IEEJ TRANSACTIONS ON ELECTRICAL AND ELECTRONIC ENGINEERING
IEEJ Trans 2020; **15**: 1169–1175
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI:10.1002/tee.23177

Paper

Nonlinear Stochastic Optimal Control with Input Saturation Constraints Based on Path Integrals

Satoshi Satoh*[a], Member
Hilbert J. Kappen**, Non-member

This paper is concerned with a finite-time nonlinear stochastic optimal control problem with input saturation as a hard constraint on the control input. The stochastic Hamilton-Jacobi-Bellman (SHJB) equation associated with this problem is derived, and a concrete solution method is also presented by extending an iteration framework of path integral stochastic optimal control. This is the first result that provides an optimal feedback control rigorously satisfying the saturation constraint as a solution to the SHJB equation. © 2020 Institute of Electrical Engineers of Japan. Published by Wiley Periodicals LLC.

Keywords: nonlinear stochastic system; stochastic optimal control; input saturation; nonlinear control; path integral

Received 21 September 2019; Revised 13 March 2020

1. Introduction

Guarantee of optimal performance under system nonlinearities, actuator limitations and stochastic uncertainties is an important subject in practical control systems, and is a challenging problem in control theory. Input saturation due to actuator limitations possibly causes performance degradation or instability in control systems [1]. Recently, there have been several compensation methods proposed for a class of nonlinear systems. In Ref. [2] a nonsingular terminal sliding mode controller for a nonlinear second-order system is proposed. The stability analysis is performed for a nonlinear fractional-order systems in Ref. [3]. A class of nonlinear time-delay systems is considered in Ref. [4], where a local state feedback controller and stability analysis with a Lyapunov-Krasovskii functional are provided. Further, not only stabilization of an equilibrium point but also disturbance attenuation is considered in Refs [5,6]. The authors in Ref. [5] investigate a constrained adaptive control for a class of uncertain multi-input multi-output nonlinear systems using the backstepping technique. Also, the gain scheduling scheme is utilized for an exponentially stabilizing controller for switched systems with parametric uncertainties in Ref. [6]. Meanwhile, the optimal control has been commonly adopted to achieve optimal performance of control systems. The authors in Ref. [7] investigate a nonlinear optimal control problem with input saturation. They newly derive a Hamilton-Jacobi-Bellman (HJB) equation, which takes the input saturation into account, and provide its solution method based on the stable manifold theory. This result enables one to obtain the optimal control for a nonlinear system, which rigorously satisfies the input constraint. Besides, model predictive control (MPC) [8,9] has been widely accepted in various fields as an effective control method handling optimization with multiple constraints. Robust MPC is further studied against disturbances [10].

Whereas, aforementioned methods are all deterministically formulated. Stochastic uncertainties always exist to a greater or lesser degree in practical systems, and they cause unexpected quality fluctuations and performance degradation. Regarding this, we consider a nonlinear stochastic system, whose dynamics is described by the stochastic differential equation [11–14], and adopt stochastic control theory [15–17] so as to take stochastic influences into account in analysis and controller design. So far, we have proposed several stochastic control methods: stochastic passivity-based control [18], stochastic bounded stability analysis and synthesis [19], performance analysis of visual motion observer [20] and nonlinear stochastic optimal control [21,22]. Particularly in Refs. [21,22], an iterative solution method for a finite-time nonlinear stochastic optimal control problem is provided based on path integral analysis [23,24]. We call this method ISOC-PI shortly, which stands for iterative stochastic optimal control based on path integrals. The purpose of this paper is to formulate a nonlinear stochastic optimal control problem with input saturation as a hard constraint on the control input based on the argument in [7], and moreover to provide a concrete solution to this problem by extending ISOC-PI [22].

In this paper, first, we newly derive a stochastic Hamilton-Jacobi-Bellman (SHJB) equation associated with the nonlinear stochastic optimal control problem with input saturation as a hard constraint. While the deterministic HJB equation provided in Ref. [7] is a nonlinear partial differential equation (PDE) of first order, the present SHJB equation is a second-order nonlinear PDE. Since the solution method in Ref. [7] strongly depends on the geometric structure of the deterministic HJB equation, it is not applicable to the SHJB equation involving second-order partial derivatives. Therefore, second, we present a method, which provides not only the solution to the SHJB equation but also the corresponding optimal control. In order to derive the optimal

[a] Correspondence to: Satoshi Satoh. E-mail: s.satoh@ieee.org

*Graduate School of Engineering, Osaka University, 2-1, Yamadaoka, Suita Osaka 565-0871, Japan

**Department of Biophysics, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

control from the SHJB equation, the partial derivatives of the solution are necessary in general [25,26]. In our framework, an iterative procedure is given as a class of Cauchy problems for linear PDEs, to which the Feynman-Kac formula [14,26] can be applied. Thus, we provide explicitly the stochastic representation of the solution to each iteration procedure by using the Feynman-Kac formula. Moreover, by applying the path integral analysis [22–24] to the resultant solution, the stochastic representations of the partial derivatives of the solution are also obtained. This enables us to explicitly provide the stochastic representation of the corresponding suboptimal control at each iteration without using any numerical differentiation of the solution, which is undesirable particularly in dealing with stochastic systems. From a theoretical point of view, since the resultant stochastic representations are rigorous in the derivation, it is an advantage that any approximation is not required. On the other hand, since those representations form the expectations, they are in general approximately computed by statistical sampling methods. We can select an appropriate sampling method to consider the trade-off between the required accuracy and the computational cost.

In comparison with other optimal control approaches dealing with stochastic uncertainties, a local quadratic approximation-based trajectory optimization [27] has been extended to the stochastic case [28–31]. However, none of them deal with optimal control with input constraints. Although only in Ref. [28], input constraints are partially taken into consideration, the method reduces the gain parameter in an *ad hoc* manner, when the input exceeds the limit. Therefore, the optimality in control does not hold any more. Besides, stochastic MPC has also been extensively studied recently, e.g., [32–35]. Since stochastic MPC methods repeatedly solve an open-loop optimal control problem, they are capable of dealing with various constraints such as state and chance constraints as well as input constraints. On the contrary, they strongly rely on numerical computation, and thus they do not obtain an explicit representation of the optimal control. Moreover, none of them provide the optimal feedback control strictly considering the input saturation as a solution to the corresponding SHJB equation. Finally, let us summarize the main contributions and features of this paper.

- This paper considers a finite-time nonlinear stochastic optimal control problem with input saturation as a hard constraint on the control input. We derive the SHJB equation associated with this problem, and moreover, provide a concrete iterative solution method. This is the first result that provides an optimal feedback control rigorously satisfying the saturation constraint as a solution to the SHJB equation.
- Based on the path integral analysis, stochastic representations for both the solution and the corresponding suboptimal control at each iteration are obtained without using numerical differentiation; and
- Since each iteration is executed using sampling-based computation, it is easy to parallelize, and to consider the trade-off between the required accuracy and the computational cost.

2. Nonlinear Stochastic Optimal Control with Input Saturation Constraints

In this section, first, we formulate a nonlinear stochastic optimal control problem with input saturation. Second, we derive the SHJB equation considering input saturation by extending the deterministic optimal control result in Ref. [7] Finally, we provide a concrete iterative solution to the resultant SHJB equation by modifying our ISOC-PI method in Ref. [22].

This paper considers a stochastic optimal control problem on the time interval $t \in [0, T]$ with any constant $T > 0$ with the following Itô stochastic system with input saturation:

$$\begin{aligned}\begin{pmatrix}\mathrm{d}x^u\\ \mathrm{d}x^c\end{pmatrix} &= \begin{pmatrix}f^u(x,t)\\ f^c(x,t)\end{pmatrix}\mathrm{d}t + \begin{pmatrix}0_{n-n_c\times m}\\ g^c(x,t)\end{pmatrix}\mathrm{sat}(u(t))\mathrm{d}t\\ &\quad + \begin{pmatrix}0_{n-n_c\times r}\\ h^c(x,t)\end{pmatrix}\mathrm{d}w\\ &=: f(x,t)\mathrm{d}t + g(x,t)\mathrm{sat}(u(t))\mathrm{d}t + h(x,t)\mathrm{d}w \qquad (1)\end{aligned}$$

with some functions $f:\mathbb{R}^n \times [0,\ T] \to \mathbb{R}^n$, $g:\mathbb{R}^n \times [0,\ T] \to \mathbb{R}^{n\times m}$ and $h:\mathbb{R}^n \times [0,\ T] \to \mathbb{R}^{n\times r}$. In (1), $x := (x^{u\top}, x^{c\top})^\top \in \mathbb{R}^n$ represents the state, which is supposed to be divided by the directly noise-driven state $x^c \in \mathbb{R}^{n_c}$ and the other one $x^u \in \mathbb{R}^{n-n_c}$. $w(t) \in \mathbb{R}^r$ denotes a standard Wiener process defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, such that

$$E^{\mathcal{P}}\{\mathrm{d}w(t)\mathrm{d}w(t)^\top\} = \mathrm{d}tI_r, \quad \forall t \in [0, T], \qquad (2)$$

where I_r represents the $r \times r$ identity matrix, and $E^{\mathcal{P}}\{\cdot\}$ denotes the expectation with respect to the probability measure $\mathcal{P}$. $\mathcal{F}$ is a sigma algebra of the observable random events and a filtration $\{\mathcal{F}_t\}$ represents an increasing family of σ-algebras with $\mathcal{F}_t \subset \mathcal{F}, \forall t \in [0, T]$. We suppose that $\{\mathcal{F}_t\}$ is right-continuous and complete. For the function $h(x,\ t)$, the following assumption is imposed:

Assumption 1. For all $x \in \mathbb{R}^n$ and $t \in [0,\ T]$, the matrix $h^c(x,\ t)h^c(x,\ t)^\top$ is positive definite.

Remark 1. The division of the state x into x^u and x^c in (1) enables one to relax the assumption of the positive definiteness of $h(x,\ t)h(x,\ t)^T$ to that of $h^c(x,\ t)h^c(x,\ t)^\top$.

Remark 2. Assumption 1 implies that the dimension of the noise r is more than that of the directly noise-driven state n_c. When Assumption 1 fails, a possible solution is to intentionally add an artificial noise channel so that $h^c(x,\ t)h^c(x,\ t)^\top$ becomes positive definite.

The control input is denoted by $u(t) \in U \subset \mathbb{R}^m$, where U is specified as a compact set of admissible controls. The saturation function $\mathrm{sat}:\mathbb{R}^m \to \mathbb{R}^m$ is defined as

$$\begin{aligned}\mathrm{sat}(u(t)) &= (\mathrm{sat}_1(u_1(t)), \ldots, \mathrm{sat}_m(u_m(t)))^\top,\\ \mathrm{sat}_j(u_j(t)) &:= \begin{cases}\overline{u}_j & (\overline{u}_j \le u_j(t))\\ u_j(t) & (\underline{u}_j < u_j(t) < \overline{u}_j)\\ \underline{u}_j & (u_j(t) \le \underline{u}_j)\end{cases}, \qquad (3)\end{aligned}$$

where constants $\overline{u}_j \ge 0$ and $\underline{u}_j \le 0$. $(j = 1, \ldots, m)$ represent the upper and lower limits. Note that the saturation function defined in (3) is not differentiable at $u_j = \overline{u}_j$ and $u_j = \underline{u}_j$. This paper assumes the existence and uniqueness of a strong solution to the system (1) on $[0,\ T]$. For a concrete sufficient condition for the existence and uniqueness of the strong solution to a stochastic differential equation with nonsmooth coefficients, see, e.g., [36]. Next, we consider the following functional to define the cost function to be minimized:

$$\Gamma(x_t, u_{t:T}, t) = E^{\mathcal{P}}\left\{\phi(x(T)) + \int_t^T V(x,\tau) + \frac{1}{2}u(\tau)^\top R(x,\tau)u(\tau)\mathrm{d}\tau \;\mid\; x(t) = x_t\right\}, \tag{4}$$

where sufficiently differentiable non-negative functions ϕ and V denote the terminal cost and the instantaneous cost, respectively. V is also supposed to be integrable. $R(x, t)$ represents the weight with respect to the control cost. According to Ref. [7], we impose the following assumption to $R(x, t)$, which is necessary to derive the SHJB equation with input saturation:

Assumption 2. The input weighting matrix $R(x, t)$ in (4) is not only positive definite but also diagonal, which is represented by

$$R(x,t) = \mathrm{diag}\{R_1(x,t), \ldots, R_m(x,t)\},$$

where $R_i(x, t)$'s $(i = 1, \ldots, m)$ are positive for all $x \in \mathbb{R}^n$ and $t \in [0, T]$.

Then, the nonlinear stochastic optimal control with input saturation constraints aims at finding the optimal controller $u = u^*(x, t)$ for the system (1) such that the cost function $\Gamma(x_0, u_{0:T}, 0)$ is minimized under the input saturation.

Now, we provide the following proposition, which characterizes the SHJB equation associated with this problem:

Proposition 1. Consider the system (1) and the cost function (4). Suppose that Assumptions 1 and 2 hold. Moreover, suppose that the value function defined by

$$J(x,t) := \inf_{u_{t:T}} \Gamma(x, u_{t:T}, t) \tag{5}$$

exists and is $C^{2,1}$ function. Then, the associated SHJB equation to this problem is derived as

$$\frac{\partial J}{\partial t} + V - \frac{1}{2}\frac{\partial J}{\partial x} g R^{-1} g^\top \frac{\partial J}{\partial x}^\top + \frac{\partial J}{\partial x} f + \frac{1}{2}\mathrm{tr}\left\{\frac{\partial^2 J}{\partial x^2} hh^\top\right\} + \Psi_{\mathrm{sat}}(J) = 0,\;\; J(x,T) = \phi(x), \tag{6}$$

where the specific term $\Psi_{\mathrm{sat}}(J)$ resulting from the input saturation is given by

$$\Psi_{\mathrm{sat}}(J) := \frac{1}{2}\left(\mathrm{sat}\left(-R^{-1}g^\top \frac{\partial J}{\partial x}^\top\right) + R^{-1}g^\top \frac{\partial J}{\partial x}^\top\right)^\top \times R\left(\mathrm{sat}\left(-R^{-1}g^\top \frac{\partial J}{\partial x}^\top\right) + R^{-1}g^\top \frac{\partial J}{\partial x}^\top\right). \tag{7}$$

Further, the optimal control for this problem is given by

$$u^*(x,t) = \mathrm{sat}\left(-R(x,t)^{-1} g(x,t)^\top \frac{\partial J(x,t)}{\partial x}^\top\right). \tag{8}$$

Proof. According to Bellman's principle of optimality [25,26,37], the value function $J(x, t)$ should satisfy

$$-\frac{\partial J}{\partial t} = \inf_u \left\{V + \frac{1}{2}u^\top R u + \frac{\partial J}{\partial x}(f + g\,\mathrm{sat}(u)) + \frac{1}{2}\mathrm{tr}\left\{\frac{\partial^2 J}{\partial x^2} hh^\top\right\}\right\}, \quad J(x,T) = \phi(x). \tag{9}$$

Let us explicitly write the optimal controller u^*, which minimizes the right hand side of (9). From the definition of the saturation function (3) and Assumption 2, for any $j \in \{1, \ldots, m\}$ we have

$$\frac{1}{2}u_j^2 R_j + \frac{\partial J}{\partial x}[g]_{:,j}\mathrm{sat}_j(u_j) = \begin{cases} \frac{1}{2}u_j^2 R_j + \frac{\partial J}{\partial x}[g]_{:,j}\overline{u}_j & (\overline{u}_j \le u_j) \\ \frac{1}{2}u_j^2 R_j + \frac{\partial J}{\partial x}[g]_{:,j}u_j & (\underline{u}_j < u_j < \overline{u}_j) \\ \frac{1}{2}u_j^2 R_j + \frac{\partial J}{\partial x}[g]_{:,j}\underline{u}_j & (u_j \le \underline{u}_j) \end{cases},$$

where the notation $[\cdot]_{:,j}$ is utilized to denote the jth column of a matrix $[\cdot]$. Thus, we can calculate the gradient with respect to u of the inside of the parenthesis on the right hand side of (9), and by setting it to zero, the jth component of the optimal control is given by

$$u_j^* = \begin{cases} \overline{u}_j & (\overline{u}_j \le u_j) \\ -\frac{1}{R_j}[g]_{:,j}^\top \frac{\partial J}{\partial x}^\top & (\underline{u}_j < u_j < \overline{u}_j) \\ \underline{u}_j & (u_j \le \underline{u}_j) \end{cases}. \tag{10}$$

[Correction added on 13 July 2020, after first online publication: Equation 10 has been amended.].

Equation (10) can be rewritten as

$$u_j^* = \mathrm{sat}_j\left(-\frac{1}{R_j}[g]_{:,j}^\top \frac{\partial J}{\partial x}^\top\right). \tag{11}$$

Thus, from (3) and (11), we explicitly provide the optimal control u^* as (8). By substituting u^* given in (8) into (9), the SHJB equation with input saturation constraints is obtained as the Cauchy problem for a PDE given by (6) and (7). The assertion of the proposition follows.

Remark 3. In the SHJB equation given by Proposition 1, the existence of the term $\Psi_{\mathrm{sat}}(J)$ in (7) is an essential difference from the conventional SHJB equation without input saturation constraints.

Since the deterministic solution method [7] essentially depends on the geometrical structure of HJB equation, which is a first-order PDE, it is not applicable to the resultant equation (6) involving the second-order partial derivatives. Therefore, this paper provides a concrete solution method to the SHJB equation with input saturation (6) by extending ISOC-PI proposed in Ref. [22]. We modify the iteration procedure of ISOC-PI (Eq. (24) in [22]) as follows:

$$\frac{\partial J_{(i)}}{\partial t} + \frac{\partial J_{(i)}}{\partial x}\widehat{f}_{(i)} + \frac{1}{2}\mathrm{tr}\left\{\frac{\partial^2 J_{(i)}}{\partial x^2} hh^\top\right\} + \widehat{V}_{\mathrm{sat}(i)} = 0, \quad J_{(i)}(x,T) = \phi(x). \tag{12}$$

Here, the ith sample generating dynamics $\widehat{f}_{(i)}$ is the same as ISOC-PI in Ref. [22], which is given by

$$\widehat{f}_{(i)} := f - gR^{-1}g^\top \frac{\partial J_{(i-1)}}{\partial x}^\top.$$

In contrast, the ith instantaneous cost of ISOC-PI is replaced with a new one $\widehat{V}_{\mathrm{sat}(i)}$, which is defined using the notation Ψ_{sat} in (7) as

$$\widehat{V}_{\mathrm{sat}(i)} := \widehat{V}_{(i)} + \Psi_{\mathrm{sat}}(J_{(i-1)}) = V(x,t) + \frac{1}{2}\frac{\partial J_{(i-1)}}{\partial x} gR^{-1}g^\top \frac{\partial J_{(i-1)}}{\partial x}^\top + \Psi_{\mathrm{sat}}(J_{(i-1)}). \tag{13}$$

Since the proposed modified iteration procedure (12) is of the form of a Cauchy problem, to which the Feynman-Kac formula is applicable, the stochastic representation of the solution $J_{(i)}$ is given as (14). Then, by applying the path integral analysis to (14), the stochastic representation of its partial derivative is also obtained as (15). The resultant equations below correspond to a modification of Theorem 1 in Ref. [22]:

$$J_{(i)}(x,t) = E^{p_{(i)}(\xi_{t:T}|x,t)}\{\widehat{S}_{\mathrm{sat}(i)}(\xi_{t:T})\}, \tag{14}$$

$$\widehat{S}_{\mathrm{sat}(i)}(\xi_{t:T}) := \int_t^T \widehat{V}_{\mathrm{sat}(i)}(\xi(\tau),\tau)\mathrm{d}\tau + \phi(\xi(T)),$$

$$\frac{\partial J_{(i)}(x,t)}{\partial x^c}^{\top}\mathrm{d}t = (h^c(x,t)h^c(x,t)^{\top})^{-1}h^c(x,t) \times E^{p_{(i)}(\xi_{t:T}|x,t)}\{\widehat{S}_{\mathrm{sat}(i)}(\xi_{t:T})\mathrm{d}w(t)\}. \tag{15}$$

Here, $p_{(i)}(\xi_{t:T}|x,\ t)$ represents the probability that a sample path $\xi_{t:T}$ is realized under the ith sample generating dynamics:

$$\mathrm{d}\xi = \widehat{f}_{(i)}(\xi,t)\mathrm{d}t + h(\xi,t)\mathrm{d}w$$

with $\xi(t) = x$ on $[t, T]$, and $E^{p_{(i)}(\xi_{t:T}|x,t)}\{\cdot\}$ denotes the expectation over all the possible sample paths. In order to utilize the result in (15), we rewrite the input term with saturation in the system (1) as

$$\mathrm{sat}(u(t))\mathrm{d}t \equiv \widetilde{\mathrm{sat}}(u(t)\mathrm{d}t),$$

where another saturation function $\widetilde{\mathrm{sat}}$ is defined as

$$\widetilde{\mathrm{sat}}(u(t)\mathrm{d}t) = (\widetilde{\mathrm{sat}}_1(u_1(t)\mathrm{d}t),\ \ldots,\widetilde{\mathrm{sat}}_m(u_m(t)\mathrm{d}t))^{\top},$$

$$\widetilde{\mathrm{sat}}_j(u_j(t)\mathrm{d}t) := \begin{cases} \overline{u}_j\mathrm{d}t & (\overline{u}_j\mathrm{d}t \le u_j(t)\mathrm{d}t) \\ u_j(t)\mathrm{d}t & (\underline{u}_j\mathrm{d}t < u_j(t)\mathrm{d}t < \overline{u}_j\mathrm{d}t) \\ \underline{u}_j\mathrm{d}t & (u_j(t)\mathrm{d}t \le \underline{u}_j\mathrm{d}t) \end{cases}. \tag{16}$$

From (8), (15) and (16), the corresponding suboptimal controller at the ith iteration, namely $u^*_{(i)}(x,t)$, is obtained as

$$u^*_{(i)}(x,t)\mathrm{d}t = \widetilde{\mathrm{sat}}\left(-R(x,t)^{-1}g^c(x,t)^{\top}(h^c(x,t)h^c(x,t)^{\top})^{-1} \times h^c(x,t)E^{p_{(i)}(\xi_{t:T}|x,t)}\{\widehat{S}_{sat(i)}(\xi_{t:T})\mathrm{d}w(t)\}\right). \tag{17}$$

In the theoretical view point, the present stochastic representations in (14) and (15) are useful and important, since they are derived without any approximation. In the computational view point, since the expectations cannot be computed analytically in general, several statistical sampling methods can be used for efficient computation. Since as the number of samples increases, the calculation error can be arbitrarily reduced, the trade-off between the computational accuracy and cost can be considered. For a concrete computation algorithm of the solution (14), its partial derivative (15) and the suboptimal controller (17) using a basic Monte Carlo sampling, see Algorithm 1 in Ref. [22] In Ref. [22], convergence analysis of Algorithm 1 is also conducted such that the iteration procedure converges and the SHJB equation is satisfied, where computational error at each iteration due to finite samples is adequately considered. The proposed iteration procedure (12) is easily adapted to the algorithm. For more efficient computation with importance sampling, see Ref. [38].

Table 1. Physical parameters

m	Mass of the link	1.0 kg
l	Length of the link	1.0 m
l_c	Length to the center of gravity	0.50 m
I	Inertia of the link	8.3×10^{-2} kgm^2
g	Gravity acceleration	9.8 m/s^2

3. Numerical Example

This section exhibits a numerical example to demonstrate the validity of the proposed method. Here, we consider a one-link robot manipulator moving on a vertical plane. The joint angle of the link is denoted by θ, and the control torque applied to the joint is denoted by u, respectively. The physical parameters of this apparatus are summarized in Table 1. We aim at the optimal swing-up control of the manipulator with input saturation rigorously satisfied. We consider the dynamics of the robot with the joint torque saturation of the form (1) as

$$\mathrm{d}x = \begin{pmatrix} x_2 \\ \frac{ml_cg\sin x_1}{ml_c^2+I} \end{pmatrix}\mathrm{d}t + \begin{pmatrix} 0 \\ \frac{1}{ml_c^2+I} \end{pmatrix}\mathrm{sat}(u)\mathrm{d}t + \begin{pmatrix} 0_{1\times 2} \\ \frac{1}{ml_c^2+I}(1-0.1x_2) \end{pmatrix}\mathrm{d}w, \tag{18}$$

where the state is defined by $x := (\theta, \dot{\theta})^{\top}$. In the dynamics model (18), persistent system noise and uncertainty in viscous friction are supposed. The upper and lower limits of the saturation function (3) are chosen as $\overline{u} = 0.5$ N and $\underline{u} = -0.5$ N, respectively. Regarding to the control objective, the cost function (4) is set as $\phi(x(T)) = 1/2x(T)\,\mathrm{diag}\{10, 0.5\}x(T)$, $V\ (x,\tau) = 1/2x(\tau)^{\top}\mathrm{diag}\{5, 0.3\}x(\tau)^{\top}$ and $R = 0.8$. The terminal time is $T = 2.0$ s, the initial condition is $x(0) = (-\pi,\ 0)^{\top}$, which represents the pendant position.

To solve the formulated nonlinear stochastic optimal control problem, we execute the proposed iteration procedure (12) based on Algorithm 1 in Ref. [22]. To generate the initial function for the iteration, we solve the LQG problem for the linearized system of (18) around the origin with the same cost function, and obtain a value function J_{LQG} and a corresponding LQG controller u_{LQG}. Then, we set the initial function as $J_{(0)}(x,\ t) = J_{LQG}(x,\ t)$. From this choice, we will easily show that the resultant controller has a better performance than the LQG one. We execute nine iterations of the proposed method, and this set of iterations is referred to as one optimization trial. We repeat 50 optimization trials. The total sample number using Monte Carlo at each iteration is $N_{(i)} = 30\,000$ ($i = 1,\ \ldots,\ 9$). Not only with LQG method, we also compare the behaviors and performance of the proposed method with those of the conventional ISOC-PI method [22].

The simulation results of the proposed method are shown in Figs 1 to 3. The thin line with error bars in Fig. 1 shows the mean and standard deviation of the cost function of 50 optimization trials for 9 iterations. This figure also exhibits the history of the realized costs in one of the 50 trials in the thick line. From Fig. 1, the cost function decreases and roughly converges. Figures 2 and 3 show the results before the iteration and those at the last iteration, respectively. The top, middle and bottom figures of each figure show the time responses of θ, $\dot{\theta}$ and u, respectively. Two horizontal lines in the each bottom figure represent the input

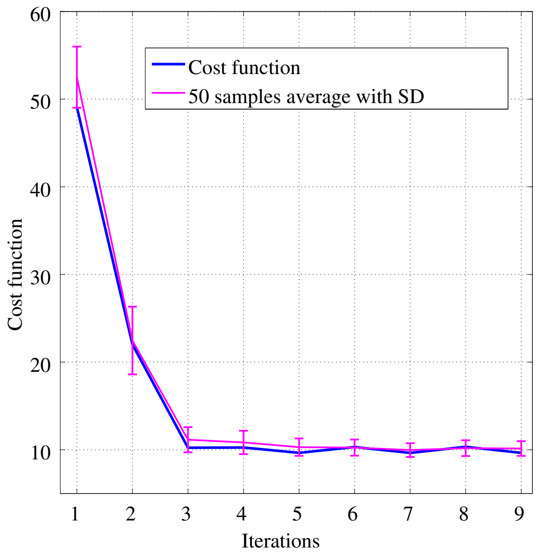


Fig 1. Mean and standard deviation of the cost function of 50 optimization trials along the iteration (thin line with error bars), and the history of the realized cost in one of the 50 trials (thick line)

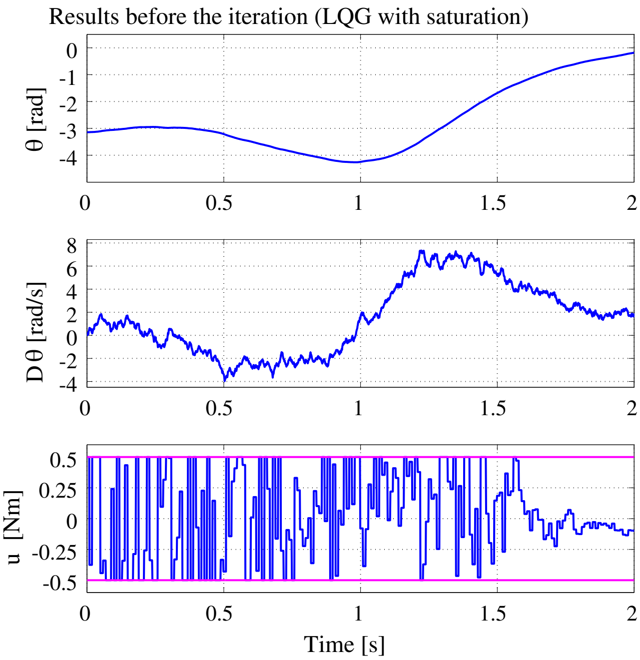


Fig 2. Time responses of θ, $\dot{\theta}$ and the initial controller u as the truncated LQG controller due to saturation

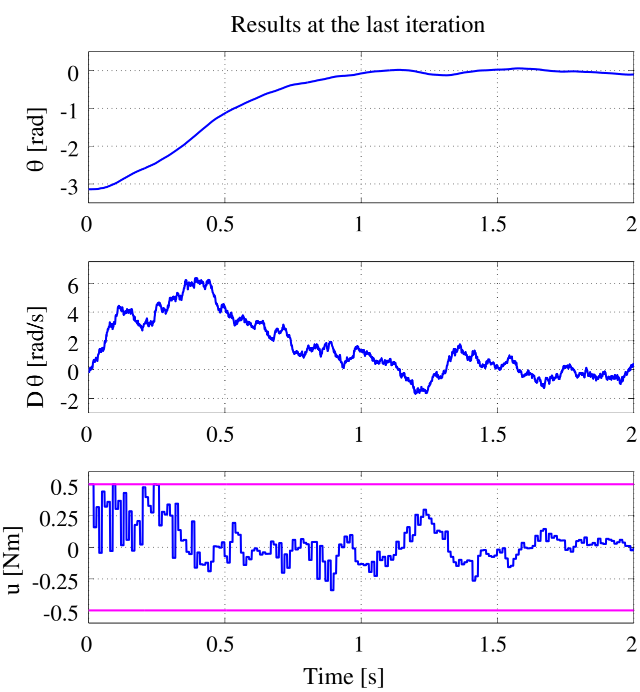


Fig 3. Time responses of θ, $\dot{\theta}$ and u of the proposed controller at the last iteration

saturation. Since we utilized the LQG controller as the initial controller of the iteration, the results shown in Fig. 2 imply the performance of the truncated LQG control due to saturation. On the contrary, the results in Fig. 3 show the performance of the proposed nonlinear stochastic optimal controller with rigorously considering input saturation. From the comparison between Figs 2 and 3, the proposed controller has a better performance than the LQG controller in that the joint angle θ reaches the desired position faster and remains nearby for a longer period, and moreover the amount of the control input u is smaller.

In order to demonstrate the effectiveness of considering the input saturation, we also executed the conventional ISOC-PI method[22] under the same conditions except for taking the input constraint into account, whose results are shown in Figs 4 to 6. From those figures, the cost function decreases and roughly converges, and a better performance than the LQG controller is achieved as well in the case of the proposed method. However, compared to the present method, Figs 1 and 4 imply that ISOC-PI has tendency that the mean and standard deviation of the costs are bigger than the present method. This is because the control inputs generated by ISOC-PI tend to be larger, since there is no input saturation. Although ISOC-PI eventually achieves a similar performance to the present method, it cannot guarantee that the input saturation is satisfied. Particularly, there are still a few moments when the input exceeds the limit in Fig. 6. Those results exhibit the validity of the proposed method.

4. Conclusion

This paper considers a finite-time nonlinear stochastic optimal control problem with input saturation as a hard constraint on the control input. The associated SHJB equation has been first derived. Moreover, we have provided a concrete solution method to this problem by extending our previous result of ISOC-PI in [22]. The proposed method can iteratively obtain stochastic representations for both the solution to the SHJB equation and the corresponding optimal control by utilizing path integral analysis. This avoids using numerical differentiation in deriving the optimal control from partial derivatives of the solution to the SHJB equation. From numerical simulations of comparisons of the proposed method with LQG and ISOC-PI [22], both of which cannot deal with input saturation, it has been demonstrated that the proposed

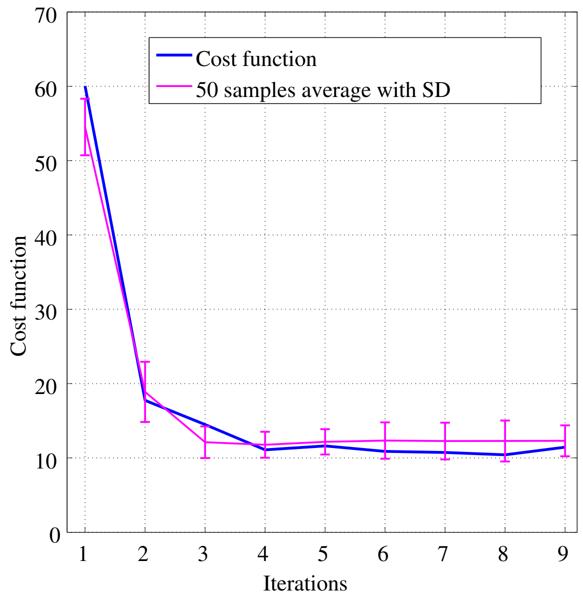


Fig 4. Mean and standard deviation of the cost function of 50 optimization trials along the iteration of ISOC-PI[22] (thin line with error bars), and the history of the realized cost in one of the 50 trials (thick line)

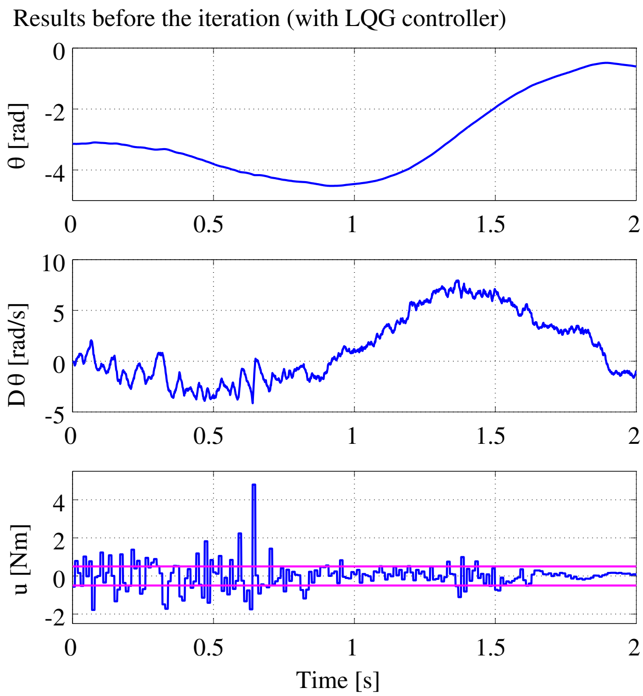


Fig 5. The results before the iteration of ISOC-PI[22] with LQG controller

method achieves better performance with the input saturation constraint rigorously satisfied. This paper is the first one that provides an optimal feedback control rigorously satisfying the saturation constraint as a solution to the corresponding SHJB equation.

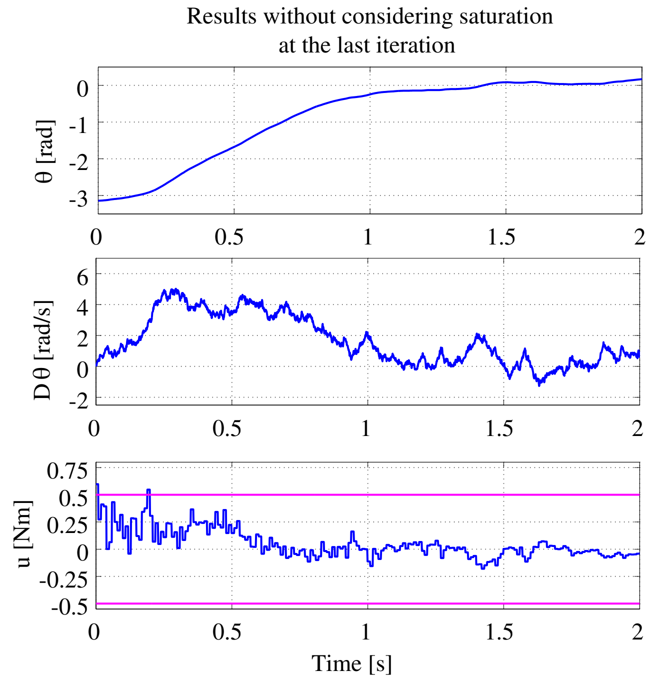


Fig 6. The last iteration results of ISOC-PI [22]

Acknowledgment

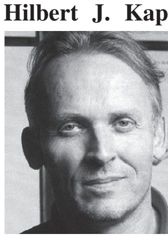
This work was supported by Japan Society for the Promotion of Science KAKENHI Grant Numbers JP17H03282 and JP18K04202.

References

(1) Hu T, Lin Z. *Control Systems with Actuator Saturation: Analysis and Design*. Birkhäuser: Boston, MA; 2001.
(2) Ding S, Zheng WX. Nonsingular terminal sliding mode control of nonlinear second-order systems with input saturation. *International Journal of Robust and Nonlinear Control* 2016; **26**(9):1857–1872.
(3) Shahri ESA, Alfi A, Machado JAT. Stability analysis of a class of nonlinear fractional-order systems under control input saturation. *International Journal of Robust and Nonlinear Control* 2018; **28**(7):2887–2905.
(4) Rehan M, Iqbal N, Hong KS. Delay-range-dependent control of nonlinear time-delay systems under input saturation. *International Journal of Robust and Nonlinear Control* 2015; **26**(8):1647–1666.
(5) Chen M, Ge SS, Ren B. Adaptive tracking control of uncertain MIMO nonlinear systems with input constraints. *Automatica* 2011; **47**(3):452–465.
(6) Wang Q, Wu Z, Shi P, Xue A. Robust control for switched systems subject to input saturation and parametric uncertainties. *Journal of Franklin Institute* 2017; **354**(16):7266–7279.
(7) Fujimoto R, Sakamoto N. The stable manifold approach for optimal swing up and stabilization of an inverted pendulum with input saturation. In *Proceedings of 18th IFAC World Congress*; 2011; 8046–8051.
(8) Mayne DQ, Rawlings JB, Rao CV, Scokaert POM. Constrained model predictive control: Stability and optimality. *Automatica* 2000; **36**(6):789–814.
(9) Rossiter JA. *Model-Based Predictive Control: A Practical Approach*. CRC Press: Boca Raton, FL; 2003.
(10) Bemporad A, Morari M. Robust model predictive control: a survey. In *Robustness in Identification and Control*. Garulli A, Tesi A (eds). Springer: London; 1999; 207–226.

(11) Itô K. On a formula concerning stochastic differentials. *Nagoya Mathematical Journal* 1951; **3**:55–65.
(12) Gihman I, Skorohod A. *Stochastic Differential Equations*. Berlin Heidelberg: Springer-Verlag; 1972.
(13) Ikeda N, Watanabe S. *Stochastic Differential Equations and Diffusion Processes*. 2nd ed.: Amsterdam/Tokyo, North-Holland/Kodansha: Kodansha; 1989.
(14) Øksendal B. *Stochastic Differential Equations, an Introduction with Applications*. 5th ed. Springer-Verlag: Berlin Heidelberg New York; 1998.
(15) Kushner HJ. *Stochastic Stability and Control*. New York: Academic Press; 1967.
(16) Has'minskii RZ. *Stochastic Stability of Differential Equations*. Sijthoff and Noordhoff: Maryland; 1980.
(17) Mao X. *Stability of Stochastic Differential Equations with Respect to Semimartingales (Pitman Research Notes Mathematics Series 251)*. New York: Longman Scientific and Technical; 1991.
(18) Satoh S, Fujimoto K. Passivity based control of stochastic port-Hamiltonian systems. *IEEE Transactions on Automatic Control* 2013; **58**(5):1139–1153.
(19) Satoh S, Saeki M. Bounded stabilization of stochastic port-Hamiltonian systems. *International Journal of Control* 2014; **87**(8):1573–1582.
(20) Yamauchi J, Nakano S, Satoh S, Hatanaka T, Fujita M. Stochastic performance analysis of visual motion observer and experimental verifications. In *Proceedings of 10th Asian Control Conference*; 2015; 797–802.
(21) Satoh S, Saeki M. A unified approach to nonlinear stochastic control based on path integral analysis. In *Proceedings of SICE International Symposium on Control Systems 2015*, (USB)712–1; 2015.
(22) Satoh S, Kappen HJ, Saeki M. An iterative method for nonlinear stochastic optimal control based on path integrals. *IEEE Transactions on Automatic Control* 2017; **62**(1):262–276.
(23) Kappen HJ. Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment* 2005; **2005**:P11011.
(24) Kappen HJ. An introduction to stochastic control theory, path integrals and reinforcement learning. In *Proceedings of 9th Granada Seminar on Computational Physics: Cooperative Behavior in Neural Systems*; 2007; 149–181.
(25) Yong J, Zhou XY. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer-Verlag: New York, NY; 1999.
(26) Pham H. *Continuous-Time Stochastic Control and Optimization with Financial Applications*. Berlin Heidelberg: Springer-Verlag; 2009.
(27) Jacobson D, Mayne D. *Differential Dynamic Programming*. Elsevier Publ. Co., Inc: New York; 1970.
(28) Li W, Todorov E. Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system. *International Journal of Control* 2007; **80**(9):1439–1453.
(29) Theodorou E, Tassa Y, Todorov E. Stochastic differential dynamic programming. In *Proceedings of American Control Conference*; 2010; 1125–1132.
(30) Satoh S, Kamiike T, Saeki M. Nonlinear stochastic H-infinity control using local quadratic approximation. In *Proceedings of SICE International Symposium on Control Systems 2016*, (USB)2A1-5; 2016.
(31) Sun W, Theodorou EA, Tsiotras P. Stochastic game theoretic trajectory optimization in continuous time. In *Proceedings of 55th IEEE Conerence on Decision and Control*; 2016; 6167–6172.
(32) Siva E, Goulart P, Maciejowski J, Kantas N. Stability of model predictive control using Markov Chain Monte Carlo optimisation. In *Proceedings of European Control Conference*; 2009; 2851–2856.
(33) Stahl D, Hauth J. Pf-mpc: Particle filter-model predictive control. *Systems & Control Letters* 2011; **60**(8):632–643.
(34) Mesbah A. Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems Magazine* 2016; **36**(6):30–344.
(35) Lorenzen M, Dabbene F, Tempo R, Allgöwer F. Constraint-tightening and stability in stochastic model predictive control. *IEEE Transactions on Automatic Control* 2017; **62**(7):3165–3177.
(36) Bahlali K, Mezerdi B, Ouknine Y. The maximum principle for optimal control of diffusions with non-smooth coefficients. *Stochastics and Stochastics Reports* 1996; **57**:303–316.
(37) Fleming WH, Soner HM. *Controlled Markov Processes and Viscosity Solutions*. 2nd ed. Springer-Verlag: New York, New York; 2006.
(38) Gómez V, Thijssen S, Symington A, Hailes S, Kappen HJ. Real-time stochastic optimal control for multi-agent quadrotor swarms. In *Proceedings of 26th International Conference on Automated Planning and Scheduling*; 2016; 468–476.

Satoshi Satoh (Member) received the B.Sc., M.Sc. and Ph.D. degrees in engineering from Nagoya University, Japan, in 2005, 2007 and 2010, respectively. He is currently an associate professor of Graduate School of Engineering, Osaka University, Japan. From 2010 to 2017, he was an assistant professor of Faculty of Engineering, Hiroshima University, Japan. He held a visiting research position at Radboud University Nijmegen, The Netherlands in 2011. His research interests include nonlinear deterministic and stochastic control problems and their application to robotics and space engineering. He received the second International Symposium on Swarm Behavior and Bioinspired Robotics Best Paper Award in 2017, the 10th Asian Control Conference the Best Paper Award in 2015, SICE Young Author's Award in 2010, and IEEE Robotics and Automation Society Japan Chapter Young Award in 2008.



Hilbert J. Kappen (Non-member) received the Ph.D. degree in particle physics in 1987 from Rockefeller University, New York, NY. From 1987 until 1989, he was a Scientist at the Philips Research Laboratories, Eindhoven, The Netherlands. Presently, he is Professor of physics at SNN University of Nijmegen, The Netherlands, conducting research in machine learning and computational neuroscience. He is author of over 150 publications.