DISCONTINUOUS REFLECTION AND A CLASS OF SINGULAR STOCHASTIC CONTROL PROBLEMS FOR DIFFUSIONS*

JIN MA†

Abstract. We study two kinds of Discontinuous Reflecting Problem (DRP for short), defined by Chaleyat-Maurel et al. [3] and Dupuis and Ishii [5] (reduced to the one-dimensional case) and the related Stochastic Differential Equations with Discontinuous Paths and Reflecting Boundary Conditions (SDEDR for short). We compare the properties of the solutions to the two DRP's as well as the two SDEDR's. Some comparison theorems, either for the solutions of different kinds of SDEDR's or for the same kind of SDEDR but with different data, are derived. As an application, we consider a class of finite-fuel singular stochastic control problems for diffusions. With the new approach, the complete class of admissible controls can now be obtained as a direct consequence of our comparison theorems without any extra conditions. The existence of the optimal control (in a wider sense) is derived.

Key words. S.D.E. with discontinuous paths, S.D.E. with reflecting boundary conditions, Skorohod Problem, Singular stochastic control.

1. Introduction. Let $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ be a complete probability space with filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the *usual conditions*, (*i.e.*, $\mathcal{F}_t$ is right-continuous in t and $\mathcal{F}_0$ contains all the P-null sets in $\mathcal{F}$). Assume that an r-dimensional $\mathcal{F}_t$-Brownian motion $\{B_t : t \geq 0\}$ is given on this probability space. Let $a : \mathbf{R}^d \to \mathbf{R}^d$; $\sigma : \mathbf{R}^d \to \mathbf{R}^{d\otimes r}$ be two functions which are assumed to be smooth. Consider the following stochastic differential equation: for $X_t = (X_t^1, ..., X_t^d)$,

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s + \xi_t + K_t, \tag{1.1}$$

where X_0 is an $\mathcal{F}_0$-measurable $\mathbf{R}^d$-valued random variable, independent of the Brownian motion $\{B_t : t \geq 0\}$ and $P(X_0^1 \geq 0) = 1$. The process $\{\xi_t : t \geq 0\}$ is $\mathcal{F}_t$-adapted, $\mathbf{R}^d$-valued, càglàd (*i.e.*, left-continuous with right-limit) and of locally bounded variation paths. Finally, $\{K_t : t \geq 0\}$ is a *local-time-like* process to assure that $X_t^1 \geq 0$. Roughly speaking, $K = (K^1, 0, ..., 0)$ where K^1 is an $\mathcal{F}_t$-adapted process whose paths are nondecreasing, flat off the set $\{X_t^1 = 0\}$, and has a jump whenever X^1 attempts a jump across the origin. The equation (1.1) is called the *Stochastic Differential Equation with Discontinuous Paths and Reflecting Boundary Conditions* and will be denoted by SDEDR (or by SDEDR(X_0, ξ) when X_0 and ξ are concerned) throughout this paper.

The SDEDR of this form is actually motivated by a class of singular stochastic control problems, in which the system is given by (1.1) with dimension one; $X_0 \equiv x \geq 0$ is the initial state; ξ is the "control process"; and K is the "reflecting" process generated automatically so as to prevent X from becoming negative. The objective is to minimize the following cost function:

$$J(\xi; T, x) = E[\int_0^T h(t, X_t)dt + \int_{[0,T)} f(t)d\check{\xi}_t + g(X_T)], \tag{1.2}$$

* This work is a part of the author's Ph.D dissertation at the University of Minnesota.

† School of Mathematics, University of Minnesota, Minneapolis, MN 55455.

where $T > 0$ is some fixed duration. $\{\check{\xi}_t : t \geq 0\}$ is the total variation process of ξ with $\check{\xi}_T \leq y < +\infty$ and h, f, g are some nice functions. (The precise formulation of the control problem and the conditions for h, f, g will be given in §6). The singular stochastic control problems of this kind have been studied by Karatzas and Shreve (1985) [9] and (1986) [10], El Karoui and Karatzas (1988) [8] in the case when $a \equiv 0, \sigma \equiv 1$, namely, the diffusion

$$X_t = x + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s \tag{1.3}$$

is simply a Brownian motion starting from $x \geq 0$. In these works, the connection between the singular stochastic control and the optimal stopping are established, and the existence of the optimal control is proved in different versions.

It was on the way of trying to generalize the above results to a more general diffusion case that we found some interesting problems in *Discontinuous Reflecting Problem* (DRP for short) and SDEDR themselves, which led to this paper. Our main results can be briefly summarized as follows. Firstly, contrary to the main property of the Skorohod problem considered by Dupuis and Ishii [5] (abbreviated as DRP(II) in the sequel), we show by presenting a counterexample (in Appendix 1) that the DRP defined by Chaleyet-Maurel et al. [3] (abbreviated as DRP(I) in the sequel) does not have the Lipschitz continuity under the sup-norm. However, despite this fact, we note that the known weaker "Lipschitz continuity" of DRP(I) (see (3.2)) derived in [3] would be good enough as far as only the existence and pathwise uniqueness of SDEDR are concerned. We discuss this topic in sections 3 and 4. Secondly, in section 5 we establish some basic relationship between the solutions of two kinds of SDEDR's, as well as some comparison theorems. Our main comparison results are Proposition 5.4 and Corolary 5.5, which facilitate the proof of the comparison theorems both for the solutions to the different kinds of SDEDR's and for that to the same kind of SDEDR but with different data. As far as we know, such comparison theorems do not exist in the literature. Thirdly, we apply the above results to the singular stochastic control problems. As a matter of fact, one of our original concerns is to construct a complete class of admissible controls, as was done in [8], [9] and [10], but without the extra technical condition on the jumps of the controls. The removal of such condition was carried out in a spacial case by Baldursson [2] after some clever but laborious work of approximations. However, it turns out that the DRP(II) will exactly do the job. We prove this, as a byproduct of our comparison theorems without much extra work, in section 6. Finally, we prove the existence of the optimal control for such control problem in section 7. In a subsequent paper of the author, the control problem will be further developed.

Since the SDEDR in a half-space as we formulated at the begining of the section is essentially a one-dimensional reflecting problem, and our main interest, either in the control problems or in the comparison theorems, is also limited to the one dimensional case, we hereafter assume that $d = r = 1$ throughout the paper without further specification.

2. Definitions, Notations and the Preliminary Results. Let us denote by D (resp. $\overline{D}$) the space of all real-valued functions defined on $[0, \infty)$ which are left-continuous with right limits (resp. right-continuous with left limits). We will call a function in D *càglàd*, and that in $\overline{D}$ *càdlàg* as usual. The space D is isometric to $\mathbf{R}\times\overline{D}$ in an obvious way. Thus the Skorohod topology on the space $\overline{D}$ (cf. [6, p.117]) would also make the space D a *Polish space*. To avoid unnecessary fussiness in notion, we will still call this topology on D the Skorohod topology. Let $\hat{D}$ be the subspace of D consisting of all the elements in D with paths of locally bounded variation. It can be shown that $\hat{D}$ is a Borel set in D (one is referred to [12] for the details concerning above facts).

We now turn to the probabilistic set-up. In this paper, all the probability spaces will be defined together with the filtration satisfying the *usual conditions*. The probability space is subject to change if necessary, especially in the control problems. However, we always assume that it is rich enough to carry a one-dimensional, $\{\mathcal{F}_t\}$-Brownian motion.

Let $W = C([0,\infty);\mathbf{R})$ with the usual sup-norm, and let W_0 be the collection of those elements $x \in W$ such that $x(0) = 0$. The *canonical space* is defined by

$$\Omega = \mathbf{R}\times W_0 \times D, \tag{2.1}$$

$$\mathcal{F} = \mathcal{B}(\mathbf{R})\times\mathcal{B}(W_0)\times\mathcal{B}(D), \tag{2.2}$$

$$\mathcal{F}_t = \mathcal{B}(\mathbf{R})\times\mathcal{B}_t(W_0)\times\mathcal{B}_t(D), \qquad t \ge 0. \tag{2.3}$$

where, for a function space E, $\mathcal{B}_t(E)$ is the σ-field generated by the elements in E of the form $x(\cdot\wedge t)$. We denote the generic element of Ω by $\omega = (y, w, \zeta)$; define the *canonical process* (X_0, B, ξ) by $(X_0, B_t, \xi_t)(\omega) = (y, w(t), \zeta(t))$, for $t \ge 0, \omega \in \Omega$; and the usual P-augmentation of $\mathcal{F}, \{\mathcal{F}_t\}$ by $\mathcal{F}^P$, $\{\mathcal{F}_t^P\}$, respectively.

Let $\mathcal{M}$ be the collection of all probability measures on the canonical space $(\Omega,\mathcal{F})$ satisfying the following conditions: (1) $P\circ\pi_3^{-1}(\hat{D}) = 1$, where π_3 is the (third) coordinate projection mapping from Ω to D. (2) Under P, the second component of the canonical process B is an $\mathcal{F}_t^P$-Brownian motion. It is not hard to prove that, for any given probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P};\tilde{\mathcal{F}}_t)$, if a triple $(\tilde{X}_0,\ \tilde{B},\ \tilde{\xi})$ is given so that $\tilde{X}_0$ is an $\tilde{\mathcal{F}}_0$-measurable, $\mathbf{R}$-valued random variable; $\{\tilde{B}_t; t \ge 0\}$ is an $\tilde{\mathcal{F}}_t$-Brownian motion; and $\{\tilde{\xi}_t; t \ge 0\}$ is an $\tilde{\mathcal{F}}_t$-adapted process whose paths are locally of bounded variation, then the law of the triple $(\tilde{X}_0,\ \tilde{B},\ \tilde{\xi})$ belongs to $\mathcal{M}$.

Finally, it will be convenient to define the following spaces. Let A be the collection of all functions $\zeta \in \hat{D}$ such that ζ is nondecreasing and $\zeta(0) = 0$. For given $T, y > 0$, denote $A(T) = \{\zeta \in A; \zeta(t) = \zeta(T+), t > T\}$; $A(T,y) = \{\zeta \in A(T); \zeta(T+) \le y\}$; $B(T) = \{\zeta \in \hat{D}; \zeta(t) = \zeta(T+), t > T\}$; $B(T,y) = \{\zeta \in B(T); \zeta(T+) \le y\}$.

Let $(\Omega,\mathcal{F},P;\mathcal{F}_t)$ be any given probability space. We denote the space of measurable, adapted, D-valued paths processes by $\mathcal{D}_P$ and also define, for given $T > 0, y > 0$, $\mathcal{A}_P$ (resp. $\mathcal{A}_P(T), \mathcal{A}_P(T,y), \mathcal{B}_P(T), \mathcal{B}_P(T,y), \hat{\mathcal{D}}_P$) to be the processes $\xi \in \mathcal{D}_P$ such that the sample paths $\xi(\omega) \in A$ (resp. $A(T), A(T,y), B(T), B(T,y), \hat{D}$) for P-almost every $\omega \in \Omega$. To simplify notation, we often drop the subscript "P" for above-mentioned spaces of processes and simply denote them by $\mathcal{A}, \hat{\mathcal{D}}$,..., and so on if the underlying probability space is clear in the context. Moreover, if we denote the total variation process of $\xi \in \hat{\mathcal{D}}$ by $\check{\xi}$, we further assume that, for any $\xi \in \hat{\mathcal{D}}$, there exist $\xi^+, \xi^- \in \mathcal{A}$ such that $\xi = \xi^+ - \xi^-, \check{\xi} = \xi^+ + \xi^-$.

3. Discontinuous Reflecting Problems. We now review the Discontinuous Reflecting Problems (DRP) defined by Chaleyat-Maurel et al. [3] and by Dupuis and Ishii [5]. We shall compare the basic properties of these two DRP's. Note that we are dealing here with deterministic functions.

Consider the space D defined in §2. Denote, for each $Y \in D$, $\Delta Y_t = Y_{t+} - Y_t,\ t \geq 0$ and $S_Y = \{t \geq 0 : |\Delta Y_t| > 0\}$. For a nondecreasing function $K \in D$, we write $K_t^d = \sum_{0 \leq s < t} \Delta K_s$, then $K_t = K_t^c + K_t^d$, where K^c is called the *continuous part* of K.

Let $Y \in D$. We first give the definition of DRP associated with Y defined by Chaleyat-Maurel et al. [3]. We denote it by DRP(I) (or DRP(I; Y) if Y needs to be specified).

Definition 3.1. *Let $Y \in D$, $Y_0 \geq 0$. We call the pair $(X, K) \in D^2$ a solution to DRP(I;Y) if the following are satisfied:*

(i) $X = Y + K$;

(ii) $X_t \geq 0$, $\forall t \geq 0$;

(iii) K *is nondecreasing,* $K_0 = 0$; *and*

(a) $\int_0^\infty X_s dK_s^c = 0$,

(b) $\Delta K_t = 2X_{t+}$, *for all* $t \in S_K$. □

This definition of the DRP has been used to define the singular stochastic control problems by Karatzas and Shreve [9, 10] and El Karoui and Karatzas [8], and will also be used to define our control model. Here we summarize some of the principal results of DRP(I) which will be useful in our paper (one is referred to [3] for details).

Proposition 3.2. *For every $Y \in D$, there exists a unique solution $(X, K) \in D^2$ to the DRP(I;Y), which satisfies the following properties: Denote*

$$A_t = \max[0, \sup_{0 \leq s \leq t} \{-Y_s\}]; \tag{3.1}$$

then

(1) $A_t \leq K_t \leq A_t + A_t^d, \qquad 0 \leq t < \infty$;

(2) $X_{t+} = |X_t + \Delta Y_t| = |Y_{t+} + K_t|, \qquad 0 \leq t < \infty$;

(3) $S_K = \{t \geq 0; X_t + \Delta Y_t < 0\}$;

(4) For any $t \in S_K$, $K_{t+} + K_t = 2A_{t+}$, *and* $\Delta K_t \leq 2\Delta A_t$;

(5) K *is continuous if and only if* $K \equiv A$; □

The existence and the uniqueness of the solution to DRP(I; Y) validates the following definition.

Definition 3.3. *Let $Y \in D$, and let (X, K) solves DRP(I; Y). The mapping $\Gamma^I : D \to D$ defined by $\Gamma^I(Y) = X$ is called the solution mapping of DRP(I; Y).* □

The DRP has been studied recently by Dupuis and Ishii [5] in another version, which will be called DRP(II) (or DRP(II; Y)) in this paper. We note that the original definition in [5] concerns a general region in $\mathbf{R}^n$; in the one-dimensional càglàd case with domain $G = [0, \infty)$, it takes the following form:

Definition 3.4. *Let $Y \in D$, $Y_0 \geq 0$. We call the pair $(X, K) \in D^2$ a solution to DRP(II; Y) if the following are satisfied:*

(i) $X = Y + K$;

(ii) $X_t \geq 0,\ t \geq 0$;

(iii) K is nondecreasing, $K_0 = 0$; and $K_t = \int_{[0,t)} 1_{\{X_{s+}=0\}} dK_s$, $t \geq 0$. □

For DRP(II), there are also some similar results as those in Proposition 3.2 for DRP(I). Here we only verify those which will be useful in our discussion later.

Proposition 3.5. *For every $Y \in D$ there exists a unique solution $(X, K) \in D^2$ to the DRP(II;Y) which satisfies the following properties:*

(1) *For any $t \in S_K$, $X_{t+} = 0$;*

(2) *$S_K = \{t \geq 0; X_t + \Delta Y_t < 0\}$;*

(3) *$\Delta K_t = |X_t + \Delta Y_t|$, for all $t \in S_K$.*

(4) *If K is continuous, then $K \equiv A$, where A is defined by (3.1), and the solutions of DRP(I) and DRP(II) coincide.*

Proof. The existence and uniqueness are proved in [5]; we need only check the assertions (1)–(4).

(1) By Definition 3.4-(iii), for any $t \geq 0$, $\int_0^t X_{s+} dK_s = \int_0^t X_{s+} 1_{\{X_{s+}=0\}} dK_s = 0$. Hence $\Delta K_t > 0$ implies $X_{t+} = 0$.

(2) Since $X_{t+} = X_t + \Delta X_t = X_t + \Delta Y_t + \Delta K_t$ and by (1), $t \in S_K$ implies that $0 < \Delta K_t = -(X_t + \Delta Y_t)$. Conversely, if $X_t + \Delta Y_t < 0$, then $\Delta K_t = X_{t+} - (X_t + \Delta Y_t) \geq -(X_t + \Delta Y_t) > 0$, *i.e.* $t \in S_K$.

(3) Follows immediately from the above discussion.

(4) Since if K is continuous, then it is easily checked that conditions (iii) of both DRP(II) and DRP(I) become the same, the result follows from Proposition 3.2-(5). □

Also we can define a solution mapping, say Γ^{II}, for DRP(II). The two DRP's have some obvious differences, especially when X attempts a jump across the origin; but among others, the Lipschitz continuity of the solution mapping is probably the most essential one. The main result in [5] is that Γ^{II} is Lipschitz continuous under the uniform topology in D (and then under the Skorohod topology). However, our counterexample in Appendix 1 shows that the same Lipschitz continuity need not hold for Γ^I. Therefore, special consideration is necessary when one studies the problems involving DRP(I) since the existing recipe depending on the "strong" Lipschitz continuity of the solution mapping may not work.

Nevertheless, it turns out that the solution mapping of DRP(I) still has a weaker "Lipschitz continuity" (see Proposition 3.6 below). Let us consider a given probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$. If a process $Y \in \mathcal{D}$ is given on this space, we may construct the solution to DRP(I; $Y(\omega)$), say $(X(\omega), K(\omega))$, for each $\omega \in \Omega$. It follows from the results of [3] that the processes X, K are all in $\mathcal{D}$ and are pathwise unique. We shall still call such pair $(X, K) \in \mathcal{D}^2$ the solution to DRP(I; Y) and still denote $X = \Gamma^I(Y)$. The solution mapping Γ^I then enjoys the following property (cf. [3]).

Proposition 3.6. *Suppose that on a probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$, two semimartingales Y, $\hat{Y} \in \mathcal{D}$ are given with the decomposition $Y = Y_0 + M + A$ and $\hat{Y} = \hat{Y}_0 + \hat{M} + \hat{A}$ respectively; assume that $Y_{0+} = \hat{Y}_{0+}$. Let $(\Gamma^I(Y), K^I(Y))$ be the solution* *to DRP(I; Y) and $(\Gamma^I(\hat{Y}), K^I(\hat{Y})$ be that of DRP(I; $\hat{Y}$)), then there exists a constant $C > 0$ such that for any stopping time $\tau \geq 0$, one has*

$$E\left[\sup_{0\leq t\leq\tau}|\Gamma^I(Y)_t - \Gamma^I(\hat{Y})_t|^2\right] + E\left[\sup_{0\leq t\leq\tau}|K^I(Y)_t - K^I(\hat{Y})_t|^2\right] \leq CE\left[[M-\hat{M}, M-\hat{M}]_\tau + \left(\int_{[0,\tau)}|d(A-\hat{A})_t|\right)^2\right]. \tag{3.2}$$

□

In particular, if $\tau \equiv T > 0$ and Y, $\hat{Y} \in \hat{D}$ are all deterministic, then (3.2) can be reduced to

$$\sup_{0\leq t\leq T}|\Gamma^I(Y)_t - \Gamma^I(\hat{Y})_t| + \sup_{0\leq t\leq T}|K^I(Y)_t - K^I(\hat{Y})_t| \leq C\int_0^T |d(Y-\hat{Y})_s|.$$

This "Lipschitz continuity" is obviously weaker than the one satified by Γ^{II}. However, with the counterexample in Appendix 1, we believe that this is the farthest one can go. The relationship between the two problems is further discussed in Proposition 5.2.

4. Stochastic Differential Equations with Discontinuous Paths and Reflecting Boundary Conditions (SDEDR). Let $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ be a given probability space. Consider the following SDEDR described in §1:

$$X_t = X_0 + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s + \xi_t + K_t. \tag{4.1}$$

We assume that the functions a and σ satisfy the following conditions:

There exists a constant $C_1 > 0$ such that

$$|a(x) - a(y)| + |\sigma(x) - \sigma(y)| \leq C_1|x-y|, \qquad \text{for all } x, y \in \mathbf{R}; \tag{4.2}$$

$$\sigma(0) \neq 0. \tag{4.3}$$

Let X_0 be an $\mathcal{F}_0$-measurable real random variable, and ξ be a given element in $\hat{\mathcal{D}}$. We denote the equation (4.1) with respect to DRP(I) (resp. DRP(II)) with given data X_0, ξ by SDEDR(I; X_0, ξ) (resp. SDEDR(II; X_0, ξ)). The precise definitions of *solutions* to SDEDR(I; X_0, ξ) and SDEDR(II;X_0, ξ) are the following.

Definition 4.1. *A pair of processes $(X, K) \in \mathcal{D}^2$ is called a solution to SDEDR(I; X_0, ξ) (4.1), denoted by $(X^{X_0,\xi}(I), K^{X_0,\xi}(I))$, if K is nondecreasing with $K_0 = 0$ and P-almost surely, (X, K) satisfies (4.1) and*

(I-i) $X_t \geq 0$, $t \geq 0$;

(I-ii) $\int_0^\infty X_s dK_s^c = 0$;

(I-iii) $\Delta K_t = 2X_{t+}$, *for all* $t \in S_K$. □

Accordingly, we have

Definition 4.2. *A pair of processes $(X, K) \in \mathcal{D}^2$ is called a solution to SDEDR(II; X_0, ξ) (4.1), denoted by $(X^{X_0,\xi}(II), K^{X_0,\xi}(II))$, if K is nondecreasing with $K_0 = 0$ and P-almost surely, (X, K) satisfies (4.1) and*

(II-i) $X_t \geq 0,\ t \geq 0$;

(II-ii) $\int_0^\infty X_{s+} dK_s = 0$;

(II-iii) $\Delta K_t = |X_t + \Delta\xi_t|$, *for all* $t \in S_K$. □

When the context is clear, we drop the indices X_0 and ξ from the notation.

To derive the existence and (pathwiwe) uniqueness of the SDEDR (4.1), we will apply the method used in [1]. Namely, we first consider the *unrestricted* equation without reflection corresponding to (4.1):

$$Y_t = X_0 + \int_0^t (a(\Gamma(Y)_s)ds + \int_0^t (\sigma(\Gamma(Y)_s)dB_s + \xi_t, \tag{4.4}$$

where Γ can be either Γ^I or Γ^{II}. Since $Y \in \mathcal{D}$ implies $\Gamma(Y) \in \mathcal{D}$, equation (4.4) is well-defined as long as the solution is sought in the space $\mathcal{D}$ with certain moment conditions (*e.g.*, in the space $\mathcal{H}^2$ defined in Appendix 2). If Y is the solution of (4.4), then we let $X = \Gamma(Y)$ so that $X = Y + K$, where K is the process in Definition 3.1. Hence

$$\begin{aligned} X_t - K_t \ = \ Y_t \ &= \ X_0 + \int_0^t a(\Gamma(Y)_s)ds + \int_0^t \sigma(\Gamma(Y)_s)dB_s + \xi_t \\ &= \ X_0 + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s + \xi_t, \end{aligned}$$

and (X, K) solves the equation (4.1).

Since the "Lipschitz continuity" of Γ^{II} facilitates the proof of the existence and pathwise uniqueness of the equation (4.4), the method will certainly work for SDEDR(II). Therefore in the rest of the section, we only discuss SDEDR(I) without specification; the analogues for SDEDR(II) are trivially true. We shall first state an existence and uniqueness theorem (Theorem 4.3); its proof is based on the "Lipschitz" property (3.2) for Γ^I (a summary of the proof is provided in Appendix 2; one is referred to [12] for complete details). Next, we give a theorem (Theorem 4.4) concerning a stronger version of the solution depending only on the distribution of (X_0, B, ξ). The derivation of such a stronger version is quite a routine task (see, for example, [15, V.10] or [12, II.4]); the details are also omited.

We remark that the existence and uniqueness of the SDEDR similar to (4.1) was also studied in [3] under the condition that $X_0 + \xi_t \geq 0,\ t \geq 0,\ a.s.$. By modifying the function space there on which the Fixed Point Theorem is applied, one can show that the theorem is also true in our case. Nevertheless, our recipe seems simpler.

Theorem 4.3. *Under the assumptions (4.2) and (4.3), the SDEDR (4.1) has a pathwise unique solution $(X, K) \in \mathcal{D}^2$. Moreover, if Y is the solution to (4.4), then the pair (X, K) solves the DRP(Y).* □

Theorem 4.4. *Let $(\Omega, \mathcal{F})$, $\{\mathcal{F}_t\}$ be the canonical space defined by (2.1)–(2.3), and $(X_0, B, \xi) : \Omega \to \Omega$ be the canonical process. Let $\mathcal{M}$ be the subspace of probability measures on $(\Omega, \mathcal{F})$ defined in §2. Assume that the conditions (4.2) and (4.3) hold. Then for any $P \in \mathcal{M}$ there exists a pair of functions $(F_P, G_P) : \Omega = \mathbf{R} \times W_0 \times D \to$*

$D \times D$ such that for all $t \geq 0$, $(F_P, G_P)^{-1}(\mathcal{B}_t(D \times D)) \subseteq \mathcal{F}_t^P$, and $(X, K) \triangleq (F_P, G_P)$ solves the SDEDR (4.1) on the probability space $(\Omega, \mathcal{F}^P, P; \mathcal{F}_t^P)$.

Moreover, let $(\tilde{\Omega}, \tilde{F}, \tilde{P}; \tilde{\mathcal{F}}_t)$ be any probability space on which the triple $(\tilde{X}_0, \tilde{B}, \tilde{\xi})$ has the joint distribution P ($\in \mathcal{M}$), then $(\tilde{X}, \tilde{K}) \triangleq (F_P, G_P)(\tilde{X}_0, \tilde{B}, \tilde{\xi})$ solves SDEDR (4.1) on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}; \tilde{\mathcal{F}}_t)$. □

To end this section, we present a moment estimate for the solution (X, K) to the SDEDR (4.1) and the solution Y to the unrestricted equation (4.4).

Theorem 4.5. *Suppose that a, σ satisfy (4.2) and (4.3); and that (X, K), Y are the solutions to the equations (4.1), (4.4) respectively on some probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$. Then for any $T > 0$ and $m \geq 1$, there exists a constant $C_{m,T}$ depending only on m and T, such that, if Z denotes X, K or Y; and $Z_t^* = \sup_{0 \leq s \leq t} |Z_s|$, then*

$$E(Z_T^*)^{2m} \leq C_{m,T}(1 + E|X_0|^{2m} + E\check{\xi}_T^{2m}). \tag{4.5}$$

Proof. To simplify notation, we henceforth denote any constant depending only on m and T by a generic one, $C_{m,T}$, which may vary line by line. Furthermore, since either $E|X_0|^{2m} = \infty$ or $E\check{\xi}_T^{2m} = \infty$ makes the theorem trivial, we will assume that $E|X_0|^{2m} < \infty$, and $E\check{\xi}_T^{2m} < \infty$.

Let (X, K) be the solution to SDEDR (4.1) and Y be that of (4.4). We first prove (4.5) for $Z = X$. By (1) of Proposition 3.2, we have $K_t \leq A_t + A_t^d \leq 2A_t$ and therefore

$$X_t^* \leq 3Y_t^*, \qquad \text{for } t \in [0, T], \tag{4.6}$$

(see also [3, Proposition 12]). Hence, recall (4.4), we have that

$$\begin{aligned} E(X_t^*)^{2m} &\leq 3^{2m} E(Y_t^*)^{2m} \\ &\leq C_{m,T} \left\{ E|X_0|^{2m} + E\int_0^t |a(X_s)|^{2m} ds + E\left[\sup_{0 \leq u \leq t} \left| \int_0^u \sigma(X_s) dB_s \right| \right]^{2m} + E\check{\xi}_T^{2m} \right\}. \end{aligned} \tag{4.7}$$

Applying the Burkholder-Davis-Gundy inequality (cf. [15, IV.42]) to the third term in {...} on the right hand side above, and using the conditions (4.2) and (4.3), we have from (4.7) that

$$E(X_t^*)^{2m} \leq C_{m,T}\left[1 + E|X_0|^{2m} + E\check{\xi}_T^{2m} + \int_0^t E(X_s^*)^{2m} ds\right].$$

Thus, Gronwall's inequality leads to that

$$E(X_T^*)^{2m} \leq C_{m,T}[1 + E|X_0|^{2m} + E\check{\xi}_T^{2m}]. \tag{4.8}$$

Namely, (4.5) is proved for $Z = X$. Now by (4.6) and (4.7),

$$E(Y_T^*)^{2m} \leq C_{m,T}\left[1 + E|X_0|^{2m} + E\check{\xi}_T^{2m} + \int_0^T E|(Y_s^*)|^{2m} ds\right].$$

So by the Gronwall inequality again, (4.5) is also true for $Z = Y$. Finally, since $K_t = X_t - Y_t$, we see that (4.5) is true for $Z = K$ as well. □

5. Relations Between Two SDEDR's and Comparison Theorems. We are now ready for the main results of this paper. We shall give in this section the explicit formulae that relate the solutions of two SDEDR's to one another, and some comparison results as well. We assume that the coefficients a and σ satisfy the conditions (4.2) and (4.3), and that the probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ is given. For the sake of simplicity, we henceforth assume that the initial state $X_0 \equiv x \in \mathbf{R}$.

To begin with, we give a proposition analogous to Proposition 5.3 in [8]. Since the proof is quite similar to that proposition, we omit it (see [12] for detail).

Proposition 5.1. *Suppose $\theta, L \in \mathcal{A}$ and Y_t is the solution to the SDE:*

$$Y_t = x + \int_0^t a(Y_s)ds + \int_0^t \sigma(Y_s)dB_s - \theta_t + L_t. \tag{5.1}$$

Suppose $Y_t \geq 0$ for all $t \geq 0$, a.s.P. Then there exist ζ, $K \in \mathcal{A}$ such that $L = \zeta + K$ and

(i) $\zeta_t = L_t^d + \int_0^t 1_{\{Y_s > 0\}} dL_s^c$; $K_t = \int_0^t 1_{\{Y_s=0\}} d\theta_s^c + \frac{1}{2}\Lambda_t^0(Y)$, where $\Lambda_\cdot^0(Y)$ is the local time of Y at zero;

(ii) (Y, K) solves the SDEDR (I; x, $\zeta - \theta$).

Remark 5.1 Recall that Y is càglàd. By "local time of Y" we will always mean the local time of $\overline{Y}$, the càdlàg modification of Y. □

5.1. The basic relationship. Recall that, for $x \geq 0$, $\xi \in \hat{\mathcal{D}}$, the solutions to SDEDR(I; x, ξ) and SDEDR(II; x, ξ) are denoted by $(X^{x,\xi}(I), K^{x,\xi}(I))$, $(X^{x,\xi}(II), K^{x,\xi}(II))$ respectively. We have the following proposition:

Proposition 5.2. *Let $x \geq 0$ and $\xi \in \hat{\mathcal{D}}$ be given.*

(1) *There exists an $\eta \in \mathcal{A}$ such that $X_t^{x,\xi}(I) = X_t^{x,\xi+\eta}(II)$, $t \geq 0$, a.s.P, and such that $K^{x,\xi+\eta}(II)$ is continuous. More precisely, we have*

$$\eta_t = (K^{x,\xi}(I))_t^d, \qquad t \geq 0, \ a.s.P. \tag{5.2}$$

(2) *There exists an $\eta \in \hat{\mathcal{D}}$ such that $X_t^{x,\xi}(II) = X_t^{x,\xi+\eta}(I)$, $t \geq 0$, a.s.P, and such that $K^{x,\xi+\eta}(I)$ is continuous. More precisely, if we denote $L_t = \xi_t^+ + K_t^{x,\xi}(II)$, then*

$$\eta_t = K_t^{x,\xi}(II) - \int_0^t 1_{\{X_s^{x,\xi}(II)=0\}} dL_s^c. \tag{5.3}$$

Proof. (i) Denote $X = X^{x,\xi}(I)$; $K = K^{x,\xi}(I)$; we have

$$\begin{aligned} X_t &= x + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s + \xi_t + K_t \\ &= x + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s + \xi_t + K_t^d + K_t^c. \end{aligned}$$

Set $\eta = K^d \in \mathcal{A}$. We claim that $K_t^{x,\xi+\eta}(II) = K_t^c$, $t \geq 0$, a.s.P. Indeed, by the definition of the solution to SDEDR(I; x, ξ), we have $\int_0^\infty X_s dK_s^c = 0$. It is then easily checked that, for each $t \geq 0$,

$$K_t^c = \int_0^t 1_{\{X_s=0\}} dK_s^c. \tag{5.4}$$

Since almost surely, the set $\{s : X_s = 0\}$ differs from the set $\{s : X_{s+} = 0\}$ only by countably many points, and the measure dK^c does not charge any countable set, (5.4) becomes $K_t^c = \int_0^t 1_{\{X_{s+}=0\}} dK_s^c,\ t \ge 0,\ a.s.P$. By Definition 4.1, $K^c = K^{x,\xi+\eta}(II)$ and $X = X^{x,\xi+\eta}(II)$, this proves (1).

(ii) We now denote $X = X^{x,\xi}(II)$; $K = K^{x,\xi}(II)$; $\theta_t = \xi_t^-$; $L_t = \xi_t^+ + K_t$, and apply Proposition 5.1. It is readily seen that if we set $\zeta_t = L_t^d + \int_0^t 1_{\{X_s>0\}} dL_s^c$; $\tilde{K}_t = \int_0^t 1_{\{X_s=0\}} d(\xi^-)_s^c + \frac{1}{2}\Lambda_t^0(X)$, then $X = X^{x,\zeta-\xi^-}(I)$; $\tilde{K}_t = K^{x,\zeta-\xi^-}(I)$. A little computation shows that

$$\zeta_t = L_t - \int_0^t 1_{\{X_s=0\}} dL_s^c = \xi_t^+ + K_t - \int_0^t 1_{\{X_s=0\}} dL_s^c,$$

so that $\zeta_t - \xi_t^- = \xi_t + K_t - \int_0^t 1_{\{X_s=0\}} dL_s^c$. Since $\tilde{K}$ is continuous and nondecreasing, setting $\eta_t = K_t - \int_0^t 1_{\{X_s=0\}} dL_s^c$, we proved (2). □

5.2. The comparison theorems. Consider the stochastic differential equations:

$$X_t^i = x_0^i + \int_0^t a^i(X_s^i)ds + \int_0^t \sigma(X_s^i)dB_s + \xi_t^i + K_t^i, \qquad i = 1, 2, \tag{5.5}$$

where $\xi^i \in \hat{\mathcal{D}}, K^i \in \mathcal{A},\ i = 1, 2$ are the same as those in the previous sections. Trying to get the minimal conditions for the comparison theorem, we will start from a necessary and sufficient condition. All of our comparison theorems are actually its corollaries.

First of all, let us give a lemma which is a modification of Le Gall's observation (cf. [11, Lemma 1.0], [15, V.39]). The proof is standard and easy, we omit it.

Lemma 5.3. *Let Y be a (càglàd) semimartingale satisfying $\sum_{0<s\le t} |\Delta Y_s| < \infty$ a.s., for each $t > 0$, and $\Lambda_\cdot^0(Y)$ be its local time at zero. Suppose there exists a function $\rho : \mathbf{R}^+ \to \mathbf{R}^+$ which is increasing, such that for every $\delta > 0$, $\int_0^\delta \rho(u)^{-1}du = \infty$.*

Then $\Lambda_\cdot^0(Y) \equiv 0$ provided $\int_0^t \rho(Y_s)^{-1} 1_{\{Y_s>0\}} d[Y]_s^c < \infty,\ a.s.P.$ □

Now let $x_0^1 \ge x_0^2 \ge 0$ and $\eta^1, \eta^2 \in \hat{\mathcal{D}}$ be given, and let $X^i,\ i = 1, 2$ be the solutions to the SDE's:

$$X_t^i = x_0^i + \int_0^t a(X_s^i)ds + \int_0^t \sigma(X_s^i)dB_s + \eta_t^i, \qquad i = 1, 2. \tag{5.6}$$

To simplify notation, we use the following convention in the sequel: for any function $\nu \in \hat{D}$, we denote $\tilde{\nu}$ to be the signed measure generated by ν, namely, for any Borel set $E \subseteq [0, \infty)$, $\tilde{\nu}(E) = \int_E d\nu_t$. If H is a Borel set in $[0, \infty)$ and $\nu, \mu \in \hat{D}$, then $\tilde{\mu} \le \tilde{\nu}$ on H means that for any Borel set $A \subseteq H$, $\tilde{\mu}(A) \le \tilde{\nu}(A)$, or equivalently, $\int_0^t 1_H d\mu_s \le \int_0^t 1_H d\nu_s$, for any $t \ge 0$. We will denote this by $1_H d\mu_t \le 1_H d\nu_t,\ t \ge 0$.

Proposition 5.4. *Suppose X^1 and X^2 are defined by (5.6) with $x_0^1 \ge x_0^2 \ge 0$. Then $P\{X_t^1 \ge X_t^2,\ t \ge 0\} = 1$ if and only if the following conditions hold: for a.e. $\omega \in \Omega$,*

(a) On the set $\mathcal{H}^1(\omega) \triangleq \{t : X_t^2(\omega) > X_t^1(\omega)\}$, one has $d\eta_t \le 0$, where $\eta_t = \eta_t^2 - \eta_t^1$. Namely, $1_{\mathcal{H}^1(\omega)} d\eta_t(\omega) \le 0,\quad t \ge 0$.

(b) On the set $\mathcal{H}^2(\omega) \triangleq \{t : X_t^1(\omega) \ge X_t^2(\omega)\}$, one has $X_{t+}^1(\omega) \ge X_{t+}^2(\omega)$.

Proof. The necessity of the proposition is obvious since if $X_t^1 \ge X_t^2$ for all $t \ge 0,\ a.s.P$, then $\mathcal{H}^1 = \emptyset,\ a.s.P$, and (b) always holds. So we need only prove the sufficiency.

Let $Y = X^2 - X^1$ and set $\rho(x) = |x|^2$ to be the function in Lemma 5.3. It is easily checked, by using the condition (4.2) for the function σ, that $\int_0^t \rho(Y_s)^{-1} 1_{\{Y_s>0\}} d[Y]_s^c \leq C^2 t < \infty,\ t \geq 0,\ a.s.P$. Further, since $\Delta Y_t = \Delta \eta_t$ and η is of locally bounded variation, we have $\sum_{0\leq s<t} \Delta|Y_t| < \infty$ for any $t \geq 0$. Thus Y satisfies all the conditions in Lemma 5.3 and hence $\Lambda^0(Y) \equiv 0$.

Denote $H_t = a(X_t^2) - a(X_t^1);\ G_t = \sigma(X_t^2) - \sigma(X_t^1)$. By Tanaka's Formula (see for example, [3, §2.1]; or [2, Appendix] for the càglàd version), we have

$$
\begin{aligned}
(5.7) \qquad Y_t^+ - Y_0^+ &= \int_{[0,t)} 1_{\{Y_s>0\}} dY_s + \frac{1}{2}\Lambda_t^0(Y) \\
&\quad + \sum_{0\leq s<t} [1_{\{Y_sY_{s+}<0\}}|Y_{s+}| + 1_{\{Y_s=0\}}(Y_{s+})^+] \\
&= \int_0^t 1_{\{Y_s>0\}} H_s ds + \int_0^t 1_{\{Y_s>0\}} G_s dB_s + \int_{[0,t)} 1_{\{Y_s>0\}} d\eta_s \\
&\quad + \sum_{0\leq s<t} 1_{\{Y_sY_{s+}<0\}}|Y_{s+}| + \sum_{0\leq s<t} 1_{\{Y_s=0\}}(Y_{s+})^+.
\end{aligned}
$$

For each $t \geq 0,\ \omega \in \Omega$, define the random time

$$
\tau_t(\omega) = \begin{cases} \sup\{s \in [0,t] : Y_s^+(\omega) = 0\}, & \text{if } Y_t^+(\omega) > 0; \\ t, & \text{if } Y_t^+(\omega) = 0. \end{cases}
$$

By the left continuity of the paths of Y^+, we have $\tau_t(\omega) < t$ if $Y_t^+(\omega) > 0$, and $Y^+_{\tau_t(\omega)}(\omega) = 0$ for all $(t,\omega) \in [0,\infty) \times \Omega$. Moreover, for any $t \geq s \geq 0$, we denote $I(t,s) = I(t) - I(s)$, where $I(t) = \int_0^t 1_{\{Y_s>0\}} G_s dB_s$, a continuous local martingale. Then we can rewrite (5.7) pathwise as

$$
\begin{aligned}
(5.8) \qquad Y_t^+ &= \int_{\tau_t}^t 1_{\{Y_s>0\}} H_s ds + I(t,\tau_t) + \int_{[\tau_t,t)} 1_{\{Y_s>0\}} d\eta_s \\
&\quad + \sum_{\tau_t\leq s<t} [1_{\{Y_sY_{s+}<0\}}|Y_{s+}| + 1_{\{Y_s=0\}} Y_{s+}^+] \\
&= I_1 + I_2 + I_3,
\end{aligned}
$$

where $I_1 = \int_{\tau_t}^t 1_{\{Y_s>0\}} H_s ds + I(t,\tau_t);\ I_2 = \int_{[\tau_t,t)} 1_{\{Y_s>0\}} d\eta_s;\ I_3 = \sum_{\tau_t\leq s<t}[1_{\{Y_sY_{s+}<0\}}|Y_{s+}| + 1_{\{Y_s=0\}}(Y_{s+})^+]$.

Clearly, condition (a) implies that $I_2 \leq 0, a.s.P$, since $\{s : Y_s > 0\} = \mathcal{H}^1$. Furthermore, the condition (b) gives that $Y_{s+}^+ = 0$ whenever $Y_s^+ = 0$. By definition of τ_t, $1_{\{Y_sY_{s+}<0\}} = 0$ for all $s \in (\tau_t, t)$ and at τ_t, we have $Y_{\tau_t}^+ = 0$ so that $Y_{\tau_t+}^+ = 0$, which implies that $Y_{\tau_t} Y_{\tau_t+} \geq 0$, so that $1_{\{Y_sY_{s+}<0\}} \equiv 0$ for $s \in [\tau_t, t)$, and thus $I_3 = 0,\ a.s.$

Thus, for any given $T > 0$ and $0 \leq t \leq T$, (5.8) gives

$$
\begin{aligned}
(5.9) \qquad & E[Y_t^+]^2 \leq E[I_1]^2 \leq 2\left\{E\left[\int_{\tau_t}^t 1_{\{Y_s>0\}} H_s ds\right]^2 + E\left[I(t,\tau_t)\right]^2\right\} \\
&\leq C_T \left\{E\left[\int_0^t 1_{\{Y_s>0\}} |X_s^2 - X_s^1|^2 ds\right] + E\left[\sup_{0\leq u\leq t}\left|\int_u^t 1_{\{Y_s>0\}} G_s dB_s\right|^2\right]\right\},
\end{aligned}
$$

where C_T is some constant depending only on T. (We will denote all such constant by a same one which may vary line by line). Since

$$\begin{aligned} E\sup_{0\le u\le t}\left|\int_u^t 1_{\{Y_s>0\}}G_s dB_s\right|^2 &\le 2\left\{E\left|\int_0^t 1_{\{Y_s>0\}}G_s dB_s\right|^2 + E\sup_{0\le u\le t}\left|\int_0^u 1_{\{Y_s>0\}}G_s dB_s\right|^2\right\} \\ &\le C_T E\int_0^t 1_{\{Y_s>0\}}|G_s|^2 ds \\ &\le C_T E\int_0^t 1_{\{Y_s>0\}}|X_s^2 - X_s^1|^2 ds, \end{aligned}$$

where the second step is due to Doob's inequality, we then have from (5.9) that

$$E[Y_t^+]^2 \le C_T \int_0^t E[Y_s^+]^2 ds. \tag{5.10}$$

So the Gronwall inequality gives $E[Y_t^+]^2 = 0$, $0 \le t \le T$, which leads to $Y_t^+ = 0$, for all $t \ge 0$ a.s.P, since Y^+ is càglàd. Thus we proved the proposition. □

A direct consequence of Proposition 5.4 is the following comparison theorem. This is sufficent for the proof of our main result of this section, provided the jumps of ξ are "ordered" (*i.e.*, if $\{T_n\}_{n=0}^\infty$ are stopping times whose graphs $\{[[T_n]]\}_0^\infty$ are disjoint, and the union of $[[T_n]]$'s equals the set of jumps of ξ (the existence of such T_n's can be found in Dallacherie [4]), then $0 < T_0 < T_1 < \cdots$). However, if the jumps of ξ are not ordered, this comparison theorem does not yield our main result. But nevertheless, we present this comparison theorem as a corollary of Proposition 5.4 because of its neat form. As far as we know, it is new.

Corollary 5.5. **(Comparison Theorem)** *Let $0 < T \le \infty$ be given and $0 \le \tau_1 < \tau_2 \le T$ be two stopping times. Suppose $\xi^1, \xi^2 \in \hat{\mathcal{D}}$ and (X^1, K^1) and (X^2, K^2) are the corresponding solutions of the SDEDR(I; x^i, ξ^i) (or SDEDR(II; x^i, ξ^i)) $i = 1, 2$, respectively. Suppose that*

(1) *$X^1_{\tau_1} \ge X^2_{\tau_1}$, a.s.$P$;*

(2) *For $P-$a.e. $\omega \in \Omega$, $\xi^1_t(\omega) - \xi^2_t(\omega)$ is nondecreasing on the interval $[\tau_1(\omega), \tau_2(\omega))$;*

(3) *K^2 is continuous on $[\tau_1, \tau_2)$, a.s.P.*

Then for a.e. $\omega \in \Omega$, $X^1_t(\omega) \ge X^2_t(\omega)$, $\tau_1(\omega) \le t < \tau_2(\omega)$.

Proof. We split the proof into three steps.

(i) First assume that $\tau_1 = 0$; $\tau_2 = T$. Let $\eta^i = \xi^i + K^i$, $i = 1, 2$; $\eta = \eta^2 - \eta^1$, $\xi = \xi^2 - \xi^1$, $K = K^2 - K^1$ and still let $Y = X^2 - X^1$. Fix ω in the set with probability one such that conditions (1)–(3) hold. Observe that $\mathcal{H}^1(\omega) = \{t : Y_t(\omega) > 0\} \subseteq \{t : X^2_t(\omega) > 0\}$, by condition (3) and the definition of K^2, $1_{\mathcal{H}^1}(\omega) dK^2_t(\omega) = 0$, $t \ge 0$. Therefore, we have, $P-a.e.$,

$$1_{\mathcal{H}^1} d\eta_t = 1_{\mathcal{H}^1} d(\xi_t + K_t) = 1_{\mathcal{H}^1} d\xi_t - 1_{\mathcal{H}^1} dK^1_t \le 0, \ t \ge 0,$$

by condition (ii) and the fact that K^1 is nondecreasing. Hence the condition (a) of Proposition 5.4 is satisfied.

The next observation is that $\Delta Y_t = \Delta\xi_t - \Delta K^1_t \le 0$, for all $t \ge 0$, since K^2 is continuous. This leads to the condition (b) of Proposition 5.4 and the result follows.

(ii) Next assume that $\tau^1 = 0$; $0 \le \tau = \tau_2 \le T$. Let $(\xi^i)^\tau_t = \xi^i_{t\wedge\tau}$, $t \ge 0$, and let $(\tilde{X}^{i,\tau}, \tilde{K}^{i,\tau})$ be the solutions to SDEDR(I; $x^i, (\xi^i)^\tau$) (or SDEDR(II; $x^i, (\xi^i)^\tau$)), $i = 1, 2$. Applying step (i) to $\tilde{X}^{1,\tau}, \tilde{X}^{2,\tau}$, and noting that $X^{i,\tau}_t = X^i_t$, $0 \le t < \tau$, $a.s.P$, $i = 1, 2$, by the pathwise uniqueness of the solutions of SDEDR, we see that the conclusion is also true in this case.

(iii) Finally, let $0 \le \tau_1 \le \tau_2 \le T$ be general. Observe that the random variable $\tau_2 - \tau_1$ is an $\{\mathcal{F}_{\tau_1+t}\}_{t\ge 0}$-stopping time. We consider the "shifted processes": for $t \ge 0$, $i = 1, 2$, let $\tilde{X}^i_t = X^i_{\tau_1+t}$; $\tilde{\xi}^i_t = \xi^i_{\tau_1+t} - \xi^i_{\tau_1}$; $\tilde{K}^i_t = K^i_{\tau_1+t} - K^i_{\tau_1}$, $\tilde{B}_t = B_{\tau_1+t} - B_{\tau_1}$; and $\tilde{\mathcal{F}}_t = \mathcal{F}_{\tau_1+t}$, $t \ge 0$. Then $\tilde{B}$ is an $\tilde{\mathcal{F}}_t$-Brownian motion such that, as one can easily check, on the space $(\Omega, \mathcal{F}, P; \tilde{\mathcal{F}}_t)$, the conditions of step (ii) above are satisfied by $\tilde{X}^i, \tilde{\xi}^i, i = 1, 2$ with B replaced by $\tilde{B}$ and τ replaced by $\tau_2 - \tau_1$, therefore the result follows from step (ii). □

Remark 5.3 (1) Some direct consequences can be drawn from Corollary 5.5. For example, if ξ^2 is continuous then condition (3) can be dropped. However, condition (2) is not removable (one can easily find a counterexample to show this). In the special case $\xi^1 \equiv \xi^2 \equiv 0$, we get the comparison theorem for reflected diffusions. (The case in which $a = 0$, $\sigma = 1$ is well known).

(2) By applying the same technique as that used in Ikeda and Watanabe [7, VI.1] or Rogers and Williams [15, V.43], one can easily extend Proposition 5.1 and Corollary 5.5 to more general cases in which the drift coefficients are different or even "non-Markovian". □

5.3. Main results. Now let $x \ge 0$ and $\xi \in \hat{\mathcal{D}}$ be given. The first main result of this section is the following:

Theorem 5.6. $X^{x,\xi}_t(I) \ge X^{x,\xi}_t(II)$, $t \ge 0$, $a.s.P$.

Proof. Let $\eta^1 = \xi + K^{x,\xi}(I)$; $\eta^2 = \xi + K^{x,\xi}(II)$, we will check the conditions (a) and (b) of Proposition 5.4.

(a) Denote $X^1 = X^{x,\xi}(I)$, $K^1 = K^{x,\xi}(I)$ and $X^2 = X^{x,\xi}(II)$, $K^2 = K^{x,\xi}(II)$, then $\eta \triangleq \eta^2 - \eta^1 = K^2 - K^1 \triangleq K$. By the definition of SDEDR(II),

$$
\begin{aligned}
(5.11)\qquad K^2_t &= \int_{[0,t)} 1_{\{X^2_{s+}=0\}} dK^2_s = \int_0^t 1_{\{X^2_{s+}=0\}} d(K^2_s)^c + \sum_{0\le s<t} 1_{\{X^2_{s+}=0\}} \Delta K^2_s \\
&= \int_0^t 1_{\{X^2_s=0\}} d(K^2_s)^c + \sum_{0\le s<t} 1_{\{X^2_{s+}=0\}} \Delta K^2_s.
\end{aligned}
$$

Note again that $\mathcal{H}^1 = \{t : X^2_t > X^1_t\} \subseteq \{t : X^2_t > 0\}$, hence $1_{\{X^2_t=0, X^2_t>X^1_t\}} \equiv 0$. By (5.11),

$$
\begin{aligned}
1_{\mathcal{H}^1} d\eta_t = 1_{\{X^2_t>X^1_t\}} d(\eta_t) &= 1_{\{X^2_t>X^1_t\}} (dK^2_t - dK^1_t) \\
&= 1_{\{X^2_{t+}=0,\ X^2_t>X^1_t\}} \Delta K^2_t - 1_{\{X^2_t>X^1_t\}} dK^1_t \\
(5.12)\qquad &= 1_{\{X^2_{t+}=0,\ X^2_t>X^1_t\}} [\Delta K^2_t - \Delta K^1_t] - 1_{\{X^2_t>X^1_t\}} d(K^1_t)^c \\
&\quad - 1_{\{X^2_{t+}>0,\ X^2_t>X^1_t\}} \Delta K^1_t \\
&\le 1_{\{X^2_{t+}=0,\ X^2_t>X^2_t\}} [\Delta K^2_t - \Delta K^1_t], \quad t \ge 0.
\end{aligned}
$$

If t is such that $X^2_{t+} = 0$, $X^2_t > X^1_t$, then $t \in S_{X^2}$ and $X^2_t + \Delta\xi_t \le 0$, since K^2 is nondecreasing. Hence $X^1_t + \Delta\xi_t < X^2_t + \Delta\xi_t \le 0$; and by Proposition 3.1-(4), $t \in S_{K^1}$. Therefore, by Definition 4.1, 4.2 and Proposition 3.2,

$$\Delta K^2_t = |X^2_t + \Delta\xi_t| < |X^1_t + \Delta\xi_t| = \frac{1}{2}\Delta K^1_t.$$

It then follows from (5.12) that $1_{\mathcal{H}^1} d\eta_t = 1_{\{X^2_t > X^1_t\}} d\eta_t \le 0$. $t \ge 0$, $a.s.P$. Condition (a) of Proposition 5.4 is verified. We will now verify (b).

Let $\omega \in \Omega$ and $X^1_t(\omega) \ge X^2_t(\omega)$. Consider $X^2_t(\omega) + \Delta\xi_t(\omega)$.

(Case 1) If $X^2_t + \Delta\xi_t \ge 0$, then $X^1_t + \Delta\xi_t \ge 0$, so by Propositions 3.2 and 3.5, t is the continuity point of both K^1 and K^2. Hence $X^1_{t+} = X^1_t + \Delta\xi_t \ge X^2_t + \Delta\xi_t = X^2_{t+}$.

(Case 2) If $X^2_t + \Delta\xi_t < 0$, then $t \in S_{K^2}$ and by proposition 3.5, we always have $X^2_{t+} = 0 \le X^1_{t+}$.

So in either case, we have $X^1_{t+}(\omega) \ge X^2_{t+}(\omega)$. This varifies condition (b) of Proposition 5.4, and the result follows. □

The next theorem is the most important one for us to construct complete class of admissible control processes in the following sections.

Theorem 5.7. *For any given $x \ge 0$ and $\xi \in \hat{\mathcal{D}}$, there exists $\gamma \in \mathcal{A}$ such that a.s.P.*

(i) $X^{x,\xi}_t(I) \ge X^{x,-\gamma}_t(I), \quad t \ge 0;$

(ii) $K^{x,-\gamma}(I)$ *is continuous;*

(iii) $d\gamma_t \le d\xi^-_t, \quad t > 0.$

Proof. Let $x \ge 0$ and $\xi \in \hat{\mathcal{D}}$ be given. Consider $(X^{x,-\xi^-}(II), K^{x,-\xi^-}(II))$. By Proposition 5.2, there exists an $\eta \in \hat{\mathcal{D}}$ such that $X^{x,-\xi^-}_t(II) = X^{x,-\xi^-+\eta}_t(I), \quad t \ge 0$. Set $\gamma = \xi^- - \eta$. We shall prove that this γ satisfies (i) —(iii) above and is nondecreasing.

(i) The proof is basically the same as that of Theorem 5.6 except for the little change that we only use the $-\xi^-$ part in SDEDR(II). The extra part ξ^+ therefore requires some consideration. For simplicity, denote $X^1_t = X^{x,\xi}_t(I)$; $X^2_t = X^{x,-\xi^-+\eta}_t(I) = X^{x,-\xi^-}_t(II)$; $K^1_t = K^{x,\xi}_t(I)$; $K^2_t = K^{x,-\xi^-+\eta}_t(I)$; $K^3_t = K^{x,-\xi^-}_t(II)$; $\xi^1_t = \xi_t$; $\xi^2_t = -\xi^-_t + \eta_t$. Here η is the process determined by Proposition 5.2, which can be writen down in its explicit form (cf. (5.3)): $\eta_t = K^3_t - \int_0^t 1_{\{X^2_s=0\}} dL^c_s$, where L now is simply K^3. Thus, by the definition of SDEDR(II), we have

$$\begin{aligned}\eta_t &= K^3_t - \int_0^t 1_{\{X^2_s=0\}} d(K^3_s)^c = K^3_t - \int_0^t 1_{\{X^2_{s+}=0\}} d(K^3_s)^c \\ &= \sum_{0 \le s < t} 1_{\{X^2_{s+}=0\}} \Delta K^3_s.\end{aligned} \tag{5.13}$$

We now check the condition (a),(b) of Proposition 5.4 (bearing in mind the difference of the notations here and those in Proposition 5.4).

(a): As in Theorem 5.6, if we let $\tilde{\xi} = \xi^2 - \xi^1$ and $K = K^2 - K^1$, and note that K^2 is continuous and flat on the set $\{t : X_t > 0\}$, we shall have by analogy with (5.12),

$$1_{\mathcal{H}^1} d\eta_t = 1_{\{X^2_t > X^1_t\}}[d(\tilde{\xi}_t + K_t)] \le 1_{\{X^2_t > X^1_t, X^2_{t+}=0\}}(\Delta K^3_t - \Delta K^1_t - \Delta\xi^+_t).$$

Again, from Proposition 3.5, for any $t \in S_{K^3}$ such that $X^2_t > X^1_t$, and $X^2_{t+} = 0$, we have the following cases:

(Case 1) If $X_t^1 - \Delta\xi_t^- + \Delta\xi_t^+ \ge 0 \ge X_t^2 - \Delta\xi_t^- > X_t^1 - \Delta\xi_t^-$, then

$$\Delta\xi_t^+ \ge |X_t^2 - \Delta\xi_t^-| = \Delta K_t^3;$$

(Case 2) If $0 \ge X_t^2 - \Delta\xi_t^- > X_t^1 - \Delta\xi_t^- + \Delta\xi_t^+ \ge X_t^1 - \Delta\xi_t^-$, then $t \in S_{K^1}$ and

$$\Delta K_t^3 = |X_t^2 - \Delta\xi_t^-| < |X_t^1 - \Delta\xi_t^- + \Delta\xi_t^+| = \frac{1}{2}\Delta K_t^1 < \Delta K_t^1;$$

(Case 3) Finally, if $0 > X_t^1 - \Delta\xi_t^- + \Delta\xi_t^+ \ge X_t^2 - \Delta\xi_t^- > X_t^1 - \Delta\xi_t^-$, then $t \in S_{K^1} \cap S_{K^3}$ and

$$\Delta K_t^3 = |X_t^2 - \Delta\xi_t^-| < |X_t^1 - \Delta\xi_t^- + \Delta\xi_t^+| + \Delta\xi_t^+ = \frac{1}{2}\Delta K_t^1 + \Delta\xi_t^+.$$

Hence in any case we have $\Delta K_t^3 - \Delta K_t^1 - \Delta\xi_t^+ \le 0$ (if t is a continuity point of K^3, this is trivially true), which leads to

$$1_{\mathcal{H}^1} d(\tilde{\xi}_t + K_t) = 1_{\{X_t^2 > X_t^1\}} d(\tilde{\xi}_t + K_t) \le 0, \quad t \ge 0, \ a.s.P.$$

Condition (a) is thus proved, whereas condition (b) follows from the same argument as in Theorem 5.6. Therefore, by Proposition 5.4, we obtain Part (i). Part (ii) follows from the construction of γ and Proposition 5.2–(2).

We combine together the proof of (iii) and that γ is nondecreasing. First, note that $d\gamma_t = d(\xi_t^- - \eta_t) = 1_{\{X_{t+}^2 = 0\}}[\Delta\xi_t^- - \Delta K_t^3] + d(\xi_t^-)^c + 1_{\{X_{t+}^2 > 0\}}\Delta\xi_t^-$. Thus

$$1_{\{X_{t+}^2 = 0\}}[\Delta\xi_t^- - \Delta K_t^3] \le d\gamma_t \le d\xi_t^-. \tag{5.14}$$

Clearly, the second inequality in (5.14) gives the part(iii). As for the left most term, it is nonnegative if t is a continuity point of K^3. If $t \in S_{K^3}$, we have $\Delta K_t^3 = |X_t^2 - \Delta\xi_t^-| \le \Delta\xi_t^-$, since $X_t^2 \ge 0$. Therefore $1_{\{X_{t+}^2 = 0\}}[\Delta\xi_t^- - \Delta K_t^3] \ge 0$ for all $t \ge 0$, which leads to that $d\gamma_t \ge 0$, $t \ge 0$. The proof is now complete. □

6. The Singular Stochastic Control Problem with Finite Fuel. From now on, we shall restrict ourselves only to the SDEDR(I). Suppose that on some probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ carrying a one-dimensional $\{\mathcal{F}_t\}$-Brownian Motion $\{B_t; t \ge 0\}$, we have the following stochastic system:

$$X_t = x + \int_0^t a(X_s)ds + \int_0^t \sigma(X_s)dB_s + \xi_t + K_t, \tag{6.1}$$

with the *cost function*:

$$J(\xi; T, x) = E[\int_0^T h(t, X_t^{x,\xi})dt + \int_{[0,T)} f(t)d\check{\xi}_t + g(X_T^{x,\xi})]. \tag{6.2}$$

We impose the following conditions on the functions a, σ, h, f, g:

(6-i) The coefficients a, σ satisfy the conditions (4.2) and (4.3).

(6-ii) The function $h : [0,T] \times [0,\infty) \to [0,\infty)$ is of class $C^{1,1}$; it plays the rôle of a "running cost per unit time" on the state. We assume that $h(t,\cdot)$ is nondecreasing and $h_x(t,0) \geq 0$ for all $t \geq 0$.

(6-iii) The function $f : [0,T] \to [0,\infty)$, which is assumed to be Lipschitz continuous, represents a cost of controlling effort per unit time.

(6-iv) The function $g : [0,\infty) \to [0,\infty)$ is convex, nondeceasing, and of class C^1; it represents a terminal cost on the state.

(6-v) For any $(t,x) \in [0,T] \times [0,\infty)$,

$$0 \leq h_x(t,x) + |h_t(t,x)| + g'(x) \leq C(1+x^m). \tag{6.3}$$

for some $C > 0, m \geq 1$, and

$$\sup_{x\geq 0} g'(x) \leq \inf_{0\leq t\leq T} f(t). \tag{6.4}$$

Remark 6.1. The condition (6.4) is inherited from [8] to facilitate the arguments (*e.g.*, in the proof of Theorem 7.3) which are analogous to those in [8]. □

If on a given probability space, there exists a process $\xi^* \in \mathcal{B}(T,y)$ such that

$$J(\xi^*;T,x) = \inf_{\xi\in\mathcal{B}(T,y)} J(\xi;T,x), \tag{6.5}$$

we say that ξ^* is an optimal control in the *strict sense.* (In [9, 10], an optimal control was sought in the strict sense, but in [8], the optimal control could be realized on a different probability space). In this paper, we shall discuss the existence of the optimal control in a "wide sense" which is basically the same as that in [8].

Let us now assume that $(\Omega,\mathcal{F},\mathcal{F}_t)$ is the canonical space defined by (2.1)–(2.3).

Definition 6.1. *Let $x \geq 0,\ y > 0.\ T > 0$ be given. Define the set $\mathcal{M}(T,x,y)$ to be the subset of $\mathcal{M}$ consisting of all $P \in \mathcal{M}$ such that the canonical process (X_0,B,ξ) satisfies:*

(1) $P(X_0 = x) = 1$.

(2) $P(\xi \in B(T,y)) = 1$. □

Theorem 4.4 tells us that for each $P \in \mathcal{M}(T,x,y)$, there exists a pair $(F_P,\ G_P) \in \mathcal{D}_P^2$ such that $X = F_P, K = G_P$ solves the SDEDR (6.1). By a Monotone-Class argument, we can show that there is a function $I_P^T : \Omega \to \mathbf{R}$ which is $\mathcal{F}_T^P/\mathcal{B}(\mathbf{R})$-measurable so that P-almost surely,

$$I_P^T(x,B,\xi) = \int_0^T h(t,F_P(x,B,\xi)_t)dt + \int_{[0,T)} f(t)d\check{\xi}_t + g(F_P(x,B,\xi)_T), \tag{6.6}$$

and hence the cost function (6.2) can be writen as

$$J(P;T,x) \triangleq E^P[I_P^T(x,B,\xi)]; \tag{6.7}$$

and the value function will be

$$V(T;x,y) = \inf_{P\in\mathcal{M}(T,x,y)} J(P;T,x). \tag{6.8}$$

If there exists a probability space, on which one can find a process $\xi^* \in \mathcal{B}(T,y)$ such that the joint distribution of the triple (x, B, ξ^*), denoted by P^*, satisfies

$$J(P^*; T, x, y) = V(T; x, y),$$

then we call such a ξ^* (together with the canonical probability space corresponding to it) the optimal control in the "wide" sense.

It is easily seen that this model is a generalization of the work of Karatzas and Shreve [9], and El Karoui and Karatzas [8] to the general "diffusion" type with finite fuel. In light of the results of [9] and [8], the following *Complete Class* of admissible controls is of essential importance.

Definition 6.2. *For each $x \geq 0,\ y > 0,\ T > 0$, $\mathcal{D}_{can}(T,x,y) \subseteq \mathcal{M}(T,x,t)$ ("can" for canonical) is the set consisting of all the elements in $\mathcal{M}(T,x,y)$ for which the condition (2) in Definition 6.1 is replaced by*

(2') $P(\xi \in B(T,y),\ \xi^+ \equiv 0) = 1$;

and such that on any realization of $P \in \mathcal{D}_{can}(T,x,y)$, the solution $X^{x,-\xi^-}$ to the equation (6.1) satisfies

$$X_t^{x,-\xi^-} \geq \Delta\xi_t^- \geq 0, \quad \textit{for all}\ \ t \geq 0,\ a.s. \tag{6.9}$$

or equivalently, the corresponding $K^{x,-\xi^-}$ is continuous, a.s.. □

In the sequel, we will simply denote the process ξ^- by ξ if there is no confussion. The following proposition is a direct consequence of Theorem 5.7.

Proposition 6.3. *For any $P \in \mathcal{M}(T,x,y)$, there exists a $\tilde{P} \in \mathcal{D}_{can}(T,x,y)$ such that*

$$J(\tilde{P}; T, x, y) \leq J(P; T, x, y).$$

Proof. We consider the canonical space $(\Omega, \mathcal{F}^P, P; \mathcal{F}^P)$. Then for $P-a.e.\ \omega \in \Omega$, $\xi(\omega) \in B(T,y)$. Let $X^{x,\xi}$ be the solution to (6.1) on this probability space (remember it is the solution to SDEDR(I)). By Theorem 5.7, there exists a process $\gamma \in \mathcal{A}(T,y)$ such that P-almost surely, $X_t^{x,-\gamma} \leq X_t^{x,\xi}$; $d\gamma_t \leq d\check{\xi}_t^- \leq d\check{\xi}_t$ for $t \geq 0$ (hence $\gamma_T \leq \check{\xi}_T \leq y$), and $K^{x,-\gamma}$ is continuous. Let $\tilde{P}$ be the joint distribution of the triple $(x, B, -\gamma)$ back on the canonical space $(\Omega, \mathcal{F})$, then $\tilde{P} \in \mathcal{D}_{can}(T,x,y)$ and the result follows from the assumptions on f, g and h. □

We end up with the following result for the "completeness" of the class $\mathcal{D}_{can}(T,x,y)$.

Proposition 6.4.

$$V(T; x, y) = \inf_{P \in \mathcal{D}_{can}(T,x,y)} J(P; T, x).$$

□

7. Existence of Optimal Control. We take a *minimizing sequence* as follows. Let $x \geq 0,\ y > 0,\ T > 0$ be given, Let $(\Omega^{(n)}, \mathcal{F}^{(n)}, P^{(n)}; \mathcal{F}_t^{(n)})$ be a sequence of probability spaces, on which a sequence of processes $\{B^{(n)}, \xi^{(n)}\}$ is defined such that $\tilde{P}^{(n)}$, the joint distribution of $(x, B^{(n)}, \xi^{(n)})$ under $P^{(n)}$, belongs to $\mathcal{D}_{can}(T; x, y)$ and such that

$$\lim_{n\to\infty} E^{(n)} I^T_{P^{(n)}}(x, B^{(n)}, \xi^{(n)}) = V(T; x, y).$$

Define $Y^{1,(n)} = \int_0^t a(X_s^{(n)})ds$, $Y^{2,(n)} = \int_0^t \sigma(X_s^{(n)})dB_s^{(n)}$, and consider the law (under $P^{(n)}$) of the quintuple $(Y^{1,(n)},\ Y^{2,(n)},\ B^{(n)},\ \xi^{(n)},\ K^{(n)})$ on the space $S = W_0 \times W_0 \times W_0 \times D \times D$. With the Skorohod topology in D, S is a complete, separable metric space. We shall need the following lemmas.

Lemma 7.1. *The laws of the quintuples* $(Y^{1,(n)}, Y^{2,(n)}, B^{(n)}, \xi^{(n)}, K^{(n)})$*;* $n = 1, 2, \ldots$ *are tight.*

Proof. It is sufficient to show that the marginal distributions are tight. The tightness of the triples $(B^{(n)}, \xi^{(n)}, K^{(n)})$; $n = 1, 2, \ldots$ is proved in [8, Proposition 7.4]; thus we only need verify that the first two marginals are tight.

1. $\{Y^{1,(n)}\}_{n=0}^{\infty}$ are tight. By the criterion of tightness of continuous processes (see for instance, [7, I-theorem 4.2]), it suffices to verify that for every $T > 0$, $\epsilon > 0$,

$$\lim_{h\downarrow 0} \sup_n P^{(n)} \left\{ \sup_{s,t\in[0,T];|t-s|\le h} |Y_t^{1,(n)} - Y_s^{1,(n)}| > \epsilon \right\} = 0, \tag{7.1}$$

as $Y_0^{(n)} = 0,\ n = 1, 2, \cdots$. Let $T > 0$, $\epsilon > 0$, $n \ge 1$ be fixed. Since

$$|Y_t^{1,(n)} - Y_s^{1,(n)}|^2 = \left| \int_s^t a(X_s^{(n)})ds \right|^2 \le 2TC^2(1 + \sup_{0\le s\le T} |X_s^{(n)}|^2)h$$

for all $s, t \in [0, T]$, $|t - s| \le h$, $a.s.P^{(n)}$, the Chebyshev inequality and the moment estimate (4.5) for $X^{(n)}$ lead to

$$P^{(n)} \left\{ \sup_{s,t\in[0,T];|t-s|\le h} |Y_t^{1,(n)} - Y_s^{1,(n)}| > \epsilon \right\} \le \epsilon^{-2} 2TC^2[1 + C_{2,T}(1 + |x|^2 + y^2)]h.$$

Thus, (7.1) holds and the laws of $\{Y^{1,(n)}\}_{n=1}^{\infty}$ are tight.

2. We verify that $Y^{2,(n)}$'s also satisfy (7.1). Let $T > 0$, $\epsilon > 0$ be given, and $n \ge 1$, $0 < h < 1$ be fixed. Define a partition of $[0, T]$: $0 = t_0 < t_1 < t_2 < \cdots < t_N = T$, so that $t_i - t_{i-1} = h$, $i = 1, 2, \ldots, N - 1$ and $t_N - t_{N-1} \le h$. Clearly, $T \le Nh \le T + h < T + 1$. It is not hard to check that the set

$$A \triangleq \{ \sup_{s,t\in[0,T];|t-s|\le h} |Y_t^{2,(n)} - Y_s^{2,(n)}| > \epsilon \} \subseteq \bigcup_{k=0}^{N-1} \{ \sup_{0\le t\le 2h} |Y_{(t_k+t)\wedge T}^{2,(n)} - Y_{t_k}^{2,(n)}| > \frac{\epsilon}{2} \}.$$

Therefore,

$$P^{(n)}(A) \le \sum_{k=0}^{N-1} P^{(n)} \left\{ \sup_{0\le t\le 2h} |Y_{(t_k+t)\wedge T}^{2,(n)} - Y_{t_k}^{2,(n)}| > \frac{\epsilon}{2} \right\}. \tag{7.2}$$

Now for each fixed k, $Z_t^k \triangleq Y_{(t_k+t)\wedge T}^{2,(n)} - Y_{t_k}^{2,(n)}$ is an $\mathcal{F}_t^k \triangleq \mathcal{F}_{t_k+t}$-martingale and $[Z^k]_t = \int_{t_k}^{(t_k+t)\wedge T} \sigma^2(X_s^{(n)})ds$, so by applying the Burkholder-Davis-Gundy inequality, and noting (4.2), (4.5), we obtain that for any $m > 1$,

$$E^{(n)} \left[\sup_{0\le t\le 2h} Z_t^k \right]^{2m} \le E^{(n)}([Z^k]_{2h})^m \le C_{m,T}(1 + |x|^{2m} + y^{2m})h^m, \tag{7.3}$$

where $C_{m,T}$ is a constant depending only on m, T. Therefore, the Chebyshev's inequality, together with (7.2) and (7.3), leads to

$$P^{(n)}(A) \leq \sum_{k=0}^{N-1} \epsilon^{-2m} 2^{2m} E^{(n)} \left[\sup_{0 \leq t \leq 2h} Z_t^k \right]^{2m} \leq C_T(\epsilon)(1 + |x|^{2m} + y^{2m}) N h^m.$$

Since $Nh < T + 1$ and $m > 1$, we get $\lim_{h \downarrow 0} \sup_n P^{(n)}(A) = 0$. Noting that $Y_0^{2,(n)} = 0,\ n = 1, 2, \cdots$, the lemma is proved. □

We can now apply the Skorohod Theorem (see, for example, [6, p.102]), to get a probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ on which is defined a sequence of adapted processes $\{(\hat{Y}^{1,(n)}, \hat{Y}^{2,(n)}, \hat{B}^{(n)}, \hat{\xi}^{(n)}, \hat{K}^{(n)})\}_{n=1}^{\infty}$, so that

(i) The quintuples $(Y^{1,(n)}, Y^{2,(n)}, B^{(n)}, \xi^{(n)}, K^{(n)})$ and $(\hat{Y}^{1,(n)}, \hat{Y}^{2,(n)}, \hat{B}^{(n)}, \hat{\xi}^{(n)}, \hat{K}^{(n)})$ are identical in law for $n = 1, 2, ...$;

(ii) the sequence $\{(\hat{Y}^{1,(n)}, \hat{Y}^{2,(n)}, \hat{B}^{(n)}, \hat{\xi}^{(n)}, \hat{K}^{(n)})\}$ converges, almost surely, to the quintuple $(Y^1, Y^2, B, \theta, L) \in W_0 \times W_0 \times W_0 \times D \times D$, as $n \to \infty$.

To simplify notation, we now drop " ˆ " for this sequence. Note that each $B^{(n)}$ is a Brownian motion under P, one can easily check that so is B. Also note that on the probability space $(\Omega, \mathcal{F}, P; \mathcal{F}_t)$ the processes $X_t^{(n)} = x + Y_t^{1,(n)} + Y_t^{2,(n)} - \xi^{(n)} + K_t^{(n)}$ will converge, almost surely, to a process $X \in \mathcal{D}$ in the Skorohod topology.

Lemma 7.2. *The process X satisfies the S.D.E:*

$$X_t = x + \int_0^t a(X_s) ds + \int_0^t \sigma(X_s) dB_s - \theta_t + L_t. \tag{7.4}$$

Proof. The proof is technical and lengthy but not of special importance in this paper, we therefore omit it. The details can be found in [12, II.7]. □

We can now follow the argument of El Karoui and Karatzas [8, p.241] line by line to derive our result. Here we only give the outline for completeness (one is referred to [8, p.241-242] for detail). First, note that the process X is nonnegative and is the solution to the SDE (7.4), so by Proposition 5.1, we can find processes $\zeta,\ K \in \mathcal{A}$ such that (X, K) solves the SDEDR$(x, \zeta - \theta)$. By using Proposition 12.1 in [8], Dominated Convergence Theorem, Fatou's lemma and the conditions on the functions $g,\ f,\ h$ (6-ii)–(6-v), one shows that, with P denoting the joint distribution of $(x, B, -\theta + \zeta)$,

$$\begin{aligned}
J(P; T, x) &= E\left[\int_0^T h(t, X_t) dt + \int_{[0,T)} f(t) d\theta_t + g(Y_T) \right] \\
&\leq \liminf_{n \to \infty} E\left[\int_0^T h(t, X_t^{(n)}) dt + \int_{[0,T)} f(t) d\xi_t^{(n)} + g(X_T^{(n)}) \right] \\
&= \lim_{n \to \infty} E^{(n)} \left[I_{P^{(n)}}^T (x, B^{(n)}, \xi^{(n)}) \right] \\
&= V(T, x, y).
\end{aligned}$$

Finally, by applying Proposition 6.3 and noting that $\theta_T \leq y$, we find a $P^* \in \mathcal{D}_{can}(T; x, y)$ such that

$$J(P^*; T, x) \leq J(P; T, x) \leq V(T, x, y).$$

Moreover, Theorem 5.7 enables us actually to find a process $\xi^* \in \mathcal{A}(T,y)$ on the same probability space on which B, θ and L are defined, such that P^* is the joint distribution of $(x, B, -\xi^*)$, *i.e.*, $-\xi^*$ is the optimal control for our singular stochastic control problem. We have proved the following theorem:

Theorem 7.3. *Under the condition (6-i)–(6-v), for any $x \geq 0,\ y > 0$, there exist a probability measure P^* on the canonical base $(\Omega, \mathcal{F})$, $P^* \in \mathcal{D}_{can}(T,x,y)$, such that*

$$E^{P^*}\left[I^T_{P^*}(x,B,\xi)\right] = V(T,x,y).$$

More precisely, there exists a probability space $(\Omega', \mathcal{F}', P'; \mathcal{F}'_t)$ on which we can find a Brownian motion B and a process $\xi^* \in \mathcal{B}(T,y)$ satisfing (6.9) on this probability space and such that the joint distribution of the triple $(x, B, -\xi^*) = P^*$. The process ξ^* is called the optimal control. □

8. Appendix 1. (A Counterexample). To show that the mapping Γ^I of Definition 3.3 is not Lipschitz continuous under the uniform topology on D, it suffices to construct for any integer $N > 0$ two functions $Y_1^N(\cdot),\ Y_2^N(\cdot) \in D$ such that for some $T > 0$,

$$\sup_{0\leq t\leq T} |\Gamma^I(Y_1^N)(t) - \Gamma^I(Y_2^N)(t)| > N \sup_{0\leq t\leq T} |Y_1^N(t) - Y_2^N(t)|. \tag{8.1}$$

To simplify the notation, we denote $\Gamma = \Gamma^I$. Fix $N > 0$. Define

$$f(t;\ m,\ (n,n+2]) = (m-2)1_{(n,n+1]}(t) + m1_{(n+1,n+2]}(t), \quad m \in \mathbf{Z},\ n \in \mathbf{N}.$$

For $t \in [0,\ 4(N+1)]$, let

$$\begin{aligned} Y_1^N(t) &= \sum_{k=0}^N f(t;\ a_k^1,\ (4k+1,\ 4k+3]) + \sum_{k=0}^N b_k^1 \cdot 1_{(4k+3,\ 4(k+1)+1]}(t) \\ &= \sum_{k=0}^N \left[(a_k^1 - 2)1_{(4k+1,\ 4k+2]}(t) + a_k^1 \cdot 1_{(4k+2,\ 4k+3]}(t)\right] \\ &\quad + \sum_{k=0}^N b_k^1 \cdot 1_{(4k+3,\ 4(k+1)+1]}(t); \end{aligned} \tag{8.2}$$

$$\begin{aligned} Y_2^N(t) &= -1_{(0,1]}(t) - 2\cdot 1_{(1,3]}(t) + \sum_{k=1}^N f(t;\ a_k^2,\ (4k-1,\ 4k+1]) \\ &\quad + \sum_{k=1}^N b_k^2 \cdot 1_{(4k+1,\ 4k+3]}(t), \\ &= -1_{(0,1]}(t) - 2\cdot 1_{(1,3]}(t) \\ &\quad + \sum_{k=1}^N \left[(a_k^2 - 2)1_{(4k-1,4k]}(t) + a_k^2 \cdot 1_{(4k,4k+1]}(t)\right] \\ &\quad + \sum_{k=1}^N b_k^2 \cdot 1_{(4k+1,\ 4k+3]}(t), \end{aligned} \tag{8.3}$$

where,

$$\begin{cases} a_k^1 = -(1+10k+8\sum_{i=1}^{k-1} i); \\ b_k^1 = a_k^1 - (5+4k), \qquad\qquad k=0,1,...N; \end{cases} \tag{8.4}$$

$$\begin{cases} a_k^2 = -(5+14(k-1)+8\sum_{i=1}^{k-2} i); \\ b_k^2 = a_k^2 - (7+4(k-1)), \qquad\qquad k=1,...,N. \end{cases} \tag{8.5}$$

(Here, we define $\sum_{i=1}^{j} i = 0$ for $j<1$). For $t \ge 4(N+1)$, let $Y_i^N(t) = Y_i^N(4(N+1))$, $i = 1,2$.

Let $\delta Y^N = Y_1^N - Y_2^N$. It is easily seen that $|\delta Y^N(t)| = 1$ for $t \in [0,\ 4]$. Furthermore, one can check that, for $k = 1,...,N$,

$$\delta Y^N(t) = \begin{cases} b_{k-1}^1 - a_k^2, & t \in (4k,\ 4k+1]; \\ (a_k^1 - 2) - b_k^2, & t \in (4k+1,\ 4k+2]; \\ a_k^1 - b_k^2, & t \in (4k+2,\ 4k+3]; \\ b_k^1 - (a_{k+1}^2 - 2), & t \in (4k+3,\ 4k+4]. \end{cases}$$

After a little algebra, by virtue of (8.4), (8.5), one can see that $|\delta Y^N(t)| \equiv 1$ for $t \in (4,\ 4(N+1)]$ (and so for $t \in [0,\infty)$). Namely, we have $\sup_{0\le t\le\infty} |\delta Y^N(t)| = 1$.

However, solving DRP(Y_1^N), (Y_2^N) by using (8.2), (8.3) and Definition 3.1, one has

$$\begin{aligned} \Gamma(Y_1^N)(t) &= \sum_{k=0}^{N} \Big[(4k+3)1_{(4k+1,\ 4k+2]}(t) + (4k+5)1_{(4k+2,\ 4k+3]}(t)\Big]; \\ \Gamma(Y_2^N)(t) &= 1_{[0,1)}(t) + \sum_{k=0}^{N} \Big[(4k+1)1_{(4k-1,4k]}(t) + (4k+3)1_{(4k,4k+1]}(t)\Big], \end{aligned}$$

(drawing a picture would be very helpful). Therefore, for $t \in [0,\ 4(N+1)]$,

$$|\Gamma(Y_1^N)(t) - \Gamma(Y_2^N)(t)| = \begin{cases} 1 & t \in [0,\ 1] \\ 4k+3 & t \in (4k,\ 4k+1] \\ 4k+5 & t \in (4k+2,\ 4(k+1)].\ \text{for } k=1,...,N. \end{cases}$$

So, Y_1^N, Y_2^N obviously meet our requirement (8.1). □

9. Appendix 2 (Sketch of the Proof of Theorem 4.3). We first introduce some spaces of semimartingales that were defined in Protter [14]. Let $\mathcal{SM}$ denotes the space of all semimartingales with paths in D. Then any $Y \in \mathcal{SM}$ admits a decomposition $Y_t = Y_0 + M_t + A_t$, $t \ge 0$, where M is a local martingale and A an adapted process with paths in D and of locally bounded variation, such that $M_0 = A_0 = 0$ (one should note that this decomposition is unique only when A is *natural*). It is easily seen that $Y \in \mathcal{SM}$ if and only if its càdlàg modification is indistinguishable with a *classical semimartingale* (cf. [14, III]). We denote the elements in $\mathcal{SM}$ with $Y_0 = 0$ by $\mathcal{SM}_0$.

Let us define two subspaces of $\mathcal{SM}$. First, for $Y \in \mathcal{SM}$, denote $Y_t^* = \sup_{0\le s\le t} |Y_s|$, $t \ge 0$, and $Y_\infty^* = \lim_{t\nearrow\infty} Y_t^*$. For $1 \le p \le \infty$, define $\|Y\|_{S^p} = \|Y_\infty^*\|_{L^p}$ and $S^p = \{Y \in$

$\mathcal{SM} : \|Y\|_{S^p} < \infty\}$. Furthermore, let $Y \in \mathcal{SM}_0$ with the decomposition $Y = M + A$, such that $M_{0+} = A_{0+} = 0$, define

$$j_p(M, A) = \|[\overline{M}, \overline{M}]_\infty^{1/2} + \int_0^\infty |d\overline{A}_s|\|_{L^p}, \tag{9.1}$$

$$\|Y\|_{H_0^p} = \inf_{Y=M+A} j_p(M, A). \tag{9.2}$$

where $\overline{M}$ and $\overline{A}$ are the càdlàg modifications of M and A respectively, and the infimum is taken over all possible decompositions of Y. Define $H_0^p = \{Y \in \mathcal{SM}_0 : Y_0 = Y_{0+} = 0, \ \|Y\|_{H^p} < \infty\}$. It is known (cf. [14, pp 188-189 and Theorem IV.2.1]) that the space H_0^2 is a Banach space. Furthermore, if Y is a general element in $\mathcal{SM}$ with the decomposition $Y = Y_0 + M + A$, we can write

$$Y_t = 1_{\{t=0\}} Y_0 + 1_{\{t>0\}} Y_{0+} + \tilde{Y}_t, \quad t \geq 0, \tag{9.3}$$

where $Y_{0+} = M_{0+} + A_{0+}$ and $\tilde{Y}_0 = \tilde{Y}_{0+} = 0$. It is now easy to construct an injective mapping $\phi : L^2 \oplus L^2 \oplus \mathcal{H}_0^2 \to \mathcal{SM}$ in an obvious way so that the $\mathcal{H}^2 \triangleq \text{Range}(\phi)$ is a Banach space with the norm

$$\|Y\|_{\mathcal{H}^2} = \|Y_0\|_{L^2} + \|Y_{0+}\|_{L^2} + \|\tilde{Y}\|_{H_0^2}.$$

In general, we shall define for any $1 \leq p \leq \infty$ and $Y \in \mathcal{SM}$,

$$\|Y\|_{\mathcal{H}^p} = \|Y_0\|_{L^p} + \|Y_{0+}\|_{L^p} + \|\tilde{Y}\|_{H_0^p}. \tag{9.4}$$

The following propositions are modifications of those in [14, V.2]:

Proposition 9.1. *For any $1 \leq p \leq \infty$, there exists a constant $C_p > 0$, such that for any $Y \in \mathcal{SM}$, $\|Y\|_{S^p} \leq C_p \|Y\|_{\mathcal{H}^p}$.* □

Proposition 9.2. (Emery's Inequality) *Let Y be a semimartingale, $Y_0 = Y_{0+} = 0$, $H \in \mathcal{D}$ and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ ($1 \leq p \leq \infty;\ 1 \leq q \leq \infty$), one has*

$$\| \int_0^\cdot H_s d\overline{Y}_s \|_{\mathcal{H}^r} \leq \|H\|_{S^p} \|Y\|_{\mathcal{H}^q},$$

where $\overline{Y}$ is the càdlàg modification of Y. □

Sketch of the proof of Theorem 4.3.

Step 1. ("Local" existence and uniqueness). Let $T > 0$ be given. Suppose $E(|X_0|^2) < \infty$, $E(\check{\xi}_T^2) < \infty$. For the given $T > 0$, define the semimartingales $Z_t^1 = t \wedge T$ and $Z_t^2 = B_{t\wedge T},\ t \geq 0$. Define a mapping $\Lambda : \mathcal{H}^2 \to \mathcal{SM}$ by

$$\Lambda(Y)_t = X_0 + \int_0^t a(\Gamma(Y)_s) dZ_s^1 + \int_0^t \sigma(\Gamma(Y)_s) dZ_s^2 + \xi_t^T, \quad t \geq 0,\ Y \in \mathcal{H}^2, \tag{9.5}$$

where $\xi_t^T = \xi_{t\wedge T}$ for $t \geq 0$ and $\Gamma = \Gamma^I$. By using Proposition 9.1 and 9.2, one can show that Λ is a contraction mapping from the Banach space $\mathcal{H}^2$ into itself provided

$T > 0$ is small enough. Therefore, the Fixed Point Theorem leads to the existence and uniqueness of a semimartingale $Y \in \mathcal{H}^2$ satisfying the equation (4.4) for $t \in [0,\ T]$.

Step 2. To get the "global" solution, we use an induction argument. Namely, we prove that for each positive integer n, there exists a unique semimartingale $Y^{(n)} \in \mathcal{H}^2$ such that $Y^{(n)}$ satisfy (4.4) on $[0, nT]$. Step 1 shows that this is true for $n = 1$. A standard "path–shifting" and Monotone class argument, together with the uniqueness of DRP, enables one to extend the solution to $[0, (n+1)T]$ if it exists on $[0, nT]$, which finishes the proof of the induction step.

Finally, since for $n > m$, the semimartingale $(Y^{(n)})^{(m)}_{\cdot} \triangleq Y^{(n)}_{\cdot \wedge mT} \in \mathcal{H}^2$ satisfies (4.4) on $[0, mT]$, the uniqueness of such element in $\mathcal{H}^2$ gives that $Y^{(n)}_t = Y^{(m)}_t$, for all $t \in [0, mT]$ $a.s.P$. Hence the family $\{Y^{(n)}\} \subseteq \mathcal{H}_2$ will be consistant and then one can "patch" them up by defining, for every $t \in [0, \infty)$, $Y_t = Y^{(n)}_t$ if $t \in [0, nT]$. Then it is easily seen that Y will be the global solution as we want. $\square$

Acknowledgement: I would like to thank my advisor Professor Naresh Jain for his constant encouragement and generous guidance. I also thank Professor I. Karatzas for providing reference [2] and [8]; especially for his generous, systematical comments on the first draft of this paper. My thanks are also due to Dr. Roger Tribe with whom I enjoyed many helpful discussions, and to the late Professor Steven Orey for providing a preprint of [5]. Finally, I thank the referee for the helpful comments.

REFERENCES

[1] R.F. Anderson and S. Orey, *Small Random Pertubation of Dynamical Systems with Reflecting Boundary,* Nagoya Math. J. Vol.60, (1976), pp. 189-126.

[2] F.M. Baldursson, *Topics in singular stochastic control and optimal stopping,* Doctoral dissertation. Dept. of Statistics, Columbia University, New York (1985).

[3] M. Chaleyat-Maurel, N. El Karoui and B. Marchal, *Réflexion discontinue et systèmes stochastiques,* Ann. Probab. **8** (1980), pp. 1049-1067.

[4] C. Dellacherie and P-A. Meyer, *Probabilités et Potentiel—Chapitres V à VIII,* Hermann (1980).

[5] P. Dupuis and H. Ishii, *On Lipschitz Continuity of the Solution Mapping to the Skorohod Problem, with Applications,* Stochastics and Stochastic Reports, Vol. 35, pp. 31-62 (1991).

[6] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence,* John Wiley & Sons. Inc. (1986)

[7] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes,* North-Holland (1981).

[8] N. El Karoui and I. Karatzas, *Probabilistic Aspects of Finite-Fuel, Reflected Follower Problems,* Acta Applicandae Mathematicae **11** (1988), pp. 223-258.

[9] I. Karatzas and S.E. Shreve, *Connection between Optimal Stopping and Singular Stochastic Control II: Reflected follower Problems,* SIAM J. Control & Optim. **23** (1985), pp. 433-451.

[10] I. Karatzas and S.E. Shreve, *Equivalent Models for Finite-Fuel Stochastic Control,* Stochastics **18** (1986), pp. 245-276.

[11] J. F. Le Gall, *Applications du temps local aux equations différentielles stochastiques unidimensionelles,* Séminaire de Probabilités XVII: Lecture Notes in Math. Vol. **986**, (1983), Springer, pp. 15–31.

[12] J. Ma, *Topics in singular stochastic control and related stochastic differential equations,* Ph.D. Dissertation. Department of Mathematics, University of Minnesota, (in process).

[13] P.A. Meyer, *Un Cours Sur Les Intégrales Stochastiques,* Séminaire de Probabilités X: Lecture Notes in Math. Vol. **511** (1976), Springer.

[14] P. Protter, *Stochastic Integration and Differential Equations, A New Approach,* Springer-Verlagr, Berlin-Heideberg-New York (1990).

[15] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales, Vol 2: Itô Calculus,* John Wiley & Sons. Ltd. (1987).