

Dynamic Programming for Mean-Field Type Control

Mathieu Lauriere, Olivier Pironneau

► **To cite this version:**

Mathieu Lauriere, Olivier Pironneau. Dynamic Programming for Mean-Field Type Control. Journal of Optimization Theory and Applications, Springer Verlag, 2015, 169 (3), pp.902-924. <10.1007/s10957-015-0785-x>. <hal-01558745>

HAL Id: hal-01558745

<http://hal.upmc.fr/hal-01558745>

Submitted on 9 Jul 2017

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.



Dynamic Programming for Mean-Field Type Control

Mathieu Laurière · Olivier Pironneau*

Communicated by Nizar Touzi

JOTA VOL.169, No.3, JUNE 2016

Abstract We investigate a model problem for optimal resource management. The problem is a stochastic control problem of mean-field type. We compare an Hamilton-Jacobi-Bellman fixed point algorithm to a Steepest Descent method issued from calculus of variations.

For mean-field type control problems, stochastic dynamic programming requires adaptation. The problem is reformulated as a distributed control problem by using the Fokker-Planck equation for the probability distribution of the stochastic process; then, an extended Bellman's principle is derived by a different argument than the one used by P.L. Lions. Both algorithms are compared numerically.

Keywords Stochastic control · mean-field game · Hamilton-Jacobi-Bellman

PACS 02.50.-r · 02.60.Cb · 2.60.Pn

Mathematics Subject Classification (2000) 35Q93 · 37N35 · 49J20

1 Introduction

Stochastic control has been studied extensively over the past five decades [1,2,3,4,5], and yet there is a renewed interest in economy and finance due to mean-field games [6,7,8,9]. Mean-field games give rise to mean-field type stochastic control problems [10] which involve not only the Markov process of the state of the system, some statistics of the process like means and variance, in the cost function or in the stochastic differential equation (SDE). For these problems, optimality conditions are derived either by stochastic calculus of variation [11] or by stochastic dynamic programming [12,13] and justified in the quadratic case by classical arguments [14,15],

Mathieu Laurière
LJLL, Université Denis Diderot (Paris 7)

*Olivier Pironneau (corresponding author)
LJLL, Université Pierre et Marie Curie (Paris 6).
E-mail: Olivier.Pironneau@UPMC.fr

but not so classical one in the general case for the fundamental reason that Bellman's principle does not apply in its original form [12, 16].

Several authors have generalized Dynamic Programming using Wasserstein distance to define derivatives with respect to measures. Others have studied the existence of a conceptual HJB equation [16, 17]. These results certainly overlap and precede our analysis but our point of view is different : it is pragmatic, so we sacrifice mathematical rigor to explicit expressions, and numerical, as we wish to compare solutions obtained by HJB to standard optimal control by calculus of variations. We have not tried to specify regularity of data for existence and differentiability of solutions. The results are stated formally only but with an intuitive feeling that they could be justified later with appropriate assumptions as in e.g. [18, 19] *if* the behavior at infinity of the solution of the HJB equation is known, which is a major riddle.

Before proceeding further, note that a direct simulation of the problem with the stochastic differential equation approximated by Monte-Carlo is too costly and not competitive with the methods that we pursue below. Indeed, the cost function of the optimization problem involves means of stochastic quantities and Monte-Carlo methods would require large numbers of evaluations of the SDE, embedded in forward-backward time loops. Faced with the same problem, Garnier *et al.* [7] and Chan *et al.* [20] came to the same conclusion.

In this article, pursuing a preliminary study published in [21], we apply the dynamic programming argument to the value functional as in [22], but instead of using the probability measure of the stochastic process, we use its probability density function (PDF). Hence, though less general, the mathematical argument will be simpler. Of course this is at the cost of several regularity assumptions, such as the existence of a regular PDF at all times. However, our analysis here is strongly motivated by the numerical solutions of these control problems, and hence assuming regularity is not a real practical limitation.

Once the problem is reformulated with the Fokker-Planck equation [1, 23], it becomes a somewhat standard exercise to find the optimality necessary conditions by a calculus of variations. So this article begins likewise in Section 3. Then, in Section 4, a similar result is obtained by using dynamic programming, and the connection with the previous approach and with stochastic dynamic programming is established. In Section 5 a problem introduced in [24] involving profit-optimizing oil producers is defined and studied for existence and optimality, and two algorithms are proposed together with a semi-analytical method based on a Riccati solution. The paper ends with a numerical Section which implements the three methods and compare them.

2 The Problem

Consider a stochastic differential system of d -equations

$$dX_t = u(X_t, t)dt + \sigma(X_t, t, u(X_t, t))dW_t, \quad t \in [0, T], \quad (1)$$

where u takes values in \mathbb{R}^d , σ in $\mathbb{R}^{d \times k}$ and W_t is a k -vector of independent Brownian motions. Assumptions for which a strong solution is known to exist once the distribution of X_0 is known to be in $L^2 \cap L^\infty$ are given in [25] (see also [26], Proposition 4

which applies when $\sigma\sigma^T$ is uniformly positive definite:

$$\begin{aligned} u \in L^2([0, T[, L^2_{loc}(\mathbb{R}^d))^d, \frac{u}{1+|x|} \in L^1([0, T[, L^1 + L^\infty(\mathbb{R}^d))^d, \\ \nabla \cdot u \in L^1([0, T[, L^\infty(\mathbb{R}^d)), \end{aligned} \quad (2)$$

and with $\tilde{\sigma}(x, t) := \sigma(x, t, u(x, t))$:

$$\tilde{\sigma} \in L^\infty([0, T[, L^\infty_{loc}(\mathbb{R}^d))^{d \times k}, \frac{\tilde{\sigma}}{1+|x|} \in L^2([0, T[, L^2 + L^\infty(\mathbb{R}^d))^{d \times k} \quad (3)$$

Then the PDF of X_t satisfies

$$\rho \in L^\infty([0, T[, L^2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)) \cap L^2([0, T[, H^1(\mathbb{R}^d)), \sigma^T \nabla \rho \in L^2([0, T[, L^2(\mathbb{R}^d))$$

and is the unique solution of the Fokker-Planck equation,

$$\partial_t \rho + \nabla \cdot (u\rho) - \nabla^2 : \left(\frac{1}{2} \sigma \sigma^T \rho\right) = 0, \quad \rho|_{t=0} = \rho_0(x), \quad (4)$$

Conversely, with (2)(3), there is a unique solution to (4) which is the PDF of a Markov process which satisfies (1).

Remark 2.1 The assumption $\sigma\sigma^T > 0$ can be replaced by assuming the conditions (2) for $u - \frac{1}{2} \nabla \cdot (\tilde{\sigma} \tilde{\sigma}^T)$ rather than u , but it implies some regularity on the second derivatives of u .

Consider the stochastic optimization problem

$$\min_{u \in \mathcal{U}_d} J(u) := J_0(u) \quad (5)$$

subject to (1) with ρ_0 given, and

$$J_\tau(u) := \int_\tau^T \mathbf{E}[\tilde{H}(X_t, t, u(X_t, t), \rho(X_t, t), \mathbf{E}[\tilde{h}(X_t, t, u(X_t, t))])] dt + \mathbf{E}[G(X_T, \mathbf{E}[g(X_T)])] \quad (6)$$

where \tilde{h} , g , \tilde{H} , G are C^1 functions taking values in $\mathbb{R}^r, \mathbb{R}^s, \mathbb{R}$ and \mathbb{R} , respectively. Assume also that \tilde{H} and G are bounded from below and

$$\mathcal{U}_d = \{u \text{ verifies (2)} : u(x, t) \in \mathcal{V}_d \forall x, t\} \text{ for some } \mathcal{V}_d \subset \mathbb{R}^d.$$

As a first approach to solve such problems, Andersson *et al.* [11] proposed to use a Stochastic Calculus of Variations; the necessary optimality conditions is a forward-backward stochastic differential system, which is numerically very hard because the backward volatility for the adjoint equation is one of the unknowns [11, 27].

A second approach is to use Stochastic Dynamic Programming (SDP), but an important adaptation needs to be made. Usually SDP uses the remaining cost function

$$V(\tau, x) = \min_{u \in \mathcal{U}_d} J_\tau(u) \text{ subject to (1), with } X_\tau = x.$$

The Bellman equation is derived by saying that $u_t, t > \tau$ is a solution only if, together with (1) and $X_\tau = x$,

$$V(\tau, x) = \min_{u \in \mathcal{U}_d} \mathbf{E} \left[\int_{\tau}^{\tau+\delta\tau} \tilde{H}(X_t, t, u(X_t, t), \rho(X_t, t), \mathbf{E}[\tilde{h}(X_t, t, u(X_t, t))]) dt + V(\tau + \delta\tau, X_{\tau+\delta\tau}) \right].$$

However, the above is not true unless $\tilde{h} = 0$, $g = 0$; as in [22], one has to work with

$$V(\tau; \rho_\tau(\cdot)) = \min_{u \in \mathcal{U}_d} J_\tau(u) \text{ subject to (1), with } \rho_\tau \text{ given,}$$

where V is a pointwise function of $\tau \in [0, T]$ and has functional dependence on $\rho_\tau(\cdot)$, i.e., depends on $\{x \mapsto \rho_\tau(x), \forall x \in \mathbb{R}^d\}$.

A third approach is to work with the deterministic version of the problem. With sufficient regularity (see [28] for weaker assumptions), with the Fokker-Planck partial differential equation for $\rho(x, t) := \rho_t(x)$; the problem is equivalent to the deterministic distributed control problem,

$$\begin{aligned} \min_{u \in \mathcal{U}_d} J &= \int_0^T \int_{\mathbb{R}^d} \tilde{H}(x, t, u(x, t), \rho(x, t), \chi(t)) \rho(x, t) dx dt + \int_{\mathbb{R}^d} G(x, \xi) \rho(x, T) dx, \\ \text{where } \chi(t) &:= \int_{\mathbb{R}^d} \tilde{h}(x, t, u(x, t), \rho(x, t)) \rho(x, t) dx, \quad \xi := \int_{\mathbb{R}^d} g(x) \rho(x, T) dx \\ \text{and } \partial_t \rho + \nabla \cdot (u \rho) - \nabla^2 : (\mu \rho) &= 0, \quad \rho|_{t=0} = \rho_0(x), \end{aligned} \quad (7)$$

where $\mu_{ij} = \frac{1}{2} \sum_k \sigma_{ik} \sigma_{jk}$ and ∇^2 is the $d \times d$ matrix operator of element ∂_{ij} . The notation $A : B$ stands for $\sum_{i,j=1}^d A_{ij} B_{ij}$ and $\nabla \cdot u$ stands for $\sum_{i=1}^d \partial_i u_i$.

Remark 2.2 Note that the problem is equivalent to the Stochastic control problem only if ρ_0 is in \mathbf{P} , the set of positive real-valued functions with measure 1. However, the deterministic control problem still makes sense even if it is not the case and $\rho \in L^2([0, T], H^1(\mathbb{R}^d))$ only. We will use this in the proof of Proposition 4.3.

Remark 2.3 Existence of solutions for (6) or (7) requires more assumptions on \tilde{H}, \tilde{h}, G and g to make sure that J is lower semi-continuous with respect to u in a norm such as $L^2([0, T], H^1(\mathbb{R}^d))$ which implies (2) and \mathcal{U}_d closed. However since there is no term containing a gradient of u in J , a Tikhonov regularization seems to be needed for these problems to be well posed.

3 Calculus of Variations

Deriving optimality conditions by calculus of variations is fairly standard, as we shall show in this section.

Proposition 3.1 *A control u is optimal for (7) only if for $\forall t \in]0, T[, \forall v \in \mathcal{V}_d$,*

$$\int_{\mathbb{R}^d} \left(\tilde{H}'_u + \tilde{h}'_u \int_{\mathbb{R}^d} \tilde{H}'_\chi \rho dx + \nabla \rho^* - \mu'_u : \nabla^2 \rho^* \right) (v - u) \rho dx \geq 0 \quad (8)$$

$$\text{with } \partial_t \rho^* + u \nabla \rho^* + \mu : \nabla^2 \rho^* = - \left[\tilde{H}'_\rho \rho + \tilde{H} + (\tilde{h}'_\rho \rho + \tilde{h}) \int_{\mathbb{R}^d} \tilde{H}'_\chi \rho dx \right], \quad (9)$$

$$\rho^*|_T = g \int_{\mathbb{R}^d} G'_\xi \rho|_T dx + G,$$

where $\tilde{H}'_u, \tilde{H}'_\rho, \tilde{H}'_\chi, \tilde{h}'_u, \tilde{h}'_\rho$ and G'_ξ are partial derivative in the classical sense.

Proof Recall that (7) is in $Q =]-\infty, +\infty[^d \times]0, T]$. The Fokker-Planck equation (7) being set in \mathbf{P} , it contains the ‘‘boundary condition’’ $\forall t \in]0, T], \lim_{|x| \rightarrow \infty} \rho(x, t) = 0$.

Consider an admissible variation $\lambda \delta u$, i.e. $u + \lambda \delta u \in \mathcal{U}_d$ for all $\lambda \in [0, 1]$. Such a variation induces a variation $\lambda \delta \rho$ of ρ given by

$$\partial_t \delta \rho + \nabla \cdot (u \delta \rho + \rho \delta u + \lambda \delta u \delta \rho) - \nabla^2 : (\mu \delta \rho + \mu'_u \delta u (\rho + \lambda \delta \rho)) = 0, \quad (10)$$

with $\delta \rho|_{t=0} = 0$ and where μ'_u is evaluated at $x, t, u + \theta \delta u$ for some $\theta \in [0, \lambda]$. We assume that there is enough regularity for the solution of the Fokker-Planck equation in (7) to depend continuously on the data u, μ . Then (10) at $\lambda = 0$ defines $\delta \rho$. Also, up to higher order terms,

$$\begin{aligned} \delta \chi &= \int_{\mathbb{R}^d} [(\tilde{h}'_u \delta u + \tilde{h}'_\rho \delta \rho) \rho + \tilde{h} \delta \rho], \quad \delta \xi = \int_{\mathbb{R}^d} [g \delta \rho|_T] \\ \delta J &= \int_Q [(\tilde{H}'_u \delta u + \tilde{H}'_\rho \delta \rho + \tilde{H}'_\chi \delta \chi) \rho + \tilde{H} \delta \rho] + \int_{\mathbb{R}^d} [G'_\xi \delta \xi \rho|_T + G \delta \rho|_T] \\ &= \int_Q [(\tilde{H}'_u + \tilde{h}'_u \int_{\mathbb{R}^d} [\tilde{H}'_\chi \rho]) \rho \delta u] + \int_Q [(\tilde{H}'_\rho \rho + \tilde{H} + (\tilde{h}'_\rho \rho + \tilde{h}) \int_{\mathbb{R}^d} [\tilde{H}'_\chi \rho]) \delta \rho] \\ &\quad + \int_{\mathbb{R}^d} [(\int_{\mathbb{R}^d} [G'_\xi \rho|_T] g + G) \delta \rho|_T]. \end{aligned}$$

Define an adjoint state ρ^* by (9). Then (9) multiplied by $\delta \rho$ and integrated on Q gives

$$\begin{aligned} &\int_Q \delta \rho \left[\tilde{H}'_\rho \rho + \tilde{H} + (\tilde{h}'_\rho \rho + \tilde{h}) \int_{\mathbb{R}^d} [\tilde{H}'_\chi \rho] \right] \\ &= - \int_Q \delta \rho [\partial_t \rho^* + u \nabla \rho^* + \mu : \nabla^2 \rho^*] \\ &= \int_Q [\rho^* (\partial_t \delta \rho + \nabla \cdot (u \delta \rho) - \nabla^2 : (\mu \delta \rho))] - \int_{\mathbb{R}^d} \rho^* \delta \rho|_0^T \\ &= - \int_Q [\rho^* \nabla \cdot (\rho \delta u - \nabla \cdot (\mu'_u \delta u \rho))] - \int_{\mathbb{R}^d} [(g \int_{\mathbb{R}^d} [G'_\xi \rho|_T] + G) \delta \rho|_T]. \end{aligned}$$

Hence $\delta J = \int_Q \left[(\tilde{H}'_u + \tilde{h}'_u \int_{\mathbb{R}^d} [\tilde{H}'_\chi \rho]) \rho \delta u \right] - \int_Q [\rho^* \nabla \cdot (\rho \delta u - \nabla \cdot (\mu'_u \delta u \rho))]$

$$= \int_Q \left[\left(\tilde{H}'_u + \tilde{h}'_u \int_{\mathbb{R}^d} [\tilde{H}'_\chi \rho] + \nabla \rho^* - \nabla^2 \rho^* : \mu'_u \right) \rho \delta u \right].$$

The rest follows by saying that u is a minimum. \square

Remark 3.1 The ‘‘boundary conditions at infinity’’ on ρ^* in (9) are problematic. The PDE is to be understood in weak form in the dual of \mathbf{P} , i.e., for all $v \in \mathbf{P}$:

$$\int_Q [v (\partial_t \rho^* + u \nabla \rho^*) - \nabla v \cdot (\mu \nabla \rho^*)] = - \int_Q \left[v [\tilde{H}'_\rho \rho + \tilde{H} + (\tilde{h}'_\rho \rho + \tilde{h}) \int_{\mathbb{R}^d} \tilde{H}'_\chi \rho dx] \right]. \quad (11)$$

Then, the decay of v at infinity will balance the potential growth of ρ^* , and the integrations by parts in the proof above will have no terms at infinity in $\|x\|$. However, the uniqueness of ρ^* is not guaranteed. We could have avoided the problem by working in $[-L, L]^d \times [0, T]$ rather than $\mathbb{R}^d \times [0, T]$ and imposing $\rho^* = 0$ on the boundary of $[-L, L]^d$, but then the solution depends strongly on L ; this problem will be rediscussed in the numerical section (Section 5).

4 Dynamic Programming

For notational clarity consider the more general case, where H, G are functionals of $\rho_t(\cdot)$. Let ρ be solution of (7) initialized at τ by $\rho_\tau(\cdot)$ and let J and V be defined as :

$$\begin{aligned} J(\tau; \rho_\tau, u) &:= \int_\tau^T \int_{\mathbb{R}^d} H(x, t, u(x, t); \rho_t) \rho_t(x) dx dt + \int_{\mathbb{R}^d} G(x; \rho|_T) \rho|_T dx, \\ V(\tau; \rho_\tau) &:= \min_{u \in \mathcal{U}_d} J(\tau; \rho_\tau, u). \end{aligned} \quad (12)$$

By the Markovian property, $\rho_t(x) := \rho(x, t)$, for $t > \tau$, is also the PDF of X_t given by (1) knowing its PDF ρ_τ at τ .

Remark 4.1 In this section, H is a pointwise function of $u(x, t) \in \mathbb{R}^d$, but the theory can be extended to the case where H is a functional of $u(\cdot, \cdot) : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$.

Assuming that J is bounded from below, V is finite and we prove the following version of *Bellman's principle of optimality* :

Proposition 4.1 *If the problem is regular, then for any $\tau \in [0, T]$ and any $\rho_\tau \in \mathbf{P}$, we have :*

$$\begin{aligned} V(\tau; \rho_\tau) &= \min_{u \in \mathcal{U}_d} \left\{ \int_\tau^{\tau+\delta\tau} \int_{\mathbb{R}^d} H(x, t, u(x, t); \rho_t) \rho_t(x) dx dt + V(\tau + \delta\tau; \rho_{\tau+\delta\tau}) \right\}, \\ &\text{subject to } \rho_t, \text{ given by (7) on } [\tau, \tau + \delta\tau], \text{ initialized by } \rho_\tau \text{ at time } \tau. \end{aligned}$$

Proof Denote the infimum of the right-hand side by $\bar{V}(\tau; \rho_\tau)$. For any $\varepsilon > 0$, there exists $u \in \mathcal{U}_d$ such that, if ρ_t is the solution of (7) with control u :

$$\begin{aligned} V(\tau; \rho_\tau) + \varepsilon &> J(\tau; \rho_\tau, u) = \int_\tau^T \int_{\mathbb{R}^d} H(x, t, u(x, t); \rho_t) \rho_t dx dt \\ &+ \int_{\mathbb{R}^d} G(x; \rho|_T) \rho|_T dx = \int_\tau^{\tau+\delta\tau} \int_{\mathbb{R}^d} H \rho_t + \int_{\tau+\delta\tau}^T \int_{\mathbb{R}^d} H \rho_t \\ &+ \int_{\mathbb{R}^d} G \rho|_T \geq \int_\tau^{\tau+\delta\tau} \int_{\mathbb{R}^d} H \rho_t + V(\tau + \delta\tau; \rho_{\tau+\delta\tau}) \geq \bar{V}(\tau; \rho_\tau). \end{aligned}$$

Conversely, given $u \in \mathcal{U}_d$ and $\varepsilon > 0$, there exists a control $\tilde{u} \in \mathcal{U}_d$, which coincides with u on $\mathbb{R}^d \times [\tau, \tau + \delta\tau]$, such that:

$$J(\tau + \delta\tau; \tilde{\rho}_{\tau+\delta\tau}, \tilde{u}) \leq V(\tau + \delta\tau; \tilde{\rho}_{\tau+\delta\tau}) + \varepsilon,$$

where $\tilde{\rho}_t$ is the solution of (7) at t with control \tilde{u} starting with ρ_τ at time τ . Hence :

$$\begin{aligned} V(\tau; \rho_\tau) &= V(\tau; \tilde{\rho}_\tau) \leq J(\tau; \tilde{\rho}_\tau, \tilde{u}) \\ &= \int_\tau^{\tau+\delta\tau} \int_{\mathbb{R}^d} H(x, t, u(x, t); \tilde{\rho}_t) \tilde{\rho}_t(x) dx dt + J(\tau + \delta\tau; \tilde{\rho}_{\tau+\delta\tau}, \tilde{u}) \\ &\leq \int_\tau^{\tau+\delta\tau} \int_{\mathbb{R}^d} H(x, t, u(x, t); \tilde{\rho}_t) \tilde{\rho}_t(x) dx dt + V(\tau + \delta\tau; \tilde{\rho}_{\tau+\delta\tau}) + \varepsilon. \end{aligned}$$

To conclude, let $\varepsilon \rightarrow 0$ and take the infimum over $u \in \mathcal{U}_d$. □

From now on, we assume that H and V are Fréchet differentiable with respect to ρ .

Remark 4.2 The correct mathematical tool for this differentiability is the Wasserstein distance and the differentiability with respect to the probability measure rather than to its density (see e.g. [13, 22]). Our approach is more pragmatic.

We denote the Fréchet derivatives by $H'_\rho(x, \tau; \rho)$ and $V'_\rho(\tau; \rho)$. Thus, and similarly for V , $H'_\rho(x, \tau; \rho)$ denotes the linear application $\mathbf{L}^2 \mapsto \mathbb{R}$ such that :

$$H(x, \tau; \rho + v) = H(x, \tau; \rho) + H'_\rho(x, \tau; \rho) \cdot v + o(\|v\|_2), \quad \forall v \in \mathbf{L}^2,$$

where $\mathbf{L}^2 := L^2(\mathbb{R}^d)$. Moreover, we denote with a prime the Riesz representative of the Fréchet derivative with respect to ρ . For instance, $V' : [0, T] \times \mathbf{L}^2 \rightarrow \mathbf{L}^2$ is defined by :

$$\int_{\mathbb{R}^d} V'(\tau; \rho)(x) v(x) dx := V'_\rho(\tau; \rho) \cdot v \quad \forall v \in \mathbf{L}^2.$$

Proposition 4.2 (*HJB minimum principle*). *Assuming that V' is smooth enough, the following holds :*

$$\begin{aligned} 0 &= \min_{v \in \mathcal{V}_d} \int_{\mathbb{R}^d} \left[H(x, \tau, v; \rho_\tau) + H'_\rho(x, \tau, v; \rho_\tau) \cdot \rho_\tau \right. \\ &\quad \left. + \partial_\tau V'(\tau; \rho_\tau) + v \cdot \nabla V'(\tau; \rho_\tau) + \mu : \nabla^2 V'(\tau; \rho_\tau) \right] \rho_\tau dx. \end{aligned} \quad (13)$$

Note: As usual, ∇ is with respect to x .

Proof A first order approximation of the time derivative in the Fokker-Planck equation gives

$$\delta_\tau \rho := \rho_{\tau+\delta\tau} - \rho_\tau = \delta\tau \left[\nabla^2 : (\mu_\tau \rho_\tau) - \nabla \cdot (u_\tau \rho) \right] + o(\delta\tau). \quad (14)$$

As V is assumed to be smooth, we have :

$$V(\tau + \delta\tau; \rho_{\tau+\delta\tau}) = V(\tau; \rho_\tau) + \partial_\tau V(\tau; \rho_\tau) \delta\tau + V'_\rho(\tau; \rho_\tau) \cdot \delta_\tau \rho + o(\delta\tau). \quad (15)$$

Using (15) and the mean value theorem for the time integral, Bellman's principle yields , up to $o(\delta\tau)$,

$$V(\tau; \rho_\tau) = \min_{u \in \mathcal{U}_d} \left\{ \delta\tau \int_{\mathbb{R}^d} H \rho_\tau dx + V(\tau; \rho_\tau) + \partial_\tau V(\tau; \rho_\tau) \delta\tau + V'_\rho(\tau; \rho_\tau) \cdot \delta_\tau \rho \right\}. \quad (16)$$

The terms $V(\tau; \rho_\tau)$ cancel; divided by $\delta\tau$ and combined with (14) and letting $\delta\tau \rightarrow 0$, (16) gives

$$0 = \min_{u \in \mathcal{U}_d} \left\{ \int_{\mathbb{R}^d} H \rho_\tau dx + \partial_\tau V(\tau; \rho_\tau) + V'_\rho(\tau; \rho_\tau) \cdot [\nabla^2 : (\mu_\tau \rho_\tau) - \nabla \cdot (u_\tau \rho)] \right\}. \quad (17)$$

To finalize the proof we need the following proposition to relate V to V'_ρ .

Proposition 4.3 *Given $\tau \in [0, T]$ and an initial $\rho_\tau \in \mathbf{P}$, let $\hat{u} \in \mathcal{U}_d$ and $\hat{\rho}$ denote an optimal control and the corresponding solution of (7), then:*

$$\begin{aligned} \int_{\mathbb{R}^d} V'(\tau; \rho_\tau) \rho_\tau dx &= V(\tau; \rho_\tau) + \int_\tau^T \int_{\mathbb{R}^d} \left(H'_\rho(x, t, \hat{u}; \hat{\rho}_t) \cdot \hat{\rho}_t \right) \hat{\rho}_t dx dt \\ &\quad + \int_{\mathbb{R}^d} \left(G'_\rho(x; \hat{\rho}_T) \cdot \hat{\rho}_T \right) \hat{\rho}_T dx. \end{aligned} \quad (18)$$

Proof Note that the Fokker-Planck equation implies the existence of a semi-group operator \mathbf{G} such that, for all $\tau \leq t$, $\rho_t = \mathbf{G}(t - \tau) * \rho_\tau$. Let $(\hat{u}_t)_{t \in [0, T]}$ be the optimal control and $(\hat{\rho}_t)_{t \in [0, T]}$ the corresponding solution of (7) and (12). Then :

$$\begin{aligned} V(\tau; \hat{\rho}_\tau) &= \int_\tau^T \int_{\mathbb{R}^d} H(x, t, \hat{u}; \mathbf{G}(t - \tau) * \hat{\rho}_\tau) \mathbf{G}(t - \tau) * \hat{\rho}_\tau dx dt \\ &\quad + \int_{\mathbb{R}^d} G(x; \hat{\rho}_T) \hat{\rho}_T dx. \end{aligned} \quad (19)$$

By the optimality of \hat{u} and $\hat{\rho}$, this can be Fréchet-differentiated with respect to ρ by computing $\frac{d}{d\lambda} [V(\tau; \hat{\rho}_\tau + \lambda v) - V(\tau; \hat{\rho}_\tau)]$. The result is :

$$\begin{aligned} V'_\rho(\tau; \hat{\rho}_\tau) \cdot v &= \int_\tau^T \int_{\mathbb{R}^d} H(x, t, \hat{u}; \mathbf{G}(t - \tau) * \hat{\rho}_\tau) \mathbf{G}(t - \tau) * v \\ &\quad + \int_{\mathbb{R}^d} G(x; \hat{\rho}_T) \mathbf{G}(T - \tau) * v dx \\ &\quad + \int_\tau^T \int_{\mathbb{R}^d} \left(H'_\rho(x, t, \hat{u}; \mathbf{G}(t - \tau) * \hat{\rho}_\tau) \cdot [\mathbf{G}(t - \tau) * v] \right) \mathbf{G}(t - \tau) * \hat{\rho}_\tau \\ &\quad + \int_{\mathbb{R}^d} \left(G'_\rho(x; \mathbf{G}(T - \tau) * \hat{\rho}_\tau) \cdot [\mathbf{G}(T - \tau) * v] \right) \mathbf{G}(T - \tau) * \hat{\rho}_\tau. \end{aligned}$$

Taking $v = \hat{\rho}_\tau$ leads to (18).

One may object however that such choice for v is not admissible because being a variation of ρ_τ it has to have zero measure, but we discussed this in Remark 2.2. \square

End of proof of Proposition 4.2 Differentiating (18) with respect to τ leads to

$$\partial_\tau V(\tau; \rho_\tau) = \int_{\mathbb{R}^d} \partial_\tau V'(\tau; \rho_\tau) \rho_\tau dx + \int_{\mathbb{R}^d} \left(H'_\rho(x, \tau, \hat{u}_\tau; \rho_\tau) \cdot \rho_\tau \right) \rho_\tau dx,$$

where \hat{u}_τ is the optimal control at time τ . Now, let us use (17), rewritten as

$$\begin{aligned} 0 &= \min_{u_\tau} \left\{ \int_{\mathbb{R}^d} \left(H(x, \tau, u_\tau(x); \rho_\tau) + H'_\rho(x, \tau, u_\tau(x); \rho_\tau) \cdot \rho_\tau \right) \rho_\tau dx \right. \\ &\quad \left. + \int_{\mathbb{R}^d} \left(\partial_\tau V'(\tau; \rho_\tau) \rho_\tau + V'(\tau; \rho_\tau) [\nabla^2 : (\mu_\tau \rho_\tau) - \nabla \cdot (u_\tau \rho_\tau)] \right) dx \right\}. \end{aligned}$$

Integrating by parts the last two terms leads to (13). \square

Remark 4.3 Note that (18) and (12) imply :

$$\int_{\mathbb{R}^d} V'_T \hat{\rho}_T dx = V(T, \hat{\rho}_T) = \int_{\mathbb{R}^d} (G + g \int_{\mathbb{R}^d} \partial_\xi G \hat{\rho}_T dx) \hat{\rho}_T dx. \quad (20)$$

Remark 4.4 By taking $\rho_\tau = \delta(x - x_0)$, the Dirac mass at x_0 , the usual HJB principle is found if $h \equiv g \equiv 0$.

Proposition 4.4 (*Hamilton-Jacobi-Bellman equation*) Denote an optimal control \hat{u} . When $\mathcal{V}_d = \mathbb{R}^d$, (13) in Proposition 4.2 gives

$$\int_{\mathbb{R}^d} (H + H' \cdot \hat{\rho}_\tau + \partial_\tau V' + \hat{\mu} : \nabla_x^2 V' + \hat{u} \cdot \nabla_x V') \hat{\rho}_\tau dx = 0 \quad (21)$$

$$\text{and} \quad \nabla_u H + \nabla_u H' \cdot \hat{\rho}_\tau + \nabla_x V' + \nabla_u \mu : \nabla^2 V' = 0, \quad (22)$$

where the second equation is in fact the first order optimality condition in (13).

Remark 4.5 When the Hamiltonian depends on the distribution ρ_t only through the local value $\rho_t(x)$ and the average of a fixed function, we can make explicit the link with the calculus of variations (see Section 3). More precisely, let us assume that $H = \tilde{H}(x, t, u, \rho_t(x), \chi(t))$ with $\chi(t) = \int_{\mathbb{R}^d} h(x, t, u(x, t), \rho_t(x)) \rho_t(x) dx$; recall that in this case, $\partial_\rho H$ and $\partial_\rho h$ denote derivative with respect to a real parameter. Then, for any $v \in \mathbf{L}^2$:

$$H'_\rho(x, \tau, u; \rho_\tau) \cdot v = v \partial_\rho \tilde{H} + \left(\int_{\mathbb{R}^d} \partial_\chi \tilde{H} v dx \right) h + \left(\int_{\mathbb{R}^d} \partial_\chi \tilde{H} \rho_\tau dx \right) v \partial_\rho h. \quad (23)$$

In particular, for $v = \rho_\tau$ we have :

$$H'_\rho(x, \tau, u; \rho_\tau) \cdot \rho_\tau = \rho_\tau \partial_\rho \tilde{H} + \left(\int_{\mathbb{R}^d} \partial_\chi \tilde{H} \rho_\tau dx \right) (h + \rho_\tau \partial_\rho h). \quad (24)$$

Then, for the optimal \hat{u} and $\hat{\rho}$, (21) yields

$$\partial_\tau V' + \hat{\mu} : \nabla_x^2 V' + \hat{u} \cdot \nabla_x V' = -[\tilde{H} + \hat{\rho} \partial_\rho \tilde{H} + (h + \hat{\rho} \partial_\rho h) \int_{\mathbb{R}^d} \partial_\chi \tilde{H} \hat{\rho} dx]. \quad (25)$$

The link with Section 3 is established : (9) and (25) coincide with $V' = \rho^*$.

Remark 4.6 Note that the adjoint equation, (25), is set in $\mathbb{R}^d \times [0, T]$ with a right-hand side which is unbounded as $x \rightarrow \pm\infty$. Existence of solutions is doable in the finite case $\Omega \times [0, T]$ with Ω a bounded open set and $V|_{\partial\Omega} = 0$, but is a riddle otherwise. It is also a source of numerical difficulties because, as \mathbb{R}^d is approximated by $] -L, L[^d$, numerical boundary conditions compatible with the (unknown) behavior at infinity of V' need to be provided.

5 An Academic Example: Production of an Exhaustible Resource

Following [24], we consider a continuum of producers exploiting an oil field. Each producer's goal is to maximize his profit, knowing the price of oil; however, this price is influenced by the quantity of oil available on the market, which is the sum of all that the producers have decided to extract at a given time. Hence, while each producer does not affect the price of oil, because each producer solves the same optimization problem, in the end the global problem must take into account the market price as a function of oil availability. For a better understanding of the relation between the individual game and the global game, the reader is referred to [10].

5.1 Notations

Let X_0 be the initial oil reserve and X_t be the quantity of oil left in the field at time t , as seen by a producer. It is modeled by

$$dX_t = -a_t dt + \sigma X_t dW_t, \quad X_0 \text{ given by its PDF,} \quad (26)$$

where $a_t dt$ is the quantity extracted by the producer in the time interval $[t, t + dt]$, (so a_t is the extraction rate), and W is a standard real valued Brownian motion reflecting the incertitude of the producer about the remaining reserve; $\sigma > 0$ is a volatility parameter, assumed constant. We suppose that $a_t := a(X_t, t)$ is a deterministic function of t and X_t , meaning by this that the producers apply a feedback law to control their production.

We denote by C the cost of oil extraction, which is function of a and assumed to be $C(a) := \alpha a + \beta a^2$, for some positive constants α and β . The price of oil is assumed to be $p_t := \kappa e^{-bt} (\mathbf{E}(a_t))^{-c}$, with positive κ, b and c . This contains the macro-economic assumption that p_t is a decreasing function of the mean production because scarcity of oil increases its price and conversely. It also says that in the future oil will be cheaper because it will be slowly replaced by renewable energy.

Note that by construction X_t takes only positive values and ought to be bounded by, say L , the maximum estimate of the reservoir content. However, nothing in the model enforces these constraints.

5.2 Model

Each producer optimizes his integrated profit up to time T , discounted by the interest rate r ; however he wishes also to drain the oil field, i.e., achieve $X_T = 0$. Thus his goal is to maximize over $a(\cdot, \cdot) \geq 0$ the functional :

$$J(a) := \mathbf{E} \left[\int_0^T (p_t a_t - C(a_t)) e^{-rt} dt \right] - \gamma \mathbf{E}[|X_T|^\eta], \quad \text{subject to (26);} \quad (27)$$

γ and η are penalization parameters.

Replacing p and C by their expressions gives

$$J(a) = \mathbf{E} \left[\int_0^T (\kappa e^{-bt} (\mathbf{E}[a_t])^{-c} a_t - \alpha a_t - \beta (a_t)^2) e^{-rt} dt \right] - \gamma \mathbf{E}[|X_T|^\eta].$$

Obviously, J being the mean of a function of $\mathbf{E}[a_t]$, it is a mean-field type stochastic control problem.

5.3 Remarks on the Existence of Solutions

Denoting $\bar{a}_t := \mathbf{E}[a_t]$, J is:

$$J(a) = \int_0^T (\kappa e^{-bt} (\bar{a}_t)^{1-c} - \alpha \bar{a}_t - \beta \mathbf{E}[a_t^2]) e^{-rt} dt - \gamma \mathbf{E}[|X_T|^\eta].$$

Since $\mathbf{E}[a_t^2] \geq \bar{a}_t^2$, J is bounded from above,

$$J(a) \leq \int_0^T (\kappa e^{-bt} (\bar{a}_t)^{1-c} - \alpha \bar{a}_t - \beta \bar{a}_t^2) e^{-rt} dt.$$

Assume that $c < 1$. Then, the maximum of the right-hand side is attained when a is such that, $a(1-c)e^{-bt}(\bar{a}_t)^{-c} = \alpha + 2\beta\bar{a}_t$, $\forall t$. Hence the maximum value

$$J \leq \int_0^T (c\alpha + (1+c)\beta\bar{a}_t) \frac{\bar{a}_t e^{-rt}}{1-c} dt$$

provides an upper bound for J , so long as \bar{a}_t is upper bounded on $[0, T]$. Furthermore when the problem is converted into a deterministic optimal control problem with the Fokker-Planck equation, it is seen that the function is upper semi-continuous; so a maximum exists.

A counter example: Assume $c > 1$. For simplicity, suppose that $a(t) = |\tau - t|$ for some $\tau \in]0, T[$. Then, $\int_0^T a_t^{1-c} dt = +\infty$; hence the problem is not well posed, an obvious consequence of the fact that the model makes the price infinite too fast if nobody extract oil.

Linear feedback: If we search for a in the class of linear feedbacks $a(x, t) = w(t)x$ where w is a deterministic function of time only, then (26) has an analytical solution

$$X_t = X_0 \exp \left(- \int_0^t w(\tau) d\tau - \frac{\sigma^2}{2} t + \sigma(W_t - W_0) \right), \quad (28)$$

and the first and second moments of a_t are

$$\mathbf{E}[a_t] = \mathbf{E}[X_0] \tilde{w}_t, \quad \mathbf{E}[a_t^2] = \mathbf{E}[X_0^2] \tilde{w}_t^2 e^{\sigma^2 t}, \quad \text{with } \tilde{w}_t := w(t) e^{-\int_0^t w(\tau) d\tau}.$$

Then, for $\eta = 2$, the problem reduces to maximizing over $\tilde{w}_t \geq 0$

$$J(\tilde{w}_t) = \int_0^T (\kappa e^{-bt} \mathbf{E}[X_0]^{1-c} \tilde{w}_t^{1-c} - \alpha \mathbf{E}[X_0] \tilde{w}_t - \beta \mathbf{E}[X_0^2] \tilde{w}_t^2 e^{\sigma^2 t}) e^{-rt} dt - \gamma \mathbf{E}[X_0^2] e^{\sigma^2 T - 2 \int_0^T w(\tau) d\tau}. \quad (29)$$

5.4 Dynamic Programming in Absence of Constraint

To connect to Section 2 let us work with $u = -a$, the reserve depletion rate. For the time being, we shall ignore the constraints $L \geq X_t \geq 0$ and $u \leq 0$; so $\mathcal{V}_d = \mathbb{R}$. Moreover we shall work with $\eta = 2$ and comment on $\eta > 2$ at the end.

Recall that $\rho(\cdot, t)$, the PDF of X_t , is given by the Fokker-Planck equation :

$$\partial_t \rho - \frac{\sigma^2}{2} \partial_{xx} (x^2 \rho) + \partial_x (\rho u) = 0 \quad (x, t) \in \mathbb{R} \times]0, T], \quad (30)$$

with initial condition : $\rho|_{t=0} = \rho_0$ given. Now $\bar{u}_t := \int_{\mathbb{R}} u_t \rho_t dx = \mathbf{E}[-a_t]$ and

$$\tilde{J}(\tau; \rho_\tau, u) := \int_\tau^T \int_{\mathbb{R}} \left(\kappa e^{-bt} (-\bar{u}_t)^{-c} u_t - \alpha u_t + \beta u_t^2 \right) e^{-rt} \rho_t dx dt \quad (31)$$

$$+ \int_{\mathbb{R}} \gamma |x|^\eta \rho(x, T) dx \quad \text{subject to (30) with } \rho|_{t=\tau} = \rho_\tau. \quad (32)$$

The goal is now to minimize \tilde{J} with respect to u . Define also

$$V(\tau; \rho_\tau) := \min_u \tilde{J}(\tau; \rho_\tau, u).$$

Application of the Results of Section 2. In this example we have : $\mathcal{V}_d = \mathbb{R}$ and

$$H(x, t; u_t, \rho_t) = (\kappa e^{-bt} \chi_t^{-c} u(x, t) - \alpha u(x, t) + \beta u(x, t)^2) e^{-rt},$$

$$\chi_t = - \int_{\mathbb{R}} u_t \rho_t dx, \text{ hence } h(x, t, u, \rho) = -u(x, t) \text{ and } G(x, \chi_T) = G(x) = \gamma |x|^\eta.$$

By Proposition 4.2 and Remark 4.5, we have $V'_T = \gamma |x|^\eta$, and V' satisfies

$$\partial_t V' + \frac{\sigma^2 x^2}{2} \partial_{xx} V' + u \partial_x V' = - \left[H + \rho \partial_\rho H + (h + \rho \partial_\rho h) \int_{\mathbb{R}} \partial_\chi H \rho dx \right] \quad (33)$$

$$= - \left(\kappa (1-c) e^{-bt} (-\bar{u})^{-c} u - \alpha u + \beta u^2 \right) e^{-rt}. \quad (34)$$

because $\partial_\rho H = \partial_\rho h = 0$ and $\partial_\chi H = c \kappa e^{-bt} \chi^{-c-1} u$. Moreover, by (22),

$$-\partial_x V' = \left(\kappa (1-c) e^{-bt} (-\bar{u})^{-c} - \alpha + 2\beta u \right) e^{-rt}, \quad (35)$$

giving:

$$u(x, t) = \frac{1}{2\beta} \left[\alpha - e^{rt} \partial_x V' - \kappa (1-c) e^{-bt} (-\bar{u})^{-c} \right]. \quad (36)$$

Now, using (35) to eliminate $\partial_x V'$ in (34) leads to

$$\partial_t V' + \frac{\sigma^2 x^2}{2} \partial_{xx} V' = \beta e^{-rt} u^2. \quad (37)$$

Finally, using (36) in (37) and the definition of $(-\bar{u})_t$ yields :

$$\partial_t V' + \frac{\sigma^2 x^2}{2} \partial_{xx} V' = \frac{e^{-rt}}{4\beta} \left(\alpha - e^{rt} \partial_x V' - \kappa (1-c) e^{-bt} (-\bar{u})^{-c} \right)^2. \quad (38)$$

Note that this equation for V' depends only on \bar{u} and not on u . Nevertheless (36)–(38) is a rather complex partial-integro-differential system.

A Fixed-Point Algorithm. We can now sketch a numerical method to solve the problem:

Algorithm 1.

Choose $\omega \in (0, 1)$. initialize u_0 , set $i = 0$.

1. Compute ρ_i by solving (30) ;
2. Compute $\bar{u}_i = \int_{\mathbb{R}} u_i \rho_i$;
3. Compute V_i' by (38);
4. Compute \bar{u}_{i+1} by (36) and set $u_{i+1} = u_i + \omega(\bar{u}_{i+1} - u_i)$;

If not converged, set $i = i + 1$ and go back to step 1.

Although it seems to work numerically in many situations, as we shall see below, nothing is known on the convergence of this fixed-point type algorithm; three points need to be clarified:

1. Equation (30) is nonlinear and existence of solution is unclear.
2. A relevant stopping criteria for Algorithm 1 is yet to be found.
3. Even if the Fokker-Planck equation (30) is set on $\mathbb{R}^+ \times]0, T[$ instead of Q as discussed below in the numerical section (Section 5), there are difficulties. Because the second order term vanishes at $x = 0$, a weak formulation ought to be in the weighted Sobolev space and would be to find ρ , with

$$\rho \in \mathbf{H} = \{v \in L^2(\mathbb{R}^+) : x\partial_x v \in L^2(\mathbb{R}^+)\} \text{ such that } \forall v \in \mathbf{H} \\ \int_{\mathbb{R}^+} \left[v \partial_t \rho dx + (x\sigma^2 - u)\rho \partial_x v + x^2 \sigma^2 \partial_x \rho \partial_x v \right] dx = 0, \text{ for almost all } t. \quad (39)$$

Theorem 2.2 in [29] asserts existence and uniqueness of ρ , provided that there exists u_M such that $u(x, t) < xu_M, \forall t$. However, this oil-resource model doesn't impose $u(0, t) = 0$, and consequently there is a singularity at that point.

5.5 Calculus of Variations on the Deterministic Control Problem

To find the optimality conditions for (31), let us introduce an adjoint ρ^* satisfying

$$\partial_t \rho^* + \frac{\sigma^2 x^2}{2} \partial_{xx} \rho^* + u \partial_x \rho^* = e^{-rt} (\alpha - \beta u - \kappa(1-c)e^{-bt}(-\bar{u})^{-c})u. \quad (40)$$

in $\mathbb{R} \times]0, T[$ and $\rho^*|_T = \gamma|x|^\eta$. Then,

$$\delta J = - \int_Q \left(e^{-rt} (\alpha - 2\beta u - \kappa(1-c)e^{-bt}(-\bar{u})^{-c}) - \partial_x \rho^* \right) \rho \delta u dx dt + o(\|\delta u\|). \quad (41)$$

In other words

$$\text{Grad}_u J = - \left(e^{-rt} (\alpha - 2\beta u - \kappa(1-c)e^{-bt}(-\bar{u})^{-c}) - \partial_x \rho^* \right) \rho. \quad (42)$$

Algorithm. We apply the steepest descent method with varying step size:

Algorithm 2.

Initialize a_0 and $i = 0$; choose $0 < \varepsilon \ll 1$;
DO

1. Compute ρ_i by (30) with $\rho_{i|t=0}$ given;
2. Compute $\bar{u}_i = \int_{\mathbb{R}} u_i \rho_i dx$;
3. Compute ρ_i^* by (40);
4. Compute $\text{Grad}_u J$ by (42);
5. Compute a feasible descent step $\mu_i \in \mathbb{R}$ by Armijo rule [30]
6. Set $u_{i+1} := u_i - \mu_i \text{Grad}_u J$, $i = i + 1$;

WHILE ($\|\text{Grad}_u J\| > \varepsilon$);

For a convergence analysis, here the situation is somewhat better: both the state and the adjoint equations (30), (40) are linear and the only problem is the asymptotic behavior of u . Note also that one could use a conjugate-gradient algorithm at little additional computational cost.

After discretization by a variation method such as finite element methods, convergence to a local minima could probably be established by the techniques of control theory (see e.g. [30]) because the problem is finite dimensional and the gradient of the cost function is exact [29]. However, convergence of the solution of the discrete problem to the solution of the continuous problem is open and even more difficult than existence.

5.6 The Riccati Equation when $\eta = 2$

Even though the problem is not linear-quadratic, when $\eta = 2$ we can still look for V' solution of (38) in the form $V'(x, t) = P_t x^2 + z_t x + s_t$. where P_t, z_t, s_t are functions of time only.

Identification of all terms proportional to x^2 in (37) gives,

$$\dot{P}_t + \sigma^2 P_t = \frac{e^{rt}}{4\beta} P_t^2, \quad P_T = \gamma.$$

For clarity, let $Q_t = e^{rt} P_t$ and $\mu = \sigma^2 - r$. Then, the above is:

$$\dot{Q}_t + \mu Q_t - \frac{Q_t^2}{4\beta e^{rt}} = 0, \quad Q_T = \gamma e^{rT} \Rightarrow \frac{\dot{Q}_t}{Q_t} + \frac{\dot{Q}_t}{4\beta e^{rt} \mu - Q_t} = -\mu.$$

As long as $4\beta e^{rt} \mu - Q_t > 0$ it leads to

$$\frac{Q_t}{Q_t - 4\beta e^{rt} \mu} = \frac{\gamma e^{(T-t)\mu}}{\gamma - 4\beta \mu} \Rightarrow P_t = \frac{4\beta \mu \gamma e^{(T-t)\mu}}{\gamma e^{(T-t)\mu} - \gamma + 4\beta \mu}. \quad (43)$$

Then, u is found by (36). In particular $\partial_x u = -\frac{1}{8\beta} \partial_{xx} V' = -\frac{1}{4\beta} P_t$. However, the Fokker-Planck equation must be solved numerically to compute \bar{u} .

Note that $\gamma < 4\beta \mu$ implies $4\beta e^{rt} \mu - Q_t > 0$ and also $P_t > 0$.

Remark 5.1 By identification of the terms of order 1 and 0 in x , equations for z and s are found:

$$\begin{aligned} \dot{z} &= -\frac{1}{\beta}(\alpha - \delta(1-c)e^{-bt}(-\bar{u})^{-2c})^2 e^{rt} P_t^2, \quad z_T = 0, \\ \dot{s} &= -\frac{e^{-rt}}{\beta}(\alpha - \delta(1-c)e^{-bt}(-\bar{u})^{-c} - e^{rt} z_t)^2, \quad s_T = 0. \end{aligned}$$

Remark 5.2 Note that $u_t = 2xP_t + z_t$ is not a linear feedback function as in (28) above.

Remark 5.3 This explicit feedback solution is smooth and adapted to the stochastic process (26), so it should be a solution of (27) if it exists (recall that (27) is not convex). Furthermore it has a behavior at infinity which is compatible with (2).

5.7 Numerical Implementation

To implement Algorithms 1 & 2, we need to localize the PDE. As $x < 0$ makes no sense for this application, we shall work on $Q_L = [0, L] \times [0, T]$ instead of $\mathbb{R} \times [0, T]$; a stopping time for the event $X_t = 0$ would be better, but too costly. At $x = L$, we set $\rho(L, t) = 0, \forall t$, which makes sense when L is large.

Assigning boundary conditions to (38) and (40) is problematic. Our numerical tests show that the computations depend strongly on L when it is not done correctly. When $\eta = 2$, we know that V' and ρ^* have asymptotically the same behavior as $P_t x^2$, giving $\frac{1}{2}\sigma^2 x^2 \partial_x V' = \sigma^2 x^3 P_t = \sigma^2 x V'$, a relation which can be used as a boundary condition in the weak form of the equation (and similarly for ρ^*): find $\rho \in H^1(Q_L)$ with V'_T given and

$$\begin{aligned} \int_{Q_L} [-v \partial_t V' + \frac{\sigma^2}{2} \partial_x (vx^2) \partial_x V'] dx dt + \int_0^T \sigma^2 L V'(L, t) v(L, t) dt \\ + \int_{Q_L} \frac{e^{-rt}}{4\beta} (\alpha - e^{rt} \partial_x V' - \kappa(1-c)e^{-bt}(-\bar{u})^{-c})^2 v dx dt = 0, \end{aligned} \quad (44)$$

for all $v \in H^1(Q_L)$ with $v_T = 0$.

To solve this non-linear PDE we use the fact that it is embedded into an iterative loop in Algorithm 1, and semi-linearise it by evaluating the square term in the last integral as a product of the same, where one factor is evaluated at the previous iteration.

To discretise it we have used a space-time finite element method of degree 1 over triangles covering Q_L . Admittedly it is an unusual method! However, it is somewhat similar to a central difference method and it is feasible because the problem is one dimensional in space and because it allows exact conservativity and exact duality with respect to time and space in the integrations by parts. It handles also automatically the storage of ρ, u, V, \bar{u} at all times and solve the backward (and forward) equation at all time steps by a single linear system. The linear systems are solved with the library MUMPS as implemented in `freefem++` [31].

5.8 Results with Algorithm 1

We used 50 points to discretize $(0, L)$, $L = 10$ and 50 time steps for $[0, T]$, $T = 5$. The following values have been used: $\alpha = 1$, $\beta = 1$, $\gamma = 0.5$, $\kappa = 1$, $b = 0.1$, $r = 0.05$, $\sigma = 0.5$ and $c = 0.5$. The initial condition ρ_0 is a Gaussian curve centred at $x = 5$ with volatility 1. We initialized u by $u_0 = -\alpha/(2\beta)$. A local minimum u_e is known from the Riccati equation; the error $\|u - u_e\|$ is used as a stopping criteria in Algorithm 1. We chose $\omega = 0.5$.

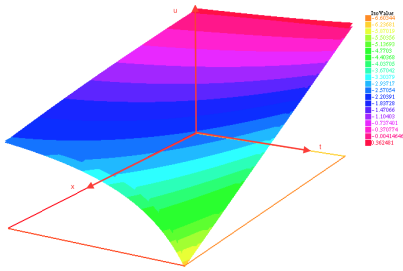


Fig. 1 Optimal $x, t \rightarrow u(x, t)$ and the Riccati solution slightly below

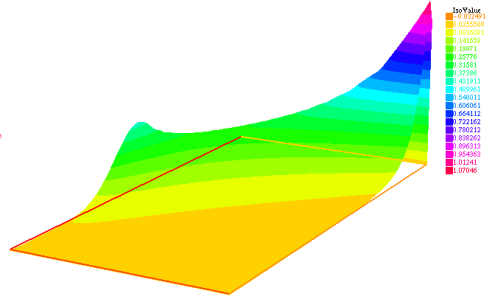


Fig. 2 PDF of resource X_t : $x, t \rightarrow \rho(x, t)$

Results with Algorithm 1. Figure 1 shows the optimal control as a function of (x, t) . For each t the control is linear in x , as predicted by the Riccati equation; the quadratic part of the Riccati solution of Section 5.6 is also plotted, and a small difference is seen on the plot (two surfaces close to each other are displayed). Figure 2 shows the evolution in time of the PDF ρ for all $x > 0$ of the resource distribution X_t . At time 0 it is a Gaussian distribution centered at $x = 5$; at time T the distribution is concentrated around $x = 0.5$, so most producers have pumped 90% of the oil available to them.

Table 1 Algorithm 1. Convergence history: H^1 -error, $\int (\partial_x u - \partial_x u_e)^2 dx dt$ versus iteration number k

k	1	2	3	4	5	6	7	8	9	10
Error	1035	661.2	8.605	44.7	3.27	0.755	0.335	0.045	0.015	0.003

All above is obtained with $L = 10$, but there is almost no difference with $L = 40$. Figures 3, 4 present the evolution of production $(-\bar{u}_t)$ and price $p_t = \kappa e^{-bt} (-\bar{u}_t)^{-c}$.

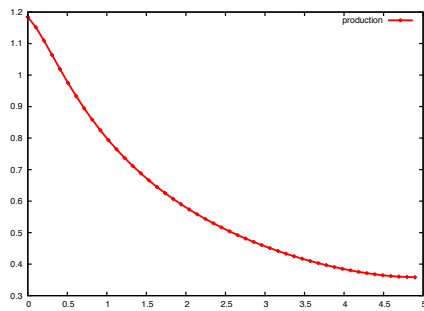


Fig. 3 Total (instantaneous) production vs time

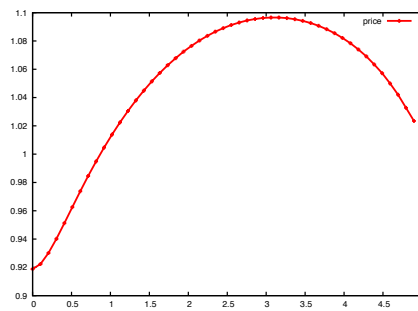


Fig. 4 Oil Price versus time

5.9 Results with Algorithm 2

The performance of descent methods were disappointing. It generated many different solutions depending on the initial value for u . If $u_0 = u_e$, then the algorithm decreases the cost functions by introducing small oscillations, a strategy which is clearly mesh dependent. If $u_0 = -0.5$, then the solution of Figure 5 is found after 10 iterations of

Table 2 Algorithm 2. Convergence history: Values of J and $\|Grad_u J\|$ versus iteration number k

k	0	1	2	3	..	9
J	0.7715	0.2834	0.2494	0.1626		0.0417
$\ GradJ\ $	0.003395	0.001602	0.000700	0.000813		0.000794

steepest descent. The convergence history is given in Table 2. The results are shown on Figure 6. Note the oscillations of ρ near $t = T$.

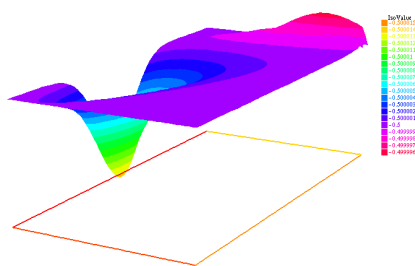


Fig. 5 Another solution u

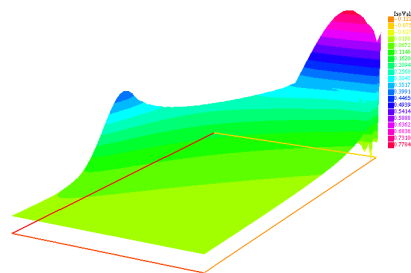


Fig. 6 The corresponding ρ

5.10 The case $\eta \neq 2$

The following parameters are now changed : $\eta = 3$ and u_0 is initialized by the Riccati solution. Algorithm 1 converges in a few iterations to a solution, but $\omega = 0.05$ is required for convergence. Algorithm 2 gives a different solution. Both adjoint states are shown on Figure 7. When $\eta = 4$, Algorithm 1 diverges, while Algorithm 2 converges

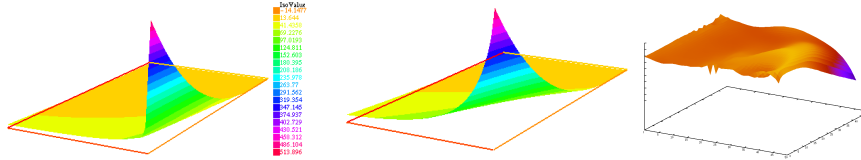


Fig. 7 The case $\eta = 3$; ρ^* computed with Algorithm 1 (left) and ρ^* computed with Algorithm 2 (center). The two algorithms computed different solutions when initialized with the Riccati solution. Right: solution u computed with Algorithm 2 when $\eta = 4$.

to the solution shown on Figure 7.

5.11 Linear Feedback Solution

Using automatic differentiation of computer programs by operator overloading in C++ and initializing a steepest descent method with the linear part of the Riccati solution of Section 5.6 and the same parameters as above, we obtained the $w(t)$ displayed on Figure 8, very close to the Riccati solution. To understand why the Riccati solution may not be the best solution we plotted $\lambda \rightarrow J^d(\lambda) := J(w_t^d + \lambda h_t)$, $\lambda \in]-0.5, +0.5[$, where h_t is an approximate $w_t - \nabla J(w_t^d)$. Figure 8 shows that there are three local minima and two local maxima, and w_t^d is only a local minimum.

Fig. 8 A solution computed by a steepest descent method to maximize (29) showing the feedback coefficient w versus time (left figure), compared with the feedback coefficient of the Riccati solution (solid line). The plot in the center shows J_λ^d as a function of $\lambda \in]-0.5, +0.5[$; the Riccati solution is at $\lambda = 0$. At $\lambda = \pm 0.12$ lies absolute minima of $J^d(\cdot)$, but it can be seen only by zooming in the central figure (right plot). Observe that the absolute minimum is very shallow and hard to find; it appears also to be mesh dependent.

6 Conclusions

Stochastic mean-field type control is numerically hard. Even a simple academic toy problem gives difficulties. The first difficulty which this paper had to deal with is

the extension of the HJB dynamic programming approach. The second difficulty is related to the well-posedness of the HJB or adjoint equation because it is set in an infinite domain in space. The third difficulty is the lack of proof of convergence of the two algorithms suggested here, an HJB-based fixed point and a steepest descent based on calculus of variations. When it converges the fixed point method is preferred but there is no known way to guarantee convergence even to a local minimum; as for the steepest descent we found it somewhat hard to use, mainly because it generates irregular oscillating solutions; some bounds on the derivatives of u need to be added in penalised form in the criteria. Numerically both algorithms are cursed by the asymptotic behaviour of the solution of the adjoint state at infinity. So, when possible, the Riccati semi-analytical feedback solution is the best. Finally, but this applies only to this toy problem, the pure feedback solution is nearly optimal, easy to compute and stable. Note also that this semi-analytical solution is a validation test, since it has been recovered by both algorithms.

Acknowledgements The authors are grateful to Yves Achdou, Alain Bensoussan and Olivier Guéant for useful discussions.

References

1. Fleming W.H., Soner H.M. *Controlled Markov Processes and Viscosity Solutions*. Stochastic Modelling and Applied Probability series vol 25, Springer 2006.
2. Kushner H.J. : *Stochastic Stability and Control*. Academic Press, New York (1967)
3. Øksendal B., Sulem A. : *Applied Stochastic Control of Jump Diffusions*. Springer (2005)
4. Touzi N. : *Optimal Stochastic Control, Stochastic Target Problems and Backward SDE*. Field Inst. Monographs 29, Springer (2013)
5. Yong J., Zhou X. Y. : *Stochastic Control. Applications of Math. series vol 43*. Springer (1999)
6. Carmona R., Fouque J.-P., Sun L.-H. : Mean-field games and systemic risk. *Communications in Mathematical Sciences*, Vol. 13(4), p. 911-933 (2015)
7. Garnier J., Papanicolaou G., Yang T.-W. : Large Deviations for a Mean-Field Model of Systemic Risk, *SIAM J. Finan. Math.*, 4(1), 151-184 (2013)
8. Lasry J.M., Lions P.L. : Mean-Field Games. *Japan. J. Math.* 2, 229-260 (2007)
9. Shen M., Turinici G. : Liquidity Generated By Heterogeneous Beliefs and Costly Estimations. Vol. 7, No 2, pp. 349-361 (June 2012).
10. Carmona R., Delarue F., Lachapelle A. : Control of Mc-Kean-Vlasov mean field games. *Math. Financ. Econ*, 7, 131-166 (2013).
11. Andersson D., Djehiche B. : A Maximum Principle for SDEs of Mean-Field Type. *Dyn. Games Appl.*, 3:537-552 (2013)
12. Bensoussan A., Frehse J., Yam S.C.P. : The Master Equation in Mean-Field Theory. *The Master Equation in Mean-Field Theory. Journal des Mathématiques Pures et Appliquées*, Vol. 103, No. 6, p. 1441-1474 (2015).
13. Carmona R., Delarue F. : The master equation for large population equilibriums. *Stochastic Analysis and Applications 2014*, Editors : D. Crisan, B. Hambly, T. Zariphopoulou. Springer (2014).
14. Bensoussan A., Frehse J. : Control and Nash Games with Mean-Field Effect. *Chinese Annal Math. Series B.* 34B(2), 161-192 (2013)
15. Bensoussan A., Frehse J., Yam S.C.P. : *Mean-field Games and Mean-field Type Control*. Springer Briefs in Math. 2014.
16. Kolokoltsov V., Troeva M., Yang W. : On the Rate of Convergence for the Mean-Field Approximation of Controlled Diffusions with Large Number of Players. *Dynamic Games and Applications* June 2014, Volume 4, Issue 2, pp 208-230.
17. Kolokoltsov V., Yang W. : Existence of solutions to path-dependent kinetic equations and related forward-backward systems. *Open Journal of Optimization* Vol 2, pp. 39-44, (2013)

18. Achdou Y., Laurière M. : On the system of partial differential equations arising in mean field type control, DCDS A (September 2015).
19. Gangbo W., Swiech A. : Optimal Transport and Large Number of Particles. *Discret. Cont. Dynam. Syst.– A* 34 1397 – 1441 (2014)
20. Chan P., Sircar R. : Bertrand & Cournot Mean-field Games, to appear in *Applied Mathematics & Optimization* (2015).
21. Laurière M., Pironneau O. : Dynamic Programming for Mean-Field Type Control. *C.R. Acad. Sci. Serie I*, Vol 352 - N^o 9 p707-713 (sept. 2014)
22. Lions P.L. : Mean-Field Games. Cours au Collège de France (2007-2008) http://www.college-de-france.fr/site/pierre-louis-lions/course-2007-2008_1.htm
23. Annunziato M., Borzi A., Nobile F., Tempone R. On the Connection between the Hamilton-Jacobi-Bellman and the Fokker-Planck Control Frameworks, *Applied Mathematics*, 2014, 5, 2476-2484.
24. Gueant O., Lasry M., Lions P.L. : Mean Field Games and Applications. *Paris-Princeton Lectures on Mathematical Finance. Lecture Notes in Math.* Springer (2011).
25. Le Bris, C., Lions P.L, Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients. *Comm. Partial Differential Equations*, 33, 1272-1317 , 2008.
26. Porretta A. : Weak solutions to Fokker-Planck equations and Mean Field Games, *Arch. Rational Mech. Anal.* 216, 1-62 (2015).
27. Bally V., Pagès G., Printems J. : A Stochastic Quantization Method for Nonlinear Problems. *Monte Carlo Methods Appl.* 7 (12), 2134 (2001)
28. Neufeld A., Nutz M. : Nonlinear Lévy processes and their Characteristics. *Transactions of the American Mathematical Society*, forthcoming. See also arXiv:1401.7253v1.
29. Achdou Y., Pironneau O. : *Computation Methods for Option Pricing*, SIAM Frontiers in Math (2005)
30. Polak E. : *Optimization Algorithms and Consistent Approximations*, Springer New York (1997)
31. Hecht F. : New Development in freefem++, *J. Numer. Math.* vol 20, no 3-4, pp 251-265 (2012).