



**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-03269700v2

**HAL Id: hal-03269700**

**<https://hal.science/hal-03269700v2>**

Preprint submitted on 9 Nov 2021 (v2), 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‡

November 8, 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 exponent 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 jumps. We illustrate this approach in the context of a display advertising auction problem.

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

## 1 Introduction

Let  $N$  be a random point process with predictable compensator  $\lambda\nu(de)dt$ , for some finite Borel measure  $\nu$  on  $\mathbb{R}$ ,  $\lambda > 0$ , 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  $\alpha$  belongs to the set  $\mathcal{A}$  of predictable controls with values in some given set  $\mathbb{A}$ . Then, under mild assumptions, the value of the control problem

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

with  $N_t := N(\mathbb{R}, [0, t])$ ,  $t \geq 0$ , solves the integro-differential equation

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

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 approximate both the value function  $V$  and the associated optimal control.

However, (1.1) is non-local and obtaining a precise approximation of the solution is highly time consuming as soon as the intensity  $\lambda$  of  $N$  is large. This is the case, for instance, for ad-auctions on the web, see e.g. [16], 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, when  $\lambda$  is very large, it is tempting to approximate the original jump diffusion control problem by its asymptotic as  $\lambda \rightarrow \infty$ . In this paper, we consider the diffusive limit approximation. Namely, if one takes  $\lambda$  of the form  $\lambda = 1/\epsilon$ , with  $\epsilon$  small, 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  $\epsilon V$  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 [16], which corresponds to considering a fluid limit, but this is less precise.

For such a specification of the coefficients  $(\lambda, b)$ , the existence of a diffusive limit is expected, see e.g. [18] 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 [17, Section 3], which considers the limit of discrete time zero-sum games, or by applying weak-convergence results. In particular an important literature on this subject exists within the insurance and queueing network literatures, see e.g. [3, 13, 14]. 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 below.

In Section 3, we verify that the above intuition is correct. Unlike [17], we do not simply rely on the stability of viscosity solutions. Neither do we rely on the weak convergence of the underlying process.

The reason is that weak convergence does not give access to the convergence speed in optimal control problems. Instead, we directly study 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 of (1.2) are uniformly Lipschitz in space and under a uniform ellipticity condition. By a second order Taylor expansion, 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 this rate 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. To achieve higher orders of convergence, this approach can be generalised to a system of non-linear PDEs, upon its existence.

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. Our numerical experiments show that our approximation permits a considerable gain in computation time.

For ease of exposition, we shall restrict to situations where the controlled process is of dimension one. This fact will be used explicitly only to derive our regularity results in Section 3.2. Similar results can be obtained in higher dimension, by using standard regularity results for parabolic partial differential equations, see e.g. [19, 20].

## 2 The pure-jump optimal control problem

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

### 2.1 Definition

Let  $\Omega = \mathbb{D}$  denote the space of one dimensional càdlàg functions on  $\mathbb{R}_+$  and  $\mathcal{M}(\mathbb{R} \times \mathbb{R}_+)$  denote the collection of positive finite measures on  $\mathbb{R} \times \mathbb{R}_+$ . Consider a measure-valued map  $N : \mathbb{D} \mapsto \mathcal{M}(\mathbb{R} \times \mathbb{R}_+)$  and a probability measure  $\mathbb{P}$  on  $\mathbb{D}$  such that  $N$  is a continuous real-valued  $\mathbb{R}$ -marked point process with compensator  $\lambda \nu(de)dt$ , in which  $\lambda > 0$  and  $\nu$  is a probability measure on  $\mathbb{R}$ . See e.g. [9]. 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  $\mathcal{A} := \cup_{t \geq 0} \mathcal{A}^t$ . Throughout this paper, unless otherwise stated we will work on the filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , where  $\mathcal{F} = \mathcal{F}_T^0$  for  $T > 0$  given and  $\mathbb{F} = \mathbb{F}^0$ .

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

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

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

$$(t, x, \alpha) \in [0, T] \times \mathbb{R} \times \mathcal{A} \mapsto J(t, x; \alpha) := \mathbb{E} \left[ \int_t^T r(X_{s^-}^{t,x,\alpha}, \alpha_s) dN_s \right], \quad (2.2)$$

together with the value function

$$V(t, x) := \sup_{\alpha \in \mathcal{A}^t} J(t, x; \alpha), \quad (t, x) \in [0, T] \times \mathbb{R}. \quad (2.3)$$

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. Moreover,  $(b, r)$  is bounded.*

**Remark 2.1.** *Note that boundedness of the coefficients  $b$  and  $r$  is not essential in the following arguments. One could assume only linear growth in space, 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 the theory of viscosity solutions.

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 supersolution (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  $C^1$  functions  $\varphi : [0, T) \times \mathbb{R} \mapsto \mathbb{R}$  such that  $(t, x)$  attains a minimum (resp. maximum) of  $U - \varphi$  on  $[0, T) \times \mathbb{R}$  we have

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

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

**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 super- and a subsolution in the viscosity sense 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, as  $|V(t, \cdot)| \leq (T - t) \|r\|_\infty$ .  $\square$

We next prove the existence of an optimal Markovian control. In the following, we denote by  $\mathfrak{A}$  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  $\tau_i^t$  is the  $i$ -th jump of  $N$  after time  $t$  for  $i \geq 1$ , with  $\tau_0^t := t$ , and  $\hat{a} \in \mathfrak{A}$  satisfies

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

*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 principle in (2.6), the continuity of  $V$ , and the definition of  $\hat{a}$  above,

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

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  $\omega \in \Omega$ ,

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

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

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

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

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

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  $\hat{\alpha}'(t, x) = \hat{a}(t, x) 1_{\{(t, x) \in E^c\}} + (x - \iota) 1_{\{(t, x) \in E\}}$ , for  $\iota > 0$  small. 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 integro-differential equation (2.4) is non-local and the computational cost of its numerical resolution increases as  $\lambda$  grows. On the other hand, we can expect that our pure-jump problem admits a diffusive limit as  $\lambda \rightarrow \infty$  which is, by its local nature, 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 the candidate for 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 take as  $\lambda$  the intensity

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

so that it is large for  $\epsilon > 0$  small. To ensure the existence of a diffusive limit, we need to assume that the jump coefficient  $b$  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 (3.3) below. Problems where this scaling of coefficient is appropriate involve many jumps of small relative size, with a variance of the same order as their drift over time.

Likewise, we consider the value function

$$V_\epsilon(t, x) := \sup_{\alpha \in \mathcal{A}^t} J_\epsilon(t, x; \alpha) \text{ with } J_\epsilon(t, x; \alpha) := \frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s-}^{t,x,\alpha}, \alpha_s) dN_s \right], \quad (t, x) \in [0, T] \times \mathbb{R}. \quad (3.2)$$

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  $V_\epsilon$ , together with the associated optimal policy, can be approximated by considering its diffusive limit as  $\epsilon \rightarrow 0$ . The coefficients of the associated Brownian diffusion 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  $\bar{\mathbb{P}}$  be a probability measure on  $\mathbb{D}$  and let  $W$  be a stochastic process such that  $W$  is a  $\bar{\mathbb{P}}$ -Brownian motion, 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  $\bar{\mathcal{A}}^t$  be the collection of  $\bar{\mathbb{F}}^t$ -predictable processes. Given  $\bar{a} \in \bar{\mathcal{A}}^t$ , we can then define  $\bar{X}^{t,x,\bar{a}}$  as the unique strong solution of

$$\bar{X}^{t,x,\bar{a}} = x + \int_t^\cdot \mu(\bar{X}_s^{t,x,\bar{a}}, \bar{a}_s) ds + \int_t^\cdot \sigma(\bar{X}_s^{t,x,\bar{a}}, \bar{a}_s) dW_s. \quad (3.3)$$

The candidate diffusive limit problem is then defined as

$$\bar{V}(t, x) := \sup_{\bar{a} \in \bar{\mathcal{A}}^t} \bar{\mathbb{E}} \left[ \int_t^T r(\bar{X}_s^{t,x,\bar{a}}, \bar{a}_s) ds \right], \quad (t, x) \in [0, T] \times \mathbb{R}$$

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 its 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*

$$\partial_t \bar{V} + \sup_{\bar{a} \in \bar{\mathbb{A}}} \left( \mu(\cdot, \bar{a}) \partial_x \bar{V} + \frac{1}{2} \sigma^2(\cdot, \bar{a}) \partial_{xx}^2 \bar{V} + r(\cdot, \bar{a}) \right) = 0, \quad \text{on } [0, T] \times \mathbb{R}, \quad (3.4)$$

$$\bar{V}(T, \cdot) = 0, \quad \text{on } \mathbb{R}. \quad (3.5)$$

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

$$\frac{1}{2}\eta |q - q'| \leq |F(x, p, q) - F(x, p, q')| \leq \frac{1}{2} \|\sigma\|_\infty^2 |q - q'| \quad (3.6)$$

$$vF(x, 0, 0) \leq \|r\|_\infty (1 + |v|^2) \quad (3.7)$$

$$\begin{aligned} |F(x, p, q) - F(x', p', q')| &\leq (|p| \|\mu\|_{\text{Lip}} + |q| \|\sigma\|_\infty \|\sigma\|_{\text{Lip}} + \|r\|_{\text{Lip}}) |x - x'| \\ &\quad + \|\mu\|_\infty |p - p'| + \frac{1}{2} \|\sigma\|_\infty^2 |q - q'| \end{aligned} \quad (3.8)$$

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 (3.4) 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 [20, Theorem 14.24], (3.6), (3.7) and (3.8), upon mollifying  $F$  and then passing to the limit.

b) We now 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  $h \in (0, T - t)$ ,

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

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

$$\|\partial_t \bar{V}_n\| \leq \|r\|_\infty. \quad (3.9)$$

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  $(\mu, \sigma)$ :

$$\bar{\mathbb{E}} \left[ \int_t^T |\bar{X}_s^{t, x+h, \bar{\alpha}} - \bar{X}_s^{t, x, \bar{\alpha}}| ds \right] \leq C_1 |h| \quad (3.10)$$

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  $\varphi(\bar{X}^{t,x+h,\bar{\alpha}})$  or  $\varphi(\bar{X}^{t,x,\bar{\alpha}})$  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

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

for some positive constants  $C_2$  and  $C'_2$  independent of  $n$ . Combined with (3.10), this leads to

$$\left\| \partial_x \bar{V}_n \right\|_{\infty} \leq \|r\|_{\text{Lip}} (C_1 + C'_2). \quad (3.11)$$

The fact that  $\bar{V}_n$  solves (3.4) combined with (3.6), (3.9) and (3.11) then proves that

$$\left\| \partial_{xx}^2 \bar{V}_n \right\|_{\infty} \leq C_3 \quad (3.12)$$

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

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

In view of (3.9), this implies that

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

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

$$\left| \partial_x \bar{V}_n(t', x') - \partial_x \bar{V}_n(t, x) \right| \leq C \left( |t' - t|^{\frac{\beta}{2}} + |x' - x|^{\beta} \right), \quad \text{for } (t', x') \in [0, T] \times \mathbb{R}, \quad (3.14)$$

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,

$$\left| \partial_{xx}^2 \Delta_h \bar{V}_n \right| \leq 2\eta^{-1} \left( \left| \partial_t \Delta_h \bar{V}_n \right| + |A_h| \left| \partial_x \Delta_h \bar{V}_n \right| + |C_h| h^{1-\beta} \right)$$

and we conclude from (3.13)-(3.14) that

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

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

$$|\partial_{xx}^2 \bar{V}_n(t', x) - \partial_{xx}^2 \bar{V}_n(t, x)| \leq C |t' - t|^{\frac{\beta}{2}}, \quad x \in (-n, n), \quad t, t' < T \quad (3.16)$$

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 [20, 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 (3.4)-(3.5). 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

$$\bar{a} \in \operatorname{argmax}_{a \in \mathbb{A}} \left( \mu(\cdot, a) \partial_x \bar{V} + \frac{1}{2} \sigma^2(\cdot, a) \partial_{xx}^2 \bar{V} + r(\cdot, a) \right) \quad \text{on } [0, T] \times \mathbb{R}, \quad (3.17)$$

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 [19, 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 [20, 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 Example 3.8 in Section 3.5 below.

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

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

in which

$$\delta r_\epsilon := \epsilon^{-1} (\bar{V}(\cdot, \cdot + b_\epsilon) - \bar{V}) - b_1 \partial_x \bar{V} - \frac{1}{2} |b_2|^2 \partial_{xx}^2 \bar{V} \quad (3.18)$$

satisfies

$$\|\delta r_\epsilon\|_\infty \leq C_K^\epsilon \epsilon^{\frac{\beta}{2}} \quad (3.19)$$

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,

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

where  $\delta r_\epsilon$  is a continuous function, defined in (3.18), satisfying

$$\begin{aligned} |\delta r_\epsilon| &\leq \frac{1}{2} \left( \epsilon \|b_1\|_\infty^2 + 2\sqrt{\epsilon} \|b_1 b_2\|_\infty \right) \left( \|\partial_{xx} \bar{V}\|_\infty + K(\epsilon \|b_1\|_\infty + \sqrt{\epsilon} \|b_2\|_\infty)^\beta \right) \\ &\quad + \frac{1}{2} K \epsilon^{\frac{\beta}{2}} (\sqrt{\epsilon} \|b_1\|_\infty + \|b_2\|_\infty)^{2+\beta}. \end{aligned} \quad (3.20)$$

By the left-hand side of Assumption 2, the above simplifies to

$$I = \epsilon \int \left( b_1(x, a, e) \partial_x \bar{V}(t, x) + \frac{1}{2} |b_2|^2(x, a, e) \partial_{xx} \bar{V}(t, x) + \delta r_\epsilon(x, a, e) \right) \nu(de).$$

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

$$\mu \partial_x \bar{V} + \frac{1}{2} \sigma^2 \partial_{xx}^2 \bar{V} + r = \frac{1}{\epsilon} \int (\bar{V}(\cdot, \cdot + b_\epsilon(\cdot, e)) - \bar{V}(t, x) + \epsilon(r - \delta r_\epsilon)) \nu(de). \quad (3.21)$$

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

$$\begin{cases} \partial_t \bar{V} + \sup_{a \in \mathbb{A}} \frac{1}{\epsilon} \int (\bar{V}(\cdot, \cdot + b_\epsilon(\cdot, a, e)) - \bar{V} + \epsilon(r(\cdot, a) - \delta r_\epsilon(\cdot, a, e))) \nu(de) = 0, & \text{on } [0, T) \times \mathbb{R}, \\ \bar{V}(T, \cdot) = 0, & \text{on } \mathbb{R}. \end{cases} \quad (3.22)$$

Applying Proposition 2.2 (with the appropriate coefficients), 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 (3.20). □

### 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 (3.2) can be constructed by considering a measurable map  $\bar{a} : [0, T) \times \mathbb{R} \mapsto \mathbb{A}$  satisfying

$$\bar{a} \in \operatorname{argmax}_{\bar{a} \in \mathbb{A}} \left( \mu(\cdot, \bar{a}) \partial_x \bar{V} + \frac{1}{2} \sigma^2(\cdot, \bar{a}) \partial_{xx}^2 \bar{V} + r(\cdot, \bar{a}) \right) \quad \text{on } [0, T) \times \mathbb{R}, \quad (3.23)$$

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  $(t, x) \in [0, T) \times \mathbb{R}$  and  $\epsilon > 0$ ,  $\bar{\alpha}^{t,x}$  is  $\epsilon^{\frac{\beta}{2}}$ -optimal for  $V_\epsilon$ . Namely,*

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

*Proof.* It follows from Proposition 3.1, (3.23) and (3.21) 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 (3.5) 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 (3.19). □

### 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  $(t, x, a) \in [0, T) \times \mathbb{R} \times \mathbb{A} \mapsto \epsilon^{-\frac{\beta}{2}} \delta r_\epsilon(t, x, a, e)$  is continuous, uniformly in  $(e, \epsilon) \in \mathbb{R} \times (0, 1)$ .*
- b. *The pointwise limit*

$$r_1 := \lim_{\epsilon \rightarrow 0} \int \epsilon^{-\frac{\beta}{2}} \delta r_\epsilon(\cdot, e) \nu(de), \tag{3.24}$$

*is well-defined on  $[0, T) \times \mathbb{R}$ .*

- 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*

$$\begin{cases} \partial_t \varphi + \max_{\bar{a} \in \mathbb{A}_0} \left( \mu(\cdot, \bar{a}) \partial_x \varphi + \frac{1}{2} \sigma^2(\cdot, \bar{a}) \partial_{xx}^2 \varphi + r_1(\cdot, \bar{a}) \right) = 0, & \text{on } [0, T) \times \mathbb{R} \\ \varphi(T, \cdot) = 0 & \text{on } \mathbb{R} \end{cases} \tag{3.25}$$

*and (3.25) 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 (3.19) in Theorem 3.3. The right-hand side term in (3.24) therefore admits a limsup and a liminf. The condition (3.24) 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 bounded third-order space derivative  $\partial_{xxx}^3 \bar{V}$ , then one easily checks that  $\beta = 1$  and  $r_1 = \frac{1}{6} \int |b_2(\cdot, e)|^3 \nu(\mathrm{d}e) \partial_{xxx}^3 \bar{V}$ , by a simple Taylor expansion.*
- c) *Also note that existence and uniqueness in (3.25) 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. (3.27) below holds with

$$\bar{V}_\epsilon^{(1)} := \bar{V} + \epsilon^{\frac{\beta}{2}} \delta \bar{V}^{(1)}. \quad (3.26)$$

**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*

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\beta}{2}} |V_\epsilon(t, x) - \bar{V}_\epsilon^{(1)}(t, x)| = 0. \quad (3.27)$$

*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 constant  $\delta\beta > 0$  such that*

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\delta\beta}{2}} \left\| \int \left( \epsilon^{-\frac{\beta}{2}} \delta r_\epsilon(\cdot, e) - r_1(\cdot, e) \right) \nu(\mathrm{d}e) \right\|_\infty < \infty, \quad (3.28)$$

*then the control defined by*

$$\check{\alpha}_s^{t,x} = \check{\alpha}(s, X_{s-}^{t,x, \check{\alpha}^{t,x}}), \quad s \in [t, T] \quad (3.29)$$

*with*

$$\check{\alpha} \in \operatorname{argmax}_{\bar{a} \in \mathbb{A}_0} \left\{ \mu(\cdot, \bar{a}) \partial_x \delta \bar{V}^{(1)} + \frac{1}{2} \sigma(\cdot, \bar{a})^2 \partial_{xx}^2 \delta \bar{V}^{(1)} + r_1(\cdot, \bar{a}) \right\}, \quad \text{on } [0, T] \times \mathbb{R}, \quad (3.30)$$

*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

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

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 (3.25). 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  $k > \epsilon_n^{\frac{1}{2}}(\|b_1\|_\infty + \|b_2\|_\infty)$ , the viscosity subsolution property of  $V_{\epsilon_n}$ , applying Proposition 2.2 to the test function  $\bar{V} + \epsilon_n^{\frac{\beta}{2}}\varphi$ , implies that

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

for some  $\bar{a}_n \in \mathbb{A}$ . Recalling (3.18)-(3.19),  $\bar{a}_n$  converges to some element of  $\bar{a} \in \mathbb{A}_0(t_o, x_o)$  as  $n$  goes to infinity, after possibly passing to a subsequence. By Proposition 3.1 again combined with (3.22),

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

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 (3.25) 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 (3.28) holds. Using (3.28) and the same arguments as in the proof of Theorem 3.3 lead to

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

in which

$$\delta r_\epsilon^{(1)} := \frac{1}{\epsilon} \left( \delta \bar{V}^{(1)}(\cdot, \cdot + b_\epsilon) - \delta \bar{V}^{(1)} \right) + \epsilon^{-\frac{\beta}{2}} \delta r_\epsilon - b_1 \partial_x \delta \bar{V}^{(1)} - \frac{1}{2} |b_2|^2 \partial_{xx}^2 \delta \bar{V}^{(1)} - r_1.$$

Moreover, direct computations using the above and (3.22) show that  $\bar{V}_\epsilon^{(1)}$  defined in (3.26) solves

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

on  $[0, T) \times \mathbb{R}$ , where  $\check{a}$  is defined as in (3.30). Together with (3.31), this implies that, for  $\check{\alpha}^{t,x}$  defined as in (3.29), we have

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

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

**Remark 3.7.** *If the limit in (3.24) 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 (3.25) 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 a similar approach can still 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 neighbourhood 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

$$\partial_t \delta \bar{f} + \sup_{\bar{a} \in \mathbb{A}_0} \left( \delta \bar{f} \{ \gamma \mu(\bar{a}) + \frac{1}{2} \gamma(\gamma - 1) \sigma^2(\bar{a}) \} + r_1(\cdot, \bar{a}) \right) = 0$$

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

$$(t, \bar{a}) \in [0, T] \times \mathbb{A} \mapsto r_1(t, \bar{a}) := \gamma(\gamma - 1) \ell \left( \int (b_1 b_2)(\bar{a}, e) \nu(\mathrm{d}e) \right) \bar{f}(t)$$

for some (explicit) continuous map  $\ell$  with linear growth. In particular, this shows that the convergence rate in  $\epsilon^{\frac{1}{2}}$  proved in Theorem 3.3 is sharp.

### 3.6 Higher order expansions

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

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

$$\begin{aligned} \partial_t \delta\bar{V}^{(i)} + \mu(\cdot, \check{\alpha}_\epsilon) \partial_x \delta\bar{V}^{(i)} + \frac{1}{2} \sigma(\cdot, \check{\alpha}_\epsilon)^2 \partial_{xx}^2 \delta\bar{V}^{(i)} + r_i(\cdot, \check{\alpha}_\epsilon) &= 0, \quad \text{on } [0, T] \times \mathbb{R}, \\ \delta\bar{V}^{(i)}(T, \cdot) &= 0 \quad \text{on } \mathbb{R}, \end{aligned}$$

in which  $\check{\alpha}_\epsilon$  is a Borel measurable map such that

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

with

$$\bar{V}_\epsilon^{(i_o)} := \delta\bar{V}^{(0)} + \sum_{j=1}^{i_o} \epsilon^{\frac{\beta_{j-1}}{2}} \delta\bar{V}^{(j)}, \quad \beta_i := \sum_{j=0}^i \delta\beta_j \quad \text{for } i \leq i_o,$$

and, using the conventions  $\delta\beta_{-1} := 0$  and  $\delta r_\epsilon^{(-1)} := r$ , for  $0 \leq i \leq i_o$ ,

$$\begin{aligned} \delta r_\epsilon^{(i)} &:= \frac{1}{\epsilon} (\delta\bar{V}^{(i)}(\cdot, \cdot + b_\epsilon) - \delta\bar{V}^{(i)}) + \epsilon^{-\frac{\delta\beta_{i-1}}{2}} \delta r_\epsilon^{(i-1)} - b_1 \partial_x \delta\bar{V}^{(i)} - \frac{1}{2} |b_2|^2 \partial_{xx}^2 \delta\bar{V}^{(i)} - r_i \\ r_i &:= r 1_{\{i=0\}} + 1_{\{i>0\}} \lim_{\epsilon \rightarrow 0} \int \epsilon^{-\frac{\delta\beta_{i-1}}{2}} \delta r_\epsilon^{(i-1)}(\cdot, e) \nu(\mathrm{d}e) \quad \text{for } i \leq i_o. \end{aligned} \tag{3.32}$$

The limits in (3.32) are well-defined on  $[0, T] \times \mathbb{R}$ , and

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\delta\beta_{i_o}}{2}} \left\| \int (\epsilon^{-\frac{\beta_{i_o-1}}{2}} \delta r_\epsilon^{(i_o-1)} - r_{i_o})(\cdot, e) \nu(\mathrm{d}e) \right\|_\infty < \infty. \tag{3.33}$$

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

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-\frac{\beta_{i_o}}{2}} |V_\epsilon - \bar{V}_\epsilon^{(i_o)}|(t, x) < \infty.$$

Moreover, the control defined by

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

satisfies

$$\frac{1}{\lambda_\epsilon} \mathbb{E} \left[ \int_t^T r(X_{s-}^{t,x,\check{\alpha}^{t,x}}, \check{\alpha}_s^{t,x}) dN_s \right] \geq V_\epsilon(t, x) - C\epsilon^{\frac{\beta_{i_0}}{2}}, \quad \text{for all } \epsilon > 0,$$

for some constant  $C > 0$ .

*Proof.* With the above construction

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

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

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

on  $[0, T) \times \mathbb{R}$  for all  $a \in \mathfrak{A}$ . By (3.33) and the same arguments as in the proof of Theorem 3.3,

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

so that the required result follows by verification. □

## 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 modeled 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 [16], we face in any case the fact that auctions are issued almost continuously which corresponds to a very small time step 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.* [21–23], and real-world auctions can take very complex formats. For simplicity, we consider an auctioneer which has implemented a lazy second price auction [23, 24] 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 price 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 [11]. Consequently, when there are many bidders a control will have 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, it 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, 10, 12, 15]. 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  $r_0 \in \mathbb{R}_+$ . The reserve price process  $X^{t,x,\alpha}$  is then defined from these coefficients as in (2.1), with  $b := b_\epsilon = \epsilon b_1 + \sqrt{\epsilon} b_2$  and  $\lambda = \lambda_\epsilon = \epsilon^{-1}$ . 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 prejudice to the agent, since the reserve price barely affects his profits. Therefore there is no need to compete with and control the reserve price. 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 over the value margin it has relative to other buyers. Finally, if the reserve price is above  $v$ , there is no possible profit so no reason to take part in the auction by bidding  $a > 0$ .

When the reserve price is dynamic, a good control seeks to maximise profit while pushing the reserve price to the non-competitive domain. You can see this in effect on figure 1. In the non-competitive regime (left), starting at a reserve price of 0.15, this policy recovers 85% of the best possible income of the static setting, where the reserve price is 0 for all  $t$ , and the average price is  $0.5 - \mathbb{E}[B] = 0.35$ . In the competitive regime (centre), 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.

## 4.2 Numerical implementation

Adapting (2.4) and (3.4), we normalise the horizon to 1, and allow the reserve price 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 a negative price  $x$ , 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  $X_t^{0,x,\alpha}$  descending below  $v$  by time  $T$  and generating any revenue decreases due

---

<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.

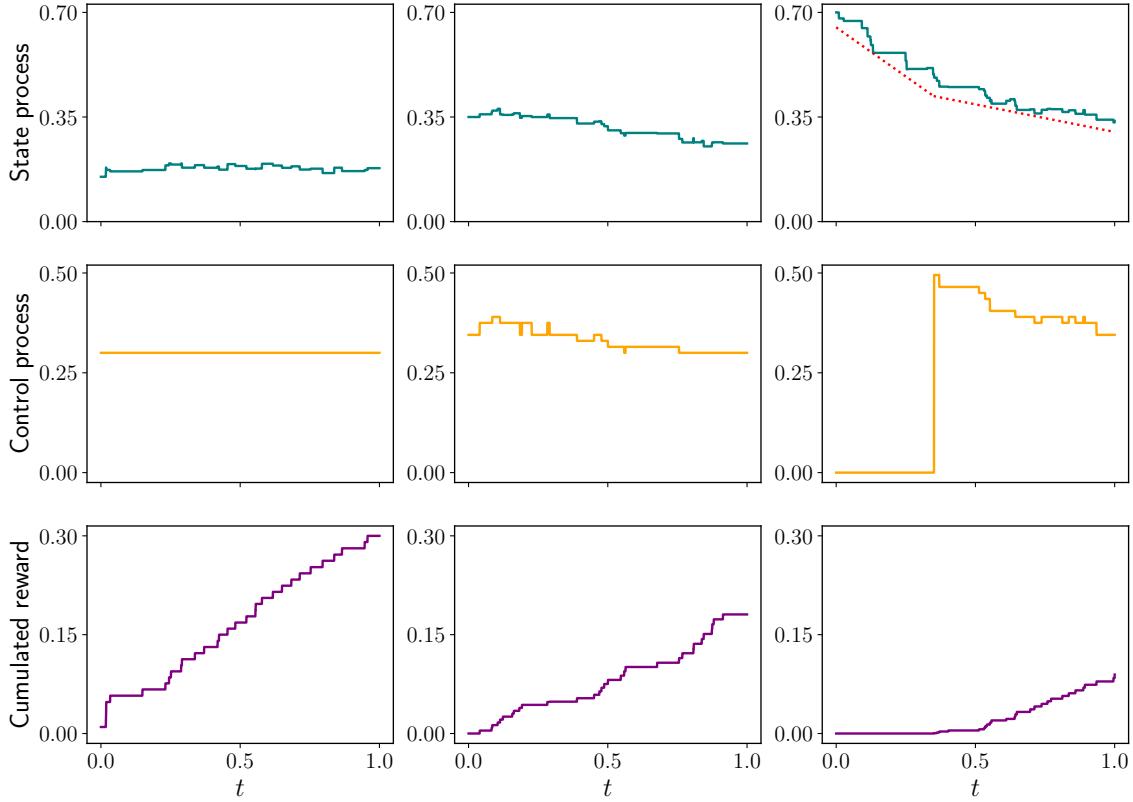


Figure 1: Selected sample realisations of the system for  $\epsilon = 10^{-1.5}$ , starting from  $x = 0.15$  (left),  $x = 0.3$  (centre), and  $x = 0.7$  (right).

to the noise. If  $\epsilon = 1$ , the first jump can always take  $X_t^{0,x,\alpha}$  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  $V_n(x_i)$  the output of the solver at time  $t_n \in \mathcal{M}_t$  and position  $x_i \in \mathcal{M}_x$ . For the pure jump problem, we



explicitly compute:

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

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  $\mathcal{M}_t = \{kd_t; k = 0, \dots, \lfloor 1/d_t \rfloor\}$ ,  $\mathcal{M}_x = \{-0.5 + kd_x; k = 0, \dots, \lfloor 3.5/d_x \rfloor\}$ , with  $d_x = 10^{-2}$ ,  $d_t = d_x^2$  and solve recursively

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

where  $\delta_x^u$  and  $\delta_{xx}$  are the uplift first order and centred second order finite differences on  $\mathcal{M}_x$  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 becomes prohibitively expensive quickly. In our example problem, the noise is supported on a bounded interval of size  $\sqrt{\epsilon}$ , so we can 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  $\epsilon$ . Even though we computed the diffusive limit to a very high precision, and with an explicit scheme, for  $\epsilon$  of order of  $10^{-3}$  the CPU time spent to is already 6 times higher in the pure-jump problem. Note that, in the pure-jump case, if the control were to intervene in a non-linear way we might need to also refine the control mesh with  $\epsilon$ , 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  $\epsilon$ .

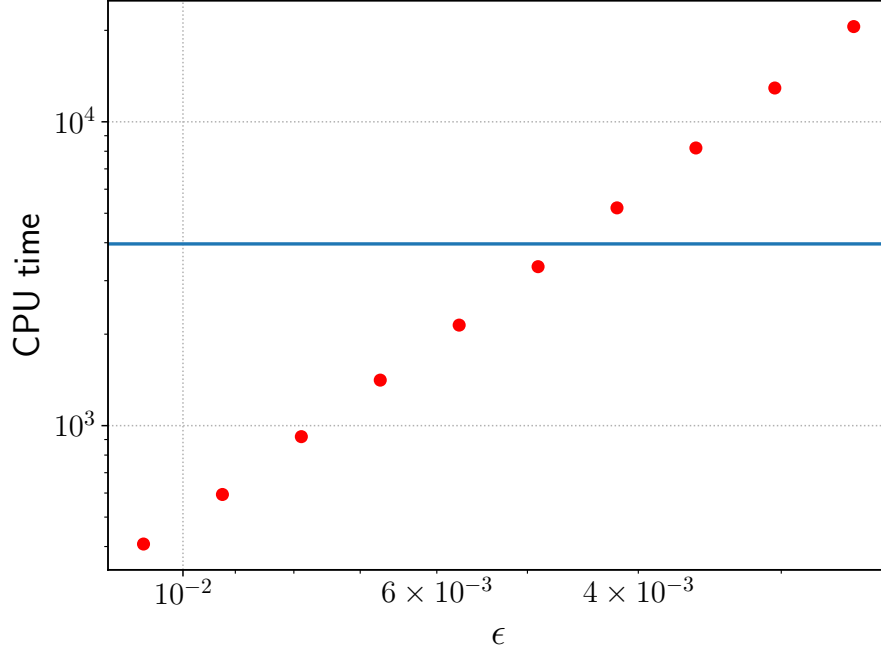


Figure 2: Numerical cost for  $V_\epsilon$  (log scales).

## 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  $X^{t,x,\alpha}$  defined by

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

and 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

$$b_n = \frac{T}{n} b_1 + \sqrt{\frac{T}{n}} b_2, \quad \text{with } \mathbb{E}[b_2(\cdot, \xi_1^n)] = 0,$$

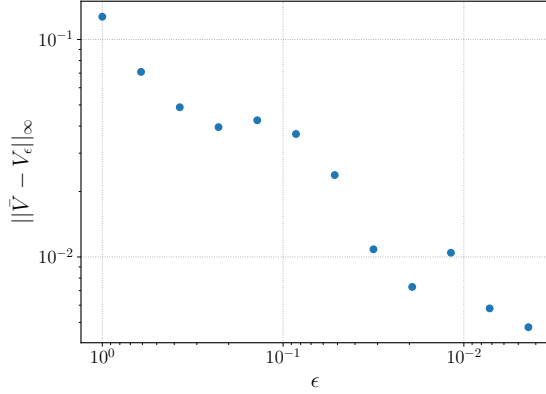


Figure 3: Limit value function error relative to  $V_\epsilon$ , at  $t = 0$  (log scales).

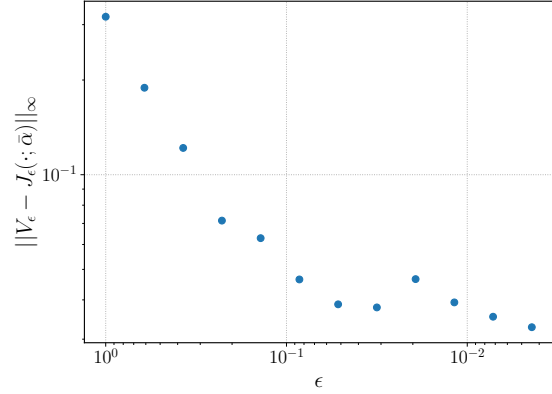


Figure 4: Limit policy error relative to  $V_\epsilon$ , at  $t = 0$  (log scales).

one would obtain the same diffusive limit as in Section 3.3 when letting  $n \rightarrow \infty$ . Namely, the same arguments as in [17, Section 3] combined with Proposition 3.1 and the fact that comparison holds for (3.4) 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

$$\|\delta r_n\|_\infty \leq C n^{-\frac{\beta}{2}} \quad (5.1)$$

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

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

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) and (5.1) 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

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

which provides the expected result.

Likewise, the Markovian control defined through (3.23) 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  $L^1$ -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] P. Brémaud. *Point processes and queues: martingale dynamics*, Springer series in statistics, 50, Springer, 1981.
- [10] 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.
- [11] J. Bulow and P. Klemperer. Auctions versus negotiations. *American Economic Review*, 86(1):180–94, 1996.
- [12] 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.
- [13] H. Chen, D. Yao. *Fundamentals of Queueing Networks*. Stochastic Modeling and Applied Probability, 2001.
- [14] 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.
- [15] 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.
- [16] J. Fernandez-Tapia, O. Guéant, and J-M. Lasry. Optimal Real-Time Bidding Strategies. *Applied Mathematics Research eXpress*, 2017(1):142–183, 09 2016.
- [17] 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.
- [18] J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.
- [19] O. A. Ladyzhenskaia, V. Solonnikov, and N. N. Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.
- [20] G. M. Lieberman. *Second order parabolic differential equations*. World scientific, 1996.
- [21] R. B. Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.
- [22] 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.

- [23] 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.
- [24] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.