$, the state-dependent control policies are obtained from the RH controller and the proposed

**Table 1** State-dependent service times

| State | RH | $\lambda = 0.25$ | | $\lambda = 0.5$ | | $\lambda = 1.0$ | | $\lambda = 2.0$ | |
|---|---|---|---|---|---|---|---|---|---|
| | | IPA | IMC | IPA | IMC | IPA | IMC | IPA | IMC |
| 1 | 1.7386 | 1.2786 | 1.2932 | 0.9997 | 1.0355 | 0.7062 | 0.7479 | 0.4552 | 0.4915 |
| 2 | 0.9365 | 0.6801 | 0.7179 | 0.5245 | 0.5792 | 0.3418 | 0.4144 | 0.2040 | 0.2612 |
| 3 | 0.5811 | 0.4242 | 0.4380 | 0.3211 | 0.3456 | 0.1806 | 0.2346 | 0.0868 | 0.1312 |
| 4 | 0.3693 | 0.3257 | 0.3204 | 0.2810 | 0.1998 | 0.1038 | 0.1184 | 0.0204 | 0.0446 |

**Table 2**  Performance comparisons

|                  | $\lambda = 0.25$ | $\lambda = 0.5$ | $\lambda = 1.0$ | $\lambda = 2.0$ |
|------------------|------------------|-----------------|-----------------|-----------------|
| RH cost          | 9.9147           | 11.0791         | 12.7984         | 14.0421         |
| IPA cost         | 9.6704           | 10.2806         | 11.2017         | 12.2787         |
| IPA improvement %| 2.46             | 7.21            | 12.48           | 12.56           |
| IMC cost         | 9.6716           | 10.2790         | 11.1856         | 12.2569         |
| IMC improvement %| 2.45             | 7.22            | 12.60           | 12.71           |

methods based on IPA and IMC. During each policy determination, the stochastic approximation algorithms used 1,000 iterations with a step-size of $\epsilon_n = 0.025/n$. IPA utilized sample paths of 10,000 jobs for gradient estimation, while IMC truncation was performed for $K = 15$ and the forward difference method employed service time increments of 0.01. The first four elements of these policies at each rate are presented in Table 1. Note that the RH controller policy does not depend on the arrival rate $\lambda$.

Note that both IPA and IMC selected state-dependent service times that are smaller than the RH controller selection. The reason for such selection is that the former two methods consider the possibility of additional jobs in the same busy period. By Assumption 1, such selection increases the process costs, however, it decreases the system-time costs of the current jobs and the possible future jobs that are expected to join the same busy period.

In order to assess the cost effects, ten sample paths of length 10,000 jobs are generated for each rate. (Note that we could also employ Eq. 20 to obtain the cost estimates without simulation.) The infinite horizon average costs obtained by averaging over these sample paths are compared in Table 2.

As supported by the results in Table 2, for small arrival rates, our methods are expected to yield incremental performance improvements over the RH controllers as the "idleness after current jobs" assumption of the RH controllers is frequently satisfied. This improvement should increase with the increasing arrival rate, as the probability of idleness after current jobs decreases.

## 5 Conclusions

We considered a stochastic single stage hybrid manufacturing system with a certain cost structure, for which the RH controller implemented a state-dependent control policy. Since the RH controller worked under the assumption that an idle period would follow the current set of jobs, the possibility of additional jobs in the same busy period motivated us to search for better state-dependent control policies.

This paper presented ways to obtain better state-dependent policies by improving the RH solution via stochastic approximation methods. The necessary gradient estimation was performed by an IPA estimator under an arbitrary arrival process. Under a Poisson arrival process, an IMC approach was also shown to be feasible for gradient estimation. A numerical example showed the potential of these methods under different Poisson arrival processes.

A complementary topic of ongoing research is the arrival process estimation that can be combined with the methods proposed in this paper to yield online algorithms. In order to give a flavor, an adaptive exponential smoothing method such as the

one in Trigg and Leach ([1967](#)) can be employed to supply the rate of the Poisson arrival process to the proposed methods. There is also the possibility of removing the Poisson arrival assumption and estimating the interarrival time distribution from the observations, under which case we resort to the IPA method only.

## References

Cassandras CG, Lafortune S (2007) Introduction to discrete event systems. Springer

Cassandras CG, Lygeros J (2006) Stochastic hybrid systems (automation and control engineering). CRC Press

Gokbayrak K, Cassandras CG (2000a) A hierarchical decomposition method for optimal control of hybrid systems. In: Proceedings of 39th IEEE conference on decision and control, pp 1816–1821

Gokbayrak K, Cassandras CG (2000b) Hybrid controllers for hierarchically decomposed systems. In: Proc. of 2000 hybrid system control conference, pp 117–129

Bryson AE, Ho YC (1975) Applied optimal control. Hemisphere Publishing Co

Cassandras CG, Mookherjee R (2003a) Receding horizon control for a class of hybrid systems with event uncertainties. In: Proc. of 2003 American control conf, pp 5197–5202

Cassandras CG, Mookherjee R (2003b) Properties of receding horizon controllers for some hybrid systems with event uncertainties. In: Proc. 2003 IFAC conference on analysis and design of hybrid systems, pp 413–418

Cassandras CG, Mookherjee R (2003c) Receding horizon optimal control for some stochastic hybrid systems. In: Proc. of 42nd IEEE conference decision and control, pp 2162–2167

Kushner HJ, Yin GG (2003) Stochastic approximation and recursive algorithms and applications. Springer-Verlag, New York

Glasserman P (1991) Gradient estimation via perturbation analysis. Kluwer Academic Publishers

Ho Y, Cao X (1991) Perturbation analysis of discrete event dynamic systems. Kluwer Academic Publishers

Suri R, Zazanis MA (1988) Perturbation analysis gives strongly consistent sensitivity estimates for the M/G/1 queue. Manag Sci 34(1):39–64

L'Ecuyer P, Glynn PW (1994) Stochastic optimization by simulation: Convergence proofs for the GI/G/1 queue in steady-state. Manag Sci 40(11):1562–1578

Chong EKP, Ramadge PJ (1992) Convergence of recursive optimization algorithms using ipa derivative estimates. Discrete Event Dyn Syst, Theory and Applications 1:339–372

Fu MC (1990) Convergence of the stochastic approximation algorithm for the gi/g/1 queue using infinitesimal perturbation analysis. J Optim Theory Appl 65:149–160

Zazanis MA, Suri R (1994) Perturbation analysis of the GI/GI/1 queue. Queueing Syst 18(3):199–248

Kendall DG (1951) Some problems in the theory of queues. J R Stat Soc B (Methodological) 13(2):151–185

Wolff RW (1982) Poisson arrivals see time averages. Oper Res 30(2):223–231

Foster FG (1953) On the stochastic matrices associated with certain queuing processes. Ann Math Stat 24(3):355–360

Harris CM (1967) Queues with state-dependent stochastic service rates. Oper Res 15(1):117–130

Heyman DP, Sobel MJ (2003) Stochastic models in operations research, vol 1. Dover Publications

Pepyne DL, Cassandras CG (1998) Modeling, analysis, and optimal control of a class of hybrid systems. Discrete Event Dyn Syst, Theory and Applications 8(2):175–201

Pepyne DL, Cassandras CG (2000) Optimal control of hybrid systems in manufacturing. Proc. IEEE 88(7):1108–1123

Cassandras CG, Pepyne DL, Wardi Y (2001) Optimal control of a class of hybrid systems. IEEE Trans Automat Contr 46(3):398–415

Wardi Y, Cassandras CG, Pepyne DL (2001) A backward algorithm for computing optimal controls for single-stage hybrid manufacturing systems. Int J Prod Res 39(2):369–393

Cho YC, Cassandras CG, Pepyne DL (2001) Forward decomposition algorithms for optimal control of a class of hybrid systems. Int J Robust Nonlinear Control 11:497–513

Zhang P, Cassandras CG (2002) An improved forward algorithm for optimal control of a class of hybrid systems. IEEE Trans Automat Contr 47(10):1735–1739

Robbins H, Monro S (1951) A stochastic approximation method. Ann. Math. Stat 22:400–407

Gross D, Shortle JF, Thompson JM, Harris CM (2008) Fundamentals of queueing theory, 4th edn. Wiley
Little JDC (1961) A proof for the queueing formula l=λw. Oper Res 9(3):383–387
Trigg DW, Leach AG (1967) Exponential smoothing with an adaptive response rate. Oper Res Q 18(1):53–59



**Kagan Gokbayrak**  was born in Istanbul, Turkey, in 1972. He received the B.S. degrees in mathematics and in electrical engineering from Bogazici University, Istanbul, in 1995, the M.S. degree in electrical and computer engineering from the University of Massachusetts, Amherst, in 1997, and the Ph.D. degree in manufacturing engineering from Boston University, Boston, MA, in 2001. From 2001 to 2003, he worked as a Network Planning Engineer at Genuity, Inc., Burlington, MA. Since 2003, he has been a faculty member in the Industrial Engineering Department, Bilkent University, Ankara, Turkey. He specializes in the areas of optimal control of discrete-event and hybrid systems, stochastic optimization, and computer simulation, with applications to inventory, manufacturing and communication systems.