



**HAL**  
open science

# Diffusive limit approximation of pure-jump optimal stochastic control problems

Marc Abeille, Bruno Bouchard, Lorenzo Croissant

► **To cite this version:**

Marc Abeille, Bruno Bouchard, Lorenzo Croissant. Diffusive limit approximation of pure-jump optimal stochastic control problems. 2021. hal-03269700v1

**HAL Id: hal-03269700**

**<https://hal.science/hal-03269700v1>**

Preprint submitted on 24 Jun 2021 (v1), last revised 7 Nov 2022 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Diffusive limit approximation of pure-jump optimal stochastic control problems

Marc Abeille\*, Bruno Bouchard†, Lorenzo Croissant‡

June 23, 2021

## Abstract

We consider the diffusive limit of a typical pure-jump Markovian control problem as the intensity of the driving Poisson process tends to infinity. We show that the convergence speed is provided by the Hölder constant of the Hessian of the limit problem, and explain how correction terms can be constructed. This provides an alternative efficient method for the numerical approximation of the optimal control of a pure-jump problem in situations with very high intensity of jump. We illustrate this approach in the context of a display advertising auction problem.

**Keywords:** Diffusive limit, stochastic optimal control, online auctions.

## 1 Introduction

Let  $N$  a compound Poisson process with predictable compensator  $\lambda\nu(de)dt$  and let  $X^{t,x,\alpha}$  be the solution of

$$X^{t,x,\alpha} := x + \int_t^\cdot \int b(X_{s^-}^{t,x,\alpha}, \alpha_s, e)N(de, ds),$$

---

\*Criteo AI Lab. m.abeille@criteo.com

†CEREMADE, Université Paris-Dauphine, PSL, CNRS. bouchard@ceremade.dauphine.fr.

‡CEREMADE, Université Paris-Dauphine, PSL, CNRS, and Criteo AI Lab. croissant@ceremade.dauphine.fr.

in which  $\alpha \in \mathcal{A}$  is a predictable control with values in some given set  $\mathbb{A}$ . Then, under mild assumptions, the value of the control problem

$$V(t, x) = \sup_{\alpha \in \mathcal{A}} \frac{1}{\lambda T} \mathbb{E} \left[ \int_t^T r(X_{s^-}^{t,x,\alpha}, \alpha_s) dN_s \right],$$

solves the integro-differential equation

$$\partial_t V + \sup_{a \in \mathbb{A}} \left( \int V(\cdot, \cdot + b(\cdot, a, e)) \nu(de) - V + r(\cdot, a) \right) \lambda = 0 \quad (1.1)$$

with boundary condition  $V(T, \cdot) = 0$ , possibly in the sense of viscosity solutions. From this characterization, standard numerical schemes follow, that allow one to estimate both the value function  $V$  and the associated optimal control.

However, (1.1) is non-local and obtaining a precise estimation of the solution is highly time consuming as soon as the intensity  $\lambda$  of  $N$  is large. This is for instance the case for auctions on the web, see e.g. [14], that are posted almost in continuous time, and on which one would typically like to apply reinforcement learning technics based on the resolution of (1.1) for the current estimation of the parameters, leading to a possibly large number of resolutions for different sets of parameters. On the other hand, typical statistics show that it is possible to pass to a diffusive limit, see Section 4. Namely, if one takes  $\lambda$  of the form  $\lambda = 1/\epsilon$  and  $b = \epsilon b^1 + \sqrt{\epsilon} b^2$  with  $\int b^2(\cdot, e) \nu(de) = 0$ , then a second order Taylor expansion on (1.1) implies that  $V/\lambda$  converges as  $\epsilon \rightarrow 0$  to the solution  $\bar{V}$  of

$$\partial_t \bar{V} + \sup_{\bar{a} \in \mathbb{A}} \left( \int b^1(\cdot, \bar{a}, e) \nu(de) \partial_x \bar{V} + \frac{1}{2} \int |b^2|^2(\cdot, \bar{a}, e) \nu(de) \partial_{xx}^2 \bar{V} + r(\cdot, \bar{a}) \right) = 0, \quad \bar{V}(T, \cdot) = 0. \quad (1.2)$$

The advantage of the above is that it is now a local equation which can be solved in a much more efficient way. Note that another possibility is to consider a first order expansion as in [14], which corresponds to considering a fluid limit, but is less precise.

For such a specification of the coefficients, the existence of a diffusive limit is expected, see e.g. [16] for general results on the convergence of stochastic processes. For control problems, the convergence of the value function can be proved by using the stability of viscosity solutions as in [15, Section 3], that considers the limit of discrete time zero-sum games, or by applying weak-convergence results. In particular an important literature on this subject exists in the insurance literature, see e.g. [3, 12]. However, it seems that there is no general result on the speed of convergence in the case of a (generic) optimal control problem as defined in Section 2.

In Section 3, we verify that the above intuition is correct. Unlike [15], we do not simply rely on the stability of viscosity solutions. We neither rely on the weak converge of the underlying process. The reason is that it does not give access to the convergence speed in optimal control problems. Instead,

we study directly the regularity of the solution to (1.2). Thanks to its vanishing terminal condition (otherwise it should be assumed smooth enough), we show that  $\partial_{xx}^2 \bar{V}$  is uniformly  $\beta$ -Hölder in space, for some  $\beta \in (0, 1]$ , whenever the coefficients are uniformly Lipschitz in space and under a uniform ellipticity condition. This allows us to pass from (1.2) to (1.1) up to an error term of order  $\epsilon^{\frac{\beta}{2}}$ , and therefore provides the required convergence rate. In general it can not be improved. As a by-product, we obtain an easy way to construct an  $\epsilon^{\frac{\beta}{2}}$ -optimal control for the original pure-jump control problem. We then study the limit  $\epsilon^{-\frac{\beta}{2}}(V - \bar{V})$  as  $\epsilon \rightarrow 0$ . Under mild assumptions, we show that it solves a (possibly non-linear) PDE. This provides a first error correction term. The later can be generalised upon existence to a system of non-linear PDEs to achieve higher orders of convergence.

As an example of application, we consider in Section 4 a simplified repeated online auction bidding problem, where a buyer seeks to maximise his profit when facing both competition and a seller who adapts his price to incoming bids. Even for values of  $\epsilon$  of 8 to 10 orders of magnitude larger than in reality, the numerical resolution of (1.1) was up to 6 times slower than the resolution of the diffusive limite PDE.

## 2 The pure-jump optimal control problem

In this section, we begin by providing the defining of our pure-jump control problem, and state the well-known link with the associated Hamilton-Jacobi-Bellman equation. The properties stated below and elementary but will be useful for the derivation of our main approximation result of Section 3.

### 2.1 Definition

Let  $\mathbb{D}$  denote the space of one dimensional càdlàg processes and  $\mathcal{M}(\mathbb{R}_+ \times \mathbb{R})$  denote the collection of positive finite measures on  $\mathbb{R}_+ \times \mathbb{R}$ . Consider a random measure  $N : \mathbb{D} \mapsto \mathcal{M}(\mathbb{R}_+ \times \mathbb{R})$  and let  $\mathbb{P}$  be a probability measure on  $\mathbb{D}$  under which  $N$  admits the compensator  $\lambda \nu(de)dt$ , in which  $\lambda > 0$  and  $\nu$  is a probability measure on  $\mathbb{R}$ . For ease of notations, we set  $N_t := N(\mathbb{R}, [0, t])$  for  $t \geq 0$ .

Let  $\mathbb{F}^t = (\mathcal{F}_s^t)_{s \geq t}$  be the  $\mathbb{P}$ -augmentation of the filtration generated by  $(N_s - N_t)_{s \geq t}$ . Given a compact subset  $\mathbb{A}$  of  $\mathbb{R}$ , we let  $\mathcal{A}^t$  be the collection of  $\mathbb{F}^t$ -predictable processes with values in  $\mathbb{A}$ . For ease of notations, we also define  $\mathcal{A} := \cup_{t \geq 0} \mathcal{A}^t$ . Through this paper, unless otherwise stated we will work in the filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ .

We now consider a bounded measurable map  $(x, a, e) \in \mathbb{R} \times \mathbb{A} \times \mathbb{R} \mapsto b(x, a, e)$ . Given  $(t, x) \in \mathbb{R}_+ \times \mathbb{R}$

and  $\alpha \in \mathcal{A}$ , we define the càdlàg process  $X^{t,x,\alpha}$  as the solution of

$$X^{t,x,\alpha} := x + \int_t^\cdot \int b(X_{s^-}^{t,x,\alpha}, \alpha_s, e) N(de, ds). \quad (2.1)$$

Given  $T > 0$ , a bounded measurable map  $(x, a) \in \mathbb{R} \times \mathbb{A} \mapsto r(x, a)$ , we consider the expected gain function

$$(t, x, \alpha) \in \mathbb{R}_+ \times \mathbb{R} \times \mathcal{A} \mapsto J(t, x; \alpha) := \mathbb{E} \left[ \int_t^T r(X_{s^-}^{t,x,\alpha}, \alpha_s) dN_s \right], \quad (2.2)$$

together with the value function

$$V(t, x) = \sup_{\alpha \in \mathcal{A}^t} J(t, x; \alpha). \quad (2.3)$$

All throughout the paper, we make the following standard assumption, which will in particular ensure that  $V$  is the unique (bounded) viscosity solution of the associated Hamilton-Jacobi-Bellman equation, see Proposition 2.2 below.

**Assumption 1.** *For each  $e \in \mathbb{R}$ ,  $(x, a) \in \mathbb{R} \times \mathbb{A} \mapsto (b(x, a, e), r(x, a))$  is continuous.*

**Remark 2.1.** *Note that boundedness of the coefficients is not essential in the following arguments. One could simply assume linear growth, uniformly in the control. We make the above (strong) assumptions to avoid unnecessary complexities.*

## 2.2 Dynamic programming equation and optimal Markovian control

Let us now state the well-known characterization of  $V$  in terms of viscosity solution.

As usual, we say that a lower-semicontinuous (resp. upper-semicontinuous) locally bounded map  $U : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$  is a viscosity super-solution (resp. subsolution) of

$$\partial_t \varphi + \sup_{a \in \mathbb{A}} \left( \int \varphi(\cdot, \cdot + b(\cdot, a, e)) \nu(de) - \varphi + r(\cdot, a) \right) \lambda = 0, \quad \text{on } [0, T] \times \mathbb{R}, \quad (2.4)$$

if for all  $(t, x) \in [0, T] \times \mathbb{R}$  and all  $C^1$ -function  $\varphi : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$  such that  $(t, x)$  reaches a minimum (resp. maximum) of  $U - \varphi$  on  $[0, T] \times \mathbb{R}$  we have

$$\delta \left\{ \partial_t \varphi(t, x) + \sup_{a \in \mathbb{A}} \left( \int U(t, x + b(x, a, e)) \nu(de) - U(t, x) + r(x, a) \right) \lambda \right\} \leq 0$$

with  $\delta = 1$  (resp.  $\delta = -1$ ).

**Proposition 2.2.**  $V$  is a continuous and bounded viscosity solution of (2.4) such that

$$\lim_{t' \uparrow T, x' \rightarrow x} V(t', x') = 0, \quad x \in \mathbb{R}. \quad (2.5)$$

Moreover, comparison holds for (2.4) in the class of bounded functions.

*Proof.* The argument being standard, we only sketch it. Fix  $h \in (0, T - t]$ . Let  $\tau_1^t$  be the first jump of  $N$  after time  $t$ . Denote by  $V_*$  and  $V^*$  the lower- and upper-semicontinuous envelopes of  $V$ , i.e.

$$V_*(t', x') := \liminf_{(s, y) \rightarrow (t', x')} V(s, y), \quad V^*(t', x') := \limsup_{(s, y) \rightarrow (t', x')} V(s, y).$$

It follows from the same arguments as in [8] that  $V$  satisfies the (weak) dynamic programming principle

$$\begin{aligned} & \sup_{\alpha \in \mathcal{A}^t} \mathbb{E} \left[ V_*(\tau_1^t \wedge h, X_{\tau_1^t \wedge h}^{t, x, \alpha}) + r(X_{\tau_1^t -}^{t, x, \alpha}, \alpha_{\tau_1^t}) 1_{\{\tau_1^t \leq h\}} \right] \\ & \leq V(t, x) \\ & \leq \sup_{\alpha \in \mathcal{A}^t} \mathbb{E} \left[ V^*(\tau_1^t \wedge h, X_{\tau_1^t \wedge h}^{t, x, \alpha}) + r(X_{\tau_1^t -}^{t, x, \alpha}, \alpha_{\tau_1^t}) 1_{\{\tau_1^t \leq h\}} \right]. \end{aligned} \quad (2.6)$$

Following [8] again and using [6, Lemma 22], this implies that  $V_*$  and  $V^*$  are respectively a viscosity super- and a subsolution of (2.4). Since  $r$  and  $b$  are bounded, the map  $(t, x) \mapsto (1 + x^2)e^{-Ct}$  is also a viscosity supersolution of the above, as soon as  $C > 0$  is large enough. Standard arguments then imply that comparison holds for the above Hamilton-Jacobi-Bellman equation in the class of bounded functions (or even with linear growth), and therefore that  $V_* = V^*$ , meaning that  $V$  is continuous. The continuity at  $T$  follows immediately from the fact that  $r$  is bounded.  $\square$

We next prove the existence of an optimal Markovian control. In the following, we denote by  $\mathbb{A}$  the collection  $\mathbb{A}$ -valued Borel maps on  $\mathbb{R}_+ \times \mathbb{R}$ .

**Proposition 2.3.** For all  $(t, x) \in \mathbb{R}_+ \times \mathbb{R}$ , there exists  $\hat{\alpha}[t, x] \in \mathcal{A}^t$  such that  $V(t, x) = J(t, x; \hat{\alpha}[t, x])$ . It takes the form

$$\hat{\alpha}[t, x] = \sum_{i \geq 0} 1_{(\tau_i^t, \tau_{i+1}^t]} \hat{\alpha}(\cdot, X_{\tau_i^t}^{t, x, \hat{\alpha}[t, x]})$$

in which  $\tau_0^t \equiv t$ ,  $\tau_i^t$  is the  $i$ -th jump of  $N$  after time  $t$  for  $i \geq 1$ , and  $\hat{\alpha} \in \mathbb{A}$  satisfies

$$\hat{\alpha}(t', x') \in \operatorname{argmax}_{a \in \mathbb{A}} \left\{ \int V(t', x' + b(x', a, e)) \nu(de) + r(x', a) \right\}, \quad (t', x') \in \mathbb{R}_+ \times \mathbb{R}.$$

*Proof.* Since  $V, b$  and  $r$  are continuous, by Proposition 2.2 and Assumption 1, and since  $\mathbb{A}$  is compact, we can find a Borel measurable map  $(t, x) \mapsto \hat{a}(t, x)$  such that  $\hat{a}(t, x)$  belongs to  $\operatorname{argmax}\{\int V(t, x + b(x, a, e))\nu(de) + r(x, a), a \in \mathbb{A}\}$  for all  $(t, x)$ , see e.g. [4, Proposition 7.33, p.153]. Let us fix  $(t_0, x_0) \in [0, T] \times \mathbb{R}$ . By the dynamic programming in (2.6), the continuity of  $V$ , and the definition of  $\hat{a}$  above,

$$\begin{aligned} V(t_0, x_0) &= \sup_{\alpha \in \mathcal{A}^{t_0}} \mathbb{E} \left[ V(\tau_1^{t_0} \wedge T, X_{\tau_1^{t_0} \wedge T}^{t_0, x_0, \alpha}) + r(X_{\tau_1^{t_0} \wedge T}^{t_0, x_0, \alpha}, \alpha_{\tau_1^{t_0}}) 1_{\{\tau_1^{t_0} \leq T\}} \right] \\ &= \sup_{\alpha \in \mathcal{A}^{t_0}} \mathbb{E} \left[ \left( \int V(\tau_1^{t_0} \wedge T, x_0 + b(x_0, \alpha_{\tau_1^{t_0} \wedge T}, e)) \nu(de) + r(x_0, \alpha_{\tau_1^{t_0}}) \right) 1_{\{\tau_1^{t_0} \leq T\}} \right] \\ &= \mathbb{E} \left[ \left( \int V(\tau_1^{t_0} \wedge T, x_0 + b(x_0, \hat{\alpha}_{\tau_1^{t_0} \wedge T}, e)) \nu(de) + r(x_0, \hat{\alpha}_{\tau_1^{t_0}}) \right) 1_{\{\tau_1^{t_0} \leq T\}} \right] \\ &= \mathbb{E} \left[ V(\tau_1^{t_0} \wedge T, X_{\tau_1^{t_0} \wedge T}^{t_0, x_0, \hat{\alpha}}) + r(X_{\tau_1^{t_0} \wedge T}^{t_0, x_0, \hat{\alpha}}, \hat{\alpha}_{\tau_1^{t_0}}) 1_{\{\tau_1^{t_0} \leq T\}} \right] \end{aligned}$$

in which  $\hat{\alpha} = \hat{a}(\cdot, x_0) 1_{(t_0, \tau_1^{t_0}]}$ .

For ease of notations, we now set  $\vartheta_1 := \tau_1^{t_0} \wedge T$  and  $X_1 := X_{\tau_1^{t_0} \wedge T}^{t_0, x_0, \hat{\alpha}}$ . By the same reasoning as above, we have, for a fixed  $\omega \in \Omega$ ,

$$V(\vartheta_1(\omega), X_1(\omega)) = \mathbb{E} \left[ V(\tau_1^{\vartheta_1(\omega)} \wedge T, X_{\tau_1^{\vartheta_1(\omega)} \wedge T}^{\vartheta_1(\omega), X_1(\omega), \hat{\alpha}(\omega)}) + r(X_{\tau_1^{\vartheta_1(\omega)} \wedge T}^{\vartheta_1(\omega), X_1(\omega), \hat{\alpha}(\omega)}, \hat{\alpha}_{\tau_1^{\vartheta_1(\omega)}(\omega)}(\omega)) 1_{\{\tau_1^{\vartheta_1(\omega)} \leq T\}} \right]$$

in which

$$\hat{\alpha}(\omega) = \hat{a}(\cdot, x_0) 1_{(t_0, \tau_1^{t_0}(\omega)]} + \hat{a}(\cdot, X_1(\omega)) 1_{(\tau_1^{t_0}(\omega), \tau_1^{\vartheta_1(\omega)})}.$$

The right-hand side of the above coincides  $\mathbb{P}$ -a.e. with

$$\begin{aligned} &\mathbb{E} \left[ V(\tau_1^{\vartheta_1} \wedge T, X_{\tau_1^{\vartheta_1} \wedge T}^{\vartheta_1, X_1, \hat{\alpha}}) + r(X_{\tau_1^{\vartheta_1} \wedge T}^{\vartheta_1, X_1, \hat{\alpha}}, \hat{\alpha}_{\tau_1^{\vartheta_1}}) 1_{\{\tau_1^{\vartheta_1} \leq T\}} \middle| \mathcal{F}_{\vartheta_1} \right] \\ &= \mathbb{E} \left[ V(\tau_2^{t_0} \wedge T, X_{\tau_2^{t_0} \wedge T}^{t_0, x_0, \hat{\alpha}}) + r(X_{\tau_2^{t_0} \wedge T}^{t_0, x_0, \hat{\alpha}}, \hat{\alpha}_{\tau_2^{t_0}}) 1_{\{\tau_2^{t_0} \leq T\}} \middle| \mathcal{F}_{\tau_1^{t_0} \wedge T} \right]. \end{aligned}$$

Let us now complete the definition of  $\hat{\alpha}$  by letting

$$\hat{\alpha} = \sum_{i \geq 0} 1_{(\tau_i^{t_0}, \tau_{i+1}^{t_0}]} \hat{a}(\cdot, X_{\tau_i^{t_0}}^{t_0, x_0, \hat{\alpha}}).$$

By iterating the above procedure, we have

$$V(t_0, x_0) = \mathbb{E} \left[ V(\tau_n^{t_0} \wedge T, X_{\tau_n^{t_0} \wedge T}^{t_0, x_0, \hat{\alpha}}) + \int_{t_0}^{\tau_n^{t_0} \wedge T} r(X_{s-}^{t_0, x_0, \hat{\alpha}}, \hat{\alpha}_s) dN_s \right], \quad n \geq 1.$$

Since  $\tau_n^{t_0} \rightarrow \infty$   $\mathbb{P}$ -a.s. as  $n \rightarrow \infty$ , it now follows from the dominated convergence theorem that

$$V(t_0, x_0) = \mathbb{E} \left[ \int_{t_0}^T r(X_{s-}^{t_0, x_0, \hat{\alpha}}, \hat{\alpha}_s) dN_s \right].$$

□

**Remark 2.4.** *In the example of application of Section 4 below, the function  $r$  is of the form  $r(x, a) = 1_{\{a \geq x\}} \rho(x, a)$  in which  $\rho$  is a continuous function and  $\mathbb{A} = \mathbb{R}_+$ . Although  $r$  is not continuous, this does not alter the result of Proposition 2.2 because  $\sup_{a \in \mathbb{A}} \{\varphi(t, x, a) + 1_{\{a \geq x\}} \rho(x, a)\} = \sup_{a \in \mathbb{A}} \{\varphi(t, x, a) + 1_{\{a > x\}} \rho(x, a)\}$  as soon as  $\varphi$  is continuous in its  $a$ -argument. This applies to  $(t', x', a) \mapsto \varphi(t', x', a) := \int V(t', x' + b(x', a, e)) \nu(de)$ . The argmax in Proposition 2.3 is also achieved as soon as  $\rho \geq 0$ . If  $\rho$  can take negative values, then the max is not achieved at points  $(t, x)$  in  $E := \{(t', x') : \sup_{a \in \mathbb{A}} \{\varphi(t', x', a) + 1_{\{a \geq x'\}} \rho(x', a)\} = \varphi(t', x', x')\}$ , but in this case one can construct a sequence of  $\iota$ -optimal controls, for all  $\iota > 0$ . For instance, if  $\varphi$  and the coefficients are uniformly Lipschitz in  $a$ , we can choose them of the form  $\hat{\alpha}^\iota(t, x) = \hat{\alpha}(t, x) 1_{\{(t, x) \in E^c\}} + (x - \iota)$ , for all  $\iota > 0$ . The same analysis holds for the results presented in Section 3 below.*

### 3 Diffusive approximation

As already mentioned, the characterization of Propositions 2.2 and 2.3 allows one to estimate numerically the value function and the associated optimal control. However, the intro-differential equation 2.4 is non-local and its resolution turns out to be highly time consuming as soon as the intensity  $\lambda$  is high. On the other hand, we can expect that our pure-jump problem admits a diffusive limit as  $\lambda \rightarrow \infty$  which, by nature, is much easier to solve numerically, and can serve as a good proxy of the original problem as soon as  $\lambda$  is large enough.

In this section, we begin by defining the diffusion control problem that is candidate to be the diffusive limit of our pure-jump problem. We then study the regularity of the corresponding value function, from which we will be able to derive our main approximation result, see Theorem 3.3 below, and construct approximate optimal controls, see Proposition 3.4. Finally, we identify a first order correction term in Subsection 3.5, which is extended to higher orders in Subsection 3.6.

#### 3.1 The candidate diffusive limit

Given  $\epsilon \in (0, 1)$ , we now replace the intensity  $\lambda$  by

$$\lambda_\epsilon := \epsilon^{-1}$$



so that it is large for  $\epsilon$  small. To ensure the existence of a diffusive limit, we need to assume that the drift  $b$  introduced in Section 2 is of the form

$$b^\epsilon = \epsilon b^1 + \sqrt{\epsilon} b^2$$

for two bounded measurable maps  $b^1, b^2 : \mathbb{R} \times \mathbb{A} \times \mathbb{R} \mapsto \mathbb{R}$ , each satisfying Assumption 1 (with  $b^i$  in place of  $b$ ,  $i = 1, 2$ ), and with  $b^2$  satisfying the additional Assumption 2.

**Assumption 2.** *The function  $b^2$  satisfies:*

$$\int b^2(x, a, e) \nu(de) = 0 \text{ for all } (x, a) \in \mathbb{R} \times \mathbb{A}, \text{ and } \inf_{(x, a) \in \mathbb{R} \times \mathbb{A}} \int |b^2(x, a, e)|^2 \nu(de) =: \eta > 0. \quad (3.1)$$

In the above, the coefficient  $b^1$  should be interpreted as a drift term while  $b^2$  is a volatility. The respective scaling in  $\epsilon$  and  $\sqrt{\epsilon}$ , together with Assumption 2 are required to ensure that our pure-jump problem actually admits a diffusive limit of the form (3.2) below.

Likewise, we consider the value function

$$V^\epsilon(t, x) := \sup_{\alpha \in \mathcal{A}^t} \frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s^-}^{t, x, \alpha}, \alpha_s) dN_s \right], \quad (t, x) \in [0, T] \times \mathbb{R}.$$

Note that the scaling by  $1/\lambda_\epsilon$  means that (up to a constant factor  $T - t$ ) we consider the gain by average unit of actions on the system. Indeed  $\mathbb{E}[N_T - N_t] = \lambda_\epsilon(T - t)$  and the control applies only at jump times of  $N$ . Note that we omit the dependence of  $N$  on  $\epsilon$ , for ease of notations.

We shall see that the above, together with the associated optimal policy, can be approximated by considering its diffusive limit as  $\epsilon \rightarrow 0$ . The coefficients of the associated Brownian SDE are given by:

$$\mu(x, a) := \int b^1(x, a, e) \nu(de), \quad \sigma(x, a) := \left( \int |b^2(x, a, e)|^2 \nu(de) \right)^{\frac{1}{2}}, \quad (x, a) \in \mathbb{R} \times \mathbb{A}.$$

From now on, we assume that they satisfy the following.

**Assumption 3.** *The maps  $x \in \mathbb{R} \mapsto \mu(x, a)$ ,  $x \in \mathbb{R} \mapsto \sigma(x, a)$  and  $x \in \mathbb{R} \mapsto r(x, a)$  are Lipschitz, uniformly in  $a \in \mathbb{A}$ , with respective Lipschitz constants  $\|\mu\|_{\text{Lip}}$ ,  $\|\sigma\|_{\text{Lip}}$  and  $\|r\|_{\text{Lip}}$ .*

More precisely, let  $W$  be a  $\bar{\mathbb{P}}$ -Brownian motion defined on  $\mathbb{D}$ , let  $\bar{\mathbb{F}}^t = (\bar{\mathcal{F}}_s^t)_{s \geq 0}$  be the  $\bar{\mathbb{P}}$ -augmentation of the filtration generated by  $(W_{\cdot \vee t} - W_t)$ , and let  $\bar{\mathcal{A}}^t$  the collection of  $\bar{\mathbb{F}}^t$ -predictable processes. Given  $\bar{\alpha} \in \bar{\mathcal{A}}^t$ , we can define  $\bar{X}^{t, x}$  as the unique strong solution of

$$\bar{X}^{t, x, \bar{\alpha}} = x + \int_t^\cdot \mu(\bar{X}_s^{t, x, \bar{\alpha}}, \alpha_s) ds + \int_t^\cdot \sigma(\bar{X}_s^{t, x, \bar{\alpha}}, \bar{\alpha}_s) dW_s. \quad (3.2)$$

The candidate diffusive limit problem is then defined as

$$\bar{V}(t, x) := \sup_{\bar{\alpha} \in \bar{\mathcal{A}}^t} \bar{\mathbb{E}} \left[ \int_t^T r(\bar{X}_s^{t,x,\bar{\alpha}}, \bar{\alpha}_s) ds \right]$$

where  $\bar{\mathbb{E}}$  is the expectation operator under  $\bar{\mathbb{P}}$ .

### 3.2 Regularity properties

We first prove that  $\bar{V}$  is a smooth solution of the associated Hamilton-Jacobi-Bellman equation. Most importantly, its second order space derivative is  $\beta$ -Hölder continuous, for some  $\beta \in (0, 1]$ . This will allow us, in Section 3.3 below, to prove that it actually coincides with the diffusive limit of  $V^\epsilon$  as  $\epsilon$  vanishes. The precise value of the Hölder exponent  $\beta$  will be further discussed in Remark 3.2 below.

**Proposition 3.1.**  *$\bar{V}$  belongs to  $C_b^{1,2}([0, T] \times \mathbb{R})$  and is the unique bounded solution of*

$$\partial_t \bar{V} + \sup_{\bar{a} \in \bar{\mathbb{A}}} \left( \mu(\cdot, \bar{a}) \partial_x \bar{V} + \frac{1}{2} \sigma^2(\cdot, \bar{a}) \partial_{xx}^2 \bar{V} + r(\cdot, \bar{a}) \right) = 0, \text{ on } [0, T] \times \mathbb{R}, \quad (3.3)$$

$$\bar{V}(T, \cdot) = 0, \text{ on } \mathbb{R}. \quad (3.4)$$

Moreover, there exists  $\beta \in (0, 1]$ , such that  $\partial_{xx}^2 \bar{V}$  is  $\beta$ -Hölder continuous in space.

*Proof.* a) We first show that  $\bar{V} \in C_b^{1,2}([0, T] \times \mathbb{R}) \cap C^0([0, T] \times \mathbb{R})$ . Let us set

$$F(x, p, q) := \sup_{\bar{a} \in \bar{\mathbb{A}}} \left( \mu(x, \bar{a}) p + \frac{1}{2} \sigma^2(x, \bar{a}) q + r(x, \bar{a}) \right), \quad (x, p, q) \in \mathbb{R}^3,$$

and observe that, by Assumptions 2 and 3,

$$\frac{1}{2} \eta |q - q'| \leq |F(x, p, q) - F(x, p, q')| \leq \frac{1}{2} \|\sigma\|_\infty^2 |q - q'| \quad (3.5)$$

$$v F(x, 0, 0) \leq \|r\|_\infty (1 + |v|^2) \quad (3.6)$$

$$\begin{aligned} |F(x, p, q) - F(x', p', q)| &\leq (|p| \|\mu\|_{\text{Lip}} + |q| \|\sigma\|_\infty \|\sigma\|_{\text{Lip}} + \|r\|_{\text{Lip}}) |x - x'| \\ &\quad + \|\mu\|_\infty |p - p'| + \frac{1}{2} \|\sigma\|_\infty^2 |q - q'| \end{aligned} \quad (3.7)$$

for all  $(x, x', p, p', q, q', v) \in \mathbb{R}^7$ .

For  $n \geq 1$ , existence of a  $C^{1,2}([0, T] \times \mathbb{R})$  solution  $\bar{V}_n$  to (3.3) on  $[0, T] \times (-n, n)$  with boundary condition  $\bar{V}_n = 0$  on  $([0, T] \times \{-n, n\}) \cup (\{T\} \times [-n, n])$  follows from [18, Theorem 14.24], (3.5), (3.6) and (3.7), upon mollifying  $F$  and then passing to the limit.

b) We now first provide uniform estimates on the gradients. Note that

$$\bar{V}_n(t, x) = \sup_{\bar{\alpha} \in \bar{\mathcal{A}}^t} \bar{\mathbb{E}} \left[ \int_t^{T \wedge \tau_n^{t, x, \bar{\alpha}}} r(\bar{X}_s^{t, x, \bar{\alpha}}, \bar{\alpha}_s) ds \right]$$

where

$$\tau_n^{t, x, \bar{\alpha}} := \inf\{s \geq t : \bar{X}_s^{t, x, \bar{\alpha}} \notin (-n, n)\}.$$

It follows that, for  $h \in (0, T - t)$ ,

$$\bar{V}_n(t + h, x) = \sup_{\bar{\alpha} \in \bar{\mathcal{A}}^t} \bar{\mathbb{E}} \left[ \int_t^{(T-h) \wedge \tau_n^{t, x, \bar{\alpha}}} r(\bar{X}_s^{t, x, \bar{\alpha}}, \alpha_s) ds \right]$$

which readily implies that  $|\bar{V}_n(t + h, x) - \bar{V}_n(t, x)| \leq h \|r\|_\infty$ , and therefore

$$\|\partial_t \bar{V}_n\| \leq \|r\|_\infty. \quad (3.8)$$

Similarly, for  $h \in (-1, 1)$  such that  $x + h \in [-n, n]$ ,

$$|\bar{V}_n(t, x + h) - \bar{V}_n(t, x)| \leq \|r\|_{\text{Lip}} \sup_{\bar{\alpha} \in \bar{\mathcal{A}}^t} \bar{\mathbb{E}} \left[ \int_t^T |\bar{X}_s^{t, x+h, \bar{\alpha}} - \bar{X}_s^{t, x, \bar{\alpha}}| ds + |\tau_n^{t, x+h, \bar{\alpha}} - \tau_n^{t, x, \bar{\alpha}}| \right].$$

The first term is handled by using the uniform Lipschitz continuity in space of  $(\mu, \sigma)$ :

$$\bar{\mathbb{E}} \left[ \int_t^T |\bar{X}_s^{t, x+h, \bar{\alpha}} - \bar{X}_s^{t, x, \bar{\alpha}}| ds \right] \leq C_1 |h| \quad (3.9)$$

in which  $C_1 > 0$  does not depend on  $n$ . As for the second term, Assumption 3 and our boundedness assumptions allow us to apply [7, Theorem 2.2] with  $\pi = 0$ ,  $r = n$  and for  $P$  of the form  $\varphi(\bar{X}_s^{t, x+h, \bar{\alpha}})$  or  $\varphi(\bar{X}_s^{t, x, \bar{\alpha}})$  for a smooth function  $\varphi$  such that  $\varphi(y) = y + n$  for  $y \in [-n, -\sqrt{n}]$  and  $\varphi(y) = n - y$  for  $y \in [\sqrt{n}, n]$ . It implies that

$$\bar{\mathbb{E}} \left[ |\tau_n^{t, x+h, \bar{\alpha}} - \tau_n^{t, x, \bar{\alpha}}| \right] \leq C_2 \bar{\mathbb{E}} \left[ \left| \bar{X}_{\tau_n^{t, x+h, \bar{\alpha}} \wedge \tau_n^{t, x, \bar{\alpha}}}^{t, x+h, \bar{\alpha}} - \bar{X}_{\tau_n^{t, x+h, \bar{\alpha}} \wedge \tau_n^{t, x, \bar{\alpha}}}^{t, x, \bar{\alpha}} \right| \right] \leq C'_2 |h|$$

for some positive constants  $C_2$  and  $C'_2$  independent of  $n$ . Combined with (3.9), this leads to

$$\|\partial_x \bar{V}_n\|_\infty \leq C_1 + C'_2. \quad (3.10)$$

The fact that  $\bar{V}_n$  solves (3.3) combined with (3.5), (3.8) and (3.10) then proves that

$$\|\partial_{xx}^2 \bar{V}_n\|_\infty \leq C_3 \quad (3.11)$$

for some  $C_3 > 0$  that does not depend on  $n$ .

c) We now prove the uniform Hölder continuity of the gradients and second derivatives. Upon mollifying  $F$ , we can assume that it is  $C^1$ . Given a neighbourhood  $\mathcal{O} \subset [0, T] \times [-n, n]$  of a point  $(t, x)$ , we derive as in [1, Section 3.1] that there exists  $C > 0$  and  $\beta \in (0, 1]$ , that depend only on the ellipticity constant  $\eta$  and the Lipschitz constants of  $F$  with respect to its second and third arguments, such that

$$|\partial_t \bar{V}_n(t', x') - \partial_t \bar{V}_n(t, x)| \leq C \left( |t' - t|^{\frac{\beta}{2}} + |x' - x|^\beta \right) \sup_{\mathcal{O}} \|\partial_t \bar{V}_n\|, \quad \text{for } (t', x') \in \mathcal{O}.$$

In view of (3.8), this implies that

$$|\partial_t \bar{V}_n(t', x') - \partial_t \bar{V}_n(t, x)| \leq C \left( |t' - t|^{\frac{\beta}{2}} + |x' - x|^\beta \right) \|r\|_\infty, \quad \text{for } (t', x') \in [0, T] \times \mathbb{R}. \quad (3.12)$$

Up to changing  $\beta \in (0, 1)$ , one can prove similarly that

$$|\partial_x \bar{V}_n(t', x') - \partial_x \bar{V}_n(t, x)| \leq C \left( |t' - t|^{\frac{\beta}{2}} + |x' - x|^\beta \right), \quad \text{for } (t', x') \in [0, T] \times \mathbb{R}, \quad (3.13)$$

for some  $C > 0$  that does not depend on  $n$ . We now set  $\Delta^h \bar{V}_n := h^{-\beta} (\bar{V}_n(\cdot, \cdot + h) - \bar{V}_n)$ ,  $h \in \mathbb{R}$ . It satisfies a linearized equation of the form

$$0 = \partial_t \Delta^h \bar{V}_n + A^h \partial_x \Delta^h \bar{V}_n + B^h \partial_{xx}^2 \Delta^h \bar{V}_n + C^h h^{1-\beta}$$

at every point  $(t, x) \in [0, T] \times \mathbb{R}$  such that  $x + h \in (-n, n)$ , in which, by Assumption 3 and the estimates in b) above,  $(A^h, C^h)_{h>0}$  is uniformly bounded and  $\inf_{h>0} \inf_{[0, T] \times \mathbb{R}} B^h \geq \eta/2$ . Hence,

$$|\partial_{xx}^2 \Delta^h \bar{V}_n| \leq 2\eta^{-1} (|\partial_t \Delta^h \bar{V}_n| + |A^h| |\partial_x \Delta^h \bar{V}_n| + |C^h| h^{1-\beta})$$

and we conclude from (3.12)-(3.13) that

$$|\partial_{xx}^2 \bar{V}_n(t, x') - \partial_{xx}^2 \bar{V}_n(t, x)| \leq C |x' - x|^\beta, \quad x, x' \in (-n, n), \quad t < T \quad (3.14)$$

for some  $C > 0$  independent on  $n$ . If we now set  $\Delta^h \bar{V}_n = h^{-\frac{\beta}{2}} (\bar{V}_n(\cdot + h, \cdot) - \bar{V}_n)$ , then the same type of arguments leads to

$$|\partial_{xx}^2 \bar{V}_n(t', x) - \partial_{xx}^2 \bar{V}_n(t, x)| \leq C |t' - t|^{\frac{\beta}{2}}, \quad x \in (-n, n), \quad t, t' < T \quad (3.15)$$

for some  $C > 0$  independent on  $n$ .

d) It follows from steps b) and c) that  $(\bar{V}_n)_{n \geq 1}$  is uniformly bounded in  $H_{2+\beta}([0, T] \times \mathbb{R})$ , as defined in [18, Section IV.1]. By the Arzelà-Ascoli theorem, it admits a subsequence that converges in  $H_{2+\beta}(B)$ , for any compact set  $B \subset [0, T] \times \mathbb{R}$ . We can pass to the limit in the equation, so that the limit solves (3.3)-(3.4). By comparison, the limit is  $\bar{V}$  and it shares the same upper-bound in  $H_{2+\beta}([0, T] \times \mathbb{R})$  as  $(\bar{V}_n)_{n \geq 1}$ .

□

**Remark 3.2.** (a) Let  $\bar{a} : [0, T] \times \mathbb{R} \mapsto \mathbb{A}$  be a measurable map satisfying

$$\bar{a} \in \operatorname{argmax}_{a \in \mathbb{A}} \left( \mu(\cdot, a) \partial_x \bar{V} + \frac{1}{2} \sigma^2(\cdot, a) \partial_{xx}^2 \bar{V} + r(\cdot, a) \right) \quad \text{on } [0, T] \times \mathbb{R}, \quad (3.16)$$

see e.g. [4, Proposition 7.33, p.153]. Assume that there exists  $\beta_o \in (0, 1)$  such that  $r(\cdot, \bar{a})$  belongs to  $H_{\beta_o}([0, T] \times \mathbb{R})$ , then we can take  $\beta = \beta_o$ . This follows from [17, Section IV.14, p390].

- (b) If  $(\mu(\cdot, \bar{a}), \sigma(\cdot, \bar{a}), r(\cdot, \bar{a}))$  has more regularity, one can obviously obtain more regularity on  $\bar{V}$  by, for instance, differentiating the associated partial differential equation.
- (c) In the case where  $\sigma$  does not depend on its  $a$ -argument, then one can appeal to [18, Theorem 12.16] to deduce that we can take  $\beta = 1$ . This follows from the Lipschitz continuity of  $F$ .

### 3.3 Convergence speed toward the diffusive limit

We now exploit the Hölder regularity stated above to prove that  $V^\epsilon$  converges to  $\bar{V}$  at a rate  $\epsilon^{\frac{\beta}{2}}$  as  $\epsilon$  vanishes. We shall see in Section 3.4 below that it provides an  $\epsilon^{\frac{\beta}{2}}$ -optimal control for the pure-jump problem. In general, it can not be improved, see Section 3.5.

**Theorem 3.3.** For all  $(t, x) \in [0, T] \times \mathbb{R}$ ,

$$|V^\epsilon - \bar{V}|(t, x) \leq \sup_{\alpha \in \mathcal{A}^t} \mathbb{E} \left[ \int_t^T \int \delta r^\epsilon(X_{s^-}^{t,x,\alpha}, \alpha_s, e) \nu(\mathrm{d}e) \mathrm{d}s \right]$$

in which

$$\delta r^\epsilon := \epsilon^{-1} (\bar{V}(\cdot, \cdot + b^\epsilon) - \bar{V}) - b^1 \partial_x \bar{V} - \frac{1}{2} |b^2|^2 \partial_{xx}^2 \bar{V} \quad (3.17)$$

satisfies

$$\|\delta r^\epsilon\|_\infty \leq C_K^\epsilon \epsilon^{\frac{\beta}{2}} \quad (3.18)$$

with

$$C_K^\epsilon := \frac{1}{2} \left( \epsilon^{1-\frac{\beta}{2}} \|b^1\|_\infty^2 + 2\epsilon^{\frac{1-\beta}{2}} \|b^1 b^2\|_\infty \right) \left( \|\partial_{xx} \bar{V}\|_\infty + K(\epsilon \|b^1\|_\infty + \sqrt{\epsilon} \|b^2\|_\infty)^\beta \right) + \frac{1}{2} K(\sqrt{\epsilon} \|b^1\|_\infty + \|b^2\|_\infty)^{2+\beta},$$

where  $K > 0$  is the Hölder constant of  $\partial_{xx}^2 \bar{V}$  with respect to its space variable.

In particular,

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\beta}{2}} \|V^\epsilon(t, \cdot) - \bar{V}(t, \cdot)\|_\infty \leq \frac{1}{2} (T - t) K (\|b^2\|_\infty)^{2+\beta}, \quad t < T.$$

*Proof.* Since  $\bar{V} \in C_b^{1,2}([0, T] \times \mathbb{R})$  and  $\partial_{xx}^2 \bar{V}$  is  $\beta$ -Hölder continuous in space,

$$\begin{aligned} I &:= \int (\bar{V}(t, x + b^\epsilon(x, a, e)) - \bar{V}(t, x)) \nu(de) \\ &= \int \left( b^\epsilon(x, a, e) \partial_x \bar{V}(t, x) + \frac{1}{2} |b^\epsilon|^2(x, a, e) (\partial_{xx}^2 \bar{V}(t, x) + O(|b^\epsilon(x, a, e)|)) \right) \nu(de) \end{aligned}$$

in which  $O$  is a continuous function satisfying  $|O(x')| \leq K |x'|^\beta$  for  $x' \in \mathbb{R}$ . By the left-hand side of Assumption 2, the above simplifies to

$$I = \epsilon \int \left( b^1(x, a, e) \partial_x \bar{V}_\theta(t, x) + \frac{1}{2} |b^2|^2(x, a, e) \partial_{xx}^2 \bar{V}_\theta(t, x) + \delta r^\epsilon(x, a, e) \right) \nu(de),$$

where  $\delta r^\epsilon$ , defined in (3.17), is a continuous function satisfying

$$\begin{aligned} |\delta r^\epsilon| &\leq \frac{1}{2} \left( \epsilon \|b^1\|_\infty^2 + 2\sqrt{\epsilon} \|b^1 b^2\|_\infty \right) \left( \|\partial_{xx} \bar{V}\|_\infty + K(\epsilon \|b^1\|_\infty + \sqrt{\epsilon} \|b^2\|_\infty)^\beta \right) \\ &\quad + \frac{1}{2} K \epsilon^{\frac{\beta}{2}} (\sqrt{\epsilon} \|b^1\|_\infty + \|b^2\|_\infty)^{2+\beta}. \end{aligned} \quad (3.19)$$

Recalling the definitions of  $\mu$  and  $\sigma$  and returning to the definition of  $I$ , this is otherwise stated as

$$\mu \partial_x \bar{V} + \frac{1}{2} \sigma^2 \partial_{xx}^2 \bar{V} + r = \frac{1}{\epsilon} \int (\bar{V}(\cdot, \cdot + b^\epsilon(\cdot, e)) - \bar{V}(t, x) + \epsilon(r - \delta r^\epsilon)) \nu(de). \quad (3.20)$$

Combined with Proposition 3.1, this shows that  $\bar{V}$  is a smooth solution of

$$\begin{cases} \partial_t \bar{V} + \sup_{a \in \mathbb{A}} \frac{1}{\epsilon} \int (\bar{V}(\cdot, \cdot + b^\epsilon(\cdot, a, e)) - \bar{V} + \epsilon(r(\cdot, a) - \delta r^\epsilon(\cdot, a, e))) \nu(de) = 0, & \text{on } [0, T] \times \mathbb{R}, \\ \bar{V}(T, \cdot) = 0, & \text{on } \mathbb{R}. \end{cases} \quad (3.21)$$

Applying Proposition 2.2, this implies that

$$\bar{V}(t, x) = \sup_{\alpha \in \mathcal{A}^t} \mathbb{E} \left[ \int_t^T \int \epsilon(r - \delta r^\epsilon)(X_{s-}^{t,x,\alpha}, \alpha_s, e) N(de, ds) \right],$$

so that, by the definition of  $V^\epsilon$ ,

$$\begin{aligned} |V^\epsilon - \bar{V}|(t, x) &\leq \sup_{\alpha \in \mathcal{A}^t} \mathbb{E} \left[ \int_t^T \int \epsilon |\delta r^\epsilon|(X_{s-}^{t,x,\alpha}, \alpha_s, e) N(de, ds) \right] \\ &= \sup_{\alpha \in \mathcal{A}^t} \mathbb{E} \left[ \int_t^T \int |\delta r^\epsilon|(X_{s-}^{t,x,\alpha}, \alpha_s, e) \nu(de) ds \right]. \end{aligned}$$

We conclude by appealing to (3.19). □

### 3.4 Construction of a $\epsilon^{\frac{\beta}{2}}$ -optimal control for the pure-jump problem

We now show that an  $\epsilon^{\frac{\beta}{2}}$ -optimal control for (2.3) can be constructed by considering a measurable map  $\bar{a} : [0, T) \times \mathbb{R} \mapsto \mathbb{A}$  satisfying

$$\bar{a} \in \operatorname{argmax}_{\bar{a} \in \mathbb{A}} \left( \mu(\cdot, \bar{a}) \partial_x \bar{V} + \frac{1}{2} \sigma^2(\cdot, \bar{a}) \partial_{xx}^2 \bar{V} + r(\cdot, \bar{a}) \right) \quad \text{on } [0, T) \times \mathbb{R}, \quad (3.22)$$

see e.g. [4, Proposition 7.33, p.153], and define  $\bar{\alpha}^{t,x} \in \mathcal{A}^t$  by

$$\bar{\alpha}_s^{t,x} = \bar{a}(s, X_{s-}^{t,x, \bar{\alpha}^{t,x}}), \quad s \in [t, T],$$

recall (2.1). As it is driven by a compound Poisson process, the couple  $(X^{t,x, \bar{\alpha}^{t,x}}, \bar{\alpha}^{t,x})$  is well-defined.

**Proposition 3.4.** *For all  $(t, x) \in [0, T) \times \mathbb{R}$ ,  $\bar{\alpha}^{t,x}$  is  $\epsilon^{\frac{\beta}{2}}$ -optimal. Namely,*

$$\frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s-}^{t,x, \bar{\alpha}^{t,x}}, \bar{\alpha}_s^{t,x}) dN_s \right] \geq V^\epsilon(t, x) - 2(T-t)C_K^\epsilon \epsilon^{\frac{\beta}{2}}, \quad \text{for all } \theta \in [0, 1].$$

*Proof.* It follows from Proposition 3.1, (3.22) and (3.20) that

$$\begin{aligned} \partial_t \bar{V} + \frac{1}{\epsilon} \int (\bar{V}(\cdot, \cdot + b^\epsilon(\cdot, \bar{a}, e)) - \bar{V} + \epsilon r(\cdot, \bar{a})) \nu(de) &\geq -\|\delta r^\epsilon\|_\infty, \\ \partial_t \bar{V} + \sup_{a \in \mathbb{A}} \frac{1}{\epsilon} \int (\bar{V}(\cdot, \cdot + b^\epsilon(\cdot, a, e)) - \bar{V} + \epsilon r(\cdot, a)) \nu(de) &\leq \|\delta r^\epsilon\|_\infty, \end{aligned}$$

so that applying Itô's Lemma and using (3.4) leads to

$$\begin{aligned} \bar{V}(t, x) - (T-t) \|\delta r^\epsilon\|_\infty &\leq \frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s-}^{t,x, \bar{\alpha}^{t,x}}, \bar{\alpha}_s^{t,x}) dN_s \right] \\ \bar{V}(t, x) + (T-t) \|\delta r^\epsilon\|_\infty &\geq \sup_{\alpha \in \mathcal{A}^t} \frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s-}^{t,x, \alpha}, \alpha_s) dN_s \right] = V^\epsilon(t, x). \end{aligned}$$

We conclude by appealing to (3.18). □

### 3.5 First order correction term

Under additional conditions, one can exhibit a first order correction term to improve the convergence speed in Theorem 3.3 and Proposition 3.4. From now on, we assume the following.

**Assumption 4.**

a. The map  $(t, x, a) \in [0, T) \times \mathbb{R} \times \mathbb{A} \mapsto \epsilon^{-\frac{\beta}{2}} \delta r^\epsilon(\cdot, e)$  is continuous, uniformly in  $(e, \epsilon)$  over  $\mathbb{R} \times (0, 1)$ .

b. The limit

$$r^1 := \lim_{\epsilon \rightarrow 0} \int \epsilon^{-\frac{\beta}{2}} \delta r^\epsilon(\cdot, e) \nu(de), \quad (3.23)$$

is well-defined.

c. Given

$$\mathbb{A}^0 := \operatorname{argmax}_{\bar{a} \in \mathbb{A}} \left( \mu(\cdot, \bar{a}) \partial_x \bar{V} + \frac{1}{2} \sigma^2(\cdot, \bar{a}) \partial_{xx}^2 \bar{V} + r(\cdot, \bar{a}) \right),$$

comparison holds in the sense of bounded discontinuous viscosity super- and subsolutions for

$$\begin{cases} \partial_t \varphi + \max_{\bar{a} \in \mathbb{A}^0} \left( \mu(\cdot, \bar{a}) \partial_x \varphi + \frac{1}{2} \sigma^2(\cdot, \bar{a}) \partial_{xx}^2 \varphi + r^1(\cdot, \bar{a}) \right) = 0, & \text{on } [0, T) \times \mathbb{R} \\ \varphi(T, \cdot) = 0 & \text{on } \mathbb{R} \end{cases} \quad (3.24)$$

and (3.24) admits a (unique) bounded viscosity solution, denoted by  $\delta \bar{V}^1$ .

**Remark 3.5.** Let us comment the above:

(a) Note that  $r^1$  is bounded, see (3.18) in Theorem 3.3. The right-hand side term in (3.23) therefore admits a limsup and a liminf. The condition (3.23) implies that the limit is actually well-defined. This point will be further discussed in Remark 3.7 below.

(b) If  $\bar{V}$  admits a continuous third-order space derivative  $\partial_{xxx}^3 \bar{V}$ , then one easily checks that  $r^1 = \frac{1}{6} \int |b^2(\cdot, e)|^3 \nu(de) \partial_{xxx}^3 \bar{V}$ , by a simple Taylor expansion.

(c) Also note that existence and uniqueness in (3.24) holds as soon as  $(t, x) \in [0, T) \times \mathbb{R} \mapsto \mathbb{A}^0(t, x)$  is regular enough, which can be checked on a case-by-case basis.

Under the above conditions,  $\delta \bar{V}^1$  is the first order term in the difference  $V^\epsilon - \bar{V}$ , i.e. (3.26) below holds with

$$\bar{V}^{1, \epsilon} := \bar{V} + \epsilon^{\frac{\beta}{2}} \delta \bar{V}^1. \quad (3.25)$$

**Theorem 3.6.** Let Assumption 4 hold. Then, for all  $(t, x) \in [0, T] \times \mathbb{R}$ ,

$$\lim_{\epsilon \downarrow 0} \epsilon^{-\frac{\beta}{2}} (V^\epsilon - \bar{V})(t, x) = \delta \bar{V}^1(t, x)$$

and therefore

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\beta}{2}} |V^\epsilon(t, x) - \bar{V}^{1, \epsilon}(t, x)| = 0. \quad (3.26)$$



If in addition  $\delta\bar{V}^1$  is  $C^{1,2}([0, T] \times \mathbb{R})$  and  $\partial_{xx}^2 \delta\bar{V}^1$  is  $\delta\beta$ -Hölder continuous in space for some  $\delta\beta > 0$  such that

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\delta\beta}{2}} \left\| \int \left( \epsilon^{-\frac{\beta}{2}} \delta r^\epsilon(\cdot, e) - r^1(\cdot, e) \right) \nu(\mathrm{d}e) \right\|_\infty < \infty, \quad (3.27)$$

then the control defined by

$$\check{\alpha}_s^{t,x} = \check{\alpha}(s, X_{s-}^{t,x, \check{\alpha}^{t,x}}), \quad s \in [t, T] \quad (3.28)$$

with

$$\check{\alpha} \in \operatorname{argmax}_{\bar{a} \in \mathbb{A}^0} \left\{ \mu(\cdot, \bar{a}) \partial_x \delta\bar{V}^1 + \frac{1}{2} \sigma(\cdot, \bar{a})^2 \partial_{xx}^2 \delta\bar{V}^1 + r^1(\cdot, \bar{a}) \right\}, \quad \text{on } [0, T] \times \mathbb{R}, \quad (3.29)$$

satisfies

$$\frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s-}^{t,x, \check{\alpha}^{t,x}}, \check{\alpha}_s^{t,x}) \mathrm{d}N_s \right] \geq V^\epsilon(t, x) - o(\epsilon^{\frac{\beta}{2}})$$

in which  $o$  is a continuous function satisfying  $\lim_{y \downarrow 0} y^{-1} o(y) = 0$ .

*Proof.* We split the proof in two steps.

a. Let us set  $W^\epsilon := \epsilon^{-\frac{\beta}{2}}(V^\epsilon - \bar{V})$  and consider its relaxed semi-limits

$$W^*(t, x) := \limsup_{(t', x') \rightarrow (t, x), \epsilon \downarrow 0} W^\epsilon(t', x'), \quad W_*(t, x) := \liminf_{(t', x') \rightarrow (t, x), \epsilon \downarrow 0} W^\epsilon(t', x').$$

Note that Theorem 3.3 ensures that the above are well-defined and bounded. We claim that  $W^*$  and  $W_*$  are respectively bounded sub- and supersolutions of (3.24). For brevity, we will only include the proof of the subsolution property, the supersolution property is proved similarly. Fix  $\varphi \in C_b^{1,2}$  and let  $(t_o, x_o) \in [0, T] \times \mathbb{R}$  achieve a maximum of  $W^* - \varphi$  on a ball  $B_k := \{(t, x) \in [0, T] \times \mathbb{R} : |t_o - t| \leq (T - t_o)/2, |x_o - x| \leq k\} \subset [0, T] \times \mathbb{R}$ , for some  $k > 0$ . Then, there exist a sequence  $(t_{\epsilon_n}, x_{\epsilon_n})_{\epsilon_n}$  such that  $\epsilon_n \rightarrow 0$ ,  $W^{\epsilon_n}(t_{\epsilon_n}, x_{\epsilon_n}) \rightarrow W^*(t_o, x_o)$ , and such that  $(t_{\epsilon_n}, x_{\epsilon_n})$  is a maximum of  $W^{\epsilon_n} - \varphi$  in the interior of  $B_{2k}$ , see e.g. [2, Lemma 6.1]. For  $k > \epsilon^{\frac{1}{2}}(\|b^1\|_\infty + \|b^2\|_\infty)$ , the viscosity subsolution property of  $V^\epsilon$ , applying Proposition 2.2 to the test function  $\bar{V} + \epsilon^{\frac{\beta}{2}}\varphi$ , implies that

$$0 \leq \partial_t (\bar{V} + \epsilon_n^{\frac{\beta}{2}} \varphi)(t_{\epsilon_n}, x_{\epsilon_n}) + \frac{1}{\epsilon_n} \left( \int \left( \bar{V} + \epsilon_n^{\frac{\beta}{2}} \varphi \right) (t_{\epsilon_n}, x_{\epsilon_n} + b^\epsilon(x_{\epsilon_n}, \bar{a}_n, e)) \nu(\mathrm{d}e) - \left( \bar{V} + \epsilon_n^{\frac{\beta}{2}} \varphi \right) (t_{\epsilon_n}, x_{\epsilon_n}) + \epsilon_n r(x_{\epsilon_n}, \bar{a}_{\epsilon_n}) \right)$$

for some  $\bar{a}_n \in \mathbb{A}$ . Recalling (3.17)-(3.18),  $\bar{a}_n$  converges to some element of  $\bar{\mathbb{A}} \in \mathbb{A}^0(t_o, x_o)$  as  $n$  goes to infinity, after possibly passing to a subsequence. By Proposition 3.1 again combined with (3.21),

$$0 \leq \partial_t \varphi(t_{\epsilon_n}, x_{\epsilon_n}) + \frac{1}{\epsilon_n} \int \left( \varphi(t_{\epsilon_n}, x_{\epsilon_n} + b^\epsilon(x_{\epsilon_n}, \bar{a}_n, e)) - \varphi(t_{\epsilon_n}, x_{\epsilon_n}) + \epsilon_n \epsilon_n^{-\frac{\beta}{2}} \delta r^{\epsilon_n}(x_{\epsilon_n}, \bar{a}_n, e) \right) \nu(\mathrm{d}e).$$

Sending  $n \rightarrow \infty$  and using parts a. and b. of Assumption 4 together with Assumption 2, this leads to

$$0 \leq \partial_t \varphi(t_o, x_o) + \mu(x_o, \bar{a}) \partial_x \varphi(t_o, x_o) + \frac{1}{2} \sigma(x_o, \bar{a}) \partial_{xx}^2 \varphi(t_o, x_o) + r^1(x_o, \bar{a}),$$

so that the required subsolution property is proved on  $[0, T) \times \mathbb{R}$ . The fact that  $W^*(T, \cdot) \leq 0$  follows from the last assertion of Theorem 3.3. By comparison,  $W := W^* = W_*$  is the unique bounded viscosity solution of (3.24) and is therefore equal to  $\delta \bar{V}^1$ .

b. We now assume that  $\delta \bar{V}^1$  is  $C^{1,2}([0, T) \times \mathbb{R})$  and that  $\partial_{xx}^2 \delta \bar{V}^1$  is  $\delta\beta$ -Hölder continuous in space for some  $\delta\beta > 0$  such that (3.27) holds. Using (3.27) and the same arguments as in the proof of Theorem 3.3 lead to

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\delta\beta}{2}} \left\| \int \delta r^{1,\epsilon}(\cdot, e) \nu(de) \right\|_\infty < \infty, \quad (3.30)$$

in which

$$\delta r^{1,\epsilon} := \frac{1}{\epsilon} (\delta \bar{V}^1(\cdot, \cdot + b^\epsilon) - \delta \bar{V}^1) + \epsilon^{-\frac{\beta}{2}} \delta r^\epsilon - b^1 \delta \bar{V}^1 - \frac{1}{2} |b^2|^2 \partial_{xx}^2 \delta \bar{V}^1 - r^1.$$

Moreover, direct computations show that  $\bar{V}^{1,\epsilon}$  defined in (3.25) solves

$$\partial_t \bar{V}^{1,\epsilon} + \frac{1}{\epsilon} \int \left( \bar{V}^{1,\epsilon}(\cdot, \cdot + b^\epsilon(\cdot, \check{a}, e)) - \bar{V}^{1,\epsilon}(t, x) - \epsilon \epsilon^{\frac{\beta}{2}} \delta r^{1,\epsilon}(\cdot, \check{a}, e) \right) \nu(de) + r(\cdot, \check{a}) = 0$$

on  $[0, T) \times \mathbb{R}$ , where  $\check{a}$  is defined as in (3.29). Together with (3.30), this implies that, for  $\check{\alpha}^{t,x}$  defined as in (3.28), we have

$$\frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s-}^{t,x,\check{\alpha}^{t,x}}, \check{\alpha}_s^{t,x}) dN_s \right] \geq \bar{V}^{1,\epsilon}(t, x) - \epsilon^{\frac{\beta}{2}} O(\epsilon),$$

in which  $O$  is a continuous function with  $O(0) = 0$ . On the other hand, it follows from Step a. that  $|V^\epsilon(t, x) - \bar{V}^{1,\epsilon}(t, x)| \leq o(\epsilon^{\frac{\beta}{2}})$ .  $\square$

**Remark 3.7.** *If the limit in (3.23) is not defined, one can still define its limsup and its liminf (recall that it is bounded). Let us denote them by  $r^{1*}$  and  $r_*^1$  respectively. Then,  $W^*$  defined in the above proof is simply a viscosity sub-solution of (3.24) with  $r^{1*}$  in place of  $r^1$ . Similarly,  $W_*$  is a viscosity super-solution of the same equation but with  $r_*^1$  in place of  $r^1$ . This still provides asymptotic upper- and lower-bounds for  $\epsilon^{-\frac{\beta}{2}}(V^\epsilon - \bar{V})$ .*

**Example 3.8.** *To illustrate the above, we consider a toy model in which explicit solutions can be derived. Although it does not satisfy our general assumptions, e.g. of boundedness and Hölder regularity in space, we shall see that the above methodology can be applied. We consider the dynamics*

$$X^{t,x,\alpha} = x + \int_t^\cdot X_{s-}^{t,x,\alpha} \int (\epsilon b^1(\alpha_s, e) + \sqrt{\epsilon} b^2(\alpha_s, e)) N(ds, de),$$

in which  $b^1$  and  $b^2$  are bounded and continuous with respect to their first argument, uniformly in the second one. For  $\gamma \in (0, 1]$ , the value function is defined as

$$V^\epsilon(t, x) = \sup_{\alpha \in \mathcal{A}^t} \frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T \int |X_{s^-}^{t,x,\alpha}|^\gamma r(\alpha_s) dN_s \right],$$

for some continuous function  $r$ . Then, one easily checks that  $\bar{V}(t, x) = \bar{f}(t)x^\gamma$  in which  $\bar{f}$  solves

$$\partial_t \bar{f} + \sup_{\bar{a} \in \mathbb{A}} \left( \bar{f} \{ \gamma \mu(\bar{a}) + \frac{1}{2} \gamma(\gamma - 1) \sigma^2(\bar{a}) \} + r(\bar{a}) \right) = 0, \text{ on } [0, T) \times \mathbb{R},$$

with  $\bar{f}(T) = 0$ . Because  $x^\gamma$  factorizes, the Hölder constant of  $\partial_{xx}^2 \bar{V}$  can be considered around  $x = 1$ . Since the third-order space derivative of  $\bar{V}$  is bounded in a neighborhood of 1, Theorem 3.3 applies with  $\beta = 1$ . The convergence rate is therefore of order  $\epsilon^{\frac{1}{2}}$ . Moreover, by direct computations, the first order correction term is of the form  $\delta \bar{V}^1(t, x) = \delta \bar{f}(t)x^\gamma$  where  $\delta \bar{f} \not\equiv 0$  solves

$$\partial_t \delta \bar{f} + \sup_{\bar{a} \in \mathbb{A}^1} \left( \delta \bar{f} \{ \gamma \mu(\bar{a}) + \frac{1}{2} \gamma(\gamma - 1) \sigma^2(\bar{a}) \} + r^1(\cdot, \bar{a}) \right) = 0$$

with  $\delta \bar{f}(T) = 0$ , in which

$$(t, \bar{a}) \in [0, T] \times \mathbb{A} \mapsto r^1(t, \bar{a}) := \gamma(\gamma - 1) \ell \left( \int (b^1 b^2)(\bar{a}, e) \nu(de) \right) \bar{f}(t)$$

for some (explicit) continuous map  $\ell$  with linear growth.

### 3.6 Higher order expansions

To conclude this section, note that higher order expansions can be obtained, upon existence to associated systems of parabolic equations. Namely, let us assume the following.

**Assumption 5.** *There exists  $(\delta \beta_i)_{i=0, \dots, i_\circ} \subset (0, 1]^{i_\circ}$  together with  $C^{1,2}([0, T] \times \mathbb{R}) \cap C^0([0, T] \times \mathbb{R})$  functions  $(\delta \bar{V}^i)_{i=0, \dots, i_\circ}$  such that, for  $i = 0, \dots, i_\circ$ ,  $\partial_{xx}^2 \delta \bar{V}^i$  is  $\delta \beta_i$ -Hölder in space and  $\delta \bar{V}^i$  solves*

$$\partial_t \delta \bar{V}^i + \mu(\cdot, \bar{a}^\epsilon) \partial_x \delta \bar{V}^i + \frac{1}{2} \sigma(\cdot, \bar{a}^\epsilon)^2 \partial_{xx}^2 \delta \bar{V}^i + r^i(\cdot, \bar{a}^\epsilon) = 0, \text{ on } [0, T) \times \mathbb{R},$$

$$\delta \bar{V}^i(T, \cdot) = 0 \text{ on } \mathbb{R},$$

in which  $r^0 = r$ ,

$$\bar{a}^\epsilon \in \operatorname{argmax}_{\bar{a} \in \mathbb{A}} \left\{ \mu(\cdot, \bar{a}) \partial_x \bar{V}^{i_\circ, \epsilon} + \frac{1}{2} \sigma(\cdot, \bar{a})^2 \partial_{xx}^2 \bar{V}^{i_\circ, \epsilon} + r(\cdot, \bar{a}) \right\}$$

with

$$\bar{V}^{i_0, \epsilon} = \delta \bar{V}^0 + \sum_{j=1}^{i_0} \epsilon^{\beta_{i-1}} \delta \bar{V}^i, \quad \beta_i = \sum_{j=0}^i \delta \beta_j \text{ for } i \leq i_0,$$

and, using the convention  $\delta \beta^{-1} = 0$  and  $\delta r^{-1, \epsilon} \equiv 0$ ,

$$\begin{aligned} \delta r^{i, \epsilon} &:= \frac{1}{\epsilon} \left( \delta \bar{V}^i(\cdot, \cdot + b^\epsilon) - \delta \bar{V}^i \right) + \epsilon^{-\frac{\delta \beta_{i-1}}{2}} \delta r^{i-1, \epsilon} - b^1 \delta \bar{V}^i - \frac{1}{2} |b^2|^2 \partial_{xx}^2 \delta \bar{V}^i - r^i \\ r^i &:= r 1_{\{i=0\}} + 1_{\{i>0\}} \lim_{\epsilon \rightarrow 0} \int \epsilon^{-\frac{\delta \beta_{i-1}}{2}} \delta r^{i-1, \epsilon}(\cdot, e) \nu(de) \text{ for } i \leq i_0. \end{aligned}$$

The limits in (3.23) are well-defined on  $[0, T] \times \mathbb{R}$ , and

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\delta \beta_{i_0}}{2}} \left\| \int (\epsilon^{-\frac{\beta_{i_0}-1}{2}} \delta r^{i_0-1, \epsilon} - r^{i_0})(\cdot, e) \nu(de) \right\|_\infty < \infty. \quad (3.31)$$

**Proposition 3.9.** *Let Assumption 5 hold. Then, for all  $(t, x) \in [0, T] \times \mathbb{R}$ ,*

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\beta_{i_0}} |V^\epsilon - \bar{V}^{i_0, \epsilon}|(t, x) < \infty.$$

Moreover, the control defined by

$$\check{\alpha}_s^{t, x} = \check{a}^\epsilon(s, X_{s-}^{t, x, \check{\alpha}^{t, x}}), \quad s \in [t, T],$$

where  $\check{a}^\epsilon$  is a Borel measurable map such that

$$\check{a}^\epsilon \in \operatorname{argmax}_{\bar{a} \in \mathbb{A}^{i_0, \epsilon}} \left( \mu(\cdot, \bar{a}) \partial_x \bar{V}^{i_0} + \frac{1}{2} \sigma(\cdot, \bar{a})^2 \partial_{xx}^2 \bar{V}^{i_0} + r^{i_0}(\cdot, \bar{a}) \right),$$

satisfies

$$\frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s-}^{t, x, \check{\alpha}^{t, x}}, \check{\alpha}_s^{t, x}) dN_s \right] \geq V^\epsilon(t, x) - C \epsilon^{\frac{\beta_{i_0}}{2}},$$

for some constant  $C > 0$ .

*Proof.* With the above construction

$$\partial_t \bar{V}^{i_0, \epsilon} + \frac{1}{\epsilon} \int \left( \bar{V}^{i_0, \epsilon}(\cdot, \cdot + b^\epsilon(\cdot, \check{a}^\epsilon, e)) - \bar{V}^{i_0, \epsilon}(t, x) - \epsilon \epsilon^{\frac{\beta_{i_0}-1}{2}} \delta r^{i_0, \epsilon}(\cdot, \check{a}^\epsilon, e) \right) \nu(de) + r(\cdot, \check{a}^\epsilon) = 0$$

on  $[0, T] \times \mathbb{R}$ , while

$$\partial_t \bar{V}^{i_0, \epsilon} + \frac{1}{\epsilon} \int \left( \bar{V}^{i_0, \epsilon}(\cdot, \cdot + b^\epsilon(\cdot, a, e)) - \bar{V}^{i_0, \epsilon}(t, x) - \epsilon \epsilon^{\frac{\beta_{i_0}-1}{2}} \delta r^{i_0, \epsilon}(\cdot, a, e) \right) \nu(de) + r(\cdot, a) \leq 0$$

on  $[0, T] \times \mathbb{R}$  for all  $a \in \mathbb{A}$ . By (3.31) and the same arguments as in the proof of Theorem 3.3,

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\delta \beta_{i_0}}{2}} \left\| \int \delta r^{i_0, \epsilon}(\cdot, e) \nu(de) \right\|_\infty < \infty,$$

so that the required result follows by verification.  $\square$

## 4 Application to an auction problem

Repeated online auction bidding are typical problems in which the real value of the parameters  $b, r$  and  $\nu$  are unknown and on which reinforcement learning techniques are applied. The latter requires to estimate, very quickly, the optimal control for different sets of parameters. Being modelled as a discrete time problem, with fixed auction times, or more realistically in the form of a pure-jumps problem as in Section 2, see also [14], we face in any case the fact that auctions are issued almost continuously which correspond to a very large time horizon in the discrete-time version or to a very large intensity in the pure-jump modelling. The numerical cost of a precise estimation of the optimal control is too important to combine it with a reinforcement learning approach.

### 4.1 Model and description of the optimal policy

We consider here a simple auction problem motivated by online advertising systems. A single ad-campaign is provided several opportunities to buy ad-space to display its ad over the course of the day. These ad spaces arrive at random, according to the point process  $N$ , since they are dependent on users from specific targeted audiences loading a website. In real-world display advertising, the kind encountered on the sides of web-pages, these opportunities take the form of an auction between several bidders and an ad-exchange platform.

The format of the auction used is critical to the strategic behaviour of bidders and the revenue of the seller. There is a large amount of literature in auction theory on the subject, see *e.g.* [19–21], and real-world auctions can take very complex formats. For simplicity, we consider an auctioneer which has implemented a lazy second price auction [21, 22] with individualised reserve price. In this format, our bidding agent wins the ad-slot if it submits a bid above its (henceforth the) reserve price and the competition, and if it wins it pays the maximum between the reserve and the competition. For a given reserve price  $x$ , a bid  $a \in (0, +\infty)$  and a random competition bid  $B \geq 0$  following a smooth probability distribution  $F_B$ , the expected payoff  $r(x, a)$  for an auction is thus expressible through a simple integration by parts as

$$r(x, a) = \mathbb{E}[(v - x \vee B)1_{a \geq x \vee B}] = 1_{a \geq x} \left( (v - a)F_B(a) + \int_x^a F_B(b)db \right), \quad (4.1)$$

in which  $v$  is the value of the ad-slot for the bidder.

As the right hand side of (4.1) highlights, reserve prices are a mechanism put in place by sellers to compensate for lack of competition, which would drive down the price and their profits. It is well established that a reserve price is not as profitable as increasing the number of participants by

one [10]. Consequently, when there are many bidders a control has little effect on the system. To clearly demonstrate the use of controlling the reserve price, we study a strongly asymmetric setting, where the agent has a value  $v = 0.5$  much higher than the competition, which we take uniform on  $(0, 0.3)$ . In this setting, he is directly competing against the seller for his extra value above the average competition. For the purpose of this example, we do not want to go to the limit of this asymmetry, the posted price auction where there is no competition, as it could lead the control problem to degeneracy, such as negative prices and difficult boundary conditions.

There is a large literature on revenue maximisation algorithms in online auctions, or how to set the reserve price to maximise revenue, such as [5, 9, 11, 13]. For the sake of simplicity, in this example, we will model the dynamics of the reserve price using a simple mean reverting process:

$$b^1(x, a, e) = \kappa a + (1 - \kappa)r_0 - x \text{ and } b^2(x, a, e) = e \text{ with } \nu \sim \text{Unif}(-0.1, 0.1),$$

with  $\kappa \in (0, 1)$  and  $r_0 > 0$ . This corresponds to setting a minimum reserve price  $(1 - \kappa)r_0$ , and tracking the agent’s bid with aggressiveness measured by  $\kappa$ . Setting  $r_0 = 0.15$  as the monopoly price of the competition guarantees the seller a better revenue against the competition, while  $\kappa a$  allows him to pursue the agent’s extra value. We set  $\kappa = 1/2$ , for a balance between prudence and aggression.

The control problem consists in maximising the static auction revenue, while considering the impact bids have on the system. In the static auction format, we can identify three domains the reserve price can be in: “non-competitive”, “competitive”, and “unprofitable”. When the reserve price is below the competition’s average<sup>1</sup> there is essentially no competition to the agent, since the reserve price barely affects his profits. On the other hand, when the reserve price is in the range between 0.3 and  $v = 0.5$ , the reserve price becomes the dominant term in  $r$  and the agent has to compete with the seller for the value margin he has above other buyers. Finally, if the reserve price is above  $v$ , there is no profit so no reason to take part in the auction.

When the reserve price is dynamic, a good control seeks to maximise profit while pushing the reserve price to “non-competitive”. You can see this in effect on figure 1. In the non-competitive regime (left), starting at  $X_0 = 0.15$ , this policy recovers 85% of the best possible income of the static setting, where  $X_t = 0$  for all  $t$  and the average price is  $0.5 - \mathbb{E}[B] = 0.35$ . In the competitive regime (center), the policy bids just above the reserve price to apply a downwards pressure until it reaches the non-competitive domain again. Finally, in the unprofitable regime (right), the agent boycotts the auction by bidding 0, bringing down the price. Notice how when the agents stops boycotting there is an inflection point in the downwards trend of the price, schematically represented by the dotted line.

---

<sup>1</sup>Recall that the competition here models the distribution of the maximum bid of all other participants, so this is the average of the maximum of other participants’ bids.

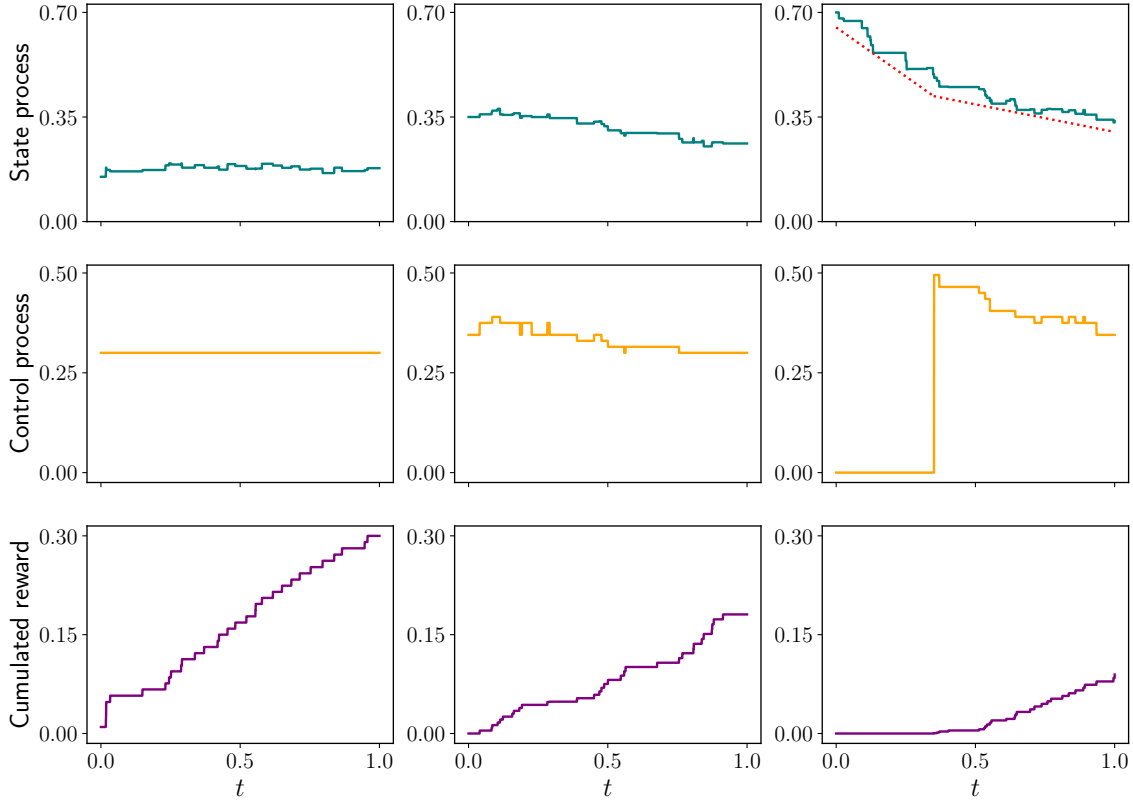


Figure 1: Selected sample realisations of the system for  $\epsilon = 10^{-1.5}$ , starting from 0.15 (left), 0.3 (center), and 0.7 (right).

## 4.2 Numerical implementation

Adapting (2.4) and (3.3), we normalise the horizon to 1, and allow the reserve price  $X$  to vary in  $\mathbb{R}$ , and  $a$  in  $\mathbb{R}_+$ . This allows us to easily set boundary conditions for the equation. When an auction happens with  $X < 0$ , the price is set by the competition, which will be a.s. positive. So as  $X \rightarrow -\infty$ , the reserve price becomes irrelevant and the value converges to the value of a single auction without reserve price due to the renormalisation of  $r^\epsilon$ . Conversely, for  $\epsilon < 1$ , as  $X \rightarrow +\infty$ , the probability of  $X$  descending below  $v$  and generating any revenue decreases due to the noise. If  $\epsilon = 1$ , the first jump can always take  $X$  to  $(1 - \kappa)r_0$  regardless of start value. In both cases, a Dirichlet boundary condition is appropriate at  $[0, 1) \times \{-\infty, +\infty\}$ . In numerical resolution, we will use Dirichlet boundary conditions equal to 0 on  $[0, 1) \times (-0.5, 3)$ . Given this domain for the reserve price, we can set the controls on an even mesh in  $[0, 3]$ , of fineness 0.01.

We solve both problems numerically with an explicit finite difference solver, and for simplicity a Riemann sum using the same mesh for the numerical integration part. This formulation is equivalent to a Markov Chain control problem. Let  $M_t = \{k\Delta_t; k = 0, \dots, \lfloor 1/\Delta_t \rfloor\}$ ,  $M_x = \{-0.5 + k\Delta_x; k = 0, \dots, \lfloor 3.5/\Delta_x \rfloor\}$  be the space and time meshes, with sizes,  $\Delta_x = \epsilon^{3/2}/2$ ,  $\Delta_t = \Delta_x^{2/3}$ . Denote  $V_n(x_i)$  be the output of the solver at time  $t_n \in M_t$  and position  $x_i \in M_x$ . For the pure jump problem, we explicitly compute:

$$V_n^\epsilon(x_i) = V_{n+1}^\epsilon(x_i) + \frac{\Delta_t}{\epsilon} \sup_{a \in \mathbb{A}_n} \sum_{x_j \in M_x} V_{n+1}^\epsilon(x_j) f_{x_i, a}^{\nu, \epsilon}(x_j) \Delta_x - V_{n+1}^\epsilon(x_i) + r(x_i, a)$$

where  $f_{x, a}^{\nu, \epsilon}$  is the transition kernel induced by  $b^1(x, a, \cdot)$ ,  $b^2(x, a, \cdot)$ , and  $\nu$ . For the diffusion, we consider meshes  $M_t = \{kd_t; k = 0, \dots, \lfloor 1/d_t \rfloor\}$ ,  $M_x = \{-0.5 + kd_x; k = 0, \dots, \lfloor 3.5/d_x \rfloor\}$ , with  $d_x = 10^{-2}$ ,  $d_t = d_x^2$  and solve recursively

$$\bar{V}_{n-1}(x_i) = \bar{V}_n(x_i) + d_t \sup_{a \in \mathbb{A}_n} \left( (\kappa a + (1 - \kappa)r_0 - x) \delta_x^u V_n(x_i) + \frac{1}{2} \sigma^2 \delta_{xx} V_n(x - i) + r(x_i, a) \right)$$

where  $\delta_x^u$  and  $\delta_{xx}$  are the uplift first order and centred second order finite differences on  $M_x$  respectively. We took  $\mathbb{A}_n = \{10^{-2}k; k = 0, \dots, 300\}$  in both cases.

To give some insight into the complexity trade-off, see that when  $\epsilon$  is large, there are relatively few jumps so the time iteration won't require many steps to get an accurate solution. This scaling is indicated by the  $\Delta_t/\epsilon$  term. At the same time, the jumps are large so even a coarse mesh in  $x$  will be sufficient for the numerical integration to approach the integral. Unfortunately as  $\epsilon \rightarrow 0$ , one must refine both the time mesh, linearly with  $1/\epsilon$ , and the integration mesh which is paid quadratically due to the non-local nature of the equation. In practice, this makes computations grow at a super-cubic rate with  $\epsilon$ , which is rapidly prohibitively expensive. In our example problem, the noise is supported on a bounded interval of size  $\sqrt{\epsilon}$ , so we save some computation time, but figure 2 shows the computation cost (pictured with dots) still grows super-quadratically and overcomes the cost of our accurate diffusion mesh (solid horizontal line) even for large  $\epsilon$ . Note that we computed the diffusive limit to a very high precision and with an explicit scheme, and that even for  $\epsilon$  of order of  $10^{-3}$  the cost is already 6 times higher. Note that if the control were to intervene in a non-linear way, we might need to refine the control mesh with  $\epsilon$  in the pure-jump case, further increasing the computational burden.

Beyond gains in computation, Figure 3 verifies that Proposition 3.3 holds with meaningful constants in finite time on this problem. Figure 3 shows that the error is very low even for large values of  $\epsilon$ , and decreases at the correct rate of  $\epsilon^{-1/2}$ . Likewise, Figure 4 shows the rate of Proposition 3.4 also holds even for large  $\epsilon$ .



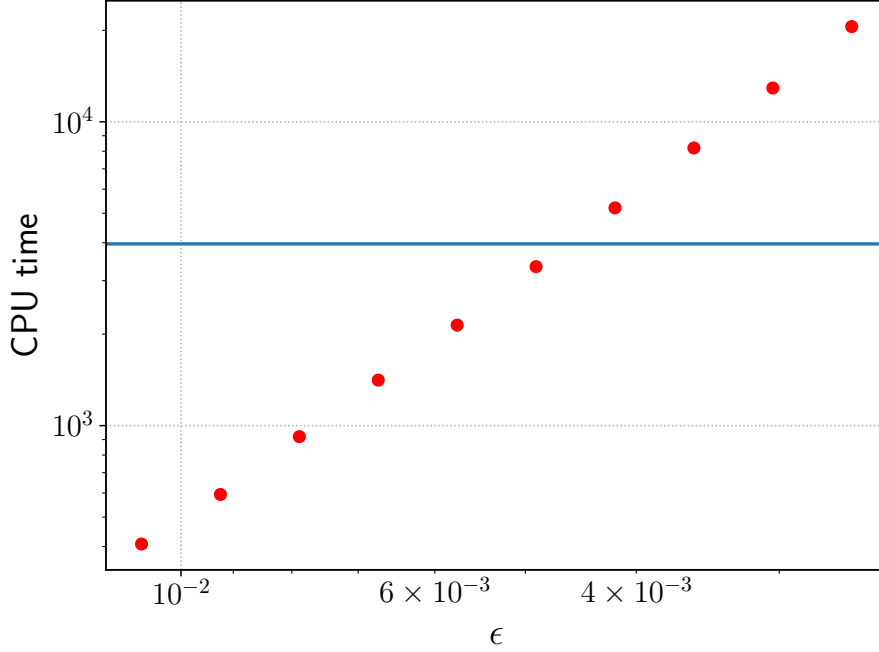


Figure 2: Numerical cost for  $V^\epsilon$  (log scales).

## 5 Remark on the diffusive limit of discrete time problems

Instead of considering the diffusive limit of a continuous time pure-jump problem, one could similarly consider a sequence of pure discrete time problems with actions at time  $t_i^n := iT/n$ ,  $i \leq n$ :

$$V^n(t, x) := \sup_{\alpha \in \mathcal{A}} \frac{T}{n} \mathbb{E} \left[ \sum_{i=0}^n 1_{\{t_i^n \geq t\}} r(X_{t_i^n}^{t, x, \alpha}, \alpha_{t_i^n}) \right],$$

with

$$X^{t, x, \alpha} := x + 1_{[t, T]} \sum_{i=0}^n 1_{\{t_i^n \in (t, \cdot]\}} b(X_{t_i^n}^{t, x, \alpha}, \alpha_{t_i^n}, \xi_i^n)$$

in which  $(\xi_i^n)_{i \geq 1}$  is i.i.d. following the distribution  $\nu$  and  $\mathcal{A}$  is the collection of  $\mathbb{A}$ -valued processes that are predictable with respect to the  $\mathbb{P}$ -augmented filtration generated by  $\sum_{i=0}^n 1_{\{t_i^n \in [0, \cdot]\}} \xi_i^n$ .

Upon taking  $b$  of the form

$$b^n = \frac{T}{n} b^1 + \sqrt{\frac{T}{n}} b^2, \quad \text{with } \mathbb{E}[b^2(\cdot, \xi_1^n)] = 0,$$

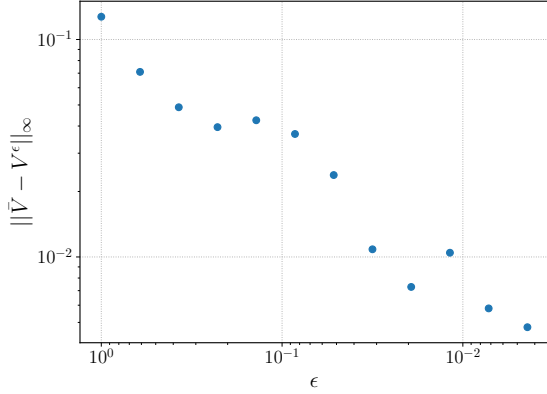


Figure 3: Limit value function error relative to  $V^\epsilon$ , at  $t = 0$  (log scales).

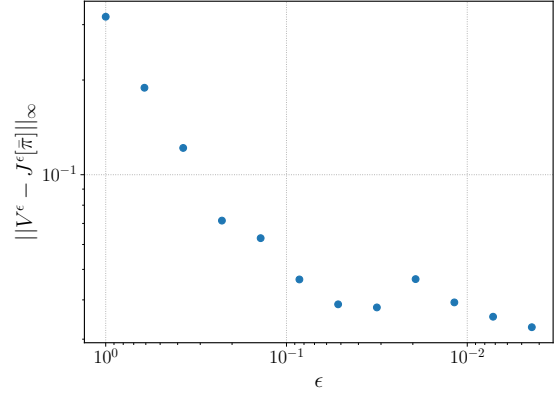


Figure 4: Limit policy error relative to  $V^\epsilon$ , at  $t = 0$  (log scales).

one would obtain the same diffusive limit as in Section 3.3 when letting  $n \rightarrow \infty$ . Namely, the same arguments as in [15, Section 3] combined with Proposition 3.1 and the fact that comparison holds for (3.3) imply that  $\lim_{n \rightarrow \infty} V^n$  is well-defined and is equal to  $\bar{V}$ .

One can also check that the convergence holds at a speed  $n^{-\frac{\beta}{2}}$ . Let us sketch the proof. First, the same arguments as in the proof of Theorem 3.3 imply that

$$\delta r^n := \frac{n}{T} (\bar{V}(\cdot, \cdot + b^n) - \bar{V}) - b^1 \partial_x \bar{V} - \frac{1}{2} |b^2|^2 \partial_{xx} \bar{V}$$

satisfies

$$\|\delta r^n\|_\infty \leq C n^{-\frac{\beta}{2}}$$

for some  $C > 0$  independent on  $n$ . Thus, by Proposition 3.1

$$0 = \partial_t \bar{V}(t, x) \frac{T}{n} + \sup_{a \in \mathbb{A}} \int \left( \bar{V}(t, x + b^n(x, a, e)) - \bar{V}(t, x) + \frac{T}{n} (r(x, a) - \delta r^n(x, \bar{a}, e)) \right) \nu(de)$$

so that

$$\begin{aligned} \bar{V}(t_i^n, x) &= \sup_{a \in \mathbb{A}} \mathbb{E} \left[ \int_{t_i^n}^{t_{i+1}^n} \partial_t \bar{V}(t_i^n, x) ds + \bar{V}(t_{i+1}^n, x + b^n(x, a, \xi_{i+1}^n)) + \frac{T}{n} (r(x, a) - \delta r^n(x, a, \xi_{i+1}^n)) \right] \\ &= \sup_{a \in \mathbb{A}} \left( \mathbb{E} \left[ \bar{V}(t_{i+1}^n, x + b^n(x, a, \xi_{i+1}^n)) + \frac{T}{n} r(x, a) \right] \right. \\ &\quad \left. + \mathbb{E} \left[ \int_{t_i^n}^{t_{i+1}^n} (\partial_t \bar{V}(t_i^n, x) - \partial_t \bar{V}(s, x + b^n(x, a, \xi_{i+1}^n)) - \delta r^n(x, a, \xi_{i+1}^n)) ds \right] \right). \end{aligned}$$

We then use (3.13) to obtain that

$$\bar{V}(t_i^n, x) = \sup_{a \in \mathbb{A}} \mathbb{E} \left[ \bar{V}(t_{i+1}^n, x + b^n(x, a, \xi_{i+1}^n)) + \frac{T}{n} r(x, a) + \int_{t_i^n}^{t_{i+1}^n} \varpi^n(s, x, a) ds \right]$$

in which  $|\varpi^n| \leq Cn^{-\frac{\beta}{2}}$ , for some  $C > 0$  independent of  $n$ . It follows that

$$\bar{V}(t_i^n, x) = \sup_{\alpha \in \mathcal{A}} \frac{T}{n} \mathbb{E} \left[ \sum_{i=0}^n 1_{\{t_i^n \geq t\}} \left( r(X_{t_i^n}^{t, x, \alpha}, \alpha_{t_i^n}) + \int_{t_i^n}^{t_{i+1}^n} \varpi^n(s, x, a) ds \right) \right],$$

which provides the expected result.

Likewise, the Markovian control defined through (3.22) can be shown to be  $n^{-\frac{\beta}{2}}$ -optimal for  $V^n$ , see the proof of Proposition 3.4.

## References

- [1] B. Andrews. Fully nonlinear parabolic equations in two space variables. *arXiv preprint math/0402235*, 2004.
- [2] G. Barles. *Solution de viscosités des équations d’Hamilton Jacobi*, volume 17 of *Mathématiques et Applications*. Springer Verlag, 1994.
- [3] N. Bäuerle. Approximation of optimal reinsurance and dividend payout policies. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 14(1):99–113, 2004.
- [4] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control. The Discrete-Time Case*. Academic Press, New York, 1978.
- [5] A. Blum, V. Kumar, A. Rudra, and F. Wu. Online learning in online auctions. *Theoretical Computer Science*, 324(2-3):137–146, 2004.
- [6] B. Bouchard. Stochastic targets with mixed diffusion processes and viscosity solutions. *Stochastic processes and their applications*, 101(2):273–302, 2002.
- [7] B. Bouchard, S. Geiss, and E. Gobet. First time to exit of a continuous itô process: General moment estimates and l1-convergence rate for discrete time approximations. *Bernoulli*, 23(3):1631–1662, 2017.
- [8] B. Bouchard and N. Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization*, 49(3):948–962, 2011.

- [9] S. Bubeck, N. R. Devanur, Z. Huang, and R. Niazadeh. Multi-scale online learning: Theory and applications to online auctions and pricing. *The Journal of Machine Learning Research*, 20(1):2248–2284, 2019.
- [10] J. Bulow and P. Klemperer. Auctions versus negotiations. *American Economic Review*, 86(1):180–94, 1996.
- [11] N. Cesa-Bianchi, C. Gentile, and Y. Mansour. Regret minimization for reserve prices in second-price auctions. *IEEE Transactions on Information Theory*, 61(1):549–564, 2014.
- [12] A. Cohen and V. R. Young. Rate of convergence of the probability of ruin in the cramér–lundberg model to its diffusion approximation. *Insurance: Mathematics and Economics*, 93:333–340, 2020.
- [13] L. Croissant, M. Abeille, and C. Calauzenes. Real-time optimisation for online learning in auctions. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 2217–2226. PMLR, 13–18 Jul 2020.
- [14] J. Fernandez-Tapia, O. Guéant, and J-M. Lasry. Optimal Real-Time Bidding Strategies. *Applied Mathematics Research eXpress*, 2017(1):142–183, 09 2016.
- [15] W. H. Fleming and P. E. Souganidis. On the existence of value functions of two-player, zero-sum stochastic differential games. *Indiana University Mathematics Journal*, 38(2):293–314, 1989.
- [16] J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.
- [17] O. A. Ladyzhenskaia, V. Solonnikov, and N. N. Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.
- [18] G. M. Lieberman. *Second order parabolic differential equations*. World scientific, 1996.
- [19] R. B. Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.
- [20] M. Ostrovsky and M. Schwarz. Reserve prices in internet advertising auctions: A field experiment. In *Proceedings of the 12th ACM Conference on Electronic Commerce*, EC '11, pages 59–60, New York, NY, USA, 2011. Association for Computing Machinery.
- [21] R. Paes Leme, M. Pal, and S. Vassilvitskii. A field guide to personalized reserve prices. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pages 1093–1102, Republic and Canton of Geneva, CHE, 2016. International World Wide Web Conferences Steering Committee.

- [22] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.