

APPROXIMATION OF THE SNELL ENVELOPE AND AMERICAN OPTIONS PRICES IN DIMENSION ONE

VLAD BALLY¹ AND BRUNO SAUSSEREAU²

Abstract. We establish some error estimates for the approximation of an optimal stopping problem along the paths of the Black–Scholes model. This approximation is based on a tree method. Moreover, we give a global approximation result for the related obstacle problem.

Mathematics Subject Classification. 49L20, 60G40, 65M15, 91B28.

Received February 15, 2001. Revised November 13, 2001.

INTRODUCTION

The purpose of this article is to estimate the rate of convergence of the approximation scheme of an optimal stopping problem along Brownian paths, when the Brownian Motion is approximated by a random walk.

We consider a standard Brownian Motion $(B_t)_{t \geq 0}$ defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and we denote

$$X_t^x = x + b t + a B_t, \quad a, b, x \in \mathbb{R}, \quad a > 0.$$

Given a function h which is assumed to be (at least) Lipschitz-continuous, our aim is to compute

$$Y_t^x = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}(h(\tau, X_\tau^x) | \mathcal{F}_t),$$

where $\mathcal{T}_{t,T}$ is the set of all the stopping times taking values in $[t, T]$. This is the Snell Envelope of the process $(h(t, X_t^x))_{t \in [0, T]}$.

Our results apply to standard American Options. Recall, for example, that an American Put on a stock is the right to sell one share at a specified price K (called the exercise price) at any instant until a given future date T (called the expiration date or date of maturity). In the Black–Scholes model, the value at time t ($0 \leq t \leq T$) of such an option is given by the Y_t^x defined above with $h(t, x) := e^{-rt}(K - e^x)^+$. The constant r is the interest rate (assumed to be positive), the diffusion coefficient a is the so-called volatility (assumed to be positive also) and the drift coefficient b , in this case, is equal to $r - a^2/2$.

We give here a discretization scheme based on a tree method, which is a variant of the tree method (see (1.20)) studied in [8–11]. These authors study a binomial approximation based on the discretization of the Brownian

Keywords and phrases: Dynamic programming, snell envelope, optimal stopping.

¹ Université du Maine, Laboratoire Statistiques et Processus, BP. 535, 72017 Le Mans Cedex, France;
e-mail: bally@ccr.jussieu.fr

² Université de Franche-Comté, Laboratoire de Mathématiques de Besançon, 16 route de Gray, 25000 Besançon, France;
e-mail: bruno.saussereau@math.univ-fcomte.fr

motion $(B_t)_{t \in \mathbb{R}^+}$. So they take $(t, y) \mapsto h(t, y) = e^{-rt}(K - e^{x+(r-a^2/2)t+ay})_+$ which is slightly different from the function h considered here. An alternative to the tree method is to approach the numerical solution of the optimal stopping problem based on the discretization of the variational inequality satisfied by the value function. We refer to [1] for the error estimates for finite difference schemes. A comparison of existing methods for American Options is given in [5].

Let us present our algorithm. We assume that the function h is of the special form $h(t, x) = e^{-rt}f(x)$ (which covers the applications in mathematical finance). We fix $\varepsilon > 0$ and we define the sequence of functions \hat{u}_k by

$$\begin{aligned} \hat{u}_{k(T)}(x + i\varepsilon) &= f(x + i\varepsilon), \quad -k(T) \leq i \leq k(T), \\ &\vdots \\ \hat{u}_k(x + i\varepsilon) &= \max \left\{ f(x + i\varepsilon), \gamma_\varepsilon^+ \hat{u}_{k+1}(x + i\varepsilon + \varepsilon) + \gamma_\varepsilon^- \hat{u}_{k+1}(x + i\varepsilon - \varepsilon) \right\}, \quad -k \leq i \leq k, \end{aligned}$$

where $k(T)$ is an integer depending on ε such that $k(T) \sim a^2 T / \varepsilon^2$. The constants γ_ε^+ and γ_ε^- appear from exact computations related to the first exit time for the Brownian motion with drift. In Section 1 we give their interpretations and their exact values. Let us give here just an expansion in powers of ε which is sufficient for concrete calculations:

$$\begin{aligned} \gamma_\varepsilon^+ &= \frac{1}{2} + \frac{b}{2a^2}\varepsilon - \frac{r}{2a^2}\varepsilon^2 - \left(\frac{b^3}{6a^6} + \frac{br}{2a^4} \right) \varepsilon^3 + \left(\frac{5r^2}{12a^4} + \frac{rb}{12a^6} \right) \varepsilon^4 + O(\varepsilon^5) \\ \gamma_\varepsilon^- &= \frac{1}{2} - \frac{b}{2a^2}\varepsilon - \frac{r}{2a^2}\varepsilon^2 + \left(\frac{b^3}{6a^6} + \frac{br}{2a^4} \right) \varepsilon^3 + \left(\frac{5r^2}{12a^4} + \frac{rb}{12a^6} \right) \varepsilon^4 + O(\varepsilon^5). \end{aligned}$$

We prove that

$$|Y_0^x - \hat{u}_0(x)| \leq c\delta_\varepsilon. \quad (0.1)$$

The values of c and δ_ε depend on the regularity of the function f (see Th. 1.2). If f is Lipschitz continuous then $\delta_\varepsilon = \sqrt{\varepsilon}$, if f is a linear combination of convex functions then $\delta_\varepsilon = \varepsilon\sqrt{-\ln(\varepsilon)}$ and if f is twice differentiable then $\delta_\varepsilon = \varepsilon$. The complexity of the algorithm is of order $\varepsilon^{-3}\sqrt{-\ln \varepsilon}$. The estimation in $\delta_\varepsilon = \varepsilon\sqrt{-\ln(\varepsilon)}$ is new, as far as we know.

The idea which leads to our algorithm is the following. We define a sequence of stopping times by $\tau_0 = 0$ and $\tau_{k+1} = \inf \{t > \tau_k : |X_t^x - X_{\tau_k}^x| \geq \varepsilon\}$. The chain $(X_{\tau_k}^x)_{k \in \mathbb{N}}$ lives on the grid $\{x + i\varepsilon, i \in \mathbb{Z}\}$ and has a very simple structure because if we define $\sigma_k = \tau_{k+1} - \tau_k$ and $\delta_k = X_{\tau_{k+1}}^x - X_{\tau_k}^x$, then $(\sigma_k, \delta_k)_{k \in \mathbb{N}}$ are independent random variables and we know their laws (this permits to compute explicitly γ_ε^+ and γ_ε^-). On the other hand, using the law of large numbers, $\tau_{k(T)} \sim T$ if $k(T) \sim a^2 T / \varepsilon^2$. So we replace the Snell envelope $(Y_t^x)_{0 \leq t \leq T}$ of $(h(t, X_t^x))_{0 \leq t \leq T}$ by the discrete Snell envelope $(Q_k^{\varepsilon, x})_{0 \leq k \leq k(T)}$ of $(h(\tau_k, X_{\tau_k}^x))_{0 \leq k \leq k(T)}$. The algorithm described above represents the computation of $Q_0^{\varepsilon, x}$ using the dynamical programming principle (see Sect. 1). Note the analogy of our idea (to produce a chain which approximates $(X_t^x)_{t \geq 0}$) and that of [11] (to embed a chain in the Brownian motion).

In order to control the error we employ the basic evaluation $\mathbb{E}|T - \tau_{k(T)}|^2 \leq C\varepsilon^2$ proved in Section 1. This immediately leads to (0.1) in the case of a Lipschitz functions f . If f is twice differentiable we use Ito's formula on the intervals $[\tau_k, \tau_{k+1}]$ in order to improve our evaluation. If f is a linear combination of convex functions (this is the more interesting case) we use the generalization of the Ito's formula for such functions and then local time comes into play. In this case, we have to prove (see Appendix 1) a more refined inequality concerning the increments of the local time.

In the second section, we prove a global approximation result. It is well known that we may write $Y_t^x = u(t, X_t^x)$ where $u(t, x)$, $(t, x) \in [0, T] \times \mathbb{R}^d$ is the solution of the obstacle problem

$$\begin{cases} (\partial_t + (a^2/2)\partial_{x^2} + b \partial_x) u(t, x) = 0, & \text{on } u > h \\ (\partial_t + (a^2/2)\partial_{x^2} + b \partial_x) u(t, x) \leq 0, & \text{on } u = h \\ u(T, x) = h(T, x). \end{cases}$$

We prove that the function $\hat{u}_k(x)$ constructed above provides an uniform approximation for $u(t, y)$. More precisely, we prove that

$$|u(t_k, y) - e^{-rt} \hat{u}_k(\bar{y})| \leq C\delta_\varepsilon$$

where $t_k = k\varepsilon^2/a^2$, $0 \leq k \leq a^2T/\varepsilon^2$, and \bar{y} is the projection of y on the grid $\{x + i\varepsilon, -t_k \leq i \leq t_k\}$ constructed for the approximation of Y_0^x . The constant C depends on a, b, T and on the regularity of the obstacle (see Th. 2.1). The above approximation permits also to obtain an approximation of the continuation region for the obstacle problem and of the first and last optimal exercise time for an American Option.

A detailed analysis of the constants appearing in our evaluation is rather heavy, so we leave it out here and send the interested reader to [14]. In this paper we will just emphasize the structure of these constants in the following sense:

- C will designate constants which depend on a, b and T . It may change from line to line;
- $C_i(h)$, $i = 1, 2, \dots$ designates constants which depend on the obstacle h and its structure will be specified in each case. This is interesting because one can see how the regularity of h appears;
- $C_i(h, x)$, $i = 1, 2, \dots$ designates constants which depend on the obstacle h and on the starting point x . We are especially interested in the dependence on x because it comes into play in Section 2 when we discuss the uniform approximation of the solution of the variational inequality.

1. APPROXIMATION OF Y_0^x

In this section, we give the discretization scheme for computing Y_0^x . In the next section, we will construct a global approximation scheme for the solution of the corresponding obstacle problem.

We fix $\varepsilon > 0$ and we define

$$\tau_0 = 0, \quad \tau_{k+1} = \inf \{t > \tau_k : |X_t^x - X_{\tau_k}^x| \geq \varepsilon\},$$

and we denote $\sigma_k = \tau_{k+1} - \tau_k$. The following proposition is proved in [12] ((3.6), p. 100).

Proposition 1.1. *For every $k \in \mathbb{N}$, $(\tau_{k+1} - \tau_k, X_{\tau_{k+1}}^x - X_{\tau_k}^x)$ is independent of \mathcal{F}_{τ_k} .*

Remark that $(\sigma_k)_{k \in \mathbb{N}}$ is a sequence of independent identically distributed random variables. Using the scaling property and the strong Markov property, σ_k has the same law that $\varepsilon^2 \sigma^\varepsilon$ where

$$\sigma^\varepsilon := \inf \left\{ t > 0 : \left| \frac{b\varepsilon}{a} t + W_t \right| \geq \frac{1}{a} \right\}$$

with $(W_t)_{t \geq 0}$ a standard Brownian Motion. The important fact is the knowledge of the law of σ^ε . Indeed, we have (see p. 233 in [4]),

$$\mathbb{E}(\exp(-\lambda \sigma^\varepsilon)) = \frac{2 \cosh\left(\frac{b\varepsilon}{a^2}\right) \sinh\left(\frac{1}{a} \sqrt{2\lambda + \frac{b^2 \varepsilon^2}{a^2}}\right)}{\sinh\left(\frac{2}{a} \sqrt{2\lambda + \frac{b^2 \varepsilon^2}{a^2}}\right)}.$$

So we can compute $\alpha_\varepsilon := \mathbb{E}\sigma^\varepsilon$ and $\beta_\varepsilon^2 := \mathbb{E}(\sigma^\varepsilon - \alpha_\varepsilon)^2$ and show that

$$\begin{aligned}\alpha_\varepsilon &= \frac{1 - \exp\left(-\frac{2|b|\varepsilon}{a^2}\right)}{\varepsilon|b|(1 + \exp\left(-\frac{2|b|\varepsilon}{a^2}\right))} \xrightarrow{\varepsilon \rightarrow 0} \frac{1}{a^2} \\ \beta_\varepsilon^2 &\xrightarrow{\varepsilon \rightarrow 0} \frac{2}{3a^4}.\end{aligned}$$

Anyway, it is easy to see that $\sup_{\varepsilon \leq 1} (\alpha_\varepsilon + \beta_\varepsilon) \leq C < \infty$.

For each time t we denote

$$k(t) := \left\lceil \frac{t}{\alpha_\varepsilon \varepsilon^2} \right\rceil,$$

where $\lceil \cdot \rceil$ is the integer part. We prove that $\tau_{k(t)} \xrightarrow{\varepsilon \rightarrow 0} t$. We write

$$\begin{aligned}\mathbb{E}|t - \tau_{k(t)}|^2 &= \mathbb{E} \left| -\left(t - \alpha_\varepsilon \varepsilon^2 k(t)\right) + \sum_{k=0}^{k(t)} (\sigma_k - \alpha_\varepsilon \varepsilon^2) \right|^2 \\ &= \left(t - \alpha_\varepsilon \varepsilon^2 k(t)\right)^2 + \sum_{k=0}^{k(t)} \mathbb{E}(\sigma_k - \alpha_\varepsilon \varepsilon^2)^2 \\ &= \left(t - \alpha_\varepsilon \varepsilon^2 k(t)\right)^2 + \varepsilon^4 \sum_{k=0}^{k(t)} (\mathbb{E}\sigma^\varepsilon - \alpha_\varepsilon)^2 \\ &= \left(t - \alpha_\varepsilon \varepsilon^2 k(t)\right)^2 + k(t) \varepsilon^4 \beta_\varepsilon^2.\end{aligned}$$

So we finally obtain that

$$\mathbb{E}|t - \tau_{k(t)}|^2 \leq \alpha_\varepsilon^2 \varepsilon^4 + \frac{t \beta_\varepsilon^2 \varepsilon^2}{\alpha_\varepsilon} \leq C \varepsilon^2. \quad (1.1)$$

The approximation scheme is defined by

$$Q_k^{\varepsilon, x} := \sup_{\nu \in \mathcal{S}_{k, k(T)}} \mathbb{E} \left(h(\tau_\nu, X_{\tau_\nu}^x) \mid \mathcal{F}_{\tau_k} \right),$$

where $\mathcal{S}_{k, k(T)}$ is the family of all the $(\mathcal{F}_{\tau_i})_{0 \leq i \leq k(T)}$ -stopping times with values in the set $\{k, \dots, k(T)\}$.

Note that τ_ν is a stopping time with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$. Indeed if $\nu \in \mathcal{S}_{k, k(T)}$, then $\{\tau_\nu \leq t\} = \bigcup_k \{\tau_k \leq t, \nu = k\}$ and since $\{\nu = k\} \in \mathcal{F}_{\tau_k}$ and τ_k is a stopping time, we have $\{\tau_k \leq t, \nu = k\} \in \mathcal{F}_t$. We shall give later an algorithm which permits to compute $Q_0^{\varepsilon, x}$.

For the moment, we want to evaluate the error $Y_0^x - Q_0^{\varepsilon, x}$. Except for Lipschitz continuous functions h , we shall also consider the following more special class of functions.

H.1. For any $t \in \mathbb{R}^+$ and $x \in \mathbb{R}$, $h(t, x) = \varphi(t)f(x)$ with

- i) φ belongs to $C_b^1(\mathbb{R}^+)$;
- ii) there exists \tilde{f} with finite variation such that $f(x) = f(0) + \int_0^x \tilde{f}(u) du$;
- iii) we denote by μ the signed measure associated to \tilde{f} via its representation as a difference of two increasing functions and we assume that μ has finite total variation, denoted by $\|\mu\|$.

The simplest example of a function satisfying **H.1** is $h(t, x) = \varphi(t)f(x)$ with $f \in C_b^2(\mathbb{R})$. In this case, $\mu(dx) = \frac{\partial^2 f(x)}{\partial x^2} dx$. Another interesting example is $h(t, x) = \exp(-rt)(K - \exp(x))_+$ which corresponds to a put. Then, $\tilde{f}(x) = -\exp(x)\mathbf{1}_{x \leq \ln(K)}$ and $\mu(dx) = -\exp(x)\mathbf{1}_{x \leq \ln(K)} dx + K\delta_{\ln(K)}(dx)$. Remark that we have $\|f\|_\infty = \|\tilde{f}\|_\infty = K$ and $\|\mu\| = 2K$.

More generally, one may prove that a function f satisfies **H.1** *ii*) and *iii*) if and only if it is a linear combination of convex functions (see [13], p. 23).

For a function h satisfying **H.1**, we have the following extension of Itô's rule:

$$h(t, X_t^x) = h(0, X_0^x) + \int_0^t \varphi'(s)f(X_s^x)ds + \int_0^t \varphi(s)\tilde{f}(X_s^x)dX_s^x + \varphi(t) \int_{\mathbb{R}} L_t^{x,y} \mu(dy), \quad (1.2)$$

where $L_t^{x,y}$ denotes the local time of $(X_t^x)_{t \geq 0}$ at the point $y \in \mathbb{R}$ (see for example [4] for the case of h depending only on x and apply the classical integration by parts formula with the function $\varphi(\cdot)$ for the general case).

We are now able to give the main result of this section.

Theorem 1.2. *Suppose that h is Lipschitz continuous. Then*

$$|Q_0^{\varepsilon,x} - Y_0^x| \leq C_1(h) \sqrt{\varepsilon}, \quad (1.3)$$

where $C_1(h) = C [h]_1$ is the Lipschitz constant of h .

Suppose that h satisfies **H.1**, then

$$|Q_0^{\varepsilon,x} - Y_0^x| \leq C_2(h, x) \varepsilon \sqrt{-\ln \varepsilon}, \quad (1.4)$$

where

$$C_2(h, x) = C \left(\|\varphi'\|_\infty (|f(x)| + [f]_1 + \|\mu\|) + \|\varphi\|_\infty (\|\mu\| + |\tilde{f}(0)|) \right). \quad (1.5)$$

Suppose that $h \in C_b^{1,2}(\mathbb{R}^+ \times \mathbb{R})$ then

$$|Q_0^{\varepsilon,x} - Y_0^x| \leq C_3(h) \varepsilon, \quad (1.6)$$

where $C_3(h) = C \|\partial_s h + b\partial_x h + \frac{a^2}{2}\partial_{x^2} h\|_\infty$.

Remark 1.3. By (1.5), one remarks that $C_2(h, x)$ is independent of x if f is bounded. This is the case when we apply this result to an American Put. Then $h(t, x) = e^{-rt}f(x)$ with $f(x) = (K - e^x)_+$ which satisfies $\|f\|_\infty = \|\tilde{f}\|_\infty = K$.

Remark 1.4. Notice that the convergence rate in (1.6) is the same as in [11] because our time discretization is of order $n = \varepsilon^{-2}$.

Note also that in [8], the author proved that in the case of a put option (which corresponds to the case (1.4) in our theorem), the speed of convergence is better: $(\sqrt{\ln n}/n)^{\frac{4}{5}}$ (see also the results in [9]). In fact numerical experiments (see [5]) show that for a put, the speed of convergence is $1/n$. So for some particular payoff functions, one may obtain better results.

Proof of Theorem 1.2.

Step 1: The first part of this proof is devoted to state estimates analogous to those in Theorem 1.2 for the following intermediary quantity.

For a stopping time τ , we denote $\tau^T = \tau \wedge T$ and we define

$$\bar{Q}_0^{\varepsilon,x} = \sup_{\nu \in \mathcal{S}_{0,k(T)}} \mathbb{E} \left(h(\tau_\nu^T, X_{\tau_\nu^T}^x) \right).$$

Note first that for every $\nu \in \mathcal{S}_{0,k(T)}$ one has $|\tau_\nu^T - \tau_\nu| \leq |T - \tau_{k(T)}|$, so that by (1.1)

$$\|\tau_\nu - \tau_\nu^T\|_2 \leq C \varepsilon. \quad (1.7)$$

- We prove that, if h is Lipschitz continuous

$$|\bar{Q}_0^{\varepsilon,x} - Q_0^{\varepsilon,x}| \leq C_{1,1}(h) \sqrt{\varepsilon} \quad \text{with } C_{1,1}(h) = C [h]_1. \quad (1.8)$$

For every $\nu \in \mathcal{S}_{0,k(T)}$

$$\begin{aligned} \left| \mathbb{E} \left(h(\tau_\nu, X_{\tau_\nu}^x) - h(\tau_\nu^T, X_{\tau_\nu^T}^x) \right) \right| &\leq [h]_1 \left(\mathbb{E} |\tau_\nu^T - \tau_\nu| + \mathbb{E} |X_{\tau_\nu^T}^x - X_{\tau_\nu}^x| \right) \\ &\leq [h]_1 \left((1 + |b|) \mathbb{E} |\tau_\nu^T - \tau_\nu| + a \sqrt{\mathbb{E} |\tau_\nu^T - \tau_\nu|} \right). \end{aligned}$$

By (1.7), it follows that (1.8) holds true.

- We prove now that if $h \in C_b^{1,2}(\mathbb{R}^+ \times \mathbb{R})$ then

$$|\bar{Q}_0^{\varepsilon,x} - Q_0^{\varepsilon,x}| \leq C_{3,1}(h) \varepsilon \quad \text{with } C_{3,1}(h) = C \left\| \partial_s h + b \partial_x h + \frac{a^2}{2} \partial_{x^2} h \right\|_\infty. \quad (1.9)$$

Using Itô's formula one gets

$$\mathbb{E} \left(h(\tau_\nu, X_{\tau_\nu}^x) - h(\tau_\nu^T, X_{\tau_\nu^T}^x) \right) = \mathbb{E} \int_{\tau_\nu^T}^{\tau_\nu} \left(\partial_s h + b \partial_x h + \frac{a^2}{2} \partial_{x^2} h \right) (s, X_s^x) ds + \mathbb{E} \int_{\tau_\nu^T}^{\tau_\nu} \partial_x h(s, X_s^x) \sigma dB_s.$$

Since $\partial_x h$ is bounded and τ_ν is integrable, the integral with respect to B is a stochastic integral (not only a local martingale), so its expectation vanishes. It follows that

$$\left| \mathbb{E} \left(h(\tau_\nu, X_{\tau_\nu}^x) - h(\tau_\nu^T, X_{\tau_\nu^T}^x) \right) \right| \leq \left\| \partial_s h + b \partial_x h + \frac{a^2}{2} \partial_{x^2} h \right\|_\infty \mathbb{E} |\tau_\nu - \tau_\nu^T|,$$

so (1.9) follows from (1.7).

- Suppose now that h satisfies **H.1**. We prove that

$$\begin{aligned} |\bar{Q}_0^{\varepsilon,x} - Q_0^{\varepsilon,x}| &\leq \varepsilon \sqrt{-\ln \varepsilon} C_{2,1}(h, x) \quad \text{with} \\ C_{2,1}(h, x) &= C'_{2,1}(h) |f(x)| + C''_{2,1}(h) \quad \text{where} \\ C'_{2,1}(h) &= C \|\varphi'\|_\infty \quad \text{and} \\ C''_{2,1}(h) &= C \|\varphi'\|_\infty ([f]_1 + \|\mu\|) + C \|\varphi\|_\infty (\|\mu\| + |\tilde{f}(0)| + 1). \end{aligned} \quad (1.10)$$

We apply the formula (1.2) between 0 and τ_ν , and between 0 and τ_ν^T . Subtracting the two formulas we obtain

$$\begin{aligned} & \left| \mathbb{E} \left(h(\tau_\nu, X_{\tau_\nu}^x) - h(\tau_\nu^T, X_{\tau_\nu^T}^x) \right) \right| \leq A_1 + A_2 \text{ with} \\ A_1 &= \mathbb{E} \mathbf{1}_{[\tau_\nu \geq T]}(\omega) \left| \int_T^{\tau_\nu} (\varphi'(s) f(X_s^x) + b\varphi(s) \tilde{f}(X_s^x)) ds \right| \\ A_2 &= \mathbb{E} \left| \int_{\mathbb{R}} (\varphi(\tau_\nu) L_{\tau_\nu}^{x,y} - \varphi(\tau_\nu^T) L_{\tau_\nu^T}^{x,y}) \mu(dy) \right|. \end{aligned}$$

We first estimate A_1 . We write

$$A_1 \leq \|\tau_{k(T)} - T\|_{\mathbb{L}^2} \left(\|\varphi'\|_\infty \left\| \sup_{s \in [0, \tau_{k(T)}]} |f(X_s^x)| \right\|_{\mathbb{L}^2} + |b| \|\varphi\|_\infty \left\| \sup_{s \in [0, \tau_{k(T)}]} |\tilde{f}(X_s^x)| \right\|_{\mathbb{L}^2} \right).$$

By **H.1** $|\tilde{f}(y) - \tilde{f}(z)| = |\mu(\{u : y \leq u \leq z\})| \leq \|\mu\|$. Hence

$$\|\tilde{f}\|_\infty \leq \|\mu\| + |\tilde{f}(0)|$$

Note that

$$|f(X_s^x)| \leq |f(x)| + [f]_1 |X_s^x - x| \leq |f(x)| + [f]_1 (|b|s + \sigma |B_S|).$$

So, using (1.1), we obtain

$$A_1 \leq \varepsilon C \left(\|\varphi'\|_\infty (|f(x)| + [f]_1) + |b| \|\varphi\|_\infty (\|\mu\| + |\tilde{f}(0)|) \right).$$

The term A_2 can be treated as follows. We write

$$\begin{aligned} A_2 &\leq A_{2,1} + A_{2,2} \text{ with} \\ A_{2,1} &= \mathbb{E} \left| \varphi(\tau_\nu) \int_{\mathbb{R}} L_{\tau_\nu}^{x,y} - L_{\tau_\nu^T}^{x,y} \mu(dy) \right| \text{ and} \\ A_{2,2} &= \mathbb{E} \left| (\varphi(\tau_\nu) - \varphi(\tau_\nu^T)) \int_{\mathbb{R}} L_{\tau_\nu^T}^{x,y} \mu(dy) \right|. \end{aligned}$$

We begin with the simplest term. Since $L_{\cdot}^{x,y}$ does not grow up to $\tau_x^y :=$ the first hitting time of y by X_{\cdot}^x , we have $L_t^{x,y} = L_{t-\tau_y^x}^{y,y} \mathbf{1}_{\tau_y^x \leq t} \leq L_t^{y,y}$ ($\mathbf{1}$ denotes the indicator function). Then using (1.7)

$$\begin{aligned} A_{2,2} &\leq \|\varphi'\|_\infty \|\tau_\nu - \tau_\nu^T\|_{\mathbb{L}^2} \int_{\mathbb{R}} \|L_{\tau_\nu^T}^{x,y}\|_{\mathbb{L}^2} |\mu|(dy) \\ &\leq \|\varphi'\|_\infty \|T - \tau_{k(T)}\|_{\mathbb{L}^2} \int_{\mathbb{R}} \|L_{\tau_\nu^T}^{y,y}\|_{\mathbb{L}^2} |\mu|(dy) \\ &\leq \varepsilon C \|\varphi'\|_\infty \|\mu\| \sup_{y \in \mathbb{R}} \|L_T^{y,y}\|_{\mathbb{L}^2}. \end{aligned}$$

Using the Tanaka's formula, one gets easily that

$$\|L_T^{y,y}\|_{\mathbb{L}^2}^2 \leq 4b^2 T^2 + 4a^2 T + 4 \sup_{y \in \mathbb{R}} \mathbb{E} |X_T^y - y|^2 \leq 12(b^2 T^2 + a^2 T),$$

and this ends the estimation of $A_{2,2}$.

For $A_{2,1}$, we will need the following lemma (the proof is given in the Appendix 1).

Lemma 1.5. *With the above notations we have,*

$$\mathbb{E}\mathbf{1}_{[\tau_{k(T)} \geq T]} (L_{\tau_{k(T)}}^{x,y} - L_T^{x,y}) \leq C \varepsilon \sqrt{-\ln \varepsilon}. \quad (1.11)$$

We assume (1.11). For $\nu \in \mathcal{S}_{0,k(T)}$ we have

$$\begin{aligned} \mathbb{E} \left| L_{\tau_\nu}^{x,y} - L_{\tau_\nu^T}^{x,y} \right| &= \mathbb{E}\mathbf{1}_{[\tau_\nu \geq T]} (L_{\tau_\nu}^{x,y} - L_T^{x,y}) \\ &\leq \mathbb{E}\mathbf{1}_{[\tau_{k(T)} \geq T]} (L_{\tau_{k(T)}}^{x,y} - L_T^{x,y}) \leq C \varepsilon \sqrt{-\ln \varepsilon}. \end{aligned}$$

It follows that $A_{2,1} \leq \|\varphi\|_\infty \int_{\mathbb{R}} \mathbb{E} \left| L_{\tau_\nu}^{x,y} - L_{\tau_\nu^T}^{x,y} \right| \mu(dy) \leq C \varepsilon \sqrt{-\ln \varepsilon}$.

Step 2: In order to prove Theorem 1.2, it remains to estimate $\bar{Q}_0^{\varepsilon,x} - Y_0^x$.

Let $\tau \in \mathcal{T}_{0,T}$ and let $\nu = \inf\{i : \tau_i \geq \tau\}$. We have $\{\nu > i\} = \bigcap_{j \leq i} \{\tau_j \leq \tau\}$ and since $\{\tau_j \leq \tau\} \in \mathcal{F}_{\tau_j} \subseteq \mathcal{F}_{\tau_i}$ for $i \geq j$, ν is a stopping time with respect to the filtration $(\mathcal{F}_{\tau_k})_k$. So, τ_ν is also a $(\mathcal{F}_t)_t$ -stopping time, and $T \wedge \tau_\nu$ also. Due to these remarks, one gets that

$$\bar{Q}_0^{\varepsilon,x} = \sup_{\nu \in \mathcal{S}_{0,k(T)}} \mathbb{E}h(\tau_\nu^T, X_{\tau_\nu^T}^x) \leq Y_0^x. \quad (1.12)$$

Given a stopping time τ , we denote $\tau' = \inf\{t > \tau : |X_t^x - X_\tau^x| \geq 2\varepsilon\}$ and using the strong Markov property we get

$$\mathbb{E}(\tau_\nu \wedge T - \tau) \leq \mathbb{E}(\tau_\nu - \tau) \leq \mathbb{E}(\tau' - \tau) = 4\alpha_{2\varepsilon}\varepsilon^2 \leq \frac{6}{a^2} \varepsilon^2. \quad (1.13)$$

The above inequality will play the same role in the sequel that (1.7) played in the first step.

We fix τ and we denote by ν the corresponding stopping time defined as above. We have to compare $\mathbb{E}h(\tau, X_\tau^x)$ and $\mathbb{E}h(\tau_\nu^T, X_{\tau_\nu^T}^x)$. The same computations as for (1.8) and (1.9) (using (1.13) instead of (1.7)) yield

- if h is Lipschitz continuous

$$\left| \mathbb{E}h(\tau, X_\tau^x) - \mathbb{E}h(\tau_\nu^T, X_{\tau_\nu^T}^x) \right| \leq \varepsilon C_{1,2}(h) \text{ with } C_{1,2}(h) = C[h]_1$$

- if $h \in C_b^{1,2}(\mathbb{R}^+ \times \mathbb{R})$ then the same reasoning (based on Itô's formula) as for (1.7) gives

$$\left| \mathbb{E}h(\tau, X_\tau^x) - \mathbb{E}h(\tau_\nu^T, X_{\tau_\nu^T}^x) \right| \leq \varepsilon^2 C_{3,2}(h) \text{ with } C_{3,2}(h) = C \left\| \partial_s h + b \partial_x h + \frac{a^2}{2} \partial_{x^2} h \right\|_\infty.$$

So we have proved that

$$Y_0^x = \sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E}h(\tau, X_\tau^x) \leq \sup_{\nu \in \mathcal{S}_{0,k(T)}} \mathbb{E}h(\tau_\nu^T, X_{\tau_\nu^T}^x) + \varepsilon C_{i,2}(h) = \bar{Q}_0^{\varepsilon,x} + \varepsilon C_{i,2}(h)$$

with $i = 1$ (respectively 3) if h is Lipschitz continuous (respectively $C_b^{1,2}$).

This, together with (1.12) proves that

- if h is Lipschitz continuous

$$\left| \bar{Q}_0^{\varepsilon,x} - Y_0^x \right| \leq \varepsilon C_{1,2}(h) \text{ with } C_{1,2}(h) = C [h]_1$$

- if $h \in C_b^{1,2}(\mathbb{R}^+ \times \mathbb{R})$

$$|\bar{Q}_0^{\varepsilon,x} - Y_0^x| \leq \varepsilon^2 C_{3,2}(h) \quad \text{with} \quad C_{3,2}(h) = C \left\| \partial_s h + b \partial_x h + \frac{a^2}{2} \partial_{x^2} h \right\|_{\infty}.$$

If h satisfies **H.1** we have to give more details.

Using the strong Markov property and the scaling property we get $\mathbb{E}(\tau' - \tau)^2 = 16\varepsilon^4 \mathbb{E}\sigma_{2\varepsilon}^2$ and recall that $\mathbb{E}\sigma_{\varepsilon}^2 \leq C$. So we have

$$\mathbb{E}(\tau' - \tau)^2 \leq C \varepsilon^4. \quad (1.14)$$

Arguing as in the proof of (1.10), one may prove that

$$\begin{aligned} & \left| \mathbb{E}h(\tau, X_{\tau}^x) - \mathbb{E}h(\tau_{\nu}^T, X_{\tau_{\nu}^T}^x) \right| \leq \tilde{C}\varepsilon^2 + \tilde{A} \quad \text{with} \\ & \tilde{A} = \sup_{\tau \in \mathcal{T}_{0,T}} \mathbb{E} \left| \int_{\mathbb{R}} \varphi(T \wedge \tau_{\nu})(L_{T \wedge \tau_{\nu}}^{x,y} - L_{\tau}^{x,y}) \mu(dy) \right| \quad \text{and} \\ & \tilde{C} = C \left(\|\varphi'\|_{\infty} (|f(x)| + [f]_1) + \|\varphi\|_{\infty} (\|\mu\| + |\tilde{f}(0)|) + \|\varphi'\|_{\infty} \|\mu\| \right). \end{aligned}$$

We write

$$\tilde{A} \leq \|\varphi\|_{\infty} \|\mu\| \sup_{\tau \in \mathcal{T}_{0,T}} \sup_{y \in \mathbb{R}} \mathbb{E}(L_{\tau_{\nu}}^{x,y} - L_{\tau}^{x,y}).$$

Using Tanaka's formula, one gets

$$\frac{1}{2} L_{\tau_{\nu}}^{x,y} = (X_{\tau_{\nu}}^x - y)^+ - (x - y)^+ + \int_0^{\tau_{\nu}} a \mathbf{1}_{[X_s > y]} dB_s - \int_0^{\tau_{\nu}} b \mathbf{1}_{[X_s > y]} ds$$

and the analogous formula for τ . Since $|(z - y)^+ - (z' - y)^+| \leq |z - z'|$, it yields

$$\mathbb{E}(L_{\tau_{\nu}}^{x,y} - L_{\tau}^{x,y})^2 \leq 36(b^2 \mathbb{E}|\tau_{\nu} - \tau|^2 + a^2 \mathbb{E}|\tau_{\nu} - \tau|),$$

so by (1.13) and (1.14)

$$\tilde{A} \leq \varepsilon C \|\varphi\|_{\infty} \|\mu\|.$$

Then the same reasoning as above yields

$$\begin{aligned} & |\bar{Q}_0^{\varepsilon,x} - Y_0^x| \leq \varepsilon C_{2,2}(h, x) \quad \text{with} \\ & C_{2,2}(h, x) = C \left(\|\varphi'\|_{\infty} (|f(x)| + [f]_1) + \|\varphi\|_{\infty} (\|\mu\| + |\tilde{f}(0)|) + \|\varphi'\|_{\infty} \|\mu\| \right). \end{aligned} \quad (1.15)$$

The Algorithm

In order to give an explicit algorithm for computing $Q_0^{\varepsilon,x}$, we have to consider functions h satisfying **H.1** which depend on the time in a special way, namely $h(t, x) = \exp(-rt)f(x)$. This is sufficient for applications in Mathematical Finance.

From the dynamic programming principle, we already know that

$$\begin{aligned} Q_{k(T)}^{\varepsilon,x} &= \exp(-r\tau_k(T))f(X_{\tau_k(T)}^x) \\ &\vdots \\ Q_k^{\varepsilon,x} &= \max\left(\exp(-r\tau_k)f(X_{\tau_k}^x), \mathbb{E}(Q_{k+1}^{\varepsilon,x} | \mathcal{F}_{\tau_k})\right). \end{aligned} \quad (1.16)$$

Our aim is to compute $Q_k^{\varepsilon,x}$. To this end, we first write $\tau_{k+1} = \tau_k + \sigma_k$ and $X_{\tau_{k+1}}^x = X_{\tau_k}^x + \delta_k$. The important fact is that $(\delta_k, \sigma_k)_{k \in \mathbb{N}}$ are independent of \mathcal{F}_{τ_k} (see Prop. 1.1) and we know that (see [4], p. 233)

$$\mathbb{E}\left[\exp(-r\sigma_k)\mathbf{1}_{\delta_k=\varepsilon}\right] = \frac{\exp\left(\frac{b\varepsilon}{a^2}\right)\sinh\left(\frac{\varepsilon}{a}\sqrt{2r+\frac{b^2}{a^2}}\right)}{\sinh\left(\frac{2\varepsilon}{a}\sqrt{2r+\frac{b^2}{a^2}}\right)} := \gamma_\varepsilon^+ \quad \text{and} \quad (1.17)$$

$$\mathbb{E}\left[\exp(-r\sigma_k)\mathbf{1}_{\delta_k=-\varepsilon}\right] = \frac{\exp\left(-\frac{b\varepsilon}{a^2}\right)\sinh\left(\frac{\varepsilon}{a}\sqrt{2r+\frac{b^2}{a^2}}\right)}{\sinh\left(\frac{2\varepsilon}{a}\sqrt{2r+\frac{b^2}{a^2}}\right)} := \gamma_\varepsilon^-. \quad (1.18)$$

We define now

$$\hat{Q}_k^{\varepsilon,x} = \exp(r\tau_k)Q_k^{\varepsilon,x},$$

and we multiply with $\exp(r\tau_k)$ in (1.16) in order to obtain the recurrence relation

$$\begin{aligned} \hat{Q}_{k(T)}^{\varepsilon,x} &= f(X_{\tau_k(T)}^x) \\ &\vdots \\ \hat{Q}_k^{\varepsilon,x} &= \max\left\{f(X_{\tau_k}^x), \mathbb{E}\left(\exp(-r\sigma_k)\hat{Q}_{k+1}^{\varepsilon,x} | \mathcal{F}_{\tau_k}\right)\right\}. \end{aligned}$$

We want to construct a sequence of functions \hat{u}_k such that $\hat{Q}_k^{\varepsilon,x} = \hat{u}_k(X_{\tau_k}^x)$. Suppose that this is true. Then the above recurrence relation yields

$$\begin{aligned} \hat{u}_k(X_{\tau_k}^x) &= \max\left\{f(X_{\tau_k}^x), \mathbb{E}(\exp(-r\sigma_k)\hat{u}_{k+1}(X_{\tau_k}^x + \delta_k) | \mathcal{F}_{\tau_k})\right\} \\ &= \max\left\{f(X_{\tau_k}^x), \mathbb{E}\left[e^{-r\sigma_k}\mathbf{1}_{\delta_k=\varepsilon}\right]\hat{u}_{k+1}(X_{\tau_k}^x + \varepsilon) + \mathbb{E}\left[e^{-r\sigma_k}\mathbf{1}_{\delta_k=-\varepsilon}\right]\hat{u}_{k+1}(X_{\tau_k}^x - \varepsilon)\right\} \\ &= \max\left\{f(X_{\tau_k}^x), \gamma_\varepsilon^+\hat{u}_{k+1}(X_{\tau_k}^x + \varepsilon) + \gamma_\varepsilon^-\hat{u}_{k+1}(X_{\tau_k}^x - \varepsilon)\right\}. \end{aligned}$$

This leads us to define

$$\begin{aligned} \hat{u}_{k(T)}(x + i\varepsilon) &= f(x + i\varepsilon), \quad -k(T) \leq i \leq k(T), \\ &\vdots \\ \hat{u}_k(x + i\varepsilon) &= \max\left\{f(x + i\varepsilon), \gamma_\varepsilon^+\hat{u}_{k+1}(x + i\varepsilon + \varepsilon) + \gamma_\varepsilon^-\hat{u}_{k+1}(x + i\varepsilon - \varepsilon)\right\}, \quad -k \leq i \leq k. \end{aligned} \quad (1.19)$$

Then $\hat{Q}_k^{\varepsilon,x} = \hat{u}_k(X_{\tau_k}^x)$ and in particular for $k = 0$

$$\hat{u}_0(x) = \hat{Q}_0^{\varepsilon,x} = Q_0^{\varepsilon,x}.$$

Remarks. Notice that the algorithm is analogous to the one proposed by [11] except for the contribution of γ_ε^+ and γ_ε^- . In [11], the authors proceed as follows. Given a terminal time T ($T = 1$ for simplicity), they approximate the Brownian motion by a standard binomial random walk which can be embedded in the paths of Brownian motion. The approximation of Y_0^x , $x = 0$, is then given by $U_0(0)$ where the sequence $(U_k(\cdot))_{0 \leq k \leq n}$ is defined by

$$\begin{aligned} U_n \left(\frac{i}{\sqrt{n}} \right) &= h \left(1, \frac{i}{\sqrt{n}} \right), \quad -n \leq i \leq n, \\ &\vdots \\ U_k \left(\frac{i}{\sqrt{n}} \right) &= \max \left(h \left(\frac{k}{n}, \frac{i}{\sqrt{n}} \right), 0.5 U_{k+1} \left(\frac{i}{\sqrt{n}} + \frac{1}{\sqrt{n}} \right) + 0.5 U_{k+1} \left(\frac{i}{\sqrt{n}} - \frac{1}{\sqrt{n}} \right) \right), \end{aligned} \quad (1.20)$$

$$-k \leq i \leq k,$$

where $h(t, u) = e^{-rt} f(x + bt + au)$.

In [11], the authors prove that for $f \in C^{1,2}([0, T] \times \mathbb{R})$, one has $|U_0 - Y_0^0| \leq C/\sqrt{n}$, and in [8], one considers the particular case of a put and proves that $|U_0 - Y_0^0| \leq C(\sqrt{\ln n}/n)^{\frac{4}{5}}$.

In our framework, we do not give a binomial approximation of the Brownian motion but of the diffusion itself (including the drift part). This leads us to use the function $h(t, x) = \varphi(t)f(x)$. So we do not work with the same function h as Lamberton–Rogers in [11]. We can compute explicitly the coefficient of our binomial approximation. Whereas Lamberton and Rogers have $\gamma_\varepsilon^+ = \gamma_\varepsilon^- = 1/2$, in our frame we have

$$\begin{aligned} \gamma_\varepsilon^+ &= \frac{1}{2} + \frac{b}{2a^2}\varepsilon - \frac{r}{2a^2}\varepsilon^2 - \left(\frac{b^3}{6a^6} + \frac{br}{2a^4} \right) \varepsilon^3 + \left(\frac{5r^2}{12a^4} + \frac{rb}{12a^6} \right) \varepsilon^4 + O(\varepsilon^5) \\ \gamma_\varepsilon^- &= \frac{1}{2} - \frac{b}{2a^2}\varepsilon - \frac{r}{2a^2}\varepsilon^2 + \left(\frac{b^3}{6a^6} + \frac{br}{2a^4} \right) \varepsilon^3 + \left(\frac{5r^2}{12a^4} + \frac{rb}{12a^6} \right) \varepsilon^4 + O(\varepsilon^5). \end{aligned}$$

Complexity of the algorithm.

It is clear that the complexity of the algorithm depends on the number of points $k = k(T)$ that we consider at the first stage. So we have to take advantage of the fact that the diffusion process does not go too far away from its starting point x . More precisely, we have

$$\mathbb{P} \left[\sup_{s \in [0, T]} |X_s^x - x| \geq c_\varepsilon \right] \leq \mathbb{P} \left[\sup_{s \in [0, T]} |B_s| \geq \frac{c_\varepsilon + |b| T}{a} \right],$$

and by Bernstein's inequality (see [12], Eq. (3.16), p. 145),

$$\begin{aligned} \mathbb{P} \left[\sup_{s \in [0, T]} |X_s^x - x| \geq c_\varepsilon \right] &\leq \exp \left(-\frac{(c_\varepsilon + |b| T)^2}{2a^2 T} \right) \\ &\leq \exp \left(-\frac{|b|^2}{2a^2} \right) \times \exp \left(-\frac{c_\varepsilon^2}{2a^2 T} \right). \end{aligned}$$

So if we take

$$c_\varepsilon = \sqrt{-2a^2 T \ln \varepsilon}, \quad (1.21)$$

we get $\mathbb{P}[\sup_{s \in [0, T]} |X_s^x - x| \geq c_\varepsilon] \leq \varepsilon$. This permits to use the function $\tilde{h}(t, y) = h(t, y) \mathbf{1}_{|y-x| \leq \sqrt{-2a^2 T \ln \varepsilon}}$ in the above algorithm. Then the number of operations to be performed is of the order $\frac{c_\varepsilon}{\varepsilon} k(T)$ which is of the order

$\varepsilon^{-3}\sqrt{-\ln \varepsilon}$ (instead of ε^{-4} in the initial algorithm) corresponding to an error of:

- $\sqrt{\varepsilon}$ for a Lipschitz continuous function,
- $\varepsilon\sqrt{-\ln \varepsilon}$ for a function satisfying **H.1** and
- ε for a regular function.

2. GLOBAL APPROXIMATION

The aim of this section is to prove that the function \hat{u} constructed in the previous section provides an approximation of the solution u of the PDE associated to our obstacle problem. Let us be more precise. We fix the time horizon T and the obstacle of special form $h(t, x) = \exp(-rt)f(x)$ and we denote by $u_{T,h}$ the unique solution of the PDE

$$\begin{cases} (\partial_t + L) u_{T,h}(t, x) = 0, & \text{on } u_{T,h} > h \\ (\partial_t + L) u_{T,h}(t, x) \leq 0, & \text{on } u_{T,h} = h \\ u_{T,h}(T, x) = h(T, x), \end{cases} \quad (2.1)$$

where L is the infinitesimal generator of the diffusion process X , namely, $L = (a^2/2)\partial_{x^2} + b\partial_x$. The equation (2.1) can be rewritten as

$$\begin{aligned} \min(u_{T,h}(t, x) - h(t, x), -\partial_t u_{T,h}(t, x) - Lu_{T,h}(t, x)) &= 0, \quad (t, x) \in [0, T] \times \mathbb{R} \\ u_{T,h}(T, x) &= h(T, x). \end{aligned}$$

Our aim is to give a uniform approximation of $u_{T,h}(\cdot, \cdot)$ coming from the previous algorithm constructed for $Y_0^x = u_{T,h}(0, x)$. We will prove that the values computed on the grid $\{x + i\varepsilon, -t_k \leq i \leq t_k\}$ will provide an approximation of $u(t_k, y)$ for $0 \leq k \leq a^2T/\varepsilon^2$ and y in the above grid.

The solution of the above PDE has to be understood in a weak sense. In [6] one considers solutions in viscosity sense and in [2], the variational sense is considered and it corresponds to variational inequalities (see [3] for the analytical approach to optimal stopping problems by means of variational inequalities). We will precise later the sense given to (2.1). The relation between the process $(Y_t^x)_{0 \leq t \leq T}$ and u is given by $Y_t^x = u(t, X_t^x)$. We shall obtain an algorithm which permits to compute Y_t^x for $0 \leq t \leq T$, by means of the previous algorithm constructed for Y_0^x .

We fix now $\varepsilon > 0$ such that $\frac{T}{\varepsilon^2 \alpha_\varepsilon}$ is an integer (of course we may find ε as small as we want, having this property). We denote

$$N := \frac{T}{\varepsilon^2 \alpha_\varepsilon}.$$

For $n \in \{0, \dots, N\}$ we define $t_n := n(\varepsilon^2 \alpha_\varepsilon)$ (remark that $t_N = T$) and we note that

$$k(t_n) = n \quad \text{and} \quad k(T - t_n) = N - n.$$

Let x be the starting point of the diffusion process for which we have already constructed the function \hat{u}_n on a triangular grid. For $y \in \mathbb{R}$, we define \bar{y} the projection of y on the grid $\{x + i\varepsilon, -n \leq i \leq n, 0 \leq n \leq N\}$, that is

$$\bar{y} = i\varepsilon \text{ if } y \in \left[\left(i - \frac{1}{2}\right)\varepsilon, \left(i + \frac{1}{2}\right)\varepsilon \right).$$

We are now able to give our result.

Theorem 2.1. *Let \hat{u} be defined in (1.19). For every $n = 0, \dots, N$ and every $y \in [x - n\varepsilon, x + n\varepsilon]$*

$$|u_{T,h}(t_n, y) - \exp(-rt_n)\hat{u}_n(\bar{y})| \leq K\delta_\varepsilon$$

where

$$\begin{aligned}\delta_\varepsilon &= \sqrt{\varepsilon} \text{ and } K = \left(C_1(h) + \frac{4[h]_1}{\sqrt{\varepsilon}} \right) \text{ if } h \text{ is Lipschitz continuous,} \\ \delta_\varepsilon &= \varepsilon\sqrt{-\ln \varepsilon} \text{ and } K = \left(C_2(h, y) + \frac{4[h]_1}{\sqrt{-\ln \varepsilon}} \right) \text{ if } h \text{ satisfies } \mathbf{H.1}, \\ \delta_\varepsilon &= \varepsilon \text{ and } K = (C_3(h) + 4[h]_1) \text{ if } h \text{ is } C^{1,2},\end{aligned}$$

and $C_1(h)$, $C_2(h, y)$ and $C_3(h)$ are the constants appearing in Theorem 1.2. In particular, under $\mathbf{H.1}$, if f is bounded then the evaluation is uniform with respect to y .

Before proving this, we need some preliminary results about the solution of the PDE (2.1). We recall the definition of viscosity solution.

Definition 2.2. a) $u_{T,h} \in C([0, T] \times \mathbb{R}; \mathbb{R})$ is called a viscosity sub-solution of (2.1) if $u_{T,h}(T, x) \leq h(T, x)$, $x \in \mathbb{R}$ and moreover for any $\varphi \in C^{1,2}([0, T] \times \mathbb{R})$ and every $(t, x) \in [0, T] \times \mathbb{R}$ which is a local maximum of $u_{T,h} - \varphi$,

$$\min(u_{T,h}(t, x) - h(t, x), -\partial_t \varphi(t, x) - L\varphi(t, x)) \leq 0$$

b) $u_{T,h} \in C([0, T] \times \mathbb{R}; \mathbb{R})$ is called a viscosity super-solution of (2.1) if $u_{T,h}(T, x) \geq h(T, x)$, $x \in \mathbb{R}$ and moreover for any $\varphi \in C^{1,2}([0, T] \times \mathbb{R})$ and every $(t, x) \in [0, T] \times \mathbb{R}$ which is a local minimum of $u_{T,h} - \varphi$,

$$\min(u_{T,h}(t, x) - h(t, x), -\partial_t \varphi(t, x) - L\varphi(t, x)) \geq 0$$

c) $u_{T,h} \in C([0, T] \times \mathbb{R}; \mathbb{R})$ is called a viscosity solution of (2.1) if it is both a viscosity sub- and super-solution.

It is proved in [6] that the equation (2.1) has a unique viscosity solution.

Given $t > 0$, we introduce the function $(s, x) \mapsto h^t(s, x) := \exp(-rt) \times h(s, x) = h(t + s, x)$.

Lemma 2.3. For any $t < T$, $x \in \mathbb{R}$ we have $u_{T,h}(t, x) = u_{T-t, h^t}(0, x)$.

Proof. We denote $v(s, x) = u_{T,h}(t + s, x)$ for $0 \leq s \leq T - t$ and $w(s, x) = u_{T-t, h^t}(s, x)$ for $0 \leq s \leq T - t$. We shall prove that $\forall 0 \leq s \leq T - t$, $v(s, x) = w(s, x)$ and in particular $s = 0$ gives the result. In order to do this, we prove that these two functions solve the same following PDE in viscosity sense:

$$\begin{cases} (\partial_s + L)z(s, x) = 0, & 0 \leq s \leq T - t, \text{ on } z > h^t \\ (\partial_s + L)z(s, x) \leq 0, & 0 \leq s \leq T - t, \text{ on } z = h^t \\ z(T - t, x) = h(T, x), \end{cases} \quad (2.2)$$

with $z \in C([0, T - t] \times \mathbb{R}; \mathbb{R})$.

We prove that both v and w are viscosity sub-solution of (2.2). By definition, w solves (2.2) so in particular it is a sub-solution. It remains to deal with v . First we have $v(T - t, x) = u_{T,h}(t + T - t, x) = h(T, x)$. Let $\varphi \in C^{1,2}([0, T - t] \times \mathbb{R})$ and $(s, x) \in [0, T - t] \times \mathbb{R}$ which is a local maximum of $v - \varphi$. We define $\varphi^t \in C^{1,2}([t, T] \times \mathbb{R})$ by $\varphi^t(r, x) = \varphi(r - t, x)$ for $(r, x) \in [t, T] \times \mathbb{R}$. Now since $(t + s, x)$ is a local maximum of $u_{T,h} - \varphi^t$ and $u_{T,h}$ solves (2.1) in the viscosity sense, we have

$$\min(u_{T,h}(t + s, x) - h(t + s, x), -\partial_t \varphi^t(t + s, x) - L\varphi^t(t + s, x)) \leq 0,$$

which reads

$$\min(v(s, x) - h^t(s, x), -\partial_t \varphi(s, x) - L\varphi(s, x)) \leq 0.$$

So v is a viscosity sub-solution of (2.2).

The same arguments show that v is also a viscosity super-solution of (2.2) so the lemma is proved. \square

We shall finally need the following lemma:

Lemma 2.4. *Suppose h Lipschitz continuous. Then for any $0 \leq t \leq T$ and any $x \in \mathbb{R}$, we have*

$$|u_{T,h}(t, x) - u_{T,h}(t, \bar{x})| \leq 4[h]_1 \varepsilon.$$

Proof. We use the following expression of the solution of the obstacle problem as the value function of an optimal stopping problem:

$$u_{T,h}(t, x) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}h(\tau, X_\tau^x).$$

In this frame we have

$$\begin{aligned} |u_{T,h}(t, x) - u_{T,h}(t, y)| &\leq \left| \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}h(\tau, X_\tau^x) - \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}h(\tau, X_\tau^y) \right| \\ &\leq \sup_{\tau \in \mathcal{T}_{t,T}} |\mathbb{E}h(\tau, X_\tau^x) - \mathbb{E}h(\tau, X_\tau^y)| \\ &\leq [h]_1 \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}|X_\tau^x - X_\tau^y| \\ &\leq [h]_1 |x - y|, \end{aligned}$$

and the result follows. \square

Now we can turn to the proof of Theorem 2.1.

Proof. Remind that we have denoted $t_n = n \frac{T}{N}$.

For $n = 0, \dots, N$ and $y \in [x - n\varepsilon, x + n\varepsilon]$, we introduce

$$\begin{aligned} \bar{u}_{k(t_N - t_n)}(N - n, y + i\varepsilon) &= f(y + i\varepsilon), \quad -(N - n) \leq i \leq N - n, \\ &\vdots \\ \bar{u}_{k(t_N - t_n)}(k, y + i\varepsilon) &= \max \left\{ f(y + i\varepsilon), \gamma_\varepsilon^+ \bar{u}_{k(t_N - t_n)}(k + 1, y + i\varepsilon + \varepsilon) \right. \\ &\quad \left. + \gamma_\varepsilon^- \bar{u}_{k(t_N - t_n)}(k + 1, y + i\varepsilon - \varepsilon) \right\}, \quad -k \leq i \leq k. \end{aligned} \quad (2.3)$$

Clearly

$$\hat{u}_n(y) = \bar{u}_{k(t_N - t_n)}(0, y). \quad (2.4)$$

Using Lemma 2.3 and (2.4)

$$\begin{aligned} u_{T,h}(t_n, y) - e^{-rt_n} \hat{u}_n(\bar{y}) &= u_{T,h}(t_n, y) - u_{T,h}(t_n, \bar{y}) \\ &\quad + u_{T,h}(t_n, \bar{y}) - u_{T-t_n, h^{t_n}}(0, \bar{y}) \\ &\quad + u_{T-t_n, h^{t_n}}(0, \bar{y}) - e^{-rt_n} \bar{u}_{k(T-t_n)}(0, \bar{y}) \\ &\quad + e^{-rt_n} \bar{u}_{k(T-t_n)}(0, \bar{y}) - e^{-rt_n} \hat{u}_n(\bar{y}) \\ &= u_{T,h}(t_n, y) - u_{T,h}(t_n, \bar{y}) + u_{T-t_n, h^{t_n}}(0, \bar{y}) - e^{-rt_n} \bar{u}_{k(T-t_n)}(0, \bar{y}) \\ &= a_1 + a_2. \end{aligned}$$

By Lemma 2.4, it holds that $|a_1| \leq 4[h]_1 \varepsilon$. The estimation of a_2 follows from Theorem 1.2. \square

An Improvement. Remind that we have fixed $\varepsilon > 0$ such that $\frac{T}{\varepsilon^2 \alpha_\varepsilon} := N$ is an integer. We moreover denote α_N the corresponding value of α_ε .

As in the previous section, we can take advantage of the fact that the diffusion process does not go far away from its starting point. Remind that $c_\varepsilon = \sqrt{-2a^2 T \ln \varepsilon}$ (see (1.21)) and we now denote

$$C_N := \sqrt{-a^2 T \ln \frac{T}{N \alpha_N}}.$$

This suggest using $\tilde{h}(t, y) = h(t, y) \mathbf{1}_{|y-x| \leq c_N}$ instead of h . We define the intervals

$$I_n^N = \left[x - n \left(\frac{T}{N \alpha_N} \right)^{1/2}, x + n \left(\frac{T}{N \alpha_N} \right)^{1/2} \right] \cap [x - c_N, x + c_N].$$

This finally leads us to the

Proposition 2.5. *For every $n = 0, \dots, N$ and every $x \in I_n^N$ we define \hat{u}_n by (1.19) replacing $h(t, y) = e^{-rt} f(y)$ by $\tilde{h}(t, y) = e^{-rt} f(y) \mathbf{1}_{|y-x| \leq c_N}$. Then we have for any $y \in I_n^N$*

$$\sup_{n=0, \dots, N} |u_{T,h}(t_n, y) - \exp(-rt_n) \hat{u}_n(\bar{y})| \leq K \delta_\varepsilon \text{ with}$$

$$K = \left(C_1(h) + \frac{4[h]_1}{\sqrt{\varepsilon}} \right) \text{ and } \delta_\varepsilon = \sqrt{\varepsilon} \text{ if } h \text{ is Lipschitz continuous,}$$

$$K = \left(C_2(h, x) + \frac{4[h]_1}{\sqrt{-\ln \varepsilon}} \right) \text{ and } \delta_\varepsilon = \varepsilon \sqrt{-\ln \varepsilon} \text{ if } h \text{ satisfies } \mathbf{H.1},$$

$$K = (C_3(h) + 4[h]_1) \text{ and } \delta_\varepsilon = \varepsilon \text{ if } h \text{ is } C^{1,2},$$

where $C_1(h)$, $C_2(h, x)$ and $C_3(h)$ are defined in Theorem 2.1.

In particular if f is bounded under $\mathbf{H.1}$, then all the approximations are uniform with respect to y .

3. THE OPTIMAL STOPPING TIME

In the PDE's language, the complementary of the set

$$\Lambda = \{(t, x) \in [0, T] \times \mathbb{R} : u(t, x) \leq h(t, x)\}$$

is called the continuation set or the continuation region. The first optimal stopping time is the first time when the diffusion hits Λ . So, as long as we are in Λ^c we continue (but there may be other optimal stopping times later). Our aim is to give an approximation of this set in order to obtain approximations for the optimal stopping time. We introduce first

$$\Lambda_\delta^- = \{(t, x) \in [0, T] \times \mathbb{R} : u(t, x) - \delta \leq h(t, x)\},$$

$$\Lambda_\delta^+ = \{(t, x) \in [0, T] \times \mathbb{R} : u(t, x) + \delta \leq h(t, x)\}.$$

Clearly $\Lambda_\delta^+ \subseteq \Lambda \subseteq \Lambda_\delta^-$. We consider now the function \hat{u} constructed in the previous section and define

$$\hat{\Lambda}_\delta^- = \{(t_n, x), n = 0, \dots, N, x \in I_n^N : \hat{u}(t, \bar{x}) - \delta \leq h(t, x)\},$$

$$\hat{\Lambda}_\delta^+ = \{(t_n, x), n = 0, \dots, N, x \in I_n^N : \hat{u}(t, \bar{x}) + \delta \leq h(t, x)\}.$$

As an immediate consequence of Corollary 2.5 we get:

Proposition 3.1. *Assume that $h(t, u) = \varphi(t)f(u)$.*

$$\text{Then for any } \delta > \begin{cases} 2 \left(C_1(h) + \frac{4[h]_1}{\sqrt{\varepsilon}} \right) \varepsilon & \text{if } h \text{ is Lipschitz continuous,} \\ 2 \left(C_2(h, x) + \frac{4[h]_1}{\sqrt{-\ln \varepsilon}} \right) \varepsilon \sqrt{-\ln \varepsilon} & \text{if } h \text{ satisfies } \mathbf{H.1}, \\ 2(C_3(h) + 4[h]_1) \varepsilon & \text{if } h \text{ is } C^{1,2}, \end{cases}$$

one has $\hat{\Lambda}_{\delta/2}^- \subseteq \Lambda_\delta^- \subseteq \hat{\Lambda}_{2\delta}^-$ and $\hat{\Lambda}_{\delta/2}^+ \subseteq \Lambda_\delta^+ \subseteq \hat{\Lambda}_{2\delta}^+$.

Define now $\tau_f^* = \inf\{t : (t, X_t) \in \Lambda\}$ the first optimal stopping time and $\tau_l^* = \sup\{t : (t, X_t) \in \Lambda\}$ the last time when one may exercise in an optimal way (of course τ_l^* is no more a stopping time). The discrete versions of these times are given by

$$\begin{aligned} \tau_{f,\delta}^- &= \inf\{t : (t, X_t) \in \hat{\Lambda}_\delta^-\}, & \tau_{f,\delta}^+ &= \inf\{t : (t, X_t) \in \hat{\Lambda}_\delta^+\}, \\ \tau_{l,\delta}^- &= \sup\{t : (t, X_t) \in \hat{\Lambda}_\delta^-\}, & \tau_{l,\delta}^+ &= \sup\{t : (t, X_t) \in \hat{\Lambda}_\delta^+\}. \end{aligned}$$

Then, as an immediate consequence of the above Proposition we have:

Proposition 3.2. *Under the hypotheses of the above Proposition*

$$\tau_{f,\delta}^- \leq \tau_f^* \leq \tau_{f,\delta}^+ \quad \text{and} \quad \tau_{l,\delta}^+ \leq \tau_l^* \leq \tau_{l,\delta}^-.$$

APPENDIX A: PROOF OF LEMMA 1.5

We prove the estimation (1.11). We write

$$\begin{aligned} \mathbb{E}\mathbf{1}_{[\tau_{k(T)} \geq T]} \left(L_{\tau_{k(T)}}^{x,y} - L_T^{x,y} \right) &= I + J \text{ with} \\ I &= \mathbb{E}\mathbf{1}_{[\tau_{k(T)} \geq T]} \mathbf{1}_{[\tau_{k(T)} < T + K\varepsilon\sqrt{-\ln \varepsilon}]} \left(L_{\tau_{k(T)}}^{x,y} - L_T^{x,y} \right) \\ J &= \mathbb{E}\mathbf{1}_{[\tau_{k(T)} \geq T]} \mathbf{1}_{[\tau_{k(T)} \geq T + K\varepsilon\sqrt{-\ln \varepsilon}]} \left(L_{\tau_{k(T)}}^{x,y} - L_T^{x,y} \right). \end{aligned}$$

Here

$$K = \frac{2T\beta_\varepsilon}{\sqrt{k(T)}\alpha_\varepsilon\varepsilon}.$$

The reason of this choice of K will become clear in the proof. For the moment we note that since $k(T) \geq T/(\alpha_\varepsilon\varepsilon^2) - 1$, we have

$$K^2 \leq \frac{4T^2\beta_\varepsilon^2}{T\alpha_\varepsilon - \alpha_\varepsilon\varepsilon^2} \leq C.$$

Step 1: We first deal with I .

Since the local time is an increasing process, we have

$$I = \mathbb{E}\mathbf{1}_{[\tau_{k(T)} \geq T]} \mathbf{1}_{[\tau_{k(T)} < T + K\varepsilon\sqrt{-\ln \varepsilon}]} \left(L_{T + K\varepsilon\sqrt{-\ln \varepsilon}}^{x,y} - L_T^{x,y} \right) \leq \mathbb{E}L_{T + K\varepsilon\sqrt{-\ln \varepsilon}}^{x,y} - \mathbb{E}L_T^{x,y}.$$

We denote $(s, x, z) \mapsto p(s, x, z)$ the transition density function of the diffusion process X and we recall that $\mathbb{E}_x L_t^{x,y} = \int_0^t p(s, x, y) ds$ (see [4], p. 21). It yields

$$\begin{aligned} I &\leq \int_T^{T+K\varepsilon\sqrt{-\ln\varepsilon}} p(s, x, y) ds \\ &\leq K\varepsilon\sqrt{-\ln\varepsilon} \sup_{s \in [T+\infty[} |p(s, x, y)| \\ &\leq K\varepsilon\sqrt{-\ln\varepsilon} \sup_{s \in [T+\infty[} \frac{1}{\sqrt{2\pi a^2 s}} \exp(-(y-x-bs)^2/(2a^2 s)) \\ &\leq C \varepsilon\sqrt{-\ln\varepsilon}. \end{aligned}$$

Step 2: In order to evaluate J , we first use the Cauchy–Schwartz inequality and get

$$\begin{aligned} J &\leq J_1 \times J_2 \text{ with} \\ J_1 &= \left\| \mathbf{1}_{[\tau_k(T) \geq T]} (L_{\tau_k(T)}^{x,y} - L_T^{x,y}) \right\|_{\mathbb{L}^2} \\ J_2 &= \left(\mathbb{P} \left[\tau_k(T) - K\varepsilon\sqrt{-\ln\varepsilon} \geq T \right] \right)^{\frac{1}{2}}. \end{aligned}$$

• We first prove that

$$J_1 \leq C \sqrt{\varepsilon}. \quad (3.1)$$

Using Tanaka's formula, one gets

$$\frac{1}{2} L_{\tau_k(T)}^{x,y} = (X_{\tau_k(T)}^x - y)^+ - (x - y)^+ + \int_0^{\tau_k(T)} a \mathbf{1}_{[X_s > y]} dB_s - \int_0^{\tau_k(T)} b \mathbf{1}_{[X_s > y]} ds$$

and one writes down the analog of the above formula with T instead of $\tau_k(T)$. Since $|(z-y)^+ - (z'-y)^+| \leq |z-z'|$, it yields

$$J_1^2 \leq 36 (b^2 \mathbb{E}|\tau_k(T) - T|^2 + a^2 \mathbb{E}|\tau_k(T) - T|),$$

so using (1.7) we get (3.1).

• We end this computations with the estimate

$$J_2 \leq C \sqrt{\varepsilon}.$$

We recall that σ_k has the same law that $\varepsilon^2 \sigma^\varepsilon$ for $k = 0, \dots, k(T)$ and they are independent. So $\tau_k(T) - T$ has the same law as $\alpha_\varepsilon \varepsilon^2 k(T) - T + \sum_{k=0}^{k(T)-1} \varepsilon^2 (\sigma_k^\varepsilon - \alpha_\varepsilon)$ where σ_k^ε , $k = 0, \dots, k(T)$, is a sequence of independent

random variables which have the same law as σ^ε . So we may write that

$$\begin{aligned} J_2^2 &\leq \mathbb{P} \left(\alpha_\varepsilon \varepsilon^2 k(T) - T + \sum_{k=0}^{k(T)-1} \varepsilon^2 (\sigma_k^\varepsilon - \alpha_\varepsilon) \geq K \varepsilon \sqrt{-\ln \varepsilon} \right) \\ &\leq \mathbb{P} \left(\sum_{k=0}^{k(T)-1} \varepsilon^2 (\sigma_k^\varepsilon - \alpha_\varepsilon) \geq K \varepsilon \sqrt{-\ln \varepsilon} \right) \\ &\leq \mathbb{P} \left(\sum_{k=0}^{k(T)-1} \frac{T}{\alpha_\varepsilon} \frac{1}{k(T)} (\sigma_k^\varepsilon - \alpha_\varepsilon) \geq K \varepsilon \sqrt{-\ln \varepsilon} \right) \\ &= \mathbb{P} \left(\sum_{k=0}^{k(T)-1} \frac{\sigma_k^\varepsilon - \alpha_\varepsilon}{\beta_\varepsilon \sqrt{k(T)}} \geq \frac{\sqrt{k(T)} \alpha_\varepsilon}{T \beta_\varepsilon} K \varepsilon \sqrt{-\ln \varepsilon} \right). \end{aligned}$$

We denote $\gamma_\varepsilon = \mathbb{E}|\sigma^\varepsilon - \alpha_\varepsilon|^3$. We apply the Berry–Esseen theorem ([7], p. 542) and get

$$J_2^2 \leq \frac{1}{\sqrt{2\pi}} \int_{\frac{\sqrt{k(T)} \alpha_\varepsilon}{T \beta_\varepsilon} K \varepsilon \sqrt{-\ln \varepsilon}}^{\infty} \exp(-x^2/2) dx + \frac{3\gamma_\varepsilon}{\beta_\varepsilon^3 \sqrt{k(T)}}.$$

From the choice of K , the lower bound of the above integral is equal to $2\sqrt{-\ln \varepsilon}$ so

$$J_2^2 \leq \frac{1}{\sqrt{2\pi}} \int_{2\sqrt{-\ln \varepsilon}}^{\infty} \exp(-x^2/2) dx + \frac{3\gamma_\varepsilon}{\beta_\varepsilon^3 \sqrt{k(T)}}.$$

For any $\lambda \in \mathbb{R}$ one may write

$$\int_{\lambda}^{\infty} \exp(-x^2/2) dx \leq C e^{-\frac{\lambda^2}{4}}.$$

Applying this inequality with $\lambda = 2\sqrt{-\ln \varepsilon}$, we get

$$J_2^2 \leq C \varepsilon + \frac{3\gamma_\varepsilon}{\beta_\varepsilon^3 \sqrt{k(T)}} \leq C \varepsilon.$$

Now the proof of (1.11) is complete.

REFERENCES

- [1] C. Baiocchi and G.A. Pozzi, Error estimates and free-boundary convergence for a finite-difference discretization of a parabolic variational inequality. *RAIRO Anal. Numér./Numer. Anal.* **11** (1977) 315-340.
- [2] V. Bally, M.E. Caballero and B. Fernandez, Reflected BSDE's, PDE's and Variational Inequalities. *J. Theoret. Probab.* (submitted).
- [3] A. Bensoussans and J.-L. Lions, *Applications of the Variational Inequalities in Stochastic Control*. North Holland (1982).
- [4] A.N. Borodin and P. Salminen, *Handbook of Brownian Motion Facts and Formulae*. Birkhauser (1996).
- [5] M. Broadie and J. Detemple, American option valuation: New bounds, approximations, and a comparison of existing methods. *Rev. Financial Stud.* **9** (1995) 1211-1250.
- [6] N. El Karoui, C. Kapoudjan, E. Pardoux, S. Peng and M.C. Quenez, Reflected Solutions of Backward Stochastic Differential Equations and related Obstacle Problems for PDE's. *Ann. Probab.* **25** (1997) 702-737.

- [7] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. II. John Wiley and Sons (1966).
- [8] D. Lamberton, Error Estimates for the Binomial Approximation of American Put Options. *Ann. Appl. Probab.* **8** (1998) 206-233.
- [9] D. Lamberton, *Brownian optimal stopping and random walks*, Preprint 03/98. Université de Marne-la-Vallée (1998).
- [10] D. Lamberton and G. Pagès, Sur l'approximation des réduites. *Ann. Inst. H. Poincaré Probab. Statist.* **26** (1990) 331-335.
- [11] D. Lamberton and C. Rogers, *Optimal Stopping and Embedding*, Preprint 17/99. Université de Marne-la-Vallée (1999).
- [12] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*. Springer Verlag, Berlin Heidelberg (1991).
- [13] A.W. Roberts and D.E. Varberg, *Convex Functions*. Academic Press, New York (1973).
- [14] B. Sausseureau, *Sur une classe d'équations aux dérivées partielles*. Ph.D. Thesis of the University of Le Mans, France (2000).