JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 32, No. 3, May–June 2009

Engineering Notes

ENGINEERING NOTES are short manuscripts describing new developments or important results of a preliminary nature. These Notes should not exceed 2500 words (where a figure or table counts as 200 words). Following informal review by the Editors, they may be published within a few months of the date of receipt. Style requirements are the same as for regular contributions (see inside back cover).

Computational Nonlinear Stochastic Control

Mrinal Kumar,* S. Chakravorty,† and John L. Junkins‡
Texas A&M University, College Station, Texas 77843

DOI: 10.2514/1.37128

I. Introduction

NUMEROUS fields of science and engineering present the problem of uncertainty propagation through nonlinear dynamic systems [1]. One may be interested in the determination of the response of engineering structures: beams/plates/entire buildings under random excitation (in structure mechanics [2]), the motion of particles under the influence of stochastic force fields (in particle physics [3]), or the computation of the prediction step in the design of a Bayesian filter (in filtering theory) [4,5], among others. All these applications require the study of the time evolution of the probability density function (PDF) $p(t, \mathbf{x})$, corresponding to the state $\mathbf{x}$ of the relevant dynamic system. The PDF is given by the solution to the Fokker–Planck–Kolmogorov equation (FPE), which is a partial differential equation (PDE) in the PDF of the system, defined by the underlying dynamic system's parameters. In this paper, approximate solutions of the FPE are considered, and subsequently leveraged for the design of controllers for nonlinear stochastic dynamic systems. In the past few years, the authors have developed a generalized multiresolution meshless finite element method (FEM) methodology, partition of unity FEM (PUFEM), using the recently developed global local orthogonal mappings methodology to provide the partition of unity functions [6] and the orthogonal local basis functions for the solution of the FPE. The PUFEM is a Galerkin projection method, and the solution is characterized in terms of a finite dimensional representation of the Fokker–Planck operator underlying the problem. The methodology is also highly amenable to parallelization [7–10].

Though the FPE is invaluable in quantifying the uncertainty evolution through nonlinear systems, perhaps its greatest benefit may be in the stochastic analysis, design, and control of nonlinear systems. In the context of nonlinear stochastic control, Markov decision processes (MDPs) have long been one of the most widely used methods for discrete time stochastic control. However, the dynamic programming (DP) equations underlying the MDPs suffer from the curse of dimensionality [11–13]. Various approximate dynamic programming methods have been proposed in the past several years for overcoming the curse of dimensionality [13–17], and can be broadly categorized under the category of functional reinforcement learning. These methods are essentially a model-free method of approximating the optimal control policy in stochastic optimal control problems. These methods generally fall under the category of value function approximation methods [13], policy gradient/approximation methods [15,16], and actor-critic methods [14,17]. These methods attempt to reduce the dimensionality of the DP problem through a compact parametrization of the value functions (with respect to a policy or the optimal function) and the policy function. The difference in the methods is mainly through the parametrization that is employed to achieve the aforementioned goal, and range from nonlinear function approximators such as neural networks [14] to linear approximation architectures [13,18]. These methods learn the optimal policies by repeated simulations on a dynamic system and thus can take a long time to converge to a good policy, especially when the problem has continuous state and control spaces. In contrast, the methodology proposed here is model-based and uses the finite dimensional representation of the underlying parametrized diffusion operator to parametrize both the value function as well as the control policy in the stochastic optimal control problem. Considering the low-order finite dimensional controlled diffusion operator allows us to significantly reduce the dimensionality of the planning problem, while providing a computationally efficient recursive method for obtaining progressively better control policies.

The literature on computational methods for solving continuous time stochastic control problems is relatively sparse when compared to the discrete time problem. One of the approaches is through the use of locally consistent Markov decision processes [19]. In this approach, the continuous controlled diffusion operator is approximated by a finite dimensional Markov chain that satisfies certain local consistency conditions, namely that its drift and diffusion coefficients match that of the original process locally. The resulting finite state MDP is solved by standard DP techniques, such as value iteration and policy iteration. The method relies on a finite difference discretization and thus can be computationally very intensive in higher-dimensional spaces. In another approach [20,21], the diffusion process is approximated by a finite dimensional Markov chain through the application of generalized cell-to-cell mapping [22]. However, even this method suffers from the curse of dimensionality because it involves discretizing the state space into a grid, which becomes increasingly infeasible as the dimension of the system grows. Also, finite difference and finite element methods have been applied directly to the nonlinear Hamilton–Jacobi–Bellman (HJB) partial differential equation [23,24]. The method proposed here differs in that it uses policy iteration in the original infinite dimensional function space, along with a finite dimensional representation of the controlled diffusion operator to solve the problem. Considering a lower-order approximation of the underlying operator results in a significant reduction in dimensionality of the computational problem. Using the policy iteration algorithm typically results in having to solve a sequence of a few linear equations (typically less than five) before practical convergence is obtained, as opposed to solving a high-dimensional nonlinear equation if the original nonlinear HJB equation is solved.

The literature for solving deterministic optimal control problems in continuous time is relatively mature when compared to its stochastic counterpart. The method of successive approximations/

Received 13 February 2008; revision received 10 December 2008; accepted for publication 16 December 2008.

*Ph.D. Student, Department of Aerospace Engineering; mrinal@neo.tamu.edu. http://people.tamu.edu/~mrinal. Student Member AIAA.
†Assistant Professor, Faculty of Aerospace Engineering. Member AIAA.
‡Royce E. Weisenbaker Distinguished Professor, Aerospace Engineering. Fellow AIAA.

policy iteration has widely been used to solve deterministic optimal control problems. Various different methods have been proposed for solving the policy evaluation step in the policy iteration algorithms that include Galerkin-based methods [25,26] and neural-network-based methods [27–31], among others. The methodology outlined here can be viewed as an extension of this extensive body of work to the stochastic optimal control problem. However, it should be noted that the extension is far from trivial as the stochastic optimal control problem, especially the control of degenerate diffusions, has pathologies that have to be treated carefully to devise effective computational methods for solving such problems. We would like to mention that we are interested in the feedback control of dynamic systems here, that is, in the solution of the HJB equation, as opposed to solving the open-loop optimal control problem based on the Pontryagin minimum principle [32] that results in a two-point boundary value problem involving the states and costates of the dynamic system. Please see [32–34] for more on this approach to solving the optimal control problem. Also, we would like to state here that, to the best of our knowledge, there is no stochastic equivalent to the two-point boundary value problems that result from the Pontryagin minimum principle being applied to the deterministic optimal control problem.

II. Forward and Backward Kolmogorov Equations

In this section, we shall introduce the forward Kolmogorov or the Fokker–Planck equation and the backward Kolmogorov equation (BKE). Whereas the former is central to the propagation of uncertainty and nonlinear filtering of dynamic systems, the latter is paramount in the solution of stochastic control problems.

The FPE and the BKE operators are, respectively, given by the following expressions:

$$L_f(p) = -\sum_i \frac{\partial}{\partial x_i} f_i p + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial x_i x_j}(GG^t)_{ij} p \tag{1}$$

$$L_b(V) = \sum_i f_i \frac{\partial}{\partial x_i} V + \frac{1}{2}\sum_{i,j} (GG^t)_{ij} \frac{\partial^2}{\partial x_i x_j} V \tag{2}$$

given the nonlinear stochastic multidimensional system

$$\dot{X} = F(X) + G(X)W \tag{3}$$

$$F(X) = [f_1(X), \ldots, f_n(X)] \tag{4}$$

where W is multidimensional white noise. Note that in most multidimensional systems, especially mechanical ones, the matrix G is rank deficient because the input to the systems is at only the velocity level variables, and hence so is the matrix GG^t. In general, a diffusion operator may be written as

$$\mathcal{D}(V) \equiv \sum_i a_i(X)\frac{\partial V}{\partial x_i} + \frac{1}{2}\sum_{i,j} d_{ij}(X)\frac{\partial^2 V}{\partial x_i x_j} \tag{5}$$

where the $A(x) = [a_1(X), \ldots, a_N(X)]$ is a drift vector and $D(X) = [d_{ij}(X)]$ is a diffusion matrix. In the case when $D(x)$ is not strictly positive definite for all x, the diffusion process is termed degenerate (it is also sometimes said that the diffusion process violates the uniform ellipticity condition). For diffusions arising out of stochastic dynamic systems, that is, the BKE operator, it follows that they are degenerate diffusions due to the rank deficiency of G. In general, the theoretical treatment of the solution of such equations is made very difficult by the degeneracy because classical solutions to PDEs governed by such operators may not exist. In these cases, the generalized notion of viscosity solutions needs to be introduced [35–37].

The FPE and the BKE are paramount in the propagation of uncertainty through, and the filtering and control of, stochastic nonlinear systems. As aforementioned, though the theoretical treatment of these operators in the multidimensional case is very difficult, we may nevertheless obtain a finite dimensional, that is, a matrix representation of the preceding operators to practically solve these problems, at least in an approximate fashion within the resulting finite dimensional space.

The Galerkin projection method is a general method to solve partial differential equations and is often used to solve the Fokker–Planck–Kolmogorov (FPK) equation. Let $p(t, x)$ represent the time varying solution to the FPK equation. Let $\{\phi_1(x), \ldots, \phi_n(x), \ldots\}$ represent a set of basis functions on some compact domain D. The various choices of the basis functions lead to different methodologies. The Galerkin method makes the assumption that the PDF $p(t, x)$ can be written as a linear combination of the basis functions $\phi_i(x)$ with time varying coefficients, that is,

$$p(t, x) = \sum_{i=1}^{\infty} c_i(t)\phi_i(x) \tag{6}$$

The key to obtaining the solution to the PDF $p(t, x)$ is to find the coefficients $c_i(t)$ in the preceding equation. The Galerkin method achieves this by truncating the expansion of $p(t, x)$ for some finite N and projecting the equation error onto the set of basis functions $\{\phi_i(x)\}$, that is,

$$p(t, x) \approx \sum_{i=1}^{N} c_i(t)\phi_i(x) \tag{7}$$

and

$$\left\langle \frac{\partial \sum_{i=1}^{N} c_i(t)\phi_i(x)}{\partial t} - L_f\left(\sum_{i=1}^{N} c_i(t)\phi_i(x)\right), \phi_j(x)\right\rangle = 0 \tag{8}$$

where $\langle ., . \rangle$ represents the inner product on $L_2(D)$, the space of Lebesgue integrable functions on the domain D, and L_f represents the FPE operator. Note that the aforementioned represents a finite set of differential equations in the coefficients $\{c_i(t)\}$, which can be solved for the coefficients $c_i(t)$, given by

$$\sum_{i=1}^{N} \dot{c}_i\langle\phi_i, \phi_j\rangle - \sum_{i=1}^{N} c_i\langle L_f(\phi_i), \phi_j\rangle = 0, \quad \forall\ j \tag{9}$$

Let

$$M = [\langle\phi_i, \phi_j\rangle] \tag{10}$$

$$K = [\langle L_f(\phi_i), \phi_j\rangle] \tag{11}$$

The pair (M, K) represents the finite dimensional approximation of the FPE operator. If the basis functions are mutually orthogonal, then the stiffness matrix K represents the finite dimensional representation of the FPE operator. An eigendecomposition of the pair (M, K) yields the approximation of the eigenstructure of the infinite dimensional FPE operator. The uncertainty propagation problem, as well as the prediction step of the filtering problem, involve the solution of the linear differential equation

$$M\dot{C} + KC = 0 \tag{12}$$

over a finite or infinite time interval. Given the eigenstructure of the preceding linear time invariant system has been found and stored a priori, the preceding equation can be trivially solved in real time. Moreover, if the FPK equation is asymptotically stable, it can be shown after some work that the eigenvalues of the FPE operator are all real and nonpositive. In the following section, we shall show how the FPE/BKE operator may be used to solve the stochastic nonlinear control problem.

The Galerkin method is not the only method that is available to obtain the finite dimensional representation of the FPE/BKE operator. The generalized Galerkin residual projection method can be used to obtain the representation. In the generalized procedure, the residual (or equation error) obtained by substituting the approximate solution into the equation of interest is projected onto a set of generalized functions instead of the basis functions themselves. For instance, collocation/pseudospectral methods involve projecting the equation error onto a set of delta functions located at a specially

selected set of collocation points [33,34]. A full comparison of these methods for the computational solution of the FPE/BKE equations is beyond the scope of this paper. However, we would like to mention that the generalized FEM-based methods developed by the authors [7–10] have consistently outperformed other competing Galerkin residual projection methods.

III. Nonlinear Stochastic Control

An iterative approach to solving the Hamilton–Jacobi–Bellman equation is presented now for the optimal control problems of nonlinear stochastic dynamic systems. Consider the following dynamic system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u} + \mathbf{h}(\mathbf{x})\mathbf{w}(t); \qquad \mathbf{x}(t=0) = \mathbf{x}_0 \tag{13}$$

where $\mathbf{u}$ is the control input vector, and $\mathbf{w}$ represents white noise excitation with intensity Q. We are interested in minimizing the following discounted cost functional over the infinite horizon:

$$J(\mathbf{x}_0) = E\left[\int_0^\infty \{l[\varphi(t)] + \|\mathbf{u}[\varphi(t)]\|_R^2\} e^{-\beta t}\,\mathrm{d}t\right] \tag{14}$$

where $\varphi(t)$ is a trajectory that solves Eq. (13), l is often referred to as the state-penalty function, and β is a given discount factor. The infinite horizon problem results in a stationary, that is, time invariant control law. However, for the case of systems with additive noise, that is, when $h(x)$ is independent of x, as would be the case for the instance in the linear quadratic Gaussian problem, the cost to go is undefined because the trajectories of the system never decay to zero. Hence, in that case, the discounting is a practical step to ensure a bounded cost-to-go function. The discount factor may also be interpreted as a finite horizon over which the control law tries to optimize the system performance. For this problem, the optimal control law $\mathbf{u}^*(\mathbf{x})$ is known to be given in terms of a *value function* $V^*(\mathbf{x})$ as follows:

$$\mathbf{u}^*(\mathbf{x}) = -\frac{1}{2}R^{-1}\mathbf{g}^T\frac{\partial V^*}{\partial \mathbf{x}}(\mathbf{x}) \tag{15}$$

where the value function $V^*(x)$ solves the following well-known (stationary) Hamilton-Jacobi-Bellman equation:

$$\frac{\partial V^{*^T}}{\partial \mathbf{x}}\mathbf{f} + \frac{1}{2}\mathbf{h}\mathbf{Q}\mathbf{h}^T\frac{\partial^2 V^*}{\partial \mathbf{x}^2} + l - \frac{1}{4}\frac{\partial V^*}{\partial \mathbf{x}}\mathbf{g}R^{-1}\mathbf{g}^T\frac{\partial V^*}{\partial \mathbf{x}} - \beta V^* = 0$$
$$V^*(\mathbf{0}) = 0 \tag{16}$$

The preceding framework [solving for the optimal control law $u^*(x)$] is known as the $\mathcal{H}_2$ control paradigm. Note that designing the optimal control in the preceding framework requires solving a nonlinear PDE [Eq. (16)], which is, in general, a very difficult problem. It is, however, possible to restructure the preceding equations so that the central problem can be reduced to solving a sequence of linear PDEs (a much easier proposition). This is achieved via the substitution of Eq. (15) into the quadratic term involving the gradient of the value function in the HJB [Eq. (16)]. Doing so gives us the following equivalent form of the HJB:

$$\frac{\partial V^{*^T}}{\partial \mathbf{x}}(\mathbf{f} + \mathbf{g}\mathbf{u}) + \frac{1}{2}\mathbf{h}\mathbf{Q}\mathbf{h}^T\frac{\partial^2 V^*}{\partial \mathbf{x}^2} + l + \|\mathbf{u}\|_R^2 - \beta V^* = 0 \tag{17}$$

The preceding form of the HJB is known as the *generalized* HJB. Notice that the substitution discussed previously has converted the nonlinear PDE to a linear PDE in the value function $V^*(\mathbf{x})$. Equation (17) forms the core of a policy iteration algorithm in which the optimal control law can be obtained by iterating upon an initial stabilizing control policy as follows:

1) **Let** $u^{(0)}$ be an initial stabilizing control law (policy) for the dynamic system [Eq. (13)], that is, the FPE corresponding to the closed loop under $u^{(0)}$ is asymptotically stable.

2) **For** $i = 0$ to ∞

Solve for $V^{(i)}$ from

$$\frac{\partial V^{(i)^T}}{\partial \mathbf{x}}\left(\mathbf{f} + \mathbf{g}\mathbf{u}^{(i)}\right) + \frac{1}{2}\mathbf{h}\mathbf{Q}\mathbf{h}^T\frac{\partial^2 V^{(i)}}{\partial \mathbf{x}^2} + l + \|\mathbf{u}^{(i)}\|_R^2 - \beta V^{(i)} = 0 \tag{18}$$

Update policy as

$$\mathbf{u}^{(i+1)}(\mathbf{x}) = -\frac{1}{2}R^{-1}\mathbf{g}^T\frac{\partial V^{(i)}}{\partial \mathbf{x}}(\mathbf{x}) \tag{19}$$

3) **End [Policy Iteration]**

The convergence of the preceding algorithm has been proven for the deterministic case ($\mathbf{Q} = 0$) in [25,26] and is also true for the stochastic case [38]. However, conditions for the existence of classical solutions to the linear elliptic PDE in the policy evaluation step [Eq. (18)] and the asymptotic stability of the closed-loop systems under the resulting control policies, in the sense that the associated FPE is stable, have not been obtained to the best knowledge of the authors. In this paper, we shall assume that the policy evaluation step admits a classical solution as well as that the sequence of control policies generated asymptotically stabilize the closed-loop system. In fact, these assumptions are borne out by the numerical examples considered in the next section. The great advantage of using the policy iteration algorithm over directly trying to solve the nonlinear HJB equation is that the algorithm typically converges in very few steps (five or less).

Let us take a closer look at Eq. (18): writing it in operator form, we have

$$(\mathcal{L} - \beta)V = q \tag{20}$$

where

$$\mathcal{L} = (\mathbf{f} + \mathbf{g}\mathbf{u})^T\frac{\partial}{\partial \mathbf{x}} + \frac{1}{2}\mathbf{h}\mathbf{Q}\mathbf{h}^T\frac{\partial^2}{\partial \mathbf{x}^2} \tag{21}$$

$$q = -(l + \|\mathbf{u}\|_R^2) \tag{22}$$

Notice that the operator $\mathcal{L}(\cdot)$ is identical to the BKE operator of the closed loop under control policy u. The approximation for the value function is then written in terms of basis functions in the following manner:

$$\hat{V} = \sum_{i=1}^{N} c_i\phi_i(\mathbf{x}) \tag{23}$$

where $\phi_i(\mathbf{x})$ could be local or global basis functions depending on the nature of the approximation being used. In the current paper, we use global polynomials (i.e., $\phi(\cdot)_i$ are polynomials supported on Ω). Then, following the Galerkin approach, we project the residual error resulting from plugging the approximation [Eq. (23)] into the generalized HJB [Eq. (18)] onto the space of polynomials on Ω as follows:

$$\sum_{i=1}^{N} c_i\int_\Omega (\mathcal{L} - \beta)[\phi_i(\mathbf{x})]\phi_j(\mathbf{x})\,\mathrm{d}\Omega = \int_\Omega q\phi_j\,\mathrm{d}\Omega$$
$$j = 1, 2, \ldots, N \tag{24}$$

Equation (24) is equivalent to a linear system of algebraic equations in c_i:

$$\mathbf{K}\mathbf{c} = \mathbf{f} \tag{25}$$

where

$$K_{ij} = \int_\Omega (\mathcal{L} - \beta)[\phi_j]\phi_i\,\mathrm{d}\Omega \tag{26}$$

$$f_i = \int_\Omega q\phi_i\,\mathrm{d}\Omega \tag{27}$$

Note that

$$K = K_b - \beta I \tag{28}$$

where K_b is the finite dimensional representation of the backward Kolmogorov (BK) operator. Some notes on the admissibility of basis functions (ϕ_i) are due here. Notice that the BK operator involves only derivatives of the unknown function $V(x)$. It is thus clear that, when the discount factor β is zero, it is not possible to include a complete set of basis functions without making the stiffness matrix **K** singular ($\phi \equiv 1$ causes a rank deficiency of one). However, so far as solving Eq. (25) is concerned, it is not important to have the constant basis function in the approximation space. In theory, a nontrivial β allows us to include $\phi \equiv 1$ in the basis set; but, in practice, it would require a large discount factor before K is reasonably well conditioned for inversion, which in turn would take the solution far from being optimal because then the control policy becomes very shortsighted.

The preceding formulation can also be carried out on local subdomains by choosing ϕ_i to be supported on subdomains, forming a cover for the global domain Ω. Then, the projection integrals (24) would be evaluated over the local subdomains. This is the basic approach in finite element methods and meshless finite element methods, where the local subdomains overlap each other. The authors have developed a method based on the partition of unity finite element method to solve the forward (Fokker–Planck) and backward Kolmogorov equations [7–10]. Finally, we would like to note that the policy iteration algorithm can be used either offline or online because the sequence of controllers obtained are monotonically improving with respect to the defined cost function. Hence, a particular controller can be used on the system until its improvement is obtained by solving the policy evaluation step, followed by the policy improvement step, which can then be implemented on the system.

IV. Numerical Example

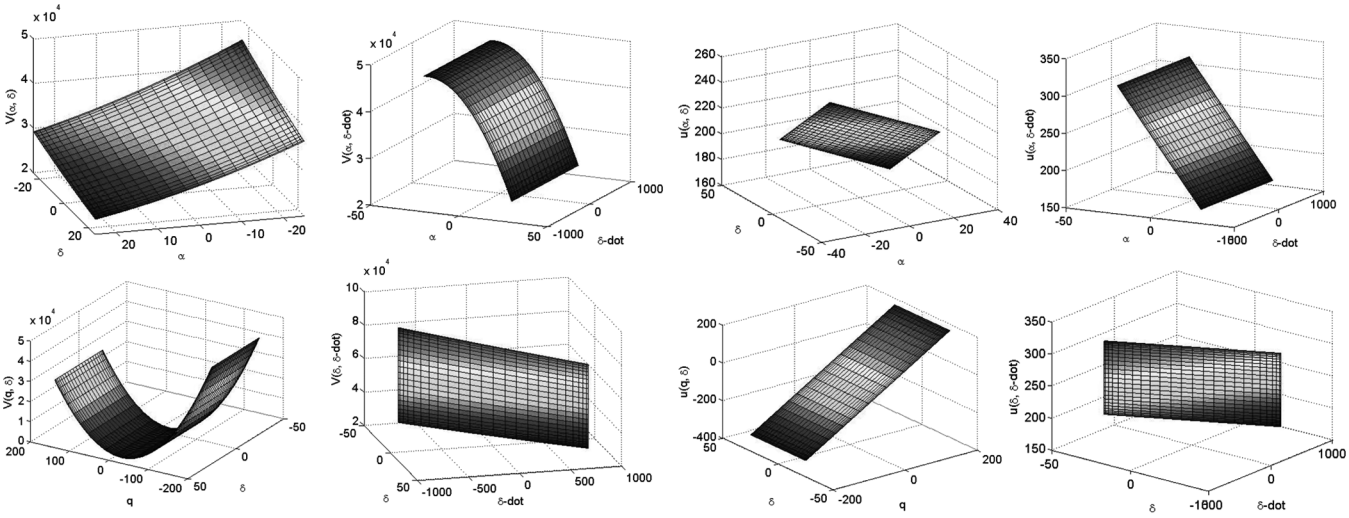
We consider a four-dimensional system that models the pitch motion control for a missile autopilot. This model is the same as that considered in [25,26] with a noise term added to all kinetic equations:

$$\dot{q} = \frac{M_y}{I_y} + g_q w_1(t), \qquad M_y = C_m(\alpha, \delta) QSd \tag{29}$$

$$\dot{\alpha} = \frac{\cos^2\alpha}{mU} F_z + q + g_\alpha w_2(t), \qquad F_z = C_n(\alpha, \delta) Sd \tag{30}$$

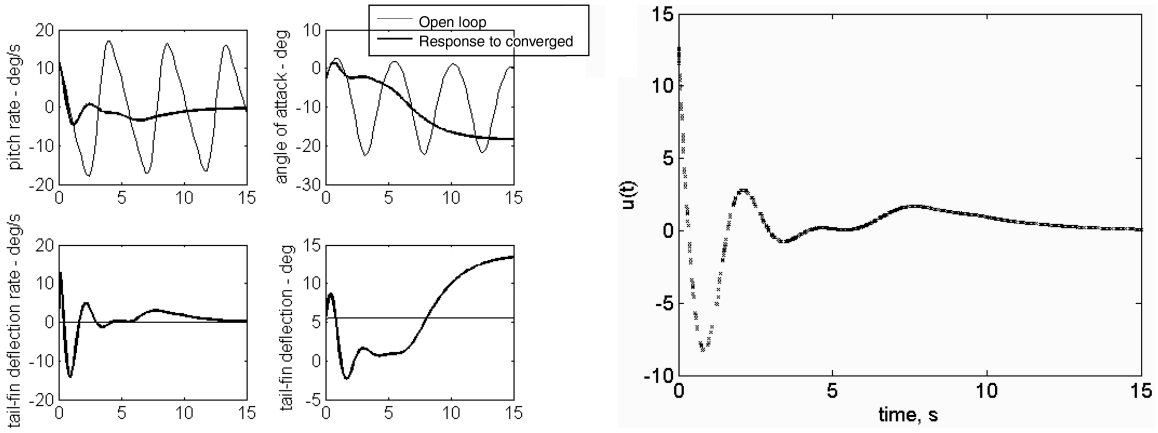
$$\ddot{\delta} = -2\zeta\omega_n \dot{\delta} + \omega_n^2(\delta_c - \delta) + g_\delta w_3(t) \tag{31}$$

The control input in the preceding equations is $\delta_c(t)$, which is the commanded tail-fin deflection, whereas $\delta(t)$ denotes the actual tail-fin deflection. Pitch rate and angle of attack are denoted by q and α,



a) Converged value function surface **b) Converged control input surface**

Fig. 1 Converged results for the missile pitch controller, showing various sections of the 4-D state space.



a) Optimal path obtained, shown alongside open-loop trajectory **b) Optimal control law obtained**

Fig. 2 System response of the missile pitch controller.

respectively. Independent components of a three-dimensional white noise process are given by $w_q(t)$, $w_\alpha(t)$, and $w_\delta(t)$. The aerodynamic coefficients are given by the following nonlinear functions:

$$C_m(\alpha, \delta) = b_1\alpha^3 + b_2\alpha|\alpha| + b_3\alpha + b_4\delta \tag{32}$$

$$C_n(\alpha, \delta) = a_1\alpha^3 + a_2\alpha|\alpha| + a_3\alpha + a_4\delta \tag{33}$$

The values of the various constants can be found in Beard et al. [25]. The cost function to be minimized is given by

$$J(\mathbf{x}_0) = E\left[\int_0^\infty \frac{1}{2}(q - q_{ss})^2 + 25(\alpha - \alpha_{ss})^2 + \left(\frac{F_z - F_{zss}}{m}\right)^2 + \frac{1}{2}\|u - \delta_{ss}\|_R^2 e^{-\beta t}\,\mathrm{d}t\right] \tag{34}$$

where the subscript ss denotes steady-state values. We use the same starting stabilizing control as that used in Beard et al.:

$$u^{(0)} = 0.08(q + q_{ss}) + 0.38(\alpha + \alpha_{ss}) + 0.37\frac{F_z}{m} \tag{35}$$

A complete basis of global quadratic shape functions was used to approximate the value function in the policy iteration algorithm. Various sections of the converged value function and optimal control surface are shown in Fig. 1. The optimal trajectories are shown in Fig. 1. Figure 2 shows the system response to the optimal control law obtained. Notice that the pitch rate settles to zero and the angle of attack acquires the desired steady-state value.

V. Conclusions

In this paper, we have outlined a computational methodology for solving the nonlinear stochastic optimal control based on the policy iteration algorithm and a finite dimensional approximation of the controlled Fokker–Planck–Kolomogorov/diffusion operator. The computational methodology was also tested on a number of test cases where it was shown to have satisfactory performance. The next step in our research will be to extend these methods to more complex stochastic systems such as jump diffusion processes and stochastic hybrid systems. We envisage using a hierarchical hybrid methodology to solve the aforementioned problems, wherein the methods outlined in this paper would form the lower, continuous level of the hierarchy.

References

[1] Risken, H., *The Fokker Planck Equation: Methods of Solution and Applications*, Springer Series in Synergetics, Springer-Verlag, Berlin, 1996.

[2] Polidori, D. C., and Beck, J. L., "Approximate Solutions for Nonlinear Vibration Problems," *Probabilistic Engineering Mechanics*, Vol. 11, No. 3, 1996, pp. 179–185. doi:10.1016/0266-8920(96)00011-2

[3] Wang, M. C., and Uhlenbeck, G., "On the Theory of Brownian Motion II," *Reviews of Modern Physics*, Vol. 17, Nos. 2–3, 1945, pp. 323–342. doi:10.1103/RevModPhys.17.323

[4] Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[5] Doucet, A., deFrietas, J. F. G., Gordon, N. J. (eds.), *Sequential Monte Carlo Methods in Practice*, Springer–Verlag, New York, 2001.

[6] Junkins, J. L., Singla, P., Griffith, T. D., and Henderson, T., "Orthogonal Global/Local Approximation in n-Dimensions: Applications to Input/Output Approximation," *6th International Conference on Dynamics and Control of Systems and Structures in Space*, Inst. of Electrical and Electronics Engineers, New York, 2004; also *Computer Methods in Engineering Sciences* (to be published).

[7] Kumar, M., Singla, P., Junkins, J., and Chakravorty, S., "A Multi-Resolution Meshless Approach to Steady State Uncertainty Determination in Nonlinear Dynamic Systems," *38th IEEE Southeastern Symposium on Systems Theory*, Inst. of Electrical and Electronics Engineers, New York, March 2006, pp. 344–348.

[8] Kumar, M., Singla, P., Chakravorty, S., and Junkins, J. L., "The Partition of Unity Finite Element Approach to the Stationary Fokker–Planck Equation," AIAA Paper 2006-6285, Aug. 2006.

[9] Chakravorty, S., "A Homotopic Galerkin Approach to the Solution of the Fokker-Planck-Kolmogorov Equation," *Proceedings of the 2006 American Control Conference*, Inst. of Electrical and Electronics Engineers, New York, 2006.

[10] "A Semianalytic Meshless Approach to the Fokker–Planck Equation with Application to Hybrid Systems," *Proceedings of the 2007 IEEE Conference on Decision and Control*, Inst. of Electrical and Electronics Engineers, New York, 2007, pp. 3078–3083.

[11] Bellman, R. E., *Dynamic Programming*, Princeton Univ. Press, Princeton, NJ, 1957.

[12] Bertsekas, D. P., *Dynamic Programming and Optimal Control*, Vols. 1–2, Athena Scientific, Belmont, MA, 2000.

[13] Bertsekas, D. P., and Tsitsiklis, J. N., *Neuro-Dynamic Programming*, Athena Scientific, Belmont, MA, 1996.

[14] Werbos, P., "Approximate Dynamic Programming for Real-Time Control and Neural Modeling," *Handbook of Intelligent Control*, edited by P. White and W. Sofge, Van Nostrand Reinhold, New York, 1992.

[15] Marbach, P., "Simulation Based Optimization of Markov Reward Processes," Ph.D. Thesis, Massachusetts Inst. of Technology, Boston, 1999.

[16] Baxter, J., and Bartlett, P., "Infinite Horizon Policy-Gradient Estimation," *Journal of Artificial Intelligence Research*, Vol. 15, No. 4, 2001, pp. 319–350. doi:10.1016/S0954-1810(01)00028-0

[17] Sutton, R. S., MacAllester, D., Singh, S., and Mansour, Y., "Policy Gradient Methods for Reinforcement Learning with Function Approximation," *Advances in Neural Information Proceedings Symposium*, MIT Press, Cambridge, MA, 2000, pp. 1057–1063.

[18] Lagoudakis, M., and Parr, R., "Least Squares Policy Iteration," *Journal of Machine Learning Research*, Vol. 4, Dec. 2003, pp. 1107–1149. doi:10.1162/jmlr.2003.4.6.1107

[19] Kushner, H. J., and Dupuis, P., *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer, New York, 2001.

[20] Crespo, L. G., and Sun, J. Q., "Stochastic Optimal Control of Nonlinear Systems via Short-Time Gaussian Approximation and Cell Mapping," *Nonlinear Dynamics*, Vol. 28, Nos. 3–4, 2002, pp. 323–342. doi:10.1023/A:1015600430713

[21] Crespo, L. G., and Sun, J. Q., "Stochastic Optimal Control via Bellman's Principle," *Automatica*, Vol. 39, Dec. 2003, pp. 2109–2114.

[22] Hsu, C. S., *Cell to Cell Mapping: A Method for the Global Analysis of Nonlinear Systems*, Springer–Verlag, New York, 1987.

[23] Hanson, F. B., "Techniques in Computational Stochastic Dynamic Programming," *Control and Dynamical Systems: Advances in Theory and Applications*, edited by Leondes, C. T., Vol. 76, Academic Press, New York, 1996.

[24] Hanson, F. B., "Computational Dynamic Programming on a Vector Multiprocessor," *IEEE Transactions on Automatic Control*, Vol. 36, April 1991, pp. 507–511.

[25] Beard, R., Saridis, G., and Wen, J., "Galerkin Approximation Of The Generalized Hamilton-Jacobi-Bellman Equation," *Automatica*, Vol. 33, No. 12, 1997, pp. 2159–2177. doi:10.1016/S0005-1098(97)00128-3

[26] Beard, R., Saridis, G., and Wen, Y., "Approximate Solutions to the Time-Invariant Hamilton-Jacobi-Bellman Equation," *Journal of Optimization Theory and Applications*, Vol. 96, No. 3, 1998, pp. 589–626. doi:10.1023/A:1022664528457

[27] Lewis, F. L., *Neural Network Control of Robotic Manipulators and Nonlinear Systems*, Taylor and Francis, Philadelphia, 1999.

[28] Kahlaf, M. A., and Lewis, F. L., "Nearly Optimal State Feedback Control of Constrained Nonlinear System Using a Neural Network HJB," *Annual Reviews in Control*, Vol. 28, No. 2, 2004, pp. 239–251. doi:10.1016/j.arcontrol.2004.07.002

[29] Abu-Khalaf, M., Huang, J., and Lewis, F. L., *Nonlinear $H2/H\text{-}\infty$ Constrained Feedback Control: A Practical Approach Using Neural Networks*, Springer–Verlag, New York, June 2006.

[30] Padhi, R., and Balakrishnan, S. N., "Proper Orthogonal Decomposition Based Optimal Neurocontrol Synthesis of a Chemical Reactor Process Using Approximate Dynamic Programming," *Neural Networks*, Vol. 16, No. 5–6, 2003, pp. 719–728. doi:10.1016/S0893-6080(03)00131-X

[31] Ferrari, S., "Algebraic and Adaptive Learning in Neural Control Systems," Ph.D. Thesis, Princeton Univ., Princeton, NJ, 2002.

[32] Bryson, A. E., and Ho, Y., *Applied Optimal Control*, Hemisphere, Washington, D.C., 1975.

[33] Ross, I. M., and Fahroo, F., "Pseudospectral Knotting Methods for Solving Nonsmooth Optimal Control Problems," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 3, 2004, pp. 397–405.
doi:10.2514/1.3426

[34] Ross, I. M., and Fahroo, F., "Issues in real-time computation of optimal control," *Mathematical and Computer Modelling*, Vol. 43, Nos. 9–102006, pp. 1172–1188.
doi:10.1016/j.mcm.2005.05.021

[35] Fleming, W. H., and Soner, H., *Controlled Markov Processes and Viscosity Solutions*, Springer, New York, 2001.

[36] Crandall, M. G., and Lions, P. L., "Viscosity Solutions of Hamilton-Jacobi Equations," *Transactions of the American Mathematical Society*, Vol. 277, No. 1, 1983, pp. 1–42.
doi:10.2307/1999343

[37] Crandall, M. G., Ishii, H., and Lions, P. L., "User's Guide to Viscosity Solutions of Second Order Partial Differential Equations," *Bulletin of the American Mathematical Society*, Vol. 27, No. 1, July 1992, pp. 1–67.
doi:10.1090/S0273-0979-1992-00266-5

[38] Puterman, M., *Markov Decision Process*, Wiley-Interscience, New York, 1994.