International Journal of Innovative Computing, Information and Control
Volume **5**, Number **4**, April **2009**

ICIC International ©2009 ISSN 1349-4198
pp. 1–07-306

ON A CLASS OF STOCHASTIC IMPULSIVE OPTIMAL PARAMETER SELECTION PROBLEMS

Chun Min Liu[1], Zhi Guo Feng[2] and Kok Lay Teo[1]

[1]Department of Mathematics and Statistics
Curtin University of Technology
Perth, 6845, Australia
chun-min.liu@postgrad.curtin.edu.au; k.l.teo@curtin.edu.au

[2]College of Mathematics and Computer Science
Chongqing Normal University
Chongqing, 400047, P.R.C.
z.feng.scholar@gmail.com

Received November 2007; revised March 2008

Abstract. *This paper considers a class of stochastic optimal parameter selection problems described by linear Ito stochastic differential equations with state jumps subject to probabilistic constraints on the state, where the times at which the jumps occurred as well as their heights are decision variables. We show that this constrained stochastic impulsive optimal parameter selection problem is equivalent to a deterministic impulsive optimal parameter selection problem subject to continuous state inequality constraints, where the times at which the jumps occurred as well as their heights remain as decision variables. Then, by introducing a time scaling transform, we show that this constrained deterministic impulsive optimal parameter selection problem is transformed into an equivalent constrained deterministic impulsive optimal parameter selection problem with fixed jump times. A constraint transcription technique is then used to approximate the continuous state inequality constraints by a sequence of canonical inequality constraints. This leads to a sequence of approximate deterministic impulsive optimal parameter selection problems subject to canonical inequality constraints. For each of these approximate problems, we derive the gradient formulas of the cost function and the constraint functions. On this basis, an efficient computational method is developed.*
Keywords: Stochastic impulsive optimal parameter selection problem, Deterministic impulsive optimal parameter selection problem, Probabilistic constraints, Time scaling transformation, Constraint transcription technique, Canonical inequality constraints, Gradient based optimization technique

1. Introduction.

Basic theory of Ito stochastic differential equations driven by Wiener processes and counting processes (for example Poisson processes) and their many important applications can be found in [2], [5] and [14]. In [16], a class of optimal control problems described by linear Ito stochastic differential equations driven by counting processes is considered and studied. In [19], a class of stochastic optimal control problems is considered, where the dynamical system is described by Ito stochastic differential equations driven by Wiener processes. It is shown that this class of stochastic optimal control problems is equivalent to a class of optimal control problems involving linear parabolic partial differential equations. However, numerical solution methods available in the literature (see, for example, [10] and [23]) for solving such deterministic optimal control problems with dynamics being described by partial differential equations are only applicable for small dimensional problems.

The optimal parameter selection problems occur in many dynamic optimization models where the controls are restricted to be constant functions of time. Examples of these can be found in a number of parameter identification problems, controller parameter design problems, as well as economic and industrial management type problems. Furthermore, it plays a fundamental role in the numerical computation of optimal control problems. To be more specific, after the control parameterization (see [22]), all optimal control problems essentially reduced to optimal parameter selection problems. Thus, the solvability of optimal parameter selection problem is crucial for generating numerical solution methods to many complex optimal control problems. In [1] and [19], respective necessary conditions for optimality are derived for deterministic and stochastic optimal parameter selection problems. Computational methods for solving deterministic optimal parameter selection problems are reported in [4], [15] and [21], where the system dynamic is described by ordinary differential equations in [4] and [21], while the system dynamic is described by parabolic partial differential equation in [15].

Optimal filtering problems and optimal fusion problems can be formulated as specific stochastic optimal control problems. These problems have been extensively studied in the literature, see for example, [3], [6], [17] and [18]. In particular, an optimal fusion problem is considered in [6], where the measurement data are obtained from multiple sensors. It is shown that this optimal fusion problem is equivalent to a deterministic optimal control problem. Optimal filtering problems with multiple sensors are also considered in [7] and [8].

In [9], a class of optimal parameter selection problems governed by a linear Ito stochastic differential equation is considered. The aim is to minimize the expected value of the cost function subject to some probabilistic constraints on the state. It is first shown that this problem is equivalent to a deterministic optimal parameter selection problem subject to continuous state inequality constraints. The continuous state inequality constraints are then transformed into equivalent equality constraints using a constraint transcription given in [20]. However, as pointed out in Remark 6.6.5 of [22], the equality constraints obtained by this constraint transcription fail to satisfy any constraint qualification. Thus, constraint violation cannot be avoided in numerical computation if this constraint transcription is used.

In this paper, we consider a general class of stochastic optimal control problems, where the system dynamics are described by linear Ito stochastic differential equations with state jumps occurring at various time points. Many natural and man-made systems do exhibit the phenomenon of jumps occurring at various time points along their trajectories. Examples include drug administration in cancer chemotherapy, insulin injection, and native forest ecosystems management, just to name a few. This problem covers the one considered in [9] as a very special case for which no state jumps are allowed.

The rest of the paper is organized as follows. In Section 2, we formulate the stochastic impulsive optimal parameter selection problem subject to probabilistic constraints. In Section 3, we show that this problem with probabilistic constraints is equivalent to a deterministic impulsive optimal parameter selection problem subject to continuous state inequality constraints. In Section 4, a time scaling transform (see [13]) is applied to map the variable jump times into pre-fixed jump times in a new time scale. A constraint transcription technique reported in [22] is used in Section 5 to approximate the continuous state inequality constraints as a sequence of inequality constraints. This leads to a sequence of approximate deterministic impulsive optimal parameter selection problems subject to inequality constraints. For each of these approximate problems, we derive the gradient formulas of the cost function and the constraint functions. On this basis, an efficient computational method is developed for solving each of these approximate problems. For illustration, an example is solved using the proposed method in Section 6.

2. **Problem Statement.** Consider an impulsive dynamical system described by linear Ito stochastic differential equations defined on a fixed time interval $(0,T]$.

$$d\boldsymbol{\xi}(t) = \boldsymbol{A}(t,\boldsymbol{\delta})\boldsymbol{\xi}(t)dt + \boldsymbol{B}(t,\boldsymbol{\delta})dt + \boldsymbol{D}(t,\boldsymbol{\delta})d\boldsymbol{w}(t) \tag{2.1.a}$$

$$\boldsymbol{\xi}(0) = \boldsymbol{\xi}^0 \tag{2.1.b}$$

$$\boldsymbol{\xi}(\tau_i^+) = \boldsymbol{J}^i\boldsymbol{\xi}(\tau_i^-) + \boldsymbol{\Delta}_i + \boldsymbol{\gamma}^i, \quad i = 1,\dots,m. \tag{2.1.c}$$

Here, $\boldsymbol{\xi}(t) = [\xi_1(t),\dots,\xi_n(t)]^\top \in \mathbb{R}^n$ is the state vector; $\boldsymbol{\delta} = [\delta_1,\dots,\delta_r]^\top \in \mathbb{R}^r$ is the system parameter vector; $\boldsymbol{\xi}^0 = [\xi_1^0,\dots,\xi_n^0]^\top \in \mathbb{R}^n$ is the initial state vector which is Gaussian with mean $\boldsymbol{\mu}^0$ and covariance matrix $\boldsymbol{\Psi}^0$; and $\boldsymbol{w}(t) = [w_1(t),\dots,w_d(t)]^\top \in \mathbb{R}^d$ is a Wiener process with mean $\mathbf{0}$ and covariance matrix given by

$$\mathcal{E}\{\boldsymbol{w}(t_1)(\boldsymbol{w}(t_2))^\top\} = \int_0^{\min\{t_1,t_2\}} \boldsymbol{I}ds,$$

where $\mathcal{E}\{\cdot\}$ denotes the mathematical expectation and $\boldsymbol{I}$ is the identity matrix. Equations (2.1.c) are conditions on the state jumps, where $\boldsymbol{J}^i \in \mathbb{R}^{n\times n}$, $i = 1,\dots,m$, are given coefficient matrices, $\tau_1,\dots,\tau_m$, are the time points at which the state jumps are occurred, $\boldsymbol{\Delta}_i$, $i = 1,\dots,m$, are Gaussian vectors with mean $\mathbf{0}$ and covariance matrices $\boldsymbol{K}^i$, $i = 1,\dots,m$, and $\boldsymbol{\gamma}^i = [\gamma_1^i,\dots,\gamma_n^i]^\top$, $i = 1,\dots,m$, are the magnitude vectors of the jumps. Let $\boldsymbol{\tau} = [\tau_1,\dots,\tau_m]^\top$.

We assume that the following conditions are satisfied.

(i). $\boldsymbol{A}(t,\boldsymbol{\delta}) \in \mathbb{R}^{n\times n}$, $\boldsymbol{B}(t,\boldsymbol{\delta}) \in \mathbb{R}^n$ and $\boldsymbol{D}(t,\boldsymbol{\delta}) \in \mathbb{R}^{n\times d}$ are continuously differentiable with respect to all their arguments.

(ii). The Wiener process $\boldsymbol{w}(t)$ and the random vectors $\boldsymbol{\xi}^0$, $\boldsymbol{\Delta}_i$, $i = 1,\dots,m$, are mutually independent.

The probabilistic state constraints given below arise naturally when the state is required to stay within a given acceptable region with a given degree of confidence for all $t \in [0,T]$.

$$Prob\{a_k \le (\boldsymbol{c}^k)^\top\boldsymbol{\xi}(t) \le b_k\} > \alpha_k, \ \ \forall t \in [0,T], \ \ k = 1,\dots,N, \tag{2.2}$$

where $\boldsymbol{c}^k$, $k = 1,\dots,N$, are $n-$vectors, and a_k, b_k, α_k, $k = 1,\dots,N$, are real constants.

Define

$$\boldsymbol{\Omega} = \{\boldsymbol{\delta} \in \mathbb{R}^r : h_j(\boldsymbol{\delta}) \le 0, \quad j = 1,\dots,M\}, \tag{2.3}$$

where h_j, $j = 1,\dots,M$, are continuously differentiable functions of the parameter $\boldsymbol{\delta}$. Let $\boldsymbol{h} = [h_1,\dots,h_M]^\top$.

For the jump time vector $\boldsymbol{\tau} = [\tau_1,...,\tau_m]^\top$, it is assumed, without loss of generality, that

$$0 < \tau_1 < \cdots < \tau_m < T. \tag{2.4}$$

Let $\boldsymbol{\mathcal{T}}$ be the set of all those $\boldsymbol{\tau} = [\tau_1,\dots,\tau_m]^\top$ which satisfy (2.4). For brevity in notation, we denote $\tau_0 = 0$ and $\tau_{m+1} = T$.

Let $\boldsymbol{\Gamma}$ be the set of all those magnitude vectors $\boldsymbol{\gamma} = [(\boldsymbol{\gamma}^1)^\top,\dots,(\boldsymbol{\gamma}^m)^\top]^\top$ such that

$$\underline{\gamma}_j^i \le \gamma_j^i \le \overline{\gamma}_j^i, \ i = 1,\dots,m; \ \ j = 1,\dots,n. \tag{2.5}$$

An element $(\boldsymbol{\delta},\boldsymbol{\tau},\boldsymbol{\gamma}) \in \boldsymbol{\Omega}\times\boldsymbol{\mathcal{T}}\times\boldsymbol{\Gamma}$ is said to be a feasible parameter vector if it satisfies the probabilistic state constraints (2.2). Let $\boldsymbol{\mathcal{D}}$ be the class of all such feasible parameter vectors. We may now state our problem formally as follows.

Problem (P1)**.** *Given the dynamical system (2.1), find a feasible parameter vector* $(\boldsymbol{\delta}, \boldsymbol{\tau}, \boldsymbol{\gamma}) \in \boldsymbol{\mathcal{D}}$*, such that the cost function*

$$g_0(\boldsymbol{\delta}, \boldsymbol{\tau}, \boldsymbol{\gamma}) = \varphi(\boldsymbol{\gamma}) + \mathcal{E}\{(\boldsymbol{\xi}(T))^{\intercal}\boldsymbol{S}_2(\boldsymbol{\delta})\boldsymbol{\xi}(T) + (\boldsymbol{S}_1(\boldsymbol{\delta}))^{\intercal}\boldsymbol{\xi}(T) + \boldsymbol{S}_0(\boldsymbol{\delta}) + \sum_{i=1}^{m+1}\int_{\tau_{i-1}}^{\tau_i}[(\boldsymbol{\xi}(t))^{\intercal}\boldsymbol{Q}_2(t,\boldsymbol{\delta})\boldsymbol{\xi}(t) + (\boldsymbol{Q}_1(t,\boldsymbol{\delta})^{\intercal}\boldsymbol{\xi}(t) + \boldsymbol{Q}_0(t,\boldsymbol{\delta})]dt\} \quad (2.6)$$

is minimized, where $\varphi(\boldsymbol{\gamma})$ *is a penalty term to prevent high jumps, and* $\boldsymbol{S}_2(\boldsymbol{\delta}) \in \mathbb{R}^{n\times n}$ *and* $\boldsymbol{Q}_2(t,\boldsymbol{\delta}) \in \mathbb{R}^{n\times n}$ *are positive semi-definite matrices which are continuously differentiable with respect to their respective arguments, while* $\boldsymbol{S}_1(\boldsymbol{\delta})$ *and* $\boldsymbol{Q}_1(t,\boldsymbol{\delta})$ *(respectively,* $\boldsymbol{S}_0(\boldsymbol{\delta})$ *and* $\boldsymbol{Q}_0(t,\boldsymbol{\delta})$*) are* $n-$*vector valued functions (respectively, real-valued functions) which are also continuously differentiable with respect to their respective arguments.*

Problem (P1) is a stochastic impulsive optimal parameter selection problem with probabilistic constraints. We shall show that it is equivalent to a deterministic optimal parameter selection problem subject to continuous state inequality constraints. A numerical computational method will be developed for solving this equivalent constrained deterministic optimal parameter selection problem. This is to be done in several stages as detailed below.

3. **Deterministic Transformation.** In this section, we shall show that the stochastic impulsive optimal parameter selection problem (P1) can be transformed into a deterministic impulsive optimal parameter selection problem.

For each $\boldsymbol{\delta}$, it is clear from (2.1) that the solution of system (2.1), for $t \in (\tau_{i-1}, \tau_i)$ with $i = 1, \ldots, m$, is given by

$$\boldsymbol{\xi}(t \mid \boldsymbol{\delta}) = \boldsymbol{\Phi}(t, \tau_{i-1} \mid \boldsymbol{\delta})\boldsymbol{\xi}(\tau_{i-1}^+) + \int_{\tau_{i-1}}^{t}\boldsymbol{\Phi}(t,s \mid \boldsymbol{\delta})\boldsymbol{B}(s,\boldsymbol{\delta})ds + \int_{\tau_{i-1}}^{t}\boldsymbol{\Phi}(t,s \mid \boldsymbol{\delta})\boldsymbol{D}(s,\boldsymbol{\delta})d\boldsymbol{w}(s), \quad (3.1)$$

where $\boldsymbol{\Phi}(t,s \mid \boldsymbol{\delta}) \in \mathbb{R}^{n\times n}$ is the principal solution matrix of the homogeneous system:

$$\frac{\partial\boldsymbol{\Phi}(t,s)}{\partial t} = \boldsymbol{A}(t,\boldsymbol{\delta})\boldsymbol{\Phi}(t,s), \quad t > s \quad (3.2a)$$

$$\boldsymbol{\Phi}(s,s) = \boldsymbol{I}. \quad (3.2b)$$

Theorem 3.1. *The process* $\{\boldsymbol{\xi}(t) : t \geq 0\}$ *is a Gaussian process with mean and covariance matrix given, for* $t \in (\tau_{i-1}, \tau_i)$ *with* $i = 1, ..., m$*, by*

$$\boldsymbol{\mu}(t \mid \boldsymbol{\delta}) = \mathcal{E}\{\boldsymbol{\xi}(t)\} = \boldsymbol{\Phi}(t, \tau_{i-1} \mid \boldsymbol{\delta})\boldsymbol{\mu}(\tau_{i-1}^+) + \int_{\tau_{i-1}}^{t}\boldsymbol{\Phi}(t,s \mid \boldsymbol{\delta})\boldsymbol{B}(s,\boldsymbol{\delta})ds, \quad (3.3)$$

and

$$\boldsymbol{\Psi}(t \mid \boldsymbol{\delta}) = \boldsymbol{\Phi}(t,\tau_{i-1} \mid \boldsymbol{\delta})\boldsymbol{\Psi}(\tau_{i-1}^+ \mid \boldsymbol{\delta})(\boldsymbol{\Phi}(t,\tau_{i-1} \mid \boldsymbol{\delta}))^{\intercal} + \int_{\tau_{i-1}}^{t}\boldsymbol{\Phi}(t,s \mid \boldsymbol{\delta})\boldsymbol{D}(\tau,\boldsymbol{\delta})(\boldsymbol{D}(s,\boldsymbol{\delta}))^{\intercal}(\boldsymbol{\Phi}(t,s \mid \boldsymbol{\delta}))^{\intercal}ds, \quad (3.4)$$

respectively. Here, at $t = \tau_i$ *with* $i = 1, ..., m$*, the mean and the covariance matrix of* $\boldsymbol{\xi}(\tau_i^+)$ *are*

$$\boldsymbol{\mu}(\tau_i^+) = \boldsymbol{J}^i\boldsymbol{\mu}(\tau_i^-) + \boldsymbol{\gamma}^i, \quad (3.5)$$

and

$$\boldsymbol{\Psi}(\tau_i^+) = \boldsymbol{J}^i\boldsymbol{\Psi}(\tau_i^-)(\boldsymbol{J}^i)^\intercal + \boldsymbol{K}^i, \tag{3.6}$$

respectively.

Proof: Since $\boldsymbol{\xi}^0$ is a Gaussian vector and the linear transformation of a Gaussian is Gaussian, it follows from (3.1) with $i = 1$ that $\{\boldsymbol{\xi}(t) : 0 \leq t \leq \tau_1\}$ is a Gaussian process. Its mean and covariance matrix are given, for $t \in (0, \tau_1)$, by

$$\begin{aligned}\boldsymbol{\mu}(t \mid \boldsymbol{\delta}) = \mathcal{E}\{\boldsymbol{\xi}(t)\} &= \boldsymbol{\Phi}(t, 0 \mid \boldsymbol{\delta})\mathcal{E}\{\boldsymbol{\xi}^0\} + \int_0^t \boldsymbol{\Phi}(t, s \mid \boldsymbol{\delta})\boldsymbol{B}(s, \boldsymbol{\delta})ds \\ &= \boldsymbol{\Phi}(t, 0 \mid \boldsymbol{\delta})\boldsymbol{\mu}^0 + \int_0^t \boldsymbol{\Phi}(t, s \mid \boldsymbol{\delta})\boldsymbol{B}(s, \boldsymbol{\delta})ds \end{aligned} \tag{3.7}$$

and

$$\begin{aligned}\boldsymbol{\Psi}(t \mid \boldsymbol{\delta}) =& \boldsymbol{\Phi}(t, \tau_{i-1} \mid \boldsymbol{\delta})\boldsymbol{\Psi}(\tau_{i-1}^+ \mid \boldsymbol{\delta})(\boldsymbol{\Phi}(t, \tau_{i-1} \mid \boldsymbol{\delta}))^\intercal \\ &+ \int_0^t \boldsymbol{\Phi}(t, s \mid \boldsymbol{\delta})\boldsymbol{D}(\tau, \boldsymbol{\delta})(\boldsymbol{D}(s, \boldsymbol{\delta}))^\intercal(\boldsymbol{\Phi}(t, s \mid \boldsymbol{\delta}))^\intercal ds, \end{aligned} \tag{3.8}$$

respectively.

Note that $\boldsymbol{\xi}(\tau_1^-)$ is a Gaussian vector. Since the linear transformation of Gaussian is Gaussian, it follows from (2.1c) with $i = 1$ that $\boldsymbol{\xi}(\tau_1^+)$ is Gaussian with the mean and covariance matrix given by

$$\boldsymbol{\mu}(\tau_1^+) = \mathcal{E}\{\boldsymbol{\xi}(\tau_1^+)\} = \mathcal{E}\{\boldsymbol{J}^1\boldsymbol{\xi}(\tau_1^-)\} + \boldsymbol{\gamma}^1 = \boldsymbol{J}^1\boldsymbol{\mu}(\tau_1^-) + \boldsymbol{\gamma}^1, \tag{3.9}$$

and

$$\begin{aligned}\boldsymbol{\Psi}(\tau_1^+) &= \mathcal{E}\{[\boldsymbol{\xi}(\tau_1^+) - \boldsymbol{\mu}(\tau_1^+)][\boldsymbol{\xi}(\tau_1^+) - \boldsymbol{\mu}(\tau_1^+)]^\intercal\} \\ &= \mathcal{E}\{[\boldsymbol{J}^1\boldsymbol{\xi}(\tau_1^-) - \boldsymbol{J}^1\boldsymbol{\mu}(\tau_1^-) + \boldsymbol{\Delta}_1][\boldsymbol{J}^1\boldsymbol{\xi}(\tau_1^-) - \boldsymbol{J}^1\boldsymbol{\mu}(t_1^-) + \boldsymbol{\Delta}_1]^\intercal\} \\ &= \mathcal{E}\{[\boldsymbol{J}^1\boldsymbol{\xi}(\tau_1^-) - \boldsymbol{J}^1\boldsymbol{\mu}(\tau_1^-)][\boldsymbol{J}^1\boldsymbol{\xi}(\tau_1^-) - \boldsymbol{J}^1\boldsymbol{\mu}(\tau_1^-)]^\intercal\} + \mathcal{E}\{\boldsymbol{\Delta}_1\boldsymbol{\Delta}_1^\intercal\} \\ &= \boldsymbol{J}^1\mathcal{E}\{[\boldsymbol{\xi}(\tau_1^-) - \boldsymbol{\mu}(\tau_1^-)][\boldsymbol{\xi}(\tau_1^-) - \boldsymbol{\mu}(\tau_1^-)]^\intercal\}(\boldsymbol{J}^1)^\intercal + \boldsymbol{K}^1 \\ &= \boldsymbol{J}^1\boldsymbol{\Psi}(\tau_1^-)(\boldsymbol{J}^1)^\intercal + \boldsymbol{K}^1. \end{aligned} \tag{3.10}$$

respectively.

By the same token, we can show that for $i = 2$, the process $\{\boldsymbol{\xi}(t) : t \in [\tau_1, \tau_2]\}$ is a Gaussian process with the mean and covariance matrix given, for $t \in (\tau_1, \tau_2)$, by

$$\boldsymbol{\mu}(t \mid \boldsymbol{\delta}) = \boldsymbol{\Phi}(t, \tau_1 \mid \boldsymbol{\delta})\boldsymbol{\mu}(\tau_1^+) + \int_{\tau_1}^t \boldsymbol{\Phi}(t, s \mid \boldsymbol{\delta})\boldsymbol{B}(s, \boldsymbol{\delta})ds \tag{3.11}$$

and

$$\begin{aligned}\boldsymbol{\Psi}(t \mid \boldsymbol{\delta}) =& \boldsymbol{\Phi}(t, \tau_1 \mid \boldsymbol{\delta})\boldsymbol{\Psi}(\tau_1^+ \mid \boldsymbol{\delta})(\boldsymbol{\Phi}(t, \tau_1 \mid \boldsymbol{\delta}))^\intercal \\ &+ \int_{\tau_1}^t \boldsymbol{\Phi}(t, s \mid \boldsymbol{\delta})\boldsymbol{D}(\tau, \boldsymbol{\delta})(\boldsymbol{D}(s, \boldsymbol{\delta}))^\intercal(\boldsymbol{\Phi}(t, s \mid \boldsymbol{\delta}))^\intercal ds, \end{aligned} \tag{3.12}$$

respectively.

At $t = \tau_2$, it follows from (2.1c) with $i = 2$ that the mean and the covariance matrix of $\boldsymbol{\xi}(\tau_2^+)$ are

$$\boldsymbol{\mu}(\tau_2^+) = \mathcal{E}\{\boldsymbol{\xi}(\tau_2^+)\} = \boldsymbol{J}^2\boldsymbol{\mu}(\tau_2^-) + \boldsymbol{\gamma}^2, \tag{3.13}$$

and

$$\begin{aligned}\boldsymbol{\Psi}(\tau_2^+) &= \mathcal{E}\{[\boldsymbol{\xi}(\tau_2^+) - \boldsymbol{\mu}(\tau_2^+)][\boldsymbol{\xi}(\tau_2^+) - \boldsymbol{\mu}(\tau_2^+)]^\intercal\} \\ &= \boldsymbol{J}^2\boldsymbol{\Psi}(\tau_2^-)(\boldsymbol{J}^2)^\intercal + \boldsymbol{K}^2, \end{aligned} \tag{3.14}$$

respectively.

The process can be repeated for $i = 3, .., m$. This completes the proof.

From (3.7), it follows that, for $t \in (\tau_{i-1}, \tau_i)$ with $i = 1, 2, ..., m$, $\boldsymbol{\mu}(t)$ is the solution of the following system of differential equations:

$$d\boldsymbol{\mu}(t)/dt = \boldsymbol{A}(t, \boldsymbol{\delta})\boldsymbol{\mu}(t) + \boldsymbol{B}(t, \boldsymbol{\delta}) \tag{3.15a}$$

with initial condition

$$\boldsymbol{\mu}(0) = \boldsymbol{\mu}^0, \tag{3.15b}$$

and jump conditions

$$\boldsymbol{\mu}(\tau_i^+) = \boldsymbol{J}^i \boldsymbol{\mu}(\tau_i^-) + \boldsymbol{\gamma}^i. \tag{3.15c}$$

Similarly, it follows from (3.8) that, for $t \in (\tau_{i-1}, \tau_i)$ with $i = 1, 2, ..., m$, $\boldsymbol{\Psi}(t \mid \boldsymbol{\delta})$ is the solution of the following matrix differential equation:

$$d\boldsymbol{\Psi}(\boldsymbol{t})/dt = \boldsymbol{A}(t, \boldsymbol{\delta})\boldsymbol{\Psi}(t) + \boldsymbol{\Psi}^\intercal(t)\boldsymbol{A}(t, \boldsymbol{\delta}) + \boldsymbol{D}(t, \boldsymbol{\delta})(\boldsymbol{D}(t, \boldsymbol{\delta}))^\intercal \tag{3.16a}$$

with initial condition

$$\boldsymbol{\Psi}(0) = \boldsymbol{\Psi}^0, \tag{3.16b}$$

and jump conditions

$$\boldsymbol{\Psi}(\tau_i^+) = \boldsymbol{J}^i \boldsymbol{\Psi}(\tau_i^-)(\boldsymbol{J}^i)^\intercal + \boldsymbol{K}^i. \tag{3.16c}$$

Consider the cost function (2.6). Since $\mathcal{E}\{\boldsymbol{\xi}(t)\boldsymbol{\xi}^\intercal(t)\} = \boldsymbol{\Psi}(t) + \boldsymbol{\mu}(t)(\boldsymbol{\mu}(t))^\intercal$, it follows that

$$\begin{aligned} g_0(\boldsymbol{\delta}, \boldsymbol{\tau}, \boldsymbol{\gamma}) = {} & \varphi(\boldsymbol{\gamma}) + \boldsymbol{S}_2(\boldsymbol{\delta})[\boldsymbol{\Psi}(T \mid \boldsymbol{\delta}) + \boldsymbol{\mu}(T \mid \boldsymbol{\delta})(\boldsymbol{\mu}(T \mid \boldsymbol{\delta}))^\intercal] + \boldsymbol{S}_1^\intercal(\boldsymbol{\delta})\boldsymbol{\mu}(T) + \boldsymbol{S}_0(\boldsymbol{\delta}) \\ & + \sum_{i=1}^{m+1} \int_{\tau_{i-1}}^{\tau_i} \{\text{trace}[\boldsymbol{Q}_2(t, \boldsymbol{\delta})(\boldsymbol{\Psi}(t \mid \boldsymbol{\delta}) + \boldsymbol{\mu}(t, \boldsymbol{\delta})(\boldsymbol{\mu}(t, \boldsymbol{\delta}))^\intercal)] \\ & + \boldsymbol{Q}_1(t, \boldsymbol{\delta})^\intercal \boldsymbol{\mu}(t \mid \boldsymbol{\delta}) + \boldsymbol{Q}_0(t, \boldsymbol{\delta})\}dt. \end{aligned} \tag{3.17}$$

We now consider the probabilistic state constraint (2.2). Since $\boldsymbol{\xi}(t)$ is Gaussian with mean $\boldsymbol{\mu}(t)$ and covariance $\boldsymbol{\Psi}(t)$, it is clear that for each $k = 1, \ldots, N$, the scalar product $(\boldsymbol{c}^k)^\intercal \boldsymbol{\xi}(t)$ is Gaussian with the mean $(\boldsymbol{c}^k)^\intercal \boldsymbol{\mu}(t)$ and covariance $(\boldsymbol{c}^k)^\intercal \boldsymbol{\Psi}(t)\boldsymbol{c}^k$. Thus, for each $k = 1, \ldots, N$, (2.2) is equivalent to

$$\begin{aligned} & q_k(t, \boldsymbol{\mu}(t \mid \boldsymbol{\delta}), \boldsymbol{\Psi}(t \mid \boldsymbol{\delta})) \\ = & \alpha_k - \int_{a_k}^{b_k} \frac{1}{(2\pi(\boldsymbol{c}^k)^\intercal \boldsymbol{\Psi}(t \mid \boldsymbol{\delta})\boldsymbol{c}^k)^{1/2}} \exp\{\frac{-(y - (\boldsymbol{c}^k)^\intercal \boldsymbol{\mu}(t \mid \boldsymbol{\delta}))^2}{2(\boldsymbol{c}^k)^\intercal \boldsymbol{\Psi}(t \mid \boldsymbol{\delta})\boldsymbol{c}^k}\}dy \le 0, \end{aligned} \tag{3.18}$$

for all $t \in [0, T]$. These constraints are continuous state inequality constraints.

Now, we have transformed the stochastic optimal parameter selection problem into a deterministic optimal parameter selection problem defined as follows.

Problem (P2). *Given the dynamical system (3.15a)-(3.15c) and (3.16a)-(3.16c), and the continuous state inequality constraints (3.18), find a feasible parameter* $(\boldsymbol{\delta}, \boldsymbol{\tau}, \boldsymbol{\gamma}) \in \boldsymbol{\Omega} \times \boldsymbol{\mathcal{T}} \times \boldsymbol{\Gamma}$*, such that the cost function (3.17) is minimized.*

We now summarize the results obtained so far below as a theorem.

Theorem 3.2. *Problem (P1) is equivalent to Problem (P2).*

4. **Time Scaling Transformation.** Problem (P2) is a deterministic impulsive optimal parameter selection problem subject to continuous state inequality constraints, where the jump times are decision variables to be determined optimally. This will encounter difficulty in numerical calculation when solving the impulsive dynamical system with varying jump times. In this section, we will use a time scaling transform reported in [13] to map these variable jump times into fixed knots in a new time scale.

We consider a new time variable s which varies from 0 to $m+1$. We re-scale $t \in [0, T]$ into $s \in [0, m+1]$. The transformation from $t \in [0, T]$ to $s \in [0, m+1]$ is defined by the differential equation

$$dt(s)/ds = \upsilon(s) = \sum_{i=1}^{m+1} \upsilon_i \chi_{[i-1,i]}(t) \tag{4.1a}$$

$$t(0) = 0, \tag{4.1b}$$

where $\upsilon_i = \tau_i - \tau_{i-1}$. Let $\boldsymbol{\Upsilon}$ be the set of all those $\boldsymbol{\upsilon} = [\upsilon_1, \ldots, \upsilon_{m+1}]^\intercal \in \mathbb{R}^{m+1}$ such that

$$\upsilon_i \geq 0, \ \ i = 1, \ldots, m+1. \tag{4.2}$$

Obviously, the following constraint must also be satisfied.

$$\sum_{i=1}^{m+1} \upsilon_i = T. \tag{4.3}$$

Denote $\hat{\boldsymbol{\mu}}(s) = \boldsymbol{\mu}(t(s))$, and $\hat{\boldsymbol{\Psi}}(s) = \boldsymbol{\Psi}(t(s))$. Then, (3.15) and (3.16) are transformed into

$$d\hat{\boldsymbol{\mu}}(s)/ds = \upsilon(s)[\boldsymbol{A}(t(s), \boldsymbol{\delta})\hat{\boldsymbol{\mu}}(s) + \boldsymbol{B}(t(s), \boldsymbol{\delta})] \tag{4.4a}$$

$$\hat{\boldsymbol{\mu}}(0) = \boldsymbol{\mu}^0 \tag{4.4b}$$

$$\hat{\boldsymbol{\mu}}(i^+) = J^i \hat{\boldsymbol{\mu}}(i^-) + \boldsymbol{\gamma}^i, \quad i = 1, \ldots, m, \tag{4.4c}$$

and

$$d\hat{\boldsymbol{\Psi}}(s)/ds = \upsilon(s)[\boldsymbol{A}(t(s), \boldsymbol{\delta})\hat{\boldsymbol{\Psi}}(s) + \hat{\boldsymbol{\Psi}}^\intercal(s)\boldsymbol{A}(t(s), \boldsymbol{\delta}) + \boldsymbol{D}(t(s), \boldsymbol{\delta})(\boldsymbol{D}(t(s), \boldsymbol{\delta}))^\intercal] \tag{4.5a}$$

$$\hat{\boldsymbol{\Psi}}(0) = \boldsymbol{\Psi}^0. \tag{4.5b}$$

$$\hat{\boldsymbol{\Psi}}(i^+) = \boldsymbol{J}^i \hat{\boldsymbol{\Psi}}(i^-)(\boldsymbol{J}^i)^\intercal + \boldsymbol{K}^i, \quad i = 1, \ldots, m. \tag{4.5c}$$

The cost function (3.17) is transformed into

$$\begin{aligned} &\hat{g}_0(\boldsymbol{\delta}, \boldsymbol{\upsilon}, \boldsymbol{\gamma}) \\ &= \hat{\Phi}_0(\hat{\boldsymbol{\mu}}(m+1), \hat{\boldsymbol{\Psi}}(m+1), \boldsymbol{\delta}, \boldsymbol{\gamma}) + \sum_{i=1}^{m+1} \int_{i-1}^{i} \hat{\mathcal{L}}_0(t(s), \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s), \boldsymbol{\delta}, \boldsymbol{\upsilon}, \boldsymbol{\gamma}), \end{aligned} \tag{4.6}$$

where

$$\begin{aligned} \hat{\Phi}_0(\hat{\boldsymbol{\mu}}(m+1), \hat{\boldsymbol{\Psi}}(m+1), \boldsymbol{\delta}, \boldsymbol{\gamma}) &= \varphi(\gamma) + \boldsymbol{S}_1^\intercal(\boldsymbol{\delta})\hat{\boldsymbol{\mu}}(m+1) + \boldsymbol{S}_0(\boldsymbol{\delta}) \\ &+ \mathrm{trace}\{\boldsymbol{S}_2(\boldsymbol{\delta})[\hat{\boldsymbol{\Psi}}(m+1 \mid \boldsymbol{\delta}) + \hat{\boldsymbol{\mu}}(m+1|\boldsymbol{\delta})(\hat{\boldsymbol{\mu}}(m+1 \mid \boldsymbol{\delta}))^\intercal]\} \end{aligned}$$

and

$$\begin{aligned} \hat{\mathcal{L}}_0(t(s), \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s), \boldsymbol{\delta}, \boldsymbol{\upsilon}, \boldsymbol{\gamma}) = \upsilon_i\{&\mathrm{trace}[\boldsymbol{Q}_2(t(s), \boldsymbol{\delta})(\hat{\boldsymbol{\Psi}}(s \mid \boldsymbol{\delta}) + \hat{\boldsymbol{\mu}}(s, \boldsymbol{\delta})(\hat{\boldsymbol{\mu}}(s, \boldsymbol{\delta}))^\intercal)] \\ &+ \boldsymbol{Q}_1(t(s), \boldsymbol{\delta})^\intercal \hat{\boldsymbol{\mu}}(s|\boldsymbol{\delta}) + \boldsymbol{Q}_0(t(s), \boldsymbol{\delta})\}. \end{aligned}$$

For the continuous state inequality constraints (3.18), they are transformed into

$$\begin{aligned}&\hat{q}_k(s, \hat{\boldsymbol{\mu}}(s \mid \boldsymbol{\delta}), \hat{\boldsymbol{\Psi}}(s \mid \boldsymbol{\delta}))\\ =&\alpha_k - \int_{a_k}^{b_k} \frac{1}{(2\pi(\boldsymbol{c}^k)^\intercal \hat{\boldsymbol{\Psi}}(s \mid \boldsymbol{\delta})\boldsymbol{c}^k)^{1/2}} \exp\{\frac{-(y-(\boldsymbol{c}^k)^\intercal \hat{\boldsymbol{\mu}}(s \mid \boldsymbol{\delta}))^2}{2(\boldsymbol{c}^k)^\intercal \hat{\boldsymbol{\Psi}}(s \mid \boldsymbol{\delta})\boldsymbol{c}^k}\}dy \leq 0, \end{aligned} \tag{4.7}$$

for all $s \in [0, m+1]$, where $k = 1, ..., N$.

Then, after this time scaling transformation, Problem (P2) is equivalent to

Problem (P3). *Given the dynamical system (4.1), (4.4) and (4.5), find a feasible parameter from* $(\boldsymbol{\delta}, \boldsymbol{\upsilon}, \boldsymbol{\gamma}) \in \boldsymbol{\Omega} \times \boldsymbol{\Upsilon} \times \boldsymbol{\Gamma}$ *such that the cost function (4.6) is minimized subject to the constraints (4.3) and the continuous state inequality constraints (4.7).*

5. **Constraint Transcription.** From the continuous state inequality constraints (4.7), we see that these inequality constraints are to be satisfied for all $s \in [0, m+1]$. They are extremely difficult to deal with directly. We shall use a constraint transcription technique introduced in [12] to approximate these continuous state inequality constraints.

For each $k = 1, \ldots, N$, the continuous state inequality constraint (4.7) is equivalent to

$$G_k(\boldsymbol{\delta}, \boldsymbol{\upsilon}, \boldsymbol{\gamma}) = \sum_{i=1}^{m+1} \int_{i-1}^{i} \max\{\hat{q}_k(s, \hat{\boldsymbol{\mu}}(s \mid \boldsymbol{\delta}), \hat{\boldsymbol{\Psi}}(s \mid \boldsymbol{\delta})), 0\} ds = 0. \tag{5.1}$$

Then, Problem (P3) is equivalent to

Problem (P4). *Problem (P3) with the continuous state inequality constraints (4.7) replaced by their respective equality constraints (5.1).*

However, the equality constraints (5.1) are non-differentiable. We shall use the constraint transcription method to construct, for each $k = 1, ..., N$, a smoothing function $\hat{\mathcal{L}}_{k,\varepsilon}(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s))$ to approximate the non-smooth function $\max\{\hat{q}_k(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s)), 0\}$ in (5.1), where

$$\begin{aligned}&\hat{\mathcal{L}}_{k,\varepsilon}(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s))\\ =&\begin{cases} 0 & \text{if } \hat{q}_k(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s)) < -\varepsilon \\ (\hat{q}_k(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s)) + \varepsilon)^2/4\varepsilon & \text{if } -\varepsilon \leq \hat{q}_k(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s)) \leq \varepsilon \\ \hat{q}_k(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s)) & \text{if } \hat{q}_k(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s)) > \varepsilon \end{cases}. \end{aligned} \tag{5.2}$$

For any $\varepsilon > 0$, $\hat{\mathcal{L}}_{k,\varepsilon}(s, \hat{\boldsymbol{\mu}}(s), \hat{\boldsymbol{\Psi}}(s))$, $k = 1, \ldots, N$, are continuously differentiable and they do not always fail to satisfy the constraint qualifications (see Chapter 3 of [22]).

Now, the equality constraints (5.1) are approximated by

$$\begin{aligned}&\hat{G}_{k,\varepsilon,\beta}(\boldsymbol{\delta}, \boldsymbol{\upsilon}, \boldsymbol{\gamma})\\ =&\beta + \sum_{i=1}^{m+1} \int_{i-1}^{i} \hat{\mathcal{L}}_{k,\varepsilon}(s, \hat{\boldsymbol{\mu}}(s \mid \boldsymbol{\delta}), \hat{\boldsymbol{\Psi}}(s \mid \boldsymbol{\delta})) ds \leq 0, \quad k = 1, \ldots, N, \end{aligned} \tag{5.3}$$

where $\beta > 0$ is the parameter to adjust the feasibility of the solution, while $\varepsilon > 0$ is the parameter to adjust the accuracy of the solution.

Problem (P4) with (5.1) replaced by (5.3) is denoted as Problem (P4(ε, β)).

For each $\varepsilon > 0$ and $\beta > 0$, Problem (P4(ε, β)) is an optimal parameter selection problem subject to canonical inequality constraints, where $\hat{\boldsymbol{\Psi}}(s)$ is determined by a system of differential equations in matrix form. We shall re-define the variables of the systems of differential equations (2.1.a) and (2.1.c) and rewrite these systems together as a system of standard ordinary differential equations in vector form.

Let $\boldsymbol{x}(s)$ be the vector formed by $t(s)$, $\hat{\boldsymbol{\mu}}(s)$ and the independent components of the matrix $\hat{\boldsymbol{\Psi}}(s)$, *i.e.*,

$$\boldsymbol{x}(s) = [t(s), \hat{\boldsymbol{\mu}}^{\intercal}(s), \hat{\psi}_{11}(s), \dots, \hat{\psi}_{1n}(s), \hat{\psi}_{22}(s), \dots, \hat{\psi}_{2n}(s), \dots, \hat{\psi}_{nn}(s)]^{\intercal}. \tag{5.4}$$

Let $\boldsymbol{\sigma} = (\boldsymbol{\delta}, \boldsymbol{\upsilon}, \boldsymbol{\gamma})$ and let $\boldsymbol{f}$ be the corresponding vector obtained from the right hand sides of (4.1a), (4.4a) and (4.5a). Furthermore, let Φ_0, $\mathcal{L}_0$ and $\mathcal{L}_{i,\varepsilon}$, $i = 1, ...N$, be obtained from $\hat{\Phi}_0$, $\hat{\mathcal{L}}_0$ and $\hat{\mathcal{L}}_{k,\varepsilon}$, $k = 1, ...N$, respectively, with $t(s)$, $\hat{\boldsymbol{\mu}}(s)$ and $\hat{\boldsymbol{\Psi}}(s)$ appropriately replaced by $\boldsymbol{x}(s)$.

Then, for each ε and β, Problem (P4(ε, β)) is equivalent to

Problem (P5(ε, β))**.** *Given the dynamical system*

$$d\boldsymbol{x}(s)/ds = \boldsymbol{f}(s, \boldsymbol{x}(s), \boldsymbol{\sigma}) \tag{5.5a}$$

$$\boldsymbol{x}(0) = \boldsymbol{x}^0 \tag{5.5b}$$

$$\boldsymbol{x}(i^+) = \boldsymbol{\psi}^i(\boldsymbol{x}(i^-), \boldsymbol{\sigma}), \quad i = 1, \dots, m, \tag{5.5c}$$

where $\boldsymbol{x}^0$ and $\boldsymbol{\psi}^i$ are obtained from (4.1b), (4.4b), (4.5b) and (4.4c), (4.5c), respectively, find a feasible parameter $\boldsymbol{\sigma} \in \boldsymbol{\Omega} \times \boldsymbol{\Upsilon} \times \boldsymbol{\Gamma}$, such that the cost function

$$\hat{g}_0(\boldsymbol{\sigma}) = \Phi_0(\boldsymbol{x}(m+1 \mid \boldsymbol{\sigma}), \boldsymbol{\sigma}) + \sum_{i=1}^{m+1} \int_{i-1}^{i} \mathcal{L}_0(s, \boldsymbol{x}(s \mid \boldsymbol{\sigma}), \boldsymbol{\sigma}) ds, \tag{5.6}$$

is minimized subject to the constraints (4.3) and

$$\hat{G}_{k,\varepsilon}(\boldsymbol{\sigma}) = \beta + \sum_{i=1}^{m+1} \int_{i-1}^{i} \mathcal{L}_{k,\varepsilon}(s, \boldsymbol{x}(s \mid \boldsymbol{\sigma}), \boldsymbol{\sigma}) ds \leq 0, \quad k = 1, \dots, N. \tag{5.7}$$

To solve Problem (P5(ε, β)) as a mathematical programming problem, we need the gradients of cost function and constraint functions. They can be obtained by using similar idea as that given for Theorem 5.2.1 of [22]. Details of these gradients are presented below in the following two theorems.

Theorem 5.1. *The gradient of the cost function (5.5) with respect to $\boldsymbol{\sigma}$ are given by*

$$\nabla_{\boldsymbol{\sigma}} \hat{g}_0(\boldsymbol{\sigma}) = \frac{\partial \Phi_0(\boldsymbol{x}(m+1), \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} + \sum_{i=1}^{m} (\boldsymbol{\lambda}^0(i^+))^{\intercal} \frac{\partial \boldsymbol{\psi}^i(\boldsymbol{x}(i^-), \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} + \sum_{i=1}^{m+1} \int_{i-1}^{i} \frac{\partial H_0(s, \boldsymbol{x}, \boldsymbol{\lambda}^0, \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} ds, \tag{5.8}$$

where the Hamiltonian H_0 is defined by

$$H_0(s, \boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) \quad = \mathcal{L}_0(s, \boldsymbol{x}(s), \boldsymbol{\sigma}) + (\boldsymbol{\lambda}(s))^{\intercal} \boldsymbol{f}(s, \boldsymbol{x}(s), \boldsymbol{\sigma}), \tag{5.9}$$

and $\boldsymbol{\lambda}^0(s)$ is the co-state determined by the following differential equations

$$\frac{d\boldsymbol{\lambda}(s)}{ds} = -\left[\frac{\partial H_0(s, \boldsymbol{x}(s), \boldsymbol{\lambda}(s), \boldsymbol{\sigma})}{\partial \boldsymbol{x}}\right]^{\intercal}, \tag{5.10a}$$

with terminal conditions:

$$\boldsymbol{\lambda}(m+1) = \left[\frac{\partial \boldsymbol{\Phi}_0(\boldsymbol{x}(m+1), \boldsymbol{\sigma})}{\partial \boldsymbol{x}}\right]^{\intercal}, \tag{5.10b}$$

and jump conditions:

$$\boldsymbol{\lambda}(i^-) = \left[\frac{\partial \boldsymbol{\psi}^i(\boldsymbol{x}(i^-), \boldsymbol{\sigma})}{\partial \boldsymbol{x}}\right]^{\intercal} \boldsymbol{\lambda}(i^+). \tag{5.10c}$$

We can also obtain the gradient of constraint functions in the same way.

Theorem 5.2. *For each $k = 1, ..., N$, the gradient of the constraint function (5.7) with respect to $\boldsymbol{\sigma}$ are given by*

$$\nabla_{\boldsymbol{\sigma}} \hat{G}_k(\boldsymbol{\sigma}) = \sum_{i=1}^{m} (\boldsymbol{\lambda}^k(i^+))^\intercal \frac{\partial \boldsymbol{\psi}^i(\boldsymbol{x}(i^-), \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} + \sum_{i=1}^{m+1} \int_{i-1}^{i} \frac{\partial H_k(s, \boldsymbol{x}, \boldsymbol{\lambda}^k, \boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} ds, \tag{5.11}$$

where the Hamiltonian H_k is defined by

$$H_k(s, \boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\sigma}) \quad = \mathcal{L}_{k,\varepsilon}(s, \boldsymbol{x}(s), \boldsymbol{\sigma}) + (\boldsymbol{\lambda}(s))^\intercal \boldsymbol{f}(s, \boldsymbol{x}(s), \boldsymbol{\sigma}), \tag{5.12}$$

and $\boldsymbol{\lambda}^k(s)$ is the co-state determined by the following differential equations

$$\frac{d\boldsymbol{\lambda}(s)}{ds} = - \left[\frac{\partial H_k(s, \boldsymbol{x}(s), \boldsymbol{\lambda}(s), \boldsymbol{\sigma})}{\partial \boldsymbol{x}} \right]^\intercal, \tag{5.13a}$$

with terminal conditions:

$$\boldsymbol{\lambda}(m+1) = \boldsymbol{0}, \tag{5.13b}$$

and jump conditions:

$$\boldsymbol{\lambda}(i^-) = \left[\frac{\partial \boldsymbol{\psi}^i(\boldsymbol{x}(i^-), \boldsymbol{\sigma})}{\partial \boldsymbol{x}} \right]^\intercal \boldsymbol{\lambda}(i^+). \tag{5.13c}$$

6. **Numerical Example.** In this section, we solve an example using our proposed methods as follows. Consider the dynamic system defined on $(0, 1]$ with the coefficients given by

$$\boldsymbol{A} = \begin{pmatrix} 0.6+\delta & \delta \\ 1.5 & -0.5-\delta \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 2+\sin(t) \\ -2-\cos(t) \end{pmatrix}, \boldsymbol{D} = \begin{pmatrix} 0.5 & 0.1 \\ 0.1 & 0.3 \end{pmatrix}.$$

The mean and the covariance matrix of the initial state are

$$\boldsymbol{\mu}^0 = \begin{pmatrix} -0.1 \\ -0.5 \end{pmatrix}, \boldsymbol{\Psi}^0 = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.2 \end{pmatrix}.$$

The coefficients of two jump functions are

$$\boldsymbol{J}^i = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{K}^i = \begin{pmatrix} 0.36 & 0 \\ 0 & 0.36 \end{pmatrix}, \quad \forall i = 1, 2.$$

The coefficients of the cost function are

$$\boldsymbol{S}_2 = \begin{pmatrix} 2.5 & 0 \\ 0 & 2.5 \end{pmatrix}, \boldsymbol{S}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{S}_0 = 0,$$

$$\boldsymbol{Q}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{Q}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{Q}_0 = 0,$$

$$\varphi(\boldsymbol{\gamma}) = \sum_{i=1}^{2} \frac{1}{2} (\boldsymbol{\gamma}^i)^\intercal \boldsymbol{\gamma}^i.$$

There is one probabilistic state constraint with the coefficients given by

$$\boldsymbol{c}^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, a = -3, b = 3, \alpha = 0.9.$$

We apply the solution procedure presented in previous sections to solve this example, where the corresponding version of the Problem (P5(ε, β)) is solved using the optimal control software package MISER3.3 (see [11]). The optimal parameter obtained is $\delta^* = -0.692009$. The first jump appears at time $\tau_1^* = 0.28796$ with the corresponding magnitude vector $\boldsymbol{\gamma}^{1,*} = (-1.01096, 1.40555)^\intercal$ and the second jump appear at time

$\tau_2^* = 0.69132$ with the corresponding magnitude vector $\boldsymbol{\gamma}^{2,*} = [-1.43826, 0.756040]^\top$. The optimal cost function value obtained is $g_0^* = 11.2821699$.

For the simulation of $\xi(t)$, we have obtained 500 samples in Matlab. The results are given in Figure 1.

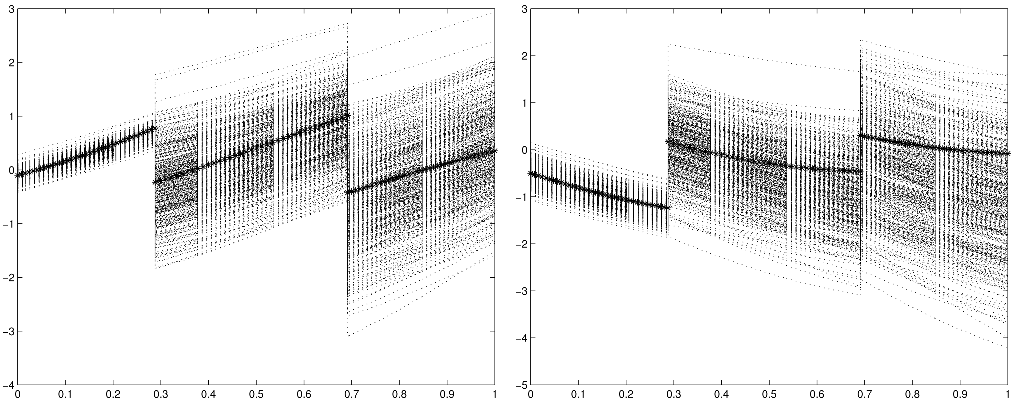

FIGURE 1. * line: $\mu_1(t)$ and $\mu_2(t)$; dotted line: 500 samples of $\xi_1(t)$ and $\xi_2(t)$.

7. **Conclusion.** In this paper, a class of stochastic optimal parameter selection problems involving an impulsive dynamical system subject to probabilistic constraints on the state is considered. This problem contains the one considered in [9] as a special case for which no state jumps are allowed. This problem can also be considered as the stochastic version of the optimal parameter selection problem governed by impulsive dynamic systems. We have shown that this stochastic optimal impulsive parameter selection problem with probabilistic constraints is equivalent to a deterministic impulsive optimal parameter selection problem with continuous state inequality constraints. A numerical method was developed for solving this equivalent constrained deterministic impulsive optimal parameter selection problem. From the numerical study through solving a numerical example, we see that the solution method is effective.

Acknowledgment. This project is supported by a research grant from the Australian Research Council, and National Natural Science Foundation of China under Grant 60704003. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] N. U. Ahmed and N. D. Georgenas, On optimal parameter selection, *IEEE Trans. Automat. Control*, vol.18, pp.313-314, 1973.

[2] S. Aberkane, J. C. Ponsart and D. Sauter, Output feedback H_∞ control of a class of stochastic hybrid systems with wiener process via convex analysis, *International Journal of Innovative Computing, Information and Control*, vol.2, no.6, pp.1179-1196, 2006.

[3] M. Basin, E. Sanchez and R. Martinez-Zuñiga, Optimal linear filtering for systems with multiple state and observation delays, *International Journal of Innovative Computing, Information and Control*, vol.3, no.5, pp.1309-1320, 2007.

[4] J. Dolezal, On the solution of optimal control problems involving parameters and general boundary conditions, *Kybernetika*, vol.17, pp.71-81, 1981.

[5] A. Friedman, *Stochastic Differential Equations and Applications*, Dover Publications, 2006.

[6] Z. G. Feng, K. L. Teo, N. U. Ahmed, Y. Zhao and W. Y. Yan, Optimal fusion of sensor data for Kalman filtering, *Discrete and Continuous Dynamical Systems*, vol.14, pp.483-503, 2006.

[7] Z. G. Feng, K. L. Teo and V. Rehbock, Hybrid method for a general optimal sensor scheduling problem in discrete time, *Automatica*, appear in 2008.

[8] Z. G. Feng, K. L. Teo and Y. Zhao, Branch and bound method for sensor scheduling in discrete time, *Journal of Industrial and Management Optimization*, vol.1, no.4, pp.499-512, 2005.

[9] C. J. Goh, K. L. Teo, On constrained stochastic optimal parameter selection problems, *Bulletin of the Australian Mathematical Society*, vol.41, pp.393-405, 1990.

[10] C. S. Huang, S. Wang and K. L. Teo, Solving Hamilton-Jacobi-Bellman equations by a modified method of characteristics, *Nonlinear Analysis, Theory, Methods and Applications*, vol.40, pp.279-293, 2000.

[11] L. S. Jennings, M. E. Fisher, K. L. Teo and C. J. Goh, *MISER3 Optimal Control Software Version 3.0, Theory and User Manual*, Website http://www.cado.uwa.edu.au/miser.

[12] L. S. Jennings and K. L. Teo, A computational algorithm for functional inequality constrained optimization problems, *Automatica*, vol.26, pp.371-375, 1990.

[13] H. W. J. Lee, K. L. Teo, V. Rehbock and L. S. Jennings, Control parametrization enhancing technique for time optimal control problems, *Dynamc Systems and Applications*, vol.6, pp.243-262, 1997.

[14] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, Springer-Verlag, New York, 2003.

[15] D. W. Reid and K. L. Teo, Optimal parameter selection of parabolic systems, *Math. Oper. Res.*, vol.5, pp.467-474, 1980.

[16] R. Situ, *Theory of Stochastic Differential Equations with Jumps and Applications: Mathematical and Analytical Techniques with Applications to Engineering*, Springer, New York, 2005.

[17] A. Tanikawa, On a smoother for descrete-time linear stochastic systems with unknown disturbances, *International Journal of Innovative Computing, Information and Control*, vol.2, no.5, pp.907-916, 2006.

[18] Y. Takeuchi, Optimazation of linear observations for the stationary Kalman filter based on a generalized water filling theorem, *International Journal of Innovative Computing, Information and Control*, vol.4, no.1, pp.907-916, 2008.

[19] K. L. Teo and N. U. Ahmed, Optimal feedback control for a class of stochastic systems, *Internat. J. Systems Sci.*, vol.5, pp.357-365, 1974.

[20] K. L. Teo and C. J. Goh, A simple computational procedure for optimization problems with functional inequality constraints, *IEEE Trans. on Automatic Control*, vol.32, pp.940-941, 1987.

[21] K. L. Teo and C. J. Goh, A computational method for combined optimal parameter selection and optimal control problems with general constraints, *J. Austral. Math. Soc. Ser. B*, vol.30, pp.350-364, 1989.

[22] K. L. Teo, C. J. Goh and K. H. Wong, *A Unified Computational Approach to Optimal Control Problems*, Longman Scientific and Technical, 1991.

[23] S. Wang, L. S. Jennings and K. L. Teo, Numerical solution of Hamilton-Jacobi-Bellman equations by an upwind finite volume method, *Journal of Global Optimization*, vol.27, pp.177-192, 2003.