Second-order necessary conditions for stochastic optimal control problems

Xu Zhang

School of Mathematics, Sichuan University, China

The 14th Workshop on Markov Processes and Related Topics

July 16–20, 2018 · Sichuan University, Chengdu, China

This talk is based on the following joint-papers with Haisen Zhang:

[1] H. Zhang and X. Zhang. *Pointwise second-order necessary conditions for stochastic optimal controls, Part I: The case of convex control constraint.* SIAM J. Control Optim. 53 (2015), 2267–2296.

[2] H. Zhang and X. Zhang. *Pointwise second-order necessary conditions for stochastic optimal controls, Part II: The general case.* SIAM J. Control Optim. 55 (2017), 2841–2875.

[3] H. Zhang and X. Zhang. *Second-order necessary conditions for stochastic optimal control problems.* SIAM Rev. 60 (2018), 139–178. (SIGEST article)

- Motivation
- The convex control constraint case
- The general case

• Motivation

Elementary facts

Consider a minimizer x_0 of a smooth function $f(\cdot)$ in a set $G \subset \mathbb{R}^n$, i.e., x_0 satisfies

$$f(x_0) = \inf_{x \in G} f(x). \tag{2.1}$$

If a nonzero vector $\ell \in \mathbb{R}^n$ is admissible (i.e., there is a $\delta > 0$ so that $x_0 + s\ell \in G$ for any $s \in [0, \delta]$), then one has the following first-order necessary condition:

$$0 \leq \lim_{s \to 0^+} \frac{f(x_0 + s\ell) - f(x_0)}{s} = \langle f_x(x_0), \ell \rangle. \tag{2.2}$$

When $\langle f_x(x_0), \ell \rangle = 0$, i.e., (2.2) degenerates, then one can obtain further a second-order necessary condition as follows:

$$0 \leq 2 \lim_{s \to 0^+} \frac{f(x_0 + s\ell) - f(x_0)}{s^2} = \langle f_{xx}(x_0)\ell, \ell \rangle. \tag{2.3}$$

In particular,
(i) If G is open, then

$$f_x(x_0) = 0, \qquad f_{xx}(x_0) \geq 0.$$

(ii) If G is convex, by (2.2), one has

$$0 \leq \langle f_x(x_0), x - x_0 \rangle, \quad \forall\, x \in G. \tag{2.4}$$

When $f_x(x_0) = 0$, then it follows from (2.3) that

$$0 \leq \langle f_{xx}(x_0)(x - x_0), x - x_0 \rangle, \quad \forall\, x \in G. \tag{2.5}$$

Clearly, compared to the first-order necessary condition (2.2)/(2.4), the second-order necessary condition (2.3)/(2.5) can be used to single out the possible minimizer x_0 from a smaller subset of G.

From the above analysis on the minimization problem (2.1), it is easy to see the following:

1) Usually, one has to impose more regularity on the data (say C^2 for $f(\cdot)$) for the second-order necessary condition than that for the first-order (for which C^1 for $f(\cdot)$ is enough);
2) The derivation of the second-order necessary condition is probably more complicated than that of the first-order situation;
3) Usually, in order to establish the second-order necessary condition, one needs to assume that the first-order condition degenerates in some sense.

Very similar phenomenons happen when one establishes the optimality conditions for optimal control problems, though generally it turns out to be much more difficult than that for the above minimization problem.

Stochastic optimal control

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space, $W(\cdot)$ be a $1-d$ standard Wiener process, $\mathbb{F}$ be the natural filtration. Consider a controlled stochastic differential equation

$$\begin{cases} dx(t) = b(t, x(t), u(t))dt + \sigma(t, x(t), u(t))dW(t), \quad t \in [0, T], \\ x(0) = x_0 \in \mathbb{R}^n, \end{cases}$$

and the following cost functional

$$J(u(\cdot)) = \mathbb{E}\Big[\int_0^T f(t, x(t), u(t))dt + h(x(T))\Big].$$

Goal: Find necessary conditions for $\bar{u}(\cdot)$ to minimize $J(u(\cdot))$ over $\mathcal{U}_{ad}$:

$$\inf J(u(\cdot)), \ s.t. \ u(\cdot) \in \mathcal{U}_{ad},$$

$$\mathcal{U}_{ad} := \Big\{u : \Omega \times [0, T] \to U \ \Big| \ u(\cdot) \text{ is } \mathbb{F}-\text{adapted}\Big\}, \ U \subset \mathbb{R}^m.$$

Stochastic maximum principle

Define Hamiltonian

$$\begin{aligned} H(t,x,u,p,q) &= \langle p, b(t,x,u)\rangle + \langle q, \sigma(t,x,u)\rangle - f(t,x,u), \\ &\qquad (t,x,u,p,q) \in [0,T]\times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^n. \end{aligned}$$

Introduce the first-order adjoint equation $(b_x(t) = b_x(t,\bar{x}(t),\bar{u}(t)),\ldots)$

$$\begin{cases} dp_1(t) = -\Big[b_x(t)^\top p_1(t) + \sigma_x(t)^\top q_1(t) - f_x(t)\Big]dt + q_1(t)dW(t), \\ p_1(T) = -h_x(\bar{x}(T)), \end{cases}$$

and the second-order adjoint equation $(H_{xx}(t) = H_{xx}(t,\bar{x}(t),\bar{u}(t),p_1(t),q_1(t)))$

$$\begin{cases} dp_2(t) = -\Big[b_x(t)^\top p_2(t) + p_2(t)b_x(t) + \sigma_x(t)^\top p_2(t)\sigma_x(t) \\ \qquad\qquad +\sigma_x(t)^\top q_2(t) + q_2(t)\sigma_x(t) + H_{xx}(t)\Big]dt + q_2(t)dW(t), \\ p_2(T) = -h_{xx}(\bar{x}(T)). \end{cases}$$

Stochastic maximum principle

Denote

$$
\begin{aligned}
&\mathbb{H}(t,x,u)\\
&= H(t,x,u,p_1(t),q_1(t)) - H(t,x,\bar u(t),p_1(t),q_1(t))\\
&\qquad +\frac{1}{2}\left\langle p_2(t)\big(\sigma(t,x,u)-\sigma(t,x,\bar u(t))\big),\sigma(t,x,u)-\sigma(t,x,\bar u(t))\right\rangle,\\
&\qquad\qquad\qquad\qquad (t,x,u)\in[0,T]\times\mathbb{R}^n\times U.
\end{aligned}
$$

Theorem (S. Peng, 1990)

Let $(\bar x(\cdot),\bar u(\cdot))$ be an optimal pair. Then,

$$
\mathbb{H}(t,\bar x(t),v)\le 0,\qquad \forall\ v\in U,\ a.e.\ (\omega,t)\in\Omega\times[0,T]. \tag{2.6}
$$

Stochastic maximum principle

For $(t, x, u) \in [0, T] \times \mathbb{R}^n \times U$, denote

$$\begin{aligned} \hat{H}(t, x, u) \;:=\; & H(t, x, u, p_1(t), q_1(t)) \\ & -\frac{1}{2}\langle p_2(t)\sigma(t, \bar{x}(t), \bar{u}(t)), \sigma(t, \bar{x}(t), \bar{u}(t))\rangle \\ & +\frac{1}{2}\Big\langle p_2(t)\big(\sigma(t, x, u) - \sigma(t, \bar{x}(t), \bar{u}(t))\big), \\ & \qquad\qquad \sigma(t, x, u) - \sigma(t, \bar{x}(t), \bar{u}(t))\Big\rangle. \end{aligned}$$

Then, (2.6) can be rewritten as

$$\hat{H}(t, \bar{x}(t), \bar{u}(t)) = \max_{v \in U} \hat{H}(t, \bar{x}(t), v), \; a.e. \; (\omega, t) \in \Omega \times [0, T]. \quad (2.7)$$

Stochastic maximum principle

For $(t,x,u)\in[0,T]\times\mathbb{R}^n\times U$, denote

$$\begin{aligned}\hat{H}(t,x,u) \quad := \quad & H(t,x,u,p_1(t),q_1(t)) \\ & -\frac{1}{2}\left\langle p_2(t)\sigma(t,\bar{x}(t),\bar{u}(t)),\sigma(t,\bar{x}(t),\bar{u}(t))\right\rangle \\ & +\frac{1}{2}\Big\langle p_2(t)\big(\sigma(t,x,u)-\sigma(t,\bar{x}(t),\bar{u}(t))\big), \\ & \qquad\qquad \sigma(t,x,u)-\sigma(t,\bar{x}(t),\bar{u}(t))\Big\rangle.\end{aligned}$$

Then, (2.6) can be rewritten as

$$\hat{H}(t,\bar{x}(t),\bar{u}(t))=\max_{v\in U}\hat{H}(t,\bar{x}(t),v),\ a.e.\ (\omega,t)\in\Omega\times[0,T]. \quad (2.7)$$

The correction term q_2 doses not appear in the first-order necessary condition.

Stochastic maximum principle

By the first- and second-order necessary conditions in classical optimization theory, when U is convex, b, σ and f are sufficiently smooth, $\bar{u}(\cdot)$ satisfies

$$\big\langle \hat{H}_u(t,\bar{x}(t),\bar{u}(t)), v-\bar{u}(t)\big\rangle \leq 0, \quad \forall\, v\in U,\ a.e.\ (\omega,t)\in\Omega\times[0,T]. \tag{2.8}$$

Moreover, if

$$\hat{H}_u(t,\bar{x}(t),\bar{u}(t)) = 0,\ a.e.\ (\omega,t)\in\Omega\times[0,T],$$

then

$$\big\langle \hat{H}_{uu}(t,\bar{x}(t),\bar{u}(t))(v-\bar{u}(t)), v-\bar{u}(t)\big\rangle \leq 0, \quad \forall\, v\in U,\ a.e.\ (\omega,t)\in\Omega\times[0,T]. \tag{2.9}$$

Example (A)

Let $T = 1$, $U = [-1, 1]$. *Consider a controlled system*

$$\begin{cases} dx(t) = u(t)dt + u(t)dW(t), & t \in [0, 1], \\ x(0) = 0, \end{cases}$$

with the cost functional

$$J(u(\cdot)) = \frac{1}{2}\mathbb{E}\int_0^1 |u(t)|^2 dt - \frac{1}{2}\mathbb{E}\,|x(1)|^2.$$

For this optimal control problem, the Hamiltonian is defined by

$$\begin{aligned} H(t, x, u, y_1, z_1) &= y_1 u + z_1 u - \frac{1}{2}u^2, \\ & \quad (t, x, u, y_1, z_1) \in [0, 1] \times \mathbb{R} \times U \times \mathbb{R} \times \mathbb{R}. \end{aligned}$$

Let $(\bar{x}(t), \bar{u}(t)) \equiv (0,0)$. It is easy to check that

$$(p_1(t), q_1(t)) \equiv (0,0), \qquad (p_2(t), q_2(t)) \equiv (1,0).$$

Then,

$$\begin{aligned}
\hat{H}(t,\bar{x},\bar{u}(t)) = \hat{H}(t,\bar{x},v) &= 0, \ \forall\, v \in U, \ \forall (\omega,t) \in \Omega \times [0,1], \\
\hat{H}_u(t,\bar{x}(t),\bar{u}(t)) &= 0, \ \forall\, v \in U, \ \forall (\omega,t) \in \Omega \times [0,1], \\
\hat{H}_{uu}(t,\bar{x}(t),\bar{u}(t)) &= 0, \ \forall\, v \in U, \ \forall (\omega,t) \in \Omega \times [0,1].
\end{aligned}$$

Therefore, $(\bar{x}(t), \bar{u}(t)) \equiv (0,0)$ satisfies the Pontryagin-type maximum condition (2.7), conditions (2.8) and (2.9). Nevertheless, choosing $\hat{u}(t) \equiv -1$, we see that

$$-\frac{1}{2} = J(\hat{u}(\cdot)) < J(\bar{u}(\cdot)) = 0.$$

That is, $\bar{u}(t) \equiv 0$ is not an optimal control !

High order necessary conditions for deterministic problems

Several books for deterministic problems.

[1] D. J. Bell and D. H. Jacobson, *Singular Optimal Control Problems*, Academic Press, London-New York, 1975.

[2] D. J. Clements and B. D. O. Anderson, *Singular Optimal Control: the Linear-Quadratic Problem*, Springer-Verlag, Berlin-New York, 1978.

[3] R. F. Gabasov and F. M. Kirillova, *Singular Optimal Controls*, Izdat. "Nauka", Moscow, 1973.

[4] H.-W. Knobloch, *Higher Order Necessary Conditions in Optimal Control Theory*, Springer-Verlag, Berlin-New York, 1981.

[5] N. P. Osmolovskii and H. Maurer, *Applications to Regular and Bang-Bang Control. Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control*, SIAM, Philadelphia, PA, 2012.

Stochastic singular control

Definition

- *We call* $\bar{u}(\cdot) \in \mathcal{U}_{ad}$ *a singular control in the classical sense, if* $\bar{u}(\cdot)$ *satisfies for a.e.* $(\omega, t) \in \Omega \times [0, T]$

$$\hat{H}_u(t, \bar{x}(t), \bar{u}(t)) = 0, \quad \hat{H}_{uu}(t, \bar{x}(t), \bar{u}(t)) = 0. \tag{2.10}$$

Stochastic singular control

Definition

- *We call* $\bar{u}(\cdot) \in \mathcal{U}_{ad}$ *a singular control in the classical sense, if* $\bar{u}(\cdot)$ *satisfies for a.e.* $(\omega, t) \in \Omega \times [0, T]$

$$\hat{H}_u(t, \bar{x}(t), \bar{u}(t)) = 0, \quad \hat{H}_{uu}(t, \bar{x}(t), \bar{u}(t)) = 0. \tag{2.10}$$

- *We call* $\bar{u}(\cdot) \in \mathcal{U}_{ad}$ *a singular control in the sense of Pontryagin's maximum principle on a control subset* $V \subset U$ *if* V *is nonempty and for a.e.* $(\omega, t) \in \Omega \times [0, T]$

$$\hat{H}(t, \bar{x}(t), \bar{u}(t)) = \hat{H}(t, \bar{x}(t), v), \ \forall \ v \in V. \tag{2.11}$$

When an optimal control is singular, one needs to establish the second-order necessary condition

- To distinguish it from other singular control;
- To provide new information for solving it numerically.

The known results

1. N. I. Mahmudov and A. E. Bashirov (1995), S. Tang (2010): a pointwise second-order maximum principle; the control region is allowed to be nonconvex, but **the diffusion term is independent of the control**.
2. J. F. Bonnans and F. J. Silva (2012): an integral-type second-order necessary condition; **the control region is convex**, the diffusion term depends on the control.

However,

1. When the diffusion term contains the control, one has to develop new technique.
2. Instead of the integral-type condition, one hopes to obtain **pointwise conditions**.

The purpose of our work

To establish some *pointwise* second-order necessary conditions for stochastic optimal controls in the general cases.

1. Convex control constraints: The pointwise second-order necessary conditions for stochastic singular optimal controls in the classical sense.
2. Nonconvex control constraints: The pointwise second-order necessary conditions for stochastic singular optimal controls in the sense of Pontryagin's maximum principle.

As we shall see, there exists essential difficulty to establish the pointwise second-order necessary conditions for stochastic optimal controls when the diffusion term contains the control, **even if the control region is convex!**

• The convex control constraint case

Let U be a nonempty, convex subset of $\mathbb{R}^m$. Define the variational equations

$$\begin{cases} dy_1(t) = \Big[b_x(t)y_1(t) + b_u(t)v(t)\Big]dt \\ \qquad\qquad + \Big[\sigma_x(t)y_1(t) + \sigma_u(t)v(t)\Big]dW(t), \quad t \in [0,T], \\ y_1(0) = 0, \end{cases}$$

and

$$\begin{cases} dy_2(t) = \Big[b_x(t)y_2(t) + y_1(t)^\top b_{xx}(t)y_1(t) + 2v(t)^\top b_{xu}(t)y_1(t) \\ \qquad +v(t)^\top b_{uu}(t)v(t)\Big]dt + \Big[\sigma_x(t)y_2(t) + y_1(t)^\top \sigma_{xx}(t)y_1(t) \\ \qquad +2v(t)^\top \sigma_{xu}(t)y_1(t) + v(t)^\top \sigma_{uu}(t)v(t)\Big]dW(t), \quad t \in [0,T], \\ y_2(0) = 0, \end{cases}$$

where $v(\cdot) = u(\cdot) - \bar{u}(\cdot), \quad u(\cdot) \in \mathcal{U}_{ad}$.

Denote by $\mathcal{T}_{\mathcal{U}_{ad}}(\bar{u}(\cdot))$ the closure of $\{u(\cdot) - \bar{u}(\cdot) \mid u(\cdot) \in \mathcal{U}_{ad}\}$ in $L^2_{\mathbb{F}}(\Omega; L^2(0,T;\mathbb{R}^m))$.

Theorem (Bonnans and Silva, 2012)

If $(\bar{x}(\cdot), \bar{u}(\cdot))$ *is an optimal pair, then*

$$\mathbb{E}\int_0^T \langle H_u(t), v(t)\rangle\, dt \le 0, \qquad \forall\ v(\cdot) \in \mathcal{T}_{\mathcal{U}_{ad}}(\bar{u}(\cdot)).$$

In addition, if $v(\cdot) \in \mathcal{T}_{\mathcal{U}_{ad}}(\bar{u}(\cdot))$ *satisfying* $\mathbb{E}\int_0^T \langle H_u(t), v(t)\rangle\, dt = 0$, *then*

$$\begin{aligned}\mathbb{E}\int_0^T \Big[&\langle H_{xx}(t)y_1(t), y_1(t)\rangle + 2\langle H_{xu}(t)y_1(t), v(t)\rangle \\ &+ \langle H_{uu}(t)v(t), v(t)\rangle \Big] dt + \mathbb{E}\langle h_{xx}(\bar{x}(T))y_1(T), y_1(T)\rangle \le 0. \quad (3.1)\end{aligned}$$

The second-order terms with respect to $y_1(\cdot)$ in (3.1) can be eliminated by introducing the second-order adjoint process $(p_2(\cdot), q_2(\cdot))$.

The second-order terms with respect to $y_1(\cdot)$ in (3.1) can be eliminated by introducing the second-order adjoint process $(p_2(\cdot), q_2(\cdot))$.
Define

$$\begin{aligned}\mathbb{S}(t,x,u,p_1,q_1,p_2,q_2) :=& H_{xu}(t,x,u,p_1,q_1) + b_u(t,x,u)^\top p_2 \\ &+ \sigma_u(t,x,u)^\top q_2 + \sigma_u(t,x,u)^\top p_2 \sigma_x(t,x,u),\end{aligned}$$

and denote

$$\mathbb{S}(t) = \mathbb{S}(t,\bar{x}(t),\bar{u}(t),p_1(t),q_1(t),p_2(t),q_2(t)), \qquad t \in [0,T].$$

Lemma (H. Zhang and X. Zhang, 2015)

If $\bar{u}(\cdot)$ is an singular optimal control in the classical sense, then

$$\mathbb{E}\int_0^T \langle \mathbb{S}(t) y_1(t), v(t)\rangle \, dt \le 0, \ \forall \ v(\cdot) = u(\cdot) - \bar{u}(\cdot), \ u(\cdot) \in \mathcal{U}_{ad}. \quad (3.2)$$

The second-order terms with respect to $y_1(\cdot)$ in (3.1) can be eliminated by introducing the second-order adjoint process $(p_2(\cdot), q_2(\cdot))$.
Define

$$\begin{aligned}\mathbb{S}(t,x,u,p_1,q_1,p_2,q_2) :=& H_{xu}(t,x,u,p_1,q_1) + b_u(t,x,u)^\top p_2 \\ &+ \sigma_u(t,x,u)^\top q_2 + \sigma_u(t,x,u)^\top p_2 \sigma_x(t,x,u),\end{aligned}$$

and denote

$$\mathbb{S}(t) = \mathbb{S}(t, \bar{x}(t), \bar{u}(t), p_1(t), q_1(t), p_2(t), q_2(t)), \qquad t \in [0,T].$$

Lemma (H. Zhang and X. Zhang, 2015)

If $\bar{u}(\cdot)$ is an singular optimal control in the classical sense, then

$$\mathbb{E}\int_0^T \langle \mathbb{S}(t) y_1(t), v(t)\rangle \, dt \le 0, \ \forall\ v(\cdot) = u(\cdot) - \bar{u}(\cdot),\ u(\cdot) \in \mathcal{U}_{ad}. \quad (3.2)$$

The correction part q_2 appears in condition (3.2) !

The main difficulty

Define

$$u(t) = \begin{cases} v, & t \in E_\varepsilon, \\ \bar{u}(t), & t \in [0,T) \setminus E_\varepsilon, \end{cases}$$

$v \in U$, $E_\varepsilon = [\tau, \tau+\varepsilon)$, $\tau \in [0,T)$, $\varepsilon > 0$, $\tau + \varepsilon \leq T$. The solution $y_1(\cdot)$ to the first-order variational equation with respect to

$$v(\cdot) = u(\cdot) - \bar{u}(\cdot) = (v - \bar{u}(\cdot))\chi_{E_\varepsilon}(\cdot)$$

enjoys the following explicit representation:

$$\begin{aligned} y_1(t) =& \Phi(t) \int_0^t \Phi(s)^{-1} \big(b_u(s) - \sigma_x(s)\sigma_u(s)\big)(v - \bar{u}(s))\chi_{E_\varepsilon}(s)ds \\ &+ \Phi(t) \int_0^t \Phi(s)^{-1}\sigma_u(s)(v - \bar{u}(s))\chi_{E_\varepsilon}(s)dW(s), \end{aligned} \tag{3.3}$$

where $\Phi(\cdot)$ solves the following stochastic differential equation:

$$\begin{cases} d\Phi(t) = b_x(t)\Phi(t)dt + \sigma_x(t)\Phi(t)dW(t), & t \in [0,T], \\ \Phi(0) = I_n. \end{cases} \tag{3.4}$$

Substituting (3.3) into (3.2), we obtain a "bad" term

$$
\begin{aligned}
&\mathbb{E}\int_\tau^{\tau+\varepsilon}\Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)(v-\bar u(s))dW(s), v-\bar u(t)\Big\rangle dt\\
\leq\ & \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}\Big|\big(\mathbb{S}(t)\Phi(t)\big)^\top(v-\bar u(t))\Big|^2 dt\Big]^{\frac12}\cdot\\
&\qquad\qquad \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}\int_\tau^t\Big|\Phi(s)^{-1}\sigma_u(s)(v-\bar u(s))\Big|^2 dsdt\Big]^{\frac12}\\
=\ & O(\varepsilon^{\frac32}),\quad \text{as } \varepsilon\to 0^+. \qquad (3.5)
\end{aligned}
$$

Substituting (3.3) into (3.2), we obtain a "bad" term

$$\begin{aligned}
& \mathbb{E}\int_\tau^{\tau+\varepsilon}\Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)(v-\bar{u}(s))dW(s), v-\bar{u}(t)\Big\rangle dt \\
\leq\ & \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}\Big|\big(\mathbb{S}(t)\Phi(t)\big)^\top(v-\bar{u}(t))\Big|^2 dt\Big]^{\frac{1}{2}}\cdot \\
& \qquad \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}\int_\tau^t\Big|\Phi(s)^{-1}\sigma_u(s)(v-\bar{u}(s))\Big|^2 dsdt\Big]^{\frac{1}{2}} \\
=\ & O(\varepsilon^{\frac{3}{2}}), \quad \text{as } \varepsilon\to 0^+. \qquad (3.5)
\end{aligned}$$

We cannot simply divide (3.5) by ε^2 and compute the limit when $\varepsilon\to 0^+$.

Substituting (3.3) into (3.2), we obtain a “bad” term

$$\begin{aligned}
&\mathbb{E}\int_\tau^{\tau+\varepsilon}\Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)(v-\bar u(s))dW(s), v-\bar u(t)\Big\rangle dt\\
\le\ & \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}\Big|\big(\mathbb{S}(t)\Phi(t)\big)^\top(v-\bar u(t))\Big|^2 dt\Big]^{\frac12}\cdot\\
&\qquad \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}\int_\tau^t\Big|\Phi(s)^{-1}\sigma_u(s)(v-\bar u(s))\Big|^2 dsdt\Big]^{\frac12}\\
=\ & O(\varepsilon^{\frac32}),\quad \text{as } \varepsilon\to 0^+. \qquad (3.5)
\end{aligned}$$

We cannot simply divide (3.5) by ε^2 and compute the limit when $\varepsilon\to 0^+$.

How about dividing (3.5) by $\varepsilon^{\frac32}$?

Substituting (3.3) into (3.2), we obtain a "bad" term

$$
\begin{aligned}
&\mathbb{E}\int_\tau^{\tau+\varepsilon}\Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)(v-\bar u(s))dW(s), v-\bar u(t)\Big\rangle dt \\
\leq\ & \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}\Big|\big(\mathbb{S}(t)\Phi(t)\big)^\top(v-\bar u(t))\Big|^2 dt\Big]^{\frac12}\cdot \\
&\qquad\qquad \Big[\mathbb{E}\int_\tau^{\tau+\varepsilon}\int_\tau^t \Big|\Phi(s)^{-1}\sigma_u(s)(v-\bar u(s))\Big|^2 dsdt\Big]^{\frac12} \\
=\ & O(\varepsilon^{\frac32}),\quad \text{as } \varepsilon\to 0^+. \qquad (3.5)
\end{aligned}
$$

We cannot simply divide (3.5) by ε^2 and compute the limit when $\varepsilon\to 0^+$.
How about dividing (3.5) by $\varepsilon^{\frac32}$?

$$
\begin{aligned}
\lim_{\varepsilon\to 0^+}\frac{1}{\varepsilon^{\frac32}}\mathbb{E}\int_\tau^{\tau+\varepsilon}\Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)\ \cdot \\
(v-\bar u(s))dW(s), v-\bar u(t)\Big\rangle dt = 0.
\end{aligned}
$$

We get nothing !

The essential difficulty is caused by the Itô integral term. If it can be replaced by a Lebesgue integral term, the difficulty can be overcome.

By the martingale representation theorem, for any $v \in U$, there exists $\phi_v(\cdot,\cdot)$ such that

$$\mathbb{S}(t)^\top (v - \bar{u}(t)) = \mathbb{E}\left[\mathbb{S}(t)^\top (v - \bar{u}(t))\right] + \int_0^t \phi_v(s,t)dW(s),$$
$$a.e.\ (\omega,t) \in \Omega \times [0,T].$$

Then,

$$\begin{aligned}
&\mathbb{E}\int_\tau^{\tau+\varepsilon} \Big\langle \mathbb{S}(t)\Phi(t)\int_\tau^t \Phi(s)^{-1}\sigma_u(s)(v-\bar{u}(s))dW(s), v-\bar{u}(t)\Big\rangle dt \\
=&\mathbb{E}\int_\tau^{\tau+\varepsilon}\int_\tau^t \Big\langle \Phi(\tau)\Phi(s)^{-1}\sigma_u(s)(v-\bar{u}(s)), \phi_v(s,t)\Big\rangle dsdt \\
&+\mathbb{E}\int_\tau^{\tau+\varepsilon}\int_\tau^t \Big\langle \sigma_x(s)\sigma_u(s)(v-\bar{u}(s)), \mathbb{S}(\tau)^\top\big((v-\bar{u}(\tau))\big)\Big\rangle dsdt + o(\varepsilon^2), \\
&\qquad (\varepsilon \to 0^+).
\end{aligned}$$

Denote

$$\begin{aligned}&\frac{1}{2}\partial_\tau^+\big(\mathbb{S}(\tau)^\top(v-\bar u(\tau));\sigma_u(\tau)(v-\bar u(\tau))\big)\\ =&\limsup_{\theta\to 0^+}\frac{1}{\varepsilon^2}\mathbb{E}\int_\tau^{\tau+\varepsilon}\int_\tau^t\Big\langle\Phi(\tau)\Phi(s)^{-1}\sigma_u(s)(v-\bar u(s)),\phi_v(s,t)\Big\rangle dsdt.\end{aligned}$$

Theorem (H. Zhang and X. Zhang, 2015)

If $\bar u(\cdot)$ is a singular optimal control in the classical sense, then for any $v\in U$, it holds that

$$\begin{aligned}&\mathbb{E}\ \langle\mathbb{S}(\tau)b_u(\tau)(v-\bar u(\tau)),v-\bar u(\tau)\rangle\\ &+\partial_\tau^+\big(\mathbb{S}(\tau)^\top(v-\bar u(\tau));\sigma_u(\tau)(v-\bar u(\tau))\big)\le 0,\ \ a.e.\ \tau\in[0,T].\end{aligned}$$

Denote by $\mathbb{D}^{1,2}(\mathbb{R}^n)$ the subspace of $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ whose elements are Malliavin differentiable, by $\mathcal{D}.\xi$ the Malliavin derivative of a random variable $\xi \in \mathbb{D}^{1,2}(\mathbb{R}^n)$, and by $\mathbb{L}^{1,2}(\mathbb{R}^n)$ the space of processes $\varphi \in L^2([0,T]\times\Omega;\mathbb{R}^n)$ such that

(i) $\varphi(t,\cdot) \in \mathbb{D}^{1,2}(\mathbb{R}^n)$, for a.e. $t \in [0,T]$;

(ii) $(\omega,t,s) \to D_s\varphi(\omega,t)$ admits an $\mathcal{F}\otimes\mathscr{B}([0,T]\times[0,T])$-measurable version; and

(iii)
$$|||\varphi|||_{1,2} := \mathbb{E}\int_0^T |\varphi(t)|^2 dt + \mathbb{E}\int_0^T\int_0^T |D_s\varphi(t)|^2 ds dt < +\infty.$$

In addition, write

$$\mathbb{L}^{1,2}_{2^+}(\mathbb{R}^n) = \Big\{\varphi(\cdot) \in \mathbb{L}^{1,2}(\mathbb{R}^n) \Big| \ \exists\ \mathcal{D}^+\varphi(\cdot) \in L^2([0,T]\times\Omega;\mathbb{R}^n) \text{ so that } f_\varepsilon(s) := \sup_{s<t<(s+\varepsilon)\wedge T} \mathbb{E}\,\big|\mathcal{D}_s\varphi(t) - \mathcal{D}^+\varphi(s)\big|^2 < \infty,\ a.e.\ s\in[0,T], \ f_\varepsilon(\cdot) \text{ is measurable on } [0,T] \text{ for any } \varepsilon>0, \text{ and } \lim_{\varepsilon\to 0^+}\int_0^T f_\varepsilon(s)ds = 0\Big\}.$$

Similarly for $\mathbb{L}^{1,2}_{2^-}(\mathbb{R}^n)$. Denote

$$\mathbb{L}^{1,2}_{2}(\mathbb{R}^n) = \mathbb{L}^{1,2}_{2^+}(\mathbb{R}^n) \cap \mathbb{L}^{1,2}_{2^-}(\mathbb{R}^n).$$

For any $\varphi(\cdot)\in\mathbb{L}^{1,2}_{2}(\mathbb{R}^n)$, denote $\nabla\varphi(\cdot) = \mathcal{D}^+\varphi(\cdot) + \mathcal{D}^-\varphi(\cdot)$.
Denote by $\mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R}^n)$ the set of all $\mathbb{F}$-adapted processes in $\mathbb{L}^{1,2}_{2}(\mathbb{R}^n)$.

Under proper regularity assumptions, we can refine our result.

$$
\begin{aligned}
(C1) \qquad & \bar{u}(\cdot) \in \mathbb{L}^{1,2}_{\mathbb{F},2}(\mathbb{R}^m) \cap L^\infty(\Omega \times [0,T]; \mathbb{R}^m), \\
& \mathbb{S}(\cdot) \in \mathbb{L}^{1,2}_{\mathbb{F},2}(\mathbb{R}^{m\times n}) \cap L^\infty(\Omega \times [0,T]; \mathbb{R}^{m\times n}).
\end{aligned}
$$

Theorem (H. Zhang and X. Zhang, 2015)

Let (C1) hold. If $\bar{u}(\cdot)$ is a singular optimal control in the classical sense, then for a.e. $(\omega, t) \in \Omega \times [0,T]$, it holds that

$$
\begin{aligned}
& \langle \mathbb{S}(\tau) b_u(\tau)(v - \bar{u}(\tau)), v - \bar{u}(\tau) \rangle \\
& \quad + \langle \nabla \mathbb{S}(\tau) \sigma_u(\tau)(v - \bar{u}(\tau)), v - \bar{u}(\tau) \rangle \\
& \quad - \langle \mathbb{S}(\tau) \sigma_u(\tau)(v - \bar{u}(\tau)), \nabla \bar{u}(\tau) \rangle \\
& \le 0, \quad \forall\, v \in U. \qquad (3.6)
\end{aligned}
$$

In example (A), $\bar{u}(t) \equiv 0$ is a singular control in the classical sense, but not an optimal control.
Now, we show that $\bar{u}(t) \equiv 0$ does not satisfy the second-order necessary optimality condition (3.6). Actually,

$$\mathbb{S}(t) \equiv 1, \quad \nabla\mathbb{S}(t) \equiv 0, \quad \nabla\bar{u}(t) \equiv 0, \quad \forall(\omega, t) \in \Omega \times [0,1].$$

Let $v = 1$, we find that

$$\begin{aligned}
&\langle \mathbb{S}(\tau) b_u(\tau)(v - \bar{u}(\tau)), v - \bar{u}(\tau)\rangle \\
&\quad + \langle \nabla\mathbb{S}(\tau)\sigma_u(\tau)(v - \bar{u}(\tau)), v - \bar{u}(\tau)\rangle \\
&\quad - \langle \mathbb{S}(\tau)\sigma_u(\tau)(v - \bar{u}(\tau)), \nabla\bar{u}(\tau)\rangle \\
&= 1 > 0, \quad \forall(\omega, \tau) \in \Omega \times [0,1],
\end{aligned}$$

which contradicts to the condition (3.6).

Example (B)

Let $U = [-1,1] \times [-1,1]$. *Consider a linear controlled system*

$$\begin{cases} dx(t) = Bu(t)dt + Du(t)dW(t), & t \in [0,T], \\ x(0) = 0, \end{cases} \tag{3.7}$$

with the cost functional

$$J(u(\cdot)) = \frac{1}{2}\mathbb{E}\left\langle Gx(T), x(T)\right\rangle,$$

where

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad G = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

In this example, the Hamiltonian is defined by

$$H(t,x,u,p_1,q_1) = \langle p_1, Bu\rangle + \langle q_1, Du\rangle\,, \qquad (t,x,u,p_1,q_1) \in [0,T]\times\mathbb{R}^2\times U\times\mathbb{R}^2\times\mathbb{R}^2.$$

Clearly, $(\bar{x}(t),\bar{u}(t)) \equiv (0,0)$ is an optimal pair, and

$$(p_1(t),q_1(t)) \equiv (0,0), \quad (p_2(t),q_2(t)) \equiv (-G,0).$$

It is easy to see that

$$\hat{H}_u(t,\bar{x}(t),\bar{u}(t)) \equiv 0, \quad \hat{H}_{uu}(t,\bar{x}(t),\bar{u}(t)) \equiv 0.$$

Therefore, $\bar{x}(t) \equiv 0$ is a singular optimal control.
On the other hand, for any $v \in U$ and a.e. $(\omega,t) \in \Omega\times[0,T]$,

$$\begin{aligned}&\langle \mathbb{S}(\tau)b_u(\tau)(v-\bar{u}(\tau)), v-\bar{u}(\tau)\rangle\\ &\quad + \langle \nabla\mathbb{S}(\tau)\sigma_u(\tau)(v-\bar{u}(\tau)), v-\bar{u}(\tau)\rangle\\ &\quad - \langle \mathbb{S}(\tau)\sigma_u(\tau)(v-\bar{u}(\tau)), \nabla\bar{u}(\tau)\rangle = -\big\langle B^\top GBv, v\big\rangle \le 0.\end{aligned}$$

That is, the necessary condition (3.6) holds.

• The general case

Let U be a nonempty subset of $\mathbb{R}^m$. U needs not to be convex. Let $(\bar{x}(\cdot), \bar{u}(\cdot))$ be an optimal pair, $u(\cdot) \in \mathcal{U}_{ad}$ be an admissible control, $E_\varepsilon \subset [0, T]$ be a measurable set with measure $|E_\varepsilon| = \varepsilon$, $\varepsilon \in (0, T)$. Define

$$u^\varepsilon(t) = \begin{cases} u(t), & t \in E_\varepsilon, \\ \bar{u}(t), & t \in [0, T] \setminus E_\varepsilon, \end{cases}$$

Let $x^\varepsilon(\cdot)$ be the state with respect to the control $u^\varepsilon(\cdot)$ and let $\delta x(\cdot) = x^\varepsilon(\cdot) - \bar{x}(\cdot)$.

It is known that $\|\delta x\|_{\infty,2} \equiv \left[\mathbb{E}\left(\sup_{t\in[0,T]} |\delta x(t, \cdot)|^2\right)\right]^{\frac{1}{2}} \leq C\varepsilon^{\frac{1}{2}}$.

Therefore, to obtain the first-order necessary optimality condition for optimal controls, the cost functional should be expanded up to order two, and two variational equations and two adjoint equations need to be introduced.

Naturally, to establish the second-order necessary optimality condition for optimal controls, the cost functional should be expanded up to order *four*, and *four* variational equations and *four* adjoint equations need to be introduced.
To simplify the notations, in what follows we only consider the 1-dimensional cases. The high-dimensional cases can be discussed similarly.
Now, we define the following four variational equations:

$$\begin{cases} dy_1^\varepsilon(t) = b_x(t)y_1^\varepsilon(t)dt + \big[\sigma_x(t)y_1^\varepsilon(t) + \delta\sigma(t)\chi_{E_\varepsilon}(t)\big]dW(t), \\ y_1^\varepsilon(0) = 0, \end{cases} \tag{4.1}$$

$$\begin{cases} dy_2^\varepsilon(t) = \big[b_x(t)y_2^\varepsilon(t) + \frac{1}{2}b_{xx}(t)y_1^\varepsilon(t)^2 + \delta b(t)\chi_{E_\varepsilon}(t)\big]dt \\ \qquad\qquad + \big[\sigma_x y_2^\varepsilon(t) + \frac{1}{2}\sigma_{xx}(t)y_1^\varepsilon(t)^2 + \delta\sigma_x(t)y_1^\varepsilon(t)\chi_{E_\varepsilon}(t)\big]dW(t), \\ y_2(0) = 0, \end{cases} \tag{4.2}$$

$$\begin{cases} dy_3^\varepsilon(t) = \Big[b_x(t)y_3^\varepsilon(t) + \frac{1}{2}b_{xx}(t)\big(2y_1^\varepsilon(t)y_2^\varepsilon(t) + y_2^\varepsilon(t)^2\big) \\ \qquad\qquad\qquad + \frac{1}{6}b_{xxx}(t)y_1^\varepsilon(t)^3 + \delta b_x(t)y_1^\varepsilon(t)\chi_{E_\varepsilon}(t)\Big]dt \\ \qquad\quad + \Big[\sigma_x(t)y_3^\varepsilon(t) + \frac{1}{2}\sigma_{xx}(t)\big(2y_1^\varepsilon(t)y_2^\varepsilon(t) + y_2^\varepsilon(t)^2\big) \\ \qquad\qquad\qquad + \frac{1}{6}\sigma_{xxx}(t)y_1^\varepsilon(t)^3 + \delta\sigma_x(t)y_2^\varepsilon(t)\chi_{E_\varepsilon}(t) \\ \qquad\qquad\qquad\qquad + \frac{1}{2}\delta\sigma_{xx}(t)y_1^\varepsilon(t)^2\chi_{E_\varepsilon}(t)\Big]dW(t), \\ y_3^\varepsilon(0) = 0, \end{cases} \tag{4.3}$$

and

$$\begin{cases} dy_4^\varepsilon(t) = \Big[b_x(t)y_4^\varepsilon(t) + \frac{1}{2}b_{xx}(t)\big(2y_1^\varepsilon(t)y_3^\varepsilon(t) + 2y_2^\varepsilon(t)y_3^\varepsilon(t) + y_3^\varepsilon(t)^2\big) \\ \qquad + \frac{1}{6}b_{xxx}(t)\big(3y_1^\varepsilon(t)^2y_2^\varepsilon(t) + 3y_1^\varepsilon(t)y_2^\varepsilon(t)^2 + y_2^\varepsilon(t)^3\big) \\ \qquad + \frac{1}{24}b_{xxxx}(t)y_1^\varepsilon(t)^4 + \delta b_x(t)y_2^\varepsilon(t)\chi_{E_\varepsilon}(t) \\ \qquad + \frac{1}{2}\delta b_{xx}(t)y_1^\varepsilon(t)^2\chi_{E_\varepsilon}(t)\Big]dt \\ \qquad + \Big[\sigma_x(t)y_4^\varepsilon(t) + \frac{1}{2}\sigma_{xx}(t)\big(2y_1^\varepsilon(t)y_3^\varepsilon(t) + 2y_2^\varepsilon(t)y_3^\varepsilon(t) + y_3^\varepsilon(t)^2\big) \\ \qquad + \frac{1}{6}\sigma_{xxx}(t)\big(3y_1^\varepsilon(t)^2y_2^\varepsilon(t) + 3y_1^\varepsilon(t)y_2^\varepsilon(t)^2 + y_2^\varepsilon(t)^3\big) \\ \qquad + \frac{1}{24}\sigma_{xxxx}(t)y_1^\varepsilon(t)^4 + \delta\sigma_x(t)y_3^\varepsilon(t)\chi_{E_\varepsilon}(t) \\ \qquad + \frac{1}{2}\delta\sigma_{xx}(t)(2y_1^\varepsilon(t)y_2^\varepsilon(t) + y_2^\varepsilon(t)^2)\chi_{E_\varepsilon}(t) \\ \qquad + \frac{1}{6}\delta\sigma_{xxx}(t)y_1^\varepsilon(t)^3\chi_{E_\varepsilon}(t)\Big]dW(t), \\ y_4^\varepsilon(0) = 0. \end{cases} \tag{4.4}$$

Corresponding to variational equations (4.1)–(4.4), we introduce the following four adjoint equations

$$\begin{cases} dp_1(t) = -\Big[b_x(t)p_1(t) + \sigma_x(t)q_1(t) - f_x(t)\Big]dt + q_1(t)dW(t), \\ p_1(T) = -h_x(\bar{x}(T)), \end{cases} \tag{4.5}$$

$$\begin{cases} dp_2(t) = -\Big[2b_x(t)p_2(t) + \sigma_x(t)^2 p_2(t) + 2\sigma_x(t)q_2(t) \\ \qquad\qquad +H_{xx}(t)\Big]dt + q_2(t)dW(t), \\ p_2(T) = -h_{xx}(\bar{x}(T)), \end{cases} \tag{4.6}$$

$$\begin{cases} dp_3(t) = -\Big[3b_x(t)p_3(t) + 3\sigma_x^2(t)p_3(t) + 3\sigma_x(t)q_3(t) \\ \qquad\qquad +3b_{xx}(t)p_2(t) + 3\sigma_{xx}(t)q_2(t) \\ \qquad\qquad +3\sigma_x(t)\sigma_{xx}(t)p_2(t) + H_{xxx}(t)\Big]dt + q_3(t)dW(t), \\ p_3(T) = -h_{xxx}(\bar{x}(T)), \end{cases} \tag{4.7}$$

and

$$
\begin{cases}
dp_4(t) = -\Big[4b_x(t)p_4(t) + 6\sigma_x^2(t)p_4(t) + 4\sigma_x(t)q_4(t) \\
\qquad\qquad +6b_{xx}(t)p_3(t) + 6\sigma_{xx}(t)q_3(t) + 12\sigma_x(t)\sigma_{xx}(t)p_3(t) \\
\qquad\qquad +4b_{xxx}(t)p_2(t) + 4\sigma_x(t)\sigma_{xxx}(t)p_2(t) + 3\sigma_{xx}^2(t)p_2(t) \\
\qquad\qquad +4\sigma_{xxx}(t)q_2(t) + H_{xxx}(t)\Big]dt + q_4(t)dW(t), \\
p_4(T) = -h_{xxxx}(\bar{x}(T)).
\end{cases}
\tag{4.8}
$$

Denote

$$
\begin{aligned}
&\mathrm{S}(t,x,u,p_2,q_2) = p_2 b(t,x,u) + q_2\sigma(t,x,u), \\
&\qquad\qquad (t,x,u,p_2,q_2) \in [0,T]\times\mathbb{R}\times U\times\mathbb{R}\times\mathbb{R}; \\
&\mathrm{T}(t,x,u,p_3,q_3) = p_3 b(t,x,u) + q_3\sigma(t,x,u), \\
&\qquad\qquad (t,x,u,p_3,q_3) \in [0,T]\times\mathbb{R}\times U\times\mathbb{R}\times\mathbb{R}; \\
&\mathbb{H}(t,x,u) = H(t,x,u,p_1(t),q_1(t)) - H(t,x,\bar{u}(t),p_1(t),q_1(t)) \\
&\qquad\qquad +\frac{1}{2}p_2(t)\big(\sigma(t,x,u)-\sigma(t,x,\bar{u}(t))\big)^2, \\
&\qquad\qquad (t,x,u) \in [0,T]\times\mathbb{R}\times U;
\end{aligned}
$$

$$
\begin{aligned}
& \widetilde{\mathbb{S}}(t,x,u) \\
= \ & \mathbb{H}_x(t,x,u) + \mathrm{S}(t,x,u,p_2(t),q_2(t)) - \mathrm{S}(t,x,\bar{u}(t),p_2(t),q_2(t)) \\
& + p_2\sigma_x(t,x,\bar{u}(t))\big(\sigma(t,x,u)-\sigma(t,x,\bar{u}(t))\big) \\
& + \frac{1}{2}p_3\big(\sigma(t,x,u)-\sigma(t,x,\bar{u}(t))\big)^2, \ (t,x,u)\in[0,T]\times\mathbb{R}\times U; \\
& \mathbb{T}(t,x,u) \\
= \ & \widetilde{\mathbb{S}}_x(t,x,u) + \mathrm{S}_x(t,x,u,p_2(t),q_2(t)) - \mathrm{S}_x(t,x,\bar{u}(t),p_2(t),q_2(t)) \\
& + \mathrm{T}(t,x,u,p_3(t),q_3(t)) - \mathrm{T}(t,x,\bar{u}(t),p_3(t),q_3(t)) \\
& + p_2\sigma_x(t,x,\bar{u}(t))\big(\sigma_x(t,x,u)-\sigma_x(t,x,\bar{u}(t))\big) \\
& + p_3\big(\sigma(t,x,u)-\sigma(t,x,\bar{u}(t))\big)\big(\sigma_x(t,x,u)-\sigma_x(t,x,\bar{u}(t))\big) \\
& + 2p_3\sigma_x(t,x,\bar{u}(t))\big(\sigma(t,x,u)-\sigma(t,x,\bar{u}(t))\big) \\
& + \frac{1}{2}p_4\big(\sigma(t,x,u(t))-\sigma(t,x,\bar{u}(t))\big)^2, \ (t,x,u)\in[0,T]\times\mathbb{R}\times U.
\end{aligned}
$$

Lemma (H. Zhang and X. Zhang, 2015)

Let $(\bar{x}(\cdot),\bar{u}(\cdot))$ *be an optimal pair, then*

$$\begin{aligned} &J(u^\varepsilon)-J(\bar{u})\\ =&-\mathbb{E}\int_0^T\Big[\mathbb{H}(t,\bar{x}(t),u(t))+\widetilde{\mathbb{S}}(t,\bar{x}(t),u(t))(y_1^\varepsilon(t)+y_2^\varepsilon(t))\\ &+\frac{1}{2}\mathbb{T}(t,\bar{x}(t),u(t))y_1^\varepsilon(t)^2\Big]\chi_{E_\varepsilon}(t)dt+o(\varepsilon^2),\ (\varepsilon\to 0^+). \end{aligned} \tag{4.9}$$

Since $\|y_1^\varepsilon\|_{\infty,2}\le C\varepsilon^{\frac{1}{2}}$, there exists a $\frac{3}{2}$-order term in (4.9). Similar to the convex control constraints cases, we can reformulate this term into a 2-order term.

$(C2)$ For any $v \in V$, $\widetilde{\mathbb{S}}(\cdot, \bar{x}(\cdot), v) \in \mathbb{L}^{1,2}_{2,\mathbb{F}}(\mathbb{R})$, and the map $v \mapsto \nabla\widetilde{\mathbb{S}}(\tau, \bar{x}(\tau), v)$ is continuous on V for a.e. $\tau \in [0, T]$.

Theorem (H. Zhang and X. Zhang, 2015)

Let (C2) hold. If $\bar{u}(\cdot)$ is a singular optimal control in the sense of Pontryagin-type maximum principle on $V \subset U$, then

$$\begin{aligned} &\widetilde{\mathbb{S}}(\tau, \bar{x}(\tau), v)(b(\tau, \bar{x}(\tau), v) - b(\tau, \bar{x}(\tau), \bar{u}(\tau))) \\ &\quad + \nabla\widetilde{\mathbb{S}}(\tau, \bar{x}(\tau), v)(\sigma(\tau, \bar{x}(\tau), v) - \sigma(\tau, \bar{x}(\tau), \bar{u}(\tau))) \\ &\quad + \frac{1}{2}\mathbb{T}(\tau, \bar{x}(\tau), v)(\sigma(\tau, \bar{x}(\tau), v) - \sigma(\tau, \bar{x}(\tau), \bar{u}(\tau)))^2 \\ &\leq 0, \ \forall \ \ v \in V, \ a.s. \end{aligned} \tag{4.10}$$

In example (A), $\bar{u}(t) \equiv 0$ is a singular control in the sense of Pontryagin maximum principle on the whole control region U but not an optimal control.

We show that $\bar{u}(t) \equiv 0$ does not satisfy the second-order necessary optimality condition (4.10). Actually, for any $v \in U$

$$\widetilde{\mathbb{S}}(t, \bar{x}(t), v) = v, \quad \mathbb{T}(t, \bar{x}(t), v) = 0, \quad \forall\ (\omega, t) \in \Omega \times [0, 1].$$

Let $v = 1$, we have

$$\begin{aligned} &\widetilde{\mathbb{S}}(\tau, \bar{x}(\tau), v)\delta b(\tau) + \nabla\widetilde{\mathbb{S}}(\tau, \bar{x}(\tau), v)\delta\sigma(\tau) \\ &\qquad\qquad + \frac{1}{2}\mathbb{T}(\tau, \bar{x}(\tau), v)\delta\sigma(\tau)^2 = 1 > 0, \end{aligned}$$

which contradicts to the condition (4.10).

Example (C)

Let $U = \{-1,\ 0,\ 1\}$. *Consider the control system*

$$\begin{cases} dx(t) = (u(t)-1)dt + (x(t)-u(t))dW(t), & t \in [0,1], \\ x(0) = 1, \end{cases}$$

and the cost functional

$$J(u(\cdot)) = \frac{1}{24}\mathbb{E}\,|x(1)-1|^4.$$

Obviously, $(\bar{x}(\cdot), \bar{u}(\cdot)) \equiv (1,1)$ is the optimal pair. The solutions to the corresponding four adjoint equations with respect to $(\bar{x}(\cdot), \bar{u}(\cdot))$ are

$$\begin{array}{ll} (p_1(t), q_1(t)) = (0,0), & (p_2(t), q_2(t)) = (0,0), \\ (p_3(t), q_3(t)) = (0,0), & (p_4(t), q_4(t)) = (-e^{6-6t}, 0), \\ & \qquad (\omega, t) \in \Omega \times [0,1]. \end{array}$$

Then, we have

$$\begin{aligned}&\mathbb{H}(t,\bar{x}(t),v)=0,\quad \widetilde{\mathbb{S}}(t,\bar{x}(t),v)=0,\\&\mathbb{T}(t,\bar{x}(t),v)=-\frac{1}{2}e^{6-6t}(v-1)^2,\quad \forall v\in U,\ \ \forall\ (\omega,t)\in\Omega\times[0,1],\end{aligned}$$

and

$$\begin{aligned}&\mathbb{T}(t,\bar{x}(t),v)\big(\sigma(\tau,\bar{x}(\tau),v)-\sigma(\tau,\bar{x}(\tau),\bar{u}(\tau))\big)^2\\&=-\frac{1}{2}e^{6-6t}(v-1)^4\leq 0,\qquad \forall v\in U,\ \ \forall\ (\omega,t)\in\Omega\times[0,1].\end{aligned}$$

Therefore, $\bar{u}(t)\equiv 1$ satisfies the second-order necessary condition (4.10).

Further works:

1) Other results for second order necessary conditions.

   [a] H. Frankowska, H. Zhang and X. Zhang. *First and second order necessary conditions for stochastic optimal controls.* J. Differential Equations. 262 (2017), 3689–3736.

2) The second-order necessary condition for stochastic optimal controls with endpoint/state constraints.

   [b] H. Frankowska, H. Zhang and X. Zhang. *Stochastic optimal control problems with control and initial-final states constraints.* SIAM J. Control Optim. 56 (2018), 1823–1855.

   [c] H. Frankowska, H. Zhang and X. Zhang. *Necessary optimality conditions for local minimizers of stochastic optimal control problems with state constraints.* Trans. Amer. Math. Soc. Accepted.

3) The regularity of stochastic optimal controls.
4) Using second order necessary condition to establish numerical methods for solving stochastic optimal controls.
5) Higher order necessary conditions, and applications.
6) The same problems but for stochastic controlled evolution equations in infinite dimensions.

[d] Q. Lü, H. Zhang and X. Zhang. *Second order optimality conditions for optimal control problems of stochastic evolution equations.* Preprint.

Thank you !