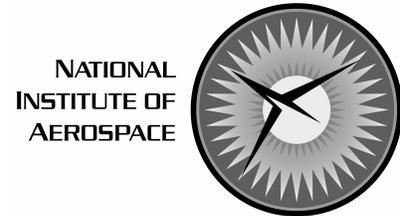
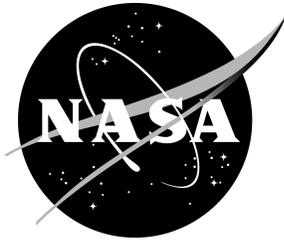


NASA/CR-2003-212419  
NIA Report No. 2003-04



## Stochastic Optimal Control via Bellman's Principle

*Luis G. Crespo*  
*National Institute of Aerospace, Hampton, Virginia*

*Jian Q. Sun*  
*University of Delaware, Newark, Delaware*

## The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

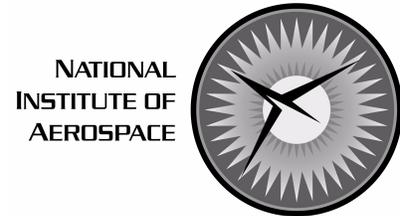
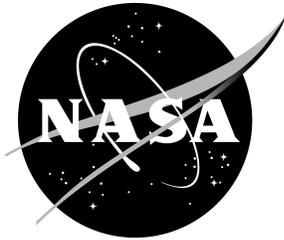
- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.
- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at <http://www.sti.nasa.gov>
- Email your question via the Internet to [help@sti.nasa.gov](mailto:help@sti.nasa.gov)
- Fax your question to the NASA STI Help Desk at (301) 621-0134
- Telephone the NASA STI Help Desk at (301) 621-0390
- Write to:  
NASA STI Help Desk  
NASA Center for Aerospace Information  
7121 Standard Drive  
Hanover, MD 21076-1320

NASA/CR-2003-212419  
NIA Report No. 2003-04



## **Stochastic Optimal Control via Bellman's Principle**

*Luis G. Crespo*

*National Institute of Aerospace, Hampton, Virginia*

*Jian Q. Sun*

*University of Delaware, Newark, Delaware*

National Aeronautics and  
Space Administration

Langley Research Center  
Hampton, Virginia 23681-2199

Prepared for Langley Research Center  
under Contract NCC-1-02043

---

December 2003

---

Available from the following:

NASA Center for AeroSpace Information (CASI)  
7121 Standard Drive  
Hanover, MD 21076-1320  
(301) 621-0390

National Technical Information Service (NTIS)  
5285 Port Royal Road  
Springfield, VA 22161-2171  
(703) 487-4650

# STOCHASTIC OPTIMAL CONTROL VIA BELLMAN'S PRINCIPLE

Luis G. Crespo\* and Jian Q. Sun†

## ABSTRACT

This paper presents a method 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 Bellman's Principle of optimality, the Gaussian closure and the Short-time Gaussian approximation. Examples include a system with a state-dependent diffusion term, a system in which the infinite hierarchy of moment equations cannot be analytically closed, and an impact system with a elastic boundary. The uncontrolled and controlled dynamics are studied by creating a Markov chain with a control dependent transition probability matrix via the Generalized Cell Mapping method. In this fashion, both the transient and stationary controlled responses are evaluated. The results show excellent control performances.

**keywords:** Stochastic optimal control, Bellman's principle, Cell mapping, Gaussian closure.

## 1 INTRODUCTION

Optimal control of stochastic nonlinear dynamic systems is an active area of research due to its relevance to many engineering applications. This is a very difficult problem to study, particularly when the system is strongly nonlinear and there are constraints on the states and the control. Optimal feedback controls for systems under white-noise random excitations may be studied by the Pontryagin maximum principle, Bellman's principle of optimality and the Hamilton-Jacobi-Bellman (HJB) equation. When the control and the state are bounded, the direct solution of the HJB equation faces exigent difficulties since it is multidimensional, nonlinear and defined in a domain that in general might not be simply connected. The theory of viscosity solutions, first introduced by Crandall *et. al.* [21], provides a convenient framework for studying the HJB equation. Very few closed form solutions to this problem have been found so far. Dimentberg *et al.* found the analytical solutions of the optimal control of a linear spring mass oscillator with Lagrange and Mayer cost functionals in a region of the phase space [2, 8].

Given the intrinsic complexity of the problem, we must resort to numerical methods to find approximate control solutions [1, 11]. While some numerical methods of solution to the HJB equation are known, they usually require knowledge of the boundary/asymptotic behavior of the solution in advance [2]. Relatively few problems are known today which are solved through the use of the HJB equation. Zhu *et al.* proposed a strategy for optimal

---

\*National Institute of Aerospace, 144 Research Drive, Hampton, VA 23669, USA, email: [lgcrespo@nianet.org](mailto:lgcrespo@nianet.org)

†Department of Mechanical Engineering, University of Delaware, Newark, DE 19711, USA, email: [sun@me.udel.edu](mailto:sun@me.udel.edu)

feedback control of randomly excited structural systems based on the stochastic averaging method for quasi-Hamiltonian systems and the HJB equation [22].

The cell mapping methods, originally developed by Hsu, have been applied to the optimal control problem of deterministic systems [9]. Strategies to solve optimal control problems with fixed final time [4] and fixed final state [5] terminal conditions have been developed and applied. The Generalized Cell Mapping (GCM) method is a very effective tool for studying the global behavior of strongly nonlinear stochastic system. The GCM method is integrated with the Short-time Gaussian Approximation (STGA) scheme to provide a very efficient way for constructing Markov chains that describe the global dynamics of the system [10, 16, 18]. In this paper, these tools are extended to the stochastic control problem. Specifically, we use Bellman's Principle of optimality and the STGA to generate global control solutions to stochastic control problems with fixed-state terminal conditions and state and/or control constraints.

The current method, that involves both analytical and numerical steps, offers several advantages. It can handle strongly nonlinear and non-smooth systems, can include various state and control constraints and leads to global solutions. Former developments on the subject can be found in [3], where numerical and analytical solutions were successfully compared. In this paper several extensions to that work are proposed and evaluated. The content of this paper is organized as follows. Section 2 briefly introduces the problem formulation and the Bellman's principle. The analytical and numerical parts of the solution method are presented in Section 3. Examples are provided in Section 4, where the following systems are controlled: a nonlinear system with a state-dependent diffusion part, a system with non-analytically closeable terms and an impact system with a reflective boundary condition. Some conclusions are stated in Section 5.

## 2 STOCHASTIC OPTIMAL CONTROL

### 2.1 Problem Formulation

Consider a system governed by stochastic differential equation (SDE) in the Stratonovich sense:

$$d\mathbf{x}(t) = \mathbf{m}(\mathbf{x}(t), \mathbf{u}(t))dt + \sigma(\mathbf{x}(t), \mathbf{u}(t))d\mathbf{B}(t) \quad (1)$$

where  $\mathbf{x}(t) \in \mathbf{R}^n$  is the state vector,  $\mathbf{u}(t) \in \mathbf{R}^m$  is the control input,  $\mathbf{B}(t)$  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. By using Itô's formula [14], we obtain the corresponding Itô SDE:

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

The associated Fokker-Planck-Kolmogorov (FPK) equation for the conditional probability density function (PDF)  $\rho(\mathbf{x}, t_0 | \mathbf{x}_0, t_0)$  is given by:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial \mathbf{x}} \left[ \rho \left( \mathbf{m}(\mathbf{x}, \mathbf{u}) + \frac{1}{2} \frac{\partial \sigma(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \sigma(\mathbf{x}, \mathbf{u})^T \right) \right] + \frac{1}{2} \frac{\partial^2}{\partial \mathbf{x}^2} [\rho \sigma(\mathbf{x}, \mathbf{u}) \sigma(\mathbf{x}, \mathbf{u})^T] \quad (3)$$

Define the cost functional as:

$$J(\mathbf{u}, \mathbf{x}_0, t_0, T) = E \left[ \phi(\mathbf{x}(T), T) + \int_{t_0}^T L(\mathbf{x}(t), \mathbf{u}(t))dt \right] \quad (4)$$

where  $[t_0, T]$  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 optimal control problem is to find a control  $\mathbf{u}(t)$  within a set  $\mathbf{U} \subset \mathbf{R}^m$  on the time interval  $[t_0, T]$  that drives the system from a given initial condition  $\mathbf{x}(t_0) = \mathbf{x}_0$  to the target set defined by  $\Psi(\mathbf{x}(T), T) = 0$  such that the cost  $J$  is minimized.

## 2.2 Bellman's Principle of optimality

Let  $(\mathbf{x}^*, \mathbf{u}^*)$  be an optimal control solution pair over the time interval  $[t_0, T]$  subject to the initial condition  $\mathbf{x}(t_0) = \mathbf{x}_0$ . Let  $\hat{t}$  be a time instant such that  $t_0 \leq \hat{t} \leq T$ . Then,  $(\mathbf{x}^*, \mathbf{u}^*)$  is still the optimal control solution pair from  $[\hat{t}, T]$  subject to the initial condition  $\mathbf{x}(\hat{t}) = \mathbf{x}^*(\hat{t})$ .

Let  $V(\mathbf{x}_0, t_0, T) = J(\mathbf{u}^*, \mathbf{x}_0, t_0, T)$  be the so-called value function or optimal cost function. Bellman's principle of optimality can be stated as [21]:

$$V(\mathbf{x}_0, t_0, T) = \inf_{\mathbf{u} \in \mathbf{U}} E \left[ \int_{t_0}^{\hat{t}} L(\mathbf{x}(t), \mathbf{u}(t)) dt + \int_{\hat{t}}^T L(\mathbf{x}^*(t), \mathbf{u}^*(t)) dt + \phi(\mathbf{x}^*(T), T) \right] \quad (5)$$

where  $t_0 \leq \hat{t} \leq T$ . Consider the optimal problem of the system starting from  $\mathbf{x}_i$  in the time interval  $[i\tau, T]$  where  $\tau$  is a discrete time step. Define an incremental cost and an accumulative cost as:

$$J_\tau = E \left[ \int_{i\tau}^{(i+1)\tau} L(\mathbf{x}(t), \mathbf{u}(t)) dt \right] \quad (6)$$

$$J_T = E \left[ \phi(\mathbf{x}^*(T), T) + \int_{(i+1)\tau}^T L(\mathbf{x}^*(t), \mathbf{u}^*(t)) dt \right] \quad (7)$$

where  $(\mathbf{x}^*(t), \mathbf{u}^*(t))$  is the optimal solution pair over the time interval  $[(i+1)\tau, T]$ . Then, Bellman's principle of optimality can be restated as:

$$V(\mathbf{x}_i, i\tau, T) = \inf_{\mathbf{u} \in \mathbf{U}} \{J_\tau + J_T\} \quad (8)$$

The incremental cost  $J_\tau$  is the cost for the system to march one time step forward starting from a deterministic initial condition  $\mathbf{x}_i$ . The system lands on an intermediate set of the state variables. The accumulative cost  $J_T$  is the cost for the system to reach the target set starting from this intermediate set, and is calculated through the accumulation of incremental costs over several short time intervals between  $(i+1)\tau$  and  $T$ .

Bellman's principle of optimality as stated in Equation (8) suggests that one can obtain a local solution of the optimal control problem over a short time interval  $\tau$  to form the global solution provided that certain continuity conditions on the solution are satisfied. In this paper, we shall impose such conditions in the probability sense. This is explained later in the paper. The global solution consists of all the intermediate solutions that are constructed backward in time starting from the terminal condition  $\phi(\mathbf{x}(T), T)$  at time  $T$ .

## 3 THE SOLUTION APPROACH

### 3.1 Background

We need to evaluate the expected values in Equations (6) and (7) for a given control. For this we need the conditional probability density function  $\rho(\mathbf{x}, \tau | \mathbf{x}_0, 0)$  which satisfies the

FPK equation associated with the process. For a given feedback control law  $\mathbf{u} = \mathbf{f}(\mathbf{x}(t))$ , the response  $\mathbf{x}(t)$  is a stationary Markov process [12]. For a very small  $\tau$ , this density function  $\rho(\mathbf{x}, \tau | \mathbf{x}_0, 0)$  is known to be approximately Gaussian within an error of order  $O(\tau^2)$  [14].

When  $\rho(\mathbf{x}, \tau | \mathbf{x}_0, 0)$  is Gaussian, only the first and second order moments of  $\mathbf{x}(t)$  need to be evaluated in order to completely specify the PDF. We can readily derive the moment equations from the Itô equation (2). When the system is nonlinear, the moment equations can be closed by applying the Gaussian closure method [12, 17], which is consistent with the short-time Gaussian approximation.

The short-time conditional probability density function  $\rho(\mathbf{x}, \tau | \mathbf{x}_0, 0)$  gives a probabilistic description of the system response  $\mathbf{x}(t)$ . The short-time solution has been implemented in two ways in the literature. The first one is the path integral [13, 19, 20], which treats the phase space as a continuum. The second approach is the cell mapping method, which discretizes the phase space into small regions, called cells. Since the path integral often has to be evaluated numerically, discretization of the phase space is inevitable.

### 3.2 The backward search algorithm

The backward solution process starts from the last segment over the time interval  $[T - \tau, T]$ . Since the terminal condition for the last segment of the local solutions is specified, we can obtain a family of local optimal solutions for all possible initial conditions  $\mathbf{x}(T - \tau)$ . Then, repeat this process to obtain the next segment of the local optimal solution over the time interval  $[T - 2\tau, T]$  subject to the continuity condition at  $t = T - \tau$ .

In general, the optimal control in the time interval  $[i\tau, T]$  is determined by minimizing the sum of the incremental cost and the accumulative cost in Equation (8) leading to  $V(\mathbf{x}_i, i\tau, T)$  subject to the continuity condition  $\mathbf{x}((i+1)\tau) = \mathbf{x}_{i+1}^*$  where  $\mathbf{x}_{i+1}^*$  is a set of initial conditions used in the problem with time interval  $[(i+1)\tau, T]$ .  $\mathbf{x}_{i+1}^*$  is a random variable. The equality  $\mathbf{x}((i+1)\tau) = \mathbf{x}_{i+1}^*$  has to be interpreted in a probabilistic sense. In this paper, it is interpreted in the sense of maximum probability. Quantitatively, this is done as follows.

In theory, the conditional PDF  $\rho(\mathbf{x}, \tau | \mathbf{x}_0, 0)$  of a diffusion process even for a very short time  $\tau$  will cover the entire phase space. Let  $\Omega$  be the extended target set to be defined later such that  $\mathbf{x}_{i+1}^* \in \Omega$ . For a given control, we define  $P_\Omega$  as:

$$P_\Omega = \int_{\mathbf{x} \in \Omega} \rho(\mathbf{x}, \tau | \mathbf{x}_i, 0) d\mathbf{x} \quad (9)$$

$P_\Omega$  represents the probability that the system enters the extended target set  $\Omega$  in time  $\tau$  when it starts at  $\mathbf{x}_i$  with probability one. 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  $P_\Omega$  is maximal among all the initial conditions under consideration.

Since we shall use a cellular structure of the phase space, we consider a finite region  $\mathbb{D} \subset \mathbf{R}^n$  and discretize  $\mathbb{D}$  into a countable number of small cells. Let  $\mathbf{U}$  denote a countable set consisting of all the admissible controls  $\mathbf{u}_i$  ( $i = 1, 2, \dots, I$ ). We shall assume that the control is constant over each time interval  $[i\tau, (i+1)\tau]$ . Let  $\Omega \subset \mathbf{R}^n$  denote the set of cells that form the target set defined by  $\Psi(\mathbf{x}(T), T) = 0$ . As the backward solution steps proceed, the set  $\Omega$  will be expanded. Assume that the terminal cost  $J_T$  is  $E[\phi(\mathbf{x}(T), T)]$ . The backward algorithm for searching the optimal control solutions is stated as follows:

1. Find all the cells that surround the target set  $\Omega$ . Denote the corresponding cell centers as  $\mathbf{z}_j$ .
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  $\mathbf{z}_j$ . Let us 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  $P_\Omega$  for all candidate pairs.  $\mathbf{z}_k^*$  is an image cell of  $\mathbf{z}_j$  in  $\Omega$ , and  $\mathbf{u}_i^*$  is the optimal control identified for  $\mathbf{z}_k^*$  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  $P_\Omega < \Theta \max\{P_\Omega\}$ , where  $0 < \Theta < 1$  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  $\Omega$  by including the cells  $\mathbf{z}_j^*$ .
7. Repeat the search from Step (1) to Step (6) until all the cells in the entire region  $\mathbb{D}$  are processed or until the original initial condition  $\mathbf{x}_0$  of the optimal control problem is reached.

As a result, the optimal control solution for all the cells covered by  $\Omega$  is found. The choice of image cells, i.e.  $\mathbf{x}_{i+1}^*$ , could certainly be biased. This however, is avoided by using (i) non-uniform integration times such that the growth of  $\Omega$  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  $P_\Omega$ . These considerations led to the same global control solution regardless of the cell size [5].

The resulting dynamics of the conditional PDF is simulated using the Generalized Cell Mapping Method (GCM) [6]. Notice that if a simulation is done for a long time all probability will eventually leave  $\mathbb{D}$ . This is a consequence of using a finite computational domain to model a diffusion process, therefore all numerical simulations face this problem. While some probability is leaking out from  $\mathbb{D}$ , probability mass is also being brought back from its complement. In this study the computational domain is set such that the leakage of probability during the transient controlled response is very small.

## 4 EXAMPLES

### 4.1 State dependent diffusion terms

Consider the dynamic system:

$$\ddot{x} + \mu(g + \ddot{v})\text{sgn}(\dot{x}) + 2\zeta\dot{x} + \omega_o^2x + \varepsilon x^3 = \ddot{f} + u(t), \quad (10)$$

where  $x(t)$  describes the horizontal sliding motion of a mass block placed on a moving foundation with rough contact surface and  $u(t)$  is a bounded force satisfying  $|u| \leq \hat{u} = 1$  [15, 16].  $\zeta$  is the viscous damping coefficient,  $\mu$  is the dry friction damping coefficient,  $g$  is the gravitational acceleration,  $\omega_o$  is the natural frequency of the linear system and  $\varepsilon$  is

the nonlinear stiffness coefficient. Assume that  $\ddot{f}$  and  $\ddot{v}$  are correlated Gaussian white noise processes such that:

$$\begin{aligned} E[\ddot{f}] &= 0, \quad E[\ddot{v}] = 0, \quad E[\ddot{v}(t)\ddot{f}(t')] = 2D_{vf}\delta(t-t'), \\ E[\ddot{v}(t)\ddot{v}(t')] &= 2D_v\delta(t-t'), \quad E[\ddot{f}(t)\ddot{f}(t')] = 2D_f\delta(t-t'). \end{aligned}$$

Let  $x_1 = x$  and  $x_2 = \dot{x}$ . The cost functional to be minimized is defined as:

$$J = E \left[ \int_{t_0}^{\infty} (\alpha |\mathbf{x}|^2 + \beta u^2) dt \right]. \quad (11)$$

The control objective is to drive the system from any arbitrary initial condition to the origin of the phase space while  $J$  is minimized. Following the rules of the Itô calculus, we convert Equation (10) into a set of stochastic differential equations (SDEs) in the Itô sense:

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= [-\mu g \text{sgn}(x_2) - 2\zeta x_2 - \omega_o^2 x_1 - \varepsilon x_1^3 + \mu^2 D_v \text{sgn}(x_2) \text{sgn}'(x_2) + \mu D_{vf} \text{sgn}'(x_2) + u] dt \\ &\quad + [2\mu^2 D_v \text{sgn}^2(x_2) + 4\mu D_{vf} \text{sgn}(x_2) + 2D_f]^{1/2} dB(t), \end{aligned} \quad (12)$$

where  $B(t)$  is a unit Wiener process satisfying  $E[B(t)] = 0$ ,  $E[B(t)B(t')] = t - t'$  where  $t > t'$ . Details of the derivation of Equation (12) are given in Appendix A. An infinite hierarchy of moment equations for the state variables can be derived from the Itô equation. Defining  $m_{ij} = E[x_1^i x_2^j]$ , and applying the Gaussian closure method by using the expressions in Appendix C, we obtain a set of nonlinear differential equations for the first and second order moments:

$$\begin{aligned} \dot{m}_{10} &= m_{01} \quad \dot{m}_{20} = 2c_{12} + 2m_{10}\dot{m}_{10} \\ \dot{m}_{01} &= \mu g \text{sgn}(m_{01}) \text{erf}(|m_{01}|/\sqrt{2}\sigma_2) - 2\zeta m_{01} - \omega_o^2 m_{10} - \varepsilon m_{10}(3\sigma_1^2 + m_{10}^2) \\ &\quad + \mu D_{vf} (2/\sqrt{2\pi}\sigma_2) \exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) + u \\ \dot{m}_{11} &= \sigma_2^2 - 2\zeta c_{12} - \omega_o^2 \sigma_1^2 - 3\varepsilon \sigma_1^2 (\sigma_1^2 + m_{10}^2) - \mu g \sqrt{2/\pi} (c_{12}/\sigma_2) \exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) \\ &\quad - \mu D_{vf} (c_{12} m_{01} / \sqrt{2/\pi} \sigma_2^3) \exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) + m_{10} \dot{m}_{01} + \dot{m}_{10} m_{01} \\ \dot{m}_{02} &= -4\zeta \sigma_2^2 - 2\omega_o^2 c_{12} - 6\varepsilon c_{12} (\sigma_1^2 + m_{10}^2) - 2\mu g \sqrt{2/\pi} \sigma_2 \exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) \\ &\quad - \mu D_{vf} \sqrt{2/\pi} (m_{01}/\sigma_2) \exp(-\frac{1}{2}(m_{01}/\sigma_2)^2) + 2\mu^2 D_v + 2D_f + 2m_{01} \dot{m}_{01} \\ &\quad + 4\mu D_{vf} \text{sgn}(m_{01}) \text{erf}(|m_{01}|/\sqrt{2}\sigma_2) \end{aligned} \quad (13)$$

where  $c_{12} = m_{11} - m_{01}m_{01}$ ,  $\sigma_1^2 = m_{20} - m_{10}^2$  and  $\sigma_2^2 = m_{02} - m_{01}^2$ . Here,  $c_{12}$  is the covariance,  $\sigma_1^2$  and  $\sigma_2^2$  are the variance of  $x_1$  and  $x_2$  respectively.

The initial conditions required to integrate Equations (13) from  $t = 0$  to  $t = \tau$  are specified by the coordinates of a cell center  $(x_1, x_2)$ , i.e.  $m_{10}(0) = x_1$ ,  $m_{01}(0) = x_2$ ,  $m_{20}(0) =$

$x_1^2$ ,  $m_{02}(0) = x_2^2$  and  $m_{11}(0) = x_1x_2$ . After obtaining the time evolution of the first and second order moments, we construct a joint Gaussian probability density function (PDF) for  $(x_1, x_2)$  from time  $t = 0$  to  $t = \tau$ . With this PDF, the cost functional and other response statistics over one time step  $\tau$  can be readily calculated. The system is parametrically and externally excited and diffusion term is a nonlinear function of the state.

### Numerical solutions and discussion

We consider a region  $\mathbb{D} = [-2, 2] \times [-2, 2]$ . It 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 in Equation (11) is specified using  $\alpha = \beta = 0.5$  and the control set is uniformly discretized into 11 levels:  $u \in \{-1, -0.8, \dots, 1\}$ .

The reader must notice that there is a region on the  $x_1$ -axis about the origin, where the mean trajectories of the uncontrolled response are trapped due to the effect of dry friction. In Equation (13), when the term  $\mu g \text{sgn}(m_{01}) \text{erf}(|m_{01}|/\sqrt{2}\sigma_2)$  is dominant, a never ending sequence of changes in the sign of  $x_2$  takes place. This term causes the response to switch indefinitely about the trapping region without having a net displacement. Figure 1 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. Jumps in the velocity are still present, but the trapping region on the  $x_1$ -axis is not. These jumps show that the velocity just before and after reaching maximum elongation of the spring differ in magnitude.

Now we evaluate the time evolution of the system response starting from a uniformly distributed initial condition. This initial condition allows us to study the control performance on the entire computational domain. The out crossing of any boundary of  $\mathbb{D}$  is an irreversible process in the sense that  $\bar{\mathbb{D}}$  acts as a sink cell [10]. The reader should notice that such boundaries are leaking probability and eventually will absorb all of it. This however, is a mere consequence of having a bounded computational domain. In order to circumvent this difficulty, we define stationarity as the state in which the leakage of probability is stable and considerably small, e.g. both the probability of the system remaining in  $\mathbb{D}$  and the expected value of the Lagrangian approach to a constant value with very small changes thereafter. Time evolutions starting from this pseudo-stationary state will be governed by the diffusion component of the dynamics.

The stationary PDF of the uncontrolled system response is shown in Figure 2. Stationarity is reached after 10 time units when the 67% of the probability remains in  $\mathbb{D}$ . Figure 3 shows the stationary PDF for the controlled response. The controlled response keeps 82% of the probability in  $\mathbb{D}$  and reaches stationarity in about 7 time units. In order to evaluate the control performance, the leakage of probability to the outside of  $\mathbb{D}$  must be taken into consideration. We propose the following quantity to evaluate the control performance:

$$J_{\mathbb{D}}(t) = \frac{E[L(\mathbf{x}(t), \mathbf{u}(t))]}{P_{\mathbb{D}}(t)} \quad (14)$$

where  $P_{\mathbb{D}}(t)$  is the probability of the system of being in  $\mathbb{D}$  as defined in Equation (9). Figure 4 shows the time evolutions of  $J_{\mathbb{D}}$  and  $P_{\mathbb{D}}$  for both the uncontrolled and controlled system responses. By comparison we conclude that (i) the controlled response reaches the target with less cost, (ii) the stationary PDF of the controlled system is more highly concentrated

around the desired target set, (iii) a faster convergence to the stationary PDF takes place and (iv) a higher percentage of the probability  $P_{\mathbb{D}}(t)$  is kept in the domain  $\mathbb{D}$ .

## 4.2 Non-analytically closeable terms

The closure methods can readily handle polynomials. However, for other types of nonlinearities the closure might not only require tedious and lengthy integrations [16], but also might lead to expressions that don't admit a closed form. This problem prevents us from integrating the system of moment equations even numerically. In order to overcome this difficulty we make use of the cellular structure of the state space by approximating such nonlinearities with Taylor expansion. Once the approximations are available, the infinite hierarchy of moments can be readily closed. We illustrate this approach next.

The example is concerned with the problem of driving a boat in a vortex field. The boat moves on the  $(x_1, x_2)$  plane with a constant velocity relative to the water. Let the control  $u$  be the heading angle with respect to the positive  $x_1$ -axis. The velocity field of the water is given by  $v(x_1, x_2) = ar/(br^2 + c)$  where  $r = \sqrt{x_1^2 + x_2^2}$  is the distance to the origin, i.e. the center of the vortex. The equation of motion of the boat is described by the SDE equations in the Stratonovich sense:

$$\begin{aligned}\dot{x}_1 &= \cos(u) - vx_2/r + \alpha v w_1, \\ \dot{x}_2 &= \sin(u) + vx_1/r + \alpha v w_2,\end{aligned}\tag{15}$$

where  $\alpha$  is a constant,  $w_1$  and  $w_2$  are correlated Gaussian white noise processes with zero mean such that  $E[w_1(t)w_1(t')] = 2D_1\delta(t-t')$ ,  $E[w_2(t)w_2(t')] = 2D_2\delta(t-t')$  and  $E[w_1(t)w_2(t')] = 2D_{12}\delta(t-t')$ . The control objective is to drive the boat from an initial condition to the target set defined by  $\Psi$  such that the cost function:

$$J = E \left[ \int_{t_0}^T \lambda v(x_1, x_2) dt \right],\tag{16}$$

is minimized. Here, we assume that the risk of navigating on the vortex is proportional to the tangential component of the vortex velocity.  $\lambda$  is a scaling constant. The cost in Equation (16) represents the cumulative risk associated with the selected path. The derivation of the Itô equations is shown in Appendix B.

Consider a special case when the white noise processes are uncorrelated. The first and second order moments of the state variables are derived as:

$$\begin{aligned}\dot{m}_{10} &= \cos(u) - E \left[ \frac{x_2 v}{r} \right] + \alpha^2 D_1 E \left[ v \frac{\partial v}{\partial x_1} \right] \\ \dot{m}_{01} &= \sin(u) + E \left[ \frac{x_1 v}{r} \right] + \alpha^2 D_2 E \left[ v \frac{\partial v}{\partial x_2} \right] \\ \dot{m}_{20} &= 2m_{10} \cos(u) - 2E \left[ \frac{x_1 x_2 v}{r} \right] + 2\alpha^2 D_1 E \left[ \frac{x_1 v}{r} \frac{\partial v}{\partial x_1} \right] + 2\alpha^2 D_1 E [v^2] \\ \dot{m}_{02} &= 2m_{01} \sin(u) + 2E \left[ \frac{x_1 x_2 v}{r} \right] + 2\alpha^2 D_2 E \left[ \frac{x_2 v}{r} \frac{\partial v}{\partial x_2} \right] + 2\alpha^2 D_2 E [v^2] \\ \dot{m}_{11} &= \cos(u)m_{01} - E \left[ \frac{x_2^2 v}{r} \right] + \alpha^2 D_1 E \left[ \frac{x_2 v}{r} \frac{\partial v}{\partial x_1} \right] + \sin(u)m_{10} + E \left[ \frac{x_1^2 v}{r} \right] + \alpha^2 D_2 E \left[ \frac{x_1 v}{r} \frac{\partial v}{\partial x_2} \right]\end{aligned}\tag{17}$$

Several expected values such as  $E \left[ \frac{x_2 v}{r} \right]$  in the above moment equations cannot be analytically expressed by the lower order moments. That is to say, the Gaussian closure method cannot be used to analytically close the infinite hierarchy of the moment equations.

Consider an initially delta PDF function to evolve over a short mapping time step  $\tau$ . The one step PDF function at  $t = \tau$  will still be concentrated in a small region with a mean not far away from the initial condition [14]. This fact suggests us to expand the non-analytically closeable functions about the initial condition, i.e. a cell center, in a Taylor series. The Taylor series provides an accurate representation of these nonlinear functions in a small neighborhood where the one step PDF function at  $t = \tau$  is defined. The polynomials of the Taylor series allow us to analytically express the higher order moments in terms of the lower order ones according to the Gaussian closure method.

Let  $f(x_1, x_2)$  denote a non-analytically closeable function such as  $\frac{x_2 v}{r}$  in Equation (17). A second order Taylor approximation  $\tilde{f}(x_1, x_2)$  is given by:

$$\tilde{f}(x_1, x_2) = \beta_1 + \beta_2 x_1 + \beta_3 x_2 + \beta_4 x_1^2 + \beta_5 x_1 x_2 + \beta_6 x_2^2, \quad (18)$$

where the coefficients are calculated from the following conditions:

$$\begin{aligned} f(x_1, x_2)|_{\mathbf{z}} &= \tilde{f}(x_1, x_2)|_{\mathbf{z}}, & f_{x_1 x_2}(x_1, x_2)|_{\mathbf{z}} &= \tilde{f}_{x_1 x_2}(x_1, x_2)|_{\mathbf{z}}, \\ f_{x_2}(x_1, x_2)|_{\mathbf{z}} &= \tilde{f}_{x_2}(x_1, x_2)|_{\mathbf{z}}, & f_{x_1}(x_1, x_2)|_{\mathbf{z}} &= \tilde{f}_{x_1}(x_1, x_2)|_{\mathbf{z}}, \\ f_{x_1 x_1}(x_1, x_2)|_{\mathbf{z}} &= \tilde{f}_{x_1 x_1}(x_1, x_2)|_{\mathbf{z}}, & f_{x_2 x_2}(x_1, x_2)|_{\mathbf{z}} &= \tilde{f}_{x_2 x_2}(x_1, x_2)|_{\mathbf{z}}. \end{aligned}$$

In these equations,  $\mathbf{z}$  stands for the state coordinates of the initial cell center and the subindices of the function refer to the partial derivative with respect to the state variable.

The Taylor expansion is for each initial cell  $\mathbf{z}$ , and is used only to integrate the moment equations over a short time interval  $(0, \tau)$ .

## Numerical solutions and discussion

The region  $\mathbb{D} = [-2, 2] \times [-2, 2]$  is discretized with 1089 square cells. A set of 30 evenly spaced angles values is taken as the control set, i.e.  $\mathbb{U} = \{-\pi, -14\pi/15, \dots, 14\pi/15\}$ . Notice that the set  $\mathbb{U}$  represents an unbounded circular set of controls. The system parameters are chosen as follows:  $\lambda = 1$ ,  $a = 15$ ,  $b = 10$ ,  $c = 2$ ,  $\alpha = 1$ ,  $D_1 = 0.05$  and  $D_2 = 0.05$ . Let  $\Omega$  be the set of cells that form a discrete representation of the target set  $\Psi = \{x_1 = 2, x_2\}$ .  $\Omega$  consists of the cells in the rightmost column of the discretized domain  $\mathbb{D}$ .

Figure 5 shows the mean vector field of the vortex. A trajectory of the boat freely moving in the vortex starting from  $\mathbf{x} = (0, -1)$  is superimposed. The center of the vortex attracts probability due to the state dependence of the diffusion term and not to the existence of an attracting point at the origin. This behavior does not have a corresponding counterpart in the deterministic case. The time evolutions for  $E[L]$  and the first and second order moments are shown in Figure 6.

The vector field for the mean of the controlled response is shown in Figure 7. The trajectory of the boat under the optimal navigation starting from the same initial condition as in Figure 5 is superimposed. The corresponding time evolutions are shown in Figure 8. In the process, the control keeps the system in  $\mathbb{D}$  with probability one. A discontinuity in the vector field exists in spite of having an unbounded control set. Such a discontinuity implies

a dichotomy in the long term behavior of the controlled response depending on the initial state of the system.

While from the initial conditions above the discontinuity, the control will guide the boat against the velocity field of the vortex, from the initial conditions under, the control will move the boat in the direction of the current. Figure 9 shows the controlled response of a system starting from the deterministic initial condition  $\mathbf{x} = (0.8, -0.32)$  after 3 time units. The two-peak density function occurs due to the discontinuity in the global control solution. Due to the diffusion effect of the random excitation, probability density functions that reach the discontinuity will bifurcate into two parts.

### 4.3 Singular boundary conditions

A vibro-impact system with a one-sided rigid barrier subject to a Gaussian white noise excitation is considered. The impact is assumed to be perfectly elastic. The control of the vibro-impact system is then subject to a state constraint given by the one-sided rigid barrier. Such a constraint makes the solution of optimal control problem very difficult to obtain analytically.

A transformation of the state variables is used to effectively remove the non-smooth state constraint. The optimal control problem with fixed terminal conditions is accordingly transformed to the new domain and solved with the proposed method. Then, the control solution is transformed back to the physical domain, where it is evaluated using the GCM method.

In the transformed domain the solution satisfies the constraints. However, the control constraint may be violated when the solution is transformed to the physical domain. For systems with optimal bang-bang solutions, the inverse transformation preserves the control bounds, and therefore all constraints are satisfied.

The equation of motion for the vibro-impact system is given by:

$$\ddot{y} + \theta(y^2 - 1)\dot{y} + 2\zeta\Omega\dot{y} + \Omega^2y = u(t) + w(t), \quad (19)$$

subject to the impact condition:

$$\dot{y}(t_{impact}^-) = -\dot{y}(t_{impact}^+) \quad \text{at } y(t_{impact}) = -h,$$

where  $t_{impact}$  is the time instant at which the impact occurs,  $y(t)$  is the displacement,  $w(t)$  is a Gaussian white noise process satisfying  $E[w(t)] = 0$ ,  $E[w(t)w(t + \xi)] = 2D\delta(\xi)$  and  $u(t)$  is a bounded control satisfying  $|u| \leq \hat{u}$ .

Equation (11) with  $\beta = 0$  is used here. The control objective is to drive the system from any arbitrary initial condition to the origin of the phase space, i.e.  $\Psi_{(y,\dot{y})} = (0, 0)$ , while the energy of the system is minimized. The computational domain is chosen to be  $\mathbb{D}_y = [-h, a] \times [-v, v]$ . As suggested in reference [7], we introduce the following transformation  $y = |x| - h$ . This leads to the SDE for  $x$ :

$$\ddot{x} + \theta(x^2 - 2h|x| + h^2 - 1)\dot{x} + 2\zeta\Omega\dot{x} + \Omega^2x = \text{sgn}(x)(u(t) + w(t) + h\Omega^2). \quad (20)$$

Now, the transformed domain is given by  $\mathbb{D}_x = [-2h, 2a] \times [-v, v]$ , the target set is  $\Psi_{(x,\dot{x})} = (\pm h, 0)$ , and the cost functional is given by:

$$J_x = E \left[ \int_{t_0}^{\infty} [(|x| - h)^2 + \dot{x}^2] dt \right] \quad (21)$$

Let  $x_1 = x$  and  $x_2 = \dot{x}$ . The Itô SDE of the system is given by:

$$\begin{aligned} dx_1 &= x_2 dt \\ dx_2 &= (-\theta x_1^2 x_2 + 2\theta h |x_1| x_2 - (\theta \Gamma + 2\zeta \Omega) x_2 - \Omega^2 x_1 + \text{sgn}(x_1) \Delta) dt + \text{sgn}(x_1) \sqrt{2D} dB \end{aligned} \quad (22)$$

where  $B(t)$  is a unit Wiener process,  $\Delta \equiv u + h\Omega^2$  and  $\Gamma \equiv h^2 - 1$ . The first and second order moment equations of the state variables are given by:

$$\begin{aligned} \dot{m}_{10} &= m_{01}, \quad \dot{m}_{20} = 2m_{11} \\ \dot{m}_{01} &= -\theta m_{21} + 2\theta h E[|x_1| x_2] - (\theta \Gamma + 2\zeta \Omega) m_{01} - \Omega^2 m_{10} + E[\text{sgn}(x_1)] \Delta, \\ \dot{m}_{11} &= m_{02} - \theta m_{31} + 2\theta h E[|x_1| x_1 x_2] - (\theta \Gamma + 2\zeta \Omega) m_{11} - \Omega^2 m_{20} + E[x_1 \text{sgn}(x_1)] \Delta \end{aligned} \quad (23)$$

$$\begin{aligned} \dot{m}_{02} &= -2\theta m_{22} + 4\theta h E[|x_1| x_2^2] - 2(\theta \Gamma + 2\zeta \Omega) m_{02} - 2\Omega^2 m_{11} + 2E[x_2 \text{sgn}(x_1)] \Delta \\ &\quad + 2DE[\text{sgn}(x_1)^2] \end{aligned} \quad (24)$$

Formulas in Appendix C can be used to close the above moment equations according to the Gaussian closure method. Next, we apply the current method and solve for optimal controls in the state domain  $\mathbb{D}_x$  with the cost function  $J_x$  and the target set  $\Psi_{(x,\dot{x})}$ . After obtaining the global optimal control solution  $u^*(x_1, x_2)$ , we transform it back to the original domain leading to the optimal control of the original system  $u^*(y, \dot{y})$ . Notice that in general, this process leads to controls that don't preserve the control constraints in  $\mathbb{D}_y$ .

The optimal control in both domains is bang-bang. The bang-bang control is fully determined by the switching curves in the phase space. From the solution  $u^*(x_1, x_2)$ , we find numerical approximations of such curves and then we transform them back to the physical state space.

## Numerical solutions and discussion

In the numerical example, we have chosen:  $\hat{u} = 1$ ,  $h = -1$ ,  $a = 1$ ,  $v = 1$ ,  $\theta = 0$ ,  $\zeta = 0$ ,  $\Omega = 1$  and  $D = 0.1$ . The transformed state space is  $\mathbb{D}_x = [-2, 2] \times [-2, 2]$ . 849 square cells are used to discretize  $\mathbb{D}_x$ . The transformed target set is formed by the cells that contain the points  $\Psi_{(x,\dot{x})} = (\pm 1, 0)$ . Since the optimal control is bang-bang, we have  $\mathbb{U} = \{-1, 1\}$ .

The uncontrolled response of the system is marginally stable about the origin since  $\theta = 0$  and  $\zeta = 0$ . When the system impacts at  $y = h$ , it is reflected back in such a way that  $y$  remains the same and the sign of  $\dot{y}$  is reversed. The crossing of any other boundary is an irreversible process in the sense that  $\bar{\mathbb{D}}_y$  acts as a sink cell [10].

To demonstrate the behavior of the uncontrolled system, we consider an uniformly distributed initial condition in  $\mathbb{D}_y = [-0.86, -0.78] \times [-0.64, -0.57]$ . The PDF of the response after 3 time units is shown in Figure 10. At this time, only 22% of the probability remains inside of  $\mathbb{D}_y$  indicating the strong effect of diffusion on the marginally stable system. It should be noted that even the probability that remains inside  $\mathbb{D}_y$  does not concentrate around the target. Time evolutions of the relevant responses are shown in Figure 11.

The global optimal control solution obtained in the domain  $\mathbb{D}_x$  is shown in Figure 12, which indicates the existence of switching curves of the bang-bang optimal control. Numerical solutions of the switching curves are obtained with the curve fitting method by using this control solution. These curves are superimposed in Figure 12. The switching curves as well as the control solutions are mapped back to the physical domain  $\mathbb{D}_y$  as shown in Figure 13.

It is interesting to point out the qualitative differences in the controlled response for states in the third quadrant of  $\mathbb{D}_y$ . While for some states, marked with crosses, the control speeds up the system favoring impact in others avoids it.

Consider now the same uniformly distributed initial condition in  $\mathbb{D}_y$ . The controlled system is found to converge to the steady state PDF shown in Figure 14 in 4 time units with 98% of the probability remaining in  $\mathbb{D}_y$ . As discussed in the previous example, the switching curves in this example will split a uni-modal PDF of the response into a bi-modal one during the transient. In particular, in the third quadrant of  $\mathbb{D}_y$  just before impact, part of the probability flow speeds up while the other slows down, thus splitting the PDF. The time evolutions of the cost and moments of the response are shown in Figure 15. As before, the controlled response (i) converges to the target set with high probability, (ii) minimizes the cost and (iii) maximizes the probability of staying in the computational domain.

## 5 CONCLUSIONS

This paper presents a study on the optimal control problem of nonlinear stochastic systems using Bellman's principle of optimality, the short-time Gaussian approximation and the Generalized Cell Mapping. Control problems of several challenging nonlinear systems with fixed final state terminal conditions subject to state and control constraints are studied to demonstrate the effectiveness of the approach. In particular, nonlinear systems with a state dependent diffusion part and non-analytically closeable terms are considered. A novel strategy based on multiple local Taylor expansions is proposed to handle the non-closeable terms of the differential equations for the moments of the state variables so that the Gaussian closure method can be applied. The optimal control of a nonlinear vibro-impact oscillator is studied by using transformation of variables. In the process, the dynamics, the cost functional, the admissible state space, and the target set are transformed to a new domain. Once the global feedback solution is obtained, we transform it back to the physical domain where it is evaluated using GCM. In all numerical examples, the method leads to controls with excellent performance. Computational improvements and extensions to higher dimensional processes deserve further investigation. A rigorous study of convergence and stability of the method is still elusive at this time, and should be pursued in the future. Since the solution of optimal control problems of non-linear stochastic systems are difficult to obtain and rare in the literature, this investigation offers a valuable contribution to the field.

## REFERENCES

- [1] D. Anderson, J. Towneill, and R. Pletcher. *Computational Fluid Mechanics and Heat Transfer*. Hemisphere, New York, 1984.
- [2] 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.
- [3] L. G. Crespo. Deterministic and stochastic optimal controls and cell mapping. *University of Delaware, Doctoral Dissertation*, 2002.

- [4] L. G. Crespo and J. Q. Sun. An improved strategy to solve the fixed final time optimal control problem via simple cell mapping. In *Proceedings of Conference on Nonlinear Vibrations Stability and Dynamics of Structures*, Blacksburg, Virginia, 2000.
- [5] 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.
- [6] 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.
- [7] M. Dimentberg. *Statistical Dynamics of Nonlinear and Time-varying Systems*. Research Studies Press, Tauton, U.K., 1988.
- [8] M. Dimentberg, D. Iourtchenko, and A. Bratus. Optimal bounded control of steady-state random vibrations. *Probabilistic Engineering Mechanics*, 15:381–386, 2000.
- [9] C. S. Hsu. A discrete method of optimal control based upon the cell state space concept. *Journal of Optimization Theory and Applications*, 46(4):547–569, 1985.
- [10] C. S. Hsu. *Cell-to-cell Mapping, A Method of Global Analysis for Nonlinear Systems*. Springer-Verlag, New York, 1987.
- [11] H. J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, New York, 2001.
- [12] Y. K. Lin and G. Q. Cai. *Probabilistic Structural Dynamics: Advanced Theory and Applications*. McGraw-Hill, New York, 1995.
- [13] A. Naess and V. Moe. Efficient path integration methods for nonlinear dynamic systems. *Probabilistic Engineering Mechanics*, 15:221–231, 2000.
- [14] H. Risken. *The Fokker-Planck Equation - Methods of Solution and Applications*. Springer-Verlag, New York, 1984.
- [15] L. Su and G. Ahmadi. A comparison study of performances of various isolation systems. *Earthquake Engineering and Structural Dynamics*, 19:21–23, 1990.
- [16] 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.
- [17] J. Q. Sun and C. S. Hsu. Cumulant-neglect closure method for asymmetric non-linear systems driven by Gaussian white noise. *Journal of Sound and Vibration*, 135:338–345, 1989.
- [18] J. Q. Sun and C. S. Hsu. The generalized cell mapping method in nonlinear random vibration based upon short-time Gaussian approximation. *Journal of Applied Mechanics*, 57:1018–1025, 1990.

- [19] B. Tombuyses and T. Aldemir. Continuous cell-to-cell mapping. *Journal of Sound and Vibration*, 202(3):395–415, 1997.
- [20] M. F. Wehner and W. G. Wolfer. Numerical evaluation of path-integral solutions to Fokker-Planck equations III: Time and functionally dependent coefficients. *Physical Review A*, 35(4):1795–1801, 1987.
- [21] J. Yong and X. Y. Zhou. *Stochastic Controls, Hamiltonian Systems and HJB Equations*. Springer-Verlag, New York, 1999.
- [22] W. Q. Zhu, Z. G. Ying, and T. T. Soong. An optimal nonlinear feedback control strategy for randomly excited structural systems. *Nonlinear Dynamics*, 24:31–51, 2001.

## Appendix A

Details on the derivation of the equations for the problem studied in Section 4.1 are presented here. From Equation (10), we identify:

$$\sigma = \begin{bmatrix} 0 & 0 \\ 1 & -\mu \operatorname{sgn}(x_2) \end{bmatrix} \mathbf{Q} = 2 \begin{bmatrix} D_f & D_{vf} \\ D_{vf} & D_v \end{bmatrix} \quad (25)$$

The Wong-Zakai correction term is given by  $\Delta \mathbf{m} = [0, \Delta m]^T$ , where

$$\Delta m = \frac{1}{2} \sigma_{2s} Q_{s2} \frac{\partial}{\partial x_2} \sigma_{22} = \frac{1}{2} (2D_{vf} + 2D_v \mu \operatorname{sgn}(x_2)) \mu \operatorname{sgn}'(x_2) \quad (26)$$

Repeated indices are used to indicate summation. The diffusion term in the FPK equation is given by:

$$\sigma \mathbf{Q} \sigma^T = \begin{bmatrix} 0 & 0 \\ 0 & 2D_f + 4\mu D_{vf} \operatorname{sgn}(x_2) + 2\mu^2 D_v \operatorname{sgn}^2(x_2) \end{bmatrix} \quad (27)$$

An equivalent single Wiener process leading to the same FPK equation has the diffusion term  $\sigma_e = [0, \sigma_e]^T$  with  $\sigma_e = (2D_f + 4\mu D_{vf} \operatorname{sgn}(x_2) + 2\mu^2 D_v \operatorname{sgn}^2(x_2))^{1/2}$ . Hence, the Itô SDE is given by  $d\mathbf{x}(t) = (\mathbf{m}(\mathbf{x}, \mathbf{u}) + \Delta \mathbf{m}) dt + \sigma_e dB(t)$ , which is Equation (12).

## Appendix B

Details on the derivation of the equations for the problem studied in Section 4.2 are presented here. From Equation (15) we find

$$\sigma = \alpha \begin{bmatrix} v & 0 \\ 0 & v \end{bmatrix} \mathbf{Q} = 2 \begin{bmatrix} D_1 & D_{12} \\ D_{12} & D_2 \end{bmatrix} \quad (28)$$

where  $v$  is an arbitrary function of the state variables. The Wong-Zakai correction term is given by  $\Delta m_j = \frac{1}{2} \sigma_{ks} Q_{sl} \frac{\partial}{\partial x_k} \sigma_{jl}$ . After some manipulations we obtain:

$$\begin{aligned} \Delta m_1 &= \alpha^2 v (D_1 \frac{\partial v}{\partial x_1} + D_{12} \frac{\partial v}{\partial x_2}) \\ \Delta m_2 &= \alpha^2 v (D_2 \frac{\partial v}{\partial x_2} + D_{12} \frac{\partial v}{\partial x_1}) \end{aligned} \quad (29)$$

The diffusion term in the FPK equation is given by  $\sigma \mathbf{Q} \sigma^T = 2\alpha^2 v^2 \mathbf{Q}$ . An equivalent Wiener process leading to the same FPK equation has the diffusion terms:

$$\begin{aligned}\sigma_{e1} &= \left[ 2\alpha^2 \left( D_1 v^2 + D_{12} \int \frac{\partial v^2}{\partial x_2} dx_1 \right) \right]^{1/2} \\ \sigma_{e2} &= \left[ 2\alpha^2 \left( D_2 v^2 + D_{12} \int \frac{\partial v^2}{\partial x_1} dx_2 \right) \right]^{1/2}\end{aligned}\quad (30)$$

Hence, the Itô SDE is given by:

$$\begin{aligned}dx_1 &= \left( \cos(u) - \frac{x_2 v}{\sqrt{x_1^2 + x_2^2}} + \alpha^2 v (D_1 \frac{\partial v}{\partial x_1} + D_{12} \frac{\partial v}{\partial x_2}) \right) dt + \sigma_{e1} dB_1 \\ dx_2 &= \left( \sin(u) + \frac{x_1 v}{\sqrt{x_1^2 + x_2^2}} + \alpha^2 v (D_2 \frac{\partial v}{\partial x_2} + D_{12} \frac{\partial v}{\partial x_1}) \right) dt + \sigma_{e2} dB_2\end{aligned}\quad (31)$$

For uncorrelated white noise processes, the moment equations are:

$$\begin{aligned}\dot{m}_{10} &= \cos(u) - E \left[ \frac{x_2 v}{r} \right] + \alpha^2 D_1 E \left[ v \frac{\partial v}{\partial x_1} \right] \\ \dot{m}_{01} &= \sin(u) + E \left[ \frac{x_1 v}{r} \right] + \alpha^2 D_2 E \left[ v \frac{\partial v}{\partial x_2} \right] \\ \dot{m}_{11} &= \cos(u) m_{01} - E \left[ \frac{x_2^2 v}{r} \right] + \alpha^2 D_1 E \left[ \frac{x_2 v}{r} \frac{\partial v}{\partial x_1} \right] + \sin(u) m_{10} + E \left[ \frac{x_1^2 v}{r} \right] + \alpha^2 D_2 E \left[ \frac{x_1 v}{r} \frac{\partial v}{\partial x_2} \right] \\ \dot{m}_{20} &= 2m_{10} \cos(u) - 2E \left[ \frac{x_1 x_2 v}{r} \right] + 2\alpha^2 D_1 E \left[ \frac{x_1 v}{r} \frac{\partial v}{\partial x_1} \right] + 2\alpha^2 D_1 E [v^2] \\ \dot{m}_{02} &= 2m_{01} \sin(u) + 2E \left[ \frac{x_1 x_2 v}{r} \right] + 2\alpha^2 D_2 E \left[ \frac{x_2 v}{r} \frac{\partial v}{\partial x_2} \right] + 2\alpha^2 D_2 E [v^2]\end{aligned}\quad (32)$$

In the example, we used Taylor expansions for the functions  $v$  and  $v/\sqrt{x_1^2 + x_2^2}$  up to the second order. After substituting the polynomial approximations into Equation (32), the expected values of the polynomial functions in terms of lower order moments are obtained.

## Appendix C

The standard joint Gaussian conditional probability density function of  $x_1$  and  $x_2$  is given by:

$$p(x_1, x_2, \tau | x_{1(0)}, x_{2(0)}, 0) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} \exp\left(\frac{-[\alpha_1^2 - 2\rho\alpha_1\alpha_2 + \alpha_2^2]}{2(1-\rho^2)}\right)\quad (33)$$

where  $\rho = \frac{E[(x_1 - m_1)(x_2 - m_2)]}{\sigma_1\sigma_2}$ ,  $m_i = E[x_i]$ ,  $\sigma_i^2 = E[(x_i - m_i)^2]$ ,  $\alpha_i = \left(\frac{x_i - m_i}{\sigma_i}\right)$  for  $i = 1, 2$ . In these expressions,  $m_i$  and  $\sigma_i$  and  $\rho$  are the means, standard variations and correlation coefficient of the density function evaluated at time  $t = \tau$  when the system starts from

the deterministic initial condition  $\mathbf{x} = (x_1, x_2)^T$ . Notice that for real variables,  $\rho$  always lies between  $-1$  and  $1$ . Recall that expected values of an arbitrary function of  $x_1$  and  $x_2$  using the Gaussian closure, can be computed from:

$$E[f(x_1, x_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \frac{f(x_1, x_2)}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} \exp\left(\frac{-[\alpha_1^2 - 2\rho\alpha_1\alpha_2 + \alpha_2^2]}{2(1-\rho^2)}\right) \right] dx_1 dx_2 \quad (34)$$

The following expressions have been found.

$$E[\text{sgn}(x_i)] = \text{sgn}\left(\frac{m_i}{\sigma_i\sqrt{2}}\right) \text{erf}\left(\frac{|m_i|}{\sigma_i\sqrt{2}}\right) \quad E[\text{sgn}'(x_i)] = \frac{2}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{m_i^2}{2\sigma_i^2}\right) \quad (35)$$

$$E[x_j \text{sgn}(x_i)] = \frac{2\rho\sigma_1\sigma_2}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{m_i^2}{2\sigma_i^2}\right) + m_j E[\text{sgn}(x_i)] \quad E[\text{sgn}(x_i)\text{sgn}'(x_i)] = 0 \quad (36)$$

$$E[(x_j - m_j)E[\text{sgn}(x_i)]] = 0, \quad E[(x_j - m_j)x_j^3] = 3\sigma_j^2(\sigma_j^2 + m_j^2) \quad (37)$$

$$E[x_j \text{sgn}'(x_i)] = \frac{-m_i\rho\sigma_1\sigma_2}{\sigma_i^3\sqrt{2\pi}} \exp\left(-\frac{m_i^2}{2\sigma_i^2}\right) + m_j E[\text{sgn}'(x_i)] \quad E[\text{sgn}(x_i)^2] = 1 \quad (38)$$

$$E[x_j \text{sgn}(x_i)\text{sgn}'(x_i)] = 0 \quad E[x_i \text{sgn}(x_i)] = \frac{2\sigma_i}{\sqrt{2\pi}} \exp\left(-\frac{m_i^2}{2\sigma_i^2}\right) + m_i E[\text{sgn}(x_i)] \quad (39)$$

$$E[x_i \text{sgn}(x_i)\text{sgn}'(x_i)] = 0, \quad E[(x_j - m_j)x_j^3] = 3\rho\sigma_1\sigma_2(\sigma_i^2 + m_i^2) \quad (40)$$

$$E[x_i \text{sgn}'(x_i)] = 0, \quad E[x_i^3] = m_i(3\sigma_i^2 + m_i^2) \quad (41)$$

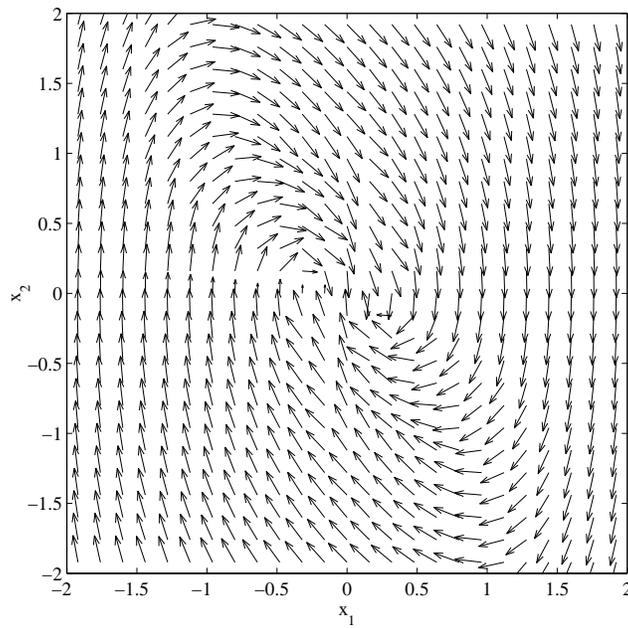


Figure 1: *Vector field of the expected trajectories of the controlled response.*

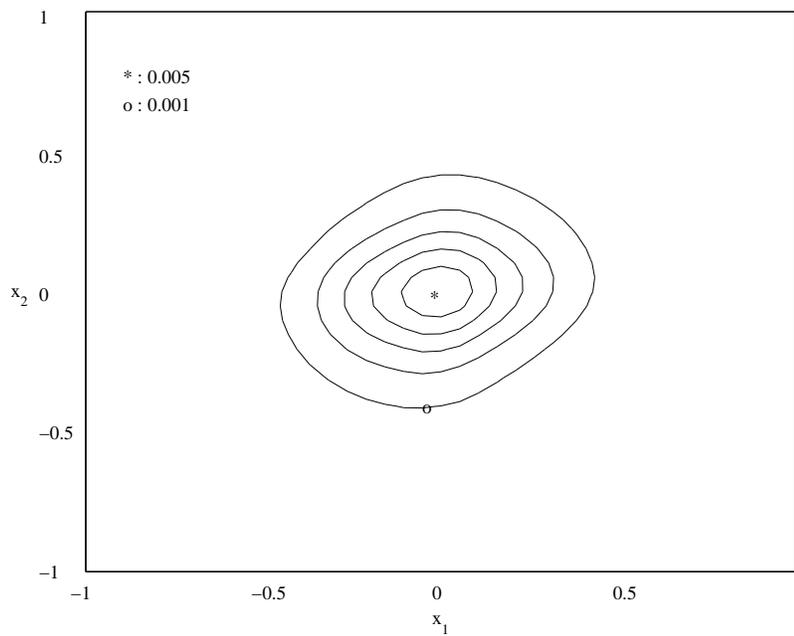


Figure 2: *Contours of the stationary PDF of the uncontrolled response of the mass block system with dry friction damping.*

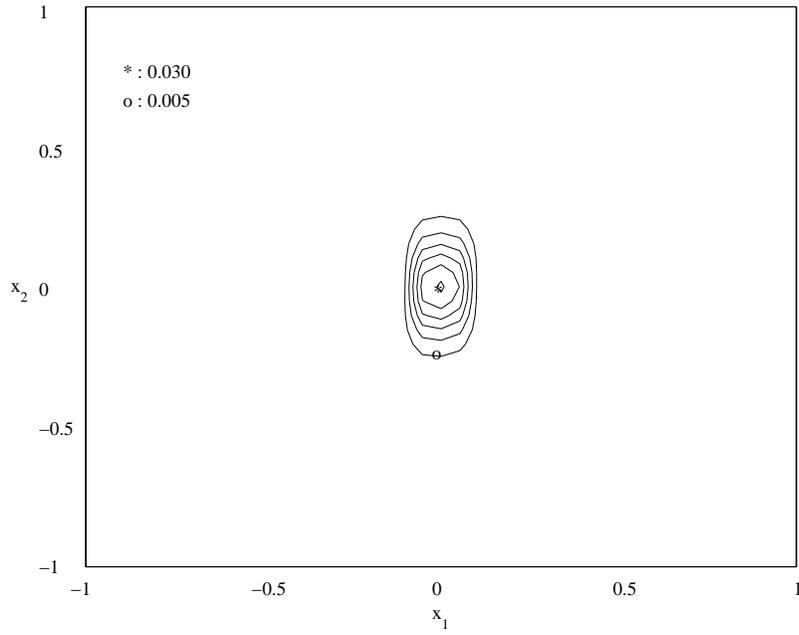


Figure 3: *Contours of the stationary PDF for the controlled response of the mass block system with dry friction damping.*

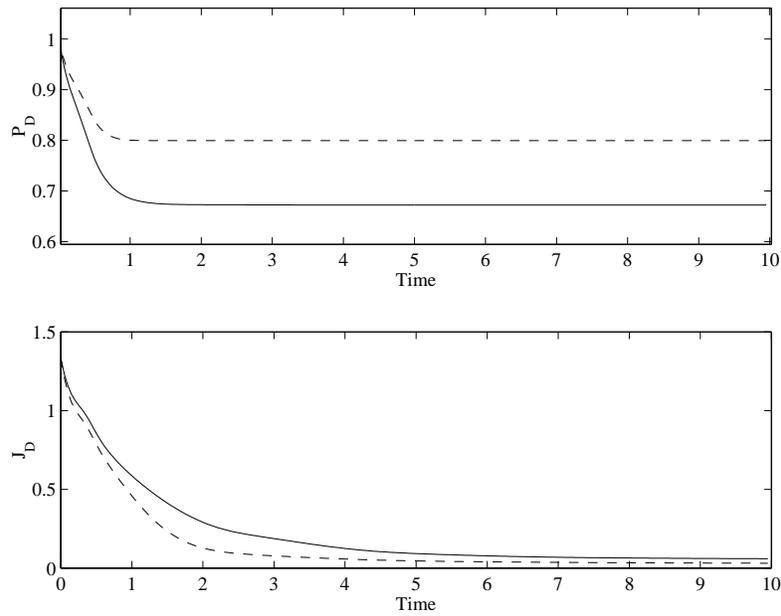


Figure 4: *Time evolutions of the scaled cost  $J_D$  and the probability  $P_D$  of the system with dry friction damping. Uncontrolled: (—) and controlled: (---).*

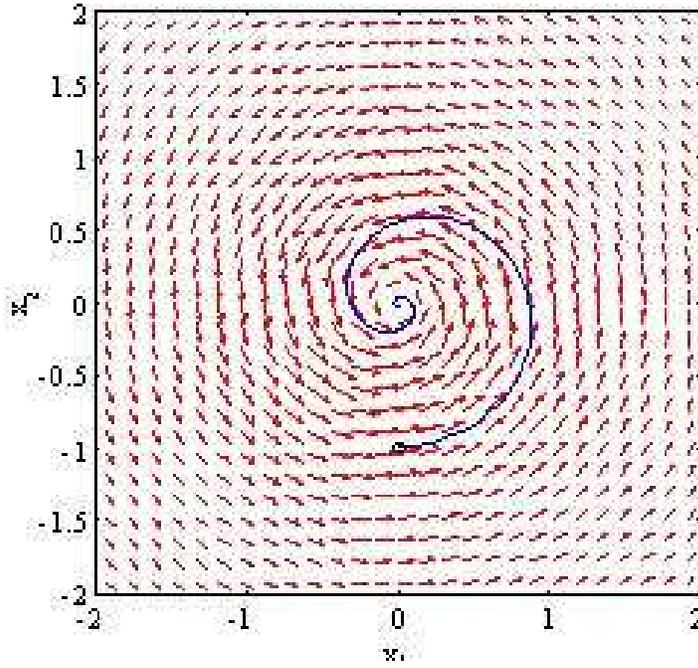


Figure 5: *Expected vector field of the uncontrolled trajectories and a single trajectory starting from  $\mathbf{x} = (0, -1)$ .*

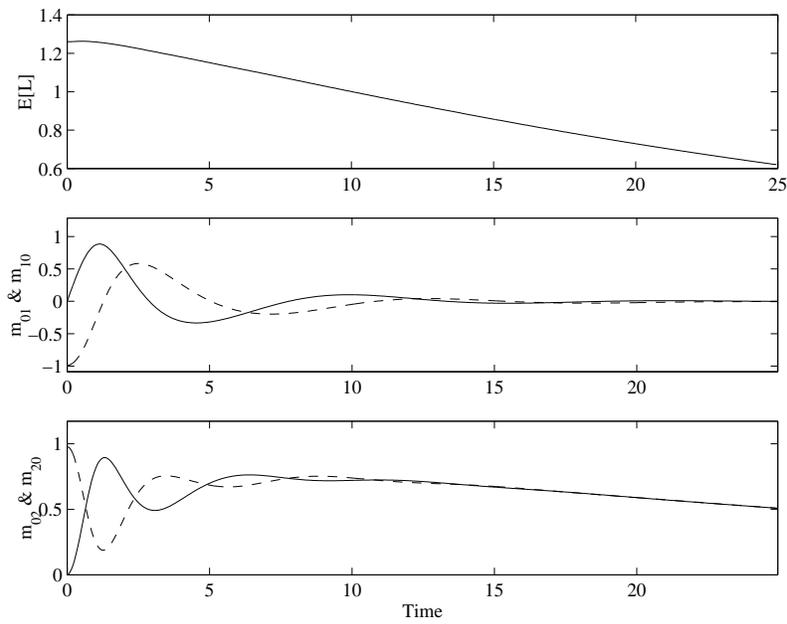


Figure 6: *Time evolutions of  $E[L]$  and the first two order moments of the state variables of the uncontrolled trajectories of the boat navigating in the vortex field. Dashed lines: moments of  $x_1$ . Continuous lines: moments of  $x_2$ .*

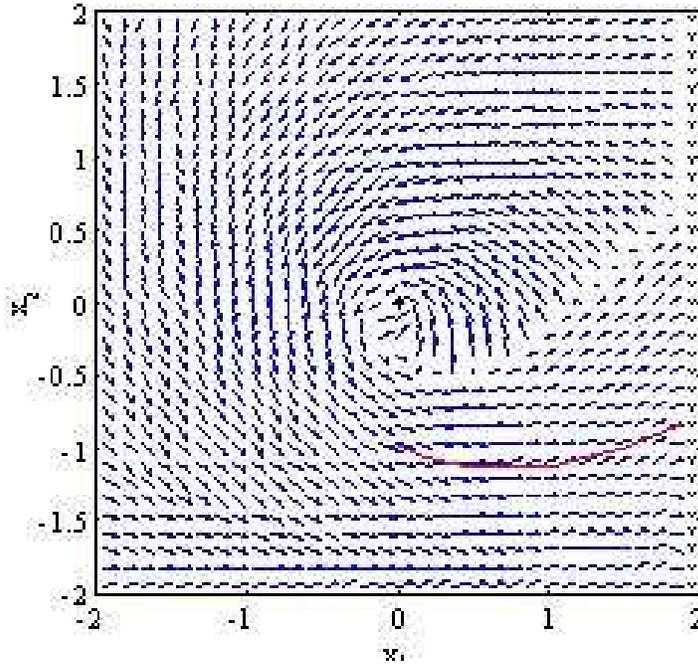


Figure 7: *Expected vector field of the controlled trajectories and single controlled trajectory starting from  $\mathbf{x} = (0, -1)$ . The target cells are marked with crosses.*

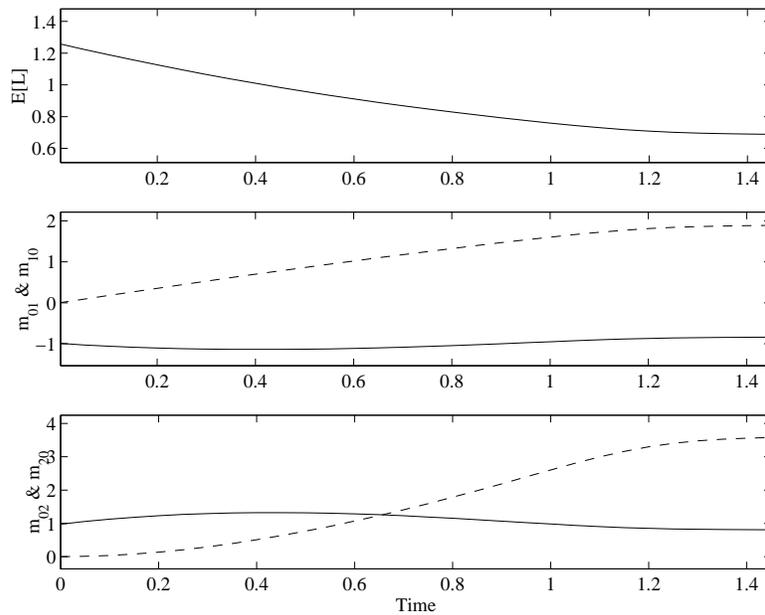


Figure 8: *Time evolutions of  $E[L]$  and the first two order moments of the state variables of the controlled trajectories of the boat navigating in the vortex field. Dashed lines: moments of  $x_1$ . Continuous lines: moments of  $x_2$ .*

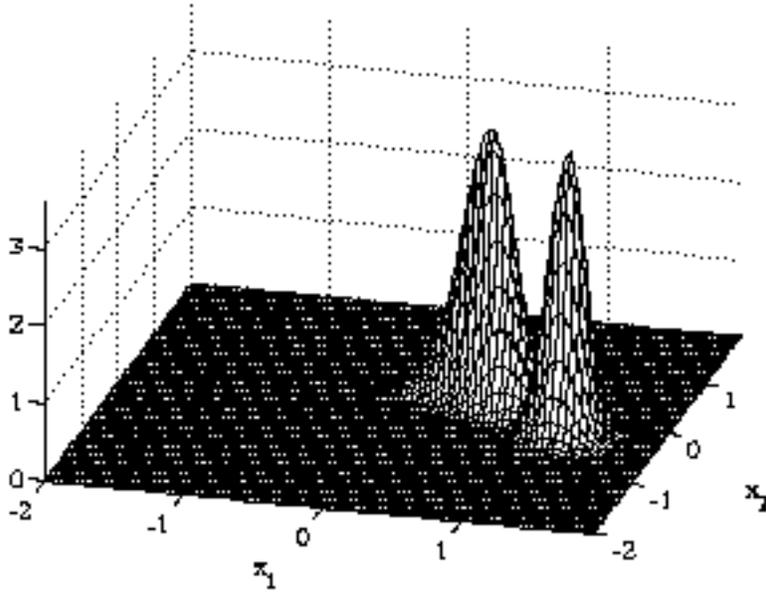


Figure 9: *The PDF of the controlled trajectories of the boat after 0.3 time units for the initial condition  $\mathbf{x}(0) = (0.8, -0.32)$ . The initial condition is close to the discontinuity in optimal controls.*

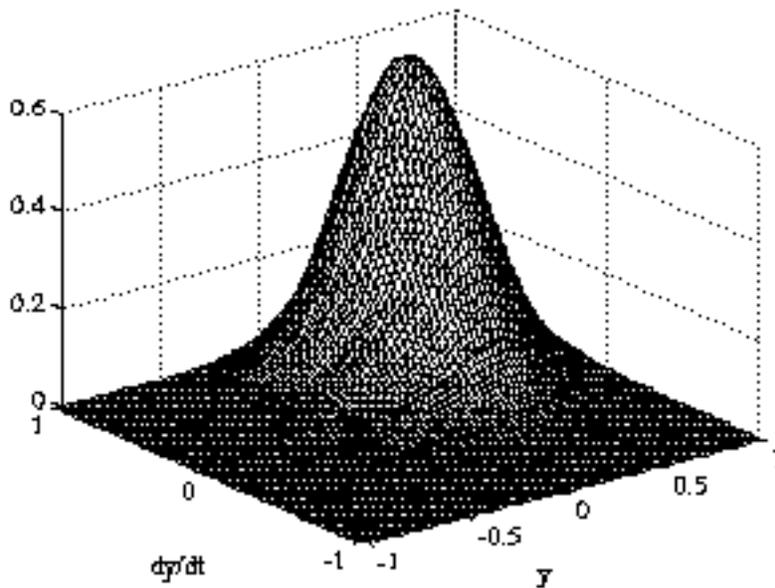


Figure 10: *The PDF of the uncontrolled response of the system with an impact barrier at  $y = -1$  after 3 time units. The system starts from the initial condition  $\mathbf{y} = [-0.7, -0.7]$ .*

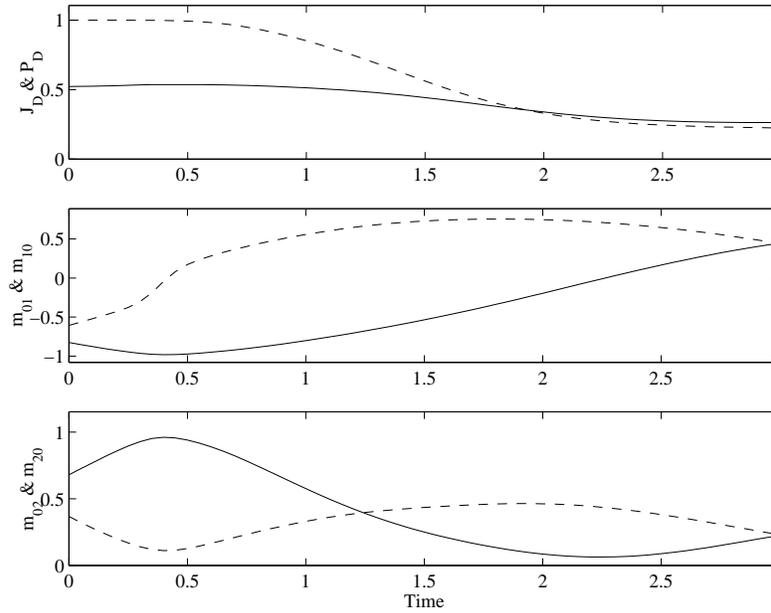


Figure 11: *Time evolutions of  $P_{\mathbb{D}}$  (---),  $J_{\mathbb{D}}$  (—) and the moments  $m_{10}$  (- · -),  $m_{01}$  (---),  $m_{20}$  (- · -), and  $m_{02}$  (---) of the uncontrolled response of the vibro-impact system.*

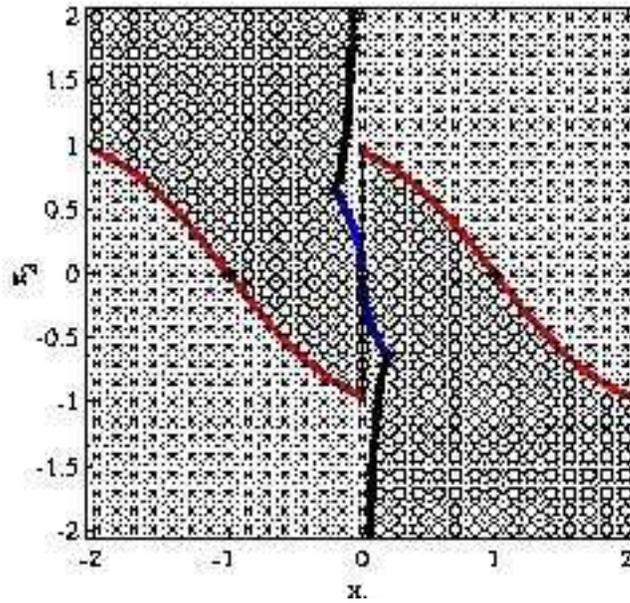


Figure 12: *Global optimal control solution of the vibro-impact system in the transformed domain  $\mathbb{D}_x$ . Cells marked with circles denote the regions where the optimal control is  $u^* = \hat{u}$ , and cells marked with crosses represent the regions where the optimal control is  $u^* = -\hat{u}$ . The continuous lines represent the switching curves*

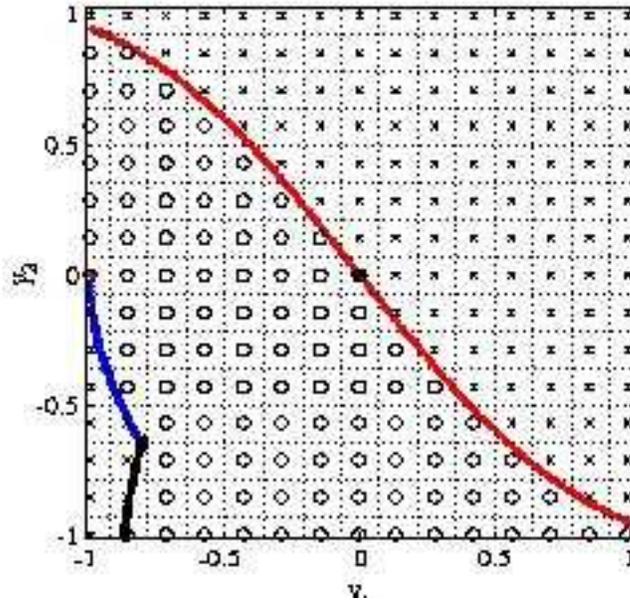


Figure 13: *Global optimal control solution of the vibro-impact system in the original domain  $\mathbb{D}_y$ . The legends are the same as in Figure 12.*

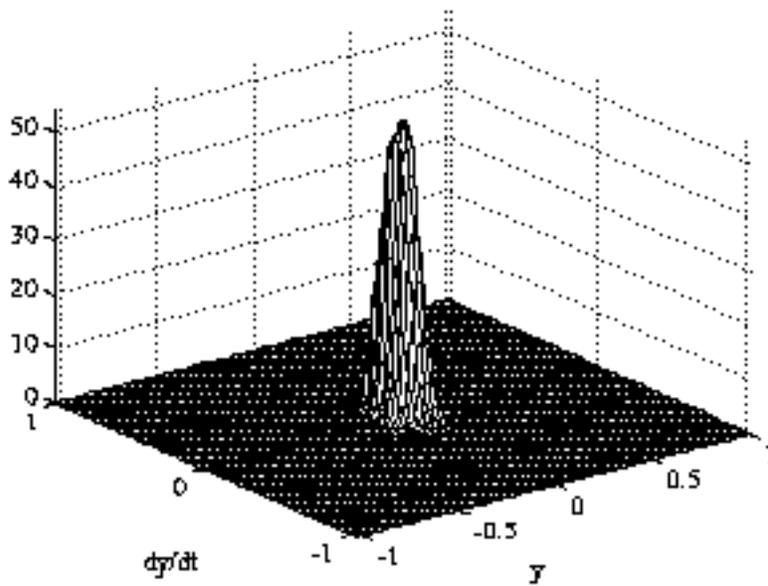


Figure 14: *Stationary PDF of the controlled response of the vibro-impact system.*

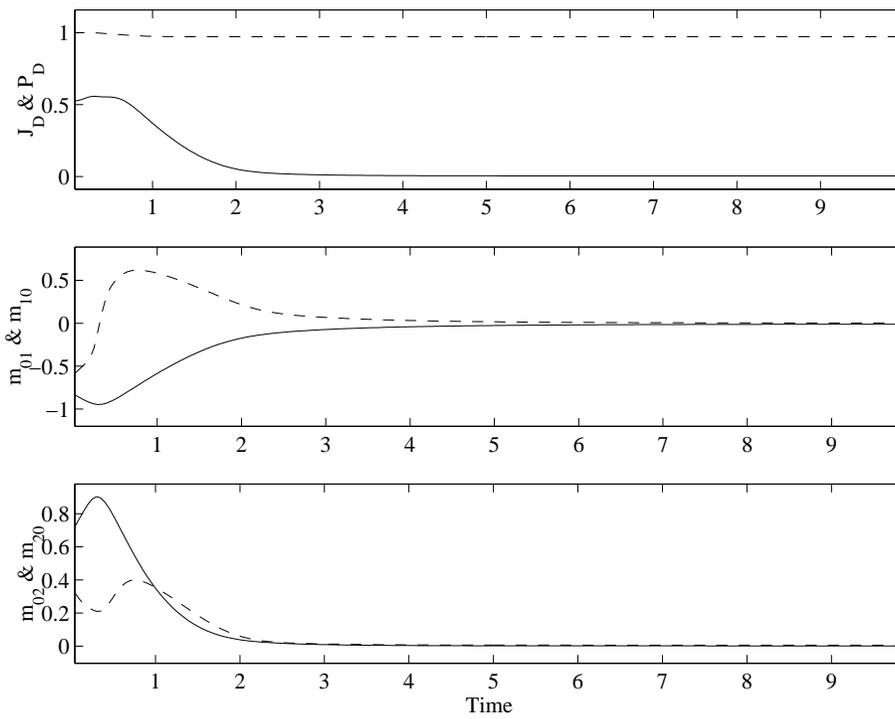


Figure 15: Time evolutions of  $P_D$  (---),  $J_D$  (—) and the moments  $m_{10}$  (-·-),  $m_{01}$  (---),  $m_{20}$  (-·-), and  $m_{02}$  (---) of the controlled response of the vibro-impact system.

**REPORT DOCUMENTATION PAGE**

*Form Approved  
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b>			<b>2. REPORT TYPE</b>		<b>3. DATES COVERED (From - To)</b>	
<b>4. TITLE AND SUBTITLE</b>					<b>5a. CONTRACT NUMBER</b>	
					<b>5b. GRANT NUMBER</b>	
					<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b>					<b>5d. PROJECT NUMBER</b>	
					<b>5e. TASK NUMBER</b>	
					<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>					<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>					<b>10. SPONSORING/MONITOR'S ACRONYM(S)</b>	
					<b>11. SPONSORING/MONITORING REPORT NUMBER</b>	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b>						
<b>13. SUPPLEMENTARY NOTES</b>						
<b>14. ABSTRACT</b>						
<b>15. SUBJECT TERMS</b>						
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>	
<b>a. REPORT</b>	<b>b. ABSTRACT</b>	<b>c. THIS PAGE</b>			<b>19b. TELEPHONE NUMBER (Include area code)</b>	