Pricing of gas swing contracts: a viscosity solution approach

Enrico Edoli*[1,2] and Tiziano Vargiolu†[1]

[1]Department of Pure and Applied Mathematics, via Trieste 63, University of Padova, I-35121 Padova, Italy
[2]Finalyst SAS, via Trentino 28, I-30010 Campolongo Maggiore, Italy

September 30, 2012

Abstract

In this paper we model in continuous time a gas swing contract in the spirit of [4], with one additional state variable corresponding to a stochastic strike price. Since in real contracts the strike is a market index which is updated monthly, this results in a mixed discrete-continuous stochastic control problem. We reduce this to the usual continuous time situation by adding a state variable, corresponding to an index rolled-over in continuous time. Then we analyse the viscosity solution of the resulting Hamilton-Jacobi-Bellman (HJB) equation, deriving results on existence, uniqueness and smoothness of the solution. Finally, we present two numerical methods to derive the price of a swing contract. The first one is based on finite differences applied to the HJB equation. The second one is based on Least Square Monte Carlo and does not use the HJB equation directly, focusing directly on a discrete-time approximation of the continuous time problem. While we apply radial basis approximation in the least square algorithm to avoid dimension issues, we also present a case when, under some assumptions or approximation, the dimension issue can be avoided by a proper quantization of the state space.

Contents

*enricoedoli@gmail.com
†vargiolu@math.unipd.it

1 Problem Framework

The physical nature of commodities, energy and gas in particular, such as their limited storability or their strong link with complex customers' patterns of consumption, had led to design a lot of supply contracts which allow flexibility of delivery. In particular, in gas markets many long-terms contracts (for 10 years or more) are embedded with flexibility of delivery options known as *swing* or *take-or-pay*. Such contracts allow the option holder to withdraw every day a quantity of gas subject to daily, as well as periodic (usually monthly or annual), minimum and maximum constraints. This flexibility addresses the need to hedge a frequent demand fluctuation which in practice is impossible to foresee in the long period, being linked to exogenous variable such as weather, economic scenario, changes in heating technology and power production and so on.

The correct valuation of these type of contracts is important for at least two reasons: first of all, thanks to the liberalization of energy markets, the price of such contracts is no more set by regulators under the assumption of cost recovery, as in the old regulated markets, but it is negotiated between agents and it is mainly related to the financial risks underlying the contracts. On the other hand, most of the existing contracts include *renegotiation* clauses which permits to adjust the contract according to developments in the markets. So it is very important for both contract parties to have methodologies to understand which impact contract's parameters have on the price. Finally, from the point of view of a profit maximizing agent, the flexibilities embedded in the contract, i.e. the possibility to decide how much quantity of gas to withdraw every day, should be used not only to manage demand fluctuation, but also if possible to make profit against local market price.

The structure of long term gas agreements is pretty standardized in Europe. The strike price, which is the price paid by the owner of the contract to the seller of the commodity, typically depends upon a basket of crude and refined oil products, which is averaged through time in order to smooth undesired volatility effects; for more details we refer the interested reader to [2, Section 3.1]. Since oil products are traded in US dollars, oil related indexes are also expressed in US dollars, thus typical market risk factors perceived by European importers are represented both by USD/EUR exchange risk, and price differential between import cost in Euros I_t and local market prices P_t settled daily by local gas market exchanges. Clearly, the future prices I_t and P_t are not known when pricing the contract so they have to be assumed as stochastic variables. It is also natural to assume that the optimal withdrawn quantity should be also linked in some way to prices, or at least to their expected future value.

Thus, pricing and hedging of swing contracts has to be performed dynamically through time, has to take into account the stochastic dynamics of both market and strike index prices and volume constraints and has to suggest an optimal withdrawal policy which should maximise the expected revenues of the contract. This is exactly the practical description of a so-called *stochastic optimal control problem*.

The scope of this work is to investigate this stochastic optimal control problem. We formalize the mathematical problem taking into account both the stochastic nature and the monthly structure also of the strike price as well as local market price. Then, by using the theory of viscosity solution, we find out some properties of the value function such as existence, uniqueness and smoothness. We then apply numerical schemes to find out the price of some typical contracts.

1.1 Index price modelling

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t, \mathbb{P})$ be a filtered probability space and W_1, W_2 two correlated Brownian motions with correlation ρ defined on Ω. Let $[0, T]$ be a fixed interval on which the swing contract is defined. This interval is then divided into subperiod $\{[T^m_{n-1}, T^m_n]\}^{m=1,\ldots,D}_{n=1,\ldots,12}$, where $[T^m_{n-1}, T^m_n]$ represents the interval covered by month n of year m. We will suppress the superscript m when not necessary.

Managing a swing contract basically leads to deal with, at least, two prices: one is the so called contract price, which is the price the buyer of the swing option pays to the seller for the withdrawal of a unitary quantity of gas. Let this price be $I(t)$. The second one is the gas spot price that the buyer can use to sell the gas to the market. Let us denote this price with $P(t)$. In practice, the owner of a swing contract can buy gas at the price $I(t)$ and then, eventually, sell this gas at the price $P(t)$, realizing a profit (or loss) equal to $P(t) - I(t)$ for each unitary quantity of gas withdrawn.

Before approaching the modelling of price dynamics, a short description of how contract price behaves could be explanatory.

The price the buyer of the swing option pays to the seller typically depends upon a basket of crude and refined oil products traded every day in the market. This oil-linked pricing scheme is pretty typical for European gas markets since most of the gas arriving to Europe is a complementary output of oil extraction; this feature is not typical in the US gas market, since American gas production is almost totally disjoint from oil one.

Typically, the price of this basket of oil related products is averaged through time in order to smooth undesired volatility effects. The averaging rule is related to a triplet of numbers (ξ_1, ξ_2, ξ_3) respectively denoting the number of months composing the backward looking average of prices, the number of months prior to delivery that should not be included in the averaging process and the number of months between one index calculation and the following (almost always equal to one, but also three and six are common). More formally, if we denote by $I(t)$ the price of the index at time t, we have that $I(t)$ is a piecewise constant function on the intervals $[T_n, T_{n+k})$ for some n, with a jump at time T_{n+k}. If $B(t) = (B_1(t) \dots B_b(t))$ denotes the vector whose components are the price at time t of the oil-related products in the basket, α is a vector of weights, and we define the set $\mathcal{I}(\xi) \subseteq \mathbb{N}$ as

$$\mathcal{I}(\xi) = \{k | k = \nu \cdot \xi, \nu \in \mathbb{N}, k + \xi \leqslant 12\}$$

we can express the index price $I(t)$ as the weighted average

$$I(t) = (T_{n-\xi_2} - T_{n-\xi_2-\xi_1})^{-1} \int_{T_{n-\xi_2-\xi_1}}^{T_{n-\xi_2}} \alpha B(s) ds, \quad \forall t \in [T_n, T_{n+\xi_3}), \forall n \in \mathcal{I}(\xi_3) \tag{1}$$

Notice that the following relationship holds:

$$I(t) \equiv I(T_n), \qquad \forall t \in [T_n, T_{n+\xi_3}), \forall n \in \mathcal{I}(\xi_3) \tag{2}$$

It follows that the index price $I(t)$ could be modelled in two ways. The first obvious way, given the identity in Eq. (2), is to model the sequence of monthly prices $\{I_{T_n^m}\}_n^m$ as a discrete sequence of random variables. The second way is a little more sophisticated. We can assume that the index price has itself a *spot* continuous time dynamics. If we use a different parametrization of the couple (ξ_1, ξ_2) by introducing ℓ_1, ℓ_2 which represent the length of the averaging window and the length prior to delivery on which this window ends, we can rewrite Eq. (1) as

$$I(t) = \ell_1^{-1} \int_{T_n-\ell_1-\ell_2}^{T_n-\ell_2} \alpha B(s) ds, \quad \forall t \in [T_n, T_{n+\xi_3}), \forall n \in \mathcal{I}(\xi_3)$$

We may take into account that ℓ_1, ℓ_2 should be functions of T_n because the length of the month is not equal for every month. This is barely an improvement and to avoid huge notations we don't care about this. Now we can re-define the index price $I(t)$ using its *spot value*

$$I(t) = \ell_1^{-1} \int_{t-\ell_1-\ell_2}^{t-\ell_2} \alpha B(s) ds \tag{3}$$

Notice that Eq. (1) and Eq. (3) give the same value at the points $\{T_n^m\}_n^m$, but have different values for others t. We then use this *index spot value* in this way: at the beginning of every month n, at time T_n, we fix the strike price for the swing contract as the realixed index price $I_{T_n} = \hat{i}$. This will be the fixed index price paid by the buyer of the contract for month n, coerently with the behaviour of Eq. (1); on the other hand for the instants $t \in [0, T] \setminus \{T_n^m\}_n^m$ we have a dynamics coherent with the one in Eq. (3) and not a constant one as in Eq. (1).

Unfortunately, the definition in (3) is clearly non-Markovian, being an average on past values of $B(t)$. Here we make the following assumption: the dynamics in (3) can be approximated by a new markovian one, solution of the following SDE

$$dI(t) = \mu_i(t, I(t))dt + \sigma_i(t, I(t))dW_i(t) \tag{4}$$

for some functions $\mu_i, \sigma_i : [0, T] \times \mathbb{R} \to \mathbb{R}$.

In contrast with the contract index price, the spot price $P(t)$ is directly traded on local market and it changes (at least) once a day, depending on the liquidity of the local market. So we make the following (continuous time) assumption for the dynamics of the spot price $P(t)$:

$$dP(t) = \mu_p(t, P(t))dt + \sigma_p(t, P(t))dW_p(t) \tag{5}$$

for some functions $\mu_p, \sigma_p : [0, T] \times \mathbb{R} \to \mathbb{R}$.

We will specify other assumptions on the functions $\mu_p, \mu_i, \sigma_p, \sigma_i$ later.

1.2 One year problem

In this and in the following sections we deal with the problem of finding the value of a one year contract. For a standard swing contract, this is not a restriction or a simplification of our problem: even if the contract is written over a longer period of time, in the absence of constraints between two different years (such as make-up, carry forward, ...) the problem of pricing and manage the contract is independent for every year. In fact, ordinary swing contract permits to the owner to buy in every sub-period a quantity of gas, which we denote by $u(t)$, bounded between a minimum (mDQ) and maximum (MDQ) level which usually reflect physical effective transportation capacity limitations; thus for every instant t

$$\mathrm{mDQ} \leqslant u(t) \leqslant \mathrm{MDQ} \qquad \forall t \in [0, T] \tag{6}$$

In addition, *for every contractual year*, minimum and maximum quantities are also established, called respectively minimum annual quantity (mAQ) and annual contract quantity (ACQ). If we introduce the cumulated quantity $z_m(t)$ for year m, at time t

$$Z^m(t) = \int_{T_0^m}^{t \wedge T_{12}^m} u(s)ds$$

we have the constraints

$$\mathrm{mAQ} \leqslant Z^m(T_{12}^m) \leqslant \mathrm{ACQ} \qquad \forall m = 1, \dots, D$$

but also the relationship

$$Z^m(T_0^m) = 0 \qquad \forall m = 1, \dots, D \tag{7}$$

Thus the admissible area for the control $u(t)$ is exactly the same for every year, and it is given by

$$\mathcal{A}_m = \{u \in [\mathrm{mDQ}, \mathrm{MDQ}] \quad \text{s.t.} \quad \mathrm{mAQ} \leqslant Z^m(T_{12}^m) \leqslant \mathrm{ACQ}\} \qquad \forall m = 1, \dots, D$$

Sometimes the bounds on mAQ and ACQ can be overridden, but a penalty is paid. In this case

$$\mathcal{A}_m = \{u \in [\mathrm{mDQ}, \mathrm{MDQ}]\} \qquad \forall m = 1, \dots, D$$

We will concentrate on the last case. In both cases, if no other inter-temporal constraints are imposed to the problem (for instance make-up clauses), this fact and Equation (7) lead to notice that the pricing problem is exactly the same in every year, and can be faced separately year by year. So, from now on, we focus on a one-year problem.

Let $[0, T]$ be the reference interval of the year and let $\{[T_{n-1}, T_n]\}_{n=1,\dots,12}$ be the sequence of intervals describing every month, with $T_0 = 0$ and $T_{12} = T$.

We notice that Eq. (6) forces the buyer of the contract to buy, during a year, *at least* the quantity $\mathrm{mDQ} \cdot T$. This quantity, called the *take-or-pay* quantity, has to be paid, and may safely not be taken in consideration in our optimization, i.e. we can always consider a decomposition of a swing contract in the same spirit of [3, Section 2]. We let

$$u(t) \in U = [0, \bar{u}]$$

with $\bar{u} = \mathrm{MDQ} - \mathrm{mDQ}$.

To keep a general view, we also let

$$Z(t) = \int_0^t u(s) ds \tag{8}$$

Penalties are often imposed if the constraint $Z(T) \in \left[\underline{M}, \bar{M}\right]$ is not satisfied. An example of such penalties can be given by the function

$$\Psi(z) = \begin{cases} \boldsymbol{p} \cdot \underline{M} & z \in (-\infty, 0) \\ \boldsymbol{p} \cdot \left[(z - \bar{M})^+ + (\underline{M} - z)^+\right] & z \in [0, \bar{u}T] \\ \boldsymbol{p} \cdot (\bar{u}T - \bar{M}) & z \in (\bar{u}T, +\infty) \end{cases} \tag{9}$$

where $\boldsymbol{p} > 0$ is a proportional amount paid if the yearly constraints are not satisfied. Other kinds of penalty functions can be considered, but in any case, from a mathematical point of view, we can not assume that those functions are neither C^2 or C^1. A more realistic assumption could be the continuity and the polynomial growth of the function Ψ, so we make such assumptions. Finally, notice that the piecewise definition is only a mathematical trick used to have a continuous and bounded function on the whole space $\mathbb{R}$. This will be an important assumption for Theorem 6. In practice, thanks to the physical constraint $u_t \in [0, \bar{u}]$, at any time $t \in [0, T]$ the cumulated quantity z always lies in the interval $[0, \bar{u}T]$ and so the maximal possible final penalty is given by $\boldsymbol{p} \cdot \left[(z - \bar{M})^+ + (\underline{M} - z)^+\right]$.

By introducing the function

$$\varphi(t) = \max\{T_n | T_n \leqslant t\}$$

and defining

$$\hat{I}(t) = I(\varphi(t)) \tag{10}$$

we can now write our value function: we want to maximize the expected value of the discounted profit and loss i.e. we are interested in finding the contract value $V^1(0, X_0)$ at the beginning of the year

$$V^1(0, X_0) = \sup_{u \in \mathcal{A}} \mathbb{E}\left[\int_0^T e^{-rs}(P_s - \hat{I}_s)u_s ds + e^{-rT}\Psi(Z_T)\right] \tag{11}$$

where, for the sake of notation, we write the states as a four dimensional vector X_t:

$$X_t = \left(P_t I_t \hat{I}_t Z_t\right)^T \in \mathbb{R}^4$$

where the superscript T stands for the transposed. For a fixed interval $t \in (T_n, T_{n+1})$ the dynamics of X_t is

$$\begin{aligned} dX_t &= f(t, X_t, u_t)dt + \Sigma(t, X_t, u_t)dW(t) = \\ &= \begin{pmatrix} \mu_p(t, P_t) \\ \mu_i(t, I_t) \\ 0 \\ u(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_p(t, P_t) & 0 \\ \sigma_i(t, I_t)\rho & \sigma_i(t, I_t)\sqrt{1-\rho^2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \end{aligned} \tag{12}$$

where W_1, W_2 are two *uncorrelated* Brownian motion, linked to W_p, W_i by the relationship

$$\begin{cases} W_p(t) = W_i(t) \\ W_i(t) = \rho W_1(t) + \sqrt{1-\rho} W_2(t) \end{cases}$$

The contract value at terminal time T is the penalty function

$$V^{13}(t, x) \equiv \Psi(z)$$

Remark 1 The function V requires two separate arguments for the index part. The first argument I_t represents the *index spot value*. This price is neither traded nor really used in the contract, but it becomes useful to predict the future strike price I_{T_n} using the (assumed) Markov property of I_t. The second argument $\hat{I}_t$, represents the *present (traded) value* at time t of the index, that is the strike price of the swing option for month n. This is the realized price of the index at the beginning of the month.

We now use the dynamic programming principle on months. Taking into account that the realized value of the index for month n is $I_{\psi(t)} = I_{T_n} = \hat{\imath}$ and it is known for $t \geqslant T_n$, we can define in every month n a value function $V^n(t, x)$ which represents contract's value during month n, when the index strike price $\hat{\imath}$ is known and fixed. Let us define:

$$\begin{aligned} V^{13}(t, x) &= \Psi(z) \qquad \forall (t, x) \in [0, T] \times \mathcal{S} \\ V^n(t, x) &= \sup_{u \in \mathcal{A}} \mathbb{E}_{t,x}\left[\int_t^{T_n} e^{-r(s-T_{n-1})}(P_s - \hat{\imath})u_s ds \right. \\ &\qquad \left. + e^{-r(T_n - T_{n-1})}V^{n+1}(T_n, X_{T_n})\right] \qquad \begin{array}{rcl} n & = & 1, \ldots, 12 \\ t & \in & [T_{n-1}, T_n] \end{array} \end{aligned} \tag{13}$$

where in $\mathbb{E}_{t,x}[\cdot]$ is the expectation with respect to $\mathbb{P}_{t,x}$ which is the probability under which X has the dynamics given by Eq. (12) with initial condition $X_t = x$. Formally

$$\mathbb{E}_{t,x}[\phi(X(s))] = \int_{\mathbb{R}^n} \phi(y)\mathbb{P}_{t,x}(s, dy)$$

where

$$\mathbb{P}_{t,x}(s, B) = \mathbb{P}(X(s) \in B | X(t) = x)$$

for every measurable function ϕ and every B in the σ-algebra of Borel sets of $\mathbb{R}^n$.

We notice that

$$V^n(T_n, \cdot) = V^{n+1}(T_n, \cdot)$$

At this stage we have no hint about the smoothness of the functions V^n, so in what follows we derive the HJB equation for V^n using the theory of viscosity solutions and proving the required smoothness for the functions V^n and their derivatives.

For every $n = 1, \dots, 12$ we introduce the following notations:

$$\begin{aligned} L^n(t, x, u) &= -e^{-r(t-T_{n-1})}(p - \hat{\imath})u, \quad s \in [T_{n-1}, T_n] \\ \psi^n(x) &= -e^{-r(T_n - T_{n-1})}V^{n+1}(T_n, p, i, \hat{\imath}, 0) \end{aligned} \tag{14}$$

and substitute them in the function V^n , rewriting Eq. (13) as

$$\begin{aligned} V^{13}(t, x) &= 0 \qquad \forall (t, x) \in [0, T] \times \mathcal{S} \\ V^n(t, x) &= -\inf_{u \in \mathcal{A}} \mathbb{E}_{t,x} \left[\int_t^{T_n} -e^{-r(s-T_{n-1})}(P_s - \hat{\imath})u_s ds \right. \\ &\qquad \left. - e^{-r(T_n - T_{n-1})}V^{n+1}(T_n, X_{T_n}) \right] = \\ &= -\inf_{u \in \mathcal{A}} J^n(t, x; u) \end{aligned} \tag{15}$$

having defined the functions $J^n(t, x; u)$ as

$$J^n(t, x; u) = \mathbb{E}_{t,x} \left[\int_t^{T_n} L^n(s, X_s, u_s) ds + \psi^n(X_{T_n}) \right] \tag{16}$$

From now on in this section, we mainly apply, and when necessary extend, the results of [11]. There, the general problem faced has as value function of the form

$$V(t, x) = \inf_{u \in \mathcal{A}} J(t, x; u) \tag{17}$$

We should introduce a new sequence of value functions $\mathcal{V}^n(t, x) = -V^n(t, x)$ for which the results in [11] hold or, as an alternative, take always into account the negative sign. To avoid involved notation we will still write V^n instead of $-V^n$ and when necessary we will come back to the original problem by doing the sign substitution.

2 Main theoretical results about viscosity solution

Let us start this section with a definition which will be very important in the rest of the paper.

Definition 1 *We say that a function $f(x)$ and its derivatives until the k-th order have* polynomial growth*, and we indicate it with $f \in C_p^k(\mathbb{R}^n)$, if for all $i = 1, \dots, k$ there exist C, m such that:*

$$|f^{(i)}(x)| \leqslant C(1 + |x|^m) \qquad \forall x \in \mathbb{R}^n$$

The method of dynamic programming provides a powerful tool for approaching the problem in Eq. (17). When the value function $V(t, x)$ of the problem is smooth enough, it can be proved that it is a solution of a non-linear equation, known as the dynamic programming equation or Hamilton-Jacobi Bellman equation (see for example [23, Chapter 19]). However, in general (and in particular in our case) the value function is not smooth enough to satisfy the HJB equation in the classical sense, or we have no hints, at this stage, that the value function is smooth. A weaker formulation of solution to this equation is necessary if we want to pursue the method of dynamic programming. Crandall and Lions provided in [9] such a weak formulation which they called *viscosity solution*. In this and in the following section we introduce the definition of such a solution and we face the problem of pricing a swing contract with the theory arising from viscosity solution in order to show that the value function satisfies a dynamic programming equation and has some regularity.

This section is devoted to the definition of viscosity solutions of a general class of partial differential equation; here we prove some general results which will be used in the rest of this work. We end the section showing the links between viscosity solution and stochastic optimal control problems.

Let O be an open subset of $\mathbb{R}^n$ and define

$$Q = [0, T) \times O, \quad \overline{Q} = [0, T] \times \bar{O}, \quad Q_0 = [0, T) \times \mathbb{R}^n, \quad \overline{Q}_0 = [0, T] \times \mathbb{R}^n$$

Let

- $C(\overline{Q})$ be the set of continuous real valued functions defined on $\overline{Q}$
- $C^{1,2}(Q)$ be the set of all real valued functions on Q which are once continuously differentiable in the first variable and twice continuously differentiable in their second argument
- $C_p(Q)$ be the set of all real valued function on Q with polynomial growth.

Consider an equation of the kind

$$-V_t(t, x) + \mathcal{H}(t, x, D_x V(t, x), D_x^2 V(t, x)) = 0 \tag{18}$$

with $\mathcal{H}$ a continuous real valued function defined on the space $Q \times \mathbb{R}^n \times \mathcal{S}^n$ (here $\mathcal{S}^n$ is the set of all $n \times n$ symmetric matrices) such that

$$\mathcal{H}(t, x, p, A + B) \leqslant \mathcal{H}(t, x, p, A)$$

for all $(t,x) \in Q$, $p \in \mathbb{R}^n$, $A, B \in \mathcal{S}^n$ with $B \geqslant 0$. We introduce the following definition of viscosity solution:

Definition 2 *We say that a function $V \in C\left(\overline{Q}\right)$ is*

- *a viscosity subsolution of Eq. (18) if for each $v \in C^{1,2}(Q)$*

$$-v_t(\bar{t},\bar{x}) + \mathcal{H}(t,x,D_x v(\bar{t},\bar{x}), D_x^2(\bar{t},\bar{x})) \leqslant 0$$

  at every $(\bar{t},\bar{x}) \in Q$ which is a local maximum of $V - v$ on Q

- *a viscosity supersolution of Eq. (18) if for each $v \in C^{1,2}(Q)$*

$$-v_t(\bar{t},\bar{x}) + \mathcal{H}(t,x,D_x v(\bar{t},\bar{x}), D_x^2(\bar{t},\bar{x})) \geqslant 0$$

  at every $(\bar{t},\bar{x}) \in Q$ which is a local minimum of $V - v$ on Q

- *a viscosity solution of Eq. (18) if it is both a viscosity subsolution and a viscosity supersolution of Eq. (18) in Q*

We call reference probability system ν a 4-uple $\nu = (\Omega, (\mathcal{F}_s)_{s\in[t,T]}, \mathbb{P}, W)$ where $(\Omega, \mathcal{F}_T, \mathbb{P})$ is a probability space, and W is a Brownian motion adapted to the filtration $(\mathcal{F}_s)_{s\in[t,T]}$. Given a compact set $U \subseteq \mathbb{R}^{n_1}$, we denote by

$$\mathcal{A}_{t,\nu} = \{u \quad \text{s.t.} \quad u \text{ is a progressively measurable } U\text{-valued process defined on } \nu\}$$

We then suppose that X is a $\mathbb{R}^n$-valued process governed by the stochastic differential equation

$$dX_s = f(s, X_s, u_s)dt + \Sigma(s, X_s, u_s)dW_s, \qquad s \in [t,T] \tag{19}$$

given the initial data $X(t) = x$ and with $u \in \mathcal{A}_{t,\nu}$. We make the following

Assumption 1 *We assume that:*

a) U is compact

b) f, Σ are continuous on $\overline{Q}_0 \times U$ and $f(\cdot,\cdot,u)$ and $\sigma(\cdot,\cdot,u)$ are of class $C^1\left(\overline{Q}_0\right)$ for each $u \in U$;

c) f, Σ are itz with respect to their second argument, i.e. there exists a constant $L \geqslant 0$ such that for all $t \in [0,T]$ and for all $u \in U$ the following hold:

$$|f(t,x,u) - f(t,y,u)| \leqslant L|x-y|$$
$$|\sigma(t,x,u) - \sigma(t,y,u)| \leqslant L|x-y|$$

d) for suitable C_1, C_2

$$|f_t| + |D_x f| \leqslant C_1, \qquad\qquad |\sigma_t| + |D_x\sigma| \leqslant C_1$$
$$|f(t,0,u)| + |\sigma(t,0,u)| \leqslant C_2, \quad \forall u \in U$$

Remark 2 Assumption 1*(c)* and 1*(d)* lead to (both for f and Σ):

$$|f(t,x,u)| \leqslant |f(t,0,u)| + |f(t,x,u) - f(t,0,u)| \leqslant C_2 + C|x| \leqslant \hat{C}(1+|x|)$$

for a suitable constant $\hat{C}$, i.e. the drift and the volatility have linear growth in x.

Let us consider the following general optimal control problem. We want to choose a control $\tilde{u} \in \mathcal{A}_{t,\nu}$ which minimize the function

$$J(t,x;u) = \mathbb{E}^{\nu}_{t,x}\left[\int_t^T L(s,X_s,u)ds + \psi(X_T)\right]$$

where $\mathbb{E}^{\nu}_{t,x}$ is the expectation with respect to $\mathbb{P}^{\nu}_{t,x}$ which is the probability under which X has the dynamics given by Eq. (19) with initial condition $X_t = x$ and L and ψ are *continuous* functions with polinomial growth, i.e. such that

$$|L(t,x,u)| \leqslant C_3(1+|x|^m) \tag{20}$$

$$|\psi(x)| \leqslant C_3(1+|x|^m) \tag{21}$$

for suitable constants $C_3 > 0$ and $m \geqslant 0$. Then we consider, for a fixed probability system ν, the infimum of J among all $u \in \mathcal{A}_{t,\nu}$:

$$V_\nu(t,x) = \inf_{u\in\mathcal{A}_{t,\nu}} J(t,x;u) \tag{22}$$

and finally we define the *value function* as:

$$V(t,x) = \inf_{\nu} V_\nu(t,x) \tag{23}$$

This problem is linked, and in this section we will detail this link, to a partial differential equation of the kind of Eq. (18), called *dynamic programming equation* or *Hamilton-Jacobi-Bellman equation (HJB equation)*, obtained by imposing

$$\mathcal{H}(t,x,p,A) = \sup_{u\in U}\left\{-f(t,x,u)\cdot p - \frac{1}{2}\operatorname{tr}(A\cdot(\Sigma\Sigma')(t,x,u)) - L(t,x,u)\right\} \tag{24}$$

and the boundary condition

$$V(T,x) = \psi(x) \quad \forall x \in \mathbb{R}^n$$

One important tool in proving that V is a viscosity solution of Eq. (18), with $\mathcal{H}$ as in Eq. (24), is the so called *dynamic programming* property for the value function.

Definition 3 *We say that a function W has property (DP) (dynamic programming) if for every reference probability system ν, for every control $u \in \mathcal{A}_{t,\nu}$ and every stopping time θ taking values in $[t,T]$ we have*

$$W(t,x) \leqslant \mathbb{E}_{t,x}\left[\int_t^\theta L(s,X_s,u_s)ds + W(\theta,X_\theta)\right]$$

and for every $\delta > 0$ there exists a ν and a control $u \in \mathcal{A}_{t,\nu}$ such that

$$W(t,x) + \delta \geqslant \mathbb{E}_{t,x}\left[\int_t^\theta L(s,X_s,u_s)ds + W(\theta,X_\theta)\right]$$

for every stopping time θ taking values in $[t,T]$.

2.1 Existence

The next result shows the links between the notion of viscosity solution of HJB equation and the corresponding stochastic optimal control problem.

Theorem 1 *If Assumption 1 holds, Q and the control set U are bounded and*

i. $\psi \in C^2(\mathbb{R}^n)$, *i.e. the final condition ψ is a continuous function, twicely continuously differentiable on $\mathbb{R}^n$*

ii. the running cost L is a continuous function on $\overline{Q}_0 \times U$ and has polynomial growth on its second argument, i.e.

$$|L(t,x,u)| \leqslant C_3(1+|x|^m) \quad \forall (t,u) \in [0,T] \times U$$

for some $C_3 \geqslant 0, m \geqslant 0$

then:

a) $V \in C(\bar{Q}_o)$, *i.e. the value function $V(t,x)$ defined in Eq. (23) is a continuous function on $\bar{Q}_0$*

b) property (DP) holds for the value function $V(t,x)$ defined in Eq. (23)

c) $V = V_\nu$ for every reference probability system ν

d) V is a viscosity solution of Eq. (18), with $\mathcal{H}$ defined in Eq. (24), in Q_0

Proof For *(a-c)* see [11, Theorem 7.1 pag. 178]. For *(d)* see [11, Corollary 3.1 pag. 209]. □

Summing up, under the hypothesis of Theorem 1, the PDE (18) can be used to find out a solution of our problem. Unfortunately, in our case (and in a lot of other cases arising from financial application) the functions $V^n(t,x)$, which are both the value function for month n but also the final condition for the problem at month $n-1$, are far from being bounded, mainly because the spread $P_t - I_t$ is not bounded. In this case we can not use Theorem 1. What we want to do in the following is to prove an extension of Theorem 1, which uses only the polynomial growth of the final condition.

Theorem 2 *If Assumption 1 holds and $\psi \in C_p(Q_0)$, then V is a viscosity solution Eq. (18), with $\mathcal{H}$ defined in Eq. (24). Moreover, $V = V_\nu$ for every reference probability system ν.*

In order to prove this theorem, we state and prove two intermediate results. The first one states that if the value function has property (DP) and it is a continuous function with polynomial growth, then it is a viscosity solution of the HJB equation. Let us remark that the weak condition $\psi \in C_p(Q_0)$ ensures (in the same way of the next Proposition 1) only the polynomial growth of V and nor the property (DP) nor the continuity.

Theorem 3 *If property (DP) holds for the value function V and $V \in C_p(Q_0)$ then V is a viscosity solution of Eq. (18), with $\mathcal{H}$ defined in Eq. (24), in Q_0.*

Proof See [11, Theorem 5.1, pag 72] $\square$

Lemma 1 *If $\psi \in C_p(\mathbb{R}^n)$, then there exists a sequence $(\psi_m)_{m\in\mathbb{N}}$ in $C_p^2(\mathbb{R}^n)$ such that $\psi_m \to \psi$ uniformly on compact sets. Moreover, there exists $C, k \geqslant 0$ such that*

$$|\psi_m(x)| \leqslant C(1 + |x|^k)$$

uniformly with respect to m.

Proof For $m \in \mathbb{N}$, let us define the sequence of functions $(\rho_m)_{m\in\mathbb{N}} \subseteq C^\infty(\mathbb{R}^n)$ such that $\rho_m \geqslant 0$, $\rho_m(y) = 0$ if $|y| \geqslant \frac{1}{m}$ and

$$\int_{\mathbb{R}^n} \rho_m(y)dy = 1$$

Now introduce the sequence $(\psi_m)_{m\in\mathbb{N}}$ as

$$\psi_m(x) = (\psi * \rho_m)(x) = \int_{\mathbb{R}^n} \psi(y)\rho_m(x-y)dy$$

Then $\psi_m \in C^\infty(\mathbb{R}^n)$. Moreover, being $\psi \in C_p(\mathbb{R}^n)$ then ψ is uniformly continuous on each compact set K: for all $\varepsilon > 0$ there exists $\delta > 0$ (which depends on ε and K) such that

$$|\psi(x-y) - \psi(x)| < \varepsilon, \quad \forall x \in K, \quad |y| \leqslant \delta$$

Then for all $x \in K$, $m > \frac{1}{\delta}$

$$\begin{aligned}
|\psi_m(x) - \psi(x)| &= \left| \int_{\mathbb{R}^n} \psi(x-y)\rho_m(y)dy - \psi(x) \right| = \\
&= \left| \int_{\mathbb{R}^n} (\psi(x-y) - \psi(x))\rho_m(y)dy \right| \leqslant \\
&\leqslant \int_{|y|\leqslant\frac{1}{m}} |(\psi(x-y) - \psi(x))|\rho_m(y)dy \leqslant \\
&\leqslant \varepsilon \int_{|y|\leqslant\frac{1}{m}} \rho_m(y)dy = \varepsilon
\end{aligned}$$

So $\psi_m \to \psi$ uniformly on compact sets. Moreover because $\psi \in C_p(\mathbb{R}^n)$

$$\begin{aligned}
|\psi_m(x)| &\leqslant \int_{|y|\leqslant\frac{1}{m}} |\psi(x-y)|\rho_m(y)dy \leqslant \\
&\leqslant \int_{|y|\leqslant\frac{1}{m}} C(1 + (|x| + |y|)^k)|\rho_m(y)dy \leqslant \\
&\leqslant C(1 + (|x| + \frac{1}{m})^k) \int_{|y|\leqslant\frac{1}{m}} \rho_m(y)dy \leqslant C(1 + (|x|+1)^k)
\end{aligned}$$

so also ψ_m has polynomial growth and the uniform estimate holds. $\square$

Let us now recall that in [11, Appendix D] the following inequality is proved:

$$\mathbb{E}_{t,x}[\|X_\cdot\|_\infty^m] =\leqslant \xi_m(1+|x|^m) \tag{25}$$

which holds $\forall m \geqslant 0$, with ξ_m constant depending only on $T-t$ and on C_1, C_2 of Assumption 1. Finally, using the Markov inequality we get

$$\mathbb{P}_{t,x}\{\|X\|_\infty \geqslant M\} \leqslant \frac{\varsigma}{M}(1+|x|) \tag{26}$$

We can now prove Theorem 2.

Proof We would like to use the result of Theorem 1 applied to the value function V. In order to do this, we need to prove that V is a continuous function with property (DP).

Let $(\psi_m)_{m\in\mathbb{N}}$ be a sequence in $C_p^2(\mathbb{R}^n)$ such that $\psi_m \to \psi$ uniformly on compact sets, as described in Lemma 1. Let $V_{m,\nu}$ and V_m the corresponding value functions, i.e. the value functions of stochastic optimal control problems with final conditions ψ_m. Let V the value function with final condition ψ.

Thanks to Theorem 1 we know that $V_{m,\nu} = V_m$ for every reference probability system, property (DP) holds for V_m and V_m are continuous functions.

We now prove that $V_m \to V$ uniformly on compact sets and so that V is continuous. By definition of V and V_m, for each $\delta > 0$ there exists ν and $u \in \mathcal{A}_{t,\nu}$ such that

$$\begin{aligned}
&V(t,x) + \delta - V(t,x) \leqslant \\
&\mathbb{E}_{t,x}\left[\int_t^T L(s,X_s,u)ds + \psi_m(X_T)\right] - V(t,x) \leqslant \\
&\leqslant \mathbb{E}_{t,x}[\psi_m(X_T) - \psi(X_T)] = \\
&= \mathbb{E}_{t,x}[(\psi_m(X_T) - \psi(X_T))(\mathbf{1}_{|X_T|\leqslant M} + \mathbf{1}_{|X_T|>M})] \leqslant \\
&\leqslant \underbrace{\|\psi_m - \psi\|_{B(0,M)}}_{=I_1} + \underbrace{\mathbb{E}_{t,x}[(\psi_m(X_T) - \psi(X_T))\mathbf{1}_{|X_T|>M}]}_{=I_2}
\end{aligned}$$

where $\|\cdot\|_{B(0,M)}$ denotes the sup norm in $B(0,M)$. An analogous inequality holds for $V(t,x) - V_m(t,x) + \delta$.

We have that, for all $M \geqslant 0$, $I_1 \to 0$ as $m \to \infty$ thanks to the uniform convergence of $(\psi_m)_{m\in N}$ on compact sets. Since ψ and ψ_m have polynomial growth, using Jensen's inequality, the well known inequality $2xy \leqslant x^2 + y^2$ and the ones in Eq. (25 - 26), we obtain:

$$
\begin{aligned}
I_2 &= \mathbb{E}_{t,x}[(\psi_m(X_T) - \psi(X_T))\mathbf{1}_{|X_T|>M}] \leqslant \\
&\leqslant \mathbb{E}_{t,x}[|\psi_m(X_T)| + |\psi(X_T)|\mathbf{1}_{|X_T|>M}] \\
&\leqslant \mathbb{E}_{t,x}[2C(1+|X_T|^k)\mathbf{1}_{|X_T|>M}]) \leqslant \\
&\leqslant (\mathbb{E}_{t,x}[4C^2(1+|X_T|^k)^2]\mathbb{E}_{t,x}[\mathbf{1}_{|X_T|>M}]) = \\
&= (\mathbb{E}_{t,x}[4C^2(1+|X_T|^k)^2]\mathbb{P}_{t,x}\{|X_T| > M\}) \leqslant \\
&\leqslant (4C^2\mathbb{E}_{t,x}[1+2\|X_\cdot\|_\infty^k + \|X_\cdot\|_\infty^{2k}]\mathbb{P}_{t,x}\{\|X_\cdot\|_\infty > M\}) \leqslant \\
&\leqslant \left(8C^2(1+\mathbb{E}_{t,x}[\|X_\cdot\|_\infty^{2k}])\frac{\varsigma}{M}(1+|x|)\right) \leqslant \\
&\leqslant \left(\frac{C_1}{M}(1+|x|^{2k})(1+|x|)\right)^{\frac{1}{2}} \leqslant \frac{C_2}{M}(1+|x|^{k+1})
\end{aligned}
$$

so I_2 can be made arbitrarily small by choosing a suitable M. This imply that $V_m \to V$ on compact sets, hence V is continuous.

We now prove that property (DP) holds for V. Given an arbitrary stopping time θ

$$
\begin{aligned}
&\left|\mathbb{E}_{t,x}\left[\int_t^\theta L(s,X_s,u_s)ds + V(\theta,X_\theta) - \int_t^\theta L(s,X_s,u_s)ds - V_m(\theta,X_\theta)\right]\right| \leqslant \\
&\leqslant |\mathbb{E}_{t,x}[V(\theta,X_\theta) - V_m(\theta,X_\theta)]| \leqslant \\
&\leqslant \mathbb{E}_{t,x}[(\mathbf{1}_{|X_\theta|\leqslant M} + \mathbf{1}_{|X_\theta|>M})|V(\theta,X_\theta) - V_m(\theta,X_\theta)|] \leqslant \\
&\leqslant \underbrace{\|V - V_m\|_{B(0,M)}}_{=I_3} + \underbrace{\mathbb{E}_{t,x}[\mathbf{1}_{|X_\theta|>M}|V(\theta,X_\theta) - V_m(\theta,X_\theta)|])}_{=I_4}
\end{aligned}
$$

We just proved that that $I_3 \to 0$ as $m \to \infty$. Let us remember that we are assuming that the running cost L and the final condition ψ has polynomial growth in x. This implies that, by its definition, also V has polynomial growth in its second argument. Combined with the results in Lemma 1 and using the same strategy used for I_2, we get:

$$
\begin{aligned}
I_4 &= \mathbb{E}_{t,x}[\mathbf{1}_{|X_\theta|>M}|V(\theta,X_\theta) - V_m(\theta,X_\theta)|]) \leqslant \\
&\leqslant (\mathbb{P}_{t,x}[|X_\theta| > M]\mathbb{E}_{t,x}[4C^2(1+|X_\theta|^k)^2]) \leqslant \\
&\leqslant (\mathbb{P}_{t,x}[\|X_\cdot\|_\infty > M]C_1(1+\mathbb{E}_{t,x}[\|X_\cdot\|_\infty^2])) \leqslant \frac{C_2}{M}(1+|x|^{k+1})
\end{aligned}
$$

also I_4 can be made arbitrarily small by choosing a suitable M and x in a given compact set. Summing up, for each $\delta > 0$ there exist M and m such that $I_3 <$ and $I_4 <$. Finally, thanks to property (DP) of V_m, there exists a ν and a control $u \in \mathcal{A}_{t,\nu}$ such that

$$
V_m(t,x) + \frac{\delta}{3} \geqslant \mathbb{E}_{t,x}\left[\int_t^\theta L(s,X_s,u_s)ds + V_m(\theta,X_\theta)\right]
$$

for every stopping time θ taking values in $[t,T]$. By putting together these three inequalities, we obtain property (DP) for V.

In conclusion, we have proved that V is continuous, has polynomial growth, and has property (DP). From Theorem 1 we can conclude. $\square$

2.2 Uniqueness

Theorem 4 *Let us assume the hypothesis in Assumption 1 and in addition that*

i. the running cost $L(t, x, u)$ and the final condition $\psi(x)$ are continuous functions with quadratic growth, i.e. $m \leqslant 2$ in Eq. (20) and (21)

Let V_1, V_2 be two viscosity solution of problem (18), with $\mathcal{H}$ as in Eq. (24) and the final condition $V(T, x) = \psi(x)$, having quadratic growth, i.e.

$$|V_i(t, x)| \leqslant C(1 + |x|^2) \quad \forall (t, x) \in Q_0, \quad i = 1, 2$$

Then $V_1 = V_2$.

Proof The proofs follows from Theorem 2.1, Corollary 2.1 and Remark 2.2(iii) in [16]. □

2.3 First Derivative

In this subsection we present a general result which gives the existence of the first derivative for a general control problem. We continue to assume Assumption 1 and the result found in Eq. (29). In addition, we need the following stronger assumption on L^n and ψ^n:

Assumption 2 *We assume that:*

i. L^n is continuous on $\overline{Q}_0 \times U$, $L^n(\cdot, \cdot, u) \in C^1(\bar{Q}_0)$ for each $u \in U$ and:

$$|L^n_t| + |L^n_x| \leqslant C_4(1 + |x|^\ell) \tag{27}$$

ii. ψ^n is locally Lipschitz

Let us now introduce the definition of different quotients $\Delta^h_\xi V^n$, which is fundamental in this section because in order to prove existence and smoothness of $V^n(t, x)$, we first need bounds for those quotients, and then a general result allows to conclude the existence of the derivatives $V^n_x(t, x) = D_x V^n \in L^p_{\text{loc}}$ for $p > 1$.

Definition 4 *We call* ***difference quotients*** *of the function $f(t, x)$ of size h and direction ξ the quantities:*

$$\Delta^h_\xi f(t, x) = \frac{f(t, x + h\xi) - f(t, x)}{h}$$

where $\xi \in \mathbb{R}^n$ is a direction, i.e. it is such that $|\xi| = 1$.

As for the existence of the solution, a lot of results on the existence of the derivatives are available for the case $\psi \in C^2(\mathbb{R}^n)$, for instance [11, Lemma 8.1, pag 183], but this is not our case. So we now extend the results to the case where $\psi \in C^0_p(\mathbb{R}^n)$ and it is also locally Lipschitz. This result can be used in a straightforward manner for $n = 12$ using the piecewise linear definition of $\Psi(x)$ and needs to be adapted by induction for $n = 1, \ldots, 11$. We state the lemma for a generic ψ and L.

Lemma 2 *If Assumptions 1 and 2 hold and the first derivative of the final condition $\psi_x(x)$ exists a.s. and has polynomial growth*

$$|\psi_x| \leqslant C_4(1+|x|^k) \tag{28}$$

then there exists M_1 which depends on C_1, C_2, C_4, k, T such that for all directions ξ

$$|\Delta_\xi^h J| \leqslant M_1(1+|x|^k)$$

for every $h \in (0,1]$.

Proof Given $(t, x_0) \in Q_0$, let $(X_z)_{z\in[t,T]}$ be the solution of

$$dX_s = f(s, X_s, u_s)dt + \sigma(s, X_s, u_s)dW_s, \qquad s \in [t,T]$$

with the initial condition $X_t = x_0$ and $(X_s^h)_{s\in[t,T]}$ the solution with initial condition $X_t = x_0 + h\xi$. Also, let $\Delta^h X_s = \frac{X_s^h - X_s}{h}$. Since L and ψ are Lipschitz, then their restriction to each line segment $\{X_s^\lambda | X_s^\lambda = (1-\lambda)X_s + \lambda X_s^h, \lambda \in [0,1]\}$ is absolutely continuous and the Fundamental Theorem of Calculus holds (see [12, pag. 102]), so we have

$$\begin{aligned}\Delta_\xi^h J(t,x;u) &= \mathbb{E}\left[\frac{1}{h}\int_t^T (L(s, X_s^h, u_s) - L(s, X_s, u_s))ds + \frac{1}{h}(\psi(X_T) - \psi(X_T^h))\right] = \\ &= \mathbb{E}\left[\int_t^T \int_0^1 L_x(s, X_s^\lambda, u_s)\cdot \Delta^h X_s d\lambda\right] + \mathbb{E}\left[\int_0^1 \psi_x(X_T^\lambda)\cdot \Delta^h X_T d\lambda\right]\end{aligned}$$

By Equation (27)

$$\left|\int_0^1 L_x(s, X_s^\lambda, u_s)\right| \leqslant \int_0^1 C_4(1+|X_s^\lambda|^k)d\lambda \leqslant M(1+|X_s|^k + |X_s^h|^k)$$

By Equation (28)

$$\left|\int_0^1 \psi_x(X_T^\lambda)d\lambda\right| \leqslant \int_0^1 C_4(1+|X_T^\lambda|^k)d\lambda \leqslant M(1+|X_T|^k + |X_T^h|^k)$$

By Cauchy-Schwartz

$$|\Delta_\xi^h J| \leqslant 2M\left(\mathbb{E}\left[\int_t^T (1+|X_T|^k + |X_T^h|^k)^2\right]\right)(\mathbb{E}[|\Delta X_T^h|^2])$$

We bound the first term on the right hand side using (25) with $m = 2k$ and $x = x_0, x_0 + h\xi$. We also have that $\mathbb{E}[|\Delta X_T^h|^2] \leqslant B$ (see [11], Appendix D) where B depends on bounds for $|f_x|$ and $|\sigma_x|$ and the costant C_1 on Assumption 1. Since $|\xi| = 1$ and $0 < h \leqslant 1$

$$1 + |x|^{2k} + |x+h\xi|^{2k} \leqslant C_k(1+|x|^{2k})$$

for suitable C_k. □

The following Theorem gives the existence of $V_x(t,x)$ and it is stated for generic final condition ψ and running cost L.

Theorem 5 *If Assumption 1, 2 and Equation (28) hold, then $V_x(t,x)$ exists and it is in $L^p_{\text{loc}}(Q_0)$ for every $p \in (1,\infty]$. Moreover*

$$|V_x(t,x)| \leqslant M_1(1+|x|^k)$$

for almost every $(t,x) \in Q_0$, where M_1 depends on C_1, C_2, C_3, C_4, k, T.

Proof We take a generic open bounded set B. Then by Lemma 2 we have

$$|\Delta_\xi^h J(t,x,u)| \leqslant M_1(1+|x|^k) \qquad \forall (t,x) \in B$$

Since these bounds are the same for all controls u, we obtain that

$$|\Delta_\xi^h V(t,x)| \leqslant M_1(1+|x|^k)$$

Then we take $p > 1$ and an open set A such that $B \subseteq A$ and $\text{dist}(B, \partial A) < \min\{1, T\}$ and we have that $\Delta_\xi^h V(t,x) \in L^p(B)$ and

$$\|\Delta_x^h V\|_{L^p(B)} \leqslant M_3 \|1 + |x|^k\|_{L^p(B)}$$

for all $h \in (0, \min\{1,T\})$, where M_3 depends on M_1 and M_2. This implies (see [15], pag 246-248) that $V_x(t,x) \in L^p(B)$ and $\|V_x\|_{L^p(B)} \leqslant \|M_3(1+|x|^k)\|_{L^p(B)}$. Moreover, $V_x(t,x)$ is also the derivative in the Sobolev sense. In fact, for each $\varphi \in C_0^\infty((0,T) \times \mathbb{R}^n)$

$$\int \varphi V_i = \int \lim_{h\to 0} \varphi \Delta_i^h V = \lim_{h\to 0} \int \varphi \Delta_i^h V = -\lim_{h\to 0} \int V \Delta_i^h \varphi = -\int V \varphi_i$$

and the conclusion follows. $\square$

3 Viscosity solution for swing contracts

In this section we apply results in Section 2 to the one year problem presented in Section 1. We continue to make the hypothesis in Assumption 1.

3.1 Existence

First of all, we notice that the functions $L^n(t,x,u)$ as defined is Eq. (14), being a linear function of p and $\hat{i}$, has polynomial growth in its second argument x, i.e.

$$|L^n(t,x,u)| \leqslant C_3(1+|x|^m) \quad \forall t \in [0,T], \forall u \in U, \forall n = 1, \dots, 12 \tag{29}$$

In particular, being $L^n(t,x,u)$ a linear function of x, we can assume quadratic growth, i.e. $m \leqslant 2$. The same for $\Psi(x)$ defined in Eq. (9). As a consequence, also the functions $V^n(t,x)$ has quadratic growth, as proved by the following proposition.

Proposition 1 *The functions $V^n(t,x)$, as defined in formula (15), have quadratic growth for all $n = 1, \dots, 13$.*

Proof By backward induction. $V^{13}(t,x) = \Psi(z)$ has quadratic growth.

Now assume that $V^{n+1}(t,x)$ has quadratic growth, i.e. for some constants B_{n+1} and $m_{n+1} \leqslant 2$ we have

$$|V^{n+1}(t,x)| \leqslant B_{n+1}(1+|x|^{m_{n+1}}) \tag{30}$$

Using notation in Eq. (14), the result in Eq. (29), the inductive hypothesis in Eq. (30) and the inequalities (25) we get:

$$\begin{aligned}
|J^n(t,x,u)| &= \left|\mathbb{E}_{t,x}\left[\int_t^{T_n} L^n(s,X_s,u_s)ds + \psi^n(X_{T_n})\right]\right| \leqslant \\
&\leqslant \mathbb{E}_{t,x}\left[\int_t^{T_n} |L^n(X_s,u_s)ds| + \left|V^{n+1}(T_n,X_{T_n})\right|\right] \leqslant \\
&\leqslant \mathbb{E}_{t,x}\left[\int_t^{T_n} C_3(1+|X_s|^m)ds + B_{n+1}(1+|X_{T_n}|^{m_{n+1}})\right] \leqslant \\
&\leqslant \mathbb{E}_{t,x}\left[\int_t^{T_n} C_3(1+||X.||_\infty^m)ds + B_{n+1}(1+||X.||_\infty^{m_{n+1}})\right] = \\
&= C_3(T_{n+1}-t)(1+\mathbb{E}_t[||X.||_\infty^m]) + B_{n+1}(1+\mathbb{E}_t[||X.||_\infty^{m_{n+1}}]) \leqslant \\
&\leqslant \varsigma + \xi_m(1+|x|^m) + B_{n+1}B_{m_{n+1}}(1+|x|^{m_{n+1}}) \leqslant \\
&\leqslant B_n(1+|x|^{m_n})
\end{aligned} \tag{31}$$

with $m_n = \max\{m, m_{n+1}\} \leqslant 2$ being both $m \leqslant 2$ (see Eq. (29)) and $m_{n+1} \leqslant 2$ (inductive hypothesis). B_n is a suitable constant which depends only on $T_n - t$, C_3, B_{n+1}. □

We prove now that V^{12} is continuous and has property (DP). Then by induction we will extend the same results also to other functions V^n for $n = 1, \ldots, 11$

Theorem 6 *If Assumptions 1 hold, then $V^{12}(t,x) \in C([T_{11},T] \times \mathbb{R}^3)$ and has property (DP). Moreover, $V^{12} = V^{12}_\nu$ for every reference probability system.*

Proof Being $\Psi(z)$ a bounded and uniformly continuous function, we can apply Corollary 7.1, pag 181, in [11].

Alternatively, we can use the same idea in the proof of Theorem 2 since $\Psi(z)$ has polynomial growth. □

Now the main result for $V^{12}(t,x)$.

Theorem 7 *The function $V^{12}(t,x)$ is a viscosity solution of the dynamic programming equation:*

$$-V_t^{12} + \mathcal{H}(t,x,D_xV^{12},D_x^2V^{12}) = 0 \tag{32}$$

in $[T_{11},T) \times \mathbb{R}^4$, where $\mathcal{H}$ in this case reads:

$$V_t + \frac{1}{2}\sigma_p^2V_{pp} + \rho\sigma_p\sigma_iV_{pi} + \frac{1}{2}\sigma_i^2V_{ii} + \mu_pV_p + \mu_iV_i - \sup_{u\in U}\{e^{-r(t-T_{11})}(p-\hat{\imath}-V_z)u\} = 0 \tag{33}$$

Proof See [11, Corollary 3.1, pag 209]. Alternatively, we can use Theorem 2 since $\Psi(z)$ has polynomial growth. $\square$

Remark 3 We now come back to our original problem with the right minus sign by substituting the function V with $-V$ in Eq. (33). Thus the correct equation satisfied by the value function in formula (15) is:

$$\begin{aligned} & V_t + \frac{1}{2}\sigma_p^2 V_{pp} + \rho\sigma_p\sigma_i V_{pi} + \frac{1}{2}\sigma_i^2 V_{ii} + \mu_p V_p + \mu_i V_i + \\ & \quad + \sup_{u \in U}\{e^{-r(t-T_{n-1})}(p - \hat{i} + V_z)u\} = 0 \\ & V^{12}(T_{12}, x) = e^{-r(T_{12}-T_{11})}\Psi(x) \end{aligned} \tag{34}$$

Note 1 By substituting $\tilde{V}(t,x) = e^{-r(t-T_n)}V(t,x)$ in Eq. (34) we obtain an equation analogous to the one in [4]:

$$V_t - rV + \frac{1}{2}\sigma_p^2 V_{pp} + \rho\sigma_p\sigma_i V_{pi} + \frac{1}{2}\sigma_i^2 V_{ii} + \mu_p V_p + \mu_i V_i + \sup_{u \in U}\{(p - \hat{i} + V_z)u\} = 0$$

We now extend the results for $V^{12}(t,x)$ also to the other value functions V^n for $n = 1, \dots, 11$.

Theorem 8 *For every $n = 1, \dots, 12$ the function V^n is a viscosity solution of the HJB equation (33) in $[T_n, T) \times \mathbb{R}^4$, with V replaced by V^n, and with final condition:*

$$V^n(T_n, x) = \psi^n(x) = e^{-r(T_n - T_{n-1})}V^{n+1}(T_n, x)$$

Moreover, $V^n = V^n_\nu$ for every reference probability system.

Proof Going backward in time, by induction, Theorem 7 states that $V^{12}(t,x)$ is a viscosity solution of Eq. (33). Let us suppose that $V^{n+1}(t,x)$ is a viscosity solution of Eq. (33). Thanks to Proposition 1, $V^{n+1}(t,x) \in C_p(Q_0)$. By Theorem 2, the function $V(t,x) = V^n(t,x)$ is the solution of

$$-V_t + \mathcal{H}(t, x, DV, D^2V) = 0$$

with boundary condition

$$V^n(T_n, x) = e^{-r(T_n - T_{n-1})}V^{n+1}(T_n, x)$$

$\square$

3.2 Uniqueness

We want to apply Theorem 4 to our problem.

To do this, we notice that the control set U is bounded, the running cost functions $L^n(t,x,u)$ are continuous with quadratic growth in x and thanks to Proposition 1 and Theorem 6 the final conditions

$$\psi^n(x) = e^{-r(T_n-T_{n-1})}V^{n+1}(T_n,p,i,\iota,0)$$

are also continuous functions with quadratic growth in x. Assumptions (1) are supposed to be satisfied. We can apply Theorem 4.

3.3 First derivative

In this section we prove that the first derivatives of our value functions, V_x^n, exist.

We notice that condition *(i)* of Assumption 2 is verified for all $n=1,\ldots,12$ in our case, that is when L^n is the one in Equation (14). Also condition *(ii)* can be easily proved:

Proposition 2 *For every $n=1,\ldots,12$, the value functions $V^n(t,x)$ are Lipschitz.*

Proof For all $u\in U$, $L^n(t,x,u)$ are Lipschitz for every $n=1,\ldots,12$:

$$\begin{aligned}
&|L^n(t,x_1,u)-L^n(t,x_2,u)| = \\
&= |-e^{-r(t-T_{n-1})}(p_1-\widehat{i_1})u + e^{-r(t-T_{n-1})}(p_2-\widehat{i_2})u| \leqslant \\
&\leqslant |(p_2-p_1)+(\hat{i}_2-\hat{i}_1)|\bar{u} \leqslant \\
&\leqslant (|p_2-p_1|+|\hat{i}_2-\hat{i}_1|)\bar{u} \leqslant \\
&\leqslant \|x_1-x_2\|_1\bar{u}
\end{aligned}$$

Starting from $n=12$, the final condition Ψ defined in Eq. (9) is a piecewise linear function and so it is Lipschitz. This implies that for all control u also $J^{12}(t,x;u)$ is Lipschitz and so $V^{12}(t,x)$ is Lipschitz. Backward induction on n completes the proof. □

The main result of this section is the following theorem.

Theorem 9 *For every $n=1,\ldots,12$, if Assumptions 1 and 2 hold, then the derivatives $V_x^n(t,x)$ exist, they are in $L^p_{\text{loc}}(Q_0)$ for every $p\in(1,+\infty]$ and for almost every $(t,x)\in Q_0$ we have*

$$|V_x^n(t,x)| \leqslant M_1(1+|x|^\ell)$$

where M_1 depends on C_1,C_2,C_3,C_4,k,T.

Moreover, we prove that also the derivatives have polynomial growth. Again, backward induction is the key to prove that the value function has a first derivative. We start with two corollaries.

Corollary 1 *There exists M_1 which depends on C_1, C_2, C_4, k, T such that for all directions ξ*

$$|\Delta_\xi^h J^{12}| \leqslant M_1(1 + |x|^k)$$

for every $h \in (0, 1]$.

Proof Apply Lemma 2 with the final condition $\Psi(x)$ defined in Eq. (9) which is a piecewise linear Lipschitz function with $\Psi_x(x)$ piecewise constant defined almost everywhere. □

Corollary 2 *The first derivative V_x^{12} exists and it is in $L^p_{\mathrm{loc}}(Q_0)$ for every $p \in (1, \infty]$. Moreover*

$$|V_x^{12}(t, x)| \leqslant M_1(1 + |x|^k)$$

for almost every $(t, x) \in Q_0$, where M_1 depends on C_1, C_2, C_3, C_4, k, T.

Proof Thanks to the result of Corollary 1 we can apply Theorem 5 to $V^{12}(t, x)$ with the final condition $\Psi(x)$. □

Now we prove the main result of this section, i.e. Theorem 9.

Proof We know that $\psi^n(x)$ are Lipschitz thanks to Proposition 2. Let us suppose that for $n \leqslant 11$ the derivatives ψ_x^n exists and satisfies

$$|\psi_x^n(x)| = |V_x^{n+1}(T_n, x)| \leqslant M_1(1 + |x|^\ell) \tag{35}$$

This is true for $n = 11$, as proved in Corollary 2. We apply Lemma 2 to bound the difference quotients

$$|\Delta_\xi^h J^n(t, x; u)| \leqslant \tilde{M}(1 + |x|^\ell)$$

and then apply Theorem 5 to obtain the existence of $V_x^n(t, x) \in L^p_{\mathrm{loc}}$ and its polinomyal growth:

$$|V_x^n(t, x)| \leqslant M_n(1 + |x|^\ell)$$

This completes the proof. □

3.3.1 Existence of the optimal control

We proved that the first derivatives V_x^n exists. Coming back to our HJB equation in (34), we can state that also a candidate for the optimal control is a.s. well defined. In fact, a straight calculation leads from (34) to:

$$u^* = u^*(t, x) = \begin{cases} 0 & \text{if} \quad p - \hat{i} + V_z(t, x) \leqslant 0 \\ \bar{u} & \text{if} \quad p - \hat{i} + V_z(t, x) > 0 \end{cases} \tag{36}$$

Remark 4 As observed in [4], the candidate optimal control in Eq. (36) has a nice economical interpretation. In fact, the marginal value V_z says how much the contract value falls down if we increase the cumulted withdrawn quantity z, i.e. if we decide to exercise the swing option (i.e. to buy gas). What this control says is that we have to exercise the option only if the spread payoff $p - \hat{i}$ (which is the marginal profit we face if we exercise) dominates the lost option value V_z.

By inserting the candidate optimal control (36) into HJB Equation (34), we obtain a linear partial differential equation for the value function V for which classical smoothness results hold (see, for instance, [18]).

4 Numerical methods

This section focuses on numerical methods to find the price of swing contracts. First of all, in Section 4.1 we introduce a more concrete dynamics for the prices: we use particular cases of the model in [22] which are rather standard models for energy prices (see for example [13, Chapter 23.3] and [20]). In Section 4.2 a finite difference method for Eq. (34) is presented. Section 4.3 deals with a popular method used among practitioners: Least Square Monte Carlo, which works on a discrete version of the value function in Eq. (11), and does not use the HJB equation. This method is not accurate as finite differences, but it is easy to implement, even if it suffers of some drawbacks we will discuss later.

All the algorithms we present work in discrete time. In the whole chapter we assume that the time intervals $[0, T]$ and $[T_n, T_{n+1}]$ are discretized into appropriate sequences which will be defined time by time when necessary. For the finite difference algorithm, also the intervals on which the prices lie has to be bounded and discretized, while the other method takes advantage from Monte Carlo simulations of path prices.

4.1 Price dynamics

We assume that the log-prices of the spot gas price $\mathcal{P}_t = \log P_t$ and spot index price $\mathcal{I}_t = \log I_t$ follow the mean reverting dynamics

$$d\mathcal{P}(t) = \theta_p(\mu_p - \mathcal{P}(t))dt + \sigma_p dW_p(t)$$
$$d\mathcal{I}(t) = \theta_i(\mu_i - \mathcal{I}(t))dt + \sigma_i dW_i(t)$$

whose solutions at time s, given the states $\mathcal{P}(t)$ and $\mathcal{I}(t)$ at time $t < s$, are

$$\mathcal{P}(s) = (\mathcal{P}(t) - \mu_p)e^{-\theta_p(s-t)} + \mu_p + \sigma_p \int_t^s e^{\theta_p(u-s)} dW_p(u) \tag{37}$$

$$\mathcal{I}(s) = (\mathcal{I}(t) - \mu_p)e^{-\theta_i(s-t)} + \mu_i + \sigma_i \int_t^s e^{\theta_i(u-s)} dW_i(u) \tag{38}$$

The processes W_p and W_i are two Brownian motions with mutual correlation ρ. The realizations of the log-prices are defined using the notation:

$$\mathfrak{p}_t = \log(p_t) = \log(P_t(\omega)) = \mathcal{P}_t(\omega)$$
$$\mathfrak{i}_t = \log(i_t) = \log(I_t(\omega)) = \mathcal{I}_t(\omega)$$

We suppress the subscript t when clear from the context.

The conditional mean and variance for the log-processes $\mathcal{P}(t)$ and $\mathcal{I}(t)$ can be derived from Equations (37-38)

$$\begin{aligned}
m_{\mathfrak{p}}(t,\mathfrak{p},s) &= \mathbb{E}_{t,x}[\mathcal{P}(s)] = (\mathcal{P}(t)-\mu_p)e^{-\theta_p(s-t)} + \mu_p = (\mathfrak{p}-\mu_p)e^{-\theta_p(s-t)} + \mu_p \\
\nu_{\mathfrak{p}}(t,s) &= \mathrm{Var}_{t,x}[\mathcal{P}(s)] = \sigma_p^2 \int_t^s e^{2\theta_p(u-s)}du = \frac{\sigma_p^2}{2\theta_p}(1-e^{2\theta_p(t-s)}) \\
m_{\mathfrak{i}}(t,\mathfrak{i},s) &= \mathbb{E}_{t,x}[\mathcal{I}(s)] = (\mathcal{I}(t)-\mu_i)e^{-\theta_i(s-t)} + \mu_i = (\mathfrak{i}-\mu_i)e^{-\theta_i(s-t)} + \mu_i \\
\nu_{\mathfrak{i}}(t,s) &= \mathrm{Var}_{t,x}[\mathcal{I}(s)] = \sigma_i^2 \int_t^s e^{2\theta_i(u-s)}du = \frac{\sigma_i^2}{2\theta_i}(1-e^{2\theta_i(t-s)})
\end{aligned}$$

For the price processes $P(t)$ and $I(t)$ we obtain:

$$\begin{aligned}
dP_t &= d\exp\{\mathcal{P}(t)\} = P_t\left(\left(\theta_p(\mu_p - \log(P_t)) + \frac{1}{2}\sigma_p^2\right)dt + \sigma_p dW_p(t)\right) \\
dI_t &= d\exp\{\mathcal{I}(t)\} = I_t\left(\left(\theta_i(\mu_i - \log(I_t)) + \frac{1}{2}\sigma_i^2\right)dt + \sigma_i dW_i(t)\right)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_{t,x}[P(s)] &= \mathbb{E}_{t,x}[e^{\mathcal{P}(s)}] = \\
&= \exp\left\{\mathbb{E}_{t,x}[\mathcal{P}(s)] + \frac{1}{2}\mathrm{Var}_{t,x}[\mathcal{P}(s)]\right\} = \\
&= \exp\left\{m_{\mathfrak{p}}(t,\mathfrak{p},s) + \frac{1}{2}\nu_{\mathfrak{p}}(t,s)\right\} = \\
&= \exp\left\{(\log(p_t)-\mu_p)e^{-\theta_p(s-t)} + \mu_p + \frac{\sigma_p^2}{4\theta_p}[1-e^{2\theta_p(t-s)}]\right\}
\end{aligned}$$

Finally, we calculate the conditional joint density $f_{\mathcal{P},\mathcal{I}}$ of the log-price random vector $(\mathcal{P},\mathcal{I})$ at time t given the realization $\mathfrak{p}$ and $\mathfrak{i}$

$$\begin{aligned}
g_{\mathcal{P}}(\mathfrak{p};t,x,s) &= \frac{\mathfrak{p} - m_{\mathfrak{p}}(t,\mathfrak{p},s)}{\nu_{\mathfrak{p}}(t,s)} \\
g_{\mathcal{I}}(\mathfrak{i};t,x,s) &= \frac{\mathfrak{i} - m_{\mathfrak{i}}(t,\mathfrak{i},s)}{\nu_{\mathfrak{i}}(t,s)} \\
f_{\mathcal{P},\mathcal{I}}(\mathfrak{p},\mathfrak{i};t,x,s) &= \frac{e^{\frac{-1}{2(1-\rho^2)}((g_{\mathcal{P}}(\mathfrak{p};t,x,s))^2 + (g_{\mathcal{I}}(\mathfrak{i};t,x,s))^2 - 2\rho g_{\mathcal{P}}(\mathfrak{p};t,x,s)g_{\mathcal{I}}(\mathfrak{i};t,x,s))}}{2\pi\nu_p(t,s)\nu_i(t,s)\sqrt{1-\rho^2}}
\end{aligned} \tag{39}$$

4.2 Finite difference alghoritm

Finite difference methods are numerical methods for approximating the solutions to differential equations, which use finite differences to approximate derivatives. In our case, we build a numerical scheme, using a finite difference method, for the HJB equation (34) to find out an approximation of the value functions V^n.

The first step to build such an algorithm is to bound and discretize all the intervals on which the arguments of V^n lie. This methology requires bounds also in the price dimensions, so we assume that we have chosen appropriate intervals $\mathfrak{P} = [p_{\min}, p_{\max}]$ and $\mathfrak{I} = [i_{\min}, i_{\max}]$ such that the processes P_t and I_t are unlikely to be outside that intervals. This can be a reasonable assumption if we use, for instance, processes such as the ones in Eq. (37-38) which exibits mean reversion: we can assume that $p_{\min}, i_{\min}$ are so small that the dominating behavior of the log-price process until the time of maturity T is to increase due to mean reversion, while $p_{\max}, i_{\max}$ are so large that the process is dominated by a decreasing behavior. Let us introduce the notation used.

$$\begin{array}{lll}
\mathfrak{T}_n = [T_{n-1}, T_n] & \to \quad T_{n-1} = t_1 < \ldots < t_\nu < \ldots < t_{N_t} = T_n & \delta_t = t_{\nu+1} - t_n \\
\mathfrak{P} = [p_{\min}, p_{\max}] & \to \quad p_{\min} = p_1 < \ldots < p_m < \ldots < p_{N_p} = p_{\max} & \delta_p = p_{m+1} - p_m \\
\mathfrak{I} = [i_{\min}, i_{\max}] & \to \quad i_{\min} = i_1 < \ldots < i_l < \ldots < i_{N_i} = i_{\max} & \delta_i = i_{m+1} - i_m \\
\mathfrak{Z} = [0, \bar{u}T] & \to \quad 0 = z_1 < \ldots < z_r < \ldots < z_{N_z} = \bar{u}T & \delta_z = z_{r+1} - z_r = \bar{u}
\end{array}$$

We notice that, having choosen such compact intervals $\mathfrak{P}$ and $\mathfrak{I}$, Assumption 1 are satisfied also for the price processes (37-38).

The covariance matrix $A(t,x) = (\Sigma\Sigma')(t,x)$ is given by:

$$A(t, X_t) = (A_{ij}(t, X_t))_{i,j\in\{1,\ldots,4\}} = \begin{pmatrix} \sigma_p^2(t, P_t) & \sigma_i(t, I_t)\sigma_p(t, P_t)\rho & 0 & 0 \\ \sigma_i(t, I_t)\sigma_p(t, P_t)\rho & \sigma_i^2(t, I_t) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Being $A_{ij}(t,x) \geqslant 0$ we can follow [14, Section 5.3.1 and 12.1] and use the following finite difference approximation. We use the notation $V_\nu^n(m, l, \hat{l}, r)$ for the approximation of $V^n(t_\nu, p_m, i_l, i_{\hat{l}}, z_r)$ and we suppress the superscript n when not needed for the sake of notation. The HJB equations becomes:

$$
\begin{aligned}
\frac{1}{\delta_t} V_\nu(m,l,\hat{l},r) = & \left(\frac{1}{\delta_t} - \frac{\sigma_p^2}{\delta_p} + \frac{\rho\sigma_p\sigma_i}{\delta_i} - \frac{\sigma_i^2}{\delta_i} - \frac{\mu_p}{\delta_p} - \frac{\mu_i}{\delta_i}\right) V_{\nu+1}(m,l,\hat{l},r) + \\
& + \left(\frac{\sigma_p^2}{2\delta_p} - \frac{\rho\sigma_p\sigma_i}{2\delta_i}\right)(V_{\nu+1}(m+1,l,\hat{l},r) + V_{\nu+1}(m-1,l,\hat{l},r)) + \\
& + \left(\frac{\sigma_i^2}{2\delta_i} - \frac{\rho\sigma_p\sigma_i}{2\delta_i}\right)(V_{\nu+1}(m,l+1,\hat{l},r) + V_{\nu+1}(m,l-1,\hat{l},r)) + \\
& + \frac{\rho\sigma_p\sigma_i}{2\delta_i\delta_p}(V_{\nu+1}(m+1,l+1,\hat{l},r) + V_{\nu+1}(m-1,l-1,\hat{l},r)) + \\
& + \frac{\mu_p}{\delta_p} V_{\nu+1}(m+1,l,\hat{l},r) + \\
& + \frac{\mu_i}{\delta_i} V_{\nu+1}(m,l+1,\hat{l},r) + \\
& + \frac{\rho\sigma_p\sigma_i}{2\delta_i}(V_{\nu+1}(m+1,l+1,\hat{l},r) + V_{\nu+1}(m-1,l-1,\hat{l},r)) + \\
& + e^{-r(t_n - T_{n-1})}\left(p - \hat{i} + \frac{[V_{\nu+1}(m,l,\hat{l},r+1) - V_{\nu+1}(m,l,\hat{l},r)]}{\delta_z}\right) \cdot \\
& \quad u^*\left(m,\hat{l},\left[\frac{V_{\nu+1}(m,l,\hat{l},r+1) - V_{\nu+1}(m,l,\hat{l},r)}{\delta_z}\right]\right)
\end{aligned}
\tag{40}
$$

Remark 5 An important feature of the Kushner scheme we presented in (40) is that the discretized HJB equation is iself the dynamic programming equation for a suitable defined stochastic control problem for Markov chains. This fact is used in [14] to prove the convergence of the discrete value function to $V^n(t,x)$. Another proof, which make use of the viscosity solution, can be found in [11, Chapter IX, Sections 4-5].

4.2.1 Boundary conditions

In order to implement the numerical scheme in (40), we need some boundary conditions. The key point of this subsection is the remark that the spot index price I, that is an average in past values, has a mean reversion whose speed should be significantly lower than the mean reversion of the spot price P.

Boundary conditions on $\mathfrak{P}$ and $\mathfrak{I}$. Regarding the boundary conditions on p and $\hat{i}$, the key idea is to use the mean-reversion behaviour of prices to determine how the holder will optimally use her optionality, i.e., determine the optimal control u_s^*.

When $p = p_{\max}$, being the mean reversion of the spot P higher than the mean reversion of the index I, we can assume that in the future this spread is likely to decrease. In this view, even if $p_{\max} - \hat{i} < 0$, the optimal operational behavior should be to use as much of the swing option as possible until $z \leqslant \overline{M}$. If we denote

$$\tau_1 = \tau_1(t,z) = \min\left\{t + \frac{\overline{M} - z}{\overline{u}}, T_n\right\}$$

then we can assume the boundary condition

$$\begin{aligned} &V^n(t, p_{\max}, i, \hat{i}, z) = \\ &= \overline{u}\mathbf{1}_{z<\bar{M}}\mathbb{E}_{t,x}\left[\int_t^{\tau_1} e^{-r(s-T_{n-1})}(P_s - \hat{i})ds\right] \\ &\quad + \mathbb{E}_{t,x}\left[e^{-r(T_n-T_{n-1})}V^{n+1}\left(T_n, P_{T_n}, I_{T_n}, \hat{I}_{T_n}, z + \overline{u}(\tau_1 - t)\mathbf{1}_{z<\bar{M}}\right)\right] \end{aligned} \tag{41}$$

When $p = p_{\min}$ the spread $p - \hat{i}$ is expected to increase. This implies that the optimal operational behavior when $p = p_{\min}$ should be to wait as long as possible before exercise. Then, by introducing

$$\tau_2 = \tau_2(t,z) = \max\left\{t, T - \frac{\underline{M} - z}{\overline{u}}\right\}$$

and using the convention $\int_a^b f(x)dx = 0$ if $a \geqslant b$, we can assume

$$\begin{aligned} &V^n(t, p_{\min}, i, \hat{i}, z) = \\ &= \overline{u}\mathbf{1}_{z<\bar{M}}\mathbb{E}_{t,x}\left[\int_{\tau_2}^{T_n} e^{-r(s-T_{n-1})}(P_s - \hat{i})ds\right] \\ &\quad + \mathbb{E}_{t,x}\left[e^{-r(T_n-T_{n-1})}V^{n+1}\left(T_n, P_{T_n}, I_{T_n}, \hat{I}_{T_n}, z + \mathbf{1}_{z<\bar{M}}\int_{\tau_2}^{T_n} \overline{u}ds\right)\right] \end{aligned} \tag{42}$$

We next calculate the stochastic integrals (41-42) defined by the boundary conditions on the truncated boundary. We have, for $a \leqslant b$

$$\begin{aligned} &\mathbb{E}_{t,x}\left[\int_a^b e^{-r(s-T_{n-1})}P_s ds\right] = \\ &= \int_a^b e^{-r(s-T_{n-1})}\mathbb{E}_{t,x}[P_s]ds = \\ &= \int_a^b \exp\left\{-r(s-T_{n-1}) + (\mathcal{P}(t) - \mu_p)e^{-\theta_p(s-t)} + \mu_p + \frac{\sigma_p^2}{4\theta_p}[1 - e^{2\theta_p(t-s)}]\right\} ds = \\ &= \exp\left\{\mu_p + \frac{\sigma_p^2}{4\theta_p} + rT_{n-1}\right\}\int_a^b \exp\left\{-rs + (\mathcal{P}(t) - \mu_p)e^{-\theta_p(s-t)} - \frac{\sigma_p^2}{4\theta_p}e^{2\theta_p(t-s)}\right]\right\} ds = \\ &= g(a, b, \mathcal{P}(t)) \end{aligned}$$

Let us define for any $a, b \in \mathbb{R}$ the function:

$$\begin{aligned}
E(a,b,p,\hat{i}) &= \mathbb{E}_{t,x}\left[\int_a^b e^{-r(s-T_{n-1})}(P_s - \hat{i})ds\right] \\
&= \mathbf{1}_{a\leqslant b}\left(g(a,b,\log(p)) - \hat{i}\int_a^b e^{-r(s-T_{n-1})}ds\right) \\
&= \mathbf{1}_{a\leqslant b}\left(g(a,b,\log(p)) + \frac{\hat{i}}{r}e^{rT_{n-1}}(e^{-rb} - e^{-ra})\right)
\end{aligned}$$

For the boundaries on $\mathfrak{I}$, as a first approximation, we take a linear interpolation, i.e. we set:

$$\begin{aligned}
V^n(t,p,i_{\min},\hat{i},z) &= V^n(t,p_{\min},i_{\min},\hat{i},z)+ \\
&+ \frac{V^n(t,p_{\max},i_{\min},\hat{i},z) - V^n(t,p_{\min},i_{\min},\hat{i},z)}{p_{\max}-p_{\min}}(p-p_{\min}) \\
V^n(t,p,i_{\max},\hat{i},z) &= V^n(t,p_{\min},i_{\max},\hat{i},z)+ \\
&+ \frac{V^n(t,p_{\max},i_{\min},\hat{i},z) - V^n(t,p_{\min},i_{\min},\hat{i},z)}{p_{\max}-p_{\min}}(p-p_{\min})
\end{aligned}$$

Boundary conditions on $\mathfrak{T}_n$. The terminal condition at time T_n is well known:

$$V^n(T_n,x) = e^{-r(T_n-T_{n-1})}V^{n+1}(T_n,x)$$

Boundary conditions on $\mathfrak{Z}$. The only boundary condition on the set $[0,\bar{u}T]$ needed for our numerical scheme is on the right boundary $\bar{u}T$. This quantity can be reached only at the end of the year and in this case the contract value is given by $\Psi(\bar{u}T)$. This implies

$$V^n(t,p,i,\hat{i},\bar{u}T) = \Psi(\bar{u}T)$$

Notice that, in principle, the feasible support for the variable z depends on the month n. In particular, we can restrict the solution of V^n to the interval $z \in [0,\bar{u}T_n]$, but we have no hint about the boundary condition $V^n(t,p,i,\hat{i},\bar{u}T_n)$. To avoid this problem, we let $\mathfrak{Z}$ be the same for every n.

Example 1 Let us give a first example for $n = 12$. We have, for the relevant functions needed:

$$\begin{aligned}
V^{12}(T_{12},p,i,\hat{i},z) &= e^{-r(T_{12}-T_{11})}\Psi(z) \\
V^{12}(t,p,i,\hat{i},\bar{u}T) &= e^{-r(T_{12}-T_{11})}\Psi(\bar{u}T) \\
V^{12}(t,p_{\max},i,\hat{i},z) &= \bar{u}\mathbf{1}_{z<\bar{M}}E(t,\tau_1,p_{\max},\hat{i})+ \\
&\quad + e^{-r(T_{12}-T_{11})}\Psi(z+\mathbf{1}_{z<\bar{M}}(\tau_1-t)\bar{u}) \\
V^{12}(t,p_{\min},i,\hat{i},z) &= \bar{u}\mathbf{1}_{z<\bar{M}}E(\tau_2,T_{12},p_{\min},\hat{i})+ \\
&\quad + e^{-r(T_{12}-T_{11})}\Psi(z+\mathbf{1}_{z<\bar{M}}\mathbf{1}_{\tau_2<T_{12}}(T_{12}-\tau_2)\bar{u})
\end{aligned}$$

Notice that, in this case, $V^{12}(t, p_{\max}, i, \hat{\imath}, z)$ and $V^{12}(t, p_{\max}, i, \hat{\imath}, z)$ do not depend on i. This is intuitive, because during the last month the knowledge of the *spot index* value give no extra information from the market.

We can now use the numerical scheme in (40) to find out an approximation of V^{12}.

Example 2 Once $V^{n+1}(t, p, i, \hat{\imath}, z)$ is known, the relevant boundary conditions on V^n reads

$$
\begin{aligned}
V^n(T_n, p, i, \hat{\imath}, z) &= e^{-r(T_n - T_{n-1})} V^{n+1}(T_n, p, i, \hat{\imath}, z) \\
V^n(t, p, i, \hat{\imath}, \overline{u}T) &= e^{-r(T_n - T_{n-1})} \Psi(\bar{u}T) \\
V^n(t, p_{\max}, i, \hat{\imath}, z) &= \bar{u}\mathbf{1}_{z<\bar{M}} E(t, \tau_1, p_{\max}, \hat{\imath}) + \\
&\quad + e^{-r(T_n - T_{n-1})} \mathbb{E}_{t,x}[V^{n+1}(T_n, P_{T_n}, I_{T_n}, \hat{I}_{T_n}, Z_{T_n})] = \\
&= \bar{u}\mathbf{1}_{z<\bar{M}} E(t, \tau_1, p_{\max}, \hat{\imath}) + \\
&\quad + e^{-r(T_n - T_{n-1})} \mathbb{E}_{t,x}[V^{n+1}(T_n, P_{T_n}, I_{T_n}, \hat{I}_{T_n}, z + \mathbf{1}_{z<\bar{M}}(\tau_1 - t)\bar{u})] \\
&= \bar{u}\mathbf{1}_{z<\bar{M}} E(t, \tau_1, p_{\max}, \hat{\imath}) + \\
&\quad + \int_{\mathbb{R}^2} V^{n+1}(T_n, e^x, e^y, e^y, z + \mathbf{1}_{z<\bar{M}}(\tau_1 - t)\bar{u}) f_{\mathcal{P},\mathcal{I}}(x, y; t, x, T_n) dx dy
\end{aligned} \tag{43}
$$

$$
\begin{aligned}
V^n(t, p_{\min}, i, \hat{\imath}, z) &= \bar{u}\mathbf{1}_{z<\bar{M}} E(\tau_2, T_n, p_{\min}, \hat{\imath}) + \\
&\quad + e^{-r(T_n - T_{n-1})} \mathbb{E}_{t,x}[V^{n+1}(T_n, P_{T_n}, I_{T_n}, \hat{I}_{T_n}, z + \mathbf{1}_{z<\bar{M}}\mathbf{1}_{\tau_2<T_n}(T_n - \tau_2)\bar{u})] = \\
&= \bar{u}\mathbf{1}_{z<\bar{M}} E(\tau_2, T_n, p_{\min}, \hat{\imath}) + \\
&\quad + \int_{\mathbb{R}^2} V^{n+1}(T_n, e^x, e^y, e^y, z + \mathbf{1}_{z<\bar{M}}\mathbf{1}_{\tau_2<T_n}(T_n - \tau_2)\bar{u}) f_{\mathcal{P},\mathcal{I}}(x, y; t, x, T_n) dx dy
\end{aligned} \tag{44}
$$

4.3 A Least Square Monte Carlo algorithm

The Least Square Monte Carlo (LSMC) approach was originally developed by Longstaff and Schwartz [17] for valuing American options. Today, it is widely used also in the energy field to evaluate structured products, see for instance [5],[6] and [8] for application of LSMC to Virtual Storage structured products and [24] for applications to Virtual Power Plant structured products. A summary of existing research on swing option valuation can be found in [1].

The Least Square Monte Carlo works with a discrete time version of the problem (11), which we present in the next Eq. (47). To reduce dimensionality, it does not take care about the spot index price I_t, and works only with the traded price $\hat{I}_t$ with monthly granularity, eventually stretched to daily granularity as we did in Formula (1), if the time step of the algorithm is daily. In practice, Least Square Monte Carlo is losing the information given every day about the knowledge of the *index spot price*. As a consequence, the algorithm does not need to distinguish value functions between months, as did before.

To avoid huge notation, in this section we set the risk free r to be 0.

Assumption 3 *In this section, the notation V^n stands for the value function calculated for path number n in a Monte Carlo environment, where $\boldsymbol{N}$ paths for stochastic dynamics have been simulated.*

The LSMC method is based on the intuition that conditional expectations in the dynamic programming pricing algorithm can be replaced by its orthogonal projection on some space generated by set of basis functions of the present state, obtained using Monte Carlo simulations and least-squares regressions to estimate numerically said ortjogonal projection. Let us introduce the key idea in whole generality, and then focus on the swing problem. Time interval $[0, T]$ is now discretized into a sequence $\{t_j\}_{j=0,\dots,N_T}$ with

$$0 = t_0 < t_2 < \ldots < t_{N_T} = T$$

and where $t_{j+1} - t_j$ represents, most of the times, one day. If X is the state process (underlying the general control problem) adapted to the filtration $\{\mathcal{F}_t\}_t$, given the realization at time t_j denoted by $X_{t_j} = x_j$, the key idea of the LSMC algorithm is to replace in the dynamic programmig equation in discrete time

$$\begin{aligned} V(t_j, x_j) &= \sup_{u_j}\{L(t_j, x_j, u_j) + \mathbb{E}[V(t_{j+1}, X_{j+1})|\mathcal{F}_{t_j}]\} \\ X_{j+1} &= f(x_j, u_j, W_{j+1}) \end{aligned}$$

the conditional expectation $\mathbb{E}[V(t_{j+1}, X_{j+1})|\mathcal{F}_{t_j}]$ with

$$\mathbb{E}[V(t_{j+1}, X_{j+1})|\mathcal{F}_{t_j}] = \sum_{\xi=1}^{+\infty} \alpha_\xi^{j+1} f_\xi(x_j, u_j) \tag{45}$$

where f_ξ are functions taken from a basis of a functional space (polynomials of degree ξ, Laguerre polynomials, radial basis functions, ...) and $\alpha_\xi^{j+1} \in \mathbb{R}$.

From a computational point of view, we can not work with infinite sums and so a first choice need to be done on the number of basis function we want to use. Let N_ξ be this number.

4.3.1 Least Square Monte Carlo for swing problem

Let us now focus on a swing problem. As said, here we consider only the spot gas price $P(t)$ and monthly index price $\hat{I}(t)$. The value function in discrete time for this problem is (see [10])

$$V(t_j, p_j, \hat{i}_j, z_j) = \sup_u \mathbb{E}_{t_j, x_j}\left[\sum_{k=j}^{N_T} (P_k - \hat{I}_k) u_k + \Psi(Z_T)\right] \tag{46}$$

Using the dynamic programming principle, which states that if a control is optimal on a whole sequence of periods than it has to be optimal on every single period, we obtain in discrete time:

$$
\begin{aligned}
V(t_j, p, \hat{i}, z) &= \sup_{u_j,\dots,u_{N_T}} \mathbb{E}_j \left[\sum_{k=j}^{N_T} (P_k - \hat{I}_k) u_k + \Psi(Z_{N_T}) \right] = \\
&= \sup_{u_j} \{(p - \hat{i}) u_j + \mathbb{E}_j[V(t_{j+1}, P_{j+1}, \hat{I}_{j+1}, Z_{j+1})]\} = \\
&= \sup_{u_j} \{(p - \hat{i}) u_j + \mathbb{E}_j[V(t_{j+1}, P_{j+1}, \hat{I}_{j+1}, z + u_j)]\}
\end{aligned} \tag{47}
$$

where the notation $\mathbb{E}_j$ stands for $\mathbb{E}_{t_j,x_j}[\cdot]$.

The Dynamic Programming Principle in (47) and the least square regression are now used as follows by the LSMC algorithm. After having simulated $\boldsymbol{N}$ paths for the price dynamics $\{p^n(t_j), \hat{i}^n(t_j)\}_{j=1,\dots,N_T}^{i=1,\dots,\boldsymbol{N}}$, which we will denote with p_j^n and $\hat{i}_j^n$, the algorithm goes backward in time.

Algorithm 1

For every $t = T, T-1, \dots, 1$:

→ if $t_j = T$ (i.e. $j = N_T$), set for every path n

$$V^n(T, p_{N_T}^n, \hat{i}_{N_T}^n, z) = \Psi(z), \quad \forall i = 1, \dots, \boldsymbol{N}$$

→ if $t_j < T$ find out the optimal control $\tilde{u}_j^n$ and the value function V^n for every path $n = 1, \dots, \boldsymbol{N}$ with the maximization

$$V^n(t_j, p_j^n, \hat{i}_j^n, z) = \sup_u \left\{ (p_j^n - \hat{i}_j^n) u + \sum_{\xi=1}^{N_\xi} \alpha_\xi^{j+1} f_\xi(p_j^n, \hat{i}_j^n, z + u) \right\}$$

→ if $t_j > 0$, calculate the coefficients α_ξ^j by minimizing the norm

$$\min_{\{\alpha_\xi^n\}_{\xi=1,\dots,N_\xi} \subset \mathbb{R}} \sum_{n=1}^{\boldsymbol{N}} \left\| V^n(t_j, p_j^n, \hat{i}_j^n, z) - \sum_{\xi=1}^{N_\xi} \alpha_\xi^j f_\xi(p_{j-1}^n, \hat{i}_{j-1}^n, z) \right\| \tag{48}$$

→ if $t_j = 0$ then $V^1(0, p_0^1, \hat{i}_0^1, 0)$ is the contract value

While the LSMC algorithm is very flexible, it may, on the other hand, be influenced by many user's choices which are capable of influencing the pricing procedure. For instance, choices regarding the type and the number N_ξ of basis functions as well as the number $\boldsymbol{N}$ of Monte Carlo simulations used. These choices can be critical: as shown in [19], while for some type of derivatives (such as the American put) the LSMC approach is very robust, for more complex derivatives the number and the type of basis functions can slightly affect option prices.

4.3.2 Radial Basis Functions Approximation

In general, from Formula (48) it is evident that f_ξ shoud be of the form $f_\xi : \mathbb{R}^3 \to \mathbb{R}$ for a general swing problem. A 3-dimensional fitting may be computationally challenging. In order to speed up the algorithm, sometimes we can simplify this point. We notice that, in absence of any particular constraints linked to single price path, the spread $P_j - \hat{I}_j$ is, in practice, the key quantity on which decisions are taken. We can define a new random variable

$$S_j = P_j - \hat{I}_j$$

and write the problem (47) as:

$$V(t_j, s, z) = \sup_{u_j} \{ s u_j + \mathbb{E}_j [V(t_{j+1}, S_{j+1}, z + u_j)] \} \tag{49}$$

From a numerical point of view, we have reduced our state space to a 2-dimensional one, and now for every t_j a surface has to be fitted.

To avoid dimension problems, a good general idea can be to adopt numerical methods based on *radial basis functions*. They are well known for their dimensional blindness, which potentially allows to use them for solving very high dimensional problems. This dimensional blindness comes directly from the definition of RBF.

A new approach based on radial basis function approximation applied to Least Square Monte Carlo problems has been recently proposed in [7]. This approach may be very promising in solving the curse of dimensionality arising from the pricing of energy structured products. In this subsection we apply exactly the same ideas of [7] to swing contract. Our final aim is to compare the finite difference scheme presented in Section 4.2 with a more practitioner algorithm as the LSMC ones, in order to deduce if they give similar results or if one perform better than the other in terms of computational time and accuracy of the solution.

For the convenience of the readers, to fix notation, definitions, and general ideas, we briefly re-propose in what follows the problem of interpolation and approximation with RBF exactly as done in [7]. Next we apply the RBF approximation to our swing problem and present the algorithm in two dimensions. This algorithm is the basic framework of more sophisticated algorithms, where additional state dimensions may included, or different strategies to sampling centres can be used, as proposed in [7]. it gives the idea of the features of radial basis approach. Finally, next subsection 4.3.3 presents a particular case when the two dimensional regression can be replaced by a simple one-dimensional regression jointly with the use of a proper quantization of the cumulated quantity space.

Definition 5 *A function $\Phi_c : \mathbb{R}^n \to \mathbb{R}$ is called* ***radial*** *provided there exists a univariate real valued function $\phi : [0, +\infty) \to \mathbb{R}$ whose value depends only on the distance from some point c, called centre, so that:*

$$\Phi_c(x) = \phi(\|x - c\|)$$

The norm is usually the Euclidean norm, although other distance functions are also possible.

There are a lot of different choices for the radial functions ϕ. Some are globally supported like the Gaussian and the generalized multiquadratic while others are compactly supported, such as the family of Wendland functions. Throughout this paper we use the same RBF proposed in [7], i.e. a specific kind of this type of Wendland functions defined in $\mathbb{R}^3$, which have smoothness of order 2 and can be used in problems up to and including three dimensions. The functional forms of such radial function is

$$\phi(r) = ((1 - \varepsilon r)^+)^4 (4\varepsilon r + 1) \tag{50}$$

This choice is given, first of all, by the property of those functions, which are sufficiently smooth for our problem, but not too smooth. As noted in [7], with a higher order of smoothness we would risk to over-fit the problem, and with globally supported functions, like Gaussians (which are infinitely smooth) the regression matrices would have very large conditioning numbers, and would be harder to invert. Finally, those functions have given good results for storage structured products: it is straightforward to use them as first benchmark also for swing.

Interpolation and approximation with RBF: introduction. The problem of RBF *interpolation and approximation* is posed as follows. Let $\{c_j\}_{j=1,\dots,M} \subset \mathbb{R}^n$ a chosen set of centers for our basis function ϕ and let $f : \mathbb{R}^n \supseteq \Omega \to \mathbb{R}$ be the function we want to interpolate/approximate. Let us suppose we have measured the sequence $\{y_i\}_{i=1,\dots,N} \subset \mathbb{R}$ whose values are realizations of f at a set of N distinct locations $\{x_i\}_{i=1,\dots,N} \subset \Omega$

$$y_i = f(x_i) \quad \forall i = 1, \dots, N \tag{51}$$

We want to find a function $s_f : \Omega \to \mathbb{R}$ who has a RBF expansion such as

$$s_f(x) = \sum_{j=1}^{M} \alpha_j \Phi_{c_j}(x) = \sum_{j=1}^{M} \alpha_j \phi(\|x - c_j\|) \qquad x \in \mathbb{R}^n \tag{52}$$

where $\{\alpha_j\}_{j=1,\dots,M} \subset \mathbb{R}$ are the regressor coefficients and ϕ is a RBF applied to the center points c_j and locations x_i. We force the conditions

$$s_f(x_i) = f(x_i) \quad \forall i = 1, \dots, N$$

which can be rewritten using Equations (51) and (52) as

$$y_i = \sum_{j=1}^{M} \alpha_j \phi(\|x_i - c_j\|) \quad \forall i = 1, \dots, N \tag{53}$$

Formula (53) is a system of linear equalities with N equations (one for every measure (x_i, y_i)) and M unknowns (the coefficients α_j) that can be expressed in matrix notation

$$Y = \mathbf{\Phi}\alpha \tag{54}$$

where $Y \in \mathbb{R}^N$ is the vector of observations $Y = (y_1, \ldots, y_N)^T$, $\alpha \in \mathbb{R}^M$ is the vector of interpolation coefficients $\alpha = (\alpha_1, \ldots, \alpha_M)^T$ and $\mathbf{\Phi} \in \mathbb{R}^{N \times M}$ is a matrix resulting from applying the RBF ϕ to every entry of the distance matrix D, whose entries are the Euclidean norm of the data sites against the center points

$$D = \begin{pmatrix} \|x_1 - c_1\| & \cdots & \|x_1 - c_j\| & \cdots & \|x_1 - c_M\| \\ \vdots & & \vdots & & \vdots \\ \|x_i - c_1\| & \cdots & \|x_i - c_j\| & \cdots & \|x_i - c_M\| \\ \vdots & & \vdots & & \vdots \\ \|x_N - c_1\| & \cdots & \|x_N - c_j\| & \cdots & \|x_N - c_M\| \end{pmatrix}$$

It is clear that, whenever $N = M$, the matrix $\mathbf{\Phi}$ is square and we can perform an interpolation while when $N > M$ the system in Eq. (54) is over-determined. Being D generated by a set of distinct center points, it is always of full column rank M. If (54) is over-determined, we can solve it by means of linear least square minimization, i.e. find the solution α^* such that

$$\alpha^* = \arg\min_{\alpha} \|Y - \mathbf{\Phi}\alpha\| \tag{55}$$

From linear algebra, a possible solution to (55) can be found using the Moore-Penrose inverse $\mathbf{\Phi}^+$ of the matrix $\mathbf{\Phi}$ (see [21]) which leads to

$$\alpha^* = \mathbf{\Phi}^+ Y$$

Application to the swing problem. Coming back to our swing problem, the RBF approximation can be used in Algorithm 4.3.1 to find out the coefficients in Eq. (48). The algorithm rewrites as

Algorithm 2

Let $[0, \bar{u}T]$ be discretized into a sequence $\{z_k\}_{k=1,\ldots,N_z}$. Let $\{c_\xi\}_{\xi=1,\ldots,M} \subset \mathbb{R}^2$ be the sequence of centers we have chosen[1] for the Radial Basis Functions.

For every $j = N_T, N_T - 1, \ldots, 1$:

→ if $t_j = T$ set $\forall n = 1, \ldots, \boldsymbol{N}$ and $\forall k = 1, \ldots, N_z$

$$V^n(T, s^n_{N_T}, z_k) = \Psi(z_k)$$

→ if $t_j < T$ find out the optimal control $\tilde{u}^n_j$ and the value function V^n for every

[1] Notice that we need centers both in the spread dimension as well as in the cumulated quantity dimension, i.e. every center c_ξ takes the form $c_\xi = (c^1_\xi, c^2_\xi)^T$. A first choice for c^1 can be chosen using the simulated path as an equispaced grid with M^1 points in the interval $[\min_{n,j}\{p^n_j\}, \max_{n,j}\{p^n_j\}]$ while for c^2 we can choose an equispaced grid of M^2 points in $[0, \bar{u}T]$, resulting in a total of $M = M^1 \cdot M^2$ centers. Other choices are possible, in particular for the cumulated quantity. As an example, in [7], a non equispaced grid with higher density of points on the boundaries is used. Also time-dependent centers can be used, at the expense of an increased effort in programming.

path n and for every[2] $k = 1, \ldots, N_z$ with a numerical maximization

$$V^n(t_j, s_j^n, z_k) = \max_u \left\{ s_j^n u + \sum_{\xi=1}^{N_\xi} \alpha_{j+1}^\xi \phi \left(\left\| \begin{pmatrix} s_j^n \\ z_k + u \end{pmatrix} - c_\xi \right\| \right) \right\} \tag{56}$$

$\rightarrow$ if $t_j > 0$ define the vector Y_j and the matrix $\mathbf{\Phi}_{j-1}$ as

$$\mathbf{\Phi}_{j-1} = \begin{pmatrix} \left\| \begin{pmatrix} s_{j-1}^1 \\ z_1 \end{pmatrix} - c_1 \right\| & \cdots & \left\| \begin{pmatrix} s_{j-1}^1 \\ z_1 \end{pmatrix} - c_M \right\| \\ \vdots & & \vdots \\ \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_1 \end{pmatrix} - c_1 \right\| & \cdots & \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_1 \end{pmatrix} - c_M \right\| \\ \left\| \begin{pmatrix} s_{j-1}^1 \\ z_2 \end{pmatrix} - c_1 \right\| & & \left\| \begin{pmatrix} s_{j-1}^1 \\ z_2 \end{pmatrix} - c_M \right\| \\ \vdots & & \vdots \\ \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_2 \end{pmatrix} - c_1 \right\| & \cdots & \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_2 \end{pmatrix} - c_M \right\| \\ \vdots & & \vdots \\ \left\| \begin{pmatrix} s_{j-1}^n \\ z_k \end{pmatrix} - c_1 \right\| & & \left\| \begin{pmatrix} s_{j-1}^n \\ z_k \end{pmatrix} - c_M \right\| \\ \vdots & & \vdots \\ \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_{N_k} \end{pmatrix} - c_1 \right\| & & \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_{N_k} \end{pmatrix} - c_M \right\| \end{pmatrix} \tag{57}$$

$$Y_j = \begin{pmatrix} V^1(t_j, s_j^1, z_1) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^{\boldsymbol{N}}, z_1) \\ V^1(t_j, s_j^1, z_2) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^{\boldsymbol{N}}, z_2) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^n, z_k) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^{\boldsymbol{N}}, z_{N_k}) \end{pmatrix} \tag{58}$$

and calculate the regression coefficients $\alpha_j = (\alpha_j^1, \ldots, \alpha_j^M)^T$ by solving the over-determined system

$$Y_j = \mathbf{\Phi}_{j-1} \alpha_j \tag{59}$$

[2] We may restrict the calculation only to the values z_k feasible at time t_j, i.e. $z_k \leqslant \bar{u} t_j$

For instance, you can use the Moore-Penrose pseudoinverse of $\mathbf{\Phi}_{j-1}$ and compute

$$\alpha_j = \arg\min_{\alpha} \|Y_j - \mathbf{\Phi}_{j-1}\alpha_j\| = \mathbf{\Phi}_{j-1}^{+} Y_j \tag{60}$$

$\rightarrow$ if $t_j = 0$ the contract value is $V^1(0, p_0^1, \hat{\imath}_0^1, 0)$

4.3.3 Reduction to one dimension: cumulated quantity discretization

Even though the system of equations (59) should not require a long solution time, we have to notice that Algorithm 4.3.2 requires $N_T \cdot \boldsymbol{N} \cdot N_z$ numerical maximizations (non-linear most of the times, and sometimes integer), coming out from Formula (56). They may require a not negligible amount of time. With some stronger assumptions, or by means of some approximation of our problem, the maximization in (56) can be avoided. The key result is the following.

Theorem 10 *Let us consider a general swing problem in discrete time defined on the interval* $[0, T]$ *and with the constraints*

$$u_j \in [\underline{u}, \bar{u}]$$
$$z_T \in \left[\underline{M}, \overline{M}\right]$$

If the quantity

$$\mathcal{K} = \frac{\overline{M} - \underline{M}}{\bar{u} - \underline{u}} \tag{61}$$

is an integer number, then there exists an optimal bang-bang Markovian control u_j^**, i.e. for all* $j = 1, \dots, N_T$ *we have* $u_j^* = \overline{u}$ *or* $u_j^* = \underline{u}$*.*

Proof See [3]. □

Thanks to Theorem 10, when assumption in Eq. (61) is satisfied, we can focus our attention only to bang-bang optimal controls of the form $u_j^* \in \{0, \bar{u}\}$. In other words, we can discretize in a suitable way the interval $[0, \bar{u}T]$ on which z_j lies. This leads to a binomial tree for the cumulated quantity because the optimal values for z have the form:

$$z = a\underline{u} + b\overline{u} \qquad a, b \in \mathbb{N}$$

Let us suppose we have such tree, i.e. we have a suitable sequence $\{z_k\}_k$ of values for the cumulated quantity. For instance, if $\underline{u} = 0$, at a first glance we can define $z_k = k\bar{u}$. Then the maximization in (56) becomes a maximization between only two possible values, and the two dimensional regression falls into a one dimensional regression. Algorithm 4.3.2 changes in this way. Having lost a dimension, now the centers $\{c_\xi\}_{\xi=1,\dots,M}$ of our RBF lie in $\mathbb{R}$ and the vector Y_j in (58) and the matrix $\mathbf{\Phi}_j$ in (57) read as

$$\mathbf{\Phi}_j = \begin{pmatrix} \mathbf{\Phi}_j^1 \\ \vdots \\ \mathbf{\Phi}_j^n \\ \vdots \\ \mathbf{\Phi}_j^{\boldsymbol{N}} \end{pmatrix} = \begin{pmatrix} |s_j^1 - c_1| & \cdots & |s_j^1 - c_M| \\ \vdots & |s_j^n - c_\xi| & \vdots \\ |s_j^{\boldsymbol{N}} - c_1| & \cdots & |s_j^{\boldsymbol{N}} - c_M| \end{pmatrix}$$

$$Y_j^k = \begin{pmatrix} V^1(t_j, s_j^1, z_k) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^{\boldsymbol{N}}, z_k) \end{pmatrix}$$

where $\mathbf{\Phi}_j^n$ stands for n-th row vector of matrix $\mathbf{\Phi}_j$. Notice that $\mathbf{\Phi}_j$ does not depend on k, while Y_j does it. Also the coefficients α_j now depend on k, and for them we use the notation α_j^k and compute them in the same way of Formula (60)

$$\alpha_j^k = \arg\min_{\alpha} \|Y_j^k - \mathbf{\Phi}_{j-1}\alpha_j^k\| = \mathbf{\Phi}_{j-1}^+ Y_j^k$$

Formula (56) can now be rewritten. Being $\mathbf{\Phi}_j$ independent from k, we can calculate it only one time for every time step: at time t_{j+1} we compute $\mathbf{\Phi}_j$, and then we re-use its rows at time t_j being

$$\begin{aligned} V^n(t_j, s_j^n, z_k) &= \max\{s_j^n \bar{u} + \mathbb{E}_j[V(t_{j+1}, S_{j+1}, z_{k+1})], \mathbb{E}_j[V(t_{j+1}, S_{j+1}, z_k)]\} \\ &= \max_u\{s_j^n \bar{u} + \mathbf{\Phi}_j^n \alpha_{j+1}^{k+1}, \mathbf{\Phi}_j^n \alpha_{j+1}^k\} \end{aligned}$$

4.4 Comparison between algorithms

Work in progress...

References

[1] O. Lindqvist A. Løland. Valuation of commodity-based swing options: a survey. *Note SAMBA/38/80*, 2008.

[2] F. Aasche, P. Osmundsen, and R. Tveterås. European market integration for gas? Volume flexibility and political risk. *Energy Economics*, 24(3):249–265, 2002.

[3] O. A. Bardou, S. Bouthemy, and G. Pagès. When are swing options bang-bang and how to use it. *ArXiv e-prints*, may 2007.

[4] F.E. Benth, J. Lempa, and T.K. Nilssen. On optimal exercise of swing options in electricity markets. *Journal of energy markets*, 4(4):3–28, 2011.

[5] A. Boogert and C. de Jong. Gas storage valuation using a Monte Carlo method. Birkbeck Working Papers in Economics and Finance 0704, Birkbeck, Department of Economics, Mathematics and Statistics, 2007.

[6] A. Boogert and C. de Jong. Gas storage valuation using a multifactor price process. *Journal of Energy Market*, 4(4), 2011.

[7] A. Boogert and D. Mazieres. A radial basis function approach to gas storage valuation. *Submitted*, 2011.

[8] R. Carmona and M. Ludkovski. Valuation of energy storage: an optimal switching approach. *Quantitative Finance*, 10(4):359–374, 2010.

[9] M.G. Crandall and P.L. Lions. Viscosity solution of Hamilton-Jacobi equations. *Trans. A.M.S.*, 277:1–42, 1984.

[10] E. Edoli, S. Fiorenzani, S. Ravelli, and T. Vargiolu. Modeling and valuing make-up clauses in gas swing contracts. *Energy Economics*, submitted, 2012.

[11] W. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions.* Springer, second edition, 2006.

[12] G.Folland. *Real Analysis: modern techniques and their applications.* Wiley, 1984.

[13] J. C. Hull. *Options, Futures, and Other Derivatives.* Prentice Hall, seventh edition, 2006.

[14] H.J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time.* Springer, second edition, 2001.

[15] L.C.Evans. *Partial differential equations.* Berkeley Mathematics Lecture Notes, 1994.

[16] F.Da Lio and O. Ley. Uniqueness results for second order Bellman-Isaacs equations under quadratic growth assumptions and applications. *SIAM Journal of Control Optimization*, 45(1):74–106, 2006.

[17] F. A. Longstaff and E. S. Schwartz. Valuing american options by simulation: a simple least-squares approach. *Review of Financial Studies*, 14(1):113–147, 2001.

[18] A. Lunardi. *Analytic Semigroups and Optimal Regularity in Parabolic Problems.* Springer, 2003.

[19] M. Moreno and J. R. Navas. On the robustness of least-squares Monte Carlo (LSM) for pricing American derivatives. Economics Working Papers 543, Department of Economics and Business, Universitat Pompeu Fabra, Apr 2001.

[20] E. I. Ronn P. Jaillet and S. Tompaidis. Valuation of commodity-based swing options. *Manage. Sci.*, 50(7):909–921, July 2004.

[21] R. Penrose. On best approximate solution of linear matrix equations. *Proceedings of the Cambridge Philosophical Society*, pages 17–19, 1956.

[22] Eduardo S. Schwartz and James E. Smith. Short-term variations and long-term dynamics in commodity prices. *Manage. Sci.*, 46(7):893–911, July 2000.

[23] T.Björk. *Arbitrage Theory in Continuous Time.* Oxford Finance, 2004.

[24] C. Tseng and G. Barz. Short-term generation asset valuation: a real options approach. *Operations Research*, 50:297–310, 2002.