SIAM J. CONTROL OPTIM.
Vol. 45, No. 2, pp. 668–698

MAXIMUM PRINCIPLE FOR SINGULAR STOCHASTIC CONTROL PROBLEMS*

F. DUFOUR† AND B. MILLER‡

Abstract. In this paper, an optimal singular stochastic control problem is considered. For this model, a general stochastic maximum principle is obtained by using a time transformation. This is the first version of the stochastic maximum principle that covers nonlinear cases.

Key words. nonlinear stochastic systems, optimal control, singular control, maximum principle, time change

AMS subject classifications. 49J30, 49N25, 93E20

DOI. 10.1137/040612403

1. Introduction.

In this paper a maximum principle is proved for stochastic singular controls. On a filtered probability space $\big(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}\big)$ the set of admissible controls is defined as the class of right continuous with left-hand limits (*corlol*) processes $\big\{(u_t, v_t)\big\}$, $\{\mathcal{F}_t\}$-progressively measurable satisfying the following conditions: $\{v_t\}$ is increasing, $v_T \leq M$ P-a.s., and $u_t \in K$, $|u_t| \leq 1$, where K is a subset of $\mathbb{R}^r$, and $T > 0$ is the finite horizon. Assuming that on the probability space $\big(\Omega, \mathcal{F}, P\big)$ there exists a standard Brownian motion $\{W_t\}$, the state process is defined by the following stochastic differential equation for any admissible control $\big\{(u_t, v_t)\big\}$:

$$x(t) \doteq \zeta + \int_0^t A(s, x(s))ds + \int_{[0,t]} B(s)u(s)dv(s) + \int_0^t D(s, x(s))dW(s), \tag{1.1}$$

and the cost is given by

$$J[C] \doteq E_P\left[G\bigg(\int_0^T |u(s)|dv(s), x(T)\bigg)\right], \tag{1.2}$$

where the functions A, B, D, and G are deterministic.

In recent years, singular stochastic control problems have received considerable attention. The connection between singular control problem and optimal stopping problem has been studied by many authors including Alvarez [1, 2], Boetius [5, 6], Boetius and Kohlmann [7], Chow, Menaldi, and Robin [10], Dufour and Miller [14], El Karoui and Karatzas [16, 17], Karatzas [28, 29], and Karatzas and Shreve [31, 32, 33].

*Received by the editors July 28, 2004; accepted for publication (in revised form) November 28, 2005; published electronically May 12, 2006. This research was supported by a CNRS/Russian Academy of Sciences cooperation (PECO/NEI 9570) and in part by the Nonlinear Control Network and by Russian Basic Investigation Foundation grant 02-01-00361.
http://www.siam.org/journals/sicon/45-2/61240.html
†Corresponding author. MAB, Université Bordeaux I, 351 cours de la Liberation, 33405 Talence Cedex, France (dufour@math.u-bordeaux1.fr).
‡Institute for Information Transmission Problems, 19 Bolshoy Karetny per., Moscow 127994, Russia (bmiller@iitp.ru). Current address: Centre of Modelling of Stochastic Systems, Monash University, 3800 Victoria, Australia (boris.miller@sci.monash.edu.au).

Results on the dynamic programming principle can be found in Boetius [6], Haussmann and Suo [25], Fleming and Soner [22], and Zhu [43]. Sufficient conditions for the existence of optimal singular control for general nonlinear models have been obtained in Dufour and Miller [14] and Haussmann and Suo [24]. Explicit problems have been solved, for example, in [11] and [30]. The authors do not pretend to present here an exhaustive panorama of the literature relative to singular control problems. However, the interested reader may consult Boetius [6] for a survey on stochastic singular control problems including theoretical results and applications.

There is a very extensive literature on the maximum principle for stochastic classical control (see, for example, [3], [4], [9], [20], [21], [26], [36], [38], [41], [42], and the references therein; this list of references is not exhaustive).

In this paper we have chosen to define the control strategy as a pair of processes $\{u(t), v(t)\}$, where $|u(t)| \leq 1$, and $v(t)$ is increasing and right continuous. An equivalent description of the singular stochastic control problem would have been to choose a process $\{\zeta(t)\}$ of bounded variation which acts on the dynamic state through the term $\int_{[0,t]} B(s)d\zeta(s)$. The connection between these two definitions of the control process $\{u(t), v(t)\}$ and $\{\zeta(t)\}$ ($\zeta(t) = \int_{[0,t]} u(s)dv(s)$) has been discussed by many authors (see, for example, [11], [22, page 318], [35], and [43]).

To the best knowledge of the authors the stochastic maximum principle for singular controls was only considered by Cadenillas and Haussmann [8]. In their paper Cadenillas and Haussmann used a different approach and different hypotheses that are presented now in order to bring to the fore the main differences between their results and ours. In [8], the control process is described by a process $\{\zeta(t)\}$ (see the above discussion) of bounded variation, and they do not impose any L^p bounds on the control while we assume that the class of admissible controls $\{v(t)\}$ is such that $v(T) \leq M$ for a constant M. However, in [8] the state process must satisfy a linear stochastic differential equation (the functions A and D are assumed to be linear with respect to the state variable $\{x(t)\}$), and the cost function is convex. In our work, we suppose that the state process is defined by a general nonlinear stochastic differential equation, and we do not impose a convexity hypothesis on the cost function (see assumptions A(1)–A(3) in the next section). In many aspects the results obtained in [8] and here are different and complementary.

In general terms, the approach we used to obtain the maximum principle for singular control problems can be divided into three steps. The first step is to convert the original singular control problem into a classical control problem by using a special time transformation. In order to be concise, the description of this method is briefly presented, and only the important properties, which we need here, are given. Although this technique of time change is similar to the one already described in [14], it must be pointed out that the approach used here presents some technical differences that are explained in section 3. The second step is to derive the maximum principle for the auxiliary control problem (see Theorem 4.5). Important properties of the adjoint variables (see Theorem 4.4 and in particular (4.13)) that will be used to obtain the singular maximum principle are derived. The last step consists of recovering from the auxiliary maximum principle the original state and control variables by using a time change, thus giving a maximum principle for the singular control problem (see Theorem 5.9). The form of the maximum principle we obtained turns out to be different from the one derived in [8] since the adjoint variables have a singular part (see Definition 5.1 and in particular the second term of the right-hand side of (5.1)) and since the optimal singular control maximizes an Hamiltonian a.s. with respect to the Doleans-Dade measure generated by the optimal control $\{v(t)\}$ (see the detailed discussion before Theorem 5.9).

REMARK 1.1. *It must be pointed out that our time change technique cannot be used to solve singular control problems in the context of infinite horizon.*

The finite fuel constraint $(v(T) \leq M)$ *is a crucial assumption in order to derive the maximum principle by using our method of time transformation. Indeed, in* [14], *it has been shown that our approach of time change enables us to convert a nonlinear singular control problem into an auxiliary control problem where the control variables are of the classical type and where the controller must choose a stopping time. However, since it is very difficult to derive a general maximum principle combining classical control and optimal stopping, we need to impose the finite fuel constraint* $(v(T) \leq M)$ *to convert the singular control problem into the classical control problem where the horizon is fixed. In* [8], *Cadenillas and Haussmann do not have to impose a finite fuel constraint, but they used different hypotheses (see the discussion above for a comparison between their approach and ours).*

Nevertheless, we have shown in [14] *that our method of time transformation can be applied to study the existence of optimal singular control in a general context where the finite fuel constraint is not a necessary assumption.*

Singular control problems have been studied under various hypotheses (for example, the finite horizon case can be found in [7, 16, 28, 29, 31]*; the infinite horizon case in* [7, 11, 28, 33]*; with a finite fuel constraint in* [16, 28, 33]*).*

The paper is organized as follows. In section 2, we formulate the singular control problem. The time change and the auxiliary control problem is briefly described in section 3. Section 4 deals with the auxiliary maximum principle and its properties. In section 5, the main results are obtained and in particular the stochastic maximum principle for singular controls (see Theorem 5.9). A simple example is presented in section 6 to illustrate the set of necessary conditions that must be satisfied by an optimal solution. In the last section, we make some comments about possible generalizations of our work.

NOTATION.

The Lebesgue measure on $\mathbb{R}$ is denoted by λ.

$\mathbb{N}_N$ is the set of the first N integers, that is, $\mathbb{N}_N = \{1, \dots, i, \dots, N\}$.

$\mathbb{N}^* \doteq \{k \in \mathbb{N} \,:\, k > 0\}$ and $\mathbb{R}_+ \doteq \{x \in \mathbb{R} \,:\, x \geq 0\}$.

If V is a vector, V_i denotes the ith component of V.

If M is a matrix, M_i denotes a vector given by the ith column of the matrix M, and M_{ij} is the element corresponding to the ith row and the jth column.

$(^\top)$ denotes the transpose operation.

For $x \in \mathbb{R}^k$, $|x|$ denotes its Euclidean norm and for a matrix $A \in \mathbb{R}^{k\times d}$ the norm of A is defined by $|A| = \sqrt{\operatorname{tr}\big[AA^\top\big]}$.

For $K \subset \mathbb{R}^k$, $B_1(K) \doteq \{x \in K : |x| \leq 1\}$, and $S_1(K) \doteq \{x \in K : |x| = 1\}$.

$\mathcal{S}_n$ denotes the set of all $(n \times n)$ real symmetric matrices.

$0_n \in \mathbb{R}^n$ is the zero vector.

The indicator function of a set A is defined as $I_A(x)$.

The function δ defined on $\mathbb{N} \times \mathbb{N}$ is such that $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

For $x \in \mathbb{R}$, x^+ is defined by $x^+ = \frac{1}{x}$ if $x \neq 0$ and $x^+ = 0$ if $x = 0$.

If V is a metric space, then $\mathcal{B}(V)$ denotes its associated borel σ-field.

A filtered probability space $\big(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}\big)$ is said to satisfy the usual hypotheses if the probability space $\big(\Omega, \mathcal{F}, P\big)$ is complete and if the filtration $\{\mathcal{F}_t\}$ is right continuous and if every $\mathcal{F}_t$ contains all P-null sets of $\mathcal{F}$.

Let $(\Omega,\, \mathcal{F},\, P,\, \{\mathcal{F}_t\})$ be a filtered probability space and $\{\theta_t\}$ be a $[0,1]$-valued, $\{\mathcal{F}_t\}$-progressively measurable process. Suppose that V denotes any of the spaces $\mathbb{R}^k$, $\mathbb{R}^{k\times d}$, or $\mathcal{S}_k$.
Then $L^2\big(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}; [0,T]; V\big)$ (respectively, $L^2_\theta(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}; [0,T]; V)$) denotes the set of V-valued processes $\{x(t)\}$ which are $\{\mathcal{F}_t\}$-progressively measurable and satisfy $E_P[\int_0^T |x(s)|^2 ds] < +\infty$ (respectively, $E_P[\int_0^T |x(s)|^2[1-\theta(s)]ds] < +\infty$).
If $\mathbb{F}$ denotes the filtered probability space $\big(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}\big)$, then $L^2\big(\mathbb{F}; [0,T]; V\big)$ (respectively, $L^2_\theta(\mathbb{F}; [0,T]; V)$) is used to write in a more compact form $L^2\big(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}; [0,T]; V\big)$ (respectively, $L^2_\theta(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}; [0,T]; V)$).
Moreover, $L^2\big(\Omega, \mathcal{F}, P; V\big)$ denotes the set of V-valued random variables X defined on the probability space $\big(\Omega, \mathcal{F}, P\big)$ such that $E_P\big[|X|^2\big] < +\infty$.
Let $\big(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}\big)$ be a filtered probability space; if $\{A(t)\}$ is a *corlol*, adapted process of finite variation on each interval $[0,t]$, then dA denotes the measure associated with the distribution function $\{A(t)\}$. For $\{H(t)\}$, a progressively measurable process, the integral process $\int_0^t H(s)dA(s)$ is denoted by $H \cdot A_t$. If $\{A(t)\}$ is an increasing *corlol*, adapted process, the measure defined on $(\mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F})$ by $E_P[\int_0^{+\infty} I_C(s)dA(s)]$ for $C \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ is denoted by $\mathcal{M}_A$.
Let $\big(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}\big)$ be a filtered probability space satisfying the usual hypotheses and supporting a standard m-dimensional Brownian motion $\{W_t\}$. Then $\{\mathcal{F}^W_t\}$ denotes the augmentation of the natural filtration generated by $\{W_t\}$.

In order to define the state processes, let us introduce the following data:

- T and M are fixed real numbers.
- K is a subset of $\mathbb{R}^r$.
- ζ is a fixed vector in $\mathbb{R}^n$.
- $A \,:\, \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$.
- $B \,:\, \mathbb{R} \to \mathbb{R}^{n\times r}$.
- $D \,:\, \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$.
- $G \,:\, \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$.
- $N \,:\, \mathbb{R} \to \mathbb{R}^2$ such that $N(t) = \binom{t-T}{t^2-T^2}$.

Let us introduce the following notation:

$\mathcal{A} \,:\, \mathbb{R} \times \mathbb{R}^n \times B_1(K) \times [0,1] \longrightarrow \mathbb{R}^{n+2}$ is defined by

$$\mathcal{A}(t,x,u,z) \doteq \begin{pmatrix} 1-z \\ A(t,x)(1-z) + zB(t)u \\ z|u| \end{pmatrix}$$

and

$\mathcal{D} \,:\, \mathbb{R} \times \mathbb{R}^n \times [0,1] \longrightarrow \mathbb{R}^{n+2}$ is defined by

$$\mathcal{D}(t,x,z) \doteq \begin{pmatrix} 0 \\ D(t,x)\sqrt{1-z} \\ 0 \end{pmatrix}$$

$(\forall (t,x,u,z,p,q,P) \in \mathbb{R} \times \mathbb{R}^n \times B_1(K) \times [0,1] \times \mathbb{R}^{2+n} \times \mathbb{R}^{(2+n)\times m} \times \mathbb{R}^{n\times n})$,

$$\mathcal{H}(t,x,u,z,p,q) \doteq \mathcal{A}(t,x,u,z)^\top p + \mathrm{tr}\big[\mathcal{D}(t,x,z)^\top q\big] \tag{1.3}$$

$(\forall(t,x,u,z,p^1,p^2,p^3,q,P,\psi) \in \mathbb{R}\times\mathbb{R}^n\times B_1(K)\times[0,1]\times\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^{n\times m}\times\mathbb{R}^{n\times n}\times\mathbb{R}^3)$,

$$J(t,x,p^2,q) \doteq A(t,x)^\top p^2 + \operatorname{tr}\big[D(t,x)^\top q\big], \tag{1.4}$$

$$\begin{aligned} H(t,x,u,z,p^1,p^2,p^3,q,P,\psi) \doteq &(\psi_2 p^1 - \psi_1 - 2T\psi_3)(1-z) \\ &+ \psi_2\big[(1-z)A(t,x) + zB(t)u\big]^\top p^2 \\ &+ \psi_2 z|u|p^3 + \psi_2\sqrt{1-z}\operatorname{tr}\big[D(t,x)^\top q\big] \\ &+ \frac{\psi_2}{2}\operatorname{tr}[D(t,x)^\top P D(t,x)](1-z). \end{aligned} \tag{1.5}$$

The following assumptions will be used in the paper:

(A1) The maps A, B, D, and G are C^2.

(A2) The first and second derivatives of A, B, D and the second derivative of G are bounded. The maps $A(t,x)$, $B(t,x)$, $D(t,x)$ are bounded by $C(1+|t|+|x|)$. The first derivative of $G(w,x)$ is bounded by $C(1+|w|+|x|)$.

(A3) $(\forall x\in\mathbb{R}^n)$, $(\forall(w_1,w_2)\in\mathbb{R}\times\mathbb{R})$ if $w_1\le w_2$, then $G(w_1,x)\le G(w_2,x)$.

In the rest of the paper, the derivative of the function B will be denoted by B_t, and the partial derivatives of the function A (respectively, G, $\mathcal{H}$, and J) with respect to the first variable will be denoted by A_t (respectively, $\mathcal{H}_t$ and J_t) and with respect to the second variable will be denoted by A_x (respectively, $\mathcal{H}_x$ and J_x). For $j\in\mathbb{N}_m$, D_{jt} (respectively, D_{jx}) denotes the partial derivative of the function D_j with respect to the first variable (respectively, the second variable). The partial derivative of G with respect to the first variable will be denoted by G_w and with respect to the second variable will be denoted by G_x. The second partial derivatives of the function G (respectively, $\mathcal{H}$ and J) with respect to the second variable will be denoted by G_{xx} (respectively, $\mathcal{H}_{xx}$ and J_{xx}).

2. Problem statement. In this section, we formulate the original singular stochastic control problem presented in the introduction using the formulation described in [18] and in [23].

DEFINITION 2.1. *A singular control is defined by the following term:*

$$C \doteq (\Omega,\mathcal{F},P,\{\mathcal{F}_t\},\{u(t),v(t)\},\{W(t)\},\{x(t)\}),$$

where

(i) $(\Omega,\mathcal{F},P)$ *is a complete probability space with a right continuous complete filtration* $\{\mathcal{F}_t\}$*;*

(ii) $\{W(t)\}$ *is a standard* m*-dimensional* $\{\mathcal{F}_t\}$*-Brownian motion;*

(iii) $\{u(t),v(t)\}$ *is a* $B_1(K)\times\mathbb{R}_+$*-valued,* corlol, $\{\mathcal{F}_t\}$*-progressively measurable process such that* $\{v(t)\}$ *is increasing and satisfies*

$$v(T)\le M; \tag{2.1}$$

(iv) $\{x(t)\}$ *is an* $\mathbb{R}^n$*-valued,* corlol, $\{\mathcal{F}_t\}$*-progressively measurable process such that* $(\forall t\in[0,T])$

$$\begin{aligned} x(t) \doteq \zeta &+ \int_0^t A(s,x(s))ds + \int_{[0,t]} B(s)u(s)dv(s) \\ &+ \int_0^t D(s,x(s))dW(s), \end{aligned} \tag{2.2}$$

and $x(0-)=\zeta$*.*

We write $\mathfrak{C}$ *for the set of controls satisfying the previous conditions.*

The cost is given by

$$J[C] \doteq E_P\left[G\left(\int_0^T |u(s)|dv(s), x(T)\right)\right]. \tag{2.3}$$

The set $\mathfrak{C}^a$ *of admissible controls is defined by*

$$\mathfrak{C}^a \doteq \{C \in \mathfrak{C} : J[C] < \infty\}. \tag{2.4}$$

The singular control problem is defined by the minimization of $J[C]$ on $\mathfrak{C}^a$. Assuming the existence of an optimal singular control $\widetilde{C}$, the aim of the paper is to derive necessary conditions for $\widetilde{C}$ to be optimal in terms of variational inequalities (see the maximum principle presented in Theorem 5.9).

3. The auxiliary control problem. In this section, it is shown that the original singular control problem can be converted into a classical control problem by using a time transformation (see Propositions 3.4, 3.5, and 3.7). We used the technique previously described in [14]. These results are presented here with minimal details in order to be concise. The interested reader may consult [14] to have a complete description of this approach. However, it must be pointed out that the model under consideration in [14] and the one presented here are different. Indeed, in order to derive a stochastic maximum principle one needs to have real-valued state constraints (see (3.26) and the definition of N), whereas the constraints used in [14] to derive sufficient conditions for the existence of an optimal control can take infinite value. Moreover, an important difference is that one needs to have a special measurability property for the auxiliary control problem (see the last part of Proposition 3.7) to ensure the existence of the adjoint variables defined by backward stochastic differential equations (see Corollary 4.2).

Assume the existence of an optimal singular control denoted by

$$\widetilde{C} \doteq \left(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}_t\}, \{\widetilde{u}(t), \widetilde{v}(t)\}, \{\widehat{W}(t)\}, \{\widetilde{x}(t)\}\right)$$

such that $\{\widetilde{u}(t), \widetilde{v}(t)\}$ is $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$-progressively measurable.

REMARK 3.1. *The existence problem for singular stochastic control has already been studied under general hypotheses in many papers (see, for example,* [14] *and* [24] *and the references therein). As just stated, it is assumed that the optimal control* $\widetilde{C}$ *satisfies an assumption related to the measurability of the singular control* $\{\widetilde{u}(t), \widetilde{v}(t)\}$ *with respect to the filtration generated by the noise* $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$. *It must be pointed out that this hypothesis is classical (see, for example, assumption* $(S0)$ *on page* 114 *in* [41] *or the definition of the underlying filtration on page* 967 *in* [38]*).*

With the next proposition, we show how it is possible to construct an optimal singular control $\widehat{C}$ satisfying $\widehat{v}(T) = M$ from the optimal singular control $\widetilde{C}$.

PROPOSITION 3.2. *The control* $\widehat{C}$ *defined by*

$$\widehat{C} \doteq \left(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}_t\}, \{\widehat{u}(t), \widehat{v}(t)\}, \{\widehat{W}(t)\}, \{\widehat{x}(t)\}\right), \tag{3.1}$$

where

$$\widehat{v}(t) = \widetilde{v}(t) I_{[0,T[} + (M - \widetilde{v}(T) + \widetilde{v}(t)) I_{[T,+\infty[}, \tag{3.2}$$

$$\widehat{u}(t) = \widetilde{u}(t)\left[\frac{\widetilde{v}(T) - \widetilde{v}(T-)}{M - \widetilde{v}(T-)} I_{[T,+\infty[\times\{\widetilde{v}(T)<M\}} + I_{[T,+\infty[\times\{\widetilde{v}(T)=M\}}\right]$$
$$+ \widetilde{u}(t) I_{[0,T[}, \tag{3.3}$$

is optimal. Moreover, $\widehat{v}(T) = M$*, and* $\{\widehat{u}(t), \widehat{v}(t)\}$ *is a* $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$*-progressively measurable process.*

Proof. From the definition of $\{\widehat{u}(t)\}$ and $\{\widehat{v}(t)\}$, we have that $\widehat{v}(T) = M$, and

$$(\forall t \in [0,T[),\quad \widehat{u}(t) = \widetilde{u}(t),\quad \widehat{v}(t) = \widetilde{v}(t),\quad \text{and}\quad \widehat{u}(T)\Delta\widehat{v}(T) = \widetilde{u}(T)\Delta\widetilde{v}(T); \tag{3.4}$$

consequently, we obtain that $\widehat{x}(T) = \widetilde{x}(T)$, and $\int_{[0,T]} |\widehat{u}(t)| d\widehat{v}(t) = \int_{[0,T]} |\widetilde{u}(t)| d\widetilde{v}(t)$, showing that $J[\widehat{C}] = J[\widetilde{C}]$. Clearly, $\{\widehat{u}(t), \widehat{v}(t)\}$ is a $B_1(K)\times\mathbb{R}_+$-valued, *corlol*, $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$-progressively measurable process implying that $\widehat{C}$ is optimal. □

REMARK 3.3. *Now we will work with the optimal control* $\widehat{C}$ *for technical reasons. However, by using Propositions* 3.2 *and* 5.8*, a general stochastic maximum principle will be derived in terms of the optimal control* $\widetilde{C}$ *giving the full generality to our result (see Theorem* 5.9*).*

PROPOSITION 3.4. *Denote the process* $\{t + \widehat{v}(t)\}$ *by* $\{\widehat{\Gamma}(t)\}$*. Let* $\{\eta^*(t)\}$ *be the right inverse of* $\{\widehat{\Gamma}(t)\}$*. Then* $\{\eta^*(t)\}$ *is a continuous time change such that the probability space* $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}_{\eta^*(t)}^{\widehat{W}}\})$ *satisfies the usual hypotheses. Moreover, there exists a* $[0,1]$*-valued,* $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$*-progressively measurable process* $\{\widehat{z}(t)\}$ *such that*

$$\widehat{v}(t) = \int_{[0,t]} \widehat{z}(s) d\widehat{\Gamma}(s). \tag{3.5}$$

Define the $B_1(K)\times[0,1]$*-valued,* $\{\widehat{\mathcal{F}}_{\eta^*(t)}^{\widehat{W}}\}$*-progressively measurable process* $\{(\alpha^*(t), \theta^*(t))\}$ *by*

$$\alpha^*(t) = \widehat{u}(\eta^*(t)) \quad \text{and} \quad \theta^*(t) = \widehat{z}(\eta^*(t)). \tag{3.6}$$

Then

$$\eta^*(t) = \int_0^t (1-\theta^*(s)) ds, \tag{3.7}$$

$$J[\widehat{C}] = E_{\widehat{P}}\Big[G\big(\mu^*(T+M), \xi^*(T+M)\big)\Big], \tag{3.8}$$

$$E_{\widehat{P}}\Big[N\big(\eta^*(T+M)\big)\Big] = 0_2, \tag{3.9}$$

where the processes $\{\xi^*(t)\}$ *and* $\{\mu^*(t)\}$ *are solutions of the following equations:*

$$\begin{aligned} \xi^*(t) \doteq \zeta &+ \int_0^t A(\eta^*(s), \xi^*(s))(1-\theta^*(s)) ds + \int_0^t \theta^*(s) B(\eta^*(s)) \alpha^*(s) ds \\ &+ \int_0^t D(\eta^*(s), \xi^*(s)) \sqrt{(1-\theta^*(s))(1-\theta^*(s))^+} d\widehat{W}(\eta^*(s)), \end{aligned} \tag{3.10}$$

$$\mu^*(t) \doteq \int_0^t |\alpha^*(s)| \theta^*(s) ds. \tag{3.11}$$

Moreover,

$$\widehat{x}(t) = \xi^*(\widehat{\Gamma}(t)) \quad \text{and} \quad \mu^*(T+M) = \int_{[0,T]} |\widehat{u}(s)| d\widehat{v}(s) \tag{3.12}$$

and

$$\widehat{v}(t) = \int_0^{\widehat{\Gamma}_t} \theta^*(s) ds \quad \text{and} \quad \int_{[0,t]} |\widehat{u}(s)| d\widehat{v}(s) = \int_0^{\widehat{\Gamma}(t)} |\alpha^*(s)| \theta^*(s) ds. \tag{3.13}$$

Proof. Following Proposition 3.1, Proposition 3.2, and Theorem 4.2 in [14], the result can be obtained. □

Let $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P}, \{\widetilde{\mathcal{F}}_t\})$ be a filtered probability space satisfying the usual hypotheses and supporting a standard m-dimensional Brownian motion $\{\widetilde{V}_t\}$. Define by $(\Omega, \mathcal{G}, Q, \{\mathcal{G}_t\})$ the usual augmentation of the filtered probability space $\{\widehat{\Omega} \times \widetilde{\Omega}, \widehat{\mathcal{F}} \otimes \widetilde{\mathcal{F}}, \widehat{P} \otimes \widetilde{P}, \widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(t)} \otimes \widetilde{\mathcal{F}}_t\}$.

A random variable $\widehat{X}$ defined on $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P})$ may be viewed as defined on $(\Omega, \mathcal{G}, Q)$ by setting $\overline{X}(\widehat{\omega}, \widetilde{\omega}) = \widehat{x}(\widehat{\omega})$ for $(\widehat{\omega}, \widetilde{\omega}) \in \widehat{\Omega} \times \widetilde{\Omega}$ and similarly for a random variable defined on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$. Consequently, let us introduce on $(\Omega, \mathcal{G}, Q, \{\mathcal{G}_t\})$ the following processes:

$$\begin{cases} \overline{\alpha}(t, \widehat{\omega}, \widetilde{\omega}) \doteq \alpha^*(t, \widehat{\omega}), \quad \overline{\theta}(t, \widehat{\omega}, \widetilde{\omega}) \doteq \theta^*(t, \widehat{\omega}), \quad \overline{\eta}(t, \widehat{\omega}, \widetilde{\omega}) \doteq \eta^*(t, \widehat{\omega}), \\ \overline{\xi}(t, \widehat{\omega}, \widetilde{\omega}) \doteq \xi^*(t, \widehat{\omega}), \quad \overline{\mu}(t, \widehat{\omega}, \widetilde{\omega}) \doteq \mu^*(t, \widehat{\omega}), \quad \overline{W}(t, \widehat{\omega}, \widetilde{\omega}) \doteq \widehat{W}(t, \widehat{\omega}), \\ \qquad \widetilde{W}(t, \widehat{\omega}, \widetilde{\omega}) \doteq \widetilde{V}(t, \widetilde{\omega}). \end{cases} \tag{3.14}$$

PROPOSITION 3.5. *On $(\Omega, \mathcal{G}, Q, \{\mathcal{G}_t\})$, the process $\{\overline{V}(t)\}$ defined by*

$$\overline{V}(t) \doteq \int_0^t \sqrt{(1 - \overline{\theta}(s))^+} d\overline{W}(\overline{\eta}(s)) + \int_0^t \sqrt{1 - (1 - \overline{\theta}(s))(1 - \overline{\theta}(s))^+} d\widetilde{W}(s) \tag{3.15}$$

is a standard m-dimensional $\{\mathcal{G}_t\}$-Brownian motion. The process $\{\overline{\xi}(t), \overline{\eta}(t), \overline{\mu}(t)\}$ is the unique solution of the following equations:

$$\begin{aligned} \overline{\xi}(t) = \zeta &+ \int_0^t A(\overline{\eta}(s), \overline{\xi}(s))[1 - \overline{\theta}(s)]ds + \int_0^t B(\overline{\eta}(s))\overline{\alpha}(s)\overline{\theta}(s)ds \\ &+ \int_0^t D(\overline{\eta}(s), \overline{\xi}(s))\sqrt{1 - \overline{\theta}(s)} d\overline{V}(s), \end{aligned} \tag{3.16}$$

$$\overline{\eta}(t) = \int_0^t [1 - \overline{\theta}(s)]ds, \tag{3.17}$$

$$\overline{\mu}(t) = \int_0^t |\overline{\alpha}(s)|\overline{\theta}(s)ds. \tag{3.18}$$

Moreover,

$$J[\overline{C}] = E_Q\Big[G\big(\overline{\mu}(T + M), \overline{\xi}(T + M)\big)\Big], \tag{3.19}$$

$$E_Q\Big[N\big(\overline{\eta}(T + M)\big)\Big] = 0_2. \tag{3.20}$$

Proof. From Theorem 2.75 in [27], it follows that $\{\overline{V}(t)\}$ is a standard m-dimensional $\{\mathcal{G}_t\}$-Brownian motion. Using Theorem 6 in [39, page 194], we obtain that the solution $\{\xi(t), \eta(t), \mu(t))\}$ of (3.16)–(3.18) exists and is unique. From the definition of $\{\overline{V}(t)\}$, it is easy to show that $\{\xi(t)\}$ is the unique solution of the following equation:

$$\begin{aligned} \xi(t) \doteq \zeta &+ \int_0^t A(\overline{\eta}(s), \xi(s))(1 - \overline{\theta}(s))ds + \int_0^t B(\overline{\eta}(s))\overline{\alpha}(s)\overline{\theta}(s)ds \\ &+ \int_0^t D(\overline{\eta}(s), \xi(s))\sqrt{(1 - \overline{\theta}(s))(1 - \overline{\theta}(s))^+} d\overline{W}(\overline{\eta}(s)). \end{aligned} \tag{3.21}$$

However, combining the fact that the process $\{\xi^*(t)\}$ is the solution of (3.10) and Proposition 10.46 in [27], it follows that $\{\overline{\xi}(t)\}$ satisfies (3.16). Therefore, $\{\overline{\xi}(t)\}$ is the unique solution of (3.16). Moreover, it is clear from their definitions that the processes $\{\mu(t)\}$ and $\{\overline{\mu}(t)\}$ (respectively, $\{\eta(t)\}$ and $\{\overline{\eta}(t)\}$) are indistinguishable.

Finally, (3.19) and (3.20) follow easily from (3.8) and (3.9) and the definition of the probability Q. □

On the probability space $(\Omega, \mathcal{G}, Q)$, define the filtration $\mathcal{J}_t \doteq \widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(t)} \otimes \{\emptyset, \widetilde{\Omega}\}$.

The set of auxiliary control $\mathcal{E}$ is the set of $\{\mathcal{J}_t\}$-progressively measurable processes defined on $(\Omega, \mathcal{G}, Q, \{\mathcal{G}_t\})$ and taking their value in $B_1(K) \times [0,1]$. For any $\{(\alpha(t), \theta(t))\}$ in $\mathcal{E}$, the auxiliary state process $(\eta(t), \xi(t), \mu(t))$ is defined on $(\Omega, \mathcal{G}, Q, \{\mathcal{G}_t\})$ by

$$\eta(t) \doteq \int_0^t (1-\theta(s))ds, \tag{3.22}$$

$$\begin{aligned} \xi(t) \doteq \zeta &+ \int_0^t A(\eta(s), \xi(s))(1-\theta(s))ds + \int_0^t B(\eta(s))\alpha(s)\theta(s)ds \\ &+ \int_0^t D(\eta(s), \xi(s))\sqrt{(1-\theta(s))}d\overline{V}(s), \end{aligned} \tag{3.23}$$

$$\mu(t) \doteq \int_0^t |\alpha(s)|\theta(s)ds. \tag{3.24}$$

Note that for any $\{(\alpha(t), \theta(t))\}$ in $\mathcal{E}$, the previous system admits a unique solution. Moreover, we have $E_Q[G(\mu(T+M), \xi(T+M))] < \infty$.

The associated cost functional is defined by

$$\mathcal{M}[\alpha, \theta] \doteq E_Q\Big[G\big(\mu(T+M), \xi(T+M)\big)\Big]. \tag{3.25}$$

DEFINITION 3.6. *The set of admissible auxiliary control $\mathcal{E}_{ad}$ is defined by the set of processes $\{(\alpha(t), \theta(t))\} \in \mathcal{E}$ such that the corresponding auxiliary state process $\{(\eta(t), \xi(t), \mu(t))\}$ satisfies the following constraint:*

$$E_Q\Big[N\big(\eta(T+M)\big)\Big] = 0_2. \tag{3.26}$$

The auxiliary control problem is to minimize the cost (3.25) over $\mathcal{E}_{ad}$.

PROPOSITION 3.7. *The auxiliary control process $\{(\overline{\alpha}(t), \overline{\theta}(t))\}$ defined by* (3.14) *is optimal. Moreover, $\{(\overline{\alpha}(t), \overline{\theta}(t))\}$ and the corresponding optimal auxiliary state $\{(\overline{\eta}(t), \overline{\xi}(t), \overline{\mu}(t))\}$ are $\{\mathcal{J}_t\}$-progressively measurable processes.*

Proof. Using Proposition 3.5, and following the same arguments as in the proof of Theorem 4.6 in [14], the result can be obtained. □

LEMMA 3.8. *Denote by $\{\tau_n\}_{n\in\mathbb{N}^*}$ the sequence of $\{\widehat{\mathcal{F}}^{\widehat{W}}_t\}$-stopping times which exhausts the jumps of $\{\widehat{v}(t)\}$. Then $\{\widehat{\Gamma}(\tau_n)\}_{n\in\mathbb{N}^*}$ and $\{\widehat{\Gamma}(\tau_n-)\}_{n\in\mathbb{N}^*}$ are sequences of $\{\widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(t)}\}$-stopping times.*

Proof. Since $\{\widehat{\Gamma}(t)\}$ is a $\{\widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(t)}\}$-progressively measurable process, then there exists a sequence of $\{\widehat{\mathcal{F}}^{\widehat{W}}_t\}$-stopping times, $\{\tau_n\}_{n\in\mathbb{N}^*}$, exhausting the jumps of $\{\widehat{\Gamma}(t)\}$. Since $\{\eta^*(t)\}$ is a time change on $\big(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}^{\widehat{W}}_t\}\big)$ and using item *a*) of Lemma 10.5 in [27], it follows that $\{\widehat{\Gamma}(\tau_n)\}_{n\in\mathbb{N}^*}$ is a sequence of $\{\widehat{\mathcal{F}}_{\eta^*(t)}\}$-stopping times. Moreover, using the fact that $\{\widehat{\Gamma}(\tau_n-) > t\} = \{\eta^*(t) < \tau_n\}$, it follows that $\{\widehat{\Gamma}(\tau_n-)\}_{n\in\mathbb{N}^*}$ is a sequence of $\{\widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(t)}\}$-stopping times. □

For the rest of the paper, we shall use the following notation for the different filtered probability space under consideration:

$$\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*} \doteq \Big(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(t)}\}\Big), \quad \mathbb{G} \doteq \Big(\Omega, \mathcal{G}, Q, \{\mathcal{G}_t\}\Big), \quad \mathbb{J} \doteq \Big(\Omega, \mathcal{G}, Q, \{\mathcal{J}_t\}\Big).$$

4. The maximum principle for the auxiliary control problem. In this section, a maximum principle is obtained for the auxiliary singular control problem. Some important properties are derived to render possible the time change in order to obtain the maximum principle for the original singular control problem presented in the next section (see Theorem 5.9). In Proposition 4.1 and Corollary 4.2 the existence of a general backward stochastic differential equations is established. In particular, this result states there exists an $\{\mathcal{J}_t\}$-adapted solution, labeled $(Y(t), X(t))$, to (4.4), although the backward equation is driven by a Wiener process, labeled $\{\overline{V}(t)\}$, such that $\mathcal{G}^{\overline{V}}_t \not\subset \mathcal{J}_t$ and $\mathcal{J}_t \not\subset \mathcal{G}^{\overline{V}}_t$. However, more importantly, a special expression is obtained for $X(t)$ (see (4.5)) that will be crucial in what follows. Using these results, Theorem 4.4 states the existence of the adjoint variables for the auxiliary control variables with an important property shown in (4.13). The stochastic maximum principle for the auxiliary control problem is then obtained in terms of these adjoint variables (see Theorem 4.5).

PROPOSITION 4.1. *On the filtered probability space $\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}$, let us assume that the function $f : \widehat{\Omega} \times [0, T+M] \times \mathbb{R}^k \times \mathbb{R}^{k\times m} \to \mathbb{R}^k$ is $\mathcal{M} \otimes \mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^{k\times m})$-measurable, where $\mathcal{M}$ denotes the progressive σ-field, and satisfies the following:*

(i) $f(.,0,0) \in L^2(\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}; [0,T+M]; \mathbb{R}^k)$.

(ii) *There exists a constant $L > 0$ such that* $(\forall (y, \overline{y}, z, \overline{z}) \in \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^{k\times m} \times \mathbb{R}^{k\times m})$

$$|f(t,y,z) - f(t,\overline{y},\overline{z})| \le L\,[|y-\overline{y}| + |z-\overline{z}|] \qquad \lambda \otimes \widehat{P}\text{-}a.s.$$

Then for any given $Y^* \in L^2(\widehat{\Omega}, \widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(T+M)}, \widehat{P}; \mathbb{R}^k)$, the backward stochastic differential equation

$$\begin{aligned} Y^*(t) = Y^* - &\int_t^{T+M} f(s, Y^*(s), [1-\theta^*(s)]Z^*(s))ds \\ &- \int_t^{T+M} Z^*(s) d\widehat{W}(\eta^*(s)) \end{aligned} \tag{4.1}$$

admits a unique solution in the following class of processes:

$$\Big(\{Y^*(t)\}, \{Z^*(t)\}\Big) \in L^2\Big(\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}; [0,T+M]; \mathbb{R}^k\Big) \times L^2_{\theta^*}\Big(\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}; [0,T+M]; \mathbb{R}^k\Big)$$

with $\{Y^*(t)\}$ continuous.

Proof. Clearly, the function f satisfies hypotheses (5.8) and (5.9) in [15, page 29], where the underlying filtered probability space is given by $\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}$. Now, using Theorem 6.1 in [15], it follows that there exists a unique tiple $\big(\{Y^*(t)\}, \{Z^*(t)\}, \{N^*(t)\}\big)$ such that $(\forall t \in [0, T+M])$

$$\begin{aligned} Y^*(t) = Y^* - \int_t^{T+M} f(s, Y^*(s), [1-\theta^*(s)]Z^*(s))ds - \int_t^{T+M} Z^*(s) d\widehat{W}(\eta^*(s)) \\ - \int_t^{T+M} dN^*(s), \end{aligned} \tag{4.2}$$

where

$$\{Y^*(t)\} \in L^2\Big(\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*};[0,T+M];\mathbb{R}^k\Big),$$
$$\{Z^*(t)\} \in L^2_{\theta^*}\Big(\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*};[0,T+M];\mathbb{R}^{k\times m}\Big),$$

and $\{N^*(t)\}$ is a *corlol*, $\mathbb{R}^k$-valued, $\{\widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(t)}\}$-martingale satisfying

$$(\forall (i,j) \in \mathbb{N}_k \times \mathbb{N}_k), \quad (\forall t \in [0,T+M]), \quad [\widehat{W}_i(\eta^*(.)), N^*_j]_t = 0 \tag{4.3}$$

with $E_{\widehat{P}}[\mathrm{tr}[[N^*,N^*]_{T+M} - [N^*,N^*]_0]] < +\infty$.

Combining the optional stopping theorem, the martingale representation theorem (see Theorem 4.15 in [34, p. 182]), and Proposition (1.5) in [40, p. 181], we obtain that there exists a process $\{\Phi(t)\} \in L^2_{\theta^*}(\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*};[0,T+M];\mathbb{R}^{k\times m})$ such that

$$N^*(t) = E_{\widehat{P}}\left[N^*(T)\right] + \int_0^t \Phi(s)d\widehat{W}(\eta^*(s)).$$

Combining the previous equation, (3.7), and (4.3), we obtain that $\forall (i,j) \in \mathbb{N}_k \times \mathbb{N}_k$, $\forall t \in [0,T+M]$, $\int_0^t \Phi_{ij}(s)(1-\theta^*(s))ds = 0$. Consequently $(\forall t \in [0,T+M])$, $\int_t^{T+M} dN^*(s) = 0$, and the result follows. □

Note that this result is not an immediate consequence of general results for backward stochastic differential equations (see, for example, [15]) since we need in particular to show that the last term in (4.2) cancels.

The next result demonstrates the existence, for a general backward stochastic differential equation driven by the Wiener process $\{\overline{V}(t)\}$ and defined on the probability space $(\Omega,\mathcal{G},Q,\{\mathcal{G}_t\})$, of an $\{\mathcal{J}_t\}$-adapted solution satisfying (4.5). This result will be used to show the existence of the adjoint variables defined on the same probability space.

COROLLARY 4.2. *On the filtered probability space* $\mathbb{G}$*, let us assume that the function* $g : \Omega \times [0,T+M] \times \mathbb{R}^k \times \mathbb{R}^{k\times m} \to \mathbb{R}^k$ *is* $\mathcal{N}\otimes\mathcal{B}(\mathbb{R}^k)\otimes\mathcal{B}(\mathbb{R}^{k\times m})$*-measurable, where* $\mathcal{N}$ *denotes the progressive* σ*-field when the probability space* $(\Omega,\mathcal{G},Q)$ *is equipped with the filtration* $\{\mathcal{J}_t\}$ *and satisfies the following:*

(i) $g(.,0,0) \in L^2(\mathbb{J};[0,T+M];\mathbb{R}^k)$.

(ii) *There exists a constant* $L > 0$ *such that* $(\forall (y,\overline{y},z,\overline{z}) \in \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^{k\times m} \times \mathbb{R}^{k\times m})$

$$|g(t,y,z) - g(t,\overline{y},\overline{z})| \leq L\left[|y-\overline{y}| + |z-\overline{z}|\right] \qquad \lambda\otimes Q\text{-a.s.}$$

Then for any given $Y \in L^2(\Omega,\mathcal{J}_{T+M},Q;\mathbb{R}^k)$ *the backward stochastic differential equation*

$$Y(t) = Y - \int_t^{T+M} g\left(s, Y(s), \sqrt{1-\overline{\theta}(s)}X(s)\right)ds - \int_t^{T+M} X(s)d\overline{V}_s \tag{4.4}$$

admits a unique solution in the following class of processes:

$$\{Y(t),X(t)\} \in L^2\Big(\mathbb{G};[0,T+M];\mathbb{R}^k\times\mathbb{R}^{k\times m}\Big)$$

with $\{Y(t)\}$ *continuous.*

Moreover, $\{Y(t), X(t)\} \in L^2(\mathbb{J}; [0, T+M]; \mathbb{R}^k \times \mathbb{R}^{k\times m})$*, and the process* $\{X(t)\}$ *can be written in the following form:*

$$X(t) = \sqrt{1-\overline{\theta}(t)}Z(t), \tag{4.5}$$

where $\{Z(t)\} \in L^2_{\overline{\theta}}(\mathbb{J}; [0, T+M]; \mathbb{R}^{k\times m})$.

Proof. For (y, z) fixed in $\mathbb{R}^k \times \mathbb{R}^{k\times m}$, the process $\{g(t, x, y)\}$ is $\{\mathcal{J}_t\}$-progressively measurable. Since $\mathcal{J}_t = \widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(t)} \otimes \{\emptyset, \widetilde{\Omega}\}$

$$(\forall \widehat{\omega} \in \widehat{\Omega}), \quad (\forall (\widetilde{\omega}_1, \widetilde{\omega}_2) \in \widetilde{\Omega} \times \widetilde{\Omega}), \quad g(\widehat{\omega}, \widetilde{\omega}_1, t, x, y) = g(\widehat{\omega}, \widetilde{\omega}_2, t, x, y).$$

For $\widetilde{\omega}_1$ fixed in $\widetilde{\Omega}$, let us define the function $f : \widehat{\Omega} \times [0, T+M] \times \mathbb{R}^k \times \mathbb{R}^{k\times m} \to \mathbb{R}^k$ such that $f(\widehat{\omega}, t, x, y) = g(\widehat{\omega}, \widetilde{\omega}_1, t, x, y)$. Then f is $\mathcal{M} \otimes \mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^{k\times m})$-measurable since $\mathcal{N}$ is isomorphic to $\mathcal{M} \otimes \{\emptyset, \widetilde{\Omega}\}$, where $\mathcal{M}$ has been defined in Proposition 4.1. Clearly, f so defined satisfies items (i) and (ii) of Proposition 4.1. In the same manner, define Y^* by $Y^*(\widehat{\omega}) \doteq Y(\widehat{\omega}, \widetilde{\omega}_1)$, and so $Y^* \in L^2(\widehat{\Omega}, \widehat{\mathcal{F}}^{\widehat{W}}_{\eta^*(T+M)}, \widehat{P}; \mathbb{R}^k)$.

Therefore, from Proposition 4.1 we can claim that there exists a unique solution to the following equation:

$$\begin{aligned} Y^*(t) = Y^* &- \int_t^{T+M} f(s, Y^*(s), [1-\theta^*(s)]Z^*(s))ds \\ &- \int_t^{T+M} Z^*(s) d\widehat{W}(\eta^*(s)). \end{aligned} \tag{4.6}$$

It is easy to show that

$$\int_t^{T+M} Z^*(s) d\widehat{W}(\eta^*(s)) = \int_t^{T+M} Z^*(s)\sqrt{(1-\theta^*(s))(1-\theta^*(s))^+} d\widehat{W}(\eta^*(s)). \tag{4.7}$$

Now define $Y(\widehat{\omega}, \widetilde{\omega}, t) \doteq Y^*(\widehat{\omega}, t)$, $Z(\widehat{\omega}, \widetilde{\omega}, t) \doteq Z^*(\widehat{\omega}, t)$. Following the same arguments as in the proof of Proposition 3.5 and combining (4.6) and (4.7), we obtain

$$Y(t) = Y - \int_t^{T+M} g(s, Y(s), [1-\overline{\theta}(s)]Z(s))ds - \int_t^{T+M} \sqrt{1-\overline{\theta}(s)}Z(s)d\overline{V}_s. \tag{4.8}$$

Consequently, we have that $(\{Y(t)\}, \{X(t)\})$ is the solution of (4.4), where

$$X(t) \doteq Z(t)\sqrt{1-\overline{\theta}(t)}.$$

Clearly,

$$\{Y(t), X(t)\} \in L^2\Big(\mathbb{J}; [0, T+M]; \mathbb{R}^k \times \mathbb{R}^{k\times m}\Big).$$

Then we have shown the existence of a solution with the desired property (see (4.5)). The uniqueness follows exactly as in [37]. ∎

We will need the following result in the proof of the next theorem.

LEMMA 4.3. *$\forall (t, x, u, z) \in \mathbb{R} \times \mathbb{R}^n \times B_1(K) \times [0, 1]$, $(p^1, p^2, p^3) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}$, and $(q^1, q^2, q^3) \in \mathbb{R}^{1\times m} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{1\times m}$,*

$$\mathcal{H}_t(t, x, u, z, p, q) = \big[A_t(t, x)(1-z) + zB_t(t)u\big]^\top p^2 + \sum_{j=1}^m D_{jt}(t, x)^\top q_j^2 \sqrt{1-z},$$

$$\mathcal{H}_x(t, x, u, z, p, q) = \big[A_x(t, x)(1-z)\big]^\top p^2 + \sum_{j=1}^m D_{jx}(t, x)^\top q_j^2 \sqrt{1-z},$$

where

$$p = \begin{pmatrix} p^1 \\ p^2 \\ p^3 \end{pmatrix}$$

and

$$q = \begin{pmatrix} q^1 \\ q^2 \\ q^3 \end{pmatrix}.$$

Proof. One may easily show that these two equalities follow directly from the definition of $\mathcal{H}$ (see (1.3)). ∎

With the previous results, we may now obtain in the following theorem a crucial property given by (4.13) of the solution of the backward stochastic differential equations (4.9)–(4.12) which renders possible a time change.

THEOREM 4.4. *On the filtered probability space* $\mathbb{G}$*, the system of backward stochastic differential equations*

$$d\overline{p}(t) = -\begin{pmatrix} \mathcal{H}_t(\overline{\eta}(t), \overline{\xi}(t), \overline{\alpha}(t), \overline{\theta}(t), \overline{p}(t), \overline{q}(t)) \\ \mathcal{H}_x(\overline{\eta}(t), \overline{\xi}(t), \overline{\alpha}(t), \overline{\theta}(t), \overline{p}(t), \overline{q}(t)) \\ 0 \end{pmatrix} dt + \overline{q}(t) d\overline{V}(t), \tag{4.9}$$

with

$$\overline{p}(T+M) = -\begin{pmatrix} 0 \\ G_x\big(\overline{\mu}(T+M), \overline{\xi}(T+M)\big) \\ G_w\big(\overline{\mu}(T+M), \overline{\xi}(T+M)\big) \end{pmatrix} \tag{4.10}$$

and

$$\begin{aligned} d\overline{P}(t) = & -A_x(\overline{\eta}(t), \overline{\xi}(t))^\top \overline{P}(t)(1-\overline{\theta}(t))dt - \overline{P}(t) A_x(\overline{\eta}(t), \overline{\xi}(t))(1-\overline{\theta}(t))dt \\ & - \sum_{j=1}^m \big[D_{jx}(\overline{\eta}(t), \overline{\xi}(t))\big]^\top \overline{P}(t) D_{jx}(\overline{\eta}(t), \overline{\xi}(t))(1-\overline{\theta}(t))dt \\ & - \sum_{j=1}^m \Big(\big[D_{jx}(\overline{\eta}(t), \overline{\xi}(t))\big]^\top \overline{Q}^j(t) + \overline{Q}^j(t) D_{jx}(\overline{\eta}(t), \overline{\xi}(t)) \Big) \sqrt{1-\overline{\theta}(t)}dt \\ & - \mathcal{H}_{xx}(\overline{\eta}(t), \overline{\xi}(t), \overline{\alpha}(t), \overline{\theta}(t), \overline{p}(t), \overline{q}(t))dt + \sum_{j=1}^m \overline{Q}^j(t) d\overline{V}_j(t), \end{aligned} \tag{4.11}$$

with

$$\overline{P}(T+M) = -G_{xx}\big(\overline{\mu}(T+M), \overline{\xi}(T+M)\big), \tag{4.12}$$

admits a unique solution in the following class of processes:

$$\Big\{\overline{p}(t), \overline{q}(t), \overline{P}(t), \big(\overline{Q}^j(t)\big)_{j\in\mathbb{N}_m}\Big\} \in L^2\Big(\mathbb{G}; [0, T+M]; \mathbb{R}^{2+n} \times \mathbb{R}^{(2+n)\times m} \times \mathcal{S}^n \times \Big[\mathcal{S}^n\Big]^m\Big)$$

with $\{\overline{p}(t), \overline{P}(t)\}$ *continuous.*

Moreover,

$$\begin{cases} \overline{p}(t,\widehat{\omega},\widetilde{\omega}) = p^*(t,\widehat{\omega}), & \overline{q}(t,\widehat{\omega},\widetilde{\omega}) = \sqrt{1-\theta^*(t,\widehat{\omega})}q^*(t,\widehat{\omega}), \\ \overline{P}(t,\widehat{\omega},\widetilde{\omega}) = P^*(t,\widehat{\omega}), & \overline{Q}^j(t,\widehat{\omega},\widetilde{\omega}) = \sqrt{1-\theta^*(t,\widehat{\omega})}Q^{*j}(t,\widehat{\omega}) \end{cases} \tag{4.13}$$

$\forall(\widehat{\omega},\widetilde{\omega}) \in \widehat{\Omega}\times\widetilde{\Omega}$ *and* $\forall j \in \mathbb{N}_m$*, where* $\{p^*(t), P^*(t), q^*(t), \big(Q^{*j}(t)\big)_{j\in\mathbb{N}_m}\}$ *is the unique solution in the following class of processes:*

$$\{p^*(t), P^*(t)\} \in L^2\Big(\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}; [0,T+M]; \mathbb{R}^{2+n}\times\mathcal{S}^n\Big),$$

$$\{q^*(t), \big(Q^{*j}(t)\big)_{j\in\mathbb{N}_m}\} \in L^2_{\theta^*}\Big(\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}; [0,T+M]; \mathbb{R}^{(2+n)\times m}\times\Big[\mathcal{S}^n\Big]^m\Big)$$

with $\{p^*(t), P^*(t)\}$ *continuous of the system of backward stochastic differential equations defined on the filtered probability space* $\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}$*:*

$$dp^*(t) = -\begin{pmatrix} \mathcal{H}_t(\eta^*(t),\xi^*(t),\alpha^*(t),\theta^*(t),p^*(t),\sqrt{1-\theta^*(t)}q^*(t)) \\ \mathcal{H}_x(\eta^*(t),\xi^*(t),\alpha^*(t),\theta^*(t),p^*(t),\sqrt{1-\theta^*(t)}q^*(t)) \\ 0 \end{pmatrix} dt$$

$$+\, q^*(t)d\widehat{W}(\eta^*(t)) \tag{4.14}$$

with

$$p^*(T+M) = -\begin{pmatrix} 0 \\ G_x\big(\mu^*(T+M),\xi^*(T+M)\big) \\ G_w\big(\mu^*(T+M),\xi^*(T+M)\big) \end{pmatrix} \tag{4.15}$$

and

$$\begin{aligned} dP^*(t) = &-A_x(\eta^*(t),\xi^*(t))^\top P^*(t)(1-\theta^*(t))dt - P^*(t)A_x(\eta^*(t),\xi^*(t))(1-\theta^*(t))dt \\ &-\sum_{j=1}^m [D_{jx}(\eta^*(t),\xi^*(t))]^\top P^*(t)D_{jx}(\eta^*(t),\xi^*(t))(1-\theta^*(t))dt \\ &-\sum_{j=1}^m \left([D_{jx}(\eta^*(t),\xi^*(t))]^\top Q^{*j}(t) + Q^{*j}(t)D_{jx}(\eta^*(t),\xi^*(t))\right)(1-\theta^*(t))dt \\ &-\mathcal{H}_{xx}(\eta^*(t),\xi^*(t),\alpha^*(t),\theta^*(t),p^*(t),\sqrt{1-\theta^*(t)}q^*(t))dt \\ &+\sum_{j=1}^m Q^{*j}(t)d\widehat{W}(\eta^*(t)) \end{aligned} \tag{4.16}$$

with

$$P^*(T+M) = -G_{xx}\big(\mu^*(T+M),\xi^*(T+M)\big). \tag{4.17}$$

Proof. Using Lemma 4.3, (4.9) can be written in the form

$$\overline{p}(t) = -\overline{p}(T+M) - \int_t^{T+M} g(s,\overline{p}(s),\sqrt{1-\overline{\theta}(s)}\overline{q}(s))ds - \int_t^{T+M}\overline{q}(s)d\overline{V}_s,$$

where the $\mathbb{R}^{n+2}$-valued function g is defined by

$$\begin{aligned} &g(t,p,q) \\ &= -\begin{pmatrix} \big[A_t(\overline{\eta}(t),\overline{\xi}(t))(1-\overline{\theta}(t)) + \overline{\theta}(t)B_t(\overline{\eta}(t))\overline{\alpha}(t)\big]^\top p^2 + \sum_{j=1}^m D_{jt}(\overline{\eta}(t),\overline{\xi}(t))^\top q_j^2 \\ \big[A_x(\overline{\eta}(t),\overline{\xi}(t))(1-\overline{\theta}(t))\big]^\top p^2 + \sum_{j=1}^m D_{jx}(\overline{\eta}(t),\overline{\xi}(t))^\top q_j^2 \\ 0 \end{pmatrix}, \end{aligned}$$

for

$$p = \begin{pmatrix} p^1 \\ p^2 \\ p^3 \end{pmatrix}$$

with $p^1 \in \mathbb{R}$, $p^2 \in \mathbb{R}^n$, and $p^3 \in \mathbb{R}$ and for

$$q = \begin{pmatrix} q^1 \\ q^2 \\ q^3 \end{pmatrix},$$

where $q^1 \in \mathbb{R}^{1\times n}$, $q^2 \in \mathbb{R}^{n\times n}$, and $q^3 \in \mathbb{R}^{1\times n}$. Moreover, from (4.10), it follows that $\widehat{p}(T+M) \in L^2(\Omega, \mathcal{J}_{T+M}, Q; \mathbb{R}^k)$. Now, using the hypotheses (A1)–(A2) on the data A, B, and D, it is easy to show that (4.9) satisfies the hypotheses of Corollary 4.2. Therefore, the existence and the uniqueness results for (4.9) are a straightforward consequence of Corollary 4.2. Using similar arguments, the existence and the uniqueness results can be obtained for (4.11). By using Proposition 4.1, it can be shown that (4.14)–(4.17) admit a unique solution. The last part of the theorem (see (4.13)) follows easily from Corollary 4.2 and its proof. □

Now we give the maximum principle for the auxiliary control problem.

THEOREM 4.5. *There exist* $\psi \in S_1(\mathbb{R}^3)$ *such that* $(\forall(\alpha,\theta) \in B_1(K) \times [0,1])$

$$\begin{aligned} &H(\eta^*(t), \xi^*(t), \alpha, \theta, p^{*1}(t), p^{*2}(t), p^{*3}(t), r^*(t), P^*(t), \psi) \\ &\quad \leq H(\eta^*(t), \xi^*(t), \alpha^*(t), \theta^*(t), p^{*1}(t), p^{*2}(t), p^{*3}(t), r^*(t), P^*(t), \psi) \end{aligned} \tag{4.18}$$

$\lambda \otimes \widehat{P}$*-a.s. on* $[0, T+M] \times \widehat{\Omega}$*, where* $\{p^{*1}(t)\}$ *(respectively,* $\{p^{*2}(t)\}$*,* $\{p^{*3}(t)\}$*) is an* $\mathbb{R}$ *(respectively,* $\mathbb{R}^n$*,* $\mathbb{R}$*)-valued process, and* $\{q^{*1}(t)\}$ *(respectively,* $\{q^{*2}(t)\}$*,* $\{q^{*3}(t)\}$*) is an* $\mathbb{R}^{1\times m}$ *(respectively,* $\mathbb{R}^{n\times m}$*,* $\mathbb{R}^{1\times m}$*)-valued process with*

$$\left(\left\{ \begin{pmatrix} p^{*1}(t) \\ p^{*2}(t) \\ p^{*3}(t) \end{pmatrix} \right\}, \left\{ \begin{pmatrix} q^{*1}(t) \\ q^{*2}(t) \\ q^{*3}(t) \end{pmatrix} \right\}, \{P^*(t)\}, \left(\{Q^{*j}\}\right)_{j\in\mathbb{N}_m} \right)$$

solutions of (4.14)–(4.17) *and*

$$r^*(t) \doteq \left[q^{*2}(t) - P^*(t) D(\eta^*(t), \xi^*(t)) \right] \sqrt{1-\theta^*(t)}. \tag{4.19}$$

Proof. From assumptions (A1)–(A2) the hypothesis (3) of Theorem 5 in [38] is satisfied. However, one hypothesis of Theorem 5 in [38] is not satisfied here in the sense that we do not require the control processes $\{\overline{\alpha}(t)\}$, $\{\overline{\theta}(t)\}$ to be $\{\mathcal{F}_t^{\overline{V}}\}$-adapted. As pointed out in [41] (see page 114 and the top of page 116), this hypothesis may appear crucial in order to ensure the existence of the adjoint variables. However, although we required the processes to be $\{\mathcal{J}_t\}$-adapted, the proof of Theorem 5 in [38] remains unchanged since we have shown in Corollary 4.2 the existence and the uniqueness of the adjoint variables with the desired property.

Due to the particular block structure of the matrix $\mathcal{D}$ and from Theorem 5 in [38], it follows that there exist $\psi \in S_1(\mathbb{R}^3)$ such that $(\forall(\alpha,\theta) \in B_1(K) \times [0,1])$ the variational inequality can be written in the following form:

$$\begin{aligned}
&(1-\theta)\overline{p}^{1\psi}(t) + (1-\theta)A(\overline{\eta}(t),\overline{\xi}(t))^\top\overline{p}^{2\psi}(t) + \theta\big[B(\overline{\eta}(t))\alpha\big]^\top\overline{p}^{2\psi}(t) + \theta|u|\overline{p}^{3\psi}(t)\\
&\qquad + \sqrt{1-\theta}\,\mathrm{tr}\big[D(\overline{\eta}(t),\overline{\xi}(t))^\top\overline{r}^{\psi}(t)\big] + (1-\theta)\,\mathrm{tr}\big[D(\overline{\eta}(t),\overline{\xi}(t))^\top\overline{P}^{\psi}(t)D(\overline{\eta}(t),\overline{\xi}(t))\big]\\
&\le (1-\overline{\theta}(t))\overline{p}^{1\psi}(t) + (1-\overline{\theta}(t))A(\overline{\eta}(t),\overline{\xi}(t))^\top\overline{p}^{2\psi}(t) + \overline{\theta}(t)\big[B(\overline{\eta}(t))\overline{\alpha}(t)\big]^\top\overline{p}^{2\psi}(t)\\
&\qquad + \overline{\theta}(t)|\overline{\alpha}(t)|\overline{p}^{3\psi}(t) + \sqrt{1-\overline{\theta}(t)}\,\mathrm{tr}\big[D(\overline{\eta}(t),\overline{\xi}(t))^\top\overline{r}^{\psi}(t)\big]\\
&\qquad + (1-\overline{\theta}(t))\,\mathrm{tr}\big[D(\overline{\eta}(t),\overline{\xi}(t))^\top\overline{P}^{\psi}(t)D(\overline{\eta}(t),\overline{\xi}(t))\big]
\end{aligned} \tag{4.20}$$

$\lambda\otimes P$-a.s. on $[0,T+M]\times\Omega$ with

$$\overline{r}^{\psi}(t) = \overline{q}^{2\psi}(t) - \overline{P}^{\psi}(t)D(\overline{\eta}(t),\overline{\xi}(t))\sqrt{1-\overline{\theta}(t)}, \tag{4.21}$$

and where $\{\overline{p}^{1\psi}(t)\}$ (respectively, $\{\overline{p}^{2\psi}(t)\}$, $\{\overline{p}^{3\psi}(t)\}$) is an $\mathbb{R}$ (respectively, $\mathbb{R}^n$, $\mathbb{R}$)-valued process, and $\{\overline{q}^{1\psi}(t)\}$ (respectively, $\{\overline{q}^{2\psi}(t)\}$, $\{\overline{p}^{3\psi}(t)\}$) is an $\mathbb{R}^{1\times m}$ (respectively, $\mathbb{R}^{n\times m}$, $\mathbb{R}^{1\times m}$)-valued process such that

$$\left(\left\{\begin{pmatrix}\overline{p}^{1\psi}(t)\\ \overline{p}^{2\psi}(t)\\ \overline{p}^{3\psi}(t)\end{pmatrix}\right\},\left\{\begin{pmatrix}\overline{q}^{1\psi}(t)\\ \overline{q}^{2\psi}(t)\\ \overline{q}^{3\psi}(t)\end{pmatrix}\right\}\right)$$

are the first-order adjoint variables solution of the following backward stochastic differential equation (see equation (21) in [38]):

$$d\overline{p}^{\psi}(t) = -\begin{pmatrix}\mathcal{H}_t(\overline{\eta}(t),\overline{\xi}(t),\overline{\alpha}(t),\overline{\theta}(t),\overline{p}^{\psi}(t),\overline{q}^{\psi}(t))\\ \mathcal{H}_x(\overline{\eta}(t),\overline{\xi}(t),\overline{\alpha}(t),\overline{\theta}(t),\overline{p}^{\psi}(t),\overline{q}^{\psi}(t))\\ 0\end{pmatrix}dt + \overline{q}^{\psi}(t)d\overline{V}(t) \tag{4.22}$$

with

$$\overline{p}^{\psi}(T+M) = -\begin{pmatrix}\psi_1 N_{1t}\big(\overline{\eta}(T+M)\big) + \psi_3 N_{2t}\big(\overline{\eta}(T+M)\big)\\ \psi_2 G_x\big(\overline{\mu}(T+M),\overline{\xi}(T+M)\big)\\ \psi_2 G_w\big(\overline{\mu}(T+M),\overline{\xi}(T+M)\big)\end{pmatrix} \tag{4.23}$$

and where the second-order adjoint variable is the solution of the following equation (see equation (22) in [38]):

$$\begin{aligned}
d\overline{P}^{\psi}(t) &= -A_x(\overline{\eta}(t),\overline{\xi}(t))^\top\overline{P}^{\psi}(t)(1-\overline{\theta}(t))dt - \overline{P}^{\psi}(t)A_x(\overline{\eta}(t),\overline{\xi}(t))(1-\overline{\theta}(t))dt\\
&\quad - \sum_{j=1}^{m}\big[D_{jx}(\overline{\eta}(t),\overline{\xi}(t))\big]^\top\overline{P}^{\psi}(t)D_{jx}(\overline{\eta}(t),\overline{\xi}(t))(1-\overline{\theta}(t))dt\\
&\quad - \sum_{j=1}^{m}\Big(\big[D_{jx}(\overline{\eta}(t),\overline{\xi}(t))\big]^\top\overline{Q}^{\psi,j}(t) + \overline{Q}^{\psi,j}(t)D_{jx}(\overline{\eta}(t),\overline{\xi}(t))\Big)\sqrt{1-\overline{\theta}(t)}dt\\
&\quad -\mathcal{H}_{xx}(\overline{\eta}(t),\overline{\xi}(t),\overline{\alpha}(t),\overline{\theta}(t),\overline{p}^{\psi}(t),\overline{q}^{\psi}(t))dt + \sum_{j=1}^{m}\overline{Q}^{\psi,j}(t)d\overline{V}_j(t)
\end{aligned} \tag{4.24}$$

with

$$\overline{P}^{\psi}(T+M) = -\psi_2\, G_{xx}\big(\overline{\mu}(T+M),\overline{\xi}(T+M)\big). \tag{4.25}$$

From Theorem 3.1 in [37], the solutions of (4.22)–(4.25) exist and are unique. Since $E_P\big[N(\overline{\eta}(T+M))\big] = 0_2$, we have $N_{1t}(\overline{\eta}(T+M)) = 1$, and $N_{2t}(\overline{\eta}(T+M)) = 2T$ (see the definition of N). Therefore, the terminal condition (4.23) is given by

$$\overline{p}^{\psi}(T+M) = -\begin{pmatrix} \psi_1 + 2T\psi_3 \\ \psi_2 G_x\big(\overline{\mu}(T+M), \overline{\xi}(T+M)\big) \\ \psi_2 G_w\big(\overline{\mu}(T+M), \overline{\xi}(T+M)\big) \end{pmatrix}. \tag{4.26}$$

Simple calculations show that

$$\left(\left\{\psi_2\overline{p}(t) + \begin{pmatrix} -\psi_1 - 2T\psi_3 \\ 0 \\ 0 \end{pmatrix}\right\}, \{\psi_2\overline{q}(t)\}\right)$$

is the solution of (4.22) with the terminal condition (4.26). By uniqueness, we obtain that

$$\overline{p}^{\psi}(t) = \psi_2\overline{p}(t) + \begin{pmatrix} -\psi_1 - 2T\psi_3 \\ 0 \\ 0 \end{pmatrix}, \qquad \overline{q}^{\psi}(t) = \psi_2\overline{q}(t). \tag{4.27}$$

Moreover, combining (4.27) and the definition of $\mathcal{H}$ (see (1.3)), it is easy to obtain that

$$\mathcal{H}_{xx}(\overline{\eta}(t), \overline{\xi}(t), \overline{\alpha}(t), \overline{\theta}(t), \overline{p}^{\psi}(t), \overline{q}^{\psi}(t)) = \psi_2\mathcal{H}_{xx}(\overline{\eta}(t), \overline{\xi}(t), \overline{\alpha}(t), \overline{\theta}(t), \overline{p}(t), \overline{q}(t)).$$

Consequently, $(\{\psi_2\overline{P}(t)\}, (\{\psi_2\overline{Q}^j(t)\})_{j\in\mathbb{N}_m})$ is the solution of (4.24) with the terminal condition (4.25). By uniqueness, we obtain that

$$\overline{P}^{\psi}(t) = \psi_2\overline{P}(t), \qquad \overline{Q}^{\psi,j}(t) = \psi_2\overline{Q}^j(t). \tag{4.28}$$

Now using Theorem 4.4 (see (4.13)) and (4.27), (4.28), it follows that

$$\overline{p}^{\psi}(t,\widehat{\omega},\widetilde{\omega}) = \psi_2 p^*(t,\widehat{\omega}) + \begin{pmatrix} -\psi_1 - 2T\psi_3 \\ 0 \\ 0 \end{pmatrix}, \qquad \overline{P}^{\psi}(t,\widehat{\omega},\widetilde{\omega}) = \psi_2 P^*(t,\widehat{\omega}), \tag{4.29}$$

$$\overline{q}^{\psi}(t,\widehat{\omega},\widetilde{\omega}) = \psi_2\sqrt{1-\theta^*(t,\widehat{\omega})}\,q^*(t,\widehat{\omega}), \qquad \overline{Q}^{\psi,j}(t,\widehat{\omega},\widetilde{\omega}) = \psi_2\sqrt{1-\theta^*(t,\widehat{\omega})}\,Q^{*j}(t,\widehat{\omega}), \tag{4.30}$$

and from (4.21) it implies that

$$\begin{aligned} \overline{r}^{\psi}(t,\widehat{\omega},\widetilde{\omega}) &= \psi_2\sqrt{1-\theta^*(t,\widehat{\omega})}\Big[q^*(t,\widehat{\omega}) - P^*(t,\widehat{\omega})D(\eta^*(t,\widehat{\omega}), \xi^*(t,\widehat{\omega}))\Big] \\ &= \psi_2 r^*(t,\widehat{\omega}), \end{aligned} \tag{4.31}$$

where we have used the definitions of $\{\overline{\eta}(t)\}$, $\{\overline{\xi}(t)\}$ (see (3.14)).

Now using the definitions of the processes $\{\overline{\alpha}(t)\}$ and $\{\overline{\theta}(t)\}$, the definition of H (see (1.5)), and (4.30), (4.31), it follows that the variational inequality (4.20) can be written on the filtered probability space $\widehat{\mathbb{F}}_{\eta^*}^{\widehat{W}}$ as $(\forall(\alpha,\theta) \in B_1(K)\times[0,1])$

$$\begin{aligned} H(\eta^*(t), \xi^*(t), &\alpha, \theta, p^{*1}(t), p^{*2}(t), p^{*3}(t), r^*(t), P^*(t), \psi) \\ &\le H(\eta^*(t), \xi^*(t), \alpha^*(t), \theta^*(t), p^{*1}(t), p^{*2}(t), p^{*3}(t), r^*(t), P^*(t), \psi) \end{aligned}$$

$\lambda \otimes \widehat{P}$-a.s. on $[0, T+M]\times\widehat{\Omega}$, showing the result. $\square$

5. The singular maximum principle and adjoint variables. In this section we first characterize the adjoint variables for the original control problem. Using the important property given by (4.13), it is possible to obtain the adjoint variables for the original control problem as shown in Theorem 5.2 by using a time change. The interesting feature of these adjoint variables is that the first component is the solution of a singular backward equation (see (5.1)). Then some technical results (Lemma 5.3, Proposition 5.4, Corollary 5.5, Lemma 5.7) are derived. Proposition 5.8 is particularly important to describe the connection between the optimal controls $\widetilde{C}$ and $\widehat{C}$. Finally, the stochastic maximum principle for the original singular control problem in the general case is obtained and presented in Theorem 5.9.

In the following definition, we introduce the backward stochastic differential equations that will be satisfied by the adjoint variables for the original control problem. Note the special form of the second term of the right-hand side of (5.1) that gives a singular part for these adjoint variables.

DEFINITION 5.1. *Let $C \in \mathfrak{C}^a$ be a singular control*

$$C \doteq (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}, \{u(t), v(t)\}, \{W(t)\}, \{x(t)\})$$

such that $\{u(t), v(t)\}$ is $\{\mathcal{F}_t^W\}$-progressively measurable. If the system of backward stochastic differential equations

$$\begin{aligned} p^1(t) = &\int_t^T A_t(s, x(s))^\top p^2(s)ds + \int_{]t,T]} p^2(s)^\top B_t(t)u(s)dv(s) \\ &+ \int_t^T \sum_{j=1}^m D_{jt}(t, x(t))^\top q_j^2(s)dt - \int_t^T q^1(s)dW(s), \end{aligned} \tag{5.1}$$

$$\begin{aligned} p^2(t) = &-G_x\left(\int_0^T |u(s)|dv(s), x(T)\right) + \int_t^T A_x(s, x(s))^\top p^2(s)ds \\ &+ \int_t^T \sum_{j=1}^m D_{jx}(s, x(s))^\top q_j^2(s)dt - \int_t^T q^2(s)dW(s), \end{aligned} \tag{5.2}$$

$$p^3(t) = -G_w\left(\int_0^T |u(s)|dv(s), x(T)\right) - \int_t^T q^3(s)dW(s), \tag{5.3}$$

and

$$\begin{aligned} P(t) = &-G_{xx}\left(\int_0^T |u(s)|dv(s), x(T)\right) + \int_t^T \Big[A_x(s, x(s))^\top P(s) - P(s)A_x(s, x(s))\Big]ds \\ &+ \sum_{j=1}^m \int_t^T D_{jx}(s, x(s))^\top P(s) D_{jx}(s, x(s))ds + \int_t^T J_{xx}(s, x(s), p^2(s), q^2(s))ds \\ &+ \sum_{j=1}^m \int_t^T \big[D_{jx}(s, x(s))^\top Q^j(s) + Q^j(s)D_{jx}(s, x(s))\big]ds - \int_t^T \sum_{j=1}^m Q^j(s)dW(s), \end{aligned}$$

admits a solution in the class of processes

$$\begin{gathered} \Big\{p^1(t), p^2(t), p^3(t)\Big\} \in L^2\Big(\mathbb{F}^W; [0,T]; \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}\Big), \\ \Big\{q^1(t), q^2(t), q^3(t)\Big\} \in L^2\Big(\mathbb{F}^W; [0,T]; \mathbb{R}^{1\times m} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{1\times m}\Big), \\ \Big(\{P(t)\}, \big(\{Q^j(t)\}\big)_{j\in\mathbb{N}_m}\Big) \in L^2\Big(\mathbb{F}^W; [0,T]; \mathcal{S}^n\Big) \times \Big[L^2\Big(\mathbb{F}^W; [0,T]; \mathcal{S}^n\Big)\Big]^m, \end{gathered}$$

with $\{p^1(t), p^2(t), p^3(t), P(t)\}$ corlol, *and where* $\mathbb{F}^W \doteq (\Omega, \mathcal{F}, P, \{\mathcal{F}_t^W\})$, *then* $((\{p^i(t)\}, \{q^i(t)\})_{i\in\mathbb{N}_3}, \{P(t)\}, (\{Q^j(t)\})_{j\in\mathbb{N}_m})$ *are called the adjoint variables associated with the control* C*. The solution is said to be unique if the solution of the previous system is unique in this class of processes.*

By using a time transformation, we show that we can obtain the adjoint variables for the original optimal control from $(\{p^*(t)\}, \{q^*(t)\}, \{P^*(t)\}, (\{Q^{*j}(t)\})_{j\in\mathbb{N}_m})$.

THEOREM 5.2. *Define the following processes:*

$$\hat{p}(t) \doteq p^*(\widehat{\Gamma}(t)), \quad \hat{q}(t) \doteq q^*(\widehat{\Gamma}(t)), \quad \widehat{P}(t) \doteq P^*(\widehat{\Gamma}(t)), \quad \widehat{Q}^j(t) \doteq Q^{*j}(\widehat{\Gamma}(t)), \tag{5.4}$$

for $j \in \mathbb{N}_m$*, where* $(\{p^*(t)\}, \{q^*(t)\}, \{P^*(t)\}, (\{Q^{*j}\})_{j\in\mathbb{N}_m})$ *are solutions of* (4.14)–(4.17)*. Write* $\hat{p}(t)$ *in the form*

$$\hat{p}(t) = \begin{pmatrix} \hat{p}^1(t) \\ \hat{p}^2(t) \\ \hat{p}^3(t) \end{pmatrix},$$

where $\hat{p}^1(t) \in \mathbb{R}$, $\hat{p}^2(t) \in \mathbb{R}^n$*, and* $\hat{p}^3(t) \in \mathbb{R}$ *and similarly*

$$\hat{q}(t) = \begin{pmatrix} \hat{q}^1(t) \\ \hat{q}^2(t) \\ \hat{q}^3(t) \end{pmatrix},$$

where $\hat{q}^1(t) \in \mathbb{R}^{1\times n}$, $\hat{q}^2(t) \in \mathbb{R}^{n\times n}$*, and* $\hat{q}^3(t) \in \mathbb{R}^{1\times n}$.

Then $((\{\hat{p}^i(t)\}, \{\hat{q}^i(t)\})_{i\in\mathbb{N}_3}, \{\widehat{P}(t)\}, (\{\widehat{Q}^j(t)\})_{j\in\mathbb{N}_m})$ *are the unique adjoint variables associated with the control* $\widehat{C}$.

Proof. Combining Lemma 4.3 and (4.14), (3.7), we obtain

$$\begin{aligned} p^{*2}(t) = &- G_x\big(\mu^*(T+M), \xi^*(T+M)\big) + \int_t^{T+M} A_x(\eta^*(s), \xi^*(s))^\top p^{*2}(s) d\eta^*(s) \\ &+ \int_t^{T+M} \sum_{j=1}^m D_{jx}(\eta^*(s), \xi^*(s))^\top q_j^{*2}(s) d\eta^*(s) - \int_t^{T+M} q^{*2}(s) d\widehat{W}(\eta^*(s)). \end{aligned} \tag{5.5}$$

However, from (3.12), it follows that

$$G_x\big(\mu^*(T+M), \xi^*(T+M)\big) = G_x\bigg(\int_{[0,T]} |\widehat{u}(s)| d\widehat{v}(s), \widehat{x}(T)\bigg).$$

Now, using the fact that $\widehat{\Gamma}(T) = T + M$, (5.5) becomes

$$\begin{aligned} p^{*2}(t) = &-G_x\big(\mu^*(T+M), \xi^*(T+M)\big) + \int_t^{\widehat{\Gamma}(T)} A_x(\eta^*(s), \xi^*(s))^\top p^{*2}(s) d\eta^*(s) \\ &+ \int_t^{\widehat{\Gamma}(T)} \sum_{j=1}^m D_{jx}(\eta^*(s), \xi^*(s))^\top q_j^{*2}(s) d\eta^*(s) - \int_t^{\widehat{\Gamma}(T)} q^{*2}(s) d\widehat{W}(\eta^*(s)). \end{aligned}$$

Consequently, the value of the process $\{p^{*2}(t)\}$ at time $\widehat{\Gamma}(t)$ is given by $\{p^{*2}(\widehat{\Gamma}(t))\}$ and satisfies

$$\begin{aligned} p^{*2}(\widehat{\Gamma}(t)) = &- G_x\bigg(\int_{[0,T]} |\widehat{u}(s)| d\widehat{v}(s), \widehat{x}(T)\bigg) + \int_{\widehat{\Gamma}(t)}^{\widehat{\Gamma}(T)} A_x(\eta^*(s), \xi^*(s))^\top p^{*2}(s) d\eta^*(s) \\ &+ \int_{\widehat{\Gamma}(t)}^{\widehat{\Gamma}(T)} \sum_{j=1}^m D_{jx}(\eta^*(s), \xi^*(s))^\top q_j^{*2}(s) d\eta^*(s) - \int_{\widehat{\Gamma}(t)}^{\widehat{\Gamma}(T)} q^{*2}(s) d\widehat{W}(\eta^*(s)). \end{aligned} \tag{5.6}$$

Clearly, $\{\widehat{\Gamma}(t)\}$ is a time change on the filtered probability space $\widehat{\mathbb{F}}^{\widehat{W}}_{\eta^*}$. According to Definition (1.3) in [40, p. 180], $\{\eta^*(t)\}$ is clearly a $\{\widehat{\Gamma}(t)\}$-continuous process since $\{\eta^*(t)\}$ is the right inverse of $\{\widehat{\Gamma}(t)\}$. Recalling $\eta^*(\widehat{\Gamma}(s)) = s$, $\xi^*(\widehat{\Gamma}(s)) = \widehat{x}(s)$ and applying Proposition (1.4) in [40, p. 180] to the second and the third terms of the right-hand side of (5.6), it follows that

$$\begin{aligned} p^{*2}(\widehat{\Gamma}(t)) = &-G_x\left(\int_{[0,T]} |\widehat{u}(s)| d\widehat{v}(s), \widehat{x}(T)\right) + \int_t^T A_x(s,\widehat{x}(s))^\top p^{*2}(\widehat{\Gamma}(s))ds \\ &+ \int_t^T \sum_{j=1}^m D_{jx}(s,\widehat{x}(s))^\top q_j^{*2}(\widehat{\Gamma}(s))ds - \int_{\widehat{\Gamma}(t)}^{\widehat{\Gamma}(T)} q^{*2}(s) d\widehat{W}(\eta^*(s)). \end{aligned}$$

Finally, applying Proposition (1.5) in [40, p. 181] to the last term of the right-hand side of the previous equation gives

$$\begin{aligned} p^{*2}(\widehat{\Gamma}(t)) = &-G_x\left(\int_{[0,T]} |\widehat{u}(s)| d\widehat{v}(s), \widehat{x}(T)\right) + \int_t^T A_x(s,\widehat{x}(s))^\top p^{*2}(\widehat{\Gamma}(s))ds \\ &+ \int_t^T \sum_{j=1}^m D_{jx}(s,\widehat{x}(s))^\top q_j^{*2}(\widehat{\Gamma}(s))ds - \int_t^T q^{*2}(\widehat{\Gamma}(s)) d\widehat{W}(s). \end{aligned} \tag{5.7}$$

Using similar arguments, it is easy to obtain that

$$p^{*3}(\widehat{\Gamma}(t)) = -G_w\left(\int_{[0,T]} |\widehat{u}(s)| d\widehat{v}(s), \widehat{x}(T)\right) - \int_t^T q^3(\widehat{\Gamma}(s)) d\widehat{W}(s), \tag{5.8}$$

$$\begin{aligned} P^*(\widehat{\Gamma}(t)) = &-G_{xx}\left(\int_{[0,T]} |\widehat{u}(s)| d\widehat{v}(s), \widehat{x}(T)\right) + \int_t^T J_{xx}(s,\widehat{x}(s), p^{*2}(\widehat{\Gamma}(s)), q^{*2}(\widehat{\Gamma}(s)))ds \\ &+ \int_t^T \Big[A_x(s,\widehat{x}(s))^\top P^*(\widehat{\Gamma}(s)) - P^*(\widehat{\Gamma}(s))A_x(s,\widehat{x}(s))\Big]ds \\ &+ \sum_{j=1}^m \int_t^T D_{jx}(s,\widehat{x}(s))^\top P^*(\widehat{\Gamma}(s)) D_{jx}(s,\widehat{x}(s))ds \\ &+ \sum_{j=1}^m \int_t^T \big[D_{jx}(s,\widehat{x}(s))^\top Q^{*j}(\widehat{\Gamma}(s)) + Q^{*j}(\widehat{\Gamma}(s)) D_{jx}(s,\widehat{x}(s))\big]ds \\ &- \int_t^T \sum_{j=1}^m Q^{*j}(\widehat{\Gamma}(s)) d\widehat{W}(s). \end{aligned} \tag{5.9}$$

Now for $\{p^{*1}(\widehat{\Gamma}(t))\}$, we have from Lemma 4.3 and (4.14)

$$\begin{aligned} p^{*1}(\widehat{\Gamma}(t)) = &\int_{\widehat{\Gamma}(t)}^{T+M} A_t(\eta^*(s), \xi^*(s))^\top p^{*2}(s) d\eta^*(s) - \int_{\widehat{\Gamma}(t)}^{T+M} q^{*1}(s) d\widehat{W}(\eta^*(s)) \\ &+ \int_{\widehat{\Gamma}(t)}^{T+M} \sum_{j=1}^m D_{jt}(\eta^*(s), \xi^*(s))^\top q_j^{*2}(s) d\eta^*(s) \\ &+ \int_{\widehat{\Gamma}(t)}^{T+M} p^{*2}(s)^\top B_t(\eta^*(s))\alpha^*(s)\theta^*(s)ds. \end{aligned} \tag{5.10}$$

Except for the last term, the same reasoning establishes

$$p^{*1}(\widehat{\Gamma}(t)) = \int_t^T A_t(s,\widehat{x}(s))^\top p^{*2}(\widehat{\Gamma}(s))ds + \int_t^T \sum_{j=1}^m D_{jt}(s,\widehat{x}(s))^\top q_j^{*2}(\widehat{\Gamma}(s))ds$$
$$- \int_t^T q^{*1}(\widehat{\Gamma}(s))d\widehat{W}(s) + \int_{\widehat{\Gamma}(t)}^{\widehat{\Gamma}(T)} p^{*2}(s)^\top B_t(\eta^*(s))\alpha^*(s)\theta^*(s)ds. \tag{5.11}$$

Note from (5.5) that $(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)])$ $p^{*2}(t)$ is constant $\widehat{P}$-a.s. on the set $\{\tau_i \leq T\}$, and so $(\forall t \in [0, T+M])$ $p^{*2}(t) = p^{*2}(\widehat{\Gamma}(\eta^*(t)))$. Consequently, combining the previous equality and (3.6), the last term of (5.11) becomes

$$\int_{\widehat{\Gamma}(t)}^{\widehat{\Gamma}(T)} p^{*2}(s)^\top B_t(\eta^*(s))\alpha^*(s)\theta^*(s)ds = \int_{\widehat{\Gamma}(t)}^{\widehat{\Gamma}(T)} p^{*2}(\widehat{\Gamma}(\eta^*(s)))^\top B_t(\eta^*(s))\widehat{u}(\eta^*(s))d\widetilde{v}(s),$$

where $\widetilde{v}(t) = \int_0^t \theta^*(s)ds$.

Now using item (55.2) in Theorem 55 in [13, p. 132], it follows that

$$\int_{\widehat{\Gamma}(t)}^{\widehat{\Gamma}(T)} p^{*2}(s)^\top B_t(\eta^*(s))\alpha^*(s)\theta^*(s)ds = \int_{]t,T]} p^{*2}(\widehat{\Gamma}(s))^\top B_t(s)\widehat{u}(s)d\widetilde{v}(\widehat{\Gamma}(s)). \tag{5.12}$$

However, $\widehat{v}(t) = \int_0^{\widehat{\Gamma}_t} \theta^*(s)ds = \widetilde{v}(\widehat{\Gamma}(t))$ (see (3.13)). Consequently, combining (5.12) and (5.11) gives

$$p^{*1}(\widehat{\Gamma}(t)) = \int_t^T A_t(s,\widehat{x}(s))^\top p^{*2}(\widehat{\Gamma}(s))ds + \int_t^T \sum_{j=1}^m D_{jt}(s,\widehat{x}(s))^\top q_j^{*2}(\widehat{\Gamma}(s))ds$$
$$- \int_t^T q^{*1}(\widehat{\Gamma}(s))d\widehat{W}(s) + \int_{]t,T]} p^{*2}(\widehat{\Gamma}(s))^\top B_t(s)\widehat{u}(s)d\widehat{v}(s). \tag{5.13}$$

In conclusion, from (5.7), (5.8), (5.9), and (5.13), we obtain that

$$\Big(\big(\{p^{*i}(\widehat{\Gamma}(t))\}, \{q^{*i}(\widehat{\Gamma}(t))\}\big)_{i\in\mathbb{N}_3}, \{P(\widehat{\Gamma}(t))\}, \big(\{Q^{*j}(\widehat{\Gamma}(t))\}\big)_{j\in\mathbb{N}_m}\Big),$$

the adjoint variables associated with the control $\widehat{C}$. By linearity, the uniqueness is easily obtained for $(\{\widehat{P}(t)\}, (\{\widehat{Q}^j(t)\})_{j\in\mathbb{N}_m})$, and $(\{\widehat{p}^2(t)\}, \{\widehat{q}^2(t)\})$ implying the uniqueness for $(\{\widehat{p}^1(t)\}, \{\widehat{q}^1(t)\})$, and $(\{\widehat{p}^3(t)\}, \{\widehat{q}^3(t)\})$, which shows the result. □

We present now some technical results that we will need to obtain the maximum principle.

LEMMA 5.3. *For $\psi \in \mathbb{R}^3$, and $(\alpha, \theta) \in B_1(K) \times [0,1]$, write*

$$L(t,\alpha,\theta) = H(\eta^*(t), \xi^*(t), \alpha, \theta, p^{*1}(t), p^{*2}(t), p^{*3}(t), r^*(t), P^*(t), \psi)$$
$$- H(\eta^*(t), \xi^*(t), \alpha^*(t), \theta^*(t), p^{*1}(t), p^{*2}(t), p^{*3}(t), r^*(t), P^*(t), \psi). \tag{5.14}$$

Then, $\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]$,

$$\begin{aligned} L(t,\alpha,\theta) &= \psi_2\big[B(\tau_i)(\theta\alpha - \widehat{u}(\tau_i))\big]^\top \hat{p}^2(\tau_i) + \psi_2(\theta|\alpha| - |\widehat{u}(\tau_i)|)\hat{p}^3(\tau_i) \\ &\quad + (1-\theta)\Bigg\{-\psi_1 - 2T\psi_3 + \psi_2\Big[\hat{p}^1(\tau_i-) - [B(\tau_i)\widehat{u}(\tau_i)]^\top (t - \widehat{\Gamma}(\tau_i-))\hat{p}^2(\tau_i) \\ &\quad + A\Big(\tau_i, \widehat{x}(\tau_i-) + B(\tau_i)\widehat{u}(\tau_i)(t-\widehat{\Gamma}(\tau_i-))\Big)^\top \hat{p}^2(\tau_i) \\ &\quad + \frac{1}{2}\operatorname{tr}\Big[D\Big(\tau_i, \widehat{x}(\tau_i-) + B(\tau_i)\widehat{u}(\tau_i)(t-\widehat{\Gamma}(\tau_i-))\Big)^\top \widehat{P}(\tau_i) \\ &\quad \times D\Big(\tau_i, \widehat{x}(\tau_i-) + B(\tau_i)\widehat{u}(\tau_i)(t-\widehat{\Gamma}(\tau_i-))\Big)\Big]\Big]\Bigg\} \end{aligned}$$

$\widehat{P}$*-a.s. on* $\{\tau_i \le T\}$.

Proof. Clearly, we have $\widehat{\Gamma}(\tau_i-) < \widehat{\Gamma}(\tau_i)$ $\widehat{P}$-a.s. on $\{\tau_i \le T\}$. Since $\{\eta^*(t)\}$ is the right inverse of $\{\widehat{\Gamma}(t)\}$, it gives that

$$(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]), \quad \eta^*(t) = \tau_i \quad \widehat{P}\text{-a.s. on } \{\tau_i \le T\}, \tag{5.15}$$

and so with (3.6) $(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)])$,

$$\theta^*(t) = \widehat{z}(\tau_i) \quad \text{and} \quad \alpha^*(t) = \widehat{u}(\tau_i) \quad \widehat{P}\text{-a.s. on } \{\tau_i \le T\}. \tag{5.16}$$

Therefore, using (3.7)

$$(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]), \quad \theta^*(t) = 1 \quad \widehat{P}\text{-a.s. on } \{\tau_i \le T\}, \tag{5.17}$$

and by (4.19)

$$(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]), \quad r^*(t) = 0 \quad \widehat{P}\text{-a.s. on } \{\tau_i \le T\}. \tag{5.18}$$

Combining Lemma 3.8 and (5.15) we have $(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)])$

$$\int_{\widehat{\Gamma}(\tau_i-)}^t q(s)dW(\eta^*(s)) = 0 \quad \widehat{P}\text{-a.s. on } \{\tau_i \le T\}. \tag{5.19}$$

Write $p^*(t)$ in the form

$$p^*(t) = \begin{pmatrix} p^{*1}(t) \\ p^{*2}(t) \\ p^{*3}(t) \end{pmatrix},$$

where $p^{*1}(t) \in \mathbb{R}$, $p^{*2}(t) \in \mathbb{R}^n$, and $p^{*3}(t) \in \mathbb{R}$.

For $\theta = 1$, simple calculations give that $\forall (\eta, \xi, \alpha, p, q) \in \mathbb{R} \times \mathbb{R}^n \times B_1(K) \times \mathbb{R}^{2+n} \times \mathbb{R}^{(2+n)\times m}$

$$\mathcal{H}_x(\eta,\xi,\alpha,\theta,p,q) = 0 \quad \text{and} \quad \mathcal{H}_t(\eta,\xi,\alpha,\theta,p,q) = \Big(0\big|\big[B_t(\eta)\alpha\big]^\top\big|0\Big)p.$$

It follows from (4.14), (5.15)–(5.17), and (5.19) that $\widehat{P}$-a.s. surely on $\{\tau_i \le T\}$, $\forall t$ in $[\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]$,

$$p^*(t) = \begin{pmatrix} p^{*1}(\widehat{\Gamma}(\tau_i-)) - \int_{\widehat{\Gamma}(\tau_i-)}^t \big[B_t(\tau_i)\widehat{u}(\tau_i)\big]^\top p^{*2}(s)ds \\ p^{*2}(\widehat{\Gamma}(\tau_i)) \\ p^{*3}(\widehat{\Gamma}(\tau_i)) \end{pmatrix}$$

and so

$$p^*(t) = \begin{pmatrix} \hat{p}^1(\tau_i-) - \left[B_t(\tau_i)\widehat{u}(\tau_i)\right]^\top \hat{p}^2(\tau_i)(t - \widehat{\Gamma}(\tau_i-)) \\ \hat{p}^2(\tau_i) \\ \hat{p}^3(\tau_i) \end{pmatrix}, \tag{5.20}$$

by using the definition of $\hat{p}(t)$.

Using similar arguments, we obtain that $\widehat{P}$-a.s. on $\{\tau_i \leq T\}$

$$(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]), \quad P^*(t) = P^*(\widehat{\Gamma}(\tau_i)) = \widehat{P}(\tau_i) \tag{5.21}$$

and from (3.10) and (3.12)

$$\begin{aligned}(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]), \quad \xi^*(t) &= \xi^*(\widehat{\Gamma}(\tau_i-)) + \int_{\widehat{\Gamma}(\tau_i-)}^{t} \theta^*(s) B(\eta^*(s))\alpha^*(s)ds \\ &= \widehat{x}(\tau_i-) + B(\tau_i)\widehat{u}(\tau_i)\big(t - \widehat{\Gamma}(\tau_i-)\big).\end{aligned} \tag{5.22}$$

Therefore, combining (1.3), (1.5), (5.15), and (5.18), we obtain that $\forall t$ in $[\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)])$,

$$\begin{aligned}&H(\eta^*(t), \xi^*(t), \alpha, \theta, p^{*1}(t), p^{*2}(t), p^{*3}(t), r^*(t), P^*(t), \psi) = (1-\theta)\Bigg\{-\psi_1 - 2T\psi_3 \\ &+ \psi_2\left[p^{*1}(t) + A\big(\tau_i, \xi^*(t)\big)^\top p^{*2}(t) + \frac{1}{2}\operatorname{tr}\Big[D\big(\tau_i, \xi^*(t)\big)^\top P^*(t) D\big(\tau_i, \xi^*(t)\big)\Big]\right]\Bigg\} \\ &+ \psi_2\theta\big[B(\tau_i)\alpha\big]^\top p^{*2}(t) + \psi_2\theta|\alpha| p^{*3}(t)\end{aligned}$$

$\widehat{P}$-a.s. on $\{\tau_i \leq T\}$. Finally, the result follows using the previous equation and (5.14), (5.20), (5.21), and (5.22). □

PROPOSITION 5.4. *Let $\psi \in S_1(\mathbb{R}^3)$ such that the variational inequality* (4.18) *is satisfied. Then, for $(\alpha, \theta) \in B_1(K) \times [0,1]$*

$$(\forall t \in [\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]), \quad L(t, \alpha, \theta) \leq 0 \quad \widehat{P}\text{-a.s. on } \{\tau_i \leq T\}. \tag{5.23}$$

Proof. From Lemma 5.3, it follows that $L(t, \alpha, \theta)$ is continuous on $[\widehat{\Gamma}(\tau_i-), \widehat{\Gamma}(\tau_i)]$ $\widehat{P}$-a.s. on $\{\tau_i \leq T\}$. However, using Fubini's theorem and Theorem 4.5, we obtain that

$$\int_{\widehat{\Gamma}(\tau_i-)}^{\widehat{\Gamma}(\tau_i)} I_{\{(t,\widehat{\omega}) \in [0,T+M]\times\widehat{\Omega} : L(t,\alpha,\theta) > 0\}}(s, \widehat{\omega}) ds = 0$$

$\widehat{P}$-a.s. on $\{\tau_i \leq T\}$, giving the result. □

COROLLARY 5.5. *Let $\psi \in S_1(\mathbb{R}^3)$ such that the variational inequality* (4.18) *is satisfied. Then, for $(\alpha, \theta) \in B_1(K) \times [0,1]$,*

$$L(\widehat{\Gamma}(t), \alpha, \theta) \leq 0 \quad \mathcal{M}_{\widehat{\Gamma}}\text{-a.s. on } [0,T] \times \widehat{\Omega}. \tag{5.24}$$

Proof. Write

$$\begin{aligned}N(\alpha, \theta) &\doteq \{(t, \widehat{\omega}) \in [0, T+M] \times \widehat{\Omega} : L(t, \alpha, \theta) > 0\}, \\ \widetilde{N}(\alpha, \theta) &\doteq \{(t, \widehat{\omega}) \in [0, T] \times \widehat{\Omega} : (\widehat{\Gamma}(t), \omega) \in N(\alpha, \theta)\}.\end{aligned}$$

Then

$$\mathcal{M}_{\widehat{\Gamma}}(\widetilde{N}(\alpha,\theta)) = E_{\widehat{P}}\left[\int_{[0,T]} I_{\widetilde{N}(\alpha,\theta)}(s,\widehat{\omega})d\widehat{\Gamma}(s)\right]$$
$$= E_{\widehat{P}}\left[\int_0^{\widehat{\Gamma}(T)} I_{\widetilde{N}(\alpha,\theta)}(\eta^*(s),\widehat{\omega})ds\right]$$
$$= \lambda \otimes \widehat{P}(N^*(\alpha,\theta)), \tag{5.25}$$

where $N^*(\alpha,\theta) = \{(t,\widehat{\omega}) \in [0,T+M] \times \widehat{\Omega} : (\widehat{\Gamma}(\eta^*(t)),\omega) \in N(\alpha,\theta)\}$.

Now, if $(t,\widehat{\omega}) \in [\overset{\infty}{\underset{i=1}{\cup}} [\![\widehat{\Gamma}(\tau_i-),\widehat{\Gamma}(\tau_i)]\!]]^c$, then $\widehat{\Gamma}(\eta^*(t)) = t$. Consequently,

$$N^*(\alpha,\theta) \cap \left[\overset{\infty}{\underset{i=1}{\cup}} [\![\widehat{\Gamma}(\tau_i-),\widehat{\Gamma}(\tau_i)]\!]\right]^c \subset N(\alpha,\theta) \cap \left[\overset{\infty}{\underset{i=1}{\cup}} [\![\widehat{\Gamma}(\tau_i-),\widehat{\Gamma}(\tau_i)]\!]\right]^c.$$

However, from Theorem 4.5, $\lambda \otimes \widehat{P}(N(\alpha,\theta)) = 0$. Therefore,

$$\lambda \otimes \widehat{P}(N^*(\alpha,\theta)) = \lambda \otimes \widehat{P}\Big(N^*(\alpha,\theta) \cap \overset{\infty}{\underset{i=1}{\cup}} [\![\widehat{\Gamma}(\tau_i-),\widehat{\Gamma}(\tau_i)]\!]\Big). \tag{5.26}$$

Furthermore, writing π for the projection of $[0,T+M] \times \widehat{\Omega}$ to $\widehat{\Omega}$, we have

$$\pi\Big(N^*(\alpha,\theta) \cap [\![\widehat{\Gamma}(\tau_i-),\widehat{\Gamma}(\tau_i)]\!]\Big) \subset \{\tau_i \leq T\} \cap \{\widehat{\omega} \in \widehat{\Omega} : L(\widehat{\Gamma}(\tau_i),\alpha,\theta) > 0\}.$$

From Proposition 5.4, it follows that $\widehat{P}(\{\tau_i \leq T\} \cap \{\widehat{\omega} \in \widehat{\Omega} : L(\widehat{\Gamma}(\tau_i),\alpha,\theta) > 0\}) = 0$, and so $\lambda \otimes \widehat{P}(N^*(\alpha,\theta) \cap \cup_{i=1}^{\infty} [\![\widehat{\Gamma}(\tau_i-),\widehat{\Gamma}(\tau_i)]\!]) = 0$. Using (5.26), we have $\lambda \otimes \widehat{P}(N^*(\alpha,\theta)) = 0$, and from (5.25), it follows that $\mathcal{M}_{\widehat{\Gamma}}(\widetilde{N}(\alpha,\theta)) = 0$, giving the result. ◻

LEMMA 5.6. *Let $(\Omega,\mathcal{F},P)$ be a complete probability space with a right continuous complete filtration $\{\mathcal{F}_t\}$ and let $\{v(t)\}$ be an $\mathbb{R}_+$-valued,* corlol, *$\{\mathcal{F}_t\}$-progressively measurable increasing process. Define*

$$v'(t) = \begin{cases} \overline{\lim}_{\substack{\epsilon\to 0\\ \epsilon>0}} \dfrac{v(t+\epsilon)-v(t-\epsilon)}{2\epsilon} & \text{if the limit exists in } \mathbb{R},\\ 0 & \text{otherwise.}\end{cases}$$

Then $\{v'(t)\}$ is an $\mathbb{R}_+$-valued, $\{\mathcal{F}_t\}$-progressively measurable process.

Proof. Clearly, the $\mathbb{R}_+$-valued process $\{v'(t)\}$ exists. By using the result 14 in [12, page 141], $\{v'(t)\}$ is an $\{\mathcal{F}_t\}$-progressively measurable process. ◻

LEMMA 5.7. *Let $(\Omega,\mathcal{F},P)$ be a complete probability space with a right continuous complete filtration $\{\mathcal{F}_t\}$ and let $\{v(t)\}$ be an $\mathbb{R}_+$-valued,* corlol, *$\{\mathcal{F}_t\}$-progressively measurable increasing process such that $v(T) \leq M$. Write $\Gamma(t) = t + v(t)$ and let $\{z(t)\}$ be an $\{\mathcal{F}_t\}$-progressively measurable process such that $v(t) = \int_{[0,t]} z(s)d\Gamma(s)$. Then $z(t) = 1$ $\mathcal{M}_{v^s}$-a.s. on $[0,T] \times \Omega$, and $z(t) = \frac{v'(t)}{1+v'(t)}$ $\mathcal{M}_{v^{ac}}$-a.s. on $[0,T] \times \Omega$.*

Proof. Clearly, we have $(\forall A \in \mathcal{B}([0,T+M]))$ $dv^{ac}(A) = \int_A z(s)d\Gamma^{ac}(s)$, and so $\int_A v'(s)ds = \int_A z(s)\Gamma'(s)ds$. Now using the definition of $\{\Gamma(t)\}$ the last part of the result follows. Denote by F a set such that $\lambda \otimes P(F) = 0$ and $(\forall A \in \mathcal{B}([0,T]) \otimes \mathcal{F}_T)$ $\mathcal{M}_{\Gamma^s}(A) = \mathcal{M}_{\Gamma}(A \cap F)$. Note that $\mathcal{M}_{v^s} = \mathcal{M}_{\Gamma^s}$. Consequently, $\forall A \in F$,

$$E_P\left[\int_{[0,T]} I_A z(s)d\Gamma(s)\right] = E_P\left[\int_{[0,T]} I_A d\Gamma(s)\right],$$

and so $z(t) = 1$ $\mathcal{M}_{v^s}$-a.s. on $[0,T] \times \Omega$, showing the last result. ◻

The next proposition presents some relations between the optimal singular control $\widetilde{C}$ and $\widehat{C}$. This result is important, as it gives the full generality to our maximum principle.

PROPOSITION 5.8. *Assume the existence of an optimal singular control denoted by*

$$\widetilde{C} \doteq \Big(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}_t\}, \{\widetilde{u}(t), \widetilde{v}(t)\}, \{\widehat{W}(t)\}, \{\widetilde{x}(t)\}\Big)$$

such that $\{\widetilde{u}(t), \widetilde{v}(t)\}$ *is* $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$*-progressively measurable.*

Then the optimal control

$$\widehat{C} \doteq \Big(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}_t\}, \{\widehat{u}(t), \widehat{v}(t)\}, \{\widehat{W}(t)\}, \{\widetilde{x}(t)\}\Big)$$

defined by (3.2) *and* (3.3) *is such that on the filtered probability space* $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}_t\})$ *the processes* $\{\widetilde{x}(t)\}$ *and* $\{\widehat{x}(t)\}$ *are indistinguishable, and*

$$\mathcal{M}_{\widetilde{v}^s} \ll \mathcal{M}_{\widehat{v}^s}, \quad \mathcal{M}_{\widetilde{v}^{ac}} = \mathcal{M}_{\widehat{v}^{ac}}, \tag{5.27}$$

$$\widetilde{u}(t) = \widehat{u}(t), \quad \textit{and} \quad \widetilde{v}(t) = \widehat{v}(t) \quad \mathcal{M}_{\widetilde{v}^{ac}}\textit{-a.s. on } [0,T] \times \widehat{\Omega}. \tag{5.28}$$

The adjoint variables $((\{\widetilde{p}^i(t)\}, \{\widetilde{q}^i(t)\})_{i\in\mathbb{N}_3}, \{\widetilde{P}(t)\}, (\{\widetilde{Q}^j(t)\})_{j\in\mathbb{N}_m})$ *associated with* $\widetilde{C}$ *exist and are unique. Moreover, for* $i \in \mathbb{N}_3$, $j \in \mathbb{N}_m$,

$$\widetilde{p}^i(t) = \hat{p}^i(t), \quad \widetilde{q}^i(t) = \hat{q}^i(t), \quad \widetilde{P}(t) = \widehat{P}(t), \quad \textit{and} \quad \widetilde{Q}^j(t) = \widehat{Q}^j(t), \tag{5.29}$$

where $((\{\hat{p}^i(t)\}, \{\hat{q}^i(t)\})_{i\in\mathbb{N}_3}, \{\widehat{P}(t)\}, (\{\widehat{Q}^j(t)\})_{j\in\mathbb{N}_m})$ *are the adjoint variables associated with* $\widehat{C}$.

Proof. From the definition of $\{\widehat{u}(t)\}$ and $\{\widehat{v}(t)\}$ (see (3.2) and (3.3)), we obtain (5.28) and the fact that $\{\widehat{x}(t)\}$ and $\{\widetilde{x}(t)\}$ are indistinguishable. Moreover $(\forall t \in [0,T])$, $\widehat{v}(t) \geq \widetilde{v}(t)$ and $\widehat{v}^{ac}(t) = \widetilde{v}^{ac}(t)$, giving (5.27). Theorem 5.2 shows the existence of the adjoint variables $((\hat{p}^i(t), \hat{q}^i(t))_{i\in\mathbb{N}_3}, \widehat{P}(t), (\widehat{Q}^j(t))_{j\in\mathbb{N}_m})$ associated with $\widehat{C}$. Now (3.4) gives the existence of the adjoint variables $((\widetilde{p}^i(t), \widetilde{q}^i(t))_{i\in\mathbb{N}_3}, \widetilde{P}(t), (\widetilde{Q}^j(t))_{j\in\mathbb{N}_m})$ associated with $\widetilde{C}$ satisfying (5.29). As in the proof of Theorem 5.2, the uniqueness follows from Theorem 3.1 in [37], showing the result. ∎

We can now establish a necessary condition for any control $\widetilde{C} \in \mathfrak{C}^a$ to be optimal. In fact similarly to the maximum principle for nonsingular control problem (see, for example, [38, 41] and the references therein), we show that if a control $\widetilde{C} = (\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}_t\}, \{\widetilde{u}(t), \widetilde{v}(t)\}, \{\widehat{W}(t)\}, \{\widetilde{x}(t)\})$ is optimal, it maximizes a certain Hamiltonian a.s. with respect to the measure of Doleans-Dade generated by the optimal control $\{\widetilde{v}(t)\}$. This last property, which represents an important difference from the classical (nonsingular) control problem where the optimal control maximizes the Hamiltonian a.s. surely with respect to the product measure of the Lebesgue measure and the underlying probability, is already present in the expression of the maximum principle derived by Cadenillas and Haussmann [8] (see the last term of the inequality (66) on page 227 in [8]). Here we present a maximum principle given in terms of three variational inequalities but not in the integral from as in the work by Cadenillas and Haussmann [8]. The first two variational inequalities result from a time change of the variational inequality of the auxiliary control problem. The first one is given with respect to the measure of Doleans-Dade generated by the absolutely continuous part of $\{\widetilde{v}(t)\}$ (see (5.30)). Note that it does not depend directly on $\{\widetilde{v}(t)\}$ but on the derivative $\frac{d\widetilde{v}}{dt}(t)$. The second one (see (5.32)) is given with respect to the measure of Doleans-Dade generated by the singular part of $\{\widetilde{v}_t\}$. The last inequality can be interpreted as a necessary condition for the size of the jumps of $\{\widetilde{v}_t\}$ (see (5.33)). It is different from the first two ones because it is not obtained from a time change of the variational inequality of the auxiliary control problem (see the proof of the next theorem).

THEOREM 5.9. *Assume the existence of an optimal singular control denoted by*

$$\widetilde{C} \doteq \Big(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{P}, \{\widehat{\mathcal{F}}_t\}, \{\widetilde{u}(t), \widetilde{v}(t)\}, \{\widehat{W}(t)\}, \{\widetilde{x}(t)\}\Big)$$

such that $\{\widetilde{u}(t), \widetilde{v}(t)\}$ *is* $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$*-progressively measurable. Denote by* $\{\widetilde{\tau}_i\}_{i\in\mathbb{N}^*}$ *the sequence of* $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$*-stopping times which exhausts the jumps of* $\{\widetilde{v}(t)\}$*. Then there exist* $\psi \in S_1(\mathbb{R}^3)$ *such that* $\forall (u,z) \in B_1(K)\times[0,1]$

$$\begin{aligned}
&[\widetilde{z}(t) - z]\Big\{-\psi_1 - 2T\psi_3 + \psi_2\Big[\widetilde{p}^1(t) + \frac{1}{2}\operatorname{tr}\big[D(t,\widetilde{x}(t))^\top \widetilde{P}(t) D(t,\widetilde{x}(t))\big] \\
&\qquad + A(t,\widetilde{x}(t))^\top \widetilde{p}^2(t)\Big]\Big\} \\
&\qquad + \psi_2\Big[\big[B(t)[zu - \widetilde{z}(t)\widetilde{u}(t)]\big]^\top \widetilde{p}^2(t) + [z|u| - \widetilde{z}(t)|\widetilde{u}(t)|]\widetilde{p}^3(t) \\
&\qquad + \operatorname{tr}\big[D(t,\widetilde{x}(t))^\top \widetilde{r}(t)\big]\Big(\sqrt{1-z} - \sqrt{1-\widetilde{z}(t)}\Big)\Big] \le 0 \qquad (5.30)
\end{aligned}$$

$\mathcal{M}_{\widetilde{v}^{ac}}$*-a.s. on* $[0,T]\times\widehat{\Omega}$*, where*

$$\widetilde{r}(t) \doteq \Big[\widetilde{q}^2(t) - \widetilde{P}(t)D(t,\widetilde{x}(t))\Big]\sqrt{1-\widetilde{z}(t)}, \qquad \widetilde{z}(t) \doteq \frac{\widetilde{v}'(t)}{1+\widetilde{v}'(t)}, \qquad (5.31)$$

and

$$\begin{aligned}
&\psi_2\Big[[zu - \widetilde{u}(t)]^\top B(t)^\top \widetilde{p}^2(t) + [z|u| - |\widetilde{u}(t)|]\widetilde{p}^3(t)\Big] + (1-z)\Big\{-\psi_1 - 2T\psi_3 \\
&\quad + \psi_2\left[\widetilde{p}^1(t) + \frac{1}{2}\operatorname{tr}\left[D(t,\widetilde{x}(t))^\top \widetilde{P}(t) D(t,\widetilde{x}(t))\right] + A(t,\widetilde{x}(t))^\top \widetilde{p}^2(t)\right]\Big\} \le 0 \qquad (5.32)
\end{aligned}$$

$\mathcal{M}_{\widetilde{v}^s}$*-a.s. on* $[0,T]\times\widehat{\Omega}$*, and* $\forall i \in \mathbb{N}$ *and* $\gamma \in [0,1]$

$$\begin{aligned}
&\psi_2\Big[\big[B(\widetilde{\tau}_i)(zu - \widetilde{u}(\widetilde{\tau}_i))\big]^\top \widetilde{p}^2(\widetilde{\tau}_i) + (z|u| - |\widetilde{u}(\widetilde{\tau}_i)|)\widetilde{p}^3(\widetilde{\tau}_i)\Big] + (1-z)\Big\{-\psi_1 - 2T\psi_3 \\
&\quad + \psi_2\Big[\widetilde{p}^1(\widetilde{\tau}_i-) - \gamma\Big[B_t(\widetilde{\tau}_i)\widetilde{u}(\widetilde{\tau}_i)\Delta\widetilde{v}(\widetilde{\tau}_i)\Big]^\top \widetilde{p}^2(\widetilde{\tau}_i) \\
&\quad + A\Big(\widetilde{\tau}_i, \widetilde{x}(\widetilde{\tau}_i-) + \gamma B(\widetilde{\tau}_i)\widetilde{u}(\widetilde{\tau}_i)\Delta\widetilde{v}(\widetilde{\tau}_i)\Big)^\top \widetilde{p}^2(\widetilde{\tau}_i) \\
&\quad + \frac{1}{2}\operatorname{tr}\Big[D\Big(\widetilde{\tau}_i, \widetilde{x}(\widetilde{\tau}_i-) + \gamma B(\widetilde{\tau}_i)\widetilde{u}(\widetilde{\tau}_i)\Delta\widetilde{v}(\widetilde{\tau}_i)\Big)^\top \widetilde{P}(\widetilde{\tau}_i) \\
&\quad \times D\Big(\widetilde{\tau}_i, \widetilde{x}(\widetilde{\tau}_i-) + \gamma B(\widetilde{\tau}_i)\widetilde{u}(\widetilde{\tau}_i)\Delta\widetilde{v}(\widetilde{\tau}_i)\Big)\Big]\Big]\Big\} \le 0 \qquad (5.33)
\end{aligned}$$

$\widehat{P}$-a.s. on $\{\widetilde{\tau}_i \leq T\}$, where $((\{\widetilde{p}^i(t)\}, \{\widetilde{q}^i(t)\})_{i\in\mathbb{N}_3}, \{\widetilde{P}(t)\}, (\{\widetilde{Q}^j(t)\})_{j\in\mathbb{N}_m})$ are the adjoint variables associated with $\widetilde{C}$.

Proof. From Proposition 3.2, it follows that the control $\widehat{C}$ defined by (3.1), (3.2), and (3.3) is such that $\widehat{v}(T) = M$. Therefore, we can apply the results of this section to $\widehat{C}$ and in particular Corollary 5.5. The stochastic maximum principle can now be expressed in the following form: there exist $\psi \in S_1(\mathbb{R}^3)$ such that $\forall (u,z) \in B_1(K) \times [0,1]$

$$\begin{aligned} &H(\eta^*(\widehat{\Gamma}(t)), \xi^*(\widehat{\Gamma}(t)), \alpha, \theta, p^{*1}(\widehat{\Gamma}(t)), p^{*2}(\widehat{\Gamma}(t)), p^{*3}(\widehat{\Gamma}(t)), r^*(\widehat{\Gamma}(t)), P^*(\widehat{\Gamma}(t)), \psi) \\ &-H(\eta^*(\widehat{\Gamma}(t)), \xi^*(\widehat{\Gamma}(t)), \alpha^*(\widehat{\Gamma}(t)), \theta^*(\widehat{\Gamma}(t)), p^{*1}(\widehat{\Gamma}(t)), p^{*2}(\widehat{\Gamma}(t)), p^{*3}(\widehat{\Gamma}(t)), r^*(\widehat{\Gamma}(t)), \\ &P^*(\widehat{\Gamma}(t)), \psi) \leq 0 \end{aligned} \tag{5.34}$$

$\mathcal{M}_{\widehat{\Gamma}}$-a.s. on $[0,T] \times \widehat{\Omega}$. Since $\{\eta^*(t)\}$ is the right inverse of $\{\widehat{\Gamma}(t)\}$ which is strictly increasing, we have that $\eta^*(\widehat{\Gamma}(t)) = t$. Moreover, combining (3.6), (3.12), (4.19), and the definition of $\{\hat{p}(t)\}$, $\{\hat{q}(t)\}$, $\{\widehat{P}(t)\}$ (see (5.4)), we obtain that $r^*(\widehat{\Gamma}(t)) = \hat{q}^2(t) - \widehat{P}(t)D(t,\widehat{x}(t))\sqrt{1-\widehat{z}(t)}$. Defining $\hat{r}(t) = r^*(\widehat{\Gamma}(t))$ it follows that ($\forall (u,z) \in B_1(K) \times [0,1]$)

$$\begin{aligned} &H(t,\widehat{x}(t),u,z,\hat{p}^1(t),\hat{p}^2(t),\hat{p}^3(t),\hat{r}(t),\widehat{P}(t),\psi) \\ &\qquad - H(t,\widehat{x}(t),\widehat{u}(t),\widehat{z}(t),\hat{p}^1(t),\hat{p}^2(t),\hat{p}^3(t),\hat{r}(t),\widehat{P}(t),\psi) \leq 0 \end{aligned}$$

$\mathcal{M}_{\widehat{\Gamma}}$-a.s. on $[0,T] \times \widehat{\Omega}$.

From the definition of $\{\widehat{\Gamma}(t)\}$, we obtain that $\mathcal{M}_{\hat{v}} \ll \mathcal{M}_{\widehat{\Gamma}}$, and so by Proposition 5.8 it follows that ($\forall (u,z) \in B_1(K) \times [0,1]$)

$$\begin{aligned} &H(t,\widetilde{x}(t),u,z,\widetilde{p}^1(t),\widetilde{p}^2(t),\widetilde{p}^3(t),\hat{r}(t),\widetilde{P}(t),\psi) \\ &\qquad - H(t,\widetilde{x}(t),\widetilde{u}(t),\widehat{z}(t),\widetilde{p}^1(t),\widetilde{p}^2(t),\widetilde{p}^3(t),\hat{r}(t),\widetilde{P}(t),\psi) \leq 0 \end{aligned} \tag{5.35}$$

$\mathcal{M}_{\hat{v}}$-a.s. on $[0,T] \times \widehat{\Omega}$.

From (3.4), it follows that $\widehat{v}'(t) = \widetilde{v}'(t)$ $\mathcal{M}_{\widetilde{v}^{ac}}$-a.s. on $[0,T]\times\widehat{\Omega}$. Combining Lemma 5.7 and (5.27) and (5.28), we have that $\widehat{z}(t) = \frac{\widetilde{v}'(t)}{1+\widetilde{v}'(t)}$ $\mathcal{M}_{\widetilde{v}^{ac}}$-a.s. on $[0,T] \times \widehat{\Omega}$, and according to the definition of $\widetilde{r}(t)$ (see (5.31)), it implies that $\widehat{r}(t) = \widetilde{r}(t)$ $\mathcal{M}_{\widetilde{v}^{ac}}$-a.s. on $[0,T] \times \widehat{\Omega}$. Therefore, we obtain

$$\begin{aligned} &H(t,\widetilde{x}(t),u,z,\widetilde{p}^1(t),\widetilde{p}^2(t),\widetilde{p}^3(t),\widetilde{r}(t),\widetilde{P}(t),\psi) \\ &\qquad - H(t,\widetilde{x}(t),\widetilde{u}(t),\widetilde{z}(t),\widetilde{p}^1(t),\widetilde{p}^2(t),\widetilde{p}^3(t),\widetilde{r}(t),\widetilde{P}(t),\psi) \leq 0 \end{aligned}$$

$\mathcal{M}_{\widetilde{v}^{ac}}$-a.s. on $[0,T] \times \widehat{\Omega}$. With the definition of H (see (1.5)), simple calculations give the first inequality (5.30).

From Lemma 5.7, we have that $\widehat{z}(t) = 1$ and $\hat{r}(t) = 0$ $\mathcal{M}_{\hat{v}^s}$-a.s. on $[0,T] \times \widehat{\Omega}$. Consequently, using (5.27), (5.35) reduces to ($\forall (u,z) \in B_1(K) \times [0,1]$)

$$\begin{aligned} &\psi_2\Big[[zu-\widehat{u}(t)]^\top B(t)^\top \widetilde{p}^2(t) + [z|u| - |\widehat{u}(t)|]\widetilde{p}^3(t)\Big] + (1-z)\Big\{-\psi_1 - 2T\psi_3 \\ &+\psi_2\left[\widetilde{p}^1(t) + \frac{1}{2}\operatorname{tr}\big[D(t,\widetilde{x}(t))^\top \widetilde{P}(t) D(t,\widetilde{x}(t))\big] + A(t,\widetilde{x}(t))^\top \widetilde{p}^2(t)\right]\Big\} \leq 0 \end{aligned} \tag{5.36}$$

$\mathcal{M}_{\widetilde{v}^s}$-a.s. on $[0,T]\times\widehat{\Omega}$. Since the left-hand side of the previous inequality is continuous in (z,u), there exists a set $\mathcal{V}\in\mathcal{B}([0,T])\otimes\widehat{\mathcal{F}}_T^{\widehat{W}}$ such that $\mathcal{M}_{\widetilde{v}^s}(\mathcal{V})=0$, and $\forall(t,\widehat{\omega})\in[0,T]\times\widehat{\Omega}-\mathcal{V}$,

$$\psi_2\Big[[zu-\widehat{u}(t)]^\top B(t)^\top\widetilde{p}^2(t)+[z|u|-|\widehat{u}(t)|]\widetilde{p}^3(t)\Big]+(1-z)\Big\{-\psi_1-2T\psi_3$$

$$+\psi_2\left[\widetilde{p}^1(t)+\frac{1}{2}\operatorname{tr}\left[D(t,\widetilde{x}(t))^\top\widetilde{P}(t)D(t,\widetilde{x}(t))\right]+A(t,\widetilde{x}(t))^\top\widetilde{p}^2(t)\right]\Big\}\le 0 \tag{5.37}$$

$\forall(u,z)\in B_1(K)\times[0,1]$.

From the definition of $\widehat{u}(t)$ (see (3.3)), we can write $\widehat{u}(t)=\widetilde{u}(t)\widetilde{c}(t)$, where $\widetilde{c}(t)=I_{[0,T[}+[\frac{\widetilde{v}(T)-\widetilde{v}(T-)}{M-\widetilde{v}(T-)}I_{[T,+\infty[\times\{\widetilde{v}(T)<M\}}+I_{[T,+\infty[\times\{\widetilde{v}(T)=M\}}]$.

Note that $0\le\widetilde{c}(t)\le 1$.

Consequently, from (5.37) we obtain

$$\psi_2\Big[[zu-\widetilde{c}(t)\widetilde{u}(t)]^\top B(t)^\top\widetilde{p}^2(t)+[z|u|-\widetilde{c}(t)|\widetilde{u}(t)|]\widetilde{p}^3(t)\Big]+(1-z)\Big\{-\psi_1-2T\psi_3$$

$$+\psi_2\left[\widetilde{p}^1(t)+\frac{1}{2}\operatorname{tr}\big[D(t,\widetilde{x}(t))^\top\widetilde{P}(t)D(t,\widetilde{x}(t))\big]+A(t,\widetilde{x}(t))^\top\widetilde{p}^2(t)\right]\Big\}\le 0 \tag{5.38}$$

$\forall(u,z)\in B_1(K)\times[0,1]$. Taking $z=1$ and $u=0$ in the previous inequality, it follows that

$$-\psi_2\widetilde{c}(t)\Big[\widetilde{u}(t)^\top B(t,\widetilde{x}(t))^\top\widetilde{p}^2(t)+|\widetilde{u}(t)|\widetilde{p}^3(t)\Big]\le 0.$$

However, for $(t,\widehat{\omega})\in[0,T]\times\widehat{\Omega}-[T,T]\times\{\Delta\widetilde{v}(T)=0\}$, $\widetilde{c}(t)>0$, and so $\widetilde{c}(t)>0$ $\mathcal{M}_{\widetilde{v}^s}$-a.s. on $[0,T]\times\widehat{\Omega}$ since $\mathcal{M}_{\widetilde{v}^s}([T,T]\times\{\Delta\widetilde{v}(T)=0\})=0$.

Consequently,

$$-\psi_2\Big[\widetilde{u}(t)^\top B(t,\widetilde{x}(t))^\top\widetilde{p}^2(t)+|\widetilde{u}(t)|\widetilde{p}^3(t)\Big]$$

$$\le-\psi_2\widetilde{c}(t)\Big[\widetilde{u}(t)^\top B(t,\widetilde{x}(t))^\top\widetilde{p}^2(t)+|\widetilde{u}(t)|\widetilde{p}^3(t)\Big]. \tag{5.39}$$

Combining (5.38) and (5.39), we obtain the inequality (5.32).

From the definition of $\{\widehat{v}(t)\}$ (see (3.2)), we have that $\{\Delta\widehat{v}(T)\neq 0\}=\{\Delta\widetilde{v}(T)\neq 0\}\cup[\{\Delta\widetilde{v}(T)=0\}\cap\{\widetilde{v}(T)<M\}]$. Clearly, $\{\Delta\widetilde{v}(T)=0\}\cap\{\widetilde{v}(T)<M\}\in\widehat{\mathcal{F}}_T^{\widehat{W}}$, and so from Lemma 5.11 in [19], $T_{\{\Delta\widetilde{v}(T)=0\}\cap\{\widetilde{v}(T)<M\}}$ is an $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$-stopping time. Moreover, $[\![\widetilde{\tau}_i]\!]\cap[\![T_{\{\Delta\widetilde{v}(T)=0\}\cap\{\widetilde{v}(T)<M\}}]\!]=\emptyset$.

The sequence $\{\{\widetilde{\tau}_i\}_{i\in\mathbb{N}^*},T_{\{\Delta\widetilde{v}(T)=0\}\cap\{\widetilde{v}(T)<M\}}\}$ of $\{\widehat{\mathcal{F}}_t^{\widehat{W}}\}$-stopping times exhausts the jumps of $\{\widehat{v}(t)\}$. Denote by $\{\tau_i\}_{i\in\mathbb{N}^*}$ the sequence $\{\{\widetilde{\tau}_i\}_{i\in\mathbb{N}^*},T_{\{\Delta\widetilde{v}(T)=0\}\cap\{\widetilde{v}(T)<M\}}\}$. We can apply Proposition 5.4 to the optimal control $\widehat{C}$. Taking $t=\widehat{\Gamma}(\tau_i-)+\gamma\Delta\widehat{\Gamma}(\tau_i)$

for $\gamma \in [0,1]$, we have for $(u,z) \in B_1(K) \times [0,1]$

$$\psi_2\Big[\big[B(\tau_i)(zu-\widehat{u}(\tau_i))\big]^\top \hat{p}^2(\tau_i) + (z|u| - |\widehat{u}(\tau_i)|)\hat{p}^3(\tau_i)\Big] + (1-z)\Big\{ -\psi_1 - 2T\psi_3 + \psi_2\Big[\hat{p}^1(\tau_i-) - \gamma\,[B_t(\tau_i)\widehat{u}(\tau_i)]^\top \Delta\widehat{\Gamma}(\tau_i)\hat{p}^2(\tau_i) + A\Big(\tau_i, \widehat{x}(\tau_i-) + \gamma B(\tau_i)\widehat{u}(\tau_i)\Delta\widehat{\Gamma}(\tau_i)\Big)^\top \hat{p}^2(\tau_i) + \frac{1}{2}\operatorname{tr}\Big[D\Big(\tau_i, \widehat{x}(\tau_i-) + \gamma B(\tau_i)\widehat{u}(\tau_i)\Delta\widehat{\Gamma}(\tau_i)\Big)^\top \widehat{P}(\tau_i) \times D\Big(\tau_i, \widehat{x}(\tau_i-) + \gamma B(\tau_i)\widehat{u}(\tau_i)\Delta\widehat{\Gamma}(\tau_i)\Big)\Big]\Big]\Big\} \le 0$$

$\widehat{P}$-a.s. on $\{\tau_i \le T\}$. However, from the definition of $\widehat{\Gamma}(t)$ and using (3.2) and (3.3) we have $\widehat{u}(\tau_i)\Delta\widehat{\Gamma}(\tau_i) = \widehat{u}(\tau_i)\Delta\widehat{v}(\tau_i) = \widetilde{u}(\tau_i)\Delta\widetilde{v}(\tau_i)$. Finally, the definition of τ_i and Proposition 5.8 give the last inequality (5.33). □

6. Example. In this section, it is shown how the set of necessary conditions given by the inequalities (5.30), (5.32), and (5.33) can be derived and used for a simple worked-out example. Let us consider the following scalar system:

$$dx(t) = Ax(t) + u(t)dv(t) + dW(t), \tag{6.1}$$

with $A > 0$ and $x(0-) = 1$. The aim is to minimize the cost given by $J[C] = E[x(T)]$, subject to $v(T) \le M$ with $M > 0$. It is easy to show that the control defined by

$$\widetilde{v}(t) = \begin{cases} 0 \text{ for } t < 0, \\ M \text{ for } t \ge 0 \end{cases} \quad \text{and} \quad \widetilde{u}(t) = \begin{cases} -1 \text{ for } t = 0, \\ 0 \text{ for } t > 0 \end{cases} \tag{6.2}$$

is an optimal control.

Consequently, we have

$$\widetilde{p}^1(t) = \widetilde{q}^1(t) = 0, \quad \widetilde{p}^3(t) = \widetilde{q}^3(t) = 0, \quad \widetilde{P}(t) = \widetilde{Q}(t) = 0, \tag{6.3}$$

and

$$d\widetilde{p}^2(t) = -A\widetilde{p}^2(t)dt, \text{ with } \widetilde{p}^2(T) = -1, \text{ and } \widetilde{q}^2(t) = 0. \tag{6.4}$$

Therefore, it follows that $\widetilde{p}^2(t) = -e^{A(T-t)}$. Clearly, we have $\mathcal{M}_{\widetilde{v}^{ac}} = 0$. Since $\mathcal{M}_{\widetilde{v}^s}(.) = I_0(.)$, it is easy to show that (5.32) and (5.33) reduce to

$$\psi_2\Big(zu - \widetilde{u}(0)\Big)\widetilde{p}^2(0) + \Big(1-z\Big)\Big[-\psi_1 - 2T\psi_3 + \psi_2 A\Big\{\widetilde{x}(0-) + \gamma\widetilde{u}(0)\Delta\widetilde{v}(0)\Big\}\widetilde{p}^2(0)\Big] \le 0 \tag{6.5}$$

for $u \in [-1,1]$, $z \in [0,1]$, and $\gamma \in [0,1]$.

Consequently, applying Theorem 5.9 to this example, it follows that there exist $\psi \in S_1(\mathbb{R}^3)$ such that $\forall u \in [-1,1]$, $z \in [0,1]$, and $\gamma \in [0,1]$, the previous inequality is satisfied. The variational inequality (6.5) becomes

$$-\psi_2\Big(zu+1\Big)e^{AT} + \Big(1-z\Big)\Big[-\psi_1 - 2T\psi_3 - \psi_2 A\Big\{1 - \gamma M\Big\}e^{AT}\Big] \le 0.$$

When $M \le 1$, a straightforward calculation shows that for any $\psi_2 \in [0,1]$, there exist $\psi_1 \in [-1,1]$ and $\psi_3 \in [-1,1]$ satisfying $\psi_1 + 2T\psi_3 = \sqrt{1-\psi_2^2}$ such that the variational inequality is satisfied $\forall u \in [-1,1]$, $z \in [0,1]$, and $\gamma \in [0,1]$.

When $M > 1$, by choosing ψ_2 such that $0 < \psi_2 \le 1$, and $A\{M-1\}e^{AT} \le \frac{\sqrt{1-\psi_2^2}}{\psi_2}$, and again $\psi_1 + 2T\psi_3 = \sqrt{1-\psi_2^2}$, it follows that the variational inequality is satisfied $\forall u \in [-1,1]$, $z \in [0,1]$, and $\gamma \in [0,1]$.

7. Conclusion. Our work can be generalized in several directions: a running cost can be added to the definition of $J[C]$ (with no convexity hypothesis) and a classical control process can be added in the dynamic of the state (for example in A, D and into the running cost if it exists) as in [8]. Soft constraints with the same form of the cost G may also be added to the model (see [23, page 855] for constraints of these types in classical control problems).

REFERENCES

[1] L.H.R. Alvarez, *A class of solvable singular stochastic control problems*, Stochastics Stochastics Rep., 67 (1999), pp. 83–122.

[2] L.H.R. Alvarez, *Singular stochastic control, linear diffusions, and optimal stopping: A class of solvable problems*, SIAM J. Control Optim., 39 (2001), pp. 1697–1710.

[3] A. Bensoussan, *Lectures on stochastic control*, in Nonlinear Filtering and Stochastic Control, Lecture Notes in Math. 972, Springer-Verlag, Berlin, 1981, pp. 1–62.

[4] J.-M. Bismut, *An introductory approach to duality in optimal stochastic control*, SIAM Rev., 20 (1978), pp. 62–78.

[5] F. Boetius, *Bounded variation singular stochastic control and associated Dynkin game*, in Mathematical Finance, Trends in Mathematics, M. Kohlmann and S. Tang, eds., Birkhäuser, Basel, 2001, pp. 111–120.

[6] F. Boetius, *Singular Stochastic Control and Its Relations to Dynkin Game and Entry-Exit Problems*, Ph.D. thesis, University of Konstanz, Konstanz, Germany, 2001.

[7] F. Boetius and M. Kohlmann, *Connections between optimal stopping and singular stochastic control*, Stochastic Process. Appl., 77 (1998), pp. 253–281.

[8] A. Cadenillas and U. Haussmann, *The stochastic maximum principle for a singular control problem*, Stochastics Stochastics Rep., 49 (1994), pp. 211–237.

[9] A. Cadenillas and I. Karatzas, *The stochastic maximum principle for linear, convex optimal control with random coefficients*, SIAM J. Control Optim., 33 (1995), pp. 590–624.

[10] P.-L. Chow, J.-L. Menaldi, and M. Robin, *Additive control of stochastic linear systems with finite horizon*, SIAM J. Control Optim., 23 (1995), pp. 858–899.

[11] M.H.A. Davis and M. Zervos, *A pair of explicitly solvable singular stochastic control problems*, Appl. Math. Optim., 38 (1998), pp. 327–352.

[12] C. Dellacherie and P.A. Meyer, *Probabilités et potentiel. Chapitre* I *à* V, Hermann, Paris, 1975.

[13] C. Dellacherie and P.A. Meyer, *Probabilités et potentiel. Théorie des Martingales. Chapitre* V *à* VIII, Hermann, Paris, 1980.

[14] F. Dufour and B. Miller, *Singular stochastic control problems*, SIAM J. Control Optim., 43 (2004), pp. 708–730.

[15] N. El Karoui and S.-J. Huang, *A general result of existence and uniqueness of backward stochastic differential equations*, in Backward Stochastic Differential Equations, Pitman Res. Notes Math. Ser. 364, N. El Karoui and L. Mazliak, eds., Longman, Harlow, UK, 1997, pp. 27–36.

[16] N. El Karoui and I. Karatzas, *Probabilistic aspects of finite-fuel, reflected follower problems*, Acta Appl. Math., 11 (1988), pp. 223–258.

[17] N. El Karoui and I. Karatzas, *A new approach to the Skorohod problem, and its applications*, Stochastics Stochastics Rep., 34 (1991), pp. 57–82.

[18] N. El Karoui, H. Nguyen, and M. Jeanblanc-Picqué, *Compactification methods in the control of degenerate diffusions: Existence of an optimal control*, Stochastics, 20 (1987), pp. 169–219.

[19] R.J. Elliott, *Stochastic Calculus and Applications*, Springer-Verlag, New York, 1982.

[20] R.J. Elliott, *The optimal control of diffusions*, Appl. Math. Optim., 22 (1990), pp. 229–240.

[21] R.J. Elliott and M. Kohlmann, *The second order minimum principle and adjoint process*, Stochastics Stochastics Rep., 46 (1994), pp. 25–39.

[22] W.H. Fleming and H.M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York, 1993.

[23] U.G. Haussmann and J.P. Lepeltier, *On the existence of optimal controls*, SIAM J. Control Optim., 28 (1990), pp. 851–902.

[24] U.G. Haussmann and W. Suo, *Singular optimal stochastic controls* I. *Existence*, SIAM J. Control Optim., 33 (1995), pp. 916–936.

[25] U.G. Haussmann and W. Suo, *Singular optimal stochastic controls* II. *Dynamic programming*, SIAM J. Control Optim., 33 (1995), pp. 937–959.

[26] U.G. Haussmann, *A Stochastic Maximum Principle for Optimal Control of Diffusions*, Pitman Res. Notes Math. Ser. 151, Longman Scientific and Technical, Harlow, UK, John Wiley Sons, New York, 1986.

[27] J. Jacod, *Calcul stochastique et problèmes de martingales*, Lecture Notes in Math. 714, Springer-Verlag, Berlin, 1979.

[28] I. Karatzas, *A class of singular stochastic control problems*, Adv. in Appl. Probab., 15 (1983), pp. 225–254.

[29] I. Karatzas, *Probabilistic aspects of finite-fuel stochastic control*, Proc. Nat. Acad. Sci. U.S.A., 82 (1985), pp. 5579–5581.

[30] I. Karatzas, D. Ocone, H. Wang, and M. Zervos, *Finite-fuel singular control with discretionary stopping*, Stochastics Stochastic Rep., 71 (2000), pp. 1–50.

[31] I. Karatzas and S.E. Shreve, *Connections between optimal stopping and singular stochastic control* I. *Monotone follower problems*, SIAM J. Control Optim., 22 (1984), pp. 856–877.

[32] I. Karatzas and S.E. Shreve, *Connections between optimal stopping and singular stochastic control* II. *Reflected follower problems*, SIAM J. Control Optim., 23 (1985), pp. 433–451.

[33] I. Karatzas and S. Shreve, *Equivalent models for finite-fuel stochastic control*, Stochastics, 18 (1986), pp. 245–276.

[34] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed., Springer-Verlag, New York, 1991.

[35] L. Kruk, *Optimal policies for n-dimensional singular stochastic control problems. Part* II: *The radially symmetric case. Ergodic control*, SIAM J. Control Optim., 39 (2000), pp. 635–659.

[36] H.J. Kushner, *On the stochastic maximum principle: Fixed time of control*, J. Math. Anal. Appl., 11 (1965), pp. 78–92.

[37] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett., 14 (1990), pp. 55–61.

[38] S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim., 28 (1990), pp. 966–979.

[39] P. Protter, *Stochastic Integration and Differential Equations*, Springer-Verlag, New York, 1990.

[40] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Springer-Verlag, New York, 1999.

[41] J. Yong and X. Zhou, *Stochastic Controls. Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.

[42] X.Y. Zhou, *A unified treatement of maximum principle and dynaamic programming in stochastic controls*, Stochastics Stochastics Rep., 36 (1991), pp. 137–161.

[43] H. Zhu, *Generalized solution in singular control: The nondegenerate problem*, Appl. Math. Optim., 25 (1992), pp. 225–245.