Doubly Reflected BSDEs with Call Protection and their Approximation

Jean-François Chassagneux*, Stéphane Crépey*

Équipe Analyse et Probabilité
Université d’Évry Val d’Essonne
91025 Évry Cedex, France

December 2, 2009

Contents

1 Introduction 2

1.1 Standing Notation .............................................. 3

2 Markovian RIBSDEs 3

2.1 Diffusion Set-Up with Marker Process ...................... 4

2.2 Markovian RIBSDE ............................................ 5

2.3 Connection with Finance ..................................... 8

2.3.1 Verification Principle ................................... 8

2.3.2 Model Specifications .................................... 9

2.4 Analytic Approach ........................................... 11

3 Approximation Results 12

3.1 Approximation of the Forward Process ..................... 13

3.2 Approximation of the Upper Barrier ....................... 13

3.3 Approximation of the RIBSDE ............................... 14

3.3.1 Discretely reflected BSDEs ............................. 16

4 Proofs 17

4.1 Stability of Call Protection Switching Times .............. 17

4.1.1 Proof of Proposition 2.2 ................................ 19

4.1.2 Proof of Proposition 3.1 ................................. 19

*The research of the authors benefited from the support of Ito33 and of the ‘Chaire Risque de crédit’, Fédération Bancaire Française.
1 Introduction

In this work and in the follow-up paper [9], we consider the issue of numerical solution of a doubly reflected backward stochastic differential equation, with an upper barrier which is only active on random time intervals (doubly reflected BSDE with an intermittent upper barrier, or RIBSDE for short henceforth, where the ‘I’ in RIBSDE stands for ‘intermittent’).

From the mathematical point of view, such RIBSDEs and, in the Markovian case, the related variational inequality (VI for short henceforth) approach, were first introduced in Crépey [11]. From the point of view of financial interpretation, RIBSDEs arise as pricing equations of game options (like convertible bonds) with call protection, in which the call times of the option’s issuer are subject to constraints preventing the issuer from calling the bond on certain random time intervals. Moreover, in the standing example of convertible bonds, this protection is typically monitored at discrete times in a possibly very path-dependent way. Calls may thus be allowed or not at a given time depending on the past values of the underlying stock $S$, which leads, after extension of the state space to markovianize the problem, to highly-dimensional pricing problems. Deterministic pricing schemes are then ruled out by the curse of dimensionality, and simulation methods appear to be the only viable alternative.

The purpose of this paper is to propose a practical and mathematically justified approach to the problem of solving numerically by simulation the RIBSDEs that arise as pricing equations of game options with call protection. The main result is Theorem 3.3, which establishes convergences rates for a discrete time approximation scheme by simulation to an RIBSDE.

The practical value of this scheme will be thoroughly assessed in the follow-up paper [9]. For motivation in this regard let us simply briefly mention here that for problems in dimension up to 30, the accuracy of the simulation scheme, in cases where alternative PDE results are available and can be used as a benchmark (low dimensional problems or problems with a
high ‘nominal’ dimension, but endowed with a specific structure allowing one to reduce them
to a much lower ‘effective dimension’ so that an efficient deterministic numerical scheme
is eventually applicable), typically lies in the range of one bp ($10^{-2}\%$) to 1% of relative
error. See Table 1 which also gives computation times of the simulation scheme and of
an alternative deterministic numerical scheme for solving the related variational inequalities
(variational inequalities used to model the original problem, before any reduction of the
state space).

<table>
<thead>
<tr>
<th>$d$</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI$_d$</td>
<td>332s</td>
<td>5332s</td>
<td>44h</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>MC$_d$</td>
<td>154s</td>
<td>212s</td>
<td>313s</td>
<td>474s</td>
<td>628s</td>
</tr>
<tr>
<td>Rel Err</td>
<td>range 1 bp – 1%</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Computation times and numerical accuracy of the simulation scheme (MC$_d$) as
compared to an alternative deterministic scheme (VI$_d$) used as a benchmark.

1.1 Standing Notation

Let us be given a continuous time stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where in the financial inter-
pretation $\mathbb{P}$ denotes a risk-neutral pricing measure. We assume that the filtration $\mathbb{F}$ satisfies
the usual completeness and right-continuity conditions, and that all semimartingales are càdlàg.
Also, since our practical concern consists in pricing a contingent claim with ma-
ture $T$, we set $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ with $\mathcal{F}_0$ trivial and $\mathcal{F}_T = \mathcal{F}$. Moreover, we declare that a
process on $[0, T]$ (resp. a random variable) has to be $\mathbb{F}$-adapted (resp. $\mathbb{F}$-measurable), by
definition. By default in the sequel, all (in)equalities between random variables or processes
are to be understood $d\mathbb{P}$ – almost surely or $d\mathbb{P} \otimes dt$ – almost everywhere, respectively.

We shall denote:
• $c\Omega$, the complement of an event $\Omega \subseteq \Omega$,
• $\mathbb{N}_n = \{0, 1, \ldots, n\}$, for every non-negative integer $n$,
• $R^q$ and $R_1^\otimes q$, the set of $q$-dimensional vectors and row-vectors with real components,
• $|\cdot|_p$ for $p \in [1, +\infty)$, or simply $|\cdot|$ in case $p = 2$, the $p$-norm of an element of $R^q$ or $R_1^\otimes q$,
• in case of a time-grid $t = (t_i)_{i \leq n}$, $|t|_+ = |t| = \max_{i \leq n-1}(t_{i+1}-t_i)$ and $|t|_- = \min_{i \leq n-1}(t_{i+1}-t_i)$,
• $^T$, the transposition operator.

2 Markovian RIBSDEs

In this section we essentially recall the results of Crépey which are of later use in this paper.
2.1 Diffusion Set-Up with Marker Process

Given a \( q \)-dimensional Brownian motion \( W \), let \( X \) be the solution on \( [0, T] \) of the following SDE:

\[
X_t = X_0 + \int_0^t b(s, X_s) \, ds + \int_0^t \sigma(s, X_s) \, dW_s ,
\]

where \( X_0 \in \mathbb{R}^q \) and the coefficients \( b : [0, T] \times \mathbb{R}^q \to \mathbb{R}^q \) and \( \sigma : [0, T] \times \mathbb{R}^q \to \mathbb{R}^{q \otimes q} \) are such that \( (Hx) b, \sigma \) are \( \Lambda \)-Lipschitz continuous in \( x \), uniformly in \( t \), for some positive constant \( \Lambda \), and \( b(t, 0) \) and \( \sigma(t, 0) \) are bounded by \( \Lambda \) over \( [0, T] \).

Let the time-state space \( \mathcal{E} = [0, T] \times \mathbb{R}^q \times \mathcal{K} \) for some finite set \( \mathcal{K} \). Given a function \( u \) of three arguments \( t, x, k \) where the third argument \( k \) takes its values in a discrete set, so that \( k \) can be thought of as referring to the index of a vector or system of functions of time \( t \) and the spatial variable \( x \), we shall denote either \( u(t, x, k) \), or \( u^k(t, x) \), depending on what is more convenient in the context at hand.

Let us further be given a set \( \mathcal{T} = \{ T_0, T_1, \ldots, T_N \} \) of fixed times with \( 0 = T_0 < T_1 < \cdots < T_{N-1} < T_N = T \). On the state-space \( \mathcal{E} \), we then consider the factor process \( X = (X, H) \), where \( X \) is defined by (1), and where the \( \mathcal{K} \)-valued pure jump marker process \( H \) is supposed to be constant except for deterministic jumps at the (strictly) positive \( T_I \)s, from \( H_{T_I^-} \) to

\[
H_{T_I} = \kappa_I(X_{T_I}, H_{T_I^-}) ,
\]

for jump functions \( \kappa_I : \mathbb{R}^q \times \mathcal{K} \to \mathcal{K} \), starting from an initial condition

\[
H_0 = k \in \mathcal{K}
\]

(note that \( H \) does not jump at time \( T_0 = 0 \)).

Remark 2.1 In the financial interpretation (see section 2.3.2), the function \( u \) typically represents a pricing function, and \( \mathcal{T} \), a set of call protection monitoring times. The marker process \( H \) is used for keeping track of the path-dependence of the call protection clauses, in view of ‘markovianizing’ the model.

We suppose that the jump function \( \kappa_I \) is given as

\[
\kappa^k_I(x) = \kappa^k_{I,-} 1_{\{x \in \mathcal{O}\}} + \kappa^k_{I,+} 1_{\{x \notin \mathcal{O}\}} ,
\]

where the \( \kappa^k_{I,\pm} \in \mathcal{K} \) and where \( d \) is the algebraic distance function to an open domain \( \mathcal{O} = \{ x \in \mathbb{R}^q \mid d(x) < 0 \} \) of \( \mathbb{R}^q \).

Remark 2.2 Observe that the function \( \kappa^k_I \) is continuous outside \( \partial \mathcal{O} \). In Crépey [11], one works with ‘abstract’ functions \( \kappa^k_I \), and it is frequently assumed that a certain condition holds ‘at a point \( x \) of continuity of \( \kappa_I^k \)’. In view of the observation above and for the sake of simplicity, we shall rather postulate instead in this paper the stronger condition that \( x \notin \partial \mathcal{O} \).
One shall work under the following regularity assumption on $O$.

\((\text{Ho})\) The distance function $d$ is of class $C^0_b$.

Let us finally be given a non-decreasing sequence of stopping times $\vartheta = (\vartheta_t)_{t \in \mathbb{N}^{N+1}}$ defined by $\vartheta_0 = 0$ and, for every $l \geq 0$:

$$\vartheta_{2l+1} = \inf\{t > \vartheta_{2l}; H_t \notin K\} \land T, \vartheta_{2l+2} = \inf\{t > \vartheta_{2l+1}; H_t \in K\} \land T,$$

relatively to a given subset $K$ of $K$. Observe that the $\vartheta_t$s, to be interpreted as \textit{times of switching of call protection} in the financial interpretation, reduce to $\mathcal{F}$-valued stopping times, and that $\vartheta_{N+1} = T$. 

### 2.2 Markovian RIBSDE

We denote by \((P)\) the class of functions $u$ on $\mathbb{R}^q, [0, T] \times \mathbb{R}^q$ or $\mathcal{E}$ such that $u$ is Borel-measurable, with polynomial growth in its spatial argument $x \in \mathbb{R}^q$. Let us further be given real-valued and continuous \textit{cost functions} $g(t, x)$, $\ell(t, x)$, $h(t, x)$ and $f(t, x, y, z)$ in \((P)\), with $y \in \mathbb{R}$ and $z \in \mathbb{R}^{1 \otimes q}$ in $f$, such that:

- the \textit{running payoff function} $f(t, x, y, z)$ is Lipschitz in $(y, z)$;
- the \textit{payoff function at maturity} $g(x)$ and the \textit{put and call payoff functions} $\ell(t, x)$ and $h(t, x)$ satisfy $\ell \leq h$, $\ell(T, \cdot) \leq g \leq h(T, \cdot)$.

In the sequel, we shall sometimes use the following assumptions

\((\text{Hf})\) $\ell(t, x) = \lambda(t, x) \lor c$, for a constant $c \in \mathbb{R} \cup \{-\infty\}$ and a function $\lambda$ of class $C^{1,2}$ on $[0, T] \times \mathbb{R}^q$ such that

$$\lambda, \mathcal{G}\lambda, \partial\lambda\sigma \in (P),$$

\((\text{Hh})\) $h(t, x)$ is jointly Lipschitz in $(t, x)$.

The Markovian \textit{RIBSDE} with data

$$f(t, X_t, y, z), \xi = g(X_T), \ell(t, X_t), h(t, X_t), \vartheta,$$

\((\mathcal{E})\), is a doubly reflected BSDE (see, e.g., [13, 11]) with lower and upper barriers respectively given by, for $t \in [0, T]$,

$$L_t = \ell(t, X_t), U_t = \sum_{l=0}^{[N/2]} \mathbf{1}_{[\vartheta_{2l}, \vartheta_{2l+1})} \lor \sum_{l=1}^{[(N+1)/2]} \mathbf{1}_{[\vartheta_{2l-1}, \vartheta_{2l})} h(t, X_t).$$

With respect to standard, ‘continuously reflected’ doubly reflected BSDEs, the peculiarity of RIBSDEs is thus that the ‘nominal’ upper obstacle $h(t, X_t)$ is only active on the ‘odd’ random time intervals $[\vartheta_{2l-1}, \vartheta_{2l})$, $l > 0$.

Let us introduce the following Banach (or Hilbert, in case of $\mathcal{L}^2$ or $\mathcal{H}_p^2$) spaces of random variables or processes, where $p$ denotes here and henceforth a real number in $[1, \infty)$:

- $\mathcal{L}^p$, the space of real valued random variables $\xi$ such that

$$\|\xi\|_{\mathcal{L}^p} = \left(\mathbb{E}[|\xi|^p]\right)^{\frac{1}{p}} < +\infty;$$
• $\mathcal{S}_d^p$, for any real $p \geq 2$ (or $\mathcal{S}_d$, in case $q = 1$), the space of $\mathbb{R}^d$-valued càdlàg processes $Y$ such that

$$\|Y\|_{\mathcal{S}_d^p} := \left( \mathbb{E} \left[ \sup_{t \in [0,T]} |Y_t|^p \right] \right)^{\frac{1}{p}} < +\infty;$$

• $\mathcal{H}_q^p$ (or $\mathcal{H}^p$, in case $d = 1$), the space of $\mathbb{R}^{1,q}$-valued predictable processes $Z$ such that

$$\|Z\|_{\mathcal{H}_q^p} = \left( \mathbb{E} \left[ \int_0^T |Z_t|^2 \, dt \right] \right)^{\frac{1}{2}} < +\infty;$$

• $\mathcal{A}^2$, the space of finite variation processes $A$ with (non-decreasing) Jordan components $A^\pm \in \mathcal{S}_2$ null at time 0.

Under $(H\xi)$, one thus has $\|X\|_{\mathcal{S}^2} \leq C_\Lambda$, where from now on $C_\Lambda$ is a generic constant which depends only on $\Lambda, T, X_0$ and $q$ (in case this constant depends on some extra parameter, say $\rho$, we shall write $C_\rho^\Lambda$). Note that the values of $\Lambda$ and $C_\Lambda$ may change from line to line.

**Definition 2.3** An $(\Omega, \mathcal{F}, \mathbb{P})$-solution $\mathcal{Y}$ to $(\mathcal{E})$ is a triple $\mathcal{Y} = (Y, Z, A)$, such that:

1. $Y \in \mathcal{S}^2, Z \in \mathcal{H}_q^2, A \in \mathcal{A}^2, A^+ \text{ is continuous, and}$
   \[ \{ (\omega, t); \Delta Y \neq 0 \} \subseteq \bigcup_{i=0}^{[N/2]} [N \vartheta_2], \Delta Y = \Delta A^- \text{ on } \bigcup_{i=0}^{[N/2]} [N \vartheta_2], \]
2. $Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s) \, ds + A_T - A_t - \int_t^T Z_s \, dW_s,$ $t \in [0, T],$
3. $L_t \leq Y_t \text{ on } [0, T], Y_t \leq U_t \text{ on } [0, T]$
   \[ \text{and } \int_0^T (Y_t - L_t) \, dA^+_t = \int_0^T (U_t - Y_t) \, dA^-_t = 0, \]

where $L$ and $U$ are defined by \( \boxed{7} \), and with the convention that $0 \times \pm \infty = 0$ in (iii).

Note that this definition admits an obvious extension to a random terminal time $\theta$, instead of constant $\theta = T$ above. This extension will be used in the next results, in the special case of simply reflected and (continuously) doubly reflected BSDEs.

Also note that $(\mathcal{E})$ is implicitly parameterized by the initial condition $(t = 0, x, k)$ of $\mathcal{X}$. In the sequel we let the superscript $^t$ (whenever necessary) stand in reference to an initial condition $(t, x, k)$ of $\mathcal{X}$ (with in particular $t \in [0, T]$, rather than $t = 0$ implicitly above).\(^1\)

By application of the results of [11], one thus has,

**Proposition 2.1 (Crépey [11])** We assume $(H\xi)$.

1. The following iterative construction is well-defined, for $l$ decreasing from $N$ to $0$: $\mathcal{Y}^{l,t} = (Y^{l,t}, Z^{l,t}, A^{l,t})$ is the unique solution, with $A^{l,t}$ continuous, to the reflected BSDE with random terminal time $\vartheta_{l+1}^t$ (for $l$ even) or the doubly reflected BSDE with random terminal time $\vartheta_{l+1}^t$ (for $l$ odd) on $[t, \vartheta_{l+1}^t]$ with data

$$\begin{cases}
  f(s, X_s^t, y, z), \ Y_{\vartheta_{l+1}^t}^t, \ \ell(s, X_s^t), \ \vartheta_{l+1}^t & (l \text{ even}) \\
  f(s, X_s^t, y, z), \ \min(Y_{\vartheta_{l+1}^t}^t, h(\vartheta_{l+1}^t, X_{\vartheta_{l+1}^t}^t)), \ \ell(s, X_s^t), \ h(s, X_s^t), \ \vartheta_{l+1}^t & (l \text{ odd})
\end{cases} \quad (8)
$$

\(^1\)However it is convenient to extend all our processes to $[0, T]$ 'in a natural way' so that they live in spaces of functions defined over $[0, T]$, which do not change with $t$, see Crépey [11].
where, in case \( l = N \), \( Y_{t+1}^{l,t} \) is to be understood as \( g(X_t^l) \).

(ii) Let us define \( \mathcal{Y}^l = (Y^l, Z^l, A^l) \) on \([t, T]\) by, for every \( l = 0, \ldots, N \):

- \( (Y^l, Z^l) = (Y^{l,t}, Z^{l,t}) \) on \([\vartheta_{t+1}^l, \vartheta_{t+1}^l]\), and also at \( \vartheta_{t+1}^l = T \) in case \( l = N \),
- \( dA^l = dA^{l,t} \) on \((\vartheta_{t+1}^l, \vartheta_{t+1}^l)\),

\[
\Delta A_{\vartheta_{t+1}^l}^l = Y_{\vartheta_{t+1}^l}^{l,t} - \min\left(Y_{\vartheta_{t+1}^l}^{l,t}, h(\vartheta_{t+1}^l, X_{\vartheta_{t+1}^l}^l)\right) = \Delta Y_{\vartheta_{t+1}^l}^l (= 0 \text{ for } l \text{ odd})
\]

and \( \Delta A_T^l = \Delta Y_T^l = 0 \). So in particular

\[
Y_t^l = \begin{cases} Y_0^{0,t}, & k \in K \\ Y_1^{l,t}, & k \notin K \end{cases},
\]

where \( k \) is the index which is implicit in the condition initial \((t, x, k)\) of \( \mathcal{X} \) referred to by the superscript \( t \).

Then \( \mathcal{Y}^l = (Y^l, Z^l, A^l) \) is the unique solution to the RIBSDE \((\mathcal{E}^l)\).

Note in particular that existence and uniqueness of solutions with a continuous reflecting process \( A^{l,t} \) to the auxiliary reflected BSDEs and doubly reflected BSDEs with random terminal time that appear in point (i) above, is granted by the results of [12, 11].

One will need further stability results on \( \mathcal{Y}^l \), or, more precisely, on the \( \mathcal{Y}^{l,t} \)s. Toward this end a suitable stability assumption on \( \vartheta^l \) is needed. Our next result is thus a càdlàg property of \( \vartheta \), viewed as a random function of the initial condition \((t, x, k)\) of \( \mathcal{X}^l \).

**Proposition 2.2** At every \((t, x, k)\) in \( \mathcal{E} \), \( \vartheta \) is, almost surely:

(i) continuous at \((t, x, k)\) if \( t \notin \mathcal{F} \), and right-continuous at \((t, x, k)\) if \( t \in \mathcal{F} \),

(ii) left-limited at \((t, x, k)\) if \( t = T_I \in \mathcal{F} \) and \( x \notin \partial \mathcal{O} \).

Let us denote, for \( I = 1, \ldots, N \),

\[
\mathcal{E}_I = \mathcal{E} \cap ([T_{I-1}, T_I] \times \mathbb{R}^q \times \mathcal{K}) , \quad \mathcal{E}^*_I = \mathcal{E} \cap ([T_{I-1}, T_I] \times \mathbb{R}^q \times \mathcal{K}).
\]

By the above statements, we mean that:

- (i) \( \vartheta_{l+1}^n \to \vartheta^l \) if \((t_n, x_n, k) \to (t, x, k)\) with \( t \notin \mathcal{F} \), or, for \( t = T_I \in \mathcal{F} \), if \( \mathcal{E}_{I+1} \ni (t_n, x_n, k) \to (T_I, x, k) \);

- (ii) if \( \mathcal{E}_I^* \ni (t_n, x_n, k) \to (t = T_I, x \notin \partial \mathcal{O}, k) \), then \( \vartheta_{l+1}^n \) converges to some non-decreasing sequence \( \vartheta^l = (\vartheta_{l+1}^n)_{n \in \mathbb{N}} \) of \([0, T]\)-valued stopping times.

Observe that since the \( \vartheta_{l+1}^n \) are in fact \( \mathcal{F} \)-valued stopping times:

- the continuity statement effectively means that \( \vartheta_{l+1}^n = \vartheta^l \) for \( n \) large enough, almost surely, for every \( l = 1, \ldots, N + 1 \) and \( \mathcal{E} \ni (t_n, x_n, k) \to (t, x, k) \in \mathcal{E} \) with \( t \notin \mathcal{F} \);

- the right-continuity, resp. left-limit statement, effectively means that for \( n \) large enough \( \vartheta_{l+1}^n = \vartheta^l \), resp. \( \vartheta_{l+1}^n \), almost surely, for every \( l = 1, \ldots, N + 1 \) and \( \mathcal{E}_{I+1} \ni \), resp. \( \mathcal{E}_I^* \ni (t_n, x_n, k) \to (T_I, x, k) \in \mathcal{E} \).

**Definition 2.4** One denotes by \( \tilde{\mathcal{Y}}^l = (\tilde{\mathcal{Y}}^{l,t})_{t \in \mathbb{N}} \), with \( \tilde{\mathcal{Y}}^{l,t} = (\tilde{Y}^{l,t}, \tilde{Z}^{l,t}, \tilde{A}^{l,t}) \) and \( \tilde{A}^{l,t} \) continuous for every \( l = 0, \ldots, N \), the sequence of solutions of the BSDEs with random terminal times which is obtained by substituting \( \vartheta^l \) to \( \vartheta^t \) in the construction of \( \mathcal{Y}^l \) in Proposition 2.1(i).
Proposition 2.3 (Crépey [11]) We assume \((H_\ell)\) and \((H_h)\). Let \(Y^t = (Y^{l,t})_{l \in \mathbb{N}}\) and \(\tilde{Y}^t = (\tilde{Y}^{l,t})_{l \in \mathbb{N}}\) be defined as in Proposition 2.1(i) and Definition 2.4, respectively. Then, for every \(l = N, \ldots, 0\):

(i) One has the following bound estimate on \(Y^l_t\),
\[
\|Y^l_t\|_{S^2}^2 + \|Z^l_t\|_{H^2_q}^2 + \|A^l_t\|_{S^2}^2 \leq C(1 + |x|^q) .
\]

Moreover, an analogous bound estimate is satisfied by \(\tilde{Y}^l_t\);

(ii) \(t^n\) referring to a perturbed initial condition \((t^n, x^n, k)\) of \(X\), then:

- in case \(t \notin T\), \(Y^{l,t^n} S^2 \times H^2_q \times S^2\) converges to \(Y^{l,t}\) as \(E \ni (t^n, x^n, k) \rightarrow (t, x, k)\);

- in case \(t = T_I \in T\):
  - \(Y^{l,t^n} S^2 \times H^2_q \times S^2\) converges to \(Y^{l,t}\) as \(E \cup \{T_I\} \ni (t^n, x^n, k) \rightarrow (t, x, k)\);
  - if \(x \notin \mathcal{O}\), then \(Y^{l,t^n} S^2 \times H^2_q \times S^2\) converges to \(\tilde{Y}^{l,t}\) as \(E^* \cap \{T_I\} \ni (t^n, x^n, k) \rightarrow (t, x, k)\).

2.3 Connection with Finance

We consider in this section the special case of an affine coefficient
\[
f = f(t, x, y, z) = c(t, x) - \mu(t, x)y + \eta(t, x)z^T,
\]
for continuous bounded real-valued and \(\mathbb{R}^{1 \otimes q}\)-valued functions \(\mu(t, x)\) and \(\eta(t, x)\). In this case, it is straightforward to verify the following classical

Lemma 2.4 \(\mathcal{Y} = (Y, Z, A)\) denoting a solution to \((\mathcal{E})\), the triple
\[
\left(\beta Y, \beta (Z + Y\eta), \int_0^t \beta dA_t\right)
\]
solves the RIBSDE with data (cf. (6))
\[
\beta c(t, X_t), \beta g(X_T), \beta \ell(t, X_t), \beta h(t, X_t), \vartheta ,
\]
where the adjoint process \(\beta\) is the solution of the following linear (forward) SDE:
\[
d\beta_t = \beta_t \left(\eta(t, X_t) dW_t - \mu(t, X_t) dt\right), \quad t \in [0, T]
\]
with the initial condition \(\beta_0 = 1\). In particular, \(\beta > 0\) on \([0, T]\).

2.3.1 Verification Principle

Using Lemma 2.4, the following verification principle can be established in a standard way (see, e.g., [3, 11]). This result establishes the connection between a solution \(\mathcal{Y} = (Y, Z, A)\) of the RIBSDE \((\mathcal{E})\) with an affine coefficient \(f\) as of (12), and a related Dynkin Game, or optimal game problem (see [14]).

Let \(T_t\) and \(T_t^\theta\) respectively denote the sets of the \([t, T]\)-valued and of the \(\cup_{l \geq 0} [\vartheta_{2l-1} \vee t, \vartheta_{2l} \vee t] \cup \{T\}\)-valued stopping times, respectively. Let also \(\zeta = \tau \wedge \theta\), for any \(\tau, \theta \in T_t\).
Proposition 2.5 Let $\mathcal{Y} = (Y, Z, A)$ denote a solution to \((E)\).

(i) $Y$ is the conditional value process of the Dynkin game with cost criterion $\mathbb{E}_t(\pi^t(\tau, \theta))$ on $T_t \times T_t^0$, where $\pi^t(\tau, \theta)$ is the $\mathcal{F}_\tau$-measurable random variable defined by, with $\zeta = \tau \land \theta$,

$$
\beta_t \pi^t(\tau, \theta) = \int_t^\zeta \beta_s c(s, X_s)ds + \beta_\zeta \left( I_{\{\zeta = \tau < T\}} L_\tau + I_{\{\zeta = \theta < T\}} U_\theta + I_{\{\zeta = T\}} \xi \right).
$$

One thus has $\mathbb{P}$-almost surely, for every $t \in [0, T]$,

$$
\esssup_{\tau \in T_t} \essinf_{\theta \in T_t^0} \mathbb{E}_t \pi^t(\tau, \theta) = Y_t = \essinf_{\theta \in T_t^0} \esssup_{\tau \in T_t} \mathbb{E}_t \pi^t(\tau, \theta).
$$

More precisely, for any $t \in [0, T]$ and for any $\varepsilon > 0$, the pair of stopping times $(\tau^\varepsilon, \theta^\varepsilon) \in T_t \times T_t^0$ given by

$$
\tau^\varepsilon = \inf \left\{ u \in [t, T] ; Y_u \leq \ell(u, X_u) + \varepsilon \right\} \land T
$$

$$
\theta^\varepsilon = \inf \left\{ u \in \bigcup_{l \geq 0} \left[ \vartheta_{2l+1} \lor t, \vartheta_{2l+2} \lor t \right) ; Y_u \geq U_u - \varepsilon \right\} \land T,
$$

is an $\varepsilon$-saddle-point for this Dynkin game at time $t$, in the sense that one has, for any $(\tau, \theta) \in T_t \times T_t^0$,

$$
\mathbb{E}_t(\pi^t(\tau, \theta^\varepsilon)) - \varepsilon \leq Y_t \leq \mathbb{E}_t(\pi^t(\tau^\varepsilon, \theta)) + \varepsilon.
$$

(ii) If the component $A$ of $\mathcal{Y}$ is continuous, then the pair of stopping times $(\tau^*, \theta^*) \in T_t \times T_t^0$ obtained by setting $\varepsilon = 0$ in \((16)\), is a saddle-point of the game. One thus has in this case, for any $(\tau, \theta) \in T_t \times T_t^0$,

$$
\mathbb{E}_t(\pi^t(\tau, \theta^*)) \leq Y_t \leq \mathbb{E}_t(\pi^t(\tau^*, \theta)).
$$

2.3.2 Model Specifications

In the case of risk-neutral pricing problems in finance, the driver coefficient function $f$ is typically given as

$$
f = f(t, x, y) = c(t, x) - \mu(t, x)y,
$$

for dividend and interest-rate related functions $c$ and $\mu$. Note that $f$ in \((18)\) is affine in $y$ and does not depend on $z$. We are thus in the sub-case of \((12)\) corresponding to $\eta = 0$, and therefore, in view of \((14)\),

$$
\beta = \exp \left( - \int_0^t \mu(t, X_t)dt \right),
$$

where, in the standard risk-neutral pricing approach, $\mu(t, X_t)$ is interpreted as a short-term risk-free interest rate. Moreover, in the financial interpretation:

- $g(X_T^t)$ corresponds to a terminal payoff that is paid by the issuer to the holder at time $T$ if the contract was not exercised before $T$;
- $\ell(X_t^t)$, resp. $h(X_t^t)$, corresponds to a lower, resp. upper payoff that is paid by the issuer to the holder of the claim in the event of early termination of the contract at the initiative of the holder, resp. issuer;
- The sequence of stopping time $\vartheta$ is interpreted as a sequence of times of switching of a call protection. More precisely, the issuer of the claim is allowed to call it back (enforcing
early exercise) on the ‘odd’ (random) time intervals \([\vartheta_{2l-1}, \vartheta_{2l})\). At other times call is not possible.

The contingent claims under consideration are thus general *game contingent claims* \([16]\), covering convertible bonds, American options (and also European options) as special cases.

Now, in view of the above verification principle and of the arbitrage theory for game options (see, e.g., \([11]\)), \(\Pi = Y\) is an arbitrage price process for the game option, the arbitrage price relative to the pricing measure \(P\).

Given a suitable set of hedging instruments, \(\Pi\) is also a *bilateral super-hedging price* (see, e.g., \([11, 4]\)), in the sense that there exists a self-financing super-hedging strategy for the issuer of the claim starting from any issuer initial wealth greater then \(\Pi\) and a self-financing super-hedging strategy for the holder of the claim starting from any holder initial wealth greater than \(-\Pi\). Finally \(\Pi\) is also the infimum of the initial wealths of all the issuer’s self-financing super-hedging strategies.

**Remark 2.5** (i) Modeling the pricing problem under the historical, as opposed to the risk-neutral, probability, would lead to a ‘\(z\)-dependent’ driver coefficient function \(f\).

(ii) The standard risk-neutral pricing approach tacitly assumes a perfect, frictionless financial market. Accounting for market imperfections would lead to a nonlinear coefficient \(f\) (see, e.g., El Karoui et al. \([15]\)).

(iii) In a context of *vulnerable claims* (defaultable game options \([2]\)), it is enough, to account for counterparty risk, to work with suitably *credit-spread adjusted interest-rates* \(\mu\) and *recovery-adjusted dividend-yields* \(c\) in \([18]\), and to amend accordingly the dynamics of the factor process \(X\) (see, e.g., \([11]\)).

A rather typical specification of the terminal cost functions is given by, for constants \(\bar{P} \leq \bar{N} \leq \bar{C}\),

\[
\ell(t, x) = \bar{P} \lor S, \ h(t, x) = \bar{C} \lor S, \ g(x) = \bar{N} \lor S, \tag{19}
\]

where \(S = x_1\) denotes the first component of \(x\). Note that this specification satisfies assumptions \((H\ell)-(Hh)\), as well as all the standing assumptions of this paper. In particular, one then has (cf. \([1], [22]\)),

\[
\lambda(t, x) = x_1 = S, \ \mathcal{G}\lambda = b_1, \ \partial\lambda\sigma = \sigma_1,
\]

so that condition \([5]\) in \((Hh)\) reduces to \(b_1, \sigma_1 \in (P)\), which holds by the Lipschitz property of \(b\) and \(\sigma\).

As for \(\vartheta\), the following specifications are commonly found in the case of convertible bonds on an underlying stock \(S\).

**Example 2.6** Given a constant *trigger level* \(\bar{S}\) and a constant \(l \leq N\):

(i) \(\mathcal{K} = \mathbb{N}_l, \ K = \mathbb{N}_{l-1}\) and \(\kappa\) defined by

\[
\kappa^k_l(x) = \begin{cases} 
(k + 1) \land l, & S \geq \bar{S} \\
0, & S < \bar{S}
\end{cases}
\]

(independently of \(l\)). With the initial condition \(H_0 = 0, H_t\) then represents the number of consecutive monitoring dates \(T_j\)s with \(S_{T_j} \geq \bar{S}\) from time \(t\) backwards, capped at \(l\). Call is
possible whenever \( H_t \geq l \), which means that \( S \) has been \( \geq \bar{S} \) at the last \( l \) monitoring times; Otherwise call protection is in force;

(ii) \( \mathcal{K} = \{0, 1\}^d \) for some given integer \( d \in \{l, \ldots, N\}, \mathcal{K} = \{k \in \mathcal{K}; |k|_1 < l\} \) with \(|k|_1 = \sum_{1 \leq p \leq d} k_p\), and \( \kappa \) defined by

\[
\kappa^k(x) = (1_{S \geq \bar{S}}, k_1, \ldots, k_{d-1}).
\]

With the initial condition \( H_0 = 0_d, H_t \) then represents the vector of the indicator functions of the events \( S_{T_k} \geq \bar{S} \) at the last \( d \) monitoring dates preceding time \( t \). Call is possible whenever \(|H_t| \geq l\), which means that \( S \) has been \( \geq \bar{S} \) on at least \( l \) of the last \( d \) monitoring times; Otherwise call protection is in force.

### 2.4 Analytic Approach

The main contribution of this article consists in a simulation scheme, shown to be convergent with a certain speed, for solving the RIBSDE \((\mathcal{E})\). However, for the sake of the numerical validation of the results of the simulation scheme, it will be useful to be able to compare them with those of an alternative, deterministic numerical scheme. A deterministic scheme for the Markovian RIBSDE \((\mathcal{E})\) is based on the analytic characterization of \((\mathcal{E})\), or, more precisely, of a related value function \( u \), in terms of an associated system of VIs.

Note that the sets \( \mathcal{E}_t \)s and \( \{T\} \times \mathbb{R}^q \times \mathcal{K} \) partition \( \mathcal{E} \).

In view of introducing the value function \( u \) in Proposition 2.6, it is convenient to state the following definition.

**Definition 2.7** (i) A Cauchy cascade \((g, \mathcal{U})\) on \( \mathcal{E} \) is pair made of a terminal condition \( g \) of class \((P)\) at \( T \), along with a sequence \( \mathcal{U} = (u_I)_{1 \leq I \leq N} \) of functions \( u_I \)s of class \((P)\) on the \( \mathcal{E}_I \)s, satisfying the following jump condition, at every \( x \notin \partial \mathcal{O} \):

\[
u_I^k(T_I, x) = \begin{cases} \min(u_{I+1}(T_I, x, \kappa^k_I(x)), h(T_I, x)) & \text{if } k \notin \mathcal{K} \text{ and } \kappa^k_I(x) \in \mathcal{K}, \\ u_{I+1}(T_I, x, \kappa^k_I(x)) & \text{else}, \end{cases}
\]

where, in case \( I = N \), \( u_{I+1} \) is to be understood as \( g \).

A continuous Cauchy cascade is a Cauchy cascade with continuous ingredients \( g \) at \( T \) and \( u_I \)s on the \( \mathcal{E}_I \)s, except maybe for discontinuities of the \( u_I \)s at the points \((T_I, x)\) with \( x \notin \partial \mathcal{O} \).

(ii) The function defined by a Cauchy cascade is the function on \( \mathcal{E} \) given as the concatenation on the \( \mathcal{E}_I \)s of the \( u_I \)s, and by the terminal condition \( g \) at \( T \).

**Remark 2.8** So, for \( x \notin \partial \mathcal{O} \), \( u_I^k(t_n, x_n) \) may fail to converge to \( u_I^k(T_I, x) \) as \( \mathcal{E}_I \ni (t_n, x_n, k) \to (T_I, x, k) \).

One then has,

**Proposition 2.6** Assuming \((H\ell)\) and \((Hh)\), the state-process \( Y \) of \( \mathcal{Y} \) satisfies, \( \mathbb{P}\)–a.s.,

\[
Y_t = u(t, \mathcal{X}_t), \ t \in [0, T],
\]

for a deterministic pricing function \( u \), defined by a continuous Cauchy cascade \((g, \mathcal{U})\) on \( \mathcal{E} \) with \( \mathcal{U} = (u_I)_{1 \leq I \leq N} \).
The next step consists in deriving an analytic characterization of the value function $u$, or, more precisely, of $U = (u_I)_{1 \leq I \leq N}$, in terms of solutions to a related analytic problem. Let $\mathcal{G}$ denote the generator of $X$, so for any function $\phi = \phi(t, x)$, with $a(t, x) = \sigma(t, x) \sigma(t, x)^T$,

$$
\mathcal{G}u(t, x) = \partial_t u(t, x) + \partial u(t, x)b(t, x) + \frac{1}{2} \text{Tr}[a(t, x)\mathcal{H}u(t, x)],
$$

where $\partial u$ and $\mathcal{H}u$ denote the row-gradient and the Hessian of a function $u = u(t, x)$ with respect to $x$.

A technical difficulty comes from the potential discontinuity in $x$ of the functions $u_I$ on $\partial O$ (unless of course one is in the special case where $\kappa_{I,1} = \kappa_{I,-1}$, but this special case is far too restrictive, thinking for instance of the situation of Example 2.6 below).

It would be possible however, though we shall not develop this further here, to characterize $U$ in terms of a suitable notion of discontinuous viscosity solution \cite{10, 11} to the following Cauchy cascade of VIs:

For $l$ decreasing from $N$ to 1,

- At $t = T_I$, for every $k \in \mathcal{K}$ and $x \in \mathbb{R}^q$,

$$
\begin{align*}
 u_I^k(T_I, x) = \begin{cases} 
 \min(u_{I+1}^k(T_I, x, \kappa_I^k(x)), h(T_I, x)), & k \notin K \text{ and } \kappa_I^k(x) \in K \\
 u_{I+1}^k(T_I, x, \kappa_I^k(x)), & \text{else,} 
\end{cases}
\end{align*}
$$

with $u_{I+1}$ in the sense of $g$ in case $I = N$,

- On the time interval $[T_{I-1}, T_I)$, for every $k \in \mathcal{K}$,

$$
\begin{align*}
 \begin{cases} 
 \min \left( -\mathcal{G}u_I^k - f_{u_I^k}, u_I^k - \ell \right) = 0, & k \in K \\
 \max \left( \min \left( -\mathcal{G}u_I^k - f_{u_I^k}, u_I^k - \ell \right), u_I^k - h \right) = 0, & k \notin K 
\end{cases}
\end{align*}
$$

with for any function $\phi = \phi(t, x)$,

$$
f^\phi = f^\phi(t, x) = f(t, x, \phi(t, x)).
$$

It should also be possible to establish related convergence results for standard deterministic (like finite differences) schemes to the viscosity solution of the Cauchy cascade \eqref{eq:cauchy-cascade}–\eqref{eq:cauchy-cascade-d}. Note that \eqref{eq:cauchy-cascade}–\eqref{eq:cauchy-cascade-d} involves Card($\mathcal{K}$) equations in the $u^k$s. From a deterministic computational point of view, the Cauchy cascade \eqref{eq:cauchy-cascade}–\eqref{eq:cauchy-cascade-d} can thus be considered as a $q + d$-dimensional pricing problem, with $d = \log(\text{Card}(\mathcal{K}))$. For ‘very large’ sets $\mathcal{K}$, like for instance in the case of Example 2.6(ii), the use of deterministic schemes is thus precluded by the curse of dimensionality, and simulation schemes are the only viable alternative.

3 Approximation Results

In sections 3.1 to 3.3 we propose an approximation scheme in time for a solution $Y = (Y, Z, A)$, assumed to exist, to $(\mathcal{E})$ (for instance because assumption (H0) holds, see Proposition 2.1), and we provide an upper bound for the convergence rate of this scheme. This convergence rate is the main contribution of this article.

However, for the sake of the numerical validation of the results of the simulation scheme in the follow-up paper \cite{9}, it will be useful to be able to compare the simulation results with
those of an alternative, deterministic numerical scheme.

A deterministic scheme for the Markovian RIBSDE (E) is based on the analytic characterization of (E), or, more precisely, of a related value function \( u \), in terms of an associated system of VIs. This will be dealt with in section 2.4. The proofs are deferred to section 4.

### 3.1 Approximation of the Forward Process

When the diffusion \( X \) in (1) cannot be perfectly simulated, we use the Euler scheme approximation \( \hat{X} \) defined for a grid \( t = \{0 = t_0 < t_1 < \ldots < t_n = T\} \) of \([0,T]\), by \( \hat{X}_0 = X_0 \), and for \( i \leq n-1 \),

\[
\hat{X}_{i+1} = \hat{X}_i + b(t_i, \hat{X}_i)(t_{i+1} - t_i) + \sigma(t_i, \hat{X}_i)(W_{i+1} - W_i) .
\]

We assume \( n|t|_+ \leq \Lambda \), where \( |t|_+ = \max_{i \leq n-1} (t_{i+1} - t_i) \).

As usual, we define a continuous-time extension of \( \hat{X} \) by setting, for every \( i \leq n-1 \) and \( t \in [t_i, t_{i+1}) \),

\[
\hat{X}_t = \hat{X}_{t_i} + b(t_i, \hat{X}_{t_i})(t - t_i) + \sigma(t_i, \hat{X}_{t_i})(W_t - W_{t_i}) ,
\]

or in an equivalent differential notation, for \( t \in [0,T] \),

\[
d\hat{X}_t = b(\bar{t}, \hat{X}_\bar{t})dt + \sigma(\bar{t}, \hat{X}_\bar{t})dW_t ,
\]

where we set \( \bar{t} = \sup\{s \in t|s \leq t\} \).

Under the Lipschitz continuity assumption (Hx), one has, for every \( p \geq 1 \) (see e.g. Kloeden and Platen [17]),

\[
\| \sup_{t \leq T} |X_t - \hat{X}_t| \|_{L^p} + \max_{i \leq n} \| \sup_{t \in [t_i, t_{i+1})} |X_t - \hat{X}_{t_i}| \|_{L^p} \leq C^p_\Lambda |t|^{\frac{1}{2}} .
\]

### 3.2 Approximation of the Upper Barrier

The lower barrier is simply approximated by \( \ell(t, \hat{X}_t) \). As for the upper barrier, we first need to define the approximation of the marker process \( H \), denoted by \( \hat{H} \), which will be given by \( \hat{H}_0 = H_0 \) and \( \hat{H}_{T_i} = \kappa_I(\hat{X}_{T_i}, \hat{H}_{T_i}) \), for \( 1 \leq I \leq N \).

We then define the approximation \( \hat{\vartheta} \) of \( \vartheta \) as the sequence of \( \mathcal{I} \)-valued stopping times obtained by using \( \hat{\mathcal{X}} = (\hat{X}, \hat{H}) \) instead of \( \mathcal{X} \) in (1).

In order to control the error between the call protection switching times \( \vartheta \) and their approximation, we make the following assumption on the coefficients of \( X \), (Hxo) \( \sigma \) and \( b \) are bounded and of class \( C^2_b \) on the following set,

\[
\mathcal{Q} = \{(t, x) \in \cup_{1 \leq I \leq N-1} [T^\Lambda_{T_i}, T_i] \times \mathbb{R}^q ; |d(x)| \leq \frac{1}{\Lambda}\},
\]

where we set, for \( 1 \leq I \leq N-1 \),

\[
T^\Lambda_I = T_I - \frac{1}{\Lambda} > T_{I-1} .
\]
Moreover, for every \((t, x) \in \mathcal{Q}\),
\[
a(t, x) := (\sigma \sigma')(t, x) \geq \frac{1}{\Lambda} I_q .
\] (29)

**Proposition 3.1** Under \((H_{\epsilon})\), for every \(\varepsilon > 0\), there exists a constant \(C_\Lambda^\varepsilon\) such that for every \(l \leq N + 1\),
\[
\mathbb{E} \left[ |\vartheta_l - \tilde{\vartheta}_l| \right] \leq C_\Lambda^\varepsilon |t|^{\frac{1}{2} - \varepsilon} .
\]

### 3.3 Approximation of the RIBSDE

In the sequel, we shall use one of the following regularity assumptions:

- \((\text{Hb})\) \(h\) and \(\ell\) are \(\Lambda\)-Lipschitz continuous with respect to \((t, x)\).
- \((\text{Hb}')\) There exists a constant \(\Lambda\) and some functions \(\Lambda_1, \Lambda_2 : \mathbb{R}^q \to \mathbb{R}^{1 \otimes q}\) and \(\Lambda_3 : \mathbb{R}^q \to \mathbb{R}^+\) such that \(|\Lambda_1(x)| + |\Lambda_2(x)| + |\Lambda_3(x)| \leq \Lambda(1 + |x|^\Lambda)\), and for every \(x, y \in \mathbb{R}^q\),
\[
\ell(t, x) - \ell(t, y) \leq \Lambda_1(x)(y - x) + \Lambda_3(x)|x - y|^2 \\
h(t, y) - h(t, x) \leq \Lambda_2(x)(y - x) + \Lambda_3(x)|x - y|^2 .
\]

**Remark 3.1** Assumption \((\text{Hb}')\), which implies \((\text{Hb})\), is slightly weaker than the semi-convexity assumption of Definition 1 in Bally–Pagès [1].

Given \(\varrho = \vartheta\) or \(\hat{\vartheta}\), let the projection operator \(\mathcal{P}_\varrho\) be defined by
\[
\mathcal{P}_\varrho(t, x, y) = y + [\ell(t, x) - y]^+ - [y - h(t, x)]^+ \sum_{i=1}^{[(N+1)/2]} 1_{\{\varrho_{2i-1} \leq t \leq \varrho_{2i}\}} .
\] (30)

To tackle the reflection issue, we introduce a discrete set of reflection times defined by
\[
\tau = \{0 = r_0 < r_1 < \cdots < r_\nu = T\} \text{ with } \mathcal{F} \subseteq \tau \subseteq t \text{ and } |\tau|_+ \leq C_\Lambda |\tau|_-. \]

The idea is that, in the approximation scheme for \(\mathcal{Y}\), the reflection will operate only on \(\tau\). The components \(Y\) and \(Z\) of a solution \(\mathcal{Y} = (Y, Z, A)\) to the RIBSDE (\(\mathcal{E}\)) are thus approximated by a triplet of processes \((\hat{Y}, \bar{Y}, \bar{Z})\) defined on \(t\) by the terminal condition
\[
\hat{Y}_T = \bar{Y}_T = g(\hat{X}_T) ,
\]
and then for \(i\) decreasing from \(n - 1\) to 0,
\[
\begin{align*}
\hat{Z}_{t_i} & = \frac{1}{t_{i+1} - t_i} \mathbb{E} \left[ \hat{Y}_{t_{i+1}} (W_{t_{i+1}} - W_{t_i})' \mid \mathcal{F}_{t_i} \right] \\
\bar{Y}_{t_i} & = \mathbb{E} \left[ \hat{Y}_{t_{i+1}} \mid \mathcal{F}_{t_i} \right] + (t_{i+1} - t_i) f(t_i, \hat{X}_{t_i}, \hat{Y}_{t_i}, \bar{Z}_{t_i}) \\
\bar{Y}_{t_i} & = \hat{Y}_{t_i} 1_{\{t_i \notin \tau\}} + \mathcal{P}_\varrho(t_i, \hat{X}_{t_i}, \hat{Y}_{t_i}) 1_{\{t_i \in \tau\}} .
\end{align*}
\] (32)

By convention, we set \(\bar{Z}_T = 0\).
Using an induction argument and the Lipschitz-continuity assumption on $f, g, l, h,$ one easily checks that the above processes are square integrable. It follows that the conditional expectations are well defined at each step of the algorithm.

We also consider a piecewise time-continuous extension of the scheme. Using the martingale representation theorem, we define $\tilde{Z}$ on $[t_i, t_{i+1})$ by

$$\tilde{Y}_{t_{i+1}} = \mathbb{E}_{t_i}[\tilde{Y}_{t_{i+1}}] + \int_{t_i}^{t_{i+1}} \tilde{Z}_s dW_s.$$ 

We then define $\tilde{Y}$ on $[t_i, t_{i+1})$ by

$$\tilde{Y}_t = \tilde{Y}_{t_{i+1}} + (t_{i+1} - t) f(t_i, \tilde{X}_{t_i}, \tilde{Y}_{t_i}, \tilde{Z}_{t_i}) - \int_t^{t_{i+1}} \tilde{Z}_s dW_s$$

and we let finally, for $t \in [0, T],$

$$\tilde{Y}_t = \tilde{Y}_i \mathbf{1}_{\{t \leq t_i\}} + \mathcal{P}_{\tilde{Y}_i}(t, \tilde{X}_t, \tilde{Y}_t) \mathbf{1}_{\{t \in [t_i, t_{i+1})\}}.$$ (33)

Observe that one has, for $i \leq n - 1,$

$$\tilde{Z}_{t_i} = \frac{1}{t_{i+1} - t_i} \mathbb{E}_{t_i}\left[\int_{t_i}^{t_{i+1}} \tilde{Z}_s ds\right].$$

Under stronger assumption on the boundaries and on the regularity of the coefficients $b, \sigma,$ it is possible to obtain a better control of the convergence rate of the approximation, see Theorem 6.2 in [8] and Theorem 4.1 in [5].

**Theorem 3.2 (See [5, 8])** Set $\alpha = \frac{1}{3}$ under $(Hb), \alpha = \frac{1}{2}$ under $(Hb)^\prime,$ then the following holds

$$\max_{i \leq n-1} \sup_{t \in [t_i, t_{i+1})} \mathbb{E}\left[Y_t - \tilde{Y}_{t_i}\right]^2 + \max_{i \leq n-1} \sup_{t \in [t_i, t_{i+1})} \mathbb{E}\left[Y_t - \tilde{Y}_{t_i}\right]^2 \leq C_\Lambda |t|^\alpha.$$ 

Under stronger assumption on the boundaries and on the regularity of the coefficients $b, \sigma,$ it is possible to obtain a better control of the convergence rate of the approximation, see Theorem 6.2 in [8] and Theorem 4.1 in [5].

Regarding call protection, our main result is the following

**Theorem 3.3** We assume that $f$ does not depend on $z.$ Set $\alpha = \frac{1}{2}$ under $(Hb), \alpha = 1$ under $(Hb)^\prime,$ then the following holds

$$\max_{i \leq n-1} \sup_{t \in [t_i, t_{i+1})} \mathbb{E}\left[Y_t - \tilde{Y}_{t_i}\right]^2 + \max_{i \leq n-1} \sup_{t \in [t_i, t_{i+1})} \mathbb{E}\left[Y_t - \tilde{Y}_{t_i}\right]^2 \leq C_\Lambda |t|^\frac{\alpha}{2 - \varepsilon},$$

for $\varepsilon > 0.$

**Remark 3.2** In the ‘no call’ or ‘no call protection’ cases, convergence bounds are also available for the $Z$ component, see Theorem 6.1 in [8] and Theorem 4.1 in [5]. The ‘call protection’ case is currently under research in this regard, as more generally in regard to establishing convergence bounds on $Y$ and $Z$ in case $f$ depends on $z.$
3.3.1 Discretely reflected BSDEs

As in [8, 5], the study of the convergence of the time-discretization scheme for $\mathcal{Y}$ will be done in several steps, using a suitable concept of \textit{discretely reflected BSDEs}.

Given the grid $\tau$, recalling [31], and $\rho = \vartheta$ or $\tilde{\vartheta}$, the solution of the \textit{discretely reflected BSDE} is a triplet $(\mathcal{Y}, \mathcal{\tilde{Y}}, \mathcal{\bar{Y}})$ defined by the terminal condition

$$\mathcal{Y}_T = \mathcal{\tilde{Y}}_T = g(X_T),$$

and then (cf. [31]) for $\nu$ decreasing from $\nu - 1$ to 0 and $t \in [\tau_1, \tau_{i+1})$,

$$\begin{cases} \mathcal{\tilde{Y}}_t = \mathcal{Y}_{\tau_{i+1}} + \int_{\tau_i}^{\tau_{i+1}} f(X_u, \mathcal{\tilde{Y}}_u, \mathcal{\bar{Y}}_u)\,du - \int_{\tau_i}^{\tau_{i+1}} \mathcal{\bar{Y}}_u\,dW_u, \\ \mathcal{\bar{Y}}_t = \mathcal{\bar{Y}}_{\tau_i}1_{\{t \in \tau\}} + \mathcal{\bar{P}}_\vartheta(t, X_t, \mathcal{\tilde{Y}}_t)1_{\{t \in \tau\}}. \end{cases} \tag{34}$$

In finance, they represent game option which can be exercised only on the discrete set of times $\tau$.

Under (Hx)-(Hb), such a solution can be defined by backward induction. At each step, existence and uniqueness of a solution in $\mathcal{S}^2 \times \mathcal{H}^2_\vartheta$ follow from [15].

\textbf{Remark 3.3} (i) $\mathcal{\tilde{Y}}$ is a càdlàg process whereas $\mathcal{Y}$ is a \textit{càglàd} process. By convention, we set $Y_{\tau_-} = Y_0$.

(ii) One has, for $r \in \tau$,

$$Y_{r-} = \mathcal{P}_\vartheta(r, X_r, Y_r), \quad \mathcal{\tilde{Y}}_r = \mathcal{\bar{P}}_\vartheta(r, X_r, \mathcal{\tilde{Y}}_r). \tag{35}$$

We first present two properties of discretely reflected BSDEs which are useful to prove Theorem 3.3. We show that under suitable conditions the discretely reflected BSDE with $\rho = \vartheta$ is a ‘good’ approximation of the RIBSDE ($\mathcal{E}$). In view of Definition 2.3(i), the component $Y$ of $\mathcal{Y}$ may be discontinuous at $\vartheta_2$. This discontinuity is problematic and the fact that $\mathcal{\tilde{Y}} \subseteq \mathcal{\bar{Y}}$ will be essential to obtain the following result.

\textbf{Proposition 3.4} Let $\alpha = \frac{1}{2}$ or $\alpha = 1$ under (Hb) or (Hb)$\dagger$, respectively. If $f$ does not depend on $z$, then

$$\sup_{t \in [0, T]} \mathbb{E}\left[|Y_t - \mathcal{\tilde{Y}}_t|^2\right] + \sup_{t \in [0, T]} \mathbb{E}\left[|Y_{t-} - \mathcal{Y}_{t-}|^2\right] + \mathbb{E}\left[\int_0^T |Z_s - \mathcal{\bar{Y}}_s|^2\,ds\right] \leq |\nu|^\alpha.$$

We also give a control of the difference between the solutions $(\mathcal{Y}, \mathcal{\tilde{Y}}, \mathcal{\bar{Y}})$ and $(\mathcal{\bar{Y}}, \mathcal{\tilde{Y}}, \mathcal{\bar{Y}})$ of the two discretely reflected BSDEs with $\rho = \vartheta$ and $\tilde{\vartheta}$.

\textbf{Proposition 3.5} Let $\alpha = \frac{1}{2}$ or $\alpha = 1$ under (Hb) or (Hb)$\dagger$, respectively. If $f$ does not depend on $z$, the following holds

$$\sup_{t \in [0, T]} \mathbb{E}\left[|\mathcal{Y}_t - \mathcal{\bar{Y}}_t|^2\right] + \sup_{t \in [0, T]} \mathbb{E}\left[|\mathcal{\tilde{Y}}_t - \mathcal{\bar{Y}}_t|^2\right] + \|\mathcal{\bar{Y}}_t - \mathcal{\tilde{Y}}_t\|^2_{\mathcal{L}^{\rho}} \leq C^\varepsilon |\nu|^\alpha \sum_{t=1}^N \left(\mathbb{E}\left[|\vartheta_t - \hat{\vartheta}_t|\right]\right)^{1-\varepsilon},$$

for $\varepsilon > 0$. 

\[ \]
We conclude this section by giving a bound for the convergence rate of the scheme \([32]\) to the discretely reflected BSDE \([34]\), with \(\varphi = \vartheta\).

**Theorem 3.6** Let \(\alpha = \frac{1}{2}\) or \(\alpha = 1\) under \((Hb)\) or \((Hb)'\), respectively. If \(f\) does not depend on \(z\), the following holds

\[
\sup_{t \in [0,T]} \mathbb{E}\left[|\tilde{Y}_t^\varphi - \tilde{Y}_t^\vartheta|^2\right] + \sup_{t \in [0,T]} \mathbb{E}\left[|Y_t^\varphi - Y_t^\vartheta|^2\right] \leq C_\Lambda |t| + C_\Lambda^\varphi |t|^{\alpha - 1} \sum_{i=1}^N \left(\mathbb{E}\left[|\vartheta_i - \tilde{\vartheta}_i|\right]\right)^{1-\varepsilon},
\]

for \(\varepsilon > 0\).

4 **Proofs**

We denote by \(\chi\) a positive random variable which may change from line to line but satisfies \(\mathbb{E}[\chi^p] \leq C_\chi^p\).

4.1 **Stability of Call Protection Switching Times**

In the following, we consider two diffusions. The first one, \(X\), starts at \((t, x)\) and is the solution of the following SDE:

\[
X_s = x + \int_t^s b(s, X_s)ds + \int_t^s \sigma(s, X_s)dW_s, \quad \text{for } s \in [t, T].
\]

The second one, \(X'\), starts at \((t', x')\) and can be written:

\[
X'_s = x' + \int_{t'}^s b'_s ds + \int_{t'}^s \sigma'_s dW_s, \quad \text{for } s \in [t', T].
\]

**Remark 4.1** In practice, \(X'\) will be either the solution of a Markovian SDE with coefficient \((b, \sigma)\) and the initial condition \((t', x')\), or the continuous-time Euler scheme for the SDE \(X\) with initial condition \((t = 0, x)\).

We consider the following ‘monitoring grid’ for \(X\), resp. \(X'\):

\[
\mathcal{T}^t = \{s \in \mathcal{I} | s > t\}, \quad \text{resp. } \mathcal{T}'^t = \{s \in \mathcal{I} | s > t'\},
\]

and we let \(T^t = \inf \mathcal{T}^t\), resp. \(T'^t = \inf \mathcal{T}'^t\).

Given an initial condition \(k \in K\), let also \(X' = (X', H')\), for the market process \(H'\) defined by \(H'_0 = k\), and for every \(T_l \in \mathcal{T}^t\),

\[
H'_{T_l} = \kappa_l(X'_T, H'_{T_l})
\]

\((H'\) being constant between two dates of \(\{t'\} \cup \mathcal{T}'^t\)). Observe that \(H'\) does not jump at \(t'\).
We also consider a non-decreasing sequence of stopping times \( \vartheta' = (\vartheta'_l)_{l \in \mathbb{N}+1} \), representing call protection switching times, defined by
\[
\vartheta'_0 = t' \quad \text{and} \quad \vartheta'_{2l+1} = \inf\{t > \vartheta'_{2l}; H'_t \notin K\} \land T, \quad \vartheta'_{2l+2} = \inf\{t > \vartheta'_{2l+1}; H'_t \in K\} \land T. \tag{36}
\]
The \( \vartheta'_l \)s this effectively reduce to \( \{t'\} \cup \mathcal{T}' \)-valued stopping times, and one has \( \vartheta_{\mathbb{N}+1} = T \).

To the process \( X \), we associate two different extended factor processes \( X \) and \( \tilde{X} \). The first one, \( \mathcal{X} = (X, H) \), is defined as above, replacing \( X' \) by \( X \). Observe that \( H \) does not jump at \( t \) and that \( H_t = H'_t = k \). We also consider the sequence of call protection monitoring times \( \vartheta \), defined as in \( (36) \) with \( t \) and \( H \) instead of \( t' \) and \( H' \).

The second factor process, \( \tilde{X} = (X, \tilde{H}) \), is given by
\[
\tilde{H}_{T_l} = \kappa_l(T_l, H_{T_l})
\]
(and \( \tilde{H} \) is constant between two dates of \( \{t\} \cup \mathcal{T}' \)). Observe that, contrary to \( H \), \( \tilde{H} \) may jump at \( t \).

We also consider the corresponding call protection switching times \( \tilde{\vartheta} \) defined as in \( (36) \) with \( t \) and \( \tilde{H} \) instead of \( t' \) and \( H' \).

We are then interested in two different cases regarding the initial set of data \((t, x)\) and \((t', x')\).

**Case 1:** \( T^t = T'^t \).

**Case 2:** \( T'^t = t \) and \( x \notin \partial \mathcal{O} \).

Let us finally introduce, for every \( 0 < h < |\mathcal{T}'| \), \( T_l \in \mathcal{T}' \) and \( \delta > 0 \), the sets
\[
\Omega^l_h = \left\{ \sup_{T_l - h \leq u \leq T_l} |X_u - X_T| \leq \frac{1}{3\Lambda} \right\}, \quad \Omega^l_h = \bigcap_{T_l \in \mathcal{T}'} \Omega^l_h
\]
\[
\tilde{\Omega}^\delta = \left\{ \sup_{u \in [T', T]} |X'_u - X_u| < \delta \right\}.
\]

The proof of the following Lemma is deferred to Appendix A.

**Lemma 4.1** Assume \((Hxo)\).

(i) One has, for \( T_l \in \mathcal{T}' \),
\[
\mathbb{P}(\Omega^l_h \cap \{|d(X_{T_l})| \leq \delta\}) \leq C\Lambda \frac{\delta}{h}. \tag{37}
\]

(ii) For \( p, \rho > 0 \), and \( l \in \mathbb{N}_{N+1} \), one has,
\[
\mathbb{E}[|q_l - \vartheta'_l| \leq |t - t'| + C\Lambda \frac{\delta}{h} + C\Lambda \frac{\delta^\rho}{h} + C\Lambda \frac{\mathbb{E}[\sup_{u \in [T', T]} |X'_u - X_u|^p]}{\delta^p},
\]
with \( q = \vartheta \) in Case 1 and \( q = \tilde{\vartheta} \) in Case 2.
4.1.1 Proof of Proposition 2.2

Let in this section $X' = X^{t_n}$, for $(t_n, x_n, k) \in \mathcal{E}$ (cf. Remark 4.1).

(i) When $t_n \downarrow t$, we want to control the difference between $\vartheta^t$ and $\vartheta^{t_n}$ to prove the càd property. We shall use here the result of Case 1. First we know that

$$\mathbb{E}\left[ \sup_{u \in [0,T]} |X^{t_n}_u - X^t_u|^p \right] \leq C^p_\Lambda (|x - x_n|^p + |t - t_n|^p).$$

We then obtain, applying Lemma 4.1(ii), that

$$\mathbb{E}[\vartheta^t_t - \vartheta^{t_n}_t] \leq |t - t_n| + C^p_\Lambda \frac{\delta_{n}}{h_n} + C^p_\Lambda h_n \frac{|x - x_n|^p + |t - t_n|^p}{\delta_{n}^2}.$$

The proof is concluded by taking $\delta_{n}^2 = |x - x_n| \vee |t - t_n|^\frac{1}{2}$, $h_n = \delta_{n}$, $\rho = p = 2$ and letting $n$ go to $\infty$.

(ii) When $t_n \uparrow t$, we want to control the difference between $\tilde{\vartheta}^t$ and $\tilde{\vartheta}^{t_n}$ to prove the làg property, assuming $x \notin \partial \mathcal{O}$. Since $x_n \to x$, we have for some $n \geq 0$ that $x_n \notin \mathcal{O}$. We then argue as in (i), using this time the result of Case 2 in Lemma 4.1(i).

4.1.2 Proof of Proposition 3.1

Let in this section $\xi = \tilde{X}$ (cf. Remark 4.1). We have here that $t = t' = 0$, so we are in Case 1. Applying Lemma 4.1(ii), we thus get, in view of (28),

$$\mathbb{E}\left[ \tilde{\vartheta}_t - \tilde{\vartheta}^{t_n}_t \right] \leq C^p_\Lambda \frac{\delta_{n}}{h_n} + C^p_\Lambda h_n \frac{|X - X_n|^p + |t - t_n|^p}{\delta_{n}^2}.$$

The proof is concluded by setting $\delta = |t|^{\frac{1}{2}} - \frac{\varepsilon}{2}$, $h = \frac{\varepsilon}{2}$, $\rho = p = \frac{1}{2} - 2$, for $\varepsilon$ and $|t|$ small enough.

4.2 Proof of the BSDE Results

4.2.1 Proof of Proposition 3.4

Let for $t \leq T$,

$$\delta \tilde{Y}_t = Y_t - \tilde{Y}^0_t, \delta Y_t = Y_{t-} - \tilde{Y}^0_{t-}, \delta Z_t = Z_{t-} - \tilde{Z}^0_{t-}, \delta f_t = f(t, X_t, Y_t) - f(t, X_t, \tilde{Y}^0_t).$$

Observe that $\delta \tilde{Y}$ is continuous outside $\tau$ and that $\delta \tilde{Y}_{t-} = \delta Y_t$ for $t \in (0, T]$, so that one has by (35), for $r \in \tau$,

$$|\delta Y_r| = |Y_{r-} - \tilde{Y}^0_{r-}| \leq |\delta \tilde{Y}_r|. \tag{38}$$

Applying Itô’s formula to the càdlàg process $|\delta \tilde{Y}|^2$ and observing that the local martingale term is in fact a martingale, we compute,

$$\mathbb{E}_{\tau} \left[ |\delta \tilde{Y}|^2 + \int_t^{r_n} |\delta Z_u|^2 du \right] = \mathbb{E}_{\tau} \left[ |\delta \tilde{Y}_{r_n} - 1|^2 + 2 \int_t^{r_n} \delta \tilde{Y}_s \delta f_s ds + 2 \int_{(t,r_n)} \delta \tilde{Y}_s dA_s \right].$$
for \( t \in [r_i, r_{i+1}) \). Given (38), one thus gets by usual arguments, for \( t \in [r_i, r_{i+1}) \),

\[
\mathbb{E}_r \left[ |\delta \bar{Y}_t|^2 + \int_t^{r_{i+1}} |\delta Z_s|^2 ds \right] \leq (1 + C_\Lambda |r|) \mathbb{E}_r \left[ |\delta \bar{Y}_{r_{i+1}}|^2 + 2 \int_{(t,r_{i+1})} \delta \bar{Y}_s dA_s^+ - 2 \int_{(t,r_{i+1})} \delta \bar{Y}_s dA_s^- \right].
\]

We first study the term related to the upper barrier. One has,

\[
-\mathbb{E}_r \left[ \int_{(t,r_{i+1})} \delta \bar{Y}_s dA_s^- \right] = \mathbb{E}_r \left[ \int_{(t,r_{i+1})} (\tilde{\gamma}_s^\theta - h(s, X_s)) dA_s^- \right] = \mathbb{E}_r \left[ \int_{(t,r_{i+1})} (\tilde{\gamma}_{r_{i+1}}^\theta - h(s, X_s)) dA_s^- + \int_{(t,r_{i+1})} \int_{s}^{r_{i+1}} f(u, X_u, \tilde{\Theta}_u^\theta) du dA_s^- \right]
\]

where in particular the upper barrier minimality condition in (E) was used in the first identity. The second term is bounded by

\[
\mathbb{E}_r \left[ \chi |r| (A_{r_{i+1}}^- - A_{r_i}^-) \right] \leq \mathbb{E}_r \left[ \chi |r| (A_{r_{i+1}}^- - A_{r_i}^-) \right],
\]

since \( f \) does not depend on \( z \) and \( A^- \) is increasing. For the first term, we use the fact that \( dA^- 1_{[\theta_{2l}, \theta_{2l+1}]} = 0, 0 \leq l \leq [(N + 1)/2] \), to obtain that

\[
\mathbb{E}_r \left[ \int_{(t,r_{i+1})} (\tilde{\gamma}_{r_{i+1}}^\theta - h(s, X_s)) dA_s^- \right] = \mathbb{E}_r \left[ \sum_{l=1}^{[(N+1)/2]} \int_{(t,r_{i+1})} (\tilde{\gamma}_{r_{i+1}}^\theta - h(s, X_s)) 1_{\{\theta_{2l-1} \leq s \leq \theta_{2l}\}} dA_s^- \right] \leq \mathbb{E}_r \left[ \int_{(t,r_{i+1})} (h(r_{i+1}, X_{r_{i+1}}) - h(s, X_s)) 1_{\{\theta_{2l-1} \leq s \leq \theta_{2l}\}} dA_s^- \right] \leq \mathbb{E}_r \left[ \int_{(t,r_{i+1})} (h(r_{i+1}, X_{r_{i+1}}) - h(s, X_s)) dA_s^- \right].
\]

The proof is then concluded using the same argument as in the proof of Propositions 2.6.1 and 1.4.1 in [7].

### 4.2.2 Proof of Proposition 3.5

Let, for \( t \leq T \),

\[
\delta \tilde{\mathcal{G}}_t = \tilde{\mathcal{G}}_t^\theta - \tilde{\mathcal{G}}_t^\phi, \ \delta \mathcal{G}_t = \mathcal{G}_t^\theta - \mathcal{G}_t^\phi, \ \delta \mathcal{A}_t = \mathcal{A}_t^\theta - \mathcal{A}_t^\phi, \ \eta_t = |\delta \mathcal{G}_t|^2 - |\delta \tilde{\mathcal{G}}_t|^2, \ \delta f_t = f(t, X_t, \tilde{\mathcal{G}}_t^\phi) - f(t, X_t, \tilde{\mathcal{G}}_t^\phi).
\]

**Step 1** Applying Itô’s formula to the càdlàg process \( |\delta \tilde{\mathcal{G}}|^2 \), we compute for \( t \in [r_i, r_{i+1}) \)

\[
\mathbb{E}_r \left[ |\delta \tilde{\mathcal{G}}_t|^2 + \int_t^{r_{i+1}} |\delta \mathcal{A}_u|^2 du \right] = \mathbb{E}_r \left[ |\delta \tilde{\mathcal{G}}_{r_{i+1}}|^2 + \eta_{r_{i+1}} + 2 \int_t^{r_{i+1}} \delta \tilde{\mathcal{G}}_s \delta f_s ds \right].
\]

Usual arguments then yield that

\[
\sup_{s \in [t,T]} \mathbb{E} \left[ |\delta \tilde{\mathcal{G}}_s|^2 + |\delta \mathcal{A}_s|^2 + \int_s^T |\delta \mathcal{A}_u|^2 du \right] \leq C_\Lambda \mathbb{E} \left[ \sum_{\tau \in \mathcal{F}} \eta_{\tau} \right], \tag{39}
\]
recalling \( |\delta \mathcal{J}_s|^2 = \eta_s + |\delta \tilde{\mathcal{J}}_s|^2 \).

**Step 2** In order to study the right-hand side term of (39), we introduce the processes defined by, for \( r \in [0, T] \),

\[
I_r = \sum_{l=1}^{[(N+1)/2]} \mathbf{1}\{\vartheta_{2l-1} \leq r < \vartheta_{2l}\}, \quad \tilde{I}_r = \sum_{l=1}^{[(N+1)/2]} \mathbf{1}\{\vartheta_{2l-1} \leq r < \vartheta_{2l}\}, \quad \psi_r = 1 - I_r, \quad \tilde{\psi}_r = 1 - \tilde{I}_r.
\]

Observe that \( I = 1 \) (or \( \tilde{I} = 1 \)) means that the upper barrier is activated for reflection.

We thus compute, for \( r \in \tau \),

\[
\eta_r \leq \mathbb{E}_r[\mathcal{J}] (|h(r, X_r) - \tilde{\vartheta}_r|^+ I_r \tilde{I}_r + |h(r, X_r) - \tilde{\vartheta}_r|^+ \tilde{I}_r \psi_r).
\]

The two terms at the right-hand side of (43) are treated similarly, we thus concentrate on the first one.

**Step 3** We have to take into account the fact that a reflection date may be a deactivation date for the upper boundary, i.e., for \( r \in \tau \),

\[
\mathbb{E}_r[\mathcal{J}] (\tilde{\vartheta}_r - h(r, X_r))^+ I_r \tilde{I}_r = \mathbb{E}_r[\mathcal{J}] (\tilde{\vartheta}_r - h(r, X_r))^+ (\sum_{l=1}^{[(N+1)/2]} \mathbf{1}_{\{r = \vartheta_{2l}\}} + \sum_{l=1}^{[(N+1)/2]} \mathbf{1}_{\{r = \vartheta_{2l-1} \leq r < \vartheta_{2l}\}})
\]

(44)

**Step 3a** We study the first term in the right hand side of (44).

We obviously have that \( \mathbb{E}_r[\mathcal{J}] (|h(r, X_r) - \tilde{\vartheta}_r|^+ \leq \mathbb{E}_r[\mathcal{J}] \) for since the \( \vartheta_l \)s are \( \Xi \)-valued stopping-times,

\[
\sum_{r \in \tau} \mathbb{E}_r[\mathcal{J}] (\tilde{\vartheta}_r - h(r, X_r))^+ \tilde{I}_r \sum_{l=1}^{[(N+1)/2]} \mathbf{1}_{\{r = \vartheta_{2l}\}} \leq \sum_{r \in \tau} \mathbb{E}_r[\mathcal{J}]^2 \tilde{I}_r \sum_{l=1}^{[(N+1)/2]} \mathbf{1}_{\{r = \vartheta_{2l}, r \neq \vartheta_{2l}\}}
\]

Moreover, by definition of \( I \) and \( \tilde{I} \),

\[
\sum_{r \in \tau} \mathbb{E}_r[\mathcal{J}]^2 \tilde{I}_r \sum_{l=1}^{[(N+1)/2]} \mathbf{1}_{\{r = \vartheta_{2l}\}} = \sum_{r \in \tau} \sup_{r \in \Xi} \mathbb{E}_r[\mathcal{J}]^2 \tilde{I}_r \sum_{l=1}^{[(N+1)/2]} \mathbf{1}_{\{r = \vartheta_{2l}, r \neq \vartheta_{2l}\}} \leq \sup_{r \in \tau} \mathbb{E}_r[\mathcal{J}]^2 \sum_{r \in \Xi} \sum_{l=1}^{[(N+1)/2]} \mathbf{1}_{\{|\vartheta_{2l} - \tilde{\vartheta}_{2l}| \geq |\Xi| \}}.
\]

Using the Cauchy-Schwartz inequality with \( \frac{1}{p} = 1 - \varepsilon \), Doob’s inequality and the Markov inequality, we obtain

\[
\mathbb{E} \left[ \sum_{r \in \Xi} \sum_{l=1}^{[(N+1)/2]} \sup_{r \in \Xi} \mathbb{E}_r[\mathcal{J}]^2 \mathbf{1}_{\{|\vartheta_{2l} - \tilde{\vartheta}_{2l}| \geq |\Xi| \}} \right] \leq C_A^{\varepsilon} \sum_{l=1}^{[(N+1)/2]} \mathbb{E}[|\vartheta_{2l} - \tilde{\vartheta}_{2l}|^{1-\varepsilon}].
\]
Step 3b We now study the last term in the right hand side of (44). On the event \( \{\vartheta_{2l-1} \leq r < \vartheta_{2l}\} \), which is \( \mathcal{F}_r \)-measurable, the upper barrier is active on \( [\vartheta_{2l-1}, \vartheta_{2l}] \), thus

\[
\tilde{Y}_r^\theta - h(r, X_r) \leq \mathbb{E}_r \left[ h(r^+, X_{r^+}) - h(r, X_r) + \int_r^{r^+} |f(s, X_s, Y_s^\theta)|ds \right]
\]

where we set \( r^+ = \inf\{s \in \mathbb{R} | s > r\} \land T \). One thus gets, using (Hb) or (Hb)',

\[
[\tilde{Y}_r^\theta - h(r, X_r)]^+ 1_{\{\vartheta_{2l-1} \leq r < \vartheta_{2l}\}} \leq \mathbb{E}_r[|\tau|]\nu^\alpha .
\]

This leads to

\[
\sum_{r \in \mathbb{R}} \mathbb{E}_r[|\tau|]^{\alpha} \sum_{l=1}^{[N+1/2]} c_t \tilde{r} \sum_{l=1}^{[N+1/2]} 1_{\{\vartheta_{2l-1} \leq r < \vartheta_{2l}\}} \leq |\tau|^\alpha \mathbb{E}_r[|\tau|]^{\alpha} \sum_{r \in \mathbb{R}} \sum_{l=1}^{[N+1/2]} c_t \tilde{r} \sum_{l=1}^{[N+1/2]} 1_{\{\vartheta_{2l-1} \leq r < \vartheta_{2l}\}} .
\]

Moreover,

\[
\sum_{r \in \mathbb{R}} \sum_{l=1}^{[N+1/2]} c_t \tilde{r} 1_{\{\vartheta_{2l-1} \leq r < \vartheta_{2l}\}} \leq \sum_{r \in \mathbb{R}} \sum_{l=1}^{[N+1/2]} 1_{\{\vartheta_{2l-1} \leq r < \vartheta_{2l}\}} (1_{\{\tilde{\vartheta}_{2l-1} > r\}} + 1_{\{r < \tilde{\vartheta}_{2l}\}})
\]

We obtain combining the last inequality with (47) and using the Cauchy-Schwartz inequality with \( \frac{1}{\rho} = 1 - \varepsilon \), Doob’s inequality and the Markov inequality

\[
\mathbb{E}_r \left[ \sum_{r \in \mathbb{R}} \mathbb{E}_r[|\tau|]^{\alpha} \tilde{Y}_r^\theta - h(r, X_r) + c_t \tilde{r} \sum_{l=1}^{[N+1/2]} 1_{\{\vartheta_{2l-1} \leq r < \vartheta_{2l}\}} \right] \leq |\tau|^{\alpha-1} C^\varepsilon \sum_{l=1}^{[N+1/2]} \mathbb{E}_r \left[ |\vartheta_{2l} - \tilde{\vartheta}_{2l}|^{1-\varepsilon} \right] .
\]

Step 4 The proof is concluded by combining (39) with (44), (45) and (48).

4.2.3 Proof of Theorem 3.6

Since

\[
|\mathcal{Y}_t^\theta - \tilde{Y}_t|^2 \leq C_{\lambda} (|\mathcal{Y}_t^\theta - \tilde{\mathcal{Y}}_t|^2 + |\tilde{\mathcal{Y}}_t^\theta - \tilde{\tilde{Y}}_t|^2) \text{ and } |\tilde{\mathcal{Y}}_t^\theta - \tilde{Y}_t|^2 \leq C_{\lambda} (|\tilde{\mathcal{Y}}_t^\theta - \tilde{\mathcal{Y}}_t|^2 + |\tilde{\mathcal{Y}}_t^\theta - \tilde{\tilde{Y}}_t|^2) ,
\]

it remains to study the error between \( (\tilde{\mathcal{Y}}^\theta, \mathcal{Y}^\theta, \mathcal{Z}^\theta) \) and the continuous-time Euler scheme \( (\tilde{Y}, \tilde{Y}, \mathcal{Z}) \). We are thus going to show that

\[
\sup_{t \in [0, T]} \mathbb{E}[|\tilde{\mathcal{Y}}_t^\theta - \tilde{Y}_t|^2] + \sup_{t \in [0, T]} \mathbb{E}[|\mathcal{Y}_t^\theta - \tilde{Y}_t|^2] \leq C_{\lambda}|t| .
\]
Towards this end, arguing as in the proof of Lemma 2.1 in [8] (See also Remark 5.2 in [8]), one shows that under (Hb), for \( t \in I \), there exists \( S_t, Q_t \) in \( \mathcal{F}_t \) such that \( S_t \cap Q_t = \emptyset \) and

\[
|\tilde{Y}_t^\theta - \tilde{Y}_t^\gamma|^2 \leq |\tilde{Y}_t^\theta - \tilde{Y}_t^\gamma|^2 1_{S_t} + C \lambda |X_t - \tilde{X}_t|^2 1_{Q_t} \tag{51}
\]

Observe in particular that for \( t \notin I \), one can take \( S_t = \Omega \) and \( Q_t = \emptyset \) in (51) since, in this case, \( \tilde{Y}_t^\theta = \tilde{Y}_t^\gamma \) and \( \tilde{Y}_t = \tilde{Y}_t^\gamma \).

The proof of (50) is then similar to the proof of Proposition 5.1 (steps ia and ii) in [8]. Note that since \( f \) does not depend on \( z \) in the present case, the expression of \( B_i \) in equation (5.5) of [8] reduces to

\[
B_i = \int_{t_{i-1}}^{t_i} (|X_u - \tilde{X}_{t_{i-1}}|^2 + |\tilde{Y}_u^\theta - \tilde{Y}_{t_{i-1}}^\theta|^2) du .
\]

Observing that, for \( u \in [t_{i-1}, t_i) \),

\[
\mathbb{E}[|\tilde{Y}_u^\theta - \tilde{Y}_{t_{i-1}}^\theta|^2] \leq C \mathbb{E} \left[ \int_{t_{i-1}}^{t_i} |f(s, X_s, \tilde{Y}_s^\theta)|^2 ds + \int_{t_{i-1}}^{t_i} |\tilde{Y}_s^\theta|^2 du \right],
\]

we obtain \( \mathbb{E}[\sum_i B_i] \leq C|t| \). Inequalities (50) then follow from exactly the same arguments as in the proof of Proposition 5.1.

The proof of the theorem is concluded combining (49) and (50) with Proposition 3.5.

### 4.2.4 Proof of Theorem 3.3

Since

\[
|Y_t - \tilde{Y}_t|^2 \leq C \lambda (|Y_t - \tilde{Y}_t^\theta|^2 + |\tilde{Y}_t^\theta - \tilde{Y}_t|^2) \quad \text{and} \quad |Y_t - \tilde{Y}_t|^2 \leq C \lambda (|Y_t - \tilde{Y}_t^\theta|^2 + |\tilde{Y}_t^\theta - \tilde{Y}_t|^2),
\]

we obtain using Proposition 3.4 and Theorem 3.6 that

\[
\sup_{t \in [0, T]} \mathbb{E}[|Y_t - \tilde{Y}_t|^2] + \sup_{t \in [0, T]} \mathbb{E}[|Y_t - \tilde{Y}_t|^2] \leq C \lambda \left( |t| + |\theta|^\alpha + C_\lambda |\theta|^\alpha - 1 \sum_{i=1}^N \mathbb{E}[|\tilde{Y}_i - \tilde{Y}_i|^2] \right)^{1-\varepsilon}.
\]

Under (Hb)', the proof is concluded by using the last inequality together with Proposition 3.1 and letting \( r = \pi \).

Under (Hb), one chooses \( |\theta| \sim |t|^{1/2} \).

### 4.3 Proof of Proposition 2.6

By standard semi-group properties of \( \mathcal{X} \) and \( \mathcal{Y} \) immediately resulting from the uniqueness of the solutions to the related SDEs, one gets, for every \( I = 1, \ldots, N \) and \( T_{i-1} \leq t \leq r < T_i \) (see Crépey [11] for the detail),

\[
Y_t^r = u_I(r, \mathcal{X}_t^r) \quad \mathbb{P}-a.s. \tag{52}
\]
for a deterministic function $u_I$ on $\mathcal{E}_i^s$. In particular,
\[
Y_t^k = u^k(t, x), \quad \text{for any } (t, x, k) \in \mathcal{E},
\]
where $u$ is the function defined on $\mathcal{E}$ by the concatenation of the $u_I$s and the terminal condition $g$ at $T$. In view of \[\square\], the fact that $u$ is of class $\mathcal{P}$ then directly follows from the bound estimates \[\square\] on $Y^{0,t}$ and $Y^{1,t}$.

Let us show that the $u_I$s are continuous over the $\mathcal{E}_i^s$. Given $\mathcal{E} \ni (t_n, x_n, k) \to (t, x, k)$ with $t \notin \mathcal{S}$ or $t_n \geq T_I = t$, one decomposes by \[\square\]:
\[
|u^k(t, x) - u^k(t_n, x_n)| = |Y_t^k - Y_{t_n}^{1,t_n}| \leq \begin{cases} 
|\mathbb{E}(Y_t^{0,t} - Y_{t_n}^{0,t})| + \mathbb{E}|Y_t^{0,t} - Y_{t_n}^{0,t}|, & k \in K \\
|\mathbb{E}(Y_t^{1,t} - Y_{t_n}^{1,t})| + \mathbb{E}|Y_t^{1,t} - Y_{t_n}^{1,t_n}|, & k \notin K 
\end{cases}
\]

In either case we conclude classically by using Proposition 2.3 as a main tool, as for instance in the proof of Theorem 9.3(i) of Crèpey \[\square\] Part II, that $|u^k(t, x) - u^k(t_n, x_n)|$ goes to zero as $n \to \infty$.

It remains to show that the $u_I$s can be extended by continuity over the $\mathcal{E}_i$s, except maybe at the boundary points $(T_I, x \in \partial \mathcal{O}, k)$, and that the jump condition \[\square\] is satisfied. Given $\mathcal{E}_i^s \ni (t_n, x_n, k) \to (t = T_I, x, k)$ with $x \notin \partial \mathcal{O}$, one needs to show that $u_I(t_n, x_n) = u_I(t_n, x_n) \to u^k(T_I, x)$, where $u^k(T_I, x)$ is given by \[\square\]. We distinguish four cases.

- In case $k \notin K$ and $\kappa_I^k(x) \in K$, one has, denoting $\tilde{u}(s, y) = \min(u(s, y, \kappa_I^k(y)), h(s, y))$ and $\tilde{u}^j(s, y) = \min(u^j(s, y), h(s, y))$,
\[
|\tilde{u}^k(t, x) - u^k(t_n, x_n)| = |\tilde{u}^k(t, x) - Y_{t_n}^{1,t_n}| \leq 2\mathbb{E}|\tilde{u}^k(t, x) - \tilde{u}(t, X_t^{1,t_n})|^2 + 2\mathbb{E}|\tilde{u}(t, X_t^{1,t_n}) - Y_{t_n}^{1,t_n}|^2.
\]

By continuity of $\kappa_I^k$ at $x$, one has $\kappa_I^k(X_t^{1,t_n}) = \kappa_I^k(x) \in K$ for $X_t^{1,t_n}$ close enough to $x$, say $|X_t^{1,t_n} - x| \leq c$. In this case $t = \partial_{2n}^0$, therefore (cf. \[\square\]) $Y_t^{1,t_n} = \tilde{u}(t, X_t^{1,t_n})$. So
\[
\mathbb{E}|1_{|X_t^{1,t_n} - x| \leq c} - \tilde{u}(t, X_t^{1,t_n}) - Y_{t_n}^{1,t_n}|^2 \leq \mathbb{E}|1_{|X_t^{1,t_n} - x| \leq c} - \tilde{u}(t, X_t^{1,t_n})|^2,
\]
which can be shown to converge to zero as $n \to \infty$ by using the R2BSDE satisfied by $Y^{1,t_n}$ and the convergence of $Y^{1,t_n}$ to $\tilde{Y}^{1,t}$. Moreover $\mathbb{E}|1_{|X_t^{1,t_n} - x| > c} - \tilde{u}(t, X_t^{1,t_n}) - Y_{t_n}^{1,t_n}|^2$ goes to zero as $n \to \infty$ by the a priori estimates on $X$ and $Y^{1,t_n}$ and the continuity of $\tilde{u}$ already established over $\mathcal{E}_i^{s+1}$. Finally by this continuity and the a priori estimates on $X$ the first term in \[\square\] also goes to zero as $n \to \infty$. So, as $n \to \infty$,
\[
u^k(t, x) \to \tilde{u}^k(t, x) = \min(u(t, x, \kappa_I^k(x)), h(T_I, x)) = u^k(T_I, x).
\]

- In case $k \in K$ and $\kappa_I^k(x) \notin K$, one can show likewise, using this time $\tilde{u}^j(s, y) = u(s, y, \kappa_I^j(y))$ instead of $\tilde{u}$, $u(t, X_t^{1,t_n})$ instead of $\tilde{u}(t, X_t^{1,t_n})$ and $Y^0$ instead of $Y^1$ above, that
\[
u^k(t, x) \to \tilde{u}^k(t, x) = u^k(T_I, x)
\]
as $n \to \infty$.

- If $k, \kappa_I^k(x) \notin K$, it comes,
\[
|\tilde{u}^k(t, x) - u^k(t_n, x_n)|^2 = |\tilde{u}^k(t, x) - Y_{t_n}^{1,t_n}|^2 \leq 2\mathbb{E}|\tilde{u}^k(t, x) - u(t, X_t^{1,t_n})|^2 + 2\mathbb{E}|u(t, X_t^{1,t_n}) - Y_{t_n}^{1,t_n}|^2 \leq 2\mathbb{E}|\tilde{u}^k(t, x) - u(t, X_t^{1,t_n})|^2 + 2\mathbb{E}|Y_{T_I}^{1,t_n} - Y_{t_n}^{1,t_n}|^2,
\]
which goes to zero as $\rightarrow \infty$ by an analysis similar to (but simpler than) that of the first bullet point. Hence (55) follows.

$\bullet$ If $k, \kappa^T(x) \in K, (55)$ can be shown as in the above bullet point.

A Proof of Lemma 4.1

A.1 Proof of Part (i)

We denote by $\mathbb{D}^{1,2}$ the set of random variables $F$ which are differentiable in the Malliavin sense and such that $\|F\|_{\mathbb{D}^{1,2}} := \|F\|_{L^2}^2 + \int_0^T \|D_t F\|_{L^2}^2 dt < \infty$, where $D_t F$ denotes the Malliavin derivative of $F$ at time $t \leq T$. After possibly passing to a suitable version, an adapted process belongs to the subspace $\mathcal{L}^{1,2}_a$ of $H^2$ whenever $V_s \in \mathbb{D}^{1,2}$ for all $s \leq T$ and $\|V\|_{\mathcal{L}^{1,2}_a} := \|V\|_{L^2}^2 + \int_0^T \|D_t V\|_{L^2}^2 dt < \infty$. For a general presentation on Malliavin calculus for stochastic differential equations, see e.g. [18].

Step 1a In order to prove the result, we define $C^q_\kappa$ extensions $\bar{\sigma}$ and $\bar{b}$ of $\sigma$ and $b$ from $Q$ to $[0, T] \times \mathbb{R}^q$, such that $\bar{\sigma}$ satisfies (29) on $[0, T] \times \mathbb{R}^q$. We then introduce, for every $t \leq s \leq T$,

$$
\tilde{X}_s^t = X_t + \int_t^s \bar{b}(u, \tilde{X}_u^t) du + \int_t^s \bar{\sigma}(u, \tilde{X}_u^t) dW_u, \quad \xi_s^t = d(\tilde{X}_s^t).
$$

Since $d, \bar{b}$ and $\bar{\sigma}$ are smooth enough, we may introduce the gradient process of $\tilde{X}$ with respect to the initial condition $x$ of $X$, which is defined by, for $t \leq s \leq T$,

$$
\nabla \tilde{X}_s^t = I_q + \int_t^s \partial \bar{b}(u, \tilde{X}_u^t) \nabla \tilde{X}_u^s du + \int_t^s \partial \bar{\sigma}(u, \tilde{X}_u^t) \nabla \tilde{X}_u^s dW_u. \quad (56)
$$

Note the following standard estimates,

$$
\|\nabla \tilde{X}_s^t\|_{S^p} + \|\nabla \tilde{X}_s^t\|_{S^p}^{-1} \|_{S^p} \leq C_\Lambda. \quad (57)
$$

Observe moreover that one has, for $t \leq r \leq s \leq T$,

$$
D_r \tilde{X}_s^t = \bar{\sigma}(t, \tilde{X}_s^t) + \int_t^s \partial \bar{b}(u, \tilde{X}_u^t) D_r \tilde{X}_u^t du + \int_t^s \partial \bar{\sigma}(u, \tilde{X}_u^t) D_r \tilde{X}_u^t dW_u
$$

and therefore

$$
D_r \tilde{X}_s^t = \nabla \tilde{X}_s^t (\nabla \tilde{X}_s^t)^{-1} \bar{\sigma}(r, \tilde{X}_r^t) 1_{t \leq r \leq s}. \quad (58)
$$

Step 1b We now prove that (57) is satisfied by $\tilde{X}_T^T$, namely, for every $1 \leq I \leq N$,

$$
\mathbb{P}\{a \leq d(\tilde{X}_I^-) \leq b\} \leq C_\Lambda \frac{b - a}{h}. \quad (59)
$$

For this step, integration with respect to $dW$ has to be understood in the Skorohod sense (see e.g. [18]). Without loss of generality we fix $I \in \{1, \ldots, N\}$. For $x \in \mathbb{R}$, let

$$
x \mapsto \phi(x) = \int_{-\infty}^b 1_{\{x > z\}} dz = \int_{-\infty}^x 1_{[a,b]} dz
$$
and let $\bar{\phi}$ stand for a regularization of $\phi$. For every $T_I - h \leq r \leq T_I$, we compute

$$D_r \bar{\phi}(d(\widetilde{X}_{T_I}^{T_I-h})) = \nabla \bar{\phi}(d(\widetilde{X}_{T_I}^{T_I-h}))D_r d(\widetilde{X}_{T_I}^{T_I-h})$$

$$= \nabla \bar{\phi}(d(\widetilde{X}_{T_I}^{T_I-h}))\nabla d(\widetilde{X}_{T_I}^{T_I-h})D_r \widetilde{X}_{T_I}^{T_I-h}$$

$$= \nabla \bar{\phi}(d(\widetilde{X}_{T_I}^{T_I-h}))\nabla d(\widetilde{X}_{T_I}^{T_I-h})\nabla \widetilde{X}_{T_I}^{T_I-h}(\nabla \widetilde{X}_{T}^{T_I-h})^{-1}\sigma(r, \widetilde{X}_{T}^{T_I-h})1_{T_I-h \leq r \leq T_I}.$$  

Multiplying by $\psi_t^I = \sigma^T(r, \widetilde{X}_{T_I}^{T_I-h})\nabla \widetilde{X}_{T_I}^{T_I-h}(\nabla \widetilde{X}_{T_I}^{T_I-h})^{-1}a(r, \widetilde{X}_{T_I}^{T_I-h})^{-1}$, it comes,

$$E\left[\nabla \bar{\phi}(d(\widetilde{X}_{T_I}^{T_I-h}))\right] = h^{-1}E\left[\int_{T_I-h}^{T_I} \ E\left[\bar{\phi}(d(\widetilde{X}_{T_I}^{T_I-h}))\psi_t^I \right] dt \right]$$

$$= h^{-1}E\left[\bar{\phi}(d(\widetilde{X}_{T_I}^{T_I-h}))\int_{T_I-h}^{T_I} \psi_t^I \right].$$

Recalling that $\bar{\phi}$ is the regularization of $\phi$, the last equality leads, together with Fubini theorem and (57), to,

$$E\left[\nabla_x \phi(\widetilde{X}_{T_I}^{T_I-h})\right] = h^{-1} \sum_{k=1}^q \int_a^b \left[1_{\{\widetilde{X}_{T_I}^{T_I-h} > x\}} \int_{T_I-h}^{T_I} \psi_t^I \right] dt,$$

which concludes the proof of (59).

**Step 2** We now prove that we can replace $\widetilde{X}_{T_I}^{T_I-h}$ by $X_{T_I}$ on $\Omega^h_I$.

Let $\Omega^h_I = \{d(X_{T_I} - h) \leq \frac{1}{3\Lambda}\}$, and

$$\tau_I^h = \inf\{s \geq T_I - h ; d(X_s)b^B_I \geq \frac{1}{\Lambda}\} \wedge T.$$  

For every $u \in [T_I - h, T]$, one has,

$$\widetilde{X}_{u \wedge \tau_I^h}^{T_I-h}1_{B_I^h} = X_{u \wedge \tau_I^h}1_{B_I^h}. \quad (60)$$

Observe that $\tau_I^h > T_I - h$ and that for $T_I - h \leq u \leq T$, one has,

$$(b, \sigma)(X_{u \wedge \tau_I^h})1_{B_I^h} = (\bar{b}, \bar{\sigma})(X_{u \wedge \tau_I^h})1_{B_I^h}.$$  

Let, for $u \geq t$, $\phi(u) = E\left[|\widetilde{X}_{u \wedge \tau_I^h}^{T_I-h} - X_{u \wedge \tau_I^h}^{T_I-h}|^21_{B_I^h}\right]$. Straightforward computations based on the Itô formula yield,

$$\phi(u) \leq C_A \int_t^u \phi(s) ds,$$

Identity (60) then follows by an application of Gronwall’s Lemma.

The proof of the Lemma is then concluded by combining (60) and (59).
### A.2 Proof of Part (ii)

We consider the two different cases.

**Case 1 (i)** By definition of $\vartheta$, $\vartheta'$, we have that $\mathbb{E}[|\vartheta_0 - \vartheta'_0|] = |t - t'|$, and obviously, for $l \geq 1$,

$$\mathbb{E}[|\vartheta_l - \vartheta'_l|] = \mathbb{E}[|\vartheta_l - \vartheta'_l|1_{\Omega^h_{\cup \cap} \cup \hat{\Omega}^s}] + \mathbb{E}[|\vartheta_l - \vartheta'_l|1_{\Omega^h \cap \hat{\Omega}^s \cap \{\vartheta_l \neq \vartheta'_l\}}] . \tag{61}$$

Tchebytchev's inequality applied on $\hat{\Omega}^h$, for $T_l \in \mathcal{T}_l$ and $\hat{\Omega}^s$, and the bound $|\vartheta_l - \vartheta'_l| \leq T_l$ yield that

$$\mathbb{E}[|\vartheta_l - \vartheta'_l|1_{\Omega^h_{\cup \cap} \cup \hat{\Omega}^s}] \leq C^h_\Lambda h^p + C^p_\Lambda \mathbb{E}\left[\sup_{u \in [T_l, T]} |X_u' - X_u|^p\right], \tag{62}$$

for $p, \rho > 0$.

**Case 2 (ii)** We now work on the second term of the right-hand side of (61).

By definition of $\vartheta$, $\vartheta'$, if $k \notin K$, we have $\mathbb{E}[|\vartheta_l - \vartheta'_l|1_{\{k \notin K\}}] = |t - t'|$. We are going to prove a control between $\vartheta$ and $\vartheta'$, for $l \geq 2$, and for $l = 1$, $k \in K$.

To this end, we observe that

$$1_{\{X_{T_l} \in \mathcal{O}\}} = 1_{\{X_{T_l} \in \mathcal{O}\}}, \forall T_l \in \mathcal{T}_l \implies H = H', \tag{63}$$

thus for $l \geq 2$, $\vartheta_l = \vartheta'_l$ and if $k \in K$, $\vartheta_1 = \vartheta'_1$.

We then introduce the set

$$\Omega_1 = \bigcup_{T_l \in \mathcal{T}_l} \left((d(X_{T_l}) \geq 0) \cap (d(X_{T_l}') < 0)\right) \cup (\{d(X_{T_1}) < 0\} \cap \{d(X_{T_1}') \geq 0\}) .$$

Since $d$ is 1-Lipschitz continuous, by definition of $\hat{\Omega}^s$, we have

$$\Omega_1 \subset \bigcup_{T_l \in \mathcal{T}_l} \{|d(X_{T_l})| \leq \delta\}$$

This leads, by definition of $\Omega^h$, to

$$\Omega^h \cap \hat{\Omega}^s \cap \Omega_1 \subset \Omega , \text{ with } \Omega := \bigcup_{T_l \in \mathcal{T}_l} \Omega^h_l \cap \{|d(X_{T_l})| < \delta\} .$$

Using (63), we have that, for $l \geq 2$, $\{\vartheta_l \neq \vartheta'_l\} \subset \Omega_1$ and if $k \in K$, $\{\vartheta_1 \neq \vartheta'_1\} \subset \Omega_1$. Thus, for $l \geq 2$, $\Omega^h \cap \hat{\Omega}^s \cap \{\vartheta_l \neq \vartheta'_l\} \subset \Omega$ and if $k \in K$, $\Omega^h \cap \hat{\Omega}^s \cap \{\vartheta_1 \neq \vartheta'_1\} \subset \Omega$.

Using the result of part (i), one then gets,

$$\mathbb{E}[|\vartheta_l - \vartheta'_l|1_{\Omega^h_l \cap \hat{\Omega}^s \cap \{\vartheta_l \neq \vartheta'_l\}}] \leq C^\delta_\Lambda 1 ,$$

for $l \geq 2$ and $l = 1$, if $k \in K$. In this case, the proof is concluded combining the last inequality with (62) and (61).
Case 2 In this case, $\mathcal{T}' = \mathcal{T} \cup \{t\}$.

As in Case 1 (i) above, we compute

$$
\mathbb{E}\left[ |\tilde{\vartheta}_t - \vartheta'_t| \right] \leq C_\Lambda^{{\rho}} h^p + C_\Lambda^{{\rho}} \frac{\mathbb{E}\left[ \sup_{u \in [T, T]} |X'_u - X_u|^p \right]}{\delta^p} + \mathbb{E}\left[ |\tilde{\vartheta}_t - \vartheta'_t| 1_{\Omega^h \cap \Omega^\delta \cap \{\tilde{\vartheta}_t \neq \vartheta'_t\}} \right], \quad (64)
$$

for $l \geq 0$ and $p, \rho > 0$.

Recall that by definition of $\tilde{\vartheta}$, $\vartheta'$, $\mathbb{E}\left[ |\tilde{\vartheta}_0 - \vartheta'_0| \right] = |t - t'|$ and if $k \notin K$, $\mathbb{E}\left[ |\tilde{\vartheta}_1 - \vartheta'_1| \right] = |t - t'|$.

Regarding the last term of (64), we observe here that

$$1_{\{X_{T_I} \in \mathcal{O}\}} = 1_{\{X'_{T_I} \in \mathcal{O}\}}, \forall T_I \in \mathcal{T} \cup \{t\} \implies \tilde{H} = H'.$$

The set $\Omega_1$ is now replaced by

$$\Omega_2 = \bigcup_{T_I \in \mathcal{T} \cup \{t\}} \{d(X_{T_I}) \geq 0 \} \cap \{d(X'_{T_I}) < 0 \} \cup \{d(X_{T_I}) < 0 \} \cap \{d(X'_{T_I}) \geq 0 \}$$

The difference with the last step is that the reunion is on $\mathcal{T} \cup \{t\}$. But, since for $\delta$ small enough $\{d(X_t) < \delta\} = \emptyset$, we have

$$\Omega^h \cap \tilde{\Omega}^\delta \cap \Omega_2 \subset \Omega.$$

The proof is then concluded arguing as in the last step.

References


