UNIVERSITÀ DEGLI STUDI DI PADOVA

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MATEMATICA
SCUOLA DI DOTTORATO IN SCIENZE MATEMATICHE
Indirizzo di Matematica Computazionale - XXV ciclo

Pricing of gas swing contracts with indexed strike: a viscosity solution approach with applications

Supervisore:
Ch.mo Prof. Tiziano Vargiolu

Direttore della Scuola:
Ch.mo Prof. Paolo Dai Pra

Dottorando:
Enrico Edoli

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI MATEMATICA
SCUOLA DI DOTTORATO IN SCIENZE MATEMATICHE
Indirizzo di Matematica Computazionale - XXV ciclo

Pricing of gas swing contracts with indexed strike: a viscosity solution approach with applications

Supervisore:
Ch.mo Prof. Tiziano Vargiolu .....

Direttore della Scuola:
Ch.mo Prof. Paolo Dai Pra .....

Dottorando:
Enrico Edoli .....

To Barbara

CONTENTS

ABSTRACT

Swing contracts are structured products mostly traded on energy and gas markets, tailor-made to handle simultaneously price and volume risk arising from the modern liberalized markets. This thesis deals with some specific swing contracts relevant for the gas market.

In gas market, swing contracts are also known as *take or pay*. They are long-term supply contracts which allow flexibility of delivery: the holder of such a contract has multiple exercise rights and can decide the amount exercised as well, hedging the volume risk caused by a frequent demand fluctuation which in practice is impossible to foresee in the long period. Moreover, such type of contracts can be also seen as a strip of spread options on gas market spot price and the contractual price (called strike): in this view, they can be used to hedge the price fluctuation risk. On the other hand, even if the holder can exercise the option with a volume control, such control has however to satisfy some upper and lower limits at all times as well as a total volume, so the given flexibility need to be optimised, i.e. one must know the optimal execution of this flexibility.

Today, the correct valuation of these type of contracts is important both for trading purposes as well as for portfolio optimization. In fact, after the recent liberalization, the price of such contracts is negotiated between agents and no more set by regulators. On the other hand, the embedded flexibilities may be used not only to manage demand fluctuation, but also to make profit against local market price.

In this thesis we model, in a continuous time framework, a gas swing contract in the spirit of [6], 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 that we reduce to the usual continuous time sit-

uation by adding another state variable, corresponding to an index rolled-over in continuous time. The price of a swing contract is then equal to the value function of one sequence of Markov control problems, each one corresponding to a period between two consecutive changing dates of the index. After that, we prove that the value function of the corresponding control problem is the unique viscosity solution of the resulting Hamilton-Jacobi-Bellman (HJB) equation, and that the value function is smooth enough to ensure the existence of an optimal strategy, that we find out. Briefly spoken, this is the content of the first two chapters of this thesis where, after having framed the valuation problem as a stochastic optimal control problem, an introductory part on viscosity solutions is then applied to the particular problem at hand. This entails in some new results in the theory of viscosity solutions for parabolic nonlinear equations stemming from these swing options.

After having found the HJB equations, which become nonlinear partial differential equations, the problem is then to solve them. To do this, in the third chapter we present a finite difference (FD) method for solving the HJB equations numerically. We derive explicitly the boundary conditions for a particular model, when both gas and strike price are supposed to be in the class of the one-factor Schwartz-Smith dynamics. More in detail, we suppose that the log-prices follow Ornstein-Uhlenbeck processes driven by two correlated Brownian motions. Also, the third chapter presents the popular Least Square Monte Carlo (LSMC) algorithm, originally developed by Longstaff-Schwartz for valuing American options and here extended to the present problem. This algorithm is based on the backward solution of the discrete-time version of the control problem. It regresses the continuation value in the Bellman equation to the current available information, obtained using Monte Carlo simulations. Two critical steps using the LSMC algorithm are the choice of both type and number of the basis functions used in the regression; we approach this problem adapting to our case the radial basis function approximation introduced in [11] for storage structured products, which seems very appropriate and interesting for such applications. Several comparison examples are then numerically analysed, in particular the effect that the number of basis functions and the number of simulated paths have on the solution of LSMC as well as the efficiency of the FD method. Some conclusion on the comparison between the two algorithms ends the chapter.

The last chapter is from the published paper [17] and deals with the so-called make-up clauses, which extend the swing option previously studied by allowing to the holder of the contract more flexibility among years. From a technical point of view, a swing contract with an embedded make-up clause can not be any more split into yearly contracts, but one must consider the whole contract at once, typically lasting over several years (usually from 3 to 5). To approach the complexity of such problem, an-

other numerical method popular among practitioners is introduced for the purpose of pricing, namely *lattice of trees*. After having presented the algorithm and analysed its computational cost, the fourth chapter ends with many numerical examples testing for the swing option price's dependency on various crucial parameters.

SOMMARIO

I contratti swing nei mercati dell'energia e del gas sono prodotti strutturati creati su misura per gli operatori al fine di gestire contemporaneamente il rischio derivante dalle variazioni del prezzo di mercato e dall'incertezza volumetrica dovuta alla continua e imprevedibile fluttuazione della domanda. La presente tesi si occupa di un particolare tipo di contratto swing, frequentemente scambiato tra grossi operatori del mercato del gas naturale.

I contratti swing nei mercati del gas sono noti anche come *take-or-pay* e sono contratti di fornitura a lungo termine che permettono flessibilità nel ritiro del gas: l'acquirente possiede infatti la possibilità (ma non l'obbligo) di decidere sia *quando* sia *quanto* gas ritirare, essendo tuttavia obbligato a soddisfare dei vincoli di quantità minima e massima sia su ciascun periodo di ritiro (solitamente il giorno) sia complessivamente sull'anno. Da un lato, tale flessibilità volumetrica ben si adatta a soddisfare una domanda altalenante e imprevedibile. Dall'altro lato, un contratto swing può semplicisticamente essere visto come una strip di opzioni sullo spread tra il prezzo del gas di mercato e il prezzo contrattuale di ritiro: in quest'ottica, esso diventa un ottimo strumento di gestione del rischio derivante dalle fluttuazioni di prezzo nel breve periodo, permettendo di non esercitare, o di esercitare il minimo possibile, nei momenti avversi.

La corretta valutazione di simili contratti è oggi di grande importanza sia per ragioni di trading, essendo il prezzo di tali opzioni contrattato direttamente tra i players del mercato e non più imposto come durante il regime regolamentato, sia per ragioni di ottimizzazione di portafoglio, poiché le flessibilità volumetriche offerte possono potenzialmente essere usate anche per generare puri profitti.

Partendo da quanto esposto in [6] per i mercati energy, in questa tesi si descrive e risolve, a tempo continuo, il problema del pricing di un con-

tratto swing tipico dei mercati gas, in cui sia il prezzo di mercato quanto il prezzo contrattuale (prezzo strike, o indice) sono aleatori. In pratica, il prezzo strike viene aggiornato mensilmente, mentre il prezzo di mercato cambia con granularità almeno giornaliera: tecnicamente questo origina un problema di controllo ottimo stocastico in cui una variabile è discreta, mentre tutte le altre sono continue. Per superare questa difficoltà, e ridurre il problema ad uno classico a solo tempo continuo, si introduce una variabile di stato ad hoc, corrispondente alla dinamica a tempo continuo dell'indice. Tale variabile continua verrà successivamente campionata ad opportuni intervalli di tempo per ottenere la reale successione discreta di strikes contrattuali. Dopodiché, si dimostra che il prezzo del contratto è dato da una serie di problemi di controllo ottimo, ciascuno dei quali viene risolto all'interno di un periodo tra due cambi consecutivi dello strike. Si dimostra quindi che le funzioni valore di tali problemi di controllo ottimo sono l'unica soluzione di viscosità delle equazioni di Hamilton-Jacobi-Bellman (HJB) associate a ciascun problema, e che tali soluzioni sono sufficientemente regolari da garantire l'esistenza del controllo ottimo. Sommariamente, questo è il contenuto dei primi due capitoli, nei quali, dopo aver richiamato la nozione e i principali risultati noti sulle soluzioni di viscosità per i problemi di controllo ottimo, la teoria classica viene estesa ed applicata al problema in oggetto, portando ad alcuni nuovi risultati per le soluzioni di viscosità di equazioni paraboliche non lineari.

Il terzo capitolo è dedicato ai metodi numerici. Dopo aver ricavato l'equazione HJB il problema si sposta alla soluzione della stessa. A tal fine, si introduce uno schema di soluzione numerica basato sulle differenze finite. Tale algoritmo necessita di condizioni al contorno sui domini di soluzione, che vengono ricavate analiticamente nell'ipotesi in cui la dinamica dei prezzi di gas e indice segua un caso particolare del modello di Schwartz-Smith ad un fattore, cioè quando si suppone che il logaritmo dei prezzi sia un processo di Ornstein-Uhlenbeck guidato da due moti Browniani correlati. Uno studio empirico, attraverso un caso numerico, sulla stabilità dell'algoritmo FD al variare della discretizzazione del dominio temporale e spaziale completa la parte analitica. Al fine di confrontare i risultati ottenuti con la *best practice* in uso nelle aziende, il popolare algoritmo noto come *Least Square Monte Carlo*, originariamente sviluppato da Longstaff e Schwartz per valutare opzioni di tipo americano, viene adattato al problema del pricing di contratti swing. Questo algoritmo, molto usato tra i practitioners, risolve un'approssimazione a tempo discreto del problema originale usando la ricorsione all'indietro. Ad ogni iterazione dell'algoritmo, il valore di continuazione nell'equazione di Bellman a tempo discreto viene regredito sull'informazione presente a quell'istante, ottenuta tramite delle simulazioni Monte Carlo. Due punti in questo tipo di algoritmo risultano essere particolarmente critici: la scelta del tipo e del numero di funzioni di base usate nella regressione, nonchè il numero di

simulazioni Monte Carlo adottate. Per quanto riguarda la scelta del tipo di funzioni, seguendo e adattando al caso presente quanto sviluppato in [11], viene introdotto un metodo di regressione basato su funzioni radiali di base, che sembra ben adattarsi a problemi di pricing di contratti strutturati. Riguardo invece l'effetto della scelta del numero di basi e del numero di simulazioni usate nell'algoritmo LSMC, viene presentato uno studio empirico attraverso casi numerici.

L'ultimo capitolo riprende quanto già pubblicato in [17] ed estende la valutazione di contratti swing in cui è presente una clausola chiamata di *make-up*, che in pratica permette più flessibilità nel ritiro del gas abbassando il livello minimo di ritiro di un certo anno, e forzando in uno o più anni successivi il richiamo del gas non preso. Tecnicamente, quando tale clausola è presente, non è più possibile separare il problema di valutazione sugli anni, ma è necessario considerare l'intero intervallo temporale su cui è scritta la clausola di make-up (solitamente da 3 a 5 anni). Per affrontare la complessità del pricing di un tale contratto nel capitolo si introduce un algoritmo basato su alberi, noto come *lattice of trees*. Dopo un'accurata descrizione di tale algoritmo, anche da un punto di vista di complessità computazionale, il quarto capitolo termina con un'applicazione reale dell'algoritmo di pricing volto ad esaminare l'impatto di vari fattori di mercato e parametri contrattuali sul prezzo di un ipotetico contratto swing.

ACKNOWLEDGMENTS

I would like to express my gratitude to the following people for their conceptual and material help.

to *Stefano Fiorenzani* for having encouraged me to undertake a PhD while working, but especially for the many teachings and opportunities that he gave me during the last years.

to *Tiziano Vargiolu*, my supervisor, for its strategical support during the writing process of this thesis.

to *Enzo Fanone* for his advice on conceptual and numerical matters, and for his many suggestions on literature.

to *my family* for having supported and helped my wandering life during the last three years.

I'm also grateful to my referee for his challenging comments, which helped me to improve some parts of this Thesis.

Enrico Edoli

Chapter 1

PROBLEM FRAMEWORK

1.1 Introduction

Europe is among the largest consumer of natural gas in the world, mainly used for heating and power generation. During the last thirty years natural gas has gradually replaced almost everywhere fuel oil for heating purposes and is actually competing with coal as main fuel source for electric power generation. Hence, long term trend of natural gas demand has been historically upward sloping. The economic crisis of 2008 has strongly impacted this tendency: global gas demand fell sharply by 3% between 2008 and 2009. However, as reported in Table 1.1, the International Energy Agency (IEA) forecasts that OECD[1] gas demand would recover slowly with consumption returning to the 2008 levels by 2012 or 2013, depending on the region. In addition, recent events concerning nuclear power generation, post Fukushima's accident, are expected to provide new strength to the long term up-growing tendency of natural gas global demand. In fact, in the medium to long term, many countries are expected to reduce their nu-

[1]Current membership of OECD: Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

clear ambitions and the fuel of choice to compensate for lower nuclear will reasonably be natural gas.

	2008	2009	2010	2011	2012	2013
Europe	557	527	544	533	540	548
North America	814	800	804	803	820	835
Pacific	175	169	180	182	189	195
Total	1546	1496	1528	1518	1549	1578

Table 1.1: OECD Natural Gas Demand by Region in billion cubic meters per year. Datas for 2008-2009 are historical, for 2010 estimated, and for 2011-2013 forecasted [35].

Despite its significant consumption, Europe (meant either as OECD or European Union (EU)), has only a limited inner production compared to its consumption and the excess demand is covered by massive natural gas imports from producer countries like Russia and Algeria, as shown in Table 1.2 where the Natural Gas Imports for the EU-27[2] countries is reported.

Natural gas imports are physically delivered via pipelines or, recently, via LNG (Liquified Natural Gas) cargoes. In any case, in order both to guarantee the security of supply of such an important energy commodity whose storability is limited and fulfill the complex customers' patterns of consumption, lot of supply contracts allow *flexibility of delivery*. In particular, in gas markets many long-terms contracts (for 10 years or more) are embedded with 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. As mentioned, 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

[2]EU-27: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxemburg, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the United Kingdom.

	Russia	Norway	Algeria	Netherlands	Others	Total
TJ/y	4'524'090	4'055'038	1'868'376	1'691'445	3'619'020	15'757'969
%	28.7%	25.7%	11.9%	10.7%	23%	100.0%

Table 1.2: EU-27 Natural Gas Imports 2009 by Country of Origin. *Source: EUROSTAT, last update 15/06/2011*

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 [1, 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. We however emphasize that USD/EUR exchange rate volatility is comparatively low compared with typical spot gas price volatility (Figure 1.1).

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*.

In the recent years swing options received vast treatment in the literature (see for instance [4, 26] and references therein for what concerns gas markets, and references in [3, 2, 6, 22] for swing options in more general markets). The scope of this work is to investigate the stochastic optimal control problem from a mathematical point of view. 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.

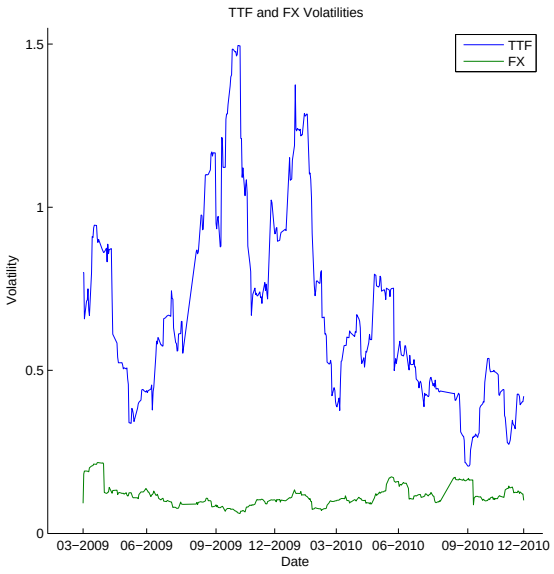


Figure 1.1: TTF (Title Transfer Facility - Netherlands gas hub) and FX (EUR/USD exchange rate) volatilities. TTF volatility is 3 to 10 times larger than that of FX

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.

The rest of the Chapter is organized as follows. We dedicate Section 1.2 to the description of how the strike price of a swing contract is made in practice, and how we can model it from a mathematical point of view. In Section 1.3 we model the pricing problem of a swing contract as a stochastic control problem, introducing the notations and the basic definitions used in the rest of the thesis.

1.2 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 $\{[T^m_{n-1}, T^m_n]\}^{m=1,\dots,D}_{n=1,\dots,12}$, 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

$$I(t) = (T_{n-\xi_2} - T_{n-\xi_2-\xi_1})^{-1} \int_{T_{n-\xi_2-\xi_1}}^{T_{n-\xi_2}} \alpha B(s) ds, \quad \forall t \in [T_n, T_{n+\xi_3}), \forall n \in \mathcal{I}(\xi_3) \tag{1.1}$$

Notice that the following relationship holds:

$$I(t) \equiv I(T_n), \qquad \forall t \in [T_n, T_{n+\xi_3}), \forall n \in \mathcal{I}(\xi_3) \tag{1.2}$$

It follows that the index price $I(t)$ could be modelled in two ways. The first obvious way, given the identity in Eq. (1.2), 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. (1.1) 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*

$$I(t) = \ell_1^{-1} \int_{t-\ell_1-\ell_2}^{t-\ell_2} \alpha B(s)ds \tag{1.3}$$

Notice that Eq. (1.1) and Eq. (1.3) 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. (1.1); 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. (1.3) and not a constant one as in Eq. (1.1).

Unfortunately, the definition in (1.3) is clearly non-Markovian, being an average on past values of $B(t)$. Here we make the following assumption: the dynamics in (1.3) can be approximated by a new markovian one, solution of the following SDE

$$dI(t) = \mu_i(t, I(t))dt + \sigma_i(t, I(t))dW_i(t) \tag{1.4}$$

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 $P(t)$:

$$dP(t) = \mu_p(t, P(t))dt + \sigma_p(t, P(t))dW_p(t) \tag{1.5}$$

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.

1.3 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

$$\mathsf{mDQ} \leqslant u(t) \leqslant \mathsf{MDQ} \qquad \forall t \in [0, T] \tag{1.6}$$

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

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

but also the relationship

$$Z^m(T_0^m) = 0 \qquad \forall m = 1, \dots, D \tag{1.7}$$

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

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

Sometimes the bounds on mAQ and ACQ can be overridden, but a penalty is paid (see for example [4]). In this case

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

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, see Chapter 4), this fact and Equation (1.7) 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. (1.6) forces the buyer of the contract to buy, during a year, *at least* the quantity $\mathsf{mDQ} \cdot T$. 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 [2, Section 2]. We let

$$u(t) \in U = [0, \bar{u}], \quad \bar{u} = \mathsf{MDQ} - \mathsf{mDQ} \tag{1.8}$$

To keep a general view, we also let

$$Z(t) = \int_0^t u(s)ds \tag{1.9}$$

$$Z(T) \in \left[\underline{M}, \bar{M}\right], \quad \underline{M} = \mathsf{mAQ}, \bar{M} = \mathsf{ACQ} \tag{1.10}$$

Penalties are often imposed if the constraints in (1.10) is not satisfied. An example of such penalties can be given by the function

$$\Psi(z) = \begin{cases} \boldsymbol{p} \cdot \underline{M} & z \in (-\infty, 0) \\ \boldsymbol{p} \cdot \left[(z - \bar{M})^+ + (\underline{M} - z)^+\right] & z \in [0, \bar{u}T] \\ \boldsymbol{p} \cdot (\bar{u}T - \bar{M}) & z \in (\bar{u}T, +\infty) \end{cases} \tag{1.11}$$

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 that we make in this thesis could be the continuity and the polynomial growth of the function Ψ, given the following definition:

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:*

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

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

$$\hat{I}(t) = I(\varphi(t)) \tag{1.12}$$

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

$$V^1(0, X_0) = \sup_{u \in \mathcal{A}} \mathbb{E}\left[\int_0^T e^{-rs}(P_s - \hat{I}_s)u_s ds + e^{-rT}\Psi(Z_T)\right] \tag{1.13}$$

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 $t \in (T_{n-1}, T_n]$ the dynamics of X_t is

$$\begin{aligned} dX_t &= f(t, X_t, u_t)dt + \Sigma(t, X_t, u_t)dW(t) = \\ &= \begin{pmatrix} \mu_p(t, P_t) \\ \mu_i(t, I_t) \\ 0 \\ u(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_p(t, P_t) & 0 \\ \sigma_i(t, I_t)\rho & \sigma_i(t, I_t)\sqrt{1-\rho^2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \end{aligned} \tag{1.14}$$

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 $I_{\psi(t)} = I_{T_n} = \hat{i}$ 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 $\hat{i}$ is known and fixed. Let us define:

$$\begin{aligned} V^{13}(t, x) &= \Psi(z) \qquad \forall (t, x) \in [0, T] \times \mathcal{S} \\ V^n(t, x) &= \sup_{u \in \mathcal{A}} \mathbb{E}_{t,x} \left[\int_t^{T_n} e^{-r(s - T_{n-1})} (P_s - \hat{i}) u_s ds \right. \\ &\quad \left. + e^{-r(T_n - T_{n-1})} V^{n+1}(T_n, P_{T_n}, I_{T_n}, I_{T_n}, Z_{T_n}) \right] \quad \begin{matrix} n = 1, \ldots, 12 \\ t \in [T_{n-1}, T_n] \end{matrix} \end{aligned} \tag{1.15}$$

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. (1.14) 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.

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

$$\begin{aligned} L^n(t, x, u) &= -e^{-r(t-T_{n-1})}(p - \hat{i})u, \quad s \in [T_{n-1}, T_n] \\ \psi^n(x) &= -e^{-r(T_n - T_{n-1})}V^{n+1}(T_n, p, i, i, z) \end{aligned} \tag{1.16}$$

and substitute them in the function V^n , rewriting Eq. (1.15) as

$$\begin{aligned} V^{13}(t, x) &= 0 \qquad \forall (t, x) \in [0, T] \times \mathcal{S} \\ V^n(t, x) &= -\inf_{u \in \mathcal{A}} \mathbb{E}_{t,x} \left[\int_t^{T_n} -e^{-r(s-T_{n-1})}(P_s - \hat{i})u_s ds \right. \\ &\qquad \left. -e^{-r(T_n - T_{n-1})}V^{n+1}(T_n, P_{T_n}, I_{T_n}, I_{T_n}, Z_{T_n}) \right] = \\ &= -\inf_{u \in \mathcal{A}} J^n(t, x; u) \end{aligned} \tag{1.17}$$

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

$$J^n(t, x; u) = \mathbb{E}_{t,x} \left[\int_t^{T_n} L^n(s, X_s, u_s) ds + \psi^n(X_{T_n}) \right] \tag{1.18}$$

From now on in this section, we mainly apply, and when necessary extend, the results of [20]. There, the general problem faced has as value function of the form

$$V(t, x) = \inf_{u \in \mathcal{A}} J(t, x; u) \tag{1.19}$$

We should introduce a new sequence of value functions $\mathcal{V}^n(t, x) = -V^n(t, x)$ for which the results in [20] 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.

Chapter 2

VISCOSITY SOLUTION FOR SWING CONTRACTS

The method of dynamic programming provides a powerful tool for approaching the problem in Eq. (1.19). 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 [8, 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 [15] such a weak formulation which they called *viscosity solution*.

Following [20], in Section 2.1 and in its subsections, we introduce the theory of such a solution and in some cases we extend the classical results in order to be able to apply the general theory to our case. Assumptions, definitions, notation used in this chapter are presented together with some classical and ah-hoc extended results: existence, uniqueness and smoothness is proved for the case when the final condition of the problem is not C^2. Section 2.2, where the theory developed is applied to the one-year problem presented Section 1.3, concludes the chapter.

2.1 Main theoretical results about viscosity solution

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

$$-V_t(t,x) + \mathcal{H}(t,x,D_xV(t,x),D_x^2V(t,x)) = 0 \tag{2.1}$$

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 $\bar{V} \in C\left(\overline{Q}\right)$ is*

- *a viscosity subsolution of Eq. (2.1) if for each $v \in C^{1,2}(Q)$*

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

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

- *a viscosity supersolution of Eq. (2.1) if for each $v \in C^{1,2}(Q)$*

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

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

- *a viscosity solution of Eq. (2.1) if it is both a viscosity subsolution and a viscosity supersolution of Eq. (2.1) 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

$$dX_s = f(s, X_s, u_s)dt + \Sigma(s, X_s, u_s)dW_s, \qquad s \in [t, T] \tag{2.2}$$

given the initial condition $X(t) = x$ 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*

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

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. (2.2) with initial condition

$X_t = x$ and L and ψ are *continuous* functions with polinomial growth, i.e. such that

$$|L(t,x,u)| \leqslant C_3(1+|x|^m) \tag{2.3}$$
$$|\psi(x)| \leqslant C_3(1+|x|^m) \tag{2.4}$$

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 $u \in \mathcal{A}_{t,\nu}$:

$$V_\nu(t,x) = \inf_{u \in \mathcal{A}_{t,\nu}} J(t,x;u) \tag{2.5}$$

and finally we define the *value function* as:

$$V(t,x) = \inf_\nu V_\nu(t,x) \tag{2.6}$$

This problem is linked, and in this section we will detail this link, to a partial differential equation of the kind of Eq. (2.1), called *dynamic programming equation* or *Hamilton-Jacobi-Bellman equation (HJB equation)*, obtained by imposing

$$\mathcal{H}(t,x,p,A) = \sup_{u \in U} \left\{ -f(t,x,u)\cdot p - \frac{1}{2}\operatorname{tr}(A\cdot(\Sigma\Sigma')(t,x,u)) - L(t,x,u) \right\} \tag{2.7}$$

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. (2.1), with $\mathcal{H}$ as in Eq. (2.7), is the so called *dynamic programming* property for the value function.

Definition 3 *We say that a function $\bar{V}$ 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*

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

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

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

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

2.1.1 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. (2.6) is a continuous function on $\bar{Q}_0$*

b) property (DP) holds for the value function $V(t,x)$ defined in Eq. (2.6)

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

d) V *is a viscosity solution of Eq. (2.1), with $\mathcal{H}$ defined in Eq. (2.7), in Q_0*

Proof For *(a-c)* see [20, Theorem 7.1 pag. 178]. For *(d)* see [20, Corollary 3.1 pag. 209]. □

Summing up, under the hypothesis of Theorem 1, the PDE (2.1) 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. (2.1), with $\mathcal{H}$ defined in Eq. (2.7). 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. (2.1), with $\mathcal{H}$ defined in Eq. (2.7), in Q_0.*

Proof See [20, Theorem 5.1, pag 72] □

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 [20, Appendix D] the following inequality is proved:

$$\mathbb{E}_{t,x}[\|X_\cdot\|_\infty^m] =\leqslant \xi_m(1+|x|^m) \tag{2.8}$$

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

$$\mathbb{P}_{t,x}\{\|X\|_\infty \geqslant M\} \leqslant \frac{\varsigma}{M}(1+|x|) \tag{2.9}$$

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. (2.8 - 2.9), 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. □

2.1.2 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. (2.3) and (2.4)

Let V_1, V_2 be two viscosity solution of problem (2.1), with $\mathcal{H}$ as in Eq. (2.7) 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]. □

2.1.3 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. (2.12). 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:

$$|L^n_t| + |L^n_x| \leqslant C_4(1 + |x|^\ell) \tag{2.10}$$

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 [20, Lemma 8.1, pag 183], but this is not our case. So we now extend the results to the case where $\psi \in C_p^0(\mathbb{R}^n)$ 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 $n = 1, \dots, 11$. 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*

$$|\psi_x| \leqslant C_4(1 + |x|^k) \tag{2.11}$$

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 [21, 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 (2.10)

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

By Equation (2.11)

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

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 (2.8) 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 [20], 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 (2.11) 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 [18], 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. □

2.2 Viscosity solution for swing contracts

In this section we apply results in Section 2.1 to the one year problem presented in Section 1.3. We continue to make the hypothesis in Assumption 1. Recall that, in this case, the contract value at month n is represented by $-V^n$, $n = 1, \ldots, 13$.

2.2.1 Existence

First of all, we notice that the functions $L^n(t,x,u)$ as defined is Eq. (1.16), being a linear function of p and $\hat{i}$, has polynomial growth in its second argument x, i.e.

$$|L^n(t,x,u)| \leqslant C_3(1+|x|^m) \quad \forall t \in [0,T], \forall u \in U, \forall n = 1,\dots,12 \tag{2.12}$$

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. (1.11). 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 (1.17), 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

$$|V^{n+1}(t,x)| \leqslant B_{n+1}(1+|x|^{m_{n+1}}) \tag{2.13}$$

Using notation in Eq. (1.16), the result in Eq. (2.12), the inductive hypothesis in Eq. (2.13) and the inequalities (2.8) we get:

$$\begin{aligned}
|J^n(t,x,u)| &= \left|\mathbb{E}_{t,x}\left[\int_t^{T_n} L^n(s,X_s,u_s)ds + \psi^n(X_{T_n})\right]\right| \leqslant \\
&\leqslant \mathbb{E}_{t,x}\left[\int_t^{T_n} |L^n(X_s,u_s)ds| + \left|V^{n+1}(T_n,X_{T_n})\right|\right] \leqslant \\
&\leqslant \mathbb{E}_{t,x}\left[\int_t^{T_n} C_3(1+|X_s|^m)ds + B_{n+1}(1+|X_{T_n}|^{m_{n+1}})\right] \leqslant \\
&\leqslant \mathbb{E}_{t,x}\left[\int_t^{T_n} C_3(1+||X_\cdot||_\infty^m)ds + B_{n+1}(1+||X_\cdot||_\infty^{m_{n+1}})\right] = \\
&= C_3(T_{n+1}-t)(1+\mathbb{E}_t[||X_\cdot||_\infty^m]) + B_{n+1}(1+\mathbb{E}_t[||X_\cdot||_\infty^{m_{n+1}}]) \leqslant \\
&\leqslant \varsigma + \xi_m(1+|x|^m) + B_{n+1}B_{m_{n+1}}(1+|x|^{m_{n+1}}) \leqslant \\
&\leqslant B_n(1+|x|^{m_n})
\end{aligned} \tag{2.14}$$

with $m_n = \max\{m, m_{n+1}\} \leqslant 2$ being both $m \leqslant 2$ (see Eq. (2.12)) 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}. □

In the following theorem we prove 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 $n = 1,\dots,11$

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, $V^{12}=V_\nu^{12}$ for every reference probability system.*

Proof Being $\Psi(z)$ a bounded and uniformly continuous function, we can apply Corollary 7.1, pag 181, in [20].

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:*

$$-V_t^{12}+\mathcal{H}(t,x,D_xV^{12},D_x^2V^{12})=0 \tag{2.15}$$

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

$$\begin{aligned}\mathcal{H}(t,x,D_xV^{12},D_x^2V^{12})=&\\ -\tfrac{1}{2}\sigma_p^2V_{pp}^{12}-\rho\sigma_p\sigma_iV_{pi^{12}}-\tfrac{1}{2}\sigma_i^2V_{ii}^{12}-\mu_pV_p^{12}-\mu_iV_i^{12}&\\ +\sup_{u\in U}\left\{\left(e^{-r(t-T_{11})}(p-\hat{i})-V_z^{12}\right)u\right\}=0&\end{aligned} \tag{2.16}$$

and Eq. (2.15) reads:

$$\begin{aligned}V_t^{12}+\tfrac{1}{2}\sigma_p^2V_{pp}^{12}+\rho\sigma_p\sigma_iV_{pi^{12}}+\tfrac{1}{2}\sigma_i^2V_{ii}^{12}+\mu_pV_p^{12}+\mu_iV_i^{12}&\\ -\sup_{u\in U}\left\{\left(e^{-r(t-T_{11})}(p-\hat{i})-V_z^{12}\right)u\right\}=0&\end{aligned} \tag{2.17}$$

Proof See [20, Corollary 3.1, pag 209]. Alternatively, we can use Theorem 2 since $\Psi(z)$ has polynomial growth. □

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 (2.17) in $[T_n,T)\times\mathbb{R}^4$, with V replaced by V^n, and with final condition:*

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

Moreover, $V^n=V_\nu^n$ 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. (2.17). Let us suppose that $V^{n+1}(t,x)$ is a viscosity solution of Eq. (2.17). 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

$$V^n(T_n,p,i,\hat{i},z)=e^{-r(T_n-T_{n-1})}V^{n+1}(T_n,p,i,i,z)$$

□

2.2.2 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

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

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

2.2.3 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 $n = 1,\dots,12$ in our case, that is when L^n is the one in Equation (1.16). Also condition *(ii)* can be easily proved:

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

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

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

Starting from $n = 12$, the final condition Ψ defined in Eq. (1.11) 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 $n = 1,\dots,12$, if Assumptions 1 and 2 hold, then the derivatives $V_x^n(t,x)$ exist, they are in $L^p_{\mathrm{loc}}(Q_0)$ 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. (1.11) 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 $L^p_{\text{loc}}(Q_0)$ 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

$$|\psi_x^n(x)| = |V_x^{n+1}(T_n, x)| \leqslant M_1(1 + |x|^\ell) \tag{2.18}$$

This is true for $n = 11$, as proved in Corollary 2. We recursively 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 $V_x^n(t, x) \in L^p_{\text{loc}}$ and its polinomyal growth:

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

This completes the proof. □

Remark 3 We now come back to our original problem with the right minus sign by substituting the function V with $-V$ in Eq. (2.17). Thus the correct equation satisfied by the value function in formula (1.17) is:

$$\begin{aligned} &V_t^n + \frac{1}{2}\sigma_p^2 V_{pp}^n + \rho\sigma_p\sigma_i V_{pi}^n + \frac{1}{2}\sigma_i^2 V_{ii}^n + \mu_p V_p^n + \mu_i V_i^n + \\ &\quad + \sup_{u \in U}\left\{\left(e^{-r(t-T_{n-1})}(p - \hat{i}) + V_z^n\right) u\right\} = 0 \quad t \in [T_{n-1}, T_n] \\ &V^n(T_n, p, i, \hat{i}, z) = e^{-r(T_n - T_{n-1})} V^{n+1}(T_n, p, i, i, z) \end{aligned} \tag{2.19}$$

By substituting the discounted value function $\tilde{V}(t,x) = e^{-r(t-T_{n-1})}V(t,x)$ in Eq. (2.19) we obtain an equation analogous to the one in [6]:

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

Existence of the optimal control

We proved that the first derivatives V_x^n exists. Coming back to our HJB equation in (2.19), we can state that also a candidate for the optimal control is a.s. well defined. In fact, a straight calculation leads from (2.19) to:

$$u^* = u^*(t,x,T_{n-1}) = \begin{cases} 0 & \text{if} \quad e^{-r(t-T_{n-1})}(p-\hat{i}) + V_z(t,x) \leqslant 0 \\ \bar{u} & \text{if} \quad e^{-r(t-T_{n-1})}(p-\hat{i}) + V_z(t,x) > 0 \end{cases} \tag{2.20}$$

Remark 4 As observed in [6], the candidate optimal control in Eq. (2.20) has a nice economical interpretation. In fact, the marginal value V_z says how much the contract value falls down if we increase the cumulated 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 (2.20) into the HJB Equation (2.19), for each fixed z and $\hat{i}$ we obtain the linear partial differential equation

$$\begin{aligned} &V_t^* + \frac{1}{2}\sigma_p^2 V_{pp}^* + \rho\sigma_p\sigma_i V_{pi}^* + \frac{1}{2}\sigma_i^2 V_{ii}^* + \mu_p V_p^* + \mu_i V_i^* + \\ &\quad + \sup_{u\in U}\left\{\left(e^{-r(t-T_{n-1})}(p-\hat{i}) + V_z^*\right)u\right\} = 0 \quad t \in [T_{n-1}, T_n] \\ &V^*(T_n,p,i,\hat{i},z) = e^{-r(T_n-T_{n-1})}V^{n+1}(T_n,p,i,i,z) \end{aligned}$$

for which $V^* = V^n$ is a viscosity solution. However, being this a uniformly parabolic linear PDE in (t,p,i) for almost every $z,\hat{i}$, it has a unique classical solution that coincides with V^n, i.e. $V^n(\cdot,\cdot,\cdot,\hat{i},z) \in C^{1,2}([0,T]\times\mathbb{R}\times\mathbb{R})$ for almost every $z,\hat{i}$ (see, for instance, [32]).

Chapter 3

NUMERICAL METHODS

This chapter focuses on numerical methods to find the price of swing contracts. First of all, in Section 3.1 we introduce a more concrete dynamics for the prices: we use particular cases of the model in [38] which are rather standard models for energy prices (see for example [23, Chapter 23.3] and [26]). In Section 3.2 a finite difference method for Eq. (2.19) is presented. Section 3.3 deals with a popular method used among practitioners: Least Square Monte Carlo, which works on a discrete version of the value function in Eq. (1.13), 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 that 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.

3.1 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

$$\mathcal{P}(s) = (\mathcal{P}(t) - \mu_p)e^{-\theta_p(s-t)} + \mu_p + \sigma_p \int_t^s e^{\theta_p(u-s)} dW_p(u) \tag{3.1}$$
$$\mathcal{I}(s) = (\mathcal{I}(t) - \mu_p)e^{-\theta_i(s-t)} + \mu_i + \sigma_i \int_t^s e^{\theta_i(u-s)} dW_i(u) \tag{3.2}$$

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 (3.1-3.2)

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

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

$$dP_t = d\exp\{\mathcal{P}(t)\} = P_t\left(\left(\theta_p(\mu_p - \log(P_t)) + \frac{1}{2}\sigma_p^2\right)dt + \sigma_p dW_p(t)\right) =$$
$$= \tilde{\mu}_p(t, P_t)dt + \tilde{\sigma}_p(t, P_t)dW_p(t) \tag{3.3}$$
$$dI_t = d\exp\{\mathcal{I}(t)\} = I_t\left(\left(\theta_i(\mu_i - \log(I_t)) + \frac{1}{2}\sigma_i^2\right)dt + \sigma_i dW_i(t)\right)$$
$$= \tilde{\mu}_i(t, I_t)dt + \tilde{\sigma}_i(t, I_t)dW_i(t) \tag{3.4}$$

and

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

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 $\mathfrak{i}$

$$\begin{aligned}g_{\mathcal{P}}(\mathfrak{p};t,x,s) &= \frac{\mathfrak{p} - m_{\mathfrak{p}}(t,\mathfrak{p},s)}{\nu_{\mathfrak{p}}(t,s)} \\ g_{\mathcal{I}}(\mathfrak{i};t,x,s) &= \frac{\mathfrak{i} - m_{\mathfrak{i}}(t,\mathfrak{i},s)}{\nu_{\mathfrak{i}}(t,s)} \\ f_{\mathcal{P},\mathcal{I}}(\mathfrak{p},\mathfrak{i};t,x,s) &= \frac{e^{\frac{-1}{2(1-\rho^2)}((g_{\mathcal{P}}(\mathfrak{p};t,x,s))^2 + (g_{\mathcal{I}}(\mathfrak{i};t,x,s))^2 - 2\rho g_{\mathcal{P}}(\mathfrak{p};t,x,s) g_{\mathcal{I}}(\mathfrak{i};t,x,s))}}{2\pi\nu_p(t,s)\nu_i(t,s)\sqrt{1-\rho^2}}\end{aligned} \tag{3.5}$$

3.2 Finite difference algorithm

Finite difference methods are numerical methods for approximating the solutions to differential equations, which use finite differences to approximate derivatives. In our case, to find out an approximation of the value functions V^n, we build a numerical scheme using a finite difference method for the HJB equation (2.19)

$$\begin{aligned}&V_t^n + \frac{1}{2}\tilde{\sigma}_p^2 V_{pp}^n + \rho\tilde{\sigma}_p\tilde{\sigma}_i V_{pi}^n + \frac{1}{2}\tilde{\sigma}_i^2 V_{ii}^n + \tilde{\mu}_p V_p^n + \tilde{\mu}_i V_i^n + \\ &\quad + \left(e^{-r(t-T_{n-1})}(p - \hat{i}) + V_z^n\right) u^*(t,x,T_{n-1}) = 0 \quad t \in [T_{n-1}, T_n] \\ &V^n(T_n,x) = e^{-r(T_n - T_{n-1})} V^{n+1}(T_n,x)\end{aligned}$$

where the coefficients $\tilde{\sigma}_p, \tilde{\mu}_p, \tilde{\sigma}_p, \tilde{\mu}_i$ are as in Formulas (3.3-3.4) and the optimal control $u^*(t,x,T_{n-1})$ has been defined in Equation (2.20).

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. (3.1-3.2) 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.

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

We notice that, having choosen such compact intervals $\mathfrak{P}$ and $\mathfrak{I}$, Assumptions 1 are satisfied also for the price processes (3.1-3.2).

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

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

where $\tilde{\sigma}_p$ and $\tilde{\sigma}_i$ are as in Formulas (3.3-3.4).

Let we follow the finite difference approximation in [29, Section 5.3.1] (let us refer to this approximation as *Kushner scheme*) and use the notation $V_\nu^n(m, l, \hat{l}, r)$ for the approximation of $V^n(t_\nu, p_m, i_l, i_{\hat{l}}, z_r)$. We suppress the

superscript n when not needed.

$$V_t(t,x) \to \tilde{V}_t := \frac{1}{\delta_t}[V_{\nu+1}(m,l,\hat{l},r) - V_\nu(m,l,\hat{l},r)]$$

$$V_p(t,x) \to \tilde{V}_p := \frac{1}{\delta_p}\begin{cases} V_{\nu+1}(m+1,l,\hat{l},r) - V_{\nu+1}(m,l,\hat{l},r) & \text{if } \tilde{\mu}_p(t,x) \geq 0 \\ V_{\nu+1}(m,l,\hat{l},r) - V_{\nu+1}(m-1,l,\hat{l},r) & \text{if } \tilde{\mu}_p(t,x) < 0 \end{cases}$$

$$V_{pp}(t,x) \to \tilde{V}_{pp} := \frac{1}{\delta_p^2}[V_{\nu+1}(m+1,l,\hat{l},r) - 2V_{\nu+1}(m,l,\hat{l},r) + V_{\nu+1}(m-1,l,\hat{l},r)]$$

$$V_i(t,x) \to \tilde{V}_i := \frac{1}{\delta_i}\begin{cases} V_{\nu+1}(m,l+1,\hat{l},r) - V_{\nu+1}(m,l,\hat{l},r) & \text{if } \tilde{\mu}_i(t,x) \geq 0 \\ V_{\nu+1}(m,l,\hat{l},r) - V_{\nu+1}(m,l-1,\hat{l},r) & \text{if } \tilde{\mu}_i(t,x) < 0 \end{cases}$$

$$V_{ii}(t,x) \to \tilde{V}_{ii} := \frac{1}{\delta_i^2}[V_{\nu+1}(m,l+1,\hat{l},r) - 2V_{\nu+1}(m,l,\hat{l},r) + V_{\nu+1}(m,l-1,\hat{l},r)]$$

$$V_{pi}(t,x) \to \tilde{V}_{pi} := \frac{1}{2\delta_i\delta_p}\begin{cases} \text{if } A_{1,2}(t,x) \geq 0: \\ \quad V_{\nu+1}(m+1,l+1,\hat{l},r) + V_{\nu+1}(m-1,l-1,\hat{l},r) \\ \quad +2V_{\nu+1}(m,l,\hat{l},r) - V_{\nu+1}(m+1,l,\hat{l},r) - V_{\nu+1}(m-1,l,\hat{l},r) \\ \quad -V_{\nu+1}(m,l+1,\hat{l},r) - V_{\nu+1}(m,l-1,\hat{l},r)] \\ \text{if } A_{1,2}(t,x) < 0: \\ \quad -V_{\nu+1}(m+1,l-1,\hat{l},r) - V_{\nu+1}(m-1,l+1,\hat{l},r) \\ \quad -2V_{\nu+1}(m,l,\hat{l},r) + V_{\nu+1}(m+1,l,\hat{l},r) + V_{\nu+1}(m-1,l,\hat{l},r) \\ \quad +V_{\nu+1}(m,l+1,\hat{l},r) - V_{\nu+1}(m,l-1,\hat{l},r)] \end{cases}$$

$$V_z(t,x) \to \tilde{V}_z := \frac{1}{\delta_z}[V_{\nu+1}(m,l,\hat{l},r+1) - V_{\nu+1}(m,l,\hat{l},r)]$$

Using such approximations we get an explicit scheme for $V_\nu(m,l,\hat{l},r)$ which lead us to rewrite the HJB equation in (2.19) as:

$$\begin{aligned} -\tilde{V}_t &= \frac{1}{2}\tilde{\sigma}_p^2\tilde{V}_{pp} + \rho\tilde{\sigma}_p\tilde{\sigma}_i\tilde{V}_{pi} + \frac{1}{2}\tilde{\sigma}_i^2\tilde{V}_{ii} + \tilde{\mu}_p\tilde{V}_p + \tilde{\mu}_i\tilde{V}_i \\ &\quad + e^{-r(t-T_{n-1})}(p - \hat{\imath} + \tilde{V}_z)u^*(t,x,T_{n-1}) \\ V_\nu(m,l,\hat{l},r) &= \delta_t\Bigg(\frac{1}{2}\tilde{\sigma}_p^2\tilde{V}_{pp} + \rho\tilde{\sigma}_p\tilde{\sigma}_i\tilde{V}_{pi} + \frac{1}{2}\tilde{\sigma}_i^2\tilde{V}_{ii} + \tilde{\mu}_p\tilde{V}_p + \tilde{\mu}_i\tilde{V}_i \\ &\quad + e^{-r(t-T_{n-1})}(p - \hat{\imath} + \tilde{V}_z)u^*(t,x,T_{n-1}) + \frac{1}{\delta_t}V_{\nu+1}(m,l,\hat{l},r)\Bigg) \end{aligned} \tag{3.6}$$

Remark 5 An important feature of the Kushner scheme we presented in (3.6) is that the discretized HJB equation is itself the dynamic programming equation for a suitable defined stochastic control problem for Markov chains. This fact is used in [29] 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 [20, Chapter IX, Sections 4-5].

3.2.1 Boundary conditions

In order to implement the numerical scheme in (3.6), we need some additional boundary conditions. The key point of subsection is the remark that the spot index price I, that is an average on 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 view of this, 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

$$\begin{aligned} &V^n(t, p_{\max}, i, \hat{i}, z) = \\ &= \overline{u}\mathbf{1}_{z<\bar{M}}\mathbb{E}_{t,x}\left[\int_t^{\tau_1} e^{-r(s-T_{n-1})}(P_s - \hat{i})ds\right] \\ &\quad + \mathbb{E}_{t,x}\left[e^{-r(T_n - T_{n-1})}V^{n+1}\left(T_n, P_{T_n}, I_{T_n}, I_{T_n}, z + \overline{u}(\tau_1 - t)\mathbf{1}_{z<\bar{M}}\right)\right] \end{aligned} \tag{3.7}$$

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

$$\begin{aligned} &V^n(t, p_{\min}, i, \hat{i}, z) = \\ &= \overline{u}\mathbf{1}_{z<\bar{M}}\mathbb{E}_{t,x}\left[\int_{\tau_2}^{T_n} e^{-r(s-T_{n-1})}(P_s - \hat{i})ds\right] \\ &\quad + \mathbb{E}_{t,x}\left[e^{-r(T_n - T_{n-1})}V^{n+1}\left(T_n, P_{T_n}, I_{T_n}, I_{T_n}, z + \mathbf{1}_{z<\bar{M}}\int_{\tau_2}^{T_n}\overline{u}ds\right)\right] \end{aligned} \tag{3.8}$$

We next calculate the stochastic integrals (3.7-3.8) defined by the boundary conditions on the truncated boundary. We have, for $a \leqslant b$

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

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

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

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

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

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

$$V^n(T_n,T_n,p,i,\hat{i},z) = e^{-r(T_n-T_{n-1})}V^{n+1}(T_n,p,i,i,z)$$

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:

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

Notice that, in this case, $V^{12}(t, p_{\max}, i, \hat{i}, z)$ and $V^{12}(t, p_{\max}, i, \hat{i}, z)$ 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 (3.6) to find out an approximation of V^{12}.

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

$$\begin{aligned}
V^n(T_n, p, i, \hat{i}, z) &= e^{-r(T_n-T_{n-1})}V^{n+1}(T_n, p, i, \hat{i}, z) \\
V^n(t, p, i, \hat{i}, \overline{u}T) &= e^{-r(T_n-T_{n-1})}\Psi(\bar{u}T) \\
V^n(t, p_{\max}, i, \hat{i}, z) &= \bar{u}\mathbf{1}_{z<\bar{M}}E(t, \tau_1, p_{\max}, \hat{i}, t) + \\
&\quad + e^{-r(T_n-T_{n-1})}\mathbb{E}_{t,x}[V^{n+1}(T_n, P_{T_n}, I_{T_n}, I_{T_n}, Z_{T_n})] = \\
&= \bar{u}\mathbf{1}_{z<\bar{M}}E(t, \tau_1, p_{\max}, \hat{i}, t) + e^{-r(T_n-T_{n-1})} \\
&\quad \cdot \mathbb{E}_{t,x}[V^{n+1}(T_n, P_{T_n}, I_{T_n}, I_{T_n}, z + \mathbf{1}_{z<\bar{M}}(\tau_1 - t)\bar{u})] \\
&= \bar{u}\mathbf{1}_{z<\bar{M}}E(t, \tau_1, p_{\max}, \hat{i}, t) + \\
&\quad + \int_{\mathbb{R}^2} V^{n+1}(T_n, e^x, e^y, e^y, z + \mathbf{1}_{z<\bar{M}}(\tau_1 - t)\bar{u}) \\
&\quad \cdot f_{\mathcal{P},\mathcal{I}}(x, y; t, x, T_n)dxdy \qquad (3.9) \\
V^n(t, p_{\min}, i, \hat{i}, z) &= \bar{u}\mathbf{1}_{z<\bar{M}}E(\tau_2, T_n, p_{\min}, \hat{i}, t) + e^{-r(T_n-T_{n-1})} \\
&\quad \cdot \mathbb{E}_{t,x}[V^{n+1}(T_n, P_{T_n}, I_{T_n}, I_{T_n}, z + \mathbf{1}_{z<\bar{M}}\mathbf{1}_{\tau_2<T_n}(T_n - \tau_2)\bar{u})] = \\
&= \bar{u}\mathbf{1}_{z<\bar{M}}E(\tau_2, T_n, p_{\min}, \hat{i}, t) \\
&\quad + \int_{\mathbb{R}^2} V^{n+1}(T_n, e^x, e^y, e^y, z + \mathbf{1}_{z<\bar{M}}\mathbf{1}_{\tau_2<T_n}(T_n - \tau_2)\bar{u}) \\
&\quad \cdot f_{\mathcal{P},\mathcal{I}}(x, y; t, x, T_n)dxdy \qquad (3.10)
\end{aligned}$$

We can use the approximated values V_1^{n+1} found with finite difference at iteration $n+1$ to calculate the integrals in (3.9-3.10) in this way:

$$\int_{\mathbb{R}^2} V^{n+1}(T_n, e^x, e^y, e^y, z + \mathbf{1}_{z<\bar{M}}(\tau_1 - t)\bar{u}) f_{\mathcal{P},\mathcal{I}}(x, y; t, x, T_n) dxdy \simeq$$
$$\simeq \sum_{m=1}^{N_p} \sum_{l=1}^{N_i} \left[V_1^{n+1} \left(m, l, l, r + \left\lceil \frac{\tau_1 - t_\nu}{\bar{u}} \right\rceil \mathbf{1}_{z_r<\bar{M}} \right) \cdot \right.$$
$$\left. f_{\mathcal{P},\mathcal{I}}(\mathfrak{p}_m, \mathfrak{i}_l; t_\nu, (\mathfrak{p}_{N_p}, \mathfrak{i}_l), T_n) \cdot (\mathfrak{p}_{m+1} - \mathfrak{p}_m)(\mathfrak{i}_{l+1} - \mathfrak{i}_l) \right]$$
$$\int_{\mathbb{R}^2} V^{n+1}(T_n, e^x, e^y, e^y, z + \mathbf{1}_{z<\bar{M}} \mathbf{1}_{\tau_2<T_n}(T_n - \tau_2)\bar{u}) f_{\mathcal{P},\mathcal{I}}(x, y; t, x, T_n) dxdy \simeq$$
$$\simeq \sum_{m=1}^{N_p} \sum_{l=1}^{N_i} \left[V_1^{n+1} \left(m, \ell, \ell, r + \left\lceil \frac{T_n - \tau_2}{\bar{u}} \right\rceil \mathbf{1}_{z_r<\bar{M}} \mathbf{1}_{\tau_2<T_n} \right) \cdot \right.$$
$$\left. f_{\mathcal{P},\mathcal{I}}(\mathfrak{p}_m, \mathfrak{i}_l; t_\nu, (\mathfrak{p}_{N_p}, \mathfrak{i}_l), T_n) \cdot (\mathfrak{p}_{m+1} - \mathfrak{p}_m)(\mathfrak{i}_{l+1} - \mathfrak{i}_l) \right]$$

The roundings $\left\lceil \frac{\tau_1 - t_\nu}{\bar{u}} \right\rceil$ and $\left\lceil \frac{T_n - \tau_2}{\bar{u}} \right\rceil$ are due to the fact that, in general, $\tau_1 - t_\nu$and $T_n - \tau_2$ may not be integer multiples of $\bar{u}$. This would imply that $z_r + \mathbf{1}_{z_r<\bar{M}}(\tau_1 - t_\nu)\bar{u} \notin \mathfrak{Z}$ or $z_r + \mathbf{1}_{z<\bar{M}} \mathbf{1}_{\tau_2<T_n}(T_n - \tau_2)\bar{u} \notin \mathfrak{Z}$, leading to the need of use some interpolation for V.

3.3 A Least Square Monte Carlo algorithm

The Least Square Monte Carlo (LSMC) approach was originally developed by Longstaff and Schwartz [31] for valuing American options. Today, it is widely used also in the energy field to evaluate structured products, see for instance [9],[10] and [14] for application of LSMC to Virtual Storage structured products and [39] for applications to Virtual Power Plant structured products. A summary of existing research on swing option valuation can be found in [30].

The Least Square Monte Carlo works with a discrete time version of the problem (1.13), which we present in the next Eq. (3.14). 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 (1.1), if the time step of the algorithm is one day. 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 among months, as done before.

To avoid a cumbersome 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 $\mathbf{N}$ paths for the 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 orthogonal 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 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

$$\mathbb{E}[V(t_{j+1}, X_{j+1})|\mathcal{F}_{t_j}] = \sum_{\xi=1}^{+\infty} \alpha_\xi^{j+1} f_\xi(x_j, u_j) \tag{3.11}$$

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, and so in practice we make use of

$$\mathbb{E}[V(t_{j+1}, X_{j+1})|\mathcal{F}_{t_j}] \simeq \sum_{\xi=1}^{+N_\xi} \alpha_\xi^{j+1} f_\xi(x_j, u_j) \tag{3.12}$$

3.3.1 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 [17])

$$V(t_j, x_j) = \sup_u \mathbb{E}_{t_j,x_j}\left[\sum_{k=j}^{N_T}(P_k - \hat{I}_k)u_k + \Psi(Z_T)\right] \tag{3.13}$$

where now $x_j = \left(p_j, \hat{i}_j, z_j\right)$.

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:

$$\begin{aligned}
V(t_j, p, \hat{i}, z) &= \sup_{u_j,\dots,u_{N_T}} \mathbb{E}_j\left[\sum_{k=j}^{N_T}(P_k - \hat{I}_k)u_k + \Psi(Z_{N_T})\right] = \\
&= \sup_{u_j,\dots,u_{N_T}} \mathbb{E}_j\left[(P_j - \hat{I}_j)u_j + \sum_{k=j+1}^{N_T}(P_k - \hat{I}_k)u_k + \Psi(Z_{N_T})\right] = \\
&= \sup_{u_j,\dots,u_{N_T}}\left\{(p-\hat{i})u_j + \mathbb{E}_j\left[\mathbb{E}_{j+1}\left[\sum_{k=j+1}^{N_T}(P_k - \hat{I}_k)u_k + \Psi(Z_{N_T})\right]\right]\right\} = \\
&= \sup_{u_j}\left\{(p-\hat{i})u_j + \mathbb{E}_j\left[\sup_{u_{j+1}\dots u_{N_T}}\mathbb{E}_{j+1}\left[\sum_{k=j+1}^{N_T}(P_k - \hat{I}_k)u_k + \Psi(Z_{N_T})\right]\right]\right\} \\
&= \sup_{u_j}\{(p-\hat{i})u_j + \mathbb{E}_j[V(t_{j+1}, P_{j+1}, \hat{I}_{j+1}, Z_{j+1})]\} = \\
&= \sup_{u_j}\{(p-\hat{i})u_j + \mathbb{E}_j[V(t_{j+1}, P_{j+1}, \hat{I}_{j+1}, z+u_j)]\}
\end{aligned} \tag{3.14}$$

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

The Dynamic Programming Principle in (3.14) 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

$$V^n(t_j, p_j^n, \hat{i}_j^n, z) = \sup_u\left\{(p_j^n - \hat{i}_j^n)u + \sum_{\xi=1}^{N_\xi}\alpha_\xi^{j+1} f_\xi(p_j^n, \hat{i}_j^n, z+u)\right\} \tag{3.15}$$

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

$$\min_{\{\alpha_\xi^n\}_{\xi=1,\dots,N_\xi} \subset \mathbb{R}} \sum_{n=1}^{\boldsymbol{N}} \left\| V^n(t_j, p_j^n, \hat{i}_j^n, z) - \sum_{\xi=1}^{N_\xi} \alpha_\xi^j f_\xi(p_{j-1}^n, \hat{i}_{j-1}^n, z) \right\| \tag{3.16}$$

$\rightarrow$ 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 [34], 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.

3.3.2 Radial Basis Functions Approximation

In general, from Formula (3.16) 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 (3.14) as:

$$V(t_j, s, z) = \sup_{u_j} \{ s u_j + \mathbb{E}_j[V(t_{j+1}, S_{j+1}, z + u_j)] \} \tag{3.17}$$

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 [11]. 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 [11] to swing contract. Our final

aim is to compare the finite difference scheme presented in Section 3.2 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 [11]. 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 [11]. It gives the idea of the features of radial basis approach. Finally, next Subsection 3.3.3 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 [11], 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 form of such radial function is

$$\phi(r) = ((1 - \varepsilon r)^+)^4(4\varepsilon r + 1) \tag{3.18}$$

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 [11], 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 $\{x_i\}_{i=1,\dots,N} \subset \Omega$

$$y_i = f(x_i) \quad \forall i = 1, \dots, N \tag{3.19}$$

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

$$s_f(x) = \sum_{j=1}^{M} \alpha_j \Phi_{c_j}(x) = \sum_{j=1}^{M} \alpha_j \phi(\|x - c_j\|) \qquad x \in \mathbb{R}^n \tag{3.20}$$

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 (3.19) and (3.20) as

$$y_i = \sum_{j=1}^{M} \alpha_j \phi(\|x_i - c_j\|) \quad \forall i = 1, \dots, N \tag{3.21}$$

Formula (3.21) 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

$$Y = \mathbf{\Phi}\alpha \tag{3.22}$$

where $Y \in \mathbb{R}^N$ is the vector of observations $Y = (y_1, \dots, y_N)^T$, $\alpha \in \mathbb{R}^M$ is the vector of interpolation coefficients $\alpha = (\alpha_1, \dots, \alpha_M)^T$ 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

$$D = (d_{ij})_{i,j} = \begin{pmatrix} \|x_1 - c_1\| & \cdots & \|x_1 - c_j\| & \cdots & \|x_1 - c_M\| \\ \vdots & & \vdots & & \vdots \\ \|x_i - c_1\| & \cdots & \|x_i - c_j\| & \cdots & \|x_i - c_M\| \\ \vdots & & \vdots & & \vdots \\ \|x_N - c_1\| & \cdots & \|x_N - c_j\| & \cdots & \|x_N - c_M\| \end{pmatrix}$$

and $\mathbf{\Phi} = (\phi(d_{ij}))_{i,j}$.

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. (3.22) is over-determined. Being D generated by a set of distinct center points, it is

always of full column rank M. If (3.22) is over-determined, we can solve it by means of linear least square minimization, i.e. find the solution α^* such that

$$\alpha^* = \arg\min_{\alpha} \|Y - \mathbf{\Phi}\alpha\| \tag{3.23}$$

From linear algebra, a possible solution to (3.23) can be found using the Moore-Penrose inverse $\mathbf{\Phi}^+$ of the matrix $\mathbf{\Phi}$, which in this case can be computed as $\mathbf{\Phi}^+ = \left(\mathbf{\Phi}^T\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^T$ (see [36]), 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 3.3.1 to find out the coefficients in Eq. (3.16). The algorithm rewrites as

Algorithm 2

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

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

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

$\rightarrow$ if $t_j < T$ find out the optimal control $\tilde{u}^n_j$ and the value function V^n for every path n and for every[2] $k = 1, \dots, N_z$ with a numerical maximization

$$V^n(t_j, s^n_j, z_k) = \max_u \left\{ s^n_j u + \sum_{\xi=1}^{N_\xi} \alpha^\xi_{j+1} \phi\left(\left\| \begin{pmatrix} s^n_j \\ z_k + u \end{pmatrix} - c_\xi \right\| \right) \right\} \tag{3.24}$$

[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 [11], 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.

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

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

$$\mathbf{\Phi}_{j-1} = \begin{pmatrix} \left\| \begin{pmatrix} s_{j-1}^1 \\ z_1 \end{pmatrix} - c_1 \right\| & \cdots & \left\| \begin{pmatrix} s_{j-1}^1 \\ z_1 \end{pmatrix} - c_M \right\| \\ \vdots & & \vdots \\ \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_1 \end{pmatrix} - c_1 \right\| & \cdots & \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_1 \end{pmatrix} - c_M \right\| \\ \left\| \begin{pmatrix} s_{j-1}^1 \\ z_2 \end{pmatrix} - c_1 \right\| & & \left\| \begin{pmatrix} s_{j-1}^1 \\ z_2 \end{pmatrix} - c_M \right\| \\ \vdots & & \vdots \\ \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_2 \end{pmatrix} - c_1 \right\| & \cdots & \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_2 \end{pmatrix} - c_M \right\| \\ \vdots & & \vdots \\ \left\| \begin{pmatrix} s_{j-1}^n \\ z_k \end{pmatrix} - c_1 \right\| & & \left\| \begin{pmatrix} s_{j-1}^n \\ z_k \end{pmatrix} - c_M \right\| \\ \vdots & & \vdots \\ \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_{N_k} \end{pmatrix} - c_1 \right\| & & \left\| \begin{pmatrix} s_{j-1}^{\boldsymbol{N}} \\ z_{N_k} \end{pmatrix} - c_M \right\| \end{pmatrix} \tag{3.25}$$

$$Y_j = \begin{pmatrix} V^1(t_j, s_j^1, z_1) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^{\boldsymbol{N}}, z_1) \\ V^1(t_j, s_j^1, z_2) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^{\boldsymbol{N}}, z_2) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^n, z_k) \\ \vdots \\ V^{\boldsymbol{N}}(t_j, s_j^{\boldsymbol{N}}, z_{N_k}) \end{pmatrix} \tag{3.26}$$

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

$$Y_j = \mathbf{\Phi}_{j-1}\alpha_j \tag{3.27}$$

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

$$\alpha_j = \arg\min_{\alpha} \|Y_j - \mathbf{\Phi}_{j-1}\alpha_j\| = \mathbf{\Phi}_{j-1}^{+} Y_j \tag{3.28}$$

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

3.3.3 Reduction to one dimension: cumulated quantity discretization

Even though the system of equations (3.27) should not require a long solution time, we have to notice that Algorithm 3.3.2 requires $N_T \cdot \boldsymbol{N} \cdot N_z$ numerical maximizations (non-linear most of the times, and sometimes integer), coming out from Formula (3.24). 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 (3.24) 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

$$\mathcal{K} = \frac{\overline{M} - \underline{M}}{\bar{u} - \underline{u}} \tag{3.29}$$

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

Proof See [2]. □

Thanks to Theorem 10, when assumption in Eq. (3.29) 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 (3.24) becomes a maximization between only two possible values, and the two dimensional regression falls into a one dimensional regression. Algorithm 3.3.2 changes in this way. Having lost one dimension, now the centers $\{c_\xi\}_{\xi=1,\ldots,M}$ of our RBF lie in $\mathbb{R}$ and the vector Y_j in (3.26) and the matrix $\boldsymbol{\Phi}_j$ in (3.25) read as

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

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

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 (3.28)

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

Formula (3.24) 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

$$\begin{aligned} V^n(t_j, s_j^n, z_k) &= \max\{s_j^n \bar{u} + \mathbb{E}_j[V(t_{j+1}, S_{j+1}, z_{k+1})], \mathbb{E}_j[V(t_{j+1}, S_{j+1}, z_k)]\} \\ &= \max_u\{s_j^n \bar{u} + \mathbf{\Phi}_j^n \alpha_{j+1}^{k+1}, \quad \mathbf{\Phi}_j^n \alpha_{j+1}^k\} \end{aligned} \tag{3.30}$$

3.4 Naïve Monte Carlo with Linear Programming

Naïve Monte Carlo is probably the easiest way of considering both the stochastic nature of price dynamics and the codependence embedded in the optimal exercise of a swing contract. The idea behind the NMC algorithm is to use in a convenient way the *deterministic optimization*. Monte Carlo algorithm proceeds as follow:

1. Discretize the time interval $[0, T]$ into a sequence $\{t_j\}_{j=1,\dots,N_T}$ with $0 = t_1 < t_2 < \dots < t_{N_T} = T$ and simulate $\boldsymbol{N}$ paths for the price dynamics $\{P^n(t_j), \hat{I}^n(t_j)\}_{j=1,\dots,N_T}^{n=1,\dots,\boldsymbol{N}}$
2. Use a *deterministic algorithm* to solve an optimization problem for every simulated path. This deterministic algorithm finds out the optimal contract's value for every path. Let this value be $V^n(0)$
3. Determine the contract value as the average over the $\boldsymbol{N}$ simulated paths:
$$V^{\text{opt}} = \frac{\sum_{n=1}^{\boldsymbol{N}} V^n(0)}{\boldsymbol{N}}$$

Regarding point 2. the deterministic algorithm can be, for instance, both a linear optimization problem or a deterministic dynamic programming algorithm. In the first case, a swing contract may be modeled as this linear problem.

Algorithm 1. Having defined the final penalty function $\Psi(z)$ as in Eq. (1.11), the linear programming problem solving the deterministic valuation of a swing contract is:

$$
\begin{aligned}
V^n(0) = \max_{u_1,\dots,u_{N_T}} \quad & \sum_{j=1}^{N_T} (P^n(t_j) - \hat{I}^n(t_j))u_j + \boldsymbol{p}(x_1 + x_2) \\
s.t. \quad & z_{N_T} = \textstyle\sum_{j=1}^{N_T} u_j \\
& x_1 \geq z_{N_T} - \bar{M} \\
& x_2 \geq \underline{M} - z_{N_T} \\
& x_1, x_2 \geq 0 \\
& u_j \in [0, \bar{u}] \qquad\qquad \forall j = 1, \dots, N_T
\end{aligned}
$$

The clear advantage of such a model is its adaptability: a lot of constraint from a wide class of swing contract can be included by adding the relative constraints. In practice, swing contracts are often characterized also by bounds on the cumulated quantity in subperiod different from the year (such as months or quarters) as well as minimum/maximum withdrawable quantity profiled for every time step j. Also different kinds of penalties and other complex clauses as carry forward and make-up can be included.

This kind of optimization is also simple to understand, fast to be solved and provides $\boldsymbol{N}$ realizations of the contract value $V^i(0)$ which in practice can be used (throught its density) to derive risk measureas such as PaR as well as optimal hedging strategies such as the minimum variance one.

On the other hand, this approach has a serious drawback: it suffers of *perfect foresight* in every path. In other words, for every simulated path this approach suppose an excess of information: the decision-maker knows in advance the price realization for every future istant t_j and he does not take his decision under an uncertainity environment. This, in conclusion, leads to an overestimation of contract's value which in practice is always used as an *upper bound* for the real contract price.

3.5 Numerical Experiments

In the previous sections we presented three algorithms to price a swing contract. Now we put the algorithm at work and compare the results ob-

tained. When not variables, the parameters used in this section are:

$$\begin{array}{rlrl} T & = 365 & \delta_t & = \frac{1}{4} \\ \mu_p & = \log(60) & \sigma_p & = \frac{0.6}{365} \\ \mu_i & = \log(40) & \sigma_i & = \frac{0.1}{365} \\ p_0 & = \log(51) & i_0 & = \log(40) \\ \rho & = 0.1 & \mathbf{p} & = -1000 \\ \mathsf{mDQ} & = 0 & \bar{u} = \mathsf{MDQ} & = 1 \\ \mathsf{ACQ} & = 20 & \mathsf{mAQ} & = 10 \\ [p_{\min}, p_{\max}] & = [1, 300] & \delta_p & = 5 \\ [i_{\min}, i_{\max}] & = [20, 60] & \delta_i & = 5 \\ [z_{\min}, z_{\max}] & = [0, 365] & \delta_z & = \frac{1}{4} \end{array} \tag{3.31}$$

The first part of this section focuses on a sort of *toy contract* of 1 month length. In other words, we decided to study the behaviours and the convergence of the three algorithms with a very short (and very unlikely!) contract. This choice was dictated by the need of *exagerate* some characteristic parameter of every numerical algorithm without exploding execution time. As an example, if we consider the finite difference algorithm, we may be interested in how the choice of δ_t and δ_i influences the valuation; for the LSMC, as said, the number of simulations and regression functions may affect contract price. After having optimized the specific behaviour of every algorithm, we focus our attention on a full 12-months contract.

Regarding the parameters choice, they have not been calibrated, but chosen in a reasonable way. A more accurate calibration will be done in next Section 4.4. Even not calibrated, those parameters are not far from real market, and in particular reflects some relationships: the gas spot price exhibits a higher volatility σ_p with respect to the index one σ_i which is an average; at the beginning of the year the contract is in the money ($p_0 > i_0$); the index is not expected to grow up ($i_0 = \mu_i$), contrary to the spot, resulting in a positive contract value with some exercise opportunity. The penalty is big enough compared to every possible spread level, and this ensure that it has its weight on the pricing procedure. Regarding contract quantities, they are realistic provided the correct unit of measure (typically, millions of cubic meters) and are such that one is forced to exercise at least 10 times, but can not exercise every day.

3.5.1 Some topics on Least Square Monte Carlo

Number of basis functions

The scope of this analysis is to understand how the number N_ξ of basis functions used in the regression formula 3.12 affects the valuation. For a fixed set of price path simulation, with $\mathbf{N} = 10^5$, we let N_ξ taking the values from 40 to 200 with a step size of 20 and we calculate with LSMC

algorithm the contract value. Figure 3.1 reports the results. We can see that, in general, the higher is N_ξ the higher is the contract value, even if the improvement in the pricing process is very small compared to contract value. When $N_\xi = 100, 120$ the contract price is more or less the same, showing a sort of *stationary behaviour* of the algorithm. We can consider one of those number as a good choice for N_ξ. A number of basis function of more or less 100 is also reasonable compared to the high number of simulations used. We may not forget that, even if, in general, the higher is the number of basis the higher is the contract value, with a very large number of basis function the regression tends to be an interpolation of the data V^n. In the limiting case when $\mathbf{N} = N_\xi$, the simulated V^n are perfectly interpolated. In the LSMC area, a perfect interpolation is equivalent to the *perfect foresight* case, i.e. the case when in every scenario we know the future price value, and we use directly the continuation value on that path rather that an approximation of the expected continuation value, such as in the Naïve Monte Carlo. This is not the objective we want to pursue, but given our choices $N_\xi = 120$ and $\mathbf{N} = 10^5$, we are sure that the interpolation is not our case. Taking into

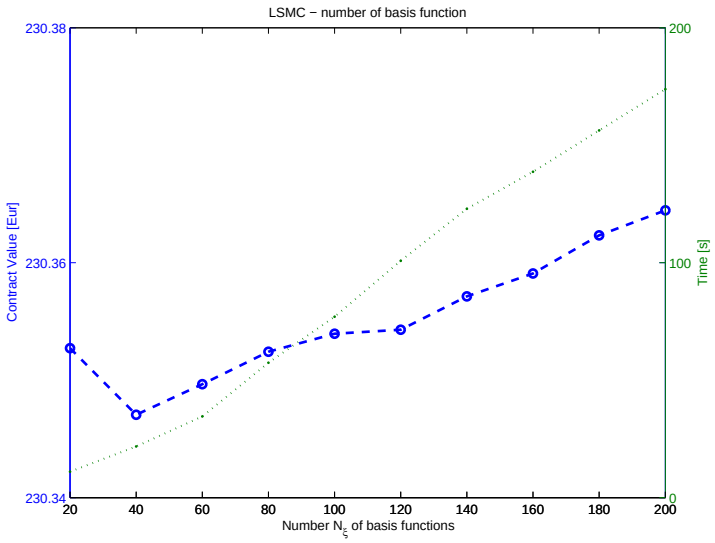


Figure 3.1: Contract value with respect to the number of basis function N_ξ. On the x axis we have the number of basis functions, while on the left y axis, in blue, we put the contract value. The y right axes, in red with dotted line, represents execution time

account also that with a higher number of basis function we have a sensible deterioration of execution times, we definitely choose

$$N_\xi = 120$$

Number of simulations

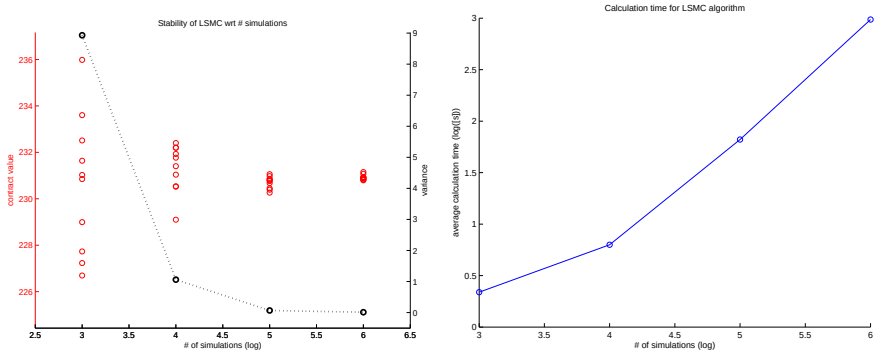
The scope of this analysis is to understand how the number **N** of simulated price paths affects the valuation. We expect that larger number of simulated paths lead to more stable results, in the sense that when the number of paths is higher, different samples lead to the same contract value. To analyse this behaviour, we let **N** taking the values in the set $\{10^n\}$ for $n = 3, 4, 5, 6$. For every fixed n, we make 10 contract valuation, every time simulating new price scenarios. Figure 3.2(a) reports the results. As expected, when **N** is large, different run of the simulations lead, more or less, to the same contract value. We can affirm that the best value for **N** should be greater or equal to 10^6. Unfortunately, 10^6 may be a very large number when dealing with longer contracts: a lot of memory is required to store the simulation, and calculation times are very long especially with large number of basis functions. We know that Monte Carlo algorithms are based on the convergence of the sampled average to the mathematical one, this is the reason why with larger set of simulated path lead to more stable results. However, the same applies if we perform some independent evaluation of the swing contract and take the average of them. In other words, our suggestion is to perform a lot of contract evaluation, say K, but with a low number of simulated path. Then fix the contract value as the average of the K values obtained. This idea is based on the fact that execution time does not perfectly grows up linearly whit the number of simulation, as shown in Figure 3.2(b), but is is faster with a few number of simulations. If we want to be more precise, with an application of the central limit theorem, we can also construct a confidence interval for the sampled contract value (with few simulations). If $\bar{X}$ is the average of the sampled valued, and s_X their standard deviation, then we may state with a confidence level α that the true contract value lies in the interval $\bar{X} \pm Q(1 - \frac{\alpha}{2})s_X$, where $Q(\cdot)$ is the quantile of the normal distribution $\mathcal{N}(0, 1)$.

To have an idea of how this trick can help execution time, the average of the 10 contract values sampled with 10^6 price paths we see in Figure 3.2(a) is 230.90. Every value has been obtained with an average effort of 15 minutes. On the other side, using only 10^4 price path simulation, we can obtain 40 samples in 8 minutes (12 secs for every evaluation) and the average of the 40 calculated contract value is 231.10, very near to 230.90!

3.5.2 Finite difference

The aim of this section is to put at work the finite difference algorithm and to compare the results with the ones of the LSMC and Naïve Monte Carlo for the short 1-month contract.

First of all, let us have a look at the behaviour of the solution with respect to the discretion of the intervals $\mathfrak{T}$ and $\mathfrak{Z}$, i.e. with respect to the length



(a) Contract value calculated 10 times for different choices of **N**. On the x axis we have the number **N** of simulations, while on the left y axis, in blue, we put the contract values. The y right axes, in red with dotted line, represents the variance of the 10 sampled contract values.

(b) Average calculation timing for different choices of **N**. Even if the the swap memory is never used, i.e. all the simulations are stored in RAM, we can see that calculation time explodes with a very large number of simulations

Figure 3.2: Some sensitivities of LSMC algorithm

of δ_t and δ_z. We expect that for smaller value of such values the algorithm tends to a better value for the contract. Following [6], we know that we have to satisfy the Courant-Friedrichs-Lewy (CFL) condition

$$\delta_t \bar{u} \leqslant \delta_z$$

required for the numerical scheme to be stable. So we parametrize the steps $\delta_t = \delta_z = \frac{1}{k}$ and let k varying from 1 to 10. Being $\bar{u} = 1$, in this way the CFL condition is satisfied. Figure 3.3 shows how the valuation is affected by those choices. We can see that the smaller is the step, the higher is contract's value, even if the changes are negligible for $k \geq 4$. On the other hand, the time required to find the solution grows up exponentially for smaller values of δ_t and δ_z, and also the memory required to store matrices into RAM becomes considerable. For the same reason, we decided not to study in a programmatically and empirical way, how the discretization on prices (δ_p and δ_i) affect the valuation: resources and time required were too big, and while the discretization on δ_z strongly affect the optimal control, the derivatives on prices are not crucial in the same way.

Table 3.1 summarizes the results found for the one-month contract. We can observe that the results provided by the LSMC and FD algorithms are very close each other, while the execution time is in favour of the FD algorithm. The perfect foresight behaviour of the NMC is clear: it overestimates the contract value, but on the other side it has the lowest execution time between the three algorithms.

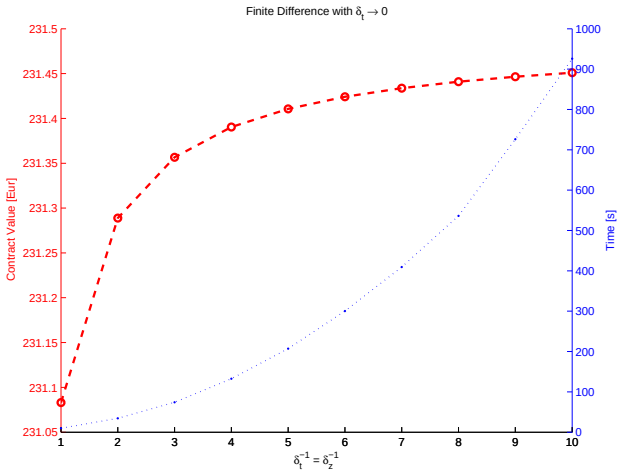


Figure 3.3: Contract value and execution time with finite differences with the relationship $\delta_t = \delta_z = \frac{1}{k}$ for $k = 1, \ldots, 10$.

Algorithm	Contract Value [Eur]	Execution Time [s]
NMC (10^4 sim)	264.73	130
FD	231.45	180
LSMC (10^5 sim)	230.90	320

Table 3.1: Summary of the main results of the three algorithms presented in this Chapter for a one-month contract.

3.5.3 One year contract

After having investigated the behaviour of the algorithms with respect to some of their characteristic parameters, now we switch our attention to a full contract of length one year. For this contract we used the following parameters, the other being fixed as in Eq. (3.31):

$$\begin{aligned} \mathsf{mDQ} &= 0 & \bar{u} = \mathsf{MDQ} &= 1 \\ \mathsf{ACQ} &= 200 & \mathsf{mAQ} &= 100 \\ N_\xi &= 120 & \mathbf{N} &= 10^4 \end{aligned}$$

Regarding the choice of $p_{\max}$ and $i_{\max}$, they can be chosen according to a probabilistic methodology, by solving $\mathbb{P}\{P_T \in [p_{\min}, p_{\max}]\} \leqslant 0.99$. However, we prefer to adapt an empiric point of view. Having done a lot of Monte Carlo simulations for the LSMC algorithm, we may simply have a look to a plot of them, and find the upper bounds as the maximum value under which all the simulated paths lie. From figure 3.4, we may conclude that the spot price is unlikely to be over a value of 300, while the monthly index price is bounded between 20 and 60. Another empirical trial when dealing with the FD algorithm, is to perform the valuation twice, where in

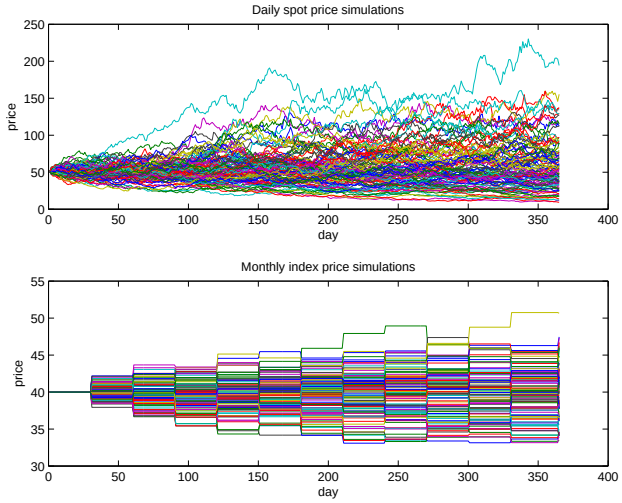


Figure 3.4: 10^5 simulations of daily spot price and monthly index price

Algorithm	Contract Value [Eur]	Execution Time [s]
NMC (10^4 sim)	5087	266
FD	4168	2480
LSMC (10^4 sim)	4032	320
LSMC (10^5 sim)	3958	3670

Table 3.2: Summary of the main results of the three algorithms presented in this Chapter for a one-year contract.

the second valuation we increase the value of $p_{\max}$ and $i_{\max}$ chosen for the first run. If, after the second valuation the contract price is the same as in the first, we can conclude that the support for p and i is large enough.

Table 3.2 shows results similar to the ones in Table 3.1, and the conclusions may be pretty much the same: the NMC overestimates contract value, but it is the fastest, while the results of the FD and LSMC are very close each other. The execution time is in favour of the FD algorithm compared to a LSMC valuation with a high number of simulation, while LSMC performs very fast with a lower number of simulated paths.

3.6 Conclusions

This chapter has been devoted to the study and implementation of three numerical methods to find the price of a swing contract. The first one, called Finite Differences (FD), is a numerical algorithm which find an approximation to the solution of the HJB equation 2.19. It is well known that, for more than three or four state variables, numerical methods such as FD

may face several difficulties being not the optimal practical choice in terms of both time of execution and implementation. In some cases, Monte Carlo methods converge to the solution more quickly than other numerical methods, require less memory and are easier to program. Is this our case? To answer this question, we developed other two methods, very popular among practitioners, based on Monte Carlo simulations. One is the so-called Least Square Monte Carlo (LSMC) which is based on backward solution of the control problem (1.13); the other one is called Naïve Monte Carlo: it is the deterministic version of the problem (1.13) repeated multiple time on Monte Carlo simulations of price paths. We have then analysed several numerical comparison examples in order to study both the performance and the accuracy of all the three methods and to find an answer to our question.

Performance

Regarding the performance, from Tables 3.1 and 3.2 we can conclude that the FD algorithm is competitive both in time and accuracy with LSMC, when one uses 10^5 path simulations. But, when used with a lower number of simulation, such as 10^4, the LSMC performs more or less 8 times faster compared to FD on longer contract, without giving up the accurancy (which in any case can be refined by repeating multiple times the valuation, as did in Figure 3.2(b)).

Practical Comparison

From a practical point of view, LSMC is *independent* of price dynamics: simulated price paths are only an *input* of the LSMC algorithm. For the FD method this is not true: the choice of the price dynamics strongly affect, for instance, boundary conditions (see Section 3.2.1). From a practitioners' point of view, an algorithm which is independent on price dynamics is really an advantage for at least two reasons:

1. it can be used to evaluate the same contract with any sophisticated models; for instance, one may want to use multi-factor models, such as the one presented in [28] where every forward/spot price is driven by two correlated Brownian motions, or introduce jumps as in [33]. With the LSMC algorithm, the only step required to implement such models is the simulations of the dynamics. No additional calculations are needed;

2. an algorithm independent of the dynamics can be easily plugged into a more complex system already existing. This is the typical situation of a big firm which already has its own models for risk management purposes, derivatives evaluation and hedging, portfolio representation and optimization. In such cases, the use of independent price

models for the same portfolio is not recommended, especially when all the prices in this portfolio have some common dependencies (such as correlation or co-integration) which can not be reproduced using different dynamics. This is always the case of energy structured products, being such prices both correlated and co-integrated (see [19, Chapter 4]).

One may object that the LSMC algorithm strongly depends on some user choices. We have (empirically) proved that for swing contracts the use of RBF as basis functions for regression together with a number of simulation greater that 10^4 are good choices that can offer a compromise between speed and accuracy, also when compared to FD. Concluding, in author's opinion, for the swing contracts presented in Section 1.3, the LSMC is not worse that a FD algorithm, and can be used in favour of FD at least when the stochastic model driving price dynamics does not allow an easy calculation of the boundary conditions.

Regarding the NMC algorithm, it is definitely the fastest algorithm and also the one which requires less programming effort. But, as expected, it overestimates the contract value (of about 25%), giving only an upper bound of it. In practice, when feasible, the NMC algorithm should be avoided in favour of LSMC, or FD. On the other hand, there exists some contract for which the NMC seems to be the only viable way of pricing, because of the complexity given by the constraints on quantity and penalty structure. As an example, in next chapter we will face a contract with an additional quantity constraint between years. Because of this clause, we cannot split the contract year by year (as did in Section 1.3) and we need to considered it at once. To evaluate such contract, we will be able to make use of a numerical approach named *lattice of trees* only thanks to the assumptions that minimum and maximum quantity constraints are imposed on months instead of days. However, if the constrains had been written on the days, probably NCM would have been the only viable way to price such a contract in a reasonable amount of time. In any case, when NMC is needed by the complexity of the problem, one has to keep in mind that the results of the valuation process is only an overestimation of the real value.

Optimal Control

Let us come back to the one-month example in order to deduce some (easy to understand) conclusions on the optimal control obtained by the three algorithms.

Regarding the FD algorithm, we obtain from the algorithm a numerical approximation to Equation (2.20), which by its nature depends on time, on the whole states of the system, $x = (p, i, \hat{i}, z)$, and also on the quantity $V_z(t, x)$. We have plotted an example of it in Figure 3.5, where we can see

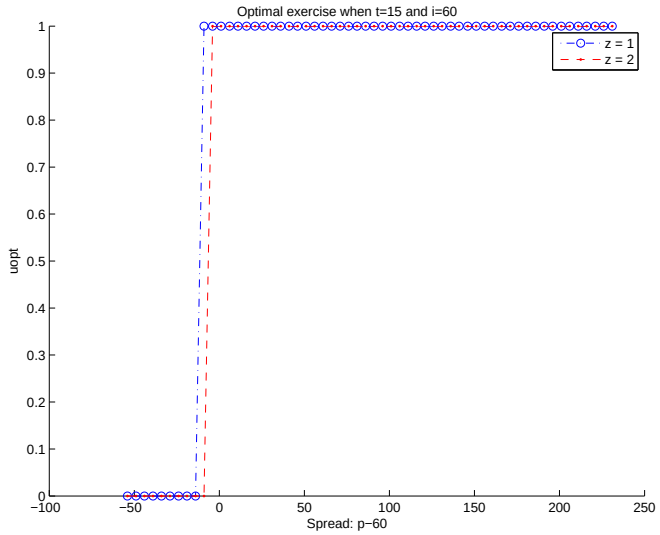
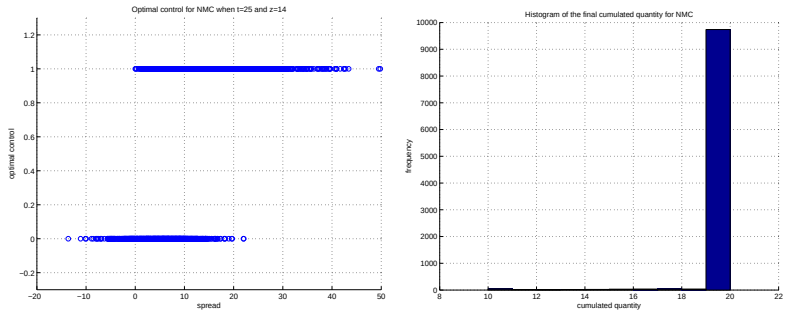


Figure 3.5: Optimal control $u^*(t, p, i, \hat{i}, z)$ obtained with FD algorithm when $t = 15, \hat{i} = 60$ and the cumulated quantity is $z = 1$ or $z = 2$. On the x axes we have the spread between gas market price p and the index $\hat{i}$.

that u^* looks like exactly as one expects: the optimal control is to exercise every time the spread is *sufficiently favorable*, taking into account also the constraint on mAQ, i.e. if we are arrived in $t = 15$ having exercised only one time (blue line, $z = 1$), then the optimal control is to wait to exercise every time the spread level between gas and index is less or equal to -14. Otherwise, if we see a spread greater or equal to -9, we may exercise the swing option. On the other case, that is when we are arrived in $t = 15$ having exercised two times (red line, $z = 2$), we can wait to withdraw even when the spread is -9, and withdraw if the spread is greater or equal to -4. In both cases (being $\delta_p = 5$) nothing is said for spread values lying between the jumps, i.e. when the spread is in $(-14, -9)$ for the case $z = 1$, or $(-9, -4)$ for the case $z = 2$.

Regarding the controls coming out from the two methods based on Monte Carlo simulations, we do not have directly an approximation of the optimal controls as a function, but we obtain a lot of samples of such function. It follows that we need to re-construct the functional form of them, when feasible, or we may treat the output with some statistical methodology. Let us start using an example from NMC algorithm analysed using statistical methodology. Having **N** simulated and optimized paths, we also have **N** optimal choices for every time t. If we assume that for any fixed t this optimal control is a function of the spread level s and the cumulated quantity z (which is an approximation of the real form of the optimal control), we can group the **N** optimized controls in classes, where



(a) Optimal control for the NMC case when $z = 14$ at time 15. We have 4149 simulation among 10^4 such that $\sum_{t=1}^{15} u^*(t) = 14$.

(b) Histogram of the 10^4 cumulated quantities at the end of contract life.

Figure 3.6: Analysis of the NMC output (optimal control and cumulated quantity at the end of contract life)

each class refers to a single value for the cumulated quantity at time t. For instance, at a given time, we may have $z = 10, 15, 10, 12$ for the first four simulations. We then collect the respectively optimal control in the three classes $(10, 12, 15)$, and put in the first class both the optimal choice of the first and third simulation). Next, we may expect that, for a fixed class, we should be able to express the optimal simulated controls as a function of the spread. Unfortunately, this is not completely true, but it is still useful; let us consider Figure 3.6(a) where this situation is plotted for $\mathbf{N} = 10^4, t = 15, z = 14$, i.e. we have plotted the optimal control of the 4149 paths (among 10^4) where $z = 14$ at time 15. From that figure we can only infer that:

1. when the spread is negative, the optimal control is to wait
2. when the spread is bigger than 20, the optimal control is to withdraw

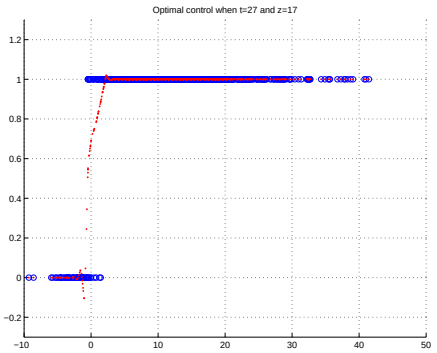
and nothing is well defined for the simulations when the spread lies in $(0, 20)$: we can observe paradoxical cases when we have both controls (0 and 1) for almost the same spread level (rounded to two decimals). Why? Because, by its nature, in the NMC the optimization is done considering the *whole* price path during the *whole* year, so the punctual level of the spread at a given point in time is meaningless if not seen inside its whole trajectory, from the beginning to the end of the year. On the other hand, the conclusions at points 1-2 can be read as a confirmation of the assumptions made in Subsection 3.2.1 for the boundary conditions, i.e. every time the spread is favourable (i.e. p is very high) the optimal control is to withdraw as much as possible, if ACQ has not been reached.

Moreover, we can infer some other statistical informations from the NMC outputs. First, we can construct, for any time, an histogram of the frequency of the cumulated quantity: this helps us to understand how *good* is the contract at any time, in particular at the end of its life. Figure 3.6(b) shows the histogram of the frequency for our example. In such case, we can confirm that the contract is, for sure, in-the-money, the quantity ACQ being the most frequent final vale for z. The optimal choices during the life of the contract were to withdraw a lot of gas. Second, we can construct statistical indicators such as the average withdrawn quantity for every month, useful to build static hedging strategies with forwards, as explained in [19, Chapter 5].

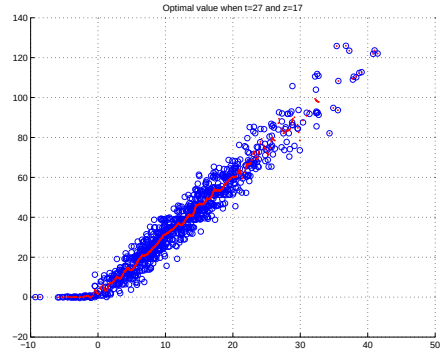
Regarding the LSMC output, an interesting use of them may be related, once again, to the boundary conditions evinced for the FD algorithm. In some sense, we can use the LSMC optimal controls (as did with the NMC ones) to validate the choices done in Section 3.2.1. By applying to u^* the same regression argument done for the value function in Equation (3.12), at every time t the $\mathbf{N}$ optimal controls obtained with LSMC can be expressed as a function. This is coherent with the theoretical assumption that the control is an *adapted process*, i.e. the decision is taken given the current information available at time t from the market. In formula, we can assume that

$$u^*(t_j, s, z) \simeq \sum_{\xi=1}^{+N_\xi} \nu_\xi^j u_\xi(s, z) \tag{3.32}$$

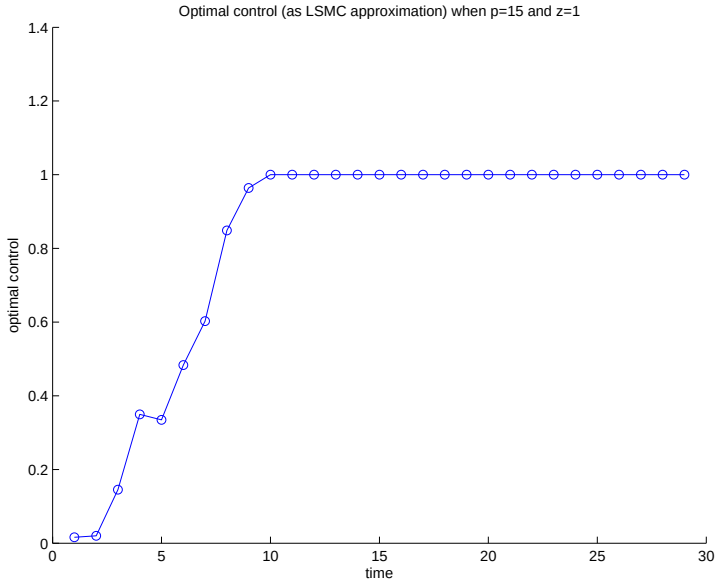
where u_ξ are some basis function and s is the spread level. The coefficients ν_ξ^j are obtained with a least-square regression, using the $\mathbf{N}$ sampled optimal decision from the maximization in Eq. (3.15) or (3.30). We then use the optimal control obtained with Eq. (3.32) to test the hypothesis done for the boundary conditions. For instance, for the one on $p_{\max}$, we can plot the graph of $u^*(\tilde{s}, z)$ for some fixed z and an appropriate high level of $\tilde{s}$. Figure 3.7(c) shows the optimal control for a one-month contract with $z = 1$ (this choice allows to have a value for the optimal control for any time $t \geqslant 1$) and $\tilde{s} = 15$ (which is a sufficient high value for the spread whose starting market value at time 0 is $s_0 = e^{p_0} - e^{i_0} = 11$). We can see that the optimal approximated control is an increasing function of time. In particular, we can read the plot in this way: at the beginning of the month, when $t = 1, 2$, a spread level of 15 is not big enough to suggest the withdrawal of gas. For $t \geqslant 9$ the optimal control, if we arrive at that time with a cumulated quantity equal to 1, is to withdraw (because we need to reach mAQ in order not to pay penalty). For intermediate times, we can affirm that the optimal control is also to withdraw (at least for $t \geqslant 5$), even if the approximated optimal control is not exactly equal to 1: this is a consequence of the approximation procedure of a $0 - 1$ optimal control with smooth functions, as shown in Figure 3.7(a). Luckily, the approximation is not so bad for the



(a) Optimal control in LSMC: simulated and regressed. To draw a 2-d plot, we put the spread in the x axis and the optimal control in the y axis.



(b) Value function in LSMC: simulated and regressed. To draw a 2-d plot, we put the spread in the x axis and the optimal control in the y axis.



(c) Optimal control for a fixed $p = \tilde{p}$ and $z = 1$

Figure 3.7: Analysis of the LSMC output

value function, as presented in Figure 3.7(b).

Summing up, we conclude that:

i. the optimal sequence of decisions coming out from the NMC optimization process can be managed only with statistical methods. They are not any numerical approximation of the optimal policy of the problem, because they are obtained taking into account the full knowledge of the simulated path;

ii. the optimal control from the LSMC can be reduced to an approximation of the optimal policy using least square arguments as done for the continuation value (in the LSMC algorithm, the optimal controls come out from the Bellman equation, so they are effectively obtained using the dynamic programming principle);

iii. the optimal control obtained with FD algorithm is a numerical approximation of the real optimal policy.

Point iii is the most powerful feature of the FD algorithm, which is able to find the best numerical approximation of the optimal control, compared to LSMC. In addition, FD algorithm has another advantage compared to LSMC: it works with the solution of the problem in continuous time (in contrast to LSMC which works in discrete time) taking into account also the dynamics of the index *between two subsequent months*. This extra-information makes the contract's holder able to predict the month-ahead strike more and more accurately as time goes on, and nor the NMC nor the LSMC are able to reproduce this particular behaviour of the indexed strike.

Chapter 4

PRICING MAKE-UP CLAUSE IN DISCRETE TIME

Until now, we have focused our attention to a classical swing contract, where *classical* stands for the fact that the structure presented in Section 1.3 can be seen as the pillar of every traded swin contract, which however in practice may contain more clauses, constraints and variants. Among those clauses, the make-up has become very important in last years. On the other hand, in Chapter 3 we have omitted a pricing technique very used among practitioners: the lattice of trees. As an extension of the results obtained so far, in this chapter we focus our attention to swing contracts with an embedded make-up clause and we use numerical schemes based on trees to evaluate them.

The chapter is organized as follows. In Section 4.1 we describe the make-up clause from an economical and qualitative point of view. In Section 4.2 we formulate the valuation problem of a generic gas swing contract in discrete time, while in Subsection 4.2.3 we introduce a particular instance of make-up clause, which will be the subject of next Section 4.3, where we mathematically frame the problem and indicate an algorithm for its solution, describing analytically both its formal representation and the various steps for the solution. We discuss the computational cost of our approach in Subsection 4.3.1 obtaining a quadratic cost with respect to the duration

in years of the make-up clause and in Subsection 4.3.2 we extend our approach to another form of make-up clause as well as to some instances of carry-forward clause. After having calibrated two mean-reverting trinomial models to market data, as explained in Section 4.4, in Section 4.5 we present a detailed example for a 3-years make-up clause. In particular, in Subsection 4.5.2 we use this contract to perform a sensitivity analysis in order to outline the key drivers for optimization and value protection given the current market scenario. Concluding remarks in Section 4.6 end the Chapter.

The results of this chapter have been published in [17].

4.1 Introduction to make-up clauses

As presented in Section 1.1, long term gas contracts in Europe have been traditionally priced using oil-linked pricing formulas as presented in Section (while, for instance, in the United States gas-to-gas competition has historically determined most of natural gas wholesale transactions). This oil indexation has its origins in the early European gas market of the 1970s, when the greatest sources of gas were the wastes of oil extraction. Since that time, sources of gas have increased, making gas markets and infrastructure much denser and open to competion. In addition, from 2008 onwards this traditional market framework has significantly changed especially for what concerns the oil-to-gas price relationship. Actually, the demand drop following the financial crisis with the subsequent economic recession associated to the significant increase in LNG (Liquefied Natural Gas) and unconventional gas supply sources flowing to Europe generated a consistent and pretty persistent oil decoupling of European gas market prices: since 2008, European gas markets are pricing systematically and significantly below indexes usually used for the strike price I (see Figure 4.1), so the spread $P - I$ has become negative.

Obviously, this market phenomenon has determined a panic situation for all the owners of classical long term gas supply contracts. Significant losses have been faced at present by pipeline importers due to this kind of oil-to-gas price decoupling; moreover, the structural market change determined an increased sense of uncertainty about European gas market future development. Interested readers may refer to [27] for a detailed and updated analysis of oil-to-gas decoupling. This new market scenario has induced many long term importers to engage a renegotiation process with their suppliers together with a more focused attention towards optimization and hedging possibilities which are naturally embedded in the current contracts. As a result, traditionally long term forward gas contracts are today often equipped with some additional volumetric flexibilities. Among those, in this market situation a new and particular importance arose for

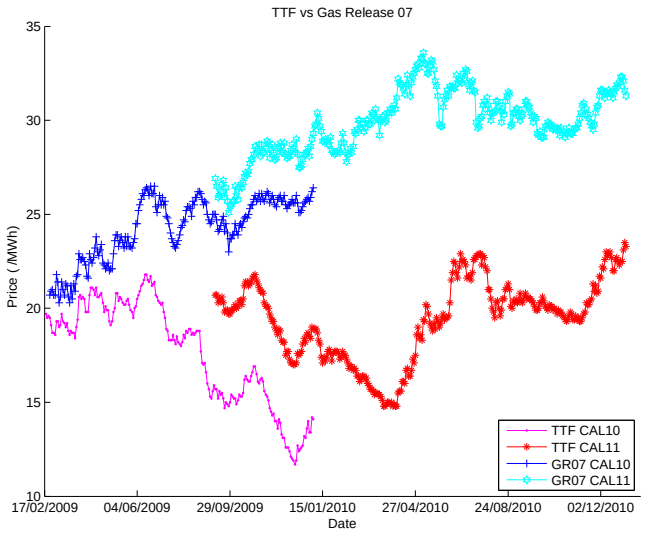


Figure 4.1: CAL10 and CAL11 (calendar forward) of TTF (Title Transfer Facility - Netherlands gas hub) and GR07 (ENI Gas Release 2007, typical gas price formula used by ENI) contracts

the so-called *make-up* and *carry forward* clauses, which flank traditional constraints as minimum and maximum withdrawal quantity established for every contract year and every contract sub-period (day or month). Basically, these clauses allow the buyer of the contract to delay or anticipate respectively the withdrawal of gas from one year to another within the full respect of sub-period capacity constraints. In particular, the introduction of make-up clauses has become very important for European long term contracts holders: in fact, in the recent oil-to-gas price decoupling situation, contracts holders were induced to delay as much as possible the gas delivery for the sake of loss minimization. Once exercised, make-up clause allows to modify the global annual minimum constraint for the year we decide to *nominate* the make-up, and the maximum constraint for the year we decide to *call it back*. With a make-up clause, contract holders can effectively postpone the delivery of gas when it is too expensive with respect to market prices, hoping that future gas prices will rise up and the exercise of the contract rights becomes again profitable.

From a mathematical point of view, the presence of make-up clauses further complicate things and introduces more complexity. Surprisingly, the quantitative literature on this topic appears scarce: in the authors' opinion, this is due to the fact that a make-up clause is worth more in a market where price decoupling is high, and the need to study such markets arose only in the last years. At a qualitative level, for instance, the make-

up clause is described in [30, 37]. An algorithm to evaluate a swing contract with the carry-forward clause using the Least Square Monte Carlo approach is presented in [22], where the authors claim that the make-up can be evaluated similarly: this is true only for make-up clauses with a single final installment, i.e. when the make-up gas is to be paid only when it is called back (which corresponds to letting $\alpha = 0$ in Rule 4 of Section 4.2.3). However, typical make-up clauses have a double-installment mechanism, i.e. the make-up gas has to be paid when nominated with a price proportion $\alpha > 0$ and when called back with a price proportion $1 - \alpha$ (as an example, [30] report $\alpha \in (0.85, 1]$ as typical proportions). To the authors' knowledge, an algorithm to properly price (and find out optimal policy) a swing contract with such a general double-installment make-up clause and where both the market and strike price are stochastic variables has never been presented so far. The aim of this Chapter is exactly to fill this literature gap, at least in the discrete time case. In particular, in this Chapter we will describe, frame and solve in discrete time the optimization issues related to the presence of make-up clauses in long term swing contracts. Finally, we use the algorithm in order to explore the value of the contract with respect to the peculiar constraints introduced by the make-up.

4.2 Problem framework

As already mentioned, the complexity introduced in the pricing problem by the make-up clause is remarkable. In practice, we will see in next Section 4.2.3 that the make-up clause led to co-dependence in exercise among years. As a result, we can not focus our attention to a one-year problem, but we are forced to adopt a larger view on the full life of the make-up clause, which in general spans from 3 to 5 years. An additional state needs to be introduced, which links one year to the following ones. Given this preamble, it is clear that the approach used in Chapter 1 fails, at least from the computational point of view of Section 3.2: with the make-up clause we should consider a five-dimensional state running with daily granularity for some years! We need to do some other approximation to our problem. The idea now is to abandon the *daily* nature of the problem, and treat it with a monthly granularity. From the point of view of the strike price process $\hat{I}$, this assumption is equivalent to model it as a discrete sequence of random variable, as introduced in Eq. (1.2), losing the information given by the moving average. From the point of view of the market gas price $P(t)$, this may be a strong restriction to our problem: we are losing all the flexibility given by daily changes to the price. But if our aim is to focus the attention on the make-up clause and to obtain a pricing in a reasonable time, this approximation is a rational way to mitigate accuracy and feasibility. So, from now on, we do not take care about the moving average $I(t)$ and we will

focus directly on a monthly indexed strike.

4.2.1 Time structure and admissible strategies

Ordinary swing contract schemes are normally defined dividing each one of the D yearly delivery periods $\{[T_{j-1}, T_j)\}_{j=1,\dots,D}$, into N sub-periods $\{[t_{j,i-1}, t_{j,i})\}_{i=1,\dots,N}^{j=1,\dots,D}$ obtaining the sequence $\{t_{j,i}\}$ such that

$$0 = T_0 = t_{1,0} < t_{1,2} < \ldots < t_{1,N} = T_1 = t_{2,0} < t_{2,1} < \ldots$$
$$\ldots < t_{j,i} < \ldots < t_{j,N} = T_j = t_{j+1,0} < \ldots < t_{D,N} = T_D$$

In particular, in every year $[T_{j-1}, T_j]$ we have the $N+1$ points $(t_{j,i})_{i=0,\dots,N}$ such that $t_{j,0} = T_j$ and $t_{j,N} = T_{j+1}$.

We are also assuming that N is also the number of exercise swing rights the holder has in every year, which can be exercised exactly at the points $t_{j,i}$, for $i = 0, \dots, N-1$ i.e. at the beginning of every sub-period. For example, in our simplified case, the decisions are taken month by month, at the beginning of every month, and so $N = 12$.

Denote by $u_{j,i}$ the quantity of gas the holder decides to buy in the sub-period $[t_{j,i}, t_{j,i+1})$, $i = 0, \dots, N-1$, and by $z_{j,j}$ the cumulated gas quantity at time $t_{j,i}$. In particular we set $z_{j,0} = 0$ for all $j = 1, \dots, D$ and

$$z_{j,i+1} = \sum_{k=0}^{i} u_{j,k} = z_{j,i} + u_{j,i} \quad \forall i \in \{0, \dots, N-1\} \tag{4.1}$$

As in the continuous time case, over each one of the N sub-periods, minimum (mDQ) and maximum (MDQ) delivery quantities are established in the contract, which usually reflect physical effective transportation capacity limitations: thus, the quantities $u_{j,i}$ are constrained by

$$\mathsf{mDQ} \leqslant u_{j,i} \leqslant \mathsf{MDQ} \qquad \forall i = 0, \dots, (N-1), \quad \forall j = 1, \dots, D \tag{4.2}$$

which is the analogous of Equation (1.6). Also in this case, *for every contractual year*, minimum and maximum quantities are also established, called respectively minimum annual quantity (mAQ) and annual contract quantity (ACQ). Two key quantities in modeling the make-up are respectively the difference between the maximum gas that the holder could physically take and his contract right, given by

$$\overline{\mathcal{M}} := N \cdot \mathsf{MDQ} - \mathsf{ACQ} \tag{4.3}$$

and the difference between the minimum gas that the holder must take by contract and the minimum which he could physically take, given by

$$\underline{\mathcal{M}} := \mathsf{mAQ} - \mathsf{mDQ} \cdot N \tag{4.4}$$

We have always non-trivial volume constraints, in the sense that

$$\overline{\mathcal{M}} > 0, \quad \underline{\mathcal{M}} > 0 \tag{4.5}$$

Thus, in the light of the discussion above, without any additional clauses and with non-trivial constraints we have

$$N \cdot \mathsf{mDQ} < \mathsf{mAQ} \leqslant z_{j,N} \leqslant \mathsf{ACQ} < N \cdot \mathsf{MDQ} \qquad \forall j = 1, \dots, D$$

To maintain a practical point of view, in this chapter we do not to use the standard decomposition as did in (1.8, 1.10), working directly with contractual quantities.

Penalty payments can be imposed if the volume constraints are exceeded in order to stimulate the buyer to respect the volumetric limits imposed as did in Equation 1.11. For the sake of notation and exposition, in this chapter we do not take into account these penalties: we take $\Psi(z) \equiv 0$.

The difference between swing contracts with trivial and non-trivial volume constraints is extremely important in the pricing and hedging of the contract itself. In fact, with non-trivial volume constraints the holder must take into account, at time $t_{j,i}$, not only the quantity $u_{j,i}$ which would be optimal for that period, but also the effects of this quantity on the future decisions that he will be allowed to take after.

This brings to model the *space of controls* in the following way. For a given year $j = 1, \dots, D$, assume that we have a final constraint $z_{j,N} \in [\underline{z}, \overline{z}]$ for some $0 \leq \underline{z} < \overline{z}$. We may have $\underline{z} = \mathsf{mAQ}, \overline{z} = \mathsf{ACQ}$ exactly as in Equation (1.10), but here we keep a general view: those bounds will be strongly affected by the make-up clause. So, in general, at a given time $t_{j,i}$, the space of controls $\mathcal{A}\left(t_{j,i}, z_{j,i}, [\underline{z}, \overline{z}]\right)$ will depend on time $t_{j,i}$, cumulated quantity $z_{j,i}$ and $[\underline{z}, \overline{z}]$.

By the constraints (4.2) and construction of $z_{j,i}$, at time $t_{j,i}$ we can restrict our attention to the case when $z_{j,i}$ satisfies the constraints

$$\mathsf{mDQ} \cdot i \leqslant z_{j,i} \leqslant \mathsf{MDQ} \cdot i \quad \forall i = 0, \dots, N$$

The problem of determining the set $\mathcal{A}_{j,i}$ is non-trivial when Eq. (4.5) holds, which translates in

$$N \cdot \mathsf{mDQ} < \underline{z} \leqslant \overline{z} < N \cdot \mathsf{MDQ}$$

(otherwise we can always reach the values in $[N \cdot \mathsf{mDQ}, N \cdot \mathsf{MDQ}]$). In this non trivial case, we are not allowed to take $u_{j,i} = \mathsf{mDQ}$ for all $i = 0, \dots, N-1$: in fact, there exists a time τ_1 such that, if we have always took this minimum for $t \leqslant \tau_1$, then for $t > \tau_1$ we have to switch to $u_{j,i} = \mathsf{MDQ}$ in order to reach $\underline{z}$. This point τ_1 is the common point between the two lines

$z = \mathsf{mDQ}(t - t_{j,0})$ and $z = \mathsf{MDQ}\,(t - t_{j,N}) + \underline{z}$, $\forall t \in [t_{j,0}, t_{j,N}]$. A simple calculation leads to

$$z_{j,i} \geqslant r_{\min}(t_{j,i}, \underline{z}) = \max\left\{\mathsf{mDQ}\,(t_{j,i} - t_{j,0})\,, \mathsf{MDQ}\,(t_{j,i} - t_{j,N}) + \underline{z}\right\}$$

Similarly, we are not allowed to take always $u_{j,i} = \mathsf{MDQ}$ either: in fact, there exists a time τ_2 such that, if we have always took this maximum for $t \leqslant \tau_2$, then for $t > \tau_2$ we have to switch to $u_{j,i} = \mathsf{mDQ}$ in order to reach, and not exceed, $\overline{z}$. The boundary for $z_{j,i}$ in this case is

$$z_{j,i} \leqslant r_{\max}(t_{j,i}, \overline{z}) = \min\left\{\mathsf{MDQ}\,(t_{j,i} - t_{j,0})\,, \mathsf{mDQ}\,(t_{j,i} - t_{j,N}) + \overline{z}\right\}$$

Figure 4.2 shows an example of the admissible area.

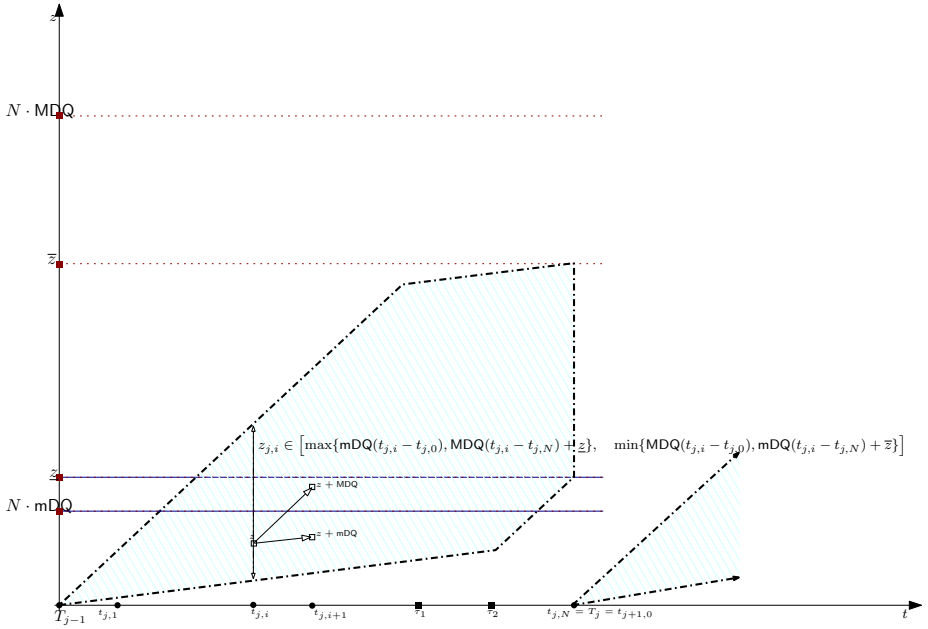


Figure 4.2: Typical admissible area for one year. Here $\underline{z} < \overline{z}$, leaving some optionality for the total intake $z_{j,N}$. If $\underline{z} = \overline{z}$ (typical of years when some make-up gas is nominated or called back), we have the constraint $z_{j,N} = \underline{z} = \overline{z}$ and the admissible region is like those in Figure 4.

In conclusion, the correct form of the space of controls $\mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}])$ at time $t_{j,i}$, given the constraint $z_{j,N} \in [\underline{z}, \overline{z}]$ and the cumulated quantity $z_{j,i} = z$, is given by

$$\mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}]) := \{u_{j,i} \in [\mathsf{mDQ}, \mathsf{MDQ}] \mid z + u_{j,i} \in [r_{\min}(t_{j,i+1}, \underline{z}), r_{\max}(t_{j,i+1}, \overline{z})]\} \tag{4.6}$$

which appears implicitly in [3, Equation 7] and is also a discretized version of the one in [6].

4.2.2 The price of a standard swing contract in discrete time

Let $P_{j,i}$ and $I_{j,i}$ be respectively the prices of gas and index in year $j = 1, \dots, D$, sub-period $[t_{j,i}, t_{j,i+1})$, $i = 0, \dots, N-1$: the contract holder has to buy the gas at the price $I_{j,i}$ and can sell it at the price $P_{j,i}$: of course with this notation we have $(P_{j,N}, I_{j,N}) = (P_{j+1,0}, I_{j+1,0})$ for each year $j = 1, \dots, D-1$. Even if $I = (I_{j,i})_{j=1,\dots,D, i=0,\dots,N}$ is a time average on several subperiod and thus has relevant memory effects, in the following we make the same assumption done in the continuous time case, i.e. that $(P_{j,i}, I_{j,i})_{j=1,\dots,D, i=0,\dots,N}$ evolves as a two-dimensional Markov process under a pricing measure $\mathbb{P}$, which is used in all the mathematical expectations that follow, while the numerical implementation that we use for the analysis of the next section will make use of the particular specification that we describe in Section 4.4. We also assume the non-arbitrage conditions

$$\mathbb{E}[e^{-rt_{j,i}} P_{j,i}] = F^P_{j,i} < +\infty, \quad \mathbb{E}[e^{-rt_{j,i}} I_{j,i}] = F^I_{j,i} < +\infty \tag{4.7}$$

where $F^P_{j,i}$, $F^I_{j,i}$ represent the forward prices of P and of I, respectively, for the delivery month i of year j.

Remark 6 Given the relative illiquidity of some new markets, the assumption that we can observe a complete forward curve for every month, as postulated by the Equation (4.7), may be very restrictive. When this is not possible, we can for instance use some extrapolation procedure to build forward curves coherently with market information and historical observation (seasonality). An elegant and useful way to obtain such a forward curve may be the one presented in [5, Chapter 7].

The objective of contract's holder is the same as in Equation (1.13), i.e. to maximize the discounted global margin of the contract (i.e. minimize the total loss), i.e. we want to calculate the value of

$$V(0, p_{1,0}, \iota_{1,0}, 0) = \sup_{u \in \mathcal{A}} \mathbb{E}\left[\sum_{j=1}^{D} \sum_{i=0}^{N-1} e^{-rt_{j,i}} u_{j,i} \left(P_{j,i} - I_{j,i} \right) \right]$$

In the absence of codependent constraints through years (i.e. without make-up), with the same arguments of Section 1.3, or also by noting that $z_{j,0} = 0$ for all $j = 1, \dots, D$, we may write:

$$V(0, p_{1,0}, \iota_{1,0}, 0) = \sum_{j=1}^{D} \sup_{u \in \mathcal{A}} \mathbb{E}\left[\sum_{i=0}^{N-1} e^{-rt_{j,i}} u_{j,i} \left(P_{j,i} - I_{j,i} \right) \right] \tag{4.8}$$

where the set $\mathcal{A}$ of *admissible controls* is defined by

$$\mathcal{A} := \{(u_{j,i})_{j,i} \text{ adapted to } (P_{j,i}, I_{j,i})_{j,i} \text{ and s.t. } u_{j,i} \in \mathcal{A}(t_{j,i}, z_{j,i}, [\mathsf{mAQ}, \mathsf{ACQ}])\}$$

where $r \geq 0$ is the risk-free annual interest rate.

It is a standard result (see e.g. [3, 4, 26]), and it will also follow as a particular case of our results in Section 4.3, that this maximisation problem can be solved by the use of the Dynamic Programming: for each year $j = 1, \ldots, D$, define the deterministic functions

$$\begin{aligned} V_j(N, p, \iota, z) &:= 0 && (4.9) \\ V_j(i, p, \iota, z) &:= \max_{u \in \mathcal{A}(t_{j,i}, z, [\mathsf{mAQ}, \mathsf{ACQ}])} \mathbb{E}_{j,i}^{p,\iota} \left[e^{-r t_{j,i}} u \left(p - \iota \right) + \right. \\ &\quad \left. + V_j(i+1, P_{j,i+1}, I_{j,i+1}, z + u) \right] \qquad \forall i < N \quad (4.10) \end{aligned}$$

where $\mathbb{E}_{j,i}^{p,\iota}$ indicates the expectation conditional to $P_{j,i} = p$ and $I_{j,i} = \iota$ (recall that, as these are Markov processes, these values are a sufficient statistics for the whole information up to subperiod i of year j). Then the original problem in Equation (4.8) is brought back to calculating

$$V(0, p_{1,0}, \iota_{1,0}, 0) = \mathbb{E} \left[\sum_{j=1}^{D} V_j(0, P_{j,0}, I_{j,0}, 0) \right]$$

4.2.3 Modeling the make-up clause

This subsection is devoted to the analytical representation of the make-up clause and its constraints. While long term contracts may have a length of 10-30 years, make-up clauses are typically written on a limited period of the contract life, often from 3 to 5 years. Given the fact that, as explained in Section 4.2.2, a contract without make-up clause can be evaluated as the sum of some yearly contract one independent from the other, we can split a contract with make-up written only on a subperiod of the whole contract life in two parts: the first part is a swing contract with a make-up clause with length equal to the original make-up clause, while the other part covers all the years when the make-up is not written. Thus, without loss of generality we can assume that the make-up clause is written on the whole contract's length, D years.

For each year $j = 1, \ldots, D$, call M_j the make-up gas nominated and U_j the make-up gas called back in year j. With this notation, we assume that the precise structure of the make-up clause follows these rules.

1. *For each year $j = 1, \ldots, D-1$, the contract holder is allowed to take $z_{j,N} <$ mAQ, provided $u_{j,i} \geq$ mDQ for all $i = 0, \ldots, N-1$.*

Thus, the make-up gas nominated in year j is

$$M_j := (\mathsf{mAQ} - z_{j,N})^+ \quad \text{and must satisfy} \quad M_j \in [0, \underline{\mathcal{M}}] \qquad (4.11)$$

where $x^+ := \max(x, 0)$ and $\underline{\mathcal{M}}$, defined in Equation (4.4), is also the maximum quantity of make-up gas that can be physically nominated in a given year.

2. *The make-up M_j nominated in year j can be called back in one or more subsequent years (the quantity M_j can be splitted and called back in more than one year). This is possible only if the ACQ quantity has been reached in that year, and of course in that year we still have to satisfy $u_{j,i} \leq$ MDQ for all $i = 1, \ldots, N$.*

Thus, the make-up gas called back in year $j = 2, \ldots, D$ is

$$U_j := (z_{j,N} - \mathsf{ACQ})^+ \quad \text{and is such that} \quad U_j \in [0, \overline{\mathcal{M}}] \tag{4.12}$$

where $\overline{\mathcal{M}}$, defined in Equation (4.3), is also the maximum quantity of make-up gas that can physically called back in a given year.

3. *It is not possible to call back make-up gas before having nominated it, and at year D all the nominated make-up gas must have been called back.*

Thus, if we define the cumulated gas debt at year j, i.e. the make-up gas not yet called back, as

$$\boldsymbol{M}_j = \sum_{k=1}^{j} M_k - \sum_{k=2}^{j} U_k = \sum_{k=1}^{j} (M_k - U_k), \tag{4.13}$$

then $U_1 = M_D = 0$, $\boldsymbol{M}_j \geqslant 0$ for all $j = 1, \ldots, D-1$ and $\boldsymbol{M}_D = 0$. Moreover,

$$\boldsymbol{M}_{j+1} = \boldsymbol{M}_j + M_j - U_j = \boldsymbol{M}_j + (\mathsf{mAQ} - z_{j,N})^+ - (z_{j,N} - \mathsf{ACQ})^+$$

Notice that conditions 2. and 3. imply, for example, that if at the beginning of the last contract year of the make-up clause we have some make-up gas not called back, i.e. $\boldsymbol{M}_{D-1} > 0$, in year d we necessarily have to reach the quantity $\mathsf{ACQ} + \boldsymbol{M}_{D-1}$.

Remark 7 More in general, for all years $j = 1, \ldots, D$, the definition of $\boldsymbol{M}_j$ implies that $\boldsymbol{M}_j \leq j \cdot \underline{\mathcal{M}}$ and $\boldsymbol{M}_j \leq (D - j) \cdot \overline{\mathcal{M}}$. By combining these two constraints, the maximum gas debt is possible at year

$$\bar{j} := \frac{D\overline{\mathcal{M}}}{\underline{\mathcal{M}} + \overline{\mathcal{M}}} \tag{4.14}$$

if $\bar{j}$ is integer, and at one of the two nearest years if $\bar{j}$ is not integer. In particular, the gas debt $\boldsymbol{M}_j$ can increase without constraints for $j < \bar{j}$ and must possibly be decreased for $j > \bar{j}$.

4. *The price of the make-up quantity nominated in year j and called back in year k, subperiod i, is defined as the weighted sum of two components respectively paid at two different times:*

   a) at time $t_{j,N}$ (i.e. at the end of year j when M_j becomes known) the buyer pays the make-up gas at the price $\alpha\overline{\Gamma}_j$ for some $\alpha \in (0,1)$ defined in the contract, where $\overline{\Gamma}_j$ is the average index price observed in year j;

   b) at time of withdrawal $t_{k,i}$, the price paid is $(1-\alpha)I_{k,i}$.

The price of make-up gas, as defined above, is associated to the gas volume $u_{k,i}$ physically delivered at time $t_{k,i}$. This means that the part $\alpha\overline{\Gamma}_j$ in (a) of the price needs to be capitalized from time $T_j = t_{j,N}$ up to time $t_{k,i}$: thus, the price $\mathcal{I}_{j,k,i}$ at time $t_{k,i}$ of the make-up gas nominated in year j and called back in year k, month i is

$$\mathcal{I}_{j,k,i} = \alpha\overline{\Gamma}_j e^{r\left(t_{k,i}-t_{j,N}\right)} + (1-\alpha)I_{k,i} \tag{4.15}$$

By discounting at time $T_0 = 0$ the price of make-up gas called back at time $t_{k,i}$, we have

$$e^{-rt_{k,i}}\mathcal{I}_{j,k,i} = \alpha\overline{\Gamma}_j e^{-rt_{j,N}} + (1-\alpha)I_{k,i}e^{-rt_{k,i}}$$

It follows that, in a year $j = 1,\dots,D$ where the make-up clause is exercised to nominate or call back gas, the residual value of the swing contract at month $i = 0,\dots,N-1$ for that year with the control policy $u_j := (u_{j,i})_{i=0,\dots,N-1}$ is given by

$$J_j(i,p,\iota,z_{j,i};u_j) := \mathbb{E}^{p,\iota}_{j,i}\left[\sum_{k=i}^{N-1} e^{-rt_{j,k}}u_{j,k}\left(P_{j,k} - A_\alpha(u_{j,k},z_{j,k})I_{j,k}\right) - e^{-rt_{j,N}}\alpha\overline{\Gamma}_j M_j\right] \tag{4.16}$$

where

$$A_\alpha(u,z) := 1-\alpha\left(1-\frac{\mathsf{ACQ}-z}{u}\right)^+ \mathbf{1}_{\{\mathsf{ACQ}-\mathsf{MDQ}<z\le\mathsf{ACQ}\}} - \alpha\mathbf{1}_{\{z>\mathsf{ACQ}\}}$$

is a pricing coefficient in the interval $[1-\alpha,1]$ for $I_{j,k}$ to accomodate the gas quality (ordinary below ACQ, called back from previous years above). Here, the main apparent difference with respect to the case when no make-up is exercised is that we can end up a year with a non-null position in the make-up gas, i.e. with $M_j - U_j \neq 0$ (notice that this notation is not ambiguous as M_j and U_j cannot be both different from zero), where the quantity $M_j - U_j$ is by definition a deterministic function of $z_{j,N}$, thus of u_j. Notice that this generalizes the payoff to be maximised in Equation

(4.8), which is reobtained for $z_{j,N} \in [\mathsf{mAQ}, \mathsf{ACQ}]$, i.e. $M_j = U_j = 0$, and setting $z_{j,0} = 0$.

As a result, the total value of the swing option with the make-up clause described above is given by

$$\sup_{u \in \mathcal{A}} \mathbb{E}\left[\sum_{j=1}^{D} J_j(0, p_{j,0}, \iota_{j,0}, 0; u_j)\right] \tag{4.17}$$

where $u = (u_{j,i})_{j=1,\dots,D, i=0,\dots,N-1}$ must now belong to the set

$$\mathcal{A} := \left\{ \begin{array}{l} (u_{j,i})_{j,i} \text{ adapted to } (P_{j,i}, I_{j,i})_{j,i} \quad s.t. \quad u_{j,i} \in [\mathsf{mDQ}, \mathsf{MDQ}], \\ \qquad U_1 = M_D = 0, \boldsymbol{M}_j \geq 0 \; \forall j = 1, \dots, D, \boldsymbol{M}_D = 0 \end{array} \right\}$$

and the functions J_j are given by Equation (4.16). Here, the constraints on U and M induce constraints on $(u_{j,i})_{j,i}$, which will be treated in detail in the next subsection.

4.3 The price of swing contracts with make-up clauses

We just saw that the set of admissible strategies for $(u_{ji})_{ji}$ for a given year $j = 1, \dots, D$ now depend on the gas debt $\boldsymbol{M}_{j-1}$ arriving from the previous years. For this reason, this quantity has to be explicitly taken into account in the evaluation of the swing contract for that year. More in detail, now we define the *value function* as

$$V_j(i, p, \iota, z, \boldsymbol{M}_{j-1}) := \sup_{u \in \mathcal{A}} \mathbb{E}_{j,i}^{p,\iota}\left[J_j(i, p, \iota, z; u_j) + \sum_{k=j+1}^{D} J_k(0, P_{k,0}, I_{k,0}, 0; u_k)\right] \tag{4.18}$$

We now build a Dynamic Programming algorithm as in [4, 26]: for each year $j = 1, \dots, D$, define the deterministic functions:

- if $j = D$, then $M_D = 0$ and $U_D = \boldsymbol{M}_{D-1}$ (recall that $\boldsymbol{M}_D \equiv 0$), so we let

$$\underline{z} := \mathsf{mAQ}\mathbf{1}_{\{\boldsymbol{M}_{D-1}=0\}} + (\mathsf{ACQ} + \boldsymbol{M}_{D-1})\mathbf{1}_{\{\boldsymbol{M}_{D-1}>0\}}, \qquad \overline{z} := \mathsf{ACQ} + \boldsymbol{M}_{D-1} \tag{4.19}$$

and define

$$V_j(N,p,\iota,z,\boldsymbol{M}_{j-1}) := 0, \tag{4.20}$$

$$V_j(i,p,\iota,z,\boldsymbol{M}_{j-1}) := \max_{u\in\mathcal{A}(t_{j,i},z,[\underline{z},\overline{z}])} \mathbb{E}_{j,i}^{p,\iota}\Big[e^{-rt_{j,i}}u\,(p-A_\alpha(u,z)\iota)+ \\ +V_j(i+1,P_{j,i+1},I_{j,i+1},z+u,\boldsymbol{M}_{j-1})\Big] \tag{4.21}$$

Notice that the functions V_j depend on $\boldsymbol{M}_{j-1}$ through $[\underline{z},\overline{z}]$.

- for $j=1,\dots,D-1$ the key quantity is $\boldsymbol{M}_{j-1}$ which is known at the beginning of the year. Assume that $\boldsymbol{M}_{j-1}\le(D-j+1)\overline{\mathcal{M}}$. For the lower bound we have two cases.
  - If $\boldsymbol{M}_{j-1}$ is admissible also for year j, i.e. if $\boldsymbol{M}_{j-1}\le(D-j)\overline{\mathcal{M}}$, then we can nominate some other make-up gas M_j as long as

    $$\boldsymbol{M}_j=\boldsymbol{M}_{j-1}+M_j\le(D-j)\overline{\mathcal{M}}\Rightarrow M_j\le(D-j)\overline{\mathcal{M}}-\boldsymbol{M}_{j-1}$$

    Taking into account the mDQ constraints, the lower bound for $z_{j,N}$ is:

    $$\underline{z}:=\mathsf{mAQ}-\min\left\{(D-j)\overline{\mathcal{M}}-\boldsymbol{M}_{j-1},\underline{\mathcal{M}}\right\} \tag{4.22}$$

  - If $\boldsymbol{M}_{j-1}$ is not admissible for year j, i.e. if $\boldsymbol{M}_{j-1}>(D-j)\overline{\mathcal{M}}$, then we must call back some make-up gas in order to obtain a final cumulated quantity $\boldsymbol{M}_j$ admissible for year j, i.e.

    $$\boldsymbol{M}_j=\boldsymbol{M}_{j-1}-U_j\le(D-j)\overline{\mathcal{M}}\Rightarrow U_j\ge\boldsymbol{M}_{j-1}-(D-j)\overline{\mathcal{M}}$$

    So the lower bound for $z_{j,N}$ is now

    $$\underline{z}:=\mathsf{ACQ}+\boldsymbol{M}_{j-1}-(D-j)\overline{\mathcal{M}} \tag{4.23}$$

For the upper bound $\overline{z}$, we do not need to distinguish between the two previous cases and let

$$\overline{z}:=\mathsf{ACQ}+\min\left\{\boldsymbol{M}_{j-1},\overline{\mathcal{M}}\right\} \tag{4.24}$$

Finally, define

$$V_j(N,p,\iota,z,\boldsymbol{M}_{j-1}) := V_{j+1}(0,p,\iota,0,\boldsymbol{M}_{j-1}+(\mathsf{mAQ}-z)^+-(z-\mathsf{ACQ})^+)+ \\ +e^{-rt_{j,N}}\alpha\Gamma_j(\mathsf{mAQ}-z)^+, \tag{4.25}$$

and for $i=N-1,\dots,0$ define $V_j(i,\cdot,\cdot,\cdot,\cdot)$ exactly as in Equation (4.21).

Remark 8 Thanks to the fact that we are assuming the admissibility of $\boldsymbol{M}_{j-1}$ for year $j-1$, we can not take care about the possibility of having an infeasible problem with $z_{j,N} > \mathsf{MDQ} \cdot N$. In fact, using the admissibility and Equation (4.3), we obtain:

$$\begin{aligned} & \mathsf{ACQ} + \boldsymbol{M}_{j-1} - (D-j)\overline{\mathcal{M}} \leq \\ \leq\ & \mathsf{ACQ} + (D-j+1)\overline{\mathcal{M}} - (D-j)\overline{\mathcal{M}} = \\ =\ & \mathsf{ACQ} + \overline{\mathcal{M}} = \mathsf{MDQ} \cdot N \end{aligned}$$

As a consequence, the final cumulated make-up quantity $\boldsymbol{M}_j$ defined as

$$\boldsymbol{M}_j = \boldsymbol{M}_{j-1} + M_j - U_j \tag{4.26}$$

is admissible for year j.

The following theorem assures the theoretical validity of the algorithm presented so far, proving that it can find out exactly the contract price as defined in (4.17). The second interesting point of the theorem is the fact that, under some conditions, we can restrict our attention only to the so called *bang-bang* controls.

Theorem 11 *The following hold:*

1. *The deterministic functions $V_j(\cdot,\cdot,\cdot,\cdot,\cdot)$, defined by the dynamic programming equations (4.20), (4.21) and (4.25) are such that $V_1(0, P_{1,0}, I_{1,0}, 0, 0)$ coincides with the value of the swing option with the make-up clause in Equation (4.17).*

2. *There exists an optimal Markovian consumption $u^*_{j,i} = u(t_{j,i}, P_{j,i}, I_{j,i}, z_{j,i}, \boldsymbol{M}_{j-1})$, where $u(\cdot,\cdot,\cdot,\cdot,\cdot)$ is given by the maximum argument in the dynamic programming equation (4.21).*

3. *If the quantities*

$$\underline{K} := \frac{\underline{\mathcal{M}}}{\mathsf{MDQ} - \mathsf{mDQ}} \qquad \textit{and} \qquad \overline{K} := \frac{\overline{\mathcal{M}}}{\mathsf{MDQ} - \mathsf{mDQ}}, \tag{4.27}$$

are integer, then there exists an optimal bang-bang Markovian consumption $u^*_{j,i}$, i.e. $u^*_{j,i} = \mathsf{mDQ}$ or $u^*_{j,i} = \mathsf{MDQ}$ for all $j = 1, \ldots, D$, $i = 0, \ldots, N-1$. Moreover, $\boldsymbol{M}_j$ turns out to be an integer multiple of $\mathsf{MDQ} - \mathsf{mDQ}$ for all $j = 1, \ldots, d$.

Proof We proceed in analogy with [4] and [2].

1. As $0 \leq t_{j,i} \leq T$, $0 \leq z \leq N \cdot \mathsf{MDQ}$, $\boldsymbol{M} \leq D \cdot \underline{\mathcal{M}}$ and $\mathbb{E}[P_{j,i}] = F^P_{j,i} < +\infty$, $\mathbb{E}[I_{j,i}] = F^I_{j,i} < +\infty$, then the assumptions ($\mathrm{F}^+$, F^-) in [7, Proposition 8.5] are satistied, so the argument follows.

2. The right-hand side of Equation (4.21) is continuous in u and $\mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}])$ is a compact set contained in $[\mathsf{mDQ}, \mathsf{MDQ}]$, thus the maximum is attained for $u \in \mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}])$ again by applying [7, Proposition 8.5].

3. As in [2], it can be proved that the functions $V_j(i, \cdot, \cdot, \cdot, \cdot)$, $j = 1, \ldots, D$, $i = 0, \ldots, N-1$ are continuous and concave on z and piecewise affine on the intervals

$$[k \cdot \mathsf{mDQ} + (i-k) \cdot \mathsf{MDQ}, (k+1) \cdot \mathsf{mDQ} + (i-k-1) \cdot \mathsf{MDQ}], \quad k = 1, \ldots, i. \tag{4.28}$$

We now prove the claim by induction on $j = 1, \ldots, D$. If $j = 1$, then $\boldsymbol{M}_0 = 0$ by definition. For a given year $j = 1, \ldots, D$, assume for now that $\boldsymbol{M}_{j-1}$ is an integer multiple of $\mathsf{MDQ} - \mathsf{mDQ}$. Then this, together with the condition $\underline{K}, \overline{K} \in \mathbb{N}$ ensures that $[r_{\min}(t_{j,i}, \underline{z}), r_{\max}(t_{j,i}, \bar{z})]$ is exactly the union of suitable intervals of the kind of Equation (4.28). Thus, if $z = k \cdot \mathsf{mDQ} + (i-k) \cdot \mathsf{MDQ}$ for some $k = 0, \ldots, i$, then the function to be maximised in Equation (4.21) is affine on u, thus its maximum point is $u^*_{j,i} = \mathsf{mDQ}$ or $u^*_{j,i} = \mathsf{MDQ}$. It can then be proved by induction that, since $z_{j,0} = 0$, the optimal $u^*_{j,i}$ is such that $z_{j,i} = k \cdot \mathsf{mDQ} + (i-k) \cdot \mathsf{MDQ}$ for some $k = 0, \ldots, i$: this also implies that $\boldsymbol{M}_j$ will be also an integer multiple of $\mathsf{MDQ} - \mathsf{mDQ}$, and the conclusion follows.

□

Remark 9 Part 3. of the theorem above is essentially a consequence of the linear structure of the payoff function in Equations (4.16–4.17): the result is that in every year j, subperiod i, the optimal quantity $u_{j,i}$ can be safely chosen to be either the maximum (MDQ) or the minimum (mDQ) admissible for that substep. This kind of control is called of *bang-bang* type, and it was already found in [4] with smoother payoffs, and studied in deep detail in [2]. Qualitatively, this is due to the fact that, if the withdrawal is profitable in the subperiod, then the better choice is the maximum quantity we can take; conversely, if the withdrawal is not profitable, then the better choice is to take the minimum quantity we can.

From a computational point of view, this point is very interesting because permits to overcome the numerical maximization in the HJB equation by using a binomial tree on the cumulated quantity, as done in next Section 4.5, where at each time step i the feasible optimal cumulated quantity z_i can be either $z_{i-1} + \mathsf{mDQ}$ or $z_{i-1} + \mathsf{MDQ}$.

4.3.1 Computational cost

As we have seen, the pricing problem for a swing option with make-up clause boils down to maximize the problem in Equation (4.18). Unfortunately, this maximization cannot be carried out by analytic means, as a

closed form for V_j is not known even in the simplest case of a standard swing option without make-up clause. Thus, this maximization must be carried out via numerical methods.

The most efficient way to do this is to assume that the quantities $\underline{K}$ and $\overline{K}$ in Equation (4.27) are integer, so that the results of Theorem 11 hold. This induces a quantization in the candidate optimal make-up gas debt $(\boldsymbol{M}_j)_{j=1,\dots,D}$: in fact, since this process at optimality has values which are multiple integers of $\mathsf{MDQ} - \mathsf{mDQ}$, we obtain that the resulting candidate optimal quantities for $\boldsymbol{M}_j$, $j = 1, \dots, D$, are a finite number. More in detail, the sequence $(\boldsymbol{M}_j)_{j=1,\dots,D}$ is bound to have a finite number of nonnegative values in each year $j = 1, \dots, D-1$, with $-\overline{\mathcal{M}} \leq \boldsymbol{M}_j - \boldsymbol{M}_{j-1} \leq \underline{\mathcal{M}}$, i.e. the increments can have at most $\underline{K} + \overline{K} + 1$ distinct values, corresponding respectively to the cases when $M_j > 0$, $U_j > 0$ and $M_j = U_j = 0$.

With this in mind, we can calculate the computational cost needed to price a D-year swing option with make-up clause, and we do this by the same backward recursion used in the Dynamic Programming algorithm used in Section 2.4. In the D-th year, we can start with $\boldsymbol{M}_{D-1}$ taking at most $\overline{K} + 1$ different values, each one of this leading to a different optimization problem: since also the values of $(z_{j,i})_{j,i}$ are quantized via the bang-bang optimal process $(u_{j,i})_{j,i}$, for each one of this optimization problem we have a total of $O(N^2)$ states which can be assumed by $(z_{j,i})_{j,i}$ at optimality[1]. Having solved the $\overline{K} + 1$ problems for the last year, we can attach the value functions thus obtained to the terminal nodes of year $D-1$: notice that also in this case, as now $\boldsymbol{M}_{D-2}$ can assume at most $2\overline{K} + 1$ distinct values, we will have to model and solve at most $2\overline{K} + 1$ distinct optimization problem, each one having as terminal condition the values of the $\overline{K} + 1$ problems of year D. These numbers do not multiply, because once that we obtain the values for the $\overline{K} + 1$ problems of year D for each possible starting state $(P_{D-1,N}, I_{D-1,N}) = (P_{D,0}, I_{D,0})$, we can take these values as terminal values to use in the computation for year $D-1$.

With this spirit, we are now ready for a result on the computational cost of the pricing of a swing option with make-up clause.

Theorem 12 *If the quantities $\underline{K}$ and $\overline{K}$, defined in Equation (4.27), are integer, then the order of distinct subproblems to be solved is $O(N^2 D^2)$.*

Proof First of all consider the 2-dimensional process $(j, \boldsymbol{M}_j)_{j=1,\dots,D}$: then the distinct states that this process can assume, at optimality, is in 1-1 correspondence with the integer solutions (x, y) of the system

$$\begin{cases} x & \geq & 0, \\ x & \leq & \overline{K}(D-y). \\ x & \leq & \underline{K}y, \end{cases} \tag{4.29}$$

[1] precisely $\leq \frac{N(N+1)}{2}$, which is the number of nodes of a complete recombining binomial tree.

In fact, if (x, y) is such a solution, then $\boldsymbol{M}_x = (\mathsf{MDQ} - \mathsf{mDQ})y$ is a possible value, at time x, of an optimal path for $\boldsymbol{M}$ by Theorem 11. Conversely, by the same theorem, if $\underline{K}$ and $\overline{K}$ are integer then $\boldsymbol{M}_x$ is a integer multiple of $\mathsf{MDQ} - \mathsf{mDQ}$. Since for each of these possible states we must solve a separate optimization problem for the corresponding year, the number of optimization subproblems for all the values of $(z_{j,i})_{j,i}$ are of order $O(N^2)$, and their total number is the sum of these, the proof boils down to find the total number $\mathbf{N}$ of integer solutions of the system (4.29). By recalling the definition of $\bar{j}$ in Equation (4.14) and the discussion below, first of all we rewrite $\bar{j}$ as

$$\bar{j} = \frac{D\overline{K}}{\underline{K} + \overline{K}}$$

Now, the region of the solutions of the system (4.29) is the union of the two triangular regions $\{(x, y) \in \mathbb{N}^2 \mid x \geq 0, x \leq \underline{K}y, y \leq \bar{j}\}$ and $\{(x, y) \in \mathbb{N}^2 \mid x \geq 0, x \leq \overline{K}(D - y), y > \bar{j}\}$. It is then easy to see that

$$\begin{aligned}
\mathbf{N} &= \sum_{\ell=0}^{[\bar{j}]}(1 + \underline{K}\ell) + \sum_{\ell=[\bar{j}]+1}^{D}(1 + \overline{K}(D - \ell)) = D + 1 + \underline{K}\sum_{\ell=1}^{[\bar{j}]}\ell + \overline{K}\sum_{\ell=[\bar{j}]+1}^{D-1}(D - \ell) = \\
&= D + 1 + \underline{K}\frac{[\bar{j}] \cdot [\bar{j} + 1]}{2} + \overline{K}\frac{[D - \bar{j} - 1] \cdot [D - \bar{j}]}{2}
\end{aligned}$$

where $[x]$ denotes the integer part of x. By noticing that for all $x > 0$ we have $D[x] \leq [Dx]$ and that

$$\underline{K} \cdot \bar{j} = \frac{D\underline{K}\overline{K}}{\underline{K} + \overline{K}} = \overline{K}(D - \bar{j})$$

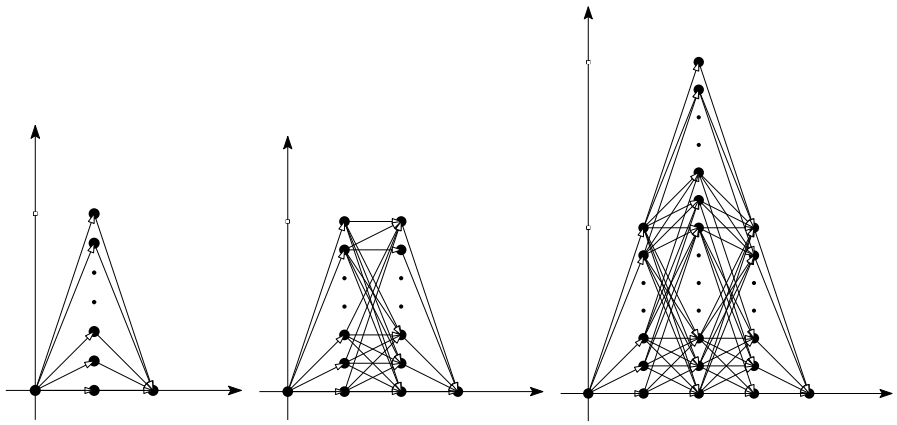
then we have

$$\mathbf{N} \leq D + 1 + \frac{1}{2}\left[\frac{D\underline{K}\overline{K}}{\underline{K} + \overline{K}}\right]([\bar{j} + 1] + [D - \bar{j} - 1]) \leq \frac{1}{2}\frac{\underline{K}\overline{K}}{\underline{K} + \overline{K}}D^2 + D + 1$$

i.e. $\mathbf{N} = O(D^2)$. By multiplying this for $O(N^2)$ (the number of subproblems for given year $j = 1, \ldots, D$ and state of make-up debt $\boldsymbol{M}_j$, we obtain that the computational cost is of order $O(N^2D^2)$, as desired □

We show in Figure 4.3 an illustration of these numbers for $D = 2, 3, 4$, by making the simplifying assumption that $\underline{K} = \overline{K} =: K$.

We also present a numerical test which validates our result of a quadratic cost in the number of years. By taking the same parameters as in Section 5, we implemented the method on a Intel i7 workstation at 3.4GHz with 8GB RAM, with the following execution times.



(a) $D = 2$ (b) $D = 3$ (c) $D = 4$

Figure 4.3: In subfigure (a), we can only obtain $\boldsymbol{M}_1 (= M_1 = U_2)$ among $K + 1$ distinct values, and the orresponding value function is then used in the final values of the optimization problem of year 1, so the total number of optimization problems to be solved is $K + 1$. In subfigure (b) we can obtain $\boldsymbol{M}_2 (= U_3)$ among $K + 1$ distinct values and then $\boldsymbol{M}_1 (= M_1)$ among $K + 1$ distinct values, so the number of optimization problems to be solved in sequence is $2K+2$. In subfigure (c) we can obtain $\boldsymbol{M}_3 (= U_4)$ among $K+1$ different values, $\boldsymbol{M}_2$ among $2K+1$ different values, and finally $\boldsymbol{M}_1 (= M_1)$ among $K + 1$ distinct values: so the number of optimization problems to be solved in sequence is now $4K + 3$.

Duration D of the make-up clause (years)	1	2	3	5	10	15
time (seconds)	0.06	0.12	0.31	1.25	7.60	23.65

Table 4.1: Execution times, in seconds, on a Intel i7 workstation at 3.4GHz with 8GB RAM. Notice that with 1 year there is no possibility *de facto* to exercise the make-up clause.

4.3.2 Extensions of the model

This section is devoted to show some extensions of the algorithm proposed so far in two directions. The first one is towards other forms of make-up clause, while the second one is directed to a possible application to carry-forward clauses. In this section we refer to ethe *genral* algorithm as to the one presented in Section 4.3, but in the following we may introduce new definitions for some quantities defined so far. These definitions will be valid only for this section.

4.3.3 Other forms of make-up clause

While different contracts may have several slightly different definition of the make-up clause, up to authors' knowledge the most negotiated variants to the make-up clause presented in Section 4.2.3 are obtained as in [30] by modifying Rule 2 as follows:

1. the make-up gas must be called back as soon as mAQ has been reached instead of after having exceeded ACQ. In this case every time that at the beginning of the year the cumulated make-up quantity $\boldsymbol{M}_{j-1}$ is positive then *all* the quantity exceeding mAQ is considered to be make-up gas called back;

2. the make-up nominated in year j can be called back only in (one or more) the subsequent ξ years, and not until the end of the contract. In this case, the nominated make-up of year j is lost if not called back before year $j + \xi + 1$. Notice that at the end of year j we have also paid this quantity with the sum $M_j \alpha \bar{\Gamma}_j$, so in this case this amount becomes a sort of penalty.

First case. Because we can call back the make-up as soon as mAQ is reached, for this case we suppose that the quantity ACQ can not be overloaded, i.e. we can never reach at the end of the year the quantity MDQ $\cdot N$. The case when also MDQ $\cdot N$ can be reached is analogous. The maximum make-up quantity we are allowed to call back every year is now

$$\overline{\mathcal{M}} = \mathsf{ACQ} - \mathsf{mAQ}$$

and again we can call it back only if $\boldsymbol{M}_{j-1} > 0$.

The algorithm for this specification of the make-up clause is similar to the one presented in Section (4.3), provided we redefine some of the quantities. Precisely, redefine

$$U_j := \mathbf{1}_{\{\boldsymbol{M}_{j-1}>0\}} \left(z_{j,N} - \mathsf{mAQ}\right)$$

which must be such that $U_j \in [0, \overline{\mathcal{M}}]$, where the maximum make-up quantity which can be called back in a given year is now defined as $\overline{\mathcal{M}} := \mathsf{ACQ} - \mathsf{mAQ}$. This of course modifies the payoff in Equation (4.16), which now becomes

$$J_j(i, p, \iota, z_{j,i}; u_j) := \mathbb{E}^{p,\iota}_{j,i}\left[\sum_{k=i}^{N-1} e^{-rt_{j,k}} u_{j,k} \left(P_{j,k} - \bar{A}_\alpha(u_{j,k}, z_{j,k}) I_{j,k}\right) - e^{-rt_{j,N}} \alpha \overline{\Gamma}_j M_j\right]$$

where

$$\bar{A}_\alpha(u, z) := 1 - \alpha \left(1 - \frac{\mathsf{mAQ} - z}{u}\right)^+ \mathbf{1}_{\{\mathsf{mAQ}-\mathsf{MDQ}<z\leq \mathsf{mAQ}, \boldsymbol{M}_{j-1}>0\}} - \alpha \mathbf{1}_{\{z>\mathsf{mAQ}, \boldsymbol{M}_{j-1}>0\}}$$

is a coefficient in $[1-\alpha, 1]$ analogous to A_α of Section 2.3.

The Dynamic Programming algorithm has to be modified as follows. Substitute Equation (4.19) with

$$\underline{z} := \mathsf{mAQ} + \boldsymbol{M}_{D-1}, \qquad \overline{z} := \mathsf{ACQ}\mathbf{1}_{\{\boldsymbol{M}_{D-1}=0\}} + (\mathsf{mAQ} + \boldsymbol{M}_{D-1})\mathbf{1}_{\{\boldsymbol{M}_{D-1}>0\}},$$

Equation (4.21) with

$$V_j(i, p, \iota, z, \boldsymbol{M}_{j-1}) := \max_{u \in \mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}])} \mathbb{E}_{j,i}^{p,\iota} \Big[e^{-rt_{j,i}} u(p - \bar{A}_\alpha(u, z)\iota) + \\ + V_j(i+1, P_{j,i+1}, I_{j,i+1}, z+u, \boldsymbol{M}_{j-1}) \Big],$$

Equation (4.23) with $\underline{z} := \mathsf{mAQ} + \boldsymbol{M}_{j-1} - (D-j)\overline{\mathcal{M}}$, Equation (4.24) with $\overline{z} := \mathsf{ACQ}$, and finally Equation (4.25) with

$$V_j(N, p, \iota, z, \boldsymbol{M}_{j-1}) := \\ V_{j+1}\Big(0, p, \iota, 0, \boldsymbol{M}_{j-1} + (\mathsf{mAQ} - z)^+ - (z - \mathsf{mAQ})^+ \mathbf{1}_{\{\boldsymbol{M}_{j-1}>0\}}\Big) + \\ + e^{-rt_{j,N}} \alpha \Gamma_j (\mathsf{mAQ} - z)^+.$$

With these substitutions, Theorem 11 still applies to this case.

Second case. In this case the key quantity $\boldsymbol{M}_{j-1}$ at the beginning of every year is no more sufficient. In each year we have to keep memory of all the past nominated quantities. The recursive algorithm proceeds as follows. Suppose that at the beginning of year j we know from the previous years the following vector:

$$\mathbb{x}_{j-1} = [x_{j-\xi}, x_{j-\xi+1}, \ldots, x_{j-1}] \tag{4.30}$$

This vector represents the *residual cumulated quantity* of make-up nominated at every year. How to obtain this vector will be clear at the end of this section, with formula (4.31). Given this vector, the cumulated make-up quantity not yet called back (useful to generate the tree) $\boldsymbol{M}_{j-1}$ is the sum of the components of the vector $\mathbb{x}_{j-1}$, which is the scalar product:

$$\boldsymbol{M}_{j-1} = \mathbb{x}_j \cdot \mathbb{1} = \sum_{\nu=1}^{\xi} x_{j-\nu}$$

With $\boldsymbol{M}_{j-1}$ we can proceed to generate the new trees quantity for year j as described in the general algorithm. Now remember that, in the general algorithm, for each final node of this tree we have a couple (M_j, U_j) as defined in Equations (4.11) and (4.3). The modified algorithm is obtained by substituting these couple of quantities with the new vector $\mathbb{x}_j$ obtained as a function of the original (M_j, U_j):

- if $M_j > 0$ we are nominating some new make-up and we loose the make-up gas nominated in year $j - \xi$. The vector $\mathbb{x}_j$ is then simply given by:
$$\mathbb{x}_j = [x_{j-\xi+1}, x_{j-\xi+2}, \ldots, x_{j-1}, M_j]$$

- if $U_j > 0$ then we are calling back some make-up. If we not call back all the quantity $x_{j-\xi}$ we will loose the (positive) quantity $(x_{j-\xi} - U_j)$. If we are calling back more that the quantity $x_{j-\xi}$ we have to update also the other residual nominated make-up quantities. By defining $\mathbb{z} = (z_1, \ldots z_k) \in \mathbb{R}^k$ and $\forall u \in \mathbb{R}$ the vector function $g_u : \mathbb{R}^k \to \mathbb{R}^k$ defined component-wise recursively by $g_u(\mathbb{z}) = (\bar{z}_1, \ldots, \bar{z}_k)$ with:
$$\begin{cases} \bar{z}_1 = z_1 \\ \bar{z}_n = z_n + \min\{\bar{z}_{n-1} - u, 0\} \end{cases}$$
we have that the new vector of residual cumulated make-up quantity can be given by:
$$\mathbb{x}_j = g_{U_j}(\mathbb{x}) \tag{4.31}$$
With this new vector we can start the algorithm for year $j + 1$.

4.3.4 Carry forward clauses

The approach used in this paper can also price another clause related to swing contracts, namely some instances of the *carry-forward* clause as described in [22]. In general, the carry-forward (CF) right gives the holder of the option the possibility to reduce mAQ (up to a contractual amount ACF called annual carry-forward) in one year if in at least one of the previous d years the total volume taken was above mAQ, while the maximum quantity which can be taken every year remains ACQ.

With a slight modification of the algorithm presented in this paper we are able to price two particular instances of the carry-forward clause presented in [22], namely when the CF rights must be used in the following year (corresponding to $d = 1$, and when CF rights do not have a deadline, this clause being called unlimited duration carry-forward (UDCF): in this latter case there is a maximum on the CF rights one can obtain in a given year, given by $\nu \cdot \mathsf{ACF}$, where $\nu > 1$ is a contractual constant.

We now present how to price a CF clause with $d = 1$. First of all, define the gas credit U_j at year $j = 1, \ldots, D$ as
$$U_j := (z_{j,12} - \mathsf{mAQ})^+$$
and $U_0 := 0$, and the minimum and maximum cumulated quantities for year $j = 1, \ldots, D$ as
$$\underline{z} := \mathsf{mAQ} - \min(U_{j-1}, \mathsf{ACF}), \qquad \overline{z} := \mathsf{ACQ} \tag{4.32}$$

Then the Dynamic Programming algorithm can be built as follows: for the last year D define

$$V_D(N, p, \iota, z, U_{D-1}) := 0, \tag{4.33}$$

$$V_D(i, p, \iota, z, U_{D-1}) := \max_{u \in \mathcal{A}(t_{j,i}, z, [\underline{z}, \overline{z}])} \mathbb{E}_{j,i}^{p,\iota} \Big[e^{-r t_{j,i}} u\,(p - \iota) + \\ + V_D(i+1, P_{j,i+1}, I_{j,i+1}, z + u, U_{D-1}) \Big] \tag{4.34}$$

while for $j = 1, \ldots, D-1$ define

$$V_j(N, p, \iota, z, U_{j-1}) := V_{j+1}(0, p, \iota, 0, (z - \mathsf{mAQ})^+), \tag{4.35}$$

and for $i = N-1, \ldots, 0$ define $V_j(i, \cdot, \cdot, \cdot, \cdot)$ exactly as in Equation (4.34). Then $V_1(0, p, \iota, 0, 0)$ gives exactly the value of the CF contract with $d = 1$.

We now present how to price a UDCF contract. First of all, redefine the cumulated gas credit U_j at year $j = 1, \ldots, D$ as

$$U_j := U_{j-1} + \min(\nu \cdot \mathsf{ACF}, z_{j,12} - \mathsf{mAQ})$$

with $U_0 := 0$: with this new definition, the gas credit U_j at the end of year j increases only if we exceed mAQ and decreases every time we do not reach mAQ. Redefine also the minimum and maximum cumulated quantities for year $j = 1, \ldots, D$ as in Equation (4.32). Then the Dynamic Programming algorithm can be built as follows: for the last year D define the value function V_D exactly as in Equations (4.33–4.34), while for $j = 1, \ldots, D-1$ define

$$V_j(N, p, \iota, z, U_{j-1}) := V_{j+1}(0, p, \iota, 0, U_{j-1} + \min(\nu \cdot \mathsf{ACF}, z_{j,12} - \mathsf{mAQ})), \tag{4.36}$$

and for $i = N-1, \ldots, 0$ define $V_j(i, \cdot, \cdot, \cdot, \cdot)$ exactly as in Equation (4.34). Then $V_1(0, p, \iota, 0, 0)$ gives now exactly the value of the UDCF contract.

Remark 10 In principle, it would be possible to price also CF clauses with $d > 1$ by introducing other state variables, but this goes beyond the scope of this paper, which is focused on make-up clauses. Also, here the computational burden could be reduced by modelling the difference process $P - I = (P_{j,i} - I_{j,i})_{j,i}$ as a single state variable as in [22]. Unfortunately, this is not possible with the make-up clause because of the delayed payment structure of the make-up gas.

4.4 Tree Prices

As in Chapter 3, we assume that the log-prices $\mathcal{P}_{j,i} := \log P_{j,i}$ and $\mathcal{I}_{j,i} := \log I_{j,i}$ follow the discretized version of the mean-reverting dynamics

$$d\mathcal{P}_t = \left(\theta_t^P - a^P X_t\right) dt + \sigma^P dW_t^P$$
$$d\mathcal{I}_t = \left(\theta_t^I - a^I Y_t\right) dt + \sigma^I dW_t^I$$

where W_t^P and W_t^I are two Brownian motions with mutual correlation ρ: these processes are particular cases of the model in [38] and are rather standard models for energy prices (see for example [23, Chapter 23.3].

In the discretized version, both $X_{j,i}$ and $Y_{j,i}$ change at the beginning of every sub-period (i.e. at the beginning of every month). This is exactly what happens for the index I, and it is an acceptable simplification for the gas price P. In particular, we discretize the prices $(P_{j,i})_{j,i}$ and $(I_{j,i})_{j,i}$ by building two trinomial trees with the procedure explained in [13, 23] and here summarized.

The first step is to build trinomial trees for X and Y by discretizing the dynamics of processes

$$dX_t^* = -aX_t^* dt + \sigma dW_t, \quad X_0^* = 0 \tag{4.37}$$

with $(a, \sigma) = (a^P, \sigma^P)$, or $(a, \sigma) = (a^I, \sigma^I)$ in the analogous specification for Y^*. The trees for these processes are symmetric around 0 and their nodes are evenly spaced in time and value at intervals of predetermined length Δt and $\Delta X^* = \sigma\sqrt{3\Delta t}$.

As usual, we denote by (i, j) the node $x_{i,j}$ in the tree for which $x_{i,j} = X_t^*$ with $t = i\Delta t$ and $X_{i\Delta t}^* = j\Delta X^*$[2]. Hull and White proved [24, 25] that the probabilities to switch from node (i, j) to node $(i + 1, k)$ are always nonnegative if $-\bar{j} \leqslant j \leqslant \bar{j}$, where $\bar{j}$ is the smallest integer greater than $0.184/(a\Delta t)$. This means that at every time step $i = 0, \ldots, N$ we have a finite number of nodes (i, j) placed at points $j\Delta X^*$ for every integer $j \in \{-j^*, \ldots, 0, \ldots, j^*\}$, with $j^* := \min\{\bar{j}, 2i - 1\}$. Thus, the total width of the tree depends on a, σ and Δt.

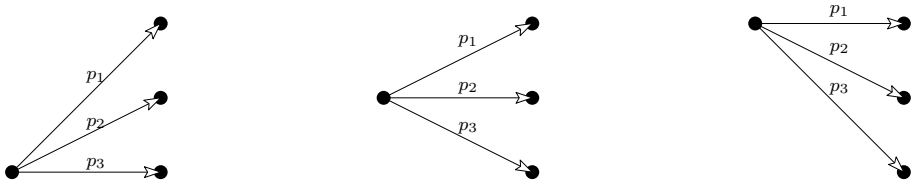


Figure 4.4: Possible branches for a tree with mean reversion.

There are three possible form of branches for a tree that include mean reversion; they are presented in Figure 4.4. The first form, type (a), (up two/up one/straigth alone) is used at the nodes such that $j = -\bar{j}$; the third form, type (c), (down two/down one/straigth alone) is used at the nodes such that $j = \bar{j}$; the standard one, type (b), (up/straight alone/down) is the one used in all the other cases, i.e. whenever $j \neq \pm\bar{j}$.

[2]Notice that in this Section the notation (i, j) is not referred to the notation "year j, month i" used until now in the thesis. Here we not distinguish between year and months, having a unique time index i that varies between 0 and $N \cdot D$. However, for sake of notation, in this section we suppose that $i = 0, \ldots, N$, being N the appropriate number.

	Branch type (a)	Branch type (b)	Branch type (c)
p_1	$\frac{1}{6} + \frac{a^2j^2\Delta t^2 + aj\Delta t}{2}$	$\frac{1}{6} + \frac{a^2j^2\Delta t^2 - aj\Delta t}{2}$	$\frac{7}{6} + \frac{a^2j^2\Delta t^2 - 3aj\Delta t}{2}$
p_2	$-\frac{1}{3} - a^2j^2\Delta t^2 - 2aj\Delta t$	$-\frac{2}{3} - a^2j^2\Delta t^2$	$-\frac{1}{3} - a^2j^2\Delta t^2 + 2aj\Delta t$
p_3	$\frac{7}{6} + \frac{a^2j^2\Delta t^2 + 3aj\Delta t}{2}$	$\frac{1}{6} + \frac{a^2j^2\Delta t^2 + aj\Delta t}{2}$	$\frac{1}{6} + \frac{a^2j^2\Delta t^2 - aj\Delta t}{2}$

Table 4.2: Probabilities on the edges for every type of branch

For every kind of branch there are different probabilities p_1, p_2, p_3 on the edges. The pedices $1, 2, 3$ refers to a clockwise enumeration of the of edges as in Figure 4.4. The probabilities are obtained matching the theoretical conditional expectation and variance and for every kind of branch they are given by the one in Table 4.2. This complete the first step, and leads to a symmetric tree around 0.

The second step is to put together the two trinomial trees in a 2-dimensional tree for (X^*, Y^*): this is done at each node in such a way to preserve the marginal distributions of X^* and Y^* and the covariance structure induced by the correlated Brownian motions W^P and W^I, as in [13, Appendix F] or as in [23].

The third step is aimed to calibrate the previous symmetric tree to the term structure F_i one has, F_i standing for the value of the forward with maturity $i\Delta t$: this step is used to incorporate into the tree mean reversion to levels different from zero, and in particular can be used here to introduce seasonality effects. This is obtained by adding a quantity α_i to the value $x_{i,j}$ of all nodes (i, j). For every step i we have a value for α_i such that:

$$\sum_j Q_{i,j} e^{\alpha_i + x_{i,j}} = F_i$$

that leads to

$$\alpha_i = \log(F_i) - \log\left(\sum_j Q_{i,j} e^{x_{i,j}}\right)$$

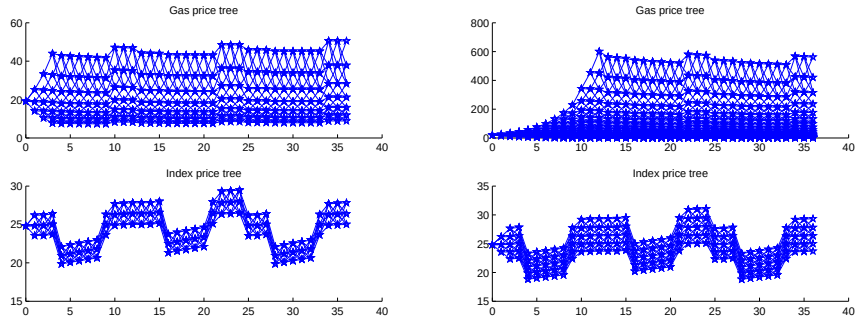
having denoting with $Q_{i,j}$ the probability to reach the node (i, j) starting from the node $(0, 0)$. Once we have the values for α_i we obtain the final tree which has, at step i, the nodes with value $e^{\alpha_i + x_{i,j}}$.

An example of two possible final results for the two trees, obtained for some values of a and σ, is plotted in Figure 4.5. Notice that the higher a^I (or a^P) is, the less nodes the respective tree have.

In order to calibrate for the parameters of Equation (4.37), we use a procedure inspired by [12].

The main idea is to use the discrete time version of the solution of Equation (4.37):

$$X^*(t) = X^*(s)e^{-a(t-s)} + \sigma e^{-at} \int_s^t e^{au} dW_u, \quad 0 \leqslant s < t, \tag{4.38}$$



(a) Strong mean reversion, low volatility: $a^P = 3, a^I = 10, \sigma^P = 0.3, \sigma^I = 0.1$

(b) Small mean reversion, high volatility: $a^P = 0.1, a^I = 0.1, \sigma^P = 0.7, \sigma^I = 0.2$

Figure 4.5: Trees for prices for different values of parameters. Notice that the higher a^I and a^I are, the less nodes the respective trees have: in subfigure (a) we have trees obtained with high a^I, a^P and few nodes in both the trees, while in subfigure (b) we have the converse situation.

which gives

$$x(t_i) = bx(t_{i-1}) + \delta\varepsilon(t_i) \tag{4.39}$$

with

$$b = e^{-a\Delta t}, \qquad \delta = \sigma\sqrt{\frac{1 - e^{-2a\Delta t}}{2a}} \tag{4.40}$$

and ε is a Gaussian white noise ($\varepsilon(t_i) \sim N(0,1)$ for all i). Then, in order to provide the maximum likelihood estimator for the parameters b and δ, perform a least squares regression of the time series $x(t_i)$ on its lagged value $x(t_{i-1})$, as in Equation (4.39). Once we have b and δ, we can invert Equation (4.40) and derive the original parameter a and σ.

4.5 Three years example

In the following we describe and analyze the case of a three years contract, i.e. the case $D = 3$ and $N = 12$. Although, as seen in Subsection 2.5, the complexity of the different kinds of control problems to be solved grows quadratically with D, we have decided to present a three years contract as an example is due to fact that the distinct qualitative combinations of years where we can nominate and/or call back make-up gas, as D grows, lead to more and more intricate combinatorical considerations, which would deviate the attention from the modelization. Thus, while a 2-year contract is not a very interesting example from a modellistic point of view, as at the end of the first year the make-up quantity is known and is exactly the quantity called back in the second year, on the other hand we think that

$D = 3$ gives the right compromise between succeeding to follow exactly what goes on in the different years and significance of the combinatorial problem.

While a longer duration of the make-up clause would not be a computational problem with our approach (see Table 4.1 above), we have to remark that typical make-up clauses are not alive during all the life of the contract, but are typically written on small sub-period spanning from 3 to 5 years. At a first glance, this choice may seem strange considering the fact that swing contracts usually have longer maturity, from 10 to 30 years. The main reason why make-up clauses have significantly shorter duration is that no seller takes the risk of giving the opportunity to move huge quantities of gas between decades. In addition, European gas market illiquidity does not allow to obtain realistic forecast of forward prices. As an example, the longest traded maturity on a very liquid market as the TTF is the 3-year ahead forward, and for longer maturities oil-related instruments are used with much fewer granularity, so the long term gas term structure is pretty flat. In such situation the decoupling is less marked and the make-up clause loses a bit its importance. For these reasons the market practice is that, if need arises for the buyer, a new make-up clause can be renegotiated in future years, so the problem of valuing make-up clauses can be split into separate problems.

Let us now concentrate on our 3-year example. Once the make-up quantity of the first year is known, in the second year there are many opportunities: one can call back some (or all) the make-up of the first year or can nominate, if possible, some other make-up that will be called back in the third year.

4.5.1 Trees quantity

First Year. By Remark 7, we must end the year with

$$0 \leq \boldsymbol{M}_1 = M_1 \leq \min(\underline{\mathcal{M}}, 2\overline{\mathcal{M}})$$

In fact, in the first year the maximum make-up quantity we can nominate has to be less than or equal to the maximum quantity we can call back in the following two years, that is $2\overline{\mathcal{M}}$, as well as to the maximum quantity we can nominate in a single year. In terms of $\underline{z}$ and $\overline{z}$, this means that $\underline{z} = \mathsf{mAQ} - \min(\underline{\mathcal{M}}, 2\overline{\mathcal{M}})$, while $\overline{z} = \mathsf{ACQ}$, because we can not call back any previous year make-up, so we have

$$z_{1,12} \in \left[\mathsf{mAQ} - \min(\underline{\mathcal{M}}, 2\overline{\mathcal{M}}), \mathsf{ACQ}\right]$$

Notice that this is in agreement with Equations (4.22) and (4.24). Figure 4.6 shows the possible actions we can perform in the first year.

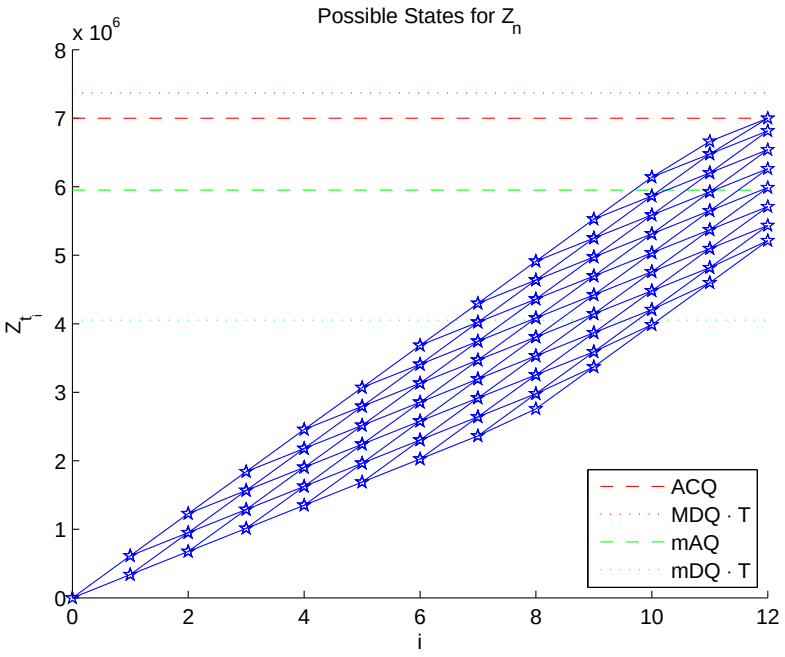


Figure 4.6: Tree quantities for the first year. In the final states where $z_{1,12} < \mathsf{mAQ}$, some make-up is nominated and has to be called back in the subsequent 2 years.

Second Year. Notice that in this year $\boldsymbol{M}_1 = M_1$: this value strongly influences the possible actions we can take in the second year:

- If $(0 \leqslant) \boldsymbol{M}_1 \leqslant \overline{\mathcal{M}}$ we can do one of the following:

  i. nominate some other make-up gas in the second year, in such a way that we are able to call back all the make-up in the third year, that is

  $$\boldsymbol{M}_1 + M_2 \leq \overline{\mathcal{M}} \Rightarrow M_2 \leq \overline{\mathcal{M}} - \boldsymbol{M}_1$$

  ii. call back some make-up gas nominated in the first year: the maximum quantity we can call back is, obviously, $\boldsymbol{M}_1$;

  iii. take a quantity of gas between mAQ and ACQ, not nominating or calling back any make-up gas.

  Summarizing, the constraints for $z_{2,12}$ in this case are

  $$z_{2,12} \in \left[\mathsf{mAQ} - \left(\overline{\mathcal{M}} - \boldsymbol{M}_1\right), \mathsf{ACQ} + \boldsymbol{M}_1\right]$$

  which are again in agreement with Equations (4.22) and (4.24). Figure 4.7(b) shows this case.

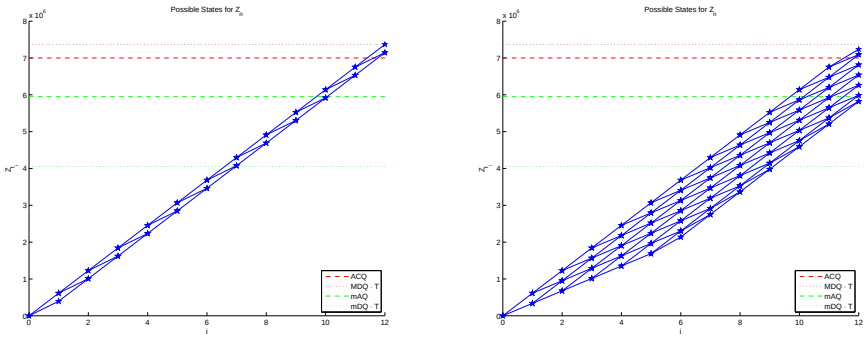
- if $\overline{\mathcal{M}} < \boldsymbol{M}_1 \left(\leqslant 2\overline{\mathcal{M}}\right)$ we must instead call back some make-up gas U_2, otherwise we are not able to arrive in T_3 having called back the whole quantity M_1. In this case the minimum U_2 we have to call back must be such that the final make-up cumulated quantity can be called back in the third year, i.e. $\boldsymbol{M}_2 \leqslant \overline{\mathcal{M}}$, which leads to:

$$\boldsymbol{M}_2 = \boldsymbol{M}_1 - U_2 \leqslant \overline{\mathcal{M}} \Rightarrow U_2 \geqslant \boldsymbol{M}_1 - \overline{\mathcal{M}}$$

Thus, the following constraints for $z_{2,12}$ hold:

$$z_{2,12} \in \left[\mathsf{ACQ} + \boldsymbol{M}_1 - \overline{\mathcal{M}}, \mathsf{ACQ} + \boldsymbol{M}_1\right]$$

which now are in agreement with Equations (4.23) and (4.24). Figure 4.7(a) shows this case.



(a) $\overline{\mathcal{M}} < \boldsymbol{M}_1 \leqslant 2\overline{\mathcal{M}}$ (b) $0 \leqslant M_1 \leq \overline{\mathcal{M}}$

Figure 4.7: Possible tree quantities for the second year. In subfigure (a), the first-year make-up gas is so much that has to be called back in both second and third year: thus, we are forced to end with $z_{2,12} > \mathsf{ACQ}$. In subfigure (b), the first-year make-up gas can be called back in a single year, so we have the choice among calling back some quantity ($z_{2,12} > \mathsf{ACQ}$), respect the constraints ($z_{2,12} \in [\mathsf{mAQ}, \mathsf{ACQ}]$) or nominate some other make-up gas ($z_{2,12} < \mathsf{mAQ}$).

Third year. The key quantity now is the residual make-up we have to call back, if any. This quantity, that is exactly $\boldsymbol{M}_2$ as defined in Equation (4.13), is the residual make-up quantity we must call back in the third year, being this the last year contract (remember we have to call back all the nominated make-up, as seen in Subsection 4.2.3).

So there are two cases for this year:

- if $\boldsymbol{M}_2 > 0$ then we have to call back the whole make-up quantity accumulated in the first two years and we have no choice for $z_{3,12}$:

$$z_{3,12} = \mathsf{ACQ} + \boldsymbol{M}_2$$

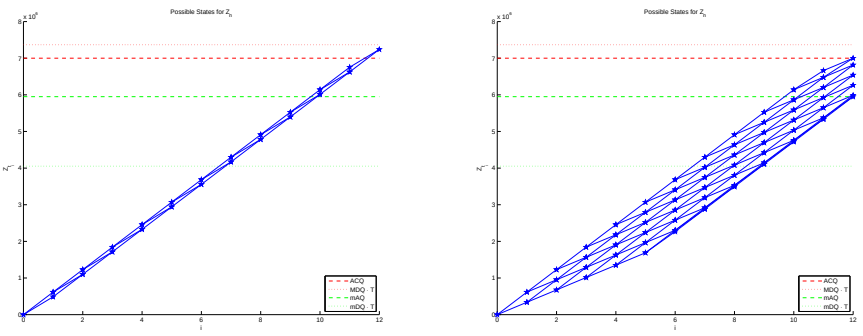
This case is shown in Figure 4.8(a)

- if $\boldsymbol{M}_2 = 0$ then, as we can not nominate any make-up, the constraints for $z_{3,12}$ are given by:

$$z_{3,12} \in [\mathsf{mAQ}, \mathsf{ACQ}]$$

This case is shown in Figure 4.8(b)

Both the cases agree with Equation (4.19).



(a) $\boldsymbol{M}_2 = U_3 > 0$ (b) $\boldsymbol{M}_2 = U_3 = 0$

Figure 4.8: Possible tree quantities for the third year: here the kind of tree totally depends on the cumulated make-up residual quantity $\boldsymbol{M}_2$ from the previous years. If $\boldsymbol{M}_2 > 0$, we are forced to put $U_3 = \boldsymbol{M}_2$ and consume $z_{3,12} = \mathsf{ACQ} + U_3 > \mathsf{ACQ}$, ending up with a tree as subfigure (a). If $\boldsymbol{M}_2 = 0$, we are forced to satisfy the constraints and consume $z_{3,12} \in [\mathsf{mAQ}, \mathsf{ACQ}]$, ending up with the tree in subfigure (b).

4.5.2 Sensitivity analysis of a three years contract with make-up clause

A swing contract is a derivative product whose value depends on two main classes of factor, namely market and volumetric. As previously explained in this paper, this kind of derivative shows an optionality value linke d to the market price dynamics of the underlying commodity (exercise or not) and an optionality value linked to the volumetric structure of the product itself (how much to allocate with the make-up clause among the years and

how much to withdraw in each subperiod). After having explained how to price a swing product on gas and how to determine the optimal exercise policy, it is now interesting to use the algorithm in order to explore and map the value of the contract with respect to some peculiar parameters of the contract and to market factors.

More in detail, we specify a trinomial dynamics for both the price P and the index I which approximates a geometric mean-reverting Ornstein-Uhlenbeck process as described in Section 4.4, and calibrate these models following [12], using historical data on TTF prices for the gas price P and the ENIGR07 formula[3] for the index price I. For ease of implementation, the average index price $\overline{\Gamma}_j$ of year j which appears in Equation (4.25) is substituted with the average of forward prices for that year. When not variable, the parameters used in this section are the ones in Table 4.3.

Parameter		*Value*	*Parameter*		*Value*
ACQ	=	$7.00 \cdot 10^6$	σ^P	=	0.6
mAQ	=	$6.00 \cdot 10^6$	a^P	=	2.95
MDQ	=	$8.75 \cdot 10^5$	σ^I	=	0.1
mDQ	=	$3.75 \cdot 10^5$	a^I	=	19.04
α	=	0.75	S	=	0
r	=	0.05	ρ	=	0

Table 4.3: Values of the parameters used for the analysis (when not variable).

We here present three analyses: the first one with respect to the volatility level σ^P of gas price, to the MDQ contract parameter, and to the level of market price decoupling. The second one is done with respect to the level of decoupling of the price term structure and to interest rates level. Finally, the third one is done with respect to correlation between P and I and level of decoupling.

The choice of these analyses have been done considering the aim of what we are pursuing, that is to analyse the flexibility given by the make-up clause in a decoupled market scenario. In view of this, we decided to change the parameters we believe to be more impactive on the value of the make-up clause. The volatility σ^P is representative of market uncertainty: in fact, σ^P is often much greater than σ^I, as the index I is calculated as a time average of a basket; as mentioned in the Introduction, this averaging is used to reduce the volatility of the index and leads also to a pretty stable value for σ^I, so changes in σ^P are likely to influence the price more than ones in σ^I. The choice of MDQ is explained by the fact that this quantity is strictly linked with the maximum make-up $\overline{\mathcal{M}}$ the owner of the con-

[3]The ENIGR07 (ENI Gas Release 2007) index is a 9-months time average of a basket of three oil-related indexes, computed as in [1, Equation (1)] or in [22, Equation (1)]

tract can call back in every year. In fact, the bigger MDQ is, the bigger $\overline{\mathcal{M}}$ becomes, and higher the possibility of the owner becomes to posticipate the calling back of the nominated make-up gas. This flexibility should increase the contract value, in particular when price decoupling is strong. We have decided not to move the minimum quantities. On one hand we set the minimum annual quantity and the minimum period quantity in such a way that the possible make-up one can nominate every year is very high ($1.5 \cdot 10^6$), so the stronger constraints are on $\overline{\mathcal{M}}$. On the other hand, we imposed the values of $\overline{K}, \underline{K}$ to be integer and we used values for MDQ in Table 4.4. The underlying idea is that any possible increase in the callable make-up quantity $\overline{\mathcal{M}} = \mathsf{mAQ} - N \cdot \mathsf{mDQ}$ is worthless if the upper bound of gas withdrawal per year $\overline{\mathcal{M}}$ is not enough to call back the nominated make-up quantity. Thus, we map the contract value for MDQ in the range between $\frac{\mathsf{ACQ}}{N} = \frac{7 \cdot 10^6}{12} \simeq 5.83 \cdot 10^5$, which reduces to the case of a standard contract without make-up clause[4], and a value big enough to ensure the withdrawal in the third year of the possible make-up gas nominated in the first and second year, i.e. bigger than $\frac{\mathsf{ACQ}+2(\mathsf{mAQ}-N \cdot \mathsf{mDQ})}{N} \simeq 8.3 \cdot 10^5$ and such that $\underline{K}, \overline{K}$ are integers.

MDQ	$\overline{\mathcal{M}}$	$\underline{\mathcal{M}}$	*Description*
$5.83 \cdot 10^5$	0	$1.5 \cdot 10^6$	No make up
$6.25 \cdot 10^5$	$5 \cdot 10^5$	$1.5 \cdot 10^6$	Low flexibility
$8.75 \cdot 10^5$	$3.5 \cdot 10^6$	$1.5 \cdot 10^6$	High flexibility

Table 4.4: Values of MDQ used in the analysis. All the other parameters, when not variable, are set as in Table 4.3.

The choice of changing MDQ and not other parameters is also a consequence of the practice: we think that the minimum annual quantity and the minimum period quantity are less negotiated than the maximum ones: the seller of the contract will never be willing to sell too much flexibility at the expense of its profits (he want to sell the physical gas), and the buyer will not pay too much for some flexibility he will probably not use in the future (he need the physical gas).

The second and third analyses mainly focus on market factors. As already stated, the make-up clause becomes profitable for the buyer of the contract only if the spread between market and index price $P_t - I_t$ is expected to be lower in the future than in the present. On the other hand, the make-up gas is paid in two different times and its price is affected by the interest rate, as seen in Eq. (4.15). Consequently, the benefits of the decoupling could be affected by high levels of interest rates, which potentially

[4] in fact, if if $\mathsf{MDQ} = \frac{\mathsf{ACQ}}{N}$ then, being not possible to call back any make-up gas before having reached ACQ, we are never able to call back any make-up gas, thus it is also impossible to nominate some.

may vanish the power of make-up clause. This is the focus of the second analysis. Also the correlation could potentially affect the benefits given by the decoupling: in principle, the decoupling should be enforced by negative correlation and weakened by positive one. This is the subject of the third analysis.

First Analysis. The first analysis studies how the contract value depends on the volatility level σ^P, on the MDQ contract parameter and on the level of decoupling. The latter is obtained by varying the initial forward prices used to calibrate the tree prices (see Section 4.4), subtracting a level S from the forward prices F^P for the first year and adding the same quantity to the forward prices for the third year, as shown in Figure 4.11. Then we let S be a parameter and see how the swing price depends on it.

We expect the swing contract value to be increasing in σ^P, with a higher dependence when there is no flexibility given either by the absence of a make-up clause or by small values of MDQ. Figure 4.9 shows exactly these qualitative intuitions. The contract value is increasing with respect to σ^P also for high values of MDQ, but the range in y axes in the figure is so large that we may not appreciate the monotonicity of the curves. This also evidences the fact that the rights given by make-up reduce the risk given by market uncertainty.

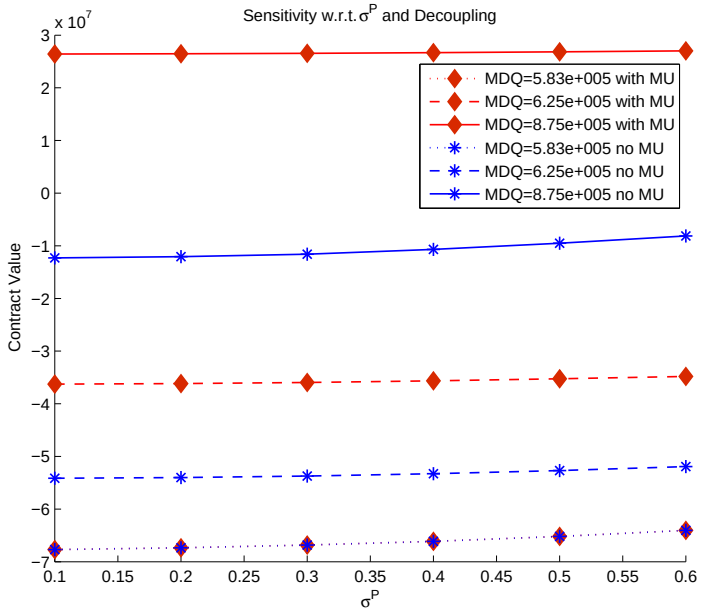


Figure 4.9: Sensitivity with respect to σ^P for three values of MDQ.

The dependence between contract value and decoupling parameter S

is presented in Figure 4.10: make-up rights are useful when market decoupling is high. In these situations, we can nominate make-up gas at the beginning of the contract life and call it back in the future, when a positive market scenario shows up.

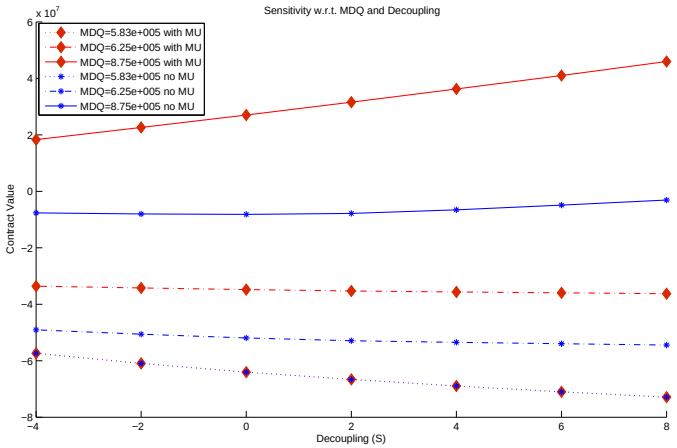
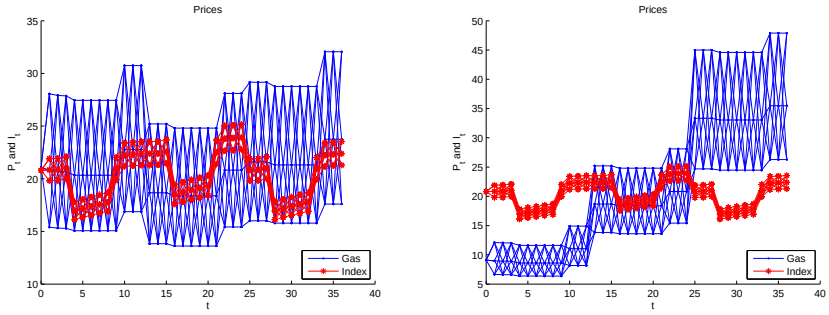


Figure 4.10: Sensitivity with respect to the decoupling S and three values of MDQ, from no make-up rights to very large flexibility. As expected, decoupling enforces make-up value.

Second Analysis. The second analysis is performed by mapping the swing value with respect to the decoupling parameter S and the interest rate r and reporting the corresponding prices in Figure 4.12.

The spirit of this analysis is that the make-up clause is exercised when a negative market scenario (typically, contractual price I higher than spot gas price P) is expected to change or disappear in the following years through a change in the slope of the index and the gas price forward term structure. On the other hand, as we saw in Subsection 4.2.3, the make-up gas nominated is paid partly immediately, and partly when the gas is withdrawn; this temporal mismatch implies a cash flow effects whose impact obviously depends also on the interest rate level: for higher interest rate levels, the benefit of the make-up clause is absorbed by the capitalization of the cost substained from the end of make-up nomination's year up to the withdrawal period. Conversely, in a standard contract without make-up clause, a higher interest rate in a market scenario with a low level of decoupling may lead to a higher contract value: in fact, if the decoupling is low, the present value of the contract in the long term, where the swing option is at or out of the money, is lower than the value in the short term, where the option is in the money. Figure 4.12 shows how any positive change in S is negatively compensated by an increase in the interest rates level.



(a) Low decoupling, swing option at the money in all the 3 years.

(b) High decoupling, swing option out of the money in the first year, at the money in the second year and in the money in the third year.

Figure 4.11: Scenarios for the term structure of gas and index prices for two levels of decoupling. In subfigure (a) make-up rights are typically not exercised, and prices are not decoupled, while in subfigure (b) typically make-up gas is declared in the first year and called back in the third year thanks to the decoupling.

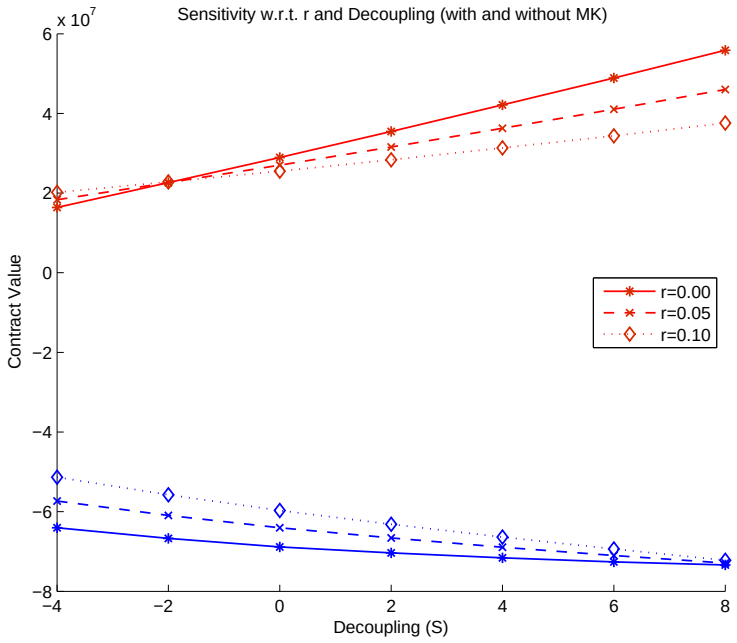
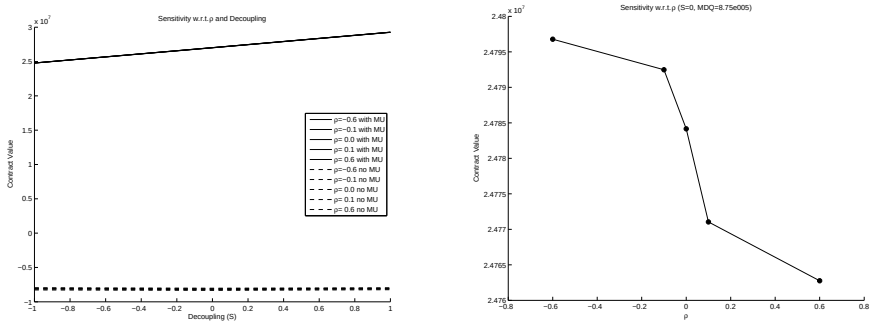


Figure 4.12: Sensitivities with respect to r and level of decoupling S in the forward prices of P. The first three cases on the top are with make-up, the last three in the bottom without.

Third Analysis. The third analysis maps the contract value with respect to the correlation ρ between the two prices P and I, and the level of decoupling S. In Figure 4.13(a) we see that decoupling knocks out correlation: in fact, the swing price's dependence on S is much greater than that on ρ, enforcing once again a strong dependence of the swing price on decoupling levels. Only a deeper analisys, performed for fixed values of decoupling, allows a better understanding of the impact of correlation: negative values of ρ leads to higher values of the contract. This is not a surprise: when ρ is negative the decoupling between prices is expected to be stronger (if P rises up then I falls down thanks to $\rho < 0$) and this increases the value of the contract. However, the changes due to correlation are still smaller than the changes due to decoupling, even for small values of S.



(a) Decoupling variable

(b) Fix value for decoupling

Figure 4.13: Contract value with respect to correlation ρ and level of decoupling. In Figure (a) the shift S vanquishes the effect of the correlation: in fact by varying ρ we obtain almost indistinguishable curves, both with or without make-up. In order to see the differences betwen curves, in Figure (b) the shift S is fixed and here we can see how correlation affecs contract value with make-up: negative values of ρ lead to higher contract values (negative ρ supports decoupling), but the stronger influence of the decoupling S is always evident.

4.6 Conclusions

The oil-to-gas price decoupling of the latest years, especially since the 2008 financial crisis onwards, made the make-up clause a very important feature embedded in most of long term gas swing deals. In this paper we describe, frame and solve the optimization issue related to the presence of a make-up clause in a swing option. As for a standard swing contract, we show that it is possible to reduce the pricing of a swing option with make-

up clause to a stochastic control problem, which can be solved using in a suitable way the Dynamic Programming algorithm. The key idea is to introduce the make-up gas debt as new state variable and incorporating it in the annual constraints on the state space. The dynamic programming is used both on every sub-period of the contract and year by year, by taking into account the gas debt at the beginning of every year. It turns out that, under some not very restrictive assumptions, the optimal withdrawal in all the single sub-periods is of bang-bang type, i.e. it is always optimal to choose in every sub-period between the minimum (mDQ) or the maximum (MDQ) possible withdrawal quantity. This induces a quantization on the number of distinct sequences of the different optimization problems we have to solve in every year, and this number is shown to be dependent on the range MDQ $-$ mDQ and on the annual upper and lower bounds (ACQ and mAQ, respecively). We prove that the total number of optimization problems to solve is quadratic with respect to the product of the duration D of the contract in years and the number N of considered sub-periods. After having described the full algorithm for a generic number of years D, we extend this algorithm to another form of make-up clause, as well as to another clause possibly present in swing contracts, namely the carry-forward.

The algorithm and its extensions are followed by a detailed description of a 3-years contract, which shows how the algorithm works in every year and how the problem is potentially complex even for small values of D. The algorithm of Section 2 is implemented on this 3-years contract, by choosing as the dynamics of P and I a suitable trinomial model with mean-reverting properties, which is calibrated to market data (in particular, TTF for the price P and ENIGR07 for the index I) as explained in Section 4.4. This implementation is then used to perform a sensitivity analisys of the price with respect to MDQ and some other market parameters, namely the volatility of the spot price σ^P, the correlation between spot price and market index ρ, the interest rates level r, and the possible decoupling of gas and index prices induced by the term structure of forward prices, introduced as a perturbation modulated by a parameter S (see Figure 4.11).

The first conclusion is that the market uncertainty given by volatility can be decreased using make-up rights: high levels of make-up rights leads to a less marked dependence in contract value in contrast with the case of a standard contract, where the dependence is more pronunciated.

The main conclusion is however that the decoupling induced by forward prices is crucial in assessing whether the make-up clause is a significant component in the price of the swing option. Figure 4.10 is clear: the higher is the value of the make-up rights, the higher is the dependence of the price on the decoupling. The slope of the contract value changes completely with high make-up rights, turning the decoupling into a favourable market behaviour. Large values of decoupling S also increase the dependence of the swing price on interest rates: in Figure 4.13 we show how the

make-up is sensible to high values of the risk free rate r. The benefits of the decoupling may be in contrast with high rates and make-up, but also in this cases a contract with make-up is always worth more than a contract without. Finally, we investigate the dependence of contract value with respect to correlation ρ. It turns out that contract value remains more or less unchanged when correlation changes, so this dependence is much less significant than the dependence on decoupling.

In conclusion, make-up clauses are a powerful tool to manage primary the new market scenario induced by decoupling and also the uncertainty given by prices. We expect that such type of contracts will be traded more frequently in the future, and here we presented a fast algorithm to price them.

Future work should mainly aim to properly take into account the discrete nature of the index price I, as opposed to the continuous evolution of gas price P, as well as the implicit non-Markovianity of I (being it a time average based on several months, its evolution has relevant memory effects).

BIBLIOGRAPHY

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

[2] O. Bardou, S. Bouthemy, and G. Pagès. When are swing options bang-bang and how to use it? *Pre-print LPMA-1141*, 2007.

[3] O. Bardou, S. Bouthemy, and G. Pagès. Optimal quantization for the pricing of swing options. *Applied Mathematical Finance*, pages 183–217, 2009.

[4] C. Barrera-Esteve, F. Bergeret, C. Dossal, E. Gobet, A. Meziou, R. Munos, and D. Reboul-Salze. Numerical methods for the pricing of swing options: a stochastic control approach. *Methodology and Computing in Applied Probability*, 8(4):517–540, 2006.

[5] F. E. Benth, J. Š. Benth, and S. Koekebakker. *Stochastic Modeling of Electricity and Related Markets*. Advanced Series on Statistical Science and Applied Probability. World Scientific, 2008.

[6] F. E. Benth, J. Lempa, and T. K. Nilssen. On optimal exercise of swing options in electricity markets. *The Journal of Energy Markets*, 4(4), Winter 2011/12.

[7] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control: The Discrete Time Case*. Harcourt Brace Jovanovich Publishers, 1978.

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

[9] 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.

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

[11] A. Boogert and D. Mazieres. A radial basis function approach to gas storage valuation. *Submitted to Journal of Energy Markets*, 2011.

[12] D. Brigo, A. Dalessandro, M. Neugebauer, and F. Triki. A stochastic processes toolkit for risk management: Geometric Brownian motion, jumps, GARCH and variance gamma models. *Journal of Risk Management in Financial Institutions*, 2:365–393, 2008.

[13] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice*. Springer Finance. Springer, 2001.

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

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

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

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

[19] S. Fiorenzani, S. Ravelli, and E. Edoli. *Handbook of Energy Trading*. Wiley, 2011.

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

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

[22] L. Holden, O. Lindqvist, and A. Løland. Valuation of long-term flexible gas contracts. *The Journal of Derivatives*, 18:75–85, 2011.

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

[24] J. C. Hull and A. D. White. Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models. *Journal of Derivatives*, 2:7–16, 1994.

[25] J. C. Hull and A. D. White. Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models. *Journal of Derivatives*, 2(2):37–48, 1994.

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

[27] M. Kanai. Decoupling the oil and the gas prices. *IFRI papers*, May 2011.

[28] R. Kiesel, G. Schindlmayr, and R. Borger. A two-factor model for the electricity forward market. *Quantitative Finance*, 9(3):279–287, 2009.

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

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

[31] 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.

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

[33] T. Meyer-Brandis and P. Tankov. Multi-factor jump-diffusion models of electricity prices. *International Journal of Theoretical and Applied Finance (IJTAF)*, 11(05):503–528, 2008.

[34] 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.

[35] OECD Publishing, http://www.oecdbookshop.org/. *IEA Natural Gas Information Statistics*, 2010. last accessed: 6/9/2011.

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

[37] A. Pokorná. Pricing of gas swing options. Master's thesis, Charles University, Prague, 2009.

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

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

(SOME) MATLAB CODE

Due to industrial privacy, we can not report the full code used in the whole thesis.

Finite Difference

```
%% Finite difference algorithm
% start finite difference
V = NaN(Nt, Np, Ni,Ni, Nz);
p_mesh = zeros(1,Np,Ni,Ni,Nz);
i_mesh = zeros(1,Np,Ni,Ni,Nz);
ihat_mesh = zeros(1,Np,Ni,Ni,Nz);
for m=1:Np
    p_mesh(:,m,:,:,:) = p_int(m);
end
for l=1:Ni
    i_mesh(:,:,l,:,:) = i_int(l);
    ihat_mesh(:,:,:,l,:) = i_int(l);
end

% boundary conditions
R = exp(-rate*T);
Tn_U = T;
Tn_L = 1;
tau1 = @(t,z)min(t+(M_overline-z)/u_bar, Tn_U);
tau2 = @(t,z)max(t, T-(M_underline-z)/u_bar);
g_integrand = @(s,p,t)exp(-rate*s+(p-mu_p)*... % p is the LOG price!
    exp(-theta_p*(s-t))-sigma_p^2*exp(2*theta_p*(t-s))/(4*theta_p));
g = @(a,b,p,t)exp(mu_p-sigma_p^2/(4*theta_p)+rate*Tn_L)*...
    quad(@(x)g_integrand(x,p,t),a,b); % p is the LOG price!
if rate==0
    E = @(a,b,p,i,t)((a<=b)*(g(a,b,log(p),t)-i*(b-a))); % p is THE price!
else
    E = @(a,b,p,i,t)((a<=b)*(g(a,b,log(p),t)-...
        (i/rate)*exp(rate*Tn_L)*(exp(-rate*b)-exp(-rate*a)))); % p is THE price!
```

```
end

h = waitbar(0,'Finite diff.', 'Name','Finite diff...', ...
            'CreateCancelBtn', 'setappdata(gcbf,''canceling'',1)');
setappdata(h,'canceling',0)


% finite difference
p_ms = p_mesh(:,2:end-1,2:end-1,:,1:end-1);
i_ms = i_mesh(:,2:end-1,2:end-1,:,1:end-1);
ihat_ms = ihat_mesh(:,2:end-1,2:end-1,:,1:end-1);
mu_p_tilde    = p_ms.*(theta_p*(mu_p-log(p_ms))+.5*sigma_p^2);
sigma_p_tilde = p_ms.*sigma_p;
mu_i_tilde    = i_ms.*(theta_i*(mu_i-log(i_ms))+.5*sigma_i^2);
sigma_i_tilde = i_ms.*sigma_i;

tic
% boundary u_bar*T
V(:, :,:,:,end) = R*PSI(z_max);

% boundary final time
V(end, :,:,:,:) = permute(repmat(R*PSI(z_int(:)), ...
                            [1, Np, Ni, Ni, 1]), [5, 2, 3, 4, 1]);
for nu=Nt-1:-1:1
    for r=1:Nz
        for lhat=1:Ni
            if getappdata(h,'canceling'); break; end
            t = t_int(nu);
            ihat = i_int(lhat);
            z = z_int(r); ind=z<M_overline;
            V(nu, end, :, lhat,r) = ...
                        u_bar*ind*E(t,tau1(t,z), p_max,ihat,t) + ...
                        R*PSI(z+ind*(tau1(t,z)-t)*u_bar);
            V(nu, 1, :, lhat,r)   = ...
                        u_bar*ind*E(tau2(t,z),Tn_U, p_min,ihat,t) + ...
                        R*PSI(z+ind*(tau2(t,z)<Tn_U)*(Tn_U-tau2(t,z))*u_bar);
        end
    end

    % boundary on I - linear interpolation
    for m=1:Np
        p = p_int(m);
        V(nu,m,1,:,:)   = V(nu,1,1,:,:) + ...
                (V(nu, end,1, :,:)-V(nu,1,1,:,:))*(p-p_min)/(p_min-p_max);
        V(nu,m,end,:,:) = V(nu,1,end,:,:) + ...
                (V(nu, end,end, :,:)-V(nu,1,end,:,:))*(p-p_min)/(p_min-p_max);
    end

    if nu<Nt
        if getappdata(h,'canceling'); break; end
        t = t_int(nu);
        V_   =  V(nu+1, 2:end-1, 2:end-1, :, 1:end-1);
        V_p_fw  = (V(nu+1, 3:end,   2:end-1, :, 1:end-1)-V_)/dp;
        V_i_fw = (V(nu+1, 2:end-1, 3:end,   :, 1:end-1)-V_)/di;
```

```
        V_p_bk  = (-V(nu+1, 1:end-2,   2:end-1, :, 1:end-1)+V_)/dp;
        V_i_bk  = (V(nu+1, 2:end-1, 1:end-2,   :, 1:end-1)+V_)/di;
        V_p = (sigma_p_tilde>=0).*V_p_fw + (sigma_p_tilde<0).*V_p_bk;
        V_i = (sigma_i_tilde>=0).*V_i_fw + (sigma_i_tilde<0).*V_i_bk;
        V_pp = (V(nu+1, 3:end,   2:end-1, :, 1:end-1) - 2*V_ + ...
                V(nu+1, 1:end-2, 2:end-1, :, 1:end-1))/(dp^2);
        V_ii = (V(nu+1, 2:end-1, 3:end,   :, 1:end-1) - 2*V_ + ...
                V(nu+1, 2:end-1, 1:end-2, :, 1:end-1))/(di^2);
        V_pi = ( (V(nu+1, 3:end,   3:end, :, 1:end-1) + ...
                  V(nu+1, 1:end-2, 1:end-2, :, 1:end-1)) - ...
                 (V(nu+1, 1:end-2, 3:end, :, 1:end-1) + ...
                  V(nu+1, 3:end,   1:end-2, :, 1:end-1)))/(4*di*dp);
        V_z  = (V(nu+1, 2:end-1, 2:end-1, :, 2:end)-V_)/dz;

        V(nu, 2:end-1, 2:end-1, :, 1:end-1) = ...
            dt*(.5*sigma_p_tilde.^2.*V_pp + .5*sigma_i_tilde.^2.*V_ii + ...
            rho*sigma_p_tilde.*sigma_i_tilde.*V_pi + ...
            mu_p_tilde.*V_p + mu_i_tilde.*V_i + ...
            (exp(-rate*(t-Tn_L))*(p_ms-ihat_ms)+V_z).*u_opt(exp(-rate*(t-Tn_L))*(p_ms-ihat_ms

        avzmnt=(Nt+1-nu)/Nt; waitbar(avzmnt,h,['Finite difference at: ' sprintf('%2.0f',100*a

        if sum(sum(sum(sum(isnan(V(nu, 2:end-1, 2:end-1, :, 1:end-1))))))
            error('NaN')
        end
    end

end
delete(h)

% valore contratto sui dati iniziali
[~,i_val_min] = min(abs(i_int-exp(i0)));
[~,p_val_min] = min(abs(p_int-exp(p0)));
disp('-------------------------------------------------')
disp(['P(0)=' num2str(p_int(p_val_min)) ', I(0)=' num2str(i_int(i_val_min))])
disp(['dp = ' num2str(dp) ' --> [p_min, p_max] = [' num2str(p_int(1)) ', ' num2str(p_int(end)
disp(['di = ' num2str(di) ' --> [i_min, i_max] = [' num2str(i_int(1)) ', ' num2str(i_int(end)
disp(['dz = ' num2str(dz) ' --> [z_min, z_max] = [' num2str(z_int(1)) ', ' num2str(z_int(end)
disp(['dt = ' num2str(dt) ' --> T = ' num2str(T) ';'])
value_fd= V(1,p_val_min,i_val_min,i_val_min,1);
disp(['Valore contratto con FD: ' separatethousands(value_fd,'''',2)])
disp(['Tempo impiegato: ' separatethousands(toc,'''',2)])
```

Price Simulations

```
%% price simulation MC
sim_lp = p0*ones(nsim,T);
sim_li = i0*ones(nsim,T);
for i=2:T
    prev_p = sim_lp(:,i-1);
    prev_i = sim_li(:,i-1);
    ep1 = randn(nsim,1);
```

```
    ep2 = rho*ep1 + sqrt(1-rho^2)*randn(nsim,1);
    sim_lp(:,i) = prev_p + theta_p*(mu_p-prev_p)*dt + sigma_p*ep1*sqrt(dt);
    sim_li(:,i) = prev_i + theta_i*(mu_i-prev_i)*dt + sigma_i*ep2*sqrt(dt);
end
sim_p = exp(sim_lp);
sim_i = exp(sim_li);
clear sim_lp sim_li

% index sampling (1 month = 30 days)
inizio_mese = 1:30:T;
inizio_mese(end) = T;
nmesi = length(inizio_mese);
for i = 1:nmesi-1
    supp = inizio_mese(i):inizio_mese(i+1)-1;
    sim_i(:,supp)=repmat(sim_i(:,supp(1)), 1,length(supp));
end
```

Naïve Monte Carlo

```
%% naive monte carlo
SP = sim_p-sim_i;
SP = [mean(SP); SP];
OB = zeros(nsim+1,1);EX=OB;
U = zeros(nsim+1,T);
X = zeros(nsim+1,2);
CUMFIN = zeros(nsim,1);
h = waitbar(0,'NMC', 'Name','Naive MC is running');
tic
OPT = optimset;
OPT.Display = 'none';
for i=1:nsim+1
    f = -[SP(i,:) 0 penalty penalty];
    A = [zeros(1,T) 1 -1 0; ...
         zeros(1,T) -1 0 -1];
    b = [M_overline; -M_underline];
    Aeq = [ones(1,T)  -1 0 0];
    beq = 0;
    LB = [mDQ*ones(T,1); -inf; 0; 0];
    UB = [MDQ*ones(T,1); +inf; +inf; +inf];
    [u, obj, exit] = linprog(f,A,b,Aeq,beq,LB,UB,[],OPT);
    OB(i)=-obj;U(i,:)=u(1:T);EX(i)=exit;
    CUMFIN(i)=u(T+1);X(i,:)=u(end-1:end);
    avzmnt = i/nsim; waitbar(avzmnt,h,['NMC is at: ' sprintf('%2.2f',100*avzmnt) '%'
end
tempo_naive = toc;
close(h)
p = profile('info');
time = max([p.FunctionTable.TotalTime]);
memory = max([p.FunctionTable.PeakMem])/(1024*1024);
disp('------------------------')
disp('- NAIVE MC -')
disp(['Valore EXTRINSIC: ' separatethousands(mean(OB),'''',2) '. Tempo: ' separatetho
```

```
disp(['Valore INTRINSIC: ' separatethousands(OB(1),'''',2)])
```

Least Square Monte Carlo

```
%% LEAST SQUARE MC 1d
nx=max([qt.nnodi]);
V = zeros(nsim,nx);
V_prev = V;
EPSI = .001;
Xc = linspace(floor(min(SP(:,end))), ceil(max(SP(:,end))), ncentersTot);
rbf = @(r)(max(1-EPSI*r ,0).^4).*(4*EPSI*r+1);
Vrg = zeros(ncentersTot, nx);
Vrg_prev = Vrg;
P = NaN(nsim,ncentersTot);
Pp = P;
h = waitbar(0,'LSMC', 'Name',['LSMC 1d is running - ' num2str(ncentersTot) ' centers']);
tic
U = NaN(nsim,nx);
for t=T:-1:1
    SSc = SP(:,t);
    if t>1
        SSp = SP(:,t-1);
        DM = abs(repmat(SSp,1,ncentersTot) - repmat(Xc,nsim,1));
        P = rbf(DM);
        Pi = pinv(P);
    end

    for i=1:qt(t).nnodi
        z_n = qt(t).nd(i).vl;
        if t==T
            V(:,i) = PSI(z_n); %condizione finale
        else
            zup = qt(t).nd(i).follower(1);
            if numel(qt(t).nd(i).follower)==2
                zdw = qt(t).nd(i).follower(2);
            else
                zdw=zup;
            end
            q_up = qt(t+1).nd(zup).vl-z_n;
            q_dw = qt(t+1).nd(zdw).vl-z_n;
            Vcont_up = SSc*q_up + Pp*Vrg_prev(:, zup);
            Vcont_dw = SSc*q_dw + Pp*Vrg_prev(:, zdw);
            [a,b] = max([Vcont_up, Vcont_dw], [], 2);
            V(:,i) = a;
            U(:,i) = double(b==1);

            if ~exist('FIG','var'); FIG=0; end
            if FIG
            figure(1), clf, hold on
            plot(SSp, U(:,i), 'o');
            plot(SSp, P*(Pi*U(:,i)), '.r');
            title(['Optimal control when t=' num2str(t) ' and z=' num2str(z_n)])
```

```
            ylim([-.3, 1.3]); grid on

            figure(2), clf, hold on
            plot(SSp, V(:,i), 'o');
            plot(SSp, P*(Pi*V(:,i)), '.r');
            grid on
            title(['Optimal value when t=' num2str(t) ' and z=' num2str(z_n)])
            end
        end

        if t>1
            Vrg(:,i) = Pi*V(:,i);
            Uregress{t,i} = Pi*U(:,i);
        end
    end
    Pp = P;
    Vrg_prev = Vrg;

    avzmnt = (T-t)/T; waitbar(avzmnt,h,['LSMC 1d is at: ' sprintf('%2.2f',100*avzmnt)
end
close(h)
Vopt_1d = full(mean(V(:,1)));
tempo_1d = toc;
disp('------------------------')
disp('- LSMC EXTRINSIC -')
disp(['Numero di centri totali: ' num2str(ncentersTot)])
disp(['Valore con RBF 1d: ' separatethousands(Vopt_1d,'''',2) '. Tempo: ' separatetho
```