Discrete-Time Drift Counteracting Stochastic Optimal Control and Intelligent Vehicle Applications

Ilya Kolmanovsky and John Michelini

Ford Research and Advanced Engineering, Ford Motor Company
2101 Village Road, Dearborn, Michigan, U.S.A.

Abstract. In this paper we present a characterization of a stochastic optimal control in the problem of maximizing expected time to violate constraints for a nonlinear discrete-time system with a measured (but unknown in advance) disturbance modeled by a Markov chain. Such an optimal control may be viewed as providing drift counteraction and is, therefore, referred as the drift counteracting stochastic optimal control. The developments are motivated by an application to an intelligent vehicle which uses an adaptive cruise control to follow a randomly accelerating and decelerating vehicle. In this application, the control objective is to maintain the distance to the lead vehicle within specified limits for as long as possible with only gradual (small) accelerations and decelerations of the follower vehicle so that driver comfort can be increased and fuel economy can be improved.

1 Introduction

In the paper we examine a stochastic optimal control problem motivated by an application of adaptive cruise control to follow a randomly accelerating and decelerating vehicle. For this application, we consider the control objective to maintain the distance to the lead vehicle within specified limits for as long as possible with only very gradual (small) accelerations and decelerations so that to improve fuel economy and increase driver comfort. This and similar application problems can be treated using methods of stochastic drift counteracting optimal control developed in [6].

The paper is organized as follows. In Section 2 we discuss a formulation of the stochastic drift counteracting optimal control problem for a nonlinear discrete-time system with measured (but unknown in advance) disturbance input modeled by a Markov chain. In Section 3 we review the theoretical results [6] pertinent to the characterization and computations of the stochastic optimal control law in this problem. We also present a result to compute expected time to violate the constraints for a fixed control policy, which may be useful in evaluating legacy control solutions. In Section 4 we discuss a simulation example illustrating the application of these methods to a vehicle following, where the lead vehicle speed trajectory is modeled by a Markov chain with known transition probabilities. Concluding remarks are made in Section 5.

Kolmanovsky I. and Michelini J. (2008).
Discrete-Time Drift Counteracting Stochastic Optimal Control and Intelligent Vehicle Applications.
In *Proceedings of the 2nd International Workshop on Intelligent Vehicle Control Systems*, pages 7-16

2 Problem Formulation

Consider a system which can be modeled by nonlinear discrete-time equations,

$$x(t+1) = f(x(t), v(t), w(t)), \tag{1}$$

where $x(t)$ is the state vector, $v(t)$ is the control vector, $w(t)$ is the vector of measured disturbances, and t is an integer, $t \in Z^+$. The system has control constraints which are expressed in the form $v(t) \in U$, where U is a given set.

The behavior of $w(t)$ is modeled by a Markov chain [3] with a finite number of states $w(t) \in W = \{w^j,\ j \in J\}$. The transition probability from $w(t) = w^i \in W$ to $w(t+1) = w^j \in W$ is denoted by $P(w^j|w^i, \bar{x})$. In our treatment of the problem, we permit this transition probability to depend on the state $x(t) = \bar{x}$. For automotive applications, modeling driving conditions using Markov chains for the purpose of applying Stochastic Dynamic Programming to determine fuel and emissions optimal powertrain operating policies has been first proposed in [4].

Our objective is to determine a control function $u(x, w)$, such that with $v(t) = u(x(t), w(t))$, a cost functional of the form,

$$J^{x_0,w_0,u} = E_{x_0,w_0}\tau^{x_0,w_0,u}(G), \tag{2}$$

is maximized. Here $\tau^{x_0,w_0,u}(G) \in Z^+$ denotes the first-time instant the trajectory of $x(t)$ and $w(t)$, denoted by $\{x^u, w^u\}$, resulting from the application of the control $v(t) = u(x(t), w(t))$, exits a prescribed compact set G. See Figure 1.

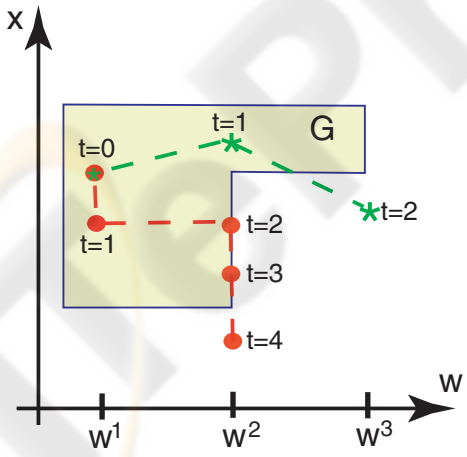

Fig. 1. The set G and two trajectories, $\{x^u, w^u\}$, exiting G at random time instants due to a random realization of $w(t)$. Here $W = \{w^1, w^2, w^3\}$. Note that one of the trajectories exits G at $t = 4$ due to the evolution of $x(t)$ alone, the other trajectory exits G at $t = 2$ due to evolution of both $x(t)$ and $w(t)$.

The specification of the set G reflects constraints existing in the system. Note that $\{x^u, w^u\}$ is a random process, $\tau^{x_0,w_0,u}(G)$ is a random variable and $E_{x_0,w_0}[\cdot]$ denotes

the expectation conditional to initial values of x and w, i.e., $x(0) = x_0, w(0) = w_0$. When clear from the context, we will omit the subscript and square brackets around E.

For continuous-time systems, under an assumption that $w(t)$ is a Wiener or a Poisson process, it can be shown [1] that determining an optimal control in this kind of a problem reduces to solving a non-smooth Partial-Differential Equation (PDE). For instance, for a first order stochastic system, $dx = (v - w_0)dt + \sigma \cdot dw$, where w_0 is a constant, w is a standard Wiener process, the control v satisfies $|v| \leq \bar{v}$, this PDE has the form,

$$\frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2} + \frac{\partial V}{\partial x}(-w_0) + |\frac{\partial V}{\partial x}|\bar{v} + 1 = 0.$$

The boundary conditions for this PDE are $V(x) = 0$ for $x \in \partial G$, where ∂G denotes the boundary of G. The optimal control has the form

$$v = \bar{v} \cdot sign(\frac{\partial V}{\partial x}).$$

Note that this optimal control is of bang-bang type.

As compared to solving the above PDE numerically, the discrete-time treatment of the problem, which is the focus of the present paper, appears to provide a more computationally tractable approach to determining the optimal control. In what follows, we will treat this discrete-time optimal control problem within the framework of optimal stopping [3] and drift counteraction [5], [6] stochastic optimal control.

3 Theoretical Results and Computations

Given a state vector, x^-, and disturbance vectors, $w^-, w^+ \in W$, we define,

$$\begin{aligned} L^u V(x^-, w^-) &\stackrel{\Delta}{=} E_{x^-,w^-}\left[V(f(x^-, u(x^-, w^-), w^-), w^+)\right] - V(x^-, w^-) \\ &= \sum_{j \in J} V(f(x^-, u(x^-, w^-), w^-), w^j) \cdot P(w^j | w^-, x^-) \\ &- V(x^-, w^-). \end{aligned} \tag{3}$$

The following theorem provides sufficient conditions for the optimal control law, $u_*(x, w)$:

Theorem 1: Suppose there exists a control function $u_*(x, w)$ and a continuous, non-negative function $V(x, w)$ such that

$$\begin{aligned} &L^{u_*} V(x, w) + 1 = 0, \text{ if } (x, w) \in G, \\ &L^u V(x, w) + 1 \leq 0, \text{ if } (x, w) \in G, u \neq u_*, \\ &V(x, w) = 0, \text{if } (x, w) \notin G. \end{aligned} \tag{4}$$

Then, u_* maximizes (2), and for all $(x_0, w_0) \in G$, $V(x_0, w_0) = J^{x_0, w_0, u_*}$, $J^{x_0, w_0, u}$ and $E[\tau^{x_0, w_0, u}(G)]$ are finite for any policy u, and the function V, satisfying (4), if exists, is unique.

Proof: The theorem follows as an immediate application of a more general result developed in [6]. More specifically, in [6], a similar result is shown for cost functionals of the form

$$J^{x_0,w_0,u} = E_{x_0,w_0}\sum_{t=0}^{\tau^{x_0,w_0,u}(G)-1} g(x(t),v(t),w(t)),$$

with $g \geq \varepsilon > 0$, of which (2) is a special case with $g = 1$. ■

The following procedure for estimating the expected time to violate constraints for a fixed control law is obtained as an immediate consequence of Theorem 1:

Corollary 1: Given a fixed control law $\bar{u}(x,w)$, suppose there exists a continuous, non-negative function $\bar{V}(x,w)$ such that

$$\begin{aligned} &L^{\bar{u}}\bar{V}(x,w)+1=0, \text{ if } (x,w)\in G,\\ &\bar{V}(x,w)=0, \text{if } (x,w)\notin G. \end{aligned} \tag{5}$$

Then, $E[\tau^{x_0,w_0,\bar{u}}(G)] = \bar{V}(x_0,w_0)$.

We next consider the application of the value iteration approach to (4), assuming, for simplicity of exposition, that f is continuous in x, and that U is compact. The proofs of subsequent results are similar to [6, 5] and are not reproduced here. We define a sequence of value functions using the following iterative process:

$$\begin{aligned} &V_0 \equiv 0\\ &V_n(x,w^i) = \max_{v\in U}\left\{\sum_{j\in J} V_{n-1}(f(x,v,w^i),w^j)P(w^j|w^i,x)+1\right\}, \text{ if } (x,w^i)\in G. \\ &\qquad n>0. \end{aligned} \tag{6}$$

This sequence of functions $\{V_n\}$ yields the following properties:

Theorem 2: Suppose the assumptions of Theorem 1 hold. Then the sequence of functions $\{V_n\}$, defined in (6), is monotonically non-decreasing and $V_n(x,w^i) \leq J^{x,w^i,u_*}$ for all n, x and w^i. Furthermore, $\{V_n\}$ converges pointwise to $V_*(x,w^i) = J^{x,w^i,u_*}$ and this convergence is uniform if J^{x,w^i,u_*} is continuous.

On the computational side, either value iterations or Linear Programming may be used to numerically approximate the solution to (4).

The value iterations (6) produce a sequence of value function approximations, V_n, at specified grid-points $x \in \{x^k,\ k \in K\}$, and a stopping criterion is $|V_n(x,w^i) - V_{n-1}(x,w^i)| \leq \epsilon$ for all $x \in \{x^k,\ k \in K\}$ and $i \in J$, where $\epsilon > 0$ is sufficiently small. In each iteration, once the values of V_{n-1} at the grid-points have been determined, linear or cubic interpolation may be employed to approximate $V_{n-1}(f(x^k,v^m,w^i),w^j)$, on the right-hand side of (6), where $v \in \{v^m,\ m \in M\}$ is a specified grid for v. Formally, the approximate value iterations can be represented as follows,

$$\begin{aligned} &V_0(x^k,w^i) \equiv 0,\\ &V_n(x^k,w^i) = \max_{v^m,m\in M}\left\{\sum_{j\in J} F_{n-1}(f(x^k,v^m,w^i),w^j)\cdot P(w^j|w^i,x^k)+1\right\},\\ &\text{where}\\ &F_{n-1}(x,w^i) = \text{Interpolant}[V_{n-1}](x,w^i) \text{ if } (x,w^i)\in G,\\ &\text{and } F_{n-1}(x,w^i) = 0 \text{ if } (x,w^i)\notin G. \end{aligned} \tag{7}$$

An alternative approach is to seek V in the form,

$$V(x, w^i) = \sum_{l \in L} \theta_l \phi_l(x, w^i),$$

where ϕ_l are specified basis functions satisfying the property that $\phi_l(x, w^i) = 0$ if $(x, w^i) \notin G$. Then relations (4) evaluated over specified grid points $x \in \{x^k, k \in K\}$, $v \in \{v^m, m \in M\}$, and $i \in J$, lead to a Linear Programming problem with respect to $\theta_l, l \in L$:

$$\begin{array}{l} \sum_{l \in L} \theta_l \sum_{k \in K, i \in J} \phi_l(x^k, w^i) \rightarrow \min, \\ \text{subject to} \\ \sum_{l \in L} \theta_l \phi_l(x^k, w^i) \geq 1 + \sum_{l \in L} \theta_l \sum_{j \in J} \phi_l(f(x^k, v^m, w^i), w^j) \cdot P(w^j | w^i, x^k), \\ k \in K, i \in J, m \in M. \end{array} \tag{8}$$

There are many aspects, such as selection of the grids and basis functions, which can be exploited to optimize the computations for specific problems. The dependence of the approximation error on the properties of the grid can be established using, for instance, techniques in Chapter 16 of [2].

Once an approximation of the value function, V_*, is available, an optimal control may be determined from the following relation:

$$\begin{array}{l} u_*(x, w^i) \in argmax_{v \in U} \left\{ 1 + \sum_{j \in J} V_*(f(x, v, w^i), w^j) P(w^j | w^i, x) \right\}, \\ \qquad \text{or} \\ u_*(x, w^i) \in argmax_{v \in U} \left\{ \sum_{j \in J} V_*(f(x, v, w^i), w^j) P(w^j | w^i, x) \right\}. \end{array} \tag{9}$$

4 Vehicle Following Example

In this section we illustrate the above developments with an example of an intelligent vehicle which uses an adaptive cruise control to follow another, randomly accelerating and decelerating vehicle. In this application, the control objective is to maintain the distance to the lead vehicle within specified limits for as long as possible with only very gradual (small) accelerations and decelerations of the follower vehicle to improve fuel economy and increase driver comfort.

The relative distance between two vehicles minus minimum acceptable distance is denoted by s [m], the velocity of the lead vehicle is denoted by v_l [mph], the velocity of the follower vehicle is denoted by v_f [mph] and ΔT is the sampling time period. Assuming that the acceleration a [mph/sec] of the follower vehicle is a control variable, the discrete-time update equations have the following form,

$$\begin{array}{l} s(t+1) = s(t) + 0.1736 \cdot \Delta T \cdot (v_l(t) - v_f(t)), \\ v_f(t+1) = v_f(t) + \Delta T \cdot a(t). \end{array} \tag{10}$$

The factor 0.1736 is introduced because the velocity units are in miles-per-hour (mph) while the distance is in meters (m). With $x = [s, v_f]^{\mathrm{T}}$, $w = v_l$, and $v = a$ as the control, (10) has the form of (1).

We consider a scenario when the vehicles are driven on a road with average speed of 55 mph, minimum speed of 46 mph and maximum speed of about 66 mph. The update period is fixed to $\Delta T = 1$ sec. The lead vehicle velocity, $w = v_l$, is modeled by a Markov chain with 20 discrete levels uniformly distributed between 46 and 66.0013 mph. The transition probabilities (see Figure 2-right) have been constructed from an experimental vehicle velocity trajectory shown in Figure 2-left.

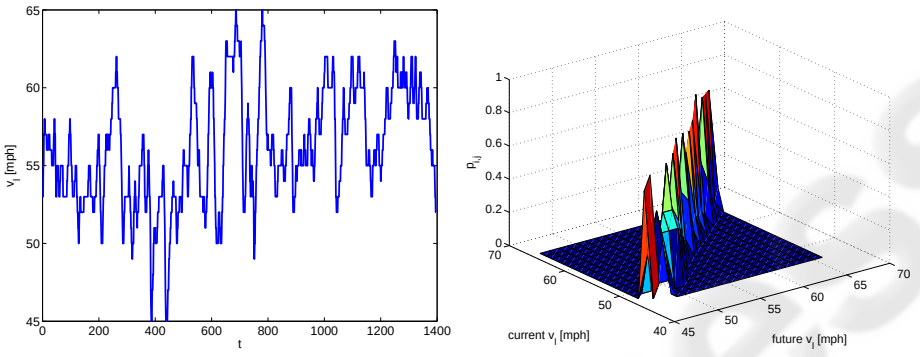

Fig. 2. Left: Experimental vehicle velocity trajectory. Right: Transition probabilities of the Markov chain model of the lead vehicle velocity.

It is desired to maintain the relative distance between two vehicles minus minimum acceptable distance in the range $s \in [0, 20]$ meters. The accelerations of the follower vehicle must be in the range $a \in [-0.5, 0.5]$ mph/sec.

An approximation of the optimal control, u_*, determined using the value iteration approach, is illustrated in Figure 3 while the value function, V_* is illustrated in Figure 4. Note that the u_* and V_* depend on three variables: s, v_f, and v_l. Hence, only the cross-sections of u_* and V_* are shown for a fixed value of v_f. Figure 5 demonstrates numerically the convergence of the value iterations. The grids used were $\{-0.5, -0.25, 0, 0.25, 0.5\}$ for a, $\{46, 47.0527, 48.1054, \cdots, 66.0013\}$ for v_f and v_l, and $\{0, 1.0526, 2.1053, \cdots, 20\}$ for s.

Figures 6 illustrates the time responses when the follower vehicle is controlled with the above approximate stochastic optimal control and when the lead vehicle velocity is a typical realization of the Markov chain trajectory. Note that the accelerations and decelerations of the lead vehicle are up to 2.1 mph/sec, well in excess of 0.5 mph/sec limit imposed on the accelerations and decelerations of the follower vehicle. Figure 7 illustrates the time responses when the lead vehicle velocity is a sequence of non-random accelerations and decelerations.

As can be observed from the plots, the velocity of the follower vehicle, controlled by stochastic drift counteracting optimal control, tracks the velocity of the lead vehicle but with smaller accelerations and decelerations, which satisfy the required limits of 0.5

mph/sec. The controller also enforces the constraints on the relative distance between the vehicles. When the lead vehicle moves at high speed, the follower vehicle increases the relative distance knowing that deceleration of the lead vehicle is more likely and acceleration is less likely. When the lead vehicle moves at low speed, the follower vehicle decreases the relative distance knowing that acceleration of the lead vehicle is more likely and deceleration is less likely. This behavior of the follower vehicle is directionally consistent with the constant headway time policy.

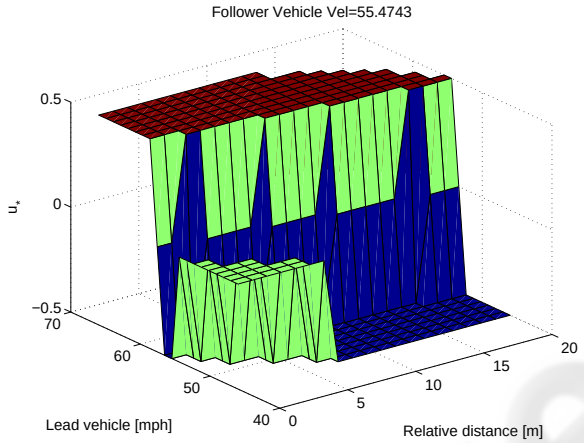

Fig. 3. A cross-section of approximate optimal control.

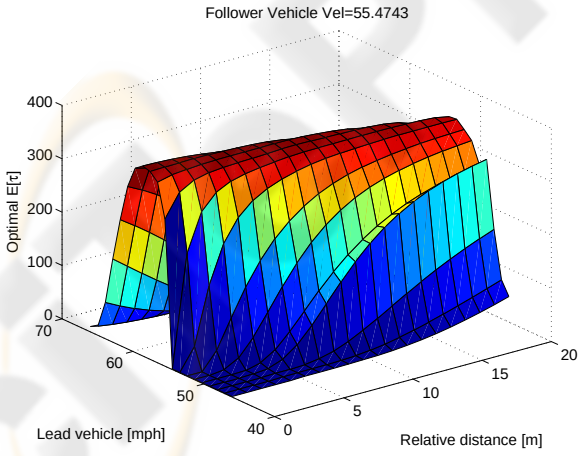

Fig. 4. A cross-section of approximate optimal value function.

Remark 1: The stochastic optimal control maximizes the expected time to violate the constraints, but it cannot entirely eliminate the possibility that the constraints are violated. If the relative distance constraints become violated, a decision needs to be

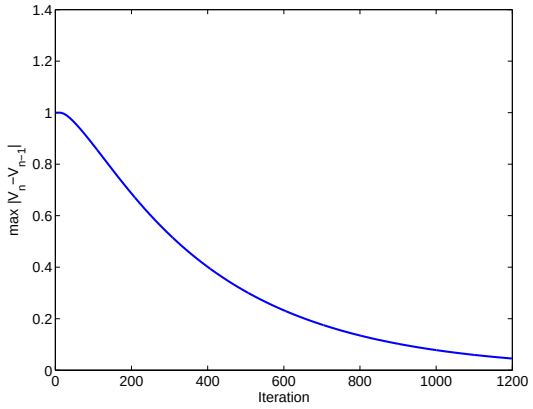

Fig. 5. Maximum of $|V_n(x, w^i) - V_{n-1}(x, w^i)|$ over $x \in \{x^k,\ k \in K\}$ and $i \in J$ as the value iterations progress (i.e., n increases).

made if to discontinue following the lead vehicle since it is too difficult to follow, or to switch to a different controller which may use larger accelerations and decelerations to bring the relative distance and the follower vehicle velocity to values appropriate to re-engage the stochastic optimal controller.

Remark 2: The transition probabilities for the lead vehicle velocity may be estimated on-line by measuring the lead vehicle velocity. Considering that on-board computing power may be limited, fast procedures to approximate u_*, once transition probabilities have been estimated, are desirable. The development of such procedures is a subject of future research.

5 Concluding Remarks

In this paper we presented a method for constructing a stochastic optimal control law in the problem of maximizing expected time to violate constraints for a nonlinear discrete-time system with a measured (but unknown in advance) disturbance modeled by a Markov chain. The resulting control law is referred to as the *stochastic drift counteracting optimal control law*.

A simulation example was considered where an intelligent vehicle follows another, randomly accelerating and decelerating lead vehicle. The control objective in this example was to control the follower vehicle acceleration to maintain the distance to the lead vehicle within specified limits and avoid high accelerations and decelerations so as to improve fuel economy and increase driver comfort. It has been shown that the behavior of the vehicle with the stochastic drift counteracting optimal control law is intuitively reasonable, e.g., the relative distance between the vehicles increases (respectively, decreased) when the lead vehicle is near its maximum (respectively, minimum) speed, as the follower vehicle expects a deceleration (respectively, acceleration) of the lead vehicle.

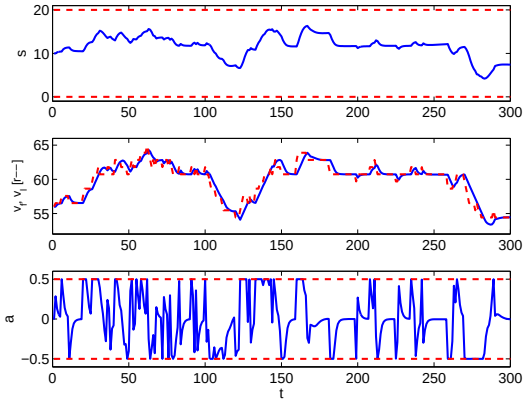

Fig. 6. Relative distance (top), follower and lead vehicle velocities (middle) and follower vehicle acceleration (bottom) in response to random lead vehicle velocity profile. Relative distance constraints and acceleration constraints are indicated on the top plot and bottom plot, respectively, by dashed lines. Dashed lines in the middle plot indicate the lead vehicle velocity.

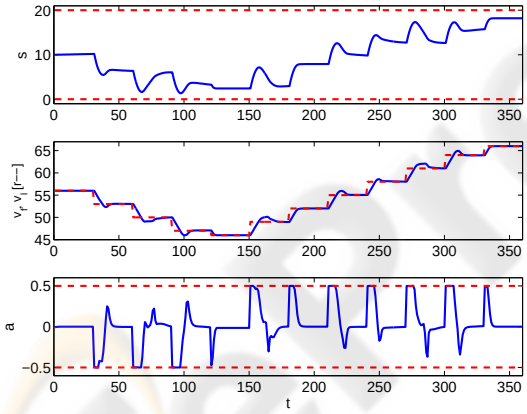

Fig. 7. Relative distance (top), follower and lead vehicle velocities (middle), and follower vehicle acceleration (bottom) in response to non-random lead vehicle velocity profile. Relative distance constraints and acceleration constraints are indicated on the top plot and bottom plot, respectively, by dashed lines. Dashed lines in the middle plot indicate the lead vehicle velocity.

More elaborate vehicle models and lead vehicle speed models can be treated similarly even though, as with any dynamic programming approach, high state dimensions present an obstruction due to "curse of dimensionality." Fast procedures for computing or approximating the stochastic optimal control law, so that it can be reconfigured on-line if the problem parameters or statistical properties of the lead vehicle velocity change, is a subject of future research. While this paper only discussed procedures suitable for off-line computations, these results are already valuable as the resulting stochastic optimal control law can be used as a benchmark for control algorithms developed other approaches, and it can yield valuable insights into the optimal behavior

desirable of the follower vehicle. Also, from Figure 3, it appears that u_* does not have a very elaborate form and so it may inspire a simpler rule-based control law which achieves a near optimal performance.

The theoretical results and computational approaches discussed in this paper can have other applications in intelligent vehicle control and manufacturing. Specifically, they may be applicable in other situations where there is a disturbance with statistical properties that can be modeled in advance (e.g., demands of the driver, changes in the environmental conditions, production orders being scheduled, etc.) while pointwise-in-time constraints on certain critical state and control variables need to be enforced. Along these lines, another example application to Hybrid Electric Vehicle (HEV) control has been discussed in [6].

References

1. Afanas'ev, V.N., Kolmanovskii, V.B., and Nosov, V.R. (1996). *Mathematical Theory of Control Systems Design*. Kluwer Academic Publishers.
2. Altman, E. (1999). *Constrained Markov Decision Processes*. Chapman and Hall/CRC.
3. Dynkin, E.B., and Yushkevich, A.A. (1967). *Markov Processes: Theorems and Problems.* Nauka, Moscow, in Russian. English translation published by Plenum, New York, 1969.
4. Kolmanovsky, I., Sivergina, I., and Lygoe, B. (2002). Optimization of powertrain operating policies for feasibility assessment and calibration: Stochastic dynamic programming approach. *Proceedings of 2002 American Control Conference.* Anchorage, AK, pp. pp. 1425–1430.
5. Kolmanovsky, I., and Maizenberg, T.L. (2002). Optimal containment control for a class of stochastic systems perturbed by Poisson and Wiener processes. *IEEE Transactions on Automatic Control.* Vol. 47, No. 12, pp. 2041–2046.
6. Kolmanovsky, I., Lezhnev, L., and Maizenberg, T.L. (2008). Discrete-time drift counteraction stochastic optimal control: Theory and application-motivated examples. *Automatica*, Vol. 44 , No. 1, pp. 177–184.