Master 2 Probabilités et Finance

Rapport de stage

Mean Field Game Theory for Gas Storage Valuation

Pierre Gasnier

Professeur :
Emmanuel Gobet
Tuteur :
Charly Godefroy

Disclaimer

Abstract

There are different methods to compute the value of a gas storage contract. One of them is to compute the optimal expected cash-flows we can get managing the storage for the duration of the contract. The aim of this report is to use this method where the price model depend on the actions of the market agents, especially other gas storages managers. We use Mean Field Games theory as introduced in 2006 by P.L. Lions to find the equilibrium state of the system and therefore obtaining the optimal control and the value function for a generic player. After presenting gas storages contracts and how to value them in general, we explore different price models and we use a linear-quadratic framework to model the storages features, presenting adapted numerical methods to solve the obtained equations. We then focus on a "hard constraints" model and the numerical method for it. We also discuss about the notion of price of anarchy. Last we introduce some extended games with multiple markets or inhomogeneous storage features, and price model beliefs among the agents.

Acknowledgements

I would like to thank BP IST for hosting me for this internship and in particular Robert Doubble and the Quantitative Analyst team.

I am grateful to my internship supervisor Charly Godefroy who gave me great advice during my internship.

I would like to thank all the team members I worked with for the friendly atmosphere and nice working conditions.

I also want to thank the professors from the M2 Probabilités et Finance of Paris VI and Ecole Polytechnique for the lectures I attended.

Contents

1 Introduction

1.1 On gas storage

As a deregulation trend has driven the energy markets and in particular the gas market since the late 80s, gas trading has become more and more important. At the same time the natural gas consumption increased and the global production hit a record high of 3 590 Billion cubic meters in 2015 (1.6% higher than in 2014). Therefore it makes sense to have a closer look at the gas market and its derivative products.

Among these products are the gas storage that are roughly speaking facilities allowing to store gas when it is cheap (usually during summer) and to pump it out and sell it when the price is high (usually during winter). A storage used to be owned by utilities for the basic purpose of balancing the supply in regards with the demands of customers. However the use of these facilities in developed countries for security of supply increased as the deregulation also spread to the gas storage contracts. Now storages are used and priced as an independent service, some of them can be operated by several companies at the same times for instance (Joint venture). These companies are usually trying to capture the moves in gas price to add value (due to weather, political decisions, ..) as well as the seasonality.

The tools developed in other asset classes to value American or Bermudan options can be used to value gas storage. Indeed, a storage contract can be viewed as an optimal stochastic problem (Swing options, which are commonly traded are degenerated gas storages). Some constraints due to the physical storage facilities must be considered when valuing such an asset.

Along this report we will study the valuation of gas storages when the price model uses the volumes of the market as an explicative variable. The volumes being the mean of the controls of the other storages, the market agents can be seen as players trying to maximize their payoff in a non-zero -sum game. We use a Mean Field Game framework to find an equilibrium state of this game. Mean Field Games are limit versions of stochastic differential games where the number of players tends to infinity, the individual influence of a player's decision becomes infinitesimal and only the mean field of players has influence on the system.

1.2 State of the art on Mean Field Games and their applications in Finance

The Mean Field Game theory was introduced by the parallel works of Lasry and Lions [35] and of Huang, Caines and Malhame [31], they show that the problem can be represented by a system of coupled Hamilton-Jacobi-Bellman and Fokker-Planck equations. Lions then introduced the concept of master equation in his lectures at the College de France [11]. Carmona and Delarue provide a probabilistic analysis of Mean Field Games in [15], showing that the solution of MFG are solution of forward-backward stochastic differential equations on McKean-Vlasov type. Bensoussan, Frehse and Yam [8] show the master equation of MFG in a stochastic framework and using the concept of representative agent optimizing against a distribution rather than taking the limit in a finite differential game as Lasry first introduced it. This is the main approach we will use in our work. Achdou and Capuzzo-Dolcetta [1] provide a numerical scheme to solve the system . Bensoussan et al. [9] shows existence and uniqueness of an equilibrium strategy in the linear-quadratic framework. Bardi [5] provides explicit solutions to some of these linear-quadratic MFG.

Mean Field Games are used in many fields, and especially in economics and finance. Graber [28] applies it to production of an exhaustible resource. Carmona et al. [18] uses the theory to model systematic risk by incorporating a game feature where each bank controls its rate of borrowing/lending to a central bank. Alasseur et al. [2] work on optimal electricity storage in smart-grid. Gomes and Saude [26] and Bagagiolo and Bauso [4] also on price formation in electricity market. Cardaliaguet and Lehalle [13] applied the theory to optimal liquidation. Then Lehalle and Mouzouni [37] extended it to portfolio of correlated assets. The articles of Alasseur and Cardaliaguet are our main inspirations for our work. Their work are actually similar because optimal storage management of gas or power can be seen as optimal liquidation starting from an empty storage.

Guo et al. [29], Hadikhanloo [30] and Cardaliaguet [12] worked on learning in Mean Field Games. Learning in game theory is the study of of dynamics of players that plays the same game repeatedly and adapt their strategies to others', similar to fictitious play. In the article of Cardaliaguet and Lehalle [13], they use learning for repeated day of trading.

Douglas et al. [23], Milstein et al. [40], Ludwig et al. [38], Carmona et al. [14], Ma et al. [39] and Delarue et al. [22] study the the systems of forward-backward stochastic differential equation and provide numerical schemes to solve them. This can be used for the FBSDE of the mean-field games.

2 Gas storage

2.1 Introduction, Purposes and Notations

As a commodity, natural gas prices are subject to time and spatial variations. Gas storage is one among several ways to capture the time variations.

Natural gas can be stored in underground facilities when demand is low and withdrawn when it is high. Because gas production is less flexible than demand, gas storages play an important role to give extra supply when there is an unexpected peak of demand (unseasonal weather, unpredicted plant disruption, political decisions ...) or to build stocks in the opposite case.

Now storage facilities are also used as financial instruments to use the predicted seasonal variations for trading purposes. That means a financial contract is sell to a customer to give him the option to inject or withdraw gas every day until the maturity. A gas storage facility is subject to physical constraints that are essentially limited injection and withdrawal rates, volume dependent (and sometimes time dependent) injection and withdrawal rates and a maximum storage capacity. There exist several kinds of gas storage facilities and the constraints depends on these types (salt caverns, depleted fields, ...).

Cost of injection and withdrawal are also associated to gas storage as there are cost to run the facility (to pump the gas in or out, plus the operating costs) which will need to be included in the valuation.

A company who owns gas storages is interested in tools to value these assets and operate them optimally.

Alongside with this physical definition, we can also consider purely financial storage contract. They replicate a physical gas storage but with greater flexibility for constraints. For example, it is possible to consider only constant injection or withdrawal rates which is not the case in the "real" world. For instance, a gas storage owner may sell some or all of its physical capacity to a counterparty in a standard contract with simple constraints. Gas storage can also be used to provide gas to utilities counterparty operating CCGT (Cycle Combined Gas Turbine). Those participants are capable to produce power when power price spikes and need to guarantee their gas supply. These gas-fired power plants are aimed to be run during peak-hours creating a higher demand for gas and subsequently higher gas price. As a result they seek to buy gas at a fixed price to produce power and sell it at a high price. Thus they are interested in buying swing contracts. A storage owner can sell such a swing and hedge it with its storage asset.

We will use the following notations in the whole document if not stated otherwise.

Gas Storage Properties	c_I	Cost of injection (/MWh)
	c_W	Cost of withdrawal (/MWh)
	α_t	Rate of injection/withdrawal at time t
	$I_{\max}$	Maximum injection rate
	$W_{\max}$	Maximum withdrawal rate
	$S_{\max}$	Storage facility capacity
	S_0	Initial volume
	$S^f_{\max}$	Final volume upper bound (0 in a standard storage contract contract)
	$S^f_{\min}$	Final volume lower bound (0 in a standard storage contract)
	S_t	Current volume at time t
	T	Maturity

If $S^f_{\max} = S^f_{\min}$, then we simplify our notations by defining $S^f = S^f_{\max} = S^f_{\min}$. Usually a storage contract starts with an initial volume equal to 0 and the gas left in stock at the end is lost and may result in penalties in some case. However, for swing options the contracts allow a certain flexibility. Furthermore, it is useful to be able to price a gas storage given a specific volume as it is required if we want to value it during its delivery period.

For the illustrations, we use the following values :

$I_{\max}$	8.197 MMBtu/day
$W_{\max}$	16.394 MMBtu/day
$S_{\max}$	1000000 MMBtu
Step	12
Injection per month	98.364 MMBtu
S^f	0
T	1

The maximum daily injection and withdrawal rates $I_{\max} = I_{\max}(t, V)$ and $W_{\max} = W_{\max}(t, V)$ are volume dependent : the more filled the storage is, the slower the injection is. These non linear profiles are called injection (resp. withdrawal) *ratchets*. In this document we will only consider constant ratchets.

2.2 Intrinsic valuation

The Lacima group published two articles about intrinsic valuations : [33], [34].

The intrinsic value of a natural gas storage is computed using only the current value of the traded futures contracts. Thus it is the maximum value that can be obtained by hedging the storage given the current forward curves. The hedging strategy is implemented at time 0 and no more actions need to be done afterwards.

This method gives a lower bound on the price of the contract. Also there is no randomness, in fact it is the 0 -volatility solution.

We suppose the cost of injection (and withdrawal) is linear in α. Once the forward curves are given, we need to solve the optimization problem

$$\max_{\alpha} \sum_t -(\alpha_t)_+(F_t + c_I) + (\alpha_t)_-(F_t - c_W) \tag{1}$$

with the constraints

$$\begin{cases} \forall t, (\alpha_t)_+ \leq I_{\max} \\ \forall t, (\alpha_t)_- \leq W_{\max} \\ \forall t, S_t \in [0, S_{\max}] \\ S_T \in [S^f_{\min}, S^f_{\max}] \end{cases} \tag{2}$$

This optimization problem can be solved using linear programming.

A particularity of this strategy is we lock the positions from the beginning we will have to take. We are perfectly hedged but we cannot adapt to the evolution of forward curves. Thus we only capture the *intrinsic* value but not the *extrinsic* value. (A rolling intrinsic method can be applied to try to benefit from some of the extrinsic value.)

2.3 Extrinsic valuation

The extrinsic value represents the difference between the real value (or premium) of the storage and the intrinsic value. We often speak about "capturing" the extrinsic value of a gas storage in addition to the locked intrinsic value for instance. In the following paragraphs we look at several ways of computing the real value, i.e. the sum of intrinsic and extrinsic values.

Real valuation of the storage also implies an associated hedging strategy which, unlike the intrinsic value strategy, is run through the whole storage activity period.

2.3.1 Basket of Spread options

The idea is very similar to the intrinsic valuation and there is an analogous improvement called the rolling Basket of Spreads. Storage is represented as a long position in a basket of calendar spread options which are hedged. As we retrieve the optionality of the storage in the spread option, we are able to capture the real value of the storage contract. That also means that, unlike intrinsic valuation, the value is not locked at the beginning.

A calendar spread option can usually be a month calendar spread option but a "daily calendar" spread option would be more accurate as it would give the opportunity to exercise every day.

Again we can solve this linear program and get the storage premium. We then hedge by injecting/withdrawing when the spread option is exercised or not.

The Basket of Spread Option offers a simple way to get an approximation of the storage value as it is possible to choose whether to exercise or not. However the volume are fixed at the beginning so there is no possibility to adjust it through the storage exercise period. This, associated to the fact that ratchets can not be included, are the main drawbacks of this methods. Furthermore using *monthly* calendar spread options loses the optionality associated to the daily exercise.

We are now interested in more general methods that could input all this parameters and flexibility to yield a closer to reality storage value.

2.3.2 Spot optimization

While the previous strategies rely on taking positions in the forward market, in this approach we model the value that can be obtained from making daily decisions of the injection and withdrawal of spot gas. This approach aims to optimize those spot trading decisions to maximise the total discounted revenue over the life of the storage contract, across all possible price paths. By using an underlying spot price model that is consistent with, and calibrated to the market forward curve, we ensure that the value obtained is consistent with the forward strategies described above. In particular, if we consider the case of zero volatility in the spot price this strategy is equivalent to the intrinsic valuation approach. To valuation is therefore solving a stochastic optimization problem.

Definition 1 (*Payoff gas storage contract*)

We introduce the continuous case and corresponding stochastic control equations. We will discretize the problem in time to solve it using dynamic programming, as the storage is discrete : Since a storage is operated daily or hourly in reality, it is not an approximation.

In the continuous case, the payoff f at time t during dt is

$$f(\alpha_t, P_t)dt = \left(-\alpha_t P_t - c_I(\alpha_t)_+ - c_W(\alpha_t)_-\right) dt$$

In the discrete case, the payoff ψ at time t_k during Δt_k is

$$f(\alpha_{t_k}, P_{t_k})\Delta t_k = \left(-\alpha_{t_k} P_{t_k} - c_I(\alpha_{t_k})_+ - c_W(\alpha_{t_k})_-\right) \Delta t_k$$

Definition 2 (*Gas storage constraints*)

Continuous case :

$$\mathcal{A} = \left\{ (\alpha_t)_{0\le t\le T} \ : \ -W_{\max}(t) \le \alpha_t \le I_{\max}(t), \ \forall t \ 0 \le S_0 + \int_0^t \alpha_s ds \le S_{\max}, \ S^f_{\min} \le S_0 + \int_0^T \alpha_t dt \le S^f_{\max} \right\} \tag{3}$$

Discrete case :

$$\mathcal{A} = \left\{ (\alpha_{t_k})_{0\le k\le N-1} \ : \ -W_{\max}(t_k) \le \alpha_{t_k} \le I_{\max}(t_k), \ \forall n \ 0 \le S_0 + \sum_{k=0}^{n-1} \alpha_{t_k}\Delta t_k \le S_{\max}, \ S^f_{\min} \le S_0 + \sum_{k=0}^{N-1} \alpha_{t_k}\Delta t_k \le S^f_{\max} \right\} \tag{4}$$

Remark 1 (*Binary exercise*)

At a given date t_k, the buyer of the storage option has an infinity of choice as he can choose the amount of gas he wants to inject withdraw between $-W_{\max}$ and $I_{\max}$, i.e. $\alpha_{t_k} \in [-W_{\max}, \ I_{\max}]$. However, in the case of constant ratchet (no time dependent neither volume dependent), Bardou et al. [6] proved it is optimal to choose either to fully inject, fully withdraw or to do nothing. This fact is well known and used among the practitioners.

2.3.3 Discrete case

We use a time discretization : $\forall k = 0, .., M,\ t_k = \frac{T}{M}k$. We take the corresponding $\mathcal{A}$ set. We then consider the discrete storage option whose value at time t_k is :

$$V\left(t_k, S_{t_k}, P_{t_k}\right) = \sup_{\alpha\in\mathcal{A}} \mathbb{E}\left[\sum_{j=k}^{M-1} f(\alpha_{t_j}, P_{t_j})\middle| P_{t_k}\right] \tag{5}$$

As mentioned earlier, this is not an approximation of the continuous case. In practice, the storage is operated daily or hourly.

2.3.4 Optimal Stochastic Control, Dynamic Programming Principle and Equation

In this section, we examine the different forms of the general valuation problem, theorems and propositions used come from Touzi [43].

The standard form of the optimal stochastic control is :

$$V(t,x) := \sup_{\alpha\in\mathcal{A}} J(t,x,\alpha) \tag{6}$$

where

$$J(t,x,\alpha) := \mathbb{E}\left[\int_t^T f(s,\alpha_s, X_s^{t,x,\alpha})ds + g(X_T^{t,x,\alpha})\right]$$

In our particular case, it writes as :

$$J(t,p,\alpha) := \mathbb{E}\left[\int_t^T f(\alpha_u, P_u^{t,x})du\right]$$

with

$$f(\alpha,p) = -\alpha p - c_I(\alpha)_+ - c_I(\alpha)_-$$

The valuation problem (5) can be written in the form of a Dynamic Programming Principle(DPP) :

- Terminal condition
$$V\left(T, S_T, P_T\right) = 0$$
- Recurrence : $\forall k = 0,\ ..,\ M-1$,

$$V\left(t_k, S_{t_k}, P_{t_k}\right) = \sup_{\alpha_{t_k}\in\mathcal{A}_{t_k}(S_{t_k})} \left[f\left(\alpha_{t_k}, P_{t_k}\right) + \mathbb{E}\left[V\left(t_{k+1}, S_{t_k} + \alpha_{t_k}\Delta_{t_k}, P_{t_{k+1}}\right) \mid P_{t_k}\right]\right] \tag{7}$$

This allows to compute the premium recursively backward. Because we consider Markovian process, it is easy to compute the conditional expectation. The most used techniques are Least-Square Monte-Carlo and trees-methods.

Remark 2

For an American call, this writes even simpler :
Terminal condition :

$$f\left(P_T\right) = \left(P_T - K\right)_+$$

Recurrence : $\forall k = 0,\ ..,\ M-1$,

$$V\left(t_k, P_{t_k}\right) = \max\ \left(f\left(P_{t_k}\right), \mathbb{E}\left[V\left(t_{k+1}, P_{t_{k+1}}\right) \mid P_{t_k}\right]\right) \tag{8}$$

We notice it is a particular case of the above DPP (7) where the continuation value corresponds to the no exercise choice $\alpha = 0$ *and the early payoff to exercising now, i.e.* $\alpha = 1$

3 Mean Field Games

In this section, we present and summarize most of the Mean Field types of problems in the literature with sketches of proof for the derivations of some equations. We will not discuss deeply of the notions of existence and uniqueness of solutions of these problems, we invite the reader to read the articles referenced for each problems for further proofs and methods of resolution.

3.1 Heuristic derivation of the Mean Field equations

Consider a non cooperative non zero sum game with N identical players with same dynamic, allowed control, and preferences with their own specific noise, the state process X is valued in $\mathbb{R}^d$, the control process α belongs to a set of admissible process $\mathcal{A}$ and is valued in $A \subset \mathbb{R}^q$. The players influence each others' dynamics and payoffs.

$$\begin{cases} dX_t^i = b(t, X_t^i, (X_t^j)_{j\neq i}, \alpha_t^i)dt + \sigma(t, X_t^i, (X_t^j)_{j\neq i}, \alpha_t^i)dW_t^i \\ J_0^i(\alpha^1, \cdots, \alpha^N) = \mathbb{E}\left[\int_0^T f(s, X_s^i, (X_s^j)_{j\neq i}, \alpha_s^i)ds + g(X_T^i, (X_T^j)_{j\neq i})\right] \end{cases} \tag{9}$$

Where b and σ are functions from $[0,T] \times \mathbb{R}^d \times (\mathbb{R}^d)^{N-1} \times A$ in $\mathbb{R}^d$ and $\mathbb{M}_{d,d}(\mathbb{R})$ respectively. f and g are the payoff functions, they are valued from $[0,T] \times \mathbb{R}^d \times (\mathbb{R}^d)^{N-1} \times A$ and $\mathbb{R}^d \times (\mathbb{R}^d)^{N-1}$, respectively, in $\mathbb{R}$.

Solving a game problem has several definition. For Mean Field Games, we use the Nash equilibrium concept. A Nash-equilibrium is α^* such that $\forall i, \forall \alpha : J_0^i(\alpha^{*,1}, \cdots, \alpha^{*,i-1}, \alpha, \alpha^{*,i+1}, \cdots, \alpha^{*,N}) \leq J_0^i(\alpha^{*,1}, \cdots, \alpha^{*,N})$. That means that no player has interest to deviate from his strategy.

Mean field games are games where the players' total influence on the others are not additive but average, so the dependency on the other players $(X_t^j)_{j\neq i}$ is the dependency in the distribution of the player across the state space $\frac{1}{N-1}\Sigma_{j\neq i}\delta_{X_t^j}(dx)$

$$\begin{cases} dX_t^i = b(t, X_t^i, \frac{1}{N-1}\Sigma_{j\neq i}\delta_{X_t^j}, \alpha_t^i)dt + \sigma(t, X_t^i, \frac{1}{N-1}\Sigma_{j\neq i}\delta_{X_t^j}, \alpha_t^i)dW_t^i \\ J_0^i(\alpha^1, \cdots, \alpha^N) = \mathbb{E}\left[\int_0^T f(s, X_s^i, \frac{1}{N-1}\Sigma_{j\neq i}\delta_{X_t^j}, \alpha_s^i)ds + g(X_T^i, \frac{1}{N-1}\Sigma_{j\neq i}\delta_{X_T^j})\right] \end{cases} \tag{10}$$

$$\begin{cases} b : [0,T] \times \mathbb{R}^d \times \mathcal{M}_1(\mathbb{R}^d) \times A \longrightarrow \mathbb{R}^d \\ \sigma : [0,T] \times \mathbb{R}^d \times \mathcal{M}_1(\mathbb{R}^d) \times A \longrightarrow \mathbb{M}_{d,d}(\mathbb{R}) \\ f : [0,T] \times \mathbb{R}^d \times \mathcal{M}_1(\mathbb{R}^d) \times A \longrightarrow \mathbb{R} \\ g : \mathbb{R}^d \times \mathcal{M}_1(\mathbb{R}^d) \longrightarrow \mathbb{R} \end{cases} \tag{11}$$

When N goes to the infinity, the discrete distribution of the other players $\frac{1}{N-1}\Sigma_{j\neq i}\delta_{X_t^j}(dx)$ tend to a continuous distribution $m_t^X(dx)$ which is deterministic in time as the noises of all the player average out in a Fokker-Planck style equation. Also it implies that the individual influence of each player on the others becomes null, only the mean field has influence. Note that from now on we will drop the X from the notation of m^X.

3.2 Traditional MFG

What follows is mainly inspired by the P. Cardaliaguet' notes of P.-L. Lions' lecture at College de France. Analysis and proofs of existence and uniqueness are also provided in these lecture notes.

The distribution (mean field) of the states of the players influence the state itself and the payoff.

A player maximize its expected payoff while knowing the mean field of players and that he has an infinitesimal influence on the mean field.

Solving the problem is finding (α^*, m) such that m is the law of X^*, X^* driven by α^* and $\forall \alpha, J(\alpha, m) \leq J(\alpha^*, m)$

Definition 3 (***State dynamic in traditional MFG***)

The dynamic of the state of a generic player is :

$$dX_t = b(t, X_t, m(t,\cdot), \alpha_t)dt + \sigma(t, X_t, m(t,\cdot), \alpha_t)dW_t$$

The associated Fokker-Planck equation is :

$$\partial_t m(t,x) - \frac{1}{2}\Sigma_{i,j}\partial_{x_i x_j}\left(m(t,x)(\sigma\sigma^T)_{ij}(t,x,m(t,\cdot),\alpha_t)\right) + div\left(m(t,x)b(t,x,m(t,\cdot),\alpha_t)\right) = 0$$

And the Hamiltonian is :

$$H(t,x,p,\gamma,m) = sup_{\alpha\in A}\frac{1}{2}Tr(\sigma\sigma^T(t,x,m,\alpha)\gamma) + b(t,x,m,\alpha).p + f(t,x,m,\alpha)$$

A generic player has to solve the following stochastic optimization problem, assuming he knows what will be the mean field $m(t,dx)$ over time :

$$\begin{cases} V_0(m) = sup_{\alpha\in\mathcal{A}}J(\alpha;m) \\ J(\alpha;m) = \mathbb{E}\left[\int_0^T f(s,X_s,m(s),\alpha_s)ds + g(X_T,m(T))\right] \end{cases} \tag{12}$$

The associated dynamic programming principle problem is :

$$\begin{cases} v(t,x;m) = sup_{\alpha\in\mathcal{A}}J(t,x,\alpha,m) \\ J(t,x,\alpha;m) = \mathbb{E}_{t,x}\left[\int_t^T f(s,X_s,m(s),\alpha_s)ds + g(X_T,m(T))\right] \end{cases} \tag{13}$$

With the associated Hamilton-Jacobi-Bellman partial differential equation system:

$$\begin{cases} \partial_t v(t,x) + H(t,x,D_x v(t,x), D_{xx}v(t,x), m(t)) = 0 \\ v(T,x) = g(x,m(T)) \end{cases} \tag{14}$$

Integrating the optimal control, the Fokker-Planck equation system becomes :

$$\begin{cases} \partial_t m(t,x) - \frac{1}{2}\Sigma_{i,j}\partial_{ij}\left[mD_\gamma H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t))\right]_{ij} + div\left[mD_p H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t))\right] = 0 \\ m(0,x) = m_0(x) = Law(X_0) \end{cases} \tag{15}$$

Definition 4 (***Mean Field Games system of HJB-FP equations***)

A Mean Field Game problem is generally described by the following system of equations:

$$\begin{cases} \partial_t v(t,x) + H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t)) = 0 \\ \partial_t m - \frac{1}{2}\Sigma_{i,j}\partial_{ij}\left[mD_\gamma H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t))\right]_{ij} + div\left[mD_p H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t))\right] = 0 \\ v(T,x) = g(x,m(T)) \\ m(0,x) = m_0 = Law(X_0) \end{cases} \tag{16}$$

Remark 3 (***Fokker-Planck equation***)

The Fokker-Planck equation is normally used in physics to describe the probability of presence of a particle in the space. Here the particles are players that have a control. The fact that there is an "infinite number" of players make that the probability of being in a certain state is, by law of large number, the proportion of players that are in that state.

Definition 5 (***Master equation***)

The game can also be described by the following HJB equation :

$$\begin{cases} \partial_t U + H(t,x,D_xU,D_{xx}U,m) + \partial_m U[\frac{1}{2}\sum_{i,j}\partial_{ij}(mD_\gamma H) - div(mD_pH)] \\ U(T,x,m) = g(x,m) \end{cases} \quad (17)$$

After solving the master equation, one can find $m(t,dx)$ with :

$$\begin{cases} \partial_t m - \frac{1}{2}\Sigma_{i,j}\partial_{ij}\left[mD_\gamma H(t,x,D_xU(t,x,m),D_{xx}U(t,x,m),m)\right]_{ij} + div\left[mD_pH(t,x,D_xU(t,x,m),D_{xx}U(t,x,m),m)\right] = 0 \\ m(0,x) = m_0(x) \end{cases} \quad (18)$$

By setting $v(t,x) = U(t,x,m(t))$, we refind that v and m satisfies the HJB-FP system.

Remark 4

The notation $\partial_f \mathcal{F}[g]$ designate the Gateaux derivative of a functional with direction g : $\partial_f \mathcal{F}[g](f) = \lim_{\epsilon\to 0}\frac{\mathcal{F}(f+\epsilon g)-\mathcal{F}(f)}{\epsilon}$.

There is existence of an equilibrium if f and g have regularizing properties (uniform boundedness and Lipschitz continuity). If in addition f and g are monotonous, the equilibrium is unique. Monotonicity is defined as :

$$\int_x (f(x,m^1) - f(x,m^2))d(m^1(x) - m^2(x)) \geq 0$$

3.3 Extended Mean Field Game

Gomes and Voskanyan [27] are the first to introduce the notion of Extended Mean Field Games, also called Mean Field Game of Controls. They provide some results of existence and uniqueness of solutions.

In this framework, the distribution of controls μ also also influences the dynamic of the players and their reward.

Definition 6 (***State dynamic in extended MFG***)

The dynamic of the state of a generic player is :

$$dX_t = b(t,X_t,m(t,\cdot),\alpha_t,\mu(t,\cdot))dt + \sigma(t,X_t,m(t,\cdot),\alpha_t,\mu(t,\cdot))dW_t$$

The associated Fokker-Planck equation is :

$$\partial_t m(t,x) - \frac{1}{2}\Sigma_{i,j}\partial_{ij}\left(m(t,x)(\sigma\sigma^T)_{ij}(t,x,m(t,\cdot),\alpha_t,\mu(t,\cdot))\right) + div\left(m(t,\cdot)b(t,\cdot,m(t,\cdot),\alpha_t,\mu(t\cdot))\right) = 0$$

And the Hamiltonian :

$$H(t,x,p,\gamma,m,\mu) = sup_{\alpha\in A}\frac{1}{2}Tr(\sigma\sigma^T(t,x,m,\alpha,\mu)\gamma) + b(t,x,m,\alpha,\mu).p + f(t,x,m,\alpha,\mu)$$

Again, the generic player solves the following optimization problem with associated Hamilton-Jacobi-Bellman

PDE, assuming they know the distributions m and μ over time :

$$\begin{cases} v(t,x) = sup_{\alpha\in\mathcal{A}} J(t,x,\alpha;m,\mu) \\ J(t,x,\alpha;m,\mu) = \mathbb{E}_{t,x}\left[\int_t^T f(s,X_s,m(s),\alpha_s,\mu(s))ds + g(X_T,m(T))\right] \\ \partial_t v(t,x) + H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t),\mu(t)) = 0 \\ v(T,x) = g(x,m(T)) \end{cases} \tag{19}$$

Plugging the optimal control in the Fokker-Planck equation gives :

$$\begin{cases} \partial_t m(t,x) - \frac{1}{2}\Sigma_{i,j}\partial_{ij}\left[m(t,x)D_\gamma H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t),\mu(t))\right]_{ij} \\ +div\left[mD_p H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t),\mu(t))\right] = 0 \\ m(0,x) = m_0(x) = Law(X_0) \end{cases} \tag{20}$$

In addition to the traditional MFG, there is the following equation coupling m amd μ

$$\mu(t) = \alpha^*(t,\cdot,D_x v(t,\cdot),D_{xx}v(t,\cdot),m(t),\mu(t))\#m(t,\cdot)$$

Definition 7 (*Extended MFG system of HJB-FP equations*)

The HJB-FP system of an extended MFG problem is :

$$\begin{cases} \partial_t v(t,x) + H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t),\mu(t)) = 0 \\ \partial_t m - \frac{1}{2}\Sigma_{i,j}\partial_{ij}\left[mD_\gamma H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t),\mu(t))\right]_{ij} + div\left[mD_p H(t,x,D_x v(t,x),D_{xx}v(t,x),m(t),\mu(t))\right] = 0 \\ \mu(t) = \alpha^*(t,\cdot,D_x v(t,\cdot),D_{xx}v(t,\cdot),m(t),\mu(t))\#m(t,\cdot) \\ v(T,x) = g(x,m(T)) \\ m(0,x) = m_0(x) \end{cases} \tag{21}$$

Remark 5

The notation $f = h\#g$ means that $f(dy)$ is the distribution of the $h(x) \in A$ with the distribution of $x \in \mathbb{R}^d$ being $g(dx)$

Definition 8 (*Master equation*)

The game can also be described by the following HJB equation :

$$\begin{cases} \partial_t U + H(t,x,D_x U,D_{xx}U,m,\mu) + \partial_m U[\frac{1}{2}\sum_{i,j}\partial_{i,j}(mD_\gamma H) - div(mD_p H)] \\ \mu = \alpha^*(t,x,D_x U,D_{xx}U,m,\mu)\#m \\ U(T,x,m) = g(x,m) \end{cases} \tag{22}$$

3.4 Mean Field type Control

Lauriere and Pironneau [36] and Bensoussan et al. [8] provides in their article some understanding of Mean Field type Control as a traditional stochastic optimization problem, with the master equation being seen a Bellman equation. Mean Field type Control are also known as Controlled McKean-Vlasov dynamics.

An overseer maximize the common interest. Changing the strategy therefore influences the mean field since by symmetry every agents have the same strategy.

Denoting m^α the mean field for strategy α . The problem is finding α^* such that $\forall\alpha, J(\alpha,m^\alpha) \leq J(\alpha^*,m^{\alpha^*})$

$$\begin{cases} J(\alpha;m) & = \mathbb{E}\left[\int_0^T f(s,X_s,m(s),\alpha_s)ds + g(X_T,m(T))\right] \\ & = \int_0^T \int_x f(s,x,m(s),\alpha(s,x))dsm(s,dx) + \int_x g(x,m(T))m(T,dx) \end{cases} \tag{23}$$

The difference with Mean Field Games is that, while no player has interest in deviating from its strategy to maximize his personal interest, they could cooperate to maximize the common interest. We cannot assume that we know the mean field and then find the optimal control corresponding to this means field because here the mean field depend on the control set by the overseer. The problem is maximizing the functional :

$$\begin{cases} \hat{J}(\alpha) = J(\alpha; m^\alpha) & = \mathbb{E}\left[\int_0^T f(s, X_s, m^\alpha(s), \alpha_s)ds + g(X_T, m^\alpha(T))\right] \\ & = \int_0^T \int_x f(s, x, m^\alpha(s), \alpha(s,x))ds m^\alpha(s, dx) + \int_x g(x, m^\alpha(T))m^\alpha(T, dx) \end{cases} \tag{24}$$

One can use the Gateaux derivative and find $\alpha(t, s)$ such that the derivative is null for any direction.

Another method is to use dynamic programming :

$$\begin{cases} \hat{J}(t, m; \alpha) = \mathbb{E}\left[\int_t^T \int_x f(s, x, m^\alpha(s), \alpha(s,x))m^\alpha(s, dx)ds + \int_x g(x, m^\alpha(T))m^\alpha(T, dx)\middle| m_t^\alpha = m\right] \\ V(t, m) = \sup_\alpha \hat{J}(t, m; \alpha) \end{cases} \tag{25}$$

The HJB equation for V is

$$\begin{cases} \partial_t V + \sup_\alpha \partial_m V[\frac{1}{2}\sum_{i,j} \partial_{ij}(m(x)(\sigma\sigma^T)_{ij}(t, x, m, \alpha(x))) - div(m(x)b(t, x, m, \alpha(x)))] + \int_x f(t, x, m, \alpha(x))m(dx) = 0 \\ V(T, m) = \int_x g(x, m)m(dx) \end{cases} \tag{26}$$

By the Riesz representation theoreme $\exists\, U : (t, x, m) \to U(t, x, m)$ s.t. $\partial_m V[\tilde{m}](t, m) = \int_x U(t, x, m)\tilde{m}(dx)$. We usually note $U = \delta_m V$. With an integration by part we can rewrite the HJB equation as :

$$\begin{cases} \partial_t V + \sup_\alpha \int_x \partial_x U D_x b(t, x, m, \alpha(t, x) + \frac{1}{2} Tr(D_{xx} U \sigma\sigma^T(t, x, m, \alpha(t, x)))m(dx) + \int_x f(t, x, m, \alpha(x))m(dx) = 0 \\ V(T, m) = \int_x g(x, m)m(dx) \end{cases} \tag{27}$$

and since the optimization can be done inside the integral :

$$\begin{cases} \partial_t V + \int_x H(t, x, D_x U, D_{xx} U, m)m(dx) = 0 \\ V(T, m) = \int_x g(x, m)m(dx) \end{cases} \tag{28}$$

Taking the derivative with respect to m we get the master equation of MFTC :

$$\begin{cases} \partial_t U + \int_x \delta_m(H(t, x, D_x U(t, x, m), D_{xx} U(t, x, m), m))m(dx) + H(t, x, D_x, D_{xx} U, m) = 0 \\ U(T, x, m) = \int_x \delta_m g(x, m)m(dx) + g(x, m) \end{cases} \tag{29}$$

3.5 Mean Field Planning Problem

Orrieri et al. [41] provide a study of Mean Field Planning Problems and are able to prove existence and uniqueness in some cases.

The problem here is to move a distribution to another while maximizing the revenue of the agents. This is a problem of optimal transport.

Definition 9 (*HJB-FP system of MFPP*)

The HJB-FP system of this kind of problem is :

$$\begin{cases} \partial_t v(t,x) + H(t,x,D_x v(t,x), D_{xx} v(t,x), m(t)) = 0 \\ \partial_t m - \frac{1}{2}\Sigma_{i,j}\partial_{ij}\left[mD_\gamma H(t,x,D_x v(t,x), D_{xx}v(t,x), m(t))_{ij}\right] + div\left[mD_p H(t,x,D_x v(t,x), D_{xx} v(t,x), m(t))\right] = 0 \\ m(0,x) = m_0(x), m(T,x) = m_T(x) \end{cases} \tag{30}$$

3.6 MFG with common noise

Carmona et al. [17] develop a theory of existence and uniqueness of solutions for Mean Field Games with common noise. In [16] Carmona and Delarue exhibit the master equation for these kinds of problems. Kolokoltsov and Troeva [32] and Firoozi et al. [24] provide further results on games with common noise.

In the previous types of problem, the randomness was specific to each player and therefore disappeared in the mean field. In the MFG with common noise type of problems, there is a systematic additional noise that drive the dynamic of the players.

$$dX_t = b(t,X_t,m_t,\alpha_t)dt + \sigma(t,X_t,m_t,\alpha_t)dW_t + \tilde{\sigma}(t,X_t,m_t,\alpha_t)dB_t$$

Here m_t denote the process $(Law(X_t|B_s, 0 \leq s \leq t))_t$. Because the distribution of the players over time is now also driven by a Brownian motion, it is a stochastic process adapted to the filtration of the common noise. It is important to note that the process is valued in the space of distribution on $\mathbb{R}^d$.

The Fokker-Planck equation becomes a stochastic differential equation :

$$\begin{cases} dm_t(x) = \left(\frac{1}{2}\Sigma_{i,j}\partial_{ij}\left[(\sigma\sigma^T + \tilde{\sigma}\tilde{\sigma}^T)_{ij}(t,x,m_t,\alpha_t)m_t(x)\right] - div\left[b(t,x,m_t,\alpha_t)m_t(x)\right]\right)dt - div[\tilde{\sigma}\,(t,x,m_t,\alpha)m_t(x)]\,dB_t \\ m_0 = Law(X_0) \end{cases} \tag{31}$$

The problem is finding (α^*, m) such that m is the law of X^* knowing B, X^* driven by α^* and $\forall \alpha, J(\alpha,m) \leq J(\alpha^*, m)$.

One can find a stochastic HJB equation coupled with the Fokker-Planck equation similarly to the classic MFG case. This leads to a system of coupled infinite dimensional forward-backward stochastic partial differential equations. The solution HJB backward equation can be represented by a deterministic function of the forward process solution of the FP equation. This function is called decoupling field and is solution of a master equation.

MFG with common noise is the framework we will mostly work in as in the gas storage problem, the price is common to every agents and the dynamic of the price is driven by a Brownian motion.

3.7 MFG with a major player

Bensoussan et al. [7] provide a general theory and results of existence and uniqueness of equilibrium in certain cases.

The difference with traditional Mean Field Games is that there is one player of which influence cannot be neglected, denoting by X^0 and α^0 its state and control :

$$\begin{cases} dX^0_t = b^0(t,X^0_t,m_t,\alpha^0_t)dt + \sigma^0(t,X^0_t,m_t,\alpha^0_t)dW^0_t \\ dX_t = b(t,X_t,X^0_t,m_t,\alpha_t,\alpha^0_t)dt + \sigma(t,X_t,X^0_t,m_t,\alpha_t,\alpha^0_t)dW_t \end{cases} \tag{32}$$

Here m_t denote the process $(Law(X_t|W^0_s, 0 \leq s \leq t))_t$. It follows the following SDE :

$$\begin{cases} dm_t(x) = \left(\frac{1}{2}\Sigma_{i,j}\partial_{x_i,x_j}\left[(\sigma\sigma^T)_{ij}(t,x,X^0_t,m_t,\alpha_t,\alpha^0_t)m_t(x)\right] - div\left[b(t,x,X^0_t,m_t,\alpha_t,\alpha^0_t)m_t(x)\right]\right)dt \\ m_0 = Law(X_0) \end{cases} \tag{33}$$

The generic player has the following reward function, depending on its strategy, the one of major player, and the mean field :

$$J(\alpha, \alpha^0, m) = \mathbb{E}\left[\int_0^T f(s, X_s, m_s, \alpha_s, \alpha_s^0)ds + g(X_T, X_T^0, m_T)\right]$$

The major player has the following reward function, depending on its strategy and the mean field :

$$J^0(\alpha^0, m) = \mathbb{E}\left[\int_0^T f^0(s, X_s^0, m_s, \alpha_s^0)ds + g^0(X_T^0, m_T)\right]$$

Suppose that for any strategy of the major player α^0 there exists a mean field $\tilde{m}^{\alpha^0}$ which is a Nash equilibrium for the other players : $\tilde{m}^{\alpha^0}$ is the mean field for a strategy α^* such that $\forall \alpha,\ J(\alpha, \alpha^0, \tilde{m}^{\alpha^0}) \leq J(\alpha^*, \alpha^0, \tilde{m}^{\alpha^0})$

Solving the problem is finding (α^{0*}) such that $\forall \alpha^0,\ J^0(\alpha^0, \tilde{m}^{\alpha^0}) \leq J^0(\alpha^{0*}, \tilde{m}^{\alpha^{0*}})$

In other terms, the problem is : first given the strategy of the major player and a mean field to find the optimal strategy for a representative player; second to find, given this major player strategy, the mean field equilibrium, then last to find the major player strategy that maximizes his interest while knowing that the mean field depends on his strategy.

Remark 6

MFG, MFC, MFPP can be traditional or extended, with our without common noise, with or without a major player.

Remark 7

One can imagine problems where players are asymmetric, that is not all players have the same state dynamic and payoff function. Actually those problems are just specific cases of the general problems described above, as one can say that the player's type is part of the state variable. The dynamic and payoff being the same for everyone but now also depends on this additional state variable, therefore having a symmetric system.

4 Spot optimization in a Mean Field Game

4.1 Description of the problem and assumptions

4.1.1 Reward functions and simplifying assumption

A game typically plays for one year $T = 1$. As we have seen before, there are constraints on the volume inside a gas storage and it must be emptied at the end of the year. The injection and withdrawal rates are bounded, boundaries can depend on the level of the storage. There are additional costs to the price paid when injecting or withdrawing, depending on the level of the storage, and the rate of injection/ withdrawal.

We'll mainly use a linear-quadratic framework in order to have existence and uniqueness of an equilibrium (see Bensoussan et al. [9] and Bardi [5]), as well as possible explicit simple solutions. Instead of having hard boundaries on the inventory and rate, there are linear-quadratic penalizations for abnormal values of inventory and control.

4.1.2 Mean field dimension of the problem

The mean field modelling comes through the mean of the injection/withdrawal rates thus we fall under the Extended Mean Field Game type. The framework is that it is the average sum of the injection rates that plays on the dynamic of the price and/or the effective paid price, representing a market impact and/or the price being function of the demand. We will present and discuss several price models inspired from traditional ones.

We denote by $\bar{\mu}_t$ the average sum of controls at time t , $\bar{\mu}_t = \int \alpha \mu_t(d\alpha)$. Also the dynamic of the inventory is : $dS_t = \alpha_t dt$. There is no noise specific to each player but there a common noise on the price so we work on the Extended Mean Field Game with common noise framework. Also there is obviously only one price common to every player. It can be seen as the state variable of the players are composed by their inventory and price, but the initial marginal distribution in price is a Dirac, and since the dynamic of the price does not involve idiosyncratic noise or control the distribution of the price across the player remains a Dirac through time.

$$dm_t(s,p) = \left(-\partial_p(m_t b_P) - \partial_s(m_t \alpha_t)\right) dt - \partial_p(m_t \sigma_P) dW_t$$

$m_0(s,p) = m_0(s)\delta_{p_0}(p)$ which can be integrated on the price variable to obtain the marginal distribution of inventories of which the dynamic is

$$dm_t(s) = -\partial_s(m_t \alpha_t) dt \tag{34}$$

Remark 8 (*Further assumptions*)

We will work assuming that every agent has the same preferences (reward functions) to have simple equations, but the problem is also totally solvable if there is a distribution of preferences across the agents. The thing is that the problem in itself assume that every player can observe the distribution of inventories at any time and the preferences of the other players and know that all other players can as well.

We also first assume the every agents have the same price model. In the same way is possible to solve the problem with different price model among the players but it at least need to assume that " every player knows the models used by the other players and knows that every other player also know all the models " .

We'll see later that in the learning framework, we do not need those assumptions.

4.2 Market impact

In the general case we assume that the mean field only plays on the drift of the price

$$dP_t = b_P(t, P_t, \bar{\mu}_t)dt + \sigma_P(t, P_t)dW_t$$

4.2.1 Reward function

Definition 10 (*Reward function*)

$$
\begin{cases}
\mathcal{A} = \{adapted\ process\ valued\ in\ A\}, A = \mathbb{R} \\
J(\alpha, \mu) = \mathbb{E}\left[\int_0^T \left[\underbrace{-P_t\alpha_t}_{(1)} \underbrace{-\frac{C}{2}\alpha_t^2}_{(2)} + \underbrace{A_1 S_t}_{(3)} \underbrace{-\frac{A_2}{2}S_t^2}_{(4)}\right] dt + \underbrace{\underbrace{P_T S_T}_{(5.1)} \underbrace{-\frac{B}{2}S_T^2}_{(5.2)}}_{(5)}\right]
\end{cases}
\tag{35}
$$

(1) corresponds to the price paid/received for the injection/withdrawal between t and $t + dt$
(2) is the cost of operations, it can also be seen as a cost of liquidity. It also replace the constraint of boundedness of the injection rate by penalizing too big injection rate
(3) is a term penalize the player for having a negative inventory, replacing the hard constraint of positivity
(4) is a term that penalize the player for having a too big inventory, replacing the hard constraint of maximum inventory
(5) we replace the hard constraint of empty storage at maturity with the assumption that the player can sell all he has left in stock at once (5.1) but with a quadratic cost of liquidity (5.2), also the greater B is, the more the player is encouraged to have a null inventory at maturity. The term (5.1) allows to have explicit solution in the Bachelier case
The rate of injection and inventory variable are no longer bounded.

Remark 9

On the terminal payoff g, one could also write $S_T(P_T - B_1 E_T - \frac{B_2}{2} S_T)$ where E_T denotes the mean of the inventories of all the players at T. The term $-B_1 E_T S_T$ being a liquidity cost, it rewards the player for having to sell while the mean of the players have to buy.

4.2.2 Bachelier market impact

$$dP_t = (f_0(t) + \nu\bar{\mu}_t)\,dt + \sigma dW_t$$

f_0 is an anticipated seasonality
It is the most simple market impact model, with soft constraints it even has a deterministic explicit solution.

4.2.3 Black-Scholes market impact

$$\frac{dP_t}{P_t} = (f_0(t) + \nu\bar{\mu}_t)\,dt + \sigma dW_t$$

f_0 is an anticipated seasonality
Black-Scholes without market impact is the most known price model in finance, this is why it is interesting to have market impact price model derived from it.

4.2.4 Clewlow-Strickland (1 factor) market impact

We assume that the dynamic of the forward contract is :

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma e^{-a(T-t)} dW_t + \nu e^{-a(T-t)}\bar{\mu}_t dt$$

Then $F_{t,T} = F_0(T)\exp\left(e^{-aT}\left(\sigma\int_0^t e^{as}dW_s + \nu\int_0^t e^{as}\bar{\mu}_s ds\right) - \frac{\sigma^2}{4a}e^{-2aT}(e^{2at}-1)\right)$. And the dynamic of $P_t = F_{t,t}$ is therefore:

$$\frac{dP_t}{P_t} = \left(\frac{F_0'(t)}{F_0(t)} + a\left(\ln(F_0(t)) - \ln(P_t)\right) + \frac{\sigma^2}{4}(1-\exp(-2at)) + \nu\bar{\mu}_t\right)dt + \sigma dW_t$$

$t \to F_0(t)$ is the forward curve at 0

Clewlow-Strickland without market impact is a model commonly use for commodities (see [20]). We should therefore aim to use market impact model derived from it.

4.3 Price is a function the demand

The demand is the sum of the average control of the players and an exogenous demand process Q_t which represents the demand of the rest of the world, the demand is algebraic. The rest of the world can be seen as agents that uses a 'classic' price model, or that do not fit in the storage model due to different kind of reward functions such as household or governments.

The price for a demand $Q_t + \bar{\mu}_t$ is a deterministic function P of that demand.

4.3.1 Reward functions

$$\begin{cases} \mathcal{A} = \{adapted\ process\ valued\ in\ A\}, A = \mathbb{R} \\ J(\alpha,\mu) = \mathbb{E}\left[\int_0^T \left[-P(Q_t+\bar{\mu}_t)\alpha_t - \frac{C}{2}\alpha_t^2 + A_1 S_t - \frac{A_2}{2}S_t^2\right]dt + P(Q_T)S_T - \frac{B}{2}S_T^2\right] \end{cases} \tag{36}$$

The reward function has the same structure than the market impact price models. The liquidation price at T is not well defined. We arbitrarily kept and wrote the term (5.1) like this but one can remove or change it, this term is exogenous so his presence does not impact the solvability of the problem.

4.3.2 Exogenous demand and price function

We use an exogenous demand process :

$$\begin{cases} dQ_t = b_Q(t,Q_t)dt + \sigma_Q(t,Q_t)dW_t \\ Q_0 = q_0 \end{cases}$$

A good modelling of this process would be an Ornstein-Uhlenbeck process with seasonal trend. For the inverse demand price function an affine price is a good compromise between realism and convenience for calculus.

$$P(q) = p_0 + p_1 q$$

4.4 General price model

More generally, one can describe the price paid by the players at a certain time as a function of a process and $\bar{\mu}$:

$$\begin{cases} P_t = P(t, X_t, \bar{\mu}_t) \\ P_T = \tilde{P}(X_T) \\ P : [0,T[\times\mathbb{R}^d \times \mathbb{R} \longrightarrow \mathbb{R} \\ \tilde{P} : \mathbb{R}^d \longrightarrow \mathbb{R} \\ dX_t = b(t,X_t,\bar{\mu}_t)dt + \sigma(t,X_t,\bar{\mu}_t)dW_t \\ X_0 \in \mathbb{R}^d \end{cases} \tag{37}$$

The reward function being

$$J(\alpha, \mu) = \mathbb{E}\left[\int_0^T \left[-P(t, X_t, \bar{\mu}_t)\alpha_t - \frac{C}{2}\alpha_t^2 + A_1 S_t - \frac{A_2}{2}S_t^2\right] dt + \tilde{P}(X_T)S_T - \frac{B}{2}S_T^2\right]$$

The previous cases falling under this general model :

Market impact

$$\begin{cases} P(t, X_t, \bar{\mu}_t) = X_t \\ \tilde{P}(X_T) = X_T \\ dX_t = b_P(t, X_t, \bar{\mu}_t)dt + \sigma_P(t, X_t)dW_t \\ X_0 = P_0 \in \mathbb{R} \end{cases} \tag{38}$$

Price is function of demand

$$\begin{cases} P(t, X_t, \bar{\mu}_t) = p_0 + p_1(X_t + \bar{\mu}_t) \\ \tilde{P}(X_T) = p_0 + p_1(X_T) \\ dX_t = b_Q(t, X_t)dt + \sigma_Q(t, X_t)dW_t \\ X_0 = Q_0 \in \mathbb{R} \end{cases} \tag{39}$$

5 System equation and resolution

5.1 Market impact models

This model is greatly inspired by the work of Cardaliaguet and Lehalle [13] on optimal liquidation in a Mean Field Game. Their work is it self an evolution of the work of Almgren and Chriss [3]. Our inspiration comes from the fact that optimal strategies for a storage can be seen as optimal liquidation starting from a null inventory, the added seasonality in the price model is what makes the players buy and sell.

5.1.1 Derivation of the equations

We are in the Mean Field Game with common noise type so we cannot use the dynamic programming principle as we would do in a 'classic' price model since the mean field m is not deterministic. We can however consider it is itself a state variable and use the dynamic programming principle. For now the average control μ is some adapted process, will use the corresponding equation to determine it.

$$J(t,s,p,m;\alpha,\mu) = \mathbb{E}_{t,s,p,m}\left[\int_t^T (-\alpha_s P_s - \frac{C}{2}\alpha_s^2 + A_1 S_s - \frac{A_2}{2}S_s^2)ds + P_T S_T - \frac{B}{2}S_T^2\right]$$

$\mathbb{E}_{t,s,p,m}[\cdot]$ means $\mathbb{E}[\cdot|S_t = s, P_t = s, m_t = m]$

Denoting $v(t,s,p,m;\mu) = \sup_{\alpha\in\mathcal{A}} J(t,s,p,m;\alpha,\mu)$, the dynamic programming principle gives :

$$\begin{cases} \partial_t v + \frac{1}{2}\sigma_P^2(t,p)\partial_{pp}v + b_P(t,s,\bar{\mu}_t)\partial_P v + \sup_{\alpha\in\mathbb{R}}\left[\alpha\partial_s v - \alpha p - \frac{C}{2}\alpha^2\right] + A_1 s - A_2\frac{s^2}{2} + \partial_m v[-\partial_s(m\tilde{\alpha})] = 0 \\ v(T,s,p,m) = ps - \frac{B}{2}s^2 \end{cases}$$

We therefore have the control $\tilde{\alpha}(t,s,p,m,\mu) = \frac{\partial_s v(t,s,p,m,\mu) - p}{C}$

Remark 10

m normally designates the process of density of the state of players, but by abuse of notation it also designates generic density argument of the functional v.

$$\begin{cases} \partial_t v + \frac{1}{2}\sigma_P^2(t,p)\partial_{pp}v + b_P(t,s,\bar{\mu}_t)\partial_P v + \frac{(\partial_s v - p)^2}{2C} + A_1 s - A_2\frac{s^2}{2} + \partial_m v[-\partial_s(m\frac{\partial_s v - p}{C})] = 0 \\ v(T,p,s) = ps - \frac{B}{2}s^2 \end{cases} \tag{40}$$

From now on we stop writing the dependency in μ for ease of notation. We assume the following separation of variables :

$$v(t,s,p,m) = h_0(t,p,m) + (p + h_1(t,p,m))s - h_2(t,p,m)\frac{s^2}{2}$$

We therefore have the following system of equation :

$$\begin{cases} \partial_t h_2 + \dfrac{1}{2}\sigma_P^2(t,p)\partial_{pp}h_2 + b_P(t,s,\bar{\mu}_t)\partial_P h_2 - \dfrac{h_2^2}{C} + A_2 + \partial_m h_2[-\partial_s\left(m\dfrac{h_1 - sh_2}{C}\right)] & = 0 \\ \partial_t h_1 + \frac{1}{2}\sigma_P^2(t,p)\partial_{pp}h_1 + b_P(t,s,\bar{\mu}_t)(1+\partial_P h_1) - \frac{h_1 h_2}{C} + A_1 + \partial_m h_1[-\partial_s\left(m\frac{h_1 - sh_2}{C}\right)] & = 0 \\ \partial_t h_0 + \frac{1}{2}\sigma_P^2(t,p)\partial_{pp}h_0 + b_P(t,s,\bar{\mu}_t)\partial_P h_0 + \frac{h_1^2}{2C} + \partial_m h_0[-\partial_s\left(m\frac{h_1 - sh_2}{C}\right)] & = 0 \\ h_2(T) = B, h_1(T) = h_0(T) = 0 \end{cases} \tag{41}$$

The first equation can be solve by assuming that $h_2(t,p,m) = h_2(t)$

$$\begin{cases} h_2' - \dfrac{h_2^2}{C} + A_2 = 0 \\ h(T) = B \end{cases}$$

Which is a Riccati equation that can be solved by writing : $\frac{h_2'}{h_2^2 - A_2 C} = \frac{1}{C}$ then $\frac{h_2'}{h_2 - \sqrt{A_2 C}} - \frac{h_2'}{h_2 + \sqrt{A_2 C}} = 2\sqrt{\frac{A_2}{C}}$ which can be integrated setting $\rho = \sqrt{\frac{A_2}{C}}$ in $\ln\left(\frac{h_2(t) - \rho C}{h_2(t) + \rho C}\right) = 2\rho(t-T) + \ln\left(\frac{B - \rho C}{B + \rho C}\right)$ which is, by setting $c_2 = \frac{B-\rho C}{B+\rho C}\exp(-2\rho T)$

$$h_2(t) = \rho C \frac{1 + c_2\exp(2\rho t)}{1 - c_2\exp(2\rho t)} \tag{42}$$

We recall that the control is

$$\alpha(t,s,p,m,\mu) = \frac{h_1(t,p,m,\mu) - h_2(t)s}{C}$$

The average control therefore satisfies this equation

$$\bar{\mu}_t = \int_s \alpha(t,S_t,P_t,m_t,\mu)m_t(ds) = \frac{h_1(t,P_t,m_t,\mu) - h_2(t)\int_s s m_t(ds)}{C}$$

Denoting by E_t the average inventory of the players, the average control is :

$$\bar{\mu}_t = \frac{h_1(t,P_t,m_t,\mu) - h_2(t)E_t}{C}$$

The process $E_t = \int s m_t(ds)$ satisfy the following SDE :
$dE_t = \int s dm_t(ds) = -\int s\partial_s(m_t\frac{h_1(t,P_t,m_t,\mu) - h_2(t)s}{C})dt = \frac{h_1(t,P_t,m_t,\mu) - h_2(t)E_t}{C}dt = \bar{\mu}_t dt$
by integration by part.

A stylizing fact is that

$$d(S_t - E_t) = -h_2(t)(S_t - E_t)dt$$

This shows that the players will try to have their inventory equal to the mean of the inventories, if it is not originally the case. This also shows that, assuming that the majority of the players are doing the best strategy, the best thing to do is following the crowd.

We can assume that the dependency of h_1 and h_0 in m is only through the mean of m , therefore :

$$\begin{cases} \partial_t h_1 + \dfrac{1}{2}\sigma_P^2\partial_{pp}h_1 + (1+\partial_p h_1)b_P\left(t,p,\dfrac{h_1(t,p,e) - eh_2(t)}{C}\right) - \dfrac{h_1 h_2}{C} + A_1 + \partial_e h_1\dfrac{h_1 - eh_2}{C} = 0 \\ h_1(T,p,e) = 0 \end{cases} \tag{43}$$

It is associated with the following Forward-Backward stochastic differential equation :

$$\begin{cases} dP_t = b_P\left(t,P_t,\dfrac{H_t - E_t H_2(t)}{C}\right)dt + \sigma(t,P_t)dW_t \\ dE_t = \frac{H_t - h_2(t)E_t}{C}dt \\ dH_t = \left(\frac{H_t h_2(t)}{C} - A_1 - b_P\left(t,P_t,\frac{H_t - h_2(t)E_t}{C}\right)\right)dt + Z_t dW_t \\ H_T = 0 \end{cases} \tag{44}$$

$H_t = h_1(t, P_t, E_t)$ being the solution of the FBSDE is equivalent to h_1 being the solution of the PDE. Also $h_0(t, P_t, E_t) = H_t^0 = \mathbb{E}_t\left[\int_t^T \frac{(H_u)^2}{2C} du\right]$.

From now on until the end of the document, we use μ_t to instead designate the mean of the controls as in this framework we do not need the entire distribution of controls.

5.1.2 Analytical solution in Bachelier case

Until now the resolution was common for every market impact price model. The affine control in the inventory is a result of the linear-quadratic framework. The average control μ_t is therefore the control of a player with the average inventory, thus easily simplifying the research of the fix point from an infinite dimension to a one dimension search. The distribution of inventories m_t only need to be accounted through its mean E_t.

We will focus on the case of the Bachelier price model where there is an explicit solution.

$$\begin{cases} dP_t = b_P\left(t, P_t, \dfrac{H_t - E_t H_2(t)}{C}\right) dt + \sigma(t, P_t) dW_t \\ dE_t = \frac{H_t - h_2(t)E_t}{C} dt \\ dH_t = \left(\frac{H_t h_2(t)}{C} - A_1 - f_0(t) - \nu \frac{H_t - h_2(t)E_t}{C}\right) dt + Z_t dW_t \\ H_T = 0 \end{cases} \tag{45}$$

Writing $E_t = E(t)$ and $H_t = h_1(t)$ gives the ODE system :

$$\begin{cases} h_1' + f_0(t) + \nu \dfrac{h_1(t) - E(t)h_2(t)}{C} - \dfrac{h_1 h_2}{C} + A_1 + = 0 \\ E'(t) = \frac{h_1(t) - E(t)h_2(t)}{C} \\ E(0) = E_0, h_1(T) = 0 \end{cases} \tag{46}$$

We therefore have :

$$\begin{aligned} E'' &= \frac{-A_1 + \frac{h_1 h_2}{C} - (f_0 + \nu \frac{h_1 - E h_2}{C}) - E' h_2 - E h_2'}{C} \\ \Longrightarrow CE'' &= -A_1 - f_0 + E\frac{h_2^2}{C} - \nu E'(t) - E h_2' \end{aligned} \tag{47}$$

Using the differential equation on h_2 we get the final system :

$$\begin{cases} CE'' + \nu E' - A_2 E = -A_1 - f_0 \\ h_1 = CE' + E h_2 \\ E(0) = E_0, CE'(T) + BE(T) = 0 \end{cases} \tag{48}$$

As it is an inhomogeneous second order linear ordinary differential equation, the solution is easily numerically computable, if not explicit, depending on the form of the seasonality f_0.

The price is therefore $P_t = P_0 + \int_0^t f_0(u) du + \nu(E(t) - E_0) + \sigma W_t$ and $h_0(t) = \frac{\int_t^T h_1^2(s) ds}{2C}$ is the solution of its related equation. The inventory of generic player is $S(t) = E(t) + (S_0 - E_0)\exp\left(-\int_0^t \frac{h_2(u)}{C} du\right)$. The expected gain of a player at time t is therefore $v(t, S_t, P_t) = h_0(t) + (P_t + h_1(t))S_t - h_2(t)\frac{S_t^2}{2}$.

5.1.3 Numerical results

For the numerical results, we take the seasonality of the form :

$$f_0(t) = K cos(2\pi t + \phi)$$

We chose the following set of parameters, (with the Bachelier model, the results do not depend on the volatility so one can take whichever value they want for σ) :

$$T = 1,\ A_1 = 5,\ A_2 = 5,\ B = 10,\ C = 1,\ P_0 = 100,\ E_0 = 0,\ K = 30,\ \phi = \frac{3\pi}{4},\ \nu = 6$$

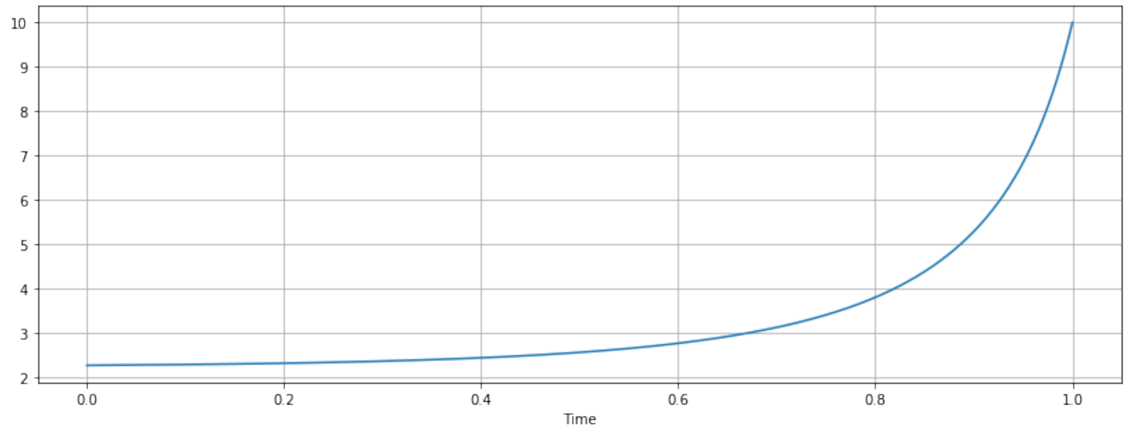

Figure 5.1: $h_2(t)$

This is the shape of h_2, the more we approach maturity the more the players they have interest in having a null inventory.

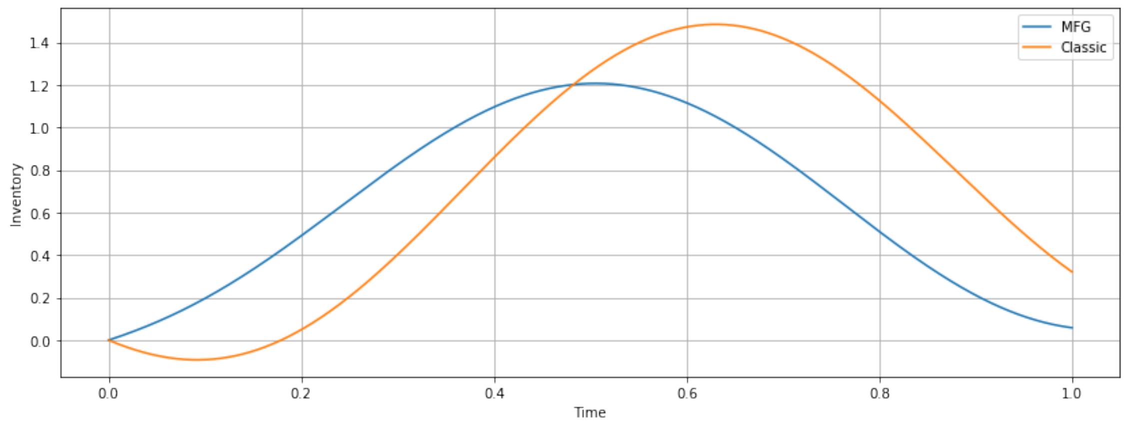

Figure 5.2: $E(t)$

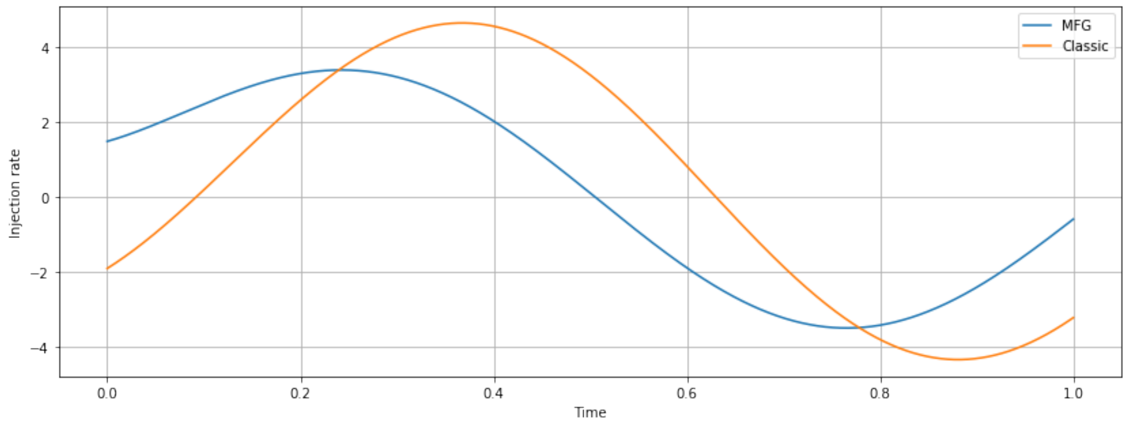

Figure 5.3: $E'(t)$

On these figures, we have plotted the profile of injection of the average player in the Mean Field Game model and in the 'classic' case that is $\nu = 0$. We can see how the players in the MFG tend to buy and sell earlier than if there was no market impact.

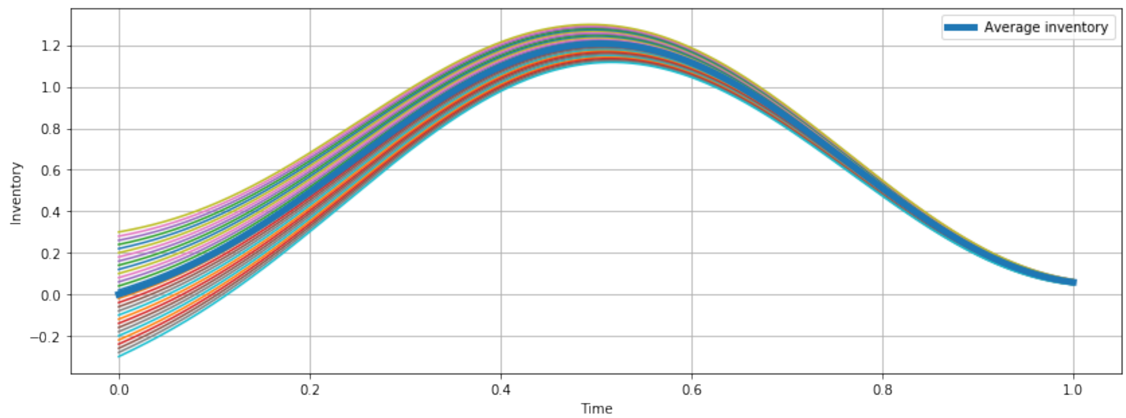

Figure 5.4: S_t of players with different initial inventories

On this figure we've simulated 30 players' trajectories with different starting inventories, and plotted the trajectory of the average inventory of the players. We can seen how their trajectories converge toward the average inventory.

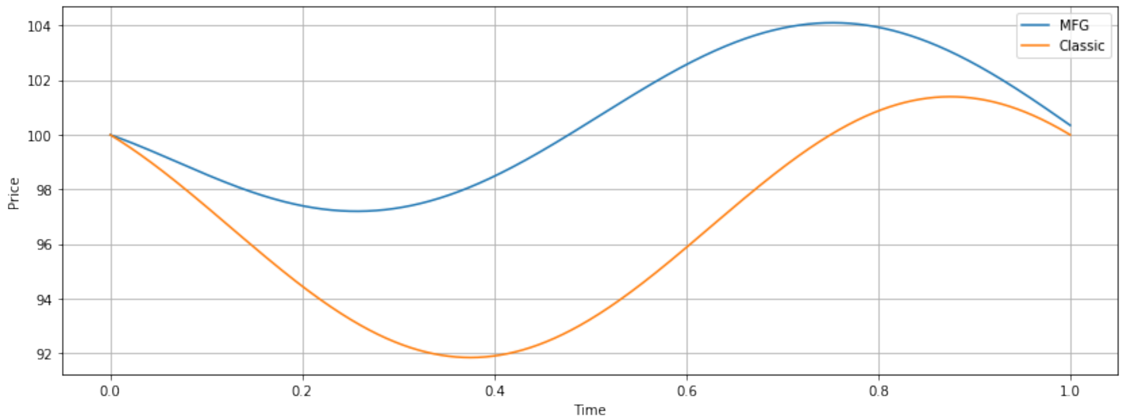

Figure 5.5: P_t with and without market impact

On this figure we can see the resulting price trajectory with the market impact compared with what the trajectory would have been without market impact.

5.2 Price function of demand model

This model comes from the work of Alasseur et al. [2]. They use the same linear quadratic model for the storage than Cardaliaguet and Lehalle. Their model is applied to power storage in smart grid.

5.2.1 Derivation of the equations

We will show and use another approach to solve the problem but both approaches are interchangeable.

$$J(\alpha,\mu) = \mathbb{E}\left[\int_0^T (-\alpha_t P(Q_t+\mu_t) - \frac{C}{2}\alpha_t^2 + A_1 S_t - \frac{A_2}{2} S_t^2)dt + P(Q_T)S_T - \frac{B}{2} S_T^2\right]$$

We use the Gateaux differentiate to find the optimal control.

$d_\beta J(\alpha,\mu) = \mathbb{E}\left[\int_0^T (-\beta_t P(Q_t+\mu_t) - C\alpha_t\beta_t + A_1 S_t^\beta - A_2 S_t S_t^\beta)dt + P(Q_T)S_T^\beta - BS_T S_T^\beta\right]$
with $S_t^\beta = \int_0^t \beta_s ds$

Let Y be the solution of the following backward-stochastic differential equation :

$$\begin{cases} dY_t = -(A_1 - A_2 S_t)dt + Z_t dW_t \\ Y_T = P(Q_T) - BS_T \end{cases} \tag{49}$$

by the Ito lemma we have :

$$d_\beta J(\alpha,\mu) = \mathbb{E}\left[\int_0^T (-P(Q_t+\mu_t) - C\alpha_t + Y_t)\beta_t dt\right]$$

If α is optimal the Gateaux differentiate is null for any β which give the following coupling equation :

$$\alpha_t = \frac{Y_t - P(Q_t + \mu_t)}{C}$$

μ_t is given by the following equation :

$$\begin{aligned} \mu_t &= \int \alpha(t, s, Q_t, E_t) m_t(ds) \\ \mu_t &= \frac{\bar{Y}_t - P(Q_t + \mu_t)}{C} \end{aligned} \tag{50}$$

The system is therefore :

$$\begin{cases} dQ_t = b_Q(t, Q_t)dt + \sigma(t, Q_t)dW_t \\ dS_t = \frac{Y_t - P(Q_t + \mu_t)}{C} dt \\ dY_t = -(A_1 - A_2 S_t)dt + Z_t dW_t \\ Y_T = P(Q_T) - BS_T \\ dE_t = \frac{\bar{Y}_t - P(Q_t + \mu_t)}{C} dt \\ d\bar{Y}_t = -(A_1 - A_2 E_t)dt + Z_t dW_t \\ \bar{Y}_T = P(Q_T) - BE_T \\ \mu_t = \frac{\bar{Y}_t - P(Q_t + \mu_t)}{C} \end{cases} \tag{51}$$

5.2.2 Analytical solution in affine price function

A natural price function is to take :

$$P(q) = p_0 + p_1 q$$

Which gives :

$$\mu_t = \frac{\bar{Y}_t - P(Q_t)}{C + p_1}$$

Assuming $\bar{Y}_t = P(Q_t) + \bar{H}_t - \bar{h_2}(t)E_t$, we have :

$$\begin{aligned} d\bar{Y}_t &= d\bar{H}_t + \left(p_1 b_Q(t, Q_t) - \bar{h_2}'(t)E_t - h_2 \frac{\bar{H}_t - \bar{h_2}(t)E_t}{C + p_1} \right) dt \\ &= -(A_1 - A_2 E_t)dt + Z_t dW_t \end{aligned} \tag{52}$$

Taking $\bar{h_2}$ satisfying :

$$\begin{cases} \bar{h_2}' - \frac{\bar{h_2}^2}{C + p_1} + A_2 = 0 \\ \bar{h_2}(T) = B \end{cases}$$

$$\bar{h_2}(t) = \bar{\rho}(C + p_1) \frac{1 + \bar{c_2} \exp(2\bar{\rho}t)}{1 - \bar{c_2} \exp(2\bar{\rho}t)} \tag{53}$$

With $\bar{\rho} = \sqrt{\frac{A_2}{C+p_1}}$ and $\bar{c_2} = \frac{B - \bar{\rho}(C+p_1)}{B + \bar{\rho}(C+p_1)} \exp(-2\bar{\rho}T)$ we have

$$\begin{cases} d\bar{H}_t = \left(\frac{\bar{h_2}(t)\bar{H}_t}{C + p_1} - A_1 - p_1 b_Q(t, Q_t) \right) dt + Z_t dW_t \\ \bar{H}_T = 0 \end{cases}$$

The solution is :

$$\bar{H}_t = \mathbb{E}_t \left[\int_t^T \exp\left(- \int_t^u \frac{\bar{h_2}(v)}{C + p_1} dv \right) (A_1 + p_1 b_Q(u, Q_u)) du \right]$$

So

$$\mu_t = \frac{\bar{H}_t - \bar{h_2}(t) E_t}{C + p_1}$$

and $dE_t = \mu_t dt = \frac{\bar{H}_t - \bar{h_2}(t)E_t}{C+p_1} dt$ gives :

$$E_t = \exp\left(- \int_0^t \frac{\bar{h_2}(u)}{C + p_1} du \right) E_0 + \frac{1}{C + p_1} \int_0^t \exp\left(- \int_u^t \frac{\bar{h_2}(v)}{C + p_1} dv \right) \bar{H}_u du$$

Now that we know the average inventory and control. We search the optimal control of a generic player, which is finding Y_t. We assume $Y_t = H_t + P(Q_t) - h_2(t) S_t$

$$\begin{aligned} dY_t &= dH_t + \left(p_1 b_Q(t, Q_t) - h_2'(t) S_t - h_2 \frac{H_t - h_2(t) S_t - p_1 \frac{\bar{H}_t - \bar{h}_2(t)E_t}{C+p_1})}{C} \right) dt \\ &= -(A_1 - A_2 S_t) dt + Z_t dW_t \end{aligned} \tag{54}$$

Taking h_2 satisfying :

$$\begin{cases} h_2' - \frac{h_2^2}{C} + A_2 = 0 \\ h_2(T) = B \end{cases}$$

We have :

$$\begin{cases} dH_t = \left(\frac{h_2(t) H_t}{C} - A_1 - p_1 b_Q(t, Q_t) - \frac{h_2(t) p_1}{C} \mu_t \right) dt + Z_t dW_t \\ H_T = 0 \end{cases}$$

Therefore :

$$H_t = \mathbb{E}_t \left[\int_t^T \exp\left(- \int_t^u \frac{h_2(v)}{C} dv \right) \left(A_1 + p_1 b_Q(u, Q_u) + \frac{h_2(u) p_1}{C} \frac{\bar{H}_u - \bar{h}_2(u) E_u}{C + p_1} \right) du \right]$$

and $dS_t = \alpha_t dt = \frac{H_t - h_2(t) S_t - p_1 \mu_t}{C} dt$ gives :

$$S_t = \exp\left(- \int_0^t \frac{h_2(u)}{C} du \right) S_0 + \frac{1}{C} \int_0^t \exp\left(- \int_u^t \frac{h_2(v)}{C} dv \right) (H_u - p_1 \mu_t) du$$

The expected payoff a time t is :

$$V_t = \mathbb{E}_t \left[\int_t^T -\alpha_u P(Q_u + \mu_u) - \frac{\alpha_u^2}{2C} + A_1 S_u - A_2 \frac{S_u^2}{2} du + P(Q_T) S_T - B \frac{S_T^2}{2} \right]$$

$$\Longrightarrow V_t = H_t^0 + (H_t + P(Q_t)) S_t - h_2(t) \frac{S_t^2}{2}$$

With $H_t^0 = \mathbb{E}_t \left[\int_t^T \frac{\left(H_u - p_1 \frac{\bar{H}_u - \bar{h}_2(u) E_u}{C+p_1} \right)^2}{2C} du \right]$

5.2.3 Numerical results

We take the exogenous demand as an Ornstein-Uhlenbeck process with a seasonality trend :

$$dQ_t = -a(Q_t - f_0(t))dt + \sigma dW_t$$

The main difficulty to have numerical results is to compute the conditional expectations. Fortunately with this demand process, most of them are computable only using deterministic quadrature methods. Only H_t^0 require a Monte-Carlo simulation to be computed.

$\mathbb{E}_t[b_Q(u, Q_u)] = \mathbb{E}_t[-a(Q_u - f_0(u))]$

$\mathbb{E}_t[Q_u] = e^{-a(u-t)}\left(Q_t + \int_t^u af(s)e^{a(s-t)}ds\right) = f_Q(u, t, Q_t)$

Therefore : $\bar{H}_t = \int_t^T \exp\left(-\int_t^u \frac{\bar{h_2}(v)}{C+p_1}dv\right)(A_1 - ap_1(f_Q(u, t, Q_t) - f_0(u)))du = \bar{h}(t, t, Q_t)$

$E_t = \exp\left(-\int_0^t \frac{\bar{h_2}(u)}{C+p_1}du\right)E_0 + \frac{1}{C+p_1}\int_0^t \exp\left(-\int_u^t \frac{\bar{h_2}(v)}{C+p_1}dv\right)\bar{h}(u, u, Q_u)du = e(t, 0, E_0,)$

$\mathbb{E}_t[\bar{H}_u] = \int_u^T \exp\left(-\int_u^v \frac{\bar{h_2}(w)}{C+p_1}dw\right)(+A_1 - ap_1(f_Q(v, t, Q_t) - f_0(v)))dv = \bar{h}(u, t, Q_t)$

$\mathbb{E}_t[E_u] = \exp\left(-\int_t^u \frac{\bar{h_2}(v)}{C+p_1}dv\right)E_t + \frac{1}{C+p_1}\int_t^u \exp\left(-\int_v^u \frac{\bar{h_2}(w)}{C+p_1}dw\right)\bar{h}(v, t, Q_t)dv = e(u, t, E_t, Q_t)$

$H_t = \int_t^T \exp\left(-\int_t^u \frac{h_2(v)}{C}dv\right)\left(A_1 - ap_1(f_Q(u, t, Q_t) - f_0(u)) + \frac{h_2(u)p_1}{c}\frac{\bar{h}(u,t,Q_t) - \bar{h}_2(u)e(u,t,E_t,Q_t)}{C+p_1}\right)du = h(t, Q_t, E_t)$

For the numerical results, we take the seasonality of the form :

$$f_0(t) = Kcos(2\pi t + \phi)$$

We chose the following set of parameters :

$$T = 1,\ A_1 = 5,\ A_2 = 5,\ B = 10,\ C = 1,\ P_0 = 100,\ p_1 = 1,\ E_0 = 0,\ a = 5,\ K = 5,\ \phi = \frac{3\pi}{4},\ \sigma = 1$$

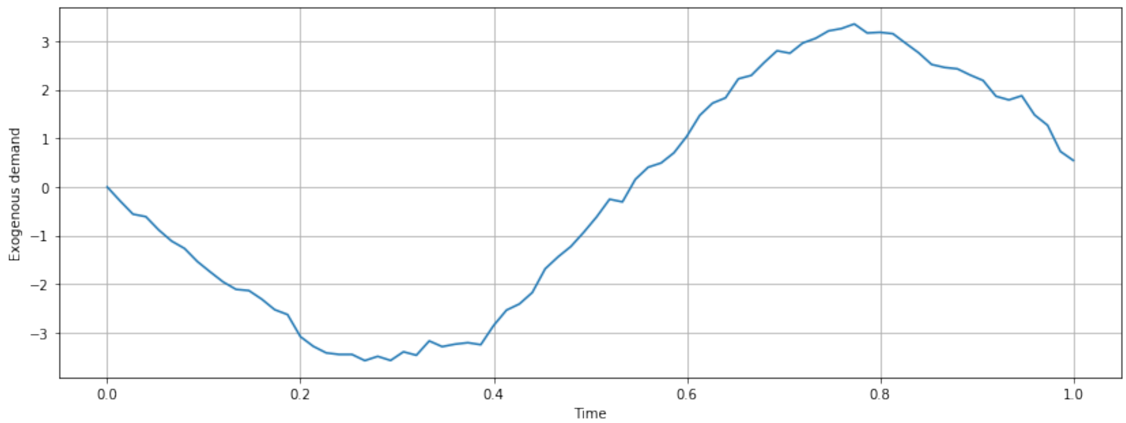

Figure 5.6: Q_t

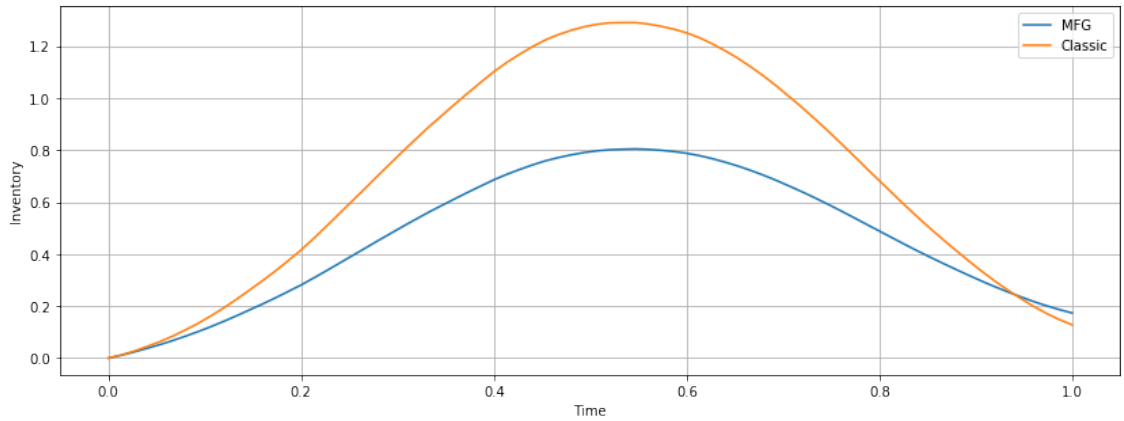

Figure 5.7: E_t classic vs MFG

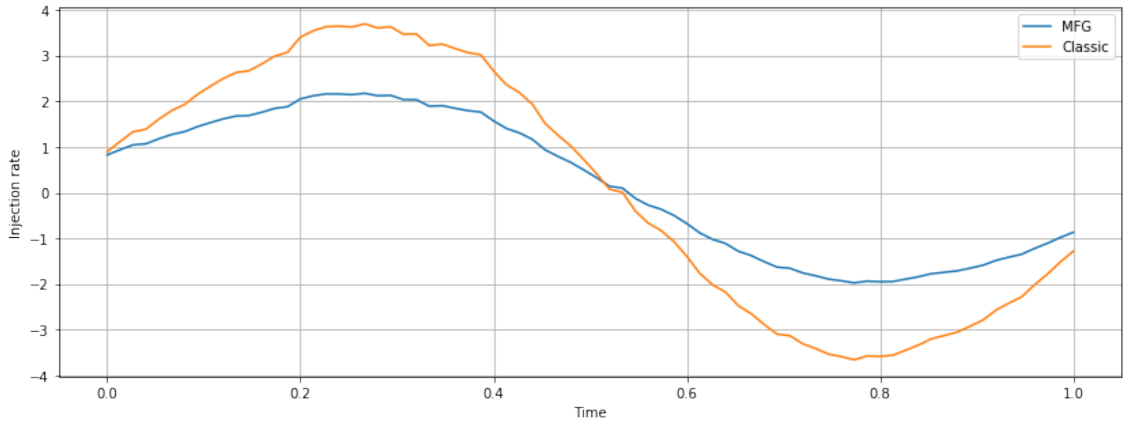

Figure 5.8: μ_t classic vs MFG

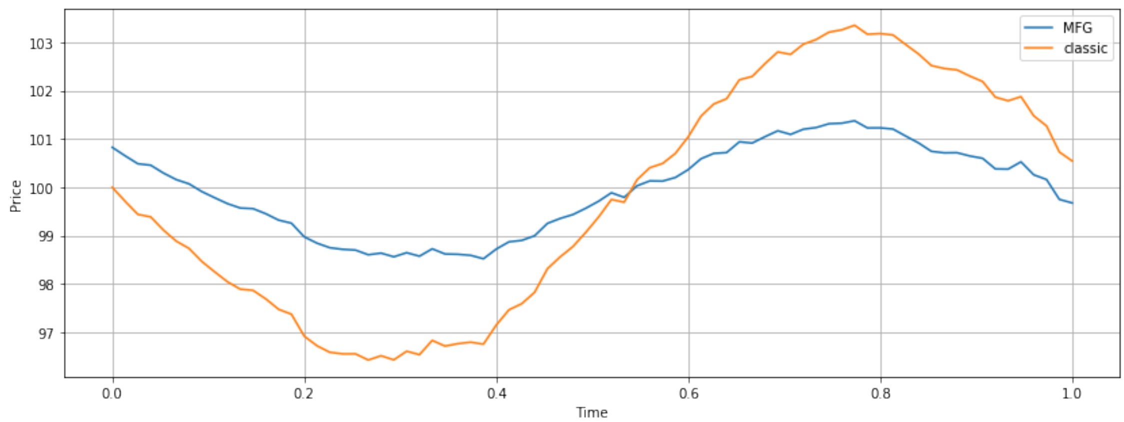

Figure 5.9: P_t classic vs MFG

We notice that, contrary to the market impact models, the MFG strategy is not in advance of the classic strategy. It seems that the players buy and sell less than what they would have if the price did not depend on μ. The interesting feature is that the price seasonality is flattened compared to the classic case.

5.3 General price model

This model is a combination of the two models from Cardaliaguet and Alasseur.

5.3.1 Derivation of the equations

We recall the payoff function :

$$J(\alpha,\mu) = \mathbb{E}\left[\int_0^T (-\alpha_t P(t,X_t,\mu_t) - \frac{C}{2}\alpha_t^2 + A_1 S_t - \frac{A_2}{2}S_t^2)dt + \tilde{P}(X_T)S_T - \frac{B}{2}S_T^2\right]$$

We use the Gateaux differentiate to find the optimal control.
$d_\beta J(\alpha,\mu) = \mathbb{E}\left[\int_0^T (-\beta_t P(t,X_t,\mu_t) - C\alpha_t\beta_t + A_1 S_t^\beta - A_2 S_t S_t^\beta)dt + \tilde{P}(X_T)S_T^\beta - BS_T S_T^\beta\right]$
with $S_t^\beta = \int_0^t \beta_s ds$

Let Y be the solution of the following backward-stochastic differential equation :

$$\begin{cases} dY_t = -(A_1 - A_2 S_t)dt + Z_t dW_t \\ Y_T = \tilde{P}(Q_T) - BS_T \end{cases} \tag{55}$$

by the Ito lemma we have :

$d_\beta J(\alpha,\mu) = \mathbb{E}\left[\int_0^T (-P(t,X_t,\mu_t) - C\alpha_t + Y_t)\beta_t dt\right]$

If α is optimal the Gateaux differentiate is null for any β which give the following coupling equation :

$\alpha_t = \frac{Y_t - P(t,X_t,\mu_t)}{C}$

μ_t is given by the following equation :

$$\begin{aligned} \mu_t &= \int \alpha(t,s,X_t,E_t)m_t(ds) \\ \mu_t &= \frac{\bar{Y}_t - P(t,X_t,\mu_t)}{C} \end{aligned} \tag{56}$$

The final Mean Field Game system is therefore :

$$\begin{cases} dX_t = b(t,X_t,\mu_t)dt + \sigma(t,X_t,\mu_t)dW_t \\ dS_t = \frac{Y_t - P(t,X_t,\mu_t)}{C}dt \\ dY_t = -(A_1 - A_2 S_t)dt + Z_t dW_t \\ Y_T = \tilde{P}(X_T) - BS_T \\ dE_t = \frac{\bar{Y}_t - P(t,X_t,\mu_t)}{C}dt \\ d\bar{Y}_t = -(A_1 - A_2 E_t)dt + Z_t dW_t \\ \bar{Y}_T = \tilde{P}(X_T) - BE_T \\ \mu_t = \frac{\bar{Y}_t - P(t,X_t,\mu_t)}{C} \end{cases} \tag{57}$$

5.3.2 Analytical solution in Bachelier permanent market impact and affine instantaneous market impact

Let's imagine that there is an economical price, or fair price, process P_t following a Bachelier dynamic with permanent market impact, and that the price paid by the players at time t is this economical price plus a spread due to the market condition : $p_1\mu_t$.

$$\begin{cases} dP_t = (f_0(t) + \nu\mu_t)dt + \sigma dW_t \\ dS_t = \frac{Y_t - P_t - p_1\mu_t}{C}dt \\ dY_t = -(A_1 - A_2 S_t)dt + Z_t dW_t \\ Y_T = P_T - BS_T \\ dE_t = \frac{\bar{Y}_t - P_t - p_1\mu_t)}{C}dt \\ d\bar{Y}_t = -(A_1 - A_2 E_t)dt + Z_t dW_t \\ \bar{Y}_T = P_T - BE_T \\ \mu_t = \frac{\bar{Y}_t - P_t - p_1\mu_t}{C} \end{cases} \tag{58}$$

Assuming $\bar{Y}_t = P_t + \bar{H}_t - \bar{h_2}(t)E_t$, we have :

$$\mu_t = \frac{\bar{H}_t - \bar{h_2}(t)E_t}{C + p_1}$$

The function h_2 satisfies :

$$\begin{cases} \bar{h_2}' - \frac{\bar{h_2}^2}{C + p_1} + A_2 = 0 \\ \bar{h_2}(T) = B \end{cases}$$

Therefore :

$$\bar{h_2}(t) = \bar{\rho}(C + p_1)\frac{1 + \bar{c_2}\exp(2\bar{\rho}t)}{1 - \bar{c_2}\exp(2\bar{\rho}t)} \tag{59}$$

With $\bar{\rho} = \sqrt{\frac{A_2}{C+p_1}}$ and $\bar{c_2} = \frac{B - \bar{\rho}(C+p_1)}{B + \bar{\rho}(C+p_1)}$

It remains :

$$\begin{cases} dE_t = \frac{\bar{H} - \bar{h}_2(t)E_t}{C + p_1}dt \\ d\bar{H}_t = \left(\frac{\bar{h}_2(t)\bar{H}_t}{C+p_1} - A_1 - f_0(t) - \nu\frac{\bar{H} - \bar{h}_2(t)E_t}{C+p_1}\right)dt + Z_t dW_t \\ \bar{H}_T = 0 \end{cases}$$

Taking $E_t = E(t)$ with E solution of this equation :

$$\begin{cases} (C + p_1)E'' + \nu E' - A_2 E = -A_1 - f_0 \\ E(0) = E_0, (C + p_1)E'(T) + BE(T) = 0 \end{cases} \tag{60}$$

and $\bar{H}_t = \bar{h}_1(t) = (C + p_1)E(t)' + E(t)\bar{h}_2(t)$, we have the solution for the FBSDE on E and $\bar{H}$. We also have $P_t = P_0 + \int_0^t f_0(u)du + \nu(E(t) - E_0) + \sigma W_t$ Now it remains to solve the system

$$\begin{cases} dS_t = \frac{Y_t - P_t - p_1\mu_t}{C}dt \\ dY_t = -(A_1 - A_2 S_t)dt + Z_t dW_t \\ Y_T = P_T - BS_T \end{cases} \tag{61}$$

Assuming $Y_t = P_t + H_t - h_2(t)S_t$ we have :

The function h_2 satisfies

$$\begin{cases} h_2{}' - \dfrac{h_2^2}{C} + A_2 = 0 \\ h_2(T) = B \end{cases}$$

Therefore :

$$h_2(t) = \rho C \frac{1 + c_2 \exp(2\rho t)}{1 - c_2 \exp(2\rho t)} \tag{62}$$

With $\rho = \sqrt{\frac{A_2}{C}}$ and $c_2 = \frac{B-\rho C}{B+\rho C}$

It remains :

$$\begin{cases} dS_t = \dfrac{H_t - h_2(t)S_t - p_1\mu_t}{C} dt \\ dH_t = \left(\frac{h_2(t)H_t}{C} - A_1 - f_0(t) - \left(\nu + \frac{h_2(t)p_1}{C}\right) E'(t)\right) dt + Z_t dW_t \\ H_T = 0 \end{cases}$$

Then

$$H_t = h_1(t) = \int_t^T \exp\left(-\int_t^u \frac{h_2(v)}{C} dv\right) \left(A_1 + f_0(u) + \left(\nu + \frac{h_2(u)p_1}{C}\right) E'(u)\right) du$$

$$S_t = \exp\left(-\int_0^t \frac{h_2(v)}{C} dv\right) S_0 + \int_0^t \exp\left(-\int_u^t \frac{h_2(v)}{C} dv\right) \frac{h_1(u) - p_1 E'(u)}{C} du$$

The expected payoff at time t is :

$$V_t = \mathbb{E}_t \left[\int_t^T \left[-\alpha_u(P_u + p_1\mu_u) - \frac{\alpha_u^2}{2C} + A_1 S_u - A_2 \frac{S_u^2}{2}\right] du + P_T S_T - B\frac{S_T^2}{2}\right]$$

$$\Longrightarrow V_t = v(t, S_t, P_t) = h_0(t) + (h_1(t) + P_t)S_t - h_2(t)\frac{S_t^2}{2}$$

With $h_0(t) = \int_t^T \frac{(h_1(u) - p_1 E'(u))^2}{2C} du$

5.3.3 Numerical results

For the numerical results, we take the seasonality of the form :

$$f_0(t) = Kcos(2\pi t + \phi)$$

We chose the following set of parameters, (the parameter σ is not relevant, we set it to zeros to see the expectation of the price) :

$$T = 1,\ A_1 = 5,\ A_2 = 5,\ B = 10,\ C = 1,\ P_0 = 100,\ p_1 = 1,\ E_0 = 0,\ a = 5,\ \nu = 6,\ K = 5,\ \phi = \frac{3\pi}{4}$$

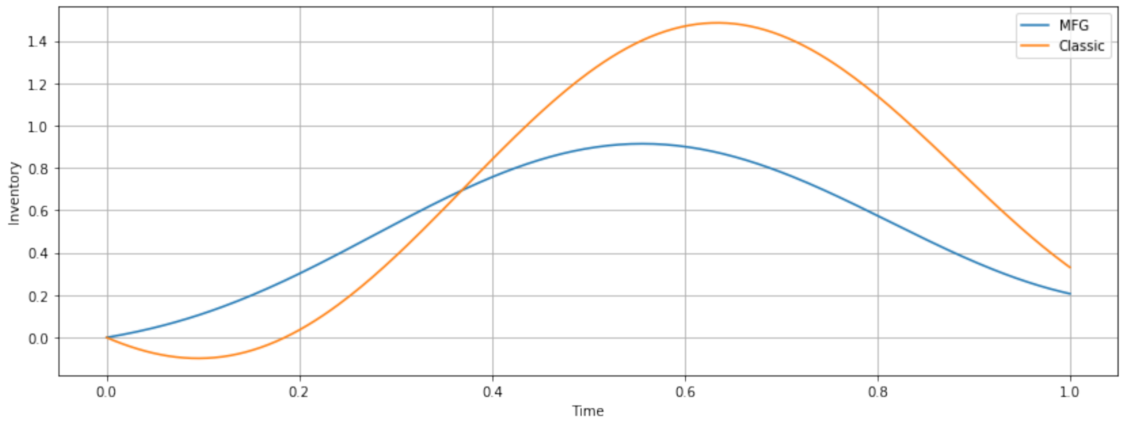

Figure 5.10: $E(t)$ classic vs MFG

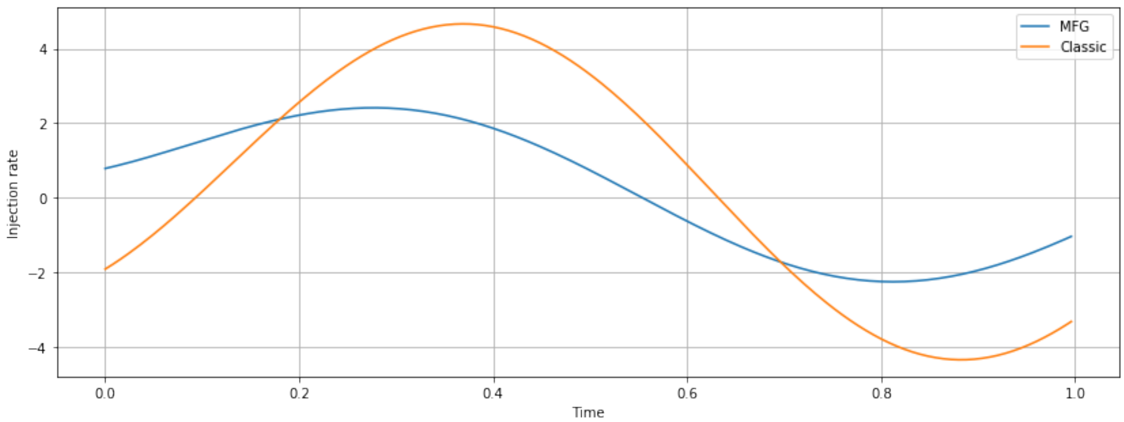

Figure 5.11: $\mu(t)$ classic vs MFG

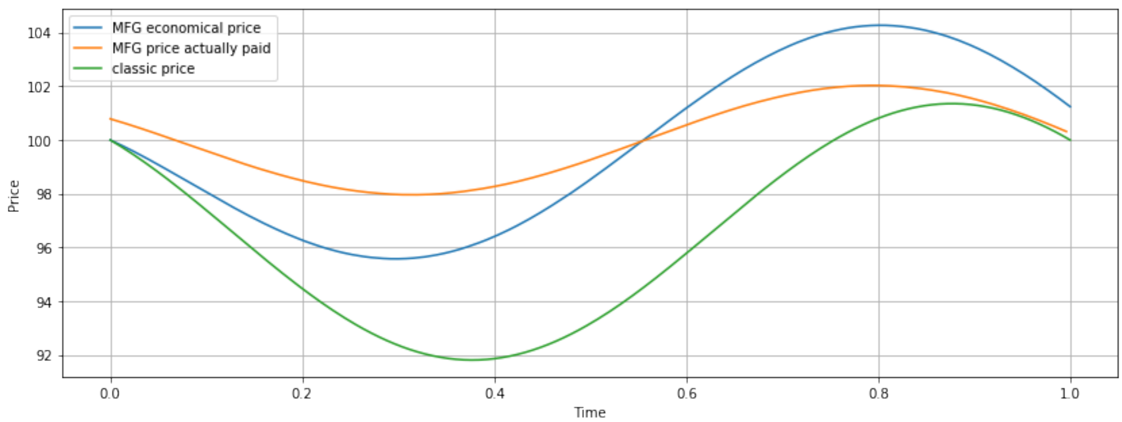

Figure 5.12: P_t classic vs MFG

We can see a combination of features from the two previous subsections. The players buy and sell less and they do it earlier than in the classic case.

6 Numerical methods to solve the system for market impact price models

Let's recall the system of equations in the general case :

$$\begin{cases} dH_t = \left[\frac{H_t h_2(t)}{C} - A_1 - b_P\left(t, P_t, \frac{H_t - E_t h_2(t)}{C}\right)\right] dt + Z_t dW_t \\ dP_t = b_P\left(t, P_t, \frac{H_t - E_t h_2(t)}{C}\right) dt + \sigma_P(t, P_t) dW_t \\ dE_t = \frac{H_t - E_t h_2(t)}{C} dt \\ H_T = 0 \end{cases} \tag{63}$$

which linked to following backward PDE :

$$\begin{cases} \partial_t h + \frac{1}{2}\sigma_P^2 \partial_{pp} h + (1 + \partial_p h) b_P\left(t, p, \frac{h(t,p,e) - e h_2(t)}{C}\right) - \frac{h h_2}{C} + A_1 + \partial_e h \frac{h - e h_2}{C} = 0 \\ h(T) = 0 \end{cases} \tag{64}$$

We have already seen that when the drift do not depend of the price itself, like in the Bachelier case, the solution for H_t and E_t is deterministic and P_t is the solution of a simple forward SDE. We therefore focus on cases where the drift of the price depend on the price.

From the literature on FBSDEs, Douglas et al. [23] and [40] et al. focus on the four step scheme developed by Ma et al. [39] which consist primarily in solving the PDE. Delarue and Menozzi [22] on the other hand, focus on solving directly the FBSDE, their method involve quantization. Ludwig et al. [38] applies this approach to stochastic control, being one of the first method involving solving the FBSDE rather than the Hamilton-Jacobi-Bellman PDE for these kinds of problems. On our side, we will focus on solving the PDE's version of the problem as the numerical methods are more simple to implement and that quantization methods are a theme that have already been treated in a previous internship in BP.

6.1 Finite differences method

The numerical method we used to solve the PDEs is the finite differences. Our main reference for this numerical scheme i the lecture notes of the lectures of Touzi at the Fields Institute [43]. Finite differences methods consist in discretizing the variable space and solve differences equations. We also need to solve the PDE on a bounded variable space in order to have a finite number of equations (therefore solvable by a computer), thus we need additional boundary conditions.

We denote by $\delta_t = \frac{T}{n}$ the time step and δ_p and δ_e the price and inventory steps. The variable grid is

$$(t_k, p_i, e_j)_{k \in [\![0;n]\!], i \in [\![-n_p-1;n_p+1]\!], j \in [\![-n_e-1;n_e+1]\!]} = (kh, p_0 + i\delta_p, e_0 + j\delta_e)_{k \in [\![0;n]\!], i \in [\![-n_p-1;n_p+1]\!], j \in [\![-n_e-1;n_e+1]\!]}$$

The discretized function $(h(t_k, p_i, e_j))_{k,i,j}$ is denoted by the table $H_{k,i,j}$. We use the following finite differences for the derivatives:

$$\partial_t h \leftarrow \Delta_t H_{k,i,j} = \frac{H_{k+1,i,j} - H_{k,i,j}}{\delta_t}$$

$$\partial_p h \leftarrow \Delta_p H_{k,i,j} = \frac{H_{k,i+1,j} - H_{k,i-1,j}}{2\delta_p}$$

$$\partial_{pp} h \leftarrow \Delta_{pp} H_{k,i,j} = \frac{H_{k,i+1,j} - 2H_{k,i,j} + H_{k,i-1,j}}{\delta_p^2}$$

$$\partial_e h \leftarrow \Delta_e H_{k,i,j} = \frac{H_{k,i,j+1} - H_{k,i,j-1}}{2\delta_e}$$

We get the following differences system

$$\left\{\begin{array}{l}\forall k\in[\![0;n-1]\!],i\in[\![-n_p;n_p]\!],j\in[\![-n_e;n_e]\!]:\Delta_t H_{k,i,j}+\frac{1}{2}\sigma^2(t_k,p_i)\Delta_{pp}H_{k,i,j}+(1+\Delta_p H_{k,i,j})b_P\left(t_k,p_i,\frac{H_{k,i,j}-e_jh_2(t_k)}{C}\right)\\ \qquad -\frac{H_{k,i,j}h_2(t_k)}{C}+A_1+\Delta_e H_{k,i,j}\frac{H_{k,i,j}-e_jh_2(t_k)}{C}=0\end{array}\right. \tag{65}$$

Solving this system is finding the zeros of a function $F:\mathbb{R}^{n(2n_p+1)(2n_e+1)}\rightarrow\mathbb{R}^{n(2n_p+1)(2n_e+1)}$. We use Newton-Raphson method to find this zero.

The initial PDE is set on an open domain $\mathbb{R}_+^*\times\mathbb{R}$ so without boundary conditions, when solve the equation on a closed bounded domain, we need to set conditions on the boundaries. We can only chose them arbitrarily since we have no clue of what are the actual solution's values at those points. Like in Thompson et al. [42] suggest, we assume that $\partial_{pp}H\xrightarrow[p\to 0]{}0$ and $\partial_{pp}H\xrightarrow[p\to\infty]{}0$. For the conditions on the lower and upper boundaries, we set $h=0$. However we have two conditions while the equation is of first order in e so this result in having spurious oscillations on the function in the e direction.

The complete system is :

$$\left\{\begin{array}{l}\forall k\in[\![0;n-1]\!],i\in[\![-n_p;n_p]\!],j\in[\![-n_e;n_e]\!]:\Delta_t H_{k,i,j}+\frac{1}{2}\sigma^2(t_k,p_i)\Delta_{pp}H_{k,i,j}+(1+\Delta_p H_{k,i,j})b_P\left(t_k,p_i,\frac{H_{k,i,j}-e_jh_2(t_k)}{C}\right)\\ \qquad -\frac{H_{k,i,j}h_2(t_k)}{C}+A_1+\Delta_e H_{k,i,j}\frac{H_{k,i,j}-e_jh_2(t_k)}{C}=0\\ \forall i,j,\ H_{n,i,j}=0\\ \forall k\in[\![0;n-1]\!]\forall j,\ \Delta_{pp}H_{k,n_p,j}=0\\ \forall k\in[\![0;n-1]\!]\forall j,\ \Delta_{pp}H_{k,-n_p,j}=0\\ \forall k\in[\![0;n-1]\!]\forall i\in[\![-n_p;n_p]\!],\ H_{k,i,\pm(n_e+1)}=0\end{array}\right. \tag{66}$$

Remark 11

This is the implicit Finite differences method : the values of $H_{k,i,j}$ cannot be computed directly, we need to solve an equation. Another method is the explicit differences method : the approximate of the time derivative is $\Delta_t H_{k,i,j}=\frac{H_{k,i,j}-H_{k-1,i,j}}{\delta_t}$. In this case at each time t_k we can compute the values $H_{k,i,j}$ explicitly from the values at time t_{k+1}. The explicit method is therefore less complex than the implicit method, the downside is that it is not stable.

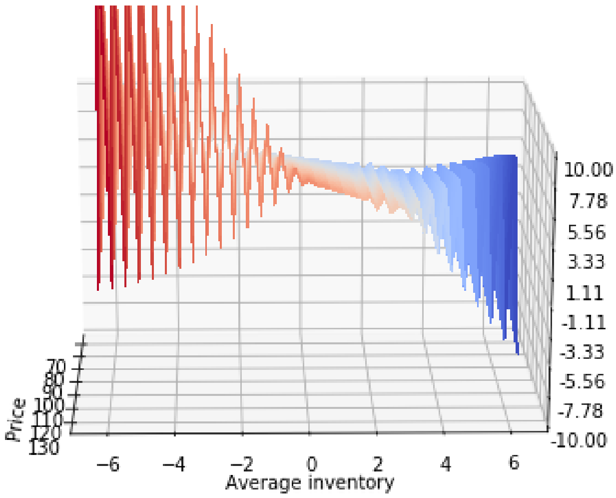

Figure 6.1: $H(t,p,e)$ at $t = \frac{T}{2}$, big spurious oscillations

On this figure we can see the spurious oscillations on the result. We use a brute-forcing algorithm to solve this issue : we repeat the algorithm replacing at each step the boundaries conditions on e like this.

$$\begin{cases} H^{m+1}[k,i,-n_e] = \dfrac{H^m[k,i,-n_e] + H^m[k,i,-n_e+1]}{2} \\ H^{m+1}[k,i,n_e] = \frac{H^m[k,i,n_e]+H^m[k,i,n_e-1]}{2} \end{cases}$$

The algorithm stop when the difference between two iterations is arbitrarily small.

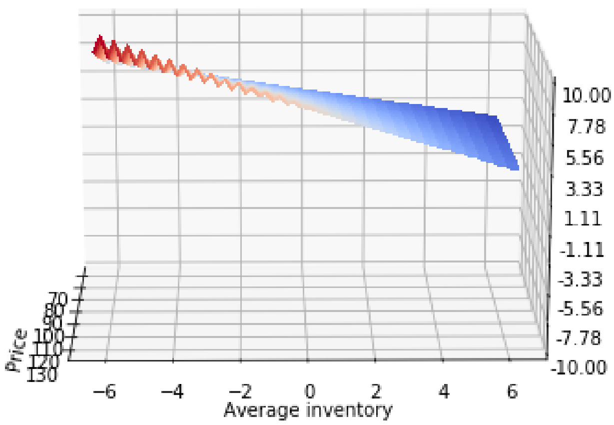

Figure 6.2: $H(t,p,e)$ at $t = \frac{T}{2}$, very small spurious oscillations

This figure shows how the oscillations of the previous figure are greatly reduced after some iterations. We are however not entirely satisfied because it require solving a non linear PDE multiple times which largely increases the complexity and calculus time of the algorithm. Another solution would be to use first order implicit or explicit approximate in e with a bounding condition on only one value of e, but for some reason this method explode.

6.1.1 Example : Black-Scholes model

The PDE to solve is :

$$\begin{cases} \partial_t h + \frac{1}{2}\sigma^2 p^2 \partial_{pp} h + (1+\partial_p h)p\left(f_0 + \nu \frac{h - eh_2}{C}\right) - \frac{hh_2}{C} + A_1 + \partial_e h \frac{h-eh_2}{C} = 0 \\ h(T) = 0 \end{cases} \tag{67}$$

The difference system is :

$$\begin{cases} \forall k \in [\![0; n-1]\!], i \in [\![-n_p; n_p]\!], j \in [\![-n_e; n_e]\!] : \Delta_t H_{k,i,j} + \frac{1}{2}\sigma^2 p_i^2 \Delta_{pp} H_{k,i,j} + (1+\Delta_p H_{k,i,j}) p_i \left(f_0(t_k) + \nu \frac{H_{k,i,j} - e_j h_2(t_k)}{C}\right) \\ \qquad - \frac{H_{k,i,j} h_2(t_k)}{C} + A_1 + \Delta_e H_{k,i,j} \frac{H_{k,i,j} - e_j h_2(t_k)}{C} = 0 \\ \forall i, j,\ H_{n,i,j} = 0 \\ \forall k \in [\![0; n-1]\!],\ \forall j, H_{k,\pm(n_p+1),j} = BC_{k,\pm(n_p+1),j} \\ \forall k \in [\![0; n-1]\!],\ \forall i, H_{k,i,\pm(n_e+1)} = BC_{k,i,\pm(n_e+1)} \end{cases} \tag{68}$$

For the numerical example, we used for the problem parameters :

$$T = 1,\ A_1 = 5,\ A_2 = 16,\ B = 2,\ C = 1,\ P_0 = 100,\ E_0 = 0,\ K = 0.3,\ \phi = \frac{3\pi}{4},\ \nu = 0.01,\ \sigma = 0.1$$

and for the numerical method parameters :

$$n = 300,\ \delta_t = T/(n-1),\ n_p = 10,\ n_e = 30,\ \delta_p = 3,\ \delta_e = 0.2,$$

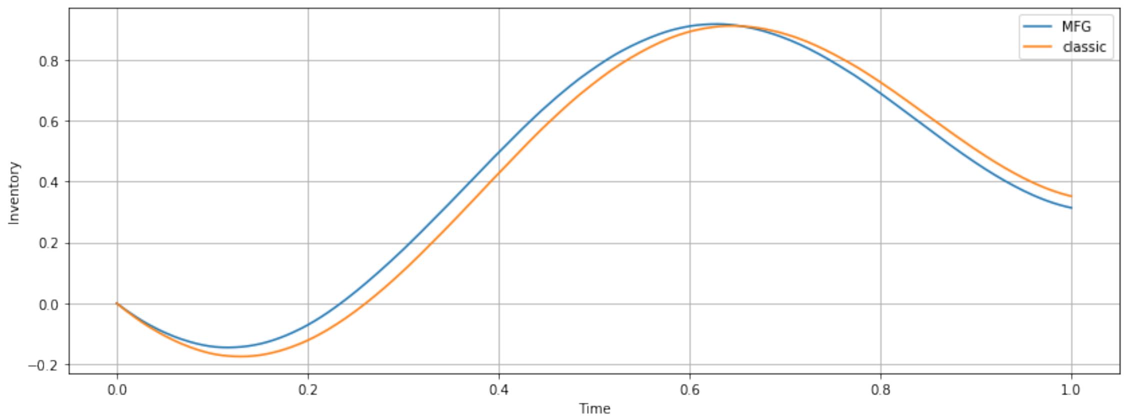

Figure 6.3: E_t with and without market impact

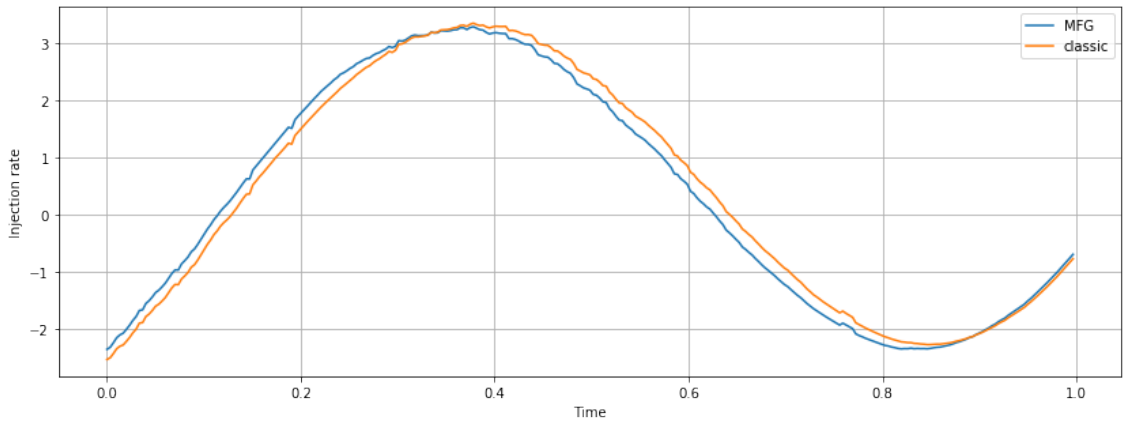

Figure 6.4: μ_t with and without market impact

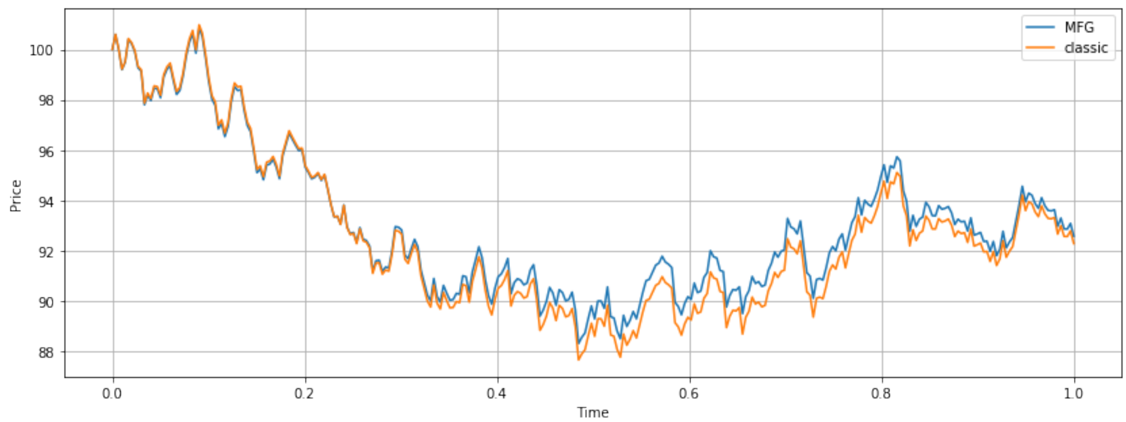

Figure 6.5: P_t with and without market impact

From the returned function $H(t,p,e)$, we simulated one trajectory for the tuple (P_t, E_t, μ_t) for the Mean Field Game and for the classical case using the same path of Brownian motion. The solution for the classical case comes from solving the forward-backward SDE directly : when $\nu = 0$ one can verify that the following expression is solution of the backward equation.

$$H_t = \mathbb{E}_{t,P_t}\left[\int_t^T \exp\left(-\int_t^s \frac{h_2(u)}{C}du\right)(A_1 + P_s f_0(s))\right]$$

The conditional expectation has closed formula because an expected Black-Scholes price in the futur knowing the price today has one. Having this exact solution closed formula of problem tied to the MFG problem allows us to be confident in the result return by the algorithm.

On these figures we can seen how in the Mean Field Game equilibrium strategy, the average player seems to buy and sell earlier, like we've seen with the Bachelier price model.

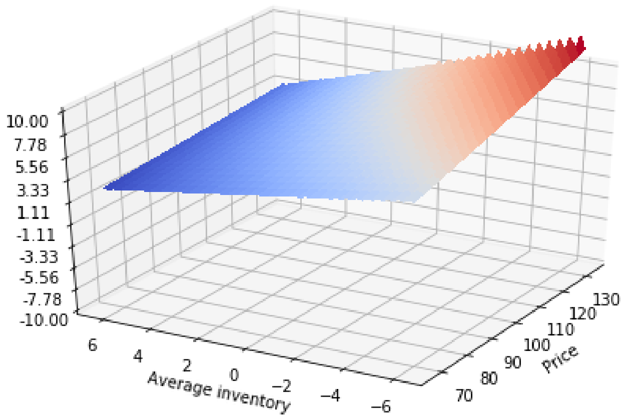

Figure 6.6: $H(t,p,e)$ at $t = \frac{T}{2}$

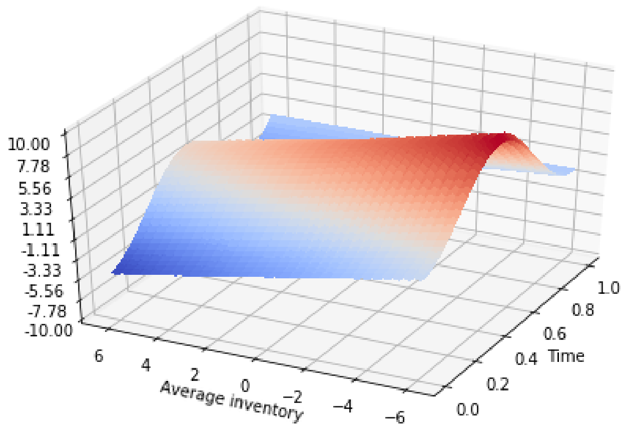

Figure 6.7: $H(t,p,e)$ at $p = p_0$

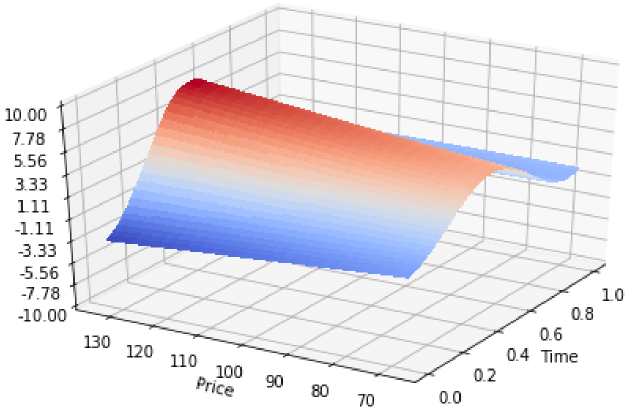

Figure 6.8: $H(t, p, e)$ at $e = 0$

On these three figure we've plotted the function H with respect to its arguments. Since the space of argument is three dimensional, we can only have 2-D projections of it to represent the function in a 3-D space. The 4-th dimension being the third argument not used, we can however use a cursor to simulate this fourth dimension. We can notice that the function H is decreasing in e. This can be expected because when E_t is large, μ_t is more likely to be low, meaning that the price is more likely to go down, therefore it is better to sell now, and vice-versa. For the monotony in price however, one would expect it to be decreasing in price because the lower the price is the more you buy. Actually the expected variation of price are proportional to the price there for when the price is expected to up at a certain moment, the more the price is already high the more you will make profit by buying it now and selling it later. This is how at some point in time, the function H is increasing in price : the more the price is high the more you want to buy.

6.1.2 Example : Clowlow-Strickland 1 factor model

The PDE to solve is :

$$\begin{cases} \partial_t h + \frac{1}{2}\sigma^2 p^2 \partial_{pp} h + (1 + \partial_p h) p \left(\frac{F_0'}{F_0} + a\left(\ln(F_0) - \ln(p)\right) + \nu \frac{h - eh_2}{C} \right) - \frac{hh_2}{C} + A_1 + \partial_e h \frac{h - eh_2}{C} = 0 \\ h(T) = 0 \end{cases} \quad (69)$$

The difference system is :

$$\begin{cases} \forall k \in [\![0; n-1]\!], i \in [\![-n_p; n_p]\!], j \in [\![-n_e; n_e]\!] : \Delta_t H_{k,i,j} + \frac{1}{2}\sigma^2 p_i^2 \Delta_{pp} H_{k,i,j} + (1 + \Delta_p H_{k,i,j}) p_i \left(\frac{F_0'(t_k)}{F_0(t_k)} + a\left(\ln(F_0(t_k)) - \ln(p_i \right. \\ \qquad - \frac{H_{k,i,j} h_2(t_k)}{C} + A_1 + \Delta_e H_{k,i,j} \frac{H_{k,i,j} - e_j h_2(t_k)}{C} = 0 \\ \forall i, j, \; H_{n,i,j} = 0 \\ \forall k \in [\![0; n-1]\!], \; \forall j, H_{k,\pm(n_p+1),j} = BC_{k,\pm(n_p+1),j} \\ \forall k \in [\![0; n-1]\!], \; \forall i, H_{k,i,\pm(n_e+1)} = BC_{k,i,\pm(n_e+1)} \end{cases} \quad (70)$$

For the numerical example, we used for the problem parameters :

$$T = 1, \; A_1 = 5, \; A_2 = 25, \; B = 5, \; C = 1, \; P_0 = 100, \; E_0 = 0, \; K = 0.3, \; \phi = \frac{\pi}{4}, \; \nu = 0.015, \; \sigma = 0.15, \; a = 5$$

and for the numerical method parameters :

$$n = 300,\ \delta_t = T/(n-1),\ n_p = 10,\ n_e = 15,\ \delta_p = 3,\ \delta_e = 0.2$$

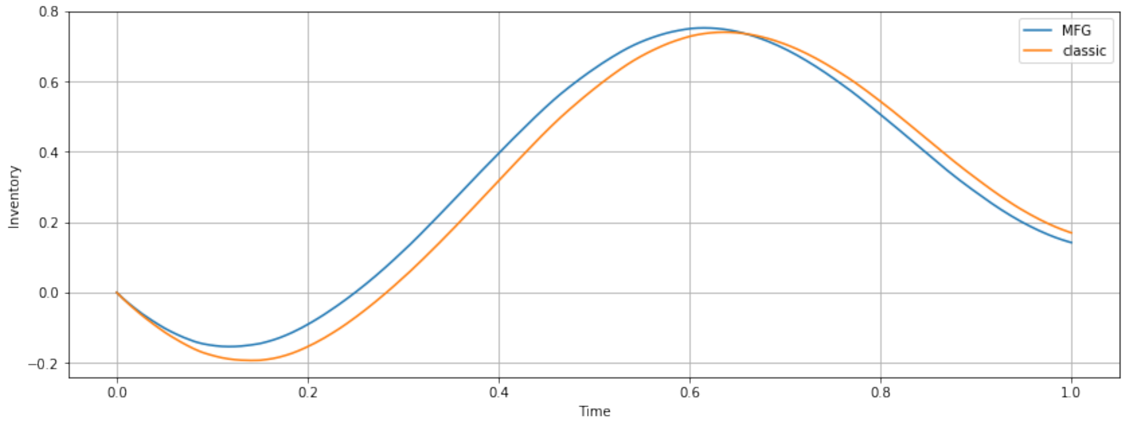

Figure 6.9: E_t with and without market impact

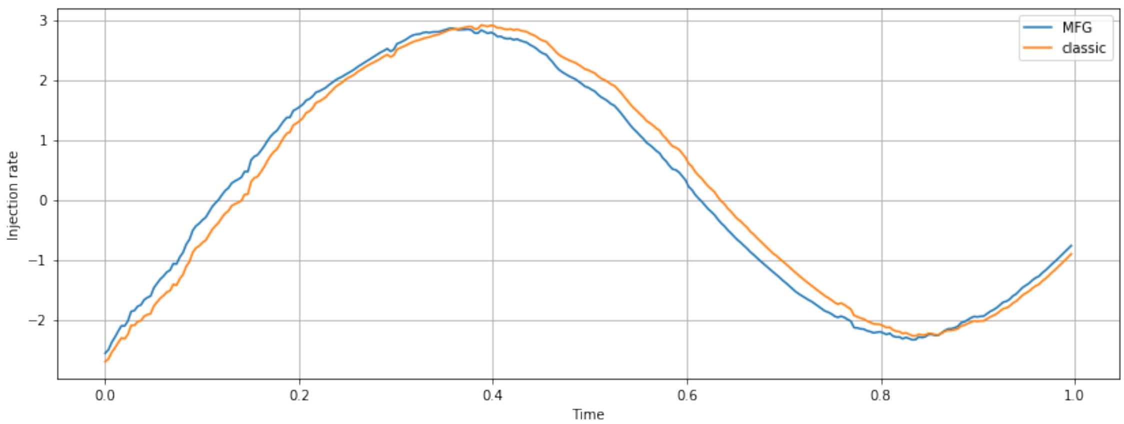

Figure 6.10: μ_t with and without market impact

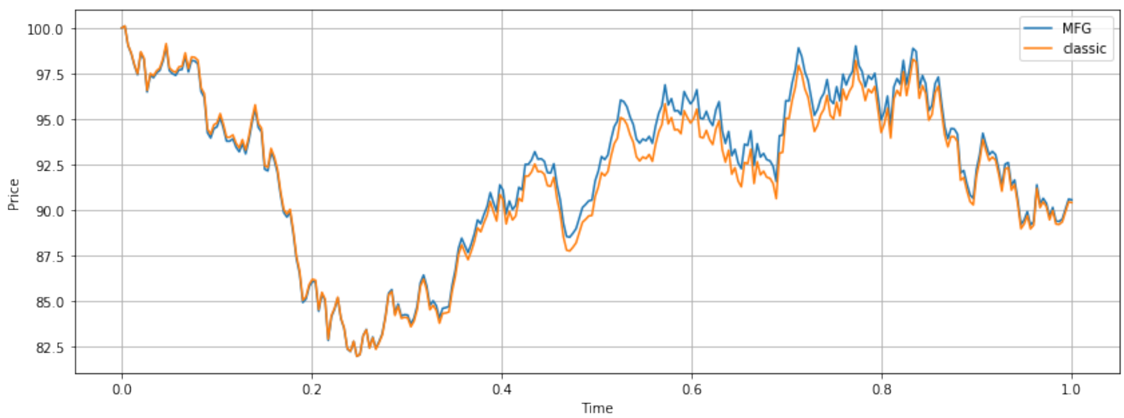

Figure 6.11: P_t with and without market impact

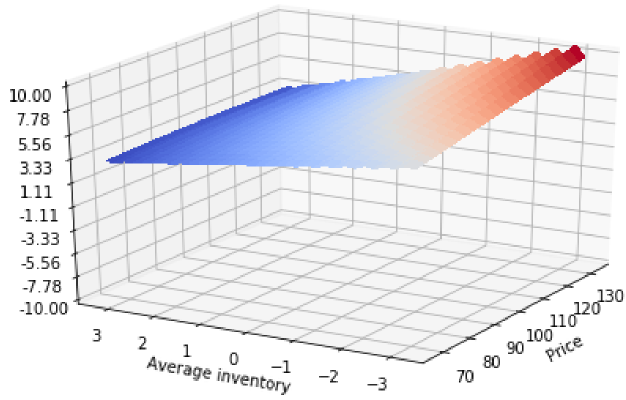

Figure 6.12: $H(t,p,e)$ at $t = \frac{T}{2}$

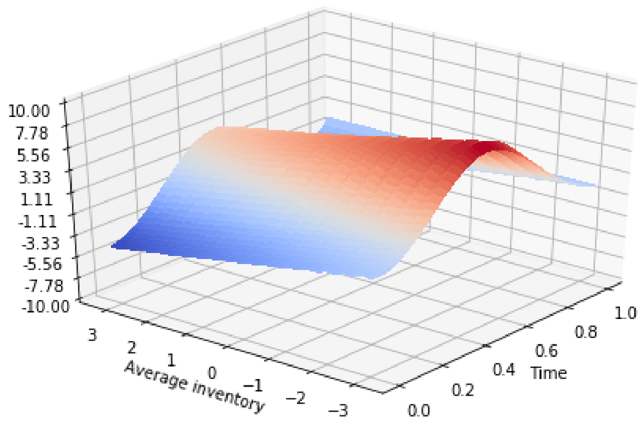

Figure 6.13: $H(t,p,e)$ at $p = p_0$

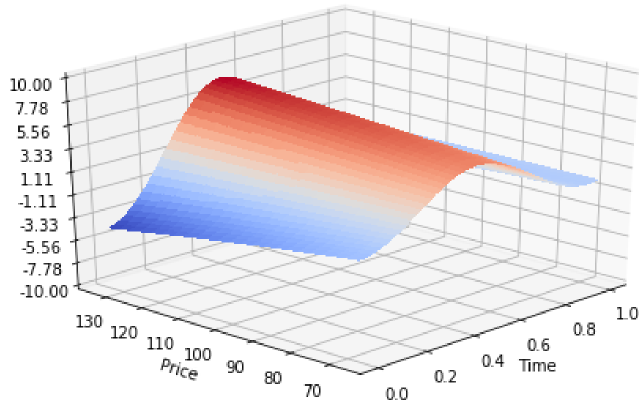

Figure 6.14: $H(t, p, e)$ at $e = 0$

On these figures, we can make the same comments as in the Black-Scholes model.

6.2 Market impact price model without noise

Assuming that $\sigma = 0$ in the price model implies that the FBSDE system becomes a FBODE system.

$$\begin{cases} h_1'(t) = \dfrac{h_1(t)h_2(t)}{C} - A_1 - b_P\left(t, P(t), \dfrac{h_1(t) - E(t)h_2(t)}{C}\right) \\ P'(t) = b_P\left(t, P(t), \frac{h_1(t)-E(t)h_2(t)}{C}\right) \\ E'(t) = \frac{h_1(t)-E(t)h_2(t)}{C} \\ h_1(T) = 0, E(0) = E_0, P(0) = P_0 \end{cases} \quad (71)$$

Those systems can be easily solved by finite differences method.

6.2.1 Example : Black-Scholes

The system becomes, after differentiating the last equation :

$$\begin{cases} h_1(t) = CE'(t) + h_2(t)E(t) \\ \frac{P'(t)}{P(t)} = f_0(t) + \nu E'(t) \\ CE''(t) + P(t)(f_0(t) + \nu E'(t)) - A_2 E + A_1 = 0 \\ CE'(T) + BE(T) = 0, E(0) = E_0, P(0) = P_0 \end{cases} \quad (72)$$

That gives :

$$P(t) = P_0 \exp\left(F(t) + \nu(E(t) - E_0)\right)$$

denoting $F(t) = \int_0^t f_0(s)ds$ and therefore it remains the following second order non-linear ODE :

$$\begin{cases} CE''(t) + P_0 \exp\left(F(t) + \nu(E(t) - E_0)\right)\left(f_0(t) + \nu E'(t)\right) - A_2 E + A_1 = 0 \\ CE'(T) + BE(t) = 0, E(0) = E_0 \end{cases} \quad (73)$$

$$T = 1,\ A_1 = 5,\ A_2 = 16,\ B = 2,\ C = 1,\ P_0 = 100,\ E_0 = 0,\ K = 0.3,\ \phi = \frac{3\pi}{4},\ \nu = 0.01$$

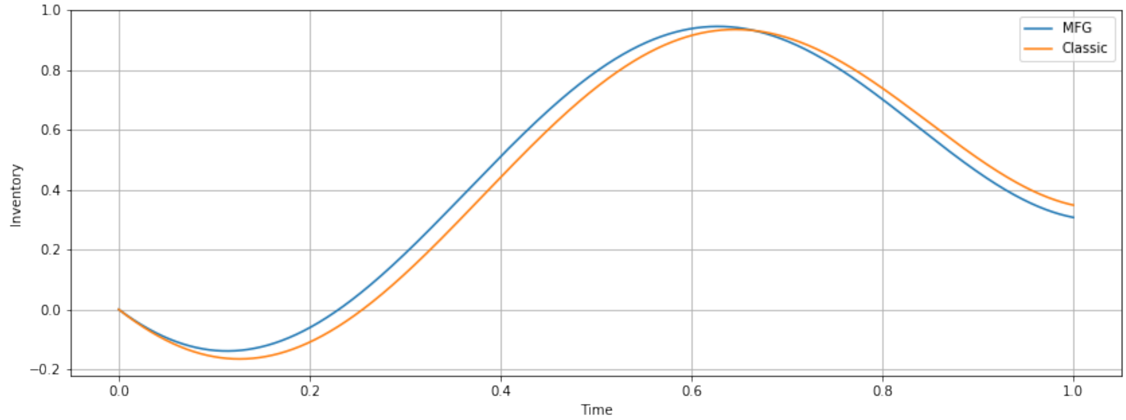

Figure 6.15: $E(t)$ with and without market impact

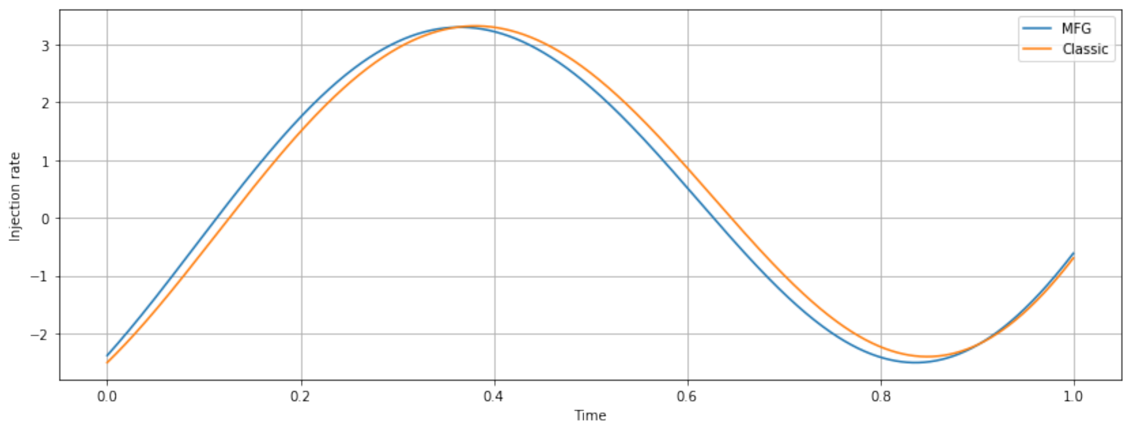

Figure 6.16: μ_t with and without market impact

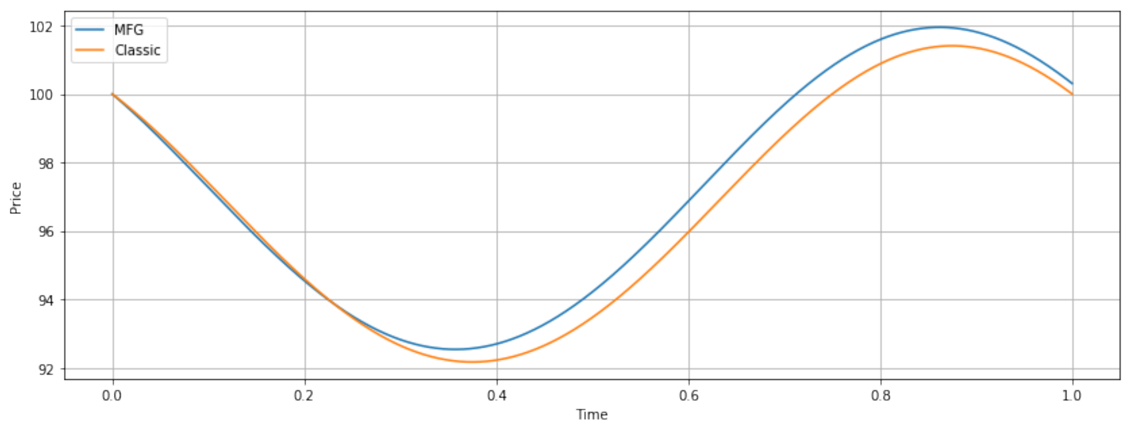

Figure 6.17: $P(t)$ with and without market impact

For the trajectories of E_t and its derivative μ_t, we find similar ones than in the case with noise, comforting us in the belief that our numerical methods gives good approximations of the real solutions.

6.2.2 Example : Clewlow-Strickland

The price in the Clewlow-Strickland price model is

$$P(t) = F_0(t)\exp\left(\nu e^{-at}\int_0^t e^{as}E'(s)ds\right)$$

and $b_P(t,p,\mu) = p\left(\frac{F_0'(t)}{F_0(t)} - a\ln(\frac{p}{F_0(t)}) + \nu\mu\right)$. The system is then :

$$\begin{cases} h_1(t) = CE'(t) + h_2(t)E(t) \\ CE''(t) + F_0(t)\exp\left(\nu e^{-at}\int_0^t e^{as}E'(s)ds\right)\left(\frac{F_0'(t)}{F_0(t)} - a\nu e^{-at}\int_0^t e^{as}E'(s)ds + \nu E'(t)\right) - A_2E + A_1 = 0 \\ CE'(T) + BE(t) = 0, E(0) = E_0 \end{cases} \quad (74)$$

$$T = 1,\ A_1 = 5,\ A_2 = 25,\ B = 5,\ C = 1,\ P_0 = 100,\ E_0 = 0,\ K = 0.3,\ \phi = \frac{\pi}{4},\ \nu = 0.015,\ a = 5$$

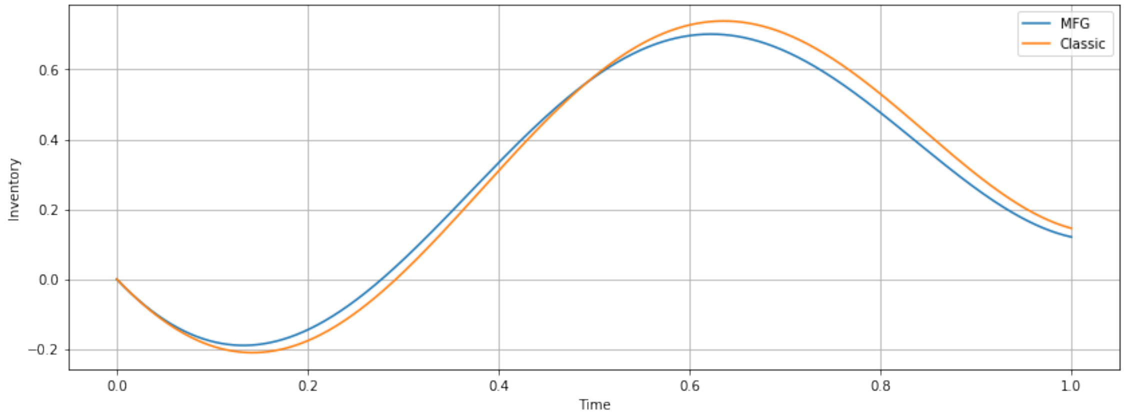

Figure 6.18: $E(t)$ with and without market impact

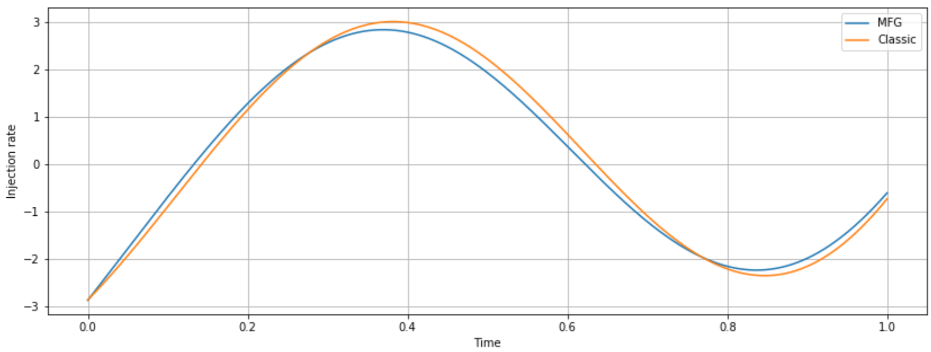

Figure 6.19: $\mu(t)$ with and without market impact

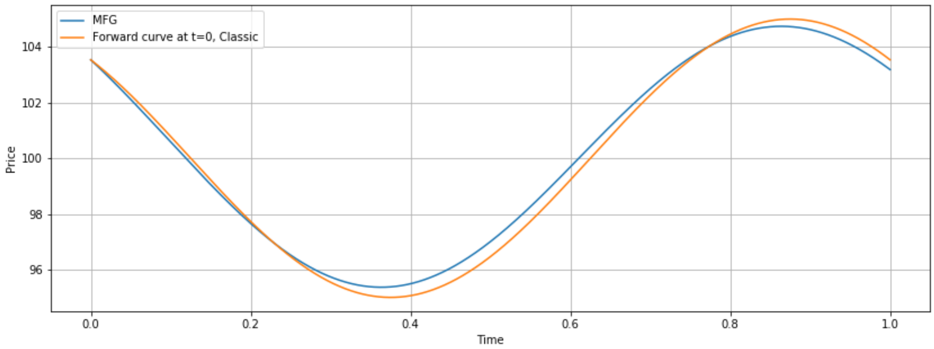

Figure 6.20: $P(t)$ with and without market impact

The same remark can be done than for the Black-Scholes model.

6.3 Learning

The solution of the MFG system describing an equilibrium , one may wonder how this configuration can be reached without the coordination of the agents. We present here a simple model to explain this phenomenon, it is inspired from fictitious play in game theory. For this we assume the game repeats the same $[0, T]$ intervals an infinite number of rounds. Round after round, market agents try to "learn" (i.e. to build an estimate of) the mean of the controls $\mu_t = \mu(t, P_t, E_t)$. It is close to what storage managers do : they try to estimate what the other storage would do in a given situation to adjust their own behaviour. The repeated games eventually converges to the Mean Field Game equilibrium, without any player knowing the others' individual payoff functions.

Cardaliaguet [12] and Hadikhanloo [30] show results of convergence of the procedure toward an equilibrium for potential Mean Field Games. Potential Mean Field Games are MFG where $f(t, \cdot)$ and g derived from potentials. A function f is said to derive from a potential if $\exists F : \mathcal{M}_1(\mathbb{R}^d) \to \mathbb{R}$ s.t. $\delta_m F = f$. One can show using Schwarz theorem that f derive from a potential if and only if

$$\forall x, y, m \ \delta_m f(x, m, y) = \delta_m f(y, m, x)$$

Guo et al. [29] worked on reinforcement learning for ad auction games.

6.3.1 Bachelier case

Cardaliaguet and Lehalle [13] applied the concept of learning to their optimal liquidation game.

The Bachelier case is simpler for learning because the controls and μ do not depend on the price, therefore they are deterministic in time. The players only need to know what will be $\mu(t)$ at each time t.

Each player a has its own prior $\mu^{a,0}$, learning rate $\pi^{a,n}$ and measure of the average control of the n^{th} game $\mu_t^{a,n,measured} = \mu_t^n + \epsilon_t^{a,n}$. At each game the use a estimator of μ : $\mu_t^{a,n,estimate}$ from which they compute their optimal control using standard stochastic optimization. Then for the next game they update their estimator of μ :

$$\mu_t^{a,n+1,estimate} = (1 - \pi^{a,n})\mu_t^{a,n,estimate} + \pi^{a,n}\mu_t^{a,n,measured}, \ \pi^{a,n} \in [0, 1]$$

The idea behind the convergence lies in the Banach fixed point theorem since the equilibrium state is such that the resulting average control $\mu(\cdot)$ induce control to the players that averages to $\mu(\cdot)$, thus $\mu(\cdot)$ is the fixed point of a certain functional.

A good choice for $\pi^{a,n}$ is

$$\pi^{a,n} = \frac{1}{(1 + n)^\gamma}$$

with $\gamma \in (0, 1)$

For the numerical example, we take the seasonality of the form :

$$f_0(t) = Kcos(2\pi t + \phi)$$

We chose the following set of parameters, (with the Bachelier model, the results do not depend on the volatility so one can take whichever value they want for σ) :

$$T = 1,\ A_1 = 5,\ A_2 = 5,\ B = 10,\ C = 1,\ P_0 = 100,\ E_0 = 0,\ K = 5,\ \phi = \frac{3\pi}{4},\ \nu = 7$$

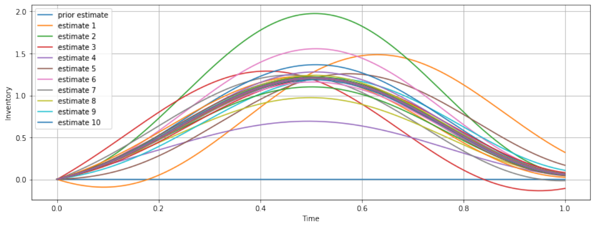

Figure 6.21: $\gamma = 0$, 74 rounds

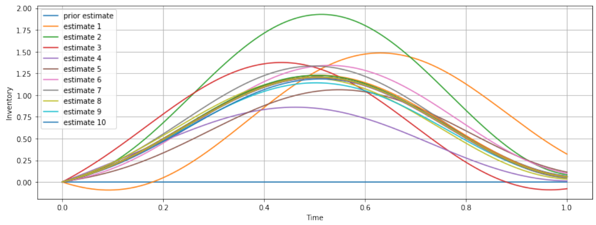

Figure 6.22: $\gamma = 0.1$, 38 rounds

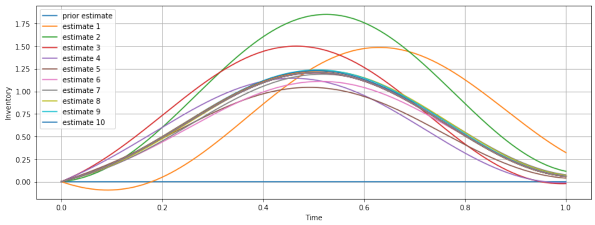

Figure 6.23: $\gamma = 0.3$, 37 rounds

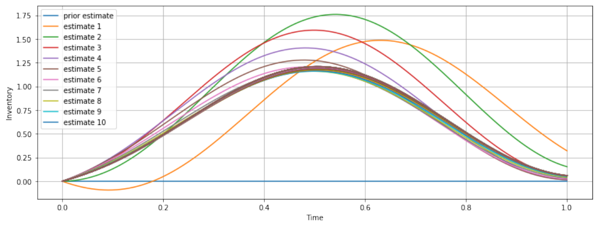

Figure 6.24: $\gamma = 0.6$, 85 rounds

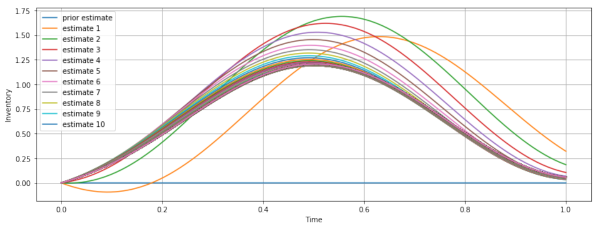

Figure 6.25: $\gamma = 0.9$, 566 rounds

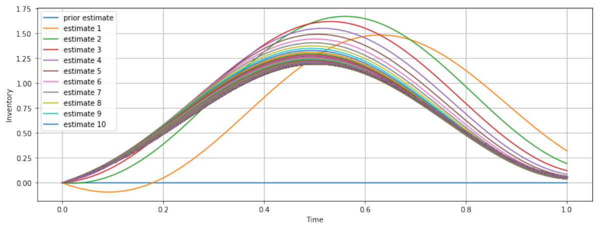

Figure 6.26: $\gamma = 1$, 1334 rounds

We simulated six sets of repeated games until the difference between two games is arbitrarily small. For each set, every player has the same learning algorithm with the same prior $\mu^0(t) = 0$, and we assumed $\epsilon^n(t) = 0$. The difference between the six sets is the power γ. Under each figure there is the number of rounds of games required to satisfy the same criteria of convergence : $d(\mu^{n+1}, \mu^n) < \epsilon$ with d a distance. We can see how γ influences how the system converges. Low enough values result in oscillations around the limit when converging, while high enough values lead to an amortized convergence. The parameter ν also plays a similar role, high enough values of ν with low enough values of γ can even lead to unstable systems. Also values of α greater than 1 can lead to a convergence but not toward the solution of the Mean Field Game equations.

6.3.2 Other cases

The average control μ now depends on t, E_t , P_t or Q_t. The learning becomes trickier because the players only 'observe' one path of Brownian motion at a time, a countable number of times, to learn the values of a function on continuous arguments space. In the Bachelier case, there were only one value to know at for each time t. Learning in Mean Field Game with common noise is described as an open problem in the literature. (The Bachelier case is a common noise problem but the optimization problem result in the controls not depending on the noise.) An idea can be to have a parametric estimator of $\mu(t,p,e)$. For example, using a family of polynomials in p, e and t and find the linear combination that best matches the observations.

Abandoning the real life analogy, we can assume that at each round the players can observe the function $\mu^n : t,p,e \to \mu^n(t,p,e)$ computed as the mean of the strategies of each player of the round n :

$$\mu^n(t,p,e) = \int_a \frac{h^{a,n}(t,p,e) - h_2(t)e}{C} \bar{m}_0(da)$$

Then like in the Bachelier case, each player construct an estimate of $\mu^{n+1}(t,p,e)$ and compute his optimal control (mainly $h^{a,n+1}(t,p,e)$) for the next round.

Numerically, computing the h^n can be done by finite differences for example. In comparison with the previous method, the PDE is linear so finding the zeros in the associated difference equation is finding the zero of a linear system, which is much simplier than using Newton-Raphson method in a non-linear equation. Loosing the computational complexity of the Newton-Raphson method, we however get the complexity from convergence in the Banach theorem.

Numerical tests show that this method of learning works and converges to the equilibrium, which is the same as the one found with finite differences. However, the method has shown to not being always reliable as some model parameters' value can make this method explode while the finite differences stay stable. Also when it converges, the learning method require more iteration of computing a Jacobian and inverting a linear system than the finite differences does, for the same criteria of convergence.

7 The hard constraints problem

7.1 Derivation of the equations

The linear-quadratic case is fine for pedagogic purposes or to show heuristic results on the behaviours of the players. However for business purposes, we need to have more realistic assumptions, especially we cannot have the inventory to be negative or exceed the maximum capacity.

$$\mathcal{A} = \left\{ (\alpha_t)_{0\leq t\leq T} \ : \ -W_{\max}(S_t) \leq \alpha_t \leq I_{\max}(S_t), \ \forall t \ 0 \leq S_0 + \int_0^t \alpha_s ds \leq S^{\max}, \ S_f^{\min} \leq S_0 + \int_0^T \alpha_t dt \leq S_f^{\max} \right\}$$

For the example we choose $S_f^{\max} = S^{\max}, S_f^{\min} = 0$, for the ratchets : $W_{\max}(s) = W\sqrt{\frac{s}{S^{\max}}}$, $I_{\max}(s) = I\sqrt{1-\frac{s}{S^{\max}}}$, and the payoff function of the form :

$$f(\alpha_t, S_t, P_t)dt = \left(-\alpha_t P_t - c_I(\alpha_t)_+ - c_W(\alpha_t)_-\right) dt$$

$$J(t,s,p,\alpha) = \mathbb{E}_{t,s,p}\left[\int_t^T f(\alpha_u, S_u, P_u)du\right]$$

With this choices of ratchets, the set of admissible control can be rewritten :

$$\mathcal{A} = \left\{ (\alpha_t)_{0\leq t\leq T} \ : \ -W_{\max}(S_t) \leq \alpha_t \leq I_{\max}(S_t) \right\}$$

The dynamic programming principle gives :

$$\begin{cases} \partial_t v + \partial_p v b_P(t,p,\mu(t,p,m)) + \frac{1}{2}\sigma^2\partial_{pp}v + \partial_m v[-\partial_s(m_t\alpha)] + H(t,s,p,\partial_s v) = 0 \\ v(T) = 0 \end{cases} \tag{75}$$

With

$$H(t,s,p,y) = \sup_{\alpha\in[-W_{\max}(s), I_{\max}(s)]} \{-\alpha p - c_I(\alpha)_+ - c_W(\alpha)_- + \alpha y\}$$

From the Hamiltonian we deduce the form of the control :

$$\alpha(t,s,p,m) = \begin{cases} I_{\max}(s) & \text{if } \partial_s v(t,s,p,m) - p > c_I \\ 0 & \text{if } \partial_s v(t,s,p,m) - p \in [-c_W, c_I] \\ -W_{\max}(s) & \text{if } \partial_s v(t,s,p,m) - p < -c_W \end{cases}$$

Like in the classic framework, we find that the control is bang-bang (see [6]). The average control μ_t is :

$$\mu_t = \mu(t, P_t, m_t) = \int \alpha(t,s,P_t,m_t) m_t(ds)$$

The dynamic of m_t is :

$$dm_t = -\partial_s(m_t\alpha(t,s,P_t,m_t))dt$$

To sum up, the problem is to find $v(t,s,p,m)$ such that :

$$\begin{cases} \partial_t v + \partial_p v b_P(t,p,\mu) + \frac{1}{2}\sigma^2\partial_{pp}v + \partial_m v[-\partial_s(m\alpha)] + (\partial_s v - p)\alpha - c_I(\alpha)_+ - c_W(\alpha)_- = 0 \\ v(T) = 0 \\ \alpha(t,s,p,m) = I_{\max}(s)\mathbb{1}(\partial_s v(t,s,p,m) - p > c_I) - W_{\max}(s)\mathbb{1}(p - \partial_s v(t,s,p,m) > c_W) \\ \mu(t,p,m) = \int \alpha(t,s,p,m)m(ds) \end{cases} \tag{76}$$

In the general case the problem seems rather impossible, even to think of a numerical method. However if we assume that the initial distribution is a finite sum of Diracs then we already have seen that the distribution remains a sum of Dirac through time. In this case we can represent the distribution variable by the Diracs positions, since the proportion for each Diracs remains the same. Denoting by $(E^i_t)_{1\leq i\leq n}$ the positions of the Diracs at time t, the system becomes.

$$\begin{cases} \partial_t v + \partial_p v b_P(t,p,\mu) + \frac{1}{2}\sigma^2\partial_{pp}v + \sum_j \partial_{e_j} v\alpha(t,e_j,p,(e_i)_{1\leq i\leq n}) + (\partial_s v - p)\alpha - c_I(\alpha)_+ - c_W(\alpha)_- = 0 \\ v(T) = 0 \\ \alpha(t,s,p,(e_i)_{1\leq i\leq n}) = I_{\max}(s)\mathbb{1}(\partial_s v(t,s,p,(e_i)_{1\leq i\leq n}) - p > C_I) - W_{\max}(s)\mathbb{1}(p - \partial_s v(t,s,p,(e_i)_{1\leq i\leq n}) > C_W) \\ \mu(t,p,(e_i)_{1\leq i\leq n}) = \sum_j \alpha(t,e_j,p,(e_i)_{1\leq i\leq n}) m_0(E^j_0) \end{cases} \tag{77}$$

Assuming that initially all players have the same inventory we have :

$$\begin{cases} \partial_t v + \partial_p v b_P(t,p,\alpha) + \frac{1}{2}\sigma^2\partial_{pp}v + (\partial_e v - p)\alpha - c_I(\alpha)_+ - c_W(\alpha)_- = 0 \\ v(T) = 0 \\ \alpha(t,e,p) = I_{\max}(e)\mathbb{1}(\partial_s v(t,e,p) - p > C_I) - W_{\max}(e)\mathbb{1}(p - \partial_s v(t,e,p) > C_W) \end{cases} \tag{78}$$

7.2 Numerical method

We'll focus on this last equation (78) where all players have the same initial inventory. This equation with $\nu = 0$ is the HJB equation for the classic problem. We will use this similarity and adapt the finite differences method from the work of Følstad [25]. He worked on various numerical method to solve the HJB equation for gas storage valuation in the classic case such as finite differences with an upwind technique, finite elements and a semi-Lagrangian time stepping method. We will use the upwind technique that require conditions when $e = 0$ or $e = S^{\max}$ that are : $\alpha(e = 0) \geq 0$ and $\alpha(e = S^{\max}) \leq 0$. Those conditions are already satisfied because of the ratchets.

Denoting $\Delta^+_x f_k = \frac{f(x_{k+1}) - f(x_k)}{d_x}$ and $\Delta^-_x f_k = \frac{f(x_k) - f(x_{k-1})}{d_x}$, the upwind technique is approximate the equation

$$f''(x) + a(x)f'(x) + b(x)f(x) = 0$$

by

$$\Delta_{xx} f_k + (a(x_k))_+ \Delta^+_x f_k - (a(x_k))_- \Delta^-_x f_k + b(x_k) f_k = 0$$

This technique is supposed to prevent the apparition of spurious oscillations. Unfortunately, we could not use it for linear-quadratic Black-Scholes or Clewlow-Strickland model, as we could not impose $\alpha \geq 0$ and $\alpha leq 0$ on the lower and upper bounds respectively.

7.2.1 Ornstein-Uhlenbeck price model

The price is an Ornstein-Uhlenbeck process.

$$dP_t = (-a(P_t - F_0(t)) + \nu\mu_t)\,dt + \sigma dW_t$$

with $F_0(t) = P_0 + K cos(2\pi t + \phi)$ We take the following values for the parameters :

$$T = 1,\ a = 30,\ \nu = 3,\ P_0 = 100,\ K = 20,\ \phi = \frac{\pi}{4},\ \sigma = 10$$

and for the storage :

$$S^{\max} = 1,\ c_I = 0.3,\ c_W = 0.1,\ I_{\max} = 5,\ W_{\max} = 10$$

We get the followings results for $y(t,p,e) = \partial_e v(t,p,e)$.

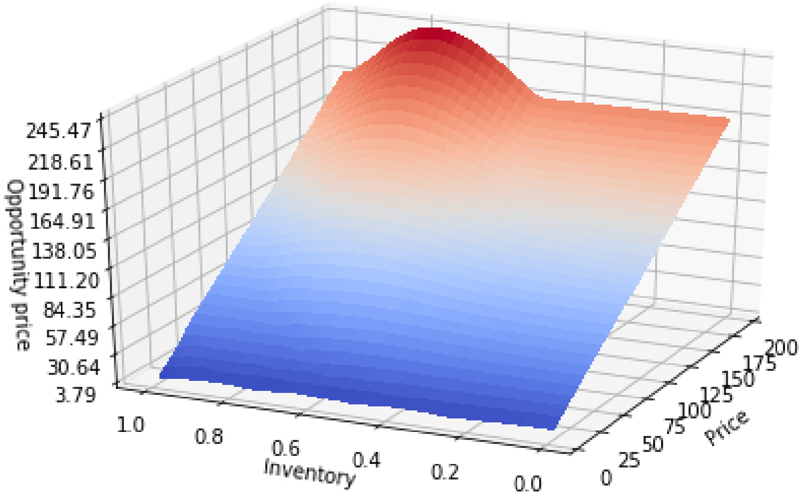

Figure 7.1: $y(t,p,e)$ at $t = \frac{T}{2}$

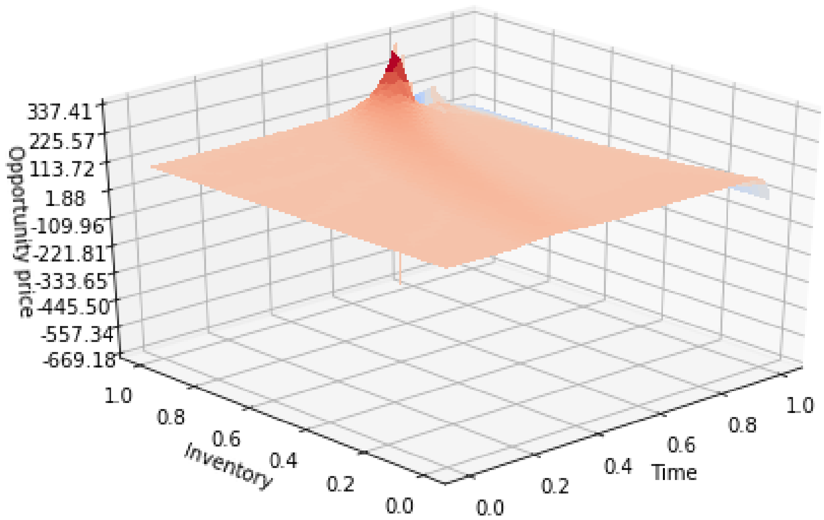

Figure 7.2: $y(t,p,e)$ at $p = P_0$

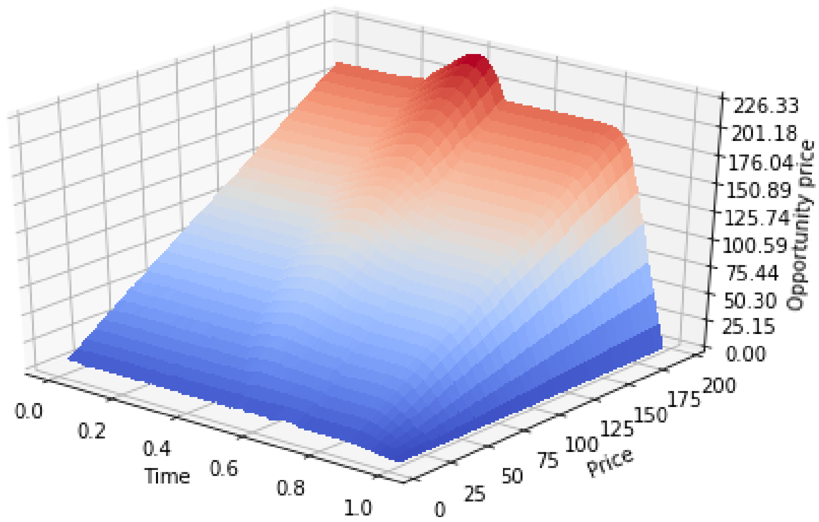

Figure 7.3: $y(t, p, e)$ at $e = \frac{S^{\max}}{2}$

And for $\alpha(t, p, e) = I_{\max}(e)\mathbb{1}(y(t, e, p) - p > C_I) - W_{\max}(e)\mathbb{1}(p - y(t, e, p) > C_W)$

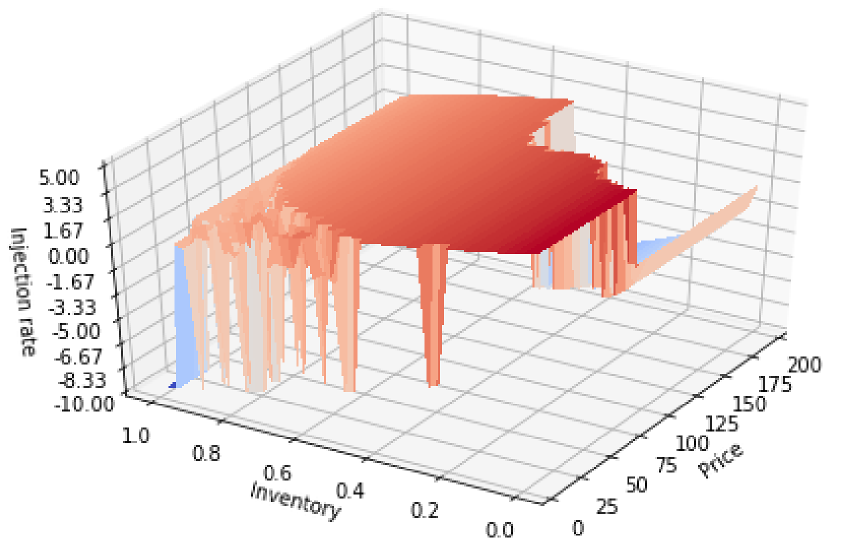

Figure 7.4: $\alpha(t, p, e)$ at $t = \frac{T}{2}$

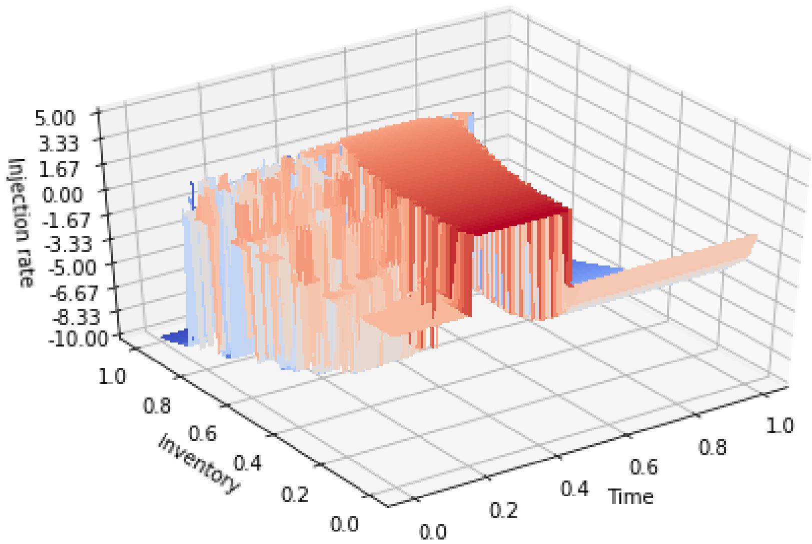

Figure 7.5: $\alpha(t, p, e)$ at $p = P_0$

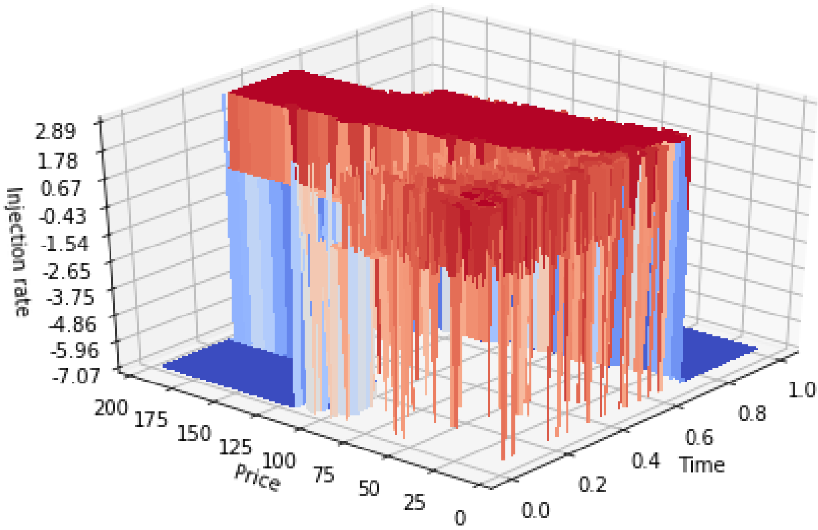

Figure 7.6: $\alpha(t, p, e)$ at $e = \frac{S^{\max}}{2}$

We can see that the numerical method does not return as smooth functions as it did for the linear-quadratic case. For the α function, we expected it to have jumps in its plot but we also expected it to be smooth outsides of jumps which is not entirely true. On y and α, we can see that the most problems occurs in the e directions, similar to the fact that we only had spurious oscillations on the e axis in the linear-quadratic case.

We plot one path of price and corresponding strategy and compare it to the classic case ($\nu = 0$). We use $E_0 = 0$, $P_0 = 100$.

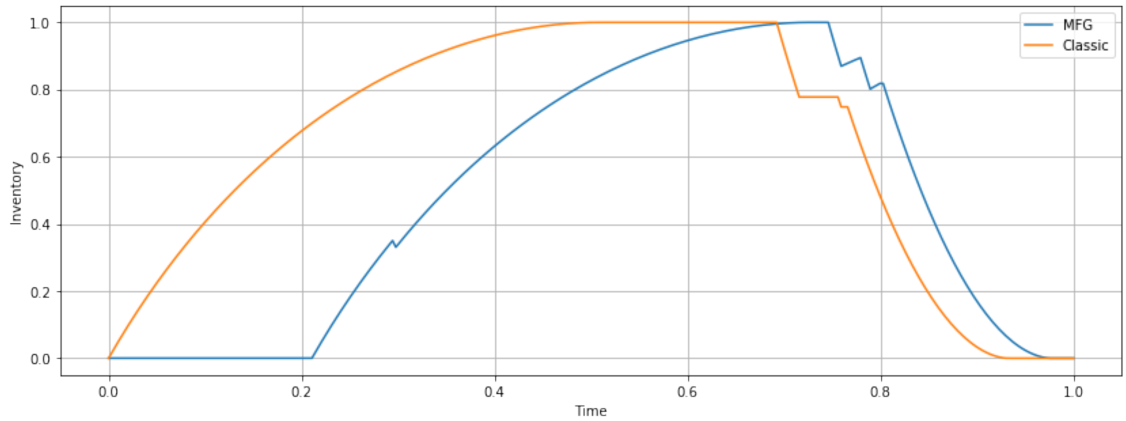

Figure 7.7: E_t

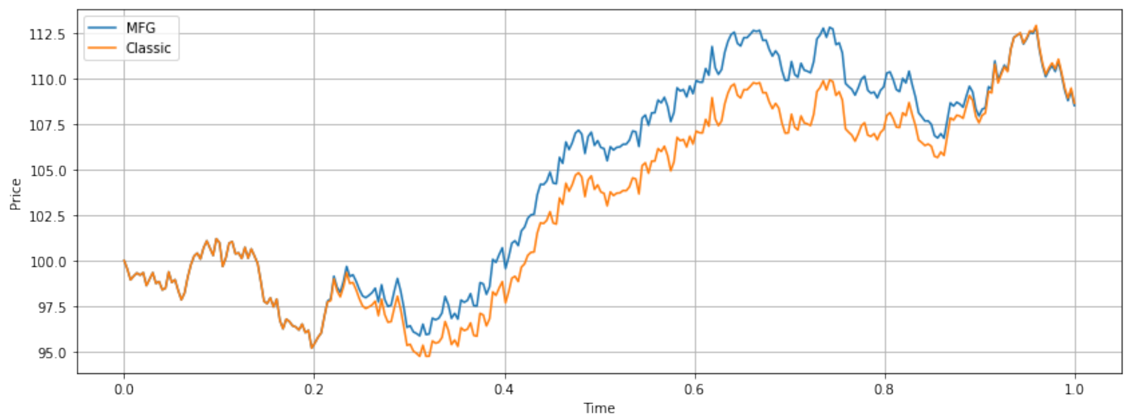

Figure 7.8: P_t

7.2.2 Clewlow-Strickland price model

We recall the stochastic differential equation in this case :

$$\frac{dP_t}{P_t} = \left(\frac{F_0'(t)}{F_0(t)} + a\left(\ln(F_0(t)) - \ln(P_t)\right) + \frac{\sigma^2}{4}(1 - \exp(-2at)) + \nu\bar{\mu}_t \right) dt + \sigma dW_t$$

with $F_0(t) = P_0 + K cos(2\pi t + \phi)$ We take the following values for the parameters :

$$T = 1,\ a = 0.3,\ \nu = 0.03,\ P_0 = 100,\ K = 20,\ \phi = \frac{\pi}{4},\ \sigma = 0.3$$

and for the storage :

$$S^{\max} = 1,\ c_I = 0.3,\ c_W = 0.1,\ I_{\max} = 5,\ W_{\max} = 10$$

We get the followings results for $y(t, p, e) = \partial_e v(t, p, e)$.

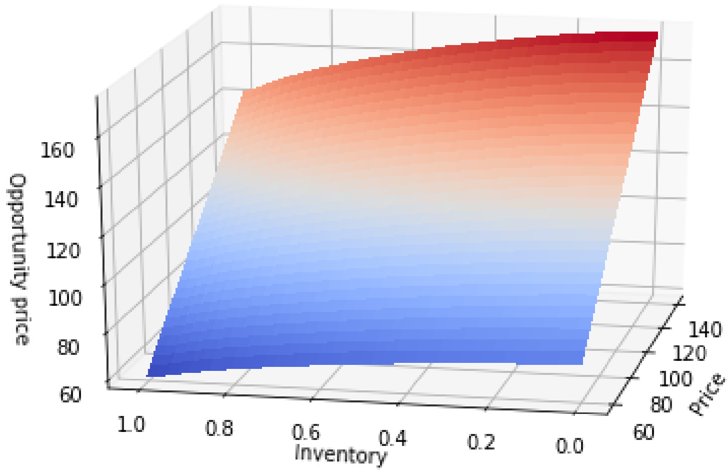

Figure 7.9: $y(t,p,e)$ at $t=\frac{T}{2}$

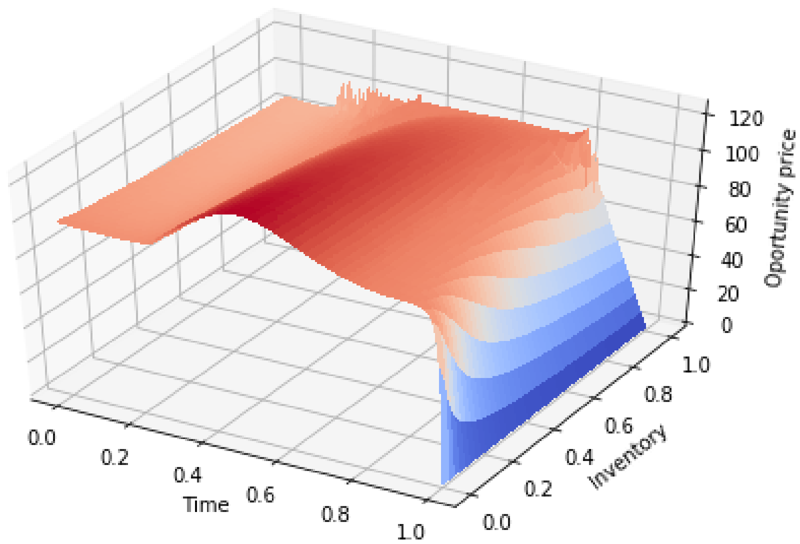

Figure 7.10: $y(t,p,e)$ at $p=P_0$

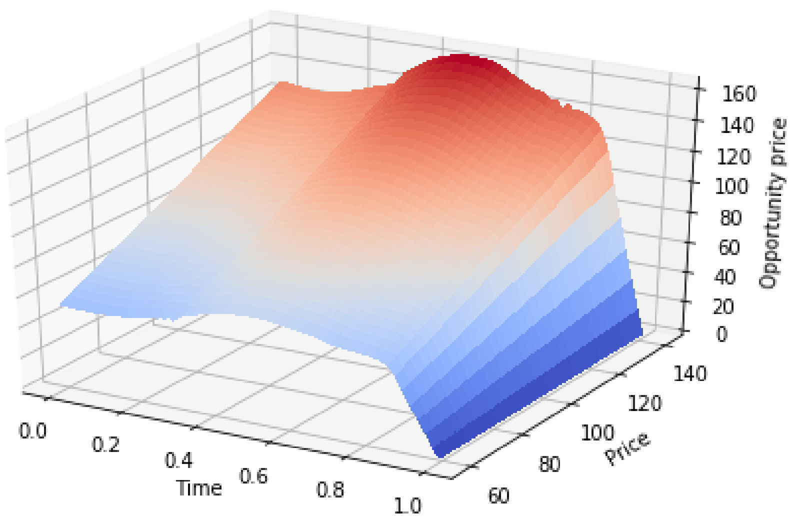

Figure 7.11: $y(t,p,e)$ at $e = \frac{S^{\max}}{2}$

And for $\alpha(t,p,e) = I_{\max}(e)\mathbb{1}(y(t,e,p) - p > C_I) - W_{\max}(e)\mathbb{1}(p - y(t,e,p) > C_W)$

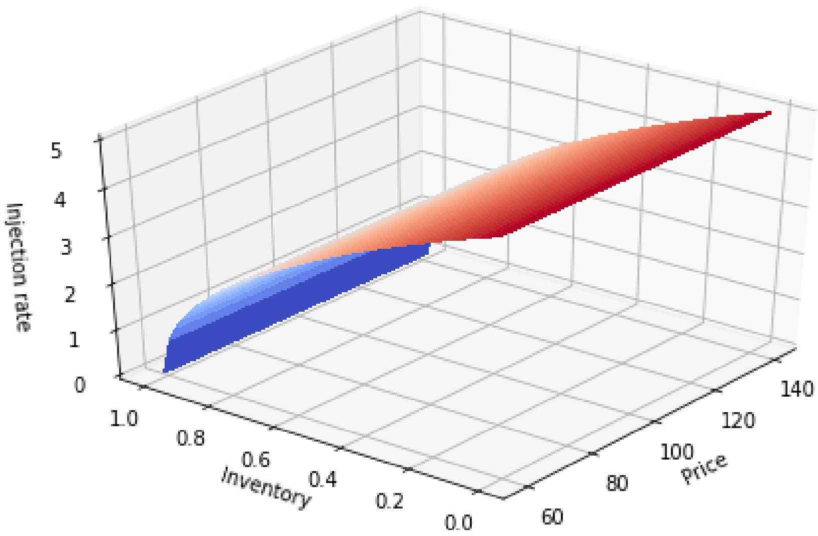

Figure 7.12: $\alpha(t,p,e)$ at $t = \frac{T}{2}$

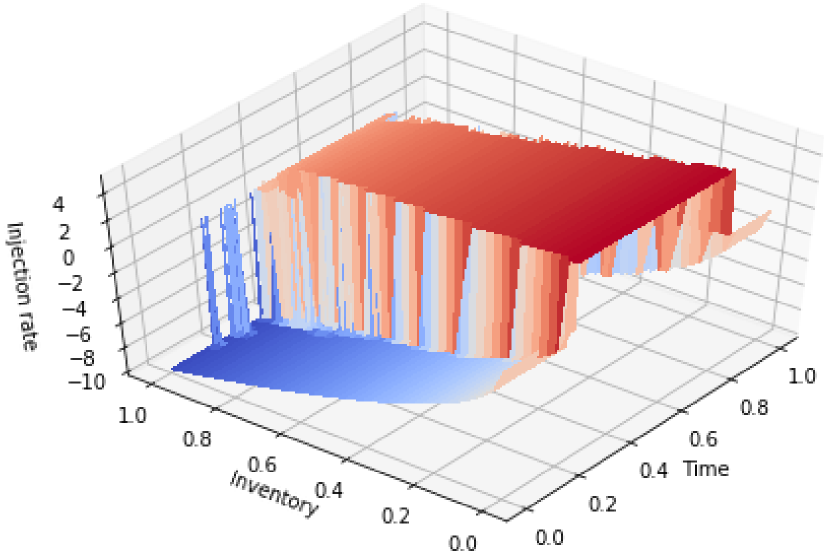

Figure 7.13: $\alpha(t, p, e)$ at $p = P_0$

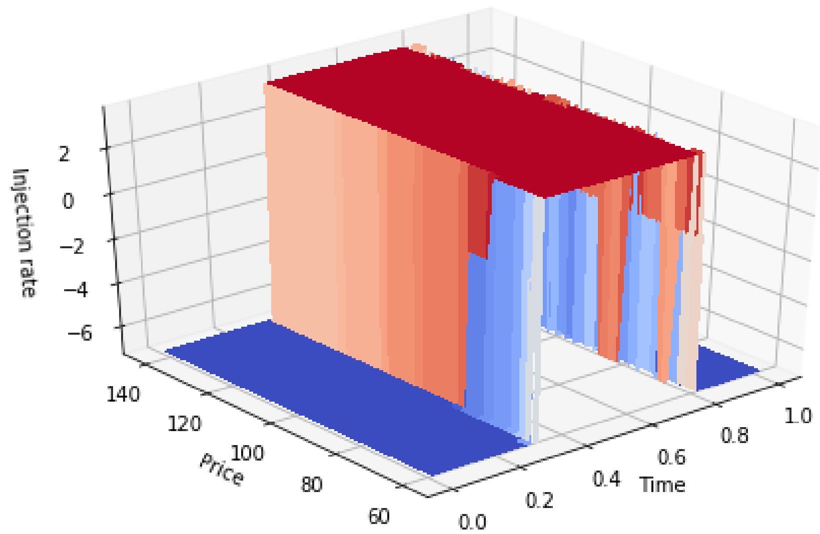

Figure 7.14: $\alpha(t, p, e)$ at $e = \frac{S^{\max}}{2}$

We plot one path of price and corresponding strategy and compare it to the classic case ($\nu = 0$). We use $E_0 = 0$, $P_0 = 100$.

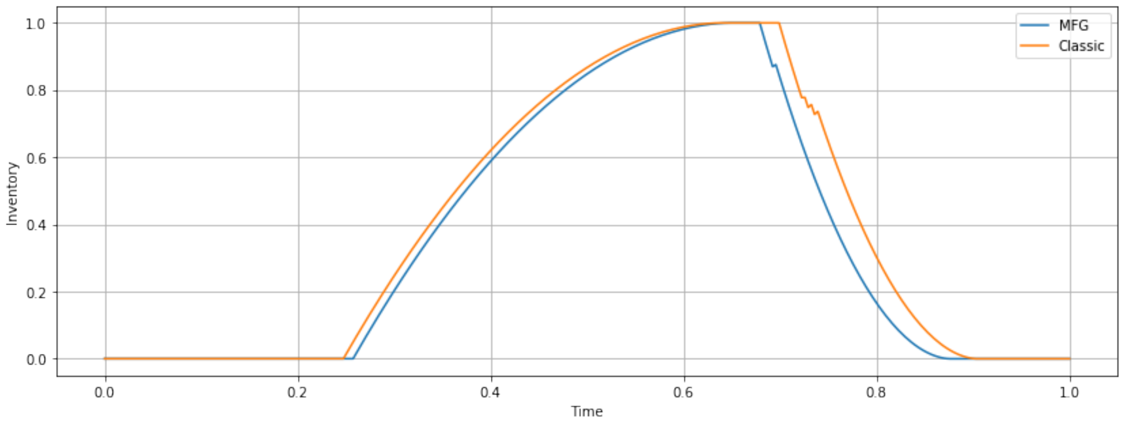

Figure 7.15: E_t

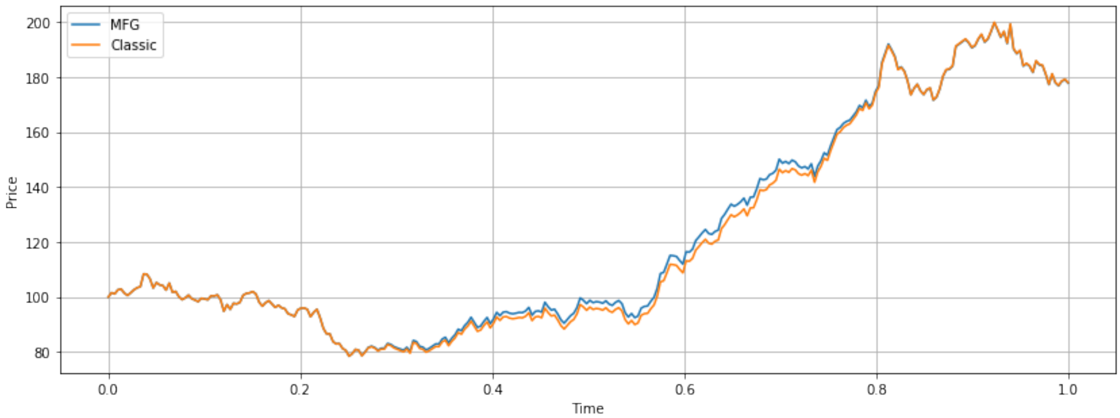

Figure 7.16: P_t

8 Calibration

8.1 Method

To use this theory in practice for gas storage valuation, we need to have ourself a price model. We try to calibrate two different models : a Bachelier price model with market impact and a model where the price is only a function of the demand. Denoting by $D_{n,n+1}$ the demand between t_n and t_{n+1}, $P_{n+m,n+m+1}$ and $\Delta P_{n+m,n+m+1}$ the mean of the prices and variation of price respectively between t_{n+m} and t_{n+m+1}, we try to find the best ν, p_1 and m that fit our data for a given model using Ordinary Least Squares regression.

$$\Delta P_{n+m,n+m+1} = \nu D_{n,n+1} + \epsilon_n$$

$$P_{n+m,n+m+1}, = p_1 D_{n,n+1} + p_0 + \epsilon_n$$

Our data is composed of the daily forward curves for the weekdays of the five last years , as well as the volumes traded for those dates for the western Europe market. The volumes traded are divided between imports, exports - pipelines and LNG -, storages withdrawal, injection, production, and consumption.

The criteria for fitness of the model is how small is the MSE (mean of the square errors).

8.2 Results

We take $t_{n+1} - t_n = 1$ week. We aggregated the consumption, export, and storage injection into a single variable called consumption, and production, import and storage withdrawal into a variable called production. We define the net consumption as the consumption minus the production. We plot the data we have.

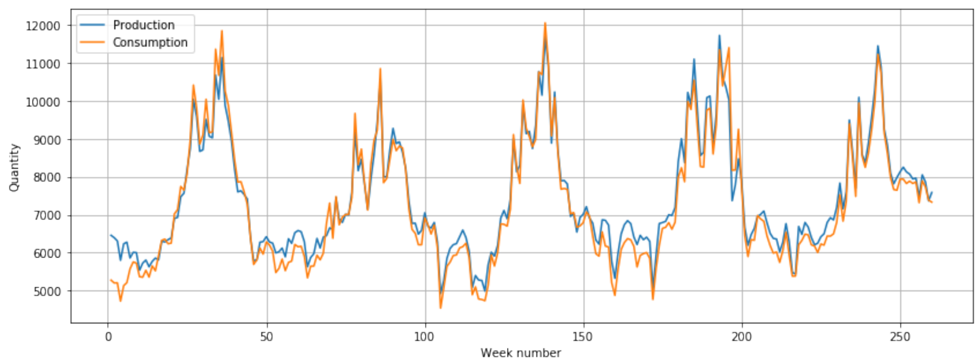

Figure 8.1: Consumption and production

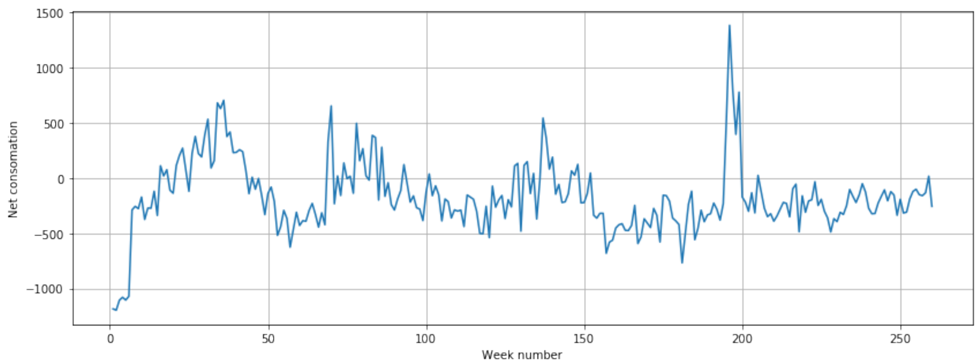

Figure 8.2: Net consumption

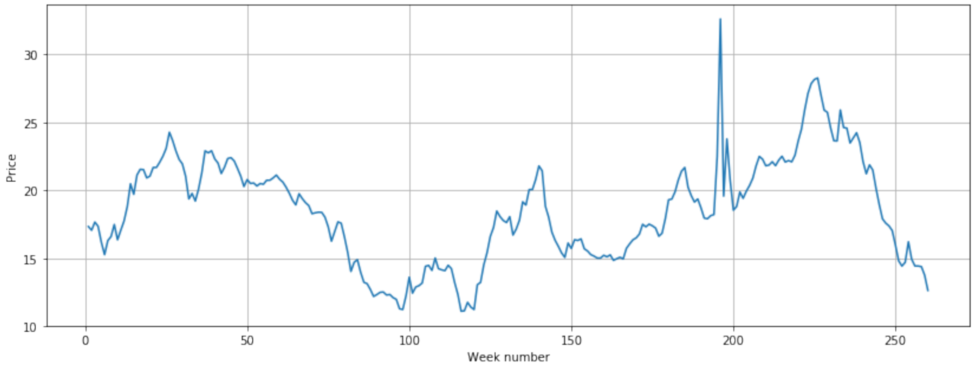

Figure 8.3: Price

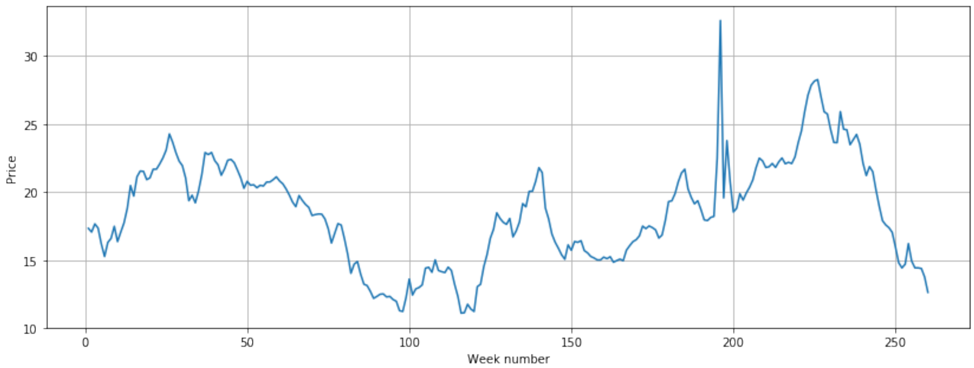

Figure 8.4: Price variation

We can first see that the curves do not match very well except for that one spike. On the following figures, we make the linear regression for $m = 0$, using the consumption and the net consumption as predictors.

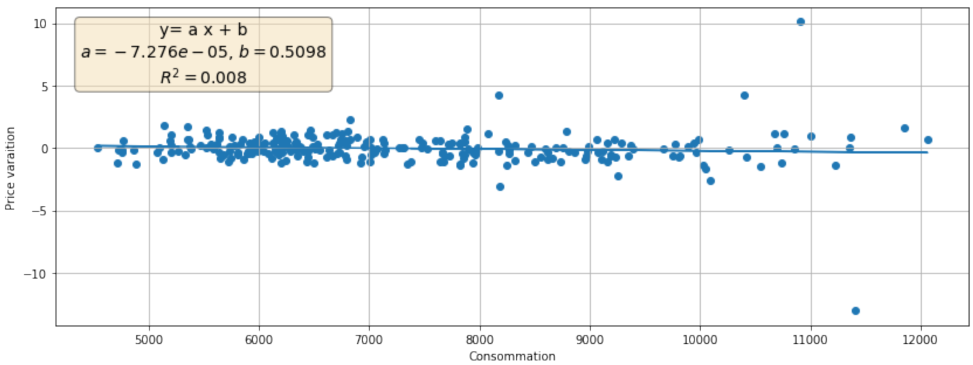

Figure 8.5: Scatter plot of the price variation and the consumption

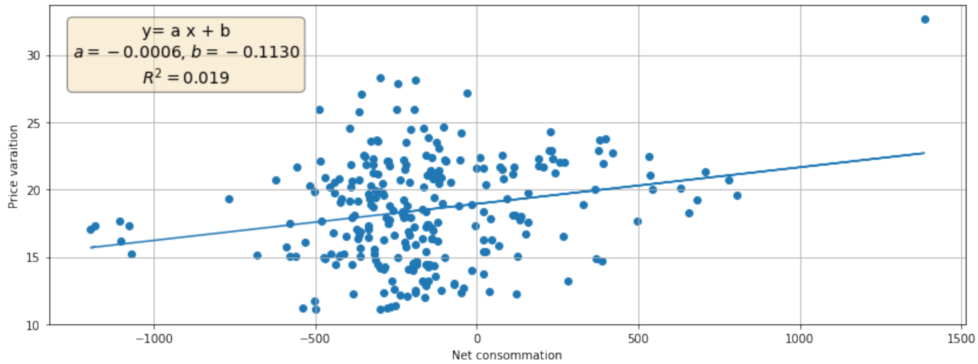

Figure 8.6: Scatter plot of the price variation and the net consumption

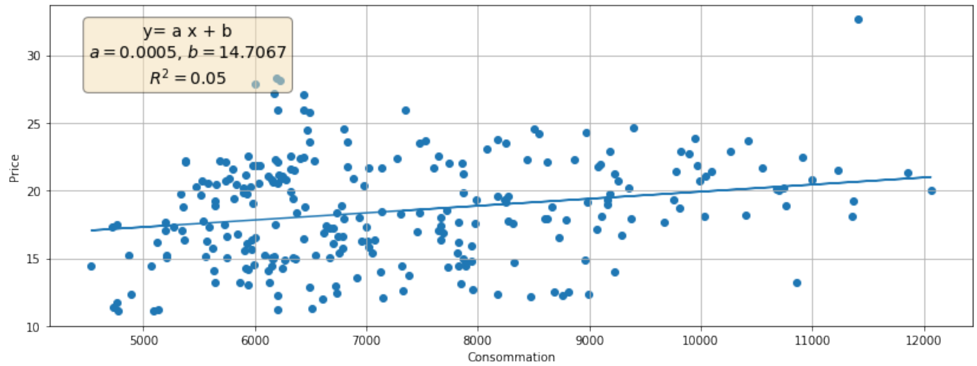

Figure 8.7: Scatter plot of the price and the consumption

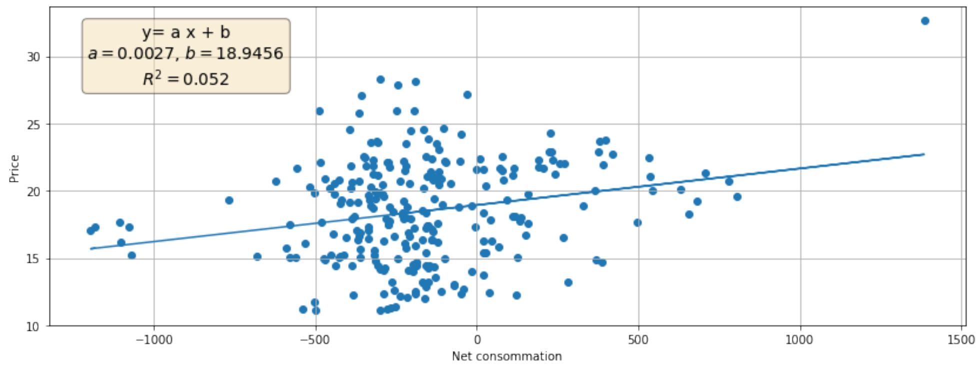

Figure 8.8: Scatter plot of the price and the net consumption

We can see visually and with the R^2 how there seems to be almost no link between the price and the consumption. We now plot the results of the regressions for different values of m.

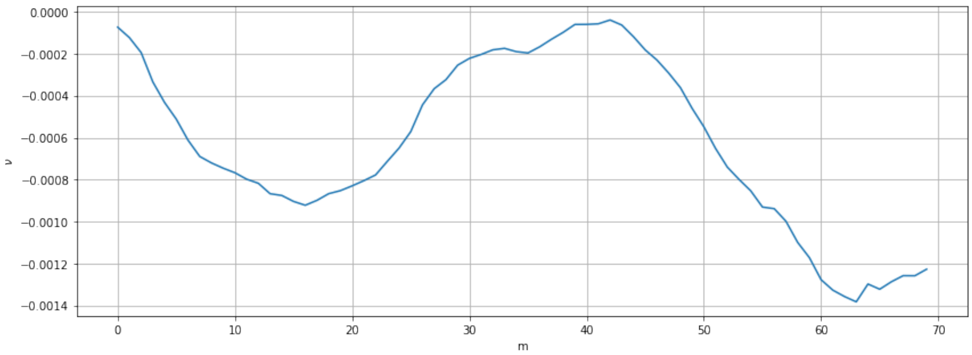

Figure 8.9: $\nu(m)$

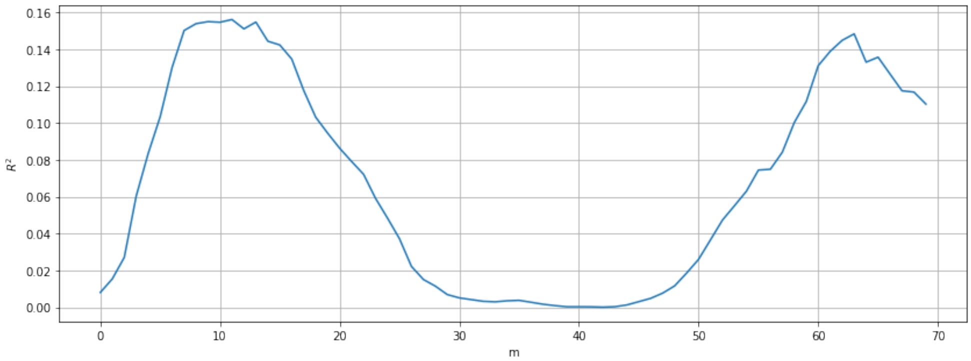

Figure 8.10: $R^2_\nu(m)$

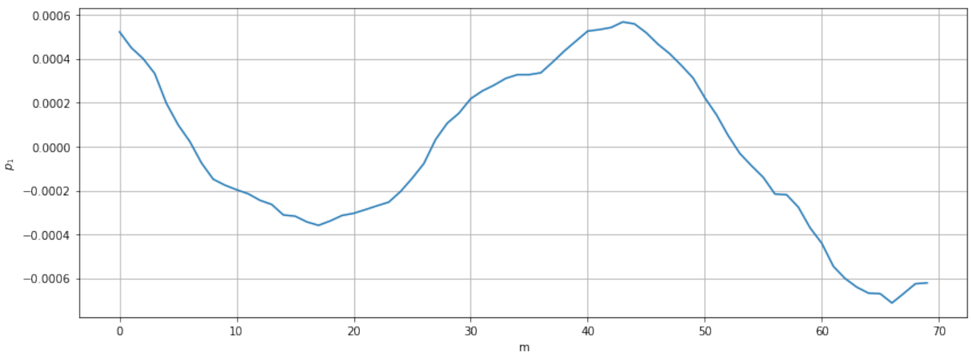

Figure 8.11: $p_1(m)$

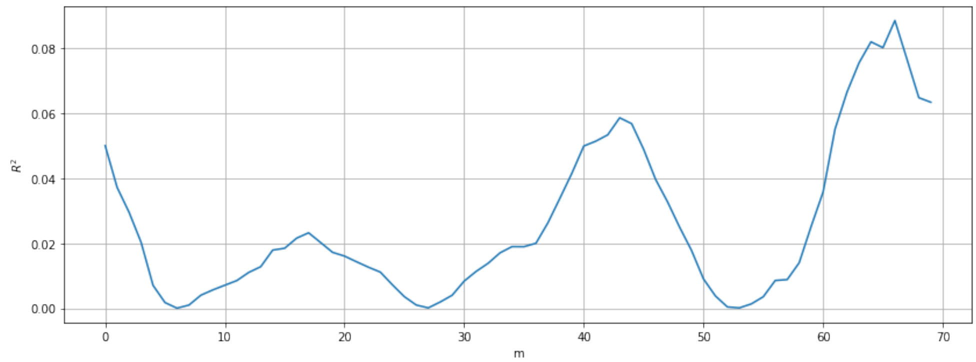

Figure 8.12: $R^2_{p_1}(m)$

We find the yearly seasonality of the consumption in these results, the price seems as much correlated to the demand of the same day that to the demand of one year in the past. This is problematic to find the best time difference between the cause and an effect, provided there is actually one. Indeed there is also the issue of knowing whether the high consumption is the result of the low prices or high prices are the results of the high consumption.

Unfortunately, it seems that these models of market impact are not accurate, at least for the data we have. Usually these kinds of models are suited for smaller time scale like hours or days. Also the model of $P = p_0 + p_1 Q_t$ from Alasseur [2] is understandable for the power market as the production has to be consumed right away and cannot really be stored or shipped away. The demand process therefore does not depend on the price, and the price adjust so that the supply matches the demand, thus having the price being an increasing function of the demand. The gas market is a lot more complicated because of all the imports and exports, the storages, the Liquid Natural Gas shipping, the political decisions, that are not present on the power market

9 Price of anarchy

The concept of the price of anarchy was introduced to quantify the inefficiency of selfish behaviour in finite number of players games. Carmona et al. [19] extended this notion to Mean Field Games. They show some interesting results for Linear Quadratic Extended Mean Field Games where explicit computations are possible. The concept is comparing the average expected payoff of a player in the mean field equilibrium with the average expected payoff of an agent when managed by a central planner in a Mean Field Type control.

9.1 Mean Field type Control for the Linear-Quadratic Bachelier case

Let's imagine that a supervisor want to maximize the common interest, he can choose the control of every agent. We fall into the Mean Field type Control framework. The average control μ_t directly depends on the control of a generic agents α_t, therefore we denote by $\bar{\alpha}_t$ instead. We want to maximize $J(\alpha, \bar{\alpha}_t)$. We denote the price by $P^{\bar{\alpha}}$ to insist on the dependency of the price on the control, which was not the case in MFG. We also denote $P^{\beta,\bar{\alpha}} = d_\beta P^{\bar{\alpha}}$. Assuming a Bachelier price model it gives $P^{\beta,\bar{\alpha}} = P^\beta = \nu E^{\bar{\beta}}$, with $E_t^{\bar{\beta}} = \mathbb{E}\left[\int_0^t \beta_s ds | W_s, s \leq t\right] = \int_0^t \bar{\beta}_S ds$

$$J(\alpha, \bar{\alpha}) = \mathbb{E}\left[\int_0^T -\alpha_t P_t^{\bar{\alpha}} - \frac{C}{2}\alpha_t^2 + A_1 S_t - \frac{A_2}{2} S_t^2 dt + P_T^{\bar{\alpha}} S_T - \frac{B}{2} S_T^2\right]$$

We differentiate J with respect to α :

$$d_\beta J(\alpha, \bar{\alpha}) = \mathbb{E}\left[\int_0^T -\beta_s P_t^{\bar{\alpha}} - \alpha_t \nu E_t^{\bar{\beta}} - C\alpha_t \beta_s + A_1 S_t^\beta - A_2 S_t S_t^\beta dt + P_T^{\bar{\alpha}} S_T^\beta + \nu E_T^{\bar{\beta}} S_T - B S_T S_T^\beta\right]$$

Using the tower property we get :

$$d_\beta J(\alpha, \bar{\alpha}) = \mathbb{E}\left[\int_0^T -\bar{\beta}_s P_t^{\bar{\alpha}} - \bar{\alpha}_t \nu E_t^{\bar{\beta}} - C\bar{\alpha}_t \bar{\beta}_t + A_1 E_t^{\bar{\beta}} - A_2 E_t E_t^{\bar{\beta}} dt + P_T^{\bar{\alpha}} E_T^{\bar{\beta}} + \nu E_T^{\bar{\beta}} E_T - B E_T E_T^{\bar{\beta}}\right]$$

Taking Y_t solution of the following BSDE :

$$\begin{cases} dY_t = -(A_1 - A_2 S_t - \nu\alpha_t)dt + Z_t \\ Y_T = P_T^{\bar{\alpha}} - (B - \nu)S_T \end{cases} \tag{79}$$

And using Ito's lemma, we get :

$$d_\beta J(\alpha, \bar{\alpha}) = \mathbb{E}\left[\int_0^T -\bar{\beta}_s P_t^{\bar{\alpha}} - C\bar{\alpha}_t \bar{\beta}_t + \bar{Y}_t \bar{\beta}_t dt\right]$$

The control α being the argument supremum, $\forall\beta$, $d_\beta J(\alpha, \bar{\alpha}) = 0$. Therefore :

$$\alpha_t = \frac{Y_t - P_t}{C}$$

It remains to find Y_t. We search it on the form $Y_t = P_t + h_1(t) - h_2(t)S_t$

$$\begin{cases} d\bar{Y}_t = \left(f_0(t) + \nu\frac{h_1(t) - h_2(t)E_t}{C} + h_1'(t) - h_2' E_t - h_2(t)\frac{h_1(t) - h_2(t)E_t}{C}\right) dt + \sigma dW_t \\ \quad = -(A_1 - A_2 E_t - \nu\frac{h_1(t) - h_2(t)E_t}{C})dt + Z_t dW_t \\ \bar{Y}_T = P_T - (B - \nu)E_T \end{cases} \tag{80}$$

Therefore :

$$\begin{cases} h_2' - \frac{h_2^2}{C} + A_2 = 0 \\ h_1' - \frac{h_2 h_1}{C} + A_1 + f_0 = 0 \\ h_2(T) = B - \nu, h_1(T) = 0 \end{cases} \tag{81}$$

Setting $\rho = \sqrt{\frac{A_2}{C}}$ and $c_2 = \frac{B-\nu-\rho C}{B-\nu+\rho C} \exp(-2\rho T)$ we have :

$$h_2(t) = \rho C \frac{1 + c_2 \exp(2\rho t)}{1 - c_2 \exp(2\rho t)} \tag{82}$$

and

$$h_1(t) = \int_t^T \exp\left(- \int_t^s \frac{h_2(u)}{C} du \right) (A_1 + f_0(s)) ds$$

To sum up :

$$\begin{cases} \alpha(t,s) = \frac{h_1(t) - h_2(t)s}{C} \\ \mu(t) = \frac{h_1(t) - h_2(t)E(t)}{C} \\ S(t) = \exp\left(- \int_0^t \frac{h_2(u)}{C} du \right) \left(S_0 + \int_0^t \exp\left(+ \int_0^u \frac{h_2(w)}{C} dw \right) \frac{h_1(u)}{C} du \right) \\ E(t) = \exp\left(- \int_0^t \frac{h_2(u)}{C} du \right) \left(E_0 + \int_0^t \exp\left(\int_0^u \frac{h_2(w)}{C} dw \right) \frac{h_1(u)}{C} du \right) \\ P_t = P_0 + \int_0^t f_0(u) du + \nu E(t) + \sigma W_t \end{cases} \tag{83}$$

The expected gain for an agent at time t is

$$V_t = \mathbb{E}_{t,s,p} \left[\int_t^T \left[-\alpha(u, S_u) P_u - \frac{C}{2} \alpha^2(t, S_u) + A_1 S_u - \frac{A_2}{2} S_u^2 \right] du + P_T S_T - \frac{B}{2} S_T^2 \right]$$

Thus

$$V_t = v(t, S_t, P_t) = h_0(t) + \delta(t) + (h_1(t) + P_t + \Delta(t)) S_t - \frac{h_2(t)}{2} S_t^2$$

With $h_0(t) = \int_t^T \frac{h_1^2(s)}{2C} ds$, $\Delta(t) = \int_t^T \exp\left(\int_t^s \frac{h_2(u)}{C} du \right) \nu E'(s) ds$ and $\delta(t) = \int_t^T \frac{h_1(s)\Delta(s)}{C} ds$. The demonstration is that both expressions of V_t are equal in T and have the same differential at any time.

We denote by $F(t) = \mathbb{E}(S(t)^2)$, we have

$$F(t) = \exp\left(- \int_0^t \frac{h_2(u)}{C} du \right) \left(F_0 + 2E_0 \int_0^t \exp\left(\int_0^u \frac{h_2(w)}{C} dw \right) \frac{h_1(u)}{C} du + \left(\int_0^t \exp\left(\int_0^u \frac{h_2(w)}{C} dw \right) \frac{h_1(u)}{C} du \right)^2 \right)$$

We average expected gain is

$$\mathbb{E}[V_0] = h_0(0) + \delta(0) + (h_1(0) + P_0 + \Delta(0)) E_0 - \frac{h_2(t)}{2} F_0$$

Remark 12

This situation also correspond a cartel of agents agreeing to follow a strategy that in the end earn them more than if there were no cartel and the state of the market were the game equilibrium. However the cartel is an unstable state in the sense that each agents has interest in betraying, the more agents are betraying the more the gains of all the players decreases.

9.2 Comparison of expected gains

We recall the expected payoff of a player in a MFG from paragraph 5.1.2

$$v^{MFG}(t, S_t, P_t) = h_0^{MFG}(t) + (P_t^{MFG} + h_1^{MFG}(t))S_t - h_2^{MFG}(t)\frac{S_t^2}{2}$$

With

$$P_t^{MFG} = P_0 + \int_0^t f_0(s)ds + \nu E^{MFG}(t) + \sigma W_t$$

$$h_2^{MFG}(t) = \rho C \frac{1 + c_2 \exp(2\rho t)}{1 - c_2 \exp(2\rho t)}$$

$\rho = \sqrt{\frac{A_2}{C}}$, $c_2 = \frac{B-\rho C}{B+\rho C} \exp(-2\rho T)$
E^{MFG} and h_1^{MFG} solutions of the system :

$$\begin{cases} C(E^{MFG})'' + \nu(E^{MFG})' - A_2 E^{MFG} = -A_1 - f_0 \\ h_1 = C(E^{MFG})' + E^{MFG} h_2^{MFG} \\ E^{MFG}(0) = E_0, C(E^{MFG})'(T) + BE^{MFG}(T) = 0 \end{cases} \tag{84}$$

and

$$h_0^{MFG}(t) = \int_t^T \frac{(h_1^{MFG}(s))^2}{2C} ds$$

The average expected payoff of the players is

$$\mathbb{E}[V_0^{MFG}] = h_0^{MFG}(0) + (P_0 + h_1^{MFG}(0))E_0 - h_2^{MFG}(0)\frac{F_0}{2}$$

Comparing the two average payoff for MFG and MFC is quite tricky analytically. By definition, the MFC strategy is optimal for maximizing the average payoff so we know that the average payoff in MFG is lower.

For the numerical example, we take the seasonality of the form :

$$f_0(t) = Kcos(2\pi t + \phi)$$

We chose the following set of parameters, (with the Bachelier model, the results do not depend on the volatility so one can take whichever value they want for σ) :

$$T = 1,\ A_1 = 5,\ A_2 = 5,\ B = 10,\ C = 1$$

$$P_0 = 100,\ E_0 = 0,\ F_0 = 0,\ K = 5,\ \phi = \frac{3\pi}{4},\ \nu = 6$$

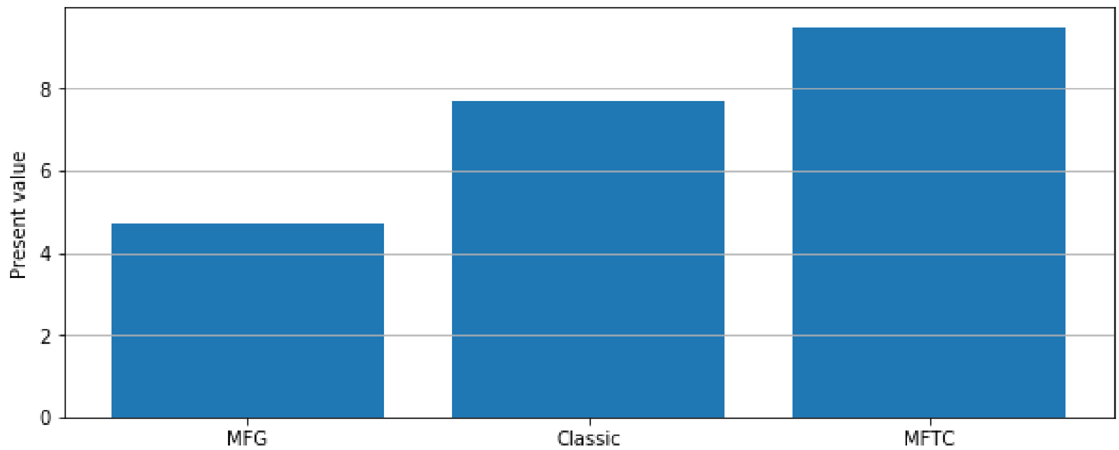

Figure 9.1: Expected payoffs in MFG and MFTC

This first figure show the expected payoff (or present value) of a player starting with a null inventory, in the MFG case and in the cartel case compared to the case where there is no market impact and the player use the optimal strategy. We can see how the Mean Field Game equilibrium significantly performs worse than the cartel situation. This shows also that the presence of a market impact allows the players to gain more than in the classical case granting that they coordinate.

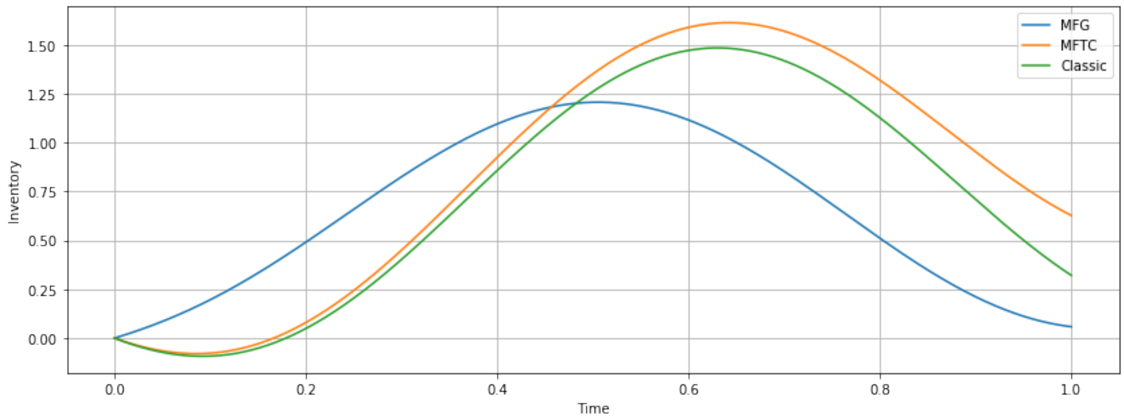

Figure 9.2: Strategy MFTC vs MFG vs classic

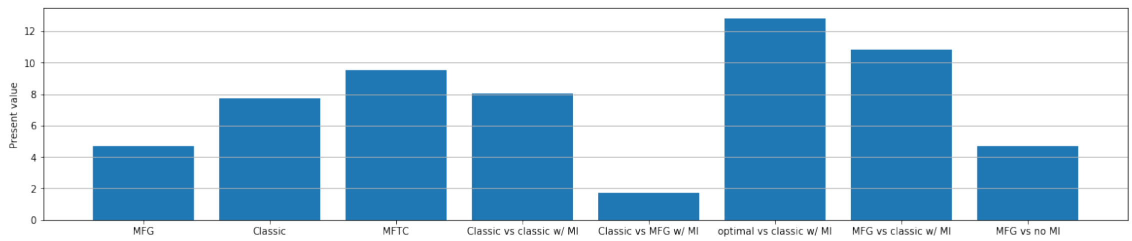

Figure 9.3: Expected payoff by strategy and environment

This last figure shows the expected payoff of players starting with a null inventory with different strategies in different environments. If there is actually no market impact, a player using the MFG strategy (last column) will always perform worse than a player with the optimal strategy that we call classic strategy (second column). In the case that there is effectively a market impact, the expected payoff of a strategy depends on the strategy of the others. We've highlighted some particular cases. If everybody uses the classic strategy (column 4), they all earn more than with the same strategy in classic situation, but not as much as if they all coordinate. If the majority of the players uses the classical strategy, a player using the MFG strategy will perform better (column 7) but not as much as the optimal strategy for this situation (column 6). If the game is at its equilibrium state, a player that keeps using the classic strategy (column 5) will significantly perform worse than the others. The optimal strategy in this situation being the MFG strategy itself.

9.3 Expected gains and learning

In what follows we simulate an evolving environment like in the learning setting with the learning algorithm simply being $\mu^{n+1,estimate} = \mu^{n,measured}$, and the mean-field does not trade on round 0. We track the present value of different strategies in this setting, one of the player always having the optimal strategy for the round, one that use the MFG strategy a each round, and one that have the same strategy than the mean field.

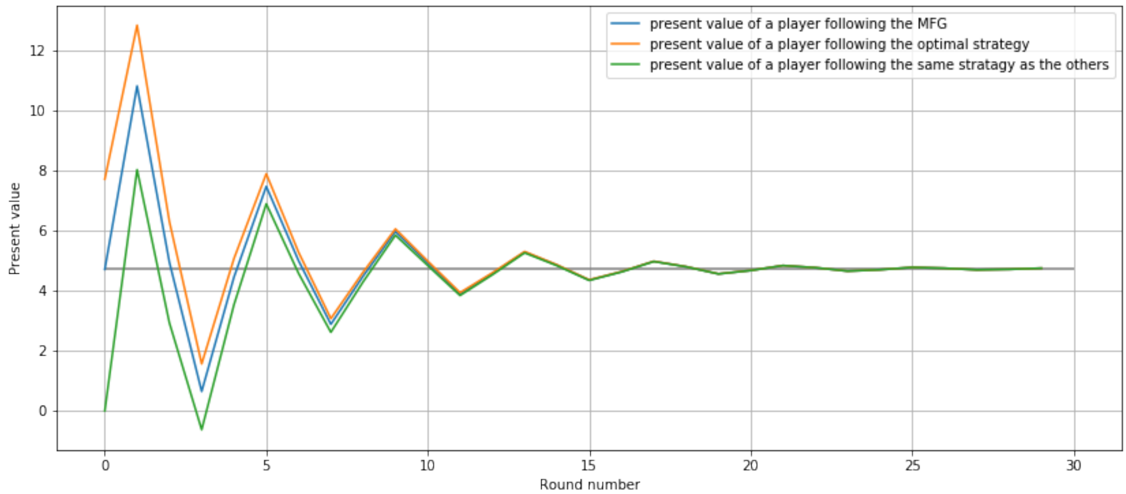

Figure 9.4: Expected payoff by strategy and environment

By definition the optimal strategy performs better than the others at every round. It is interesting to see how the payoff of the MFG strategy, which is the only of the three strategies to remain the same at every round, performs always between the two others. Sadly, we could not find ant heuristic explanation for this feature. Naturally when the the number of rounds goes to infinity, the optimal strategy and the mean field strategy converges to the MFG strategy, and the present values of the three strategies converge to the present value of the MFG strategy in a MFG equilibrium. We know that the mean field would converge to the MFG equilibrium from the results on learning, and we know that the mean field strategy for a round is actually the optimal strategy for the previous round. And of course the optimal strategy in the MFG equilibrium is the MFG strategy itself. This is why the three sequences converge toward this same point.

10 Extended models

10.1 Multiple markets

10.1.1 The model

We assume that the agents have access to multiple gas markets with different prices. They therefore have an array of injection rates, corresponding to the rates of each markets. The dynamic of the inventory is therefore :

$$dS_t = \Sigma_{k=1}^n \alpha_{t,k} dt = \alpha_t^\top \mathbb{1}_n dt$$

The reward function is :

$$\begin{aligned} J(\alpha,\mu) &= \mathbb{E}\left[\int_0^T \left[-\Sigma_{k=1}^n \alpha_{t,k} P_{t,k} - \frac{\Sigma_{k=1}^n C_k \alpha_{t,k}^2 + C_0(\Sigma_{k=1}^n \alpha_{t,k})^2}{2} + A_1 S_t - \frac{A_2}{2} S_t^2\right] dt + \frac{\Sigma_{k=1}^n P_{T,k}/C_k}{\Sigma_{k=1}^n 1/C_k} S_T - \frac{B}{2} S_T^2\right] \\ &= \mathbb{E}\left[\int_0^T \left[-\alpha_t^\top P_t - \frac{\alpha_t^\top (diag(C) + C_0 \mathbb{1}_{n,n})\alpha_t}{2} + A_1 S_t - \frac{A_2}{2} S_t^2\right] dt + \frac{\Sigma_{k=1}^n P_{T,k}/C_k}{\Sigma_{k=1}^n 1/C_k} S_T - \frac{B}{2} S_T^2\right] \end{aligned} \tag{85}$$

10.1.2 Derivation of the equations

Using Pontrayagin's maximum principle we get that

$$d_\beta J(\alpha,\mu) = \mathbb{E}\left[\int_0^T \left[-\beta_t^\top P_t - \beta_t^\top (diag(C) + C_0 \mathbb{1}_{n,n})\alpha_t + A_1 S_t^\beta - A_2 S_t S_t^\beta\right] dt + \frac{\Sigma_{k=1}^n P_{T,k}/C_k}{\Sigma_{k=1}^n 1/C_k} S_T^\beta - B S_T S_T^\beta\right]$$

where $S_t^\beta = \int_0^t \beta_s^\top \mathbb{1}_n ds$

Let Y_t be the solution of the following BSDE :

$$\begin{cases} dY_t = -(A_1 - A_2 S_t)dt + Z_t dW_t \\ Y_T = \frac{\Sigma_{k=1}^n P_{T,k}/C_k}{\Sigma_{k=1}^n 1/C_k} - B S_T \end{cases} \tag{86}$$

Then, using Ito's lemma on $\frac{\Sigma_{k=1}^n P_{T,k}/C_k}{\Sigma_{k=1}^n 1/C_k} S_T^\beta - B S_T S_T^\beta$, we have :

$$d_\beta J(\alpha,\mu) = \mathbb{E}\left[\int_0^T \beta_t^\top \left[-P_t - (diag(C) + C_0 \mathbb{1}_{n,n})\alpha_t + Y_t \mathbb{1}_n\right] dt\right]$$

α being the optimal control means

$$\forall \beta,\ \mathbb{E}\left[\int_0^T \beta_t^\top \left[-P_t - (diag(C) + C_0 \mathbb{1}_{n,n})\alpha_t + Y_t \mathbb{1}_n\right] dt\right] = 0$$

Therefore :

$$\alpha_t = (diag(C) + C_0 \mathbb{1}_{n,n})^{-1}(Y_t \mathbb{1}_n - P_t)$$

Assuming $Y_t = \frac{\Sigma_{k=1}^n P_{t,k}/C_k}{\Sigma_{k=1}^n 1/C_k} - h_2(t) S_t + H_t$, we get :

$$\begin{aligned} dY_t &= dH_t + \frac{\Sigma_{k=1}^n \frac{dP_{k,t}}{C_k}}{\Sigma_{k=1}^n \frac{1}{C_k}} + \left(-h_2'(t)S_t - h_2(t)\mathbb{1}_n^\top (diag(C)+C_0\mathbb{1}_{n,n})^{-1}\left(\left(\frac{\Sigma_{k=1}^n P_{t,k}/C_k}{\Sigma_{k=1}^n 1/C_k} - h_2(t)S_t + H_t\right)\mathbb{1}_n - P_t\right)\right)dt \\ &= -(A_1 - A_2 S_t)dt + Z_t dW_t \end{aligned}$$

and $\mu_t = (diag(C) + C_0\mathbb{1}_{n,n})^{-1}((\frac{\Sigma_{k=1}^n P_{t,k}/C_k}{\Sigma_{k=1}^n 1/C_k} - h_2(t)E_t + H_t)\mathbb{1}_n - P_t)$

(87)

Therefore, denoting $\tilde{C} = \left(\mathbb{1}_n^\top (diag(C)+C_0\mathbb{1}_{n,n})^{-1}\mathbb{1}_n\right)^{-1}$, we have :

$$\begin{cases} h_2'(t) - \dfrac{h_2^2(t)}{\tilde{C}} + A_2 = 0 \\ h_2(T) = B \\ dH_t = \left(-A_1 - \dfrac{\Sigma_{k=1}^n \frac{b_k(t,P_t,\mu_t)}{C_k}}{\Sigma_{k=1}^n \frac{1}{C_k}} + \dfrac{h_2(t)(H_t + \frac{\Sigma_{k=1}^n P_{t,k}/C_k}{\Sigma_{k=1}^n 1/C_k})}{\tilde{C}} - h_2(t)\mathbb{1}_n^\top (diag(C)+C_0\mathbb{1}_{n,n})^{-1}P_t\right)dt + Z_t dW_t \\ H_T = 0 \end{cases} \quad (88)$$

As usual, writing $H_t = h_1(t, P_t, E_t)$, gives the following PDE :

$$\begin{cases} \partial_t h_1 + b^\top \left(\nabla_p h_1 + \dfrac{diag(C)^{-1}\mathbb{1}_n}{\mathbb{1}_n^\top diag(C)^{-1}\mathbb{1}_n}\right) + \dfrac{1}{2}Tr\left(\sigma\sigma^\top D_{pp}h_1\right) - h_2\left(\dfrac{h_1 + \frac{p^\top diag(C)^{-1}\mathbb{1}_n}{\mathbb{1}_n^\top diag(C)^{-1}\mathbb{1}_n}}{\tilde{C}} - p^\top (diag(C)+C_0\mathbb{1}_{n,n})^{-1}\mathbb{1}_n\right) \\ +\partial_e h_1 \mathbb{1}_n^\top \mu(t,p,e) + A_1 = 0 \\ h_1(T) = 0 \end{cases} \quad (89)$$

Assuming $C_0 = 0$ gives :

$$\begin{cases} \partial_t h_1 + b\left(\cdot, \mu(\cdot)\right)^\top \left(\nabla_p h_1 + \dfrac{vect(C^{-1})}{\mathbb{1}_n^\top vect(C^{-1})}\right) + \dfrac{1}{2}Tr\left(\sigma\sigma^\top D_{pp}h_1\right) - \dfrac{h_2 h_1}{\tilde{C}} + \partial_e h_1 \mathbb{1}_n^\top \mu(t,p,e) + A_1 = 0 \\ h_1(T) = 0 \end{cases} \quad (90)$$

10.1.3 Symmetric Bachelier case

To have an explicit solution, let's assume a Bachelier price model like this :

$$dP_t = \left(f_0(t) + \left(\underbrace{\nu_1 I_n}_{\text{endogenous market impact}} + \underbrace{\frac{\nu_2}{n}\mathbb{1}_{n,n}}_{\text{cross market impact}}\right)\mu_t\right)dt + \sigma dW_t$$

We also need to assume that the market access cost is symmetric : $\forall k, C_k = nC \Rightarrow \tilde{C} = C$. The equation (90) becomes :

$$\begin{cases} \partial_t h_1 + \left(f_0 + \left(\nu_1 I_n + \dfrac{\nu_2}{n}\mathbb{1}_{n,n}\right)\mu_t\right)^\top \left(\nabla_p h_1 + \dfrac{\mathbb{1}_n}{n}\right) + \dfrac{1}{2}Tr\left(\sigma\sigma^\top D_{pp}h_1\right) - \dfrac{h_2 h_1}{C} + \partial_e h_1 \mathbb{1}_n^\top \mu + A_1 = 0 \\ h_1(T) = 0 \end{cases} \quad (91)$$

As ${\mathbb{1}_n}^\top \mu(t,p,e) = {\mathbb{1}_n}^\top \frac{\left(\frac{{\mathbb{1}_n}^\top}{n} p - h_2(t)e + h_1(t,p,e)\right)\mathbb{1}_n - p}{nC} = \frac{h_1(t,p,e) - h_2(t)e}{C}$. If we assume that h_1 does not depend on the price. Then the dynamic of E is deterministic and we have the following system :

$$\begin{cases} h_1'(t) + \left(f_0(t) + \left(\nu_1\mu(t) + \frac{\nu_2}{n}({\mathbb{1}_n}^\top \mu(t))\mathbb{1}_n\right)\right)^\top \frac{\mathbb{1}_n}{n} - \frac{h_2(t)h_1(t)}{C} + A_1 = 0 \\ E'(t) = {\mathbb{1}_n}^\top \mu(t) \\ h_1(T) = 0, E(0) = E_0 \end{cases} \tag{92}$$

That is

$$\begin{cases} h_1(t)' + (\nu_1 + \nu_2)\frac{h_1(t) - h_2(t)E(t)}{nC} - \frac{h_2(t)h_1(t)}{C} + A_1 + f_0(t)^\top \frac{\mathbb{1}_n}{n} = 0 \\ E'(t) = \frac{h_1(t) - h_2(t)E(t)}{C} \\ h_1(T) = 0, E(0) = E_0 \end{cases} \tag{93}$$

Denoting $\bar{x} = \left(x^\top \frac{\mathbb{1}_n}{n}\right)$, we recognize the same system than with one market :

$$\begin{cases} CE'' + \frac{\nu_1 + \nu_2}{n}E' - A_2 E = -A_1 - \bar{f}_0 \\ h_1 = CE' + Eh_2 \\ E(0) = E_0, CE'(T) + BE(T) = 0 \end{cases} \tag{94}$$

The control for the k^{th} market writes :

$$\alpha_k(t,s,p) = \frac{h(t) - h_2(t)s}{nC} + \frac{\bar{p} - p_k}{nC} = \bar{\alpha}(t,s) + \frac{\bar{p} - p_k}{nC}$$
$$\mu_k(t,p) = \frac{h(t) - h_2(t)E(t)}{nC} + \frac{\bar{p} - p_k}{nC} = \bar{\mu}(t) + \frac{\bar{p} - p_k}{nC}$$

So in this multi-market Bachelier framework, there is an average strategy exploiting the average price seasonality, added with a price spread strategy, buying more/selling less on markets where the price is lower and vice versa.

Also the dynamic of the average price is :

$$d\bar{P}_t = \left(\bar{f}_0 + (\nu_1 + \nu_2)\bar{\mu}_t)\right) dt + \bar{\sigma} dW_t$$

Thus the spread between the k^{th} price and the average price is mean reverting in an Ornstein-Uhlenbeck way :

$$d(P_{k,t} - \bar{P}_t) = \left((f_{0,k} - \bar{f}_0) - \frac{\nu_1}{nC}(P_{k,t} - \bar{P}_t)\right) dt + (\sigma_k - \bar{\sigma}) dW_t$$

Furthermore we have

$$\begin{cases} h_1'(t) + \bar{f}_0(t) + (\nu_1 + \nu_2)\bar{\mu}(t) - \frac{h_2(t)h_1(t)}{C} + A_1 = 0 \\ \bar{\mu}(t) = \frac{E'(t)}{n} h_1(T) = 0 \end{cases} \tag{95}$$

So the learning scheme is exactly the same as in the case there is only one market, using the the average $\bar{\mu}$.

10.1.4 Numerical results

For the numerical example we take two market driven by two Brownian motion. For both market we take the same seasonality of the form :

$$f_0(t) = Kcos(2\pi t + \phi)$$

We chose the following set of parameters :

$$T = 1,\ A_1 = 5,\ A_2 = 5,\ B = 10,\ C = 0.5,\ P_0 = 100,\ E_0 = 0,\ K = 5,\ \phi = \frac{3\pi}{4},\ \nu = 6,\ \kappa = 1$$

The volatility of the first market is $\sigma_1 = 8$, of the second market $\sigma_2 = 4$ and the correlation is $\rho = 0.5$

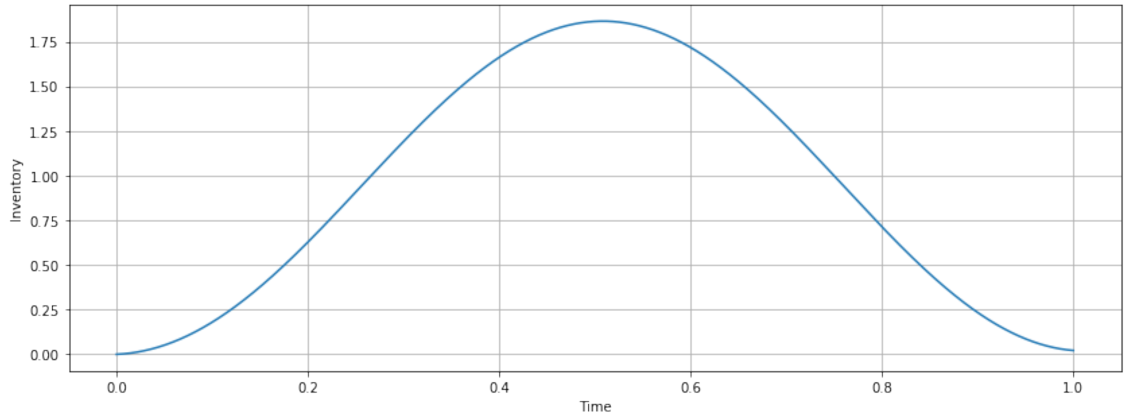

Figure 10.1: $E(t)$

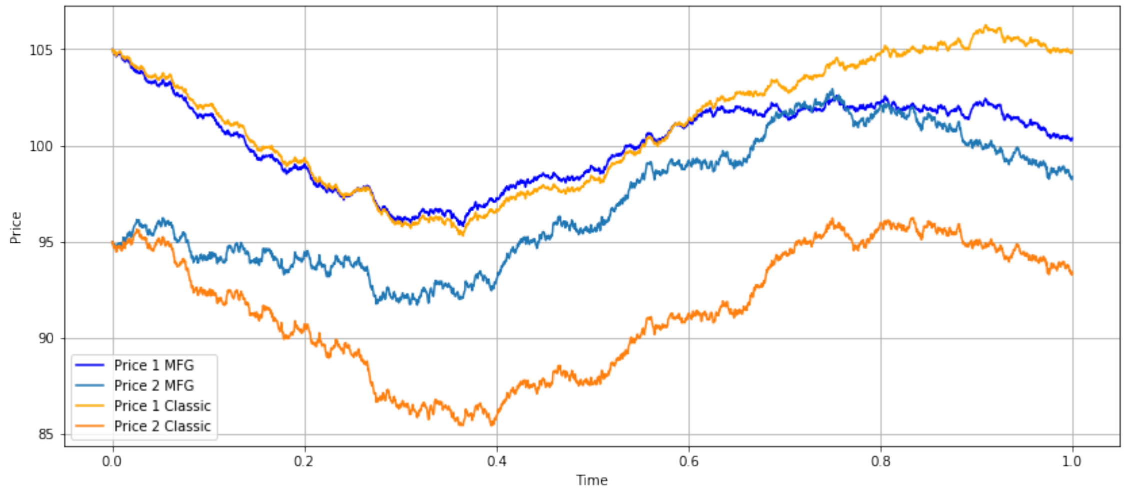

Figure 10.2: Prices with and without market impact

On this figure we have plotted the prices' trajectories of the two markets compared to what they would have been without market impact (so with the same trajectory of Brownian motion).

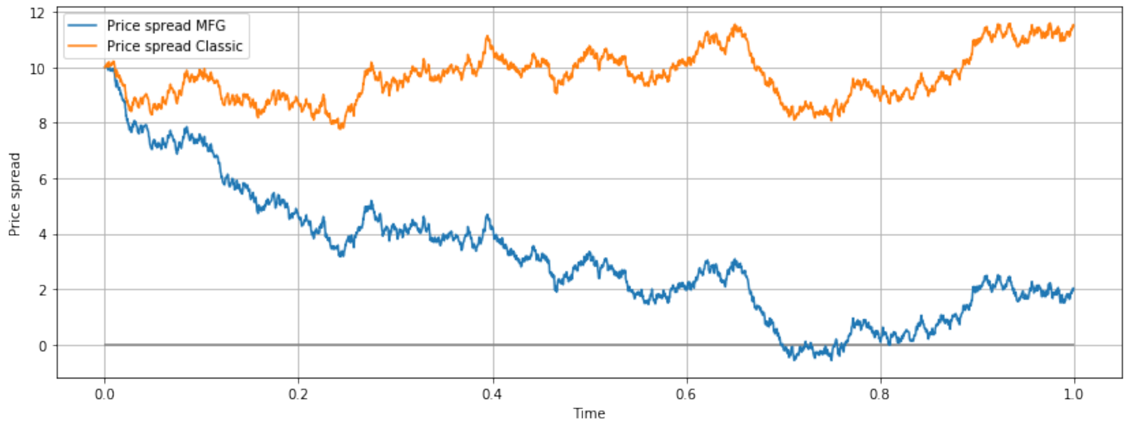

Figure 10.3: Price spread with and without market impact

We can see on this figure that price spread is mean reverting when there is market impact like we've shown in the above proof.

10.2 Inhomogeneous reward functions

Cardaliaguet and Lehalle [13] introduced this extended model for their optimal liquidation problem. We have adapted it to optimal use of a storage.

10.2.1 The model and derivation the equations

As we talked in the sections about Mean Field Games about the possibility to have different reward functions representing different storage capacities and transaction cost among the players. We will keep our linear quadratic framework to have explicit computations. We have now A_1, A_2, B and C depending on the type of agent a.

The density of player m_t now also includes the type of agents : $m_t(ds, da)$. We will assume for the sake of simplicity that preferences of the players does not evolve over time. The distribution of type of agents only depending on the initial condition $\bar{m}_0(da) = \int_s m_0(ds, da)$ Thus we can disintegrate m_t in :

$$m_t(ds, da) = m_t^a(ds)\bar{m}_0(da)$$

We denote $E_t^a = \int_s s m_t^a(ds)$, the mean of the inventories of the players of type a. And in general the stochastic process of application $E_t : a \to E_t^a$ and the state variable $e : a \to e(a)$.

Using the previous sections we have as usual :

$$\alpha_t^a = \frac{H_t^a - h_2^a s}{C^a}$$

$$\mu_t^a = \frac{H_t^a - h_2^a E_t^a}{C^a}$$

$$\mu_t = \int_a \mu_t^a \bar{m}_0(da)$$

$$\begin{cases} dP_t = b(t, P_t, \mu_t)dt + \sigma(t, P_t)dW_t \\ dH_t^a = -\left(b(t, P_t, \mu_t) + \frac{H_t^a h_2^a(t)}{C}^a - A_1^a\right) dt + Z_t^a dW_t \\ dE_t^a = \mu_t^a dt \\ h_2^{a\prime} - \frac{h_2^a}{C^a} + A_2^a = 0 \\ h_2^a(T) = B^a, H_T^a = 0 \end{cases} \tag{96}$$

We assume that $H_t^a = h^a(t, P_t, E_t)$, therefore :

$$\begin{cases} \partial_t h^a + (1+\partial_p h^a) b(t,p,\mu(t,p,e)) + \frac{1}{2}\sigma^2 \partial_{pp} h^a + \partial_e h^a [a \to \mu_t^a] - \frac{h_2^a(t) h^a(t,p,e)}{C^a} + A_1^a = 0 \\ h^a(T) = 0 \end{cases} \tag{97}$$

10.2.2 Bachelier case

In the Bachelier case, assuming that h do not depend on the price, the dynamic of E_t becomes deterministic therefore the system rewrites :

$$\begin{cases} (h^a)'(t) + f_0(t) + \nu \int_{\tilde a} \frac{h^{\tilde a}(t) - h_2^{\tilde a}(t) E^{\tilde a}(t)}{C^{\tilde a}} \bar m_0(d\tilde a) - \frac{h_2^a(t) h^a(t)}{C^a} + A_1^a = 0 \\ (E^a)'(t) = \frac{h^a(t) - h_2^a(t) E^a(t)}{C} \\ h^a(T) = 0, E^a(0) = E_0^a \end{cases} \tag{98}$$

In the most general case where there is a continuum of preferences, the system is an infinite number of coupled parametric equations. In the case where the set of preferences is finite, the system is finite. Let's take the example where there are two kinds of players, evenly distributed :

$$\bar m_0(da) = \frac{\delta_x + \delta_y}{2}(da)$$

Then the system is :

$$\begin{cases} (h^x)'(t) + f_0(t) + \nu \left(\frac{h^x(t) - h_2^x(t) E^x(t) + h^y(t)}{2C^x} + \frac{h_2^y(t) E^y(t)}{2C^y} \right) - \frac{h_2^x(t) h^x(t)}{C^x} + A_1^x = 0 \\ (h^y)'(t) + f_0(t) + \nu \left(\frac{h^x(t) - h_2^x(t) E^x(t) + h^y(t)}{2C^x} + \frac{h_2^y(t) E^y(t)}{2C^y} \right) - \frac{h_2^y(t) h^y(t)}{C^y} + A_1^y = 0 \\ (E^x)'(t) = \frac{h^x(t) - h_2^x(t) E^x(t)}{C^x} \\ (E^y)'(t) = \frac{h^y(t) - h_2^y(t) E^y(t)}{C^x} \\ h^x(t) = h^y(T) = 0, E^x(0) = E_0^x, E^y(0) = E_0^y \end{cases} \tag{99}$$

10.2.3 Numerical results

For the numerical example we took, $A_1^x = 5$, $A_2^x = 5$, $A_1^y = 10$, $A_2^y = 5$, $B^x = B^y = 10$ and $C^x = C^y = 1$. The values are taken to represent that players of type x have greater storage capacity than players of type y. The average initial inventory for each group is $E_0^x = E_0^y = 0$. For the price model we took $K = 30$, $\phi = \frac{3\pi}{4}$, $\nu = 6$, $p_0 = 100$

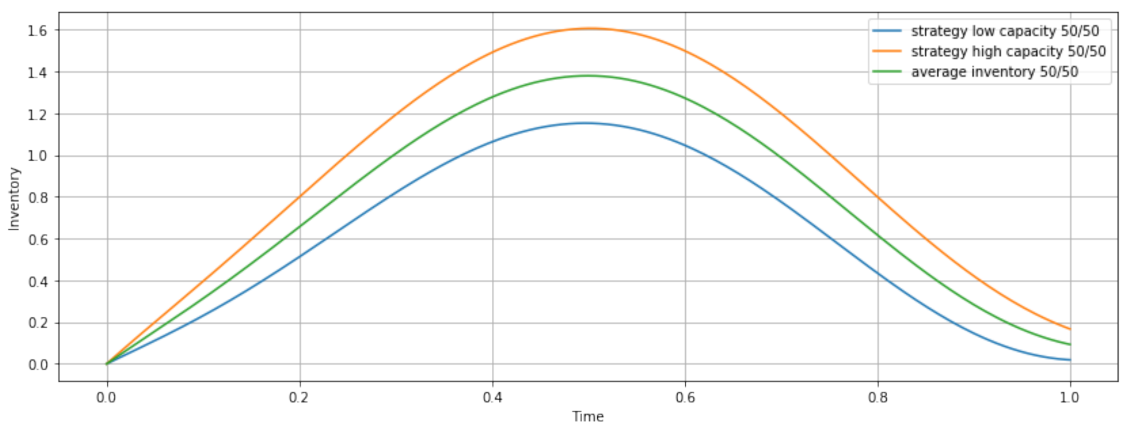

Figure 10.4: Average inventory for each group of players

We plotted the solutions for E^x and E^y in blue and yellow, and the resulting $E = \frac{E^x+E^y}{2}$ in green.

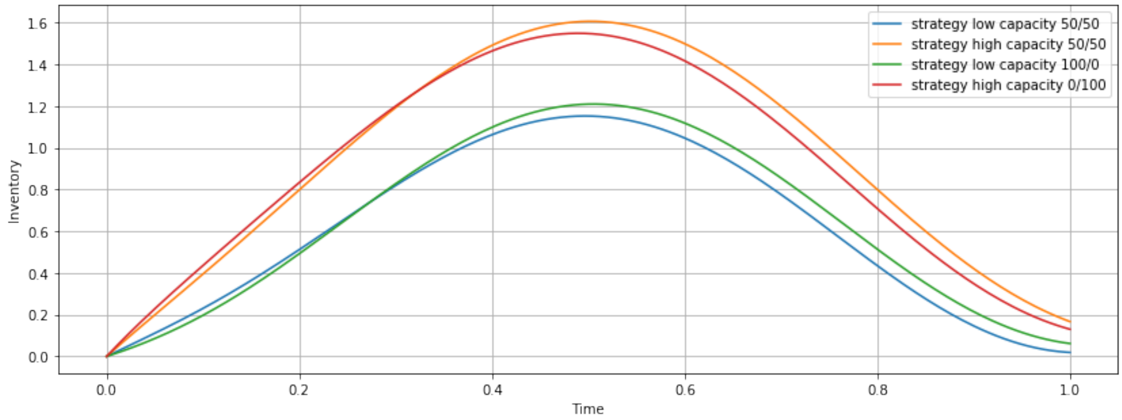

Figure 10.5: Comparison with the cases where the all the players are identical

We also plotted the solutions for the cases where the players think all other player have the same type as them, in red and green. We can see how the high capacities buy even more than when they are alone and the opposite for the low capacities. It is difficult to understand really why. We can assume that the bigger μ is the less the players buy. Then since μ is smaller in the mixed player case for the high capacities, they buy more, and vice versa for the small capacities.

10.2.4 Learning

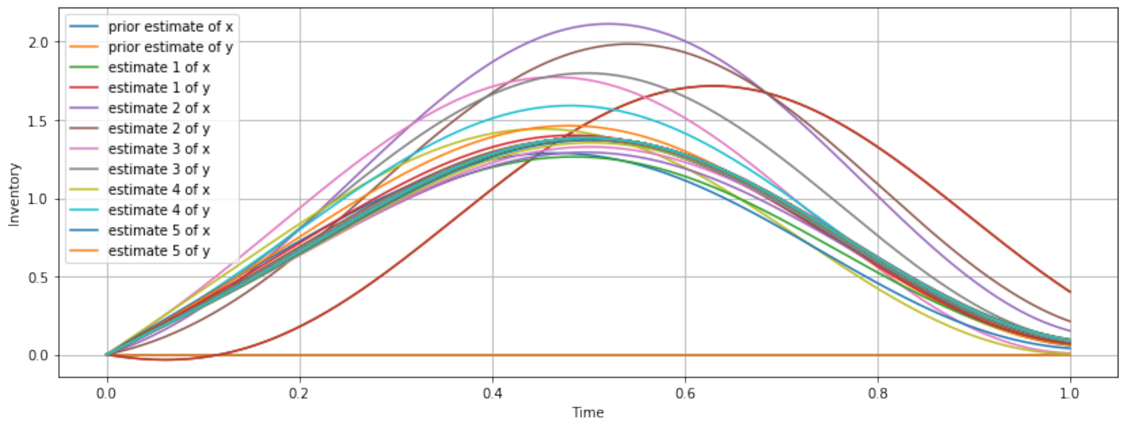

Figure 10.6: Learning trajectories

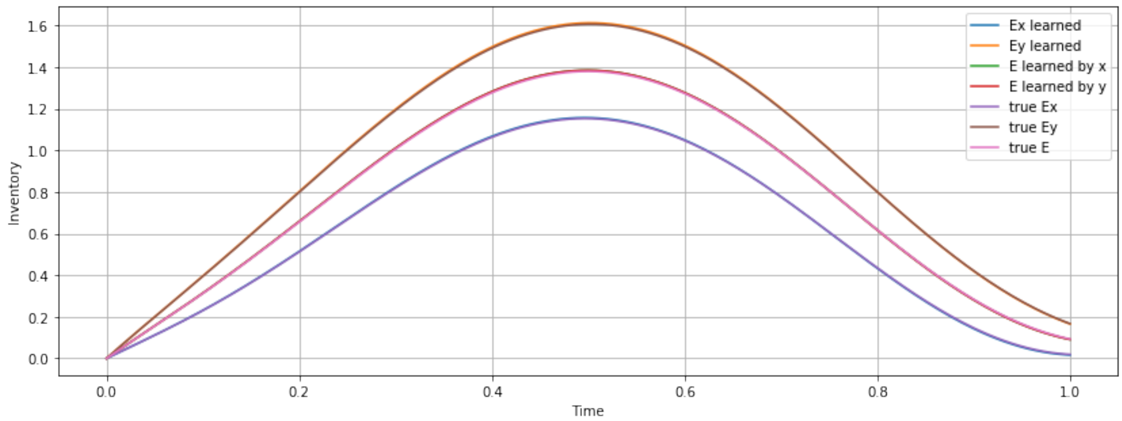

Figure 10.7: Learning results vs system solutions

Ex learned	Trajectory of players of type x computed according to their final estimate of μ
Ey learned	Trajectory of players of type y computed according to their final estimate of μ
E learned by x	Players of type x's final estimate of $E = \int_0^{\cdot} \mu(t)dt$
E learned by y	Players of type y's final estimate of $E = \int_0^{\cdot} \mu(t)dt$
true Ex	E^x as solution of the system of equations
true Ey	E^y as solution of the system of equations
true E	$E = \frac{E^x+E^y}{2}$

For the learning, each players of the same type share their learning method, that is that at any game they all have the same estimate of what will be $\mu(t)$ and therefore the same strategy. For the example, both groups have the same prior but different learning rates : $\alpha^x = 0.3$ and $\alpha^y = 0.7$. We can see that the strategy eventually converges toward the solution of the equation, as the curves coincides.

10.3 Inhomogeneous price models beliefs

The idea for this extended model comes from the inhomogeneous reward functions model. Like it is unlikely that in reality all agent are identical, it is unlikely that they use the same pricing models with same calibration.

10.3.1 Model and derivation of the equation

Let's imagine now instead that the players do not use the same price model. Each price model resulting in a strategy.

Remark 13

The players knows of the price model used by other players and are aware that other players are in the same position. Also there is no need for a 'true price model', the players' control is a function of the time , price, and distribution of inventories and beliefs .

The density of player m_t now also includes the type of agents : $m_t(ds, da)$. We will assume for the sake of simplicity that beliefs of the players does not evolve over time. The distribution of type of agents only depending on the initial condition $\bar{m}_0(da) = \int_s m_0(ds, da)$ Thus we can disintegrate m_t in :

$$m_t(ds, da) = m_t^a(ds)\bar{m}_0(da)$$

We denote $E_t^a = \int_s s m_t^a(ds)$, the average inventory of the players of type a. And in general the stochastic process of application $E_t : a \to E_t^a$ and the state variable $e : a \to e(a)$.

Using the previous sections we have as usual :

$$\alpha_t^a = \frac{H_t^a - h_2 s}{C}$$
$$\mu_t^a = \frac{H_t^a - h_2 E_t^a}{C}$$
$$\mu_t = \int_a \mu_t^a \bar{m}_0(da)$$

In this framework we still have :

$$\begin{cases} {h_2}' - \dfrac{h_2}{C} + A_2 = 0 \\ h_2(T) = B \end{cases} \tag{100}$$

$$\begin{cases} dH_t^a = -\left(b^a(t, P_t, \mu_t) + \dfrac{H_t^a h_2(t)}{C} - A_1\right) dt + Z_t^a dW_t \\ dE_t^a = \mu_t^a dt \\ H_T^a = 0 \end{cases} \tag{101}$$

We assume that $H_t^a = h^a(t, P_t, E_t)$, therefore :

$$\begin{cases} \partial_t h^a + (1 + \partial_p h^a) b^a(t, p, \mu(t, p, e)) + \dfrac{1}{2}(\sigma^a(t, p))^2 \partial_{pp} h^a + \partial_e h^a [a \to \mu_t^a] - \dfrac{h_2^a(t) h^a(t, p, e)}{C} + A_1 = 0 \\ h^a(T) = 0 \end{cases} \tag{102}$$

10.3.2 Bachelier case

In order to have explicit computations, we assume that all the beliefs are Bachelier price models with different parameters.

$$dP_t = (f_0^a(t) + \nu^a \mu_t)\, dt + \sigma^a dW_t$$

Assuming that h do not depend on the price make that the process E_t is deterministic.

$$\begin{cases} (h^a)'(t) + f_0^a(t) + \nu^a \displaystyle\int_{\tilde{a}} \frac{h^{\tilde{a}}(t) - h_2(t) E^{\tilde{a}}(t)}{C} \bar{m}_0(d\tilde{a}) - \frac{h_2(t) h^a(t)}{C} + A_1 = 0 \\ (E^a)'(t) = \frac{h^a(t) - h_2(t) E^a(t)}{C} \\ h^a(T) = 0, E^a(0) = E_0^a \end{cases} \tag{103}$$

Denoting $h = \int_a h^a \bar{m}_0(da)$, $E = \int_a E^a \bar{m}_0(da)$, $\nu = \int_a \nu^a \bar{m}_0(da)$ and $f_0 = \int_a f_0^a \bar{m}_0(da)$ we have :

$$\begin{cases} (h^a)'(t) + f_0^a(t) + \nu^a \dfrac{h(t) - h_2(t) E(t)}{C} - \dfrac{h_2(t) h^a(t)}{C} + A_1 = 0 \\ h'(t) + f_0(t) + \nu \frac{h(t) - h_2(t) E(t)}{C} - \frac{h_2(t) h(t)}{C} + A_1 = 0 \\ (E^a)' = \frac{h^a(t) - h_2(t) E^a(t)}{C} \\ E'(t) = \frac{h(t) - h_2(t) E(t)}{C} \\ h^a(T) = h(T) = 0, E^a(0) = E_0^a, E(0) = E_0 \end{cases} \tag{104}$$

Thus we have the same system than usual for h and E. Then $h^a(t) = \int_t^T \exp\left(-\int_t^s \frac{h_2(u)}{C} du\right) \left(A_1 + f_0^a(s) + \nu^a \frac{h(t) - h_2(t) E(t)}{C}\right)$ and $E^a(t) = \exp\left(-\int_0^t \frac{h_2(s)}{C} ds\right) \left(E_0^a + \int_0^t \exp\left(+\int_0^s \frac{h_2(u)}{C} du\right) \frac{h^a(s)}{C} ds\right)$.

10.3.3 Numerical results

For the numerical example we take two kinds of players evenly distributed. We take $K^x = K^y = 30$, $\phi^x = \phi^y = \frac{3\pi}{4}$, $\nu^x = 8$, $\nu^y = 4$ and $p_0 = 100$.
The values are taken to represent that players of type y are more optimistic on the market market impact (they think there is less) than the players of type x . The average initial inventory for each group is $E_0^x = E_0^y = 0$. For the storage model we took $A_1 = 5$, $A_2 = 5$, $B = 10$ and $C = 1$.

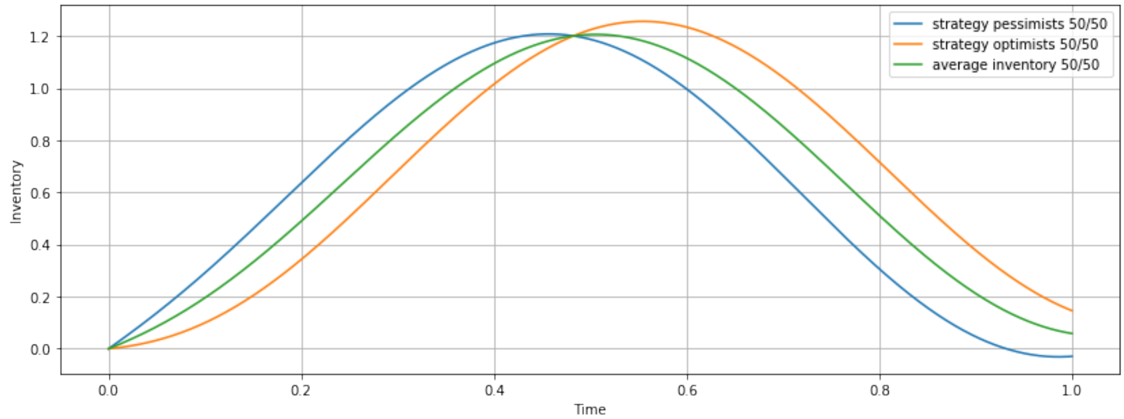

Figure 10.8: Average inventory for each group of players

We notice how pessimistic players buy and sell earlier than the optimists.

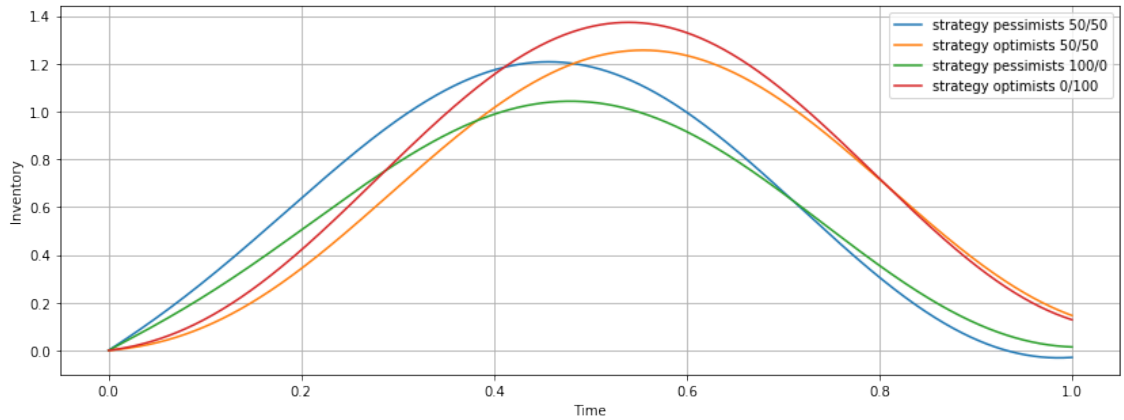

Figure 10.9: Comparison with the cases where the all the players are identical

On this figure we plotted what would have been the strategies if each group of players represented 100% of the players, allowing us to compare these strategy with the case where there is 50% of each type. We notice that the optimists trade more slowly than when they are alone, and the opposite for the pessimists. Like the inhomogeneous rewards model, it is difficult to interpret this difference.

10.3.4 Learning

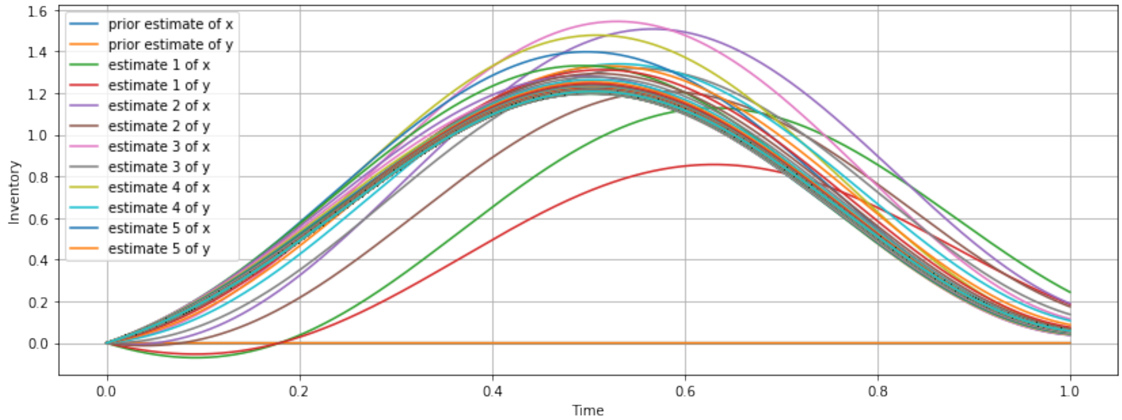

Figure 10.10: Learning trajectories

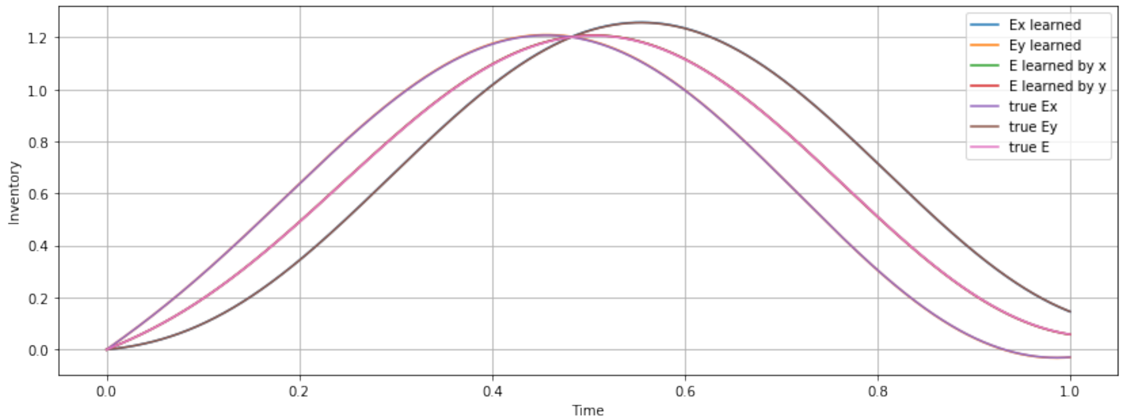

Figure 10.11: Learning results vs system solutions

Ex learned	Trajectory of players of type x computed according to their final estimate of μ
Ey learned	Trajectory of players of type y computed according to their final estimate of μ
E learned by x	Players of type x's final estimate of $E = \int_0^{\cdot} \mu(t)dt$
E learned by y	Players of type y's final estimate of $E = \int_0^{\cdot} \mu(t)dt$
true Ex	E^x as solution of the system of equations
true Ey	E^y as solution of the system of equations
true E	$E = \frac{E^x+E^y}{2}$

In this learning model, we do not make the player adapt their price model to the actual price behaviour they see at each round, they only adapt their estimate of $\mu(t)$ for the next round. If we would do a learning procedure on the price model itself, the model of each player would converge to the 'true price model' and there would be no more inhomogeneity in the models. We first show the learning strategies for a learning rate of , with different prior for the two groups. On the second graph we can see that that the final strategy of each group is the same than the ones obtained by solving the equations.

10.4 General comments

It is possible to combine all these extensions at the same time. Using the learning method is a very handy and intuitive way to solve the system when there are too many different coupled equations. It allows to replicate the real world where agents have different storages, beliefs and learning procedures.

There is also the possibility to have beliefs on the initial distribution of players. Distribution of initial inventories but also storage capacities, price model beliefs, and beliefs on the others. The difference with before is that each player has a different belief on what will be the average control $\mu(t, P_t, E_t)$. Then we already seen that the strategy of a generic player is the optimal strategy given by a classic Bellman optimization principle. The belief on the others' characteristics gives beliefs on what the others think is $\mu(t, p, e)$, from these beliefs we can deduce a μ. Therefore having a belief on the others can be resumed to having a belief on μ, from which you compute your optimal strategy belief associated with your price model.

This remark is to open the discussion concerning the hard constraints case. In the end we compute our control from the μ and our price model, the μ being solution of a PDE depending on the storage capacity (or what are the rewards function) and price model. But the storage capacity and price model of the other players can only be estimated a priori. The uncertainty on the others induce an uncertainty on μ. So it can be more suitable to estimate directly μ rather than solving numerically an equation which add uncertainty on the already existing uncertainty from the belief on the others storage capacity and price model. The difficulty to read m_t or μ_t at given time, since it require to have access to most storages' inventories.

11 Conclusion

After presenting the storage optimization problem in the classical case, we made an introduction to Mean Field theory. We then proposed some extensions of the classical storage optimization by using price models that take into account the flow of trades on the market, this resulted in Mean Field Games problems. We used a linear-quadratic storage constraints assumption for the sake of pedagogy as it allows to resume the distribution of inventory to its mean. The Bachelier price model is a simple model from which explicit solutions can be given, allowing to easily deduce stylizing facts of more complex models. For more complex model like Black-Scholes or Clewlow-Strickland, the problem resumes to solving a parabolic PDE, finite differences can be used. Learning is a promising way to solve the problems as the MFG problems are fixed point problems, but we could only use it in the Bachelier case. However it is easily adaptable to the extensions of the Bachelier cases with multiple markets and multiple storage capacities and beliefs among the players. These extended models in general can be used to model lots of different kinds of market agents, not only storage managers. The calibration we tried to do unfortunately showed that models of market impacts are not relevant for this scale of time but we are confident that the results we showed are still useful and that our model could be adapted for other problems, for example intraday trading.

12 Appendix

12.1 Heuristic derivation of the Hamilton-Jacobi-Bellman equation

Let's take a random process X :

$$dX_t = b(t, X_t, \alpha_t)dt + \sigma(t, X_t, \alpha_t)dW_t$$

We have the stochastic optimization problem :

$$\begin{cases} J(t,x,\alpha) = \mathbb{E}_{t,x}\left[\int_t^T f(s, X_s, \alpha_s)ds + g(X_T)\right] \\ v(t,x) = \sup_{\alpha\in\mathcal{A}} J(t,x,\alpha) \end{cases} \tag{105}$$

The dynamic programming principle gives :

$$v(t,x) = \sup_{\alpha\in\mathcal{A}} \mathbb{E}_{t,x}\left[v(t+h, X_{t+h}) + \int_t^{t+h} f(s, X_s, \alpha_s)ds\right]$$

Therefore, by Ito's lemma :

$$0 = \sup_{\alpha\in\mathcal{A}} \mathbb{E}_{t,x}\left[\int_t^{t+h} (\partial_t v + D_x v \cdot b + \frac{1}{2}Tr(D_{xx}v(\sigma\sigma^\top)) + f)(s, X_s, \alpha_s)ds\right]$$

Taking h to zero gives :

$$\partial_t v(t,x) + \sup_{\alpha\in A} D_x v(t,x)\cdot b(t,x,\alpha) + \frac{1}{2}Tr(D_{xx}v(t,x)(\sigma\sigma^\top)(t,x,\alpha)) + f(s,x,\alpha) = 0$$

Also, $v(T,x) = \mathbb{E}_{T,x}\left[g(X_T)\right] = g(x)$
Thus we have the HJB system :

$$\begin{cases} \partial_t v(t,x) + H(t,x,D_x v(t,x), D_{xx}v(t,x)) = 0 \\ v(T,x) = g(x) \end{cases} \tag{106}$$

With H the Hamiltonian : $H(t,x,p,\gamma) = \sup_{\alpha\in A} p\cdot b(t,x,\alpha) + \frac{1}{2}Tr(\gamma(\sigma\sigma^\top)(t,x,\alpha)) + f(s,x,\alpha)$

12.2 Heuristic derivation of the Fokker-Planck equation

Let's take a random process X :

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$

To show the Fokker-Planck equation, we compute the derivation of an expectation by two different ways :

$$\mathbb{E}[f(X_t)] = \int f(x)m_t(dx)$$

$$\Longrightarrow \partial_t\mathbb{E}[f(X_t)] = \int f(x)\partial_t m_t(dx)$$

By Ito's lemma, we have :

$$\mathbb{E}[D_x f(X_t)\cdot b(t,X_t) + \frac{1}{2}Tr(D_{xx}f(X_T)(\sigma\sigma^\top)(t,X_t)] = \int f(x)\partial_t m_t(dx)$$

$$\Longrightarrow \int (D_x f(x) \cdot b(t,x) + \frac{1}{2} Tr(D_{xx} f(x)(\sigma\sigma^\top)(t,x)) m_t(dx) = \int f(x) \partial_t m_t(dx)$$

Then by integration by parts :

$$\int f(x)(-div(b(t,x)m_t(x)) + \frac{1}{2}\sum_{i,j} \partial_{x_i,x_j}(m_t(x)(\sigma\sigma^\top)_{i,j}(t,x)))dx = \int f(x)\partial_t m_t(x)dx$$

It is true for any f, therefore we get the Fokker-Planck equation :

$$\partial_t m_t(x) + div(b(t,x)m_t(x)) - \frac{1}{2}\sum_{i,j} \partial_{x_i,x_j}(m_t(x)(\sigma\sigma^\top)_{i,j}(t,x)) = 0$$

Bibliography

[1] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games : Numerical methods. 2009.

[2] C. Alasseur, I. Ben Tahar, and A. Matoussi. An extended mean field game for storage in smart grids. 2018.

[3] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. 2000.

[4] F. Bagagiolo and D. Saude. Mean-field games and dynamic demand management in power grids. 2018.

[5] M. Bardi. Explicit solutions of some linear-quadratic mean field games. 2012.

[6] O. Bardou, S. Bouthemy, and G. Pages. When are swing options bang bang and how to use it. 2007.

[7] A. Bensoussan, M. H. M. Chaud, and S. C. P. Yam. Mean field games with a dominating player. 2014.

[8] A. Bensoussan, J. Frehse, and S. C. P. Yam. The master equation in mean field theory. 2014.

[9] A. Bensoussan, K. C. J. Sung, S. C. P. Yam, and S. P. Yung. Linear-quadratic mean field games. 2014.

[10] A. Boogert and C. de Jong. Gas storage valuation using a monte carlo method. 2006.

[11] P. Cardaliaguet. Notes on mean field game. 2012.

[12] P. Cardaliaguet and S. Hadikhanloo. Learning in mean field games: the fictitious play. 2015.

[13] P. Cardaliaguet and C.-A. Lehalle. Mean field game of controls and an application to trade crowding. 2017.

[14] R. Carmona and F. Delarue. Forward-backward stochastic differential equations and controlled mckean vlasov dynamics. 2011.

[15] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. 2011.

[16] R. Carmona and F. Delarue. The master equation for large population equilibriums. 2013.

[17] R. Carmona, F. Delarue, and D. Lacker. Mean field games with common noise. 2015.

[18] R. Carmona, J.-P. Fouque, and L.-H. Sun. Mean field games and systematic risk. 2013.

[19] R. Carmona, C. V. Graves, and Z. Tan. Price of anarchy for mean field games. 2018.

[20] L. Clewlow and S. Strickland. A multi-factor model for energy derivatives. 1999.

[21] J. C. Cox, A. R. Ross, and M. Rubinstein. Option pricing : a simplifeid approach. 1979.

[22] F. Delarue and S. Menozzi. A forward-backward stochastic algorithm for quasi-linear pdes. 2005.

[23] J. Jr. Douglas, J. Ma, and P. Protter. Numerical methods for forward-backward stochastic differential equations. 1996.

[24] D. Firoozi, P. E. Caines, and S. Jaimungal. Mean field game systems with common noise and markovian latent processes. 2018.

[25] E. Følstad. Numerical methods for valuation and optimal operation of natural gas storage. 2015.

[26] D. A. Gomes and J. Saude. A mean-field game approach to price formation in electricity markets. 2018.

[27] D. A. Gomes and V. K. Voskanyan. Extended deterministic mean-field games. 2018.

[28] J. Graber. Linear quadratic mean field type control and mean field games with common noise, with application to productuion of an exchaustible resource. 2018.

[29] X. Guo, A. Hu, R. Xu, and J. Zhang. Learning mean-field games. 2019.

[30] S Hadikhanloo. Learning in mean field games. 2018.

[31] M. Huang, R. P. Malhame, and P. E. Caines. Large population stochastic dynamic games : closed-loop mckean=vlasov systems and the nash certainty equivalence principle. 2006.

[32] V. N. Kolokoltsov and M. Troeva. On the mean field games with common noise and the mckean-vlasov spdes. 2015.

[33] Lacima. Gas storage : Overwiew and static valuation. 2008.

[34] Lacima. Gas storage : Rolling intrinsic valuation. 2009.

[35] J.-M. Lasry and P.-L. Lions. Mean field games. 2007.

[36] M. Lauriere and O. Pironneau. Dynamic programming for mean-field type control. 2014.

[37] C.-A. Lehalle and C. Mouzouni. A mean field game of portfolio trading and its consequences on perceived correlations. 2019.

[38] S. E. Ludwig, J. A. Sirignano, R. Huang, and G. Papanicolaou. A forward-backward algorithm for stochastic control problems. 2005.

[39] J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly - a four step scheme. 1994.

[40] G. N. Milstein and M. V. Tretyakov. Numerical algorythms for forward-backward stochastic differential equations. 2006.

[41] C. Orrieri, A. Porretta, and G. Savare. A variational approach to the mean field planning problem. 2018.

[42] M. Thompson, M. amd Davison and H. Rasmussen. Natural gas storage valuation and optimization : A real options application. 2009.

[43] N. Touzi. Optimal stochastic control, stochastic target problems, and backward sde. 2010.