A NUMERICAL APPROACH TO STOCHASTIC OPTIMAL CONTROL VIA DYNAMIC PROGRAMMING

Luis G. Crespo [*] **and Jian Q. Sun** [**]

[*] *National Institute of Aerospace, 144 Research Drive, Hampton VA 23666, USA*
[**] *Department of Mechanical Engineering, University of Delaware, Newark, DE 19711, USA*

Abstract: This paper presents a strategy for finding optimal controls of nonlinear systems subject to random excitations. The method is capable to generate global control solutions when state and control constraints are present. The solution is global in the sense that controls for all initial conditions in a region of the state space are obtained. The approach is based on the Bellman's Principle of optimality, the cumulant neglect closure method and the Short-time Gaussian approximation. Nonlinear problems with non-smooth terms and range control bounds are considered in the examples. The controlled system responses derived are simulated and sucessfully validated by using the Generalized Cell Mapping method.

Keywords: Stochastic control, Non-linearity, Optimality.

1. INTRODUCTION

The optimal control of stochastic systems is a difficult problem, particularly when the system is strongly nonlinear and constraints are present. Given its complexity, we usually resort to numerical methods, Kushner and Dupuis (2001). While some numerical methods of solution to the Hamilton Jacobi Bellman (HJB) equation are known, they usually require knowledge of the boundary/asymptotic behavior of the solution, Bratus et al. (2000). Numerical strategies to find deterministic optimal controls based on the Bellman's principle of optimality (BPO) are available Crespo and Sun (2000). In this paper, these ideas are extended to stochastic optimal control. The method, that involves both analytical and numerical steps, offers several advantages: (i) it can be applied to strongly nonlinear systems, (ii) it takes into account state and control constraints and (iii) it leads to global solutions, from where topological features can be extracted, e.g. switching curves. Former developments can be found in Crespo and Sun (2002) and Crespo and Sun (2003).

2. STOCHASTIC OPTIMAL CONTROL

2.1 Problem Formulation

Consider a system governed by the stochastic differential equation (SDE) in the Stratonovich sense $d\mathbf{x}(t) = \mathbf{m}(\mathbf{x}(t), \mathbf{u}(t))dt + \sigma(\mathbf{x}(t), \mathbf{u}(t))d\mathbf{B}(t)$, where $\mathbf{x}(t) \in \mathbf{R}^n$ is the state vector, $\mathbf{u}(t) \in \mathbf{R}^m$ is the control, $\mathbf{B}(t)$ is a vector of independent unit Wiener processes and the functions $\mathbf{m}(\cdot)$ and $\sigma(\cdot)$ are in general nonlinear functions of their arguments. Itô's calculus leads to

$$d\mathbf{x}(t) = \left(\mathbf{m} + \frac{1}{2}\frac{\partial \sigma}{\partial \mathbf{x}}\sigma^T\right)dt + \sigma d\mathbf{B}(t) \qquad (1)$$

⋆ lgcrespo@nianet.org, Tel.+1 (757) 864 8444

The corresponding Fokker-Planck-Kolmogorov equation (FPK) is given by

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial \mathbf{x}}\left[\rho\left(\mathbf{m} + \frac{1}{2}\frac{\partial \sigma}{\partial \mathbf{x}}\sigma^T\right)\right] + \frac{1}{2}\frac{\partial^2}{\partial \mathbf{x}^2}\left[\rho\sigma\sigma^T\right] \quad (2)$$

where $\rho(\mathbf{x}, t_0|\mathbf{x}_0, t_0)$ is the conditional probability density function (PDF) of the state. Let the cost functional be

$$J(\mathbf{u}, \mathbf{x}_0, t_0, T) = E\left[\phi(\mathbf{x}(T), T) + \int_{t_0}^{T} L(\mathbf{x}, \mathbf{u})dt\right] \quad (3)$$

where $E[\cdot]$ is the expected value operator, $[t_0, T]$ is the time interval of interest, $\phi(\mathbf{x}(T), T)$ is the terminal cost and $L(\mathbf{x}(t), \mathbf{u}(t))$ is the Lagrangian function. The problem formulation is to find the control $\mathbf{u}(t) \in \mathbf{U}$ for $t \in [t_0, T]$ in Equation (1) that drives the system from the initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$ to the target set defined by $\mathbf{\Psi}(\mathbf{x}(T), T) = 0$ such that J is minimized. The fixed final state condition leads to control solutions of the feedback type, i.e. $\mathbf{u}(\mathbf{x})$.

2.2 Bellman's Principle of Optimality

Let $V(\mathbf{x}_0, t_0, T) = J(\mathbf{u}^*, \mathbf{x}_0, t_0, T)$ be the value function or optimal cost function, Yong and Zhou (1999). The BPO is prescribed by $V(\mathbf{x}_0, t_0, T) = \inf_{\mathbf{u}\in\mathbf{U}} E[\beta]$, where

$$\beta = \int_{t_0}^{\hat{t}} L(\mathbf{x}, \mathbf{u})dt + \int_{\hat{t}}^{T} L(\mathbf{x}^*, \mathbf{u}^*)dt + \phi(\mathbf{x}^*(T), T)$$

where $t_0 \leq \hat{t} \leq T$. Consider the problem of finding the optimal control for a system starting from $\mathbf{x}_i$ in the time interval $[i\tau, T]$. Define the incremental and the accumulative costs as

$$J_\tau = E\left[\int_{i\tau}^{(i+1)\tau} L(\mathbf{x}, \mathbf{u})dt\right] \quad (4)$$

$$J_T = E\left[\phi(\mathbf{x}^*(T), T) + \int_{(i+1)\tau}^{T} L(\mathbf{x}^*, \mathbf{u}^*)dt\right] \quad (5)$$

where τ is a discrete time step, $\{\mathbf{x}^*(t), \mathbf{u}^*(t)\}$ is the optimal solution pair over the time interval $[(i+1)\tau, T]$. In this context, the BPO is given by $V(\mathbf{x}_i, i\tau, T) = \inf_{\mathbf{u}\in\mathbf{U}} \{J_\tau + J_T\}$. The incremental cost J_τ is the cost to march one time step forward starting from the deterministic initial condition $\mathbf{x}_i$. The system moves to an intermediate set of the state variables. The accumulative cost J_T is the optimal cost of reaching the target set $\mathbf{\Psi}(\mathbf{x}(T), T) = 0$ starting from this intermediate set and is calculated through the accumulation of incremental costs over time intervals between $(i+1)\tau$ and T, i.e. $V(\mathbf{x}((i+1)\tau), (i+1)\tau, T)$ for the processed state space.

At the intermediate state, the continuity condition $\mathbf{x}((i+1)\tau) = \mathbf{x}^*_{i+1}$ must be imposed in the probabilistic sense. Notice that $\mathbf{x}((i+1)\tau)$ is a random variable. To quantify the continuity condition, let $\mathbf{\Omega}$ be the extended target set such that $\mathbf{x}^*_{i+1} \in \mathbf{\Omega}$. Hence, the condition implies that the support of $\mathbf{x}((i+1)\tau)$ should mostly lie on the processed state space $\mathbf{\Omega}$. Notice however, that $\rho(\mathbf{x}, \tau|\mathbf{x}_0, 0)$ covers the entire state space no matter how small τ is. For a given control, define $P_{\mathbf{\Omega}}$ as

$$P_{\mathbf{\Omega}} = \int_{\mathbf{x}\in\mathbf{\Omega}} \rho(\mathbf{x}, \tau|\mathbf{x}_i, 0)d\mathbf{x} \quad (6)$$

then $P_{\mathbf{\Omega}}$ is the probability of reaching the extended target set $\mathbf{\Omega}$ in time τ starting from $\mathbf{x}_i$. The controlled response $\mathbf{x}(t)$ starting from a set of initial conditions $\mathbf{x}_i$ will become a candidate for the optimal solution when $P_{\mathbf{\Omega}}$ is maximal.

3. SOLUTION APPROACH

To evaluate the expected values in Equation (5), $\rho(\mathbf{x}, \tau|\mathbf{x}_0, 0)$ is needed. For a given feedback control $\mathbf{u} = \mathbf{f}(\mathbf{x})$, $\mathbf{x}(t)$ is a stationary Markov process, Lin and Cai (1995). For a small τ, $\rho(\mathbf{x}, \tau|\mathbf{x}_0, 0)$ is approximately Gaussian within an error of order $O(\tau^2)$, Risken (1984). This is usually refered to as the Short-Time Gaussian Approximation (STGA). The state dynamics can be derived from Equation (1), leading to a set of coupled differential equations. Such set is closed when the dynamics is linear. For nonlinear systems the cumulant neglect closure method (CNC) can be used for closing, Lin and Cai (1995). According to the STGA, the CNC of order two is used here to approximate the conditional density function. Since this approximation is good locally, a partition of the state space is required for accurate results. Therefore, the state space is divided into multiple regions, called cells, whose dynamics is well approximated by the STGA. A CNC of higher order will refine the dynamics of the cell. Such practice will allow for either larger cells or more accurate estimations of the costs. This feature will be useful in the cases where considerably different controls lead to comparable costs.

3.1 Backward Search Algorithm

The backward search algorithm starts from the last segment of the time interval. Since the final state is fixed, a family of optimal solutions for all initial conditions surrounding the target is easily found. The optimal control in the interval $[i\tau, T]$ is determined by minimizing the sum of the incremental and the accumulative cost leading to $V(\mathbf{x}_i, i\tau, T)$ subject to the continuity condition introduced above. The numerical procedure is presented next. Discretize a finite state region $\mathbb{D} \subset \mathbf{R}^n$ into a countable number of parts/cells. Let $\mathbf{U}$

be a set consisting of a countable number of admissible controls $\mathbf{u}_i$ for $i = 1, 2, ..., I$. The control is assumed to be constant over the time intervals. Let $\mathbf{\Omega} \subset \mathbf{R}^n$ denote the discretized version of the target set $\mathbf{\Psi}(\cdot) = 0$ and $J_T = E[\phi(\mathbf{x}(T), T)]$ be the terminal cost. In this framework, the algorithm is as follows

(1) Find all the cells that surround the target set $\mathbf{\Omega}$. Denote the corresponding cell centers $\mathbf{z}_j$.
(2) Construct the conditional probability density function $\rho(\mathbf{x}, \tau|\mathbf{z}_j, 0)$ for each control $\mathbf{u}_i$ and for all cell centers $\mathbf{z}_j$. Call every combination $(\mathbf{z}_j, \mathbf{u}_i)$ a candidate pair.
(3) Calculate the incremental cost $J_\tau(\mathbf{z}_j, \mathbf{u}_i)$, the accumulative cost $J_T(\mathbf{z}_k^*, \mathbf{u}_i^*)$ and $P_\mathbf{\Omega}$ for all candidate pairs, where $\mathbf{z}_k^* \in \Omega$ is an image cell of $\mathbf{z}_j$ and $\mathbf{u}_i^*$ is the optimal control of $\mathbf{z}_k^*$ found in previous iterations.
(4) Search for the candidate pairs that minimize $J_\tau(\mathbf{z}_j, \mathbf{u}_i) + J_T(\mathbf{z}_k^*, \mathbf{u}_i^*)$ and satisfy $P_\mathbf{\Omega} < \Theta max\{P_\mathbf{\Omega}\}$, where $0 \ll \Theta < 1$ is a factor set in advance. Denote such pairs as $(\mathbf{z}_j^*, \mathbf{u}_i^*)$.
(5) Save the minimized accumulative cost function $J_T(\mathbf{z}_j^*, \mathbf{u}_i^*) = J_\tau(\mathbf{z}_j^*, \mathbf{u}_i^*) + J_T(\mathbf{z}_k^*, \mathbf{u}_i^*)$ and the optimal pairs $(\mathbf{z}_k^*, \mathbf{u}_i^*)$.
(6) Expand the target set $\mathbf{\Omega}$ by including the cells $\mathbf{z}_j^*$.
(7) Repeat the search from Step (1) to Step (6) until the initial condition $\mathbf{x}_0$ is reached.

As a result, the optimal control for all the cells in Ω is found. The choice of image cells, i.e. $\mathbf{x}((i+1)\tau)$, could certainly by biased. This however, is avoided by using (i) non-uniform integration times such that the growth of Ω is gradual, i.e. mapping most of the probability to neighboring cells, and (ii) by restricting the potential optimal pairs to be candidate pairs with high $P_\mathbf{\Omega}$. These considerations led to consistent global control solutions regardless of the cell size, Crespo and Sun (2000). The resulting controlled dynamics of the conditional PDF is simulated using the Generalized Cell Mapping Method (GCM), Crespo and Sun (2002).

4. EXAMPLES

4.1 Non-linear Oscillator with Dry Friction

Consider the non-linear system

$$\ddot{x} + \mu(g + \ddot{v})\text{sgn}(\dot{x}) + 2\zeta\dot{x} + \omega_0^2 x + \varepsilon x^3 = \ddot{f} + u(t) \quad (7)$$

where $x(t)$ is the horizontal sliding motion of a mass block placed on a moving foundation with rough contact surface and $u(t)$ is a force satisfying $|u| \leq \hat{u} = 1$, Sun (1995). ζ is the viscous damping coefficient, μ is the dry friction damping coefficient, g is the gravitational acceleration, ω_o is the natural frequency of the linear system and ε is the non-linear stiffness coefficient. Assume that $\ddot{f}$ and $\ddot{v}$ satisfy $E[\ddot{f}] = 0$, $E[\ddot{v}] = 0$, $E[\ddot{v}(t)\ddot{f}(t')] = 2D_{vf}\delta(t-t')$, $E[\ddot{v}(t)\ddot{v}(t')] = 2D_v\delta(t-t')$ and $E[\ddot{f}(t)\ddot{f}(t')] = 2D_f\delta(t-t')$. Let $x_1 = x$, $x_2 = \dot{x}$ and the Lagrangian be $L = \alpha x_1^2 + \beta x_2^2 + \gamma u^2$ The corresponding SDE in the Stratonovich sense is given by

$$\begin{aligned} dx_1 &= x_2 dt \qquad (8)\\ dx_2 &= (-\mu g \text{sgn}(x_2) - 2\zeta x_2 - \omega_0^2 x_1 - \varepsilon x_1^3 + u)dt \\ &\quad - \mu \text{sgn}(x_2) dB_1 + dB_2 \end{aligned}$$

where B_1 and B_2 are dependent delta correlated Gaussian white noises with zero mean. Following the rules of the Itô calculus, we convert Equation (8) into a set of SDE in the Itô sense

$$\begin{aligned} dx_1 &= x_2 dt \qquad (9)\\ dx_2 &= [-\mu g \text{sgn}(x_2) - 2\zeta x_2 - \omega_0^2 x_1 - \varepsilon x_1^3 + \\ &\quad \mu^2 D_v \text{sgn}(x_2)\text{sgn}'(x_2) + \mu D_{vf}\text{sgn}'(x_2) + u]dt + \\ &\quad 2[\mu^2 D_v \text{sgn}^2(x_2) + 2\mu D_{vf}\text{sgn}(x_2) + D_f]^{1/2} dW \end{aligned}$$

where $W(t)$ is a unit Wiener process satisfying $E[W(t)] = 0$, $E[W(t)W(t')] = t - t'$ where $t > t'$. An infinite hierarchy of moment equations for the state variables can be derived from the Itô Equation applying the expected value operator and using independence. Defining $m_{nm} = E[x_1^n x_2^m]$, and using the analytical expressions for the expected values, we obtain the differential equations for the first two order moments

$$\begin{aligned} \dot{m}_{10} &= m_{01} \\ \dot{m}_{01} &= \mu g \text{sgn}(m_{01}) \,\text{erf}(|m_{01}|/\sqrt{2}\sigma_2) - 2\zeta m_{01} - \\ &\quad \omega_0^2 m_{10} - \varepsilon m_{10}(3\sigma_1^2 + m_{10}^2) + \\ &\quad \mu D_{vf}(2/\sqrt{2\pi}\sigma_2)\exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) + u, \\ \dot{m}_{11} &= \sigma_2^2 - 2\zeta c_{12} - \omega_0^2\sigma_1^2 - 3\varepsilon\sigma_1^2(\sigma_1^2 + m_{10}^2) - \\ &\quad \mu g\sqrt{2/\pi}(c_{12}/\sigma_2)\exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) - \\ &\quad \mu D_{vf}(c_{12}m_{01}/\sqrt{2/\pi}\sigma_2^3)\exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) + \\ &\quad m_{10}\dot{m}_{01} + \dot{m}_{10}m_{01} \\ \dot{m}_{20} &= 2c_{12} + 2m_{10}\dot{m}_{10} \\ \dot{m}_{02} &= -4\zeta\sigma_2^2 - 2\omega_0^2 c_{12} - 6\varepsilon c_{12}(\sigma_1^2 + m_{10}^2) - \\ &\quad 2\mu g\sqrt{2/\pi}\sigma_2\exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) - \\ &\quad \mu D_{vf}\sqrt{2/\pi}(m_{01}/\sigma_2)\exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) + \\ &\quad 2\mu^2 D_v + 2D_f + 2m_{01}\dot{m}_{01} + \\ &\quad 4\mu D_{vf}\text{sgn}(m_{01})\,\text{erf}(|m_{01}|/\sqrt{2}\sigma_2) \qquad (10) \end{aligned}$$

where $c_{12} = m_{11} - m_{01}m_{01}$ is the covariance of x_1 and x_2, $\sigma_1^2 = m_{20} - m_{10}^2$ is the variance of x_1 and $\sigma_2^2 = m_{02} - m_{01}^2$ is the variance of x_2. The initial conditions required to integrate Equations (10) from $t = 0$ to $t = \tau$ are specified by the coordinates of a cell center (x_1, x_2), i.e. $m_{10}(0) = x_1$, $m_{01}(0) = x_2$, $m_{20}(0) = 0$, $m_{02}(0) = 0$ and $m_{11}(0) = 0$. The joint probability density of the response and the corresponding costs can be then readily calculated.

Notice that the system is parametrically and externally excited and the diffusion term is state dependent. The region defined by $x_1 \in [-2, 2]$ and $x_2 \in [-2, 2]$ is discretized with $25 \times 25 = 625$ uniform cells. The parameters of the system are set as follows: $\mu = 0.05$, $\zeta = 0.1$, $\omega_0 = 1$, $\varepsilon = 1$, $D_v = 0.1$, $D_f = 0.1$, and $D_{vf} = 0$. The Lagrangian of the cost function is evaluated using $\alpha = \beta = \gamma = 0.5$ and the control set is uniformly discretized into 11 levels $u \in \{-1, -0.8, \cdots, 1\}$.

The vector field of the mean of the uncontrolled response is shown in Figure 1. There is a region on the x_1-axis close to the origin, where the mean trajectories get trapped. Such strip is highlighted in the figure with a thick line. When the term $\mu g \mathrm{sgn}(m_{01}) \,\mathrm{erf}(|m_{01}|/\sqrt{2}\sigma_2)$ in Equation (10) becomes dominant, a never ending sequence of changes in the sign of the velocity takes place. This phenomenon forces the response to switch indefinitely about the strip without having a net displacement. Starting from a uniformly distributed initial condition in $\mathbb{D} = (x_1, x_2) \in [-2, 2] \times [-2, 2]$, the time evolution of the probability density of the response is calculated. It was found that 67% of the response stays in $\mathbb{D}$, converging to th PDF in Figure 2. Stationarity is reached after 10 time units.

Figure 3 shows the vector field of the mean trajectory for the controlled response. The size of the arrows about the origin is enlarged to enable better observation. Abrupt changes in the velocity are still present about the x_1 axis are present, but the trapping strip is not. These jumps show that the velocity just before and just after reaching maximum elongation of the spring differ in magnitude. Starting from the same uniformly distributed initial condition, the time evolution of the controlled system response was studied. Figure 4 shows the corresponding stationary density function reached after 7 time units. It was found that 80% of the response stays in the admissible domain.

In order to evaluate the control performance in a global basis, the leakage of the probability to outside of the domain $\mathbb{D}$ must be taken into consideration. For PDFs which are not fully contained in $\mathbb{D}$ the calculation of the expected cost in the cellular domain is insufficient due to the use of a single cell, i.e. the sink cell, to represent $\bar{\mathbb{D}}$. With this in mind, we propose the following measure to evaluate the costs:

$$J_d(t') = \frac{E\left[L(\mathbf{x}(\mathbf{u}(t')), \mathbf{u}(t'))\right]_d}{P_d(t')} \tag{11}$$

where $E[\cdot]_d$ and P_d are the expected value operator and the probability of the response on the

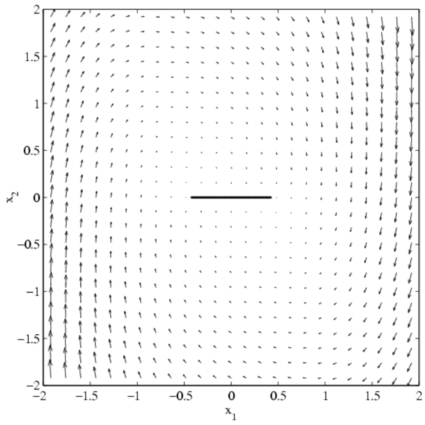

Fig. 1. Vector field of the mean trajectories of the uncontrolled response.

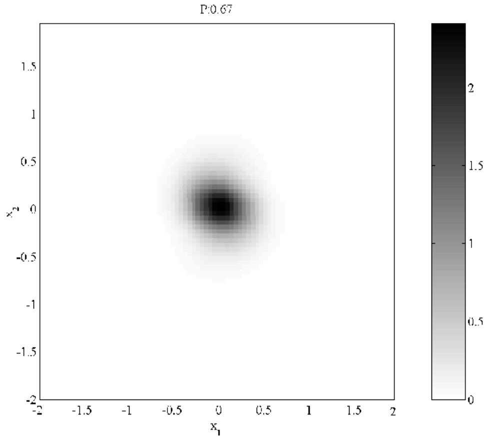

Fig. 2. Upper view of the Stationary PDF for the uncontrolled response.

domain $\mathbb{D}$ at $t = t'$, given a uniform probability distribution at $t = 0$. For cases in which the entire PDF remains within $\mathbb{D}$ at all times, the total cost in Equation (3) can be easily calculated by integrating J_d from $t' = t_0$ to $t' = T$. Figure 5 shows the time evolutions for J_d and P_d for both the uncontrolled and controlled system responses.

By comparing with the uncontrolled response, we conclude that (i) the controlled response reaches the target with minimum cost, (ii) the stationary PDF is closer to the desired target set, (iii) the convergence to the stationary PDF is faster and (iv) a higher percentage of the probability is kept in the admissible domain. Because the system starts from a uniformly distributed PDF, these results validate the efficacy of the global control solution and of the methodology.

4.2 Range Bounded Control for the Van der Pol Oscillator

Consider the system

$$\ddot{x} + \theta(x^2 - 1)\dot{x} + \Omega^2 x = u(t) + w(t) \tag{12}$$

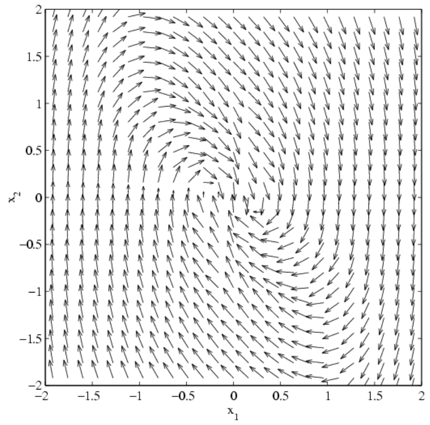

Fig. 3. Vector field of the mean trajectories of the controlled response.

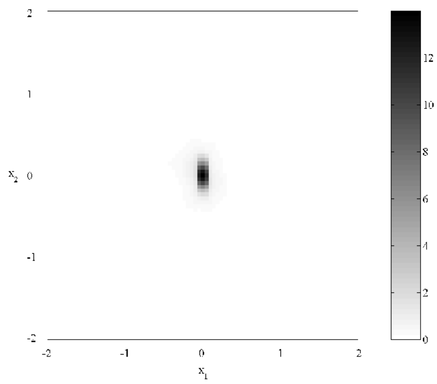

Fig. 4. Upper view of the stationary PDF for the controlled response.

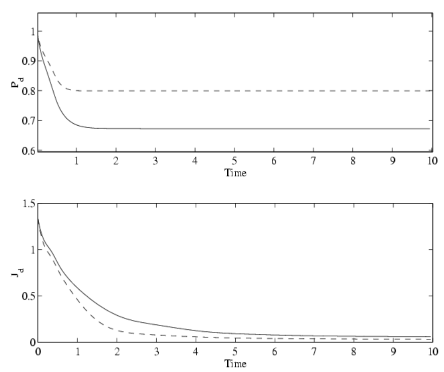

Fig. 5. Time evolution of the normalized cost J_p and Pd for the uncontrolled (—) and controlled (−−) responses.

where $w(t)$ is a Gaussian white noise process satisfying $E[w(t)] = 0$, $E[w(t)w(t+t')] = 2D\delta(t')$ and $u(t)$ is bounded with $|u| \leq \hat{u}$. The same cost function as in the previous example is considered herein. Let $x_1 = x$ and $x_2 = \dot{x}$. The SDE in the Stratonovich sense for the system is given by

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= (-\theta(x_1^2 - 1)x_2 - \Omega^2 x_1 + u)dt + dW \end{aligned}$$

where the drift and diffusion terms are given by the vector $\mathbf{m} = [x_2, -\theta(x^2-1)\dot{x} - \Omega^2 x + u]^T$ and $\sigma = [0, 1]^T$ respectively. In this case, the Wong-Zakai correction term is zero. The corresponding SDE in the Itô sense is given by

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= (-\theta(x_1^2 - 1)x_2 - \Omega^2 x_1 + u)dt + (2D)^{1/2} dW \end{aligned}$$

where $W(t)$ is a unit Weiner process satisfying $E[W(t)] = 0$ and $E[W(t)W(t+t')] = t'$. The moment equations of the state variables are

$$\begin{aligned} \dot{m}_{10} &= m_{01} \\ \dot{m}_{01} &= \theta(m_{01} - m_{21}) - \Omega^2 m_{10} + u \\ \dot{m}_{20} &= 2m_{11} \\ \dot{m}_{02} &= 2\theta(m_{02} - m_{22}) - 2\Omega^2 m_{11} + 2u m_{01} + 2D \\ \dot{m}_{11} &= \theta(m_{11} - m_{31}) - \Omega^2 m_{20} + u m_{10} + m_{02} \end{aligned} \tag{13}$$

The higher order moments are approximated using the Gaussian closure as explained. The parameters of the system are set as $\theta = 1$, $\Omega = 1$, $\alpha = 0.5$, $\beta = 0.5$, $\gamma = 0$ and $D = 0.2$. The region defined by $x_1 \in [-4, 4]$ and $x_2 \in [-4, 4]$ is discretized with $25 \times 25 = 625$ uniform cells. When $\gamma = 0$, the control is found to be bang-bang. Thus, the control set is bi-level $u \in \{-1, 1\}$. The steady state PDF with the limit cycle of the uncontrolled system is shown in Figure 6. At the peak locations of the PDF, the response moves slower along the limit cycle.

We now study the effect of the bound $\hat{u}$ in the optimal control. Several problems, satisfying $\hat{u} \leq 1$ were considered. Time evolutions of the moments and $E[L]$ taking the distribution of Figure 6 as initial condition and $\hat{u} = 0.5$ are shown in Figure 7. The corresponding steady state PDF of the controlled response is shown in Figure 8. Notice that the control is unable to break the limit cycle behavior. However, the resulting steady-state cycle is moved closer to the target than in the uncontrolled case. More importantly, the control starts building up a 'bridge' between the two regions with higher probability i.e. where the dynamics is slower, moving probability out of the limit cycle and placing it at the origin. This mechanism causes the formation of a third peak in the location of the target. For higher control bounds this third peak becomes more dominant. Results for $\hat{u} = 1$ are shown in Figure 9. The corresponding steady state PDF is shown in Figure 10. For this bound, the control solution has completely eliminated the limit cycle. This leads to a stationary PDF with means very close to the target and small covariances. However, the traces of the 'bridge' are still present.

5. CONCLUSIONS

This paper applies the methodology proposed in Crespo and Sun (2003) to study the control of

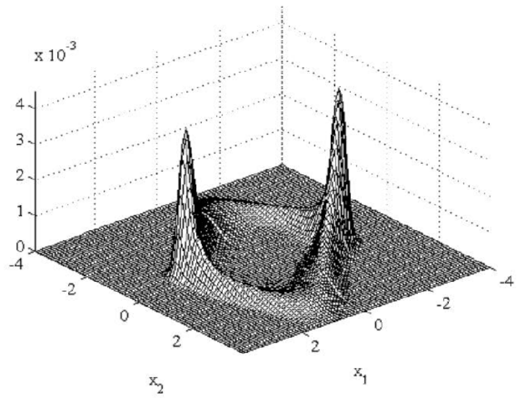

Fig. 6. Steady state PDF of the uncontrolled Van der Pol oscillator

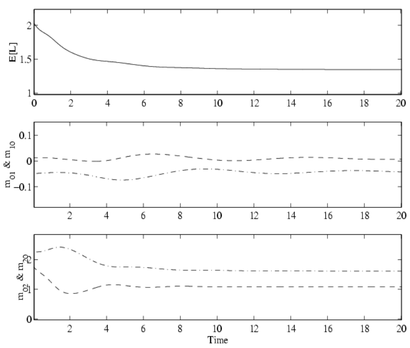

Fig. 7. Time evolution of the controlled response for $\hat{u} = 0.5$

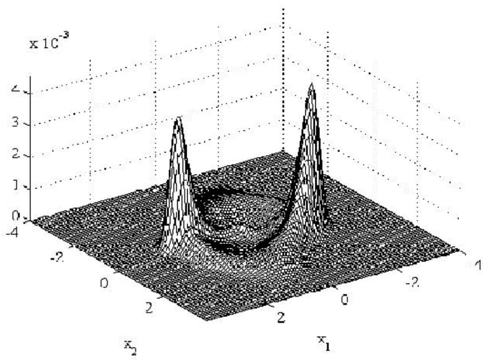

Fig. 8. Stationary PDF of the controlled response for $\hat{u} = 0.5$

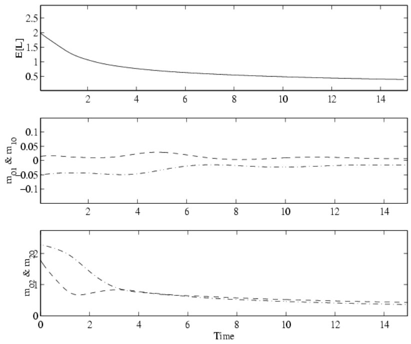

Fig. 9. Time evolution of the controlled response for $\hat{u} = 1$

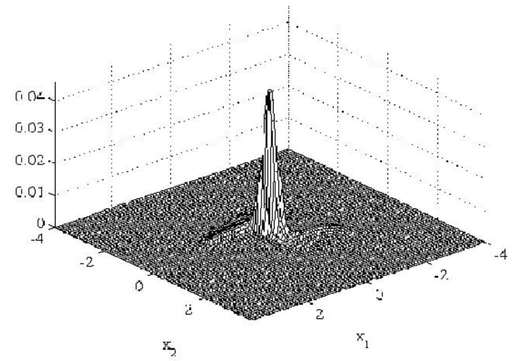

Fig. 10. Steady state PDF for $\hat{u} = 1$

a non-smooth non-linear system with parametric and external excitations and the effects of control bounds on the optimal stabilization of the Van der Pol oscillator. Even though the computational demands restrict the methodology to low state dimensional systems, the solutions provided are very difficult to obtain otherwise.

REFERENCES

A. Bratus, M. Dimentberg, and D. Iourtchenko. Optimal bounded response control for a second-order system under a white-noise excitation. *Journal of Vibration and Control*, 6:741–755, 2000.

L. G. Crespo and J. Q. Sun. Solution of fixed final state optimal control problems via simple cell mapping. *Nonlinear Dynamics*, 23:391–403, 2000.

L. G. Crespo and J. Q. Sun. Stochastic optimal control of non-linear systems via short-time Gaussian approximation and cell mapping. *Nonlinear Dynamics*, 28:323–342, 2002.

L. G. Crespo and J. Q. Sun. Stochastic optimal control via Bellman's principle. *Automatica*, 2003.

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

Y. K. Lin and G. Q. Cai. *Probabilistic Structural Dynamics: Advanced Theory and Applications*. McGraw-Hill, New York, 1995.

H. Risken. *The Fokker-Planck Equation - Methods of Solution and Applications*. Springer-Verlag, New York, 1984.

J. Q. Sun. Random vibration analysis of a non-linear system with dry friction damping by the short-time Gaussian cell mapping method. *Journal of Sound and Vibration*, 180(5):785–795, 1995.

J. Yong and X. Y. Zhou. *Stochastic Controls, Hamiltonian Systems and HJB Equations*. Springer-Verlag, New York, 1999.