AN OPTIMAL CONTROL PROBLEM WITH TERMINAL STOCHASTIC LINEAR COMPLEMENTARITY CONSTRAINTS*

JIANFENG LUO† AND XIAOJUN CHEN‡

Abstract. In this paper, we investigate an optimal control problem with a crucial ODE constraint involving a terminal stochastic linear complementarity problem (SLCP), and its discrete approximation using the relaxation, the sample average approximation (SAA) and the implicit Euler time-stepping scheme. We show the existence of feasible solutions and optimal solutions to the optimal control problem and its discrete approximation under the conditions that the expectation of the stochastic matrix in the SLCP is a Z-matrix or an adequate matrix. Moreover, we prove that the solution sequence generated by the discrete approximation converges to a solution of the original optimal control problem with probability 1 by the repeated limits in the order of $\epsilon \downarrow 0$, $\nu \to \infty$ and $h \downarrow 0$, where ϵ is the relaxation parameter, ν is the sample size and h is the mesh size. We also provide asymptotics of the SAA optimal value and error bounds of the time-stepping method. A numerical example is used to illustrate the existence of optimal solutions, the discretization scheme and error estimation.

Key words. ODE constrained optimal control problem, stochastic linear complementarity problem, sample average approximation, implicit Euler time-stepping, convergence analysis.

MSC codes. 49M25, 49N10, 90C15, 90C33

1. Introduction.

In this paper, we aim to find an optimal solution $(x, u) \in H^1(0,T)^n \times L^2(0,T)^m$ of the following optimal control problem with terminal stochastic linear complementarity constraints:

$$
\begin{aligned}
&\min_{x,u} \; \mathbb{E}[F(x(T),\xi)] + \frac{1}{2}\|x - x_d\|_{L^2}^2 + \frac{\delta}{2}\|u - u_d\|_{L^2}^2 \\
&\text{s.t.} \begin{cases} \left.\begin{array}{l} \dot{x}(t) = Ax(t) + Bu(t), \\ Cx(t) + Du(t) - f(t) \le 0, \end{array}\right\} a.e.\ t \in (0,T), \\ 0 \le x(T) \perp \mathbb{E}[M(\xi)x(T) + q(\xi)] \ge 0, \\ x(0) = x_0,\ \mathbb{E}[g(x(T),\xi)] \in K. \end{cases}
\end{aligned}
\tag{1.1}
$$

Here ξ denotes a random variable defined in the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ with support set $\Xi := \xi(\Omega) \subseteq \mathbb{R}^{\mathfrak{b}}$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{l\times n}$, $D \in \mathbb{R}^{l\times m}$, $x_0 \in \mathbb{R}^n$, and $f \in L^2(0,T)^l$, $\delta > 0$ is a scalar, $K \subseteq \mathbb{R}^k$ is a nonempty, closed and convex set, $x_d \in L^2(0,T)^n$ and $u_d \in L^2(0,T)^m$ are the given desired state and control, respectively, $F : \mathbb{R}^n \times \Xi \to \mathbb{R}$, $g : \mathbb{R}^n \times \Xi \to \mathbb{R}^k$, $M : \Xi \to \mathbb{R}^{n\times n}$ and $q : \Xi \to \mathbb{R}^n$. We assume that the expected values in (1.1) are well defined, and F and g are continuously differentiable with respect to $x(T)$ over $\mathbb{R}^n$.

Let $\|\cdot\|$ denote the Euclidean norm of a vector and a matrix. We denote $L^2(0,T)^n$ the Banach space of all quadratically Lebesgue integrable functions mapping from

---

*Submitted to the editors DATE.

Funding: This work is supported by the Hong Kong Research Grants Council grant PolyU15300021 and the CAS AMSS-PolyU Joint Laboratory in Applied Mathematics.

†CAS AMSS-PolyU Joint Laboratory of Applied Mathematics (Shenzhen), The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen, China (gstljf@polyu.edu.hk).

‡Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China (maxjchen@polyu.edu.hk).

$(0,T)$ to $\mathbb{R}^n$, which is equipped with the norm

$$\|x\|_{L^2} := \left(\int_0^T \|x(t)\|^2 dt\right)^{\frac{1}{2}}, \ \forall\, x \in L^2(0,T)^n.$$

Denote $H^1(0,T)^n$ the space whose components $x_1, \cdots, x_n : (0,T) \to \mathbb{R}$ possess weak derivatives such that the function $\dot{x} \in L^2(0,T)^n$. A suitable norm in $H^1(0,T)^n$ is defined by

$$\|x\|_{H^1} := \left(\|x\|_{L^2}^2 + \|\dot{x}\|_{L^2}^2\right)^{\frac{1}{2}}, \ \forall\, x \in H^1(0,T)^n.$$

In [2], Benita and Mehlita studied an optimal control problem with terminal deterministic nonlinear complementarity constraints, which has many interesting practical applications in multi-agent control networks. They derived some stationarity conditions and presented constraint qualifications which ensure that these conditions hold at a local optimal solution of the optimal control problem under the assumption that the feasible set is nonempty. However, sufficient conditions were not given for the existence of $x(T)$ such that the terminal deterministic nonlinear complementarity constraints

$$0 \leq \bar{H}(x(T)) \perp \bar{G}(x(T)) \geq 0, \qquad \bar{g}(x(T)) \in K, \tag{1.2}$$

hold, where $\bar{H} : \mathbb{R}^n \to \mathbb{R}^n$, $\bar{G} : \mathbb{R}^n \to \mathbb{R}^n$, and $\bar{g} : \mathbb{R}^n \to \mathbb{R}^k$. Motivated by the work of [2], we consider problem (1.1) in uncertain environment, which replaces (1.2) by stochastic terminal conditions

$$0 \leq x(T) \perp \mathbb{E}[M(\xi)x(T) + q(\xi)] \geq 0, \qquad \mathbb{E}[g(x(T),\xi)] \in K. \tag{1.3}$$

Optimal control with differential equations and complementarity constraints provides a powerful modeling paradigm for many practical problems such as the optimal control of electrical networks with diodes and/or MOS transistors [4] and dynamic optimization of chemical processes [21]. It can also be derived from the KKT conditions of a bilevel optimal control if the lower level problem is convex and satisfies a constraint qualification [18]. A series of works [5, 7, 11, 14, 25] are devoted to the study of optimal control problems with complementarity constraints. It should be noted that these papers focus on deterministic problems, where the system coefficients including system parameters and boundary/initial conditions are perfectly known. On the other hand, optimal control problems with stochastic differential equation constraints under uncertain environment have been extensively studied [17, 19, 20]. These papers investigate theory and algorithms for optimal control when the parameters in the differential equations have noise and uncertainties. However, there is very little research on optimal control with terminal stochastic complementarity constraints.

It is worth noting that the ODE constraint with a terminal complementarity problem (1.2) or a terminal stochastic linear complementarity condition (1.3) is different from the linear complementarity systems (LCS) (see for example [6]),

$$\begin{cases} \dot{x}(t) = \tilde{A}x(t) + \tilde{B}u(t), \\ 0 \leq u(t) \perp \tilde{C}x(t) + \tilde{D}u(t) \geq 0, \qquad t \in [0,T], \\ x(0) = x_0, \end{cases} \tag{1.4}$$

where $\tilde{A} \in \mathbb{R}^{n\times n}$, $\tilde{B} \in \mathbb{R}^{n\times m}$, $\tilde{C} \in \mathbb{R}^{m\times n}$ and $\tilde{D} \in \mathbb{R}^{m\times m}$ are given matrices. In the LCS (1.4), the complementarity constraint involves state and control variables and holds for the whole time interval, while in (1.1), the complementarity constraint holds for the state variable at terminal time.

The main contributions of this paper are summarized as follows. We show the existence of feasible solutions to the optimal control problem (1.1) under the conditions that $\mathbb{E}[M(\xi)]$ is a Z-matrix or an adequate matrix, which gives reasonable conditions for the existence of $x(T)$ such that (1.3) hold. Moreover, we prove the existence of feasible solutions and optimal solutions to the discrete approximation using the relaxation, the sample average approximation (SAA) and the implicit Euler time-stepping scheme under the same conditions. In the convergence analysis, we prove that the solution sequence generated by the discrete approximation converges to a solution of the original optimal control problem with probability 1 (w.p.1) by the repeated limits in the order of $\epsilon \downarrow 0$, $\nu \to \infty$ and $h \downarrow 0$, where ϵ is the relaxation parameter, ν is the sample size and h is the mesh size. We also provide asymptotics of the SAA optimal value and error bounds of the time-stepping method. These results extend the approximation error of the Euler time-stepping method of an optimal control problem with convex terminal constraints to nonconvex terminal stochastic complementarity constraints.

The paper is organised as follows: Section 2 deals with the existence of feasible solutions of problem (1.1). Section 3 studies the existence of feasible solutions of the relaxation and the SAA of (1.1) and the convergence to the original problem (1.1) as the relaxation parameter goes to zero and the sample size approaches to infinity. In Section 4, we study the convergence of the time-stepping scheme and show the convergence properties of the discrete method using the SAA and the implicit Euler time-stepping scheme. A numerical example is given in Section 5 to illustrate the theoretical results obtained in this paper. Final conclusion remarks are presented in Section 6.

1.1. Notation and assumptions. Throughout this paper we use the following notation. For a matrix $\hat{A} \in \mathbb{R}^{m\times n}$, $\hat{A}^\top$ denotes its transpose matrix, and $\hat{A}^\dagger$ is its pseudoinverse matrix. If $\hat{A}$ possesses full row rank m, we have $\hat{A}^\dagger = \hat{A}^\top(\hat{A}\hat{A}^\top)^{-1}$. Let I denote the identity matrix with a certain dimension. For a vector $z \in \mathbb{R}^n$, $\|z\|_1 = \sum_{i=1}^n |z_i|$ and $\|z\|_0 = \sum_{i=1}^n |z_i|^0$, and we set $0^0 = 0$. For a matrix $\hat{A} \in \mathbb{R}^{n\times m}$, $\|\hat{A}\|_1 = \max_{1\le j\le m} \sum_{i=1}^n |a_{ij}|$.

For sets $S_1, S_2 \subseteq \mathbb{R}^n$, we denote the distance from $v \in \mathbb{R}^n$ to S_1 and the deviation of the set S_1 from the set S_2 by $\mathrm{dist}(v, S_1) = \inf_{v'\in S_1} \|v - v'\|$, and $\mathbb{D}(S_1, S_2) = \sup_{v\in S_1} \mathrm{dist}(v, S_2)$, respectively. For sets $S_1, S_2 \subseteq H^1(0,T)^n \times L^2(0,T)^m$, we denote the distance from $(v_1, v_2) \in H^1(0,T)^n \times L^2(0,T)^m$ to S_1 by $\mathrm{dist}((v_1, v_2), S_1) = \inf_{(v_1', v_2')\in S_1}(\|v_1 - v_1'\|_{H^1} + \|v_2 - v_2'\|_{L^2})$, and the deviation of the set S_1 from the set S_2 by $\mathbb{D}(S_1, S_2) = \sup_{(v_1,v_2)\in S_1} \mathrm{dist}((v_1, v_2), S_2)$. Let $\mathcal{B}(v, \varepsilon) = \{w : \|w - v\| \le \varepsilon\}$ be the closed ball centered at v with the radius of ε. Let $\mathrm{int}S$ denote the interior of a set S. Let $[N] = \{1, 2, \ldots, N\}$.

Assumption 1.1. There exist four nonnegative measurable functions $\kappa_i(\xi)$ with $\mathbb{E}[\kappa_i(\xi)] < \infty$ $(i = 1, 2, 3, 4)$ such that for any z_1, $z_2 \in \mathbb{R}^n$,

$$|F(z_1, \xi) - F(z_2, \xi)| \le \kappa_1(\xi)\|z_1 - z_2\|, \ \|g(z_1, \xi)\| \le \kappa_2(\xi)\|z_1\|, \ a.e.\ \xi \in \Xi,$$

and

$$\|M(\xi)\| \le \kappa_3(\xi) \text{ and } \|q(\xi)\| \le \kappa_4(\xi), \ \forall\ \xi \in \Xi.$$

Assumption 1.2. The matrix $D \in \mathbb{R}^{l\times m}$ is full row rank with $l < m$ and the matrix

$$\mathcal{R} := [BY \;\; (A-BD^{\dagger}C)BY \;\; (A-BD^{\dagger}C)^2BY \;\; \cdots \;\; (A-BD^{\dagger}C)^{n-1}BY] \in \mathbb{R}^{n\times n(m-l)}$$

is also full row rank, where $Y \in \mathbb{R}^{m\times(m-l)}$ is a matrix with full column rank $m-l$ such that $DY = 0$.

2. Existence of optimal solutions of problem (1.1)**.** In this section, we first investigate the feasibility of problem (1.1). We call $(x,u) \in H^1(0,T)^n \times L^2(0,T)^m$ a feasible solution of (1.1) if it satisfies the constraints in (1.1).

For an index set $J \subseteq [n]$, let $|J|$ denote its cardinality and J^c denote its complementarity set. We denote by $q_J \in \mathbb{R}^{|J|}$ the subvector formed from a vector $q \in \mathbb{R}^n$ by picking the entries indexed by J and denote by $M_{J_1,J_2} \in \mathbb{R}^{|J_1|\times|J_2|}$ the submatrix formed from a matrix $M \in \mathbb{R}^{n\times n}$ by picking the rows indexed by J_1 and columns indexed by J_2. Let $\mathcal{J} = \{J : \mathbb{E}[M_{J,J}(\xi)] \text{ is nonsingular}\}$ and

$$\beta = \begin{cases} 1 & \text{if } \mathcal{J} = \emptyset, \\ \max\{\|(\mathbb{E}[M_{J,J}(\xi)])^{-1}\|_1 \mid J \in \mathcal{J}\} & \text{otherwise.} \end{cases} \tag{2.1}$$

A square matrix is said to be a P-matrix if all its principal minors are positive. A square matrix is said to be a Z-matrix if its off-diagonal entries are non-positive. A matrix $\mathbb{E}[M(\xi)] \in \mathbb{R}^{n\times n}$ is called column adequate if for each $z \in \mathbb{R}^n$, $z_i(\mathbb{E}[M(\xi)]z)_i \le 0$ for all $i \in [n]$ implies $\mathbb{E}[M(\xi)]z = 0$. The matrix $\mathbb{E}[M(\xi)]$ is row adequate if $\mathbb{E}[M(\xi)]^\top$ is column adequate and it is adequate if it is both column and row adequate [12]. It is known that a P-matrix is adequate and a symmetric positive semi-definite matrix is also adequate [12, Theorem 3.1.7, Theorem 3.4.4]. However, an adequate matrix may neither be a P-matrix nor a positive semi-definite matrix [12].

For a given matrix $\bar{M} \in \mathbb{R}^{n\times n}$ and a given vector $\bar{q} \in \mathbb{R}^n$, let LCP$(\bar{q},\bar{M})$ denote the LCP $0 \le z \perp \bar{M}z + \bar{q} \ge 0$ and $\mathbf{SOL}(\bar{q},\bar{M})$ denote the solution set. A vector $\bar{z} \in \mathbf{SOL}(\bar{q},\bar{M})$ is called a sparse solution of the LCP$(\bar{q},\bar{M})$ if $\bar{z}$ is a solution of the following optimization problem:

$$\begin{aligned} &\min \|z\|_0 \\ &\text{s.t.} \quad z \in \mathbf{SOL}(\bar{q},\bar{M}). \end{aligned}$$

A vector $\bar{z} \in \mathbf{SOL}(\bar{q},\bar{M})$ is called a least-element solution of the LCP$(\bar{q},\bar{M})$ if $\bar{z} \le z$ for all $z \in \mathbf{SOL}(\bar{q},\bar{M})$. If $\bar{M}$ is a Z-matrix and $\mathbf{SOL}(\bar{q},\bar{M}) \ne \emptyset$, then $\mathbf{SOL}(\bar{q},\bar{M})$ has a unique least-element solution which is the unique sparse solution of the LCP$(\bar{q},\bar{M})$[10].

Let $R_{LCP}(\bar{M})$ denote the set of all vectors $\bar{q}$ such that $\mathbf{SOL}(\bar{q},\bar{M}) \ne \emptyset$. For any $y(\bar{q}) \in \mathbf{SOL}(\bar{q},\bar{M})$, we define an index set $\bar{J} = \{i : y_i(\bar{q}) > 0\}$ and a diagonal matrix $\bar{D}$ whose diagonal elements are $(\bar{D})_{ii} = 1$ for $i \in \bar{J}$ and $(\bar{D})_{ii} = 0$ for $i \notin \bar{J}$.

Lemma 2.1. *([9, Theorem 2.2]) Let $\bar{M} \in \mathbb{R}^{n\times n}$ be a Z-matrix, $\bar{q} \in R_{LCP}(\bar{M})$, and let $y(\bar{q})$ be the least-element solution of LCP$(\bar{q},\bar{M})$. With the index set $\bar{J}$ and diagonal matrix $\bar{D}$, the following statements hold.*

(i) $\bar{M}_{\bar{J},\bar{J}}$ is nonsingular for $\bar{J} \ne \emptyset$;
(ii) $y(\bar{q}) = -(I-\bar{D}+\bar{D}\bar{M})^{-1}\bar{D}\bar{q}$;
(iii) $\|(I-\bar{D}+\bar{D}\bar{M})^{-1}\bar{D}\| \le \mathcal{L} := \max\{\|\bar{M}^{-1}_{\alpha,\alpha}\| : M_{\alpha,\alpha} \text{ is nonsingular for } \alpha \subseteq [n]\}$;
(iv) For any neighborhood $\mathcal{N}_{\bar{q}}$ of $\bar{q}$, there is a $p \in \mathcal{N}_{\bar{q}}$ such that $\mathbf{SOL}(p,\bar{M}) \ne \emptyset$. Moreover, we have $-(I-\bar{D}+\bar{D}\bar{M})^{-1}\bar{D} \in \partial y(\bar{q})$.

Lemma 2.2. *([10, Theorem 3.1]) Let $\bar{M}$ be column adequate, $\bar{q} \in R_{LCP}(\bar{M})$ and let $\bar{z}$ be a sparse solution of the LCP$(\bar{q},\bar{M})$. With the index set $\bar{J}$ and diagonal matrix $\bar{D}$, the following statements hold.*

(i) $\bar{M}_{\bar{J},\bar{J}}$ *is nonsingular for* $\bar{J} \neq \emptyset$*;*
(ii) $\bar{z} = -(I - \bar{D} + \bar{D}\bar{M})^{-1}\bar{D}\bar{q}$*;*
(iii) $\|\bar{z}\|_1 \leq L\|\bar{q}\|_1$*, where* $L = \max\{\|\bar{M}_{\alpha,\alpha}^{-1}\|_1 : \bar{M}_{\alpha,\alpha}$ *is nonsingular for* $\alpha \subseteq [n]\}$*;*
(iv) There is no another solution $z \in \boldsymbol{SOL}(\bar{q}, \bar{M})$ *with* $\alpha = \{i : z_i > 0\}$ *such that* $\alpha \subseteq \bar{J}$.

THEOREM 2.3. *Let Assumptions* 1.1 *and* 1.2 *hold. Suppose the following three conditions hold:*
(i) $\mathcal{B}(0, \beta\mathbb{E}[\kappa_2(\xi)]\|\mathbb{E}[q(\xi)]\|_1) \subseteq K$*, where* β *is defined in* (2.1)*,*
(ii) the set $\mathcal{V} := \{v \in \mathbb{R}^n \,|\, \mathbb{E}[M(\xi)v + q(\xi)] \geq 0,\ v \geq 0\}$ *is nonempty,*
(iii) $\mathbb{E}[M(\xi)]$ *is an adequate matrix or a Z-matrix.*
Then problem (1.1) *has a feasible solution* $(x, u) \in H^1(0,T)^n \times L^2(0,T)^m$*. Moreover, problem* (1.1) *admits an optimal solution if* $\mathbb{E}[F(\cdot, \xi)]$ *is bounded from below.*

Proof. According to Theorem 4.1.6 of [24], for arbitrary $p \in L^2(0,T)^l$, the following non-homogeneous differential equation

$$\begin{cases} \dot{x}(t) = (A - BD^\dagger C)x(t) + BD^\dagger p(t), \\ x(0) = x_0, \end{cases} \qquad a.e.\ t \in (0,T).$$

admits a unique solution $\bar{x} \in H^1(0,T)^n$. The matrix $\mathcal{R}$ in Assumption 1.2 possesses full row rank n and is the controllability matrix of the differential equation

$$\dot{x}(t) = (A - BD^\dagger C)x(t) + BYv(t), \tag{2.2}$$

where $v \in L^2(0,T)^{m-l}$ is an input control variable. Hence system (2.2) is a controllable system [24, Corollary 1.4.10], which implies that for any $b \in \mathbb{R}^n$, the following non-homogeneous differential equation

$$\begin{cases} \dot{x}(t) = (A - BD^\dagger C)x(t) + BYv(t), \\ x(0) = 0,\ x(T) = b - \bar{x}(T), \end{cases} \qquad a.e.\ t \in (0,T)$$

admits a solution pair $(\tilde{x}, \tilde{v}) \in H^1(0,T)^n \times L^2(0,T)^{m-l}$.

It is easy to verify that $(\tilde{x} + \bar{x}, \tilde{v})$ is a solution of the following system:

$$\begin{cases} \dot{x}(t) = (A - BD^\dagger C)x(t) + BYv(t) + BD^\dagger p(t), \\ x(0) = x_0,\ x(T) = b, \end{cases} \qquad a.e.\ t \in (0,T).$$

Let $\tilde{u}(t) = Y\tilde{v}(t) + D^\dagger(p(t) - C(\tilde{x} + \bar{x})(t))$, then we have $D\tilde{u}(t) = p(t) - C(\tilde{x} + \bar{x})(t)$. Following Lemma 7.2 in [2], Assumption 1.2 implies that $(\tilde{x} + \bar{x}, \tilde{u}) \in H^1(0,T)^n \times L^2(0,T)^m$ is a solution of the following system:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ Cx(t) + Du(t) = p(t), \qquad a.e.\ t \in (0,T). \\ x(0) = x_0,\ \ x(T) = b, \end{cases} \tag{2.3}$$

If we set $p(t) = f(t) + \tilde{p}(t)$ in (2.3) for arbitrary $\tilde{p} \in L^2(0,T)^l$ with $\tilde{p}(t) \leq 0$ and $f(t)$ in (1.1), then the following problem

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ Cx(t) + Du(t) - f(t) \leq 0, \qquad a.e.\ t \in (0,T), \\ x(0) = x_0,\ \ x(T) = b, \end{cases} \tag{2.4}$$

has a solution $(x,u) \in H^1(0,T)^n \times L^2(0,T)^m$ for any $b \in \mathbb{R}^n$.

Now we show the solution set of the following stochastic constrained LCP is nonempty,

$$\begin{cases} \min\{x(T), \mathbb{E}[M(\xi)x(T) + q(\xi)]\} = 0, \\ \mathbb{E}[g(x(T),\xi)] \in K. \end{cases} \tag{2.5}$$

Following Corollary 3.5.6 and Theorem 3.11.6 in [12], the LCP in (2.5) has a solution from the assumption that the set $\mathcal{V}$ is nonempty and $\mathbb{E}[M(\xi)]$ is adequate or a Z-matrix. Let $x^*(T)$ be a sparse solution of the LCP in (2.5). If there is no J such that $\mathbb{E}[M_{J,J}(\xi)]$ is nonsingular, that is, $\mathcal{J} = \emptyset$, then by Lemma 2.1 and Lemma 2.2, $\|x^*(T)\|_0 = \|x^*(T)\|_1 = 0$. Hence, we have

$$\|x^*(T)\| \le \beta \|\mathbb{E}[q(\xi)]\|_1. \tag{2.6}$$

If there is J such that $x^*(T)_J > 0$ and $x^*(T)_{J^c} = 0$, where J^c is the complementarity set of an index set J, from Lemmas 2.1 and 2.2, we know that $\mathbb{E}[M_{J,J}(\xi)]$ is nonsingular and $x^*(T) = -(I - \Lambda + \Lambda\mathbb{E}[M(\xi)])^{-1}\Lambda\mathbb{E}[q(\xi)]$, where Λ is a diagonal matrix with $\Lambda_{i,i} = 1$, if $i \in J$ and $\Lambda_{i,i} = 0$, if $i \in J^c$. Moreover, from $\|(I - \Lambda + \Lambda\mathbb{E}[M(\xi)])^{-1}\Lambda\| \le \max\{\|(\mathbb{E}[M_{J,J}(\xi)])^{-1}\|_1 \mid J \in \mathcal{J}\}$, we obtain (2.6) for $\mathcal{J} \neq \emptyset$.

Therefore, from Assumption 1.1 and assumption (i) of this theorem, we have

$$\|\mathbb{E}[g(x^*(T),\xi)]\| \le \mathbb{E}[\kappa_2(\xi)]\|x^*(T)\| \le \mathbb{E}[\kappa_2(\xi)]\|x^*(T)\|_1 \le \beta\mathbb{E}[\kappa_2(\xi)]\|\mathbb{E}[q(\xi)]\|_1,$$

which implies that $\mathbb{E}[g(x^*(T),\xi)] \in K$. Hence the solution set of (2.5) is nonempty.

Similar to the proof of Theorem 5.1 in [2], we can derive the existence of optimal solutions to problem (1.1) if $\mathbb{E}[F(\cdot,\xi)]$ is bounded from below. ∎

Remark 2.4. The constrained LCP (2.5) may have multiple solutions or may not have a solution. If $\mathbb{E}[M(\xi)]$ is a P-matrix, then for any $\mathbb{E}[q(\xi)]$, the LCP in (2.5) has a unique solution $x(T)$. In such case, if $\mathbb{E}[g(x(T),\xi)] \in K$, then (2.5) has a unique solution, otherwise (2.5) does not have a solution. If $\mathbb{E}[M(\xi)]$ is a Z-matrix or an adequate matrix, the LCP in (2.5) may have multiple solutions, while some solutions can be bounded by $\beta\|\mathbb{E}[q(\xi)]\|_1$. When $\mathcal{B}(0, \beta\mathbb{E}[\kappa_2(\xi)]\|\mathbb{E}[q(\xi)]\|_1) \subseteq K$, some solutions of the LCP satisfy $\mathbb{E}[g(x(T),\xi)] \in K$ and thus the constrained LCP (2.5) is solvable. See the example in Section 5.

Remark 2.5. Assumption 1.2 is also used in [2] for the case $l < m$, which allows more freedom for the system controls. If $l = m$ and D is invertible, we can write $Cx(t) + Du(t) - f(t) = -v(t)$ with $v(t) \ge 0$ for a.e. $t \in [0,T]$, where $v \in L^2(0,T)^l$. Then the solvability of (2.4) becomes to find a solution pair $(x,v) \in H^1(0,T)^n \times L^2(0,T)^l$ with $v(t) \ge 0$ satisfying

$$\begin{cases} \dot{x}(t) = (A - BD^{-1}C)x(t) + BD^{-1}f(t) - BD^{-1}v(t), \\ x(0) = x_0, \;\; x(T) = b, \end{cases} \quad a.e.\ t \in (0,T). \tag{2.7}$$

It then requires the concept of positive controllability [3, 26]. Therefore, the solution set of (2.7) is nonempty for any $b \in \mathbb{R}^n$ under the following conditions:

(i) the block matrix

$$[BD^{-1} \;\; (A - BD^{-1}C)BD^{-1} \;\cdots\; (A - BD^{-1}C)^{n-1}BD^{-1}] \in \mathbb{R}^{n\times(nm)}$$

with n submatrices in $\mathbb{R}^{n\times m}$ possesses full row rank,

(ii) there is no real eigenvector $\mathbf{w} \in \mathbb{R}^n$ of $(A-BD^{-1}C)^\top$ such that $\mathbf{w}^\top BD^{-1}\mathbf{v} \geq 0$ for any $\mathbf{v} \in \mathbb{R}^m_+$.

Then there is a finite time T_0 such that the solution set of (2.4) is nonempty for any $b \in \mathbb{R}^n$ and $T \geq T_0$. Hence we can replace Assumption 1.2 in Theorem 2.3 by these two conditions for the case that $l = m$ and D is invertible.

3. Relaxation and sample average approximation (SAA).

In this section, we apply the relaxation and the SAA approach to solve (1.1). We consider an independent identically distributed (i.i.d) sample of $\xi(\omega)$, which is denoted by $\{\xi_1, \cdots, \xi_\nu\}$, and use the following relaxation and SAA problem to approximate problem (1.1):

$$
\begin{aligned}
&\min_{x,u} \ \frac{1}{\nu}\sum_{\ell=1}^{\nu} F(x(T),\xi_\ell) + \frac{1}{2}\|x-x_d\|_{L^2}^2 + \frac{\delta}{2}\|u-u_d\|_{L^2}^2 \\
&\text{s.t.} \begin{cases} \left.\begin{array}{l} \dot{x}(t) = Ax(t) + Bu(t), \\ Cx(t) + Du(t) - f(t) \leq 0, \end{array}\right\} a.e.\ t \in (0,T), \\ \left\| \min\left\{ x(T), \dfrac{1}{\nu}\displaystyle\sum_{\ell=1}^{\nu}[M(\xi_\ell)x(T) + q(\xi_\ell)] \right\} \right\| \leq \epsilon, \\ x(0) = x_0, \ \dfrac{1}{\nu}\displaystyle\sum_{\ell=1}^{\nu} g(x(T),\xi_\ell) \in K^\epsilon := \{z \,|\, \mathrm{dist}(z,K) \leq \epsilon\}, \end{cases}
\end{aligned} \tag{3.1}
$$

where $\epsilon > 0$ is a sufficiently small number.

By saying a property holds w.p.1 for sufficiently large ν, we mean that there is a set $\Omega_0 \subset \Omega$ of $\mathcal{P}$-measure zero such that for all $\omega \in \Omega\backslash\Omega_0$ there exists a positive integer $\nu^*(\omega)$ such that the property holds for all $\nu \geq \nu^*(\omega)$.

3.1. Convergence of the relaxation and SAA.

In this subsection, we show the existence of a solution of problem (3.1), and its convergence as $\epsilon \downarrow 0$ and $\nu \to \infty$.

THEOREM 3.1. *Suppose that the conditions of Theorem 2.3 hold. Then for any* $\epsilon > 0$, *the SAA problem* (3.1) *has an optimal solution* $(x^{\epsilon,\nu}, u^{\epsilon,\nu}) \in H^1(0,T)^n \times L^2(0,T)^m$ *w.p.1 for sufficiently large* ν.

Proof. Since the solution set of the linear control system (2.4) is nonempty for any $b \in \mathbb{R}^n$, for the existence of a feasible solution to the SAA problem (3.1), it suffices to show that for any given $\epsilon > 0$ the solution set of the following system

$$
\begin{cases} \left\| \min\left\{ x(T), \dfrac{1}{\nu}\displaystyle\sum_{\ell=1}^{\nu}[M(\xi_\ell)x(T) + q(\xi_\ell)] \right\} \right\| \leq \epsilon, \\ \dfrac{1}{\nu}\displaystyle\sum_{\ell=1}^{\nu} g(x(T),\xi_\ell) \in K^\epsilon \end{cases} \tag{3.2}
$$

is nonempty w.p.1 for sufficiently large ν.

Let $x^*(T)$ be a sparse solution of the LCP in (2.5). From Theorem 2.3, we know that $x^*(T)$ satisfies (2.5). By the strong Law of Large Number, for sufficiently large ν, $x^*(T)$ is a solution of (3.2). It concludes with any given $\epsilon > 0$ that the solution set of the system (3.2) is nonempty w.p.1 for sufficiently large ν.

Since $\mathbb{E}[F(\cdot,\xi)]$ is bounded from below, we can also obtain that $\frac{1}{\nu}\sum_{\ell=1}^{\nu} F(\cdot,\xi_\ell)$ is bounded from below with sufficiently large ν. The existence of optimal solutions to problem (3.1) is similar to the proof of Theorem 2.3. ∎

We define the objective functions of problems (1.1) and (3.1), respectively as the following

$$\Phi(x,u) = \mathbb{E}[F(x(T),\xi)] + \frac{1}{2}\|x - x_d\|_{L^2}^2 + \frac{\delta}{2}\|u - u_d\|_{L^2}^2,$$

and

$$\Phi^\nu(x,u) = \frac{1}{\nu}\sum_{\ell=1}^{\nu} F(x(T),\xi_\ell) + \frac{1}{2}\|x - x_d\|_{L^2}^2 + \frac{\delta}{2}\|u - u_d\|_{L^2}^2,$$

where $(x,u) \in H^1(0,T)^n \times L^2(0,T)^m$, and $\nu > 0$.

Let $Z \subseteq R^n$ be an open set, $\bar{R} = [-\infty, \infty]$ and $\mathbb{N} = \{1, 2, 3, \cdots\}$.

Definition 3.2. *([22]) A sequence of functions $\{g^k : Z \to \bar{R},\ k \in \mathbb{N}\}$ epiconverges to $g : Z \to \bar{R}$ if for all $z \in Z$,*

(i) $\liminf_{k\to\infty} g^k(z^k) \geq g(z)$ for all $z^k \to z$, and

(ii) $\limsup_{k\to\infty} g^k(z^k) \leq g(z)$ for some $z^k \to z$.

Definition 3.3. *([16]) A function $g : \Xi \times Z \to \bar{R}$ is a random lower semicontinuous (lsc) function if g is jointly measurable in (ξ, z) and $g(\xi,\cdot)$ is lsc for every $\xi \in \Xi$.*

Definition 3.4. *([16]) A sequence of random lsc function $\{g^k : \Xi \times Z \to \bar{R}, k \in [K]\}$ epiconverges to $g : \Xi \times Z \to \bar{R}$ almost surely, if for a.e. $\xi \in \Xi$, $\{g^k(\xi,\cdot) : Z \to \bar{R}, k \in \mathbb{N}\}$ epiconverges to $g : Z \to \bar{R}$.*

Since $F(\cdot,\xi)$ is a smoothing function for a.e. $\xi \in \Xi$, following the proof of Lemma 3.5 in [8], we can have the following lemma.

Lemma 3.5. *Let $\mathcal{C}_1 \times \mathcal{C}_2$ denote a compact subset of $H^1(0,T)^n \times L^2(0,T)^m$. It holds that Φ^ν epiconverges to Φ w.p.1 over $\mathcal{C}_1 \times \mathcal{C}_2$ as $\nu \to \infty$.*

Let $\mathcal{Z}^{\epsilon,\nu}$ and $\mathcal{Z}$ denote the solution sets of (3.2) and (2.5), respectively. Let $\mathcal{S}^{\epsilon,\nu}$ and $\mathcal{S}$ be the feasible solution sets, and $\hat{\mathcal{S}}^{\epsilon,\nu}$ and $\hat{\mathcal{S}}$ be optimal solution sets of (3.1) and (1.1), respectively.

Theorem 3.6. *Suppose that the conditions of Theorem 2.3 and K is bounded. Assume that there are $\bar{\epsilon} > 0$, $\gamma > 0$ and $\eta > 0$ such that for $z \in \mathbb{R}^n_{-\bar{\epsilon}} := \{z \in \mathbb{R}^n : z_i \geq -\bar{\epsilon},\ i \in [n]\}$,*

$$\gamma + \|\mathbb{E}[g(z,\xi)]\| \geq \eta\|z\|. \tag{3.3}$$

Then it holds that $\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty} \mathbb{D}(\mathcal{Z}^{\epsilon,\nu}, \mathcal{Z}) = 0$ w.p.1, $\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty} \mathbb{D}(\mathcal{S}^{\epsilon,\nu}, \mathcal{S}) = 0$ w.p.1. and $\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty} \mathbb{D}(\hat{\mathcal{S}}^{\epsilon,\nu}, \hat{\mathcal{S}}) = 0$ w.p.1.

Proof. From Theorem 2.3, we know that $\mathcal{Z}$ is nonempty. And by Theorem 3.1, for any given $\epsilon > 0$, $\mathcal{Z}^{\epsilon,\nu}$ is nonempty w.p.1 for sufficiently large ν. Denote $\mathcal{Z}^\epsilon$ the solution set of the following problem for any given $\epsilon > 0$

$$\begin{cases} \|\min\{x(T), \mathbb{E}[M(\xi)x(T) + q(\xi)]\}\| \leq \epsilon, \\ \mathbb{E}[g(x(T),\xi)] \in K^\epsilon. \end{cases} \tag{3.4}$$

It is obvious that $\mathcal{Z} \subseteq \mathcal{Z}^\epsilon$ and then $\mathcal{Z}^\epsilon$ is nonempty for any given $\epsilon > 0$. Since K is compact, $K^{\bar{\epsilon}}$ is a compact set, which means that there is $\rho_{\bar{\epsilon}} > 0$ such that $\|y\| \leq \rho_{\bar{\epsilon}}$ for any $y \in K^{\bar{\epsilon}}$. Obviously, $\mathcal{Z}^\epsilon \subset \mathcal{Z}^{\bar{\epsilon}} \subset \mathbb{R}^n_{-\bar{\epsilon}}$ for any $\epsilon \leq \bar{\epsilon}$. By condition (3.3), for any $z \in \mathcal{Z}^\epsilon$ with $\epsilon \leq \bar{\epsilon}$,

$$\eta\|z\| \leq \|\mathbb{E}[g(z,\xi)]\| + \gamma \leq \rho_{\bar{\epsilon}} + \gamma.$$

Hence we have, for any $x(T) \in \mathcal{Z}^\epsilon$ with $\epsilon \le \bar{\epsilon}$,

$$\|x(T)\| \le \frac{\rho_{\bar{\epsilon}} + \gamma}{\eta}.$$

Similarly, by (3.3) and the strong Law of Large Number, we have that for any $z \in \mathbb{R}^n_{-\bar{\epsilon}}$

$$2\gamma + \left\| \frac{1}{\nu} \sum_{\ell=1}^{\nu} g(z, \xi_\ell) \right\| \ge \eta \|z\|$$

w.p.1 for sufficiently large ν. Since $\frac{1}{\nu}\sum_{\ell=1}^{\nu} g(x(T), \xi_\ell) \in K^{\bar{\epsilon}}$ and $\mathcal{Z}^{\epsilon,\nu} \subset \mathbb{R}^n_{-\bar{\epsilon}}$ for any $\epsilon \le \bar{\epsilon}$, we obtain that for any $x(T) \in \mathcal{Z}^{\epsilon,\nu}$ with $\epsilon \le \bar{\epsilon}$,

$$\|x(T)\| \le \frac{\rho_{\bar{\epsilon}} + 2\gamma}{\eta}$$

w.p.1 for sufficiently large ν. Therefore, for any $\epsilon \le \bar{\epsilon}$, there is a compact set $\mathcal{X}$ such that $\mathcal{Z} \subseteq \mathcal{X}$ and $\mathcal{Z}^{\epsilon,\nu} \subseteq \mathcal{X}$ w.p.1 for sufficiently large ν.

Let

$$\phi(x(T)) := \min\{x(T), \mathbb{E}[M(\xi)x(T) + q(\xi)]\} \quad \text{and} \quad \psi(x(T)) := \mathbb{E}[g(x(T), \xi)].$$

For $x(T) \in \mathcal{Z}$, $\phi(x(T)) = 0$ and $\psi(x(T)) \in K$. From (3.2), for $x(T) \in \mathcal{Z}^{\epsilon,\nu}$, there are $v^\nu \in \mathbb{R}^n, w^\nu \in \mathbb{R}^k$ with $\|v^\nu\| \le \epsilon$ and $\|w^\nu\| \le \epsilon$ w.p.1 for sufficiently large ν such that

$$\begin{aligned} \phi_\epsilon^\nu(x(T)) &:= \min\left\{ x(T), \frac{1}{\nu}\sum_{\ell=1}^{\nu}[M(\xi_\ell)x(T) + q(\xi_\ell)] \right\} + v^\nu = 0, \\ \psi_\epsilon^\nu(x(T)) &:= \frac{1}{\nu}\sum_{\ell=1}^{\nu} g(x(T), \xi_\ell) + w^\nu \in K. \end{aligned}$$

Since ϕ and ψ are continuous, and $M(\cdot), q(\cdot)$ and $g(x(T), \cdot)$ satisfy Assumption 1.1, we have ϕ_ϵ^ν and ψ_ϵ^ν converge to ϕ and ψ uniformly w.p.1, respectively on the compact set $\mathcal{X}$ as $\epsilon \downarrow 0$ and $\nu \to \infty$, that is,

$$\lim_{\epsilon\downarrow 0} \lim_{\nu\to\infty} \max_{x(T)\in\mathcal{X}} \|\phi_\epsilon^\nu(x(T)) - \phi(x(T))\| = 0, \;\; w.p.1$$

and

$$\lim_{\epsilon\downarrow 0} \lim_{\nu\to\infty} \max_{x(T)\in\mathcal{X}} \|\psi_\epsilon^\nu(x(T)) - \psi(x(T))\| = 0, \;\; w.p.1.$$

Therefore, following Theorem 5.12 in [23], $\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty} \mathbb{D}(\mathcal{Z}^{\epsilon,\nu}, \mathcal{Z}) = 0$ w.p.1.

Now we show $\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty} \mathbb{D}(\mathcal{S}^{\epsilon,\nu}, \mathcal{S}) = 0$ holds w.p.1. Note that $\mathcal{S}^{\epsilon,\nu}$ and $\mathcal{S}$ are two nonempty closed sets. Obviously, two nonempty closed sets $\mathcal{S}$ and $\mathcal{S}^{\epsilon,\nu}$ are the solution sets of problem (2.4) with terminal sets $\mathcal{Z}$ and $\mathcal{Z}^{\epsilon,\nu}$, respectively. For any $p \in L^2(0,T)^l$, the pair $(\|x\|_{H^1}, \|u\|_{L^2})$, where (x, u) is a solution of problem (2.3), is uniquely defined by the terminal point $x(T)$. In addition, it is clear that a solution (x, u) of problem (2.3) is continuous with respect to the terminal point $x(T)$. Hence, for any $(x^{\epsilon,\nu}, u^{\epsilon,\nu}) \in \mathcal{S}^{\epsilon,\nu}$ and $(x, u) \in \mathcal{S}$, we have $(x^{\epsilon,\nu}, u^{\epsilon,\nu}) \to (x, u)$ w.p.1 in the norm $\|\cdot\|_{H^1} \times \|\cdot\|_{L^2}$ when $x^{\epsilon,\nu}(T) \to x(T)$ w.p.1 as $\epsilon \downarrow 0$ and $\nu \to \infty$. It then concludes $\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty} \mathbb{D}(\mathcal{S}^{\epsilon,\nu}, \mathcal{S}) = 0$ w.p.1.

It is clear that from $\hat{\mathcal{S}} \subseteq \mathcal{S}$, $\hat{\mathcal{S}}^{\epsilon,\nu} \subseteq \mathcal{S}^{\epsilon,\nu}$ and $\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty} \mathbb{D}(\mathcal{S}^{\epsilon,\nu}, \mathcal{S}) = 0$ w.p.1, we have, for any $(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) \in \hat{\mathcal{S}}^{\epsilon,\nu}$, there is $(\hat{x}, \hat{u}) \in \mathcal{S}$ such that $(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) \to (\hat{x}, \hat{u})$ w.p.1 in the norm $\|\cdot\|_{H^1} \times \|\cdot\|_{L^2}$ as $\epsilon \downarrow 0$ and $\nu \to \infty$. In addition, according to Theorem 2.5 in [1], we obtain $(\hat{x}, \hat{u}) \in \hat{\mathcal{S}}$ by the epiconvergence of Φ^ν to Φ w.p.1, which implies $\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty} \mathbb{D}(\hat{\mathcal{S}}^{\epsilon,\nu}, \hat{\mathcal{S}}) = 0$ w.p.1. ∎

3.2. Asymptotics of the SAA optimal value. We introduce the relaxation of problem (1.1) with a parameter $\epsilon > 0$ as follows

$$\min_{x,u} \ \Phi(x,u)$$
$$\text{s.t.} \begin{cases} \left.\begin{array}{l} \dot{x}(t) = Ax(t) + Bu(t), \\ Cx(t) + Du(t) - f(t) \leq 0, \end{array}\right\} a.e.\ t \in (0,T), \\ \|\min\{x(T), \mathbb{E}[M(\xi)x(T) + q(\xi)]\}\| \leq \epsilon, \\ x(0) = x_0,\ \mathbb{E}[g(x(T),\xi)] \in K^\epsilon. \end{cases} \tag{3.5}$$

Recall that $\mathcal{Z}^\epsilon$ is the solution set of the terminal constraints of (3.5). Denote by $\mathcal{S}^\epsilon$ and $\hat{\mathcal{S}}^\epsilon$ the feasible solution set and optimal solution set of (3.5), respectively. Recall that $\mathcal{Z}$ is the solution set of (2.5), and $\mathcal{S}$ and $\hat{\mathcal{S}}$ are the feasible solution set and optimal solution set of (1.1), respectively. It is clear that $\mathcal{Z} \subseteq \mathcal{Z}^\epsilon$ and $\mathcal{S} \subseteq \mathcal{S}^\epsilon$, which mean that $\Phi(\hat{x}^\epsilon, \hat{u}^\epsilon) \leq \Phi(\hat{x}, \hat{u})$ for any $(\hat{x}^\epsilon, \hat{u}^\epsilon) \in \hat{\mathcal{S}}^\epsilon$ and $(\hat{x}, \hat{u}) \in \hat{\mathcal{S}}$. Therefore, $\mathcal{Z}^\epsilon$, $\mathcal{S}^\epsilon$ and $\hat{\mathcal{S}}^\epsilon$ are nonempty since $\mathcal{Z}$ and $\mathcal{S}$ are nonempty.

According to Theorem 3.6, we also conclude that $\mathcal{Z}^\epsilon$ and $\hat{\mathcal{S}}^\epsilon$ are compact. It can also be derived that $\lim_{\epsilon\downarrow 0} \mathbb{D}(\mathcal{Z}^\epsilon, \mathcal{Z}) = 0$, $\lim_{\epsilon\downarrow 0} \mathbb{D}(\mathcal{S}^\epsilon, \mathcal{S}) = 0$ and $\lim_{\epsilon\downarrow 0} \mathbb{D}(\hat{\mathcal{S}}^\epsilon, \hat{\mathcal{S}}) = 0$. It is clear that (3.1) is the corresponding SAA problem of (3.5). By Theorem 3.6, we conclude that $\lim_{\nu\to\infty} \mathbb{D}(\mathcal{Z}^{\epsilon,\nu}, \mathcal{Z}^\epsilon) = 0$ w.p.1, $\lim_{\nu\to\infty} \mathbb{D}(\mathcal{S}^{\epsilon,\nu}, \mathcal{S}^\epsilon) = 0$ w.p.1 and $\lim_{\nu\to\infty} \mathbb{D}(\hat{\mathcal{S}}^{\epsilon,\nu}, \hat{\mathcal{S}}^\epsilon) = 0$ w.p.1.

In the rest of this section, we study the asymptotics of optimal value of the SAA problem (3.1) for a fixed $\epsilon > 0$.

Since $\min\{x(T), \mathbb{E}[M(\xi)x(T)+q(\xi)]\} = 0$ and $\mathbb{E}[g(x(T),\xi)] \in K$ for any $x(T) \in \mathcal{Z}$, we have $\mathcal{Z} \subseteq \mathrm{int}\mathcal{Z}^\epsilon$, which means that $\mathrm{int}\mathcal{Z}^\epsilon \neq \emptyset$. Let

$$\hat{\mathcal{Z}} = \{x(T) : (x,u) \in \hat{\mathcal{S}}\} \quad \text{and} \quad \hat{\mathcal{Z}}^\epsilon = \{x(T) : (x,u) \in \hat{\mathcal{S}}^\epsilon\}.$$

Obviously, we have $\hat{\mathcal{Z}} \subseteq \mathrm{int}\mathcal{Z}^\epsilon$ and $\lim_{\epsilon\downarrow 0} \mathbb{D}(\hat{\mathcal{Z}}^\epsilon, \hat{\mathcal{Z}}) = 0$. We give the following assumptions.

Assumption 3.7. The set $\hat{\mathcal{Z}}$ is a singleton.

Assumption 3.8. (i) There exists a nonnegative measurable function $\kappa_1(\xi)$ with $\mathbb{E}[\kappa_1^2(\xi)] < \infty$ such that for any z_1, $z_2 \in \mathbb{R}^n$ and $\xi \in \Xi$,

$$|F(z_1,\xi) - F(z_2,\xi)| \leq \kappa_1(\xi)\|z_1 - z_2\|,$$

and $\mathbb{E}[F^2(z,\xi)] < \infty$ for any $z \in \mathbb{R}^n$.

(ii) The function $\mathbb{E}[F(\cdot,\xi)]$ is a strongly convex function, that is, there is a constant $\mu > 0$ such that, for any z_1, $z_2 \in \mathbb{R}^n$ and $\tau \in (0,1)$,

$$\mathbb{E}[F((1-\tau)z_1 + \tau z_2,\xi)] \leq (1-\tau)\mathbb{E}[F(z_1,\xi)] + \tau\mathbb{E}[F(z_2,\xi)] - \frac{\mu\tau(1-\tau)}{2}\|z_1 - z_2\|^2.$$

THEOREM 3.9. *Suppose that the conditions of Theorem 3.6, Assumption 3.7 and Assumption 3.8 hold. Let $(\hat{x}^\epsilon, \hat{u}^\epsilon)$ and $(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu})$ be optimal solutions of* (3.5) *and* (3.1)*, respectively. Then for sufficiently small ϵ, we have*

$$\sqrt{\nu}(\Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) - \Phi(\hat{x}^\epsilon, \hat{u}^\epsilon)) \quad \xrightarrow{D} \quad \mathcal{N}(0, \sigma^2(\hat{x}^\epsilon(T))),$$

where “ $\xrightarrow{D}$ ” denotes convergence in distribution and $\mathcal{N}(0, \sigma^2(\hat{x}^\epsilon(T)))$ denotes the normal distribution with mean 0 and variance $\sigma^2(\hat{x}^\epsilon(T)) := \mathbb{V}ar[F(\hat{x}^\epsilon(T),\xi)]$.

Proof. Since $\hat{\mathcal{Z}}$ is a singleton, $\hat{\mathcal{Z}} \subseteq \text{int}\mathcal{Z}^\epsilon$ and $\lim_{\epsilon\downarrow 0}\mathbb{D}(\hat{\mathcal{Z}}^\epsilon, \hat{\mathcal{Z}}) = 0$, we have $\hat{\mathcal{Z}}^\epsilon \subseteq \text{int}\mathcal{Z}^\epsilon$ for sufficiently small ϵ, which means that there is a convex set $\mathcal{Z}_{\mathcal{X}}$ such that $\hat{\mathcal{Z}}^\epsilon \subseteq \mathcal{Z}_{\mathcal{X}} \subseteq \mathcal{Z}^\epsilon$ for sufficiently small ϵ. We can also obtain that $\hat{\mathcal{Z}}^\epsilon$ is a singleton for sufficiently small ϵ under Assumption 3.8(ii). We argue it by contradiction. Suppose $(\hat{x}^\epsilon, \hat{u}^\epsilon)$ and $(\check{x}^\epsilon, \check{u}^\epsilon)$ are two optimal solutions of (3.5) with $\hat{x}^\epsilon(T) \neq \check{x}^\epsilon(T)$. Then $(x^\epsilon_\tau, u^\epsilon_\tau) := ((1-\tau)\hat{x}^\epsilon + \tau\check{x}^\epsilon, (1-\tau)\hat{u}^\epsilon + \tau\check{u}^\epsilon)$ with $\tau \in (0,1)$ is also a feasible solution of (3.5), since $x^\epsilon_\tau(T) \in \mathcal{Z}_{\mathcal{X}} \subseteq \mathcal{Z}^\epsilon$. Moreover,

$$\Phi(x^\epsilon_\tau, u^\epsilon_\tau) \leq (1-\tau)\Phi(\hat{x}^\epsilon, \hat{u}^\epsilon) + \tau\Phi(\check{x}^\epsilon, \check{u}^\epsilon) - \frac{\mu\tau(1-\tau)}{2}\|\hat{x}^\epsilon(T) - \check{x}^\epsilon(T)\|^2,$$

which means $\Phi(x^\epsilon_\tau, u^\epsilon_\tau) < \Phi(\hat{x}^\epsilon, \hat{u}^\epsilon)$ since $\Phi(\hat{x}^\epsilon, \hat{u}^\epsilon) = \Phi(\check{x}^\epsilon, \check{u}^\epsilon)$ and $\hat{x}^\epsilon(T) \neq \check{x}^\epsilon(T)$. It contradicts the assumption that $(\hat{x}^\epsilon, \hat{u}^\epsilon)$ is an optimal solution of (3.5), and then we know that $\hat{\mathcal{Z}}^\epsilon$ is a singleton for sufficiently small ϵ.

In the following argument, $\epsilon > 0$ is a fixed number such that $\hat{\mathcal{Z}}^\epsilon$ is singleton and $\hat{\mathcal{Z}}^\epsilon \subseteq \text{int}\mathcal{Z}^\epsilon$. Denote $\hat{\mathcal{Z}}^{\epsilon,\nu} = \{x(T) : (x,u) \in \hat{\mathcal{S}}^{\epsilon,\nu}\}$. We then obtain that $\lim_{\nu\to\infty}\mathbb{D}(\hat{\mathcal{Z}}^{\epsilon,\nu}, \hat{\mathcal{Z}}^\epsilon) = 0$ w.p.1 and $\hat{\mathcal{Z}}^\epsilon \subseteq \text{int}\mathcal{Z}^{\epsilon,\nu}$ w.p.1 for sufficiently large ν according to $\lim_{\nu\to\infty}\mathbb{D}(\mathcal{Z}^{\epsilon,\nu}, \mathcal{Z}^\epsilon) = 0$ w.p.1. Therefore, there is a $(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) \in \hat{\mathcal{S}}^{\epsilon,\nu}$ such that $\hat{x}^{\epsilon,\nu}(T) \in \text{int}\mathcal{Z}^{\epsilon,\nu}$ for sufficiently large ν, which implies that, there is a compact set $\mathcal{X}$ such that $\hat{\mathcal{Z}}^\epsilon \subseteq \mathcal{X} \subseteq \mathcal{Z}^\epsilon$ and $\hat{x}^{\epsilon,\nu}(T) \in \mathcal{X} \subseteq \mathcal{Z}^{\epsilon,\nu}$ w.p.1 for sufficiently large ν.

The solution (x,u) of ODE in (2.4) is continuous with respect to the state terminal value $x(T)$ and the pair $(\|x\|_{H^1}, \|u\|_{L^2})$ is uniquely defined by $x(T)$. Therefore, there is a compact set $\mathfrak{X}$ such that $\hat{\mathcal{S}}^\epsilon \subseteq \mathfrak{X} \subseteq \mathcal{S}^\epsilon$ and $\mathfrak{X} \subseteq \mathcal{S}^{\epsilon,\nu}$ with $\hat{\mathcal{S}}^{\epsilon,\nu} \cap \mathfrak{X} \neq \emptyset$ w.p.1 for sufficiently large ν. To derive the error of approximation for optimal value of (3.1) to that of (3.5), it suffices to investigate the error approximation for optimal value of the following problem

$$\min_{(x,u)\in\mathfrak{X}} \Phi(x,u) \tag{3.6}$$

and its SAA problem

$$\min_{(x,u)\in\mathfrak{X}} \Phi^\nu(x,u), \tag{3.7}$$

where Φ and Φ^ν are defined in (3.5) and (3.1), respectively. Clearly, $\mathfrak{X} \subseteq \mathcal{S}^\epsilon$ with $\hat{\mathcal{S}}^\epsilon \cap \mathfrak{X} \neq \emptyset$ and $\mathfrak{X} \subseteq \mathcal{S}^{\epsilon,\nu}$ with $\hat{\mathcal{S}}^{\epsilon,\nu} \cap \mathfrak{X} \neq \emptyset$ w.p.1 for sufficiently large ν mean that an optimal solution of (3.6) is an optimal solution of (3.5), and an optimal solution of (3.7) is also an optimal solution of (3.1). Therefore, according to Theorem 5.7 in [23], we can obtain that, under Assumption 3.8,

$$\begin{aligned}
&\sqrt{\nu}(\Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) - \Phi(\hat{x}^\epsilon, \hat{u}^\epsilon)) \\
&= \sqrt{\nu}(\frac{1}{\nu}\sum_{\ell=1}^{\nu} F(\hat{x}^{\epsilon,\nu}(T), \xi_\ell) + \frac{1}{2}\|\hat{x}^{\epsilon,\nu} - x_d\|^2_{L^2} + \frac{\delta}{2}\|\hat{u}^{\epsilon,\nu} - u_d\|^2_{L^2} \\
&\qquad -\mathbb{E}[F(\hat{x}^\epsilon(T), \xi)] - \frac{1}{2}\|\hat{x}^\epsilon - x_d\|^2_{L^2} - \frac{\delta}{2}\|\hat{u}^\epsilon - u_d\|^2_{L^2}) \\
&\xrightarrow{D} \inf_{(x,u)\in\hat{\mathcal{S}}^\epsilon} \mathcal{Y}(x,u),
\end{aligned}$$

where $\mathcal{Y}(x,u)$ has a normal distribution with mean 0 and variance $\mathbb{V}\text{ar}[F(x(T), \xi)]$ with $(x,u) \in \hat{\mathcal{S}}^\epsilon$. Since $\hat{\mathcal{Z}}^\epsilon = \{\hat{x}^\epsilon(T)\}$ is a singleton, $\mathcal{Y}(x,u)$ for any $(x,u) \in \hat{\mathcal{S}}^\epsilon$ has the same normal distribution with mean 0 and variance $\mathbb{V}\text{ar}[F(\hat{x}^\epsilon(T), \xi)]$. It then concludes our desired result. ∎

4. The time-stepping method. We now adopt the time-stepping method for solving problem (3.1) with a fixed sample $\{\xi_1, \dots, \xi_\nu\}$, which uses a finite-difference formula to approximate the time derivative $\dot{x}$. It begins with the division of the time interval $[0, T]$ into N subintervals for a fixed step size $h = T/N = t_{i+1} - t_i$ where $i = 0, \cdots, N-1$. Starting from $\mathbf{x}_0^\nu = x_0$, we compute two finite sets of vectors $\{\mathbf{x}_1^{\epsilon,\nu}, \mathbf{x}_2^{\epsilon,\nu}, \cdots, \mathbf{x}_N^{\epsilon,\nu}\} \subset \mathbb{R}^n$ and $\{\mathbf{u}_1^{\epsilon,\nu}, \mathbf{u}_2^{\epsilon,\nu}, \cdots, \mathbf{u}_N^{\epsilon,\nu}\} \subset \mathbb{R}^m$ in the following manner:

$$
\begin{aligned}
&\min_{\{\mathbf{x}_i, \mathbf{u}_i\}_{i=1}^N} \ \frac{1}{\nu}\sum_{\ell=1}^{\nu} F(\mathbf{x}_N, \xi_\ell) + \frac{h}{2}\sum_{i=1}^{N}\left(\|\mathbf{x}_i - x_{d,i}\|^2 + \delta\|\mathbf{u}_i - u_{d,i}\|^2\right) \\
&\text{s.t.} \begin{cases} \left.\begin{array}{l} \mathbf{x}_{i+1} - \mathbf{x}_i = hA\mathbf{x}_{i+1} + hB\mathbf{u}_{i+1}, \\ C\mathbf{x}_{i+1} + D\mathbf{u}_{i+1} - f_{i+1} \le 0, \end{array}\right\} i = 0, 1, \cdots, N-1, \\ \left\|\min\left\{\mathbf{x}_N, \dfrac{1}{\nu}\displaystyle\sum_{\ell=1}^{\nu}[M(\xi_\ell)\mathbf{x}_N + q(\xi_\ell)]\right\}\right\| \le \epsilon, \\ \dfrac{1}{\nu}\displaystyle\sum_{\ell=1}^{\nu} g(\mathbf{x}_N, \xi_\ell) \in K^\epsilon, \end{cases}
\end{aligned} \tag{4.1}
$$

where $\epsilon > 0$ is a sufficiently small number, $x_{d,i} = x_d(t_i)$, $u_{d,i} = u_d(t_i)$ and $f_i = f(t_i)$ for $i \in [N]$.

THEOREM 4.1. *Suppose that the conditions of Theorem 2.3 hold, then for any $\epsilon >$ 0, problem* (4.1) *has an optimal solution w.p.1 for sufficiently large ν and sufficiently small h.*

Proof. Theorem 3.1 has shown that the solution set of (3.2) with any $\epsilon > 0$ is nonempty w.p.1 for sufficiently large ν. About the existence of feasible solution to problem (4.1), it suffices to show that the following problem has a solution for any $b \in \mathbb{R}^n$,

$$
\begin{cases} \left.\begin{array}{l} \mathbf{x}_{i+1} = \mathbf{x}_i + hA\mathbf{x}_{i+1} + hB\mathbf{u}_{i+1}, \\ C\mathbf{x}_{i+1} + D\mathbf{u}_{i+1} - f_{i+1} \le 0, \end{array}\right\} i = 0, 1, \cdots, N-1, \\ \mathbf{x}_0 = x_0, \qquad \mathbf{x}_N = b. \end{cases} \tag{4.2}
$$

Firstly, denote $A_h = I - h(A - BD^\dagger C)$. It is obvious that all eigenvalues of A_h are $1 - h\lambda_i$ with $i \in [n]$, where λ_i with $i \in [n]$ are the eigenvalues of $A - BD^\dagger C$. We then obtain that all eigenvalues of A_h are nonzero for sufficiently small h and A_h is nonsingular. Similar to the proof of Theorem 2.3, from $\mathbf{x}_{i+1} = \mathbf{x}_i + h(A - BD^\dagger C)\mathbf{x}_{i+1} + hBD^\dagger p_{i+1}$, the following iteration with $\mathbf{x}_0 = x_0$,

$$
\mathbf{x}_{i+1} = A_h^{-1}(\mathbf{x}_i + hBD^\dagger p_{i+1}), \quad i = 0, 1, \cdots, N-1
$$

generates a solution $\{\bar{\mathbf{x}}_i\}_{i=1}^N$ of the system with $\mathbf{x}_0 = x_0$,

$$
\mathbf{x}_{i+1} = \mathbf{x}_i + hA\mathbf{x}_{i+1} + hB\mathbf{u}_{i+1}, \quad C\mathbf{x}_{i+1} + D\mathbf{u}_{i+1} = p_{i+1}, \quad i = 0, 1, \cdots, N-1,
$$

for any given $p_i \in \mathbb{R}^l$, $i = 1, \dots, N$.

From Assumption 1.2 and the nonsingularity of A_h, we know that the matrix $\tilde{\mathcal{R}}_d := [BY \ A_h BY \ \cdots \ A_h^{n-1}BY]$ has full row rank n. Hence the matrix $\mathcal{R}_d := [hA_h^{-1}BY \ h(A_h^{-1})^2BY \ \cdots \ h(A_h^{-1})^nBY]$ has full row rank n. According to Theorem

3.1.1 in [15], the system with $\mathbf{x}_0 = 0$,

$$\begin{cases} \mathbf{x}_{i+1} = A_h^{-1}(\mathbf{x}_i + hBYv_{i+1}),\ i = 0, 1, \cdots, N-1, \\ \mathbf{x}_N = b - \bar{\mathbf{x}}_N, \end{cases}$$

admits a solution $\{\tilde{\mathbf{x}}_i, \tilde{v}_i\}_{i=1}^N$ for any $b \in \mathbb{R}^n$. Therefore, $\{\tilde{\mathbf{x}}_i + \bar{\mathbf{x}}_i, \tilde{v}_i\}_{i=1}^N$ is a solution of the following equation

$$\begin{cases} \mathbf{x}_{i+1} = A_h^{-1}(\mathbf{x}_i + hBYv_{i+1} + hBD^\dagger p_{i+1}),\ i = 0, 1, \cdots, N-1, \\ \mathbf{x}_0 = x_0, \qquad \mathbf{x}_N = b. \end{cases}$$

Let $\tilde{\mathbf{u}}_i = Y\tilde{v}_i + D^\dagger(p_i - C(\tilde{\mathbf{x}}_i + \bar{\mathbf{x}}_i))$. Then it is easy to verify that $\{\tilde{\mathbf{x}}_i + \bar{\mathbf{x}}_i, \tilde{\mathbf{u}}_i\}_{i=1}^N$ is a solution of (4.2) by setting $p_i = f_i + \tilde{p}_i$ for any $\tilde{p}_i \le 0$.

Since $\mathbb{E}[F(\cdot, \xi)]$ is bounded from below, we can also obtain that $\frac{1}{\nu}\sum_{\ell=1}^{\nu} F(\cdot, \xi_\ell)$ is also bounded from below with sufficiently large ν. Similar to Theorem 5.1 in [2], we can prove a minimizing sequence tends to an optimal solution of (4.1), which shows the existence of optimal solutions to (4.1) with any $\epsilon > 0$ for sufficiently large ν and sufficiently small h. ∎

Let $\{\mathbf{x}_i^{\epsilon,\nu}, \mathbf{u}_i^{\epsilon,\nu}\}_{i=1}^N$ be a solution of (4.1). We define a piecewise linear function $x_h^{\epsilon,\nu}$ and a piecewise constant function $u_h^{\epsilon,\nu}$ on $[0, T]$ as below:

$$x_h^{\epsilon,\nu}(t) = \mathbf{x}_i^{\epsilon,\nu} + \frac{t - t_i}{h}(\mathbf{x}_{i+1}^{\epsilon,\nu} - \mathbf{x}_i^{\epsilon,\nu}), \quad u_h^{\epsilon,\nu}(t) = \mathbf{u}_{i+1}^{\epsilon,\nu}, \ \ \forall\, t \in (t_i, t_{i+1}]. \tag{4.3}$$

Denote $\hat{\mathcal{S}}_h^{\epsilon,\nu}$ the set of $(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu}) \in H^1(0,T)^n \times L^2(0,T)^m$, where $(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu})$ are defined in (4.3) based on an optimal solution $\{\hat{\mathbf{x}}_i^{\epsilon,\nu}, \hat{\mathbf{u}}_i^{\epsilon,\nu}\}_{i=1}^N$ of (4.1). Define, for any $(x, u) \in H^1(0,T)^n \times L^2(0,T)^m$, $\nu > 0$ and $h > 0$,

$$\Phi_h^\nu(x, u) = \frac{1}{\nu}\sum_{\ell=1}^{\nu} F(x(T), \xi_\ell) + \frac{h}{2}\sum_{i=1}^{N}\left(\|x(t_i) - x_{d,i}\|^2 + \delta\|u(t_i) - u_{d,i}\|^2\right).$$

THEOREM 4.2. *Suppose that the conditions of Theorem 3.6 hold, then we have*

$$\lim_{\epsilon\downarrow 0}\lim_{\nu\to\infty}\lim_{h\downarrow 0}\mathbb{D}(\hat{\mathcal{S}}_h^{\epsilon,\nu}, \hat{\mathcal{S}}) = 0, \quad \textit{w.p.1.}$$

Proof. Firstly, we show Φ_h^ν epiconverges to Φ^ν as $h \downarrow 0$ over a bounded subset $\mathcal{C}$ of $H^1(0,T)^n \times L^2(0,T)^m$. It is sufficient to prove that for any given sequences $\{h_k\}_{k=1}^\infty \downarrow 0$ and $\{(x^k, u^k)\}_{k=1}^\infty \subseteq \mathcal{C}$ with $(x^k, u^k) \to (x^*, u^*)$ as $k \to \infty$ by the norm $\|\cdot\|_{H^1} \times \|\cdot\|_{L^2}$, we have $\lim_{k\to\infty}|\Phi_k^\nu(x^k, u^k) - \Phi^\nu(x^*, u^*)| = 0$, where $\Phi_k^\nu = \Phi_{h_k}^\nu$.

By Assumption 1.1 and $x^k(T) \to x^*(T)$, we can easily get $\lim_{k\to\infty}|\Phi^\nu(x^k, u^k) - \Phi^\nu(x^*, u^*)| = 0$. Moreover, since x_d, $u_d \in L^2(0,T)^l$, there is $\bar{h} > 0$ such that $\|x_d(t) - x_d(t_i)\| \le h$ and $\|u_d(t) - u_d(t_i)\| \le h$ for any $h \in (0, \bar{h}]$ and a.e. $t \in (t_{i-1}, t_i]$. Following from the boundedness of (x^k, u^k), we can also get $|\Phi_k^\nu(x^k, u^k) - \Phi^\nu(x^k, u^k)| = O(h_k)$. Therefore, we obtain our result about epiconvergence by $|\Phi_k^\nu(x^k, u^k) - \Phi^\nu(x^*, u^*)| \le |\Phi_k^\nu(x^k, u^k) - \Phi^\nu(x^k, u^k)| + |\Phi^\nu(x^k, u^k) - \Phi^\nu(x^*, u^*)|$.

Let $\{\hat{\mathbf{x}}_i^{\epsilon,\nu}, \hat{\mathbf{u}}_i^{\epsilon,\nu}\}_{i=1}^N$ be an optimal solution of (4.1), which means the boundedness of $\{\hat{u}_{h_k}^{\epsilon,\nu}\}_{k=1}^\infty \subseteq L^2(0,T)^m$. Since $L^2(0,T)^m$ is reflexive, there is a subsequence of $\{\hat{u}_{h_k}^{\epsilon,\nu}\}$, which we may assume without loss of generality to be $\{\hat{u}_{h_k}^{\epsilon,\nu}\}$ itself, having a weak limit $\hat{u}_*^{\epsilon,\nu} \in L^2(0,T)^m$. It is easy to see that $(\hat{x}_{h_k}^{\epsilon,\nu}, \hat{u}_{h_k}^{\epsilon,\nu})$ satisfies the differential equation $\dot{x}^{\epsilon,\nu}_{h_k}(t) = A\mathbf{x}^{\epsilon,\nu}_{i+1} + Bu^{\epsilon,\nu}_{h_k}(t)$ for a.e. $t \in (t_i, t_{i+1})$ with some $i \in [N]$. Therefore, there is $\hat{x}^{\epsilon,\nu}_* \in H^1(0,T)^n$ such that $\hat{x}^{\epsilon,\nu}_h \to \hat{x}^{\epsilon,\nu}_*$ in $H^1(0,T)^n$ by $\hat{u}^{\epsilon,\nu}_h \to \hat{u}^{\epsilon,\nu}_*$ in $L^2(0,T)^m$. By [1, Theorem 2.5], we can obtain $\lim_{h\downarrow 0} \mathbb{D}(\hat{\mathcal{S}}^{\epsilon,\nu}_h, \hat{\mathcal{S}}^{\epsilon,\nu}) = 0$ with some $\epsilon > 0$ and sufficiently large ν and then $\lim_{\epsilon\downarrow 0} \lim_{\nu\to\infty} \lim_{h\downarrow 0} \mathbb{D}(\hat{\mathcal{S}}^{\epsilon,\nu}_h, \hat{\mathcal{S}}) = 0$ w.p.1. ◻

4.1. Error estimates of optimal values of problem (4.1) to problem (3.1). In this subsection, we investigate the Euler approximation of problem (3.1). Our results are related to the Euler approximation of the optimal control problem with two-point differential system [13, Theorem 5], which requires the convexity of the terminal set. However, the terminal constraint set $\mathcal{Z}^{\epsilon,\nu}$ in (3.1) is generally nonconvex due to the existence of the complementarity constraints.

We have the following theorem as our main result about the Euler approximation of problem (3.1) in this subsection.

THEOREM 4.3. *Suppose that the conditions of Theorem 2.3 hold. Let $(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu})$ be an optimal solution of (3.1), and let $(\hat{x}^{\epsilon,\nu}_h, \hat{u}^{\epsilon,\nu}_h)$ be defined in (4.3) associated with an optimal solution $\{\hat{\mathbf{x}}^{\epsilon,\nu}_i, \hat{\mathbf{u}}^{\epsilon,\nu}_i\}_{i=1}^N$ of (4.1). Then, for sufficiently small h,*

$$|\Phi^\nu_h(\hat{x}^{\epsilon,\nu}_h, \hat{u}^{\epsilon,\nu}_h) - \Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu})| = O(h). \tag{4.4}$$

To prove Theorem 4.3, we need three lemmas (Lemmas 4.4, 4.5 and 4.6).

LEMMA 4.4. *Suppose that the conditions of Theorem 2.3 hold. Let $(x^{\epsilon,\nu}_h, u^{\epsilon,\nu}_h)$ be defined in (4.3) by a feasible solution $\{\mathbf{x}^{\epsilon,\nu}_i, \mathbf{u}^{\epsilon,\nu}_i\}_{i=1}^N$ of (4.1). Then, for sufficiently small h, there is a feasible solution $(x^{\epsilon,\nu}, u^{\epsilon,\nu})$ of problem (3.1) such that*

$$\|x^{\epsilon,\nu} - x^{\epsilon,\nu}_h\|_{L^2} = O(h), \;\; \|x^{\epsilon,\nu} - x^{\epsilon,\nu}_h\|_{H^1} = O(h), \;\; \|u^{\epsilon,\nu} - u^{\epsilon,\nu}_h\|_{L^2} = O(h).$$

Proof. We denote two positive constants θ_x and θ_u such that $\max_{i\in[N]} \|\mathbf{x}^{\epsilon,\nu}_i\| \le \theta_x$ and $\max_{i\in[N]} \|\mathbf{u}^{\epsilon,\nu}_i\| \le \theta_u$. According to Theorem 4.1, there are $v_i \in \mathbb{R}^{m-l}$ and $\tilde{p}_i \le 0$ such that $\mathbf{u}^{\epsilon,\nu}_i = Yv_i + D^\dagger(f_i + \tilde{p}_i - C\mathbf{x}^{\epsilon,\nu}_i)$ for $i \in [N]$. Let $x^{\epsilon,\nu}(t)$ be the solution of the following system, for $t \in (t_i, t_{i+1}]$,

$$\begin{cases} \dot{x}^{\epsilon,\nu}(t) = (A - BD^\dagger C)x^{\epsilon,\nu}(t) + BY(v_{i+1} + a_{i+1}(t - t_i)) + BD^\dagger(\tilde{p}_{i+1} + f(t)), \\ x^{\epsilon,\nu}(0) = x_0, \;\; x^{\epsilon,\nu}(T) = \mathbf{x}^{\epsilon,\nu}_N, \end{cases}$$

where $\{a_i\}_{i=1}^N \subset \mathbb{R}^{m-l}$ fulfills

$$\begin{aligned} \mathbf{x}^{\epsilon,\nu}_N = e^{(A-BD^\dagger C)T}x_0 + \sum_{i=0}^{N-1} \Bigg[\int_{t_i}^{t_{i+1}} e^{(A-BD^\dagger C)(T-\tau)} d\tau B(Yv_{i+1} + D^\dagger \tilde{p}_{i+1}) \\ + \int_{t_i}^{t_{i+1}} e^{(A-BD^\dagger C)(T-\tau)} BD^\dagger f(\tau) d\tau + \int_{t_i}^{t_{i+1}} e^{(A-BD^\dagger C)(T-\tau)} BYa_{i+1}(\tau - t_i) d\tau \Bigg]. \end{aligned}$$

In addition, we know that $x^{\epsilon,\nu}_h$ solves the differential equation, for any $t \in (t_i, t_{i+1}]$,

$$\begin{cases} \dot{x}^{\epsilon,\nu}_h(t) = (A - BD^\dagger C)x^{\epsilon,\nu}_h(t) + BY(v_{i+1} + a_{i+1}(t - t_i)) + BD^\dagger(\tilde{p}_{i+1} + f(t)) + y(t), \\ x^{\epsilon,\nu}_h(0) = x_0, \;\; x^{\epsilon,\nu}_h(T) = \mathbf{x}^{\epsilon,\nu}_N, \end{cases}$$

where $y(t) = (A - BD^\dagger C)(\mathbf{x}^{\epsilon,\nu}_{i+1} - x^{\epsilon,\nu}_h(t)) - BYa_{i+1}(t - t_i) - BD^\dagger(f(t) - f_{i+1})$. Since $f \in L^2(0,T)^l$, there is a $h_0 > 0$ such that $\|f(t) - f_{i+1}\| \le h$ for any $h \in (0, h_0]$ and a.e. $t \in (t_i, t_{i+1}]$. Let $\tilde{f}(t) = f_{i+1}$ for $t \in (t_i, t_{i+1}]$, we then have $\|f - \tilde{f}\|_{L^2} = O(h)$. It means that $\|y\|_{L^2} = O(h)$ for any $h \in (0, h_0]$. Therefore, we have, for any $t \in (t_i, t_{i+1}]$,

$$\|x^{\epsilon,\nu} - x_h^{\epsilon,\nu}\|_{L^2}^2 \le \int_0^T \int_0^t \|\dot{x}^{\epsilon,\nu}(\tau) - \dot{x}_h^{\epsilon,\nu}(\tau)\|^2 d\tau dt = \int_0^T \int_0^t \|y(\tau)\|^2 d\tau dt \le \|y\|_{L^2}^2 T.$$

Hence, according to the definition of $\|\cdot\|_{H^1}$, we obtain that

$$\|x^{\epsilon,\nu} - x_h^{\epsilon,\nu}\|_{H^1} \le \sqrt{1+T}\|y\|_{L^2} = O(h).$$

Let $u^{\epsilon,\nu}(t) = Y(v_{i+1} + a_{i+1}(t - t_i)) + D^\dagger(\tilde{p}_{i+1} + f(t) - Cx^{\epsilon,\nu}(t))$ for any $t \in (t_i, t_{i+1}]$. It is clear that $u_h^{\epsilon,\nu}(t) = \mathbf{u}_{i+1}^{\epsilon,\nu} = Yv_{i+1} + D^\dagger(f_{i+1} + \tilde{p}_{i+1} - C\mathbf{x}_{i+1}^{\epsilon,\nu})$ for any $t \in (t_i, t_{i+1}]$. Then we have $\|u^{\epsilon,\nu} - u_h^{\epsilon,\nu}\|_{L^2} = O(h)$.

Clearly, according to the definition of $u^\nu(t)$, we can obtain that for any $t \in (t_i, t_{i+1}]$,

$$Cx^{\epsilon,\nu}(t) + Du^{\epsilon,\nu}(t) - f(t) = \tilde{p}_{i+1} \le 0,$$

which shows that $(x^{\epsilon,\nu}, u^{\epsilon,\nu})$ is a feasible solution of problem (3.1). ∎

LEMMA 4.5. *Suppose that the conditions of Theorem 2.3 hold. Let* $(x^{\epsilon,\nu}, u^{\epsilon,\nu})$ *be a feasible solution of problem* (3.1) *with* $\|x^{\epsilon,\nu}(t)\| \le \theta_x'$ *and* $\|u^{\epsilon,\nu}(t)\| \le \theta_u'$ *for a.e.* $t \in [0, T]$, *where* θ_x' *and* θ_u' *are two positive constants. Then, for sufficiently small h, there is* $(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu})$ *defined in* (4.3) *by a feasible solution* $\{\mathbf{x}_i^{\epsilon,\nu}, \mathbf{u}_i^{\epsilon,\nu}\}_{i=1}^N$ *of* (4.1), *such that*

$$\|x^{\epsilon,\nu} - x_h^{\epsilon,\nu}\|_{L^2} = O(h), \ \ \|x^{\epsilon,\nu} - x_h^{\epsilon,\nu}\|_{H^1} = O(h), \ \ \|u^{\epsilon,\nu} - u_h^{\epsilon,\nu}\|_{L^2} = O(h).$$

Proof. Let $(x^{\epsilon,\nu}, u^{\epsilon,\nu}) \in H^1(0,T)^n \times L^2(0,T)^m$ be a feasible solution of problem (3.1), then there are $v \in L^2(0,T)^{m-l}$ and $\tilde{p} \in L^2(0,T)^l$ with $\tilde{p}(t) \le 0$ for a.e. $t \in [0,T]$ such that $u^{\epsilon,\nu}(t) = Yv(t) + D^\dagger(\tilde{p}(t) + f(t) - Cx^{\epsilon,\nu}(t))$. In addition, there are $h_1 > 0$ and a piecewise constant function $\varphi_v(t) = \frac{1}{h}\int_{t_i}^{t_{i+1}} v(\tau)d\tau := \varphi_{i+1}$ for any $t \in (t_i, t_{i+1}]$ such that $\|v(t) - \varphi_v(t)\| \le h$ for a.e. $t \in (t_i, t_{i+1}]$ with $h \in (0, h_1]$. There are also $h_2 > 0$ and a piecewise constant function $\varphi_p(t) = \frac{1}{h}\int_{t_i}^{t_{i+1}} \tilde{p}(\tau)d\tau := \tilde{\varphi}_{i+1}$ for any $t \in (t_i, t_{i+1}]$ such that $\|\tilde{p}(t) - \varphi_p(t)\| \le h$ with $h \in (0, h_2]$ and $\varphi_p(t) \le 0$ for a.e. $t \in (t_i, t_{i+1}]$.

Recall $A_h = I - h(A - BD^\dagger C)$. For $i = 0, 1, \cdots, N-1$, let $\mathbf{x}_0^{\epsilon,\nu} = x_0$ and

$$\mathbf{x}_{i+1}^{\epsilon,\nu} = A_h^{-1}(\mathbf{x}_i^{\epsilon,\nu} + hBY(\varphi_{i+1} + a_{i+1}h) + hBD^\dagger(\tilde{\varphi}_{i+1} + f_{i+1})),$$

where $\{a_i\}_{i=1}^N \subset \mathbb{R}^{m-l}$ fulfills

$$x^{\epsilon,\nu}(T) = A_h^{-N}x_0 + h\sum_{i=0}^{N-1} A_h^{-(i+1)}[BY(\varphi_{i+1} + a_{i+1}h) + hBD^\dagger(\tilde{\varphi}_{i+1} + f_{i+1})].$$

Let $\mathbf{u}_i^{\epsilon,\nu} = Y(\varphi_i + a_i h) + D^\dagger(\tilde{\varphi}_i + f_i - C\mathbf{x}_i^{\epsilon,\nu})$ for $i \in [N]$. Since $\|x^{\epsilon,\nu}(t)\| \le \theta_x'$ and $\|u^{\epsilon,\nu}(t)\| \le \theta_u'$ for a.e. $t \in [0,T]$, there is a partition to $[0,T]$ such that the sequences $\{Y\varphi_i, D^\dagger\tilde{\varphi}_i\}_{i=1}^N$ and $\{\mathbf{x}_i^{\epsilon,\nu}, \mathbf{u}_i^{\epsilon,\nu}\}_{i=1}^N$ are also bounded for any given N. We denote that $\tilde{\theta}_x$ and $\tilde{\theta}_u$ are two positive constants such that $\max_{i\in[N]}\|\mathbf{x}_i^{\epsilon,\nu}\| \le \tilde{\theta}_x$ and $\max_{i\in[N]}\|\mathbf{u}_i^{\epsilon,\nu}\| \le \tilde{\theta}_u$. It is clear that $x_h^{\epsilon,\nu}$ satisfies,

$$\dot{x}_h^{\epsilon,\nu}(t) = (A - BD^\dagger C)x_h^{\epsilon,\nu}(t) + BYv(t) + BD^\dagger(\tilde{p}(t) + f(t)) + \tilde{y}(t), \ \ t \in (t_i, t_{i+1}],$$

where $\tilde{y}(t) = (A - BD^\dagger C)(\mathbf{x}_{i+1}^{\epsilon,\nu} - x_h^{\epsilon,\nu}(t)) + BY(\varphi_{i+1} + a_{i+1}h - v(t)) + BD^\dagger(\tilde{\varphi}_{i+1} + f_{i+1} - \tilde{p}(t) - f(t))$. It means that $\|\tilde{y}\|_{L^2} = O(h)$ for any $h \in (0, \min\{h_0, h_1, h_2\}]$. Hence $\|x^{\epsilon,\nu} - x_h^{\epsilon,\nu}\|_{L^2} \le \sqrt{T}\|\tilde{y}\|_{L^2} = O(h)$ and $\|x^{\epsilon,\nu} - x_h^{\epsilon,\nu}\|_{H^1} = O(h)$. Moreover, we have $\|u^{\epsilon,\nu} - u_h^{\epsilon,\nu}\|_{L^2} = O(h)$.

Obviously, from the definition of $\mathbf{u}_i^{\epsilon,\nu}$, we get $C\mathbf{x}_i^{\epsilon,\nu} + D\mathbf{u}_i^{\epsilon,\nu} - f_i = \tilde{\varphi}_i \le 0$ $(i \in [N])$, which means that $\{\mathbf{x}_i^{\epsilon,\nu}, \mathbf{u}_i^{\epsilon,\nu}\}_{i=1}^N$ is a feasible solution of (4.1). ∎

It should be noted that Lemma 4.4 implies that for any given optimal solution of (4.1) there is a feasible solution of problem (3.1) such that their distances are $O(h)$. Conversely, Lemma 4.5 means that for any given optimal solution of (3.1) there is a feasible solution of problem (4.1) such that their distances are $O(h)$. These two results will help us to prove Theorem 4.3.

Lemma 4.6. *Suppose that the conditions of Theorem 2.3 hold. Let $\{\mathbf{x}_i^{\epsilon,\nu}, \mathbf{u}_i^{\epsilon,\nu}\}_{i=1}^N$ be a feasible solution of (4.1) with $\max_{i\in[N]} \|\mathbf{x}_i^{\epsilon,\nu}\| \le \bar{\theta}_x$ and $\max_{i\in[N]} \|\mathbf{u}_i^{\epsilon,\nu}\| \le \bar{\theta}_u$, where $\bar{\theta}_x$ and $\bar{\theta}_u$ are two positive constants, and let $(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu})$ be defined in (4.3). Then, for sufficiently small h,*

$$|\Phi^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu}) - \Phi_h^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu})| = O(h).$$

Proof. Since $\{\mathbf{x}_i^{\epsilon,\nu}, \mathbf{u}_i^{\epsilon,\nu}\}_{i=1}^N$ is a bounded feasible solution of (4.1), $\Phi_h^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu})$ is bounded, which means that there is a $\theta_o > 0$ such that $\max_{i\in[N]}\{\|\mathbf{x}_i^{\epsilon,\nu} - x_{d,i}\|, \|\mathbf{u}_i^{\epsilon,\nu} - u_{d,i}\|\} \le \theta_o$. Therefore, we have $\Phi^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu}) - \Phi_h^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu}) = W_1 + W_2$, where

$$\begin{aligned} W_1 &= \frac{1}{2}\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} \left(\|x_h^{\epsilon,\nu}(t) - x_d(t)\|^2 - \|\mathbf{x}_{i+1}^{\epsilon,\nu} - x_d(t_{i+1})\|^2\right) dt, \\ W_2 &= \frac{\delta}{2}\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} \left(\|u_h^{\epsilon,\nu}(t) - u_d(t)\|^2 - \|\mathbf{u}_{i+1}^{\epsilon,\nu} - u_d(t_{i+1})\|^2\right) dt. \end{aligned}$$

Note that $x_d \in L^2(0,T)^n$ implies that there is $h_x > 0$ such that $\|x_d(t_{i+1}) - x_d(t)\| \le h$ for a.e. $t \in (t_i, t_{i+1}]$ with $h \in (0, h_x]$. Then we have

$$\begin{aligned} &|W_1| \le \frac{1}{2}\sum_{i=0}^{N-1}\int_{t_i}^{t_{i+1}} (\|x_h^{\epsilon,\nu}(t) - \mathbf{x}_{i+1}^{\epsilon,\nu}\| \\ &+\|x_d(t_{i+1}) - x_d(t)\|)\left(\|x_h^{\epsilon,\nu}(t) - \mathbf{x}_{i+1}^{\epsilon,\nu}\| + \|x_d(t_{i+1}) - x_d(t)\| + 2\|\mathbf{x}_{i+1}^{\epsilon,\nu} - x_d(t_{i+1})\|\right) dt \\ &\le \frac{1}{2}\sum_{i=0}^{N-1} (\|A\|\bar{\theta}_x + \|B\|\bar{\theta}_u + 1)((\|A\|\bar{\theta}_x + \|B\|\bar{\theta}_u + 1)h + 2\theta_o)h^2 = O(h). \end{aligned}$$

Moreover, $u_d \in L^2(0,T)^m$ implies that there is $h_u > 0$ such that $\|u_d(t_{i+1}) - u_d(t)\| \le h$ for a.e. $t \in (t_i, t_{i+1}]$ with $h \in (0, h_u]$. Then we also have $|W_2| = O(h)$ and derives our result for $h \in (0, \min\{h_x, h_u\}]$. ∎

Proof of Theorem 4.3. Since $\{\hat{\mathbf{x}}_i^{\epsilon,\nu}, \hat{\mathbf{u}}_i^{\epsilon,\nu}\}_{i=1}^N$ is an optimal solution of (4.1), there is ψ_0 such that $\max\{\|\hat{x}_h^{\epsilon,\nu} - x_d\|_{L^2}, \|\hat{u}_h^{\epsilon,\nu} - u_d\|_{L^2}\} \le \psi_0$, where $(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu})$ is defined in (4.3) associated with the sequence $\{\hat{\mathbf{x}}_i^{\epsilon,\nu}, \hat{\mathbf{u}}_i^{\epsilon,\nu}\}_{i=1}^N$. Similarly, $(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu})$ is an optimal solution of (1.1), which means that there is ψ_1 such that $\max\{\|\hat{x}^{\epsilon,\nu} - x_d\|_{L^2}, \|\hat{u}^{\epsilon,\nu} - u_d\|_{L^2}\} \le \psi_1$.

Following Lemma 4.4, there is $\bar{h} > 0$ such that for any $h \in (0, \bar{h}]$ there is $(x^{\epsilon,\nu}, u^{\epsilon,\nu})$, which is a feasible solution of (3.1) satisfying $\|x^{\epsilon,\nu} - \hat{x}_h^{\epsilon,\nu}\|_{L^2} = O(h)$ and $\|u^{\epsilon,\nu} - \hat{u}_h^{\epsilon,\nu}\|_{L^2} = O(h)$. Moreover, according to Lemma 4.5, for any $h \in (0, \bar{h}]$ there is a $\{\mathbf{x}_i^{\epsilon,\nu}, \mathbf{u}_i^{\epsilon,\nu}\}_{i=1}^N$, which is a feasible solution of (4.1), such that $\|\hat{x}^{\epsilon,\nu} - x_h^{\epsilon,\nu}\|_{L^2} = O(h)$ and $\|\hat{u}^{\epsilon,\nu} - u_h^{\epsilon,\nu}\|_{L^2} = O(h)$, where $(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu})$ is defined in (4.3) based on the sequence $\{\mathbf{x}_i^{\epsilon,\nu}, \mathbf{u}_i^{\epsilon,\nu}\}_{i=1}^N$.

Then we have $\Phi_h^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu}) \le \Phi_h^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu})$, which means

$$\begin{aligned}
&\Phi_h^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu}) - \Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) \le \Phi_h^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu}) - \Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) \\
&\le |\Phi_h^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu}) - \Phi^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu})| + |\Phi^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu}) - \Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu})|.
\end{aligned}$$

Clearly,

$$\begin{aligned}
&|\Phi^\nu(x_h^{\epsilon,\nu}, u_h^{\epsilon,\nu}) - \Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu})| \le \frac{1}{2}\|x_h^{\epsilon,\nu} - \hat{x}^{\epsilon,\nu}\|_{L^2}(\|x_h^{\epsilon,\nu} - \hat{x}^{\epsilon,\nu}\|_{L^2} + 2\|\hat{x}^{\epsilon,\nu} - x_d\|_{L^2}) \\
&+\frac{\delta}{2}\|u_h^{\epsilon,\nu} - \hat{u}^{\epsilon,\nu}\|_{L^2}(\|u_h^{\epsilon,\nu} - \hat{u}^{\epsilon,\nu}\|_{L^2} + 2\|\hat{u}^{\epsilon,\nu} - u_d\|_{L^2}) = O(h).
\end{aligned}$$

Hence, according to Lemma 4.6, we get $\Phi_h^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu}) - \Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) = O(h)$.

From $\Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) \le \Phi^\nu(x^{\epsilon,\nu}, u^{\epsilon,\nu})$, we have

$$\begin{aligned}
&\Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) - \Phi_h^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu}) \le \Phi^\nu(x^{\epsilon,\nu}, u^{\epsilon,\nu}) - \Phi_h^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu}) \\
&\le |\Phi^\nu(x^{\epsilon,\nu}, u^{\epsilon,\nu}) - \Phi^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu})| + |\Phi^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu}) - \Phi_h^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu})|
\end{aligned}$$

and $|\Phi^\nu(x^{\epsilon,\nu}, u^{\epsilon,\nu}) - \Phi^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu})| = O(h)$. It holds $\Phi^\nu(\hat{x}^{\epsilon,\nu}, \hat{u}^{\epsilon,\nu}) - \Phi_h^\nu(\hat{x}_h^{\epsilon,\nu}, \hat{u}_h^{\epsilon,\nu}) = O(h)$ and then (4.4) holds.

5. Numerical experiments.

We use the following numerical example to illustrate the theoretical results obtained in this paper.

$$\begin{aligned}
&\min_{x,u}\ (\mathbb{E}[\xi_1^2 + \xi_2] + 1)\|x(T)\|^2 + \frac{1}{2}\left(\|x\|_{L^2}^2 + \|u\|_{L^2}^2\right) \\
&\text{s.t.} \begin{cases}
\left.\begin{array}{l}
\dot{x}_1(t) = u_1(t), \\
\dot{x}_2(t) = x_2(t) - u_2(t), \\
\dot{x}_3(t) = u_3(t), \\
\dot{x}_4(t) = x_4(t) - u_4(t), \\
x_1(t) + u_2(t) \le 0, \\
x_4(t) + u_3(t) \le 0,
\end{array}\right\} a.e.\ t \in (0, T), \\
x(0) = (1,1,1,1)^\top, \quad 0 \le x(T) \perp \mathbb{E}[M(\xi)x(T) + q(\xi)] \ge 0, \\
(x_1(T) + x_3(T), (\mathbb{E}[\xi_1] + 1)(x_2(T) + x_4(T)))^\top \in \mathcal{B}(0, \sqrt{6}) \subset \mathbb{R}^2,
\end{cases}
\end{aligned} \tag{5.1}$$

where

$$q(\xi) = \begin{pmatrix} 3 + \xi_2 \\ \xi_1 \\ 1 - \xi_2 \\ \xi_1 + 1 \end{pmatrix} \quad \text{and} \quad M(\xi) = \begin{pmatrix} -2 - \xi_1 & 0 & -\xi_2 & -\xi_1 \\ 0 & \xi_2 & -1 & 0 \\ 0 & -\xi_1 & \xi_2 & 0 \\ \xi_2 - 1 & 0 & 0 & \xi_1 \end{pmatrix}.$$

We set $T = 1$, and $\xi_1 \sim \mathcal{N}(1, 0.01)$ and $\xi_2 \sim \mathcal{U}(-1, 1)$. It is easy to verify that $\mathbb{E}[M(\xi)]$ is a Z-matrix and the controllability matrix in Assumption 1.2

$$\mathcal{R} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & -1 & 0 & -1 & 0 \end{pmatrix},$$

is full row rank. We can derive that the solution set of the LCP in (5.1) is

$$\{(0,0,0,0)^\top,(1,0,0,0)^\top,(0,1,1,0)^\top,(1,1,1,0)^\top\}$$

and the solution set of the terminal constraints in (5.1) is

$$\{(0,0,0,0)^\top,(1,0,0,0)^\top,(0,1,1,0)^\top\}. \tag{5.2}$$

With $x(T)=(0,1,1,0)^\top$, we obtain an optimal solution of problem (5.1) by Maple as the following

$$\begin{aligned}
x_1^*(t) =&(-40.3067\sin(at)+0.3685\cos(at))e^{-ct}+(1.3063\sin(at)+0.6315\cos(at))e^{ct},\\
x_2^*(t) =&(17.379\sin(at)+2.4445\cos(at))e^{-ct}+(3.0042\sin(at)-1.4445\cos(at))e^{ct},\\
x_3^*(t) =&2.0488e^{-1.618t}+1.8734e^{1.618t}-0.2901e^{0.61805t}-2.6321e^{-0.61805t},\\
x_4^*(t) =&3.315e^{-1.618t}+0.46938e^{1.618t}-1.1578e^{0.61805t}-1.6266e^{-0.61805t},\\
u_1^*(t) =&(51.113\sin(at)-14.198\cos(at))e^{-ct}+(1.4471\sin(at)+1.2488\cos(at))e^{ct},\\
u_2^*(t) =&(40.3067\sin(at)-0.3685\cos(at))e^{-ct}-(1.3063\sin(at)+0.6315\cos(at))e^{ct},\\
u_3^*(t) =&-3.315e^{-1.618t}-0.46938e^{1.618t}+1.1578e^{0.61805t}+1.6266e^{-0.61805t},\\
u_4^*(t) =&8.6789e^{-1.618t}-0.2901e^{1.618t}-0.4423e^{0.61805t}-2.6321e^{-0.61805t},
\end{aligned}$$

where $a=0.34066$ and $c=1.2712$. Then we get the optimal value of problem (5.1) is 25.17501124.

It is easy to verify that Assumption 1.1, Assumption 3.7 and Assumption 3.8 hold for the functions $g(x(T),\xi)=(x_1(T)+x_3(T),(\xi_1+1)(x_2(T)+x_4(T)))^\top$ and $F(x(T),\xi)=(\xi_1^2+\xi_2+1)\|x(T)\|^2$, and random matrix $M(\xi)$ and vector $q(\xi)$. Moreover, the conditions of Theorem 3.6 hold, since $\mathbf{0}\in\mathcal{V}$, $\mathbb{E}[M(\xi)]$ is a Z-matrix, $K=\mathcal{B}(0,\sqrt{6})\subset\mathbb{R}^2$, and (3.3) can be fulfilled for $\bar\epsilon=\eta=1$ and $\gamma\ge 10$.

We apply the relaxation, the SAA scheme and the time-stepping method to problem (5.1). We use Matlab built solver *fmincon* to solve the discrete approximation problems of problem (5.1). Setting $\epsilon=0.00001$, for each pair (ν,h) with

$$\nu\in\{500,1000,2000,3000,4000\},\quad h\in\{0.008,0.005,0.004,0.002,0.001\},$$

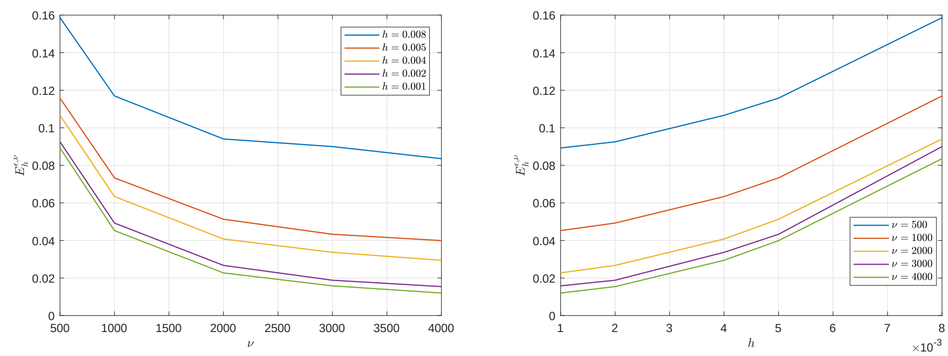
we generate i.i.d. samples $\Xi^{\nu,k}=\{\xi_1^k,\dots,\xi_\nu^k\}, k=1,\dots,10000$. We solve the discrete problem to find a solution $(x_{h,k}^{\epsilon,\nu},u_{h,k}^{\epsilon,\nu})$ using each of the samples $\Xi^{\nu,k}$, $k=1,\dots,10000$. Then we compute the optimal value of the discrete problem for each k

$$\Phi_h^{\nu,k}(x_{h,k}^{\epsilon,\nu},u_{h,k}^{\epsilon,\nu})=\frac{1}{\nu}\sum_{i=1}^{\nu}F(x_{h,k}^{\epsilon,\nu}(T),\xi_i^k)+\frac{1}{2}(\|x_{h,k}^{\epsilon,\nu}\|_{L^2}^2+\|u_{h,k}^{\epsilon,\nu}\|_{L^2}^2).$$

The errors between $\Phi(x^*,u^*)=25.17501124$ and the optimal value $\Phi_h^\nu(x_h^{\epsilon,\nu},u_h^{\epsilon,\nu})$ are estimated by

$$E_h^{\epsilon,\nu}=\frac{1}{10000}\sum_{k=1}^{10000}(\Phi(x^*,u^*)-\Phi_h^{\nu,k}(x_{h,k}^{\epsilon,\nu},u_{h,k}^{\epsilon,\nu}))^2.$$

The numerical results are shown in FIG. 1 and Table 1, which verify the convergence results in Sections 3-4.



(a) (b)

FIG. 1. *Numerical errors between optimal values of (5.1) and its discrete problems with* $\epsilon = 10^{-5}$

TABLE 1
Numerical errors $E_h^{\epsilon,\nu}$ *between optimal values of (5.1) and its discrete problems with* $\nu = 4000$

h \ ϵ	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
0.008	0.20462	0.10333	0.08904	0.08354	0.08174
0.005	0.12368	0.05318	0.04715	0.03992	0.03478
0.004	0.10163	0.04149	0.03516	0.02942	0.02852

6. Conclusions. In this paper, we study the optimal control problem with terminal stochastic linear complementarity constraints (1.1), and its relaxation-SAA problem (3.1) and the relaxation-SAA-time stepping approximation problem (4.1). We prove the existence of feasible solutions and optimal solutions to problem (1.1) in Theorem 2.3 under the assumption $\mathbb{E}[M(\xi)]$ is a Z-matrix or an adequate matrix. Under the same assumptions of Theorem 2.3, we prove the existence of feasible solutions and optimal solutions to (3.1) and (4.1). We also show the convergent properties of these two discrete problems (3.1) and (4.1) by the repeated limits in the order of the relaxation parameter $\epsilon \downarrow 0$, the sample size $\nu \to \infty$ and mesh size $h \downarrow 0$. Moreover, we provide asymptotics of the SAA optimal value and the error bound of the time-stepping method. Problem (1.1) extends optimal control problem with terminal deterministic linear complementarity constraints in [2] to stochastic problems. In [2], Benita and Mehlita derived some stationary points and constraint qualifications under the assumption that the constrained LCP (1.2) is solvable. Theorem 2.3 gives sufficient conditions for the extension of solutions of (1.3).

Acknowledgement We would like to thank the Associate Editor and two referees for their very helpful comments.

REFERENCES

[1] H. ATTOUCH AND R. J.-B. WETS, *A convergence theory for saddle functions*, Trans. Amer. Math. Soc., 280 (1983), pp. 1–41.

[2] F. BENITA AND P. MEHLITZ, *Optimal control problems with terminal complemenatrity constraints*, SIAM J. Optim., 28 (2018), pp. 3079–3104.

[3] R. F. BRAMMER, *Controllability in linear autonomous systems with positive controllers*, SIAM

J. Control Optim., 10 (1972), pp. 339–353.

[4] B. Brogliato, *Some perspectives on the analysis and control of complementarity systems*, IEEE T. Automat. Contr., 48 (2003), pp. 918–935.

[5] M. Brokate and P. Krejčí, *Optimal control of ODE systems involving a rate independent variational inequality*, Discrete Contin. Dyn. Syst. Ser. B, 18 (2013), pp. 331–348.

[6] M. K. Camlibel, J.-S. Pang, and J. Shen, *Lyapunov stability of complementarity and extended systems*, SIAM J. Optim., 17 (2007), pp. 1056–1101.

[7] X. Chen, *First order conditions for nonsmooth discretized constrained optimal control problems*, SIAM J. Control Optim., 42 (2004), pp. 2004–2015.

[8] X. Chen, R. J.-B. Wets, and Y. Zhang, *Stochastic variational inequalities: residual minimization smoothing sample average approximations*, SIAM J. Optim., 22 (2012), pp. 649–673.

[9] X. Chen and S. Xiang, *Newton iterations in implicit time-stepping scheme for differential linear complementarity systems*, Math. Program., 138 (2013), pp. 579–606.

[10] X. Chen and S. Xiang, *Sparse solutions of linear complementarity problems*, Math. Program., 159 (2016), pp. 539–556.

[11] C. Clason, Y. Deng, P. Mehlitz, and U. Prüfert, *Optimal control problems with control complementarity constraints: existence results, optimality conditions, and a penalty method*, Optim. Methods Softw., 35 (2020), pp. 142–170.

[12] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, Boston, 1992.

[13] A. L. Dontchev and W. W. Hager, *Euler approximation of the feasible set*, Numer. Funct. Anal. Optim., 15 (1994), pp. 245–261.

[14] L. Guo and J. J. Ye, *Necessary optimality conditions for optimal control problems with equilibrium constraints*, SIAM J. Control Optim., 54 (2016), pp. 2710–2733.

[15] C. Heij, A. C. M. Ran, and F. van Schagen, *Introduction to Mathematical Systems Theory: Discrete Time Linear Systems, Control and Identification (Second Edition)*, Birkhuser Cham, Gewerbestrasse, 2021.

[16] L. A. Korf and R. J.-B. Wets, *Random lsc functions: An ergodic theorem*, Math. Oper. Res., 26 (2001), pp. 421–445.

[17] P. Lambrianides, Q. Gong, and D. Venturi, *A new scalable algorithm for computational optimal control under uncertainty*, J. Comput. Phys., 420 (2020), p. 109710.

[18] P. Mehlitz and G. Wachsmuth, *Bilevel Optimal Control: Existence Results and Stationarity Conditions*, Springer International Publishing, Cham, 2020, pp. 451–484.

[19] C. Phelps, Q. Gong, J. O. Royset, C. Walton, and I. Kaminer, *Consistent approximation of a nonlinear optimal control problem with uncertain parameters*, Automatica, 50 (2014), pp. 2987–2997.

[20] C. Phelps, J. O. Royset, and Q. Gong, *Optimal control of uncertain systems using sample average approximations*, SIAM J. Control Optim., 54 (2016), pp. 1–29.

[21] A. U. Raghunathan, M. S. Diaz, and L. T. Biegler, *An MPEC formulation for dynamic optimization of distillation operations*, Comput. Chem. Eng., 28 (2004), pp. 2037–2052.

[22] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[23] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009.

[24] M. Tucsnak and G. Weiss, *Observation and Control for Operator Semigroups*, Birkhäuser, Basel, Switzerland, 2009.

[25] A. Vieira, B. Brogliato, and C. Prieur, *Quadratic optimal control of linear complementarity systems: first order necessary conditions and numerical analysis*, IEEE T. Automat. Contr., 65 (2020), pp. 1168–1175.

[26] H. Yoshida and T. Tanaka, *Positive controllability test for continuous-time linear systems*, IEEE T. Automat. Contr., 52 (2007), pp. 1685–1689.