Received: 4 March 2022 | Accepted: 11 March 2023

DOI: 10.1111/mafi.12385

ORIGINAL ARTICLE

MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics

WILEY

Equilibria of time-inconsistent stopping for one-dimensional diffusion processes

Erhan Bayraktar[1] | **Zhenhua Wang**[1] | **Zhou Zhou**[2]

[1]Department of Mathematics, University of Michigan, Ann Arbor, Michigan, USA

[2]School of Mathematics and Statistics, University of Sydney, Sydney, Australia

Correspondence
Erhan Bayraktar.
Email: erhan@umich.edu

Funding information
National Science Foundation, Directorate for Mathematical and Physical Sciences, Grant/Award Number: DMS-2106556

Abstract

We consider three equilibrium concepts proposed in the literature for time-inconsistent stopping problems, including mild equilibria (introduced in Huang and Nguyen-Huu (2018)), weak equilibria (introduced in Christensen and Lindensjö (2018)), and strong equilibria (introduced in Bayraktar et al. (2021)). The discount function is assumed to be log subadditive and the underlying process is one-dimensional diffusion. We first provide necessary and sufficient conditions for the characterization of weak equilibria. The smooth-fit condition is obtained as a by-product. Next, based on the characterization of weak equilibria, we show that an optimal mild equilibrium is also weak. Then we provide conditions under which a weak equilibrium is strong. We further show that an optimal mild equilibrium is also strong under a certain condition. Finally, we provide several examples including one showing a weak equilibrium may not be strong, and another one showing a strong equilibrium may not be optimal mild.

1 | INTRODUCTION

On a filtered probability space, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ consider the optimal stopping problem,

$$\sup_{\tau\in\mathcal{T}} \mathbb{E}[\delta(\tau)f(X_\tau)], \tag{1}$$

where $\delta(\cdot)$ is a discount function, $X = (X_t)_t$ is a time-homogeneous one-dimensional strong Markov process, and $f(\cdot)$ is a payoff function. It is well known that when $\delta(\cdot)$ is not exponential, the problem can be time-inconsistent in the sense that an optimal stopping rule obtained today may no longer be optimal from a future's perspective.

One way to deal with this time-inconsistency is to consider the precommitted strategy, that is, to derive a policy that is optimal with respect to the initial preference and stick to it over the whole planning horizon even if the preference changes later; see, for example, Agram and Djehiche (2021); Miller (2017). Another approach to address the time-inconsistency is to look for a subgame perfect Nash equilibrium; given the future selves follow the equilibrium strategy, the current self has no incentive to deviate from it. For equilibrium strategies, we refer to the works (Björk et al., 2021; Ekeland & Lazrak, 2010; Ekeland & Pirvu, 2008; He & Jiang, 2021; Ekeland & Lazrak, 2006; Hernández & Possamaï, 2020; Hamaguchi, 2021; Huang & Zhou, 2021; Wang & Yong, 2021; Wei et al., 2017) among others for time-inconsistent control, and Christensen and Lindensjö (2020a, 2020b); Ebert and Strack (2018); He and Zhou (2022); Huang and Nguyen-Huu (2018); Liang and Yuan (2021); Tan et al. (2021); Bodnariu et al. (2022); Ebert et al. (2020), and the references therein for time-inconsistent stopping.

How to properly define the notion of an equilibrium is quite subtle in continuous time. There are mainly two streams of research for equilibrium strategies of time-inconsistent stopping problems in continuous time. In the first stream of research, the following notion of equilibrium is considered.

Definition 1.1. A closed set $S \subset \mathbb{X}$ is said to be a mild equilibrium, if

$$\begin{cases} f(x) \leq J(x,S), & \forall x \notin S, \quad (2)\\ f(x) \geq J(x,S), & \forall x \in S, \quad (3) \end{cases}$$

where

$$J(x,S) := \mathbb{E}^x[\delta(\rho_S)f(X_{\rho_S})] \quad \text{with} \quad \rho_S := \inf\{t > 0 : X_t \in S\} \text{ and } \mathbb{E}^x[\cdot] = \mathbb{E}[\cdot|X_0 = x]. \tag{4}$$

This kind of equilibrium is first proposed and studied in stopping problems in the context of nonexponential discounting in Huang and Nguyen-Huu (2018). It is called *mild equilibrium* in Bayraktar et al. (2021) to distinguish from other equilibrium concepts. Mild equilibria are further considered in Huang et al. (2020) and Huang & Yu (2021) where the time inconsistency is caused by probability distortion and model uncertainty, respectively.

Note that $f(x)$ is the value for immediate stopping, and $J(x,S) = \mathbb{E}^x[\delta(\rho_S)f(X_{\rho_S})]$ is the value for continuing as ρ_S is the first time to enter S after time 0. As a result, the economic meaning of mild equilibria appears to be clear: in Equation (2) when $x \notin S$, it is better to continue and get the value J rather than to stop and get the value f. In other words, there is no incentive to deviate from the action of "continuing." The same reasoning seems to also apply to the other case $x \in S$

in Equation (3), that is, no incentive for changing the action from "stopping" to "continuing." However, this is not really captured in Equation (3) after a second thought: In the one-dimensional diffusion (and continuous-time Markov chain) setting, under some very nonrestrictive condition, we have $\rho_S = 0$ a.s., and thus Equation (3) holds trivially.[1] That is, there is no actual deviation from stopping to continuing captured in Equation (3).

Because of this issue, mild equilibria are indeed too "mild": the whole state space is always a mild equilibrium; in most of the examples provided in Huang and Nguyen-Huu (2018); Huang and Zhou (2020); Huang & Yu (2021), there is a continuum of mild equilibria. As there are often too many mild equilibria in various models, it is natural to consider the problem of equilibrium selection.

Definition 1.2. A mild equilibrium S is said to be optimal, if for any other mild equilibrium R,

$$\mathbb{E}^x[\delta(\rho_S)f(X_{\rho_S})] \geq \mathbb{E}^x[\delta(\rho_R)f(X_{\rho_R})], \quad \forall\, x \in \mathbb{X}.$$

Note that the optimality of a mild equilibrium is defined in the sense of pointwise dominance, which is a very strong condition. The existence of optimal equilibria is first established in Huang and Zhou (2019) in discrete time models. The existence result is further extended to diffusion models for one-dimensional case in Huang and Zhou (2020) and multidimensional case in Huang and Wang (2021). In particular, for the one-dimensional diffusion case, Huang and Zhou (2020) shows that under some general assumptions an optimal mild equilibrium exists and is given by the intersection of all mild equilibria (also see Lemma 4.1 below). Huang and Zhou (2020) also provide an example indicating that, in general, there may exist multiple optimal mild equilibria.

In the second stream of the research for equilibrium strategies for time-inconsistent stopping in continuous time, the following notion of equilibrium is introduced.

Definition 1.3. A closed set $S \subset \mathbb{X}$ is said to be a weak equilibrium, if

$$\begin{cases} f(x) \leq J(x, S), & \forall x \notin S, \qquad (5)\\ \liminf\limits_{\varepsilon \searrow 0} \dfrac{f(x) - \mathbb{E}^x[\delta(\rho_S^\varepsilon)f(X_{\rho_S^\varepsilon})]}{\varepsilon} \geq 0, & \forall x \in S, \qquad (6) \end{cases}$$

where

$$\rho_S^\varepsilon := \inf\{t \geq \varepsilon : X_t \in S\}.$$

The weak equilibrium concept for time inconsistent stopping is proposed in Christensen and Lindensjö (2018), and further studied in Christensen and Lindensjö (2020a); Liang and Yuan (2021); Tan et al. (2021). Obviously, as Equation (3) trivially holds for one-dimensional process, a weak equilibrium is also mild. Compared to mild equilibria, the condition (3) is replaced by Equation (6) for weak equilibria using a first-order condition. This is analog to the first-order condition criterion in time-inconsistent control. As $\rho_S^\varepsilon \geq \varepsilon > 0$, the condition (6) does capture the deviation from stopping to continuing, and is much stronger than Equation (3). However,

[1] In multidimensional setting, if $x \in S^o$, then $\rho_S = 0$, $\mathbb{P}^x$-a.s.; if $x \in \partial S$, then the identity $\rho_S = 0$ requires some regularity of ∂S, and consequently, the verification of Equation (3) on the boundary may not be trivial.

there is still a drawback for Equation (6): when the limit is equal to zero, it is possible that for all $\varepsilon > 0$, we have $f(x) < \mathbb{E}^x[\delta(\rho_S^\varepsilon) f(X_{\rho_S^\varepsilon})]$, and thus there is an incentive to deviate (see Björk et al. (2017, Remark 3.5), Huang & Zhou (2021); Bayraktar et al. (2021); He & Jiang (2021) for more details). Roughly speaking, this is similar to a critical point not necessarily being a local maximum in calculus.

Recently, Bayraktar et al. (2021) investigated the relation between the equilibrium concepts in these two streams of research we described above, and proposed an additional notion of equilibria.

Definition 1.4. A closed set $S \subset \mathbb{X}$ is said to be a strong equilibrium, if

$$\begin{cases} f(x) \le J(x, S), & \forall x \notin S, \\ \exists \varepsilon(x) > 0, \text{ s.t. } \forall \varepsilon' \le \varepsilon(x), f(x) - \mathbb{E}^x[\delta(\rho_S^{\varepsilon'}) f(X_{\rho_S^{\varepsilon'}})] \ge 0, & \forall x \in S. \end{cases} \tag{7}$$

Note that in the definition of strong equilibrium, the first-order condition (6) is replaced by a local maximum condition (7). This remedies the issue of weak equilibria mentioned in the above, and captures the economic meaning of "equilibrium" more accurately. Such kind of equilibria is also studied in Huang and Zhou (2021); He and Jiang (2021) for time inconsistent control. Obviously, a strong equilibrium must be weak. In Bayraktar et al. (2021) under continuous-time Markov chain models with nonexponential discounting, a complete relation between mild, optimal mild, weak, and strong equilibria is obtained:

$$\text{optimal mild} \subsetneq \text{strong} \subsetneq \text{weak} \subsetneq \text{mild}. \tag{8}$$

In this paper, we aim to establish the result (8) for one-dimensional diffusion models under nonexponential discounting. Compared to Bayraktar et al. (2021), the analysis in this paper is much more delicate. The proof in Bayraktar et al. (2021) crucially relies on the discrete state space of the Markov chain setting, and many critical ideas and steps therein cannot be applied in our diffusion framework, where novel approaches are needed for the characterizations of weak and strong equilibria. Here we list the main contributions of our paper as follows.

- We provide a complete characterization (necessary and sufficient conditions) of weak equilibria. As a by-product, we show that any weak equilibrium must satisfy the smooth-fit condition when the pay-off function f is smooth. This gives a much sharper result in a much more general setting as compared to the smooth-fit result obtained in Tan et al. (2021). (See Remark 3.3 for more details.) Moreover, in our paper f need not to be smooth, and our result also indicates that the smooth-fit condition is a special case of the "local convexity" property of weak equilibria. See Remark 3.4. Undoubtedly, such results related to smooth-fit condition have no correspondence in the Markov chain framework in Bayraktar et al. (2021).
- We show that an optimal mild equilibrium is also a weak equilibrium. This proves that the set of weak equilibria is not empty. In terms of the mathematical method, in Bayraktar et al. (2021), the technique for the proof of such result relies on the fact that removing a point from a stopping region changes the stopping time, which is no longer applicable in the diffusion context. A different approach is developed to overcome this difficulty.
- We provide a sufficient condition under which a weak equilibrium is also strong. The condition is easy to verify as suggested by our examples. We also show that one may remove some "inessential" part of an optimal mild equilibrium, and the remaining part is still optimal mild

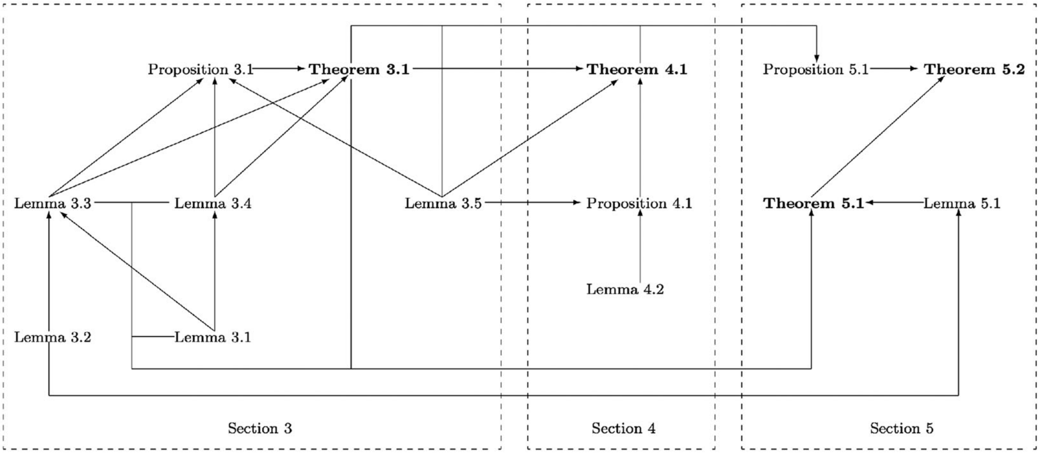

FIGURE 1 Relations between results in Sections 3–5 of this paper. $A \to B$ means that statement A is used in the proof of statement B.

(and thus weak), and in fact strong under an additional assumption. In particular, this result implies that the smallest mild equilibrium essentially has no "inessential" parts and thus is strong. See Theorem 5.2 and Remark 5.3.

The rest of the paper is organized as follows. Section 2 introduces the notation and main assumptions, as well as some auxiliary results that will be used frequently throughout the paper. In Section 3, we provide a complete characterization of a weak equilibrium. In Section 4, we show that an optimal mild equilibrium is a weak equilibrium. Next, in Section 5, we provide a sufficient condition for a weak equilibrium to be strong. We also demonstrate how to construct a strong equilibrium from an optimal mild equilibrium by removing "inessential" parts. In particular, we show that the smallest mild equilibrium is strong under a mild assumption. Finally, three examples are provided in Section 6. The first example shows that a weak equilibrium may not be strong, while the second example shows that a strong equilibrium may not be optimal mild. The final example is about finding equilibria for the stopping problem of an American put option, which is used to demonstrate the usefulness of the results in Section 5. Figure 1 summarizes relations between the results in this paper.

2 | SETUP AND SOME AUXILIARY RESULTS

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, which supports a standard Brownian motion $W = (W_t)_{t\geq 0}$. Let $X = (X_t)_{t\geq 0}$ be a one-dimensional diffusion process with the dynamics

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \tag{9}$$

and take values in an interval $\mathbb{X} \subset \mathbb{R}$. Let $\mathbb{P}^x$ be the probability measure given $X_0 = x$ and denote $\mathbb{E}^x[\cdot] = \mathbb{E}[\cdot|X_0 = x]$. Let L_t^x be the local time of X at point x up to time t. Denote by $\mathcal{T}$ the set of stopping times.

Let $\mathcal{B}$ be the family of all Borel subsets within $\mathbb{X}$. For any $A \in \mathcal{B}$, denote $A^c := \mathbb{X} \setminus A$ and $\partial A := \overline{A} \setminus A^\circ$, where A° is the interior of A and $\overline{A}$ is the closure of A under the Euclidean topology within $\mathbb{X}$. Denote $B(x, r) := (x - r, x + r) \cap \mathbb{X}$. For $A \in \mathcal{B}$, we define the first hitting and exit times

$$\rho_A := \inf\{t > 0 : X_t \in A\} \quad \text{and} \quad \tau_A := \inf\{t > 0 : X_t \notin A\} = \rho_{A^c}. \tag{10}$$

Given a stopping region $A \in \mathcal{B}$, we define the value function $V(t, x, A) : [0, \infty) \times \mathbb{X} \to \mathbb{R}$ as

$$V(t, x, A) := \mathbb{E}^x[\delta(t + \rho_A) f(X_{\rho_A})]. \tag{11}$$

Recall function J defined in Equation (4), we have $J(x, A) = V(0, x, A)$.

Denote $\mathbb{N} := \{1, 2, \ldots\}$ and $\mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}$. Given $E \in \mathcal{B}$ and $k \in \mathbb{N} \cup \{0\}$, denote by $C^{1,k}([0, \infty) \times E)$ the family of functions $v(t, x)$ that are continuously differentiable with respect to (w.r.t.) t and k-times continuously differentiable w.r.t. x when restricted to $[0, \infty) \times E$, and $C^k(E)$ the family of functions $v(x)$ that are k-times continuously differentiable when restricted to E.[2] For a function $v(t, x) : [0, \infty) \times E \to \mathbb{R}$, v_x, v_{xx} (resp. v_t) denote the first- and second-order derivatives w.r.t x (resp. the first-order derivative w.r.t. t) if the derivatives exist. Moreover, denote by $v_x(t, x-)$ (resp. $v_x(t, x+)$) the left (resp. right) derivative of v w.r.t. x at point (t, x). Similar notation applies to $v_{xx}(t, x-), v_{xx}(t, x+)$. For convenience, we denote $v_t(0, x)$ as the right derivative w.r.t. t at time $t = 0$. We further define the parabolic operator

$$\mathcal{L}v(t, x) := v_t(t, x) + \mu(x) v_x(t, x) + \frac{1}{2}\sigma^2(x) v_{xx}(t, x) \quad \text{for a function } v \in C^{1,2}([0, \infty) \times E).$$

Let us also use the following notation involving left or right derivatives w.r.t. x:

$$\mathcal{L}v(t, x\pm) := v_t(t, x) + \mu(x) v_x(t, x\pm) + \frac{1}{2}\sigma^2(x) v_{xx}(t, x\pm), \quad \forall t \geq 0.$$

We now introduce the main assumptions in this paper. The first assumption concerns μ and σ.

Assumption 2.1.

(i) $\mu, \sigma : \mathbb{X} \to \mathbb{R}$ are Lipschitz continuous.
(ii) $\sigma^2(x) > 0$ for all $x \in \mathbb{X}$.

Remark 2.2. Assumption 2.1(i) guarantees that Equation (9) has a unique strong solution given $X_0 = x \in \mathbb{X}$. Assumption 2.1(i)(ii) together imply that for any $x \in \mathbb{X}$ and $t > 0$,

$$\mathbb{P}^x\left(\min_{0 \leq s \leq t} X_s < x\right) = \mathbb{P}^x\left(\max_{0 \leq s \leq t} X_s > x\right) = 1, \text{ and thus } \rho_{\{x\}} = 0, \ \mathbb{P}^x\text{-a.s.}. \tag{12}$$

[2] Continuous differentiability is extended to the boundary in a natural way if the boundary is included in E. For example, given $[a, b] \subset \mathbb{X}$, we say $g \in C^{1,2}([0, \infty) \times [a, b])$, if $g = g_1$ on $[0, \infty) \times [a, b]$ for some $g_1 \in C^{1,2}([0, \infty) \times \mathbb{X})$.

A quick proof for Equation (12) is relegated in Appendix A.

Notice that a (time-homogeneous Markovian) stopping policy can be characterized by a stopping region $S \subset \mathbb{X}$. For $S \in \mathcal{B}$, Equation (12) implies that $\rho_S = \rho_{\overline{S}}$ $\mathbb{P}^x$-a.s. for any $x \in \mathbb{X}$. Also, $f(x) = J(x, S)$ for all $x \in \overline{S}$, and a boundary point x of S corresponds to the action "immediate stopping," not matter x belongs to S or not. Therefore, it suffices to work on stopping regions that are closed.

Definition 2.3. $S \in \mathcal{B}$ is called an admissible stopping policy, if S is closed (w.r.t the Euclidean topology within $\mathbb{X}$) and for any $x \in \partial S$, one the following two cases holds:

(a) $x \in \partial(S^\circ)$, that is, $\exists h > 0$ such that either $(x-h, x) \subset S^\circ$ and $(x, x+h) \subset S^c$, or $(x-h, x) \subset S^c$ and $(x, x+h) \subset S^\circ$;
(b) x is an isolated point, that is, $B(x, h) \setminus \{x\} \subset S^c$ for some $h > 0$.

Remark 2.4. Except cases (a) and (b), the rest situation for a boundary point $x \in \partial S$ is the following:

(c) There exist two sequences $(x_n)_{n\in\mathbb{N}} \subset S$ and $(y_n)_{n\in\mathbb{N}} \subset S^c$ such that both $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ approach to x from the left, or both approach to x from the right.[3]

Stopping regions containing boundary case (c) lack economic meaning, since it is not practical for an agent to follow a stopping policy classified as case (c). Mathematically, the regularity of $V(t, x, S)$ may also be missing when S contains boundary case (c), for example, $V_x(t, 0+, S)$ may not exist for S being the cantor set on [0,1]; this would cause serious issue to establish our main results later as they crucially rely on the regularity of $V(t, x, S)$ (e.g., the characterization of weak equilibria).

Focusing on admissible stopping policies is also well-aligned with the literature, and a stopping policy containing boundary case (c) is rarely studied in applications. For instance, all the case studies in Bodnariu et al. (2022); Tan et al. (2021); Christensen and Lindensjö (2018); Ebert et al. (2020); Huang et al. (2020) only focus on threshold-type equilibria. The results in Ebert and Strack (2018) mainly focus on two threshold stopping regions. The mild equilibria in all the examples of Huang and Nguyen-Huu (2018) have boundaries of cases only (a) and (b). All mild equilibria provided in Huang and Zhou (2020, Sections 6.1 and 6.2) are all admissible, so are the mild equilibria in the case study of Huang & Yu (2021, Section 4).

Let us also point out that all the interesting equilibria (i.e., optimal mild, weak, strong equilibria) provided in all the examples in this paper are admissible. Specifically, in the case study in Section 6.3, which can be thought of as a continuation of Huang and Zhou (2020, Sections 6.3), all the weak, strong, and optimal mild equilibria are admissible, and any mild equilibria is either admissible or has an admissible alternative (see Remark 6.6).

To sum up, focusing on cases (a) and (b) is economically meaningful, mathematically necessary, well-aligned with the literature, and general enough for applications.

[3] Indeed, let $x \in \partial S$. Suppose x does not satisfy case (c). Then there exists $h > 0$ such that either $(x-h, x) \subset S$ or $(x-h, x) \subset S^c$, so is the interval $(x, x+h)$. If $(x-h, x) \subset S^c$ and $(x, x+h) \subset S^c$, then x satisfies case (b); otherwise x satisfies case (a).

Let $\delta(\cdot) : [0, \infty) \to [0, 1]$ be a discount function that is nonincreasing, continuously differentiable, and $\delta(0) = 1$, $\delta(t) < 1$ for $t > 0$. We assume that δ satisfies the following condition.

Assumption 2.5. δ is log subadditive:

$$\delta(t+s) \geq \delta(t)\delta(s), \quad \forall s, t \geq 0. \tag{13}$$

Remark 2.6. Condition (13) can be interpreted as the so-called decreasing impatience in finance and economics. Many nonexponential discount functions, including hyperbolic, generalized hyperbolic, and pseudo-exponential discounting, satisfy Equation (13). See the discussion below (Huang and Nguyen-Huu, 2018, Assumption 3.12) for a more detailed explanation.

Recall that $\delta'(0)$ denotes the right derivative of $\delta(t)$ at $t = 0$. The following lemma is a quick result for δ and the proof is relegated in the Appendix A.

Lemma 2.7. *Let Assumption 2.5 hold. Then,*

$$\delta'(t) \geq \delta(t)\delta'(0), \quad \textit{and} \quad 1 - \delta(t) \leq |\delta'(0)|t, \quad \forall t \geq 0.$$

Let the payoff function $f(x) : \mathbb{X} \to \mathbb{R}$ be non-negative and continuous. We further assume that f satisfies the following assumptions.

Assumption 2.8. (i) For any $x \in \mathbb{X}$,

$$\lim_{t\to\infty} \delta(t) f(X_t) = 0, \quad \mathbb{P}^x - a.s., \tag{14}$$

and there exists $\zeta > 0$ such that

$$\mathbb{E}^x\left[\sup_{t\geq 0} \left(\delta(t) f(X_t)\right)^{1+\zeta}\right] < \infty. \tag{15}$$

(ii) $f(x)$ belongs to C^2 piecewisely. That is, there exists an either finite or countable set $(\theta_n)_{n\in I} \subset \mathbb{X}$, with $I \subset \mathbb{Z}$ and $\theta_n < \theta_{n+1}$ for all $n \in I$, such that $f \in C^2([\theta_n, \theta_{n+1}])$ for any $n \in I$. We also assume that $\inf_{n\in I}(\theta_{n+1} - \theta_n) > 0$ and denote

$$\mathcal{G} := \mathbb{X} \setminus \{\theta_n : n \in I\}. \tag{16}$$

Remark 2.9. The assumption (15) will be used for Lemma 3.8, which is an essential lemma for all the main results in the paper. Moreover, Equation (15) implies that

$$\mathbb{E}^x\left[\sup_{t\geq 0} \delta(t) f(X_t)\right] < \infty, \quad \forall x \in \mathbb{X}. \tag{17}$$

This together with Equation (14) guarantees the well-posedness of $V(t,x,S)$ for any stopping policy S.[4] Equations (17) and (14) will also be used for applying the dominated convergence theorem in some localization arguments in the proofs later. Furthermore, Equations (17) and (14) also ensure the existence of an optimal mild equilibrium as demonstrated in Huang and Zhou (2020, Theorem 4.12) (also see Lemma 4.1 in this paper).

Let us make an assumption on $V(t,x,S)$.

Assumption 2.10. For any admissible stopping policy S and $a, b \in \mathbb{X}$ with $a < b$ and $(a,b) \subset S^c$, $V(t,x,S)$ defined in Equation (11) (with $A = S$) belongs to $C^{1,2}([0,\infty) \times [a,b])$, and

$$\limsup_{t \searrow 0} \frac{1}{\sqrt{t}} |V_x(t, x\pm, S) - V_x(0, x\pm, S)| = 0, \quad \forall\, x \in \mathbb{X}. \tag{18}$$

Remark 2.11. It turns out that Assumption 2.10 is quite general. A sufficient condition for Assumption 2.10 is that $\delta(t)$ is a weighted discount function as shown in the lemma below. One may also directly verify this assumption given the probability density functions of exit time

$$p(x,t) := \mathbb{P}^x\left(\tau_{(c,d)} \in dt, X_{\tau_{(c,d)}} = c\right) \quad \text{and} \quad q(x,t) := \mathbb{P}^x\left(\tau_{(c,d)} \in dt, X_{\tau_{(c,d)}} = d\right) \tag{19}$$

being regular enough. For example, if X is a Brownian motion on $\mathbb{X} = \mathbb{R}$, $\delta(t) = \frac{1}{1+t}$, and $f(x) = 0 \vee x$, then we can verify that Assumption 2.10 holds by using Equation (19) for the Brownian motion. Providing a more general sufficient condition for Assumption 2.10 is out of the scope of this paper.

Lemma 2.12. *Let Assumption 2.1 hold and f be bounded on $\mathbb{X}$. Suppose $\delta(t)$ is a* weighted discount function *of the following form:*

$$\delta(t) = \int_0^\infty e^{-rt} dF(r), \tag{20}$$

where $F(r) : [0,\infty) \to [0,1]$ is a cumulative distribution function satisfying $\int_0^\infty r dF(r) < \infty$ and

$$\lim_{t \searrow 0} \frac{1}{\sqrt{t}} \int_0^\infty r(1 - e^{-rt}) dF(r) = 0. \tag{21}$$

Then, Assumption 2.10 holds.

The proof of Lemma 2.12 is included in Appendix A.

Remark 2.13. In Ebert et al. (2020), weighted discount functions are studied in detail. Tan et al. (2021) investigates weak equilibria and the smooth-fit condition for time-inconsistent stopping in a weighted discounting setting. Many discount functions, including exponential, hyperbolic,

[4] Equation (14) is used for the well-posedness of $V(t,x,S)$ because otherwise $\delta(t)f(X_t)$ is not well-defined on $\rho_S = \infty$ (unless we do some extension, e.g., by considering the upper limit $\limsup_{t\to\infty} \delta(t)f(X_t)$).

generalized hyperbolic, and pseudo-exponential discounting, satisfy Equations (20) and (21). For example, a generalized hyperbolic discount function can be written as

$$\delta(t)=\frac{1}{(1+\beta t)^{\frac{\gamma}{\beta}}}=\int_0^\infty e^{-rt}\frac{r^{\frac{\gamma}{\beta}-1}e^{-\frac{r}{\beta}}}{\beta^{\frac{\gamma}{\beta}}\Gamma(\frac{\gamma}{\beta})}dr=\int_0^\infty e^{-rt}dF(r),\quad \text{with } \frac{dF(r)}{dr}=\frac{r^{\frac{\gamma}{\beta}-1}e^{-\frac{r}{\beta}}}{\beta^{\frac{\gamma}{\beta}}\Gamma(\frac{\gamma}{\beta})},$$

where $\beta,\gamma>0$ are constants and $\Gamma(\cdot)$ is the gamma function (see Tan et al. (2021, Section 2.1)). A direct calculation shows that

$$\int_0^\infty r(1-e^{-rt})dF(r)=\int_0^\infty r(1-e^{-rt})\frac{r^{\frac{\gamma}{\beta}-1}e^{-\frac{r}{\beta}}}{\beta^{\frac{\gamma}{\beta}}\Gamma(\frac{\gamma}{\beta})}dr=\gamma-\gamma\frac{1}{(1+\beta t)^{\frac{\gamma}{\beta}+1}}\le\gamma(\gamma+\beta)t\quad\forall t>0,$$

which implies Equation (21).

The next lemma summarizes several preliminary properties of $V(t,x,S)$, which will be used to establish the main results in later sections.

Lemma 2.14. *Let Assumptions 2.1, 2.5, 2.8(ii), 2.10 hold and S be an admissible stopping policy. Then,*

(a) $V(t,x,S)$ belongs to $C^{1,2}([0,\infty)\times\overline{S^c})$, and $V(t,x,S)=\delta(t)f(x)$ for any $(t,x)\in[0,\infty)\times S$. Moreover,

$$\mathcal{L}V(t,x,S)\equiv 0,\quad \forall (t,x)\in[0,\infty)\times S^c. \tag{22}$$

(b) $\mathcal{L}V(t,x\pm,S)$ exists for all $(t,x)\in[0,\infty)\times\mathbb{X}$. For any $h>0$ and $x_0\in\mathbb{X}$ such that $\overline{B(x_0,h)}\subset\mathbb{X}$, we have that

$$\sup_{(t,x)\in[0,\infty)\times\overline{B(x_0,h)}}|\mathcal{L}V(t,x\pm,S)|<\infty.$$

The proof of Lemma 2.14 is provided in Appendix A. Throughout this paper, we will keep using the following local time integral formula provided in Peskir (2007).

Lemma 2.15. *Let $a,x_0,b\in\mathbb{R}$ with $a<x_0<b$. Suppose $g(t,y):[0,\infty)\times\mathbb{R}\to\mathbb{R}$ such that $g\in C^{1,2}((0,\infty)\times(a,x_0])$, $g\in C^{1,2}((0,\infty)\times[x_0,b))$. Then, for $X_0=x\in(a,b)$, we have that*

$$g(t,X_t)=g(0,x)+\int_0^t\frac{1}{2}(\mathcal{L}g(s,X_s-)+\mathcal{L}g(s,X_s+))ds+\int_0^t g_x(s,X_s)\sigma(X_s)\cdot 1_{\{X_s\neq x_0\}}dW_s$$
$$+\frac{1}{2}\int_0^t(g_x(s,x_0+)-g_x(s,x_0-))dL_s^{x_0}\quad\forall 0\le t\le\tau_{(a,b)}.$$

3 | CHARACTERIZATION FOR WEAK EQUILIBRIA

In this section, we provide the characterization for weak equilibria. Such characterization is critical to study of the relations between mild, weak, and strong equilibria. Below is the main result of this section.

Theorem 3.1. *Let Assumptions 2.1–2.10 hold. Suppose S is an admissible stopping policy. Then S is a weak equilibrium if and only if the followings are satisfied.*

$$V(0,x,S) \geq f(x) \quad \forall x \notin S; \tag{23}$$

$$V_x(0,x-,S) \geq V_x(0,x+,S) \quad \forall x \in S; \tag{24}$$

$$\mathcal{L}V(0,x-,S) \vee \mathcal{L}V(0,x+,S) \leq 0 \quad \forall x \in \mathbb{X}. \tag{25}$$

The proof of Theorem 3.1 will be presented in the next subsection. A consequence of Theorem 3.11 is the following smooth-fit condition of V at the boundary ∂S when f is smooth.

Corollary 3.2 (Smooth-fit condition for weak equilibria when f is smooth). *Let Assumptions 2.1–2.10 hold, and let S be an admissible stopping policy. Suppose S is a weak equilibrium. Then for any $x \in \partial S$, if $f'(x)$ exists, then $V_x(0,x-,S) = V_x(0,x+,S)$.*

Proof. Take an arbitrary $x \in \partial S$. Take $x \in \partial S$. By Theorem 3.11, it suffices to prove that $V_x(0,x-,S) \leq V_x(0,x+,S)$ for both boundary cases (a) and (b).

Recall $\mathcal{G}$ defined in Equation (16). For boundary case (a), without loss of generality, we assume $(x, x+h) \subset (S^\circ \cap \mathcal{G})$ and $(x-h, x) \subset S^c$ for some $h > 0$. Since $V(0,x,S) \geq f(x)$ on S^c by Equation (5) and $V(0,x,S) = f(x)$ on S by Lemma 2.14 (a), we have that for $\varepsilon > 0$ small enough,

$$\frac{V(0,x-\varepsilon,S) - V(0,x,S)}{\varepsilon} \geq \frac{f(x-\varepsilon) - f(x)}{\varepsilon}.$$

By the differentiability of V on $[0,\infty) \times \overline{S^c}$ (due to Lemma 2.14(a)) and existence of $f'(x)$, the above inequalities implies that

$$V_x(0,x-,S) \leq f'(x-) = f'(x+) = V_x(0,x+,S),$$

where the last equality follows from $V(0,x,S) = f(x)$ on $(x, x+h) \subset (S^\circ \cap \mathcal{G})$.

For boundary case (b), we can choose a constant $h > 0$ such that $(B(x,h) \setminus \{x\}) \subset (S^c \cap \mathcal{G})$. Then $V(0,y,S) \geq f(y)$ for all $y \in B(x,h) \setminus \{x\}$, which implies that

$$V_x(0,x-,S) \leq f'(x-) = f'(x+) \leq V_x(0,x+,S)$$

by an argument similar to that for boundary case (a). □

Remark 3.3. In Tan et al. (2021), it is shown that with the underlying process being a geometric Brownian motion, the smooth-fit condition together with some inequalities provides a weak equilibrium; in addition, the real options example in Tan et al. (2021) indicates that when smooth-fit condition fails, there is no weak equilibrium. This, however, does not indicate whether any weak equilibrium must satisfy the smooth-fit condition. Here, we are able to provide a much sharper result in a much more general setting: given f is smooth, any weak equilibrium must satisfy the smooth-fit condition, and may be constructed by the smooth-fit condition together with some other related inequalities. Let us also mention that smooth-fit result is also established in a very recent paper (Bodnariu et al., 2022) for mixed weak equilibrium under a general setting.

Remark 3.4. In our paper, the payoff function f is only required to be piecewisely smooth. The inequality in Equation (24) and Corollary 3.2 show that the smooth-fit condition is a specially case of the "local convexity" property for a weak equilibrium S: the left derivative w.r.t. x of the value function $V(0,x,S)$ must be bigger than or equal to its right derivative for *any* $x \in S$. In particular, if the payoff function is smooth at a point $x \in S$, such convexity property is reduced to the smooth-fit condition.

Remark 3.5. Suppose the discount function is exponential in the current one-dimensional diffusion context. Then Equations (23) and (25) together yield the variational inequalities. As is well known in classical optimal stopping theory, (under suitable assumptions) the optimal stopping value and strategy can be characterized by variational inequalities. Therefore, when the discount function is exponential, Theorem 3.1 indicates that any weak equilibrium is an optimal stopping region in the classical sense, so are strong and optimal mild equilibrium (as we will show later that an optimal mild equilibrium is also weak). On the other hand, a mild equilibrium is not necessarily a classical optimal stopping region, for example, the whole state space $\mathbb{X}$ is a mild equilibrium but may not be an optimal stopping region in general.

3.1 | Proof of Theorem 3.1

To characterize a weak equilibrium, one shall consider the two conditions (5) and (6) in Definition 1.3. Equation (5) is the same as Equation (23) and thus we will focus on condition (6). By V defined in Equation (11), Equation (6) can be rewritten as

$$\limsup_{\varepsilon \searrow 0} \frac{1}{\varepsilon}\left(\mathbb{E}^x[\delta(\rho_S^\varepsilon)f(X_{\rho_S^\varepsilon})] - f(x)\right) = \limsup_{\varepsilon \searrow 0} \frac{1}{\varepsilon}(\mathbb{E}^x[V(\varepsilon, X_\varepsilon, S)] - V(0,x,S)) \le 0, \quad x \in S. \tag{26}$$

Since X_ε and thus $V(\varepsilon, X_\varepsilon, S)$ are not uniformly bounded, we will apply some localization argument and restrict X within a bounded ball $B(x,h)$. Moreover, as $x \mapsto V(t,\cdot,S)$ is only piecewisely smooth, we will choose $h > 0$ small enough, such that V_x is only (possibly) discontinuous at the center of the ball $B(x,h)$, in order to apply Lemma 2.15 to V in Equation (26). By doing so, we will end up with

$$\mathbb{E}^x[V(\varepsilon, X_\varepsilon, S) - V(0,x,S)] \approx \mathbb{E}^x[V(\varepsilon \wedge \tau_{B(x,h)}, X_{\varepsilon\wedge\tau_{B(x,h)}}) - V(0,x,S)] \tag{27}$$

$$= \mathbb{E}^x\left[\int_0^{\varepsilon\wedge\tau_{B(x,h)}} \frac{1}{2}(\mathcal{L}V(s,X_s-,S) + \mathcal{L}V(s,X_s+,S))ds\right]$$

$$+ \mathbb{E}^x\left[\frac{1}{2}\int_0^{\varepsilon\wedge\tau_{B(x,h)}} (V_x(s,x+,S) - V_x(s,x-,S))dL_s^x\right]. \tag{28}$$

where the approximation in Equation (27) will be made rigorous in Lemma 3.8, which is built on Lemma 3.6, and Equation (28) is due to Lemma 2.15. Then the condition (26) boils down to comparing the two integral terms on the right-hand side (RHS) of Equation (28). This requires estimates for the expected local time $\mathbb{E}^x[L^x_{\varepsilon\wedge\tau_{B(x,h)}}]$ (see Lemma 3.9, which is built upon Lemmas 3.6

and 3.7) and the growth of $V_x(t, x\pm, S)$ w.r.t. t (see Lemma 3.10). This is the overall idea on how we obtain Theorem 3.1.

Throughout this section, we shall also take advantage of the following standard estimate for moments of diffusions (see, e.g., Karatzas and Shreve (1991, Problem 3.15 on page 306)): Given a process Z_t satisfying $dZ_t = \beta(Z_t)dt + \delta(Z_t)dW_t$ with β, δ being Lipschitz and $Z_0 = z \in \mathbb{X}$, for all $0 \le \varepsilon \le 1$ and $m \ge 1$, it holds that

$$\mathbb{E}^x\left[\sup_{0\le s\le t} |Z_s|^{2m}\right] < \infty \quad \forall t \in (0,\infty), \tag{29}$$

$$\mathbb{E}^x\left[|Z_\varepsilon - z|^{2m}\right] \le K(1+|z|^{2m})\varepsilon^m, \tag{30}$$

where K is a constant independent of ε.

We first provide two Lemmas dealing with the probability of X exiting a ball, and the first-order moment related to X over a small time horizon ε. They will be used for proofs in both the current and later sections.

Lemma 3.6. *Let Assumption 2.1 hold. For any fixed $a > 0$ we have that*

$$\mathbb{P}^x(\tau_{B(x,h)} \le \varepsilon) = o(\varepsilon^a), \quad \textit{for } \varepsilon > 0 \textit{ small enough.} \tag{31}$$

Proof. Fix $a > 0$. We invoke the "change of space" method in Peskir and Shiryaev (2006, Section 5.2). Consider the process

$$Y_t := \phi(X_t), Y_0 := \phi(x) \text{ with } \phi(y) := \int_0^y \exp\left(-\int_0^l \frac{2\mu(z)}{\sigma^2(z)}dz\right)dl. \tag{32}$$

Thanks to Assumption 2.1, ϕ is well-defined, strictly increasing, and has first and second derivatives. A direct calculation shows that $dY_t = \sigma(X_t)\phi'(X_t)dW_t$, and the exit time to $\overline{B(x,h)}$ of X_t is equivalent to the exit time of Y_t to the interval $[\phi(x-h), \phi(x+h)]$. Set $\tilde{h} := (\phi(x+h) - \phi(x)) \wedge (\phi(x) - \phi(x-h)) > 0$ and $\tilde{a} := a + 1$. Let $0 < \varepsilon \le 1$. We have that

$$\mathbb{P}^x(\tau_{B(x,h)} \le \varepsilon) \le \mathbb{P}^{Y_0}\left(\sup_{0\le t\le\varepsilon} |Y_t - Y_0| \ge \tilde{h}\right) = \mathbb{P}^{Y_0}\left(\sup_{0\le t\le\varepsilon} |Y_t - Y_0|^{2\tilde{a}} \ge \tilde{h}^{2\tilde{a}}\right). \tag{33}$$

Notice that Y is a martigale (within the interval $\overline{B(\phi(x),\tilde{h})}$), we can then apply the Doob's submartingale inequality to the RHS of Equation (33) to conclude that

$$\mathbb{P}^{Y_0}\left(\sup_{0\le t\le\varepsilon} |Y_t - Y_0|^{2\tilde{a}} \ge \tilde{h}^{2\tilde{a}}\right) \le \frac{\mathbb{E}^{Y_0}[|Y_\varepsilon - Y_0|^{2\tilde{a}}]}{\tilde{h}^{2\tilde{a}}} \le \frac{\tilde{K}(1+\phi^{2\tilde{a}}(x))\varepsilon^{\tilde{a}}}{\tilde{h}^{2\tilde{a}}}, \tag{34}$$

where the last inequality follows from Equation (30), and $\tilde{K}$ is a positive constant independent of ε. Then Equation (31) follows from Equations (33) and (34) and the fact that $\tilde{a} > a$. □

Lemma 3.7. *Let Assumption 2.1(i) hold. For $\varepsilon > 0$ small enough we have that*

$$\mathbb{E}^x[|\bar{X}_\varepsilon|] = O(\varepsilon), \quad \textit{with } \widetilde{X}_t := x + \mu(x)t + \sigma(x)W_t \textit{ and } \bar{X}_t := X_t - \widetilde{X}_t. \tag{35}$$

Proof. Throughout the proof, C will serve as a generic constant may change from line to line but is independent of ε. Let $0 < \varepsilon \leq 1$. First, we have

$$\mathbb{E}^x|\bar{X}_\varepsilon| \leq \mathbb{E}^x\left|\int_0^\varepsilon (\mu(X_t) - \mu(x))dt\right| + \mathbb{E}^x\left|\int_0^\varepsilon (\sigma(X_t) - \sigma(x))dW_t\right|. \tag{36}$$

By applying Equation (30) on X_t with $m = 1$, we have $\mathbb{E}^x[|X_\varepsilon - x|^2] \leq C\varepsilon$. This together with the Lipschitz continuity of μ implies

$$\begin{aligned}\mathbb{E}^x\left|\int_0^\varepsilon (\mu(X_t) - \mu(x))dt\right| &\leq \mathbb{E}^x\left[\int_0^\varepsilon \frac{1}{2}(1 + |\mu(X_t) - \mu(x)|^2)dt\right] \\ &\leq \frac{1}{2}\varepsilon + \frac{1}{2}\int_0^\varepsilon C\mathbb{E}^x[|X_t - x|^2]dt = O(\varepsilon).\end{aligned} \tag{37}$$

Similarly, we can estimate the second term in Equation (36) as follows:

$$\begin{aligned}\mathbb{E}^x\left|\int_0^\varepsilon (\sigma(X_t) - \sigma(x))dW_t\right| &\leq \left(\mathbb{E}^x\left[\int_0^\varepsilon (\sigma(X_t) - \sigma(x))^2 dt\right]\right)^{1/2} \\ &\leq \left(\int_0^\varepsilon C\mathbb{E}^x[|X_t - x|^2]dt\right)^{1/2} = O(\varepsilon)\end{aligned} \tag{38}$$

Then, by plugging Equations (37) and (38) into Equation (36), we have $\mathbb{E}^x[|\bar{X}_\varepsilon|] = O(\varepsilon)$. □

The next lemma concerns the approximation in Equation (27).

Lemma 3.8. *Let Assumptions 2.1 and 2.8(i) hold. Let $S \in \mathcal{B}$, $x \in \mathbb{X}$ and $h > 0$. Then, for $\varepsilon > 0$ small enough,*

$$\mathbb{E}^x[V(\varepsilon, X_\varepsilon, S)] = \mathbb{E}^x[V(\varepsilon \wedge \tau_{B(x,h)}, X_{\varepsilon\wedge\tau_{B(x,h)}}, S)] + o(\varepsilon). \tag{39}$$

Proof. Let $h > 0$ and $x \in \mathbb{X}$. Recall the constant ζ in Equation (15). We have that

$$\begin{aligned}0 \leq& \mathbb{E}^x[V(\varepsilon, X_\varepsilon, S) \cdot 1_{\{\varepsilon > \tau_{B(x,h)}\}}] \\ \leq& \left(\mathbb{E}^x\left[V^{1+\zeta}(\varepsilon, X_\varepsilon, S)\right]\right)^{\frac{1}{1+\zeta}} \cdot \left(\mathbb{E}^x\left[1^{\frac{1+\zeta}{\zeta}}_{\{\varepsilon > \tau_{B(x,h)}\}}\right]\right)^{\frac{\zeta}{1+\zeta}} \\ \leq& \left(\mathbb{E}^x\left[\sup_{t\geq 0}(\delta(t)f(X_t))^{1+\zeta}\right]\right)^{\frac{1}{1+\zeta}} \cdot \left(\mathbb{E}^x\left[1_{\{\varepsilon > \tau_{B(x,h)}\}}\right]\right)^{\frac{\zeta}{1+\zeta}} \\ \leq& O(1) \cdot \left(\mathbb{P}^x\left(\tau_{B(x,h)} \leq \varepsilon\right)\right)^{\frac{\zeta}{1+\zeta}},\end{aligned} \tag{40}$$

where the first inequality follows from $f \geq 0$, the second inequality follows from Hölder's inequality, the third inequality follows from Jensen's inequality, and the last inequality follows from

Equation (15). Applying Lemma 3.6 with $a = \frac{1+\zeta}{\zeta}$ to Equation (40), we have

$$\mathbb{E}^x\left[V(\varepsilon, X_\varepsilon, S) \cdot 1_{\{\varepsilon > \tau_{B(x,h)}\}}\right] = o(\varepsilon). \tag{41}$$

Similarly, we can show that

$$\mathbb{E}^x\left[V(\varepsilon \wedge \tau_{B(x,h)}, X_{\varepsilon \wedge \tau_{B(x,h)}}, S) \cdot 1_{\{\varepsilon > \tau_{B(x,h)}\}}\right] = o(\varepsilon).$$

This together with Equation (41) implies Equation (39). □

Recall that L_t^x is the local time of X at position x up to time t. We have the following result.

Lemma 3.9. *Let Assumption 2.1 hold. Then, for any $x \in \mathbb{X}$ and $h > 0$,*

$$\lim_{\varepsilon \searrow 0} \frac{\mathbb{E}^x\left[L^x_{\varepsilon \wedge \tau_{B(x,h)}}\right]}{\sqrt{\varepsilon}} = \sqrt{\frac{2}{\pi}} \cdot |\sigma(x)|. \tag{42}$$

Proof. Let $h > 0$ and $x \in \mathbb{X}$. Thanks to Assumption 2.1(i) and (29) (with $m = \frac{p+2}{2} > 1$), it holds for any $p, t > 0$ that

$$\mathbb{E}^x\left[\sup_{0 \le s \le t} |X_s|^p\right] \le 1 + \mathbb{E}^x\left[\sup_{0 \le s \le t} |X_s|^{p+2}\right] < \infty.$$

This enables us to apply an argument similar to the proof of Lemma 3.8 and get that

$$\mathbb{E}^x[|X_\varepsilon - x|] + o(\varepsilon) = \mathbb{E}^x\left[|X_{\varepsilon \wedge \tau_{B(x,h)}} - x|\right]. \tag{43}$$

Applying Lemma 2.15 on $[0, \varepsilon \wedge \tau_{B(x,h)}]$ with $g(t, y) := |y - x|$ and then taking expectation, and using Equation (43), we have that

$$\mathbb{E}^x[|X_\varepsilon - x|] + o(\varepsilon) = \mathbb{E}^x\left[\int_0^{\varepsilon \wedge \tau_{B(x,h)}} \operatorname{sgn}(X_s - x)\mu(X_s)ds\right] + \mathbb{E}^x\left[L^x_{\varepsilon \wedge \tau_{B(x,h)}}\right]. \tag{44}$$

By Assumption 2.1(i), the first term on the RHS of Equation (44) can be estimated as follows:

$$\left|\mathbb{E}^x\left[\int_0^{\varepsilon \wedge \tau_{B(x,h)}} \operatorname{sgn}(X_s - x)\mu(X_s)ds\right]\right| \le \sup_{y \in \overline{B(x,h)}} |\mu(y)|\varepsilon = O(\varepsilon). \tag{45}$$

As for the left-hand side (LHS) of Equation (44), by Lemma 3.7 we have that

$$\begin{aligned}\mathbb{E}^x[|X_\varepsilon - x|] &= \mathbb{E}^x[|\widetilde{X}_\varepsilon - x|] + O(\varepsilon) = \mathbb{E}^x[|\mu(x)\varepsilon + \sigma(x)W_\varepsilon|] + O(\varepsilon) \\ &= |\sigma(x)|\mathbb{E}[|W_\varepsilon|] + O(\varepsilon) = |\sigma(x)|\mathbb{E}[|W_1|]\sqrt{\varepsilon} + O(\varepsilon) = |\sigma(x)|\sqrt{\frac{2}{\pi}}\sqrt{\varepsilon} + O(\varepsilon). \end{aligned} \tag{46}$$

Then Equation (42) follows from plugging Equations (45) and (46) into Equation (44). □

Lemma 3.10. *Let Assumptions 2.1, 2.5, 2.10 hold. Let $S \in \mathcal{B}$ and $x \in \partial S$, and suppose $(x-h,x) \subset S^c$ (resp. $(x,x+h) \subset S^c$) for some $h>0$. Then,*

$$V_x(t,x-,S) \le \delta(t)V_x(0,x-,S) \quad (\textit{resp. } V_x(t,x+,S) \ge \delta(t)V_x(0,x+,S)). \tag{47}$$

Proof. Notice that Assumption 2.10 gives the existence of $V_x(t,y-,S)$ for $y \in (x-h,x]$ (resp. $V_x(t,y+,S)$ for $y \in [x,x+h)$) when $(x-h,x) \subset S^c$ (resp. when $(x,x+h) \subset S^c$). For any $y \in \mathbb{X}, t \ge 0$, by the non-negativity of f and Equation (13),

$$V(t,y,S) = \mathbb{E}^y[\delta(t+\rho_S)f(X_{\rho_S})] \ge \delta(t)\mathbb{E}^y[\delta(\rho_S)f(X_{\rho_S})] = \delta(t)V(0,y,S).$$

Suppose $(x-h,x) \subset S^c$. Then by the fact that $V(t,x,S) = \delta(t)f(x)$ (due to Lemma 2.14 (a)) and the above inequality, we have that

$$\begin{aligned} V_x(t,x-,S) &= \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon}(V(t,x,S) - V(t,x-\varepsilon,S)) \\ &\le \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon}(\delta(t)(f(x) - V(0,x-\varepsilon,S))) = \delta(t)V_x(0,x-,S). \end{aligned}$$

Similar argument is applied for the result of $V_x(t,x+,S)$ when $(x,x+h) \subset S^c$. □

Lemmas 3.9 and 3.10 together indicate that, as long as $V_x(0,x+,S) - V_x(0,x-,S) \neq 0$, the local time integral is the dominating term on the RHS of Equation (28). Thus, to make the LHS of Equation (28) nonpositive in the limit, $V_x(0,x+,S) - V_x(0,x-,S)$ shall be nonpositive. Based on this and recalling Equation (26), we now prove the necessary conditions for a weak equilibrium in the following proposition. The sufficiency part follows next.

Proposition 3.11. *Let Assumptions 2.1–2.10 hold. Suppose S is an admissible stopping policy. If S is a weak equilibrium, then*

$$\begin{cases} V_x(0,x+,S) \le V_x(0,x-,S) & \forall x \in S, \\ \mathcal{L}V(0,x+,S) \vee \mathcal{L}V(0,x-,S) \le 0 & \forall x \in \mathbb{X}. \end{cases} \tag{48}$$

Proof. We verify the first inequality in Equation (48) by contradiction. Take $x \in S$ and suppose

$$a := V_x(0,x+,S) - V_x(0,x-,S) > 0. \tag{49}$$

Recall $\mathcal{G}$ defined in Equation (16). Choose $h>0$ such that $(x-h,x) \cup (x,x+h)$ is contained in $(\mathcal{G} \cap S^\circ) \cup S^c$. By Lemma 2.14(a) and Assumption 2.8(ii), $V \in C^{1,2}([0,\infty) \times (x-h,x])$ and $V \in C^{1,2}([0,\infty) \times [x,x+h))$. Then, we can apply Lemma 2.15 to get

$$\begin{aligned} V(\varepsilon \wedge \tau_{B(x,h)}, X_{\varepsilon \wedge \tau_{B(x,h)}}, S) - V(0,x,S) &= \int_0^{\varepsilon \wedge \tau_{B(x,h)}} \frac{1}{2}(\mathcal{L}V(s,X_s-,S) + \mathcal{L}V(s,X_s+,S))ds \\ &\quad + \int_0^{\varepsilon \wedge \tau_{B(x,h)}} V_x(s,X_s,S)\sigma(X_s) \cdot 1_{\{X_s \neq x\}} dW_s \\ &\quad + \frac{1}{2}\int_0^{\varepsilon \wedge \tau_{B(x,h)}} (V_x(s,x+,S) - V_x(s,x-,S))dL_s^x. \end{aligned} \tag{50}$$

Let $\varepsilon \in (0,1)$, notice that the diffusion integrand above is bounded on $[0,1] \times \overline{B(x,h)}$. Taking expectation on both sides of Equation (50) and then applying Lemma 3.8, we have that

$$\begin{aligned}\mathbb{E}^x[V(\varepsilon, X_\varepsilon, S) - V(0,x,S)] = \mathbb{E}^x\left[\int_0^{\varepsilon\wedge\tau_{B(x,h)}} \frac{1}{2}(\mathcal{L}V(s,X_s-,S) + \mathcal{L}V(s,X_s+,S))ds\right] \\ + \mathbb{E}^x\left[\frac{1}{2}\int_0^{\varepsilon\wedge\tau_{B(x,h)}} (V_x(s,x+,S) - V_x(s,x-,S))dL_s^x\right] + o(\varepsilon).\end{aligned} \tag{51}$$

By Lemma 3.10 and Equation (49),

$$V_x(t,x+,S) - V_x(t,x-,S) \geq \delta(t)(V_x(0,x+,S) - V_x(0,x-,S)) = a\delta(t), \quad \forall t \geq 0.$$

By the above inequality and the continuity of δ, we can take $T > 0$ such that

$$V_x(s,x+,S) - V_x(s,x-,S) \geq \frac{a}{2}, \quad \forall s \in [0,T].$$

Then, for $\varepsilon \in [0, T\wedge 1]$, the second term on the RHS of Equation (51) can be estimated as follows:

$$\mathbb{E}^x\left[\frac{1}{2}\int_0^{\varepsilon\wedge\tau_{B(x,h)}} (V_x(s,x+,S) - V_x(s,x-,S))dL_s^x\right] \geq \frac{a}{4}\mathbb{E}^x[L^x_{\tau_{B(x,h)}\wedge\varepsilon}]. \tag{52}$$

By Lemma 2.14(b), we have

$$\sup_{(t,y)\in[0,1]\times\overline{B(x,h)}} |\mathcal{L}V(t,y-,S) + \mathcal{L}V(t,y+,S)| < \infty,$$

and thus the first term on the RHS of Equation (51) is of order $O(\varepsilon)$. Plugging this and Equation (52) into Equation (51) and then applying Lemma 3.9, we have

$$\liminf_{\varepsilon\searrow 0} \frac{1}{\varepsilon}\mathbb{E}^x[V(\varepsilon,X_\varepsilon,S) - V(0,x,S)] \geq O(1) + \frac{a}{4}\liminf_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\mathbb{E}^x[L^x_{\tau_{B(x,h)}\wedge\varepsilon}] = \infty,$$

which contradicts S being a weak equilibrium. Hence, $V_x(0,x+,S) - V_x(0,x-,S) \leq 0$.

Next, we verify the second inequality in Equation (48). Take $x \in \mathbb{X}$ and we consider three cases.

Case (i) $x \in S^c$. Lemma 2.14(a) shows that $\mathcal{L}V(0,x,S) = 0$.

Case (ii) $x \in \mathcal{G} \cap S^\circ$. Choose $h > 0$ such that $B(x,h) \subset \mathcal{G} \cap S^\circ$. Notice that $V(t,y,S) = \delta(t)f(y)$ for $y \in S^\circ$. Then, by Assumptions 2.1(i) and 2.8(ii), we have $V(t,y,S) \in C^{1,2}([0,\infty)\times\overline{B(x,h)})$ and $(t,y) \mapsto \mathcal{L}V(t,y,S)$ is continuous on $[0,\infty)\times\overline{B(x,h)}$. Thus,

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^{\varepsilon\wedge\tau_{B(x,h)}} \mathcal{L}V(s,X_s,S)ds = \mathcal{L}V(0,x,S), \ \mathbb{P}^x - \text{a.s.}.$$

By Lemma 2.14(b), $\sup_{(t,y)\in[0,\infty)\times\overline{B(x,h)}} |\mathcal{L}V(t,y,S)| < \infty$. Then, we can apply the dominated convergence theorem to derive

$$\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathbb{E}^x\left[\int_0^{\varepsilon\wedge\tau_{B(x,h)}} \mathcal{L}V(s,X_s,S)ds\right] = \mathcal{L}V(0,x,S). \tag{53}$$

Notice that Equation (51) is valid and the local time integral term in Equation (51) vanishes in this case. Then, Equations (51) and (53) together lead to

$$\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\mathbb{E}^x[V(\varepsilon,X_\varepsilon,S)-V(0,x,S)]=\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\mathbb{E}^x\left[\int_0^{\varepsilon\wedge\tau_{B(x,h)}}\mathcal{L}V(s,X_s,S)ds\right]=\mathcal{L}V(0,x,S).$$

Since S is a weak equilibrium, we have $\mathcal{L}V(0,x,S)\le 0$.

Case (iii) $x\in S\setminus(\mathcal{G}\cap S^\circ)$. As S is admissible, we can pick $h>0$ such that $(x-h,x)$ is contained in either $\mathcal{G}\cap S^\circ$ or S^c. By the results in Cases (i) and (ii), as well as the continuity of $x\mapsto\mathcal{L}V(0,x-,S)$ on $(x-h,x]$, we have that

$$\mathcal{L}V(0,x-,S)=\lim_{\varepsilon\searrow 0}\mathcal{L}V(0,(x-\varepsilon)-,S)\le 0.$$

Similarly, $\mathcal{L}V(0,x+,S)\le 0$. □

Proof of Theorem 3.1. The necessity is implied by Proposition 3.11. Let us prove the sufficiency.

Take $x\in S$. Since S is admissible, by Lemma 2.14 and Assumption 2.8(ii), no matter $x\in S^\circ$ or $x\in\partial S$, we can choose $h>0$ such that $V(t,x,S)\in C^{1,2}([0,\infty)\times(x-h,x])$ and $V(t,x,S)\in C^{1,2}([0,\infty)\times[x,x+h))$. By a similar argument as that for Equation (51) (with Lemmas 2.15 and 3.8 applied), we have that

$$\begin{aligned}\frac{1}{\varepsilon}(\mathbb{E}^x[V(\varepsilon,X_\varepsilon,S)]-V(0,x,S))=&\frac{1}{\varepsilon}\mathbb{E}^x\left[\int_0^{\varepsilon\wedge\tau_{B(x,h)}}\frac{1}{2}(\mathcal{L}V(s,X_s-,S)+\mathcal{L}V(s,X_s+,S))ds\right]\\&+\frac{1}{\varepsilon}\mathbb{E}^x\left[\frac{1}{2}\int_0^{\varepsilon\wedge\tau_{B(x,h)}}(V_x(s,x+,S)-V_x(s,x-,S))dL_s^x\right]+o(1)\end{aligned}\tag{54}$$

By Equation (25) and the (left/right) continuity of $(s,y)\mapsto\mathcal{L}V(s,y\pm,S)$ at $(0,x)$, for $\mathbb{P}$-a.s. $\omega\in\Omega$,

$$\limsup_{s\searrow 0}\frac{1}{2}(\mathcal{L}V(s,X_s(\omega)-,S)+\mathcal{L}V(s,X_s(\omega)+,S))\le 0,$$

which leads to

$$\limsup_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\int_0^{\varepsilon\wedge\tau_{B(x,h)}}\frac{1}{2}(\mathcal{L}V(s,X_s(\omega)-,S)+\mathcal{L}V(s,X_s(\omega)+,S))ds\le 0.\tag{55}$$

By Lemma 2.14 (b),

$$\sup_{(t,y)\in[0,1]\times\overline{B(x,h)}}|\mathcal{L}V(t,y-,S)+\mathcal{L}V(t,y+,S)|<\infty.$$

This enables us to apply Fatou's lemma for Equation (55) and get

$$\limsup_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\mathbb{E}^x\left[\int_0^{\varepsilon\wedge\tau_{B(x,h)}}\frac{1}{2}(\mathcal{L}V(s,X_s-,S)+\mathcal{L}V(s,X_s+,S))ds\right]\le 0.\tag{56}$$

By Equation (18),

$$V_x(t,x+,S)-V_x(t,x-,S)\le V_x(0,x+,S)-V_x(0,x-,S)+o(\sqrt{t}).$$

This together with Equation (24) implies that

$$\begin{aligned}
&\frac{1}{\varepsilon}\mathbb{E}^x\left[\frac{1}{2}\int_0^{\varepsilon\wedge\tau_{B(x,h)}}(V_x(s,x+,S)-V_x(s,x-,S))dL_s^x\right]\\
&\le\frac{1}{\varepsilon}\mathbb{E}^x\left[\frac{1}{2}\int_0^{\varepsilon\wedge\tau_{B(x,h)}}(V_x(0,x+,S)-V_x(0,x-,S)+o(\sqrt{\varepsilon}))dL_s^x\right]\\
&\le\frac{1}{2\varepsilon}\cdot o(\sqrt{\varepsilon})\cdot\mathbb{E}^x\left[L^x_{\varepsilon\wedge\tau_{B(x,h)}}\right]\\
&=\frac{1}{2\varepsilon}\cdot o(\sqrt{\varepsilon})\cdot O(\sqrt{\varepsilon})=o(1),
\end{aligned}\tag{57}$$

where the last line follows from Lemma 3.9. Then by Equations (54), (56), and (57), we have that

$$\limsup_{\varepsilon\searrow 0}\frac{1}{\varepsilon}(\mathbb{E}^x[V(\varepsilon,X_\varepsilon,S)]-V(0,x,S))\le 0.$$

□

4 | OPTIMAL MILD EQUILIBRIA ARE WEAK EQUILIBRIA

In this section, we show that an optimal mild equilibrium is a weak equilibrium.

To begin with, let us point out that optimal mild equilibria exist for one-dimensional diffusions. Such existence result is provided in Huang and Zhou (2020, Theorem 4.12), and we summarize it in the current context as follows.

Lemma 4.1. *Let Assumptions 2.1, 2.5, and Equations (14) and (17) hold. Then,*

$$S^* := \cap_{S\in\mathcal{E}} S \tag{58}$$

is an optimal mild equilibrium, where $\mathcal{E}$ is the set containing all mild equilibria.

Remark 4.2. Notice that Equation (12) is assumed in Huang and Zhou (2020, Theorem 4.12) for S^* being an optimal mild equilibrium, which is guaranteed by Assumption 2.1 as stated in Remark 2.2. Also, Equation (17) can be deduced from Assumption 2.8(i) as stated in Remark 2.9.

Below is the main result of this section.

Theorem 4.3. *Let Assumptions 2.1–2.10 hold and S be an admissible stopping policy. If S is an optimal mild equilibrium, then it is also a weak equilibrium.*

By Lemma 4.1 and Theorem 4.3, we have the following.

Corollary 4.4. *Let Assumptions 2.1–2.10 hold. Suppose S^* is admissible. Then S^* is also weak.*

Remark 4.5. The above corollary also provides the existence of weak equilibria (ignoring admissibility) as a by-product. Moreover, since any weak equilibrium is also mild, we can see that S^* is optimal among all mild and weak equilibria.

4.1 | Proof of Theorem 4.3

To show that an optimal mild equilibrium S is a weak equilibrium, by Theorem 3.1 it suffices to verify Equations (24) and (25) for S. Equation (24) will be proved in Proposition 4.7 by contradiction. In particular, if we assume $V_x(0, x_0+, S) - V_x(0, x_0-, S) > 0$, then a mild equilibrium *better* than S can be construct by "digging a small hole $B(x_0, h)$" out of S. The proof of Equation (25) is also carried out via contradiction by finding a *better* mild equilibrium.

Such construction of a *better* mild equilibrium requires the comparison between the expectation of a local time integral before the exit time $\tau_{B(x,h)}$ and the expectation of $\tau_{B(x,h)}$ for small h, which is stated in the following lemma.

Lemma 4.6. *Suppose Assumptions 2.1 and 2.5 hold. For $x_0 \in \mathbb{X}$, we have that*

$$\frac{\mathbb{E}^{x_0+rh}\left[\int_0^{\tau_{B(x_0,h)}} \delta(t) dL_t^{x_0}\right] \cdot h}{\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]} \xrightarrow{h\searrow 0} \frac{\sigma^2(x_0)}{1+|r|} \quad \textit{uniformly for } r \in (-1,1). \tag{59}$$

Proof. We first prove

$$\frac{\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]}{(1-r^2)h^2} \xrightarrow{h\searrow 0} \frac{1}{\sigma^2(x_0)} \quad \text{uniformly for } r \in (-1,1) \tag{60}$$

by using an argument similar to that for Christensen and Lindensjö (2020a, Lemma A.5). Pick a constant a and consider the function $g(t,z) := a(z-x_0)^2 - t$. We have that

$$\mathcal{L}g(t,z) = -1 + a\sigma^2(z) + \mu(z)2a(z-x_0).$$

By Assumption 2.1(ii), $\sigma^2(x_0) > 0$. For any constant $a > \frac{1}{\sigma^2(x_0)}$, by the continuity of $\mu(x)$ and $\sigma(x)$, we can find $h > 0$, which only depends on a, such that $\mathcal{L}g(t,z) \geq 0$ for any $z \in B(x_0,h)$. Applying Ito's formula to $g(t, X_t)$, we have that

$$a\mathbb{E}^y\left[(X_{\tau_{B(x_0,h)}} - x_0)^2\right] - \mathbb{E}^y[\tau_{B(x_0,h)}] - a(y-x_0)^2 = \mathbb{E}^y\left[\int_0^{\tau_{B(x_0,h)}} \mathcal{L}g(t,X_t)dt\right] \geq 0, \quad \forall y \in B(x_0,h).$$

For $y \in B(x_0,h)$, rewrite $y = x_0 + rh$ for some $r \in (-1,1)$. Then the above inequality leads to

$$\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}] \leq a\mathbb{E}^{x_0+rh}\left[(X_{\tau_{B(x_0,h)}} - x_0)^2\right] - a(rh)^2 = ah^2 - ar^2h^2 = a(1-r^2)h^2, \quad \forall r \in (-1,1). \tag{61}$$

Similarly, for any constant $0 < \tilde{a} < \frac{1}{\sigma^2(x_0)}$, we can find $\tilde{h}$, which only depends on $\tilde{a}$, such that $\mathcal{L}g(t,z) \leq 0$ on $B(x_0,\tilde{h})$, and

$$\mathbb{E}^{x_0+r\tilde{h}}[\tau_{B(x_0,\tilde{h})}] \geq \tilde{a}(1-r^2)\tilde{h}^2, \quad \forall r \in (-1,1). \tag{62}$$

By (61) and (62), for any $H \in (0, h \wedge \tilde{h}]$ we have that

$$\tilde{a} \leq \frac{\mathbb{E}^{x_0+rH}[\tau_{B(x_0,H)}]}{(1-r^2)H^2} \leq a, \quad \text{for all } r \in (-1,1),\ \tilde{a} \in \left(0, \frac{1}{\sigma^2(x_0)}\right) \text{ and } a > \frac{1}{\sigma^2(x_0)}.$$

Let $\tilde{a} = \frac{1}{\sigma^2(x_0)} - \varepsilon$ and $a = \frac{1}{\sigma^2(x_0)} + \varepsilon$ for any $\varepsilon > 0$ and then take $H \searrow 0$ for the above inequality. By the arbitrariness of ε, Equation (60) follows.

Next, we prove Equation (59). Consider the function $\mathfrak{g}(t,z) := \delta(t)|z - x_0|$. For $h > 0$ and $y \in B(x_0, h)$, applying Lemma 2.15 to $\mathfrak{g}(t, X_t)$, we have that

$$\begin{aligned}\mathbb{E}^y[\delta(\tau_{B(x_0,h)})|X_{\tau_{B(x_0,h)}} - x_0|] - |y - x_0| = \mathbb{E}^y\left[\int_0^{\tau_{B(x_0,h)}} \frac{1}{2}(\mathcal{L}\mathfrak{g}(t, X_t-) + \mathcal{L}\mathfrak{g}(t, X_t+))dt\right] \\ + \mathbb{E}^y\left[\int_0^{\tau_{B(x_0,h)}} \frac{1}{2}(1-(-1))\delta(t)dL_t^{x_0}\right].\end{aligned} \tag{63}$$

By Lemma 2.7, $|\delta'(t)| \leq |\delta'(0)|\delta(t) \leq |\delta'(0)|$. This implies that

$$\left|\frac{1}{2}(\mathcal{L}\mathfrak{g}(t,z-) + \mathcal{L}\mathfrak{g}(t,z+))\right| \leq |\delta'(t)| \cdot |z - x_0| + \delta(t)|\mu(z)| \leq |\delta'(0)| \cdot |z - x_0| + |\mu(z)|. \tag{64}$$

By Equations (63) and (64), we have that

$$\begin{aligned}&h\mathbb{E}^y[\delta(\tau_{B(x_0,h)})] - |y - x_0| - \mathbb{E}^y\left[\int_0^{\tau_{B(x_0,h)}} (|\delta'(0)||X_t - x_0| + |\mu(X_t)|)dt\right] \\ &\leq \mathbb{E}^y\left[\int_0^{\tau_{B(x_0,h)}} \delta(t)dL_t^{x_0}\right] \\ &\leq h\mathbb{E}^y[\delta(\tau_{B(x_0,h)})] - |y - x_0| + \mathbb{E}^y\left[\int_0^{\tau_{B(x_0,h)}} (|\delta'(0)||X_t - x_0| + |\mu(X_t)|)dt\right].\end{aligned} \tag{65}$$

Notice that $|X_t - x| \leq h$ for $t \leq \tau_{B(x_0,h)}$ and $\sup_{z \in B(x_0,1)} |\mu(z)| \leq K$ for some constant $K > 0$ that depends on x_0. Then, for $h \leq 1$, by rewriting $y = x_0 + rh$ in Equation (65) we have that

$$\begin{aligned}&h\mathbb{E}^{x_0+rh}[\delta(\tau_{B(x_0,h)})] - h|r| - (h|\delta'(0)| + K) \cdot \mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}] \\ &\leq \mathbb{E}^{x_0+rh}\left[\int_0^{\tau_{B(x_0,h)}} \delta(t)dL_t^{x_0}\right] \\ &\leq h\mathbb{E}^{x_0+rh}[\delta(\tau_{B(x_0,h)})] - h|r| + (h|\delta'(0)| + K) \cdot \mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}], \quad \forall r \in (-1,1).\end{aligned}$$

Then,

$$\begin{aligned}&\frac{h^2(\mathbb{E}^{x_0+rh}[\delta(\tau_{B(x_0,h)})] - |r|)}{\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]} - \frac{(|\delta'(0)|h^2 + Kh) \cdot \mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]}{\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]} \\ &\leq \left(h\mathbb{E}^{x_0+rh}\left[\int_0^{\tau_{B(x_0,h)}} \delta(t)dL_t^{x_0}\right]\right) \Big/ \left(\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]\right) \\ &\leq \frac{h^2(\mathbb{E}^{x_0+rh}[\delta(\tau_{B(x_0,h)})] - |r|)}{\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]} + \frac{(|\delta'(0)|h^2 + Kh) \cdot \mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]}{\mathbb{E}[\tau_{B(x_0,h)}]}, \quad \forall r \in (-1,1).\end{aligned} \tag{66}$$

By the second inequality in Lemma 2.7, it holds uniformly in $r \in (-1,1)$ that

$$|\mathbb{E}^{x_0+rh}[\delta(\tau_{B(x_0,h)})] - 1| = \mathbb{E}^{x_0+rh}[1 - \delta(\tau_{B(x_0,h)})] \le |\delta'(0)| \cdot \mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}] \xrightarrow{h\searrow 0} 0.$$

This together with Equation (60) implies that

$$\frac{h^2(\mathbb{E}^{x_0+rh}[\delta(\tau_{B(x_0,h)})] - |r|)}{\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]} \xrightarrow{h\searrow 0} \frac{\sigma^2(x_0)}{1-r^2} \cdot (1-|r|) = \frac{\sigma^2(x_0)}{1+|r|} \quad \text{uniformly for } r \in (-1,1). \tag{67}$$

Notice that $\lim_{h\searrow 0}(|\delta'(0)|h^2 + Kh) = 0$. This together with Equations (66) and (67) implies Equation (59). □

Now we are ready to deal with Equation (24) in the following proposition. The verification for Equation (25) follows next.

Proposition 4.7. *Let Assumptions 2.1–2.10 hold and S be an admissible stopping policy. If S is an optimal mild equilibrium, then*

$$V_x(0, x-, S) \ge V_x(0, x+, S) \quad \forall x \in S.$$

Proof. Notice that Assumption 2.8(ii) and Lemma 2.14 guarantees the existence of $V_x(t, x\pm, S)$ and $\mathcal{L}V(t, x\pm, S)$ for any $(t,x) \in [0,\infty) \times \mathbb{X}$. We prove the desired result by contradiction. Take $x_0 \in S$ and suppose

$$a := V_x(0, x_0+, S) - V_x(0, x_0-, S) > 0. \tag{68}$$

Recall $\mathcal{G}$ defined in Equation (16). To reach to a contradiction, we will construct a new mild equilibrium, which is strictly better than S, for each of the three cases:

(i) $x_0 \in \partial S$ for boundary case (a);
(ii) $x_0 \in \partial S$ for boundary case (b);
(iii) $x_0 \in S^\circ$.

Case (i) $x_0 \in \partial S$ for boundary case (a). Without loss of generality, we assume that $(x_0, x_0 + h_0) \subset (S^\circ \cap \mathcal{G})$ and $(x_0 - h_0, x_0) \subset S^c$ for some $h_0 > 0$. Denote $l := \sup\{y \le x_0 - h_0 : y \in S\}$, and note that l can be $-\infty$. We proceed the proof for this case in three steps.

Step 1. We show that there exists $h \in (0, h_0)$ such that,

$$\mathbb{E}^y\left[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)\right] - f(y) > 0, \quad \forall y \in (x_0 - h, x_0 + h). \tag{69}$$

Notice from Lemma 3.10 and Equation (68) that

$$V_x(t, x_0+, S) - V_x(t, x_0-, S) \ge \delta(t)(V_x(0, x_0+, S) - V(0, x_0-, S)) = a\delta(t), \quad \forall t \ge 0. \tag{70}$$

Fix $h \in (0, h_0)$ and pick an arbitrary $x \in B(x_0, h)$. For all $n \in \mathbb{N}$, write

$$\tau_n := \tau_{B(x_0,h)} \wedge n \tag{71}$$

for short. We apply Lemma 2.15 to $V(t,X_t,S)$ on $[0,\tau_n]$ and take expectation; the diffusion term vanishes under expectation due to Lemma 2.14 and continuity of σ. Then combining with Equation (70), we have that

$$\begin{aligned}&\mathbb{E}^x\left[V\left(\tau_n,X_{\tau_n},S\right)\right]-V(0,x,S)\\&\geq\mathbb{E}^x\left[\int_0^{\tau_n}\frac{1}{2}(\mathcal{L}V(t,X_t-,S)+\mathcal{L}V(t,X_t+,S))dt\right]+a\mathbb{E}^x\left[\int_0^{\tau_n}\delta(t)dL_t^{x_0}\right].\end{aligned}\tag{72}$$

By Lemma 2.14(b), we have

$$M:=\sup_{(t,y)\in[0,\infty)\times\overline{B(x_0,h_0)}}\frac{1}{2}(|\mathcal{L}V(t,y-,S)|+|\mathcal{L}V(t,y+,S)|)<\infty.$$

This together with Equation (72) implies that

$$\mathbb{E}^x[V(\tau_n,X_{\tau_n},S)]-V(0,x,S)\geq-M\mathbb{E}^x[\tau_n]+a\mathbb{E}^x\left[\int_0^{\tau_n}\delta(t)dL_t^{x_0}\right]\quad\forall n\in\mathbb{N}.\tag{73}$$

For the LHS of Equation (73), Equation (17) readily implies that

$$\lim_{n\to\infty}\mathbb{E}^{x_0}[V(\tau_n,X_{\tau_n},S)=\mathbb{E}^{x_0}[V(\tau_{B(x_0,h)},X_{\tau_{B(x_0,h)}},S)].\tag{74}$$

Indeed, set $\eta_n:=\inf\{t\geq\tau_n,X_t\in S\}$ for all $n\in\mathbb{N}$, and $\eta:=\inf\{t\geq\tau_{B(x_0,h)},X_t\in S\}$. We have $\mathbb{E}^x[V(\tau_n,X_{\tau_n},S)]=\mathbb{E}^x[\mathbb{E}^x[\delta(\eta_n)f(X_{\eta_n})|\mathcal{F}_{\tau_n}]]=\mathbb{E}^x[\delta(\eta_n)f(X_{\eta_n})]$ for all $n\in\mathbb{N}$, and $\mathbb{E}^{x_0}[V(\tau_{B(x_0,h)},X_{\tau_{B(x_0,h)}},S)]=\mathbb{E}^{x_0}[\delta(\eta)f(X_\eta)]$. As $n\to\infty$, $\eta_n\to\eta$, $\mathbb{P}^x$-a.s.. Then by Assumption 2.8, we can apply the dominated convergence theorem to get $\lim_{n\to\infty}\mathbb{E}^{x_0}[\delta(\eta_n)f(X_{\eta_n})]=\mathbb{E}^{x_0}[\delta(\eta)f(X_\eta)]$, that is, Equation (74) holds. (Note that Equation (14) is used on $\{\eta=\infty\}$.)

Applying the monotone convergence theorem to the RHS of Equation (73) and combining with Equation (74), we have that

$$\mathbb{E}^x[V(\tau_{B(x_0,h)},X_{B(x_0,h)},S)]-V(0,x,S)\geq-M\mathbb{E}^x[\tau_{B(x_0,h)}]+a\mathbb{E}^x\left[\int_0^{\tau_{B(x_0,h)}}\delta(t)dL_t^{x_0}\right].\tag{75}$$

By the arbitrariness of $x\in B(x_0,h)$,

$$\begin{aligned}&\mathbb{E}^{x_0+rh}[V(\tau_{B(x_0,h)},X_{B(x_0,h)},S)]-V(0,x_0+rh,S)\\&\geq-M\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]+a\mathbb{E}^{x_0+rh}\left[\int_0^{\tau_{B(x_0,h)}}\delta(t)dL_t^{x_0}\right],\quad\forall r\in(-1,1).\end{aligned}\tag{76}$$

By Lemma 4.6 and $|\sigma(x_0)|>0$, we can choose the above h small enough such that

$$\mathbb{E}^{x_0+rh}\left[\int_0^{\tau_{B(x_0,h)}}\delta(t)dL_t^{x_0}\right]\geq\left(\frac{M}{a}+1\right)\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}],\quad\forall r\in(-1,1).$$

Consequently, Equation (76) leads to

$$\mathbb{E}^{x_0+rh}[V(\tau_{B(x_0,h)},X_{\tau_{B(x_0,h)}},S)]-V(0,x_0+rh,S)\geq a\mathbb{E}^{x_0+rh}[\tau_{B(x_0,h)}]>0,\quad\forall r\in(-1,1),$$

which gives Equation (69).

Step 2. In the rest part of case (i), we take h such that Equation (69) holds and write $S_h := S \setminus B(x_0, h)$ for short. In this step, we prove by contradiction that

$$J(y, S_h) \geq \mathbb{E}^y[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)], \quad \forall y \in [x_0, x_0 + h). \tag{77}$$

Suppose

$$\alpha := \inf_{y \in [x_0, x_0+h]} \Big(J(y, S_h) - \mathbb{E}^y[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)] \Big) < 0. \tag{78}$$

As $x_0 + h \in S^\circ$, by Lemma 2.14(a),

$$J(x_0 + h, S_h) = f(x_0 + h) = \mathbb{E}^{x_0+h}[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)].$$

By the continuity of functions $y \mapsto J(y, S_h)$ and $y \mapsto \mathbb{E}^y[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)]$ on $[x_0, x_0 + h]$, there exists $z^* \in [x_0, x_0 + h)$ such that the infimum in Equation (78) is attained at z^*, that is,

$$J(z^*, S_h) - \mathbb{E}^{z^*}[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)] = \alpha. \tag{79}$$

Define

$$\nu := \inf\{t \geq \tau_{B(x_0,h)} \, : \, X_t \in S\} \quad \text{and} \quad A := \{X_{\tau_{B(x_0,h)}} = x_0 - h, \, X_\nu = x_0, \, \nu < \infty\}.$$

Notice that $\rho_{S_h} = \nu$, $\mathbb{P}^{z^*}$-a.s. on both sets $\{X_{\tau_{B(x_0,h)}} = x_0 + h\}$ and $\{X_{\tau_{B(x_0,h)}} = x_0 - h, X_\nu < x_0\}$. We have that

$$\begin{aligned}
J(z^*, S_h) - \mathbb{E}^{z^*}[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)] &= \mathbb{E}^{z^*}\Big[1_A \cdot \Big(\delta(\rho_{S_h}) f(X_{\rho_{S_h}}) - \delta(\nu) f(X_\nu)\Big)\Big] \\
&\geq \mathbb{E}^{z^*}\Big[1_A \delta(\nu) \cdot \Big(\mathbb{E}^{z^*}[\delta(\rho_{S_h} - \nu) f(X_{\rho_{S_h}}) \mid \mathcal{F}_\nu] - f(X_\nu)\Big)\Big] \\
&= \mathbb{E}^{z^*}[1_A \delta(\nu)] \cdot (J(x_0, S_h) - f(x_0)) \\
&> \mathbb{E}^{z^*}[1_A \delta(\nu)] \cdot \Big(J(x_0, S_h) - \mathbb{E}^{x_0}[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)]\Big) \\
&\geq \mathbb{E}^{z^*}[1_A \delta(\nu)] \cdot \alpha \\
&> \alpha,
\end{aligned}$$

where the second (in)equality follows from Equation (13) and $f \geq 0$, the third (in)equality follows from the strong Markov property of X and the fact that $X_v = x_0$ on A, the fourth (in)equality follows from Equation (69) with $y = x_0$, the fifth (in)equality follows from the definition of α in Equation (78), and the last (in)equality follows from the fact that $\nu \geq \tau_{B(x_0,h)} > 0$ and $\delta(t) < 1$ for $t > 0$. This contradicts Equation (79). Therefore, Equation (77) holds.

Step 3. Now we prove that S_h is a mild equilibrium and is strictly better than S. By Equations (69) and (77) and noticing that $(x_0, x_0 + h) \subset S^\circ$, we have

$$J(y, S_h) > f(y) = J(y, S), \quad \forall y \in [x_0, x_0 + h). \tag{80}$$

Then for any $y \in (l, x_0)$, we have that

$$
\begin{aligned}
J(y,S_h)-J(y,S) &= \mathbb{E}^y\left[1_{\{X_{\rho_S}=x_0,\,\rho_S<\infty\}}\Big(\delta(\rho_{S_h})f(X_{\rho_{S_h}})-\delta(\rho_S)f(x_0)\Big)\right]\\
&\geq \mathbb{E}^y\left[1_{\{X_{\rho_S}=x_0,\,\rho_S<\infty\}}\delta(\rho_S)\Big(\mathbb{E}^y\left[\delta(\rho_{S_h}-\rho_S)f(X_{\rho_{S_h}})|\mathcal{F}_{\rho_S}\right]-f(x_0)\Big)\right]\\
&= \mathbb{E}^y\left[1_{\{X_{\rho_S}=x_0,\,\rho_S<\infty\}}\delta(\rho_S)(J(x_0,S_h)-f(x_0))\right]\\
&\geq 0.
\end{aligned}\tag{81}
$$

where the second (in)equality follows again from Equation (13) and the non-negativity of f, the third (in)equality follows from the strong Markov property of X, and the last (in)equality follows from Equation (80) with $y=x_0$. As S is a mild equilibrium, above inequality implies

$$
J(y,S_h)\geq J(y,S)\geq f(y),\quad \forall y\in(l,x_0).\tag{82}
$$

This together with Equation (80) and the fact $J(\cdot,S_h)=J(\cdot,S)$ on $\mathbb{X}\setminus(l,x_0+h)$ implies that S_h is a mild equilibrium and is strictly better than S.

Case (ii) $x_0\in\partial S$ for boundary case (b). We denote

$$
l:=\sup\{y<x_0,y\in S\},\quad r:=\inf\{y>x_0,y\in S\},\quad \text{and}\quad \tilde{\tau}_n:=\tau_{((l,r)\cap B(x_0,n))}\wedge n \text{ for } n\in\mathbb{N}.
$$

By a similar discussion through Equations (70)–(72) (with Lemmas 2.15 and 3.10 applied), we have that

$$
\begin{aligned}
&\mathbb{E}^{x_0}[V(\tilde{\tau}_n,X_{\tilde{\tau}_n},S)]-V(0,x_0,S)\\
&\geq \mathbb{E}^{x_0}\left[\int_0^{\tilde{\tau}_n}\frac{1}{2}(\mathcal{L}V(s,X_s-,S)+\mathcal{L}V(s,X_s+,S))ds\right]+a\mathbb{E}^{x_0}\left[\int_0^{\tilde{\tau}_n}\delta(t)dL_t^{x_0}\right].
\end{aligned}\tag{83}
$$

By Lemma 2.14(a), $\mathcal{L}V(t,x,S)=0$ for any $(t,x)\in[0,\infty)\times S^c$, and thus the first term on the RHS of Equation (83) vanishes for all $n\in\mathbb{N}$. As a result, we can rewrite Equation (83) as

$$
\mathbb{E}^{x_0}[V(\tilde{\tau}_n,X_{\tilde{\tau}_n},S)]-V(0,x_0,S)\geq a\mathbb{E}^{x_0}\left[\int_0^{\tilde{\tau}_n}\delta(t)dL_t^{x_0}\right]\geq a\mathbb{E}^{x_0}\left[\int_0^{\tilde{\tau}_1}\delta(t)dL_t^{x_0}\right]>0,\quad \forall n\in\mathbb{N}.
$$

Meanwhile, similar to Equation (74), Assumption 2.8 implies that $\mathbb{E}^{x_0}[V(\tilde{\tau}_n,X_{\tilde{\tau}_n},S)]\to\mathbb{E}^{x_0}[V(\tau_{(l,r)},X_{\tau_{(l,r)}},S)]$ as $n\to\infty$. This together with the above inequality implies that

$$
J(x_0,S\setminus\{x_0\})-f(x_0)=\mathbb{E}^{x_0}[V(\tau_{(l,r)},X_{\tau_{(l,r)}},S)]-V(0,x_0,S)\geq a\mathbb{E}^{x_0}\left[\int_0^{\tilde{\tau}_1}\delta(t)dL_t^{x_0}\right]>0.\tag{84}
$$

Now set $\widetilde{S}:=S\setminus\{x_0\}$ and pick any $y\in(l,r)$. We can apply an argument similar to that in Equation (81), by using Equation (84) and replacing S_h with $\widetilde{S}$, to reach that $J(y,\widetilde{S})-J(y,S)\geq 0$. Hence, $J(y,\widetilde{S})\geq f(y)$ for $y\in(l,r)$. As $J(\cdot,\widetilde{S})=J(\cdot,S)$ on $\mathbb{X}\setminus(l,r)$, we have that $\widetilde{S}$ is a mild equilibrium. Due to Equation (84), $\widetilde{S}$ is strictly better than S.

Case (iii) $x_0\in S^\circ$. Choose $h_0>0$ such that $B(x_0,h_0)\setminus\{x_0\}\subset(\mathcal{G}\cap S^\circ)$. Following the argument in *Step 1* of case (i), we can again reach Equation (69) for some $0<h\leq h_0$, which indicates

$$
J(y,S\setminus B(x_0,h))>f(y),\quad \forall y\in B(x_0,h).
$$

As $J(\cdot, S \setminus B(x_0, h)) = J(\cdot, S)$ on $\mathbb{X} \setminus B(x_0, h)$, we have that $S \setminus B(x_0, h)$ is a mild equilibrium and is strictly better than S. □

Proof of Theorem 4.3. Thanks to Lemma 2.14(a), Theorem 3.1, and Proposition 4.7, we only need to show Equation (25) for $x \in S$. Recall $\mathcal{G}$ defined in Equation (16). Let $x_0 \in S$ and we consider three cases: (i) $x_0 \in (S^\circ \cap \mathcal{G})$, (ii) $x_0 = \theta_n \in S^\circ \setminus \mathcal{G}$ for some $n \in I$, and (iii) $x_0 \in \partial S$.

Case (i) $x_0 \in (S^\circ \cap \mathcal{G})$. We prove Equation (25) by contradiction. Suppose $\mathcal{L}V(0, x_0, S) = a > 0$. By Assumption 2.8(ii), we can choose $h > 0$ such that $V(t, x, S) = \delta(t) f(x) \in C^{1,2}(B(x_0, h) \times (0, \infty))$ and

$$\mathcal{L}V(0, x, S) = \delta'(0) f(x) + \mu(x) f'(x) + \frac{1}{2}\sigma^2(x) f''(x) \geq \frac{a}{2}, \quad \forall x \in B(x_0, h). \tag{85}$$

Then for any $(t, x) \in [0, \infty) \times B(x_0, h)$, we have that

$$\begin{aligned} \mathcal{L}V(t, x, S) =& \delta'(t) f(x) + \delta(t)\left(\mu(x) f'(x) + \frac{1}{2}\sigma^2(x) f''(x) \right) \\ \geq& \delta(t)\left(\delta'(0) f(x) + \mu(x) f'(x) + \frac{1}{2}\sigma^2(x) f''(x) \right) \geq \delta(t)\frac{a}{2}, \end{aligned} \tag{86}$$

where the first inequality above follows from Lemma 2.7 and the non-negativity of f. Let us reuse the notation τ_n defined in Equation (71). By Equation (86) and an argument similar to that for Equations (72) and (74) (notice that the local time integral in Lemma 2.15 vanishes in the current case), we have that for any $x \in B(x_0, h)$,

$$\begin{cases} \mathbb{E}^x[V(\tau_n, X_{\tau_n}, S)] - V(0, x, S) = \mathbb{E}^x\left[\int_0^{\tau_n} \mathcal{L}V(s, X_s, S)\right] \geq \mathbb{E}^x\left[\int_0^{\tau_n} \delta(t)\frac{a}{2} dt\right] > 0, \quad \forall n \in \mathbb{N}, \\ \mathbb{E}^x[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)] = \lim_{n\to\infty} \mathbb{E}^x[V(\tau_n, X_{\tau_n}, S)] \end{cases}$$

This implies that

$$\mathbb{E}^x[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S) - V(0, x, S)] \geq \mathbb{E}^x\left[\int_0^{\tau_{B(x_0,h)}} \delta(t)\frac{a}{2} dt\right] > 0, \quad \forall x \in B(x_0, h).$$

Now consider $\widetilde{S} = S \setminus B(x_0, h)$. The above inequality implies

$$J(x, \widetilde{S}) - f(x) = \mathbb{E}^x[V(\tau_{B(x_0,h)}, X_{\tau_{B(x_0,h)}}, S)] - V(0, x, S) > 0 \quad \forall x \in B(x_0, h). \tag{87}$$

Obviously, $J(\cdot, \widetilde{S}) = J(\cdot, S)$ on $\mathbb{X} \setminus B(x_0, h)$. This together with Equation (87) shows that $\widetilde{S}$ is an equilibrium and is strictly better than S, a contradiction. Hence, $\mathcal{L}V(0, x_0, S) \leq 0$, as desired.

Case (ii) $x_0 = \theta_n \in S^\circ \setminus \mathcal{G}$ for some $n \in I$. Without loss of generality, we assume $\mathcal{L}V(0, x_0+, S) = a > 0$. Then we can pick $h > 0$ such that $(x_0, x_0 + h) \subset S^\circ \cap (\theta_n, \theta_{n+1})$. By the continuity of $x \to \mathcal{L}V(0, x+, S)$ on $[x_0, x_0 + h)$ (due to Assumptions 2.1(i), 2.8(ii), and the fact that $V(t, x, S) = \delta(t) f(x)$ for $x \in S$), we can find $0 < \tilde{h} < h$ such that $\mathcal{L}V(0, y, S) \geq a/2 > 0$ for all $y \in (x_0, x_0 + \tilde{h})$. Set $\tilde{x} := (2x_0 + \tilde{h})/2$. Then $B(\tilde{x}, \tilde{h}/4) \subset (x_0, x_0 + \tilde{h}) \subset S^\circ \cap \mathcal{G}$, and a contradiction can be reached by the same argument as in case (i).

Case (iii) $x_0 \in \partial S$. For boundary case (a), suppose again that $\mathcal{L}V(0, x_0-, S) \vee \mathcal{L}V(0, x_0+, S) > 0$. Without loss of generality, we assume $(x_0, x_0 + h_0) \subset (S^\circ \cap \mathcal{G})$ and $(x_0 - h_0, x_0) \subset S^c$ for some $h_0 > 0$. By Lemma 2.14(a), $\mathcal{L}V(0, x-, S) \equiv 0$ on $(x_0 - h_0, x_0]$, and therefore, $\mathcal{L}V(0, x_0+, S) > 0$. Then the same argument as in case (ii) can be applied to get a contradiction.

For boundary case (b), Lemma 2.14(a) directly tells that $\mathcal{L}V(0,x-,S) \vee \mathcal{L}V(0,x+,S) = 0$, and the proof is complete. $\square$

5 | WHEN WEAK OR OPTIMAL MILD EQUILIBRIA ARE STRONG

After establishing the relation between optimal mild and weak equilibria, we take a further step to study whether a weak or optimal mild equilibrium is strong.

We already know that an admissible weak or optimal mild equilibrium S satisfies the two conditions (24) and (25) in Theorem 3.1. To make S a strong equilibrium, the first-order condition (6) needs to be upgraded to the local maximum condition (7). Recall the discussion at the beginning of Section 3.1. Intuitively, a sufficient condition for Equation (7) is the LHS of Equation (28) being negative for all ε small enough. As a result, if at least one of the two inequalities (24) and (25) is strict for all the points in the weak or optimal equilibrium S, then S should also be strong. To this end, let us define for any admissible $S \in \mathcal{B}$,

$$\mathfrak{S}_S := \{x \in S : \mathcal{L}V(0,x-,S \wedge \mathcal{L}V(0,x+,S) < 0\} \cup \{x \in S : V_x(0,x-,S) > V_x(0,x+,S)\}. \tag{88}$$

Theorems 5.1 and 5.2 are the main results of this section, and their proofs are provided in the next subsection. The first main result concerns when a weak equilibrium is strong.

Theorem 5.1. *Let Assumptions 2.1–2.10 hold and S be an admissible weak equilibrium. If $S = \mathfrak{S}_S$, then S is also strong.*

The next result regards the relation between optimal mild and strong equilibria.

Theorem 5.2. *Let Assumptions 2.1–2.10 hold.*

(a) For any admissible optimal mild equilibrium S, if $\mathfrak{S}_S$ is admissible and closed, then $\mathfrak{S}_S$ is a strong equilibrium.
(b) Recall S^* defined in Equation (58). We have $\overline{\mathfrak{S}_{S^*}} = S^*$. Hence, if $\mathfrak{S}_{S^*}$ is closed and admissible, then S^* is a strong equilibrium.

Remark 5.3. Theorem 5.2 indicates that S^* and $\mathfrak{S}_{S^*}$ are almost the same, and roughly speaking, S^* is a strong equilibrium possibly except some points in $\overline{\mathfrak{S}_{S^*}} \setminus \mathfrak{S}_{S^*}$. In many cases, we indeed have $S^* = \mathfrak{S}_{S^*}$, as a result of which S^* is strong. This is demonstrated in all the examples in Section 6.

Remark 5.4. Suppose Assumptions 2.1–2.10 hold and S^* is admissible. Then S^* cannot contain an isolated point at which f is continuously differentiable. Indeed, suppose x is an isolated point of $S^* = \overline{\mathfrak{S}_{S^*}}$ and f is smooth at x. Then $x \in \mathfrak{S}_{S^*}$. On the other hand, since $\mathcal{L}(0,x-,S^*) = \mathcal{L}(0,x+,S^*) = 0$ by Lemma 2.14(a), and $V_x(0,x-,S^*) = V_x(0,x+,S)$ by Corollary 3.2, we would have $x \notin \mathfrak{S}_{S^*}$, a contradiction.

5.1 | Proofs of Theorems 5.1 and 5.2

As discussed above, we aim to achieve the negativity in the RHS of Equation (28) for ε small enough; when $V_x(s, x+, S) - V_x(s, x-, S) = 0$, the integral on $\frac{1}{2}(\mathcal{L}V(s, X_s-, S) + \mathcal{L}V(s, X_s+, S))$ on the RHS of Equation (28) should be negative. Since only one of the two values $\mathcal{L}V(s, X_s\pm, S)$ is required to be negative in the definition of $\mathfrak{S}_S$, we will estimate the probability that X goes to the left/right from the starting point. Such probability estimation is provided in the following lemma.

Lemma 5.5. *Let Assumption 2.1 hold. Then,*

$$\lim_{t\searrow 0} \mathbb{P}^{x_0}(X_t > x_0) = \lim_{t\searrow 0} \mathbb{P}^{x_0}(X_t < x_0) = \frac{1}{2}, \quad \forall x_0 \in \mathbb{X}. \tag{89}$$

Proof. Let $X_0 = x_0 \in \mathbb{X}$. Recall $\widetilde{X}$ and $\bar{X}$ defined in Equation (35). Denote $R_\varepsilon := \mu(x_0)\varepsilon + \bar{X}_\varepsilon$. Then,

$$X_\varepsilon = x_0 + R_\varepsilon + \sigma(x_0)W_\varepsilon. \tag{90}$$

By Lemma 3.7, there exists some constant $C > 0$ such that for any $\varepsilon > 0$ small enough, $\mathbb{E}^{x_0}[|R_\varepsilon|] \leq C\varepsilon$, which leads to

$$\mathbb{P}^{x_0}\left(|R_\varepsilon| \geq \frac{1}{2}\varepsilon^{3/4}\right) \leq \frac{2\mathbb{E}^{x_0}[|R_\varepsilon|]}{\varepsilon^{3/4}} \leq 2C \cdot \varepsilon^{1/4}. \tag{91}$$

By Equations (90) and (91), for $\varepsilon > 0$ small enough

$$\begin{aligned}
&\mathbb{P}^{x_0}(X_\varepsilon > x_0) \geq \mathbb{P}^{x_0}\left(\sigma(x_0)W_\varepsilon > \varepsilon^{3/4}, R_\varepsilon > -\frac{1}{2}\varepsilon^{3/4}\right) \\
&\geq \mathbb{P}^{x_0}\left(\sigma(x_0)W_\varepsilon > \varepsilon^{3/4}\right) - \mathbb{P}^{x_0}\left(R_\varepsilon \leq -\frac{1}{2}\varepsilon^{3/4}\right) \geq 1 - \Phi\left(\frac{\varepsilon^{3/4}}{\sigma(x_0)\sqrt{\varepsilon}}\right) - \mathbb{P}^{x_0}\left(|R_\varepsilon| \geq \frac{1}{2}\varepsilon^{3/4}\right) \\
&\geq 1 - \Phi\left(\frac{\varepsilon^{1/4}}{\sigma(x_0)}\right) - 2C\varepsilon^{1/4} \to 1 - \Phi(0) - 0 = \frac{1}{2}, \quad \text{as } \varepsilon \searrow 0,
\end{aligned}$$

where Φ is the cumulative distribution function for the standard normal distribution. Therefore, $\liminf_{t\searrow 0} \mathbb{P}^{x_0}(X_t > x_0) \geq \frac{1}{2}$. Similarly, $\liminf_{t\searrow 0} \mathbb{P}^{x_0}(X_t < x_0) \geq \frac{1}{2}$. Thus, Equation (89) holds. $\square$

Now we are ready to prove Theorem 5.1.

Proof of Theorem 5.1. To prove the desired result, we need to verify that for any $x_0 \in S$,

$$\exists \varepsilon(x_0) > 0, \text{ s.t. } \forall \varepsilon' \leq \varepsilon(x_0), f(x_0) - \mathbb{E}^{x_0}[\delta(\rho_S^{\varepsilon'})f(X_{\rho_S^{\varepsilon'}})] \geq 0. \tag{92}$$

Since S is a weak equilibrium, by Theorem 3.1,

$$V_x(0, x_0-, S) - V_x(0, x_0+, S) \geq 0, \quad \forall x_0 \in S.$$

Recall Equation (88) and $\mathcal{G}$ defined in Equation (16). Pick $x_0 \in \mathfrak{S}_S$, and we shall verify Equation (92) for two cases: (i) $V_x(0, x_0-, S) - V_x(0, x_0+, S) > 0$, and (ii) $V_x(0, x_0-, S) - V_x(0, x_0+, S) = 0$.

Case (i) Suppose $a := V_x(0, x_0-, S) - V_x(0, x_0+, S) > 0$. By the continuity of $t \mapsto V_x(t, x_0\pm, S)$, we take $\varepsilon > 0$ small enough such that $\delta(t) > \frac{1}{2}$ for all $t \in (0, \varepsilon)$, and

$$V_x(t, x_0+, S) - V_x(t, x_0-, S) < -\frac{a}{2}, \quad \forall t \in (0, \varepsilon). \tag{93}$$

Let $h > 0$ such that both $(x_0 - h, x_0)$ and $(x_0, x_0 + h)$ belong to $\mathcal{G}$. Then for ε small enough,

$$\begin{aligned}&\mathbb{E}^{x_0}[V(\varepsilon, X_\varepsilon, S)] - V(0, x_0, S) + o(\varepsilon) = \mathbb{E}^{x_0}[V(\varepsilon \wedge \tau_{B(x_0,h)}, X_{\varepsilon\wedge\tau_{B(x_0,h)}}, S)] - V(0, x_0, S)\\ &\le \mathbb{E}^{x_0}\left[\int_0^{\varepsilon\wedge\tau_{B(x_0,h)}} \frac{1}{2}(\mathcal{L}V(s, X_s-, S) + \mathcal{L}V(s, X_s+, S))ds\right] - \frac{a}{4}\mathbb{E}^{x_0}[L^{x_0}_{\tau_{B(x_0,h)}\wedge\varepsilon}],\end{aligned} \tag{94}$$

where the first (in)equality follows from Lemma 3.8, the second (in)equality follows from Lemma 2.15 and Equation (93) (the diffusion term vanishes after taking expectation due to the boundedness of $V_x\sigma$ on $[0, \varepsilon] \times \overline{B(x_0, h)}$). By Lemma 2.14(b), there exists a constant $K > 0$ such that

$$\sup_{(t,y)\in[0,1]\times\overline{B(x_0,h)}} \frac{1}{2}|\mathcal{L}V(t, y+, S) + \mathcal{L}V(t, y-, S)| \le K,$$

Then by Lemma 3.9 and $|\sigma(x_0)| > 0$, we can take $\varepsilon_0 \in (0, 1)$ such that for any $\varepsilon \in (0, \varepsilon_0)$, $\frac{a}{4\varepsilon}\mathbb{E}^{x_0}[L^{x_0}_{\tau_{B(x_0,h)}\wedge\varepsilon}] \ge (K + 1)$ and the term $o(\varepsilon)$ in Equation (94) satisfies $|o(\varepsilon)| \le \frac{1}{2}\varepsilon$. Hence, Equation (94) leads to

$$\begin{aligned}&\mathbb{E}^{x_0}[\delta(\rho^\varepsilon_S)f(X_{\rho^\varepsilon_S})] - f(x_0) = \mathbb{E}^{x_0}[V(\varepsilon, X_\varepsilon, S)] - V(0, x_0, S)\\ &\le K\varepsilon - \frac{a}{4}\mathbb{E}^{x_0}[L^{x_0}_{\tau_{B(x_0,h)}\wedge\varepsilon}] + \frac{1}{2}\varepsilon \le -\frac{1}{2}\varepsilon, \quad \forall \varepsilon \le \varepsilon_0.\end{aligned}$$

Case (ii) Suppose $V_x(0, x_0-, S) - V_x(0, x_0+, S) = 0$. Then, by Equation (18),

$$|V_x(t, x_0+, S) - V_x(t, x_0-, S)| = o(\sqrt{t}) \quad \text{for } t > 0 \text{ small enough.}$$

This together with Lemma 3.9 leads to

$$\mathbb{E}^{x_0}\left|\frac{1}{2}\int_0^{\varepsilon\wedge\tau_{B(x_0,h)}} (V_x(s, x_0+, S) - V_x(s, x_0-, S))dL^{x_0}_s\right| = o(\sqrt{\varepsilon}) \cdot \mathbb{E}^{x_0}\left[L^{x_0}_{\varepsilon\wedge\tau_{B(x_0,h)}}\right] = o(\varepsilon). \tag{95}$$

Choose $h_0 > 0$ such that $(x_0 - h_0, x_0) \cup (x_0, x_0 + h_0)$ is contained in $(S^\circ \cap \mathcal{G}) \cup (\mathbb{X} \setminus S)$. For any $h \in (0, h_0)$, similar to Equation (94), we apply Lemmas 2.15, 3.8 and then combine with Equation (95) to get

$$\begin{aligned}\mathbb{E}^{x_0}[V(\varepsilon, X_\varepsilon, S)] - V(0, x_0, S) =&\mathbb{E}^{x_0}[V(\tau_{B(x_0,h)} \wedge \varepsilon, X_{\tau_{B(x_0,h)}\wedge\varepsilon}, S)] - V(0, x_0, S) + o(\varepsilon)\\ =&\mathbb{E}^{x_0}\left[\int_0^{\tau_{B(x_0,h)}\wedge\varepsilon} \mathcal{L}V(s, X_s, S)ds\right] + o(\varepsilon).\end{aligned} \tag{96}$$

Since $V_x(0, x_0-, S) - V_x(0, x_0+, S) = 0$ and $x_0 \in \mathfrak{S}_S$, we have

$$\mathcal{L}V(0, x_0-, S) \wedge \mathcal{L}V(0, x_0+, S) < 0. \tag{97}$$

Without loss of generality, we can assume that

$$-A := \mathcal{L}V(0, x_0+, S) < 0 \quad \text{and} \quad \mathcal{L}V(0, x_0-, S) \leq 0.$$

By the (left/right) continuity of $(t, x) \mapsto \mathcal{L}V(t, x\pm, S)$ at $(0, x_0)$, we can choose $h \in (0, h_0)$ and $\varepsilon_0 > 0$ small enough, such that for any $t \in [0, \varepsilon_0]$, $x \in (x_0, x_0 + h)$ and $y \in (x_0 - h, x_0)$,

$$\mathcal{L}V(t, x, S) = \mathcal{L}V(t, x+, S) \leq -\frac{A}{2} \quad \text{and} \quad \mathcal{L}V(t, y, S) = \mathcal{L}V(t, y-, S) \leq \frac{A}{8}. \tag{98}$$

Then for $\varepsilon \in (0, \varepsilon_0)$ small enough, the first inequality in Equation (98) implies that

$$\begin{aligned}
&\mathbb{E}^{x_0}\left[\int_0^{\tau_{B(x_0,h)}\wedge\varepsilon} \mathcal{L}V(t, X_t, S)1_{\{X_t > x_0\}}dt\right] \leq -\frac{A}{2}\mathbb{E}^{x_0}\left[\int_0^{\tau_{B(x_0,h)}\wedge\varepsilon} 1_{\{X_t > x_0\}}dt\right] \\
&= -\frac{A}{2}\mathbb{E}^{x_0}\left[\int_0^{\varepsilon} 1_{\{X_t > x_0\}}dt\right] + \frac{A}{2}\mathbb{E}^{x_0}\left[\int_{\tau_{B(x_0,h)}\wedge\varepsilon}^{\varepsilon} 1_{\{X_t > x_0\}}dt\right] \\
&= -\frac{A}{2}\int_0^{\varepsilon} \mathbb{P}^{x_0}(X_t > x_0)dt + \frac{A}{2}\mathbb{E}^{x_0}\left[(\varepsilon - \tau_{B(x_0,h)})1_{\{\tau_{B(x_0,h)} < \varepsilon\}}\right] \\
&\leq -\frac{A}{5}\varepsilon + \frac{A}{2}\mathbb{E}^{x_0}\left[(\varepsilon - \tau_{B(x_0,h)})1_{\{\tau_{B(x_0,h)} < \varepsilon\}}\right] \\
&\leq -\frac{A}{5}\varepsilon + \frac{A}{2}\varepsilon\mathbb{P}^{x_0}\left(\tau_{B(x_0,h)} < \varepsilon\right) = -\frac{A}{5}\varepsilon + \frac{A}{2}\varepsilon \cdot o(\varepsilon) \leq -\frac{A}{6}\varepsilon,
\end{aligned} \tag{99}$$

where the forth (in)equality above follows from Lemma 5.5, and the sixth (in)equality follows from Lemma 3.6. In addition, the second inequality in Equation (98) implies

$$\mathbb{E}^{x_0}\left[\int_0^{\tau_{B(x_0,h)}\wedge\varepsilon} \mathcal{L}V(t, X_t, S)1_{\{X_t < x_0\}}dt\right] \leq \frac{A}{8}\varepsilon. \tag{100}$$

Therefore, by plugging Equations (99) and (100) into Equation (96), we have that for $\varepsilon > 0$ small enough,

$$\mathbb{E}^{x_0}[\delta(\rho_S^\varepsilon)f(X_{\rho_S^\varepsilon})] - f(x_0) = \mathbb{E}^{x_0}[V(\varepsilon, X_\varepsilon, S)] - V(0, x_0, S) \leq -\frac{A}{6}\varepsilon + \frac{A}{8}\varepsilon + o(\varepsilon) < -\frac{A\varepsilon}{25},$$

and the proof is complete. □

To prepare for the proof of Theorem 5.2, let us illustrate a property of an arbitrary optimal mild equilibrium S, which says that $\mathfrak{S}_S$ actually forms the "essential" part of S, and by removing the "inessential" part from S, the remaining part is still optimal mild.

Proposition 5.6. *Let Assumptions 2.1–2.10 hold. For any admissible optimal mild equilibrium S, $\overline{\mathfrak{S}_S}$ is also optimal mild. In addition, if $\overline{\mathfrak{S}_S}$ is admissible, then $\overline{\mathfrak{S}_S}$ is a weak equilibrium.*

Proof. *Step 1*. We first characterize $S \setminus \overline{\mathfrak{S}_S}$. As S is admissible, we can write S as a union of disjoint closed intervals

$$S = \cup_{n\in\Lambda_1}[\alpha_{2n-1}, \alpha_{2n}], \text{ where } \alpha_{2n-1} \le \alpha_{2n} < \alpha_{2n+1}.^{5} \tag{101}$$

where $\Lambda_1 \subset \mathbb{Z}$ is either a finite or countable subset. Since S is closed, we have that $\overline{\mathfrak{S}_S} \subset S$. For each $n \in \Lambda_1$, by the closeness of $\overline{\mathfrak{S}_S}$, we can see that $[\alpha_{2n-1}, \alpha_{2n}] \setminus \overline{\mathfrak{S}_S}$ consists of at most countably many disjoint intervals $(I_{n_k})_k$ of the following four forms:

$$1.\ [\alpha_{2n-1}, \gamma); \quad 2.\ (\gamma', \alpha_{2n}]; \quad 3.\ (\beta, \beta'); \quad 4.\ [\alpha_{2n-1}, \alpha_{2n}]. \tag{102}$$

For each I_{n_k} of the four forms in Equation (102), we define an open interval (l_{n_k}, r_{n_k}) as follows:

$$\begin{cases} 1.\ l_{n_k} := \sup\{y < \alpha_{2n-1}, y \in \mathfrak{S}_S\}, & r_{n_k} := \gamma; \\ 2.\ l_{n_k} := \gamma', & r_{n_k} := \inf\{y > \alpha_{2n}, y \in \mathfrak{S}_S\}; \\ 3.\ l_{n_k} = \beta, & r_{n_k} := \beta'; \\ 4.\ l_{n_k} := \sup\{y < \alpha_{2n-1}, y \in \mathfrak{S}_S\}, & r_{n_k} := \inf\{y > \alpha_{2n}, y \in \mathfrak{S}_S\}, \end{cases} \tag{103}$$

and set $\sup\emptyset := \inf\mathbb{X}$ and $\inf\emptyset := \sup\mathbb{X}$ if it happens. Notice that each two of those open intervals $((l_{n_k}, r_{n_k}))_{n,k}$ are either disjoint or identical, and l_{n_k} can be $-\infty$ (resp. r_{n_k} can be ∞). Since the total number of these intervals $((l_{n_k}, r_{n_k}))_{n,k}$ is at most countable, we omit the repeating ones and re-index them as $((l_k, r_k))_{k\in\Lambda}$ such that they are disjoint and $\Lambda \subset \mathbb{Z}$ is either a finite or countable subset. Then $S \setminus (\cup_{k\in\Lambda}(l_k, r_k)) = \overline{\mathfrak{S}_S}$.

Step 2. We prove that for each $k \in \Lambda$,

$$J(x, S \setminus (l_k, r_k)) = J(x, S), \quad \forall x \in (l_k, r_k). \tag{104}$$

Fix $k \in \Lambda$. *Step 1* tells that for any $x \in (l_k, r_k)$, x either belongs to $S \setminus \overline{\mathfrak{S}_S}$ or belongs to S^c.

(1) If $x \in S^c$ or $x \in \partial S$ for boundary case (b), Lemma 2.14 tells that

$$\mathcal{L}V(t, x+, S) \equiv \mathcal{L}V(t, x-, S) \equiv 0 \quad \forall t \in [0, \infty). \tag{105}$$

(2) Suppose $x \in S^\circ \setminus \overline{\mathfrak{S}_S}$. By the fact that S is an admissible optimal mild equilibrium, Theorem 4.3 tells that S is also weak. Then, Equation (25) together with the definition of $\mathfrak{S}_S$ leads to

$$\mathcal{L}V(0, x-, S) = \mathcal{L}V(0, x+, S) = 0; \quad V(t, x, S) = \delta(t)f(x) \quad \forall t \ge 0.$$

Then by a similar argument as in Equation (86) (with $\frac{a}{2}$ replaced by 0), we reach that

$$\mathcal{L}V(t, x-, S) \wedge \mathcal{L}V(t, x+, S) \ge 0 \quad \forall t \in [0, \infty). \tag{106}$$

(3) Otherwise, $x \in \partial S \setminus \overline{\mathfrak{S}_S}$ of boundary case (a), and for this case, we can also deduce Equation (106) by a combination of cases (1) and (2).

[5] These intervals are understood as the ones restricted in $\mathbb{X}$ in a natural way, for example, one $[\alpha_{2n-1}, \alpha_{2n}]$ could be $[\alpha_{2n-1}, \infty)$ if $\sup\mathbb{X} = \infty$.

In sum, we have

$$\frac{1}{2}(\mathcal{L}V(t,x-,S)+\mathcal{L}V(t,x+,S))\geq 0 \quad \forall (t,x)\in[0,\infty)\times(l_k,r_k). \tag{107}$$

Recall $(\theta_i)_{i\in I}$ defined in Assumption 2.8(ii). By Proposition 4.7 and the definition of $\mathfrak{S}_S$, $V_x(0,\theta_i+,S)=V_x(0,\theta_i-,S)$ for each $\theta_i\in((l_k,r_k)\cap S)$. Then for any $n\in\mathbb{N}$ and $\theta_i\in(l_k,r_k)\cap B(x_0,n)$, no matter θ_i belongs to $S\setminus\overline{\mathfrak{S}_S}$ or S^c, from the fact that $V(x,t,S)=\delta(t)f(x)$ for $x\in S$ and Lemma 3.10, we have that

$$V_x(t,\theta_i+,S)-V_x(t,\theta_i-,S)\geq\delta(t)(V_x(0,\theta_i+,S)-V_x(0,\theta_i-,S))=0, \quad \forall t\geq 0. \tag{108}$$

Note that for each $n\in\mathbb{N}$ the interval $B(x_0,n)\cap(l_k,r_k)$ contains at most finite points θ_i. Now take $x_0\in(l_k,r_k)$ and denote $\tau_n := \tau_{(l_k,r_k)\cap B(x_0,n)}\wedge n$ for $n\in\mathbb{N}$. By Lemma 2.15,

$$V(\tau_n,X_{\tau_n},S)-V(0,x_0,S)=\int_0^{\tau_n}\frac{1}{2}(\mathcal{L}V(t,X_t-,S)+\mathcal{L}V(t,X_t+,S))dt$$

$$+\int_0^{\tau_n}V_x(t,X_t,S)\sigma(X_s)\cdot 1_{\{X_t\neq\theta_i,\forall i\}}dW_t+\frac{1}{2}\sum_{\theta_i\in(l_k,r_k)}\int_0^{\tau_n}(V_x(t,\theta_i+,S)-V_x(t,\theta_i-,S))dL_t^{\theta_i},$$

Taking expectation for the above and combining with Equations (107) and (108), we have that

$$\mathbb{E}^{x_0}[V(\tau_n,X_{\tau_n},S)]-V(0,x_0,S)\geq 0.$$

Similar to Equation (74), we can show that $\lim_{n\to\infty}\mathbb{E}^{x_0}[V(\tau_n,X_{\tau_n},S)]=\mathbb{E}^{x_0}[V(\tau_{(l_k,r_k)},X_{\tau_{(l_k,r_k)}},S)]$. This together with the above inequality implies that

$$J(x_0,S\setminus(l_k,r_k))-J(x_0,S)=\mathbb{E}^{x_0}[V(\tau_{(l_k,r_k)},X_{\tau_{(l_k,r_k)}},S)]-V(0,x_0,S)\geq 0.$$

By the arbitrariness of $x_0\in(l_k,r_k)$, we have $J(x,S\setminus(l_k,r_k))\geq J(x,S)$ for all $x\in(l_k,r_k)$. Meanwhile, $J(x,S\setminus(l_k,r_k))=J(x,S)$ for $x\in\mathbb{X}\setminus(l_k,r_k)$. Then by the optimality of S, $S\setminus(l_k,r_k)$ is also an optimal mild equilibrium, and thus Equation (104) follows.

Step 3. We show that $\overline{\mathfrak{S}_S}$ is optimal mild. By *Step 2*, Equation (104) holds for all $k\in\Lambda$. From the construction of the intervals $(l_k,r_k)_{k\in\Lambda}$ in Equation (103), we can see that removing one of them does not change the values of function J on the rest parts, that is, for any $k\in\Lambda$,

$$J(x,\overline{\mathfrak{S}_S})=J(x,S\setminus(l_k,r_k)) \quad \forall x\in(l_k,r_k).$$

Hence, we can conclude that for any $k\in\Lambda$,

$$J(x,\overline{\mathfrak{S}_S})=J(x,S\setminus(l_k,r_k))=J(x,S), \quad \forall x\in(l_k,r_k).$$

As $J(x,\overline{\mathfrak{S}_S})=f(x)=J(x,S)$ for all $x\in\overline{\mathfrak{S}_S}$,

$$J(x,\overline{\mathfrak{S}_S})=J(x,S), \quad \forall x\in\mathbb{X}.$$

This implies that $\overline{\mathfrak{S}_S}$ is an optimal mild equilibrium. By Theorem 4.3, if $\overline{\mathfrak{S}_S}$ is admissible then it is also a weak equilibrium. □

Thanks to Theorem 5.1 and Proposition 5.6, we are ready to prove Theorem 5.2.

Proof of Theorem 5.2. **Part (a)**: Suppose S is an optimal mild equilibrium and $\mathfrak{S}_S$ is closed and admissible. Proposition 5.6 tells that $\mathfrak{S}_S = \overline{\mathfrak{S}_S}$ is both an optimal mild and weak equilibrium. Then by Theorem 5.1, to prove that $\mathfrak{S}_S$ is strong, it is sufficient to verify that $\mathfrak{S}_S = \mathfrak{S}_{\mathfrak{S}_S}$. Notice that $\mathfrak{S}_{\mathfrak{S}_S} \subset \mathfrak{S}_S$. Take $x_0 \in \mathfrak{S}_S$ and we show $x_0 \in \mathfrak{S}_{\mathfrak{S}_S}$. If $V_x(0, x_0, \mathfrak{S}_S) > V_x(0, x_0, \mathfrak{S}_S)$, then $x_0 \in \mathfrak{S}_{\mathfrak{S}_S}$. Otherwise, $V_x(0, x_0, \mathfrak{S}_S) = V_x(0, x_0, \mathfrak{S}_S)$, and it remains to verify that

$$\mathcal{L}V(0, x_0-, \mathfrak{S}_S) \wedge \mathcal{L}V(0, x_0+, \mathfrak{S}_S) < 0. \tag{109}$$

Since both S and $\mathfrak{S}_S$ are optimal mild, we have

$$V(0, x, \mathfrak{S}_S) \equiv J(x, \mathfrak{S}_S) \equiv J(x, \overline{\mathfrak{S}_S}) \equiv J(x, S) \equiv V(0, x, S) \quad \forall x \in \mathbb{X}.$$

Then,

$$V_x(0, x_0-, S) - V_x(0, x_0+, S) = V_x(0, x_0-, \mathfrak{S}_S) - V_x(0, x_0+, \mathfrak{S}_S) = 0. \tag{110}$$

Since $x_0 \in \mathfrak{S}_S$, by the definition of $\mathfrak{S}_S$, Equation (110) leads to that

$$\mathcal{L}V(0, x_0-, S) \wedge \mathcal{L}V(0, x_0+, S) < 0. \tag{111}$$

This together with Equation (22) implies that x_0 cannot be an isolated point of S. We consider the following two cases.

(1) Suppose $x_0 \in S^\circ$. Note that $V(t, x, S) = \delta(t)f(x)$ on S. Then by Equation (111), without loss of generality, we assume $\mathcal{L}V(0, x_0+, S) = \mathcal{L}(\delta f)(0, x_0+) < 0$. By the right continuity of $x \mapsto \mathcal{L}(\delta f)(0, x_+)$ at x_0, we can find $h > 0$ small enough such that $[x_0, x_0 + h) \subset (S^\circ \cap \mathcal{G})$ (recall $\mathcal{G}$ defined in Equation 16) and

$$\mathcal{L}(\delta f)(0, x_+) < 0, \quad \forall x \in [x_0, x_0 + h). \tag{112}$$

Hence, $[x_0, x_0 + h) \subset \mathfrak{S}_S$, and thus $\mathcal{L}V(0, x_0+, \mathfrak{S}_S) = \mathcal{L}(\delta f)(0, x_+) < 0$.

(2) Otherwise, $x_0 \in \partial(S^\circ)$ for boundary case (a). Without loss of generality, we assume $(x_0, x_0 + h) \subset (S^\circ \cap \mathcal{G})$ for $h > 0$ small enough. Then by Equations (111) and (22), we again have $\mathcal{L}V(0, x_0+, S) = \mathcal{L}(\delta f)(0, x_0+) < 0$. A similar discussion as in case (1) implies $\mathcal{L}V(0, x_0+, \mathfrak{S}_S) = \mathcal{L}(\delta f)(0, x_0+) < 0$.

In sum, Equation (109) holds, and the proof of part (a) is complete.

Part (b): Lemma 4.1 indicates that S^* is an optimal mild equilibrium. Then by Proposition 5.6, $\overline{\mathfrak{S}_{S^*}} \subset S^*$ is an optimal mild equilibrium. As S^* is the smallest optimal mild equilibrium, $\overline{\mathfrak{S}_{S^*}} = S^*$. The rest statement directly follows from part (a). □

6 | EXAMPLES

In this section, we provide three examples to demonstrate our results. In the first example, we have two strong equilibria, one of which is not optimal mild. This indicates that an strong equilibrium may not be optimal mild. In the second example, we show that a weak equilibrium may not be strong. The third example is the stopping for an American put option on a geometric Brownian motion, in which we provide all three types of equilibria.

6.1 An example showing optimal mild $\subsetneqq$ strong

In this subsection, we construct an example where the set of optimal mild equilibria is strictly contained (i.e., $\subsetneqq$) in the set of strong equilibria. Let $dX_t = dW_t$ and thus X is a Brownian motion with $\mathbb{X} = \mathbb{R}$. Take discount function $\delta(t) = \frac{1}{1+\beta t}$. Let $a < b$, $0 < c < d$ such that

$$\frac{\int_0^\infty e^{-s} \frac{\sqrt{2\beta s}}{\sinh((b-a)\sqrt{2\beta s})} ds}{\sqrt{\frac{\pi\beta}{2}} + \int_0^\infty e^{-s}\sqrt{2\beta s}\coth((b-a)\sqrt{2\beta s})ds} < \frac{c}{d} < \int_0^\infty e^{-(s+(b-a)\sqrt{2\beta s})}ds. \tag{113}$$

Notice that such parameters do exist, for example, let $b - a = 1$, then for Equation (113), we have LHS $\approx 0.3952 < \frac{c}{d} < 0.4544 \approx$ RHS.

Define

$$\begin{aligned} J_b(x) :&= d\mathbb{E}^x[\delta(\rho_{\{b\}})] = d\int_0^\infty \frac{p(t)}{1+\beta t}dt = d\int_0^\infty \int_0^\infty e^{-(1+\beta t)s}p(t)dsdt \\ &= d\int_0^\infty e^{-s}\mathbb{E}^x[e^{-\beta s \rho_{\{b\}}}]ds = d\int_0^\infty e^{-s}e^{-|x-b|\sqrt{2\beta s}}ds, \quad x \in \mathbb{X}. \end{aligned} \tag{114}$$

where the second line uses the formula in Borodin and Salminen (2002, 2.0.1 on page 204). We further define

$$\begin{aligned} J_{ab}(x) :&= c\mathbb{E}^x[\delta(\rho_{\{a,b\}})\cdot 1_{\{\rho_{\{a,b\}}=a\}}] + d\mathbb{E}^x[\delta(\rho_{\{a,b\}})\cdot 1_{\{\rho_{\{a,b\}}=b\}}] \\ &= \begin{cases} c\int_0^\infty e^{-s}e^{-|x-a|\sqrt{2\beta s}}ds, & x < a, \\ c\int_0^\infty e^{-s}\frac{\sinh((b-x)\sqrt{2\beta s})}{\sinh((b-a)\sqrt{2\beta s})}ds + d\int_0^\infty e^{-s}\frac{\sinh((x-a)\sqrt{2\beta s})}{\sinh((b-a)\sqrt{2\beta s})}ds, & a \le x \le b, \\ J_b(x), & x > b. \end{cases} \end{aligned} \tag{115}$$

where the expression for J_{ab} on $[a, b]$ is obtained by the formula in Borodin and Salminen (2002, 3.0.5 (a) and (b) on page 218) combined with an argument similar to that in Equation (114). Let f be any function satisfying Assumption 2.8 such that

$$f(a) = c, f(b) = d; \quad f(x) < \min\{J_b(x), J_{ab}(x)\}, \quad \forall\, x \in \mathbb{X}\setminus\{a,b\}. \tag{116}$$

Note that

$$c < d\int_0^\infty e^{-(s+(b-a)\sqrt{2\beta s})}ds = J_b(a), \tag{117}$$

which shows such function $f(x)$ indeed exists.[6]

One can easily verify that Assumptions 2.1–2.8 hold. Moreover, Assumption 2.10 is also satisfied due to Lemma 2.12 and Remark 2.13. We have the following result.

Proposition 6.1. *$\{b\}$ is the unique optimal mild equilibrium, while both $\{b\}$ and $\{a, b\}$ are strong equilibria.*

Proof. Recall S^* defined in Equation (58). First notice that

$$J_b(x) = J(x, \{b\}) \quad \text{and} \quad J_{ab}(x) = J(x, \{a, b\}), \quad \forall\, x \in \mathbb{X}.$$

Then by Equations (116) and (117), it is easy to see that both $\{a, b\}$ and $\{b\}$ are mild equilibria. Since b is the global maximum of f, any mild equilibrium must contain b. Therefore, $\{b\}$ is the smallest mild equilibrium, that is, $S^* = \{b\}$. It then follows from Lemma 4.1 that $\{b\}$ is optimal mild. Moreover, by Equations (116) and (117) again, we have that $f(x) < J(x, \{b\})$ for any $x \neq b$, which implies that $\{b\}$ is the unique optimal mild equilibrium.

Now we verify that both $\{b\}$ and $\{a, b\}$ are strong equilibria. As for the optimal mild equilibrium $\{b\}$, a direct calculation from Equation (114) shows that

$$J'(b-, \{b\}) = d \int_0^\infty e^{-s} \sqrt{2\beta s} ds > 0,$$

and by symmetry, we have $J'(b+, \{b\}) < 0$. Then,

$$V_x(0, b-, \{b\}) - V_x(0, b+, \{b\}) = J'(b-, \{b\}) - J'(b+, \{b\}) > 0.$$

Meanwhile, by Lemma 2.14(a), we have $\mathcal{L}V(t, x\pm, \{b\}) \equiv 0$ on $\mathbb{X}$. Therefore, we have $\mathfrak{S}_{\{b\}} = \{b\}$ from Equation (88). Since $\{b\}$ is closed and admissible, Theorem 5.2(b) tells that $\{b\}$ is a strong equilibrium. Now consider the mild equilibirum $\{a, b\}$. Direct calculations from Equation (115) show that

$$J'(a-, \{a, b\}) = c \int_0^\infty e^{-s} \sqrt{2\beta s} ds = \sqrt{\frac{\pi\beta}{2}} c,$$

and for any $x \in (a, b)$,

$$J'(x, \{a, b\}) = -c \int_0^\infty e^{-s} \frac{\cosh((b-x)\sqrt{2\beta s})\sqrt{2\beta s}}{\sinh((b-a)\sqrt{2\beta s})} ds + d \int_0^\infty e^{-s} \frac{\cosh((x-a)\sqrt{2\beta s})\sqrt{2\beta s}}{\sinh((b-a)\sqrt{2\beta s})} ds. \tag{118}$$

[6] By the strong Markov property of X and Equation (117), one can easily check that $J_{ab}(x) < J_b(x)$ for $x < b$. Hence, a quick example for such f would be

$$f(x) := \begin{cases} \frac{1}{1+(a-x)} J_{ab}(x), & x \leq a, \\ \frac{1}{1+(x-a)(b-x)} J_{ab}(x), & a < x \leq b, \\ \frac{1}{1+(x-b)} J_b(x), & x > b. \end{cases}$$

By taking $x = a+$ in Equation (118) and combining with the first inequality in Equation (113), we have that

$$J'(a+,\{a,b\}) = -c\int_0^\infty e^{-s}\coth((b-a)\sqrt{2\beta s})\sqrt{2\beta s}ds + d\int_0^\infty e^{-s}\frac{\sqrt{2\beta s}}{\sinh((b-a)\sqrt{2\beta s})}ds$$
$$<\sqrt{\frac{\pi\beta}{2}}c = J'(a-,\{a,b\}).$$

By taking $x = b-$ in Equation (118) and the fact that $0 < c < d$, we have that

$$J'(b-,\{a,b\}) = -c\int_0^\infty e^{-s}\frac{\sqrt{2\beta s}}{\sinh((b-a)\sqrt{2\beta s})}ds + d\int_0^\infty e^{-s}\coth((b-a)\sqrt{2\beta s})\sqrt{2\beta s}ds$$
$$>0 > J'(b+,\{b\}) = J'(b+,\{a,b\}).$$

Hence,

$$V_x(0,x-,\{a,b\}) > V_x(0,x+,\{a,b\}), \text{ for both } x = a,b. \tag{119}$$

Meanwhile, Lemma 2.14(a) tells that $\mathcal{L}V(t,x,\{a,b\}) \equiv 0$ on $\mathbb{X}\setminus\{a,b\}$. Therefore, by Theorem 3.1, $\{a,b\}$ is a weak equilibrium. Moreover, by Equations (119) and (88), $\mathfrak{S}_{\{a,b\}} = \{a,b\}$. It then follows from Theorem 5.1 that $\{a,b\}$ is a strong equilibrium. □

6.2 | An example showing strong $\subsetneqq$ weak

In this subsection, we give an example in which a weak equilibrium is not strong, and thus {strong equilibria} $\subsetneqq$ {weak equilibria}. Let X be a geometric Brownian motion:

$$dX_t = \mu X_t dt + \sigma X_t dW_t \tag{120}$$

with $\mathbb{X} = (0,\infty)$. Let $\delta(t) = \frac{1}{1+\beta t}$ and $f(x) = x \wedge K$ for some constant $K > 0$. Assume that $\mu = \beta > 0$.

Proposition 6.2. $(0,\infty)$ *is a weak equilibrium but not strong, while* $[K,\infty)$ *is the unique optimal mild equilibrium and a strong equilibrium.*

Proof. We first verify the result for $(0,\infty)$. Notice that $V(t,x,(0,\infty)) \equiv \delta(t)f(x)$, then direct calculations show

$$\begin{cases} \mathcal{L}V(0,x,(0,\infty)) = (-\beta+\mu)x = 0, & 0 < x < K, \\ \mathcal{L}V(0,x,(0,\infty)) = -\beta K < 0, & x > K, \\ V_x(0,K-,(0,\infty)) = 1 > 0 = V_x(0,K+,(0,\infty)). \end{cases}$$

Therefore, by Theorem 3.1, $(0,\infty)$ is a weak equilibrium. For $x \in (0,K)$, we have that for $\varepsilon > 0$ small enough,

$$\mathbb{E}[\delta(\rho^\varepsilon_{(0,\infty)})f(X_{\rho^\varepsilon_{(0,\infty)}})] > \mathbb{E}\left[\delta(\varepsilon)X_\varepsilon 1_{\{X_\varepsilon \le K\}}\right] = \frac{e^{\mu\varepsilon}}{1+\mu\varepsilon}N(d_\varepsilon)\cdot x$$

$$\geq \left(1+\mu\varepsilon+\frac{1}{2}\mu^2\varepsilon^2+o(\varepsilon^2)\right)\left(1-\mu\varepsilon+\mu^2\varepsilon^2+o(\varepsilon^2)\right)\left(1-\frac{1}{\sqrt{2\pi}d_\varepsilon}e^{-d_\varepsilon^2/2}\right)\cdot x$$
$$\geq \left(1+\frac{1}{2}\mu^2\varepsilon^2+o(\varepsilon^2)\right)(1+o(\varepsilon^2))\cdot x=\left(1+\frac{1}{2}\mu^2\varepsilon^2+o(\varepsilon^2)\right)x>x,$$

where $d_\varepsilon := \frac{\ln(K/x)-(\mu+\frac{1}{2}\sigma^2)\varepsilon}{\sigma\sqrt{\varepsilon}}$. This indicates that $(0,\infty)$ is not a strong equilibrium.

Now we verify the result for $[K,\infty)$. By Ito's formula,

$$d\left(\frac{X_t}{1+\beta t}\right)=\frac{X_t}{1+\beta t}\left(-\frac{\beta}{1+\beta t}+\mu\right)dt+\frac{\sigma X_t}{1+\beta t}dW_t=\frac{\mu^2 tX_t}{(1+\mu t)^2}dt+\frac{\sigma X_t}{1+\mu t}dW_t.$$

Then, by the facts that $\rho_{[K,\infty)}>0$ $\mathbb{P}^x$-a.s. and $X_t>0$ for $x\in(0,K)$, we have that

$$J(x,[K,\infty))-f(x)=\mathbb{E}^x\left[\frac{X_{\rho_{[K,\infty)}}}{1+\beta\rho_{[K,\infty)})}\right]-x=\mathbb{E}^x\left[\int_0^{\rho_{[K,\infty)}}\frac{\mu^2 tX_t}{(1+\mu t)^2}dt\right]>0,\quad \forall x\in(0,K), \tag{121}$$

which shows that $[K,\infty)$ is a mild equilibrium. On the other hand, since $[K,\infty)$ is the set of global maxima of f, any mild equilibrium S must contain $[K,\infty)$, for otherwise, $J(x,S)<K=f(x)$ for any $x\in S^c\cap[K,\infty)$, a contradiction. Therefore, $[K,\infty)$ is the smallest mild equilibrium and thus optimal. Now for any mild equilibrium S such that $S\setminus[K,\infty)\neq\emptyset$, Equation (121) indicates that $f(x)<J(x,[K,\infty))$ on $S\setminus[K,\infty)$, which implies that S is not an optimal mild equilibrium. Hence, $[K,\infty)$ is the unique optimal mild equilibrium. Moreover, direct calculation shows that

$$\mathcal{L}V(0,x+,[K,\infty))=\mathcal{L}(\delta(t)K)=-\beta K,\quad \forall x\in[K,\infty),$$

which tells that $\mathfrak{S}_{[K,\infty)}=[K,\infty)$. Then by Theorem 5.2(b), $[K,\infty)$ is also strong. □

6.3 | Stopping of an American put option

Consider the American put example in Huang and Zhou (2020, Section 6.3). In particular, X is a geometric Brownian motion given by Equation (120) with $\mathbb{X}:=(0,\infty)$. Let $\mu\geq 0$. The payoff function is defined as $f(x):=(K-x)^+$, and $\delta(t):=\frac{1}{1+\beta t}$. We shall provide all three types of equilibria. To begin with, the following lemma summarizes the results of mild equilibria stated in Lemma 6.11, Corollary 6.13, and Proposition 6.15 in Huang and Zhou (2020).

Lemma 6.3.

(i) *If S is a mild equilibrium, then $S\cap(0,K]=(0,a]$ for some $a\in(0,K]$.*
(ii) *$S=(0,a]\subset(0,K]$ is mild equilibrium if and only if $a\geq\frac{\lambda}{1+\lambda}K$, where*

$$\lambda:=\int_0^\infty e^{-s}\left(\sqrt{\nu^2+2\beta s/\sigma^2}+\nu\right)>0,\quad \nu:=\frac{u}{\sigma^2}-\frac{1}{2}. \tag{122}$$

(iii) $S^* = (0, \frac{\lambda}{1+\lambda}K]$ *is the intersection of all mild equilibria and is the unique optimal mild equilibria.*

Following from Lemma 6.3(ii), we shall call the mild equilibria that belong to the family $\{(0,a] : a \geq \frac{\lambda}{1+\lambda}K\}$ are "type I" mild equilibria.[7] The following proposition shows that, except "type I" mild equilibria, all other mild equilibria take the same form: $(0,a] \cup D$ that satisfies a certain condition, and we shall call this family of mild equilibria "type II" mild equilibria.

Proposition 6.4. *Except the "type I" mild equilibria in Lemma 6.3(ii), all other mild equilibria take form:* $(0,a] \cup D$ *such that*

$$-(K-a)\int_0^\infty e^{-s}\left(\frac{\nu}{a} + \frac{\sqrt{\nu^2+2\beta s/\sigma^2}}{a}\cdot\frac{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}}+(a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}\right)ds \geq -1, \tag{123}$$

where D *is a closed subset of* $[K,\infty)$ *and* $b := \inf\{x \in D\}$ *satisfying* $b > a$.

Proof. Lemma 6.3(i)(ii) together imply that any mild equilibrium is either of type I or takes the form: $(0,a] \cup D$ with D being a closed subset of $[K,\infty)$. Consider a closed set of such form $S = (0,a] \cup D$ with $b := \inf\{x \in D\} > a$. When $a \geq K$, the fact that $f = 0$ on $[K,\infty)$ immediately gives that S is a mild equilibrium. Notice that $-(K-a) \geq 0$ and the integrand in the LHS of Equation (123) is positive, so Equation (123) holds.

When $a < K$, we have

$$\begin{aligned}
J(x,S) &= (K-a)\int_0^\infty \frac{p(t)}{1+\beta t}dt = (K-a)\int_0^\infty e^{-s}\mathbb{E}^x[e^{-\beta s\tau_{(a,b)}}\cdot 1_{\{X_{\tau_{(a,b)}}=a\}}]ds \\
&= (K-a)\int_0^\infty e^{-s}\left(\frac{a}{x}\right)^\nu \frac{(b/x)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(x/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}ds, \quad \forall x \in (a,b),
\end{aligned}$$

where $p(t) := \mathbb{P}^x(\tau_{(a,b)} \in dt, X_{\tau_{(a,b)}} = a)$, and the second line above follows from Borodin and Salminen (2002, 3.0.5 (a) on page 633). Direct calculations show that for any $x \in (a,b)$

$$\begin{aligned}
J'(x,S) = -(K-a)\int_0^\infty e^{-s}\Bigg(&\frac{a^\nu\nu}{x^{\nu+1}}\cdot\frac{(b/x)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(x/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}} \\
&+\frac{a^\nu\sqrt{\nu^2+2\beta s/\sigma^2}}{x^{\nu+1}}\cdot\frac{(b/x)^{\sqrt{\nu^2+2\beta s/\sigma^2}}+(x/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}\Bigg)ds,
\end{aligned} \tag{124}$$

$$\begin{aligned}
J''(x,S) = (K-a)\int_0^\infty e^{-s}\Bigg(&\frac{a^\nu(2\nu^2+\nu+2\beta s/\sigma^2)}{x^{\nu+2}}\cdot\frac{(b/x)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(x/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}} \\
&+\frac{a^\nu(2\nu+1)\sqrt{\nu^2+2\beta s/\sigma^2}}{x^{\nu+2}}\cdot\frac{(b/x)^{\sqrt{\nu^2+2\beta s/\sigma^2}}+(x/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}}-(a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}\Bigg)ds.
\end{aligned} \tag{125}$$

[7] It contains the trivial mild equilibrium $\mathbb{X}$ by setting $a = \infty$ and $\mathbb{X} = \mathbb{X} \cap (0,\infty]$.

Recall ν in Equation (122), we have that

$$\nu + \sqrt{\nu^2 + 2\beta s/\sigma^2} > 0 \quad \text{and} \quad (2\nu^2 + \nu + 2\beta s/\sigma^2) + \left((2\nu + 1)\sqrt{\nu^2 + 2\beta s/\sigma^2}\right) > 0.$$

This together with

$$0 < (b/x)^{\sqrt{\nu^2+2\beta s/\sigma^2}} - (x/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}} < (b/x)^{\sqrt{\nu^2+2\beta s/\sigma^2}} + (x/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}$$

implies that both the integrands on the RHS of Equations (124) and (125) are positive. Therefore, $J'(x,S) < 0$ and $J''(x,S) > 0$ for $x \in (a,b)$, and thus $J(x,S)$ is strictly decreasing and convex on (a,b). This together with the shape of f on (a,b) indicates that S is a mild equilibrium if and only if $J'(a+,S) \geq -1$. From Equation (124), we have

$$J'(a+,S) = -(K-a)\int_0^\infty e^{-s}\left(\frac{\nu}{a} + \frac{\sqrt{\nu^2+2\beta s/\sigma^2}}{a} \cdot \frac{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}} + (a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}} - (a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}\right)ds,$$

so S is a mild equilibrium if and only if Equation (123) holds. Notice that $J'(a+,S)$ converges to 0 when $a \nearrow K$. Then for any $b > K$, by the continuity of function $a \mapsto J'(a+,S)$, there exists a constant $a_b < K$ such that for all $a \in [a_b, K)$, Equation (123) indeed holds and S is a mild equilibrium. □

Proposition 6.5. $S^* = (0, \frac{\lambda}{1+\lambda}K]$ *is the unique weak and the unique strong equilibrium.*

Proof. We first find all weak equilibria. Since a weak equilibrium is also mild, by Proposition 6.4, it is sufficient to select weak equilibria from the two types of mild equilibria. Given a mild equilibrium S that is weak, no matter which type it is, S must not contain K. Otherwise, by Lemma 6.3(i), $(0,K] \subset S$, which together with $f = 0$ on $[K,\infty)$ implies that

$$V_x(0,K-,S) = -1 < 0 = V_x(0,K+,S),$$

which contradicts Equation (24) in Theorem 3.1.

Consider an arbitrary type I mild equilibrium $(0,a]$ with $\frac{\lambda}{1+\lambda}K \leq a < K$. By the smooth-fit condition in Corollary 3.2, $S = (0,a]$ is a weak equilibrium if and only if

$$V_x(0,a+,(0,a]) = J'(a+,(0,a]) = -1.$$

From the calculation in the proof of Lemma 6.12 in Huang and Zhou (2020), such condition is satisfied if and only if $a = \frac{\lambda}{1+\lambda}K$. Hence, $S^* = (0, \frac{\lambda}{1+\lambda}K]$ is the only weak equilibrium among the type I mild equilibria. Now pick any type II mild equilibrium $S = (0,a] \cup D$ with $b := \inf\{x \in D\}$. As $K \notin S$, we have $a < K$. Then by Equation (124), we have

$$V_x(0,b-,S) = -\frac{K-a}{b^{\nu+1}}\int_0^\infty e^{-s}\frac{2a^\nu\sqrt{\nu^2+2\beta s/\sigma^2}}{(b/a)^{\sqrt{\nu^2+2\beta s/\sigma^2}} - (a/b)^{\sqrt{\nu^2+2\beta s/\sigma^2}}}ds < 0 = V_x(0,b+,S).$$

That is, the smooth-fit condition fails at the boundary $x = b$, and hence S is not weak. In sum, $S^* = (0, \frac{\lambda}{1+\lambda}K]$ is the unique weak equilibrium.

Finally, a direct calculation shows that

$$\mathcal{L}V(0,x-,S^*) = -\beta(K-x) - \mu x < 0, \quad \forall x \in \left(0, \frac{\lambda}{1+\lambda}K\right],$$

so $S^* = \mathfrak{S}_{S^*}$. Then, by Theorem 5.1 and the fact that S^* is the unique weak equilibrium, we can conclude that S^* is the unique strong equilibrium. □

Remark 6.6. Within this example, we do not restrict equilibria to be admissible. The unique weak, strong, optimal mild equilibrium $(0, \frac{\lambda}{1+\lambda}K]$ turns out to be indeed admissible. Moreover, type I mild equilibria are all admissible, while any type II mild equilibrium $S = (0, a] \cup D$ with $b := \inf\{x \in D\} > a$ has an alternative $(0, a] \cup [b, \infty)$, which share the same J value and is admissible.

ACKNOWLEDGMENTS

E. Bayraktar is partially supported by the National Science Foundation under Grant DMS2106556 and by the Susan M. Smith chair.

DATA AVAILABILITY STATEMENT

Not applicable since no data were used.

ORCID

Erhan Bayraktar https://orcid.org/0000-0002-1926-4570
Zhou Zhou https://orcid.org/0000-0001-8092-4745

REFERENCES

Agram, N., & Djehiche, B. (2021). Reflected Backward Stochastic Volterra Integral Equations and related time-inconsistent optimal stopping problems. *Systems & Control Letters*, 155(2021), 104989.

Bayraktar, E., Zhang, J., & Zhou, Z. (2021). Equilibrium concepts for time-inconsistent stopping problems in continuous time. *Mathematical Finance*, *31*(1), 508–530.

Björk, T., Khapko, M., & Murgoci, A. (2017). On time-inconsistent stochastic control in continuous time. *Finance and Stochastics*, *21*(2), 331–360.

Björk, T., Khapko, M., & Murgoci, A. (2021). *Time-inconsistent control theory with finance applications*. *Springer finance*. Springer.

Bodnariu, A., Christensen, S., & Lindensjö, K. (2022). Local time pushed mixed stopping and smooth fit for time-inconsistent stopping problems. *arXiv preprint arXiv:2206.15124*.

Borodin, A. N., & Salminen, P. (2002). *Handbook of Brownian motion—facts and formulae*. *Probability and its applications* (2nd ed.). Birkhäuser Verlag.

Christensen, S., & Lindensjö, K. (2018). On finding equilibrium stopping times for time-inconsistent Markovian problems. *SIAM Journal on Control and Optimization*, *56*(6), 4228–4255.

Christensen, S., & Lindensjö, K. (2020a). On time-inconsistent stopping problems and mixed strategy stopping times. *Stochastic Processes and their Applications*, *130*(5), 2886–2917.

Christensen, S., & Lindensjö, K. (2020b). Time-inconsistent stopping, myopic adjustment and equilibrium stability: With a mean-variance application. In *Stochastic modeling and control*, *Banach Center Publications* (Vol. 122, pp. 53–76). Polish Academy of Sciences, Institute of Mathematics.

Ebert, S., & Strack, P. (2018). Never, ever getting started: On prospect theory without commitment. *Available at SSRN 2765550*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765550

Ebert, S., Wei, W., & Zhou, X. Y. (2020). Weighted discounting—on group diversity, time-inconsistency, and consequences for investment. *Journal of Economic Theory*, *189*, 105089.

Ekeland, I., & Lazrak, A. (2006). Being serious about non-commitment: Subgame perfect equilibrium in continuous time. *arXiv preprint math/0604264*.

Ekeland, I., & Lazrak, A. (2010). The golden rule when preferences are time inconsistent. *Mathematics and Financial Economics*, *4*(1), 29–55.

Ekeland, I., & Pirvu, T. A. (2008). Investment and consumption without commitment. *Mathematics and Financial Economics*, *2*(1), 57–86.

Hamaguchi, Y. (2021). Extended backward stochastic Volterra integral equations and their applications to time-inconsistent stochastic recursive control problems. *Mathematical Control and Related Fields*, *11*(2), 433–478.

He, X. D., & Jiang, Z. L. (2021). On the equilibrium strategies for time-inconsistent problems in continuous time. *SIAM Journal on Control and Optimization*, *59*(5), 3860–3886.

He, X. D., & Zhou, X. Y. (2022). Who are I: Time inconsistency and intrapersonal conflict and reconciliation. In *Stochastic analysis, filtering, and stochastic optimization* (pp. 177–208). Springer.

Hernández, C., & Possamaï, D. (2020). Me, myself and I: A general theory of non-Markovian time-inconsistent stochastic control for sophisticated agents. *Ann. Appl. Probab.*, *33*(2), 1196–1258 (April 2023). https://doi.org/10.1214/22-AAP1845

Huang, Y.-J., & Nguyen-Huu, A. (2018). Time-consistent stopping under decreasing impatience. *Finance and Stochastics*, *22*(1), 69–95.

Huang, Y.-J., Nguyen-Huu, A., & Zhou, X. Y. (2020). General stopping behaviors of naïve and noncommitted sophisticated agents, with application to probability distortion. *Mathematical Finance*, *30*(1), 310–340.

Huang, Y.-J., & Wang, Z. (2021). Optimal equilibria for multidimensional time-inconsistent stopping problems. *SIAM Journal on Control and Optimization*, *59*(2), 1705–1729.

Huang, Y.-J., & Yu, X. (2021). Optimal stopping under model ambiguity: A time-consistent equilibrium approach. *Mathematical Finance*, *31*(3), 979–1012.

Huang, Y.-J., & Zhou, Z. (2019). The optimal equilibrium for time-inconsistent stopping problems—the discrete-time case. *SIAM Journal on Control and Optimization*, *57*(1), 590–609.

Huang, Y.-J., & Zhou, Z. (2020). Optimal equilibria for time-inconsistent stopping problems in continuous time. *Mathematical Finance*, *30*(3), 1103–1134.

Huang, Y.-J., & Zhou, Z. (2021). Strong and weak equilibria for time-inconsistent stochastic control in continuous time. *Mathematics of Operations Research*, *46*(2), 428–451.

Karatzas, I., & Shreve, S. E. (1991). *Brownian motion and stochastic calculus*. *Graduate texts in mathematics* (2nd ed., Vol. 113). Springer-Verlag.

Liang, Z., & Yuan, F. (2021). Weak equilibriums for time-inconsistent stopping control problems. *arXiv preprint arXiv:2105.06607*.

Miller, C. W. (2017). Nonlinear PDE approach to time-inconsistent optimal stopping. *SIAM Journal on Control and Optimization*, *55*(1), 557–573.

Peskir, G. (2007). A change-of-variable formula with local time on surfaces. In *Séminaire de probabilités XL*. *Lecture notes in mathematics* (Vol. 1899, pp. 69–96). Springer.

Peskir, G., & Shiryaev, A. (2006). *Optimal stopping and free-boundary problems*. *Lectures in mathematics. ETH Zürich*. Birkhäuser Verlag.

Tan, K. S., Wei, W., & Zhou, X. Y. (2021). Failure of smooth pasting principle and nonexistence of equilibrium stopping rules under time-inconsistency. *SIAM Journal on Control and Optimization*, *59*(6), 4136–4154.

Wang, H., & Yong, J. (2021). Time-inconsistent stochastic optimal control problems and backward stochastic Volterra integral equations. *ESAIM: Control, Optimisation and Calculus of Variations*, *27*(22), 40.

Wei, Q., Yong, J., & Yu, Z. (2017). Time-inconsistent recursive stochastic optimal control problems. *SIAM Journal on Control and Optimization*, *55*(6), 4156–4201.

How to cite this article: Bayraktar, E., Wang, Z., & Zhou, Z. (2023). Equilibria of time-inconsistent stopping for one-dimensional diffusion processes. *Mathematical Finance*, *33*, 797–841. https://doi.org/10.1111/mafi.12385

APPENDIX A: PROOF FOR RESULTS IN SECTION 2

Proof (of (12) in Remark 2.2). Let $X_0 = x \in \mathbb{X}$ and $h > 0$ be small enough such that $[x-h, x+h] \subset \mathbb{X}$. Let $Y = (Y_t)_{t\geq 0}$ follows $dY_t = \hat{\mu}(Y_t)dt + \hat{\sigma}(Y_t)dW_t$ with $Y_0 = x$, where

$$\hat{\mu}(y) := \begin{cases} \mu(y), & x-h \leq y \leq x+h, \\ \mu(x-h), & y < x-h, \\ \mu(x+h), & y > x+h, \end{cases} \quad \text{and} \quad \hat{\sigma}(y) := \begin{cases} \sigma(y), & x-h \leq y \leq x+h, \\ \sigma(x-h), & y < x-h, \\ \sigma(x+h), & y > x+h. \end{cases}$$

Then by Huang et al. (2020, Lemma A.1), for any $t > 0$,

$$\mathbb{P}^x\left(\max_{0\leq s\leq t} Y_s > x\right) = \mathbb{P}^x\left(\min_{0\leq s\leq t} Y_s < x\right) = 1.$$

Note that $Y_s = X_s$ for $s \leq \tau_{B(x,h)}$. Then for a.s. $\omega \in \{\tau_{B(x,h)} > 1/n\}$,

$$\max_{0\leq s\leq t} X_s(\omega) > x \quad \text{and} \quad \min_{0\leq s\leq t} X_s(\omega) < x, \quad \forall\, t \in (0, 1/n) \text{ and thus } \forall\, t > 0.$$

Then Equation (12) follows from the arbitrariness of $n \in \mathbb{N}$. □

Proof of Lemma 2.7. By Equation (13), for any $t, r \geq 0$

$$\delta(t+r) - \delta(t) \geq \delta(t)(\delta(r) - \delta(0)),$$

This together with the differentiability of $\delta(t)$ implies that $\delta'(t) \geq \delta(t)\delta'(0)$. As $\delta'(t) \leq 0$,

$$1 - \delta(t) = \int_0^t -\delta'(s)ds \leq \int_0^t -\delta(s)\delta'(0)ds \leq \int_0^t |\delta'(0)|ds = |\delta'(0)|t.$$

□

Proof of Lemma 2.12. Take an admissible stopping policy S. Let $a, b \in \mathbb{X}$ such that $[a,b] \subset \mathbb{X}$ and $(a,b) \subset S^c$. Throughout the proof, $C > 0$ will serve as a generic constant that may change from one line to another and is independent of r.

Set $v(x,r,S) := \mathbb{E}^x[e^{-r\rho_S} f(X_{\rho_S})]$. We first provide an estimate for $|v_x(x,r,S)| + |v_{xx}(x,r,S)|$ on $[a,b]$. Assumption 2.1(i) and the boundedness of f gives the well-posedness of $v(\cdot,r,S)$ for all $r \geq 0$, and

$$\sup_{x\in[a,b], r\geq 0} |v(x,r,S)| \leq \sup_{x\in[a,b]} |v(x,0,S)| \leq C. \tag{A.1}$$

For an arbitrary $r \geq 0$, by a standard probabilistic argument, one can derive that $v(x,r,S) \in C^2([a,b])$ satisfies the following elliptic equation:

$$\begin{cases} -ru(x) + \mu(x)u'(x) + \frac{1}{2}\sigma^2(x)u''(x) = 0, & x \in (a,b), \\ u(a) = v(a,r,S), \ u(b) = v(b,r,S). \end{cases} \tag{A.2}$$

Recall the strictly increasing function $y = \phi(x)$ defined in Equation (32) and denote by ϕ^{-1} the inverse function of ϕ. Define function $\tilde{u}(y) := u(\phi^{-1}(y))$ (i.e., $u(x) = \tilde{u}(\phi(x))$) on $[\phi(a), \phi(b)]$.

Then $\tilde{u} \in C^2([\phi(a), \phi(b)])$, and Equation (A.2) leads to

$$\begin{cases} -r\tilde{u}(y) + \frac{1}{2}\tilde{\sigma}^2(y)\tilde{u}''(y) = 0, & y \in (\phi(a), \phi(b)), \\ \tilde{u}(\phi(a)) = v(a, r, S), \ \tilde{u}(\phi(b)) = v(b, r, S), \end{cases} \tag{A.3}$$

with $\tilde{\sigma}(y) := \sigma(\phi^{-1}(y))\phi'(\phi^{-1}(y))$. Then Equation (A.1) together with the maximum principle implies that

$$\sup_{y \in (\phi(a), \phi(b))} |\tilde{u}(y)| \le v(a, r, S) \vee v(b, r, S) \le C \qquad \forall r \ge 0.$$

This together with the fact that $\tilde{u}'' = 2\tilde{u}r/(\tilde{\sigma}^2)$ and uniform ellipticity of $\tilde{\sigma}$ on $(\phi(a), \phi(b))$ leads to

$$\sup_{y \in (\phi(a), \phi(b))} |\tilde{u}''(y)| \le rC \qquad \forall r \ge 0. \tag{A.4}$$

By the mean value theorem, there exists $y_0 \in (\phi(a), \phi(b))$ such that

$$|\tilde{u}'(y_0)| = \left| \frac{v(b, r, S) - v(a, r, S)}{\phi(b) - \phi(a)} \right| \le \frac{2C}{\phi(b) - \phi(a)} \qquad \forall r \ge 0. \tag{A.5}$$

Then by Equations (A.4) and (A.5), for any $y \in (\phi(a), \phi(b))$ and $r \ge 0$, we have

$$|\tilde{u}'(y)| \le |\tilde{u}'(y_0)| + \int_{y_0}^{y} |\tilde{u}''(l)| dl \le \frac{2C}{\phi(b) - \phi(a)} + \int_{\phi(a)}^{\phi(b)} rC dl = \frac{2C}{\phi(b) - \phi(a)} + rC(\phi(b) - \phi(a)).$$

Therefore,

$$\sup_{y \in (\phi(a), \phi(b))} (|\tilde{u}'(y)| + |\tilde{u}''(y)|) \le C(1 + r) \qquad \forall r \ge 0.$$

This together with the fact that

$$\sup_{x \in (a,b)} (|u'(x)| + |u''(x)|) \le \sup_{y \in (\phi(a), \phi(b))} (|\tilde{u}'(y)| + |\tilde{u}''(y)|) \cdot \sup_{x \in (a,b)} (|\phi'(x)| + |\phi'(x)|^2 + |\phi''(x)|)$$

implies

$$\sup_{x \in (a,b)} (|v_x(x, r, S)| + |v_{xx}(x, r, S)|) = \sup_{x \in (a,b)} (|u'(x)| + |u''(x)|) \le C(1 + r) \qquad \forall r \ge 0. \tag{A.6}$$

Next, we verify that $V \in C^{1,2}([0, \infty) \times [a, b])$. For any $r \ge 0$, $v_x(a+, r, S)$, $v_x(b-, r, S)$ (resp. $v_{xx}(a+, r, S)$, $v_{xx}(b-, r, S)$) all exist and satisfy the same bound as the RHS of Equation (A.6). Hence, we conclude that $v(x, r, S) \in C^2([a, b])$. By Fubini theorem, Equation (20) leads to

$$V(t, x, S) = \int_0^\infty e^{-rt} v(x, r, S) dF(r) \quad \forall x \in \mathbb{X}. \tag{A.7}$$

This, together with Equation (A.6) and the assumption $\int_0^\infty r dF(r) < \infty$, implies that $V \in C^{1,2}([0, \infty) \times [a, b])$.

Finally, we prove Equation (18) for $V_x(t, x+, S)$ (the verification for $V_x(t, x-, S)$ is similar and thus omitted). For $(t, x) \in [0, \infty) \times [a, b)$, by Equations (A.6), (A.7) and the assumption $\int_0^\infty r dF(r) < \infty$,

$$V_x(t, x+, S) = \int_0^\infty e^{-rt} v_x(x+, r, S) dF(r). \tag{A.8}$$

Now take any $x \in \mathbb{X}$. If there exists some $h > 0$ such that $(x, x+h) \subset S^c$, then by Equations (21), (A.6), and (A.8), we have that

$$\begin{aligned}|V_x(t, x+, S) - V_x(0, x+, S)| &\le \int_0^\infty |v_x(x, r, S)||e^{-rt} - 1| dF(r) \\ &\le C \int_0^\infty (1+r)(1-e^{-rt}) dF(r) = o(\sqrt{t}), \text{ as } t \to 0.\end{aligned}$$

Otherwise, since S is admissible, there exists some $\bar{h} > 0$ such that $(x, x+\bar{h}) \subset S$. Then,

$$|V_x(t, x+, S) - V_x(0, x+, S)| = |\delta(t)f(x) - \delta(0)f(x)| \le |\delta'(0)| t f(x) = o(\sqrt{t}), \text{ as } t \to 0,$$

where the above inequality follows from Lemma 2.7. In sum, Equation (18) holds for $V_x(t, x+, S)$. ☐

Proof of Lemma 2.14. **Part (a)**: Assumption 2.10 guarantees that $V(t, x, S) \in C^{1,2}([0, \infty) \times \overline{S^c})$. Equation (12) implies that $\mathbb{P}^x(\rho_S = 0) = 1$ for any $x \in S$, so $V(t, x, S) = \delta(t)f(x)$ for any $(t, x) \in [0, \infty) \times S$. Now we prove that

$$\mathcal{L}V(t, x, S) \equiv 0 \quad \forall (t, x) \in [0, \infty) \times S^c. \tag{A.9}$$

Take $(t, x_0) \in [0, \infty) \times S^c$. Since S^c is open, we can take $h > 0$ such that $[x_0 - h, x_0 + h] \subset S^c$. By Assumption 2.1(i) and $V(t, x, S) \in C^{1,2}([0, \infty) \times \overline{S^c})$, $(s, x) \mapsto \mathcal{L}V(s, x, S)$ is continuous on the compact set $[t, t+1] \times \overline{B(x_0, h)}$. Then,

$$\sup_{(s,x) \in [t,t+1] \times \overline{B(x_0,h)}} |\mathcal{L}V(s, x, S)| < \infty. \tag{A.10}$$

Applying Ito's formula to $V(t+s, X_s, S)$ and taking expectation, the diffusion term vanishes due to the boundedness of $V_x \sigma$ on $[t, t+1] \times \overline{B(x_0, h)}$, we have that

$$\mathbb{E}^{x_0}[V(t + \varepsilon \wedge \tau_{B(x_0,h)}, X_{\varepsilon \wedge \tau_{B(x_0,h)}}, S)] - V(t, x_0, S) = \mathbb{E}^{x_0}\left[\int_0^{\varepsilon \wedge \tau_{B(x_0,h)}} \mathcal{L}V(t+s, X_s, S) ds\right]. \tag{A.11}$$

Meanwhile, by the continuity of $\mathcal{L}V(s, x, S)$ on $[t, t+1] \times \overline{B(x_0, h)}$,

$$\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} \int_0^{\varepsilon \wedge \tau_{B(x_0,h)}} \mathcal{L}V(t+s, X_s, S) ds = \mathcal{L}V(t, x_0, S), \quad \mathbb{P}^{x_0}\text{-a.s..}$$

Thanks to Equation (A.10), we can apply the dominated convergence theorem to above equality and get that

$$\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\mathbb{E}^{x_0}\left[\int_0^{\varepsilon\wedge\tau_{B(x_0,h)}}\mathcal{L}V(t+s,X_s,S)ds\right]=\mathcal{L}V(t,x_0,S). \tag{A.12}$$

On the other hand, for any $\varepsilon>0$, it is obvious that

$$\mathbb{E}^{x_0}[V(t+\varepsilon\wedge\tau_{B(x_0,h)},X_{\varepsilon\wedge\tau_{B(x_0,h)}},S)]=\mathbb{E}^{x_0}[\delta(t+\rho_S)f(X_{\rho_S})]=V(t,x_0,S). \tag{A.13}$$

Then by Equations (A.11)–(A.13), we have that

$$0=\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\Big(\mathbb{E}^{x_0}[V(t+\varepsilon\wedge\tau_{B(x_0,h)},X_{\varepsilon\wedge\tau_{B(x_0,h)}},S)]-V(t,x_0,S)\Big)=\mathcal{L}V(t,x_0,S),$$

and thus Equation (A.9) holds.

Part (b): The existence of $\mathcal{L}V(t,x\pm,S)$ on $[0,\infty)\times\mathbb{X}$ follows from part (a), the differentiability of δ and Assumption 2.8(ii). Take $x_0\in\mathbb{X}$ and $h>0$ such that $[x_0-h,x_0+h]\subset\mathbb{X}$. We show that

$$\sup_{(t,x)\in[0,\infty)\times\overline{B(x_0,h)}}|\mathcal{L}V(t,x-,S)|<\infty, \tag{A.14}$$

and the result for $\mathcal{L}V(t,x+,S)$ follows from a similar argument. Let $x\in B(x_0,h)$. If $(x-h',x)\in S^c$ for some constant $h'>0$, then by the left continuity of $y\mapsto\mathcal{L}V(t,y-,S)$ at $y=x$ and Equation (22) in part (a), we have $\mathcal{L}V(t,x-,S)=0$. Otherwise, since S is admissible, there exists $\bar{h}\in(0,h)$ such that $(x-\bar{h},x)\subset S$, then part (a) tells that $V(t,x,S)=\delta(t)f(x)$ on $[0,\infty)\times(x-\bar{h},x)$, and we have that

$$\begin{aligned}|\mathcal{L}V(t,x-,S)|=&\left|\delta'(t)f(x)+\delta(t)\Big(b(x)f'(x-)+\frac{1}{2}\sigma^2(x)f''(x-)\Big)\right|\\ \le&\sup_{y\in\overline{B(x_0,h)}}\Big(|\delta'(0)||f(y)|+|b(y)f'(y-)|+\frac{1}{2}\sigma^2(y)|f''(y-)|\Big)<\infty,\end{aligned}$$

where the inequality above follows from the first inequality in Lemma 2.7, Assumptions 2.1(i) and 2.8(ii). Hence, Equation (A.14) holds. □