Optimal Stopping under Adverse Nonlinear Expectation and Related Games

Marcel Nutz* Jianfeng Zhang†

First version: December 7, 2012. This version: July 21, 2014

Abstract

We study the existence of optimal actions in a zero-sum game $\inf_\tau \sup_P E^P[X_\tau]$ between a stopper and a controller choosing a probability measure. This includes the optimal stopping problem $\inf_\tau \mathcal{E}(X_\tau)$ for a class of sublinear expectations $\mathcal{E}(\cdot)$ such as the G-expectation. We show that the game has a value. Moreover, exploiting the theory of sublinear expectations, we define a nonlinear Snell envelope Y and prove that the first hitting time $\inf\{t : Y_t = X_t\}$ is an optimal stopping time. The existence of a saddle point is shown under a compactness condition. Finally, the results are applied to the subhedging of American options under volatility uncertainty.

Keywords Controller-and-stopper game; Optimal stopping; Saddle point; Nonlinear expectation; G-expectation

AMS 2000 Subject Classification 93E20; 49L20; 91A15; 60G44; 91B28

*Department of Mathematics, Columbia University, mnutz@math.columbia.edu. Research supported by NSF Grant DMS-1208985.

†Department of Mathematics, University of Southern California, jianfenz@usc.edu. Research supported by NSF Grant DMS-1008873.

The authors thank Ioannis Karatzas and the anonymous referees for their comments.

1 Introduction

On the space of continuous paths, we study a zero-sum stochastic game

$$\inf_{\tau\in\mathcal{T}} \sup_{P\in\mathcal{P}} E^P[X_\tau] \tag{1.1}$$

between a stopper and a controller; here $X = (X_t)$ is the process to be stopped, $\mathcal{T}$ is the set of stopping times with values in a given interval $[0, T]$, and $\mathcal{P}$ is a given set of probability measures. Specifically, we are interested in the situation where $\mathcal{P}$ may be nondominated; i.e., there is no reference measure with respect to which all $P \in \mathcal{P}$ are absolutely continuous. This is the case, for instance, when $\mathcal{P}$ is the set of laws resulting from a controlled stochastic differential equation whose diffusion coefficient is affected by the control (cf. Example 3.9), whereas the dominated case would correspond to the case where only the drift is controlled. Or, in the language of partial differential equations, we are interested in the fully nonlinear case rather than the semilinear case. Technically, the nondominated situation entails that general minimax results cannot be applied to (1.1), that the cost functional $\sup_{P\in\mathcal{P}} E^P[\,\cdot\,]$ does not satisfy the dominated convergence theorem, and of course the absence of various tools from stochastic analysis. Our main results for the controller-and-stopper game include the existence of an optimal action τ^* for the stopper under general conditions and the existence of a saddle point (τ^*, P^*) under a compactness condition. Both of these results were previously known only in the case of drift control; cf. the review of literature at the end of this section.

If we introduce the sublinear expectation $\mathcal{E}(\cdot) = \sup_{P\in\mathcal{P}} E^P[\,\cdot\,]$, the stopper's part of the game can also be interpreted as the nonlinear optimal stopping problem

$$\inf_{\tau\in\mathcal{T}} \mathcal{E}(X_\tau). \tag{1.2}$$

This alternate point of view is of independent interest, but it will also prove to be useful in establishing the existence of optimal actions for the game. Indeed, we shall start our analysis with (1.2) and exploit the theory of nonlinear expectations, which suggests to mimic the classical theory of optimal stopping under linear expectation (e.g., [10]). Namely, we define the Snell envelope

$$Y_t = \inf_{\tau\in\mathcal{T}_t} \mathcal{E}_t(X_\tau),$$

where $\mathcal{T}_t$ is the set of stopping times with values in $[t, T]$ and $\mathcal{E}_t(\cdot)$ is the conditional sublinear expectation as obtained by following the construction of [26]. Under suitable assumptions, we show that the first hitting time

$$\tau^* = \inf\{t : Y_t = X_t\}$$

is a stopping time which is optimal; i.e., $\mathcal{E}(X_{\tau^*}) = \inf_{\tau\in\mathcal{T}} \mathcal{E}(X_\tau)$. Armed with this result, we return to the game-theoretic point of view and prove the existence of the value,

$$\inf_{\tau\in\mathcal{T}} \sup_{P\in\mathcal{P}} E^P[X_\tau] = \sup_{P\in\mathcal{P}} \inf_{\tau\in\mathcal{T}} E^P[X_\tau].$$

Moreover, under a weak compactness assumption on $\mathcal{P}$, we construct $P^* \in \mathcal{P}$ such that (τ^*, P^*) is a saddle point for (1.1). These three main results are summarized in Theorem 3.4. Finally, we give an application to the financial problem of pricing a path-dependent American option under volatility uncertainty and show in Theorem 5.1 that (1.2) yields the buyer's price (or subhedging price) in an appropriate financial market model.

It is worth remarking that our results are obtained by working "globally" and not, as is often the case in the study of continuous-time games, by a local-to-global passage based on a Bellman–Isaacs operator; in fact, all ingredients of our setup can be non-Markovian (i.e., path-dependent). The "weak" formulation of the game, where the canonical process plays the role of the state process, is important in this respect.

Like in the classical stopping theory, a dynamic programming principle plays a key role in our analysis. We first prove this principle for the upper value function Y rather than the lower one (sup-inf), which would be the standard choice in the literature on robust optimal stopping (but of course the result for the lower value follows once the existence of the value is established). The reason is that, due to the absence of a reference measure, the structure of the set $\mathcal{T}$ of stopping times is inconvenient for measurable selections, whereas on the set $\mathcal{P}$ we can exploit the natural Polish structure. Once again due to the absence of a reference measure, we are unable to infer the optimality of τ^* directly from the dynamic programming principle. However, we observe that in the discrete-time case, the classical recursive analysis can be carried over rather easily by exploiting the tower property of the nonlinear expectation. This recursive structure extends to the case where the processes are running in continuous time but the stopping times are restricted to take values in a given discrete set, like for a Bermudan option. To obtain the optimality of τ^*, we then approximate the continuous-time problem with such discrete ones; the key idea is to compare the first hitting times for the discrete problems with the times $\tau^\varepsilon = \inf\{t : X_t - Y_t \leq \varepsilon\}$

and exploit the $\mathcal{E}$-martingale property of the discrete-time Snell envelope. A similar approximation is also used to prove the existence of the value, as it allows to circumvent the mentioned difficulty in working with the lower value function: we first identify the upper and lower value functions in the discrete problems and then pass to the limit. Finally, for the existence of P^*, an important difficulty is that we have little information about the regularity of X_{τ^*}. Our construction uses a compactness argument and a result of [9] on the approximation of hitting times by random times that are continuous in ω to find a measure $P^\sharp$ which is optimal up to the time τ^*. In a second step, we manipulate $P^\sharp$ in such a way that for the stopper, immediate stopping after τ^* is optimal, which yields the optimal measure P^* for the full time interval. All this is quite different from the existing arguments for the dominated case.

While the remainder of this introduction concerns the related literature, the rest of the paper is organized as follows: Section 2 details the setup and the construction of the sublinear expectation. In Section 3, we state our main result, discuss its assumptions, and give a concrete example related to controlled stochastic functional/differential equations. Section 4 contains the proof of the main result, while the application to option pricing is studied in Section 5.

Literature. In terms of the mathematics involved, the study of the problem $\sup_{\tau\in\mathcal{T}}\sup_{P\in\mathcal{P}} E^P[X_\tau]$ in [9] is the closest to the present one. Although this is a control problem with discretionary stopping rather than a game, their regularity results are similar to ours. On the other hand, the proofs of the optimality of τ^* are completely different: in [9], it was relatively simple to obtain the martingale property up to τ^ε, directly in continuous time, and the main difficulty was the passage from τ^ε to τ^*, which is trivial in our case. (The existence of an optimal $P^*\in\mathcal{P}$ was not studied in [9].) Somewhat surprisingly, the conditions obtained in the present paper are weaker than the ones in [9]; in particular, for the optimality of τ^*, we do not assume that $\mathcal{P}$ is compact.

After the publication of the preprint of the present work, [6] showed the existence of the optimal stopping time and value in a case where X is not bounded (under various other assumptions). The authors go through the dynamic programming for the lower value rather than the upper one, by using approximations based on the regularity of X. The existence of a saddle point is not addressed directly and we mention that our result does not apply, because the technical assumptions of [6] preclude closedness of $\mathcal{P}$

in most cases of interest.

For the case where $\mathcal{P}$ is dominated, the problem of optimal stopping under nonlinear expectation (and related risk measures) is fairly well studied; see, in particular, [3, 4, 5, 11, 13, 17, 18, 30]. The mathematical analysis for that case is quite different. On the other hand, there is a literature on controller-and-stopper games. In the discrete-time case, [20] obtained a general result on the existence of the value. For the continuous-time problem, the literature is divided into two parts: In the non-Markovian case, only the pure drift control has been studied, cf. [18] and the references therein; this again corresponds to the dominated situation. For the nondominated situation, results exist only in the Markovian case, where the presence of singular measures plays a lesser role; cf. [2, 15, 16, 18]. In particular, [2] obtained the existence of the value for a diffusion setting via the comparison principle for the associated partial differential equation. On the other hand, [15] studied a linear diffusion valued in the unit interval with absorbing boundaries and found, based on scale-function arguments, rather explicit formulas for the value and a saddle point. Apart from such rather specific models, our results on the existence of optimal actions are new even in the Markovian case.

Regarding the literature on nonlinear expectations, we refer to [27, 29] and the references therein; for the related second order backward stochastic differential equations (2BSDE) to [8, 33, 32], and in particular to [21, 22] for the reflected 2BSDE related to our problem; whereas for the uncertain volatility model in finance, we refer to [1, 19, 31].

2 Preliminaries

In this section, we introduce the setup and in particular the sublinear expectation. We follow [26] as we need the conditional expectation to be defined at every path and for all Borel functions.

2.1 Notation

We fix $d \in \mathbb{N}$ and let $\Omega = \{\omega \in C(\mathbb{R}_+;\mathbb{R}^d) : \omega_0 = 0\}$ be the space of continuous paths equipped with the topology of locally uniform convergence and the Borel σ-field $\mathcal{F} = \mathcal{B}(\Omega)$. We denote by $B = (B_t)_{t\geq 0}$ the canonical process $B_t(\omega) = \omega_t$ and by $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ the (raw) filtration generated by B. Finally, $\mathfrak{P}(\Omega)$ denotes the space of probability measures on Ω with the topology of weak convergence. Throughout this paper, "stopping time" will refer to a finite $\mathbb{F}$-stopping time. Given a stopping time τ and $\omega, \omega' \in \Omega$, we set

$$(\omega \otimes_\tau \omega')_u := \omega_u \mathbf{1}_{[0,\tau(\omega))}(u) + \big(\omega_{\tau(\omega)} + \omega'_{u-\tau(\omega)}\big)\mathbf{1}_{[\tau(\omega),\infty)}(u), \quad u \geq 0.$$

For any probability measure $P \in \mathfrak{P}(\Omega)$, there is a regular conditional probability distribution $\{P^\omega_\tau\}_{\omega\in\Omega}$ given $\mathcal{F}_\tau$ satisfying

$$P^\omega_\tau\big\{\omega' \in \Omega : \omega' = \omega \text{ on } [0, \tau(\omega)]\big\} = 1 \quad \text{for all} \quad \omega \in \Omega;$$

cf. [34, p. 34]. We then define $P^{\tau,\omega} \in \mathfrak{P}(\Omega)$ by

$$P^{\tau,\omega}(A) := P^\omega_\tau(\omega \otimes_\tau A), \quad A \in \mathcal{F}, \quad \text{where } \omega \otimes_\tau A := \{\omega \otimes_\tau \omega' : \omega' \in A\}.$$

Given a function f on Ω and $\omega \in \Omega$, we also define the function $f^{\tau,\omega}$ by

$$f^{\tau,\omega}(\omega') := f(\omega \otimes_\tau \omega'), \quad \omega' \in \Omega.$$

We then have $E^{P^{\tau,\omega}}[f^{\tau,\omega}] = E^P[f|\mathcal{F}_\tau](\omega)$ for P-a.e. $\omega \in \Omega$.

2.2 Sublinear Expectation

Let $\{\mathcal{P}(s, \omega)\}_{(s,\omega)\in\mathbb{R}_+\times\Omega}$ be a family of subsets of $\mathfrak{P}(\Omega)$, adapted in the sense that

$$\mathcal{P}(s, \omega) = \mathcal{P}(s, \omega') \quad \text{if} \quad \omega|_{[0,s]} = \omega'|_{[0,s]},$$

and define $\mathcal{P}(\tau, \omega) := \mathcal{P}(\tau(\omega), \omega)$ for any stopping time τ. Note that the set $\mathcal{P}(0, \omega)$ is independent of ω as all paths start at the origin. Thus, we can define $\mathcal{P} := \mathcal{P}(0, \omega)$. We assume throughout that $\mathcal{P}(s, \omega) \neq \emptyset$ for all $(s, \omega) \in \mathbb{R}_+ \times \Omega$.

The following assumption, which is in force throughout the paper, will enable us to construct the conditional sublinear expectation related to $\{\mathcal{P}(s, \omega)\}$; it essentially states that our problem admits dynamic programming. We recall that a subset of a Polish space is called analytic if it is the image of a Borel subset of another Polish space under a Borel mapping (we refer to [7, Chapter 7] for background).

Assumption 2.1. Let $s \in \mathbb{R}_+$, let τ be a stopping time such that $\tau \geq s$, let $\bar\omega \in \Omega$ and $P \in \mathcal{P}(s, \bar\omega)$. Set $\theta := \tau^{s,\bar\omega} - s$.

(i) The graph $\{(P', \omega) : \omega \in \Omega,\ P' \in \mathcal{P}(\tau, \omega)\} \subseteq \mathfrak{P}(\Omega) \times \Omega$ is analytic.

(ii) We have $P^{\theta,\omega} \in \mathcal{P}(\tau, \bar\omega \otimes_s \omega)$ for P-a.e. $\omega \in \Omega$.

(iii) If $\nu : \Omega \to \mathfrak{P}(\Omega)$ is an $\mathcal{F}_\theta$-measurable kernel and $\nu(\omega) \in \mathcal{P}(\tau, \bar{\omega} \otimes_s \omega)$ for P-a.e. $\omega \in \Omega$, then the measure defined by

$$\bar{P}(A) = \iint (\mathbf{1}_A)^{\theta,\omega}(\omega')\, \nu(d\omega'; \omega)\, P(d\omega), \quad A \in \mathcal{F}$$

is an element of $\mathcal{P}(s, \bar{\omega})$.

Let us recall that a function $f : \Omega \to \overline{\mathbb{R}}$ is called upper semianalytic if $\{f > c\}$ is analytic for each $c \in \mathbb{R}$; in particular, every Borel function is upper semianalytic (cf. [7, Chapter 7]). Moreover, we recall that the universal completion of a σ-field $\mathcal{A}$ is given by $\mathcal{A}^* := \cap_P \mathcal{A}^P$, where $\mathcal{A}^P$ denotes the completion with respect to P and the intersection is taken over all probability measures on $\mathcal{A}$. Let us agree that $E^P[f] := -\infty$ if $E^P[f^+] = E^P[f^-] = +\infty$, then we can introduce the sublinear expectation corresponding to $\{\mathcal{P}(s, \omega)\}$ as follows (cf. [26, Theorem 2.3]).

Proposition 2.2. *Let $\sigma \leq \tau$ be stopping times and let $f : \Omega \to \overline{\mathbb{R}}$ be an upper semianalytic function. Then the function*

$$\mathcal{E}_\tau(f)(\omega) := \sup_{P \in \mathcal{P}(\tau,\omega)} E^P[f^{\tau,\omega}], \quad \omega \in \Omega \tag{2.1}$$

is $\mathcal{F}_\tau^*$-measurable and upper semianalytic. Moreover,

$$\mathcal{E}_\sigma(f)(\omega) = \mathcal{E}_\sigma(\mathcal{E}_\tau(f))(\omega) \quad \textit{for all} \quad \omega \in \Omega. \tag{2.2}$$

We write $\mathcal{E}(\cdot)$ for $\mathcal{E}_0(\cdot)$. We shall use very frequently (and often implicitly) the following extension of Galmarino's test (cf. [26, Lemma 2.5]).

Lemma 2.3. *Let $f : \Omega \to \overline{\mathbb{R}}$ be $\mathcal{F}^*$-measurable and let τ be a stopping time. Then f is $\mathcal{F}_\tau^*$-measurable if and only if $f(\omega) = f(\omega_{\cdot \wedge \tau(\omega)})$ for all $\omega \in \Omega$.*

The following is an example for the use of Lemma 2.3: If f and g are bounded and upper semianalytic, and g is $\mathcal{F}_t^*$ measurable, then

$$\mathcal{E}_t(f + g) = \mathcal{E}_t(f) + g$$

by (2.1), since the test shows that $g^{t,\omega} = g$. Similarly, if we also have that $g \geq 0$, then $\mathcal{E}_t(fg) = \mathcal{E}_t(f)g$. We emphasize that all these identities hold at every single $\omega \in \Omega$, without an exceptional set.

The most basic example we have in mind for $\mathcal{E}(\cdot)$ is the G-expectation of [27, 28]; or more precisely, its extension to the upper semianalytic functions. In this case, $\mathcal{P}(s, \omega)$ is actually independent of (s, ω); more general cases are discussed in Section 3.1.

Example 2.4 (G-Expectation)**.** Let $U \neq \emptyset$ be a convex, compact set of nonnegative definite symmetric $d \times d$ matrices and define $\mathcal{P}_G$ to be the set of all probabilities on Ω under which the canonical process B is a martingale whose quadratic variation $\langle B \rangle$ is absolutely continuous $dt \times P$-a.e. and

$$\frac{d\langle B \rangle_t}{dt} \in U \quad dt \times P\text{-a.e.}$$

Moreover, set $\mathcal{P}(s,\omega) := \mathcal{P}_G$ for all $(s,\omega) \in \mathbb{R}_+ \times \Omega$. Then Assumption 2.1 is satisfied (cf. [26, Theorem 4.3]) and $\mathcal{E}(\cdot)$ is called the G-expectation associated with U (where $2G$ is the support function of U). We remark that $\mathcal{P}(s,\omega)$ is weakly compact in this setup.

More generally, Assumption 2.1 is established in [26] when U is a set-valued process (i.e., a Borel set depending on t and ω). In this case, $\mathcal{P}(s,\omega)$ need not be compact and depends on (s,ω).

3 Main Results

Let $T \in (0,\infty)$ be the time horizon. For any $t \in [0,T]$, we denote by $\mathcal{T}_t$ the set of all $[t,T]$-valued stopping times. For technical reasons, we shall also consider the smaller set $\mathcal{T}^t \subseteq \mathcal{T}_t$ of stopping times which do not depend on the path up to time t; i.e.,

$$\mathcal{T}^t = \{\tau \in \mathcal{T}_t : \tau^{t,\omega} = \tau^{t,\omega'} \text{ for all } \omega,\omega' \in \Omega\}. \tag{3.1}$$

In particular, $\mathcal{T} := \mathcal{T}_0 = \mathcal{T}^0$ is the set of all $[0,T]$-valued stopping times. We define the pseudometric $\mathbf{d}$ on $[0,T] \times \Omega$ by

$$\mathbf{d}[(t,\omega),(s,\omega')] := |t-s| + \|\omega_{\cdot\wedge t} - \omega'_{\cdot\wedge s}\|_T$$

for $(t,\omega),(s,\omega') \in [0,T] \times \Omega$, where $\|\omega\|_u := \sup_{r \le u} |\omega_r|$ for $u \geq 0$ and $|\cdot|$ is the Euclidean norm.

Let us now introduce the process X to be stopped. Of course, the most classical example is $X_t = f(B_t)$ for some function $f : \mathbb{R}^d \to \mathbb{R}$. We consider a fairly general, possibly path-dependent functional $X = X(B)$; note that the canonical process B plays the role of the state process. We shall work under the following regularity assumption.

Assumption 3.1. The process $X = (X_t)_{0 \le t \le T}$ is progressively measurable, uniformly bounded, has càdlàg trajectories with nonpositive jumps, and there exists a modulus of continuity ρ_X such that

$$X_t(\omega) - X_s(\omega') \le \rho_X\big(\mathbf{d}[(t,\omega),(s,\omega')]\big) \quad \text{for all} \quad s \le t \tag{3.2}$$

and $\omega,\omega' \in \Omega$.

The subsequent assumptions are stated in a form that is convenient for the proofs; in that sense, they are in the most general form. Sufficient conditions and examples will be discussed in Section 3.1.

Assumption 3.2. There is a modulus of continuity $\rho_{\mathcal{E}}$ with the following property. Let $t \in [0,T]$, $\tau \in \mathcal{T}^t$ and $\bar\omega \in \Omega$, then for all $\omega \in \Omega$ there exists $\tau_\omega \in \mathcal{T}^t$ such that

$$|\mathcal{E}_t(X_\tau)(\bar\omega) - \mathcal{E}_t(X_{\tau_\omega})(\omega)| \leq \rho_{\mathcal{E}}(\|\bar\omega - \omega\|_t)$$

and such that $(\omega, \omega') \mapsto \tau_\omega(\omega')$ is $\mathcal{F}_t \otimes \mathcal{F}$-measurable.

We note that as X is bounded, the moduli ρ_X and $\rho_{\mathcal{E}}$ can also be taken to be bounded. In the subsequent assumption, we use the notation B^θ for the process $B_{\cdot+\theta} - B_\theta$, where θ is a stopping time.

Assumption 3.3. Let ρ' be a bounded modulus of continuity. Then there exists a modulus of continuity ρ such that

$$E^P\left[\rho'(\delta + \|B^\theta\|_\delta)\right] \leq \rho(\delta), \quad \delta \in [0,T]$$

for all $\theta \in \mathcal{T}$, $P \in \mathcal{P}(t,\omega)$ and $(t,\omega) \in [0,T] \times \Omega$.

Let us now introduce the main object under consideration, the value function ("nonlinear Snell envelope") given by

$$Y_t(\omega) := \inf_{\tau \in \mathcal{T}_t} \mathcal{E}_t(X_\tau)(\omega), \quad (t,\omega) \in [0,T] \times \Omega. \tag{3.3}$$

We shall see that Y is Borel-measurable under the above assumptions, and that $\mathcal{T}_t$ can be replaced by $\mathcal{T}^t$ without changing the value of Y_t. The following is our main result.

Theorem 3.4. *Let Assumptions 3.1, 3.2 and 3.3 hold.*

(i) There exists an optimal stopping time; namely,

$$\tau^* := \inf\{t \in [0,T] : Y_t = X_t\}$$

satisfies $\tau^* \in \mathcal{T}$ *and* $\mathcal{E}(X_{\tau^*}) = \inf_{\tau \in \mathcal{T}} \mathcal{E}(X_\tau)$.

(ii) The game has a value; that is,

$$\inf_{\tau \in \mathcal{T}} \sup_{P \in \mathcal{P}} E^P[X_\tau] = \sup_{P \in \mathcal{P}} \inf_{\tau \in \mathcal{T}} E^P[X_\tau].$$

(iii) *Suppose that $\mathcal{P}(t,\omega)$ is weakly compact for all $(t,\omega)\in[0,T]\times\Omega$. Then the game has a saddle point: there exists $P^*\in\mathcal{P}$ such that*

$$\inf_{\tau\in\mathcal{T}} E^{P_*}[X_\tau] = E^{P^*}[X_{\tau^*}] = \sup_{P\in\mathcal{P}} E^P[X_{\tau^*}].$$

Of course, weak compactness in (iii) refers to the topology induced by the continuous bounded functions, and a similar identity as in (ii) holds for the value functions at positive times (cf. Lemma 4.12). The proof of the theorem is stated in Section 4. We mention the following variant of Theorem 3.4(i) where the stopping times are restricted to take values in a discrete set $\mathbb{T}$; in this case, Assumption 3.3 is unnecessary. The proof is again deferred to the subsequent section.

Remark 3.5. Let Assumptions 3.1 and 3.2 hold. Let $\mathbb{T}=\{t_0,t_1,\dots,t_n\}$, where $n\in\mathbb{N}$ and $t_0<t_1<\cdots<t_n=T$, and consider the obvious corresponding notions like $\mathcal{T}_t(\mathbb{T})=\{\tau\in\mathcal{T}_t:\,\tau(\cdot)\in\mathbb{T}\}$ and $Y_t=\inf_{\tau\in\mathcal{T}_t(\mathbb{T})}\mathcal{E}_t(X_\tau)$. Then Y satisfies the backward recursion

$$Y_{t_n}=X_{t_n}\quad\text{and}\quad Y_{t_i}=X_{t_i}\wedge\mathcal{E}_{t_i}(Y_{t_{i+1}}),\quad i=0,\dots,n-1$$

and $\tau^*:=\inf\{t\in\mathbb{T}:\,Y_t=X_t\}$ satisfies $\mathcal{E}(X_{\tau^*})=\inf_{\tau\in\mathcal{T}(\mathbb{T})}\mathcal{E}(X_\tau)$.

3.1 Sufficient Conditions for the Main Assumptions

In the remainder of this section, we discuss the conditions of the theorem. Assumption 3.1 is clearly satisfied when $X_t=f(B_t)$ for a bounded, uniformly continuous function f. The following shows that Assumption 3.2 is trivially satisfied, e.g., for the G-expectation of Example 2.4 (a nontrivial situation is discussed in Example 3.9).

Remark 3.6. Assume that $\mathcal{P}(t,\omega)$ does not depend on ω, for all $t\in[0,T]$. Then Assumption 3.1 implies Assumption 3.2.

Proof. Let $\tau\in\mathcal{T}^t$, then $\tau^{t,\omega}=\tau^{t,\omega'}=:\theta$ for all $\omega,\omega'\in\Omega$. Moreover, taking $s=t$ in (3.2) shows that

$$|X_s(\omega)-X_s(\omega')|\le\rho_X(\|\omega-\omega'\|_s).$$

We deduce that for all $\tilde\omega\in\Omega$,

$$\begin{aligned}|(X_\tau)^{t,\omega}(\tilde\omega)-(X_\tau)^{t,\omega'}(\tilde\omega)| &= |X_{\theta(\tilde\omega)}(\omega\otimes_t\tilde\omega)-X_{\theta(\tilde\omega)}(\omega'\otimes_t\tilde\omega)|\\ &\le\rho_X(\|\omega\otimes_t\tilde\omega-\omega'\otimes_t\tilde\omega\|_{\theta(\tilde\omega)})\\ &=\rho_X(\|\omega-\omega'\|_t).\end{aligned}\tag{3.4}$$

If $\mathcal{P}(t,\cdot)=\mathcal{P}(t)$, it follows that

$$
\begin{aligned}
|\mathcal{E}_t(X_\tau)(\omega)-\mathcal{E}_t(X_\tau)(\omega')| &\leq \sup_{P\in\mathcal{P}(t)} E^P[|(X_\tau)^{t,\omega}-(X_\tau)^{t,\omega'}|] \\
&\leq \rho_X(\|\omega-\omega'\|_t);
\end{aligned}
$$

that is, Assumption 3.2 holds with $\rho_{\mathcal{E}}=\rho_X$ and $\tau_\omega=\tau$. □

The following is a fairly general sufficient condition for Assumption 3.3; it covers most cases of interest.

Remark 3.7. Suppose that for some $\alpha, c>0$, the moment condition

$$
E^P[\|B^\theta\|_\delta]\leq c\delta^\alpha, \quad \delta\in[0,T]
$$

is satisfied for all $\theta\in\mathcal{T}$, $P\in\mathcal{P}(t,\omega)$ and $(t,\omega)\in[0,T]\times\Omega$. Then Assumption 3.3 holds. In particular, this is the case if every $P\in\mathcal{P}(t,\omega)$ is the law of an Itô process

$$
\Gamma_s=\int_0^s \mu_r\,dr+\int_0^s\sigma_r\,dW_r
$$

(where W is a Brownian motion) and $|\mu|+|\sigma|\leq C$ for a universal constant C.

Proof. Let $r\in\mathbb{R}$ be such that $\rho'\leq r$. For any $a>0$, we have

$$
E^P\big[\rho'(\delta+\|B^\theta\|_\delta)\big]\leq\rho'(\delta+a)+rP\{\|B^\theta\|_\delta\geq a\}.
$$

Using that $P\{\|B^\theta\|_\delta\geq a\}\leq a^{-1}E^P[\|B^\theta\|_\delta]\leq a^{-1}c\delta^\alpha$ and choosing $a=\delta^{\alpha/2}$, we obtain that

$$
E^P\big[\rho'(\delta+\|B^\theta\|_\delta)\big]\leq\rho'(\delta+\delta^{\alpha/2})+cr\delta^{\alpha/2}=:\rho(\delta),
$$

which was the first claim.

Suppose that $B=A+M$ under P, where $|dA|+d|\langle M\rangle|\leq C\,dt$; we focus on the scalar case for simplicity. Using the Burkholder–Davis–Gundy inequalities for $M^\theta=M_{\cdot+\theta}-M_\theta$, we have

$$
\begin{aligned}
E^P[\|B^\theta\|_\delta] &\leq E^P[\|A^\theta\|_\delta]+E^P[\|M^\theta\|_\delta] \\
&\leq C\delta+c_1E^P[\|\langle M^\theta\rangle^{1/2}\|_\delta] \\
&\leq (CT^{1/2}+c_1C^{1/2})\delta^{1/2}, \quad \delta\in[0,T],
\end{aligned}
$$

where $c_1>0$ is a universal constant. □

Let us now discuss two classes of models. The first one is the main example for control in the "weak formulation;" that is, the set of controls is stated directly in terms of laws.

Example 3.8. Let U be a nonempty, bounded Borel set of $\mathbb{R}^d \times \mathbb{S}_+$, where $\mathbb{S}_+$ is the set of $d \times d$ nonnegative definite symmetric matrices. Moreover, let $\mathcal{P}$ be the set of all laws of continuous semimartingales whose characteristics are absolutely continuous (with respect to the Lebesgue measure) and whose differential characteristics take values in U. That is, $\mathcal{P}$ consists of laws of Itô processes $\int b_t \, dt + \int \sigma_t \, dW_t$, each situated on its own probability space with a Brownian motion W, where the pair $(b, \sigma\sigma^\top)$ almost surely takes values in the set U. For instance, if $d = 1$ and $U = I_1 \times I_2$ is a product of intervals, this models the case where the controller can choose the drift from I_1 and the (squared) diffusion from I_2.

The above setup (and its extension to jump processes) is studied in [24], where it is shown in particular that Assumption 2.1 holds. Moreover, we see from Remark 3.6 that Assumption 3.2 is satisfied, while Remark 3.7 shows that Assumption 3.3 holds as well. Finally, if U is compact and convex, standard results (see [35]) imply that the set $\mathcal{P}$ is weakly compact.

In the second class of models, whose formulation is borrowed from [25], the elements of $\mathcal{P}$ correspond to the possible laws of the solution to a controlled stochastic functional/differential equation (SDE). This is the main case of interest for the controller-and-stopper games in the "strong formulation" of control and as we shall see, the sets $\mathcal{P}(t, \omega)$ indeed depend on (t, ω). Note that in the setting of the strong formulation the set $\mathcal{P}$ is typically not closed and in particular not compact, so that we cannot expect the existence of a saddle point in general. For simplicity, we only discuss the case of a driftless SDE.

Example 3.9. Let U be a nonempty Borel set of $\mathbb{R}^d$ and let $\mathcal{U}$ be the set of all U-valued, progressively measurable, càdlàg processes ν. We denote by $\mathbb{S}_{++}$ the set of positive definite symmetric matrices and by $\mathbb{D}$ the Skorohod space of càdlàg paths in $\mathbb{R}^d$ starting at the origin, and consider a function

$$\sigma : \mathbb{R}_+ \times \mathbb{D} \times U \to \mathbb{S}_{++}$$

such that $(t, \omega) \mapsto \sigma(t, \Gamma(\omega), \nu_t(\omega))$ is progressively measurable (càdlàg) whenever Γ and ν are progressively measurable (càdlàg). We assume that σ is uniformly Lipschitz in its second variable with respect to the supremum norm, and (for simplicity) uniformly bounded. Moreover, we assume that σ

is a one-to-one function in its third variable, admitting a measurable inverse on its range. More precisely, there exists a function $\sigma^{inv} : \mathbb{R}_+ \times \mathbb{D} \times \mathbb{S}_{++} \to U$ such that

$$\sigma^{inv}(t, \omega, \sigma(t, \omega, u)) = u$$

for all $(t, \omega, u) \in \mathbb{R}_+ \times \mathbb{D} \times U$, and σ^{inv} satisfies the same measurability and càdlàg properties as σ. Given $\nu \in \mathcal{U}$, we consider the stochastic functional/differential equation

$$\Gamma_t = \int_0^t \sigma(r, \Gamma, \nu_r)\, dB_r, \quad t \geq 0$$

under the Wiener measure P_0 (i.e., B is a d-dimensional Brownian motion). This equation has a P_0-a.s. unique strong solution whose law is denoted by $P(\nu)$. We then define $\mathcal{P} = \{P(\nu) :\ \nu \in \mathcal{U}\}$. More generally, $\mathcal{P}(s, \omega)$ is defined as the set of laws $P(s, \omega, \nu)$ corresponding to the SDE with conditioned coefficient $(r, \omega', u) \mapsto \sigma(r + s, \omega \otimes_s \omega', u)$ and initial condition $\Gamma_0 = \omega_s$; more precisely, $P(s, \omega, \nu)$ is the law on Ω of the solution translated to start at the origin (see also [25]).

In this model, Assumption 2.1 can be verified by the arguments used in [25] and [23]; the details are lengthy but routine. Assumption 3.3 is satisfied by Remark 3.7 since σ is bounded (note that this condition can be improved by using SDE estimates). We impose Assumption 3.1 on X and turn to Assumption 3.2, which is the main interest in this example. The main problem in this regard is that we cannot impose continuity conditions on the stopping times.

Lemma 3.10. *Assumption 3.2 is satisfied in the present setting.*

While we defer the actual proof to the Appendix, we sketch here the rough idea for the case where there is no control in the SDE (i.e., U is a singleton) and thus each set $\mathcal{P}(s, \omega)$ consists of a single measure $P(s, \omega)$. Let $t \in [0, T]$, $\tau \in \mathcal{T}^t$ and $\bar{\omega} \in \Omega$; we shall construct $\tau_\omega \in \mathcal{T}^t$ such that

$$|\mathcal{E}_t(X_\tau)(\bar{\omega}) - \mathcal{E}_t(X_{\tau_\omega})(\omega)| \leq \rho_{\mathcal{E}}(\|\bar{\omega} - \omega\|_t)$$

for all $\omega \in \Omega$ (and such that $(\omega, \omega') \mapsto \tau_\omega(\omega')$ is $\mathcal{F}_t \otimes \mathcal{F}$-measurable).

Fix $\omega, \bar{\omega} \in \Omega$ and denote by $\Gamma^{t,\omega}$ and $\Gamma^{t,\bar{\omega}}$ the corresponding solutions of the SDE (translated to start at the origin); that is, we have

$$d\Gamma^{t,\omega}_u = \sigma(u+t, \omega \otimes_t \Gamma^{t,\omega})\, dB_u \quad \text{and} \quad d\Gamma^{t,\bar{\omega}}_u = \sigma(u+t, \bar{\omega} \otimes_t \Gamma^{t,\bar{\omega}})\, dB_u \quad P_0\text{-a.s.}$$

Our aim is to construct $\tau_\omega \in \mathcal{T}^t$ with the property that

$$\tau_\omega(0 \otimes_t \Gamma^{t,\omega}) = \tau(0 \otimes_t \Gamma^{t,\bar\omega}) \quad P_0\text{-a.s.}, \tag{3.5}$$

where 0 is the constant path (or any other path, for that matter). Indeed, if this identity holds, then

$$\begin{aligned}\mathcal{E}_t(X_{\tau_\omega})(\omega) &= E^{P(t,\omega)}\big[(X_{\tau_\omega})^{t,\omega}\big] \\ &= E^{P(t,\omega)}\big[X_{\tau_\omega(0\otimes_t \cdot)}(\omega \otimes_t \cdot)\big] \\ &= E^{P_0}\big[X_{\tau_\omega(0\otimes_t \Gamma^{t,\omega})}(\omega \otimes_t \Gamma^{t,\omega})\big] \\ &= E^{P_0}\big[X_{\tau(0\otimes_t \Gamma^{t,\bar\omega})}(\omega \otimes_t \Gamma^{t,\omega})\big]\end{aligned}$$

and thus

$$\begin{aligned}\big|\mathcal{E}_t(X_\tau)(\bar\omega) - &\mathcal{E}_t(X_{\tau_\omega})(\omega)\big| \\ &= \big|E^{P_0}\big[X_{\tau(0\otimes_t\Gamma^{t,\bar\omega})}(\bar\omega \otimes_t \Gamma^{t,\bar\omega})\big] - E^{P_0}\big[X_{\tau(0\otimes_t\Gamma^{t,\bar\omega})}(\omega \otimes_t \Gamma^{t,\omega})\big]\big| \\ &\leq E^{P_0}\big[\big|X_{\tau(0\otimes_t\Gamma^{t,\bar\omega})}(\bar\omega \otimes_t \Gamma^{t,\bar\omega}) - X_{\tau(0\otimes_t\Gamma^{t,\bar\omega})}(\omega \otimes_t \Gamma^{t,\omega})\big|\big] \\ &\leq E^{P_0}\big[\rho_X\big(\big\|\bar\omega \otimes_t \Gamma^{t,\bar\omega} - \omega \otimes_t \Gamma^{t,\omega}\big\|_T\big)\big] \\ &\leq \rho_X(C\|\bar\omega - \omega\|_t),\end{aligned} \tag{3.6}$$

where the last inequality follows by a standard SDE estimate as in [25, Lemma 2.6], with $C > 0$ depending only on the Lipschitz constant of σ and the time horizon T. This is the desired estimate with $\rho_{\mathcal{E}}(\cdot) = \rho_X(C\cdot)$.

To construct τ_ω satisfying (3.5), we basically require a transformation $\zeta^\omega : \Omega \to \Omega$ mapping the paths of $\Gamma^{t,\omega}$ to the corresponding paths of $\Gamma^{t,\bar\omega}$. (The dependence of ζ^ω on the fixed path $\bar\omega$ is suppressed in our notation.) Roughly speaking, this is accomplished by the solution of

$$\zeta = \int_0^\cdot \sigma(u+t, \bar\omega \otimes_t \zeta)\sigma(u+t, \omega \otimes_t B)^{-1}\, dB_u.$$

Indeed, let us suppose for the moment that a solution ζ^ω can be defined in some meaningful way and that all paths of ζ^ω are continuous. Then, formally, we have $\zeta^\omega(\Gamma^{t,\omega}) = \Gamma^{t,\bar\omega}$ and

$$\tau_\omega(\omega') := \tau(0 \otimes_t \zeta^\omega(\omega'_{\cdot+t} - \omega'_t)) \tag{3.7}$$

defines a stopping time with the desired property (3.5). In the Appendix, we show how to make this sketch rigorous and include the case of a controlled equation.

4 Proof of Theorem 3.4

Assumptions 3.1, 3.2 and 3.3 are in force throughout this section, in which we state the proof of Theorem 3.4 through a sequence of lemmas.

4.1 Optimality of τ^*

We begin with the optimality of τ^*. All results have their obvious analogues for the discrete case discussed in Remark 3.5; we shall state this separately only where necessary. We first show that $\mathcal{T}_t$ may be replaced by the set $\mathcal{T}^t$ from (3.1) in the definition (3.3) of Y_t.

Lemma 4.1. *Let $t \in [0, T]$. Then*

$$Y_t = \inf_{\tau \in \mathcal{T}^t} \mathcal{E}_t(X_\tau). \tag{4.1}$$

Proof. The inequality "$\leq$" follows from the fact that $\mathcal{T}^t \subseteq \mathcal{T}_t$. To see the reverse inequality, fix $\omega \in \Omega$ and let $\varepsilon > 0$. By the definition (3.3) of Y_t, there exists $\tau \in \mathcal{T}_t$ (depending on ω) such that

$$Y_t(\omega) \geq \mathcal{E}_t(X_\tau)(\omega) - \varepsilon.$$

Define $\theta = \tau^{t,\omega}(B^t)$, where $B^t = B_{\cdot + t} - B_t$. Clearly $\theta \geq t$, and we see from Galmarino's test that θ is a stopping time. Moreover, as a function of B^t, θ is independent of the path up to time t; that is, $\theta \in \mathcal{T}^t$. Noting that $\tau^{t,\omega} = \theta^{t,\omega}$, the definition (2.1) of $\mathcal{E}_t(\cdot)$ shows that $\mathcal{E}_t(X_\tau)(\omega) = \mathcal{E}_t(X_\theta)(\omega)$ and hence

$$Y_t(\omega) \geq \mathcal{E}_t(X_\theta)(\omega) - \varepsilon.$$

The result follows as $\varepsilon > 0$ was arbitrary. □

Lemma 4.2. *We have*

$$|Y_t(\omega) - Y_t(\omega')| \leq \rho_{\mathcal{E}}(\|\omega - \omega'\|_t) \tag{4.2}$$

for all $t \in [0, T]$ and $\omega, \omega' \in \Omega$. In particular, Y_t is $\mathcal{F}_t$-measurable.

Proof. In view of (4.1), this follows from Assumption 3.2. □

The following dynamic programming principle is at the heart of this section.

Lemma 4.3. *Let $0 \le s \le t \le T$. Then*

$$Y_s = \inf_{\tau \in \mathcal{T}^s} \mathcal{E}_s\big(X_\tau \mathbf{1}_{\{\tau < t\}} + Y_t \mathbf{1}_{\{\tau \ge t\}}\big);$$

moreover, $\mathcal{T}^s$ can be replaced by $\mathcal{T}_s$.

Proof. We first show the inequality "$\ge$". Let $\tau \in \mathcal{T}_s$. As $\tau \vee t \in \mathcal{T}_t$, we have $\mathcal{E}_t(X_{\tau \vee t}) \ge Y_t$ by the definition (3.3) of Y. Using the tower property (2.2), it follows that

$$\mathcal{E}_s(X_\tau) = \mathcal{E}_s\big(X_\tau \mathbf{1}_{\{\tau < t\}} + \mathcal{E}_t(X_{\tau \vee t}) \mathbf{1}_{\{\tau \ge t\}}\big) \ge \mathcal{E}_s\big(X_\tau \mathbf{1}_{\{\tau < t\}} + Y_t \mathbf{1}_{\{\tau \ge t\}}\big);$$

where we have used Lemma 2.3 and that all the involved random variables are upper semianalytic. In view of (4.1), taking the infimum over $\tau \in \mathcal{T}_s$ (resp. $\mathcal{T}^s$) yields the claimed inequality.

We now turn to the inequality "$\le$". Fix $\tau \in \mathcal{T}_s$ and set $\Lambda_0 := \{\tau < t\}$. Moreover, let $\varepsilon > 0$ and let $(\Lambda^i)_{i \ge 1}$ be an $\mathcal{F}_t$-measurable partition of the set $\{\tau \ge t\} \in \mathcal{F}_t$ such that the $\|\cdot\|_t$-diameter of Λ^i is smaller than ε for all $i \ge 1$. Fix $\omega^i \in \Lambda^i$. By (4.1), there exist $\tau^i \in \mathcal{T}^t$ such that

$$Y_t(\omega^i) \ge \mathcal{E}_t(X_{\tau^i})(\omega^i) - \varepsilon, \quad i \ge 1.$$

In view of Assumption 3.2 and (4.2), there exist stopping times $\tau^i_\omega \in \mathcal{T}^t$ such that

$$Y_t(\omega) \ge \mathcal{E}_t(X_{\tau^i_\omega})(\omega) - \rho(\varepsilon), \quad \omega \in \Lambda^i, \quad i \ge 1, \tag{4.3}$$

where $\rho(\varepsilon) = \varepsilon + 2\rho_{\mathcal{E}}(\varepsilon)$. Define $\hat{\tau}^i(\omega) := \tau^i_\omega(\omega)$ and

$$\bar{\tau} := \tau \mathbf{1}_{\{\tau < t\}} + \sum_{i \ge 1} \hat{\tau}^i \mathbf{1}_{\Lambda^i}.$$

In view of the measurability condition in Assumption 3.2, we then have $\bar{\tau} \in \mathcal{T}_s$ and $\bar{\tau}^{t,\omega} = (\hat{\tau}^i)^{t,\omega} = (\tau^i_\omega)^{t,\omega}$ for $\omega \in \Lambda^i$. Using also (4.3), the tower property, and $\{\tau < t\} = \{\bar{\tau} < t\}$, we deduce that

$$\begin{aligned}
\mathcal{E}_s\big(X_\tau \mathbf{1}_{\{\tau < t\}} + Y_t \mathbf{1}_{\{\tau \ge t\}}\big) &\ge \mathcal{E}_s\Bigg(X_\tau \mathbf{1}_{\{\tau < t\}} + \sum_{i \ge 1} \mathcal{E}_t(X_{\hat{\tau}^i}) \mathbf{1}_{\Lambda^i}\Bigg) - \rho(\varepsilon) \\
&= \mathcal{E}_s\Bigg(X_\tau \mathbf{1}_{\{\tau < t\}} + \sum_{i \ge 1} \mathcal{E}_t(X_{\bar{\tau}}) \mathbf{1}_{\Lambda^i}\Bigg) - \rho(\varepsilon) \\
&= \mathcal{E}_s\big(X_\tau \mathbf{1}_{\{\tau < t\}} + \mathcal{E}_t(X_{\bar{\tau}}) \mathbf{1}_{\{\tau \ge t\}}\big) - \rho(\varepsilon) \\
&= \mathcal{E}_s\big(X_\tau \mathbf{1}_{\{\tau < t\}} + X_{\bar{\tau}} \mathbf{1}_{\{\tau \ge t\}}\big) - \rho(\varepsilon) \\
&= \mathcal{E}_s(X_{\bar{\tau}}) - \rho(\varepsilon) \\
&\ge Y_s - \rho(\varepsilon).
\end{aligned}$$

As $\tau \in \mathcal{T}_s \supseteq \mathcal{T}^s$ was arbitrary, the result follows by letting ε tend to zero. □

Based on the dynamic programming principle of Lemma 4.3 and Assumption 3.3, we can now establish the path regularity of Y. The following is quite similar to [9, Lemma 4.2].

Lemma 4.4. *There exists a modulus of continuity ρ_Y such that*

$$|Y_s(\omega) - Y_t(\omega)| \leq \rho_Y(\mathbf{d}[(s,\omega),(t,\omega)])$$

for all $s,t \in [0,T]$ and $\omega \in \Omega$.

Proof. We may assume that $s \leq t$. Using Lemma 4.3, we have

$$Y_s(\omega) - Y_t(\omega) \leq \mathcal{E}_s(Y_t)(\omega) - Y_t(\omega) \leq \mathcal{E}_s\big(|Y_t - Y_t(\omega)|\big)(\omega), \quad \omega \in \Omega.$$

(The right-hand side is the conditional expectation of the random variable $\omega' \mapsto |Y_t(\omega') - Y_t(\omega)|$, evaluated at the point ω.) On the other hand, Lemma 4.3 and the subadditivity of $\mathcal{E}_s(\cdot)$ yield that

$$\begin{aligned}
Y_t(\omega) - Y_s(\omega) &= Y_t(\omega) - \inf_{\tau\in\mathcal{T}^s} \mathcal{E}_s\big(X_\tau \mathbf{1}_{\{\tau<t\}} + Y_t \mathbf{1}_{\{\tau\geq t\}}\big)(\omega)\\
&\leq \sup_{\tau\in\mathcal{T}^s} \mathcal{E}_s\big(Y_t(\omega) - X_\tau \mathbf{1}_{\{\tau<t\}} - Y_t \mathbf{1}_{\{\tau\geq t\}}\big)(\omega)\\
&= \sup_{\tau\in\mathcal{T}^s} \mathcal{E}_s\big(Y_t(\omega) - Y_t + (Y_t - X_\tau)\mathbf{1}_{\{\tau<t\}}\big)(\omega)\\
&\leq \mathcal{E}_s\big(|Y_t - Y_t(\omega)|\big)(\omega) + \sup_{\tau\in\mathcal{T}^s} \mathcal{E}_s\big((Y_t - X_\tau)\mathbf{1}_{\{\tau<t\}}\big)(\omega).
\end{aligned}$$

Combining these two estimates, we obtain that

$$|Y_s(\omega) - Y_t(\omega)| \leq \mathcal{E}_s\big(|Y_t - Y_t(\omega)|\big)(\omega) + \sup_{\tau\in\mathcal{T}^s} \mathcal{E}_s\big((Y_t - X_\tau)\mathbf{1}_{\{\tau<t\}}\big)(\omega). \quad (4.4)$$

Set $\delta := \mathbf{d}[(s,\omega),(t,\omega)]$. Then $\delta \geq t-s$ and we may use Lemma 4.2 to estimate the first term in (4.4) as

$$\begin{aligned}
\mathcal{E}_s\big(|Y_t - Y_t(\omega)|\big)(\omega) &= \sup_{P\in\mathcal{P}(s,\omega)} E^P[|Y_t^{s,\omega} - Y_t(\omega)|]\\
&\leq \sup_{P\in\mathcal{P}(s,\omega)} E^P[\rho_{\mathcal{E}}(\|(\omega\otimes_s B) - \omega\|_t)]\\
&\leq \sup_{P\in\mathcal{P}(s,\omega)} E^P[\rho_{\mathcal{E}}(\delta + \|B\|_{t-s})]\\
&\leq \sup_{P\in\mathcal{P}(s,\omega)} E^P[\rho_{\mathcal{E}}(\delta + \|B\|_{\delta})].
\end{aligned}$$

To estimate the second term in (4.4), let $\tau \in \mathcal{T}^s$. As $Y \leq X$ by the definition of Y, we deduce from (3.2) that

$$\begin{aligned}
\mathcal{E}_s\big((Y_t - X_\tau)\mathbf{1}_{\{\tau<t\}}\big)(\omega) &\leq \mathcal{E}_s\big((X_t - X_\tau)\mathbf{1}_{\{\tau<t\}}\big)(\omega) \\
&= \sup_{P\in\mathcal{P}(s,\omega)} E^P\big[\big(X_t^{s,\omega} - (X_\tau)^{s,\omega}\big)\mathbf{1}_{\{\tau^{s,\omega}<t\}}\big] \\
&\leq \sup_{P\in\mathcal{P}(s,\omega)} E^P\big[\rho_X\big(\mathbf{d}[(t,\omega\otimes_s B),(s,\omega\otimes_s B)]\big)\big] \\
&= \sup_{P\in\mathcal{P}(s,\omega)} E^P\big[\rho_X\big(|t-s| + \|B\|_{t-s}\big)\big] \\
&\leq \sup_{P\in\mathcal{P}(s,\omega)} E^P\big[\rho_X\big(\delta + \|B\|_\delta\big)\big].
\end{aligned}$$

Setting $\rho' = \rho_X \vee \rho_{\mathcal{E}}$, we conclude that

$$|Y_s(\omega) - Y_t(\omega)| \leq 2 \sup_{P\in\mathcal{P}(s,\omega)} E^P\big[\rho'\big(\delta + \|B\|_\delta\big)\big] \leq 2\rho(\delta),$$

where ρ is given by Assumption 3.3. It remains to set $\rho_Y = 2\rho$. □

Remark 4.5. We see from Lemmas 4.2 and 4.4 that Y is an adapted process with continuous paths. In view of Assumption 3.1, it follows that $X - Y$ is a càdlàg adapted process with nonpositive jumps. This implies that for every $\varepsilon \geq 0$, the hitting time $\tau^\varepsilon = \inf\{t \in [0,T] : X_t - Y_t \leq \varepsilon\}$ coincides with the contact time $\inf\{t \in [0,T] : (X_t \wedge X_{t-}) - Y_t \leq \varepsilon\}$ and, therefore, is a stopping time (of the raw filtration $\mathbb{F}$). Moreover, it implies the pointwise convergence $\tau^\varepsilon \to \tau^0 \equiv \tau^*$ for $\varepsilon \to 0$, which we will also find useful below.

Lemmas 4.2 and 4.4 also yield the following joint continuity property.

Corollary 4.6. *There exists a modulus of continuity ρ_Y such that*

$$|Y_s(\omega) - Y_t(\omega')| \leq \rho_Y(\mathbf{d}[(s,\omega),(t,\omega)]) + \rho_{\mathcal{E}}(\|\omega - \omega'\|_T)$$

for all $s, t \in [0,T]$ and $\omega, \omega' \in \Omega$. In particular, if $\theta : \Omega \to [0,T]$ is any $\|\cdot\|_T$-continuous function, then Y_θ is again continuous.

The following submartingale property is a consequence of the dynamic programming principle of Lemma 4.3 and "optional sampling."

Lemma 4.7. *Let $s \in [0,T]$ and $\tau \in \mathcal{T}_s$. Then*

$$Y_s \leq \mathcal{E}_s(Y_\tau). \tag{4.5}$$

Proof. By Lemma 4.3, we have $Y_s \le \mathcal{E}_s(Y_t)$ for any deterministic time $t \in \mathcal{T}_s$.

Step 1. We show that (4.5) holds when $\tau \in \mathcal{T}_s$ has finitely many values $t_1 < t_2 < \cdots < t_n$.

We proceed by induction. If $n = 1$, we are in the deterministic case. Suppose that the result holds for $n-1$ values; in particular, for the stopping time $\tau \vee t_2 \in \mathcal{T}_{t_1}$. Then, using the tower property,

$$\begin{aligned}\mathcal{E}_s(Y_\tau) &= \mathcal{E}_s\big(Y_{t_1}\mathbf{1}_{\{\tau=t_1\}} + Y_{\tau\vee t_2}\mathbf{1}_{\{\tau>t_1\}}\big)\\ &= \mathcal{E}_s\big(Y_{t_1}\mathbf{1}_{\{\tau=t_1\}} + \mathcal{E}_{t_1}(Y_{\tau\vee t_2})\mathbf{1}_{\{\tau>t_1\}}\big)\\ &\ge \mathcal{E}_s\big(Y_{t_1}\mathbf{1}_{\{\tau=t_1\}} + Y_{t_1}\mathbf{1}_{\{\tau>t_1\}}\big)\\ &\ge Y_s\ .\end{aligned}$$

Step 2. Let $\tau \in \mathcal{T}_s$ be arbitrary. Let $\tau_n = \inf\{t \in D_n : t \ge \tau\}$, where $D_n = \{k2^{-n}T : k = 0, \dots, 2^n\}$ for $n \ge 1$. Then each τ_n is a stopping time with finitely many values and hence

$$Y_s \le \mathcal{E}_s(Y_{\tau_n}), \quad n \ge 1 \tag{4.6}$$

by Step 1. In view of $|\tau_n - \tau| \le 2^{-n}T$, Lemma 4.4 yields that

$$|Y_{\tau_n} - Y_\tau| \le \rho_Y\big(\mathbf{d}[(\tau, B), (\tau_n, B)]\big) \le \rho_Y\big(2^{-n}T + \|B^\tau\|_{2^{-n}T}\big).$$

In particular,

$$(|Y_{\tau_n} - Y_\tau|)^{s,\omega} \le \rho_Y\big(2^{-n}T + \|(\omega\otimes_s B)^{\tau^{s,\omega}}\|_{2^{-n}T}\big) = \rho_Y\big(2^{-n}T + \|B^{\tau^{s,\omega}-s}\|_{2^{-n}T}\big)$$

and thus

$$\begin{aligned}|\mathcal{E}_s(Y_{\tau_n})(\omega) - \mathcal{E}_s(Y_\tau)(\omega)| &\le \mathcal{E}_s(|Y_{\tau_n} - Y_\tau|)(\omega)\\ &= \sup_{P\in\mathcal{P}(s,\omega)} E^P[(|Y_{\tau_n} - Y_\tau|)^{s,\omega}]\\ &\le \sup_{P\in\mathcal{P}(s,\omega)} E^P[\rho_Y(2^{-n}T + \|B^{\tau^{s,\omega}-s}\|_{2^{-n}T})].\end{aligned} \tag{4.7}$$

Note that $\tau^{s,\omega} - s$ is a stopping time as $\tau \in \mathcal{T}_s$. Thus, the right-hand side tends to zero as $n \to \infty$, by Assumption 3.3. In view of (4.6), this completes the proof. □

Next, we discuss the specifics of the discrete situation as introduced in Remark 3.5; recall that we use the same notation Y for the corresponding value function.

Lemma 4.8. *Let $\mathbb{T} = \{t_0, t_1, \dots, t_n\}$, where $n \in \mathbb{N}$ and $t_0 < \cdots < t_n = T$. Then Y is given by $Y_{t_n} = X_{t_n}$ and*

$$Y_{t_i} = X_{t_i} \wedge \mathcal{E}_{t_i}(Y_{t_{i+1}}), \quad i = 0, \dots, n-1. \tag{4.8}$$

Let $\tau^* = \inf\{t \in \mathbb{T} : Y_t = X_t\}$, then $Y_{\cdot \wedge \tau^*}$ is an $\mathcal{E}$-martingale on $\mathbb{T}$; i.e.,

$$Y_{t_i \wedge \tau^*} = \mathcal{E}_{t_i}(Y_{t_{i+1} \wedge \tau^*}), \quad i = 0, \dots, n-1, \tag{4.9}$$

and in particular $Y_0 = \mathcal{E}(X_{\tau^*})$.

Proof. Note that $X_T = Y_T$ by the definition of Y. Let $i < n$. From (the obvious discrete version of) Lemma 4.3,

$$Y_{t_i} = \inf_{\tau \in \mathcal{T}^{t_i}(\mathbb{T})} \mathcal{E}_{t_i}\big(X_\tau \mathbf{1}_{\{\tau < t_{i+1}\}} + Y_{t_{i+1}} \mathbf{1}_{\{\tau \geq t_{i+1}\}}\big).$$

For any $\tau \in \mathcal{T}^{t_i}(\mathbb{T})$, we have either $\tau \equiv t_i$ or $\tau \geq t_{i+1}$ identically; hence, the right-hand side equals $X_{t_i} \wedge \mathcal{E}_{t_i}(Y_{t_{i+1}})$, which yields (4.8).

We turn to the martingale property. Let $i < n$. On $\{t_i \geq \tau^*\}$, we have $Y_{t_{i+1} \wedge \tau^*} = Y_{t_i \wedge \tau^*}$ and hence $Y_{t_i \wedge \tau^*} = \mathcal{E}_{t_i}(Y_{t_{i+1} \wedge \tau^*})$; whereas on $\{t_i < \tau^*\}$, we have $Y_{t_i} < X_{t_i}$ and so (4.8) yields that

$$Y_{t_i \wedge \tau^*} = Y_{t_i} = \mathcal{E}_{t_i}(Y_{t_{i+1}}) = \mathcal{E}_{t_i}(Y_{t_{i+1} \wedge \tau^*}).$$

This completes the proof of (4.9), which, by the tower property, also shows that $Y_0 = \mathcal{E}(Y_{T \wedge \tau^*}) = \mathcal{E}(Y_{\tau^*}) = \mathcal{E}(X_{\tau^*})$. □

We can now prove the optimality of τ^* by approximating the continuous problem with suitable discrete ones.

Lemma 4.9. *Let $\tau^* = \inf\{t \in [0, T] : Y_t = X_t\}$. Then $Y_0 = \mathcal{E}(X_{\tau^*})$.*

Proof. For $n \geq 1$, let $\mathbb{T}_n = D_n = \{k2^{-n}T : k = 0, \dots, 2^n\}$. Given $t \in [0, T]$, we denote by $\mathcal{T}_n^t := \mathcal{T}^t(\mathbb{T}_n)$ the corresponding set of stopping times and by

$$Y_t^n := \inf_{\tau \in \mathcal{T}_n^t} \mathcal{E}_t(X_\tau)$$

the corresponding value function. In view of $\mathcal{T}_n^t \subseteq \mathcal{T}^t$, we have

$$Y^n \geq Y \quad \text{on} \quad [0, T] \times \Omega.$$

Step 1. There exists a modulus of continuity ρ such that

$$|Y^n - Y| \leq \rho(2^{-n}) \quad \text{on} \quad \mathbb{T}_n \times \Omega. \tag{4.10}$$

Indeed, let $n \geq 1$, $t \in \mathbb{T}_n$ and $\tau \in \mathcal{T}^t$. Then $\vartheta := \inf\{t \in \mathbb{T}_n : t \geq \tau\}$ is in $\mathcal{T}_n^t$ and $0 \leq \vartheta - \tau \leq 2^{-n}T$. Therefore, Assumption 3.1 yields that

$$\begin{aligned}(X_\vartheta - X_\tau)^{t,\omega} &\leq \rho_X\big(\mathbf{d}[(\vartheta^{t,\omega}, \omega \otimes_t B), (\tau^{t,\omega}, \omega \otimes_t B)]\big)\\ &\leq \rho_X\big(2^{-n}T + \|B^{\tau^{t,\omega}-t}\|_{\vartheta^{t,\omega}-\tau^{t,\omega}}\big)\\ &\leq \rho_X\big(2^{-n}T + \|B^\theta\|_{2^{-n}T}\big),\end{aligned}$$

where $\theta := \tau^{t,\omega} - t \in \mathcal{T}$, and hence

$$\begin{aligned}\mathcal{E}_t(X_\vartheta)(\omega) - \mathcal{E}_t(X_\tau)(\omega) &\leq \mathcal{E}_t(X_\vartheta - X_\tau)(\omega)\\ &\leq \sup_{P \in \mathcal{P}(t,\omega)} E^P\big[\rho_X\big(2^{-n}T + \|B^\theta\|_{2^{-n}T}\big)\big]\\ &\leq \rho(2^{-n})\end{aligned}$$

for some modulus of continuity ρ, by Assumption 3.3. As a result,

$$0 \leq Y_t^n - Y_t = \inf_{\vartheta \in \mathcal{T}_n^t} \mathcal{E}_t(X_\vartheta) - \inf_{\tau \in \mathcal{T}^t} \mathcal{E}_t(X_\tau) \leq \rho(2^{-n}).$$

Step 2. Fix $\varepsilon > 0$ and define $\tau^\varepsilon = \inf\{t \in [0,T] : X_t - Y_t \leq \varepsilon\}$. There exists a modulus of continuity ρ' such that for all n satisfying $\rho(2^{-n}) < \varepsilon$,

$$Y_0 \geq \mathcal{E}(Y_{\tau^\varepsilon}) + 2\rho(2^{-n}) + \rho'(2^{-n}).$$

Indeed, let $\tau_n^* = \inf\{t \in \mathbb{T}_n : Y_t^n = X_t\}$. As $\rho(2^{-n}) < \varepsilon$, (4.10) entails that

$$X - Y^n > 0 \quad \text{on} \quad [\![0, \tau^\varepsilon[\![\cap (\mathbb{T}_n \times \Omega);$$

that is, we have $\tau^\varepsilon \leq \tau_n^*$. Define the stopping time

$$\tau^{\varepsilon,n} = \inf\{t \in \mathbb{T}_n : t \geq \tau^\varepsilon\}.$$

Recalling that τ_n^* takes values in $\mathbb{T}_n$, we see that $\tau^\varepsilon \leq \tau_n^*$ even implies that

$$\tau^{\varepsilon,n} \leq \tau_n^*.$$

By Lemma 4.8, the process $Y^n_{\cdot \wedge \tau_n^*}$ is an $\mathcal{E}$-martingale on $\mathbb{T}_n$; in particular, using an optional sampling argument as in Step 1 of the proof of Lemma 4.7,

$$Y_0^n = \mathcal{E}(Y^n_{\tau^{\varepsilon,n} \wedge \tau_n^*}) = \mathcal{E}(Y^n_{\tau^{\varepsilon,n}}).$$

In view of (4.10), this implies that

$$Y_0 \geq \mathcal{E}(Y_{\tau^{\varepsilon,n}}) - 2\rho(2^{-n}). \tag{4.11}$$

On the other hand, an estimate similar to (4.7) and Assumption 3.3 entail that

$$|\mathcal{E}(Y_{\tau^{\varepsilon,n}}) - \mathcal{E}(Y_{\tau^\varepsilon})| \leq \sup_{P\in\mathcal{P}} E^P[\rho_Y(2^{-n}T + \|B^{\tau^\varepsilon}\|_{2^{-n}T})] \leq \rho'(2^{-n})$$

for some modulus of continuity ρ'. Together with (4.11), this yields the claim.

Step 3. Letting $n \to \infty$, Step 2 implies that

$$Y_0 \geq \mathcal{E}(Y_{\tau^\varepsilon}).$$

By Remark 4.5, we have $\tau^\varepsilon \to \tau^*$ for $\varepsilon \to 0$, and as Y has continuous paths (Lemma 4.4), it follows that $Y_{\tau^\varepsilon} \to Y_{\tau^*}$ pointwise. Thus, (an obvious version of) Fatou's lemma yields that $Y_0 \geq \mathcal{E}(Y_{\tau^*})$. Recalling the definition of τ^*, we conclude that

$$Y_0 \geq \mathcal{E}(Y_{\tau^*}) = \mathcal{E}(X_{\tau^*}) \geq \inf_{\tau\in\mathcal{T}} \mathcal{E}(X_\tau) = Y_0.$$

This completes the proof. □

Remark 4.10. The process $Y_{\cdot\wedge\tau^*}$ is a P-supermartingale for any $P \in \mathcal{P}$.

Proof. Let $0 \leq s \leq t \leq T$, where $s \in \mathbb{T}_n$ for some n, and let $\tau \in \mathcal{T}$ be such that $\tau \leq \tau^\varepsilon$. Going through Step 2 of the preceding proof with the appropriate modifications then shows that $Y_{s\wedge\tau} \geq \mathcal{E}_s(Y_\tau)$. Fix $P \in \mathcal{P}$ and note that Assumption 2.1(ii) implies $\mathcal{E}_s(Y_\tau) \geq E^P[Y_\tau|\mathcal{F}_s]$ P-a.s. Choosing $\tau = t \wedge \tau^\varepsilon$, we obtain that $Y_{s\wedge\tau^\varepsilon} \geq E^P[Y_{t\wedge\tau^\varepsilon}|\mathcal{F}_s]$ P-a.s. Now let $\varepsilon \to 0$, then Fatou's lemma yields that $Y_{s\wedge\tau^*} \geq E^P[Y_{t\wedge\tau^*}|\mathcal{F}_s]$. This shows that $Y_{\cdot\wedge\tau^*}$ is a P-supermartingale on $\cup_n \mathbb{T}_n$, and as Y is continuous, this implies the claim. □

4.2 Existence of the Value

The aim of this subsection is to show that the upper value function Y coincides with the lower one, denoted by Z below. As mentioned in the Introduction, there is an obstruction to directly proving the dynamic programming principle for Z in continuous time; namely, we are unable to perform measurable selections on the set of stopping times in the absence of a reference measure. This is related to the measurability problems that are well-known in the literature; see, e.g., [12]. In the following lemma, we consider the discrete setting and prove simultaneously the dynamic programming for the lower value and that it coincides with the upper value.

Lemma 4.11. *Let $\mathbb{T} = \{t_0, t_1, \dots, t_n\}$, where $n \in \mathbb{N}$ and $t_0 < \dots < t_n = T$, define the lower value function*

$$Z_{t_i}(\omega) := \sup_{P \in \mathcal{P}(t_i,\omega)} \inf_{\tau \in \mathcal{T}^{t_i}(\mathbb{T})} E^P[(X_\tau)^{t_i,\omega}], \quad i = 0, \dots, n,$$

and recall the upper value function Y introduced in Remark 3.5. For any $j = 0, \dots, n$, we have

$$Z_{t_j} = Y_{t_j} \tag{4.12}$$

and

$$Z_{t_i}(\omega) = \sup_{P \in \mathcal{P}(t_i,\omega)} \inf_{\tau \in \mathcal{T}^{t_i}(\mathbb{T})} E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_j\}} + Z_{t_j} \mathbf{1}_{\{\tau \geq t_j\}}\big)^{t_i,\omega}\big] \tag{4.13}$$

for all $i = 0, \dots, j$ and $\omega \in \Omega$.

Proof. We proceed by backward induction over j. As $Z_{t_n} = X_{t_n} = Y_{t_n}$, the claim is clear for $j = n$; we show the passage from $j + 1$ to j. That is, we assume that for some fixed $j < n$, we have

$$Z_{t_{j+1}} = Y_{t_{j+1}} \tag{4.14}$$

(which, in particular, entails that $Z_{t_{j+1}}$ is $\mathcal{F}_{t_{j+1}}$-measurable) and

$$Z_{t_i}(\omega) = \sup_{P \in \mathcal{P}(t_i,\omega)} \inf_{\tau \in \mathcal{T}^{t_i}(\mathbb{T})} E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_{j+1}\}} + Z_{t_{j+1}} \mathbf{1}_{\{\tau \geq t_{j+1}\}}\big)^{t_i,\omega}\big], \quad i = 0, \dots, j+1 \tag{4.15}$$

for all $\omega \in \Omega$. We first note that if $\tau \in \mathcal{T}^{t_j}(\mathbb{T})$, then either $\tau \equiv t_j$ or $\tau > t_j$ identically; therefore, (4.15) yields that

$$\begin{aligned} Z_{t_j}(\omega) &= \sup_{P \in \mathcal{P}(t_j,\omega)} \inf_{\tau \in \mathcal{T}^{t_j}(\mathbb{T})} E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_{j+1}\}} + Z_{t_{j+1}} \mathbf{1}_{\{\tau \geq t_{j+1}\}}\big)^{t_j,\omega}\big] \\ &= X_{t_j}(\omega) \wedge \sup_{P \in \mathcal{P}(t_j,\omega)} E^P[Z_{t_{j+1}}^{t_j,\omega}] \\ &= X_{t_j}(\omega) \wedge \mathcal{E}_{t_j}(Z_{t_{j+1}})(\omega). \end{aligned}$$

By the induction assumption (4.14) and the recursion (4.8) for Y, this shows that

$$Z_{t_j} = X_{t_j} \wedge \mathcal{E}_{t_j}(Z_{t_{j+1}}) = X_{t_j} \wedge \mathcal{E}_{t_j}(Y_{t_{j+1}}) = Y_{t_j},$$

which is (4.12). In particular, Z_{t_j} is $\mathcal{F}_{t_j}$-measurable.

Let us now fix $i \in \{0, \dots, j\}$ and prove the remaining claim (4.13). To this end, we first rewrite the latter equation: substituting the just obtained expression $Z_{t_j} = X_{t_j} \wedge \mathcal{E}_{t_j}(Z_{t_{j+1}})$ for Z_{t_j} in the right-hand side of (4.13), and using (4.15) to substitute Z_{t_i} on the left-hand side of (4.13), we see that our claim is equivalent to the identity

$$\begin{aligned}
&\sup_{P \in \mathcal{P}(t_i,\omega)} \inf_{\tau \in \mathcal{T}^{t_i}(\mathbb{T})} E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_{j+1}\}} + Z_{t_{j+1}} \mathbf{1}_{\{\tau \geq t_{j+1}\}}\big)^{t_i,\omega}\big] \\
&\quad = \sup_{P \in \mathcal{P}(t_i,\omega)} \inf_{\tau \in \mathcal{T}^{t_i}(\mathbb{T})} E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_j\}} + \{X_{t_j} \wedge \mathcal{E}_{t_j}(Z_{t_{j+1}})\} \mathbf{1}_{\{\tau \geq t_j\}}\big)^{t_i,\omega}\big].
\end{aligned} \tag{4.16}$$

We first show the inequality "$\geq$" in this equation. To this end, let $\omega \in \Omega$, $\tau \in \mathcal{T}^{t_i}(\mathbb{T})$ and $P \in \mathcal{P}(t_i, \omega)$. In view of (4.14), Lemma 4.2 yields that $Z_{t_{j+1}}$ is continuous and in particular upper semianalytic. Given $\varepsilon > 0$, it then follows from Assumption 2.1 and an application of the Jankov–von Neumann selection theorem similar to Step 2 of the proof of [26, Theorem 2.3] that there exists an $\mathcal{F}_{t_j - t_i}$-measurable kernel $\nu : \Omega \to \mathfrak{P}(\Omega)$ such that

$$\nu(\cdot) \in \mathcal{P}(t_j, \omega \otimes_{t_i} \cdot) \quad \text{and} \quad E^{\nu(\cdot)}\big[Z_{t_{j+1}}^{t_j, \omega \otimes_{t_i} \cdot}\big] \geq \mathcal{E}_{t_j}(Z_{t_{j+1}})^{t_i,\omega}(\cdot) - \varepsilon \tag{4.17}$$

hold P-a.s. Let $\bar{P}$ be the measure defined by

$$\bar{P}(A) = \iint (\mathbf{1}_A)^{t_j - t_i, \omega'}(\omega'')\, \nu(d\omega''; \omega')\, P(d\omega'), \quad A \in \mathcal{F};$$

then $\bar{P} \in \mathcal{P}(t_i, \omega)$ by Assumption 2.1(iii); moreover, $\bar{P}^{t_j - t_i, \cdot} = \nu(\cdot)$ P-a.s. and $\bar{P} = P$ on $\mathcal{F}_{t_j - t_i}$. In view of (4.17), we have

$$\begin{aligned}
E^{\bar{P}}\big[Z_{t_{j+1}}^{t_i,\omega} \big| \mathcal{F}_{t_j - t_i}\big](\cdot) &= E^{\bar{P}^{t_j - t_i, \cdot}}\big[Z_{t_{j+1}}^{t_j, \omega \otimes_{t_i} \cdot}\big] \\
&= E^{\nu(\cdot)}\big[Z_{t_{j+1}}^{t_j, \omega \otimes_{t_i} \cdot}\big] \\
&\geq \mathcal{E}_{t_j}(Z_{t_{j+1}})^{t_i,\omega}(\cdot) - \varepsilon \quad P\text{-a.s.}
\end{aligned}$$

Using this inequality and the tower property of $E^P[\,\cdot\,]$, we deduce that

$$\begin{aligned}
&E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_j\}} + \{X_{t_j} \wedge \mathcal{E}_{t_j}(Z_{t_{j+1}})\} \mathbf{1}_{\{\tau \geq t_j\}}\big)^{t_i,\omega}\big] \\
&\qquad \leq E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_j\}} + X_{t_j} \mathbf{1}_{\{\tau = t_j\}} + \mathcal{E}_{t_j}(Z_{t_{j+1}}) \mathbf{1}_{\{\tau \geq t_{j+1}\}}\big)^{t_i,\omega}\big] \\
&\qquad = E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_{j+1}\}} + \mathcal{E}_{t_j}(Z_{t_{j+1}}) \mathbf{1}_{\{\tau \geq t_{j+1}\}}\big)^{t_i,\omega}\big] \\
&\qquad \leq E^{\bar{P}}\big[\big(X_\tau \mathbf{1}_{\{\tau < t_{j+1}\}} + Z_{t_{j+1}} \mathbf{1}_{\{\tau \geq t_{j+1}\}}\big)^{t_i,\omega}\big] + \varepsilon.
\end{aligned}$$

As $\varepsilon > 0$, $\tau \in \mathcal{T}^{t_i}(\mathbb{T})$ and $P \in \mathcal{P}(t_i, \omega)$ were arbitrary, this implies the inequality "$\geq$" in (4.16).

It remains to show the inequality "$\leq$" in (4.16). To this end, let $\omega \in \Omega$, $\tau \in \mathcal{T}^{t_i}(\mathbb{T})$, $P \in \mathcal{P}(t_i, \omega)$ and define

$$\bar{\tau} := \tau \mathbf{1}_{\{\tau < t_j\}} + (t_j \mathbf{1}_\Lambda + t_{j+1} \mathbf{1}_{\Lambda^c}) \mathbf{1}_{\{\tau \geq t_j\}}, \quad \Lambda := \{X_{t_j} \leq \mathcal{E}_{t_j}(Z_{t_{j+1}})\}.$$

Noting that $Z_{t_j} = X_{t_j} \wedge \mathcal{E}_{t_j}(Z_{t_{j+1}})$ yields $\Lambda = \{X_{t_j} = Z_{t_j}\} \in \mathcal{F}_{t_j}$, we see that $\bar{\tau} \in \mathcal{T}_{t_i}(\mathbb{T})$. After observing that (2.1) and Assumption 2.1(ii) imply

$$\mathcal{E}_{t_j}(Z_{t_{j+1}})^{t_i,\omega} \geq E^P[Z^{t_i,\omega}_{t_{j+1}} | \mathcal{F}_{t_j - t_i}] \quad P\text{-a.s.},$$

we can then use the tower property of $E^P[\,\cdot\,]$ to obtain that

$$\begin{aligned}
E^P\big[&\big(X_{\bar{\tau}} \mathbf{1}_{\{\bar{\tau} < t_{j+1}\}} + Z_{t_{j+1}} \mathbf{1}_{\{\bar{\tau} \geq t_{j+1}\}}\big)^{t_i,\omega}\big] \\
&= E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_j\}} + \{X_{t_j} \mathbf{1}_\Lambda + Z_{t_{j+1}} \mathbf{1}_{\Lambda^c}\} \mathbf{1}_{\{\tau \geq t_j\}}\big)^{t_i,\omega}\big] \\
&\leq E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_j\}} + \{X_{t_j} \mathbf{1}_\Lambda + \mathcal{E}_{t_j}(Z_{t_{j+1}}) \mathbf{1}_{\Lambda^c}\} \mathbf{1}_{\{\tau \geq t_j\}}\big)^{t_i,\omega}\big] \\
&= E^P\big[\big(X_\tau \mathbf{1}_{\{\tau < t_j\}} + \{X_{t_j} \wedge \mathcal{E}_{t_j}(Z_{t_{j+1}})\} \mathbf{1}_{\{\tau \geq t_j\}}\big)^{t_i,\omega}\big].
\end{aligned}$$

As $\tau \in \mathcal{T}^{t_i}(\mathbb{T})$ and $P \in \mathcal{P}(t_i, \omega)$ were arbitrary, this implies the desired inequality "$\leq$" in (4.16). Here we have used the fact that, similarly as in Lemma 4.1, the left-hand side of (4.16) does not change if we replace $\mathcal{T}^{t_i}(\mathbb{T})$ by $\mathcal{T}_{t_i}(\mathbb{T})$. □

We can now show the existence of the value for the continuous-time game by an approximation argument.

Lemma 4.12. *For all $(t, \omega) \in [0, T] \times \Omega$, we have*

$$Z_t(\omega) := \sup_{P \in \mathcal{P}(t,\omega)} \inf_{\tau \in \mathcal{T}^t} E^P[(X_\tau)^{t,\omega}] = \inf_{\tau \in \mathcal{T}^t} \sup_{P \in \mathcal{P}(t,\omega)} E^P[(X_\tau)^{t,\omega}] \equiv Y_t(\omega).$$

Proof. Let $t \in [0, T]$. The inequality $Z_t \leq Y_t$ is immediate from the ordering of infima and suprema in the definitions; we prove the reverse inequality. Given $n \in \mathbb{N}$, we consider $\mathbb{T}_n = \{t\} \cup \{k 2^{-n} T : k = 0, \dots, 2^n\}$ and denote by Y^n and Z^n the corresponding upper and lower value functions as in Lemmas 4.8 and 4.11, respectively. As in Step 1 of the proof of Lemma 4.9, there exists a modulus of continuity ρ such that

$$|Y^n - Y| \leq \rho(2^{-n}) \quad \text{on} \quad \mathbb{T}_n \times \Omega.$$

Moreover, a similar argument as in the mentioned step shows that

$$|Z^n - Z| \leq \rho(2^{-n}) \quad \text{on} \quad \mathbb{T}_n \times \Omega.$$

Since $Y^n_t = Z^n_t$ by Lemma 4.11, we deduce that $|Y_t - Z_t| \leq 2\rho(2^{-n})$, and now the claim follows by letting n tend to infinity. □

4.3 Existence of P^*

An important tool in this subsection is an approximation of hitting times by continuous random times, essentially taken from [9].

Lemma 4.13. *Let $\mathcal{P}$ be weakly compact and $\tau_n = \inf\{t : X_t - Y_t \leq 2^{-n}\}$ for $n \geq 1$. There exist continuous, $\mathcal{F}_T$-measurable functions $\theta_n : \Omega \to [0,T]$ and $\mathcal{F}_T$-measurable sets $\Omega_n \subseteq \Omega$ such that*

$$\sup_{P\in\mathcal{P}} P(\Omega_n^c) \leq 2^{-n} \quad \text{and} \quad \tau_{n-1} - 2^{-n} \leq \theta_n \leq \tau_{n+1} + 2^{-n} \text{ on } \Omega_n.$$

Moreover, $\theta_n \to \tau^*$ P-a.s. for all $P \in \mathcal{P}$.

Proof. In view of Lemmas 4.2 and 4.4 and Assumption 3.1, the first claim can be argued like Step 1 in the proof of [9, Theorem 3.3]. Since $\tau_n \to \tau^*$ by Remark 4.5 and $|\tau_n - \theta_n| \leq 2^{-n}$ on Ω_n, the second claim follows via the Borel-Cantelli lemma. □

We first establish a measure $P^\sharp$ whose restriction to $\mathcal{F}_{\tau^*}$ will be used in the construction of the saddle point.

Lemma 4.14. *Let $\mathcal{P}$ be weakly compact. Then there exists $P^\sharp \in \mathcal{P}$ such that $E^{P^\sharp}[X_{\tau^*}] = Y_0$.*

Proof. As $X_{\tau^*} = Y_{\tau^*}$, we need to find $P^\sharp \in \mathcal{P}$ such that $E^{P^\sharp}[Y_{\tau^*}] \geq Y_0$; the reverse inequality is clear from Lemma 4.9. For $n \geq 1$, let τ_n and θ_n be as in Lemma 4.13. In view of Lemma 4.7 and the definition of $\mathcal{E}(\cdot)$, there exist $P_n \in \mathcal{P}$ such that

$$E^{P_n}[Y_{\tau_n}] \geq \mathcal{E}(Y_{\tau_n}) - 2^{-n} \geq Y_0 - 2^{-n}. \tag{4.18}$$

By passing to a subsequence, we may assume that $P_n \to P^\sharp$ weakly, for some $P^\sharp \in \mathcal{P}$. Recall that Y is bounded. As $\theta_n \to \tau^*$ $P^\sharp$-a.s., it follows from Corollary 4.6 that $E^{P^\sharp}[Y_{\theta_n}] \to E^{P^\sharp}[Y_{\tau^*}]$. Moreover, the weak convergence $P_m \to P^\sharp$ and Corollary 4.6 imply that $\lim_m E^{P_m}[Y_{\theta_n}] = E^{P^\sharp}[Y_{\theta_n}]$ for any fixed n. As a result, we have

$$E^{P^\sharp}[Y_{\tau^*}] = \lim_n \lim_m E^{P_m}[Y_{\theta_n}]. \tag{4.19}$$

On the other hand, for $m \geq n$, we observe that $\tau_m \geq \tau_n$ and therefore $E^{P_m}[Y_{\tau_n}] \geq E^{P_m}[Y_{\tau_m}]$ by the supermartingale property mentioned in Remark 4.10. Using also (4.18), we deduce that

$$\liminf_n \liminf_m E^{P_m}[Y_{\tau_n}] \geq \liminf_m E^{P_m}[Y_{\tau_m}] \geq Y_0.$$

Combining this with (4.19), we obtain that

$$\begin{aligned} Y_0 - E^{P^\sharp}[Y_{\tau^*}] &\leq \liminf_n \liminf_m E^{P_m}[Y_{\tau_n}] - \lim_n \lim_m E^{P_m}[Y_{\theta_n}] \\ &\leq \limsup_n \limsup_m E^{P_m}[|Y_{\tau_n} - Y_{\theta_n}|]. \end{aligned} \tag{4.20}$$

It remains to show that the right-hand side vanishes. To this end, we first note that Lemma 4.13 yields

$$\theta_{n-1} - 2^{1-n} \leq \tau_n \leq \theta_{n+1} + 2^{-n-1} \quad \text{on} \quad \Omega_{n-1} \cap \Omega_{n+1}.$$

Of course, we also have $0 \leq \tau_n \leq T$. Thus, setting

$$\psi_n = \sup \big\{ |Y_t - Y_{\theta_n}| : \, t \in [\theta_{n-1} - 2^{1-n}, \theta_{n+1} + 2^{-n-1}] \cap [0,T] \big\},$$

we have

$$\begin{aligned} E^{P_m}[|Y_{\tau_n} - Y_{\theta_n}|] &\leq E^{P_m}[\psi_n] + 4\|Y\|_\infty P_m(\Omega^c_{n-1} \cup \Omega^c_{n+1}) \\ &\leq E^{P_m}[\psi_n] + 2^{4-n}\|Y\|_\infty. \end{aligned}$$

Moreover, ψ_n is uniformly bounded, and the continuity of θ_k and Corollary 4.6 yield that ψ_n is continuous. Therefore, $E^{P_m}[\psi_n] \to E^{P^\sharp}[\psi_n]$ for each n, and we conclude that

$$\begin{aligned} \limsup_n \limsup_m E^{P_m}[|Y_{\tau_n} - Y_{\theta_n}|] &\leq \limsup_n \limsup_m E^{P_m}[\psi_n] \\ &\leq \limsup_n E^{P^\sharp}[\psi_n] = 0, \end{aligned}$$

where the last step used dominated convergence under $P^\sharp$ and the fact that $\psi_n \to 0$ $P^\sharp$-a.s. due to $\theta_n \to \tau^*$ $P^\sharp$-a.s. In view of (4.20), this completes the proof. □

The measure $P^\sharp$ already satisfies

$$\inf_{\tau \in \mathcal{T},\, \tau \leq \tau^*} E^{P^\sharp}[X_\tau] = E^{P^\sharp}[X_{\tau^*}].$$

(This follows using Remark 4.10; cf. the proof of Lemma 4.17 below.) In order to obtain a saddle point, we need to find an extension of $P^\sharp|_{\mathcal{F}^*_\tau}$ to $\mathcal{F}$ under which "after τ^*, immediate stopping is optimal." As a preparation, we first note the following semicontinuity property.

Lemma 4.15. *The function $P \mapsto \inf_{\tau \in \mathcal{T}^t} E^P[(X_\tau)^{t,\omega}]$ is upper semicontinuous on $\mathfrak{P}(\Omega)$, for all $(t,\omega) \in [0,T] \times \Omega$.*

Proof. We state the proof for $t = 0$; the general case is proved similarly. Let $P_n \to P$ in $\mathfrak{P}(\Omega)$; we need to show that

$$\limsup_{n\to\infty} \inf_{\tau\in\mathcal{T}} E^{P_n}[X_\tau] \leq \inf_{\tau\in\mathcal{T}} E^P[X_\tau].$$

To this end, it suffices to show that given $\varepsilon > 0$ and $\tau \in \mathcal{T}$, there exists $\tau' \in \mathcal{T}$ such that

$$\limsup_{n\to\infty} E^{P_n}[X_{\tau'}] \leq E^P[X_\tau] + \varepsilon. \tag{4.21}$$

Moreover, by an approximation from the right, we may suppose that $\tau = \sum_{i=1}^N t_i \mathbf{1}_{A_i}$ for some $N \in \mathbb{N}$, $t_i \in [0,T]$ and $A_i \in \mathcal{F}_{t_i}$. Given $\delta > 0$, we can find for each $1 \leq i \leq N$ a P-continuity set $D_i \in \mathcal{F}_{t_i}$ (that is, $P(\partial D_i) = 0$) satisfying $P(A_i \Delta D_i) < \delta$. Note that $D_0 := (D_1 \cup \cdots \cup D_N)^c$ is then also a P-continuity set. Define $t_0 := T$ and

$$\tau' := \sum_{i=0}^N t_i \mathbf{1}_{D_i};$$

then $\tau' \in \mathcal{T}$ and $P\{\tau \neq \tau'\} < N\delta$. As X is bounded, it follows that $E^P[|X_\tau - X_{\tau'}|] < \varepsilon$ for $\delta > 0$ chosen small enough, while

$$E^{P_n}[X_{\tau'}] \to E^P[X_{\tau'}]$$

since $X_{\tau'} = \sum_{i=0}^N X_{t_i} \mathbf{1}_{D_i}$, each X_{t_i} is bounded and continuous, and each D_i is a P-continuity set. This implies (4.21). □

We can now construct the kernel that will be used to extend $P^\sharp$.

Lemma 4.16. *Let $\mathcal{P}(t,\omega)$ be weakly compact for all $(t,\omega) \in [0,T] \times \Omega$ and let $\theta \in \mathcal{T}$. There exists an $\mathcal{F}_\theta^*$-measurable kernel $\hat{P}_\theta : \Omega \to \mathfrak{P}(\Omega)$ such that $\hat{P}_\theta(\omega) \in \mathcal{P}(\theta,\omega)$ and*

$$\inf_{\tau\in\mathcal{T}^{\theta(\omega)}} E^{\hat{P}_\theta(\omega)}[(X_\tau)^{\theta,\omega}] = \sup_{P\in\mathcal{P}(\theta,\omega)} \inf_{\tau\in\mathcal{T}^{\theta(\omega)}} E^P[(X_\tau)^{\theta,\omega}]$$

for all $\omega \in \Omega$.

Proof. For brevity, let us define

$$V(t,\omega,P) := \inf_{\tau\in\mathcal{T}^t} E^P[(X_\tau)^{t,\omega}].$$

We first fix $P \in \mathfrak{P}(\Omega)$ and note that $(t,\omega) \mapsto V(t,\omega,P)$ is Borel. To see this, we first observe that

$$V(t,\omega,P) = \inf_{\tau\in\mathcal{T}} E^P[X_{\tau(\cdot)\vee t}(\omega\otimes\cdot)]$$

by the argument of Lemma 4.1. Moreover, let $\mathcal{T}' \subseteq \mathcal{T}$ be a countable set such that for each $\tau \in \mathcal{T}$ there exist $\tau_n \in \mathcal{T}'$ satisfying $\tau_n \downarrow \tau$ P-a.s.; for instance, $\mathcal{T}'$ can be chosen to consists of stopping times of the form $\sum_{i=1}^N t_i \mathbf{1}_{A_i}$, where each t_i is dyadic and A_i belongs to a countable collection generating $\mathcal{F}_{t_i}$. Then we have

$$V(t,\omega,P) = \inf_{\tau\in\mathcal{T}^t} E^P[(X_\tau)^{t,\omega}] = \inf_{\tau\in\mathcal{T}'} E^P[X_{\tau(\cdot)\vee t}(\omega\otimes\cdot)]$$

by dominated convergence and as $(t,\omega) \mapsto E^P[X_{\tau(\cdot)\vee t}(\omega\otimes\cdot)]$ is Borel for every τ by Fubini's theorem, it follows that $(t,\omega)\mapsto V(t,\omega,P)$ is Borel.

On the other hand, we know from Lemma 4.15 that $P \mapsto V(t,\omega,P)$ is upper semicontinuous. Together, it follows that V is Borel as a function on $[0,T]\times\Omega\times\mathfrak{P}(\Omega)$; in particular, $(\omega,P)\mapsto V_\theta(\omega,P) := V(\theta(\omega),\omega,P)$ is again Borel (recall that $\mathcal{T}$ consists of $\mathbb{F}$-stopping times).

For each $(t,\omega)\in[0,T]\times\Omega$, it follows by compactness and Lemma 4.15 that there exists $\hat{P}\in\mathcal{P}(t,\omega)$ such that

$$V(t,\omega,\hat{P}) = \sup_{P\in\mathcal{P}(t,\omega)} V(t,\omega,P);$$

in particular, $P\mapsto V_\theta(\omega,P)$ admits a maximizer for each $\omega\in\Omega$. As the graph of $\mathcal{P}(\theta,\cdot)$ is analytic, the Jankov–von Neumann Theorem in the form of [7, Proposition 7.50(b), p. 184] shows that a maximizer can be chosen in a universally measurable way, which yields the claim. □

Finally, we can prove the remaining result of Theorem 3.4.

Lemma 4.17. *Let $\mathcal{P}(t,\omega)$ be weakly compact for all $(t,\omega)\in[0,T]\times\Omega$, let $P^\sharp$ be as in Lemma 4.14 and let $\hat{P}_{\tau^*}$ be as in Lemma 4.16. Then the measure defined by*

$$P^*(A) = \iint (\mathbf{1}_A)^{\tau^*,\omega}(\omega')\,\hat{P}_{\tau^*}(d\omega';\omega)\,P^\sharp(d\omega),\quad A\in\mathcal{F}$$

is an element of $\mathcal{P}$ and satisfies

$$\inf_{\tau\in\mathcal{T}} E^{P^*}[X_\tau] = E^{P^*}[X_{\tau^*}].$$

Proof. We set $\hat{P} := \hat{P}_{\tau^*}$. After replacing $\hat{P}$ with a Borel kernel ν such that $\nu = \hat{P}$ $\mathbb{P}^\sharp$-a.s., it follows from Assumption 2.1(iii) that $P^*\in\mathcal{P}$. Let $\tau\in\mathcal{T}$;

then the definition of $\hat{P}$ and Lemma 4.12 yield

$$\begin{aligned}
E^{\hat{P}(\omega)}[(X_{\tau\vee\tau^*})^{\tau^*,\omega}] &\geq \inf_{\theta\in\mathcal{T}^{\tau^*(\omega)}} E^{\hat{P}(\omega)}[(X_\theta)^{\tau^*,\omega}] \\
&= \sup_{P\in\mathcal{P}(\tau^*,\omega)} \inf_{\theta\in\mathcal{T}^{\tau^*(\omega)}} E^P[(X_\theta)^{\tau^*,\omega}] \\
&= \inf_{\theta\in\mathcal{T}^{\tau^*(\omega)}} \sup_{P\in\mathcal{P}(\tau^*,\omega)} E^P[(X_\theta)^{\tau^*,\omega}] \\
&= Y_{\tau^*}(\omega)
\end{aligned}$$

for all $\omega \in \Omega$. This means that $E^{P^*}[X_{\tau\vee\tau^*}|\mathcal{F}_{\tau^*}] \geq Y_{\tau^*}$ P^*-a.s. and thus

$$\begin{aligned}
E^{P^*}[X_\tau|\mathcal{F}_{\tau^*}] &= E^{P^*}[X_{\tau\wedge\tau^*}|\mathcal{F}_{\tau^*}]\mathbf{1}_{\{\tau<\tau^*\}} + E^{P^*}[X_{\tau\vee\tau^*}|\mathcal{F}_{\tau^*}]\mathbf{1}_{\{\tau\geq\tau^*\}} \\
&\geq X_{\tau\wedge\tau^*}\mathbf{1}_{\{\tau<\tau^*\}} + Y_{\tau^*}\mathbf{1}_{\{\tau\geq\tau^*\}} \quad P^*\text{-a.s.}
\end{aligned}$$

By Remark 4.10, $Y_{\cdot\wedge\tau^*}$ is a $P^\sharp$-supermartingale, but as $Y_0 = E^{P^\sharp}[Y_{\tau^*}]$ by Lemma 4.14, $Y_{\cdot\wedge\tau^*}$ is even a $P^\sharp$-martingale and hence a P^*-martingale. Using also that $X \geq Y$, we conclude that

$$E^{P^*}[X_\tau|\mathcal{F}_{\tau\wedge\tau^*}] \geq Y_{\tau\wedge\tau^*}\mathbf{1}_{\{\tau<\tau^*\}} + E^{P^*}[Y_{\tau^*}|\mathcal{F}_{\tau\wedge\tau^*}]\mathbf{1}_{\{\tau\geq\tau^*\}} = Y_{\tau\wedge\tau^*} \quad P^*\text{-a.s.}$$

and thus

$$E^{P^*}[X_\tau] \geq E^{P^*}[Y_{\tau\wedge\tau^*}] = E^{P^*}[Y_{\tau^*}] = E^{P^*}[X_{\tau^*}].$$

Since $\tau \in \mathcal{T}$ was arbitrary, this proves the claim. □

5 Application to American Options

In this section, we apply our main result to the pricing of American options under volatility uncertainty. To this end, we interpret B as the stock price process and assume that $\mathcal{P}$ consists of local martingale measures, each of which is seen as a possible scenario for the volatility. More precisely, following [33], we assume that $\mathcal{P}$ is a subset of $\mathcal{P}_S$, the set of all local martingale laws of the form

$$P^\alpha = P_0 \circ \left(\int_0^\cdot \alpha_u^{1/2}\, dB_u\right)^{-1},$$

where P_0 is the Wiener measure and α ranges over all locally square integrable, progressively measurable processes with values in $\mathbb{S}_{++}$. We remark that if $\mathcal{P}$ is not already a subset of $\mathcal{P}_S$, then we may replace $\{\mathcal{P}(s,\omega)\}$ by $\{\mathcal{P}(s,\omega)\cap\mathcal{P}_S\}$ without invalidating Assumption 2.1; cf. [23, Corollary 2.5].

Let $\mathbb{G} = (\mathcal{G}_t)_{0 \le t \le T}$ be the filtration defined by $\mathcal{G}_t = \mathcal{F}_t^* \vee \mathcal{N}^{\mathcal{P}}$, where $\mathcal{F}_t^*$ is the universal completion of $\mathcal{F}_t$ and $\mathcal{N}^{\mathcal{P}}$ is the collection of all sets which are $(\mathcal{F}_T, P)$-null for all $P \in \mathcal{P}$. Let H be an $\mathbb{R}^d$-valued, $\mathbb{G}$-predictable process such that $\int_0^T H_u^\top \, d\langle B \rangle_u \, H_u < \infty$ P-a.s. for all $P \in \mathcal{P}$. Then H is called an admissible trading strategy if the P-integral $\int H \, dB$ is a P-supermartingale, for all $P \in \mathcal{P}$, and we denote by $\mathcal{H}$ the set of all admissible trading strategies.

If X is an American-style option where the buyer chooses the exercise time, then it's buyer's price (or subhedging price) is given by

$$x_*(X) := \sup \bigg\{ x \in \mathbb{R} : \text{ there exist } \tau \in \mathcal{T} \text{ and } H \in \mathcal{H} \text{ such that} \\ X_\tau + \int_0^\tau H_u \, dB_u \ge x \quad P\text{-a.s. for all } P \in \mathcal{P} \bigg\}.$$

This is the supremum of all prices x such that, by using a suitable choice of hedging strategy and exercise time, the buyer will incur no loss, no matter which scenario P occurs. On the other hand, if X is a short position in an American option, so that the seller chooses the exercise time, then the corresponding sellers's price is given by

$$x^*(X) := \inf \bigg\{ x \in \mathbb{R} : \text{ there exist } \tau \in \mathcal{T} \text{ and } H \in \mathcal{H} \text{ such that} \\ x + \int_0^\tau H_u \, dB_u \ge X_\tau \quad P\text{-a.s. for all } P \in \mathcal{P} \bigg\}.$$

Clearly $x_*(X) = -x^*(-X)$, so it suffices to study one of these cases. We state the result for the seller's price x^* (because it matches the sign convention for nonlinear expectations).

Theorem 5.1. *Let Assumptions 3.1, 3.2 and 3.3 hold. Then*

$$x^*(X) = \inf_{\tau \in \mathcal{T}} \mathcal{E}(X_\tau) = \mathcal{E}(X_{\tau^*}) \quad \textit{for} \quad \tau^* = \inf\{t \in [0, T] : Y_t = X_t\},$$

and there exists $H \in \mathcal{H}$ such that $x^*(X) + \int_0^\tau H_u \, dB_u \ge X_{\tau^*}$ P-a.s. for all $P \in \mathcal{P}$; in particular, the infimum defining $x^*(X)$ is attained.

Proof. We set $x^* = x^*(X)$ and $y^* = \inf_{\tau \in \mathcal{T}} \mathcal{E}(X_\tau)$. Let $x > x^*$, then the definition of x^* yields $\tau \in \mathcal{T}$ and $H \in \mathcal{H}$ such that

$$x + \int_0^\tau H_u \, dB_u \ge X_\tau \quad P\text{-a.s. for all } P \in \mathcal{P}.$$

As H is admissible, this implies that $x \geq E^P[X_\tau]$ for all $P \in \mathcal{P}$, and thus $x \geq \mathcal{E}(X_\tau)$. In particular, $x \geq \inf_{\tau \in \mathcal{T}} \mathcal{E}(X_\tau) = y^*$. As $x > x^*$ was arbitrary, this shows that $x^* \geq y^*$.

Conversely, we have $y^* = \mathcal{E}(X_{\tau^*})$ by Theorem 3.4. Moreover, as X_{τ^*} is Borel-measurable and bounded, the (European) superhedging result stated in [23, Theorem 2.3] yields $H \in \mathcal{H}$ such that

$$\mathcal{E}(X_{\tau^*}) + \int_0^\tau H_u \, dB_u \geq X_{\tau^*} \quad P\text{-a.s. for all } P \in \mathcal{P}.$$

Thus, the definition of x^* implies that $x^* \leq \mathcal{E}(X_{\tau^*}) = y^*$. □

Remark 5.2. In view of Remark 3.5, we can show a similar result for Bermudan options; i.e., options where the exercise time can be chosen from a given set $\mathbb{T} = \{t_0, \dots, t_n\}$.

6 Appendix: Proof of Lemma 3.10

In this section, we complete Example 3.9 by showing that Assumption 3.2 is satisfied. We use the setting and notation introduced in that example.

Proof of Lemma 3.10. Let $t \in [0,T]$, $\tau \in \mathcal{T}^t$ and $\bar{\omega} \in \Omega$. Using a discretization of stochastic integrals as in [14] and the fact that the paths of B are continuous, we can define $\mathbb{F}$-progressively measurable processes A^n such that

$$A := \limsup_{n \to \infty} A^n$$

coincides P-a.s. with the usual quadratic variation process of B under P, for any semimartingale law P. Let us also define (with $\infty - \infty := -\infty$, say)

$$a_u := \limsup_{n \to \infty} n \big(A_{u+1/n} - A_u \big) \quad \text{and} \quad \hat{a}_u := a_u \mathbf{1}_{\{a_u \in \mathbb{S}_{++}\}} + 1 \mathbf{1}_{\{a_u \notin \mathbb{S}_{++}\}}; \tag{6.1}$$

then $\hat{a}$ is $\mathbb{F}_+$-progressively measurable and coincides $dt \times P$-a.s. with the squared volatility of B under P, for any $P \in \cup_\omega \mathcal{P}(t,\omega)$. Given $\omega \in \Omega$ and recalling that σ admits the inverse σ^{inv} in its third argument, we may then define the U-valued process

$$\hat{\nu}^\omega_u := \sigma^{inv}(u + t, \omega \otimes_t \cdot, \hat{a}_u^{1/2}).$$

Let $\Gamma^{t,\omega,\nu}$ denote the solution of the SDE with parameters (t,ω) and control $\nu \in \mathcal{U}$. For any $\nu \in \mathcal{U}$, we have by construction that

$$\hat{a}(\Gamma^{t,\omega,\nu}) = \sigma^2(\cdot + t, \omega \otimes_t \Gamma^{t,\omega,\nu}, \nu) \quad P_0\text{-a.s.}$$

and thus

$$\hat{\nu}^{\omega}(\Gamma^{t,\omega,\nu}) = \nu \quad P_0\text{-a.s.} \tag{6.2}$$

We emphasize that these identities indeed hold up to P_0-evanescence (rather than just $dt \times P_0$-a.s.) because $\sigma^2(\cdot + t, \omega \otimes_t \Gamma^{t,\omega,\nu})$ is right-continuous P_0-a.s. and the "derivative" in (6.1) is taken from the right. In particular, (6.2) implies that $\hat{\nu}^{\omega}$ has càdlàg paths $P(t, \omega, \nu)$-a.s. For later use, we also note that $(\omega, \omega') \mapsto \hat{\nu}^{\omega}_u(\omega')$ is $\mathcal{F}_t \otimes \mathcal{F}$-measurable.

Given two paths $\omega, \bar{\omega} \in \Omega$, let us now consider the equation

$$\zeta = \int_0^{\cdot} \sigma(u + t, \bar{\omega} \otimes_t \zeta, \hat{\nu}^{\omega}_u) \sigma(u + t, \omega \otimes_t B, \hat{\nu}^{\omega}_u)^{-1} \, dB_u. \tag{6.3}$$

Under $P(t, \omega, \nu)$, there exists an almost-surely unique strong solution $\zeta^{t,\omega,\nu}$ and it follows via (6.2) that $\zeta^{t,\omega,\nu}(\Gamma^{t,\omega,\nu}) = \Gamma^{t,\bar{\omega},\nu}$ P_0-a.s. However, we need to define the solution universally, without reference to ν. To this end, we again use a discretization as in [14] to define approximate solutions ζ^n (which are $\mathbb{F}_+$-progressively measurable and merely càdlàg, whence the need to have σ defined on $\mathbb{D}$) and set $\zeta'^{\omega} := \limsup_{n \to \infty} \zeta^n$. Since the integrand in (6.3) is $P(t, \omega, \nu)$-a.s. càdlàg, we have that ζ'^{ω} coincides with $\zeta^{t,\omega,\nu}$ $P(t, \omega, \nu)$-a.s.; cf. [14]. In particular, ζ'^{ω} is continuous $P(t, \omega, \nu)$-a.s., so that

$$\zeta^{\omega}_u := \limsup_{q \in \mathbb{Q},\, q \uparrow u} \zeta'^{\omega}_q$$

still coincides with $\zeta^{t,\omega,\nu}$ $P(t, \omega, \nu)$-a.s., while in addition being $\mathbb{F}$-progressively measurable. Moreover, $(\omega, \omega') \mapsto \zeta^{\omega}(\omega')$ is $\mathcal{F}_t \otimes \mathcal{F}$-measurable by construction. While we now have

$$\zeta^{\omega}(\Gamma^{t,\omega,\nu}) = \Gamma^{t,\bar{\omega},\nu} \quad P_0\text{-a.s.}$$

simultaneously for all $\nu \in \mathcal{U}$, as desired, we still have to elaborate on the definition of τ_{ω}. Indeed, we cannot ensure that all paths of ζ^{ω} are continuous, so that the right-hand side of (3.7) is not well defined. (We cannot simply set the irregular paths to zero as in [14] , for then the resulting process would not be $\mathbb{F}$-adapted and so τ_{ω} would not be an $\mathbb{F}$-stopping time as required.)

To simplify the notation, let $\tilde{\zeta}$ be the process defined by

$$\tilde{\zeta}^{\omega}(\omega') := 0 \otimes_t \zeta^{\omega}(\omega'_{\cdot+t} - \omega'_t), \quad \omega' \in \Omega.$$

Given $r \in [0, T]$, let $\| \cdot \|_{1/3,r}$ be the 1/3-Hölder norm for functions considered on $[0, r] \cap \mathbb{Q}$, and note that its computation involves only a countable supremum. Thus,

$$C^{\omega}_r := \big\{ \omega' \in \Omega : \, \|\tilde{\zeta}^{\omega}(\omega')\|_{1/3,r} < \infty \big\} \in \mathcal{F}_r.$$

Moreover, $\tilde{\zeta}^\omega|_{[0,r]}$ is continuous on C_r^ω, and a standard result for the path regularity of martingales shows that C_r^ω has full $P(t,\omega,\nu)$-measure for any $\nu \in \mathcal{U}$. Consider

$$D_r^\omega := \big\{\omega' \in C_r^\omega : \|\tau(\tilde{\zeta}_{\cdot\wedge r}(\omega')) \leq r\big\} \in \mathcal{F}_r.$$

By Galmarino's test, we have that $\tau(\tilde{\zeta}_{\cdot\wedge r}) = \tau(\tilde{\zeta}_{\cdot\wedge r'})$ on $D_r^\omega \cap D_{r'}^\omega$ for any $r, r' \in [0,T]$. Thus,

$$\tau_\omega(\omega') := \begin{cases} \tau(\tilde{\zeta}_{\cdot\wedge r}(\omega')) & \text{if } \omega' \in D_r^\omega, \quad r \in [0,T]; \\ T, & \text{if } \omega' \in (\cup_r D_r^\omega)^c \end{cases}$$

is well defined. To see that the Borel-measurable function τ_ω is an $\mathbb{F}$-stopping time, we observe that $\{\tau_\omega = T\} \in \mathcal{F}_T$ and, for $u < T$,

$$\{\tau_\omega = u\} = \big\{\omega' \in C_u^\omega : \tau(\tilde{\zeta}_{\cdot\wedge u}(\omega')) = u\big\} \in \mathcal{F}_u,$$

due to the fact that $\tilde{\zeta}$ is $\mathbb{F}$-adapted. In fact, we have $\tau_\omega \in \mathcal{T}^t$ by the definition of $\tilde{\zeta}$ and the condition that $\tau \geq t$. Moreover, $(\omega, \omega') \mapsto \tau_\omega(\omega')$ is $\mathcal{F}_t \otimes \mathcal{F}$-measurable by construction. Since C_r^ω has full $P(t,\omega,\nu)$-measure for any $\nu \in \mathcal{U}$, we also have

$$\tau_\omega(0 \otimes_t \Gamma^{t,\omega,\nu}) = \tau(0 \otimes_t \Gamma^{t,\bar{\omega},\nu}) \quad P_0\text{-a.s.}$$

for all $\nu \in \mathcal{U}$, and we deduce as in (3.6) that

$$\begin{aligned} &\big|\mathcal{E}_t(X_\tau)(\bar{\omega}) - \mathcal{E}_t(X_{\tau_\omega})(\omega)\big| \\ &\qquad \leq \sup_{\nu\in\mathcal{U}} E^{P_0}\big[\big|X_{\tau(0\otimes_t\Gamma^{t,\bar{\omega},\nu})}(\bar{\omega} \otimes_t \Gamma^{t,\bar{\omega},\nu}) - X_{\tau(0\otimes_t\Gamma^{t,\bar{\omega},\nu})}(\omega \otimes_t \Gamma^{t,\omega,\nu})\big|\big] \\ &\qquad \leq \rho_X(C\|\bar{\omega} - \omega\|_t) \end{aligned}$$

as desired. □

References

[1] M. Avellaneda, A. Levy, and A. Parás. Pricing and hedging derivative securities in markets with uncertain volatilities. *Appl. Math. Finance*, 2(2):73–88, 1995.

[2] E. Bayraktar and Y.-J. Huang. On the multi-dimensional controller and stopper games. *SIAM J. Control Optim.*, 51(2):1263–1297, 2013.

[3] E. Bayraktar, I. Karatzas, and S. Yao. Optimal stopping for dynamic convex risk measures. *Illinois J. Math.*, 54(3):1025–1067, 2010.

[4] E. Bayraktar and S. Yao. Optimal stopping for non-linear expectations—Part I. *Stochastic Process. Appl.*, 121(2):185–211, 2011.

[5] E. Bayraktar and S. Yao. Optimal stopping for non-linear expectations—Part II. *Stochastic Process. Appl.*, 121(2):212–264, 2011.

[6] E. Bayraktar and S. Yao. Robust optimal stopping under volatility uncertainty. *Preprint arXiv:1301.0091v5*, 2013.

[7] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control. The Discrete-Time Case.* Academic Press, New York, 1978.

[8] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir. Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Comm. Pure Appl. Math.*, 60(7):1081–1110, 2007.

[9] I. Ekren, N. Touzi, and J. Zhang. Optimal stopping under nonlinear expectation. *Stochastic Process. Appl.*, 124(10):3277–3311, 2014.

[10] N. El Karoui. Les aspects probabilistes du contrôle stochastique. In *Ecole d'été de probabilitées de Saint-Flour*, volume 876 of *Lecture Notes in Math.*, pages 73–238, Springer, Berlin, 1981.

[11] N. El Karoui, E. Pardoux, and M. C. Quenez. Reflected backward SDEs and American options. In *Numerical methods in finance*, volume 13 of *Publ. Newton Inst.*, pages 215–231. Cambridge Univ. Press, Cambridge, 1997.

[12] W. H. Fleming and P. E. Souganidis. On the existence of value functions of two-player, zero-sum stochastic differential games. *Indiana Univ. Math. J.*, 38(2):293–314, 1989.

[13] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. W. de Gruyter, Berlin, 2nd edition, 2004.

[14] R. L. Karandikar. On pathwise stochastic integration. *Stochastic Process. Appl.*, 57(1):11–18, 1995.

[15] I. Karatzas and W. D. Sudderth. The controller-and-stopper game for a linear diffusion. *Ann. Probab.*, 29(3):1111–1127, 2001.

[16] I. Karatzas and W. D. Sudderth. Stochastic games of control and stopping for a linear diffusion. In A. Hsiung, Zh. Ying, and C. H. Zhang, editors, *Random Walk, Sequential Analysis and Related Topics: A Festschrift in Honor of Y. S. Chow*, pages 100–117, World Scientific, Singapore, 2007.

[17] I. Karatzas and I.-M. Zamfirescu. Game approach to the optimal stopping problem. *Stochastics*, 77(5):401–435, 2005.

[18] I. Karatzas and I.-M. Zamfirescu. Martingale approach to stochastic differential games of control and stopping. *Ann. Probab.*, 36(4):1495–1527, 2008.

[19] T. J. Lyons. Uncertain volatility and the risk-free synthesis of derivatives. *Appl. Math. Finance*, 2(2):117–133, 1995.

[20] A. Maitra and W. D. Sudderth. The gambler and the stopper. In *Statistics, Probability and Game Theory*, volume 30 of *IMS Lecture Notes Monogr. Ser.*, pages 191–208, Inst. Math. Statist., Hayward, CA, 1996.

[21] A. Matoussi, L. Piozin, and D. Possamai. Second-order BSDEs with general reflection and Dynkin games under uncertainty. *Stochastic Process. Appl.*, 124(7):2281–2321, 2014.

[22] A. Matoussi, D. Possamai, and C. Zhou. Second order reflected backward stochastic differential equations. *Ann. Appl. Probab.*, 23(6):2420–2457, 2013.

[23] A. Neufeld and M. Nutz. Superreplication under volatility uncertainty for measurable claims. *Electron. J. Probab.*, 18(48):1–14, 2013.

[24] A. Neufeld and M. Nutz. Nonlinear Lévy processes and their characteristics. *Preprint arXiv:1401.7253v1*, 2014.

[25] M. Nutz. A quasi-sure approach to the control of non-Markovian stochastic differential equations. *Electron. J. Probab.*, 17(23):1–23, 2012.

[26] M. Nutz and R. van Handel. Constructing sublinear expectations on path space. *Stochastic Process. Appl.*, 123(8):3100–3121, 2013.

[27] S. Peng. G-expectation, G-Brownian motion and related stochastic calculus of Itô type. In *Stochastic Analysis and Applications*, volume 2 of *Abel Symp.*, pages 541–567, Springer, Berlin, 2007.

[28] S. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Process. Appl.*, 118(12):2223–2253, 2008.

[29] S. Peng. Backward stochastic differential equation, nonlinear expectation and their applications. In R. Bhatia, editor, *Proceedings of the International Congress of Mathematicians 2010, Hyderabad, India*, volume 1, pages 393–432, World Scientific, Singapore, 2011.

[30] F. Riedel. Optimal stopping with multiple priors. *Econometrica*, 77(3):857–908, 2009.

[31] A. T. Smith. American options under uncertain volatility. *Appl. Math. Finance*, 9(2):123–141, 2002.

[32] H. M. Soner, N. Touzi, and J. Zhang. Wellposedness of second order backward SDEs. *Probab. Theory Related Fields*, 153(1–2):149–190, 2012.

[33] H. M. Soner, N. Touzi, and J. Zhang. Dual formulation of second order target problems. *Ann. Appl. Probab.*, 23(1):308–347, 2013.

[34] D. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes.* Springer, New York, 1979.

[35] W. A. Zheng. Tightness results for laws of diffusion processes application to stochastic mechanics. *Ann. Inst. H. Poincaré Probab. Statist.*, 21(2):103–124, 1985.