Mean Field Game of Controls and An Application To Trade Crowding

Pierre Cardaliaguet, Charles-Albert Lehalle

Mean Field Game of Controls and An Application To Trade Crowding

Pierre Cardaliaguet* and Charles-Albert Lehalle†

September 19, 2017

Abstract

In this paper we formulate the now classical problem of optimal liquidation (or optimal trading) inside a Mean Field Game (MFG). This is a noticeable change since usually mathematical frameworks focus on one large trader facing a "background noise" (or "mean field"). In standard frameworks, the interactions between the large trader and the price are a temporary and a permanent market impact terms, the latter influencing the public price.

In this paper the trader faces the uncertainty of fair price changes too but not only. He also has to deal with price changes generated by other similar market participants, impacting the prices permanently too, and acting strategically.

Our MFG formulation of this problem belongs to the class of "extended MFG", we hence provide generic results to address these "MFG of controls", before solving the one generated by the cost function of optimal trading. We provide a closed form formula of its solution, and address the case of "heterogenous preferences" (when each participant has a different risk aversion). Last but not least we give conditions under which participants do not need to instantaneously know the state of the whole system, but can "learn" it day after day, observing others' behaviors.

1 Introduction

Optimal trading (or optimal liquidation) deals with the optimization of a trading path from a given position to zero in a given time interval. Once a large asset manager takes the decision to buy or to sell a large number of shares or contracts, he needs to implement his decision on trading platforms. He has to be fast enough so that the traded price is as close as possible to his decision price, while he needs to take care of his market impact (market impact is the way his trading pressure moves the market prices, including his own prices, a detrimental way; see [Lehalle et al., 2013, Chapter 3] for details).

The academic answers to this need goes from mean-variance frameworks (initiated by [Almgren and Chriss, 2000]) to more stochastic and liquidity driven ones (see for instance [Guéant and Lehalle, 2015]). Fine modeling of the interactions between the price dynamics and the asset manager trading process is difficult. The reality is probably a superposition of a continuous time evolving process on the price formation side and of an impulse control-driven strategy on the asset manager or trader side (see [Bouchard et al., 2011]).

The modeling of market dynamics for an optimal trading framework is sophisticated ([Guilbaud and Pha proposes a fine model of the orderbook bid and ask, [Obizhaeva and Wang, 2005] suggests a martingale relaxation of the consumed liquidity, and [Alfonsi and Blanc, 2014] uses an Hawkes

*CEREMADE, Université Paris-Dauphine, Place du Marchal de Lattre de Tassigny, 75775 Paris cedex 16 (France)

†Capital Fund Management (CFM), Paris (France) and Imperial College, London (UK).

process, among others). Part of the sophistication comes from the fact "market dynamics" are in reality the aggregation of the behaviour of other asset managers, buying or selling thanks to similar optimal schemes, behind the scene. The usual framework for optimal trading is nevertheless the one of one large privileged asset manager, facing a "mean field" (or a "background noise") made of the sum of the behaviour of other market participants.

An answer to the uncertainty on the components of an appropriate model of market dynamics (and especially of the "market impact", see empirical studies like [Bacry et al., 2015] or [Almgren et al., 2005] for details) could be to implement a robust control framework. Usually robust control consists in introducing an adversarial player in front of the optimal strategy, implementing systematically the worst decision (inside a well-defined domain) for the optimal player (see [Bernhard, 2003] for an extreme framework). Because of the adversarial player, one can hope that the uncertainty around the modeling cannot be used at his advantage by the optimal player. To authors' knowledge, it has not been proposed for optimal trading by now.

Since we know the structure of market dynamics at our time scale of interest (i.e. a mix of players implementing similar strategies), another way to obtain robust results is to model directly this game, instead of the superposition of an approximate mean field and a synthetic sequences of adversarial decisions.

Our work takes place within the framework of Mean Field Games (MFG). Mean Field Game theory studies optimal control problems with infinitely many interacting agents. The solution of the problem is an equilibrium configuration in which no agent has interest to deviate (a Nash equilibrium). The terminology and many substantial ideas were introduced in the seminal papers by Lasry and Lions [Lasry and Lions, 2006a, Lasry and Lions, 2006b, Lasry and Lions, 2007]. Similar models were discussed at the same time by Caines, Huang and Malhamé (see, e.g., [Huang et al., 2006]), who computed explicit solutions for the linear-quadratic (LQ) case (see also [Bensoussan et al., 2016]): we use these techniques in our specific framework. Applications to economics and finance were first developed in [Guéant et al., 2011]. Since these pioneering works the literature on MFG has grown very fast: see for instance the monographs or the survey papers [Bensoussan et al., 2013, Caines, 2015, Gomes and Others, 2014].

The MFG system considered here for the optimal trading differs in a substantial way from the standard ones by the fact that the mean field is not, as usual, on the position of the agents, but on their controls: this specificity is dictated by the model. Similar—more general—MFG systems were introduced in [Gomes et al., 2014] under the terminology of *extended* MFG models. In [Gomes et al., 2014, Gomes and Voskanyan, 2016], existence of solutions is proved for deterministic MFG under suitable structure assumptions. A chapter in the monograph [Carmona and Delarue, 2017] is also devoted to this class of MFG, with a probabilistic point of view. Since our problem naturally involves a degenerate diffusion, we provide a new and general existence result of a solution in this framework. Our approach is related with techniques for standard MFG of first order with smoothing coupling functions (see [Lasry and Lions, 2007] or [Cardaliaguet and Hadikhanloo, 2017]).

As the MFG system is an equilibrium configuration—in which theoretically each agent has to know how the other agents are going to play in order to act optimally—it is important to explain how such a configuration can pop up in practice. This issue is called "learning" in game theory and has been the aim of a huge literature (see, for instance, the monograph [Fudenberg and Levine, 1998]). The key idea is that an equilibrium configuration appears—without coordination of the agents—because the game has been played sufficiently many times. In MFG theory, the closely related concept of adaptative control was implemented, for infinite horizon problems, in [Kizilkale and Caines, 2013, Nourian et al., 2012]. The first explicit reference to learning in MFG theory can be found in [Cardaliaguet and Hadikhanloo, 2017]. This idea seems very meaningful for our optimal trading problem where the asset managers are lead to buy or sell daily a large number of shares or contracts. We implement the learning procedure within the simple LQ-framework: we show that, when the game has been repeated a sufficiently large number of times, the agents—without coordination and subject to measurement errors—actually implement trading strategies which are close to the one suggested by the MFG theory.

The starting point of this paper is an *optimal liquidation problem*: a trader has to buy or sell (we formulate the problem for a sell) a large amount of shares or contracts in a given interval of time (from $t = 0$ to $t = T$). The reader can think about T as typically being a day or a week. Like in most optimal liquidation (or optimal trading) problems, the utility function of the trader has three components: the state of the cash account at terminal time T (i.e. the more money the trader obtained from the large sell, the better), a liquidation value for the remaining inventory (with a penalization corresponding to a large market impact for this large "block"), and a risk aversion term (corresponding to not trading instantaneously at his decision price). As usual again, the price dynamics will be influenced by the sells of the trader (via a *permanent market impact* term: the faster the trader trades, the more he impacts the price in a detrimental way); on the top of this cost, the trader will suffer from *temporary market impact*: this will not change the public price but his own price (the reader can think about the "cost of liquidity", like a bid-ask spread cost). In their seminal paper [Almgren and Chriss, 2000], Almgren and Chriss noticed such a framework gives birth to an interesting optimization problem for the trader: on the one hand if he trades too fast he will suffer from market impact and liquidity costs on his price, but on the other hand if he trades too slow, he will suffer from a large risk penalization (the "fair price" will have time to change a detrimental way). Once expressed in a dynamic (i.e. time dependent) way, this optimization turns into a stochastic control problem. For details about this standard framework, see [Cartea et al., 2015], [Guéant, 2016] or [Lehalle et al., 2013, Chapter 3].

In the standard framework, the trader faces a *mean field*: a Brownian motion he is the only one to influence via his permanent market impact. This will no longer be the case in this paper: on the top of a Brownian motion (corresponding to the unexpected variations of the *fair price* while market participants are trading), we add the consequences of the actions of a *continuum of market participants*. Each participant has to buy or sell a number of shares or contracts q (positive for sellers and negative for buyers). This continuum is characterized by the density of the remaining inventory of participants $dm(t, q)$. A variable of paramount importance is the net inventory of all participants: $E(t) := \int_q q\, m(t, q)\, dq$.

To Authors' knowledge, only two papers are related to our approach: one by Carmona et al., using MFG for fire sales [Carmona et al., 2013], and one by Jaimungal and Nourian [Jaimungal and Nourian, 2015] for optimal liquidation of one large trader in front of smaller ones. They are nevertheless different from ours: the first one (fire sales) has not the same cost function as in optimal liquidation problems, and the second one investigates the behavior of a large trader (having to sell substantially more shares or contracts than the others, with a risk aversion) facing *a crowd of small traders* (with a lot of small inventories and no risk aversion). The topic of this second paper is more the one of a large asset manager trading in front of high frequency traders.

In this paper we assume that the public price is influenced by the permanent market impact of all market participants. Note that, conversely, all market participants face the public price thus affected. It corresponds to the day-to-day reality of traders: it is not a good news to buy while other participants are buying, but it is good to have to buy while others are selling. In such a configuration, the participants act strategically, taking into account all the information they have. As explained in Section 2, this leads to a Nash equilibrium of MFG type, in which the mean field depends on the agents' actions. This Nash MFG equilibrium can be summarized by a system of forward-backward PDEs (see (9)), coupling a (backward) Hamilton-Jacobi equation with a (forward) Kolmogorov equation. When the preferences of the agents are homogeneous, this system can be solved explicitly and displays interesting—and not completely intuitive—features (Section 3). For instance, one can notice that the coefficient affecting the permanent market impact has a strong influence on the whole system: the highest the coefficient, the fastest the market players have to drive their inventory to zero. Another interesting situation is when a participant is close to a zero inventory (or for low terminal constraints and low risk aversion). It can then be rewarding to "follow the crowd": a seller may have interest to buy for a while. These qualitative conclusions are summarized as "Stylized Facts" at the end of Section 3.

The reader might object that an equilibrium configuration as described in Section 2 or 3 is unlikely to be observed because, in practice, the market participants compute their optimal strategy for an optimal control problem, and not for a game. We address this issue in Section 4, in the case where each participant has his own risk aversion (i.e. a case of *heterogenous preferences*, in game theory terminology). We first discuss the existence and the uniqueness of a MFG Nash equilibrium in this more general framework. Then we model—in a slightly more realistic way—the day-after-day behavior of the market participants: we explain that, as market participants observe a (possibly) noisy measurement of the daily net trading speed of all investors, they can try to derive from their past observations an approximation of the permanent market impact for the next day and compute their optimal strategy accordingly. We show that, doing so, *they end up playing a Nash MFG equilibrium.* Let us underline that, in our model, the market participants do not have access to the distribution of the trading positions of the other participants; they do not necessarily have the same estimate of the permanent market impact; they are not even aware that they are "playing a game". Nevertheless, the configuration after stabilization is an MFG equilibrium.

Since we had to develop our own MFG tools to address this case (the mean field involving the *controls of the agents*, not in their state), the last part of this paper (Section 5) addresses this kind of MFG systems in a very generic way. The reader will hence find in this Section tools to handle such problems, with generic cost functions and not only with the ones of the usual optimal liquidation problem. Note that we only address here the question of well-posedness of the game with infinitely many agents: the application to games with a finite number of players, as well as the learning procedures will be developed in future works.

Acknowledgement: Authors thank Marc Abeille for his careful reading of Section 3.2.1. The first author was partially supported by the ANR (Agence Nationale de la Recherche) projects ANR-14-ACHN-0030-01 and ANR-16-CE40-0015-01.

2 Trading Optimally Within The Crowd

2.1 Modeling a Mean Field of Optimal Liquidations

A continuum of investors indexed by a decide to buy or sell a given tradable instrument. The decision is a signed quantity Q_0^a to buy (in such a case Q_0^a is negative: the investor has a negative inventory at the initial time $t = 0$) or to sell (when Q_0^a is positive). All investors have to buy or sell before a given terminal time T, each of them will nevertheless potentially trade faster or slower since each of them will be submitted to different risk aversion parameters ϕ^a and A^a. The distribution of the risk aversion parameters is independent to anything else.

Each investor will control its trading speed ν_t^a through time, in order to fullfil its goal. The price S_t of the tradable instrument is submitted to two kinds of moves: an exogenous innovation supported by a standard Wiener process W_t (with its natural probability space and the associated filtration $\mathcal{F}_t$), and the *permanent market impact* generated linearly from the buying or selling pressure $\alpha\mu_t$ where μ_t is the net sum of the trading speed of all investors (like in [Cartea and Jaimungal, 2015], but in our case μ_t is endogenous where it is exogenous in their case) and $\alpha > 0$ is a fixed parameter.

$$dS_t = \alpha\mu_t\, dt + \sigma\, dW_t. \tag{1}$$

The state of each investor is described by two variables: its inventory Q_t^a and its wealth X_t^a

(starting with $X_0^a = 0$ for all investors). The evolution of Q^a reads

$$dQ_t^a = \nu_t^a \, dt, \tag{2}$$

since for a seller, $Q_0^a > 0$ (the associated control ν^a will be mostly negative) and the wealth suffers from linear trading costs (or *temporary*, or *immediate market impact*, parametrized by κ):

$$dX_t^a = -\nu_t^a (S_t + \kappa \cdot \nu_t^a) \, dt. \tag{3}$$

Meaning the wealth of a seller will be positive (and the faster you sell –i.e. ν^a is largely negative–, the smaller the sell price).

The cost function of investor a is similar to the ones used in [Cartea et al., 2015]: it is made of the wealth at T, plus the value of the inventory penalized by a terminal market impact, and minus a running cost quadratic in the inventory:

$$V_t^a := \sup_{\nu} \mathbb{E} \left(X_T^a + Q_T^a (S_T - A^a \cdot Q_T^a) - \phi^a \int_t^T (Q_s^a)^2 \, ds \right). \tag{4}$$

We use this cost function by purpose: a lot of efforts have been made around this utility function by Cartea, Jaimungal and their different co-authors to show it can emulate (provided some changes) most of costs functions provided by brokers to dealing desks of asset managers. This specific one emulates an *Implementation Shortfall* algorithms, while it can be changed to emulate a *Percentage of Volume* or a *Volume Weighted Average Price* (see [Lehalle et al., 2013, Capter 3] for a list of common trading algorithms). Such changes in the cost function would nevertheless impact the paper and demand some complementary work.

2.2 The Mean Field Game system of Controls

The Hamilton-Jacobi-Bellman associated to (4) is

$$0 = \partial_t V^a - \phi^a \, q^2 + \frac{1}{2}\sigma^2 \partial_S^2 V^a + \alpha\mu \partial_S V^a + \sup_{\nu} \left\{ \nu \partial_q V^a - \nu(s + \kappa \, \nu) \partial_X V^a \right\} \tag{5}$$

with terminal condition

$$V^a(T, x, s, q; \mu) = x + q(s - A^a q).$$

Following the Cartea and Jaimungal's approach, we will use the following ersatz:

$$V^a = x + qs + v^a(t, q; \mu). \tag{6}$$

Thus the HJB on v is

$$-\alpha\mu \, q = \partial_t v^a - \phi^a \, q^2 + \sup_{\nu} \left\{ \nu \partial_q v^a - \kappa \, \nu^2 \right\}$$

with terminal condition

$$v^a(T, q; \mu) = -A^a q^2$$

and the associated optimal feedback is

$$\nu^a(t, q) = \frac{\partial_q v^a(t, q)}{2\kappa}. \tag{7}$$

Defining the mean field. The expression of the optimal control (i.e. trading speed) of each investor shows that the important parameter for each investor is its current inventory Q^a_t. The mean field of this framework is hence the distribution $m(t, dq, da)$ of the inventories of investors and of their preferences. Its initial distribution is fully specified by the initial targets of investors and the distribution of the (ϕ^a, A^a).

It is then straightforward to write the net trading flow μ at any time t

$$\mu_t = \int_{(q,a)} \nu^a_t(q)\, m(t, dq, da) = \int_{q,a} \frac{\partial_q v^a(t, q)}{2\kappa}\, m(t, dq, da). \tag{8}$$

Note that implicitly v^a is a function of μ, meaning we will have a fixed point problem to solve in μ.

We now write the evolution of the density $m(t, dq, da)$ of (Q^a_t). By the dynamics (2) of Q^a_t, we have

$$\partial_t m + \partial_q \left(m \frac{\partial_q v^a}{2\kappa} \right) = 0$$

with initial condition $m_0 = m_0(dq, da)$ (recall that the preference (ϕ^a, A^a) of an agent a is fixed during the period).

The full system. Collecting the above equations we find our twofolds mean field game system made of the backward PDE on v coupled with the forward transport equation of m:

$$\left\{ \begin{array}{rcl} -\alpha q\, \mu_t &=& \partial_t v^a - \phi^a\, q^2 + \dfrac{(\partial_q v^a)^2}{4\kappa} \\ 0 &=& \partial_t m + \partial_q \left(m \frac{\partial_q v^a}{2\kappa} \right) \\ \mu_t &=& \displaystyle\int_{(q,a)} \frac{\partial_q v^a(t, q)}{2\kappa}\, m(t, dq, da) \end{array} \right. \tag{9}$$

The system is complemented with the initial condition (for m) and terminal condition (for v):

$$m(0, dq, da) = m_0(dq, da), \qquad v^a(T, q; \mu) = -A^a q^2.$$

The above system is interpreted as a Nash equilibrium configuration in a game with infinitely many market participants: a (small) market participant, anticipating the net trading flow (μ_t), computes his optimal strategy by solving an optimal control problem which, after simplification, leads to the equation for v^a coupled with the terminal condition. When all market participants trade optimally, the distribution m of the inventories and preferences evolves according to the second equation, complemented with the initial for m. Then one derives the net trading flow (μ_t) as a function of m and v^a through the third equation.

3 Trade crowding with identical preferences

In this section, we suppose that all agents have identical preferences: $\phi^a = \phi$ and $A^a = A$ for all a. The main advantage of this assumption is that it leads to explicit solutions.

3.1 The system in the case of identical preferences

To simplify notation, we omit the parameter a in all expressions. We aim at solving (9) (in which a is omitted). It is convenient to set $E(t) = \mathbb{E}[Q_t] = \int_q qm(t, dq)$. Note that

$$E'(t) = \int_q q \partial_t m(t, dq),$$

so that, using the equation for m and an integration by parts:

$$E'(t) = -\int_q q\partial_q \left(m(t,q) \frac{\partial_q v(t,q)}{2\kappa} \right) dq = \int_q \frac{\partial_q v(t,q)}{2\kappa} m(t,dq). \tag{10}$$

3.2 Quadratic Value Functions

When $v(t,q)$ can be expressed as a quadratic function of q:

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

then the backward part of the master equation can be split in three parts

$$\begin{array}{rcl} 0 & = & -2\kappa h_2'(t) - 4\kappa\phi + (h_2(t))^2, \\ 2\kappa\alpha\mu(t) & = & -2\kappa h_1'(t) + h_1(t)h_2(t), \\ -(h_1(t))^2 & = & 4\kappa h_0'(t), \end{array} \tag{11}$$

One also has to add the terminal condition: as $V_T = x + q(s - Aq)$, $v(T,q) = -Aq^2$. This implies that

$$h_0(T) = h_1(T) = 0, \;\; h_2(T) = 2A. \tag{12}$$

Dynamics of the mean field. Recalling (8), we have

$$\mu(t) = \int_q \frac{\partial_q v(t,q)}{2\kappa} \, dm(q) = \int_q \frac{h_1(t) - qh_2(t)}{2\kappa} \, dm(q) = \frac{h_1(t)}{2\kappa} - \frac{h_2(t)}{2\kappa} E(t). \tag{13}$$

Moreover, by (10), we also have

$$E'(t) = \int_q m(t,q) \left(\frac{h_1(t)}{2\kappa} - \frac{h_2(t)}{2\kappa} q \right) dq = \frac{h_1(t)}{2\kappa} - \frac{h_2(t)}{2\kappa} E(t).$$

So we can supplement (11) with

$$2\kappa E'(t) = h_1(t) - E(t) \cdot h_2(t). \tag{14}$$

Summary of the system. We now collect all the equations. Recalling (13), we find:

$$\begin{cases} 4\kappa\phi = -2\kappa h_2'(t) + (h_2(t))^2, & \text{(15a)} \\ \alpha h_2(t)E(t) = 2\kappa h_1'(t) + h_1(t)\left(\alpha - h_2(t)\right), & \text{(15b)} \\ -(h_1(t))^2 = 4\kappa h_0'(t), & \text{(15c)} \\ 2\kappa E'(t) = h_1(t) - h_2(t)E(t). & \text{(15d)} \end{cases}$$

with the boundary conditions

$$h_0(T) = h_1(T) = 0, \;\; h_2(T) = 2A, \;\; E(0) = E_0,$$

where $E_0 = \int_q q m_0(q) dq$ is the net initial inventory of market participants (i.e. the expectation of the initial density m).

3.2.1 Reduction to a single equation

From now on we consider h_2 as given and derive an equation satisfied by E. By (15d), we have

$$h_1 = 2\kappa E' + h_2 E,$$

so that

$$h_1' = 2\kappa E'' + E h_2' + h_2 E'. \tag{16}$$

Plugging these expressions into (15b), we obtain

$$\begin{aligned} 0 &= h_1' + h_1 \frac{\alpha - h_2}{2\kappa} - \frac{\alpha h_2}{2\kappa} E, \\ &= 2\kappa E'' + E h_2' + h_2 E' + (2\kappa E' + h_2 E) \frac{\alpha - h_2}{2\kappa} - \alpha \frac{h_2}{2\kappa} E \\ &= 2\kappa E'' + \alpha E' + 2E \left(\frac{1}{2} h_2' - \frac{(h_2)^2}{4\kappa} \right). \end{aligned}$$

Recalling (15a) we find

$$0 = 2\kappa E'' + \alpha E' - 2\phi E.$$

The boundary conditions are $E(0) = E_0$, $h_2(T) = 2A$, $h_1(T) = 0$, where the last expression can be rewritten by taking (15d) into account. To summarize, the equation satisfied by E is:

$$\begin{cases} 0 = 2\kappa E''(t) + \alpha E'(t) - 2\phi E(t) & \text{for } t \in (0, T), \\ E(0) = E_0, \qquad \kappa E'(T) + A E(T) = 0. \end{cases} \tag{17}$$

3.2.2 Solving (17)

After some easy but tedious computation explained in Appendix A, one finds:

Proposition 3.1 (Closed form for the net inventory dynamics $E(t)$)**.** *For any $\alpha \in \mathbb{R}$, the problem* (17) *has a unique solution E, given by*

$$E(t) = E_0 a \left(\exp\{r_+ t\} - \exp\{r_- t\} \right) + E_0 \exp\{r_- t\}$$

where a is given by

$$a = \frac{(\alpha/4 + \kappa\theta - A) \exp\{-\theta T\}}{-\frac{\alpha}{2} \mathrm{sh}\{\theta T\} + 2\kappa\theta \mathrm{ch}\{\theta T\} + 2A \mathrm{sh}\{\theta T\}},$$

the denominator being positive and the constants $r_\alpha^\pm$ and θ being given by

$$r_\pm := -\frac{\alpha}{4\kappa} \pm \theta, \qquad \theta := \frac{1}{\kappa} \sqrt{\kappa\phi + \frac{\alpha^2}{16}}.$$

Moreover,

$$h_2(t) = 2\sqrt{\kappa\phi} \frac{1 + c_2 e^{rt}}{1 - c_2 e^{rt}}, \tag{18}$$

where $r = 2\sqrt{\phi/\kappa}$ and

$$c_2 = -\frac{1 - A/\sqrt{\kappa\phi}}{1 + A/\sqrt{\kappa\phi}} \cdot e^{-rT}.$$

Remark 3.2. The last needed component to obtain the optimal control using (7) is $h_1(t)$. Thanks to (16), it can be easily written from E, E' and h_2 (note h_2 is mostly negative for our sell order):

$$h_1(t) = 2\kappa \cdot E'(t) + h_2(t) \cdot E(t).$$

This gives an explicit formula for the optimal control for any value of the parameters: κ (the instantaneous market impact), ϕ (the risk aversion), α (the permanent market impact), A (the terminal penalization), E_0 (the initial net position of all participants), and T (the duration of the "game").

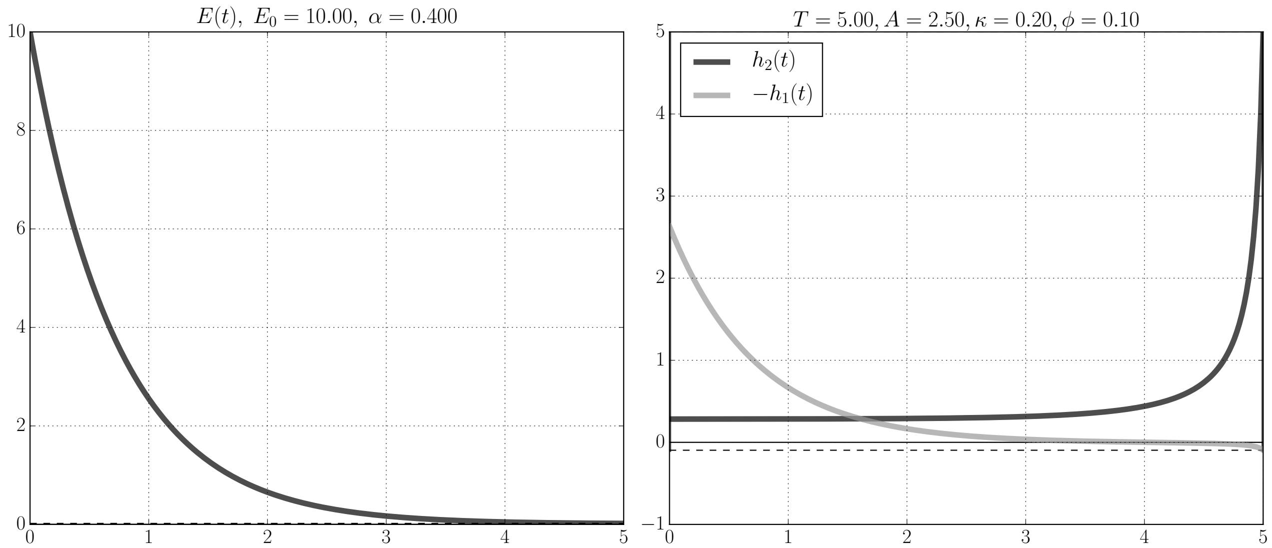

Figure 1: Dynamics of E (left) and $-h_1$ and h_2 (right) for a standard set of parameters: $\alpha = 0.4$, $\kappa = 0.2$, $\phi = 0.1$, $A = 2.5$, $T = 5$, $E_0 = 10$.

Typical Dynamics. Figure 1 shows the joint dynamics of E (left panel), $-h_1$ and h_2 (right panel) for typical values of the parameters: $\alpha = 0.4$, $\kappa = 0.2$, $\phi = 0.1$, $A = 2.5$, $T = 5$, $E_0 = 10$. As expected, $E(t)$ goes very close to 0 at $t = T$; in our reference case $E(T) = 0.02$. Looking carefully at Proposition 3.1, it can be seen the main component driving $E(T)$ to zero is $\exp r_-^\alpha T$, where $r_-^\alpha = -\alpha/4 - (\sqrt{\kappa\phi + \alpha^2/16})/\kappa$, hence:

- The best way to obtain a terminal net inventory of zero is to have a large α, or a large ϕ, or a small κ. Surprisingly, having a large A does not help that much. It mainly urges the trading very close to $t = T$ when the other parameters decrease E earlier.
- h_2 increases slowly to $2A$, while h_1 goes from a negative value to a slightly positive one.

To understand the respective roles of h_2 and h_1, one should keep in mind the optimal control is $(h_1(t) - qh_2(t))/(2\kappa)$. Having a negative h_1 increases the trading speed of a seller. That's why we draw $-h_1$ instead of h_1 on all the figures.

Since participants influences themselves via $\alpha\mu_t$ the permanent market impact coefficient times the sum of their controls (that are functions of the mean field), one can consider the lower α, the more "disconnected" players from the influence of the mean field.

Stylized Fact 1 (Influence of the mean field varies with α)**.** *Figure 2 compares the components of the optimal strategies for two values of α (the strength of the influence of the players one on each others): when players are less connected: $h_2(t)$ does not change and $h_1(t)$ is smaller, except at the end of the trading.*

Keep in mind the optimal control $\nu_t(q)$ is proportional to $h_1(t) - q\,h_2(t)$. This means when t is close to zero (i.e. start of the trading), q is close to Q_0, and hence $q\,h_2(t)$ is large compared to $h_1(t)$.

Stylized Fact 2 (Driving E to zero)**.** *A large permanent impact α, a large risk aversion ϕ and a small temporary impact κ are the main components driving the net inventory of participants E to zero.*

Another way to understand the optimal control ν is to look at its formulation not only as a function of h_1 and h_2, but at a function of E, E' and h_2:

$$\nu(t,q) = \frac{\partial_q v}{2\kappa} = \frac{1}{2\kappa}(h_1(t) - q\cdot h_2(t)) = E'(t) + \frac{E(t) - q}{2\kappa}\,h_2(t). \tag{19}$$

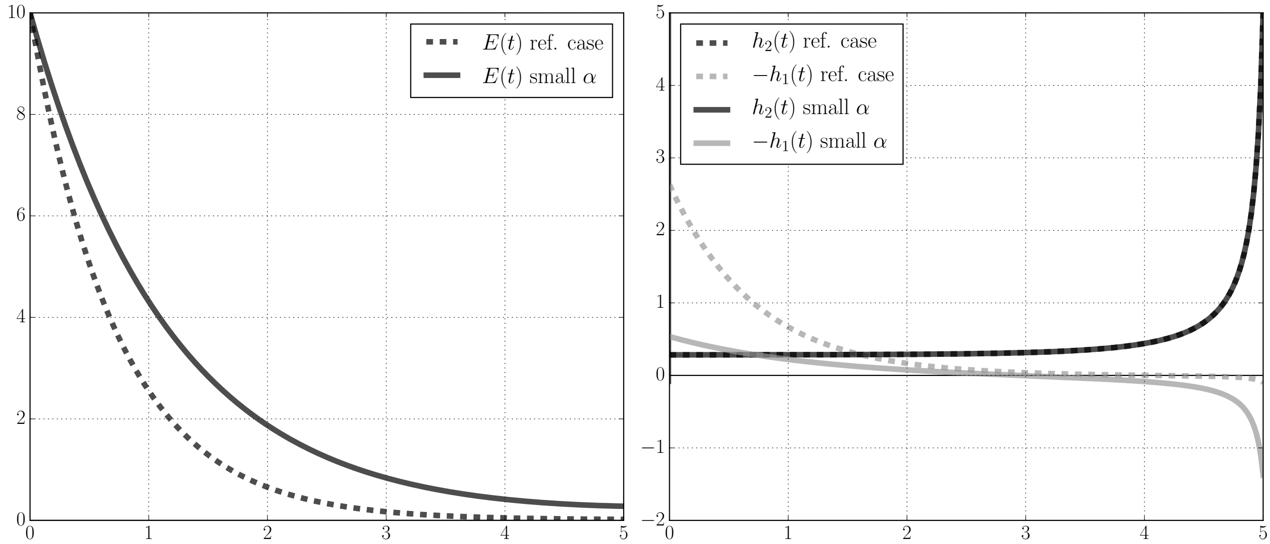

Figure 2: Comparison of the dynamics of E (left) and $-h_1$ and h_2 (right) between the "reference" parameters of Figure 1 and smaller α (i.e. $\alpha = 0.1$ instead of 0.4) such that $|h_1(0)|$ is smaller.

Keeping in mind we formulated the problem from the viewpoint of a seller: $Q_0 > 0$ is the number of shares to be sold from 0 to T, and $E_0 > 0$ means the net initial position of all participants is dominated by sellers. Since $E(t)$ decreases with time, E' is rather negative.

The upper expression of the optimal control of our seller means the larger the remaining shares to sell $q(t)$, the faster to trade, proportionally to $h_2(t)$. The influence of A is clear: $h_2(T) = 2A$ says the larger the terminal penalization, the faster to trade when T is close, for a given remaining number of shares to sell.

Expression (19) for the optimal control reads:

Stylized Fact 3 (Influence of $E(t)$ and $E'(t)$ on the optimal control)**.** *The optimal control is made of two parts: one ($-qh_2/(2\kappa)$) is proportional to the remaining quantity and independent of others' behavior; the other ($h_1 = E' + E/(2\kappa)$) increases with the net inventory of other market participants and follows their trading flow. Hence, in this framework:*

(i) it is optimal to "follow the crowd" (because of the E' term)

(ii) but not too fast (since E and E' often have an opposite sign); especially when t is close to T (because of the h_2 term in factor of E).

This pattern can be seen as a *fire sales* pattern: the trader should follow participants while they trade in the same direction. This also means when the trader's inventory is opposite to market participants' net inventory, he can afford to slow down (because the price will be better for him soon).

Stylized Fact 4 (Optimal trading speed with a very low inventory)**.** *When a participant is close to a zero inventory (i.e. q is close to zero) or for participant with low terminal constraints and low risk aversion, it can be rewarding to "follow the crowd". The dominant term is then h_1: a sign change of h_1 implies a change of trading direction for a participant with a low inventory. Nevertheless once a participant followed h_1, his (no more neglectable) inventory multiplied by h_2 drives his trading speed.*

Readers can have a look at the right panel Figure 2 to (solid grey line) observe a sign change of h_1 (around $t = 4$).

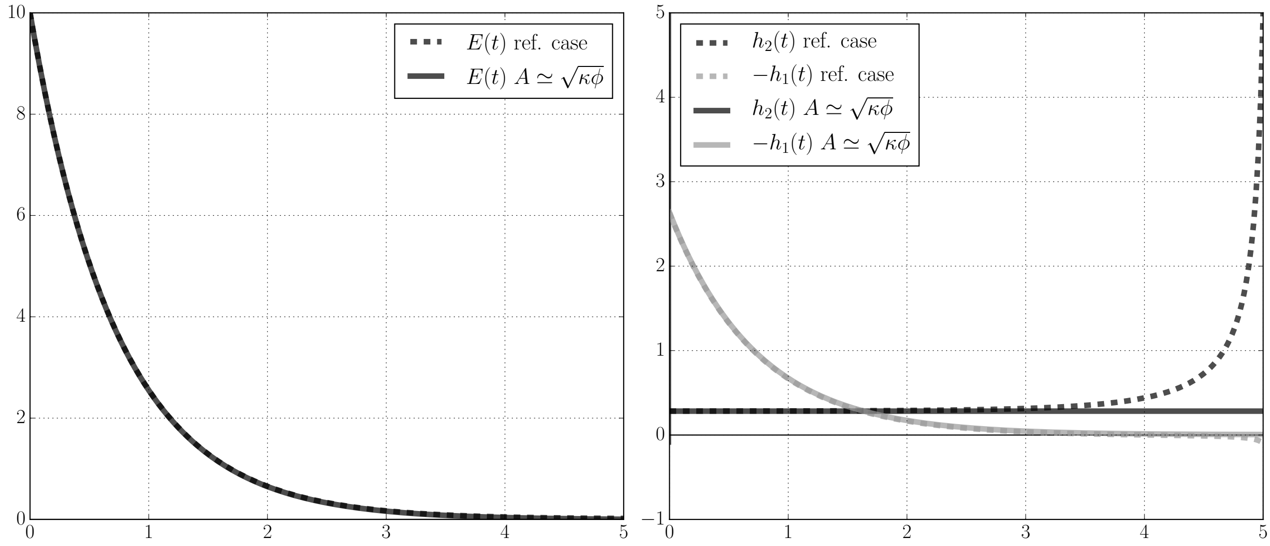

Figure 3: Comparison of the dynamics of E (left) and $-h_1$ and h_2 (right) between the "reference" parameters of Figure 1 and when $\sqrt{\kappa\phi} \simeq A$: in such a case h_2 is almost constant but E and h_1 are almost unchanged.

A specific case where h_2 is almost constant. When A is very close to $\sqrt{\kappa\phi}$, expression (18) for h_2 shows c_2 will be very close to zero, and hence $h_2(t) \simeq 2\sqrt{\kappa\phi} \simeq 2A$ for any t.

Figure 3 is an illustration of such a case: $A = \sqrt{\kappa\phi} = 0.141$. Since E is not affected too much by this change in A and remains close to zero when t is large enough, the change of h_2 slope close to T cannot affect significantly h_1.

Stylized Fact 5 (Constant h_2)**.** *When $A = \sqrt{\kappa\phi}$, h_2 is constant (no more a function of t) equals to $2A$. Hence the multiplier of q (the remaining quantity) is a constant:*

$$\nu(t,q) = E'(t) + \frac{E(t)-q}{\kappa}\,A.$$

Specific configurations of E. When the considered trader has to sell while the net initial position of all the participants is to buy (i.e. $Q_0 > 0$ and $E_0 < 0$), the multiplier h_2 of the remaining quantity q stays the same, but the constant term h_1 is turned into $-h_1$:

Stylized Fact 6 (Alone against the crowd)**.** *When the trader position does not have the same direction than the net inventory of all participants E: he has to trade slower, independently from his remaining inventory q.*

Moreover, the formulation of Stylized Fact 5 shows it is possible to change the monotony of $E(t)$ so that after a given t is no more decreases:

$$E(t)' = E_0 a \underbrace{\left(r_+ \exp\{r_+ t\} - r_- \exp\{r_- t\}\right)}_{\text{positive}} + \underbrace{E_0 r_- \exp\{r_- t\}}_{\text{negative}}.$$

For well chosen configuration of parameters the first term can be larger than the second term, for any t greater than a critical t^m such that $E'(t^m) = 0$. For the configuration of Figure 4, with $\alpha = 0.01$, $\kappa = 1.5$, $\phi = 0.03$, $A = 2.5$, $T = 5$ and $E_0 = 10$, we have $t^m \simeq 3.82$.

Going back to the meaning of this mean field game framework: it models dealing desks of asset managers receiving instructions from their portfolio managers to buy or sell large amounts of shares at the start of the day (or of the week). When $t^m < T$, it means that while the sum of initial instructions where to buy (respectively sell) this day, the "mean field" of dealing desks changed its mind: they did not strictly followed instructions. They bought (resp. sold) more

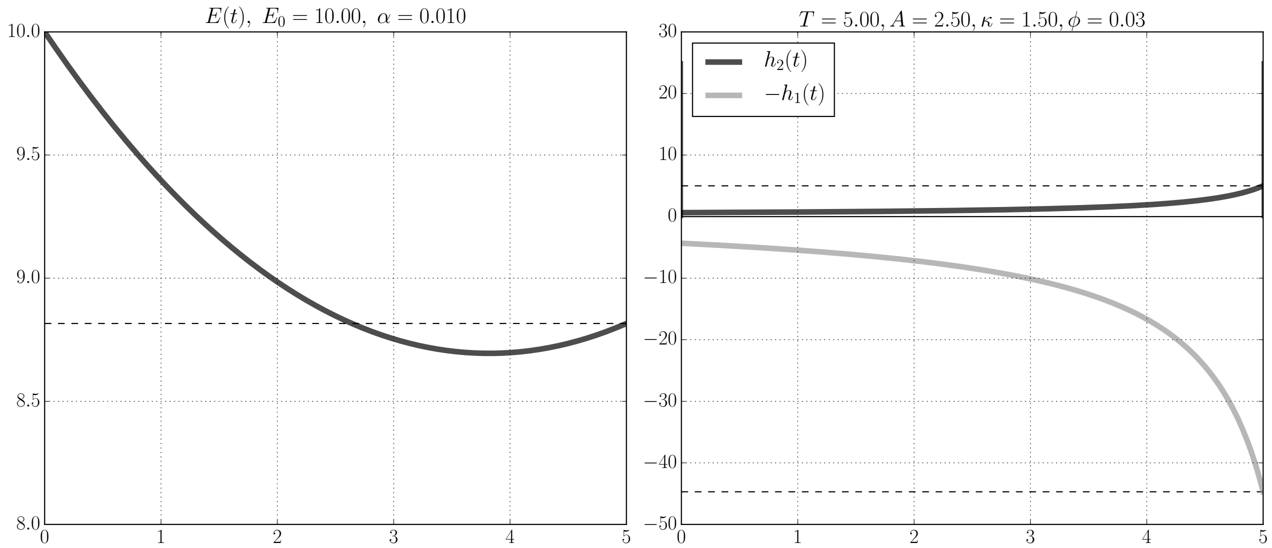

Figure 4: A specific case for which E is not monotonous: $\alpha = 0.01$, $\kappa = 1.5$, $\phi = 0.03$, $A = 2.5$, $T = 5$ and $E_0 = 10$.

than asked, and are now starting to sell back (resp. buy back) this temporary inventory to make profits by their own. Regulators could be interested by market parameters (α, κ, ϕ) allowing such configurations to appear.

Figure 5 shows configurations for which t^m exists: small values of ϕ and α and large value of κ are in favor of a small t^m. This means when the risk aversion and the permanent market impact coefficient are small while the temporary market impact is large, the slope of the net inventory of participants can have sign change.

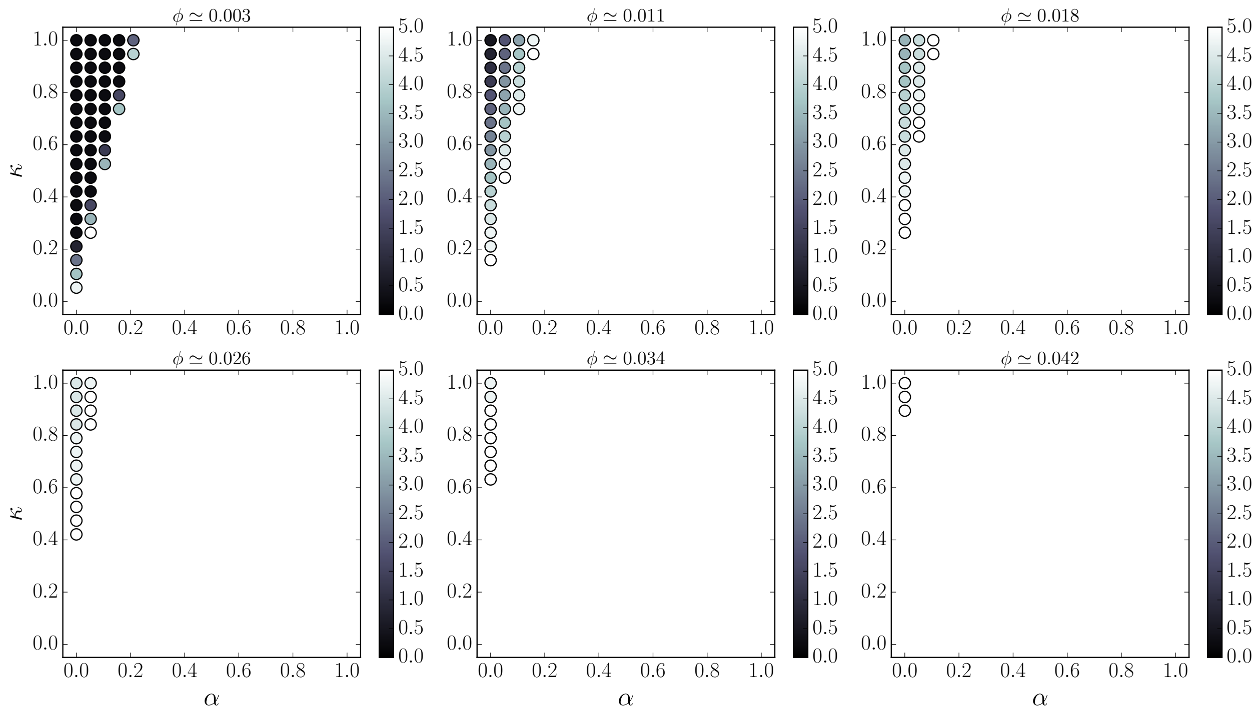

Figure 5: Numerical explorations of t^m for different values of ϕ (very small ϕ at the top left to small ϕ at the bottom right) on the $\alpha \times \kappa$ plane, when $T = 5$ and $A = 2.5$. The color circles codes the value of t^m: small values (dark color) when E changes its slope very early; large values (in light colors) when E changes its slope close to T.

4 Trade crowding with heterogeneous preferences

We now come back to our original model (9) in which each agent may be adverse to the risk a different way. In other words, the constants A^a and ϕ^a may depend on a but are fixed for 0 to T (i.e. for the day or for the week). For simplicity, we will mostly work under the condition that $A^a = \sqrt{\phi^a \kappa}$, which allows to simplify the formulae.

4.1 Existence and uniqueness for the equilibrium model

As in the case of identical agents, we look for a solution to system (9) in which the map v^a is quadratic:

$$v^a(t,q) = h_0^a(t) + q\, h_1^a(t) - q^2\, \frac{h_2^a(t)}{2},$$

and we find the relations:

$$\begin{array}{rcl} 0 &=& -2\kappa (h_2^a)' - 4\kappa\phi^a + (h_2^a)^2, \\ 2\kappa\alpha\mu(t) &=& -2\kappa (h_1^a)' + h_1^a h_2^a, \\ -(h_1^a)^2 &=& 4\kappa (h_0^a)', \end{array} \tag{20}$$

with terminal conditions

$$h_0^a(T) = h_1^a(T) = 0,\ h_2^a(T) = 2A^a. \tag{21}$$

Let $m = m(t, dq, da)$ be the repartition at time t of the wealth q and of the parameter a. As before, the net trading flow μ at any time t can be expressed as

$$\mu(t) = \int_{a,q} \frac{\partial_q v^a(t,q)}{2\kappa}\, m(t,da,dq) = \int_{a,q} \frac{1}{2\kappa}\left(h_1^a(t) - h_2^a(t) q\right)\, m(t,da,dq).$$

The measure m solves the continuity equation

$$\partial_t m + \mathrm{div}_q \left(m \frac{\partial_q v^a(t,q)}{2\kappa} \right) = 0$$

with initial condition $m_0 = m_0(da, dq)$. For later use, we set

$$\bar{m}_0(da) = \int_q m_0(da, dq).$$

As the agents do not change their parameter a over the time, we always have

$$\int_q m(t, da, dq) = \bar{m}_0(da),$$

so that we can disintegrate m into

$$m(t, da, dq) = m^a(t, dq)\bar{m}_0(da),$$

where $m^a(t, dq)$ is a probability measure in q for $\bar{m}_0-$ almost any a. Let us set

$$E^a(t) = \int_q q m^a(t, dq).$$

Then by similar argument as in the case of identical agents one has

$$(E^a)'(t) = \frac{h_1^a(t)}{2\kappa} - \frac{h_2^a(t)}{2\kappa} E^a(t). \tag{22}$$

With the notation E^a, we can rewrite μ as

$$\mu(t) = \frac{1}{2\kappa}\left(\int_a h_1^a(t)\bar{m}_0(da) - \int_a h_2^a(t)E^a(t)\bar{m}_0(da)\right). \tag{23}$$

From now on we assume *for simplicity of notation* that

$$A^a = \sqrt{\phi^a\kappa} \qquad \forall a \tag{24}$$

so that h_2^a is constant in time with $h_2^a(t) = 2A^a$. We set

$$\theta^a = \frac{h_2^a}{2\kappa}.$$

We solve the equation satisfied by h_1^a: as $h_1^a(T) = 0$, we get

$$h_1^a(t) = \alpha \int_t^T ds \exp\{\theta^a(t-s)\}\mu(s).$$

In the same way, we solve the equation for E^a in (22) by taking into account the fact that $E^a(0) = E_0^a := \int_q q\bar{m}_0(a,q)dq$ and the computation for h_1^a: we find

$$E^a(t) = \exp\{-\theta^a t\}E_0^a + \frac{\alpha}{2\kappa}\int_0^t d\tau \int_\tau^T ds \exp\{\theta^a(2\tau - t - s)\}\mu(s).$$

This can be read as the "mean field" associated to players with risk aversion parameters (ϕ^a, A^a).

Putting these relations together and summing over all risk aversions (i.e. integrating over $\bar{m}_0(da)$) we obtain by (23) that μ satisfies:

$$\begin{aligned}\mu(t) \;=\; & -\int_a \theta^a \exp\{-\theta^a t\}E_0^a\bar{m}_0(da) + \frac{\alpha}{2\kappa}\int_t^T ds\mu(s)\int_a \exp\{\theta^a(t-s)\}\bar{m}_0(da)\\ & -\frac{\alpha}{2\kappa}\int_0^t d\tau\int_\tau^T ds\mu(s)\int_a \theta^a\exp\{\theta^a(2\tau - t - s)\}\bar{m}_0(da).\end{aligned} \tag{25}$$

Proposition 4.1. *Assume that (24) holds and that the (A^a) are bounded. Then there exists $\alpha_0 > 0$ such that, for $|\alpha| \le \alpha_0$, there exists a unique solution to the fixed point relation (25).*

As a consequence, the MFG system has at least one solution obtained by plugging the fixed point μ into the relations (20).

Proof. One just uses Banach fixed point Theorem on $\Phi_\alpha : C([0,T]) \to C([0,T])$ which associates to any $\mu \in C([0,T])$ the map

$$\begin{aligned}\Phi_\alpha(\mu)(t) \;:=\; & -\int_a \theta^a \exp\{\theta^a t\}E_0^a\bar{m}_0(da) + \frac{\alpha}{2\kappa}\int_t^T ds\mu(s)\int_a \exp\{\theta^a(t-s)\}\bar{m}_0(da)\\ & -\frac{\alpha}{2\kappa}\int_0^t d\tau\int_\tau^T ds\mu(s)\int_a \theta^a\exp\{\theta^a(2\tau - t - s)\}\bar{m}_0(da).\end{aligned}$$

It is clear that Φ_α is a contraction for $|\alpha|$ small enough: indeed, given $\mu, \mu' \in C([0,T])$, we have

$$\begin{aligned}|\Phi_\alpha(\mu)(t) - \Phi_\alpha(\mu')(t)| \;\le\; & \frac{|\alpha|}{2\kappa}\int_t^T ds\, |\mu(s) - \mu'(s)| \int_a \exp\{\theta^a(t-s)\}\bar{m}_0(da)\\ & + \frac{|\alpha|}{2\kappa}\int_0^t d\tau \int_\tau^T ds\, |\mu(s) - \mu'(s)| \int_a \theta^a \exp\{\theta^a(2\tau - t - s)\}\bar{m}_0(da)\\ \le\; & C|\alpha|\|\mu - \mu'\|_\infty,\end{aligned}$$

for some constant C independent of α, because the (θ^a) are bounded. ∎

4.2 Learning other participants' net flows to (potentially) converge towards the MFG equilibrium

The solution of the MFG system describing an equilibrium configuration, 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. For this we assume the game repeats the same $[0, T]$ intervals an infinite number of *rounds*. The reader can think about 0 to T being a day (or a week), and hence each round will last a day (or a week). Round after round, market participants try to "learn" (i.e. to build an estimate of) the trading speed μ_t of equation (1).

It is close to what dealing desks of asset managers[1] are doing on financial markets: they try to estimate the buying or selling pressure exerted by other market participants to adjust their own behaviour. That for, investment banks provide to their clients (asset managers) information about the recent state of the flows on exchanges, to help them to adjust their trading behaviours. For instance, JP Morgan's corporate and investment bank has a twenty pages recurrent publication titled "*Flows and Liquidity*", with charts and tables describing different metrics like *money flows*, turnovers, and other similar metrics by asset class (equity, bonds, options) and countries. Almost each investment bank has such a publication for its clients; Barclay's one is titled "*TIC Monthly Flows*". As an example its August 2016 issue (9 pages) subtitle was "*Private foreign investors remain buyers*".

Combining these expected flows with market impact models identified on their own flows (see [Bacry et al., 2015], [Brokmann et al., 2014] or [Waelbroeck and Gomes, 2013] for academics papers on the topic), traders of dealing desks tune their optimal liquidation (or trading) schemes for the coming days. It is hence interesting to note this practice is very close to the framework we propose in this Subsection.

We also assume that, at the beginning of round n, each agent (generically indexed by a) has an estimate on $\mu^{a,n}$ on the crowd impact. Then he solves his corresponding optimal control problem:

$$\begin{array}{rcl} 0 & = & -2\kappa \dfrac{(h_2^{a,n})'}{2} - 4\kappa\phi^a + (h_2^{a,n})^2, \\ 2\kappa\alpha\mu^{a,n}(t) & = & -2\kappa (h_1^{a,n})' + h_1^{a,n} h_2^{a,n}, \\ -(h_1^{a,n})^2 & = & 4\kappa (h_0^{a,n})', \end{array} \tag{26}$$

with terminal conditions

$$h_0^{a,n}(T) = h_1^{a,n}(T) = 0,\ h_2^{a,n}(T) = 2A^a.$$

For simplicity we assume once more that

$$A^a = \sqrt{\phi^a \kappa} \qquad \forall a$$

so that $h_2^{a,n}$ is constant in time and independent of n: $h^a = 2A^a$. We set as before

$$\theta^a = \frac{h_2^a}{2\kappa}.$$

We find as previously

$$h_1^{a,n}(t) = \alpha \int_t^T ds \exp\{\theta^a (t-s)\} \mu^{a,n}(s).$$

and

$$E^{a,n}(t) = \exp\{-\theta^a t\} E_0^a + \frac{\alpha}{2\kappa} \int_0^t d\tau \int_\tau^T ds \exp\{\theta^a (2\tau - t - s)\} \mu^{a,n}(s).$$

[1]Large asset managers, like Blackrock, Amundi, Fidelity or Allianz, delegate the implementation of their investment decisions to a dedicated (internal) team: their dealing desk. This team is in charge of trading to drive the real portfolios to their targets. They are implementing on a day to day basis what this paper is modelling.

Then the net sum m^{n+1} of the trading speed of all investors at stage n is given by (23) where E^a and h_1^a are replaced by $E^{a,n}$ and $h_1^{a,n}$:

$$
\begin{aligned}
m^{n+1}(t) \;&:=\; \frac{1}{2\kappa}\left(\int_a h_1^{a,n}(t)\bar{m}_0(da) - \int_a h_2^a E^{a,n}(t)\bar{m}_0(da)\right) \\
&= -\int_a \theta^a \exp\{-\theta^a t\} E_0^a \bar{m}_0(da) + \frac{\alpha}{2\kappa}\int_t^T ds \int_a \mu^{a,n}(s)\exp\{\theta^a(t-s)\}\bar{m}_0(da) \\
&\quad -\frac{\alpha}{2\kappa}\int_0^t d\tau \int_\tau^T ds \int_a \mu^{a,n}(s)\theta^a \exp\{\theta^a(2\tau - t - s)\}\bar{m}_0(da).
\end{aligned}
\tag{27}
$$

We now model how the agents evaluate the crowd impact. We assume that each agent has his own way of evaluating the value of $m^{n+1}(t)$ and of incorporating this new value into his estimate of the crowd impact. Namely, we suppose a relation of the form

$$
\mu^{a,n+1}(t) := (1-\pi^{a,n+1})\mu^{a,n}(t) + \pi^{a,n+1}(m^{n+1}(t) + \varepsilon^{a,n+1}(t)),
$$

where $\pi^{a,n+1} \in (0,1)$ is the weight chosen by an agent a at round $n+1$ to adjust his new estimate of μ with respect to his previous belief $\mu^{a,n}(t)$ and his new estimate $m^{n+1}(t)+\varepsilon^{a,n+1}(t)$ and $\varepsilon^{a,n+1}$ is a small error term on the measured crowd impact: $\|\varepsilon^{a,n+1}\|_\infty \leq \varepsilon$.

Proposition 4.2. *Under the assumption of Proposition 4.1, let μ be the solution of the MFG crowd trading model. Assume that*

$$
\frac{1}{Cn} \leq \pi^{a,n} \leq \frac{C}{n},
$$

for some constants C. Suppose also that α is small enough: $|\alpha| \leq \alpha_1$ for some small $\alpha_1 > 0$ depending on C. Then

$$
\limsup \sup_a \|\mu^{a,n} - \mu\|_\infty \leq C\varepsilon
$$

for some constant C.

Let us note that, as the solution to the system (26) depends in a continuous way of $\mu^{a,n}$, the optimal trading strategies of the agents are close to the one corresponding to the equilibrium configuration for n sufficiently large.

Proof. In the proof the constant C might differ from line to line, but does not depend on n nor on ε. We have

$$
\begin{aligned}
\sup_a \|\mu^{a,n+1} - \mu\|_\infty \;&\leq\; \sup_a \|(1-\pi^{a,n+1})\mu^{a,n} + \pi^{a,n+1}(m^{n+1} + \varepsilon^{a,n+1}) - \mu\|_\infty \\
&\leq\; \sup_a \left((1-\pi^{a,n+1})\|\mu^{a,n} - \mu\|_\infty + \pi^{a,n+1}\|m^{n+1} - \mu\|_\infty + \pi^{a,n+1}\|\varepsilon^{a,n+1}\|_\infty\right)
\end{aligned}
$$

where, by (27):

$$
\|m^{n+1} - \mu\|_\infty \leq C\frac{\alpha}{2\kappa}\sup_a \|\mu^{a,n} - \mu\|_\infty .
$$

So

$$
\sup_a \|\mu^{a,n+1} - \mu\|_\infty \leq \sup_a \left((1-\pi^{a,n+1}) + C\frac{\alpha}{2\kappa}\pi^{a,n+1}\right) \sup_a \|\mu^{a,n} - \mu\|_\infty + \sup_a \pi^{a,n+1}\varepsilon.
$$

Thus, setting $\beta_n = \sup_a \left((1-\pi^{a,n}) + C\frac{\alpha}{2\kappa}\pi^{a,n}\right)$ and $\delta_n := \sup_a \pi^{a,n}$, we have

$$
\sup_a \|\mu^{a,n} - \mu\|_\infty \leq \sup_a \|\mu^{a,0} - \mu\|_\infty \prod_{k=1}^n \beta_k + \varepsilon \sum_{k=1}^n \delta_k \prod_{l=k+1}^n \beta_l.
$$

By our assumption on $\pi^{a,n}$, we can choose α small enough so that

$$\beta_n \leq 1 - \frac{1}{Cn} \qquad \text{and} \qquad \delta_n \leq \frac{C}{n}.$$

Then, for $1 \leq k < n$, we have

$$\ln\left(\prod_{l=k}^{n} \beta_l\right) \leq \sum_{l=k}^{n} \ln(1 - 1/(Cl)) \leq -(1/C) \sum_{l=k}^{n} \frac{1}{l} \leq -(1/C)\ln((n+1)/k).$$

Hence $\displaystyle\prod_{l=k}^{n} \beta_l \leq C\left(\frac{n}{k}\right)^{-1/C}$, which easily implies that

$$\lim_n \prod_{k=1}^{n} \beta_k = 0 \qquad \text{and} \qquad \sum_{k=1}^{n} \delta_k \prod_{l=k+1}^{n} \beta_l \leq C.$$

The desired result follows. □

5 A General Model for Mean Field Games of Controls

In this section we discuss a general existence result for a Mean Field Game of Control (or, in the terminology of [Gomes et al., 2014], an Extended Mean Field Game). As in our main application above, we aim at describing a system in which the *infinitely many small* agents control their own state and interact through a "mean field of control". By "small", we mean that the individual behavior of each agent has a negligible influence on the whole system. The requirement that they are "infinitely many" corresponds to the fact that their initial configuration is distributed according to an absolutely continuous density on the state space. The "mean field of control" consists in the *joint distribution of the agents and their instantaneous control*: this is in contrats with the standard MFGs, in which the mean field is the distribution of the positions of the agents only.

We denote by $X_t \in \mathbb{R}^d$ the individual state of a generic agent at time t and by α_t his control. The state space is the finite dimensional space $\mathbb{R}^d$, while the controls take their value in a metric space (A, δ_A). In this Section the distribution density of the pair (X_t, α_t) is denoted by μ_t. It is a probability measure on $\mathbb{R}^d \times A$. The first marginal m_t of μ_t is the distribution of the players at time t (hence a probability measure on $\mathbb{R}^d$). In the MFG of control, dynamics and payoffs depend on (μ_t) (and not only on (m_t) as in standard MFGs).

We assume that the dynamics of a small agent is a controlled SDE of the form

$$\begin{cases} dX_t = b(t, X_t, \alpha_t; \mu_t)dt + \sigma(t, X_t)dW_t \\ X_{t_0} = x_0 \end{cases}$$

where α is the control and W is a standard $D-$dimensional Brownian Motion (the Brownian Motions of the agents are independent). Note that, for simplicity of notation, the heterogeneity of the agents (i.e., the parameter a in the previous section) is encoded here in the state variable: it is a variable which is not affected by the dynamics. For this reason it is important to handle a degenerate diffusion term σ. The cost function is given by

$$J(t_0, x_0, \alpha; \mu) = \mathbb{E}\left[\int_{t_0}^{T} L(t, X_t, \alpha_t; \mu_t)\ dt + g(X_T, m_T)\right].$$

It is known that, given μ, the value function $u = u(t_0, x_0; \mu)$ of the agent, defined by

$$u(t_0, x_0) = \inf_{\alpha} J(t_0, x_0, \alpha; \mu),$$

is a viscosity solution of the HJB equation

$$\begin{cases} -\partial_t u(t,x) - \mathrm{tr}(a(t,x)D^2u(t,x)) + H(t,x,Du(t,x);\mu_t) = 0 & \text{in } (0,T)\times\mathbb{R}^d \\ u(T,x) = g(x) \quad \text{in } \mathbb{R}^d \end{cases} \tag{28}$$

where $a = (a_{ij}) = \sigma\sigma^T$ and

$$H(t,x,p;\nu) = \sup_\alpha \left\{-p\cdot b(t,x,\alpha;\nu) - L(t,x,\alpha;\nu)\right\}. \tag{29}$$

Moreover $b(t,x,\alpha^*(t,x);\mu_t) := -D_pH(t,x,Du(t,x);\mu_t)$ is (formally) the optimal drift for the agent at position x and at time t. Thus, the population density $m = m_t(x)$ is expected to evolve according to the Kolmogorov equation

$$\begin{cases} \partial_t m_t(x) - \sum_{i,j}\partial_{ij}(a_{ij}(t,x)m_t(x)) - \mathrm{div}\,(m_t(x)D_pH(t,x,Du(t,x);\mu_t)) = 0 \\ \hfill \text{in } (0,T)\times\mathbb{R}^d \\ m_0(x) = \bar{m}_0(x) \quad \text{in } \mathbb{R}^d \end{cases} \tag{30}$$

Throughout this part we assume that the map $\alpha \to b(t,x,\alpha;\nu)$ is one-to-one with a smooth inverse and we denote by $\alpha^* = \alpha^*(t,x,p;\nu)$ the map which associates to any p the unique control $\alpha^* \in A$ such that

$$b(t,x,\alpha^*;\nu) = -D_pH(t,x,p;\nu). \tag{31}$$

This means that, at each time t and position x, the optimal control of a typically small player is $\alpha^*(t,x,Du(t,x);\mu_t)$. So, in an equilibrium configuration, the measure μ_t has to be the image of the measure m_t by the map $x \to (x,\alpha^*(t,x,Du(t,x);\mu_t))$. This leads to the fixed-point relation:

$$\mu_t = (id, \alpha^*(t,\cdot,Du(t,\cdot);\mu_t))\,\sharp m_t, \tag{32}$$

where the right-hand side is the image of the measure m_t by the map $x \to (x,\alpha^*(t,x,Du(t,x);\mu_t))$. To summarize, the MFG of control takes the form:

$$\begin{cases} (i) & -\partial_t u(t,x) - \mathrm{tr}(a(t,x)D^2u(t,x)) + H(t,x,Du(t,x);\mu_t) = 0 \text{ in } (0,T)\times\mathbb{R}^d, \\ (ii) & \partial_t m_t(x) - \sum_{i,j}\partial_{ij}(a_{ij}(t,x)m_t(x)) - \mathrm{div}\,(m_t(x)D_pH(t,x,Du(t,x);\mu_t)) = 0 \\ & \hfill \text{in } (0,T)\times\mathbb{R}^d, \\ (iii) & m_0(x) = \bar{m}_0(x),\ u(T,x) = g(x,m_T) \quad \text{in } \mathbb{R}^d, \\ (iv) & \mu_t = (id,\alpha^*(t,\cdot,Du(t,\cdot);\mu_t))\,\sharp m_t \quad \text{in } [0,T]. \end{cases} \tag{33}$$

The typical framework in which we expect to have a solution is the following: u is continuous in (t,x), Lipschitz continuous in x (uniformly with respect to t) and satisfies equation (28) in the viscosity sense; m is in L^∞ and satisfies (30) in the sense of distribution.

In order to state the assumptions on the data, we need a few notations. Given a metric space (E,δ_E) we denote by $\mathcal{P}_1(E)$ the set of Borel probability measures ν on E with a finite first order moment $M_1(\nu)$:

$$M_1(\nu) = \int_E \delta_E(x_0,x)d\nu(x) < +\infty$$

for some (and thus all) $x_0 \in E$. The set $\mathcal{P}_1(E)$ is endowed with the Monge-Kantorovitch distance:

$$\mathbf{d}_1(\nu_1,\nu_2) = \sup_\phi \int_E \phi(x)d(\nu_1-\nu_2)(x) \qquad \forall \nu_1,\nu_2 \in \mathcal{P}_1(E),$$

where the supremum is taken over all 1-Lipschitz continuous maps $\phi : E \to \mathbb{R}$.

We will prove the existence of a solution for (33) under the following assumptions:

1. The terminal cost $g : \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}$ and the diffusion matrix $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d\times D}$ are continuous and bounded, uniformly bounded in C^2 in the space variable,
2. The drift has a separate form: $b(t, x, \alpha, \mu_t) = b_0(t, x, \mu_t) + b_1(t, x, \alpha)$,
3. The map $L : [0, T] \times \mathbb{R}^d \times A \times \mathcal{P}_1(\mathbb{R}^d \times A) \to \mathbb{R}$ satisfies the Lasry-Lions monotonicity condition: for any $\nu_1, \nu_2 \in \mathcal{P}_1(\mathbb{R}^d \times A)$ with the same first marginal,
$$\int_{\mathbb{R}^d\times A} (L(t, x, \alpha; \nu_1) - L(t, x, \alpha; \nu_2))d(\nu_1 - \nu_2)(x, \alpha) \geq 0,$$
4. For each $(t, x, p, \nu) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d \times A)$, there exists a unique maximum point $\alpha^*(t, x, p; \nu)$ in (31) and $\alpha^* : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d \times A) \to \mathbb{R}$ is continuous, with a linear growth: for any $L > 0$, there exists $C_L > 0$ such that
$$\delta_A(\alpha_0, \alpha^*(t, x, p; \nu)) \leq C_L(|x|+1) \qquad \forall (t, x, p, \nu) \in [0, T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{P}_1(\mathbb{R}^d\times A) \text{ with } |p| \leq L,$$
(where α_0 is a fixed element of A).
5. The Hamiltonian $H : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d \times A) \to \mathbb{R}$ is continuous; H is bounded in C^2 in (x, p) uniformly with respect to (t, ν), and convex in p.
6. The initial measure $\bar{m}_0$ is a continuous probability density on $\mathbb{R}^d$ with a finite second order moment.

The uniform bounds and uniform continuity assumptions are very strong requirements: for instance they are not satisfied in the linear-quadratic example studied before. These conditions can be relaxed in a more or less standard way; we choose not do so in order to keep the argument relatively simple.

Following [Carmona and Delarue, 2017], assumptions (2) and (3) ensure the uniqueness of the fixed-point in (32). For the sake of completeness, details are given in Lemma 5.2 below.

Theorem 5.1. *Under the above assumptions, there exists at least one solution to the MFG system of controls* (33) *for which* (μ_t) *is continuous from* $[0, T]$ *to* $\mathcal{P}_1(\mathbb{R}^d \times A)$.

As almost always the case for this kind of results, the argument of proof consists in applying a fixed point argument of Schauder type and requires therefore compactness properties. The main difficulty is the control of the time regularity of the evolving measure μ. This regularity is related with the time regularity of the optimal controls. In the first order case, it is not an issue because the optimal trajectories satisfy a Pontryagin maximum principle and thus are (at least) uniformly C^1 (see [Gomes et al., 2014, Gomes and Voskanyan, 2016]). In the second order setting, the Pontryagin maximum principle is not so simple to manipulate and a similar regularity for the controls would be much more heavy to express. We use instead the full power of semi-concavity of the value function combined with compactness arguments (see Lemma 5.4 below). The main advantage of our approach is its robustness: for instance stability property of the solution is almost straightforward.

The proof of Theorem 5.1 requires several preliminary remarks. Let us start with the fixed-point relation (32).

Lemma 5.2. *Let* $m \in \mathcal{P}_2(\mathbb{R}^d)$ *with a bounded density and* $p \in L^\infty(\mathbb{R}^d, \mathbb{R}^d)$.

- *(Existence and uniqueness.) There exists a unique fixed point* $\mu = F(p, m) \in \mathcal{P}_1(\mathbb{R}^d \times A)$ *to the relation*
$$\mu = (id, \alpha^*(t, \cdot, p(\cdot); \mu))\sharp m. \tag{34}$$
Moreover, there exists a constant C_0*, depending only on* $\|p\|_\infty$ *and on the second order moment of* m*, such that*
$$\int_{\mathbb{R}^d\times A} \left\{|x|^2 + \delta_A(\alpha_0, \alpha)\right\} \, d\mu(x, \alpha) \leq C_0.$$

- *(Stability.) Let (m_n) be a family of $\mathcal{P}_1(\mathbb{R}^d)$, with a uniformly bounded density in L^∞ and uniformly bounded second order moment, which converges in $\mathcal{P}_1(\mathbb{R}^d)$ to some m, (p_n) be a uniformly bounded family in L^∞ which converges a.e. to some p. Then $F(p_n, m_n)$ converges to $F(p,m)$ in $\mathcal{P}_1(\mathbb{R}^d \times A)$.*

The uniqueness part is borrowed from [Carmona and Delarue, 2017].

Proof. Let $L := \|p\|_\infty$. For $\mu \in \mathcal{P}_1(\mathbb{R}^d \times A)$, let us set $\Psi(\mu) := (id, \alpha^*(t, \cdot, p(\cdot); \mu))\sharp m$. Using the growth assumption on α^*, we have

$$\begin{aligned}\int_{\mathbb{R}^d\times A} |x|^2 + \delta_A^2(\alpha_0, \alpha)\, d\Psi(\mu)(x,\alpha) &= \int_{\mathbb{R}^d} |x|^2 + \delta_A^2(\alpha_0, \alpha^*(t,x,p(x);\mu))\, m(x)dx \\ &\leq \int_{\mathbb{R}^d\times A} |x|^2 + C_L^2(|x|+1)^2\, m(x)dx =: C_0.\end{aligned}$$

This leads us to define $\mathcal{K}$ as the convex and compact subset of measures $\mu \in \mathcal{P}_1(\mathbb{R}^d \times A)$ with second order moment bounded by C_0. Let us check that the map Ψ is continuous on $\mathcal{K}$. If (μ_n) converges to μ in $\mathcal{K}$, we have, for any map ϕ, continuous and bounded on $\mathbb{R}^d \times A$:

$$\int_{\mathbb{R}^d\times A} \phi(x,\alpha) d\Psi(\mu_n)(x,\alpha) = \int_{\mathbb{R}^d} \phi(x, \alpha^*(t,x,p(x);\mu_n) m(x)dx.$$

By continuity of α^*, the term $\phi(x, \alpha^*(t,x,p(x);\mu_n))$ converges a.e. to $\phi(x, \alpha^*(t,x,p(x);\mu)$ and is bounded. The mesure m having a bounded second order moment and being absolutely continuous, this implies the convergence of the integral:

$$\lim_n \int_{\mathbb{R}^d\times A} \phi(x,\alpha) d\Psi(\mu_n)(x,\alpha) = \int_{\mathbb{R}^d} \phi(x,\alpha^*(t,x,p(x);\mu)) m(x)dx = \int_{\mathbb{R}^d\times A} \phi(x,\alpha) d\Psi(\mu)(x,\alpha).$$

Thus the sequence $(\Psi(\mu_n))$ converges weakly to $\Psi(\mu)$ and, having a uniformly bounded second order moment, also converges for the $\mathbf{d}_1$ distance. The map Ψ is continuous on the compact set $\mathcal{K}$ and therefore has a fixed point by Schauder fixed point Theorem.

Let us now check the uniqueness. If there exists two fixed points μ_1 and μ_2 of (34), then, by the monotonicity condition in Assumption (3), we have

$$\begin{aligned}0 &\leq \int_{\mathbb{R}^d\times A} \{L(t,x,\alpha;\mu_1) - L(t,x,\alpha;\mu_2)\}\, d(\mu_1-\mu_2)(x,\alpha) \\ &= \int_{\mathbb{R}^d} \{L(x,\alpha_1(x);\mu_1) - L(x,\alpha_1(x);\mu_2) - L(x,\alpha_2(x);\mu_1) + L(x,\alpha_2(x);\mu_2)\}\, m(x)dx\end{aligned}$$

where we have set $\alpha_i(x) := \alpha^*(t,x,p(x);\mu_i)$ $(i=1,2)$ and $L(x,\dots) := L(t,x,p(x),\dots)$ to simplify the expressions. So

$$\begin{aligned}0 \leq \int_{\mathbb{R}^d} &\{b_1(t,x,\alpha_1(x))\cdot p(x) + L(x,\alpha_1(x);\mu_1) - b_1(t,x,\alpha_2(x))\cdot p(x) - L(x,\alpha_2(x);\mu_1) \\ &-b_1(t,x,\alpha_1(x))\cdot p(x) - L(x,\alpha_1(x);\mu_2) + b_1(t,x,\alpha_2(x))\cdot p(x) + L(x,\alpha_2(x);\mu_2)\}\, m(x)dx,\end{aligned}$$

where, by assumption (4), $\alpha_1(x)$ is the unique maximum point in the expression

$$-b_1(t,x,\alpha)\cdot p(x) - L(t,x,p(x),\alpha;\mu_1)$$

and $\alpha_2(x)$ the unique maximum point in the symmetric expression with μ_2. This implies that $\alpha_1 = \alpha_2$ $m-$a.e., and therefore, by the fixed point relation (34), that $\mu_1 = \mu_2$.

Finally we show the stability. In view of the previous discussion, we know that $(\mu_n := F(p_n, m_n))$ has a uniformly bounded second order moment and thus converges, up to a subsequence, to some $\tilde{\mu}$. We just have to check that $\tilde{\mu} = F(p,m)$, i.e., $\tilde{\mu}$ satisfies the fixed-point relation. Let ϕ be a continuous and bounded map on $\mathbb{R}^d \times A$. Then

$$\int_{\mathbb{R}^d\times A} \phi(x,\alpha) d\mu_n(x,\alpha) = \int_{\mathbb{R}^d} \phi(x,\alpha^*(t,x,p_n(x);\mu_n)) m_n(x)dx.$$

As (m_n) is bounded in L^∞ and converges to m in $\mathcal{P}_1(\mathbb{R}^d)$, it converges in $L^\infty-$weak-$*$ to m. The map ϕ being bounded and the (m_n) having a uniformly bounded second order moment, we have therefore

$$\lim_n \int_{\mathbb{R}^d} \phi(x, \alpha^*(t,x,p(x);\mu))m_n(x)dx = \int_{\mathbb{R}^d} \phi(x, \alpha^*(t,x,p(x);\mu))m(x)dx.$$

On the other hand, by continuity of α^* and a.e. convergence of (p_n), $\phi(x, \alpha^*(t,x,p_n(x);\mu_n))$ converges to $\phi(x,\alpha^*(t,x,p(x);\mu))$ a.e. and thus in L^1_{loc}. As the (m_n) are uniformly bounded in L^∞ and have a uniformly bounded second order moment and as ϕ is bounded, this implies that

$$\lim_n \int_{\mathbb{R}^d} \phi(x, \alpha^*(t,x,p_n(x);\mu_n))m_n(x)dx - \int_{\mathbb{R}^d} \phi(x, \alpha^*(t,x,p(x);\mu))m_n(x)dx = 0.$$

So we have proved that

$$\lim_n \int_{\mathbb{R}^d\times A} \phi(x,\alpha)d\mu_n(x,\alpha) = \int_{\mathbb{R}^d\times A} \phi(x,\alpha)d\mu(x,\alpha),$$

which implies that (μ_n) converges to μ in $\mathcal{P}_1(\mathbb{R}^d\times A)$ because of the second order moment estimates on (μ_n). $\square$

Next we address the existence of a solution to the Kolmogorov equation when the map u is Lipschitz continuous and has some semi-concavity property in space. We will see in the proof of Theorem 5.1 that this is typically the regularity for the solution of the Hamilton-Jacobi equation.

Lemma 5.3. *Assume that $u = u(t,x)$ is uniformly Lipschitz continuous in space with constant $M > 0$ and semi-concave with respect to the space variable with constant M and that (μ_t) is a time-measurable family in $\mathcal{P}_1(\mathbb{R}^d\times A)$. Then there exists at least one solution to the Kolmogorov equation (30) which satisfies the bound*

$$\|m_t(\cdot)\|_\infty \le \|\bar m_0\|_\infty \exp\{C_0 t\}, \tag{35}$$

where C_0 is given by

$$C_0 := C_0(M) = \sup_{(t,x)} |D^2 a(t,x)| + \sup_{(t,x,p,\nu)} \left|D^2_{xp} H(t,x,p;\nu)\right| + \sup_{(t,x,p,X,\nu)} \mathrm{Tr}\left(D^2_{pp}H(t,x,p;\nu)X\right)$$

where the suprema are taken over the $(t,x,p,X,\nu) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{S}^d\times\mathcal{P}_1(\mathbb{R}^d\times A)$ such that $|p|\le M$ and $X \le M\ Id$. Moreover, we have the second order moment estimate

$$\int_{\mathbb{R}^d} |x|^2 m_t(x)dx \le M_0 \tag{36}$$

and the continuity in time estimate

$$\sup_{s,t} \mathbf{d}_1\left(m_{n,s}), m_{n,t}\right) \le M_0|t-s|^{\frac{1}{2}} \qquad \forall s,t\in[0,T], \tag{37}$$

where M_0 depends only on $\|a\|_\infty$, the second order moment of $\bar m_0$ and on $\sup_{t,x,p,\nu}|D_pH(t,x,p;\nu)|$, the supremum being taken over the $(t,x,p,\nu)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{P}_1(\mathbb{R}^d\times A)$ such that $|p|\le M$.

Assume now that $(u_n = u_n(t,x))$ is a family of continuous maps which are Lipschitz continuous and semi-concave in space with uniformly constant M and converge locally uniformly to a map u; that $(\mu_n = \mu_{n,t})$ converges in $C^0([0,T],\mathcal{P}_1(\mathbb{R}^d\times A))$ to μ. Let m_n be a solution to the Kolmogorov equation associated with u_n, μ_n which satisfies the L^∞ bound (35). Then (m_n) converges, up to a subsequence, in $L^\infty-$weak, to a solution m of the Kolmogorov equation associated with u and μ.*

Proof. We first prove the existence of a solution. Let (u_n) be smooth approximations of u_n. Then equation

$$\left\{ \begin{array}{l} \partial_t m_t(x) - \sum_{i,j} \partial_{ij}(a_{n,ij}(t,x)m_t(x)) - \mathrm{div}\,(m_t(x)D_pH(t,x,Du_n(t,x);\mu_t)) = 0 \text{ in } (0,T)\times\mathbb{R}^d \\ m_0(x) = \bar{m}_0(x) \qquad \text{in } \mathbb{R}^d \end{array} \right.$$

has a unique classical solution m_n, which is the law of the process

$$\left\{ \begin{array}{l} dX_t = -D_pH(X_t, Du_n(t,X_t);\mu_t)dt + \sigma(t,X_t)dW_t \\ X_0 = \mathbf{x_0} \end{array} \right.$$

where $\mathbf{x_0}$ is a random variable with law $\bar{m}_0$ independent of W. By standard argument, we have therefore that

$$\sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|^2 m_{n,t}(dx) \le \mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^2\right] \le M_0,$$

where M_0 depends only on $\|a\|_\infty$, the second order moment of $\bar{m}_0$ and on $\sup_{t,x,p,\nu} |D_pH(t,x,p;\nu)|$, the supremum being taken over the $(t,x,p,\nu)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\mathcal{P}_1(\mathbb{R}^d\times A)$ such that $|p|\le M$. Moreover,

$$\mathbf{d}_1\left(m_n(s), m_n(t)\right) \le \sup_{s,t} \mathbb{E}\left[|X_t - X_s|\right] \le M_0|t-s|^{\frac{1}{2}} \qquad \forall s,t\in[0,T],$$

changing M_0 if necessary. The key step is that (m_n) is bounded in L^∞. For this we rewrite the equation for m_n as a linear equation in non-divergence form

$$\partial_t m_n - \mathrm{Tr}\left(aD^2 m_n\right) + b_n\cdot Dm_n - c_n m_n = 0$$

where b_n is a some bounded vector field and c_n is given by

$$c_n(t,x) = \sum_{i,j} \partial^2_{ij} a_{n,ij}(t,x) + \mathrm{Tr}\left(D^2_{xp}H(t,x,Du_n(t,x);\mu_t)\right) + \mathrm{Tr}\left(D^2_{pp}H(t,x,Du_n(t,x);\mu_t)D^2u_n(t,x)\right).$$

As u is uniformly Lipschitz continuous and semi-concave with respect to the $x-$variable, we can assume that (u_n) enjoys the same property, so that

$$\|Du_n\|_\infty \le M, \qquad D^2 u_n \le M\ I_d.$$

Then

$$|D^2 a_n(t,x)| + \left|D^2_{xp}H(t,x,Du_n(t,x);\mu_t)\right| + \mathrm{Tr}\left(D^2_{pp}H(t,x,Du_n(t,x);\mu_t)D^2u_n(t,x)\right) \le C_0$$

because $D^2_{pp}H \ge 0$ by convexity of H. This proves that

$$c_n(t,x) \le C_0 \qquad \text{a.e.}$$

By standard maximum principle, we have therefore that

$$\sup_x m_n(t,x) \le \sup_x \bar{m}_0(x)\exp(C_0 t) \qquad \forall t\ge 0,$$

which proves the uniform bound of m_n in L^∞. Thus (m_n) converges weakly-* in L^∞ to some map m satisfying (in the sense of distribution)

$$\left\{ \begin{array}{l} \partial_t m - \sum_{i,j} \partial_{ij}(a_{ij}m) - \mathrm{div}\,(mD_pH(t,x,Du;\mu_t)) = 0 \text{ in } (0,T)\times\mathbb{R}^d \\ m_0(x) = \bar{m}_0(x) \qquad \text{in } \mathbb{R}^d \end{array} \right.$$

This shows the existence of a solution. The proof of the stability goes along the same line, except that there is no need of regularization. □

Next we need to show some uniform regularity in time of Du where u is a solution of the Hamilton-Jacobi equation. For this, let us note that the set

$$\mathcal{D} := \{p \in L^\infty(\mathbb{R}^d),\ \exists v \in W^{1,\infty}(\mathbb{R}^d),\ p = Dv,\ \|v\|_\infty \leq M,\ \|Dv\|_\infty \leq M,\ D^2 v \leq M\ I_d\} \tag{38}$$

is sequentially compact for the a.e. convergence and therefore for the distance

$$d_{\mathcal{D}}(p_1, p_2) = \int_{\mathbb{R}^d} \frac{|p_1(x) - p_2(x)|}{(1+|x|)^{d+1}} dx \qquad \forall p_1, p_2 \in \mathcal{D}. \tag{39}$$

Lemma 5.4. *There is a modulus ω such that, for any $(\mu_t) \in C^0([0,T], \mathcal{P}_1(\mathbb{R}^d \times A))$, the viscosity solution u to (28) satisfies*

$$d_{\mathcal{D}}(Du(t_1,\cdot), Du(t_2,\cdot)) \leq \omega(|t_1 - t_2|) \qquad \forall t_1, t_2 \in [0,T].$$

Proof. Suppose that the result does not hold. Then there exists $\epsilon > 0$ such that, for any $n \in \mathbb{N}\backslash\{0\}$, there is $\mu_n \in C^0([0,T], \mathcal{P}_1(\mathbb{R}^d \times A))$ and times $t_n \in [0, T-h_n]$, $h_n \in [0, 1/n]$ with

$$d_{\mathcal{D}}(Du_n(t_n,\cdot), Du_n(t_n + h_n, \cdot)) \geq \varepsilon,$$

where u_n is the solution to (28) associated with μ_n. In view of our regularity assumption on H, the map $(x,p) \to H(t,x,p;\mu_{n,t})$ is locally uniformly bounded in C^2 independently of t and n. So, there exists a time-measurable Hamiltonian $h = h(t,x,p)$, obtained as a weak limit of $H(\cdot,\cdot,\cdot;\mu_{n,\cdot})$, such that, up to a subsequence, $\int_0^t H(s,x,p;\mu_{n,s})ds$ converges locally uniformly in x, p to $\int_0^t h(s,x,p)ds$. Note that h inherits the regularity property of H in (x,p). Using the notion of L^1-viscosity solution and Barles' stability result of L^1-viscosity solutions for the weak (in time) convergence of the Hamiltonian [Barles, 2006], we can deduce that the u_n converge locally uniformly to the L^1-viscosity solution u of the Hamilton-Jacobi equation (28) associated with the Hamiltonian h. Without loss of generality, we can also assume that (t_n) converges to some $t \in [0,T]$. Then, by semi-concavity, $Du_n(t_n,\cdot)$ and $Du_n(t_n + h_n, \cdot)$ both converge a.e. to $Du(t,\cdot)$ because $u_n(t_n,\cdot)$ and $u_n(t_n+h_n,\cdot)$ both converge to $u(t,\cdot)$ locally uniformly. As the Du_n are uniformly bounded in L^∞, we conclude that

$$d_{\mathcal{D}}(Du_n(t_n,\cdot), Du_n(t_n + h_n, \cdot)) \to 0,$$

and there is a contradiction. □

Proof of Theorem 5.1. We proceed as usual by a fixed point argument. We first solve an approximate problem, in which we smoothen the Komogorov equation, and then we pass to the limit. Let $\varepsilon > 0$ small, $\xi^\varepsilon = \varepsilon^{-(d+1)}\xi((t,x)/\varepsilon)$ be a standard smooth mollifier.

To any $\mu = (\mu_t)$ belonging to $C^0([0,T], \mathcal{P}_1(\mathbb{R}^d \times A))$ we associate the unique viscosity solution u to (28). Note that, with our assumption on H and G, u is uniformly bounded, uniformly continuous in (t,x) (uniformly with respect to (μ_t)), uniformly Lipschitz continuous and semi-concave in x (uniformly with respect to t and (μ_t)). We denote by M the uniform bound on the $L^\infty-$norm, the Lipschitz constant and semi-concavity constant:

$$\|u\|_\infty \leq M, \qquad \|Du\|_\infty \leq M, \qquad D^2 u \leq M\ I_d. \tag{40}$$

Then we consider (m_t) to be the unique solution to the (smoothened) Kolmogorov equation

$$\begin{cases} \partial_t m_t(x) - \sum_{i,j} \partial_{ij}(a_{ij}(t,x) m_t(x)) - \mathrm{div}\,(m_t(x) D_p H(t,x,Du^\varepsilon(t,x);\mu_t)) = 0 \\ \hfill \text{in } (0,T) \times \mathbb{R}^d \\ m_0(x) = \bar{m}_0(x) \qquad \text{in } \mathbb{R}^d \end{cases} \tag{41}$$

where $u^\varepsilon = u \star \xi^\varepsilon$. Following Lemma 5.3, the solution m—which is unique thanks to the space regularity of the drift—satisfies the bounds (35), (36) and (37) (which are independent of μ

and ε). Finally, we set $\tilde{\mu}_t = F(m_t, Du(t,\cdot))$, where F is defined in Lemma 5.2. From Lemma 5.2, we know that there exists $C_0 > 0$ (still independent of μ and ε) such that

$$\sup_{t\in[0,T]} \int_{\mathbb{R}^d\times A} |x|^2 + \delta_A(\alpha_0,\alpha)\, d\tilde{\mu}_t(x,\alpha) \leq C_0.$$

Our aim is to check that the map $\Psi^\varepsilon : (\mu_t) \to (\tilde{\mu}_t)$ has a fixed point. Let us first prove that $(\tilde{\mu}_t)$ is uniformly continuous in t with a modulus ω independent of (μ_t) and ε. Recall that the set $\mathcal{D}$ defined by (38) is compact. Moreover, the subset $\mathcal{M}$ of measures in $\mathcal{P}_1(\mathbb{R}^d)$ with a second moment bounded by a constant C_0 is also compact. By Lemma 5.2, the map F is continuous on the compact set $\mathcal{D}\times\mathcal{M}$, and thus uniformly continuous. On the other hand, Lemma 5.4 states that the map $t \to Du(t,\cdot)$ is continuous from $[0,T]$ to $\mathcal{D}$, with a modulus independent of (μ_t). As (m_t) is also uniformly continuous in time (recall (37)), we deduce that $(\tilde{\mu}_t := F(m_t, Du(t,\cdot))$ is uniformly continuous with a modulus ω independent of (μ_t) and ε.

Let $\mathcal{K}$ be the set of $\mu \in C^0([0,T], \mathcal{P}(\mathbb{R}^d\times A))$ with a second order moment bounded by C_0 and a modulus of time-continuity ω. Note that $\mathcal{K}$ is convex and compact and that Ψ^ε is defined from $\mathcal{K}$ to $\mathcal{K}$.

Next we show that the map $\mu \to \Psi^\varepsilon(\mu)$ is continuous on $\mathcal{K}$. Let (μ_n) converge to μ in $\mathcal{K}$. From the stability of viscosity solution, the solution u_n of (28) associated to μ_n converge locally uniformly to the solution u associated with μ. Moreover, recalling the estimates (40), the (Du_n) are uniformly bounded and converge a.e. to Du by semi-concavity. Let m_n be the solution to (41) associated with μ_n and u_n. By the stability part of Lemma 5.3, any subsequence of the compact family (m_n) converges uniformly to a solution (41). This solution m being unique, the full sequence (m_n) converges to m. By continuity of the map F, we then have, for any $t\in[0,T]$,

$$\Psi^\varepsilon(\mu_n)(t) = F(m_t, Du(t,\cdot)) = \lim_n F(m_{n,t}, Du_n(t,\cdot)) = \lim_n \Psi^\varepsilon(\mu_n)(t).$$

Since the (m_n) and (Du_n) are uniformly continuous in time, the convergence of $\Psi^\varepsilon(\mu_n)$ to $\Psi^\varepsilon(\mu)$ actually holds in $\mathcal{K}$.

So, by Schauder fixed point Theorem, Ψ^ε has a fixed point μ^ε. It remains to let $\varepsilon\to 0$. Up to a subsequence, labelled in the same way, (μ^ε) converges to some $\mu\in\mathcal{K}$. As above, the solution u^ε of (28) associated with μ^ε converges to the solution u associated with μ and Du^ε converges to Du a.e.. The solution m^ε of (41) (with μ^ε and u^ε) satisfies the estimates (35), (36) and (37). Thus the stability result in Lemma 5.3 implies that a subsequence, still denoted (m^ε), converges uniformly to a solution m of the unperturbed Kolmogorov equation (30). Recall that $\mu^\varepsilon_t = F(m^\varepsilon_t, u^\varepsilon(t,\cdot))$ for any $t\in[0,T]$. We can pass to the limit in this expression to get $\mu_t = F(m_t, u(t,\cdot))$ for any $t\in[0,T]$. Then the triple (u,m,μ) satisfies system (33). □

6 Conclusion

In this paper, we proposed a model for optimal execution or optimal trading in which, instead of having as usual a large trader in front of a neutral "background noise", the trader is surrounded by a continuum of other market participants. The cost functions of these participants have the same shape, without necessarily sharing the same values of parameters. Each player of this mean field game (MFG) is similar in the following sense: it suffers from temporary market impact, impacts permanently the price, and fears uncertainty.

The stake of such a framework is to provide robust trading strategies to asset managers, and to shed light on the price formation process for regulators.

Our framework is not a traditional MFG, but falls into the class of *extended mean filed games*, in which participants interact via their controls. We provide a generic framework to address it for a vast class of cost functions, beyond the scope of our model.

Thanks to it, we solved our model and provide insights on the influence of its parameters (temporary and permanent market impact coefficients, terminal penalization, risk aversion and duration of the game) on the obtained results. We provide the solution in a closed form and formulate a series of "stylized facts" (Stylized Fact 1 to Stylized Fact 6) describing our results. For instance we unveil three components of the optimal control: two coming from the mean field $E(t)$ and its derivative $E'(t)$ (summarized in a function $h_1(t)$), and the third one proportional to the remaning quantity to trade $q(t)$ via an increasing function $h_2(t)$. We show also how to slow down trading when the net inventory of the participants is of the opposite sign (i.e. the "market" is buying while the trader is selling, or the reverse): $h_2(t)$ is unchanged but $h_1(t)$ is changed in its opposite.

To conclude on this, we provide numerical illustrations showing market participants could end up not following their initial instructions, for some configurations of the market structure. This could help regulators to smooth such behaviours if needed.

In a second stage, we address the case of heterogenous preferences (i.e. when each agent has his own risk aversion parameter). We show the existence of an unique solution but do no more have a closed form formulation. Last but not least we study a more realistic case in which participants do not know instantaneously the optimal strategies of others, but have *to learn them*. We list in Proposition 4.2 conditions needed so that the learnt strategy is the optimal one.

A Proof of Proposition 3.1

Proof of Proposition 3.1. The discriminant of the second order equation is

$$\Delta = \alpha^2 + 16\kappa\phi$$

and the roots are

$$r_\pm = -\frac{\alpha}{4\kappa} \pm \frac{1}{\kappa}\sqrt{\kappa\phi + \frac{\alpha^2}{16}}.$$

Hence

$$E(t) = E_0 a \left(\exp\{r_+ t\} - \exp\{r_- t\}\right) + E_0 \exp\{r_- t\}$$

where $a \in \mathbb{R}$ determined by the condition

$$\kappa E'(T) + AE(T) = 0$$

and thus has to solve the relation

$$\begin{aligned}\kappa E_0 \left[a\left(r_+ \exp\{r_+ T\} - r_- \exp\{r_- T\}\right) + r_- \exp\{r_- T\}\right] \\ + AE_0 \left[a\left(\exp\{r_+ T\} - \exp\{r_- T\}\right) + \exp\{r_- T\}\right] = 0.\end{aligned}$$

There is a unique solution if $E_0 \neq 0$ and

$$\kappa\left(r_+ \exp\{r_+ T\} - r_- \exp\{r_- T\}\right) + A\left(\exp\{r_+ T\} - \exp\{r_- T\}\right) \neq 0. \tag{42}$$

Writing $r^\pm = -\dfrac{\alpha}{4\kappa} \pm \theta$ where $\theta := \frac{1}{\kappa}\sqrt{\kappa\phi + \frac{\alpha^2}{16}}$, condition (42) is equivalent to

$$\left[(-\frac{\alpha}{4} + \kappa\theta)\exp\{\theta T\} - (-\frac{\alpha}{4} - \kappa\theta)\exp\{-\theta T\}\right] + A\left[\exp\{\theta T\} - \exp\{-\theta T\}\right] \neq 0,$$

which leads to the condition

$$-\frac{\alpha}{2}\mathrm{sh}\{\theta T\} + 2\kappa\theta\mathrm{ch}\{\theta T\} + 2A\mathrm{sh}\{\theta T\} \neq 0.$$

As $\mathrm{ch}\{\theta T\} > \mathrm{sh}\{\theta T\}$, one has

$$\begin{aligned} -\frac{\alpha}{2}\mathrm{sh}\{\theta T\} + 2\kappa\theta\mathrm{ch}\{\theta T\} + 2A\mathrm{sh}\{\theta T\} \;&>\; \mathrm{sh}\{\theta T\}\left(-\frac{\alpha}{2} + 2\kappa\theta + 2A\right) \\ &= \mathrm{sh}\{\theta T\}\left(-\frac{\alpha}{2} + 2\left(\kappa\phi + \frac{\alpha^2}{16}\right)^{1/2} + 2A\right) \;\geq\; 2A\mathrm{sh}\{\theta T\} \;>\; 0. \end{aligned}$$

So condition (42) is always fulfilled and

$$\begin{aligned} a \;&=\; -\frac{\kappa r_- \exp\{r_- T\} + A\exp\{r_- T\}}{\kappa\left(r_+\exp\{r_+ T\} - r_-\exp\{r_- T\}\right) + A\left(\exp\{r_+ T\} - \exp\{r_- T\}\right)} \\ &=\; -\frac{(-\alpha/4 - \kappa\theta + A)\exp\{-\theta T\}}{-\frac{\alpha}{2}\mathrm{sh}\{\theta T\} + 2\kappa\theta\mathrm{ch}\{\theta T\} + 2A\mathrm{sh}\{\theta T\}}. \end{aligned}$$

To compute h_2, we note that it solves the following backward ordinary differential equation (15a):

$$\begin{cases} 0 &= 2\kappa \cdot h_2'(t) + 4\kappa \cdot \phi - (h_2(t))^2 \\ h_2(T) &= 2A \end{cases}$$

It is easy to check the solution is given by (18), where where $r = 2\sqrt{\phi/\kappa}$ and c_2 solves the terminal condition. Hence

$$c_2 = \frac{1 - A/\sqrt{\kappa\phi}}{1 + A/\sqrt{\kappa\phi}} \cdot e^{-rT}.$$

□

References

[Alfonsi and Blanc, 2014] Alfonsi, A. and Blanc, P. (2014). Dynamic optimal execution in a mixed-market-impact Hawkes price model.

[Almgren et al., 2005] Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct Estimation of Equity Market Impact. *Risk*, 18:57–62.

[Almgren and Chriss, 2000] Almgren, R. F. and Chriss, N. (2000). Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39.

[Bacry et al., 2015] Bacry, E., Iuga, A., Lasnier, M., and Lehalle, C.-A. (2015). Market Impacts and the Life Cycle of Investors Orders. *Market Microstructure and Liquidity*, 1(2).

[Barles, 2006] Barles, G. (2006). A new stability result for viscosity solutions of nonlinear parabolic equations with weak convergence in time. *Comptes Rendus Mathematique*, 343(3):173–178.

[Bensoussan et al., 2013] Bensoussan, A., Frehse, J., and Yam, P. (2013). *Mean field games and mean field type control theory*, volume 101. Springer.

[Bensoussan et al., 2016] Bensoussan, A., Sung, K. C. J., Yam, S. C. P., and Yung, S. P. (2016). Linear-quadratic mean field games. *Journal of Optimization Theory and Applications*, 169(2):496–529.

[Bernhard, 2003] Bernhard, P. (2003). A robust control approach to option pricing. *Applications of Robust Decision Theory and Ambiguity in Finance*, City University Press, London.

[Bouchard et al., 2011] Bouchard, B., Dang, N.-M., and Lehalle, C.-A. (2011). Optimal control of trading algorithms: a general impulse control approach. *SIAM J. Financial Mathematics*, 2(1):404–438.

[Brokmann et al., 2014] Brokmann, X., Serie, E., Kockelkoren, J., and Bouchaud, J. P (2015). Slow decay of impact in equity markets. *Market Microstructure and Liquidity*, 1(2).

[Caines, 2015] Caines, P. E. (2015). Mean field games. *Encyclopedia of Systems and Control*, pages 706–712.

[Cardaliaguet and Hadikhanloo, 2017] Cardaliaguet, P. and Hadikhanloo, S. (2017). Learning in Mean Field Games: the Fictitious Play. *ESAIM: Control, Optimisation and Calculus of Variations*, 23(2):569–591

[Carmona and Delarue, 2017] Carmona, R. and Delarue, F. (to appear). *Probabilistic Theory of Mean Field Games with Applications*. Springer.

[Carmona et al., 2013] Carmona, R., Fouque, J.-P., and Sun, L.-H. (2013). Mean Field Games and Systemic Risk. *Social Science Research Network Working Paper Series*.

[Cartea and Jaimungal, 2015] Cartea, A. and Jaimungal, S. (2015). Incorporating Order-Flow into Optimal Execution. *Social Science Research Network Working Paper Series*.

[Cartea et al., 2015] Cartea, A., Jaimungal, S., and Penalva, J. (2015). *Algorithmic and High-Frequency Trading (Mathematics, Finance and Risk)*. Cambridge University Press, 1 edition.

[Fudenberg and Levine, 1998] Fudenberg, D. and Levine, D. K. (1998). *The theory of learning in games*, volume 2. MIT press.

[Gomes and Others, 2014] Gomes, D. A. and Others (2014). Mean field games models?a brief survey. *Dynamic Games and Applications*, 4(2):110–154.

[Gomes et al., 2014] Gomes, D. A., Patrizi, S., and Voskanyan, V. (2014). On the existence of classical solutions for stationary extended mean field games. *Nonlinear Analysis: Theory, Methods & Applications*, 99:49–79.

[Gomes and Voskanyan, 2016] Gomes, D. A. and Voskanyan, V. K. (2016). Extended deterministic mean-field games. *SIAM Journal on Control and Optimization*, 54(2):1030–1055.

[Guéant, 2016] Guéant, O. (2016). *The Financial Mathematics of Market Liquidity: From Optimal Execution to Market Making*. Chapman and Hall/CRC.

[Guéant et al., 2011] Guéant, O., Lasry, J.-M., and Lions, P.-L. (2011). Mean field games and applications. In *Paris-Princeton lectures on mathematical finance 2010*, pages 205–266. Springer.

[Guéant and Lehalle, 2015] Guéant, O. and Lehalle, C.-A. (2015). General intensity shapes in optimal liquidation. *Mathematical Finance*, 25(3):457–495.

[Guilbaud and Pham, 2013] Guilbaud, F. and Pham, H. (2013). Optimal High Frequency Trading with limit and market orders. Quantitative Finance, 13(1):79–94.

[Huang et al., 2006] Huang, M., Malhamé, R. P., Caines, P. E., and Others (2006). Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252.

[Jaimungal and Nourian, 2015] Jaimungal, S. and Nourian, M. (2015). Mean-Field Game Strategies for a Major-Minor Agent Optimal Execution Problem. *Social Science Research Network Working Paper Series*.

[Kizilkale and Caines, 2013] Kizilkale, A. C. and Caines, P. E. (2013). Mean field stochastic adaptive control. *IEEE Transactions on Automatic Control*, 58(4):905–920.

[Lasry and Lions, 2006a] Lasry, J.-M. and Lions, P.-L. (2006a). Jeux à champ moyen. I - Le cas stationnaire. *Comptes Rendus Mathematique*, 343(9):619–625.

[Lasry and Lions, 2006b] Lasry, J.-M. and Lions, P.-L. (2006b). Jeux à champ moyen. II - Horizon fini et contrôle optimal. *Comptes Rendus Mathematique*, 343(10):679–684.

[Lasry and Lions, 2007] Lasry, J.-M. and Lions, P.-L. (2007). Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260.

[Lehalle et al., 2013] Lehalle, C.-A., Laruelle, S., Burgot, R., Pelin, S., and Lasnier, M. (2013). *Market Microstructure in Practice*. World Scientific publishing.

[Nourian et al., 2012] Nourian, M., Caines, P. E., Malhamé, R. P., and Huang, M. (2012). Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation. *IEEE Transactions on Automatic Control*, 57(11):2801–2816.

[Obizhaeva and Wang, 2005] Obizhaeva, A. and Wang, J. (2005). Optimal Trading Strategy and Supply/Demand Dynamics. *Social Science Research Network Working Paper Series*.

[Waelbroeck and Gomes, 2013] Waelbroeck, H. and Gomes, C. (2013). Is Market Impact a Measure of the Information Value of Trades? Market Response to Liquidity vs. Informed Trades. *Journal of Financial Markets*, 16(1):1–32.