On the dynamics of measure flows and multi-agent and mean-field financial games

Vadim Platonov

Doctor of Philosophy
University of Edinburgh
2022

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Vadim Platonov*)

To Pouzya and all Poupouzyas

Abstract

This thesis comprises two different collection of results. We present several Itô-Wentzell formulae on Wiener spaces for real-valued functional random field of Itô type that depends on measure flows. We distinguish the full- and the marginal-measure flow cases in the spirit of mean-field games without or with common noise respectively.

Second part studies several portfolio management problems featuring many-player and mean-field competition and relative performance concerns under the forward performance processes (FPP) framework. In the first problem, we focus on agents using power (CRRA) type FPPs for their investment-consumption optimisation problem under a common noise Merton market model. We solve both the many-player and mean field game providing closed-form expressions for the solutions where the limit of the former yields the latter. In our case, the FPP framework yields a continuum of solutions for the consumption component as indexed to a market parameter we coin "market-risk relative consumption preference". The parameter permits the agent to set a preference for their consumption going forward in time that, in the competition case, reflects a common market behaviour. We show the FPP framework, under both competition and no-competition, allows the agent to disentangle their risk-tolerance and elasticity of intertemporal substitution (EIS). This, in turn, allows a finer analysis on the agent's consumption "income" and "substitution" regimes, and, of independent interest, motivates a new strand of economics research on EIS under the FPP framework. We find that competition rescales the agent's perception of consumption in a non-trivial manner in addition to a time-dependent Elasticity of Conformity of the agent to the market-risk relative consumption preference. In the follow-up problem, we solve the forward-utility finite player and mean-field investment game for the agent following exponential (CARA) type FPPs. We explicitly derive best response and equilibrium strategies in the single common stock asset and the asset specialisation with common noise. As an application, we draw on the core features of the forward utility paradigm and discuss a problem of time-consistent mean-field dynamic model selection in sequential time-horizons.

Lay summary

In the first part of this thesis lay the results on generalised Itô-type expansions for the measure-dependent random fields. It is common in natural sciences that the measuring functional experiences the perturbation of the same nature as the underlying signals. These signals can be modelled as trajectories of the idiosyncratic particles. For the large number of particles the first simplification occurs, when substituting the multiple effects of them all by the one of distribution of the particles' paths.

In Lions calculus, the dynamics of the functional of measures is determined by the average trajectory across the whole field of particles. When the functional has an Itô differential (random field), the expected contribution of cross-correlation between the noises of the random field and generic particle does not appear. This phenomenon can be understood as an effect of the 'mean-field' regime, where the whole system is independent of one particle (even when the latter is treated as 'representative' of all the others). When the functional and the cloud of particles are subjected to the common noise, and conditioned on it, the cross-variation becomes non-trivial.

The second part conveys the study of financial games of investment and consumption under relative performance concerns and forward utility framework. Both the investment game under CARA preferences and the investment-consumption game under CRRA preferences are analytically solvable. We determine the finite-player Nash equilibria and explicitly derive the mean-field equilibria, the former converges to the latter.

Under the CRRA preferences, the forward performance setting allows assigning different time dynamics for utilities from wealth and consumption. Their dissimilarity is captured by the new parameter of modelling, market-risk relative consumption preference. Within this framework one can further observe the non-canonical dependency between Elasticity of Intertemporal Substitution and risk-aversion with both the absence and presence of competition.

Acknowledgements

First of all, I would like to thank Gonçalo dos Reis for his incredible help and support. I could not imagine enduring all the PhD journey without such a wonderful and caring supervisor. You were giving me huge inspiration in time of frustration, always coming up with helpful comments and suggestions in various situations, not limited to research.

I am very grateful to my collaborators, Burak Buke, Ankush Agarwal, Stefano Pagliarani and Francesco Russo, for their contribution and guidance that helped to shape my mathematical mind.

I would like to thank all the staff of the Maxwell Institute Graduate School for their enormous support throughout these years.

Special thanks go to the University of Edinburgh. Your vibrant and innovative environment is truly inspirational for all the young people studying here.

I would like to thank the healthcare professionals and essential workers, your professionalism and sacrifice helped to protect the whole planet from the Covid-19 pandemic.

Within the last 4 years, I made a lot of friends at the department and outside of it. I would like to particularly thank Jean, Arman, Dominic, Domenico, Spyros and Sergey for great time together and all the wonderful conversations we held.

I wish to thank all my Russian friends who have given me unforgettable help and encouragement during my studies.

Special thanks go to my family. To my grandmother, Nelya, for always being calm and understanding, our conversations cheered me up a lot. To my father, Dmitry, your critical and broad-minded academic approach set a stellar example for me. To my mother, Irina, for her eagerness to help and constantly pushing me forward. To my sister, Varvara, for always standing at my side and for all the tenderness during our calls. To my family-in-law, Natalia, Boris, Valentina, Vyacheslav and Igor for their tremendous endorsement and support.

Finally, huge thanks to my amazing wife, Alexandra, for all your love, patience, and unconditional belief in me. Without you, I would not have coped with this journey.

Special *post scriptum* acknowledgements go to Victor, your appearance to this world spans the new chapters of our life, way beyond those written ones.

Contents

Part I

Introduction

Chapter 1

Measure flows

1.1 Motivation and background

The last century's development of statistical physics gave inspiration for the new modelling techniques in the field of Partial Differential Equations and Stochastics. One of the examples of the latter is the so-called mean-field or McKean-Vlasov Stochastic Differential Equations, usually denoted by

$$dX_t = b(t, X_t, \mu_t)\mathrm{d}t + \sigma(t, X_t, \mu_t)\mathrm{d}W_t, \quad \mu_t = \text{Law}(X_t), \quad X_0 = \xi, \tag{1.1.1}$$

where W is a standard Brownian motion, ξ is a random variable independent of W and, b and σ are measurable functions for some stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$.

Here the measure flow $t \mapsto \mu_t$ for any $t \in [0, T]$ belongs to the space $\mathcal{P}_2(\mathbb{R}^d)$ of probability measures with finite second moments. This space is Polish (complete and separable) under the Wasserstein distance

$$W_2(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \Big(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \pi(dx, dy) \Big)^{\frac{1}{2}}, \quad \mu, \nu \in \mathcal{P}_2(\mathbb{R}^d),$$

where $\Pi(\mu, \nu)$ is the set of couplings for μ and ν such that $\pi \in \Pi(\mu, \nu)$ is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ such that $\pi(\cdot \times \mathbb{R}^d) = \mu$ and $\pi(\mathbb{R}^d \times \cdot) = \nu$. We refer to the space $\mathcal{P}_2(\mathbb{R}^d)$ as the 2-Wassertein space.

Equation (1.1.1) postulates the dynamics of the stochastic process X. For smooth functional $u : \mathbb{R} \to \mathbb{R}$ the application of the classical Itô formula gives the dynamics of $t \mapsto u(X_t)$ in terms of differential of X. One may ask on what is the expansion for $t \mapsto u(\mu_t)$, where $u : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R}$. The explicit dependence of the coefficients on the timed measure flow does not allow to apply classical Itô formula and requires to differentiate with respect to measure leading to the necessity of building calculus on the Wassertein space.

Lions derivative

To consider a variational calculus for functional of measure one requires to build a suitable differentiation operator on the 2-Wasserstein space. Among the several notions of differentiability of a functional u defined over $\mathcal{P}_2(\mathbb{R}^d)$ we follow the approach introduced by Lions in their lectures at Collège de France [76] and further developed in [18]. A comprehensive presentation can be found in the joint monograph of Carmona and Delarue [21],[22].

We consider the canonical lifting of the function $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ to $\tilde{u}$ defined as $\tilde{u} : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \ni X \to \tilde{u}(X) = u(Law(X)) \in \mathbb{R}$, where $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ is a space of square integrable random variables. We can say that u is L-differentiable at μ, if $\tilde{u}$ is Frechèt differentiable (in L^2) at some X, such that $\mu = \mathbb{P} \circ X^{(-1)}$. Denoting the gradient by $D\tilde{u}$ and using a Hilbert structure of the L^2 space, we can identify $D\tilde{u}$ as an element of its dual, L^2 itself. It was shown in [18, Theorem 6.2] that $D\tilde{u}(X)$ is a $\sigma(X)$-measurable random variable that depends only upon the μ and not upon the particular random variable X with law μ. Now

$$
\begin{tikzcd}
& u(\mu) \arrow[rr, "\textit{Lifting}"] \arrow[dd, dotted] & & \tilde{u}(X) \arrow[d, "\textit{Frechét}"] & \\
\boxed{\mathcal{P}_2(\mathbb{R}^d)} & & & D\tilde{u}(X) \arrow[d, "\textit{Structure ([18])}"] & \boxed{L_2(\Omega,\mathcal{F};\mathbb{R}^d)} \\
& \partial_\mu u(\mu,\cdot) & & Du(\mu,\cdot) \arrow[ll, "\textit{Definition}"] &
\end{tikzcd}
$$

Figure 1.1.1: Schematic construction of the Lions derivative

Theorem 6.5 from [18] proves the existence of a function $\partial_\mu u(\mu,\cdot) \in L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu; \mathbb{R}^d)$ such that $\partial_\mu u(\mu, X) = Du(X)$. Hereinafter, we define the L-derivative of u at μ as the map $\partial_\mu u(\mu,\cdot) : \mathbb{R}^d \ni v \to \partial_\mu u(\mu, v) \in \mathbb{R}^d$, satisfying $D\tilde{u}(X) = \partial_\mu u(\mu, X)$. We refer to Figure 1.1.1 for scheme of the construction.

For *deterministic* functionals of measures, extending the classical Itô formula to the so-called Itô-Lions formula, there are several approaches and results available in the literature. The classical difference of increments approach is used in [15] under a strong regularity assumption of existence of second order Fréchet derivatives. In [26] an approach using projections over empirical measures is used allowing for minimal regularity assumptions. Both approaches are neatly reviewed in [21, Chapter 5]. Linked to the existence of a regular solution to the master equation for mean-field games with common noise is the approach by [20, Appendix 6]. Their proof is carried out using Itô-Taylor type expansions (similar to [15]) and requiring the involved maps to be twice Fréchet differentiable. Lastly, another approach is to use a semi-group type approach to describe the flow of measures and obtain the necessary infinitesimal expansions see [19, Appendix A]. More recently [25] present such Itô-Lions formula for maps belonging to Sobolev spaces, [55, 87] provide also such formula for semi-martingales – these three works leave out the conditional measure-flow case.

Itô-Wentzell formula

The extension of the celebrated Itô chain rule from deterministic regular functions to random fields of Itô type was proposed originally by Wentzell [90] and later generalised in [47, 67, 69, 70, 84]. This successful result has appeared in Stochastic Partial Differential Equations problems from wellposedness to numerical methods and applications to fluid dynamics modelling [24, 32, 47, 53, 84], in stochastic regularisation problems [39], filtering [68], and mathematical Finance [2, 43, 77].

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a stochastic basis satisfying usual conditions. We first introduce the stochastic process $(X_t)_{t\in[0,T]}$ satisfying the dynamics

$$
dX_t = \beta_t \mathrm{d}t + \gamma_t \mathrm{d}W_t, \quad \text{and initial condition } X_0, \tag{1.1.2}
$$

where W is a d-dimensional Brownian motion. The involved parameters satisfy the next condition.

Assumption 1.1.1. *Let $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R})$ (X_0 is $\mathcal{F}_0$-measurable and independent of W_t, $t \in [0,T]$). Take $\beta : \Omega \times [0,T] \to \mathbb{R}^d$ and $\gamma : \Omega \times [0,T] \to \mathbb{R}^{d\times d}$ such that $(\beta_t)_{t\in[0,T]}, (\gamma_t)_{t\in[0,T]}$ are $\mathbb{F}$-progressively measurable processes and satisfy*

$$
\int_0^T (|\beta_s| + |\gamma_s|^2) ds < \infty, \quad \mathbb{P}\text{-a.s.}.
$$

We recall the Itô-Wentzell formula in the style of [19, Section A.3.1].

Theorem 1.1.2 (Itô-Wentzell). *Take $(X_t)_{t\in[0,T]}$ given by* (1.1.2) *under Assumption 1.1.1. Let a map $V : \Omega \times [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}$ be such that:*

i) Fix $x \in \mathbb{R}^d$, $(V_t(x))_{t\in[0,T]}$ is a continuous adapted process taking values in $\mathbb{R}$;

ii) Fix $t \in [0,T]$, $\omega \in \Omega$, $\mathbb{R}^d \ni x \mapsto V_t(x)$ is a $\mathcal{C}^2(\mathbb{R}^d)$-mapping with values in $\mathbb{R}$;

iii) $(V_t(x))_{t\in[0,T]}$, $x \in \mathbb{R}^d$ is a random field that admits the Itô dynamics

$$V_t(x) = V_0(x) + \int_0^t \phi_s(x)ds + \int_0^t \psi_s(x) \cdot dW_s, \quad t \in [0,T],$$

where $(\phi_t(\cdot))_{t\in[0,T]}$ and $(\psi_t(\cdot))_{t\in[0,T]}$ are $\mathbb{F}$-progressively measurable processes with values in $\mathcal{C}^2(\mathbb{R}^d,\mathbb{R})$ and $\mathcal{C}^2(\mathbb{R}^d,\mathbb{R}^d)$ respectively, such that for any compact $K \subset \mathbb{R}^d$

$$\int_0^T \big(\|\phi_s(\cdot)\|_{\mathcal{C}^1(K)} + \|\psi_s(\cdot)\|^2_{\mathcal{C}^2(K)} \big)ds < \infty \quad \mathbb{P}\text{-a.s.}$$

The maps ϕ, ψ are called the characteristics of the random field V.

Then $(V_t(X_t))_{t\in[0,T]}$ is an Itô process and it satisfies $\mathbb{P}$-a.s. the following expansion

$$\begin{aligned} V_T(X_T) - V_0(X_0) = &\int_0^T \phi_s(X_s)ds + \int_0^T \psi_s(X_s) \cdot dW_s + \int_0^T \partial_x V_s(X_s) \cdot \beta_s ds \\ &+ \int_0^T \partial_x V_s(X_s) \cdot \gamma_s dW_s + \int_0^T \frac{1}{2} \mathit{Trace}\big\{\partial^2_{xx} V_s(X_s)\, \gamma_s(\gamma_s)^\intercal\big\}ds \\ &+ \int_0^T \mathit{Trace}\big\{\partial_x \psi_s(X_s)(\gamma_s)^\intercal\big\}ds. \end{aligned}$$

The first two terms correspond to dynamics of the field $V_t(\cdot)$ within installed X_t-trajectories. The subsequent three terms correspond to the usual Itô formula. The last term is a cross-variation of the diffusion coefficient of the process with the same noise induced by the stochastic field $V_t(\cdot)$.

Proof. In this formulation, we state conditions on the differentiability of ϕ, ψ directly as opposed to the original formulation by [70, Theorem 3.3.1] where conditions over the characteristics of the driving semimartingale were given; [70, Exercise 3.1.5] closes the gap. □

Imagine now that V, ϕ and ψ are measure functionals, i.e., $V, \phi, \psi : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ and satisfy for some $\mathcal{P}_2(\mathbb{R}^d)$

$$V_t(\mu) = V_0(\mu) + \int_0^t \phi_s(\mu)\mathrm{d}s + \int_0^t \psi_s(\mu) \cdot \mathrm{d}W_s, \quad t \in [0,T].$$

Take X satisfying equation (1.1.2), and $\mu_t := \text{Law}(X_t)$. Consider the map $t \mapsto V_t(X_t)$ for any $t \in [0,T]$.

We ask the following questions:

- How would $V_t(\mu_t)$ expand? In other words, what is the corresponding Itô-Wentzell formula for the functional dependent on measure flow?
- How would a joint (in space and measure) chain rule look like?

Considering further two independent $\mathcal{F}$-Brownian motions $W^0,\ W^1$, let the random field V and its characteristics $\phi,\ \psi$ satisfy the expansion

$$V_t(\mu) = V_0(\mu) + \int_0^t \phi_s(\mu)\mathrm{d}s + \int_0^t \psi^0_s(\mu) \cdot \mathrm{d}W^0_s + \int_0^t \psi^1_s(\mu) \cdot \mathrm{d}W^1_s, \quad t \in [0,T].$$

Now let the process $(X_t)_{t\in[0,T]}$ satisfy

$$dX_t = \beta_t \mathrm{d}t + \gamma^0_t \mathrm{d}W^0_t + \gamma^0_t \mathrm{d}W^0_t, \quad \text{and initial condition } X_0.$$

We view W^0 and W^1 as *common noise* and *idiosyncratic noise* respectively. Let μ_t a law of X_t conditioned on common noise W^0 (we formally define μ_t in Section 4.2). Consider the map $t \mapsto V_t(\mu_t)$.

We ask the follow-up questions:

- What is the expansion of $V_t(\mu_t)$ under a common noise setting?
- In this case, how would a joint (in space and measure) chain rule look like?

We shall answer these questions in Part II.

1.2 Methodological perspective

To prove the chain rule for $t \mapsto V_t(\mu_t)$ as describe above we combine two techniques. The first one is the projection over empirical measures approach of [21, 26], which has the benefit of yielding lower regularity requirements on the underlying coefficients. The other approach is to use Taylor-like expansions in the vein of [69] – we argue next that this is the suitable methodology for this result.

The chain rule in the measure component first appears in [15] making use of the telescopic summation technique and building on a strong assumption of a second order Fréchet differentiability of the lifting map. To overcome the requirement of a second Fréchet derivative and reduce it to just first order Fréchet derivative (in fact the so-called *Partial-$\mathcal{C}^2$* regularity) for full measure case, the approach of empirical projection was introduced [21, 22, 26]: this is the approach we follow. In [19] the Itô-Wentzell-Lions formula is shown under the constant diffusion of the random field. The authors follow the semi-group approach and require the existence of the density.

Recently, [55] introduced the use of cylindrical polynomials approximation to build a measure chain rule for the measure flow of semimartingales, i.e., an Itô-Lions formula for semimartingales. Finally, [87] shows an Itô-Lions formula for semimartingales with jumps (the exact same result of [55]) but using the mechanisms of [15]. Concretely, they make use of a telescopic summation technique building on the functional linear derivative instead of the Lions one. This approach relies on the assumptions of growth and boundedness of the functional linear derivative and its partial derivative with respect to new spatial variable. For both [55, 87] the conditional measure flow case is left unaddressed.

We already argued that neither the proof techniques of [19] or [15] are appropriate as proofs for our results. The former requires constant diffusion coefficients to ensure existence of densities while the latter requires higher Fréchet regularity than needed. Hence the reason we follow [21, 22, 26]. Two recent works [55, 87], posterior to ours, use new techniques to prove the Itô-Lions formula for general semimartingales (*for deterministic fields*) — it is not clear if those techniques can be adapted to prove the Itô-Wentzell-Lions formulae we present in this manuscript under the same minimal regularity constraints we impose. The difficulty stems from our use of random fields while in [55, 87] the fields are deterministic.

Concretely, to prove the Itô-Wentzell-Lions formula of this manuscript with the same methodology of [55] one would require a Leibniz rule to interchange the Fréchet derivative symbol with the stochastic integral one and thus would demand further regularity assumptions on top of the existing ones – a general Leibniz rule within this framework is presently an open question[1]. Moreover, such a result is not needed in [55, 87] due to their use of deterministic fields! The telescopic summation approach from [87] has another limitation in the context of proving an Itô-Wentzell-Lions formula. One needs to expand the local difference of integrands by the application of the classical Itô-Wentzell formula which requires the existence and well-definiteness of the random fields spanned by the differentiation in measure (in sense of Lions or linear functional; see our Theorem 1.1.2). This limitation could be avoided by proving the aforementioned Leibniz rule which in turn would demand stronger regularity for the random field and its characteristics as mentioned earlier. For these reasons, we argue that the 'empirical projection' technique [21, 22, 26] is the suitable methodology.

[1] Preliminary work on exchanging Fréchet derivatives with the Lebesgue's integral has been carried in a note by O. Kammar [63]. It is unclear presently under which minimal conditions can one exchange Fréchet derivatives (and later the Lions derivatives) with the stochastic (Itô) integral.

1.3 Contribution

In Chapter 4 we propose an Itô-Wentzell formulae for random fields that embed measure-functionals in a way that is amenable to an analysis in the sense of Lions derivatives. We establish two formulae, and two further corollaries, all decoupled from the applications either in mean-field game theory in finance [21, 22, 55, 87], fluid mechanics [13, 59, 65], neuroscience modeling [45], population dynamics models [7] or further related stochastic analysis problems [34, 71] albeit motivated by them.

To the best of our knowledge we have found only one Itô-Wentzell-Lions type formula in the literature, [19, Appendix A]. Their approach is set in relation to an existing regular solution to a certain master equation for mean-field control games with common noise. Their proof is carried out via expansions of the densities of the underlying (conditional) measure flow but where the involved diffusion components are constants.

Our first result is for the full flow of measures (the measure is deterministic) while the second is for a partial flow of measures (the measure is random). Each result is then further extended to a full joint chain rule allowing for the an additional driving stochastic processes $(X_t)_{t\geqslant 0}$ having a semi-martingale expansion. In particular, we recover the results in [19, Appendix A] while finessing their assumptions, see our Remark 4.1.5 below. A by-product of our results is a clarification on the necessity of the assumptions on the classical Itô-Wentzell formula [70, Theorem 3.3.1] (see our Theorem 1.1.2 below). Namely, we prove that one can require one order of regularity less from the drift and diffusion coefficient of the random vector field to which the Itô-Wentzell formula is applied to (see our Theorem 3.2.1). This smaller result is of its own independent interest.

The usefullness of these result is manyfold. Direct applications within mean-field optimal control could be envisaged in neuroscience modeling [45]; extending the contribution of [41, 77], where the classical Itô-Wentzell formula is used to develop a consistent forward utilities of investment and consumption – introducing the relative performance concerns (see 2). Also building from [7], where a mean-field games with Fisher-Wright common noise is discussed. This model is used in the evolution of population genetics and where it would be natural to update the model to support the distributional component, making use of the results we provide in order to establish the verification procedure. In fluid dynamics these formulae would allow to expand the dynamics of driving signals against the underlying vector field [13, 59, 65].

Lastly, our work can be extended in several directions to include anticipative processes [83], general semimartingale dynamics [55, 87], path dependent functionals in combination with functional Itô calculus [29], or extensions to K-forms for SPDEs in fluid dynamics [32].

Chapter 2

Financial games

2.1 Motivation and background

Consider the market with one riskless asset and n risky securities which serve as proxies for two distinct asset classes. The price of risky asset is assumed to be of log-normal type, driven by two independent Brownian motions. Precisely, the price $(S_t^i)_{t\geqslant 0}$ of stock i traded exclusively by the i-th agent solves

$$\frac{dS_t^i}{S_t^i} = \breve{\mu}_i \mathrm{d}t + \nu_i \mathrm{d}W_t^i + \sigma_i \mathrm{d}B_t, \quad S_0^i = s_0^i > 0, \tag{2.1.1}$$

with constant market parameters $\breve{\mu}_i \in \mathbb{R},\ \sigma_i \geqslant 0$ and $\nu_i \geqslant 0$ with $\sigma_i + \nu_i > 0$. We refer the reader to [72, 73] for an in-depth motivation of the model. The one-dimensional standard Brownian motions $B, W^1, \ldots, W^n$ are independent. When $\sigma_i > 0$, the process B induces a correlation between the stocks, and thus we call B the *common noise* and W^i an *idiosyncratic noise*. The independent Brownian motions $B, W^1, \ldots, W^n$ are defined on a probability space $(\Omega, \mathbb{F}, \mathcal{F}, \mathbb{P})$ endowed with the natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}$ generated by them and satisfies the usual conditions. We recall the case of *single common stock*, where for any $i = 1, \ldots, n,\ (\breve{\mu}_i, \sigma_i) = (\breve{\mu}, \sigma),\ \nu_i = 0$, for some $\breve{\mu} \in \mathbb{R},\ \sigma > 0$ and independent of i[1].

All the agents follow one of the two different risk preferences frameworks: *constant absolute risk aversion (CARA)* or *constant relevant risk aversion (CRRA)*. These regimes are characterised by the choice of underlying utility function (exponential utility for CARA and power utility for CRRA). For clarity of presentation we explicitly define them in Chapter 6 and Chapter 7 (for CARA and CRRA respectively).

Under each of these frameworks we differently define the wealth process X^i of manager i and their strategies. Namely, for CARA preferences each agent $i = 1, \ldots, n$ trades using a self-financing strategy, $(\pi_t^i)_{t\geqslant 0}$, representing the *amount* of wealth to invest in i-th stock. At the same time under CRRA preferences each agent $i = 1, \ldots, n$ forms their portfolio by having $(\pi_t^i)_{t\geqslant 0}$ – *fraction* of wealth invested in the i-th stock and consumption policy, $(c_t^i)_{t\geqslant 0}$, representing the instantaneous rate of consumption per unit of wealth. These differences are reflected in the corresponding dynamics of the wealth processes, given by Equation 5.2.1 and 5.4.2 respectively.

In contrast with the standard Merton portfolio problem [78], each agent seeks to maximise utility not purely their own wealth, but their *relative performance* $\hat{X}$, given by

$$\text{(CARA)} \qquad \widehat{X}^i = X^i - \theta_i \overline{X}^{(-i)}, \qquad \text{where} \qquad \overline{X}^{(-i)} = \frac{1}{n-1}\sum_{k\neq i}^{n} X^k,$$

$$\text{(CRRA)} \qquad \widehat{X}^i = \frac{X^i}{\left(\widetilde{X}^{(-i)}\right)^{\theta_i}}, \qquad \text{where} \qquad \widetilde{X}^{(-i)} = \Big(\prod_{k\neq i}^{n} X^k\Big)^{\frac{1}{n-1}},$$

[1] For the sake of simplicity we work with one-dimensional stocks S^i, however the generalisation to the vector of stocks available to agent i is straightforward. Furthermore, the single stock framework will correspond to the realistic situation of a large number of agents trading in the same vector of stocks.

where θ_i reflects the *competition preference* of manager i. For CRRA preferences we define the relative consumption $\hat{c}^{(-i)}$ in a similar way.

The quantities $\overline{X}^{(-i)}$ and $\widetilde{X}^{(-i)}$ capture the average wealth of a population of managers in a arithmetic and geometric way respectively. The agent is not purely concerned about their own wealth (and consumption for CRRA), but compares it with the same of the others through the lens of their competition tolerance.

Benchmarking is critical for fund managers who need to keep the fund competitive and attractive for future investments. We refer to the excellent economic and finance motivations found in [3] and also [11, 72, 73]. In this work, we build on the structure proposed firstly in [46] and then in [48] and [11, 33, 72, 73]. Additionally, we point the reader to the beautiful introductions of [72, 73] who brought those concepts to the framework of mean field games. Further, those works also make for an excellent review of mean field games in the context of the Merton problem. Investment under performance concerns has been taken up in many variants, most recently: [33] analyse a partial information finite-player CARA utility game with performance concerns employing Forward Backward SDE (FBSDE) machinery (market model with stochastic coefficients), see [33, Table 1]. [49] works in the same context but addresses the mean field game in the full non-Markovian framework using mean field FBSDEs, see also [49, Table 1]. Aspects of investment under relative consumption are much less explored, only in [72].

Throughout we consider fund managers that trade between their individual stock and common riskless asset, the so-called problem of *asset specialisation*. This problem was addressed initially by [14, 78] and further developed in [3, 61, 72, 73, 89]. These works report on asset specialisation for a variety of reasons: asset familiarity, trading costs and constraints, reduction of learning costs or industry specialisation. Competition games, of finite-player or mean field type, in the context of asset specialisation has received much attention recently and we defer to [3, 36, 72, 73] for an overview that relates to our context. For a long-view of mean field games theory and its applications we refer to the monographs [23].

2.2 Mathematical framework

At the core of our work, are the ideas behind the *forward investment performance criteria* introduced by Musiela and Zariphopoulou [81] and [57] as a way to solve portfolio optimisation problems without the specific drawbacks of the classical utility theory. When entering the market, investors prescribe their risk profile at horizon time and therefore cannot adapt it to changes in market conditions or update risk preferences; additionally, the investment time horizon is fixed, and the portfolio is derived in respect to this temporal reference point. The forward criteria (portfolio) optimisation problem is formulated as a maximisation of the conditional expectation of a stochastic utility function across all points in time. Hence, time-horizons are arbitrary which is more realistic for practitioners. In essence, the forward *dynamic* performance map is built with the property of dynamic consistency taken as the starting point.

We recall the concept of the forward performance processes (FPP) from [79] (for simplicity we present the FPP from the perspective of wealth optimisation, the wealth-consumption FPP criteria is set out further in Chapter 5).

Definition 2.2.1 (Forward performance process). *Let X^α is a stochastic process governed by admissible control α. Let $U : \Omega \times \mathbb{R} \times [0, \infty) \to \mathbb{R}$ be an $\mathbb{F}$-progressively measurable random field. U is a* forward performance process *if*

- *For all $t \geqslant 0$ the map $x \mapsto U(x, t)$ is $\mathbb{P}$-a.s. increasing and concave;*
- *It satisfies $U(x, 0) = u_0(x)$;*
- *For all $T \geqslant t$ and each strategy, represented by α, the associated process X^α satisfies a supermartingale property*

$$\mathbb{E}[U(X_T^\alpha, T)|\mathcal{F}_t] \leqslant U(X_t^\alpha, t) \quad \mathbb{P}\text{-a.s.}$$

- *For all $T \geqslant t$ there exists a strategy, represented by α^*, for which the associated discounted wealth X^* satisfies a martingale property*

$$\mathbb{E}[U(X^*_T, T)|\mathcal{F}_t] = U(X^*_t, t) \quad \mathbb{P}\text{-}a.s.$$

The Forward performance processes capture the time evolution of such stochastic utility functions and since then much progress has been done in characterising such processes in a variety of settings and assumptions [10, 28, 41–43, 62, 75, 93]. For excellent literature overviews on the developments of FPP we point the reader to [4, 27, 62, 77]. Particularly clear is [62] who provides a near taxonomic breakdown of the literature of the forward criteria with respect to their behaviour in time along with approaches used. Here, as in [62] or [4], we focus on the class of forward investment-consumption criteria with the specific property that they are differentiable in time (a dynamics without volatility). In [4] the authors study an optimal portfolio selection problem (no consumption, no competition) under the FPP of power form in an incomplete market. Their arguments make use of HJB machinery and, like in this work or [28], an assumption of "separable power factor form" for the FPP is made (Assumption 7.3.3 below). Lastly, a general form of the utility dynamics is made with a non-zero diffusion component and the driving market model is fully stochastic multi-factor one coined "EVE correlation" models. The study of consumption with FPPs (or forward utility) was considered in [10, 28, 41]; interesting is [28] who manage the difficulty associated to having a FPP dynamics featuring a volatility component by assuming the volatility to be exogenously postulated by the agent as a market perception. Their FPP consumption optimisation problem yields non-uniqueness.

Competition within the framework of FPP is in its infancy. The concept originated in [52] for a two-player game in the CRRA context with a Merton market model, and its vision is expanded upon in [3], still within a two-player game but allowing for random coefficients in an incomplete market model. Inspired by [52] and [73], the formulation of *mean field forward performance games* and the concept of *mean field forward equilibrium* was proposed by [36]. From a competition perspective, the FPP approach reflects that agents need not all have the same time-horizons since under asset specialisation different industries have different timeframes (although quarterly reporting are common points of reference), see [3] for a full discussion.

The FPP framework is not the only theory seeking to overcome limitations of the classical utility theory. An alternative from the 90s are the so-called Epstein-Zin preferences within the recursive utility framework [44, 92] expanding on the theoretical framework of Kreps-Porteus [66].

One of cornerstones of the Epstein-Zin model was the idea of disentagling agent's risk tolerance δ and *Elasticity of Intertemporal Substitution* (EIS) defined as

$$\text{EIS} := -\frac{d\big(\partial_t c_t / c_t\big)}{d\big(\partial_t V_x(c_t z, t) / V_x(c_t z, t)\big)}.$$

for some $z > 0$[2]. The classic CRRA utility optimisation yields a strict relation between the δ and EIS, namely that $\text{EIS}^{\text{classic, [72]}} = \delta$ (e.g., see [72]). As discussed in [88] regarding EIS, higher interest rates increase the overall wealth of the consumer due to higher cash-flows in future periods. The effect that consumers spend a part of this higher future income already today is called "income effect". On the other hand, with higher interest rates a smaller fraction of today's consumption has to be saved in order to have an additional unit of consumption tomorrow. This motivation to save more today and postpone today's consumption is called the "substitution effect". Consumers with a high EIS are more willing to substitute consumption over time, which has a direct impact on the "substitution effect". Now, classic theory yields $\text{EIS}^{\text{classic, [72]}} = \delta$, however, the risk tolerance is atemporal relating how a consumer substitutes consumption across different states of the world while EIS is intertemporal relating how a consumer substitutes consumption between now and later [31]. Thus classic utility framework cannot capture how agent competition with or without performance concerns changes EIS.

Time-continuous stochastic preferences capturing "intertemporal substitution" of Hindy-Huang-Kreps [58] have been used to understand consumption optimisation [9] and Arrow-

[2]Here V can be seen as utility from consumption, and z is some fixed level of wealth.

Debreu equilibria [6]. More recently, [1] revisited the EIS discussion using [40]'s stochastic recursive utility in continuous time specified to the Kreps-Porteus [66] family. Forthcoming is [12] who recast the MFG of [73] under the Epstein-Zin/Kreps-Porteus recursive utility framework. Nonetheless, inspecting this reference one sees several of the criticisms aimed at the standard utility framework appearing again (see [3]): [16] emphasises the dependence on investment's horizon ("The investor's horizon also plays a crucial role in optimal policies") and that the underlying model is fixed throughout the investment time frame. In a nutshell, the FPP works with the forward point of view to investment (risk profile is prescribed at $t = 0$) while the recursive utility still works within the backward one (risk profile is prescribed at a horizon time $T > 0$.

We study many-player games of investment-consumption optimisation under the forward performance framework in a Merton market model featuring common-noise. In each game, the agents trade in a specific stock affected by a common market signal and seek to optimise their wealth and consumption process while having relative concerns towards the average wealth and consumption of the other agents. The core features of our model are (a) relative consumption concerns, (b) relative wealth concerns, (c) asset specialisation, (d) game competition (finite-player and mean field games) and (e) the forward performance processes (FPP) view. The former (a)–(d) have been analysed in [72] through the lens of the classical utility framework. In this work we subscribe to the elegantly argued (and supported by empirical evidence) FPP paradigm of [3] to view the earlier work [72] through its light. The arguments we make using FPPs, competition and the presence of consumption make the analysis involved and reveal elements not present in [72] (despite the similarity to (a)–(d)) or [64].

We ask the following question:

- How to establish a n-player and mean-field game-theoretical framework under CARA and CRRA FPP preferences in a Merton market model under common noise? What is a Nash equilibrium and MF-equilibrium? How to derive them?

CARA preferences are more accessible for derivation of optimality principles. We additionally investigate the wealth-consumption optimisation of CRRA framework by asking with the following questions:

- Under CRRA preferences how to derive both the utilities of investment and consumption from the single PDE? What assumption to make?
- How to interpret the results?

We shall answer these questions in Part III.

2.3 Contribution

We contribute to the literature on mean field games and FPP by also providing an explicitly solvable example. Outside linear-quadratic structures it is very uncommon, as argued in [72], and this is one of these rarities we bring here. We work with the very tractable model (2.1.1) and include common noise, heterogeneous agents, a mean field interaction through the controls in addition to the state processes and FPPs.

With this work, we investigate an n-player and mean field game for asset specialising agents who optimise investment-consumption (or investment only) under with relative-wealth and relative-consumption concerns through the lens of the FPP framework. The tractability of our setting yields findings not seen in the classic utility theory. Namely, our contributions are:

(I) We expand the concept of CARA (investment only) and CRRA FPP to a game-theoretical framework both for finite-player games and mean field games within the common-noise Merton market model. The MFG approach is based on the simpler concept work [73] for CARA FPP (wealth optimisation only), and [72] for the classical CRRA investment-consumption utility problem. For both games, we provide explicit expressions for the quantities of interest. From a methodological point of view, we solve the control problem by combining convex duality with HJB-type arguments.

(II) We introduce the dynamic model selection procedure under CARA preferences: the player can dynamically update the market model and their risk and competition parameters. We show that the utility coming from the new investments horizon is contributing towards total utility in a multiplicative way, while the latter accumulates the optimisation results from the past horizons.

(III) With the inclusion of consumption, the CRRA FPP wealth-consumption problem does not have a unique solution but a class of them indexed by a certain parameter $\kappa \in \mathbb{R}$ interpreted as *market-risk relative consumption preference*. This parameter is present in the single-agent optimisation (no competition/performance concerns) and both games settings. Under the competition environment it is common to all players and reflects how the environment weights the utility from consumption in respect to the one from wealth. The tractability of our setting, exposes the κ parameter to clearly enable an extra layer of interpretable market modelling (see Section 7.4) and, critically, when $\kappa = 0$ we recover classic utility theory results [72] also [64].

(IV) We show that without competition the CRRA FPP framework encapsulates the same key feature of the Epstein-Zin preferences recursive utility. It breaks the strict relation between risk tolerance δ and *Elasticity of Intertemporal Substitution* (EIS) of the classical utility theory, and at the same time keeps its core features of independence from the investment time-horizon and flexibility towards updating risk preferences. Namely, $\text{EIS}^{\text{no competition}(\theta=0)} = \kappa\delta$, with κ spanning another dimension of the agent's risk preference.

(V) We show that performance concerns ($\theta \in (0, 1]$) re-scale the agent's perception of consumption. Namely, for $t \geqslant 0$

$$\text{EIS}_t^{\text{with competition}(\theta\neq 0)} = P_t \times \text{EIS}^{\text{no competition}(\theta=0)}, \quad \text{where} \quad \text{EIS}^{\text{no competition}(\theta=0)} = \kappa\delta,$$

where P_t is a random stochastic process depending on risk-competition preferences, the market-risk relative consumption preference κ (that is uniform for the entire population of competitors) and equilibrium consumption. The time-dependence of P is intimately related to the use of the FPP framework with competition and discussed in detail in Section 7.4 (see Equation (7.4.2)). Similar results using Epstein-Zin's recursive utility exist [12].

Part II

Measure flows

Chapter 3

Preliminaries

3.1 Notation and Spaces

Let $\mathbb{N}$ be the set of natural numbers starting at 1, $\mathbb{R}$ denotes the real numbers. For collections of vectors in $\{x^l\}_l \in \mathbb{R}^d$, let the upper index l denote the distinct vectors, whereas the lower index the vector components, i.e. $x^l = (x^l_1, \dots, x^l_d) \in \mathbb{R}^d$ namely x^l_j denotes the j-th component of l-th vector. For $x, y \in \mathbb{R}^d$ denote the scalar product by $x \cdot y = \sum_{j=1}^d x_j y_j$; and $|x| = (\sum_{j=1}^d x_j^2)^{1/2}$ the usual Euclidean distance; and $x \otimes y$ denotes the tensor product of vectors $x, y \in \mathbb{R}^d$. Let $\mathbb{1}_A$ be the indicator function of set $A \subset \mathbb{R}^d$. For a matrix $A \in \mathbb{R}^{d\times n}$ we denote by $A^\intercal$ its transpose and its Frobenius norm by $|A| = \text{Trace}\{AA^\intercal\}^{1/2}$. Let $I_d : \mathbb{R}^d \to \mathbb{R}^d$ be the identity map.

We denote by $\mathcal{C}(A, B)$ for $A, B \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, the space of continuous functions $f : A \to B$. In terms of derivative operators and differentiable functions, ∂_t denotes the partial differential in the time parameter $t \in [0, T]$; ∂_x denotes the gradient operators in the spatial variables x in $\mathbb{R}^d$ while $\partial^2_{xx}, \partial^2_{yy}$ the Hessian operator in x or $y \in \mathbb{R}^d$.

For $p, d, m \in \mathbb{N}$ denote $\mathcal{C}^p(\mathbb{R}^d, \mathbb{R}^m)$ the space of p-times continuously differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}^m$. The space $\mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^m)$ is equipped with a collection of seminorms $\{\|g\|_{\mathcal{C}^1(K)} := \sup_{x\in K}(|g(x)| + |\partial_x g(x)|),\ g \in \mathcal{C}^p(\mathbb{R}^d)\}$, indexed by the compact subsets $K \subset \mathbb{R}^d$. The space $\mathcal{C}^2(\mathbb{R}^d)$ is equipped with a collection of seminorms $\{\|g\|_{\mathcal{C}^2(K)} := \sup_{x\in K}(|g(x)| + |\partial_x g(x)| + |\partial^2_{xx} g(x)|), g \in \mathcal{C}^p(\mathbb{R}^d)\}$, indexed by the compact subsets $K \subset \mathbb{R}^d$; we refer to $C^{1,2} = C^{1,2}([0, T]\times \mathbb{R}^d, \mathbb{R}^m)$ as the usual space of maps $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}^m$ that are once continuously differentiable in the first variable, twice so in the second variable (as in $C^2(\mathbb{R}^d, \mathbb{R}^m)$) and jointly continuous across the several derivatives.

We say that the function is locally bounded, when its restriction to every compact set is bounded.

Spaces

We recall the Wassertein space from Subsection 1.1.

Throughout set some $0 < T < +\infty$ and we work the finite time interval $[0, T]$. Let our probability space be a completion of $(\Omega, \mathbb{F}, \mathcal{F}, \mathbb{P})$ with $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ carrying a d-dimensional Brownian motion $W = (W_1, \dots, W_d)$ generating the probability space's filtration, augmented by all $\mathbb{P}$-null sets, and with an additionally sufficiently rich sub σ-algebra $\mathcal{F}_0$ independent of W. Let our probability space be an atomless Polish. We denote by $\mathbb{E}[\ \cdot\ ] = \mathbb{E}^{\mathbb{P}}[\ \cdot\ ]$ the usual expectation operator with respect to $\mathbb{P}$.

We adopt the following convention, that for d-dimensional random vector $X = (X_1, \dots, X_d)$ we understand denote $\mathbb{E}[X]$ by the d-dimensional vector $(\mathbb{E}[X_1], \dots, \mathbb{E}[X_d])$. The convenience of this notation will become apparent in the later Section 4.2.

We define $L^2(\Omega, \mathcal{F}_0, \mathbb{P}, \mathbb{R}^d)$ as the space of $\mathcal{F}_0$-measurable random variables $\xi : \Omega \to \mathbb{R}^d$ that are square integrable $\mathbb{E}^{\mathbb{P}}[|\xi|^2] < \infty$. Given two processes $(X_t)_{t\in[0,T]}$ and $(Y_t)_{t\in[0,T]}$ let $\langle X, Y\rangle_t$ or $\langle X_t, Y_t\rangle$ denote its' cross-variation up to time $t \in [0, T]$.

Lastly, for convenience we choose to work over 1-, d- and $d \times d$-dimensional spaces. This is particularly helpful in lowering complexity of the presentation of the later sections where many sequences of approximating vector-valued stochastic processes are pushed through the Itô and Itô-Wentzell formula. At the same time we emphasise that generalisation of our results to arbitrary dimensions is straightforward.

3.2 The Itô-Wentzell formula (classic)

We recall the Itô-Wentzell formula introduced in 1.1.

A close inspection of Theorem 1.1.2 and its proof ([69],[70]) reveals that the theorem holds under reduced regularity requirements. We explore this observation with our next result.

Theorem 3.2.1 (Itô-Wentzell under reduced regularity). *The conclusion of Theorem 1.1.2 still holds for $(V_t(X_t))_{t\in[0,T]}$ if in condition iii) the constraints on ϕ, ψ are replaced by:*

> $(\phi_t(\cdot))_{t\in[0,T]}$ *are* $(\psi_t(\cdot))_{t\in[0,T]}$ $\mathbb{F}$*-progressively measurable processes with values on the spaces* $\mathcal{C}^0(\mathbb{R}^d,\mathbb{R})$ *and* $\mathcal{C}^1(\mathbb{R}^d,\mathbb{R}^d)$ *respectively, such that for any compact* $K \subset \mathbb{R}^d$
>
> $$\int_0^T \big(\|\phi_s(\cdot)\|_{\mathcal{C}^0(K)} + \|\psi_s(\cdot)\|^2_{\mathcal{C}^1(K)} \big) ds < \infty \quad \mathbb{P}\text{-a.s.} \tag{3.2.1}$$

Proof. The arguments we use are classical. We mollify V, ϕ, ψ in their spatial components by convolution with a smoothing kernel and obtain a sequence (V^n, ϕ^n, ψ^n), $n \in \mathbb{N}$, such that for each $n \in \mathbb{N}$ $(\phi^n_t(\cdot))_{t\in[0,T]}$ are $(\psi^n_t(\cdot))_{t\in[0,T]}$ $\mathbb{F}$-progressively measurable processes with values in $\mathcal{C}^1(\mathbb{R}^d,\mathbb{R})$ and $\mathcal{C}^2(\mathbb{R}^d,\mathbb{R}^d)$ respectively (in fact even more due to the mollification), such that for any compact $K \subset \mathbb{R}^d$, $\mathbb{P}$-a.s.

$$\int_0^T \big(\|\phi^n_s(\cdot)\|_{\mathcal{C}^1(K)} + \|\psi^n_s(\cdot)\|^2_{\mathcal{C}^2(K)} \big) \mathrm{d}s + \sup_{\hat{n}\in\mathbb{N}} \int_0^T \big(\|\phi^{\hat{n}}_s(\cdot)\|_{\mathcal{C}^0(K)} + \|\psi^{\hat{n}}_s(\cdot)\|^2_{\mathcal{C}^1(K)} \big) \mathrm{d}s < \infty. \tag{3.2.2}$$

Lastly, $\mathbb{P}$-a.s. for $t \in [0,T]$ a.e. we have that $\phi^n_t, \psi^n_t, \partial_x \psi^n_t$ converge to $\phi_t, \psi_t, \partial_x \psi_t$ uniformly (in n) on compact sets. It is clear that V^n retains the properties of V, uniformly over n for the 0-th, 1-st and 2-nd derivative. In particular, $\mathbb{P}$-a.s. for any $t \in [0,T]$ $V^n_t, \partial_x V^n_t, \partial^2_{xx} V^n_t$ converge to $V_t, \partial_x V_t, \partial^2_{xx} V_t$ uniformly on compact sets. We conclude via Theorem 1.1.2 that $(V^n_t(X_t))_{t\in[0,T]}$ is an Itô process satisfying the expansion given. The passage to the limit as $n \to \infty$ is also argued in a classical way. First we make use of a localizing sequence $(\tau^m)_{m\in\mathbb{N}}$ over X defined as $\tau^m := \inf\{t > 0 : |X_t| > m\}$, $m \in \mathbb{N}$ which in turn allows us to make use of the uniform convergence over compacts for the maps' sequence (in n) and (3.2.1)-(3.2.2) repeatedly, i.e. we can assume that X is bounded. Arguing convergence of the Lebesgue integrals follows via continuity of the maps, integrability of the coefficients (see Assumption 3.4.1) and dominated convergence theorem taking advantage of uniform convergence over compacts given that X is assumed to take values in a bounded set. The stochastic integral terms requires an additional argument which we provide for the 2nd integral (the 1st is handled similarly),

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T} \Big| \int_0^t \partial_x V^n_s(X_s)\cdot\gamma_s \mathrm{d}W_s - \int_0^t \partial_x V_s(X_s)\cdot\gamma_s \mathrm{d}W_s \Big|^2 \Big] \\ \leqslant \mathbb{E}\Big[\int_0^T \big| \partial_x V^n_s(X_s) - \partial_x V_s(X_s) \big|^2 |\gamma_s|^2 \mathrm{d}s \Big].\end{aligned}$$

Since $\partial_x V, \partial_x V^n$ are jointly continuous in their variables and converge uniformly over compacts, X is assumed to take values in a bounded set and γ satisfies Assumption 3.4.1, then the RHS converges to zero as $n \to \infty$. $\square$

3.3 The Lions derivative

3.3.1 The Lions derivative and notational conventions

We recall the notation of L-derivative introduced in Subsection 1.1.

We always denote $\partial_\mu u$ as the version of the L-derivative that is continuous in the product topology of all components of u. Moreover, let ∂_μ^2 denote second derivative in measure and $\partial_v \partial_\mu u$ denote the derivative with respect to new variable arisen after applying derivative in measure. The notion of ∂_μ^2 is chosen in favour of $\partial_{\mu\mu}^2$, as the latter may be hinting at the linear nature of L-derivative, that is not the case at all.

When we do the lift $\tilde{\xi}$ and $\hat{\xi}$ are the lifted random variables defined over the twin stochastic spaces $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ respectively, having the same law μ. We form a new probability space $(\Omega, \mathcal{F}, \mathbb{P}) \times (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and consider random variables $\tilde{\xi}(\omega, \tilde{\omega}) = \xi(\tilde{\omega})$. Since this procedure is valid for the stochastic processes on respective stochastic bases $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{t\in[0,T]}, \tilde{\mathbb{P}})$ and $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}} = (\hat{\mathcal{F}}_t)_{t\in[0,T]}, \hat{\mathbb{P}})$, one can consider $(X_t, \tilde{X}_t, \hat{X}_t)$ as a triple of independent identically distributed processes. The same applies to a finite amount of copy spaces $(\Omega^l, \mathcal{F}^l, \mathbb{F}^l = (\mathcal{F}_t^l)_{t\in[0,T]}, \mathbb{P}^l), 1 \leqslant l \leqslant N \in \mathbb{N}$ to form a new product space and the respective tuple $(X_t, \tilde{X}_t, \hat{X}_t, X_t^1, \dots, X_t^N)$ remains mutually independent.

We will add the bases $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ and $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{\mathbb{P}})$ and further use them as an environment for model representatives of the mean-field (each living in the distinct respective space), whereas sampling from the mean-field will give us N particles living within respective spaces $(\Omega^l, \mathcal{F}^l, \mathbb{F}^l = (\mathcal{F}_t^l)_{t\in[0,T]}, \mathbb{P}^l),\ 1 \leqslant l \leqslant N$, to be used within the propagation of chaos procedures below. Hereinafter $\tilde{\mathbb{E}}$ denotes the expectation acting on the model twin space $\tilde{\Omega}$.

Over the present work we omit the re-notation after adding some new probability spaces, but will assume that adding a copy processes automatically intimates the procedure described above. The common noise setting given in Section 4 requires a slightly variation of this approach which we disclose in the proof of Theorem 4.2.2.

3.3.2 Regularity in the measure argument

In this section we recall several spaces of measure-regularity arising in the literature on Wasserstein calculus.

Definition 3.3.1. *We say the functional $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is Fully $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ if*

i) u is L-differentiable at every point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, and $\partial_\mu u : \mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d \to \mathbb{R}$ is joint-continuous at every pair $(\mu, v) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$;

ii) For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the map $v \mapsto \partial_\mu u(\mu, v) \in \mathbb{R}^d$ is $\mathbb{R}^d$-differentiable at every point $v \in \mathbb{R}^d$; and $\partial_v\partial_\mu u : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is joint-continuous at every pair $(\mu, v) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$;

iii) For any $v \in \mathbb{R}^d$, the map $\mu \mapsto \partial_\mu u(\mu, v) :\in \mathbb{R}^d$ is L-differentiable at every point $\mu \in \mathbb{R}^d$, and $\partial_\mu^2 u : \mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$ is joint-continuous at every triple $(\mu, v, v') \in \mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\times\mathbb{R}^d$.

We next restrict the regularity with respect to the space variable arising after taking measure derivative to the Supp(μ), since in our probabilistic setting the process sitting there obviously will not escape this set. This restriction comes from the interplay with the *Partial-$\mathcal{C}^2$*-regularity of [21, Chapter 5.6.4].

Definition 3.3.2. *We say the function $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is Partially $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}))$ if*

i) u is L-differentiable at every point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, such that $\partial_\mu u$ is locally bounded and joint-continuous at every pair $(\mu, v), \mu \in \mathcal{P}_2(\mathbb{R}^d),\ v \in$ Supp(μ);

ii) For any $v \in \mathbb{R}^d$, the map $\mathbb{R}^d \ni v \mapsto \partial_\mu u(\mu, v) \in \mathbb{R}$ is $\mathbb{R}^d$-differentiable at every point $v \in$ Supp(μ). Moreover, $\partial_v\partial_\mu u : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is locally bounded and joint-continuous at every pair $(\mu, v),\ \mu \in \mathcal{P}_2(\mathbb{R}^d),\ v \in$ Supp(μ).

This regularity level does not require a second Frechét derivative of the lift to exist. Looking ahead, we do not expect to receive any second-order terms in the expansion of the measure component, hence it is quite essential not to demand such a regularity (see Theorem 4.1.8 or Theorem 4.1.3 below).

For the purpose of Theorem 3.4.4 we require the regularity in all components, and we introduce the following definition.

Definition 3.3.3. *A function $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is $\mathcal{C}^{1,2,(1,1)}$ if*

i) For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ the map $[0,T] \times \mathbb{R}^d \ni (t,x) \mapsto u_t(x,\mu)$ is $\mathcal{C}^{1,2}$, and the maps $\partial_t u$, $\partial_x u$ and $\partial^2_{xx} u$ are joint-continuous at every triple $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$;

ii) For any $(t,x) \in [0,T] \times \mathbb{R}^d$, the map $\mu \mapsto u_t(x,\mu)$ is continuously L-differentiable at every point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\partial_\mu u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ is joint-continuous and locally bounded at every quadruple (t,x,μ,v), with $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d),\ v \in$ Supp(μ);

iii) For any $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, the map $v \mapsto \partial_\mu u_t(x,\mu,v)$ is continuously $\mathbb{R}^d$-differentiable at every point $v \in \mathbb{R}^d$. Moreover, its derivative $\partial_v \partial_\mu u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is continuous and locally bounded at every quadruple (t,x,μ,v), with $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d),\ v \in$ Supp(μ).

3.3.3 The Empirical projection map

We recall the concept of *empirical projection map* given in [26] which will be one of the main workhorses throughout our work.

Definition 3.3.4 (Empirical projection of a map)**.** *Given $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ and $N \in \mathbb{N}$, define the empirical projection u^N of u via $u^N : (\mathbb{R}^d)^N \to \mathbb{R}$, such that*

$$u^N(x^1,\dots,x^N) := u\big(\bar{\mu}^N\big), \quad \textit{with} \quad \bar{\mu}^N := \frac{1}{N}\sum_{l=1}^N \delta_{x^l} \quad \textit{and} \quad x^l \in \mathbb{R}^d,\ l = 1,\dots,N.$$

We recall [21, Proposition 5.91 and Proposition 5.35] which relates the spatial derivative of u^N with the L-derivative of u.

Proposition 3.3.5. *Let $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be Fully-$\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$, then, for any $N > 1$, the empirical projection u^N is $\mathcal{C}^2$ on $(\mathbb{R}^d)^N$ and for all $x^1,\dots,x^N \in \mathbb{R}^d$ we have the following differentiation rules*

$$\partial_{x^j} u^N(x^1,\dots,x^N) = \frac{1}{N}\,\partial_\mu u\Big(\frac{1}{N}\sum_{l=1}^N \delta_{x^l}, x^j\Big),$$
$$\partial_{x^k}\partial_{x^j} u^N(x^1,\dots,x^N) = \frac{1}{N}\,\partial_v\partial_\mu u\Big(\frac{1}{N}\sum_{l=1}^N \delta_{x^l}, x^j\Big)\mathbb{1}_{j=k} + \frac{1}{N^2}\,\partial^2_\mu u\Big(\frac{1}{N}\sum_{l=1}^N \delta_{x^l}, x^j, x^k\Big).$$

3.4 Itô-Lions chain rule along a full flow of measures (classic)

For the clarity of writing we again present the process $(X_t)_{t\in[0,T]}$ from Section 1.1 satisfying the dynamics

$$dX_t = \beta_t \mathrm{d}t + \gamma_t \mathrm{d}W_t, \quad \text{and initial condition } X_0, \tag{3.4.1}$$

where W is a d-dimensional Brownian motion. The involved parameters satisfy the next condition.

Assumption 3.4.1. *Let $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R})$ (X_0 is $\mathcal{F}_0$-measurable and independent of $W_t,\ t \in [0,T]$). Take $\beta : \Omega \times [0,T] \to \mathbb{R}^d$ and $\gamma : \Omega \times [0,T] \to \mathbb{R}^{d\times d}$ such that $(\beta_t)_{t\in[0,T]}, (\gamma_t)_{t\in[0,T]}$ are $\mathbb{F}$-progressively measurable processes and satisfy*

$$\int_0^T (|\beta_s| + |\gamma_s|^2) ds < \infty, \quad \mathbb{P}\text{-a.s.}.$$

Alongside $(X_t)_{t\in[0,T]}$ we introduce another process $(Y_t)_{t\in[0,T]}$ and its law $(\mu_t)_{t\in[0,T]}$. Take W as a d-dimensional Brownian motion and let $(Y_t)_{t\in[0,T]}$ satisfy the dynamics

$$dY_t = b_t \mathrm{d}t + \sigma_t \mathrm{d}W_t, \quad \text{and initial condition } Y_0, \tag{3.4.2}$$

where we denote the law of Y_t by $\mu_t := \mathbb{P} \circ Y_t^{(-1)},\ t \in [0,T]$ and the associated coefficients satisfy the below assumption.

Assumption 3.4.2. *Let $Y_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P})$ (Y_0 is $\mathcal{F}_0$-measurable and independent of $W_t,\ t \in [0,T]$). Take $b : \Omega \times [0,T] \to \mathbb{R}^d$ and $\sigma : \Omega \times [0,T] \to \mathbb{R}^{d\times d}$ such that $(b_t)_{t\in[0,T]}, (\sigma_t)_{t\in[0,T]}$ are $\mathbb{F}$-progressively measurable processes and satisfy*

$$\mathbb{E}\Big[\int_0^T |b_s|^2 + |\sigma_s|^4 ds\Big] < \infty.$$

Remark 3.4.3. *One can take "closed-loop" type dependence for the coefficients, i.e. coefficients of the form $\hat{b}_t := b_t(Y_t, \mu_t)$ and $\hat{\sigma}_t := \sigma_t(Y_t, \mu_t)$, since our setting covers all the special cases. In fact, an existence & uniqueness result for the SDE for Y allows to freeze the components inside the coefficients and with sufficient integrability the "frozen" SDE follows the dynamics* (3.4.2).

For completeness we recall the Itô-Lions formula [21, Proposition 5.102] for deterministic maps following the framework Section 3.3.1, recall that $\tilde{\mathbb{E}}$ denotes the expectation acting on the model twin space $(\tilde{\Omega}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ and let the processes $(\tilde{Y}_t, \tilde{b}_t, \tilde{\sigma}_t)_{t\in[0,T]}$ be the twin processes of $(Y_t, b_t, \sigma_t)_{t\in[0,T]}$ respectively living within.

Theorem 3.4.4. *Let $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be $\mathcal{C}^{1,2,(1,1)}$. Furthermore, for any compact $K \subset \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ we have*

$$\sup_{(t,x,\mu)\in[0,T]\times K} \Big\{ \int_{\mathbb{R}^d} \Big[|\partial_\mu u_t(x,\mu,v)|^2 + |\partial_v \partial_\mu u_t(x,\mu,v)|^2 \Big] \mu(dv) \Big\} < \infty, \quad \mathbb{P}\text{-a.s.}.$$

Take $(X_t)_{t\in[0,T]}$ given by (3.4.1) *under Assumption 3.4.1 and take μ associated to* (3.4.2) *under Assumption 3.4.2. Then $(u_t(X_t, \mu_t))_{t\in[0,T]}$ is an Itô process satisfying $\mathbb{P}$-a.s.*

$$\begin{aligned}
u_T(X_T, \mu_T) - u_0(X_0, \mu_0)
&= \int_0^T \partial_t u_s(X_s, \mu_s) ds + \int_0^T \Big[\partial_x u_s(X_s, \mu_s) \cdot \beta_s ds + \int_0^T \partial_x u_s(X_s, \mu_s) \cdot \gamma_s dW_s \\
&\quad + \int_0^T \frac{1}{2}\, \mathit{Trace}\big\{ \partial_{xx}^2 u_s(X_s, \mu_s)\ \gamma_s (\gamma_s)^\intercal \big\} ds \\
&\quad + \int_0^T \tilde{\mathbb{E}}\Big[\partial_\mu u_s(X_s, \mu_s, \tilde{Y}_s) \cdot \tilde{b}_s \Big] ds + \int_0^T \frac{1}{2}\, \tilde{\mathbb{E}}\Big[\mathit{Trace}\big\{ \partial_v \partial_\mu u_s(X_s, \mu_s, \tilde{Y}_s)\ \tilde{\sigma}_s (\tilde{\sigma}_s)^\intercal \big\} \Big] ds.
\end{aligned}$$

3.5 Probability space under conditional measure flows

Now we introduce the construction of the probability space in presence of the common noise to be used in Section 4.2.

We consider $(\Omega^0, \mathcal{F}^0, \mathbb{F}^0 = (\mathcal{F}_t^0)_{t\in[0,T]}, \mathbb{P}^0)$ and $(\Omega^1, \mathcal{F}^1, \mathbb{F}^1 = (\mathcal{F}_t^1)_{t\in[0,T]}, \mathbb{P}^1)$ atomless Polish probability spaces to be the respective completions of $(\Omega^0, \mathbb{F}^0, \mathbb{P}^0)$ and $(\Omega^1, \mathbb{F}^1, \mathbb{P}^1)$ carrying a respective d-dimensional Brownian motions $W^0 = (W_t^0)_{t\in[0,T]}$ and $W^1 = (W_t^1)_{t\in[0,T]}$ generating the probability space's filtration, augmented by all $\mathbb{P}^0$- and $\mathbb{P}^1$-null sets respectively. We augment $(\Omega^0, \mathcal{F}^0, \mathbb{F}^0 = (\mathcal{F}^0_t)_{t\in[0,T]}, \mathbb{P}^0)$ with a sufficiently rich sub σ-algebra $\mathcal{F}^0_0$ independent of W^0 and W^1. We denote by $(\Omega, \mathbb{F}, \mathbb{P})$ the completion of the product space $(\Omega^0 \times \Omega^1, \mathbb{F}^0 \otimes \mathbb{F}^1, \mathbb{P}^0 \otimes \mathbb{P}^1)$ equipped with the filtration $\mathbb{F}$ obtained by augmenting the product filtration $\mathbb{F}^0 \otimes \mathbb{F}^1$ in a right-continuous way and by completing it. We let $\mathbb{E}^0$ and $\mathbb{E}^1$ taking the expectation on the first and second space respectively. We adopt the following convention, that for d-dimensional random vector $Y = (Y_1, \dots, Y_d)$ we denote $\mathbb{E}[Y]$ by the d-dimensional vector $(\mathbb{E}[Y_1], \dots, \mathbb{E}[Y_d])$.

We define $L^2(\Omega, \mathcal{F}_0, \mathbb{P}, \mathbb{R}^d)$ as the space of $\mathcal{F}_0$-measurable random variables $\xi : \Omega \to \mathbb{R}^d$ that are square integrable $\mathbb{E}^{\mathbb{P}}[|\xi|^2] < \infty$.

3.6 Itô-Lions chain rule along a conditional flow of measures (classic)

Take measurable $(b, \sigma^0, \sigma^1) : \Omega \times [0,T] \to \mathbb{R}^d \times \mathbb{R}^{d\times d} \times \mathbb{R}^{d\times d}$ and define the following process

$$dY_t = b_t \mathrm{d}t + \sigma^0_t \mathrm{d}W^0_t + \sigma^1_t \mathrm{d}W^1_t, \text{ and initial condition } Y_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}), \tag{3.6.1}$$

and $\mu_t := \mathrm{Law}(Y_t(\omega_0, \cdot))$ for $\mathbb{P}^0$-almost any ω_0. Here $\mathrm{Law}(Y_t(\omega_0, \cdot))$ can be understood as RV from $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$ into $\mathcal{P}(\mathbb{R}^d)$ (for further details see discussion in [22, Section 4.3]).

Moreover, the involved coefficients satisfy the next conditions

Assumption 3.6.1. *$(Y_t)_{t\in[0,T]}$ satisfies Assumption 3.4.2 with $\sigma_t := \begin{pmatrix} \sigma^0_t & 0 \\ 0 & \sigma^1_t \end{pmatrix}$ and $W_t := (W^0_t, W^1_t)^\intercal$.*

Take $(X_t)_{t\in[0,T]}$ satisfying dynamics

$$dX_t = \beta_t \mathrm{d}t + \gamma^0_t \mathrm{d}W^0_t + \gamma^1_t \mathrm{d}W^1_t, \text{ and initial condition } X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}), \tag{3.6.2}$$

with coefficients satisfying

Assumption 3.6.2. *$(X_t)_{t\in[0,T]}$ satisfies Assumption 3.4.1 with $\gamma_t := \begin{pmatrix} \gamma^0_t & 0 \\ 0 & \gamma^1_t \end{pmatrix}$ and $W_t := (W^0_t, W^1_t)^\intercal$.*

We name $(W^0_t)_{t\in[0,T]}$ as a common noise affecting the whole setting, whilst $(W^1_t)_{t\in[0,T]}$ is the idiosyncratic chaos for the random field and all processes within. For the purposes of Section 4.2 we fix the common noise and derive the dynamics of the random field by conditioning on W^0.

We recall the Itô-Lions formula for the flow of marginals [22, Theorem 4.17].

First, we provide the regularity assumption as given in [22, Subsection 4.3.4].

Definition 3.6.3. *A function $u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is $\mathcal{C}^{1,2,(2)}$ if*

i) For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ the map $[0,T] \times \mathbb{R}^d \ni (t,x) \mapsto u_t(x,\mu)$ is $\mathcal{C}^{1,2}$, and the maps $\partial_t u$, $\partial_x u$ and $\partial^2_{xx} u$ are joint-continuous at every triple $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$;

ii) For any $(t,x) \in [0,T] \times \mathbb{R}^d$, the map $\mu \mapsto u_t(x,\mu)$ is continuously L-differentiable at every point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\partial_\mu u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ is joint-continuous and locally bounded at every quadruple (t,x,μ,v), with $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, $v \in$ Supp(μ);

iii) For any $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, the map $v \mapsto \partial_\mu u_t(x,\mu,v)$ is continuously $\mathbb{R}^d$-differentiable at every point $v \in \mathbb{R}^d$. Moreover, its derivative $\partial_v \partial_\mu u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is continuous and locally bounded at every quadruple (t,x,μ,v), with $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, $v \in$ Supp(μ);

iv) For any $(t,x,\mu,v) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times$ Supp(μ), the map $\nu \mapsto \partial_\mu u_t(x,\mu,v)$ is continuously L-differentiable at every point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, its derivative $\partial^2_\mu u : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is continuous and locally bounded at every quintuple (t,x,μ,v,v'), with $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, $v, v' \in$ Supp(μ);

v) *For any* $(t, x, \mu, v) \in [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times Supp(\nu)$*, the map* $x \mapsto \partial_\mu u_t(x,\mu,v)$ *is continuously* $\mathbb{R}^d$*-differentiable at every point* $x \in \mathbb{R}^d$*. Moreover, its derivative* $\partial_x\partial_\mu u : [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$ *is continuous and locally bounded at every quadruple* (t, x, μ, v)*, with* $(t, x, \mu) \in [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$*,* $v \in Supp(\mu)$*;*

Theorem 3.6.4. *Let* $u : [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ *be* $\mathcal{C}^{1,2,(2)}$*. Furthermore for any compact* $K \subset \mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ *we have*

$$\sup_{(t,x,\mu)\in[0,T]\times K}\left\{\int_{\mathbb{R}^d}\Big[|\partial_\mu u_t(x,\mu,v)|^2+|\partial_v\partial_\mu u_t(x,\mu,v)|^2+|\partial_x\partial_\mu u_t(x,\mu,v)|^2\Big]\mu(dv)\right.$$
$$\left.+\int_{\mathbb{R}^d\times\mathbb{R}^d}\Big[|\partial^2_\mu u_t(x,\mu,v,v')|^2\Big]\mu(dv)\mu(dv')\right\}<\infty,\quad \mathbb{P}\text{-a.s.}.$$

Take $(\mu_t)_{t\in[0,T]}$ *associated to* (3.6.1) *under Assumption 3.6.1. Take* $(X_t)_{t\in[0,T]}$ *to be a* d*-dimensional Itô process with dynamics* (3.6.2) *satisfying Assumption 3.6.2.*

Then $(u_t(X_t,\mu_t))_{t\in[0,T]}$ *is an Itô process satisfying* $\mathbb{P}$*-a.s.*

$$\begin{aligned}
&u_T(X_T,\mu_T)-u_0(X_0,\mu_0)\\
&\quad=\int_0^T\partial_t u_s(X_s,\mu_s)ds+\int_0^T\partial_x u_s(X_s,\mu_s)\cdot\beta_s ds+\int_0^T\partial_x u_s(X_s,\mu_s)\cdot\gamma^0_s dW^0_s\\
&\quad+\int_0^T\partial_x u_s(X_s,\mu_s)\cdot\gamma^1_s dW^1_s+\int_0^T\frac{1}{2}\,Trace\big\{\partial^2_{xx}u_s(X_s,\mu_s)(\gamma^0_s(\gamma^0_s)^\intercal+\gamma^1_s(\gamma^1_s)^\intercal)\big\}ds\\
&\quad+\int_0^T\tilde{\mathbb{E}}^1\Big[\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)\cdot\tilde{b}_s\Big]ds+\int_0^T\tilde{\mathbb{E}}^1\Big[(\tilde{\sigma}^0_s)^\intercal\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)\Big]\cdot dW^0_s\\
&\quad+\int_0^T\frac{1}{2}\,\tilde{\mathbb{E}}^1\Big[Trace\big\{\partial_v\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)(\tilde{\sigma}^0_s(\tilde{\sigma}^0_s)^\intercal+\tilde{\sigma}^1_s(\tilde{\sigma}^1_s)^\intercal)\big\}\Big]ds\\
&\quad+\int_0^T\tilde{\mathbb{E}}^1\Big[Trace\big\{\partial_x\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)\;\gamma^1_s(\tilde{\sigma}^1_s)^\intercal\big\}\Big]ds\\
&\quad+\int_0^T\frac{1}{2}\,\hat{\mathbb{E}}^1\Big[\tilde{\mathbb{E}}^1\Big[Trace\big\{\partial^2_\mu u_s(X_s,\mu_s,\tilde{Y}_s,\hat{Y}_s)\;\tilde{\sigma}^0_s(\hat{\sigma}^0_s)^\intercal\big\}\Big]\Big]ds
\end{aligned}$$

where $\tilde{\mathbb{E}}$ *denotes the expectation acting on the model twin spaces* $(\tilde{\Omega},\tilde{\mathbb{F}},\tilde{\mathbb{P}})$ *and* $(\hat{\Omega},\hat{\mathbb{F}},\hat{\mathbb{P}})$ *and let the processes* $(\tilde{Y}_t,\tilde{b}_t,\tilde{\sigma}_t)_{t\in[0,T]}$ *and* $(\hat{Y}_t,\hat{b}_t,\hat{\sigma}_t)_{t\in[0,T]}$ *be the twin processes of* $(Y_t,b_t,\sigma_t)_{t\in[0,T]}$ *respectively living within.*

Remark 3.6.5 (On measurability)**.** *The measurability of the measure expansion component is deeply discussed in* [21, Remarks 5.101 and 5.103]. *Within the present work we are interested in conditioning on the field noise, the matter of which is discussed in* [22, Section 4.3]. *We refer the reader to this monograph for comprehensive and detailed approach.*

Chapter 4

Itô-Wentzell-Lions formula

4.1 Itô-Wentzell-Lions chain rule with a full flow of measures

As it was shown in [26], one can apply an approach based on empirical projections to built the chain rule. This approach is very convenient since with it we are able require (loosely) the same regularity as in Theorem 3.4.4 above. One can notice that the second measure derivative term of the formulae appearing within measure argument expansion vanishes when applying the limit procedure. Nonetheless, in order to argue via Taylor expansions the second derivative in measure has to exist which is a very strong assumptions. We can avoid this requirement using this technique.

Let $u : \Omega \times [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a random field, satisfying the expansion

$$du_t(x,\mu) = \phi_t(x,\mu)\mathrm{d}t + \psi_t(x,\mu) \cdot \mathrm{d}W_t, \quad u_0(x,\mu) = f(x,\mu), \tag{4.1.1}$$

where $f(x,\mu) : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a deterministic function, $(W_t)_{t\in[0,T]}$ is a d-dimensional $\mathbb{F}$-Brownian motion, $(\phi,\psi) : \Omega \times [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \times \mathbb{R}^d$ are $\mathbb{F}$-progressively measurable processes.

Throughout we will work with the law $(\mu_t)_{t\in[0,T]}$ of the process $(Y_t)_{t\in[0,T]}$ given in (3.4.2) under Assumption 3.4.2. In the second portion of the section, we additionally work with $(X_t)_{t\in[0,T]}$ solution to (3.4.1) under under Assumption 3.4.1.

4.1.1 Itô-Wentzell-Lions formula for measure functionals

We start by discussing the measurability of the involved structures and for which the following remark addresses the issue for the whole manuscript.

Itô-Wentzell expansion

In this subsection we work with the Itô random field (4.1.1) and we keep $x \in \mathbb{R}^d$ at some fixed value for the whole subsection and hereinafter we will omit its presence within u, ϕ and ψ, i.e. we set

$$(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \;\; u_t(\mu) := u_t(x,\mu), \quad \phi_t(\mu) := \phi_t(x,\mu), \quad \text{and} \quad \psi_t(\mu) := \psi_t(x,\mu).$$

Similarly to the full- and partial-$\mathcal{C}^2$ maps concept in Definition 3.3.1 and 3.3.2, we introduce the concept of a *partially-$\mathcal{C}^2$ Itô random field*, describing the field's regularity in the measure component and we coin it *RF-Partially $\mathcal{C}^2$*.

Definition 4.1.1. *We say the random field $u : \Omega \times [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ given in* (4.1.1) *(for some $x \in \mathbb{R}^d$ fixed) is* RF-Partially-$\mathcal{C}^2$ *if*

i) For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $(u_t(\mu))_{t\in[0,T]}$ is a continuous adapted process taking values over $\mathbb{R}^d$ and $(\phi_t(\mu))_{t\in[0,T]}, (\psi_t(\mu))_{t\in[0,T]}$ are $\mathbb{F}$-progressively measurable processes with values in $\mathbb{R}$ and $\mathbb{R}^d$ respectively;

ii) For almost all $t \in [0,T]$, the maps $\mu \mapsto \phi_t(\mu),\ \mu \mapsto \psi_t(\mu)$ are $\mathbb{P}$-a.s. continuous in the topology induced by the Wasserstein metric for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$;

iii) For any $t \in [0,T]$ the map $\mu \mapsto u_t(\mu)$ is $\mathbb{P}$-a.s. continuous in topology, induced by Wasserstein metric and L-differentiable $\mathbb{P}$-a.s. at every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\partial_\mu u_t(\mu, v)$ is $\mathbb{P}$-a.s. joint-continuous at every triple (t,μ,v) with $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d),\ v \in Supp(\mu),\ \mathbb{P}$-a.s.;

iv) For any $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d)$ the map $v \mapsto \partial_\mu u_t(\mu, v)$ is $\mathbb{R}^d$-differentiable $\mathbb{P}$-a.s. at every $v \in Supp(\mu)$. Moreover, the map $\partial_v \partial_\mu u_t(\mu, v)$ is $\mathbb{P}$-a.s. joint-continuous at every triple (t,μ,v), with $(t,\mu) \in [0,T] \times \mathcal{P}_2(\mathbb{R}^d),\ v \in Supp(\mu),\ \mathbb{P}$-a.s..

Remark 4.1.2. *In contrast with* [21,22]*, where the local boundedness condition is present in the regularity conditions, we restrict ourselves to the continuous version of the Lions derivative from the beginning, hence local boundedness is automatically implied by the continuity.*

The main proof mechanics relies on the projection over empirical distributions technique as explored in [21,26]. Recall that $\tilde{\mathbb{E}}$ denotes the expectation acting on the model twin space $(\tilde{\Omega}, \tilde{\mathbb{F}}, \tilde{\mathbb{P}})$ and let the processes $(\tilde{Y}_t, \tilde{b}_t, \tilde{\sigma}_t)_{t\in[0,T]}$ be the twin processes of $(Y_t, b_t, \sigma_t)_{t\in[0,T]}$ respectively living within (see Section 3.3.1).

Theorem 4.1.3. *Let u be the* RF-Partially-$\mathcal{C}^2$ *Itô random field* (4.1.1) *(where $x \in \mathbb{R}^d$ is fixed and omitted throughout, also for ϕ and ψ). Assume for any compact $K \subset \mathcal{P}_2(\mathbb{R}^d)$ and for any $t \in [0,T]$ that*

$$\int_0^t \sup_{\mu \in K} \Big\{ |\phi_s(\mu)| + |\psi_s(\mu)|^2 \Big\} ds < \infty, \quad \mathbb{P}\text{-a.s.},$$

and

$$\sup_{(t,\mu)\in[0,T]\times K} \Big\{ \int_{\mathbb{R}^d} \Big[|\partial_\mu u_t(\mu, v)|^2 + |\partial_v \partial_\mu u_t(\mu, v)|^2 \Big] \mu(dv) \Big\} < \infty, \quad \mathbb{P}\text{-a.s.}. \tag{4.1.2}$$

Let $(\mu_t)_{t\in[0,T]}$ be the law of the solution to (3.4.2) *satisfying Assumption 3.4.2. Then $(u_t(\mu_t))_{t\in[0,T]}$ is an Itô process $\mathbb{P}$-a.s. satisfying the expansion*

$$\begin{aligned} u_T(\mu_T) - u_0(\mu_0) = \int_0^T \phi_s(\mu_s) ds + \int_0^T \psi_s(\mu_s) \cdot dW_s + \int_0^T \tilde{\mathbb{E}}\Big[\partial_\mu u_s(\mu_s, \tilde{Y}_s) \cdot \tilde{b}_s \Big] ds \\ + \int_0^T \frac{1}{2}\, \tilde{\mathbb{E}}\Big[\mathit{Trace}\big\{ \partial_v \partial_\mu u_s(\mu_s, \tilde{Y}_s)\ \tilde{\sigma_s}(\tilde{\sigma}_s)^\intercal \big\} \Big] ds. \end{aligned}$$

Remark 4.1.4. *Following from Theorem 3.4.4 we have that for fixed $r \in [0,T],\ t \mapsto u(r, \mu_t)$ $\mathbb{P}$-a.s. satisfies the expansion*

$$\begin{aligned} u_r(\mu_T) - u_r(\mu_0) = \int_0^T \tilde{\mathbb{E}}\Big[\partial_\mu u_r(\mu_s, \tilde{Y}_s) \cdot \tilde{b}_s \Big] ds \\ + \int_0^T \frac{1}{2}\, \tilde{\mathbb{E}}\Big[\mathit{Trace}\big\{ \partial_v \partial_\mu u_r(\mu_s, \tilde{Y}_s)\ \tilde{\sigma_s}(\tilde{\sigma}_s)^\intercal \big\} \Big] ds. \end{aligned}$$

Remark 4.1.5. *We note that our result is a generalisation of the Itô-Wentzell formula of* [19, Appendix A]*, where in contrast with our setting, the involved diffusion components are constants. Furthermore, we highlight the requirement of the square integrability on $\partial_\mu u$ and $\partial_v \partial_\mu u$ in* (4.1.2) *which is not present in* [19, Appendix A]. *The requirement is necessary for the intermediary step of W_2-convergence of the empirical measure appearing in those terms.*

Remark 4.1.6. *Here we write Trace within last term assuming the symmetry of respective matrix holding $\mathbb{P}$-a.s. for any $t \in [0,T]$. One can see that within the approximating procedure, i.e. the distance between the Hessian of the mollified empirical projection and the $\partial_v \partial_\mu u$-term is controlled through the decreasing sequence $\varepsilon_N \searrow 0$, thus the symmetry follows by approximation. See* [21, Remark 5.98] *for details.*

Proof of Theorem 4.1.3. For this proof we follow as guideline the proof of Theorem 5.99 in [21]. Let throughout $t \in [0, T]$. Recall that $\mathbb{E}^{1,\dots,N}$ denotes an expectation with respect to the product of sample twin spaces $(\Omega^1, \mathbb{F}^1, \mathbb{P}^1) \times \dots \times (\Omega^N, \mathbb{F}^N, \mathbb{P}^N)$. We again underline that we act on an atomless Polish space.

Step 1: Mollification & compactification. If the desired expansion holds true for any u - *RF-Partially* $\mathcal{C}^2$, bounded and uniformly continuous (in space and measure arguments), then the formula holds for u satisfying the conditions of the theorem. This fact is straightforward by applying a two-step mollification procedure in the vein of [21, Theorem 5.99] and which we introduce next.

Defining for any $t \in [0, T]$ the $(u \star \rho)_t(\mu) := u_t(\mu \circ \rho^{-1})$ with $\rho : \mathbb{R}^d \to \mathbb{R}^d$ smooth function with compact support, the $\mathbb{P}$-a.s. boundedness of $(u\star\rho)_t(\mu), \partial_\mu(u\star\rho)_t(\mu, v)$ and $\partial_v\partial_\mu(u\star\rho)_t(\mu, v)$ follows from $\mathbb{P}$-a.s. local boundedness of $u_t, \partial_\mu u_t(\mu, v)$ and $\partial_v\partial_\mu u_t(\mu, v)$ respectively. We also notice that $\partial_\mu(u \star \rho)_t(\mu, v)$ and $\partial_v\partial_\mu(u \star \rho)_t(\mu, v)$ are $\mathbb{P}$-a.s. joint-continuous in every triple $t \in [0, T], \mu \in \mathcal{P}^2(\mathbb{R}^d), v \in \text{Supp}(\mu)$. In order to obtain continuity over the whole space we smooth out the distribution by convolution with a Gaussian density, i.e. considering $\mu \mapsto (u \star \rho)(\mu * \phi_G)$ instead of $\mu \mapsto (u \star \rho)(\mu)$ with ϕ_G - density of standard d-dimensional Gaussian distribution $N(0, I_d)$ on $\mathbb{R}^d$ and $(\mu*\phi_G)(x) := \int_{\mathbb{R}^d} \phi_G(x-y)d\mu(y)$. Now the support of $\mu*\phi_G$ is the whole $\mathbb{R}^d$ and $\partial_\mu(u\star\rho)$ and $\partial_v\partial_\mu(u\star\rho)$ are $\mathbb{P}$-a.s. continuous at every triple $(t, \mu*\phi_G, v), t \in [0, T], v \in \mathbb{R}^d$.

Now we introduce $\phi_{\varepsilon,G}$ - Gaussian densities $N(0, \varepsilon I_d)$. Letting $\varepsilon \searrow 0$ one can see convergence of $\phi_{\varepsilon,G}$ to Dirac measure at 0 for the W_2 distance and thus convergence of $\partial_\mu(u \star \rho)_t(\mu * \phi_{\varepsilon,G}, v)$ and $\partial_v\partial_\mu(u \star \rho)_t(\mu * \phi_{\varepsilon,G}, v)$ to $\partial_\mu(u \star \rho)_t(\mu, v)$ and $\partial_v\partial_\mu(u \star \rho)_t(\mu, v)$ respectively for any $t \in [0, T], v \in \text{Supp}(\mu)$. Now picking ρ_n in a way that $(\rho_n, \partial_z\rho_n, \partial^2_{zz}\rho_n)(z) \to (z, I_d, 0)$ as $n \to \infty$, we can conclude that $\partial_\mu(u \star \rho_n)_t(\mu, v)$ and $\partial_v\partial_\mu(u \star \rho_n)_t(\mu, v)$ converge to $\partial_\mu u_t(\mu, v)$ and $\partial_v\partial_\mu u_t(\mu, v)$ $\mathbb{P}$-a.s.. One should notice that all the conditions in the theorem hold true while doing mollification. Thus we can assume that u and its first and partial second order derivatives are $\mathbb{P}$-a.s. uniformly bounded and uniformly continuous, and Y is a bounded process.

Now we are to show the well-posedness of the mollification scheme, i.e. that chain rule applied to $u_n := u \star \rho_n$ converges to the one for u. It is straightforward to verify that u_n satisfies $\mathbb{P}$-a.s. (4.1.2) uniformly in $n \geqslant 1$. We apply the dominated convergence theorem twice to conclude the $\mathbb{P}$-a.s. convergence for all the terms but the stochastic integral. To handle the latter one additionally requires an argument across the quadratic variation as written in Theorem 3.2.1 and localisation.

Step 2. Wellposedness and approximation. For a smooth compactly supported density ρ on $\mathbb{R}^d$ we define, for $n \in \mathbb{N}$, the mollified version $u^{N,n}$ of u^N (introduced in Definition 3.3.4) for any $t \in [0, T]$, any $y^1, \dots, y^N \in \mathbb{R}^d$ by

$$u^{N,n}_t(y^1, \dots, y^N) : = n^{Nd} \int_{(\mathbb{R}^d)^N} u^N_t(y^1 - z^1, \dots, y^N - z^N) \prod_{l=1}^N \rho(nz^l) \prod_{l=1}^N dz^l,$$

where ρ is a smooth and compactly supported density. We define $\phi^{N,n}, \psi^{N,n}$, in the same way as $u^{N,n}$. One can notice that $u^{N,n}_t, \phi^{N,n}_t, \psi^{N,n}_t$ are maps in $\mathcal{C}^2\big((\mathbb{R}^d)^N\big)$ and thus all derivatives up to second order exist and are regular. Furthermore, to $u^{N,n}$ one can apply the standard Itô-Wentzell formulae, since it satisfies all the conditions of Theorem 1.1.2 (verified below).

Now we describe the approximation procedure. From the properties of the Wasserstein metric for finitely supported measures with uniformly bounded second moments, we have

$$W_2\Big(\frac{1}{N}\sum_{i=1}^N \delta_{y^i}, \frac{1}{N}\sum_{i=1}^N \delta_{y^i - Z^i/n}\Big)^2 \leqslant \frac{C}{n^2},$$

where C depends on the support of ρ.

We generate the processes $\big((Y^l_t)_{t\in[0,T]}\big)_{l=1,\dots,N}$ - the independent twin processes of $(Y_t)_{t\in[0,T]}$. We underline that processes $\big(Y^l_t, b^l_t, \sigma^l_t\big)_{t\in[0,T]}$, $l = 1, \dots N$ are i.i.d. $\mathbb{P}$-a.s. and the random variables Y^l_0 are i.i.d $\mathbb{P}$-a.s. as well.

The technique is as follows: we mollify the empirical projection u^N_t and obtain $u^{N,n}_t$, this way we can take second-order derivatives and afterwards apply the "propagation of chaos" argument to approximate u_t by u_t^N, namely for any $t \in [0,T]$ one have $\mathbb{P}$-a.s.

$$\begin{aligned}
&\sup_{t\in[0,T]} \mathbb{E}^{1,\dots,N}\Big[|u_t^{N,n}(Y_t^1,\dots,Y_t^N) - u_t(\mu)|\Big] \\
&\leqslant \sup_{t\in[0,T]} \mathbb{E}^{1,\dots,N}\Big[|u_t^{N,n}(Y_t^1,\dots,Y_t^N) - u_t^N(Y_t^1,\dots,Y_t^N)|\Big] + \sup_{t\in[0,T]} \mathbb{E}^{1,\dots,N}\Big[|u_t(\bar{\mu}_t^N) - u_t(\mu_t)|\Big] \\
&\leqslant \varepsilon_n + \varepsilon_N,
\end{aligned}$$

where $(\varepsilon_k)_{k\geqslant 1}$ is a sequence of random variables $\mathbb{P}$-a.s. converging to 0, as $k \to \infty$ uniformly in time, this is seen via a propagation of chaos argument, continuity of u, dominated convergence theorem and the fact that convergence in Wasserstein metric only depends on the moments of the distribution.

By the $\mathbb{P}$-a.s. boundedness of u one can get for any $p \geqslant 1$

$$\sup_{t\in[0,T]} \mathbb{E}^{1,\dots,N}\Big[|u_t^{N,n}(Y_t^1,\dots,Y_t^N) - u_t(\mu)|^p\Big]^{\frac{1}{p}} \leqslant \varepsilon_n^{(p)} + \varepsilon_N^{(p)}, \tag{4.1.3}$$

where $(\varepsilon_k^{(p)})_{k\in\mathbb{N}}$ is a sequence converging $\mathbb{P}$-a.s. to 0.

Now we use the Proposition 3.3.5 to get for any $t \in [0,T]$, $\mathbb{P}$-a.s.

$$\begin{aligned}
\partial_{y^i} u_t^{N,n}(y^1,\dots,y^N) &= n^{Nd} \int_{(\mathbb{R}^d)^N} \partial_{y^i} u_t^N(y^1 - z^1,\dots,y^N - z^N) \prod_{l=1}^N \rho(nz^l) \prod_{l=1}^N dz^l \\
&= \frac{n^{Nd}}{N} \int_{(\mathbb{R}^d)^N} \partial_\mu u_t\Big(\frac{1}{N}\sum_{l=0}^N \delta_{y^l - z^l}, y^i - z^i\Big) \prod_{l=1}^N \rho(nz^l) \prod_{l=1}^N dz^l.
\end{aligned}$$

Applying the same argument as above we get $\mathbb{P}$-a.s., $p \geqslant 1$

$$\sup_{t\in[0,T]} \mathbb{E}^{1,\dots,N}\Big[|N\partial_{y^i} u_t^{N,n}(Y_t^1,\dots,Y_t^N) - \partial_\mu u_t(\mu_t, Y_t^i)|^p\Big]^{\frac{1}{p}} \leqslant \varepsilon_n^{(p)} + \varepsilon_N^{(p)}. \tag{4.1.4}$$

Now we differentiate once again with respect to y^i

$$\partial_{y^i}\partial_{y^i} u_t^{N,n}(y^1,\dots,y^N) = \frac{n^{Nd+1}}{N} \int_{(\mathbb{R}^d)^N} \partial_\mu u_t^N\Big(\frac{1}{N}\sum_{l=1}^N \delta_{y^l - z^l}, y^i - z^i\Big) \otimes \partial_{z^i}\rho(nz^i) \prod_{l\neq i}^N \rho(nz^l) \prod_{l=1}^N dz^l,$$

with standard tensor product operating on elements of $\mathbb{R}^d$.

To the previous identity we add and subtract a perturbation term focusing on the contribution by δ_{y^i}

$$\begin{aligned}
&N\partial^2_{y^iy^i} u_t^{N,n}(y^1,\dots,y^N) \\
&\quad = n^{Nd+1} \int_{(\mathbb{R}^d)^N} \partial_\mu u_t\Big(\frac{1}{N}\sum_{l\neq i}^N \delta_{y^l - z^l} + \frac{1}{N}\delta_{y^i}, y^i - z^i\Big) \otimes \partial_{z^i}\rho(nz^i) \prod_{l\neq i}^N \rho(nz^l) \prod_{l=1}^N dz^l \\
&\quad + n^{Nd+1} \int_{(\mathbb{R}^d)^N} \Big[\partial_\mu u_t\Big(\frac{1}{N}\sum_{l=1}^N \delta_{y^l - z^l}, y^i - z^i\Big) - \partial_\mu u_t\Big(\frac{1}{N}\sum_{l\neq i}^N \delta_{y^l - z^l} + \frac{1}{N}\delta_{y^i}, y^i - z^i\Big)\Big] \\
&\quad \otimes \partial_{z^i}\rho(nz^i) \prod_{l\neq i}^N \rho(nz^l) \prod_{l=1}^N dz^l \\
&\quad = T_{n,i}^{1,N}(y^1,\dots,y^N) + T_{n,i}^{2,N}(y^1,\dots,y^N).
\end{aligned}$$

We integrate by parts $T_{n,i}^{1,N}$ with respect to the space variable y (that appears from the derivative in measure and notice the two minus signs), use the compact support of ρ for the boundary term, and to the resulting integral term we add and subtract a $\partial_v\partial_\mu u_t^N$ over the whole empirical measure, this yields

$$
\begin{aligned}
T_{n,i}^{1,N}(y^1,\dots,y^N) &= n^{Nd}\int_{(\mathbb{R}^d)^N} \partial_v\partial_\mu u_t\Big(\frac{1}{N}\sum_{l=1}^N \delta_{y^l-z^l}, y^i-z^i\Big)\prod_{l=1}^N \rho(nz^l)\prod_{l=1}^N dz^l \\
&\qquad + n^{Nd}\int_{(\mathbb{R}^d)^N}\Big[\partial_v\partial_\mu u_t\Big(\frac{1}{N}\sum_{l\neq i}^N \delta_{y^l-z^l}+\frac{1}{N}\delta_{y^i}, y^i-z^i\Big) \\
&\qquad - \partial_v\partial_\mu u_t\Big(\frac{1}{N}\sum_{l=1}^N \delta_{y^l-z^l}, y^i-z^i\Big)\Big]\prod_{l=1}^N \rho(nz^l)\prod_{l=1}^N dz^l \\
&= T_{n,i}^{11,N}(y^1,\dots,y^N)+T_{n,i}^{12,N}(y^1,\dots,y^N).
\end{aligned}
$$

For $T_{n,i}^{11,N}$ we have, as previously due to uniform continuity of $\partial_v\partial_\mu u_t$, $\mathbb{P}$-a.s., $p\geqslant 1$

$$
\sup_{t\in[0,T]} \mathbb{E}^{1,\dots,N}\Big[|T_{n,i}^{11,N}(Y_t^1,\dots,Y_t^N)-\partial_v\partial_\mu u_t(\mu_t,Y_t^i)|^p\Big]^{\frac{1}{p}} \leqslant \varepsilon_n^{(p)}+\varepsilon_N^{(p)}. \tag{4.1.5}
$$

Uniform continuity of $\partial_v\partial_\mu u$ (in space-measure variables) together with the properties of the Wasserstein metric over finitely supported measures gives

$$
W_2\Big(\frac{1}{N}\sum_{l\neq i}^N \delta_{y^l-z^l}+\frac{1}{N}\delta_{y^i}, \frac{1}{N}\sum_{l=1}^N \delta_{y^l-z^l}\Big)^2 \leqslant \frac{1}{N}C,
$$

which in turn implies $\mathbb{P}$-a.s., $p\geqslant 1$

$$
\sup_{t\in[0,T]} \mathbb{E}^{1,\dots,N}\big[|T_{n,i}^{12,N}(Y_t^1,\dots,Y_t^N)|^p\big]^{\frac{1}{p}} \leqslant \varepsilon_N^{(p)}. \tag{4.1.6}
$$

The procedure to deal with $T_{n,i}^{12,N}$ also applies to $T_{n,i}^{2,N}$ and yields $\mathbb{P}$-a.s. for any $t\in[0,T]$

$$
\sup_{t\in[0,T]} \mathbb{E}^{1,\dots,N}\big[|T_{n,i}^{2,N}(Y_t^1,\dots,Y_t^N)|^p\big]^{\frac{1}{p}} \leqslant n\varepsilon_N^{(p)}, \tag{4.1.7}
$$

with an additional multiplicative factor n appearing after differentiating the regularisation kernel.

We say that $\phi_t(\cdot),\psi_t(\cdot)=0$ for all other t, where ϕ,ψ are not defined. Now the same technique is valid to $\phi^{N,n},\psi^{N,n}$ to get $\mathbb{P}$-a.s for almost all t

$$
\begin{aligned}
\mathbb{E}^{1,\dots,N}\Big[\phi_t^{N,n}(Y_t^1,\dots,Y_t^N)\Big] &\to \phi_t(\mu_t), \quad \text{as } N,n\to\infty, \\
\mathbb{E}^{1,\dots,N}\Big[\psi_t^{N,n}(Y_t^1,\dots,Y_t^N)\Big] &\to \psi_t(\mu_t), \quad \text{as } N,n\to\infty.
\end{aligned}
$$

Hence, $\mathbb{P}$-a.s., $p\geqslant 1$

$$
\sup_{0\leqslant t\leqslant T} \mathbb{E}^{1,\dots,N}\Big[\Big|\int_0^t \phi_s^{N,n}(Y_s^1,\dots,Y_s^N)\mathrm{d}s-\int_0^t \phi_s(\mu_s)\mathrm{d}s\Big|^p\Big]^{\frac{1}{p}} \leqslant \varepsilon_n^{(p)}+\varepsilon_N^{(p)}, \tag{4.1.8}
$$

$$
\sup_{0\leqslant t\leqslant T} \mathbb{E}^{1,\dots,N}\Big[\Big|\int_0^t \psi_s^{N,n}(Y_s^1,\dots,Y_s^N)\cdot\mathrm{d}W_s-\int_0^t \psi_s(\mu_s)\cdot\mathrm{d}W_s\Big|^p\Big]^{\frac{1}{p}} \leqslant \varepsilon_n^{(p)}+\varepsilon_N^{(p)}. \tag{4.1.9}
$$

Without loss of generality we pick the $(\varepsilon_k)_{k\in\mathbb{N}}$ the same as for u. One can notice that $\psi^{N,n}$, $\phi^{N,n}$ satisfy condition (3.2.1) of Theorem 3.2.1, due to mollification and the identification from Proposition 3.3.5.

Step 3: Applying the classical Itô-Wentzell to the approximation. Under our assumptions and the mollification argument in combination with Proposition 3.3.5, we have sufficient regularity that we can apply the standard Itô-Wentzell formula (see Theorem 1.1.2 and Theorem 3.2.1) to $u^{N,n}$ and obtain

$$\begin{aligned}
0 = u_t^{N,n}(Y_t^1, \dots, Y_t^N) &- u_t^{N,n}(Y_0^1, \dots, Y_0^N) \\
&- \int_0^t \phi_s^{N,n}(Y_s^1, \dots, Y_s^N)\mathrm{d}s - \int_0^t \psi_s^{N,n}(Y_s^1, \dots, Y_s^N) \cdot \mathrm{d}W_s \\
&- \frac{1}{N}\sum_{l=1}^N \int_0^t \partial_{y^l} u_s^{N,n}(Y_s^1, \dots, Y_s^N) \cdot b_s^l \mathrm{d}s \\
&- \frac{1}{N}\sum_{l=1}^N \int_0^t \partial_{y^l} u_s^{N,n}(Y_s^1, \dots, Y_s^N) \cdot \sigma_s^l \mathrm{d}W_s^l \\
&- \frac{1}{2N}\sum_{l=1}^N \int_0^t \mathrm{Trace}\big\{\partial^2_{y^l y^l} u_s^{N,n}(Y_s^1, \dots, Y_s^N)\ \sigma_s^l(\sigma_s^l)^\intercal\big\}\mathrm{d}s.
\end{aligned} \tag{4.1.10}$$

Note two important simplifications. Firstly, one would expect the second-derivative term to contain a Hessian, but for independent processes $Y_t^{l_1}, Y_t^{l_2}, l_1 \neq l_2$, we have $\mathrm{d}\langle Y^{l_1}, Y^{l_2}\rangle_t = \mathbb{1}_{\{l_1=l_2\}}\sigma_t^{l_1}(\sigma_t^{l_2})^\intercal \mathrm{d}t$ and hence only diagonal terms appear. Secondly, no cross-variation term $d\langle \partial_\mu u^{N,n}, Y^l\rangle_t$ appears, this is due to the independence of the field's noise W_t and noise of the particles $\{W_t^l\}_{l=1,\dots,N}$ within empirical approximation (this will not be the case in the next section).

Now we can proceed with the expected result. Define $\Delta^{N,n}$ as the difference between the RHS of (4.1.10) and the RHS of the below equation, we then have for any $t \in [0,T]$ $\mathbb{P}$-a.s. (the tautology)

$$\begin{aligned}
\Delta_t^{N,n} = u_t(\mu_t) - u_0(\mu_0) - \int_0^t \phi_s(\mu_s)\mathrm{d}s - \int_0^t \psi_s(\mu_s)\cdot \mathrm{d}W_s - \frac{1}{N}\sum_{l=1}^N \int_0^t \partial_\mu u_s(\mu_s, Y_s^l)\cdot b_s^l \mathrm{d}s \\
- \frac{1}{N}\sum_{l=1}^N \int_0^t \partial_\mu u_s(\mu_s, Y_s^l)\cdot \sigma_s^l \mathrm{d}W_s^l - \frac{1}{2N}\sum_{l=1}^N \int_0^t \mathrm{Trace}\big\{\partial_v\partial_\mu u_s(\mu_s, Y_s^l)\ \sigma_s^l(\sigma_s^l)^\intercal\big\}\mathrm{d}s.
\end{aligned}$$

It is clear that $[0,T] \ni t \mapsto \Delta_t^{N,n}$ is continuous. Moreover, collecting the inequalities (4.1.3)-(4.1.9) we have $\sup_{0\leqslant t\leqslant T} |\mathbb{E}^{1,\dots,N}\big[\Delta_t^{N,n}\big]| \leqslant \varepsilon_n + (1+n)\varepsilon_N$, $\mathbb{P}$-a.s.

We let $N \to \infty$ to get by Fatou's lemma, the law of large numbers and the joint-continuity of all derivatives with localisation argument for stochastic integral term, $\mathbb{P}$-a.s. that $\sup_{0\leqslant t\leqslant T}|\Delta_t^n| \leqslant 3\varepsilon_n$, where $\mathbb{P}$-a.s.

$$\begin{aligned}
\Delta_t^n = u_t(\mu_t) - u_0(\mu_0) - \int_0^t \phi_s(\mu_s)\mathrm{d}s - \int_0^t \psi_s(\mu_s)\cdot \mathrm{d}W_s \\
- \int_0^t \tilde{\mathbb{E}}\Big[\partial_\mu u_s(\mu_s, \tilde{Y}_s)\cdot \tilde{b}_s\Big]\mathrm{d}s - \frac{1}{2}\int_0^t \tilde{\mathbb{E}}\Big[\mathrm{Trace}\big\{\partial_v\partial_\mu u_s(\mu_s, \tilde{Y}_s)\ \tilde{\sigma}_s(\tilde{\sigma}_s)^\intercal\big\}\Big]\mathrm{d}s,
\end{aligned} \tag{4.1.11}$$

where $\Delta_t^n := \lim_{n\to\infty} \Delta_t^{N,n}$, and we applied Fubini's theorem to interchange the Lebesgue integral with the expectation. Note that to handle the stochastic integral we apply the localisation technique and use dominated convergence theorem once more. Letting $n \to \infty$ in the equation above, we conclude that $\Delta_t^n \to \Delta \equiv 0$, $\mathbb{P}$-a.s., which finishes this part of the proof. The measurability of the involved coefficients follows the guidelines set in Remark 3.6.5. $\square$

4.1.2 The joint chain rule

Now we are ready to provide a joint chain rule formula expanding the nature of the random field to support a space variable dependence, i.e. the case $t \mapsto u_t(X_t, \mu_t)$ for μ the law of (3.4.2) and X solution to (3.4.1). Let us start by inheriting the structure and properties of the setup of Theorem 4.1.3.

Definition 4.1.7. *We say the random field* $u : \Omega \times [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ *given in* (4.1.1) *is* RF-Joint-Partially-$\mathcal{C}^2$ *if* *i) For any $(x,\mu)\in\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, $(u_t(x,\mu))_{t\in[0,T]}$ is a continuous adapted process taking values in $\mathbb{R}$ and $(\phi_t(x,\mu))_{t\in[0,T]},(\psi_t(x,\mu))_{t\in[0,T]}$ are $\mathbb{F}$-progressively measurable processes with values in $\mathbb{R}$ and $\mathbb{R}^d$ respectively;*

ii) For almost any $t\in[0,T]$, the maps $(x,\mu)\mapsto\phi_t(x,\mu)$, $(x,\mu)\mapsto\psi_t(x,\mu)$ are $\mathbb{P}$-a.s. jointly-continuous in the product topology of $\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ at every pair $(x,\mu)\in\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$;

iii) For any $(t,\mu)\in[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, the map $x\mapsto u_t(x,\mu)$ is $\mathcal{C}^2(\mathbb{R}^d)$, $\mathbb{P}$-a.s. at every $x\in\mathbb{R}^d$, with $\partial_x u,\partial^2_{xx}u$ being $\mathbb{P}$-a.s. joint continuous at every triple $(t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, $\mathbb{P}$-a.s.;

iv) For almost any $t\in[0,T]$, for any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, the map $x\mapsto\psi_t(x,\mu)$ is $\mathcal{C}^1(\mathbb{R}^d)$, $\mathbb{P}$-a.s. at every $x\in\mathbb{R}^d$, with $\partial_x\psi$ being $\mathbb{P}$-a.s. joint continuous at every pair $(x,\mu)\in\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, $\mathbb{P}$-a.s.;

v) For any $(t,x)\in[0,T]\times\mathbb{R}^d$, the map $\mu\mapsto u_t(x,\mu)$ is $\mathbb{P}$-a.s. continuous in the Wasserstein metric and L-differentiable $\mathbb{P}$-a.s. at every $\mu\in\mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\partial_\mu u_t(x,\mu,v)$ is $\mathbb{P}$-a.s. joint continuous at every quadruple (t,x,μ,v), with $(t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, $v\in Supp(\mu)$, $\mathbb{P}$-a.s.;

vi) For any $(t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, the map $v\mapsto\partial_\mu u_t(x,\mu,v)$ is $\mathbb{R}^d$-differentiable $\mathbb{P}$-a.s., at every $v\in Supp(\mu)$. Moreover, $\partial_v\partial_\mu u_t(x,\mu,v)$ is $\mathbb{P}$-a.s. joint continuous at every quadruple (t,x,μ,v), with $(t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, $v\in Supp(\mu)$, $\mathbb{P}$-a.s..

Theorem 4.1.8. *Let $u:\Omega\times[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ defined by* (4.1.1) *to be* RF-Joint-Partially-$\mathcal{C}^2$. *Assume that for any compact $K\subset\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ and $t\in[0,T]$ we have*

$$\int_0^t \sup_{(x,\mu)\in K}\Big\{|\phi_s(x,\mu)|+|\psi_s(x,\mu)|^2+|\partial_x\psi_s(x,\mu)|^2\Big\}ds<\infty,\quad \mathbb{P}\text{-a.s.},\tag{4.1.12}$$

and

$$\sup_{(t,x,\mu)\in[0,T]\times K}\Big\{\int_{\mathbb{R}^d}\Big[|\partial_\mu u_t(x,\mu,v)|^2+|\partial_v\partial_\mu u_t(x,\mu,v)|^2\Big]\mu(dv)\Big\}<\infty,\quad \mathbb{P}\text{-a.s.}.\tag{4.1.13}$$

Let $(\mu_t)_{t\in[0,T]}$ be the law of the solution to (3.4.2) *satisfying Assumption 3.4.2. Let $(X_t)_{t\in[0,T]}$ be the solution process to* (3.4.1) *under Assumption 3.4.1.*

Then the process $(u_t(X_t,\mu_t))_{t\in[0,T]}$ is an Itô process $\mathbb{P}$-a.s. satisfying the dynamics

$$\begin{aligned}
&u_T(X_T,\mu_T)-u_0(X_0,\mu_0)\\
&\quad=\int_0^T\phi_s(X_s,\mu_s)ds+\int_0^T\psi_s(X_s,\mu_s)\cdot dW_s\\
&\quad+\int_0^T\partial_x u_s(X_s,\mu_s)\cdot\gamma_s dW_s+\int_0^T \mathit{Trace}\big\{\partial_x\psi_s(X_s,\mu_s)(\gamma_s)^\intercal\big\}ds\\
&\quad+\int_0^T\partial_x u_s(X_s,\mu_s)\cdot\beta_s ds+\int_0^T\frac{1}{2}\,\mathit{Trace}\big\{\partial^2_{xx}u_s(X_s,\mu_s)\;\gamma_s(\gamma_s)^\intercal\big\}ds\\
&\quad+\int_0^T\tilde{\mathbb{E}}\Big[\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)\cdot\tilde{b}_s\Big]ds+\int_0^T\frac{1}{2}\,\tilde{\mathbb{E}}\Big[\mathit{Trace}\big\{\partial_v\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)\;\tilde{\sigma}_s(\tilde{\sigma_s})^\intercal\big\}\Big]ds,
\end{aligned}\tag{4.1.14}$$

with $u_0(X_0,\mu_0)=f(X_0,\mu_0)$.

Observe that the terms of the first and the last line on the RHS of the formula are the ones from our Theorem 4.1.3, whereas the middle two arise from the standard Itô-Wentzell formulae.

Proof. In view of the proof of Theorem 4.1.3 we assume a compactification/mollification argument in the measure component has been applied. In this way we avoid a repetition of arguments.

We start by fixing a time T and let $\Delta_K = \{0 = t_0 < t_1 < \cdots < t_K = T\}$ be a partition of $[0,T]$ with modulus $|\Delta_K| = \min_{0\leqslant j\leqslant K-1} |t_{j+1} - t_j| > 0$. Then

$$\begin{aligned}
&u_T(X_T, \mu_T) - u_0(X_0)\\
&= \sum_{i=0}^{K-1} \Big[u_{t_{i+1}}(X_{t_{i+1}}, \mu_{t_{i+1}}) - u_{t_i}(X_{t_i}, \mu_{t_i})\Big]\\
&= \sum_{i=0}^{K-1} \Big[u_{t_{i+1}}(X_{t_{i+1}}, \mu_{t_{i+1}}) - u_{t_i}(X_{t_i}, \mu_{t_{i+1}})\Big] + \sum_{i=0}^{K-1} \Big[u_{t_i}(X_{t_i}, \mu_{t_{i+1}}) - u_{t_i}(X_{t_i}, \mu_{t_i})\Big] = I_1^{(K)} + I_2^{(K)}.
\end{aligned}$$

Now we see that $I_2^{(K)}$ is amenable to Remark 4.1.4 which together with the joint time-space continuity of the measure derivatives, a localisation procedure for X, applying twice the dominated convergence theorem in combination with Assumption 3.4.2 yields

$$\begin{aligned}
I_2^{(K)} &= \sum_{i=0}^{K-1} \Bigg[\int_{t_i}^{t_{i+1}} \tilde{\mathbb{E}}[\partial_\mu u_{t_i}(X_{t_i}, \mu_s, \tilde{Y}_s) \cdot \tilde{b}_s]\mathrm{d}s + \frac{1}{2}\int_{t_i}^{t_{i+1}} \tilde{\mathbb{E}}\Big[\text{Trace}\big\{\partial_v\partial_\mu u_{t_i}(X_{t_i}, \mu_s, \tilde{Y}_s)\tilde{\sigma}_s(\tilde{\sigma}_s)^\intercal\big\}\Big]\mathrm{d}s\Bigg],\\
&\to \int_0^T \tilde{\mathbb{E}}[\partial_\mu u_s(X_s, \mu_s, \tilde{Y}_s) \cdot \tilde{b}_s]\mathrm{d}s + \frac{1}{2}\int_0^T \tilde{\mathbb{E}}\Big[\text{Trace}\big\{\partial_v\partial_\mu u_s(X_s, \mu_s, \tilde{Y}_s)\ \tilde{\sigma}_s(\tilde{\sigma}_s)^\intercal\big\}\Big]\mathrm{d}s,
\end{aligned}$$

where we have taken the limit $|\Delta_K| \to 0$.

The measure increment is forward in time for $I_1^{(K)}$, however its flow is deterministic allowing to directly pass to the limit, after applying the Theorem 3.2.1, whose assumptions are satisfied, having

$$\begin{aligned}
I_1^{(K)} = \sum_{i=0}^{K-1} \Bigg[&\int_{t_i}^{t_{i+1}} \phi_s(X_s, \mu_{t_{i+1}})\mathrm{d}s + \int_{t_i}^{t_{i+1}} \psi_s(X_s, \mu_{t_{i+1}}) \cdot \mathrm{d}W_s\\
&+ \int_{t_i}^{t_{i+1}} \partial_x u_s(X_s, \mu_{t_{i+1}}) \cdot \beta_s\mathrm{d}s + \int_{t_i}^{t_{i+1}} \partial_x u_s(X_s, \mu_{t_{i+1}}) \cdot \gamma_s\mathrm{d}W_s]\\
&+ \frac{1}{2}\int_{t_i}^{t_{i+1}} \text{Trace}\big\{\partial^2_{xx} u_s(X_s, \mu_{t_{i+1}})\ \gamma_s(\gamma_s)^\intercal\big\}\mathrm{d}s + \int_{t_i}^{t_{i+1}} \text{Trace}\big\{\partial_x\psi_s(X_s, \mu_{t_{i+1}})(\gamma_s)^\intercal\big\}\mathrm{d}s\Bigg].
\end{aligned}$$

Now one can pass to the limit in $I_1^{(K)}$ as $|\Delta_K| \to 0$, by applying joint-continuity of u and its derivatives, alongside Lebesgue dominated convergence theorem, localisation procedure to deal with X, and standard quadratic variation argument to handle stochastic integral, so

$$\begin{aligned}
I_1^{(K)} \to &\int_0^T \phi_s(X_s, \mu_s)\mathrm{d}s + \int_0^T \psi_s(X_s, \mu_s) \cdot \mathrm{d}W_s\\
&+ \int_0^T \partial_x u_s(X_s, \mu_s) \cdot \beta_s\mathrm{d}s + \int_0^T \partial_x u_s(X_s, \mu_s) \cdot \gamma_s\mathrm{d}W_s]\\
&+ \frac{1}{2}\int_0^T \text{Trace}\{\partial^2_{xx} u_s(X_s, \mu_s)\ \gamma_s(\gamma_s)^\intercal\}\mathrm{d}s + \int_0^T \text{Trace}\big\{\partial_x\psi_s(X_s, \mu_s)(\gamma_s)^\intercal\big\}\mathrm{d}s.
\end{aligned} \tag{4.1.15}$$

Joining all the limits we see that (4.1.14) immediately follows. Measurability is dealt by Remark 3.6.5.

□

4.2 Itô-Wentzell-Lions chain rule with a conditional flow of measures

The setting discussed in this section is inspired by the developments in the theory of mean-field games with common noise, [19] and [22]. Since the framework evolves from that in the previous chapters we set up our probability spaces and notation anew.

For the purposes of this section the probability space inherits the structure outlined in Section 3.5. We refer to Section 3.6 for introduction of the processes and existing Itô-Lions expansion.

Let $u : \Omega \times [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a random field, satisfying the dynamics

$$du_t(x,\mu) = \phi_t(x,\mu)\mathrm{d}t + \psi^0_t(x,\mu)\cdot \mathrm{d}W^0_t + \psi^1_t(x,\mu)\cdot \mathrm{d}W^1_t, \quad u_0(x,\mu) = f(x,\mu), \tag{4.2.1}$$

where $f(x,\mu) : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is a deterministic function, $W^0 = (W^0_t)_{t\in[0,T]}$ and $W^1 = (W^1_t)_{t\in[0,T]}$ are independent d-dimensional $\mathbb{F}^0$ and $\mathbb{F}^1$-Brownian motions respectively; $(\phi,\psi^0,\psi^1) : \Omega \times [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ are $\mathbb{F}$-progressively measurable processes.

4.2.1 Itô-Wentzell-Lions formula for measure functionals

For the derivation of the expansion in measure component, and as in Theorem 4.1.3, we fix $x \in \mathbb{R}^d$ then omit its dependence, i.e.

$$u_t(\mu) := u_t(x,\mu), \quad \phi_t(\mu) := \phi_t(x,\mu), \quad \psi^0_t(\mu) := \psi^0_t(x,\mu) \quad \text{and} \quad \psi^1_t(\mu) := \psi^1_t(x,\mu).$$

Now we introduce the regularity for random field given by (4.2.1) which inherits Definition 4.1.1 and requires additionally a second-order Fréchet differentiability.

Definition 4.2.1. *We say the random field $u : \Omega \times [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ given in* (4.2.1) *(for some $x \in \mathbb{R}^d$ fixed) is* RF-Generally-$\mathcal{C}^2$ *if*

i) u is RF-Partially-$\mathcal{C}^2$ *for $\psi_t := (\psi^0_t, \psi^1_t)^\intercal$ and $W_t := (W^0_t, W^1_t)^\intercal$;*

ii) For any $(t,v) \in [0,T] \times Supp(\mu)$, the map $\mu \mapsto \partial_\mu u_t(\mu,v)$ is L-differentiable $\mathbb{P}$-a.s. at every point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\partial^2_\mu u_t(\mu,v,v')$ is $\mathbb{P}$-a.s. joint-continuous at every quadruple (t,μ,v,v'), with $(t,\mu) \in [0;T] \times \mathcal{P}_2(\mathbb{R}^d),\ v,v' \in Supp(\mu),\ \mathbb{P}$-a.s.;

iii) For almost any $t \in [0,T]$, the map $\mu \mapsto \psi^0_t(\mu)$ is L-differentiable $\mathbb{P}$-a.s. at every point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\partial_\mu\psi^0_t(\mu,v)$ is $\mathbb{P}$-a.s. joint-continuous at every pair $(\mu,v),\ \mu \in \mathcal{P}_2(\mathbb{R}^d),\ v \in Supp(\mu),\ \mathbb{P}$-a.s.

We highlight the slight abuse of notation in the way point *i)* in the above Definition 4.2.1 is formulated. This avoids re-stating a full assumption that is nonetheless clear to understand.

Theorem 4.2.2. *Let u be* RF-Generally-$\mathcal{C}^2$ *Itô random field* (4.2.1) *(where $x \in \mathbb{R}^d$ is fixed and omitted throughout, also for ϕ and ψ). Assume for any compact $K \subset \mathcal{P}_2(\mathbb{R}^d)$ we have*

$$\int_0^T \sup_{\mu\in K} \Big\{ |\phi_s(\mu)| + |\psi^0_s(\mu)|^2 + |\psi^1_s(\mu)|^2 + \int_{\mathbb{R}^d} |\partial_\mu \psi^0_s(\mu,v)|^2 \mu(dv) \Big\} ds < \infty, \quad \mathbb{P}\text{-a.s.},$$

and

$$\sup_{(t,\mu)\in[0,T]\times K} \Big\{ \int_{\mathbb{R}^d} \Big[|\partial_\mu u_t(\mu,v)|^2 + |\partial_v\partial_\mu u_t(\mu,v)|^2 \Big] \mu(dv) + \int_{\mathbb{R}^d\times\mathbb{R}^d} |\partial^2_\mu u_t(\mu,v,v'|^2 \mu(dv)\mu(dv') \Big\} < \infty, \quad \mathbb{P}\text{-a.s.}. \tag{4.2.2}$$

For almost all $\omega^0 \in \Omega^0$ take $(\mu_t)_{t\in[0,T]} := \big(Law(Y_t(\omega_0,\cdot))\big)_{t\in[0,T]}$, with Y solution to (3.6.1) *under Assumption 3.6.1.*

Then $(u_t(\mu_t))_{t\in[0,T]}$ is an Itô process $\mathbb{P}$-a.s. satisfying the expansion

$$
\begin{aligned}
u_T(\mu_T) - u_0(\mu_0) = &\int_0^T \phi_s(\mu_s)ds + \int_0^T \psi_s^0(\mu_s)\cdot dW_s^0 + \int_0^T \psi_s^1(\mu_s)\cdot dW_s^1 \\
&+ \int_0^T \tilde{\mathbb{E}}^1\Big[\partial_\mu u_s(\mu_s,\tilde{Y}_s)\cdot \tilde{b}_s\Big]ds + \int_0^T \tilde{\mathbb{E}}^1\Big[(\tilde{\sigma}_s^0)^\intercal \partial_\mu u_s(\mu_s,\tilde{Y}_s)\Big]\cdot dW_s^0 \\
&+ \int_0^T \frac{1}{2}\,\tilde{\mathbb{E}}^1\Big[\mathit{Trace}\big\{\partial_v\partial_\mu u_s(\mu_s,\tilde{Y}_s)(\tilde{\sigma}_s^0(\tilde{\sigma}_s^0)^\intercal + \tilde{\sigma}_s^1(\tilde{\sigma}_s^1)^\intercal)\big\}\Big]ds \qquad (4.2.3)\\
&+ \int_0^T \frac{1}{2}\,\hat{\mathbb{E}}^1\Big[\tilde{\mathbb{E}}^1\Big[\mathit{Trace}\big\{\partial_\mu^2 u_s(\mu_s,\tilde{Y}_s,\hat{Y}_s)\ \tilde{\sigma}_s^0(\hat{\sigma}_s^0)^\intercal\big\}\Big]\Big]ds \\
&+ \int_0^T \tilde{\mathbb{E}}^1\Big[\mathit{Trace}\big\{\partial_\mu\psi_s^0(\mu_s,\tilde{Y}_s)(\tilde{\sigma}_s^0)^\intercal\big\}\Big]ds,
\end{aligned}
$$

where the formula above $\tilde{\mathbb{E}}$ and $\hat{\mathbb{E}}$ denote the expectation acting on the model twin spaces $(\tilde{\Omega},\tilde{\mathbb{F}},\tilde{\mathbb{P}})$ and $(\hat{\Omega},\hat{\mathbb{F}},\hat{\mathbb{P}})$ respectively, and let the processes $(\tilde{Y}_t,\tilde{b}_t,\tilde{\sigma}_t)_{t\in[0,T]}$ and $(\hat{Y}_t,\hat{b}_t,\hat{\sigma}_t)_{t\in[0,T]}$ be the independent twin processes of $(Y_t,b_t,\sigma_t)_{t\in[0,T]}$ respectively living within.

One can notice two new terms appearing in contrast with the formula in Theorem 4.1.3. Whilst the $\partial_\mu^2 u$ term appears as a cross-variation of two model particles $\tilde{Y}$ and $\hat{Y}$ experiencing the same noise W^0 and is present in Theorem 3.6.4, a brand new $\partial_\mu\psi^0$ term now indicates an interaction of the field u with the model particle $\tilde{Y}$ through the same W^0.

In contrast to the proof of Theorem 4.1.3, the arguments here are far more straightforward. This is due to the fact that we now expect to receive a $\partial_\mu^2 u$ term within the expansion, so we should assume the respective regularity, whilst the same situation in the proof of Theorem 4.1.3 requires another round of mollification.

Proof of Theorem 4.2.2. Step 1. Mollification. We carry out mollification in two steps - firstly we construct the mollifying sequence and later show its convergence. As in the proof of Theorem 4.1.3, we pick a smooth function $\rho:\mathbb{R}^d\to\mathbb{R}^d$ with compact support, letting for any $t\in[0,T]$, $(u\star\rho)_t(\mu) := u_t(\mu\circ\rho^{-1})$ and for any $t\in[0,T]$ having u $\mathbb{P}$-a.s. bounded and continuous at every pair $(t,\mu)\in\mathcal{P}_2(\mathbb{R}^d)$, $\partial_\mu u$ and $\partial_v\partial_\mu u$ $\mathbb{P}$-a.s. bounded and continuous at every triple (t,μ,v) for $v\in$ Supp(μ) and $\partial_\mu^2 u$ $\mathbb{P}$-a.s. bounded and continuous at every quadruple (t,μ,v,v') for $v,v'\in$ Supp(μ), what follows from local boundedness of u and its derivatives. Now picking the sequence $(\rho_n)_{n\geqslant 1}$ in a way that $(\rho_n,\partial_x\rho_n,\partial_{xx}^2\rho_n)(x)\to(x,I_d,0)$ as $n\to\infty$, we can conclude that $(u\star\rho_n)_t(\mu),\partial_\mu(u\star\rho_n)_t(\mu,v),\partial_v\partial_\mu(u\star\rho_n)_t(\mu,v)$ and $\partial_\mu^2(u\star\rho_n)_t(\mu,v,v')$ converge $\mathbb{P}$-a.s. to $u_t(\mu),\partial_\mu u_t(\mu,v)$, $\partial_v\partial_\mu u_t(\mu,v)$ and $\partial_\mu^2 u_t(\mu,v,v')$ respectively. Thus we can assume u and its derivatives to be $\mathbb{P}$-a.s. bounded.

Again as in Theorem 4.1.3 we consider $\mu\mapsto(u\star\rho)(\mu*\phi_G)$ instead of $\mu\mapsto(u\star\rho)(\mu)$ with ϕ_G - density of standard d-dimensional Gaussian distribution $N(0,I_d)$ on $\mathbb{R}^d$ and $(\mu*\phi_G)(x) := \int_{\mathbb{R}^d}\phi_G(x-y)d\mu(y)$. Now the support of $\mu*\phi_G$ is the whole $\mathbb{R}^d$ and $\partial_\mu u,\partial_v\partial_\mu u$ and $\partial_\mu^2 u$ are $\mathbb{P}$-a.s. continuous at every triple $(t,\mu\circ\phi_G,v),t\in[0,T],v\in\mathbb{R}^d$. Installing $\phi_{\varepsilon,G}$ - d-dimensional Gaussian distribution $N(0;\varepsilon I_d)$ and letting $\varepsilon\searrow 0$, we conclude the $\mathbb{P}$-a.s. convergence of $\partial_\mu u_t(\mu*\phi_{\varepsilon,G},v)$, $\partial_v\partial_\mu u_t(\mu*\phi_{\varepsilon,G},v)$ and $\partial_\mu^2 u_t(\mu*\phi_{\varepsilon,G},v,v')$ to $\partial_\mu u_t(\mu,v)$, $\partial_v\partial_\mu u_t(\mu,v)$ and $\partial_\mu^2 u_t(\mu,v,v')$ respectively. Thus we can assume $\mathbb{P}$-a.s. uniform continuity of measure expansion terms for the whole $\mathbb{R}^d$.

Now we are to show that mollification procedure is well-posed. It is straightforward to verify that $u_n := u\star\rho_n$ satisfies $\mathbb{P}$-a.s. (4.2.2) uniformly in $n\geqslant 1$. Applying twice the dominated convergence theorem we conclude the $\mathbb{P}$-a.s. convergence for all the terms but the stochastic integral. To handle the latter one additionally requires an argument across the quadratic variation, as written in Theorem 3.2.1 and localisation.

As before we define $\phi_t=\psi_t^0=\psi_t^1:=0$, for those t where the functions are not well-defined. We copy the procedure above to conclude that ϕ,ψ^0,ψ^1 $\mathbb{P}$-a.s. have compact support.

Step 2. Approximation. By our mollification argument one can assume the $u,\partial_\mu u,\partial_v\partial_\mu u,\partial_x\partial_\mu$ and $\partial_\mu^2 u$ to be $\mathbb{P}$-a.s. bounded and $\mathbb{P}$-a.s. uniformly continuous in respective topology spaces. We construct twin processes $(Y_t^l)_{t\in[0,T]}$, $l=1,\dots,N$ of $(Y_t)_{t\in[0,T]}$ each supporting its own independent Brownian motion $(W_t^{1,l})_{t\in[0,T]}$ that generate $(\Omega^{1,l},\mathcal{F}^{1,l},\mathbb{F}^{1,l},\mathbb{P}^{1,l})$ alongside with $\mathcal{F}_0^l$, altogether forming a copy of $(\Omega^1,\mathcal{F}^1,\mathbb{F}^1,\mathbb{P}^1)$. Since the stochastic basis $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ of our initial space is constructed as a completion of $(\Omega^0\times\Omega^1,\mathcal{F}^0\otimes\mathcal{F}^1,\mathbb{F}^0\otimes\mathbb{F}^1,\mathbb{P}^0\otimes\mathbb{P}^1)$ augmented in a right-continuous way and then completed, we introduce a new product basis $(\Omega^l,\mathcal{F}^l,\mathbb{F}^l,\mathbb{P}^l)$ to be completion of $(\Omega^0\times\Omega^{1,l},\mathcal{F}^0\otimes\mathcal{F}^{1,l},\mathbb{F}^0\otimes\mathbb{F}^{1,l},\mathbb{P}^0\otimes\mathbb{P}^{1,l})$ augmented in right-continuous way and then completed. Now we copy the dynamics of $(Y_t)_{t\in[0,T]}$, as

$$dY_t^l = b_t^l\mathrm{d}t + \sigma_t^{0,l}\mathrm{d}W_t^0\mathrm{d}t + \sigma_t^{1,l}\mathrm{d}W_t^{1,l},\quad Y_0^l = Y_0^l,$$

where $Y_0^l, b_t^l, \sigma_t^{0,l}$ and $\sigma_t^{1,l}$ are copies of $Y^0, b_t, \sigma_t^0, \sigma_t^1$ respectively. Now we construct a total stochastic basis $(\Omega^{1,\dots,N},\mathcal{F}^{1,\dots,N},\mathbb{F}^{1,\dots,N},\mathbb{P}^{1,\dots,N})$, where

$$\Omega^{1,\dots,N} = \Omega^0\times\Omega^1\times\prod_{l=1}^N\Omega^{1,l},\quad \mathcal{F}^{1,\dots,N} = \mathcal{F}^0\otimes\mathcal{F}^1\otimes\bigotimes_{l=1}^N\mathcal{F}^{1,l},$$
$$\mathbb{F}^{1,\dots,N} = \mathbb{F}^0\otimes\mathbb{F}^1\otimes\bigotimes_{l=1}^N\mathbb{F}^{1,l},\quad \mathbb{P}^{1,\dots,N} = \mathbb{P}^0\otimes\mathbb{P}^1\otimes\bigotimes_{l=1}^N\mathbb{P}^{1,l},$$

where we again and finally augment the filtration in a right-continuous way and complete. We underline that processes $\big((Y_t^l)(\omega^0,\cdot), b_t^l(\omega^0,\cdot), \sigma_t^{0,l}(\omega^0,\cdot), \sigma_t^{1,l}(\omega^0,\cdot)\big)_{t\in[0,T]}$, $l=1,\dots N$ are i.i.d. $\mathbb{P}^0$-a.s.

Hereinafter while fixing the $\omega^0\in\Omega^0$, and for the sake of simplicity we will omit adding the $(\omega^0,\cdot)$ to the processes $Y_t,\ b_t,\ \sigma_t^0,\sigma_t^1$ to highlight the respective relation to ω^0, but will leave in after $\bar\mu_t^N$ as to underline the nature of this dependency.

Denoting the flow of marginals for almost all $\omega^0\in\Omega^0$ as $\bar\mu_t^N(\omega^0,\cdot) := \frac{1}{N}\sum_{l=1}^N\delta_{Y_t^l(\omega^0,\cdot)}$ for $t\in[0,T]$ and the empirical projection of u as u^N we proceed by applying Itô-Wentzell formula (Theorem 3.2.1) to u_t^N and using Proposition 3.3.5 to expand $\mathbb{P}$-a.s.

$$\begin{aligned}
u_t^N(Y_t^1,\dots,Y_t^N) - u_0^N(Y_0^1,\dots,Y_0^N) &= \int_0^t\phi_s(\bar\mu_s^N(\omega^0,\cdot))\mathrm{d}s + \int_0^t\psi_s^0(\bar\mu_s^N(\omega^0,\cdot))\cdot\mathrm{d}W_s^0\\
&+ \int_0^t\psi_s^1(\bar\mu_s^N(\omega^0,\cdot))\cdot\mathrm{d}W_s^1 + \frac{1}{N}\sum_{l=1}^N\int_0^t\partial_\mu u_s(\bar\mu_s^N(\omega^0,\cdot),Y_s^l)\cdot b_s^l\mathrm{d}s\\
&+ \frac{1}{N}\sum_{l=1}^N\int_0^t\partial_\mu u_s(\bar\mu_s^N(\omega^0,\cdot),Y_s^l)\cdot\sigma_s^{0,l}\mathrm{d}W_s^0 + \frac{1}{N}\sum_{l=1}^N\int_0^t\partial_\mu u_s(\bar\mu_s^N(\omega^0,\cdot),Y_s^l)\cdot\sigma_s^{1,l}\mathrm{d}W_s^{1,l}\\
&+ \frac{1}{2N}\sum_{l=1}^N\int_0^t\mathrm{Trace}\big\{\partial_v\partial_\mu u_s(\bar\mu_s^N(\omega^0,\cdot),Y_s^l)(\sigma_s^{0,l}(\sigma_s^{0,l})^\intercal + \sigma_s^{1,l}(\sigma_s^{1,l})^\intercal)\big\}\mathrm{d}s\\
&+ \frac{1}{2N^2}\sum_{l,l'=1}^{N,N}\int_0^t\mathrm{Trace}\big\{\partial_\mu^2 u_s(\bar\mu_s^N(\omega^0,\cdot),Y_s^l,Y_s^{l'})\ \sigma_s^{0,l}(\sigma_s^{0,l'})^\intercal\big\}\mathrm{d}s\\
&+ \frac{1}{2N^2}\sum_{l=1}^N\int_0^t\mathrm{Trace}\big\{\partial_\mu^2 u_s(\bar\mu_s^N(\omega^0,\cdot),Y_s^l,Y_s^l)\ \sigma_s^{1,l}(\sigma_s^{1,l})^\intercal\big\}\mathrm{d}s\\
&+ \frac{1}{N}\sum_{l=1}^N\int_0^t\mathrm{Trace}\big\{\partial_\mu\psi_s^0(\bar\mu_s^N(\omega^0,\cdot),Y_s^l)(\sigma_s^{0,l})^\intercal\big\}\mathrm{d}s.
\end{aligned}$$

We highlight that we do not have $\partial_\mu\psi^1$ terms due to the fact that $\langle W^1,W^{1,l}\rangle_t = 0,\ l = 1,\dots,N$, whilst one of the $\partial_\mu^2 u$ terms is summed up diagonally, due to independence of $W^i,W^j,\ i,j\in 1,\dots,N,\ i\neq j$. Taking conditional expectations on the above formula $\mathbb{E}^{1,1,\dots,N}\big[\cdot\big] := \mathbb{E}^{\mathbb{P}^{1,\dots,N}}\big[\cdot \,|\, \mathcal{F}^0\otimes\mathcal{F}^1\big]$ we have by the stochastic Fubini theorem (see [91, Theorem 3.5]) and boundedness of $\partial_\mu^2 u$ for any $t \in [0,T]$, $\mathbb{P}$-a.s.

$$\begin{aligned}
&\mathbb{E}^{1,1,\dots,N}\Big[u_t(\bar{\mu}_T^N(\omega^0,\cdot))\Big] - \mathbb{E}^{1,1,\dots N}\Big[u_0(\bar{\mu}_0^N)\Big] = \mathbb{E}^{1,1,\dots,N}\Big[\int_0^t \phi_s(\bar{\mu}_s^N(\omega^0,\cdot))\mathrm{d}s\Big]\\
&+ \mathbb{E}^{1,1,\dots,N}\Big[\int_0^t \psi_s^0(\bar{\mu}_s^N(\omega^0,\cdot))\cdot \mathrm{d}W_s^0\Big] + \mathbb{E}^{1,1,\dots,N}\Big[\int_0^t \psi_s^1(\bar{\mu}_s^N(\omega^0,\cdot))\cdot \mathrm{d}W_s^1\Big]\\
&+ \int_0^t \mathbb{E}^{1,1,\dots N}\Big[\partial_\mu u_s(\bar{\mu}_s^N(\omega^0,\cdot),Y_s^1)\cdot b_s^1\Big]\mathrm{d}s\\
&+ \int_0^t \mathbb{E}^{1,1,\dots N}\Big[(\sigma_s^{0,1})^\intercal \partial_\mu u_s(\bar{\mu}_s^N(\omega^0,\cdot),Y_s^1)\Big]\cdot \mathrm{d}W_s^0 \qquad (4.2.4)\\
&+ 0 + \frac{1}{2}\int_0^t \mathbb{E}^{1,1,\dots N}\Big[\text{Trace}\big\{\partial_v\partial_\mu u_s(\bar{\mu}_s^N(\omega^0,\cdot),Y_s^1)(\sigma_s^{0,1}(\sigma_s^{0,1})^\intercal + (\sigma_s^{1,1})(\sigma_s^{1,1})^\intercal)\big\}\Big]\mathrm{d}s\\
&+ \frac{1}{2}\int_0^t \mathbb{E}^{1,1,\dots N}\Big[\text{Trace}\big\{\partial_\mu^2 u_s(\bar{\mu}_s^N(\omega^0,\cdot),Y_s^1,Y_s^2)\ \sigma_s^{0,1}(\sigma_s^{0,2})^\intercal\big\}\Big]\mathrm{d}s\\
&+ \int_0^t \mathbb{E}^{1,1,\dots N}\Big[\text{Trace}\big\{\partial_\mu \psi_s^0(\bar{\mu}_s^N(\omega^0,\cdot),Y_s^1)(\sigma_s^{0,1})^\intercal\big\}\Big]\mathrm{d}s + O\big(\frac{1}{N}\big),
\end{aligned}$$

with $O(1/N)$ standing for the Bachmann-Landau big-O notation (sequence bounded by $\frac{C}{N}$, for some $C \geqslant 0$) which appears from the second $1/N^2$ summation term (notice the sum is over only one index).

Lifting to L_2-space and using continuity of the underlying process, as in [22, Theorem 4.14], we conclude that $\mathbb{P}^0 \otimes \mathbb{P}^1$-a.s.

$$\limsup_{N\to\infty} \mathbb{E}^{1,1,\dots N}\Big[\sup_{0\leqslant s\leqslant T} W_2\big(\bar{\mu}_s^N(\omega^0,\cdot),\mu_s(\omega^0,\cdot)\big)^2\Big] = 0.$$

Now due to the continuity in the measure-component and dominated convergence theorem we can pass to the limit in (4.2.4) (as $N\to\infty$) to conclude the formula. The convergence of stochastic integral is secured by localisation and arguing across quadratic variation. We swap the integral and expectation by stochastic Fubini theorem. Finally we rewrite the expectations in the RHS upon dependance on two model particles (living on $(\Omega^0\times\tilde{\Omega}^1, \mathcal{F}^0\otimes\tilde{\mathcal{F}}^1, \mathbb{F}^0\otimes\tilde{\mathbb{F}}^1, \mathbb{P}^0\otimes\tilde{\mathbb{P}}^1)$ and $(\Omega^0\times\hat{\Omega}^1, \mathcal{F}^0\otimes\hat{\mathcal{F}}^1, \mathbb{F}^0\otimes\hat{\mathbb{F}}^1, \mathbb{P}^0\otimes\hat{\mathbb{P}}^1)$ respectively). Measurability is again secured by Remark 3.6.5. □

4.2.2 The joint chain rule

Now we are ready to prove a joint chain rule for $u : \Omega\times[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ as given by (4.2.1). We introduce minimal regularity requirements.

Definition 4.2.3. *We say the random field $u : \Omega\times[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ given in* (4.2.1) *is* RF-Joint-Generally-$\mathcal{C}^2$ *if*

i) u *is* RF-Joint-Partially-$\mathcal{C}^2$ *for $\psi_t := (\psi_t^0,\psi_t^1)^\intercal$ and $W_t := (W_t^0, W_t^1)^\intercal$;*

ii) For almost any $t\in[0,T]$, the maps $(x,\mu)\mapsto\phi_t(x,\mu)$, $(x,\mu)\mapsto\psi_t^0(x,\mu)$, $(x,\mu)\mapsto\psi_t^1(x,\mu)$, are $\mathbb{P}$-a.s. joint-continuous in product topology of $\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ at every pair $(x,\mu)\in\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$;

iii) For almost all $t\in[0,T]$, the map $\mu\mapsto\psi_t^0(\mu)$ is L-differentiable $\mathbb{P}$-a.s. at every point $\mu\in\mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\partial_\mu\psi_t^0(\mu,v)$ is $\mathbb{P}$-a.s. joint-continuous at every pair (μ,v), $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, $v\in$ Supp(μ), $\mathbb{P}$-a.s.;

iv) For any $(t,\mu,v)\in[0,T]\times\mathcal{P}_2(\mathbb{R}^d)\times$ Supp(μ), the map $x\mapsto\partial_\mu u_t(x,\mu,v)$ is $\mathbb{R}^d$-differentiable $\mathbb{P}$-a.s. at every point $x\in\mathbb{R}^d$. Moreover, $\partial_x\partial_\mu u_t(x,\mu,v)$ is $\mathbb{P}$-a.s. joint-continuous at every quadruple (t,x,μ,v), with $(t,x,\mu)\in[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, $v\in$ Supp(μ), $\mathbb{P}$-a.s.;

v) For any $(t,x,v) \in [0,T] \times \mathbb{R}^d \times Supp(\mu)$, the map $\mu \mapsto \partial_\mu u_t(x,\mu,v)$ is L-differentiable $\mathbb{P}$-a.s. at every point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, $\partial^2_\mu u_t(x,\mu,v,v')$ is $\mathbb{P}$-a.s. joint-continuous at every quintuple (t,x,μ,v,v'), with $(t,x,\mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d),\ v,v' \in Supp(\mu),\ \mathbb{P}$-a.s..

We highlight again the slight abuse of notation in the way point i) in the above Definition 4.2.3 is formulated. This avoids re-stating a full assumption that is nonetheless clear to understand.

Theorem 4.2.4. *Let u be RF-Joint-Generally-$\mathcal{C}^2$ Itô random field* (4.2.1). *Assume for any compact $K \subset \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ that*

$$\int_0^T \sup_{(x,\mu)\in K} \Big\{ |\phi_s(x,\mu)| + |\psi^0_s(x,\mu)|^2 + |\partial_x \psi^0_s(x,\mu)|^2 + \int_{\mathbb{R}^d} |\partial_\mu \psi^0_s(x,\mu,v)|^2 \mu(dv) + |\psi^1_s(x,\mu)|^2 + |\partial_x \psi^1_s(x,\mu)|^2 \Big\} ds < \infty, \quad \mathbb{P}\text{-a.s.},$$

and

$$\sup_{(t,x,\mu)\in[0,T]\times K} \Big\{ \int_{\mathbb{R}^d} \Big[|\partial_\mu u_t(x,\mu,v)|^2 + |\partial_v \partial_\mu u_t(x,\mu,v)|^2 + |\partial_x \partial_\mu u_t(x,\mu,v)|^2 \Big] \mu(dv) + \int_{\mathbb{R}^d\times\mathbb{R}^d} |\partial^2_\mu u_t(\mu,v,v')|^2 \mu(dv)\mu(dv') \Big\} < \infty, \quad \mathbb{P}\text{-a.s.}.$$

Take $(\mu_t)_{t\in[0,T]} = \big(Law(Y_t(\omega^0,\cdot))\big)_{t\in[0,T]}$ with $(Y_t)_{t\in[0,T]}$ solution to (3.6.1) *under Assumption 3.6.1 and $(X_t)_{t\in[0,T]}$ given by* (3.6.2) *under Assumption 3.6.2.*

Then $(u_t(X_t,\mu_t))_{t\in[0,T]}$ is an Itô process $\mathbb{P}$-a.s. satisfying the expansion

$$\begin{aligned}
u_T(X_T,\mu_T) - u_0(X_0,\mu_0) &= \int_0^T \phi_s(X_s,\mu_s)ds + \int_0^T \psi^0_s(X_s,\mu_s)\cdot dW^0_s + \int_0^T \psi^1_s(X_s,\mu_s)\cdot dW^1_s \\
&+ \int_0^T \partial_x u_s(X_s,\mu_s)\cdot \beta_s ds + \int_0^T \partial_x u_s(X_s,\mu_s)\cdot \gamma^0_s dW^0_s + \int_0^T \partial_x u_s(X_s,\mu_s)\cdot \gamma^1_s dW^1_s \\
&+ \int_0^T \frac{1}{2}\, Trace\big\{\partial^2_{xx} u_s(X_s,\mu_s)(\gamma^0_s(\gamma^0_s)^\intercal + \gamma^1_s(\gamma^1_s)^\intercal)\big\} ds \\
&+ \int_0^T \tilde{\mathbb{E}}^1\Big[\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)\cdot \tilde{b}_s\Big] ds + \int_0^T \tilde{\mathbb{E}}^1\Big[(\tilde{\sigma}^0_s)^\intercal \partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)\Big]\cdot dW^0_s \\
&+ \int_0^T \frac{1}{2}\, \tilde{\mathbb{E}}^1\Big[Trace\big\{\partial_v\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)(\tilde{\sigma}^0_s(\tilde{\sigma}^0_s)^\intercal + \tilde{\sigma}^1_s(\tilde{\sigma}^1_s)^\intercal)\big\}\Big] ds \qquad (4.2.5) \\
&+ \int_0^T \frac{1}{2}\, \hat{\mathbb{E}}^1\Big[\tilde{\mathbb{E}}^1\Big[Trace\big\{\partial^2_\mu u_s(X_s,\mu_s,\tilde{Y}_s,\hat{Y}_s)\ \tilde{\sigma}^0_s(\hat{\sigma}^0_s)^\intercal\big\}\Big]\Big] ds \\
&+ \int_0^T \tilde{\mathbb{E}}^1\Big[Trace\big\{\partial_x\partial_\mu u_s(X_s,\mu_s,\tilde{Y}_s)\ \gamma^0_s(\tilde{\sigma}^0_s)^\intercal\big\}\Big] ds + \int_0^T Trace\big\{\partial_x \psi^0_s(X_s,\mu_s)(\gamma^0_s)^\intercal\big\} ds \\
&+ \int_0^T Trace\big\{\partial_x \psi^1_s(X_s,\mu_s)(\gamma^1_s)^\intercal\big\} ds + \int_0^T \tilde{\mathbb{E}}^1\Big[Trace\big\{\partial_\mu \psi^0_s(X_s,\mu_s,\tilde{Y}_s)(\tilde{\sigma}^0_s)^\intercal\big\}\Big] ds,
\end{aligned}$$

where the formula above $\tilde{\mathbb{E}}$ and $\hat{\mathbb{E}}$ denote the expectation acting on the model twin spaces $(\tilde{\Omega},\tilde{\mathbb{F}},\tilde{\mathbb{P}})$ and $(\hat{\Omega},\hat{\mathbb{F}},\hat{\mathbb{P}})$ respectively, and let the processes $(\tilde{Y}_t,\tilde{b}_t,\tilde{\sigma}_t)_{t\in[0,T]}$ and $(\hat{Y}_t,\hat{b}_t,\hat{\sigma}_t)_{t\in[0,T]}$ be the independent twin processes of $(Y_t,b_t,\sigma_t)_{t\in[0,T]}$ respectively living within.

It is interesting to mention the $\partial_x\partial_\mu u$ term, which also appears in Theorem 3.6.4, it is nothing else but the cross-variation of the process X and the model particle $\tilde{Y}$. The very last two lines contain all the possible ways of cross-interactions, namely, between the random field u, the process X and the random measure μ.

Remark 4.2.5. *According to* [30], *$\partial_x\partial_\mu u = \partial_\mu\partial_x u$, when both crossed derivatives exist and are Lipschitz. However, as one can notice in the proof, within our mollification procedure for empirically* *projected mapping the space derivatives could be swapped in the convenient way to secure the existence of the limit - desired derivative. Thus in the Definition 4.2.3 one can equally demand the existence of $\partial_\mu \partial_x u$ instead of $\partial_x \partial_\mu$. The same applies for respective derivatives for ϕ, ψ^0 and ψ^1.*

The proof used in Theorem 4.1.8 does not carry directly to this case, crucially due to the passage to the limit in (4.1.15) as the measure flow is random. Of the possible angles of attack to show the result the direct application of the empirical projection approach is the simplest. We follow it and provide alternative arguments when the passage to the limit issue arises.

Proof of Theorem 4.2.4. In view of the proof of Theorem 4.2.2 we assume a compactification & mollification argument in the measure component as been applied and hence we do not repeat its construction. Moreover, without loss of generality assume $(b_t)_{t\in[0,T]}$ and $(\sigma_t)_{t\in[0,T]}$ to be bounded.

Again as in previous theorem, we consider u^N- empirical projection of u, and construct generic $(Y_t^l)_{t\in[0,T]}$ in the same way, underlining that the processes $Y_t^l(\omega^0,\cdot)$, $b_t^1(\omega^0,\cdot)$, $\sigma_t^0(\omega^0,\cdot)$, $\sigma_t^1(\omega^0,\cdot)$, $t \in [0,T]$, $l = 1,\dots N$ are i.i.d. For $\phi_t,\ \psi_t := \left(\begin{smallmatrix} \psi_t^0 & 0 \\ 0 & \psi_t^1 \end{smallmatrix}\right)$ and $W_t := (W_t^0, W_t^1)^\intercal$ we copy the same procedure as before to have for almost all t, $\mathbb{P}$-a.s.

$$\mathbb{E}^{1,1,\dots,N}\big[\phi_t^N(X_t, Y_t^1,\dots,Y_t^N)\big] \to \phi_t(X_t,\mu_t) \quad \text{and} \quad \mathbb{E}^{1,1,\dots,N}\big[\psi_t^N(X_t, Y_t^1,\dots,Y_t^N)\big] \to \psi_t(X_t,\mu_t),$$

as $N \to \infty$, together with

$$\begin{aligned}
\mathbb{E}^{1,1,\dots,N}\Big[\int_0^t \phi_s^N(X_s, Y_s^1,\dots,Y_s^N)\mathrm{d}s\Big] &\to \int_0^t \phi_s(X_s,\mu_s)\mathrm{d}s,\\
\mathbb{E}^{1,1,\dots,N}\Big[\int_0^t \psi_s^N(X_s, Y_s^1,\dots,Y_s^N)\cdot \mathrm{d}W_s\Big] &\to \int_0^t \psi_s(X_s,\mu_s)\cdot \mathrm{d}W_s,
\end{aligned}$$

$\mathbb{P}$-a.s. as $N \to \infty$ for all $t \in [0,T]$. As before, for the sake of simplicity we omit adding the $(\omega^0,\cdot)$ to the processes $Y_t,\ b_t,\ \sigma_t^0, \sigma_t^1$, but will leave one for $\bar{\mu}_t^N$.

Since all conditions of Theorem 3.2.1 hold we apply it to $u^N_t(X_t, Y^1_t, \dots, Y^N_t)$ getting

$$
\begin{aligned}
u^N(X_T, Y^1_T, \dots, Y^N_T) - u^N(X_0, Y^1_0 \dots, Y^N_0) &= \int_0^T \phi^N_s(X_s, Y^1_s, \dots, Y^N_s)\mathrm{d}s \\
&+ \int_0^T \psi^{0,N}_s(X_s, Y^1_s, \dots, Y^N_s) \cdot \mathrm{d}W^0_s + \int_0^T \psi^{1,N}_s(X_s, Y^1_s, \dots, Y^N_s) \cdot \mathrm{d}W^1_s \\
&+ \int_0^T \partial_x u^N_s(X_s, Y^1_s, \dots, Y^N_s) \cdot \beta^1_s \mathrm{d}s + \int_0^T \partial_x u^N_s(X_s, Y^1_s, \dots, Y^N_s) \cdot \gamma^0_s \mathrm{d}W^0_s \\
&+ \int_0^T \partial_x u^N_s(X_s, Y^1_s, \dots, Y^N_s) \cdot \gamma^1_s \mathrm{d}W^1_s + \frac{1}{2}\int_0^T \mathrm{Trace}\big\{\partial_{xx} u^N_s(X_s, Y^1_s, \dots, Y^N_s)\ \gamma^0_s(\gamma^0_s)^\intercal\big\}\mathrm{d}s \\
&+ \frac{1}{2}\int_0^T \mathrm{Trace}\big\{\partial_{xx} u^N_s(X_s, Y^1_s, \dots, Y^N_s)\ \gamma^1_s(\gamma^1_s)^\intercal\big\}\mathrm{d}s \\
&+ \int_0^T \mathrm{Trace}\big\{\partial_x \psi^{0,N}_s(X_s, Y^1_s, \dots, Y^N_s)(\gamma^0_s)^\intercal\big\}\mathrm{d}s + \int_0^T \mathrm{Trace}\big\{\partial_x \psi^{1,N}_s(X_s, Y^1_s, \dots, Y^N_s)(\gamma^1_s)^\intercal\big\}\mathrm{d}s \\
&+ \sum_{l=1}^N \int_0^T \partial_{y^l} u^N_s(X_s, Y^1_s, \dots, Y^N_s) \cdot \beta^1_s \mathrm{d}s + \sum_{l=1}^N \int_0^T \partial_{y^l} u^N_s(X_s, Y^1_s, \dots, Y^N_s) \cdot \sigma^{0,l}_s \mathrm{d}W^0_s \\
&+ \sum_{l=1}^N \int_0^T \partial_{y^l} u^N_s(X_s, Y^1_s, \dots, Y^N_s) \cdot \sigma^{1,l}_s \mathrm{d}W^{1,l}_s \\
&+ \frac{1}{2}\sum_{l,l'=1}^{N,N} \int_0^T \mathrm{Trace}\big\{\partial_{y^l y^{l'}} u^N_s(X_s, Y^1_s, \dots, Y^N_s)\ \sigma^{0,l}_s(\sigma^{0,l'}_s)^\intercal\big\}\mathrm{d}s \\
&+ \frac{1}{2}\sum_{l=1}^{N} \int_0^T \mathrm{Trace}\big\{\partial_{y^l y^{l'}} u^N_s(X_s, Y^1_s, \dots, Y^N_s)\ \sigma^{1,l}_s(\sigma^{1,l}_s)^\intercal\big\}\mathrm{d}s \\
&+ \sum_{l=1}^{N} \int_0^T \mathrm{Trace}\big\{\partial_{x y^l} u^N_s(X_s, Y^1_s, \dots, Y^N_s)\ \gamma^0_s(\sigma^{0,l}_s)^\intercal\big\}\mathrm{d}s \\
&+ \sum_{l=1}^{N} \int_0^T \mathrm{Trace}\big\{\partial_{y^l} \psi^{0,N}_s(X_s, Y^1_s, \dots, Y^N_s)(\sigma^{0,l}_s)^\intercal\big\}\mathrm{d}s.
\end{aligned}
$$

We again underline that we do not have additional $\partial_{xy^l}u$ and $\partial_{y^l}\psi$ terms due to the fact that $\langle W^1, W^{1,l}\rangle_t = 0,\ l = 1, \dots, N$, at the same time diagonally summing one of $\partial^2_\mu u$, due to mutual independence of $W^i, W^j,\ i, j \in 1, \dots, N,\ i \neq j$.

Now we transform the equation according to Proposition 3.3.5, and applying $\mathbb{E}^{1,1,\dots,N}\big[\,\cdot\,\big] :=$

$\mathbb{E}\big[\,\cdot\,|\mathcal{F}^0\otimes\mathcal{F}^1\big]$, law of large numbers, Fubini theorem and boundedness of $\partial_\mu^2 u$ we get $\mathbb{P}$-a.s

$$
\begin{aligned}
&\mathbb{E}^{1,1,\dots,N}\big[u(X_T,\bar\mu_T^N)\big]-\mathbb{E}^{1,1,\dots,N}\big[u^N(X_0,\bar\mu_0^N)\big]=\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\phi_s^N(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\mathrm{d}s\\
&\quad+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\psi_s^{0,N}(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\cdot\mathrm{d}W_s^0+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\psi_s^{1,N}(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\cdot\mathrm{d}W_s^1\\
&\quad+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\partial_x u_s(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\cdot\beta_s^1\mathrm{d}s+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\partial_x u_s(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\cdot\gamma_s^0\mathrm{d}W_s^0\\
&\quad+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\partial_x u_s(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\cdot\gamma_s^1\mathrm{d}W_s^1\\
&\quad+\int_0^T\frac{1}{2}\,\mathbb{E}^{1,1,\dots,N}\Big[\mathrm{Trace}\big\{\partial_{xx}u_s(X_s,\bar\mu_s^N(\omega^0,\cdot))(\gamma_s^0(\gamma_s^0)^\intercal+\gamma_s^1(\gamma_s^1)^\intercal)\big\}\Big]\mathrm{d}s\\
&\quad+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\partial_x\psi_s^0(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\gamma_s^0\mathrm{d}s+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\partial_x\psi_s^1(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\gamma_s^1\mathrm{d}s\\
&\quad+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\partial_\mu u_s(X_s,\bar\mu_s^N(\omega^0,\cdot),Y_s^1)\cdot\beta_s^1\Big]\mathrm{d}s+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[(\sigma_s^{0,1})^\intercal\partial_\mu u_s(X_s,\bar\mu_s^N(\omega^0,\cdot))\Big]\cdot\mathrm{d}W_s^0\\
&\quad+\int_0^T\frac{1}{2}\,\mathbb{E}^{1,1,\dots,N}\Big[\mathrm{Trace}\big\{\partial_v\partial_\mu u_s(X_s,\bar\mu_s^N(\omega^0,\cdot),Y_s^1)(\sigma_s^{0,1}(\sigma_s^{0,1})+\sigma_s^{1,1}(\sigma_s^{1,1})^\intercal)\big\}\Big]\mathrm{d}s\\
&\quad+\int_0^T\frac{1}{2}\,\mathbb{E}^{1,1,\dots,N}\Big[\mathrm{Trace}\big\{\partial_\mu^2 u_s(X_s,\bar\mu_s^N(\omega^0,\cdot),Y_s^1,Y_s^2)\;\sigma_s^{0,1}(\sigma_s^{0,2})^\intercal\big\}\Big]\mathrm{d}s\\
&\quad+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\mathrm{Trace}\big\{\partial_x\partial_\mu u_s(X_s,\bar\mu_s^N(\omega^0,\cdot),Y_s^1)\;\gamma_s^0(\sigma_s^{0,1})^\intercal\big\}\Big]\mathrm{d}s\\
&\quad+\int_0^T\mathbb{E}^{1,1,\dots,N}\Big[\mathrm{Trace}\big\{\partial_\mu\psi_s^0(X_s,\bar\mu_s^N(\omega^0,\cdot),Y_s^1)(\sigma_s^{0,1})^\intercal\big\}\Big]\mathrm{d}s+O\big(\frac{1}{N}\big).
\end{aligned}
$$

We note that the expectation taken on the term in the fifth line does not charge the process $(\gamma^0(\gamma^0)^\intercal+\gamma^1(\gamma^1)^\intercal)$, we write it as it is to preserve the matrix-trace notation.

According to the conditional propagation of chaos argument, as given in Theorem 4.2.2, dominated convergence theorem (twice for the terms from the last five lines), localisation for X and joint continuity and integrability of involved terms one can conclude the convergence of the above formula to (4.2.5). We argue additionally across convergence of quadratic variation to handle the stochastic integral terms.

As before we switch to two model particles (living on $(\Omega^0\times\tilde\Omega^1,\mathcal{F}^0\otimes\tilde{\mathcal{F}}^1,\mathbb{F}^0\otimes\tilde{\mathbb{F}}^1,\mathbb{P}^0\otimes\tilde{\mathbb{P}}^1)$ and $(\Omega^0\times\hat\Omega^1,\mathcal{F}^0\otimes\hat{\mathcal{F}}^1,\mathbb{F}^0\otimes\hat{\mathbb{F}}^1,\mathbb{P}^0\otimes\hat{\mathbb{P}}^1)$ respectively) and swap the integral and expectation by stochastic Fubini theorem. Again and finally, we assert the measurability of involved terms by Remark 3.6.5.

□

Part III

Investment-consumption many-player and mean-field games under relative performance concerns

Chapter 5

Preliminaries

5.1 Notation and spaces

Throughout Part III we work on stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, satisfying standard conditions.

For functional $U : \Omega \times \mathbb{R} \times [0, \infty) \to \mathbb{R}$, U_t denotes the partial derivative in the time parameter $t \in [0, \infty)$; U_x denotes the partial derivative in the spatial variable $x \in \mathbb{R}$ whilst U_{xx} the second derivative in $x \in \mathbb{R}$.

5.2 Relative performance concerns for CARA preferences

We recall the market structure outlined in Chapter 2. For the clarity of presentation we present some concepts anew.

Agents' wealth. Each agent $i = 1, \dots, n$ trades using a self-financing strategy, $(\pi_t^i)_{t \geqslant 0}$, which represent the (discounted by the bond) amount invested in the i-th stock. The i^{th} agent's wealth $(X_t^i)_{t \geqslant 0}$ then solves

$$dX_t^i = \pi_t^i \Big(\mu_i \mathrm{d}t + \nu_i \mathrm{d}W_t^i + \sigma_i \mathrm{d}B_t \Big), \quad \text{with} \quad X_0^i = x_0^i \in \mathbb{R}^1. \tag{5.2.1}$$

We recall that the strategy is self-financing, when the agent wealth evolve from the starting capital only by agent's investment decisions in the market without any external sources of income and this evolution is described by respective SDE (5.2.1).

A portfolio strategy is said admissible if it belongs to the set $\mathcal{A}^i$, which consists of

$$\mathcal{A}^i = \Big\{ \pi^i : \mathbb{F}\text{-progressively measurable } \mathbb{R}\text{-valued processes } (\pi_t^i)_{t \geqslant 0}, \\ \text{and self-financing such that } \mathbb{E}[\int_0^t |\pi_s|^2 \mathrm{d}s] < \infty, \text{ for all } t \geqslant 0 \Big\}.$$

The agents' social interaction.

Each manager measures the performance of their strategy taking into account the policy of the other. Each agent engages in a form of social interaction (in the sense of [11,46]) that affects the agent's perception of wealth, all in an additive fashion modelled through the arithmetic average wealth of all agents (this model is largely inspired in [11, 46, 48, 73]). The way the agent assesses and optimises their relative performance is explored through Definition 5.3.1 in the latter Section 5.3. So far we introduce the *relative performance metric* of manager $i \in$

[1] We emphasise that our results can be generalised by allowing for alternative investment in the risk-free bond with rate r. We fully investigate $r \neq 0$ for the CRRA preferences, where we study its interplay with other parameters. Under CARA preferences we assume $r = 0$.

$\{1,\ldots,n\}$, denoted $\widetilde{X}^i$ is defined to be

$$\widetilde{X}^i = X^i - \theta_i \overline{X}, \quad \text{where} \quad \overline{X} := \frac{1}{n}\sum_{k=1}^n X^k \quad \text{and} \quad \theta_i \in [0,1], \tag{5.2.2}$$

where deterministic θ_i stands for the competition weight for agent i. We easily obtain a dynamics for $\overline{X}$ and $\widetilde{X}^i$, namely

$$\begin{aligned}
d\overline{X}_t &= \Big(\frac{1}{n}\sum_{k=1}^n \pi_t^k \mu_k\Big)\mathrm{d}t + \Big(\frac{1}{n}\sum_{k=1}^n \pi_t^k \nu_k \mathrm{d}W_t^k\Big) + \Big(\frac{1}{n}\sum_{k=1}^n \pi_t^k \sigma_k\Big)\mathrm{d}B_t \\
&= \overline{(\pi\mu)}_t \mathrm{d}t + \Big(\frac{1}{n}\sum_{k=1}^n \pi_t^k \nu_k \mathrm{d}W_t^k\Big) + \overline{(\pi\sigma)}_t \mathrm{d}B_t, \quad \overline{X}_0 = \overline{x}_0 = \frac{1}{n}\sum_{k=1}^n x_0^k \\
d\widetilde{X}_t^i &= \big(\pi_t^i \mu_i - \theta_i \overline{(\pi\mu)}_t\big)\mathrm{d}t + \Big(\pi_t^i \nu_i \mathrm{d}W_t^i - \theta_i\big(\frac{1}{n}\sum_{k=1}^n \pi_t^k \nu_k \mathrm{d}W_t^k\big)\Big) \\
&\qquad + \big(\pi_t^i \sigma_i - \theta_i \overline{(\pi\sigma)}_t\big)\mathrm{d}B_t, \quad \widetilde{X}_0^i = x_0^i - \theta_i \overline{x}_0,
\end{aligned} \tag{5.2.3}$$

where $\overline{x}_0$, $\overline{\pi\mu}$ and $\overline{\pi\sigma}$ are identified as averages (as seen from the 1st equation to the 2nd). Similarly to [73, Remark 2.5], it is natural to replace the average wealth $\overline{X}$ in (5.2.2) by the average over all other agents. With that in mind we define for convenience $\overline{X}^{(-i)} = \frac{1}{n-1}\sum_{k\neq i} X^k$ and $Y^{(-i)} = \frac{n}{n-1}\overline{X}^{(-i)}$. This leads us to recast (5.2.2) as

$$\widehat{X}^i = X^i - \theta_i \overline{X}^{(-i)}, \qquad \text{where} \qquad \overline{X}^{(-i)} = \frac{1}{n-1}\sum_{k\neq i} X^k. \tag{5.2.4}$$

We easily obtain a dynamics for $\widehat{X}$ and $\overline{X}^{(-i)}$, namely

$$\begin{aligned}
d\overline{X}_t^{(-i)} &= \overline{(\pi\mu)}_t^{(-i)} \mathrm{d}t + \Big(\frac{1}{n-1}\sum_{k\neq i} \pi_t^k \nu_k \mathrm{d}W_t^k\Big) + \overline{(\pi\sigma)}_t^{(-i)} \mathrm{d}B_t, \quad \overline{X}_0^{(-i)} = \overline{x}_0^{(-i)} \\
d\widehat{X}_t^i &= \big(\pi_t^i \mu_i - \theta_i \overline{(\pi\mu)}_t^{(-i)}\big)\mathrm{d}t + \Big(\pi_t^i \nu_i \mathrm{d}W_t^i - \theta_i\big(\frac{1}{n-1}\sum_{k\neq i}^n \pi_t^k \nu_k \mathrm{d}W_t^k\big)\Big) \\
&\qquad + \big(\pi_t^i \sigma_i - \theta_i \overline{(\pi\sigma)}_t^{(-i)}\big)\mathrm{d}B_t, \quad \widehat{X}_t^i = x_0^i - \theta_i \overline{x}_0^{(-i)}.
\end{aligned} \tag{5.2.5}$$

We also define the quantities

$$\widehat{\pi\sigma}^{(-i)} := \frac{1}{n}\sum_{k\neq i} \pi^k \sigma_k, \quad \overline{(\pi\mu)}^{(-i)} := \frac{1}{n}\sum_{k\neq i} \pi^k \mu_k \ \text{ and } \ \overline{(\pi\nu)^2}^{(-i)} := \frac{1}{n}\sum_{k\neq i} (\pi^k \nu_k)^2,$$

where we have the following relations between $\widehat{\pi\sigma}^{(-i)}$, $\overline{\pi\sigma}^{(-i)}$ and $\overline{\pi\sigma}$:

$$\overline{\pi\sigma}^{(-i)} = \frac{n}{n-1}\overline{\pi\sigma} - \frac{1}{n-1}\pi^i \sigma_i, \quad \overline{\pi\sigma}^{(-i)} = \frac{n}{n-1}\widehat{\pi\sigma}^{(-i)}, \tag{5.2.6}$$

and $\widehat{\pi\sigma}^{(-i)} = \overline{\pi\sigma} - \frac{1}{n}\pi^i \sigma_i$. We do not write it explicitly but we extend the same notation and relations to $\widehat{\pi\mu}^{(-i)}$, $\overline{\pi\mu}^{(-i)}$ and $\overline{\pi\mu}$.

5.3 Forward relative performance criteria for wealth optimisation

5.3.1 Forward relative performance criteria

Each manager measures the output of their relative performance metric using a forward relative one as modelled by an $\mathcal{F}_t$-progressively measurable random field $U^i : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ for $i \in \{1, \dots, n\}$. The below criteria follows those proposed in [52].

The main idea here being a formulation inspired in the first step in the usual strategy of solving a Nash game, namely the best response of an agent to the actions of all other agents. Take manager i and assume all other agents $j \neq i$ have acted with an investment policy π^j then for any strategy $\pi^i \in \mathcal{A}^i$, the process $U^i(\widehat{X}_t^i, t)$ is a (local) supermartingale, and there exists $\pi^{i,*} \in \mathcal{A}^i$ such that $U^i(\widehat{X}_t^{i,*}, t)$ is a (local) martingale where $\widehat{X}^i$ and $\widehat{X}^{i,*}$ solves (5.2.4) with strategies π^i and $\pi^{i,*}$ respectively.

This version of a relative criterion is (implicitly and) exogenously parametrised by the policies of all other managers $j \neq i$ over which there is no assumption on their optimality. In Nash-game language, we solve the so-called best response.

Definition 5.3.1 (Forward relative performance for the manager)**.** *Each manager $i \in \{1, \dots, n\}$ satisfies the following. Let $\pi^j \in \mathcal{A}^j$, for any $j \neq i$ be arbitrary but fixed admissible policies, in other words, the other managers have fixed their admissible strategies.*

An $\mathbb{F}$-progressively measurable random field $U^i(x, t)$ is a forward relative performance *for manager i if, for all $t \geqslant 0$, the following conditions hold:*

i) The mapping $x \mapsto U^i(x, t)$, is $\mathbb{P}$-a.s. strictly increasing and strictly concave;

ii) For any $\pi^i \in \mathcal{A}^i$, $U^i(\widehat{X}_t^i, t)$ is a (local) supermartingale and $\widehat{X}^i$ is the relative performance metric given in (5.2.4)*;*

iii) There exists $\pi^{i,*} \in \mathcal{A}^i$ such that $U^i(\widehat{X}_t^{i,*}, t)$ is a (local) martingale where $\widehat{X}^{i,*}$ solves (5.2.4) *with strategies $\pi^{i,*}$ being used.*

In the above definition, we do not make explicit references to the initial conditions $U^i(x, 0)$ but we assume that admissible initial data exists such that the above definition is viable. Contrary to the classical expected utility case, the forward utility process is an investor-specific input. Once it is chosen, the supermartingale and martingale properties impose conditions on the drift of the process. Under enough regularity, these conditions lead to the forward performance SPDE (see [82]).

Since we are working in a log-normal market, it suffices to study smooth relative performance criteria of zero volatility (of the forward utility map). Such processes are extensively analysed in [80] in the absence of relative performance concerns. There, a concise characterisation of the forward criteria is given along necessary and sufficient conditions for their existence and uniqueness. In that setting, the zero-volatility forward processes are always time-decreasing processes. We point to the reader that this does not have to be case if relative performance concerns are present (see also [52]).

5.4 Relative performance concerns for CRRA preferences with consumption

Mimicking the presentation of Section 5.2 we introduce the setting anew.

The market. The market environment is the same as in 5.2 with one notational change. Now the price $(S_t^i)_{t \geqslant 0}$ of stock i traded exclusively by the i-th agent solves

$$\frac{dS_t^i}{S_t^i} = \breve{\mu}_i \mathrm{d}t + \nu_i \mathrm{d}W_t^i + \sigma_i \mathrm{d}B_t, \quad S_0^i = s_0^i > 0, \tag{5.4.1}$$

with constant parameters $\breve{\mu}_i \in \mathbb{R},\ \sigma_i \geqslant 0$ and $\nu_i \geqslant 0$ with $\sigma_i + \nu_i > 0$. This is in contrast with 5.2, where we set different notation for the drift. The convenience of this notation could be understood through the next paragraphs, when we introduce the risk-free rate r.

We recall the case of *single common stock*, where for any $i = 1, \dots, n,\ (\breve{\mu}_i, \sigma_i) = (\breve{\mu}, \sigma),\ \nu_i = 0$, for some $\breve{\mu} \in \mathbb{R},\ \sigma > 0$ and independent of i.

We again point out that we contribute to the literature on mean field games and FPP by also providing an explicitly solvable example.

Agents' wealth. Each agent $i = 1, \dots, n$ trades using a self-financing strategy, $(\pi^i_t)_{t\geqslant 0}$, representing the fraction of wealth invested in the i-th stock and consumption policy, $(c^i_t)_{t\geqslant 0}$, representing the instantaneous rate of consumption per unit of wealth. The i-th agent's wealth dynamics $(X^i_t)_{t\geqslant 0}$ is given by

$$dX^i_t = rX^i_t dt + \pi^i_t X^i_t\Big(\mu_i dt + \nu_i dW^i_t + \sigma_i dB_t\Big) - c^i_t X^i_t dt, \quad \text{with} \quad X^i_0 = x^i_0 > 0,\ \mu_i = \breve{\mu}_i - r. \tag{5.4.2}$$

We interpret μ_i as an excess return. A portfolio investment-consumption strategy is deemed *admissible* if it belongs to the admissibility set $\mathcal{A}^i$,

$$\mathcal{A}^i = \Big\{(\pi^i, c^i) : \mathbb{F}\text{-progressively measurable } \mathbb{R}\times(0,\infty)\text{-valued process } (\pi^i_t, c^i_t)_{t\geqslant 0},$$
$$\text{such that } \mathbb{E}\Big[\int_0^t (|\pi^i_s|^2 + |c^i_s|^2)\mathrm{d}s\Big] < \infty,\ \text{for any } t > 0\Big\}.$$

As in [72] we do not allow a consumption rate of zero. It is also clear that for any admissible strategy we have $X^i_t > 0$ for all $t \geqslant 0$.

The agents' interaction and relative performance concerns. Each manager measures the performance of their strategy taking into account the policies of the others. Each agent engages in a form of social interaction that affects that agent's perception of wealth, all in a multiplicative fashion modelled through the geometric average wealth of all the agents, excluding themselves. The *relative performance wealth process* of manager $i \in \{1, \dots, n\}$, denoted $\widehat{X}^i$ is defined to be

$$\widehat{X}^i = \frac{X^i}{\Big(\widetilde{X}^{(-i)}\Big)^{\theta_i}}, \qquad \text{where} \qquad \widetilde{X}^{(-i)} = \Big(\prod_{k\neq i}^n X^k\Big)^{\frac{1}{n-1}}. \tag{5.4.3}$$

In the same vein we introduce the *relative consumption metric*

$$\hat{c}^i = \frac{c^i}{\Big(\tilde{c}^{(-i)}\Big)^{\theta_i}}, \quad \text{where} \quad \tilde{c}^{(-i)} = \Big(\prod_{k\neq i}^n c^k\Big)^{\frac{1}{n-1}}. \tag{5.4.4}$$

We emphasise that both relative metrics (hat and tilde) can be equivalently reformulated such that the respective geometric average include the agent themselves. This becomes an equivalent problem that can be reduced to the original one by rescaling the parameters, see [73, Remark 3.3], [72, Section 2] or [36, Section 2].

By an application of Itô's formula we obtain the dynamics for *the average performance wealth* $Y = \widetilde{X}^{(-i)}$ as follows,

$$\frac{\mathrm{d}Y_t}{Y_t} = \left(r + \overline{\mu\pi}_t^{(-i)} - \frac{1}{2}\left(\overline{\Sigma\pi^2_t}^{(-i)} - \big(\overline{\sigma\pi_t}^{(-i)}\big)^2 - \frac{1}{n-1}\overline{(\nu\pi_t)^2}^{(-i)}\right) - \bar{c}_t^{(-i)}\right)\mathrm{d}t$$
$$+ \frac{1}{n-1}\sum_{k\neq i}^n \nu_k \pi^k_t \mathrm{d}W^k_t + \overline{\sigma\pi_t}^{(-i)}\mathrm{d}B_t, \quad Y_0 = \Big(\prod_{k\neq i}^n x^k_0\Big)^{\frac{1}{n-1}},$$

where we define the helpful auxiliary quantities for $t \geqslant 0$

$$\overline{\mu\pi_t}^{(-i)} = \frac{1}{n-1}\sum_{k\neq i}^{n} \mu_k \pi_t^k, \quad \overline{(\nu\pi_t)^2}^{(-i)} = \frac{1}{n-1}\sum_{k\neq i}^{n}(\nu_k\pi_t^k)^2, \quad \overline{\sigma\pi_t}^{(-i)} = \frac{1}{n-1}\sum_{k\neq i}^{n}\sigma_k\pi_t^k$$
$$\overline{\Sigma\pi_t^2}^{(-i)} = \frac{1}{n-1}\sum_{k\neq i}^{n}\Sigma_k(\pi_t^k)^2, \quad \bar{c}_t^{(-i)} = \frac{1}{n-1}\sum_{k\neq i}^{n} c_t^k, \quad \Sigma_k = \sigma_k^2 + \nu_k^2.$$

Via Itô's formula one finds the dynamics of *relative performance wealth* $\widehat{X}^i$ to be

$$\frac{\mathrm{d}\widehat{X}_t^i}{\widehat{X}_t^i} = \xi_i \mathrm{d}t - \left(c_t^i - \theta_i \bar{c}_t^{(-i)}\right) + \left(\nu_i\pi_t^i \mathrm{d}W_t^i - \theta_i\Big(\frac{1}{n-1}\sum_{k\neq i}^{n}\nu_k\pi_t^k\mathrm{d}W_t^k\Big)\right) + \left(\sigma_i\pi_t^i - \theta_i\overline{\sigma\pi_t}^{(-i)}\right)\mathrm{d}B_t, \tag{5.4.5}$$

where

$$\xi_i = r(1-\theta) + \mu_i\pi_t^i - \theta_i\overline{\mu\pi_t}^{(-i)} + \frac{\theta_i}{2}\overline{\Sigma\pi_t^2}^{(-i)} - \frac{\theta_i^2}{2}\left(\left(\overline{\sigma\pi_t}^{(-i)}\right)^2 + \frac{1}{n-1}\overline{(\nu\pi_t)^2}^{(-i)}\right) - \theta_i\sigma_i\pi_t^i\overline{\sigma\pi_t}^{(-i)}.$$

5.5 Forward relative performance criteria for wealth-consumption optimisation

5.5.1 Forward utility of investment and consumption (classical)

We recall the classic *forward utility* formulation in presence of consumption preferences. We define a *forward dynamic utilities* in the context of the probability space $(\Omega, \mathbb{F}, \mathcal{F}, \mathbb{P})$. We denote by $u_0 : \mathbb{R} \to \mathbb{R},\ v_0 : \mathbb{R} \to \mathbb{R}$ the initial data. As a standard forward utility the forward utility of investment and consumption is constructed based on the martingale optimality principle (see [10] and [41]).

Definition 5.5.1 (Forward dynamic utilities of investment and consumption). *Let $Q, U, V : \Omega \times \mathbb{R} \times [0,\infty) \to \mathbb{R}$ be an $\mathbb{F}$-progressively measurable random fields such that $Q(x,t) = U(t,x) + \int_0^t V(cx,s)ds$ for any $c > 0$. Q is a* forward dynamic utility *if*

- *For all $t \geqslant 0$ the map $x \mapsto U(x,t),\ x \mapsto V(x,t)$ is $\mathbb{P}$-a.s. increasing and concave;*
- *It satisfies $Q(x,0) = u_0(x)$;*
- *For all $T \geqslant t$ and each self-financing strategy, represented by (π, c), the associated discounted wealth process X^π satisfies a supermartingale property*
$$\mathbb{E}[Q(X_T^\pi, T)|\mathcal{F}_t] \leqslant Q(X_t^\pi, t) \quad \mathbb{P}\text{-a.s.};$$
- *For all $T \geqslant t$ there exists a self financing strategy, represented by (π^*, c^*), for which the associated discounted wealth X^* satisfies a martingale property*
$$\mathbb{E}[Q(X_T^*, T)|\mathcal{F}_t] = Q(X_t^*, t) \quad \mathbb{P}\text{-a.s.}$$

The above definition assumes the optimiser is attained. We refer the reader to [41] for the discussion on consistency and well-posedness.

5.5.2 Forward relative performance criteria

Each manager $i \in \{1, \ldots, n\}$ measures the output of their relative performance metric using a forward relative utility as modelled by an $\mathcal{F}_t$-progressively measurable random field $Q^i : \Omega \times (0, \infty) \times [0, \infty) \to \mathbb{R}$ for $i \in \{1, \ldots, n\}$. The below criteria follows those proposed in [52] (power-type) and later in [36] (exponential-type) for a forward-performance game.

The formulation here is inspired in the first step of the usual strategy of solving Nash equilibria games, namely, the best response of an agent to the actions of all other agents. This version of a relative criterion is (implicit and) endogenously parametrised by the policies of all other managers $j \neq i$ over which no assumption on their optimality is made.

Definition 5.5.2 (Forward relative performance for the manager)**.** *Each manager $i \in \{1, \ldots, n\}$ satisfies the following conditions: let for any $j \neq i$, $(\pi^j, c^j) \in \mathcal{A}^j$ be arbitrary but fixed and admissible policies, in other words, the other managers have fixed their investment-consumption admissible strategies.*

For $(\pi^i, c^i) \in \mathcal{A}^i$ and subjective discount factor $\rho_i \geqslant 0$, define the $\mathbb{F}$-progressively measurable random field $Q^i : \Omega \times (0, \infty) \times [0, \infty) \to \mathbb{R}$

$$Q^i(x,t) := e^{-\rho_i t} U^i(x,t) + \int_0^t e^{-\rho_i s} V^i(\hat{c}^i_s x, s) ds, \tag{5.5.1}$$

where $\hat{c}^i$ is given by (5.4.4) *and $U^i, V^i : \Omega \times (0, \infty) \times [0, \infty) \to \mathbb{R}$ are two other $\mathbb{F}$-progressively measurable random fields.*

The random field Q^i is a forward relative performance process for i-th manager *if, for all $t \geqslant 0$, the following conditions hold:*

- *The mappings $x \mapsto U^i(x,t)$ and $x \mapsto V^i(x,t)$ are $\mathbb{P}$-a.s. strictly increasing and strictly concave;*
- *For any $(\pi^i, c^i) \in \mathcal{A}^i$, $Q^i(\widehat{X}^i_t, t)$ is a (local) supermartingale, where $\widehat{X}^i$ is the relative performance wealth process given in* (5.4.3)*;*
- *There exists $(\pi^{i,*}, c^{i,*}) \in \mathcal{A}^i$ such that $Q^i(\widehat{X}^{i,*}_t, t)$ is a (local) martingale, where $\widehat{X}^{i,*}$ solves* (5.4.3) *with the strategies $(\pi^{i,*}, c^{i,*})$ being used. The strategy $(\pi^{i,*}, c^{i,*})$ is said to be optimal.*

In the above definition, we do not make explicit references to the initial conditions $U^i(x, 0)$, $V^i(x, 0)$ but we assume that admissible ($\mathcal{F}_0$-measurable) initial data exists such that the above definition is viable. Contrary to the classical expected utility case, the forward performance process is a manager-specific input. Once it is chosen, the supermartingale and martingale properties impose certain conditions on the drift of the process. Under enough regularity, these conditions lead to the forward performance SPDE (see [43, 82]) which, in our case, reduces to a PDE with stochastic coefficients (see Proposition 7.1.3 below).

Remark 5.5.3. *The first term of* (5.5.1) *corresponds to the utility that the agent derives at time t from owning the amount of wealth x. The second term captures the utility the agent accumulates from time 0 up to time t from consuming at the rate of $\hat{c}^i x$. Hereinafter, and for simplicity, we call U the utility from wealth and V the utility from consumption.*

Since we are working in a log-normal market, it suffices to study smooth relative performance criteria of zero volatility (of the FPP map). Such processes are extensively analysed in [80] in the absence of relative performance concerns. There, a concise characterisation of the forward criteria is given along necessary and sufficient conditions for their existence and uniqueness. In that setting, the zero-volatility forward processes are always time-decreasing processes. We point to the reader that this does noeh have to be case if relative performance concerns are present (see also [36, 52]).

Chapter 6

Investment games under constant absolute risk aversion preferences

6.1 Best responses

Before proving the main result of the subsection, we make a standing assumption regarding the regularity of the forward utility maps

Assumption 6.1.1. *Assume that the derivatives $U_t^i(x,t)$, $U_x^i(x,t)$ and $U_{xx}^i(x,t)$ exist for $t \geqslant 0,\ x \in \mathbb{R}$, $\mathbb{P}$-a.s.*

From Assumption 6.1.1, the Itô decomposition of the forward utility map is

$$\mathrm{d}U^i(x,t) = U_t^i(x,t)\mathrm{d}t, \quad \text{for } i \in \{1,\dots,n\}. \tag{6.1.1}$$

We next derive a PDE with random coefficients and an optimal investment strategy for a smooth relative performance criteria of zero volatility of some agent i assuming that all other agents $j \neq i$ have made their investment decisions.

Proposition 6.1.2 (Best responses)**.** *Fix $i \in \{1,\dots,n\}$ and the agent's initial preference u_0^i. Assume that each manager $j \neq i$ follows $\pi^j \in \mathcal{A}^j$. Consider the PDE with stochastic coefficients for $(x,t) \in \mathbb{R} \times [0,\infty)$*

$$\begin{aligned} U_t^i = \Big(\theta_i \overline{(\pi\mu)}_t^{(-i)} - \frac{\mu_i\theta_i\sigma_i\overline{(\pi\sigma)}_t^{(-i)}}{\nu_i^2+\sigma_i^2}\Big)U_x^i + \frac{\mu_i^2}{2(\nu_i^2+\sigma_i^2)}\frac{(U_x^i)^2}{U_{xx}^i} \\ + \frac{1}{2}U_{xx}^i\Big[\Big(\theta_i\overline{(\pi\sigma)}_t^{(-i)}\Big)^2\Big(\frac{\sigma_i^2}{\nu_i^2+\sigma_i^2}-1\Big) - \frac{\theta_i^2}{n-1}\overline{(\pi\nu)^2}^{(-i)}\Big], \end{aligned} \tag{6.1.2}$$

and assume that for an admissible initial condition $U(\cdot,0) = u_0^i(\cdot)$, the PDE has a smooth solution U^i satisfying Assumption 6.1.1, such that $x \mapsto U^i(x,t)$ is strictly increasing ($U_x > 0$) and strictly concave ($U_{xx} < 0$) for each $t > 0,\ \mathbb{P}$-a.s.

Define the strategy $\pi^{i,*}$

$$\pi_t^{i,*} = \frac{1}{\nu_i^2+\sigma_i^2}\Big(\theta_i\sigma_i\overline{(\pi\sigma)}_t^{(-i)} - \mu_i\frac{U_x^i(\widehat{X}_t^{i,*},t)}{U_{xx}^i(\widehat{X}_t^{i,*},t)}\Big), \quad t > 0,$$

where $\widehat{X}^{i,*}$ solves (5.2.5) *with $\pi^{i,*}$ being used.*

If $\pi^{i,*} \in \mathcal{A}^i$ and $\widehat{X}^{i,*}$ are well-defined, then $U^i(x,t)$ is a forward utility performance process. Moreover, the policy $\pi^{i,*}$ is optimal (in the sense of Definition 5.3.1).

Remark 6.1.3. *Note that the randomness in PDE* (6.1.2) *is coming from $\pi^{\cdot}$ only.*

By direct inspection of the expression for $\pi^{i,*}$ one sees that if the *local risk tolerance function* $r^i(x,t) = r^i = Const$, for all $t > 0$ (e.g. the utility is of *Constant Absolute Risk Aversion (CARA)* type – then the optimal strategy will be constant throughout time if additionally all other agents also choose a constant strategy.

Corollary 6.1.4 (Constant strategies under CARA). *Assume that all agents $j \neq i$ invest according to constant strategies $\pi^j \in \mathbb{R}$ and that the* local risk tolerance function *r^i is constant. Then $\pi^{i,*}$ is constant.*

We now prove the previous "best responses" proposition above.

of Proposition 6.1.2. From (5.2.4) we have the dynamics of $d\widehat{X}^i$ (and hence that of $d(X^i - \theta_i \overline{X}^{(-i)})$). We now apply the Itô formula to $U^i(\widehat{X}^i_t, t) = U^i(X^i_t - \theta_i \overline{X}^{(-i)}_t, t)$,

$$\begin{aligned}
\mathrm{d}U^i(\widehat{X}^i_t, t) &= U^i_t(\widehat{X}^i_t, t)\mathrm{d}t + U^i_x(\widehat{X}^i_t, t)\mathrm{d}\widehat{X}^i_t + \frac{1}{2}U^i_{xx}(\widehat{X}^i_t, t)\mathrm{d}\langle \widehat{X}^i_t\rangle \\
&= U^i_t(\widehat{X}^i_t, t)\mathrm{d}t + U^i_x(\widehat{X}^i_t, t)\big(\pi^i_t\mu_i - \theta_i\overline{(\pi\mu)}^{(-i)}_t\big)\mathrm{d}t \\
&\quad + U^i_x(\widehat{X}^i_t, t)\Big(\pi^i_t\nu_i\mathrm{d}W^i_t - \theta_i\big(\frac{1}{n-1}\sum_{k\neq i}^{n}\pi^k_t\nu_k\mathrm{d}W^k_t\big)\Big) \\
&\quad + U^i_x(\widehat{X}^i_t, t)\big(\pi^i_t\sigma_i - \theta_i\overline{(\pi\sigma)}^{(-i)}_t\big)\mathrm{d}B_t \\
&\quad + \frac{1}{2}U^i_{xx}(\widehat{X}^i_t, t)\Big[(\pi^i_t\nu_i)^2 + \frac{\theta_i^2}{n-1}\overline{(\pi\nu)^2}^{(-i)} + \big(\pi^i_t\sigma_i - \theta_i\overline{(\pi\sigma)}^{(-i)}_t\big)^2\Big]\mathrm{d}t,
\end{aligned} \tag{6.1.3}$$

with $U^i(\widehat{X}^i_0, 0) = U^i(x^i_0 - \theta_i\overline{x}^{(-i)}_0, 0)$ and we used that the B, W^j are all i.i.d.

By Definition 5.3.1, the process $U^i(\widehat{X}^i_t, t)$ becomes a Martingale at the optimum π. Direct computations using first order conditions (∂_{π^i}"drift" $= 0$) yield

$$\begin{aligned}
& 0 + U^i_x\big(\mu_i - 0\big) + \frac{1}{2}U^i_{xx}\Big[2\pi^i\nu_i^2 + 0 + 2\big(\pi^i_t\sigma_i - \theta_i\overline{(\pi\sigma)}^{(-i)}_t\big)\sigma_i\Big] = 0 \\
\Leftrightarrow \qquad & U^i_{xx}\pi^i(\nu_i^2 + \sigma_i^2) = -U^i_x\mu_i + U^i_{xx}\theta_i\sigma_i\overline{(\pi\sigma)}^{(-i)}_t \\
\Rightarrow \qquad & \pi^i_t = \frac{1}{\nu_i^2 + \sigma_i^2}\Big(\theta_i\sigma_i\overline{(\pi\sigma)}^{(-i)}_t - \mu_i\frac{U^i_x(\widehat{X}^i_t, t)}{U^i_{xx}(\widehat{X}^i_t, t)}\Big).
\end{aligned} \tag{6.1.4}$$

Injecting the expression of π^i_t in the drift term of (6.1.3) and simplifying we arrive at the consistency condition (6.1.2), we do not carry out this step explicitly, nonetheless, using that U^i solves (6.1.2) equation (6.1.3) simplifies to (exact calculations are carried out in the Section A.1),

$$\begin{aligned}
&\mathrm{d}U^i(\widehat{X}^i_t, t) \\
&= U^i_x(\widehat{X}^i_t, t)\Big(\pi^i_t\nu_i\mathrm{d}W^i_t - \theta_i\big(\frac{1}{n-1}\sum_{k\neq i}^{n}\pi^k_t\nu_k\mathrm{d}W^k_t\big)\Big) \\
&\quad + U^i_x(\widehat{X}^i_t, t)\Big(\pi^i_t\sigma_i - \theta_i\overline{(\pi\sigma)}^{(-i)}_t\Big)\mathrm{d}B_t \\
&\quad + \frac{1}{2}U^i_{xx}(\widehat{X}^i_t, t)\frac{1}{\nu_i^2 + \sigma_i^2}\Big|\pi^i(\nu_i^2 + \sigma_i^2) - \Big(\theta_i\sigma_i\overline{(\pi\sigma)}^{(-i)}_t - \mu_i\frac{U^i_x(\widehat{X}^i_t, t)}{U^i_{xx}(\widehat{X}^i_t, t)}\Big)\Big|^2\mathrm{d}t.
\end{aligned} \tag{6.1.5}$$

The concavity assumption of $U^i(x, t)$ implies that the drift term above is non-positive and vanishes when (6.1.4) holds. We can conclude that, if $\pi^{i,*}_t = \pi^i_t \in \mathcal{A}^i$ and the associated process $\widehat{X}^{i,*}$ is well-defined (solution to (5.2.5) with $\pi^{i,*}$), the process $U^i(\widehat{X}^{i,*}_t, t)$ is a local-martingale, otherwise it is a local supermartingale. □

Examples: CARA case

Example 6.1.5 (The classic CARA case - exponential case)**.** *The exponential criterion takes as initial condition the map $U(x,0)$ ($x \in \mathbb{R}$) defined as*

$$U^i(x,0) = -e^{-x/\delta}, \quad \textit{with} \quad \delta > 0. \tag{6.1.6}$$

In this case, the local risk tolerance function $r = -U^i_x/U^i_{xx} = \delta$.

In our case accounting for social interaction between agents in the form of performance concerns, the i-th agent's utility is a function $U^i : \Omega \times \mathbb{R} \times \mathbb{R} \times [0,\infty) \to \mathbb{R}$ of both their individual wealth x and the average wealth wealth of all agents, m. The initial/starting utility map is of the form

$$U^i(x,m,0) = -\exp\Big\{ -\frac{1}{\delta_i}(x-\theta_i m)\Big\},$$

where we refer to the constants $\delta_i > 0$ and $\theta_i \in [0,1]$ as *personal risk tolerance* and *competition weight* parameters, respectively.

Example 6.1.6 (The time-monotone forward utility with starting exponential)**.** *For $i \in \{1,\dots,n\}$, let the dynamics of U^i be given by* (6.1.1) *and assume $U^i(x,0) = -e^{-x/\delta_i}$ with $\delta_i > 0$. Then the solution to the PDE* (6.1.2) *is given by*

$$U^i(x,t) = -e^{-\frac{x}{\delta_i}+f_i(t)}, \qquad \textit{with} \quad \delta_i > 0, \tag{6.1.7}$$

where $(f_i(t))_{t\geqslant 0}$ is the random map given below independent of x satisfying $f_i(0)=0$, sufficiently integrable and $t \mapsto f_i(t)$ is differentiable. Note that in this case, the local risk tolerance function *satisfies $r^i = -U^i_x/U^i_{xx} = \delta_i$.*

Injecting $U^i(x,t)$ above in (6.1.2) *yields an ODE for f_i (we omit the time variable),*

$$\begin{aligned} f_i' &= -\frac{\theta_i}{\delta_i}\Big(\overline{(\pi\mu)}^{(-i)} - \frac{\mu_i\sigma_i\overline{(\pi\sigma)}^{(-i)}}{\nu_i^2+\sigma_i^2}\Big) + \frac{\mu_i^2}{2(\nu_i^2+\sigma_i^2)} \\ &\qquad + \frac{\theta_i^2}{2\delta_i^2}\Big[\Big(\overline{(\pi\sigma)}^{(-i)}\Big)^2\Big(\frac{\sigma_i^2}{\nu_i^2+\sigma_i^2}-1\Big) - \frac{1}{n-1}\overline{(\pi\nu)^2}^{(-i)}\Big] \\ &= -\frac{\theta_i}{\delta_i}\overline{(\pi\mu)}^{(-i)} + \frac{1}{2(\nu_i^2+\sigma_i^2)}\Big(\mu_i + \frac{\theta_i}{\delta_i}\sigma_i\overline{(\pi\sigma)}^{(-i)}\Big)^2 \\ &\qquad - \frac{\theta_i^2}{2\delta_i^2}\Big[\Big(\overline{(\pi\sigma)}_t^{(-i)}\Big)^2 + \frac{1}{n-1}\overline{(\pi\nu)^2}^{(-i)}\Big] =: \eta_i. \end{aligned}$$

Hence, $f_i(t) = \int_0^t \eta_i(s)ds$. In particular, if all coefficients and strategies are constant, then (with a slight abuse of notation) $f_i(t) = t\eta_i$ for a constant η_i given by the RHS of the above ODE.

Example 6.1.7 (No performance concerns: $\theta^i = 0$)**.** *We continue to work under the time-monotone forward utility case of the previous example. Without performance concerns, i.e. $\theta_i = 0$, then $\eta_i = \frac{\mu_i^2}{2(\nu_i^2+\sigma_i^2)}$ and we recover well-known results. We have from Proposition 6.1.2 that*

$$\pi^{i,*}_{\cdot} = \frac{\mu_i\delta_i}{\nu_i^2+\sigma_i^2} \quad \textit{and} \quad U^i(x,t) = -\exp\Big\{-\frac{x}{\delta_i} + t\eta_i^{(\theta_i=0)}\Big\},$$

with the constant $\eta_i^{(\theta_i=0)} = \frac{\mu_i^2}{2(\nu_i^2+\sigma_i^2)}$.

6.1.1 The Forward Nash equilibrium

In view of the *best responses* discussed in Proposition 6.1.2 we now investigate the *simultaneous best responses* as to establish the existence of a Nash equilibrium.

Definition 6.1.8 (Forward Nash equilibrium)**.** *A forward Nash equilibrium consists of n-pairs of $\mathbb{F}$-adapted maps $(U^i, \pi^{i,*})$ such that for any $t \geqslant 0$ the following conditions hold.*

i) For any $i \in \{1,\dots,n\}$, $\pi^{i,*} \in \mathcal{A}^i$;

ii) For each player $i \in \{1,\dots,n\}$ the following holds: given the strategies $\pi^{j,*} \in \mathcal{A}^j$ (any $j \neq i$) the processes $U^i(\widehat{X}^i_t(\pi^{*,-i}),t)$ is a (local) supermartingale where $\widehat{X}^i(\pi^{*,-i})$ solves (5.2.5) *with all managers $j \neq i$ acting according to $\pi^{j,*}$;*

iii) For each player $i \in \{1,\dots,n\}$ the following holds: the process $U^i(\widehat{X}^{i,*}_t(\pi^{*,-i}),t)$ is a (local) martingale where $\widehat{X}^i(\pi^{*,-i})$ solves (5.2.5) *with* all *managers j acting according to $\pi^{j,*}$.*

If all the optimal strategies are constant we say we have a constant forward Nash equilibrium*.*

Remark 6.1.9. *The equilibrium strategies in this manuscript should be formally understood as open-loop equilibria, and our choice is motivated by better analytical tractability. At the same time, the open-loop and close-loop controls coincide in case of deterministic equilibrium, and our choice of utility framework (*(6.1.1)*,6.1.6,* (7.1.1) *and 7.2.5) guarantees that. We highlight that comparative study of open-loop and closed-loop controls under FPPs remains an open question and is outside the scope of this work.*

Under appropriate integrability conditions plus the martingale/supermartingale characterisations, we have for some agent i for any $\pi^i \in \mathcal{A}^i$

$$\begin{aligned}\mathbb{E}[U^i(\widehat{X}^{i,*}_t(\pi^{*,-i}),t)] &= \mathbb{E}[U^i(\widehat{X}^{i,*}_0(\pi^{*,-i}),0)] = \mathbb{E}[U^i(x^i_0 - \theta_i \overline{x}^{(-i)}_0,0)]\\ &= U^i(x^i_0 - \theta_i \overline{x}^{(-i)}_0,0) \geqslant \mathbb{E}[U^i(\widehat{X}^i_t(\pi^{*,-i}),t)].\end{aligned}$$

As expected, no manager can increase the expected utility of their relative performance metric by unilateral decision.

The solvability of the general forward Nash equilibrium seems very difficult for a general forward criteria as one needs to solve the following system for the $\pi^{i,*}$ (see Proposition 6.1.2, in particular (6.1.4)) and the corresponding PDEs for the U^i, $i \in \{1,\dots,n\}$:

$$\pi^{i,*}_t(\nu_i^2+\sigma_i^2) = \theta_i\sigma_i\Big(\frac{1}{n-1}\sum_{k=1,k\neq i}^{n} \pi^{k,*}_t \sigma_k\Big) - \mu_i \frac{U^i_x\big(\widehat{X}^{i,*}_t(\pi^{*,-i}),t\big)}{U^i_{xx}\big(\widehat{X}^{i,*}_t(\pi^{*,-i}),t\big)}. \tag{6.1.8}$$

Equilibrium with time-monotone forward utilities and exponential initial condition

In order to obtain explicit results we focus on the time-monotone case presented in Example 6.1.6 for which $U^i_x/U^i_{xx} = -\delta_i$. More notably, at the level at which we have formulated our problem we can easily recover the results of [73, Theorem 2.3] for which one has $U^i_x/U^i_{xx} = -\delta_i$, for any t (note their Remark 2.5).

Theorem 6.1.10. *Assume the conditions of Proposition 6.1.2 hold for all agents $i \in \{1,\dots,n\}$. Assume furthermore that agents have time-monotone forward utility U^i with initial condition* (6.1.6).

Define the quantities φ^σ_n and ψ^σ_n by

$$\varphi^\sigma_n := \frac{1}{n}\sum_{i=1}^{n} \delta_i \frac{\mu_i\sigma_i}{\nu_i^2+\sigma_i^2\big(1+\frac{\theta_i}{n-1}\big)} \quad \textit{and} \quad \psi^\sigma_n := \frac{1}{n-1}\sum_{i=1}^{n} \theta_i \frac{\sigma_i^2}{\nu_i^2+\sigma_i^2\big(1+\frac{\theta_i}{n-1}\big)}. \tag{6.1.9}$$

If $\psi^\sigma_n \neq 1$, then a constant forward Nash equilibrium *exists and is unique, with the constant optimal strategies $\pi^{i,*}$ given by*

$$\pi^{i,*}_\cdot = \frac{1}{\nu_i^2+\sigma_i^2\big(1+\frac{\theta_i}{n-1}\big)}\Big(\theta_i\sigma_i\Big(1+\frac{1}{n-1}\Big)\frac{\varphi^\sigma_n}{1-\psi^\sigma_n} + \mu_i\delta_i\Big). \tag{6.1.10}$$

The forward Nash equilibria is given by the n-pairs $\{(U^{i,*},\pi^{i,*})\}_{i=1,\dots,n}$ where the $U^{i,*}$ is the solution of (6.1.2) *(see Example 6.1.6) under the optimal constant strategies $\pi^{\cdot,*}$.*

The term η_i (see Example 6.1.6), at equilibrium, is given by

$$\eta_i = -\frac{\theta_i}{\delta_i}\Big(\Big\{\frac{n}{n-1}\overline{\pi\mu} - \frac{1}{n-1}\pi^i\mu_i\Big\} - \frac{\mu_i\sigma_i}{\nu_i^2+\sigma_i^2}\Big\{\frac{n}{n-1}\overline{\pi\sigma} - \frac{1}{n-1}\pi^i\sigma_i\Big\}\Big)$$
$$+ \frac{\mu_i^2}{2(\nu_i^2+\sigma_i^2)} + \frac{\theta_i^2}{2\delta_i^2}\Big[\Big\{\frac{n}{n-1}\overline{\pi\sigma} - \frac{1}{n-1}\pi^i\sigma_i\Big\}^2\Big(\frac{\sigma_i^2}{\nu_i^2+\sigma_i^2} - 1\Big)$$
$$- \Big\{\frac{n}{(n-1)^2}\overline{(\pi\nu)^2} - \frac{1}{(n-1)^2}(\pi^i\nu_i)^2\Big\}\Big], \tag{6.1.11}$$

where the relevant expressions for $\overline{\pi\sigma}$, $\overline{\pi\mu}$ and $\overline{(\pi\nu)^2}$ are given below in (6.1.12)*,* (6.1.13) *and* (6.1.14).

Remark 6.1.11. *We note that we do not solve the same problem studied at* [73] *but an equivalent one. However, imposing the scaling factor given by* [73, Remark 2.5] *we recover the same results as in* [73, Theorem 2.3].

Proof. Injecting the condition $U_x/U_{xx} = -\delta_i$ in (6.1.8), the system to be solved in order to ascertain the Nash equilibrium is, across $i \in \{1,\dots,n\}$,

$$\begin{aligned}\pi_t^{i,*}(\nu_i^2+\sigma_i^2) &= \theta_i\sigma_i\Big(\frac{1}{n-1}\sum_{k=1,k\neq i}^{n}\pi_t^{k,*}\sigma_k\Big) + \mu_i\delta_i\\ &= \theta_i\sigma_i\Big(\frac{n}{n-1}\overline{(\pi\sigma)}_t - \frac{1}{n-1}\pi^{i,*}\sigma_i\Big) + \mu_i\delta_i\\ \Leftrightarrow \quad \pi_t^{i,*} &= \frac{1}{\nu_i^2+\sigma_i^2\big(1+\frac{\theta_i}{n-1}\big)}\Big(\theta_i\sigma_i\frac{n}{n-1}\overline{(\pi\sigma)}_t + \mu_i\delta_i\Big).\end{aligned}$$

The final line yields the expression for $\pi^{i,*}$ as a function of the unknown $\overline{\pi\sigma}$. To determine the latter, multiply both sides by σ_i and average over $i \in \{1,\dots,n\}$, this yields a solvability condition

$$\overline{(\pi\sigma)}_t = \overline{(\pi\sigma)}_t\psi_n^\sigma + \varphi_n^\sigma \quad\Leftrightarrow\quad \overline{\pi\sigma} = \frac{\varphi_n^\sigma}{1-\psi_n^\sigma} \qquad \text{as long as} \qquad \psi_n^\sigma \neq 1. \tag{6.1.12}$$

Plugging the expression $\overline{(\pi\sigma)}$ in that for $\pi^{i,*}$ yields the result. That the optimal strategies are constant is now obvious.

It remains to derive the expression for the η_i's. Just like for $\overline{\pi\sigma}$, we obtain an expression for $\overline{\pi\mu}$ by multiplying $\pi^{i,*}$ by μ_i and averaging on both sides, we have

$$\overline{\pi\mu} = \frac{n}{n-1}\cdot\frac{\varphi_n^\sigma}{1-\psi_n^\sigma}\cdot\psi_n^\mu + \phi_n^\mu \quad \text{and} \quad \overline{\pi\mu}^{(-i)} = \frac{n}{n-1}\overline{\pi\mu} - \frac{1}{n-1}\pi^i\mu_i, \tag{6.1.13}$$

where we used (5.2.6) and the quantities $\varphi_n^\mu, \psi_n^\mu$ are defined as

$$\varphi_n^\mu := \frac{1}{n}\sum_{k=1}^{n}\delta_k\frac{\mu_k^2}{\nu_k^2+\sigma_k^2(1+\frac{\theta_k}{n-1})} \qquad \text{and} \qquad \psi_n^\mu := \frac{1}{n}\sum_{k=1}^{n}\theta_k\frac{\mu_k\sigma_k}{\nu_k^2+\sigma_k^2(1+\frac{\theta_k}{n-1})}.$$

Similarly, defining $\overline{(\pi\nu)^2} := \frac{1}{n-1}\sum_{k\neq i}(\pi_t^k\nu_k)^2$ we have

$$\overline{(\pi\nu)^2} = \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{\nu_i\theta_i\sigma_i\cdot\frac{n}{n-1}\cdot\frac{\varphi_n^\sigma}{1-\psi_n^\sigma} + \nu_i\mu_i\delta_i}{\nu_i^2+\sigma_i^2\big(1+\frac{\theta_i}{n-1}\big)}\Big)^2. \tag{6.1.14}$$

Similarly to (5.2.6), we have $\overline{(\pi\nu)^2}^{(-i)} = \frac{n}{n-1}\overline{(\pi\nu)^2} - \frac{1}{n-1}(\pi^i\nu_i)^2$. Replacing these expressions in that for η_i in Example 6.1.6 the expression in the result's statement follows. □

From the forward utility machinery one can easily recover the classical case of utility optimisation where one prescribes the utility map for the horizon time T then proceeds to optimise.

Example 6.1.12 (Recovering the classical utility problem from the forward one.)**.** *If one would start the forward utility with (for some $0 < T < \infty$)*

$$u_0^i(x) := -e^{-x/\delta_i - T\eta_i},$$

then computations like those presented yield the forward utility map $U(x,t)$ as

$$U^i(x,t) = -e^{-x/\delta_i + (t-T)\eta_i}, \quad t \in [0,T]$$

and in particular $U(x,T) = -e^{-x/\delta_i}$. In other words, our forward utility recovers as a particular case the classical exponential utility maximisation problem (discussed in [73]*).*

Corollary 6.1.13 (Single stock)**.** *Let $\mu_i = \mu > 0,\ \sigma_i = \sigma > 0$ and $\nu_i = 0$, for any $i = 1, \dots, n$. Let the constants*

$$\varphi_n^\sigma := \frac{1}{n}\sum_{i=1}^n \frac{\delta_i}{1+\frac{\theta_i}{n-1}} \quad \textit{and} \quad \psi_n^\sigma := \frac{1}{n-1}\sum_{i=1}^n \frac{\theta_i}{1+\frac{\theta_i}{n-1}}. \tag{6.1.15}$$

If $\psi_n^\sigma \neq 1$, then a constant forward Nash equilibrium *exists, with the constant optimal strategies $\pi^{i,*}$ given by*

$$\pi^{i,*}_\cdot = \frac{\mu}{\sigma^2\big(1+\frac{\theta_i}{n-1}\big)}\Big(\theta_i\Big(1+\frac{1}{n-1}\Big)\frac{\varphi_n^\sigma}{1-\psi_n^\sigma} + \delta_i\Big).$$

6.2 Forward mean-field game

By inspection of Theorem 6.1.10 one sees that the optimal strategy and forward utility map for some agent depend on that agent's specific parameters (model parameters, initial wealth, risk tolerance and performance concern) and on certain averages of the parameters of all agents. This makes a case for a MFG approach to the game.

In this section and inspired by the results in the previous one, we formalise the concept of forward mean-field Nash game. We use the concept of *type distributions* introduced in [60] and [72,73]. We follow the construction presented in the latter.

We focus on initial forward utilities at time $t = 0$ that are of exponential type,

$$U^i(x,m,0) = -\exp\Big\{-\frac{1}{\delta_i}(x-\theta_i m)\Big\},$$

where we refer to the constants $\delta_i > 0$ and $\theta_i \in [0,1]$ as *personal risk tolerance* and *competition weight* parameters, respectively.

For the n-agent game, we define for each agent $i = 1, \dots, n$ the *type vector*

$$\zeta_i := (x_0^i, \delta_i, \theta_i, \mu_i, \nu_i, \sigma_i),$$

which characterises perfectly each agent i. These *type vectors* induce an empirical measure, called the *type distribution*, which is the probability measure on the *type space*

$$\mathcal{Z}^e := \mathbb{R} \times (0,\infty) \times [0,1] \times (0,\infty) \times [0,\infty) \times [0,\infty), \tag{6.2.1}$$

given by

$$m_n(A) = \frac{1}{n}\sum_{i=1}^n 1_A(\zeta_i), \quad \text{for Borel sets } A \subset \mathcal{Z}^e.$$

Assume now that as the number of agents becomes large, $n \to \infty$, the above empirical measure m_n has a weak limit m, in the sense that $\int_{\mathcal{Z}^e} f\, dm_n \to \int_{\mathcal{Z}^e} f\, dm$ for every bounded continuous function f on $\mathcal{Z}^e$. For example, this holds almost surely if the ζ_i's are i.i.d. samples from m. Let $\zeta = (\xi, \delta, \theta, \mu, \nu, \sigma)$ denote an $\mathcal{Z}^e$-valued random variable with this limiting distribution m.

The *mean field game* (MFG) defined next allows us to derive the limiting strategy as the outcome of a self-contained equilibrium problem, which intuitively represents a game with a continuum of agents with type distribution m. Rather than directly modelling a continuum of agents, we follow the MFG paradigm of modelling a single *generic agent*, who we view as randomly selected from the population. The probability measure m represents the distribution of type parameters among the continuum of agents; equivalently, the generic agent's type vector is a random variable with law m. Heuristically, each agent in the continuum trades in a single stock driven by two Brownian motions, one of which is unique to this agent and one of which is common to all agents. We extend the Forward Nash equilibrium of Definition 6.1.8 to the MFG setting below.

6.2.1 Agents through type-distribution and the market

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F})_{t\geqslant 0}, \mathbb{P})$ be a stochastic basis supporting two independent Brownian motions $W = (W_t)_{t\geqslant 0}$ and $B = (B_t)_{t\geqslant 0}$ together with a random vector ζ having distribution m and given by

$$\zeta = (\xi, \delta, \theta, \mu, \nu, \sigma),$$

with values in the space $\mathcal{Z}^e$ defined in (6.2.1) and independent of W and B. Let $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ denote the smallest filtration satisfying the usual assumptions for which ζ is $\mathcal{F}_0$-measurable and both W and B are adapted. Let also $\mathbb{F}^B = (\mathcal{F}_t^B)_{t\in[0,T]}$ denote the natural filtration generated by the Brownian motion B.

The *generic agent's* wealth process solves

$$dX_t = \pi_t(\mu \mathrm{d}t + \nu \mathrm{d}W_t + \sigma \mathrm{d}B_t), \quad X_0 = \xi, \tag{6.2.2}$$

where the portfolio strategy must belong to the admissible set $\mathcal{A}_{\mathrm{MF}}$ of self-financing $\mathbb{F}$-progressively measurable real-valued processes $(\pi_t)_{t\geqslant 0}$ satisfying the square-integrability condition $\mathbb{E}[\int_0^T |\pi_t|^2 \mathrm{d}t] < \infty$ for any $T \in [0, \infty)$. The generic agent's initial wealth is given by ξ, whereas (μ, ν, σ) are the market parameters. In the sequel, the parameters δ and θ will affect the risk preferences of the generic agent. Each agent among the continuum will have different preference parameters and hence these six parameters are $\mathcal{F}_0$-random, and each has the exact same interpretation an in the n-player game of the earlier section.

6.2.2 The mean-field equilibrium

The formulation of the forward Nash game of Section 5.3 drives the formulation of the Mean-field game we discuss here. Recall that in the MFG-formulation the *generic agent* has no influence on the average wealth of the continuum of agents, as but one agent amid a continuum of agents. We next introduce the concept of the main object of interest the *MF-Forward relative performance equilibrium*.

We introduce the regularity requirements for the utility.

Assumption 6.2.1. *Assume that the derivatives $U_t(x,t)$, $U_x(x,t)$ and $U_{xx}(x,t)$ exist for $t \geqslant 0,\ x \in \mathbb{R},\ \mathbb{P}$-a.s.*

As in Section 5.3.1, Assumption 6.2.1 implies the Itô decomposition of map U

$$\mathrm{d}U(x,t) = U_t(x,t)\mathrm{d}t.$$

Given this market setup we next define our concept of equilibrium.

Definition 6.2.2 (MF-Forward CARA relative performance equilibrium (for the generic manager))**.** *Let $(\overline{X}_t)_{t\geqslant 0}$ be the $\mathbb{F}^B$-adapted square integrable stochastic process representing the average wealth of the continuum of agents. Let $\pi \in \mathcal{A}^{\mathrm{MF}}$ and X^π solve* (6.2.2) *with π.*

The $\mathbb{F}^{\mathrm{MF}}$-progressively measurable random field $(U(x,t))_{t\geqslant 0}$ is an MF-forward relative performance *for the* generic manager *if, for all $t \geqslant 0$, the following conditions hold:*

i) The mapping $x \mapsto U(x,t)$, is $\mathbb{P}$-a.s. strictly increasing and strictly concave;

ii) For any $\pi \in \mathcal{A}^{MF}$, $U(X_t^\pi - \theta\overline{X}_t, t)$ is a (local) supermartingale and X^π is the generic agent's *wealth process solving* (6.2.2) *for the strategy π;*

iii) There exists $\pi^* \in \mathcal{A}^{MF}$ such that $U(X^*_t - \theta\overline{X}_t, t)$ is a (local) martingale where X^* solves (6.2.2) *with π^* plugged in as the strategy;*

iv) We call π^* of point iii) a MF-equilibrium *if $\overline{X}_t = \mathbb{E}[X^*_t|\mathcal{F}^B_t]$ for all $t \geqslant 0$ where where X^* solves* (6.2.2) *with π^* plugged in as the strategy.*

We denote the triplet $(U, \pi^*, \overline{X},)$ satisfying i)-iv) the MF-Forward relative performance equilibrium. *An MF-equilibrium is* constant *if there exists an $\mathcal{F}^{MF}_0$-measurable RV π^* such that $\pi_t = \pi^*$, for all $t \geqslant 0$.*

The last point can be understood as a fixed point argument which creates a compatibility condition between the generic agent within the continuum of agents. In fact, conditionally on the BM B each agent faces an independent noise W and an independent type vector ζ. As in Mean-field games [72, 73], conditionally on B, all agents faces i.i.d. copies of the same optimisation problem. The law of large numbers suggests that the average terminal wealth of the whole population should be $\mathbb{E}[X^*_t|\mathcal{F}^B_t]$.

Our construction allows us to identify $\mathbb{E}[X^*_t|\mathcal{F}^B_t]$ with a certain dynamics and, in turn, treat this component as an additional uncontrolled state process. This avoids altogether the conceptualisation of the master equation for models with different types of agents. The latter is left for future research.

6.2.3 Solving the optimisation problem

We now present the main result of this section which is the existence of a *MF-Forward CARA relative performance equilibrium* for the generic manager according to Definition 6.2.2 within the context of time-monotone forward utilities.

From the methodological point of view, the problem is solved as before. Apply Itô formula to $U(Z^\pi_t, t)$, determine the optimal strategy π^* and the consistency condition (the PDE) for U such that the first three conditions of Definition 6.2.2 hold. The last condition, to show that π^* is indeed the MFG Forward equilibrium follows by construction as we will see.

Theorem 6.2.3. *Take a generic agent $\zeta = (\xi, \delta, \theta, \mu, \nu, \sigma)$ and assume that $\delta > 0$, $\theta \in [0, 1]$, $\mu > 0$, $\sigma \geqslant 0$, $\nu \geqslant 0$ such that $\sigma^2 + \nu^2 > 0$.*

Assume the following constants are finite

$$\psi^\sigma := \mathbb{E}\Big[\theta \frac{\sigma^2}{\nu^2 + \sigma^2}\Big], \quad \varphi^\sigma := \mathbb{E}\Big[\delta \frac{\mu\sigma}{\nu^2 + \sigma^2}\Big],$$
$$\psi^\mu := \mathbb{E}\Big[\theta \frac{\mu\sigma}{\nu^2 + \sigma^2}\Big], \quad \text{and} \quad \varphi^\mu := \mathbb{E}\Big[\delta \frac{\mu^2}{\nu^2 + \sigma^2}\Big].$$

Assume that $\psi^\sigma \neq 1$. Then there exists a unique constant MF-Forward CARA relative performance equilibrium in the sense of Definition 6.2.2.

The constant MF-equilibrium strategy is unique and is given by

$$\pi^* = \frac{1}{\nu^2 + \sigma^2}\Big(\theta\sigma \frac{\varphi^\sigma}{1 - \psi^\sigma} + \mu\delta\Big), \tag{6.2.3}$$

constrained to the identity

$$\mathbb{E}[\sigma\pi^*] = \frac{\varphi^\sigma}{1 - \psi^\sigma} < \infty.$$

The MF-forward CARA relative performance utility map under Assumption 6.2.1 is the unique solution of the PDE with stochastic coefficients

$$U_t = \theta\Big(\frac{\varphi^\sigma}{1 - \psi^\sigma} \cdot \psi^\mu + \varphi^\mu - \mu\frac{\sigma}{\nu^2 + \sigma^2} \cdot \frac{\varphi^\sigma}{1 - \psi^\sigma}\Big)U_x$$
$$+ \frac{\mu^2}{2(\nu^2 + \sigma^2)} \frac{(U_x)^2}{U_{xx}} + \frac{1}{2}U_{xx} \cdot \theta^2\Big(\frac{\varphi^\sigma}{1 - \psi^\sigma}\Big)^2\Big(\frac{\sigma^2}{\nu^2 + \sigma^2} - 1\Big). \tag{6.2.4}$$

When the initial condition is $U(x,0) = u_0(x) = -e^{-x/\delta}$, i.e. the exponential preferences, U is given explicitly by $U(x,t) = u_0(x)e^{t\eta}$ with η given by

$$\eta = -\frac{\theta}{\delta}\overline{\mu\pi} + \frac{1}{2(\nu^2+\sigma^2)}\Big(\mu + \frac{\theta}{\delta}\sigma\overline{\sigma\pi}\Big)^2 - \frac{\theta^2}{2\delta^2}\Big(\overline{\sigma\pi}\Big)^2, \tag{6.2.5}$$

where $\overline{\sigma\pi}$ and $\overline{\mu\pi}$ are given by (6.2.9) *and* (6.2.10) *respectively. If $\psi^\sigma = 1$, then there exists no constant MF-equilibrium.*

By comparing the statements of Theorem 6.1.10 and Theorem 6.2.3 (and same happens for the respective Single (common) Stock Corollaries) one easily sees that as $n \to \infty$ the strategies, weights ($\phi_n^{\cdot}$ and $\psi_n^{\cdot}$) and forward-utility map in Theorem 6.1.10 converge to the respective quantities appearing in Theorem 6.2.3.

Remark 6.2.4. *We point out that the interaction of the generic agent with the continuum is only performed through the common noise B. That can be seen by the term $\frac{n}{(n-1)^2}\overline{(\pi\nu)^2} - \frac{1}{(n-1)^2}(\pi^i\nu_i)^2$ from η_i's in* (6.1.11) *converging to zero as $n \to \infty$, as we have by* (6.2.12) *(compare with* (6.2.5)*). We can interpret it via the standard mean-field approximation, the individual's impact on the others is negligible for the infinite system.*

Remark 6.2.5. *In contrast with Remark 6.1.11, here we recover the result from* [73, Theorem 2.10] *as the scaling factors converge to 1 (as $n \to \infty$). Hence, due to space constraints we defer the reader to* [73, Section 2.3] *for the discussion of the equilibria.*

Proof. We proceed in several steps in order to construct the constant MF-equilibrium. To that end we must solve ii)-iii) in Definition 6.2.2 for a given $\overline{X}$ process associated to $\pi \in \mathcal{A}_{\text{MF}}$. Condition iv), for MF-equilibrium allows us to focus only on processes of the form $\overline{X}_t = \mathbb{E}[X_t^\pi|\mathcal{F}_t^B]$ where X^π solves (6.2.2) for a constant strategy π (i.e. $\mathcal{F}_0^{\text{MF}}$-measurable) satisfying $\mathbb{E}[\pi^2] < \infty$.

Step 0. The dynamics of the average wealth process. To solve the above problem given $(\overline{X}_t)_{t\geqslant 0}$ it suffices to restrict ourselves to processes $(\overline{X}_t)_{t\geqslant 0}$ satisfying $\overline{X}_t = \mathbb{E}[X_t^\pi|\mathcal{F}_t^B]$, $\mathbb{P}$-a.s.. We then have

$$\begin{aligned}\overline{X}_t = \mathbb{E}[X_t^\pi|\mathcal{F}_t^B] &= \mathbb{E}\Big[\xi + \int_0^t \mu\pi \mathrm{d}s + \int_0^t \nu\pi \mathrm{d}W_s + \int_0^t \sigma\pi \mathrm{d}B_s\Big|\mathcal{F}_t^B\Big] \\ &= \bar{\xi} + \int_0^t \overline{\mu\pi_s}\mathrm{d}s + \int_0^t \overline{\sigma\pi_s}\mathrm{d}B_s,\end{aligned} \tag{6.2.6}$$

where, for consistency of notation with the previous section, we denote

$$\bar{\xi} := \mathbb{E}[\xi], \quad \overline{\mu\pi} := \mathbb{E}[\mu\pi] \quad \text{and} \quad \overline{\sigma\pi} := \mathbb{E}[\sigma\pi].$$

Hence for $\pi \in \mathcal{A}^{\text{MF}}$ and as in the previous section we can define the dynamics of the process $Z^\pi = X^\pi - \theta\overline{X}$

$$\mathrm{d}Z_t^\pi = \big(\mu\pi_t - \theta\overline{\mu\pi}\big)\mathrm{d}t + \nu\pi_t\mathrm{d}W_t + \big(\sigma\pi_t - \theta\overline{\sigma\pi}\big)\mathrm{d}B_t, \quad Z_0^\pi = \xi - \theta\bar{\xi},$$

and solve the MFG Forward utility problem in Definition 6.2.2 with its help.

Hence applying Itô's formula to $U(Z_t^\pi, t)$ yields

$$\begin{aligned}\mathrm{d}U(Z_t^\pi,t) &= U_t(Z_t^\pi,t)\mathrm{d}t + U_x(Z_t^\pi,t)\mathrm{d}Z_t^\pi + \frac{1}{2}U_{xx}(Z_t^\pi,t)\mathrm{d}\langle Z_t^\pi\rangle \\ &= \Big[U_t(Z_t^\pi,t) + U_x(Z_t^\pi,t)\big(\mu\pi_t - \theta\overline{\mu\pi}\big) \\ &\qquad + \frac{1}{2}U_{xx}(Z_t^\pi,t)\Big((\nu\pi_t)^2 + \big(\sigma\pi_t - \theta\overline{\sigma\pi}\big)^2\Big)\Big]\mathrm{d}t, \\ &\qquad + U_x(Z_t^\pi,t)\nu\pi_t\mathrm{d}W_t + U_x(Z_t^\pi,t)\big(\sigma\pi_t - \theta\overline{\sigma\pi}\big)\mathrm{d}B_t,\end{aligned} \tag{6.2.7}$$

with $U(Z_0^\pi,0) = U(\xi - \theta\bar{\xi},0) = -\exp\{-(\xi - \theta\bar{\xi})/\delta\}$ and we used that the B, W are all i.i.d. Exact calculations on deriving (6.2.7) are presented in the Section A.1.

Step 1. Finding the candidate optimal strategy π^*. As before, the process $U(Z_t^\pi, t)$ becomes a Martingale at the optimum π. Direct computations using first order conditions (∂_π"drift" $= 0$) yield

$$
\begin{aligned}
&0 + U_x \cdot (\mu - 0) + \frac{1}{2} U_{xx} \Big[2\pi\nu^2 + 2\big(\sigma\pi_t - \theta\overline{\sigma\pi}\big)\sigma \Big] = 0 \\
\Rightarrow \qquad & \pi_t^* (\nu^2 + \sigma^2) = \theta\sigma\overline{\sigma\pi} - \mu \frac{U_x(Z_t^\pi, t)}{U_{xx}(Z_t^\pi, t)} = \theta\sigma\overline{\sigma\pi} + \mu\delta,
\end{aligned}
\tag{6.2.8}
$$

where we injected the CARA constraint $U_x/U_{xx} = -\delta$, for all t. By inspection it is clear that π^* is a $\mathcal{F}_0^{\mathrm{MF}}$-measurable RV which is independent of time and is well-defined as long as $\overline{\sigma\pi}$ is finite.

Step 2. The optimality of the strategy. The argument is similar to that in [73]. The original constant strategy π if a MF-equilibrium if and only if for all $t \geqslant 0$

$$
\begin{aligned}
&\mathbb{E}[X_t^\pi | \mathcal{F}_t^B] = \mathbb{E}[X_t^{\pi^*} | \mathcal{F}_t^B] \quad a.s. \\
&\qquad \Leftrightarrow \quad \bar{\xi} + \overline{\mu\pi}\, t + \overline{\sigma\pi} B_t = \bar{\xi} + \overline{\mu\pi^*}\, t + \overline{\sigma\pi^*} B_t \quad a.s.
\end{aligned}
$$

Taking expectations on both sides implies that π is a MG-equilibrium if and only if the following two conditions holds

$$
\overline{\mu\pi} = \overline{\mu\pi^*} \quad \text{and} \quad \overline{\sigma\pi} = \overline{\sigma\pi^*}.
$$

Using (6.2.8) with $U_x/U_{xx} = -\delta$ and the expressions for $\varphi^\sigma, \psi^\sigma$ one derives that

$$
\sigma\pi^* = \theta \frac{\sigma^2}{\nu^2 + \sigma^2} \overline{\sigma\pi} + \delta \frac{\mu\sigma}{\nu^2 + \sigma^2} \quad \Rightarrow \quad \overline{\sigma\pi^*} = \overline{\sigma\pi}\psi^\sigma + \varphi^\sigma,
$$

using that $\overline{\sigma\pi} = \overline{\sigma\pi^*}$ yields solvability if $\psi^\sigma = \mathbb{E}\big[\theta \frac{\sigma^2}{\nu^2+\sigma^2}\big] \neq 1$. The same procedure deals with the condition $\overline{\mu\pi} = \overline{\mu\pi^*}$. We then have

$$
\overline{\sigma\pi^*} = \overline{\sigma\pi} = \frac{\varphi^\sigma}{1 - \psi^\sigma} = \text{Const}, \tag{6.2.9}
$$

$$
\overline{\mu\pi^*} = \overline{\mu\pi} = \frac{\varphi^\sigma}{1 - \psi^\sigma} \cdot \psi^\mu + \varphi^\mu = \text{Const}. \tag{6.2.10}
$$

Injecting these identities in the expression for π^* we find (6.2.3).

For the non-solvability statement, if the equation (6.2.10) has $\psi^\sigma = 1$ and $\varphi^\sigma \neq 0$ then the equation has no solution and hence no constant MF-equilibrium exists. The case $\psi^\sigma = 1$ and $\varphi^\sigma = 0$ is impossible. Since $\mu > 0$ and $\delta > 0$ by assumption, it implies that $\sigma = 0$ and hence that $\psi^\sigma = 0$ contradicting the condition $\psi^\sigma = 1$.

Step 3. Finding the consistency PDE and the Utility map. We do not carry out this step explicitly, nonetheless, injecting the expression of π^*, $\overline{\sigma\pi}$ and $\overline{\mu\pi}$ in the drift term of (6.2.7) and simplifying, we find the necessary equation (6.2.4), i.e. the consistency condition the random field U must satisfy to that the required properties in Definition 6.2.2 hold.

Just like in Example 6.1.6, the time-monotone forward utility equation (6.2.4) can be solved and indeed one has a simplified version. We have

$$
U(x, t) = -e^{-x/\delta + t\eta}, \tag{6.2.11}
$$

where the $\mathcal{F}_0^{\mathrm{MF}}$-measurable RV η is given by (using (6.2.9) and (6.2.10))

$$
\begin{aligned}
\eta &= -\frac{\theta}{\delta}\overline{\mu\pi} + \frac{1}{2(\nu^2 + \sigma^2)} \Big(\mu + \frac{\theta}{\delta}\sigma\overline{\sigma\pi}\Big)^2 - \frac{\theta^2}{2\delta^2}\Big(\overline{\sigma\pi}\Big)^2 \\
&= -\frac{\theta}{\delta}\Big(\frac{\varphi^\sigma}{1 - \psi^\sigma} \cdot \psi^\mu + \varphi^\mu - \mu\frac{\sigma}{\nu^2 + \sigma^2} \cdot \frac{\varphi^\sigma}{1 - \psi^\sigma}\Big) \\
&\quad + \frac{\mu^2}{2(\nu^2 + \sigma^2)} + \frac{\theta^2}{2\delta^2}\Big(\frac{\varphi^\sigma}{1 - \psi^\sigma}\Big)^2 \Big(\frac{\sigma^2}{\nu^2 + \sigma^2} - 1\Big).
\end{aligned}
\tag{6.2.12}
$$

Step 4. The MFG forward utility dynamics. Injecting the consistency PDE (6.2.4) in the expression for $\mathrm{d}U(Z_t^\pi, t)$ given in (6.2.7) yields,

$$\mathrm{d}U(Z_t^\pi, t) = \frac{1}{2}\frac{U_{xx}(Z_t^\pi, t)}{(\nu^2+\sigma^2)}\Big|\pi_t(\nu^2+\sigma^2) - \Big(\theta\sigma\cdot\frac{\varphi^\sigma}{1-\psi^\sigma} + \mu\delta\Big)\Big|^2\mathrm{d}t$$
$$+ U_x(Z_t^\pi, t)\nu\pi_t\mathrm{d}W_t + U_x(Z_t^\pi, t)\Big(\sigma\pi_t - \theta\cdot\frac{\varphi^\sigma}{1-\psi^\sigma}\Big)\mathrm{d}B_t.$$

□

We close with a corollary regarding the common stock case.

Corollary 6.2.6 (Single stock)**.** *Let μ, σ, ν be deterministic with $\nu = 0, \mu, \sigma > 0$. Let the constants*

$$\varphi := \mathbb{E}[\delta] \quad \text{and} \quad \psi := \mathbb{E}[\theta]. \tag{6.2.13}$$

Then, if $\psi \neq 1$ then a constant MF-equilibrium *exists, with the constant optimal strategy π^* given by*

$$\pi^*_\cdot = \frac{\mu}{\sigma^2}\Big(\theta\frac{\varphi}{1-\psi} + \delta\Big).$$

6.2.4 Mean-field dynamic model selection with large horizons

Over the time interval $[0,\infty)$ our generic agent selects a sequence of horizon time $(T_j)_{j\in\mathbb{N}_0}$ (such that $T_0 = 0$, $T_{j+1} - T_j > 0$ and $\lim_j T_j = \infty$) on which the agent assesses and updates the market model by adjusting the model's coefficients. Comparing with (6.2.2) the agent models the stock as

$$\frac{dS_t^j}{S_t^j} = \mu_j\mathrm{d}t + \nu_j\mathrm{d}W_t + \sigma_j\mathrm{d}B_t, \quad S_{T_j} = s_j, \;\; t\in[T_j, T_{j+1}], \tag{6.2.14}$$

where the index j represents the model specification at time T_j. The associated wealth process of the *generic agent* is

$$dX_t^j = \pi_t(\mu_j\mathrm{d}t + \nu_j\mathrm{d}W_t + \sigma_j\mathrm{d}B_t), \quad X_{T_j} = \xi_j, \;\; t\in[T_j, T_{j+1}].$$

Following the earlier constructions of this section, assume that at time $T_0 = 0$ the agent starts with initial utility $u_0(x) = -e^{-x/\delta}$. Then using the results of Theorem 6.2.3, the agent's forward utility map is given by

$$U(x,t) = -e^{x/\delta}e^{t\eta_0} = u_0(x)e^{t\eta_0}, \quad t\in[T_0, T_1] = [0, T_1],$$

where η_0 is the version of (6.2.12) for the type of the agent over the time interval $[T_0, T_1]$ and all the coefficients correspond to a type ζ_0, i.e. $\eta(\zeta_0) = \eta_0$, with

$$\eta_0 = \eta(\zeta_0) := -\frac{\theta}{\delta}\overline{\mu\pi} + \frac{1}{2(\nu^2+\sigma^2)}\Big(\mu + \frac{\theta}{\delta}\sigma\overline{\sigma\pi}\Big)^2 - \frac{\theta^2}{2\delta^2}\Big(\overline{\sigma\pi}\Big)^2. \tag{6.2.15}$$

At time T_1, the generic agent assesses the previous model specification and chooses new coefficients (leading to a change in type, say from ζ_0 to ζ_1). The agent then carries out the optimisation program over $t\in[T_1, T_2]$ but starting from initial utility $U(x, T_1)$. Under the assumption of constant coefficients Theorem 6.2.3, yields,

$$U(x,t) = \Big(u_0(x)e^{T_1\eta_0}\Big)e^{(t-T_1)\eta_1}, \quad t\in[T_1, T_2],$$

where $\eta_1 = \eta(\zeta_1)$ (given by (6.2.15)) depends only on information at time T_1. Quick calculations generalise to any time horizon T_j. Assume we work on the time interval $[T_j, T_{j+1}]$. Stemming from previous calculations, it is easy to see that the initial condition for the forward utility problem is

$$U(x,T_j) = u_0(x)\prod_{k=1}^{j} e^{(T_k-T_{k-1})\eta_{k-1}}$$

(with the convention that if $j<1$ then $\prod_{k=1}^{j}\cdots = 0$) and the MFG forward utility is for all $t\in[T_j,T_{j+1}],\ j>1$ and using that $\eta_j = \eta(\zeta_j)$.

$$\begin{aligned}
U(x,t) = U(x,T_j)e^{(t-T_j)\eta_j} &= u_0(x)\prod_{k=1}^{j} e^{(T_k-T_{k-1})\eta_{k-1}}\cdot e^{(t-T_j)\eta_j},\\
&= u_0(x)\exp\Big\{T_1(\eta_0-\eta_1)+T_2(\eta_1-\eta_2)+\cdots+T_j(\eta_{j-1}-\eta_j)\Big\}e^{t\eta_j}.
\end{aligned}$$

There are two points to highlight. Firstly, the agent needs to carry information of what happened in the past in order to have time-consistency at present time. Secondly, this construction also allows the agents to change not just the model specification (μ,ν,σ) but also their type including risk parameter δ and performance-concern level θ. The initial wealth is fixed from the previous time interval.

Chapter 7

Investment-consumption games under constant relative risk aversion preferences

7.1 Best responses

We make a standing assumption regarding the regularity of the FPP map.

Assumption 7.1.1. *Assume the partial derivatives $U^i_t(x,t)$, $U^i_x(x,t)$, $U^i_{xx}(x,t)$ and $V^i_x(x,t)$, $V^i_{xx}(x,t)$ exist for all $t \geqslant 0,\ x > 0,\ \mathbb{P}$-a.s.*

The maps, $x \mapsto U^i(x,t)$ and $x \mapsto V^i(x,t)$ are strictly increasing ($U^i_x, V^i_x > 0$) and strictly concave ($U^i_{xx}, V^i_{xx} < 0$) for any $t \geqslant 0,\ x > 0,\ \mathbb{P}$-a.s.. Furthermore, $\int_0^t |U^i(x,s)|^2 ds < \infty$, for any $x > 0,\ t \geqslant 0$, $\mathbb{P}$-a.s.

From Assumption 7.1.1, for $i \in \{1,\dots,n\}$ the Itô decomposition of the forward map is

$$\mathrm{d}Q^i(x,t) = e^{-\rho_i t}U^i_t(x,t)\mathrm{d}t - \rho_i e^{-\rho_i t}U^i(x,t)\mathrm{d}t + e^{-\rho_i t}V^i(\hat{c}^i_t x,t)\mathrm{d}t, \quad Q^i(x,0) = u^i_0(x). \tag{7.1.1}$$

For posterior use, we recall the notion of Fenchel-Legendre transform applied to a random field V under the above assumption.

Definition 7.1.2. *Let $V : \Omega \times (0,\infty) \times [0,\infty) \to \mathbb{R}$ be a random field such that $x \mapsto V(x,t)$ is a $\mathbb{P}$-a.s. strictly concave function for all $t \geqslant 0$. Define the random field $\widetilde{V} : \Omega \times (0,\infty) \times [0,\infty) \to \mathbb{R}$ as $\widetilde{V}(x',t) := \sup_{x>0}\left\{V(x,t) - x'x\right\}$ for any $x' > 0$, $t \geqslant 0$ and $\omega \in \Omega$.*

Then, we call $\widetilde{V}$ the Fenchel-Legendre transform *of V.*

For a map V satisfying Assumption 7.1.1, its differentiability and strict concavity ensures $\widetilde{V}$ is always well-defined and can be computed (up to a closed form).

Here we derive a PDE with random coefficients and an optimal investment-consumption strategy for a smooth relative performance criteria of zero-volatility of some agent i assuming that all other agents $j \neq i$ have made their investment decisions.

Proposition 7.1.3 (Best responses)**.** *Fix $i \in \{1,\dots,n\}$. Assume that each manager $j \neq i$ follows*

$(\pi^j, c^j) \in \mathcal{A}^j$. *Consider the PDE with stochastic coefficients for* $(x,t) \in (0,\infty) \times [0,\infty)$ *given by*

$$
\begin{aligned}
U_t^i =&\rho_i U^i + \Bigg(\theta_i \overline{\mu\pi_t}^{(-i)} - r(1-\theta) - \frac{\theta_i\sigma_i\overline{\sigma\pi_t}^{(-i)}(\mu_i - \theta_i\sigma_i\overline{\sigma\pi_t}^{(-i)})}{\nu_i^2+\sigma_i^2} - \frac{\theta_i}{2}\overline{\Sigma\pi_t^2}^{(-i)} \\
&\qquad - \frac{\theta_i^2}{2}\left(\left(\overline{\sigma\pi_t}^{(-i)}\right)^2 + \frac{1}{n-1}\overline{(\nu\pi_t)^2}^{(-i)}\right)\Bigg) x U_x^i \qquad (7.1.2)\\
&+ \frac{(\mu_i - \theta_i\sigma_i\overline{\sigma\pi_t}^{(-i)})^2}{2(\nu_i^2+\sigma_i^2)}\frac{(U_x^i)^2}{U_{xx}^i} + \frac{1}{2}\Bigg(\left(\theta_i\overline{\sigma\pi_t}^{(-i)}\right)^2\left(\frac{\sigma_i^2}{\nu_i^2+\sigma_i^2}-1\right) - \frac{\theta_i^2}{n-1}\overline{(\nu\pi_t)^2}^{(-i)}\Bigg)x^2U_{xx}^i \\
&+ \theta_i \bar{c}_t^{(-i)}U_x^i - \widetilde{V}^i(U_x^i, t),
\end{aligned}
$$

where $\widetilde{V}^i$ *stands for Fenchel-Legendre transform of* V^i *in variable* x. *Assume that for admissible initial conditions* $U^i(\cdot,0) = u_0^i(\cdot),\ V^i(\cdot,0) = v_0^i(\cdot)$, *the PDE has a smooth solution* (U^i, V^i), *that is not necessarily unique, but satisfy Assumption 7.1.1.*

Define the strategy $(\pi^{i,*}, c^{i,*})$

$$
\pi_t^{i,*} = \frac{1}{\nu_i^2+\sigma_i^2}\Big(\theta_i\sigma_i\overline{\sigma\pi_t}^{(-i)} - \big(\mu_i - \theta_i\sigma_i\overline{\sigma\pi_t}^{(-i)}\big)\frac{U_x^i(\widehat{X}_t^{i,*},t)}{U_{xx}^i(\widehat{X}_t^{i,*},t)\widehat{X}_t^{i,*}}\Big), \qquad (7.1.3)
$$

$$
c_t^{i,*} = \frac{(V_x^i)^{-1}\Big(U_x^i(\widehat{X}_t^{i,*},t)\big(\tilde{c}_t^{(-i)}\big)^{\theta_i},t\Big)\big(\tilde{c}_t^{(-i)}\big)^{\theta_i}}{\widehat{X}_t^{i,*}}, \qquad (7.1.4)
$$

where $\widehat{X}^{i,*}$ *solves* (5.4.5) *with* $(\pi^{i,*}, c^{i,*})$ *being used.*

Then, in the sense of Definition 5.5.2, if $(\pi^{i,*}, c^{i,*}) \in \mathcal{A}^i$ *and if* $\widehat{X}^{i,*}$ *is well-defined, then* $Q^i(x,t)$ *is a forward relative performance process for manager* i *and, moreover, the policy* $(\pi^{i,*}, c^{i,*})$ *is optimal.*

Let us recall the concept of a CRRA utility map. Following [79, Section 5], we say a utility map U is of *Constant Relative Risk Aversion* (CRRA) type if the *local risk tolerance function* $r : \Omega \times (0,\infty) \times [0,\infty) \to \mathbb{R}$, given by the quotient $r(x,t) = -U_x(x,t)/U_{xx}(x,t)$ is linear in space, i.e. $r(x,t) = \delta x$, for any $\delta > 0$ and $t \geqslant 0,\ x > 0$. This is the case for the classical power utility function, see Section 7.2 next.

By direct inspection of the expression for $\pi^{i,*}$ in (7.1.3), if the local risk tolerance function satisfies $r^i(x,t) = \delta_i x$, for any $t > 0$ (e.g., a CRRA utility) then the optimal investment strategy is constant throughout time if additionally all other agents also choose a constant investment strategy.

Corollary 7.1.4 (Constant investment strategies under CRRA)**.** *Assume* U^i *and* V^i *to be of CRRA type, i.e., the* local risk tolerance function *is linear in space uniformly. Assume further that all agents* $j \neq i$ *invest according to constant investment strategies* $\pi^j \in \mathbb{R}$. *Then,* $\pi^{i,*}$ *is constant.*

At this point it is not clear that if the agents $j \neq i$ use deterministic continuous consumption strategies c^j then the optimal consumption strategy $c^{i,*}$ in (7.1.4) is also so. We next prove the "best responses" result, Proposition 7.1.3.

Proof of Proposition 7.1.3. From (5.4.3) we have access to the dynamics of $d(X_t^i\big(\overline{X}_t^{(-i)}\big)^{-\theta_i})$ and $d\widehat{X}^i$. Using (7.1.1), we apply the Itô formula to $Q^i(\widehat{X}_t^i, t) = Q^i\big(X_t^i\big(\overline{X}_t^{(-i)}\big)^{-\theta_i}, t\big)$, and obtain using the notation set before (5.4.5)

$$
\begin{aligned}
\mathrm{d}Q^i(\widehat{X}^i_t,t) &= e^{-\rho_i t}\Big[U^i_t(\widehat{X}^i_t,t)\mathrm{d}t - \rho_i U^i(\widehat{X}^i_t,t)\mathrm{d}t + U^i_x(\widehat{X}^i_t,t)\mathrm{d}\widehat{X}^i_t + \frac{1}{2}U^i_{xx}(\widehat{X}^i_t,t)\mathrm{d}\langle\widehat{X}^i_t\rangle \\
&\qquad + V^i(\hat{c}^i_t\widehat{X}^i_t,t)\mathrm{d}t\Big] \\
&= e^{-\rho_i t}\Big[U^i_t(\widehat{X}^i_t,t)\mathrm{d}t - \rho_i U^i(\widehat{X}^i_t,t)\mathrm{d}t + U^i_x(\widehat{X}^i_t,t)\Big(\mu_i\pi^i_t - \theta_i\overline{\mu\pi_t}^{(-i)} + r(1-\theta) \\
&\qquad + \frac{\theta_i}{2}\overline{\Sigma\pi^2_t}^{(-i)} + \frac{\theta_i^2}{2}\Big(\big(\overline{\sigma\pi_t}^{(-i)}\big)^2 + \frac{1}{n-1}\overline{(\nu\pi_t)^2}^{(-i)}\Big) - \theta_i\sigma_i\pi^i_t\overline{\sigma\pi_t}^{(-i)}\Big)\overline{X}^i_t\mathrm{d}t \\
&\qquad + U^i_x(\widehat{X}^i_t,t)\big(\sigma_i\pi^i_t - \theta_i\overline{\sigma\pi_t}^{(-i)}\big)\widehat{X}^i_t\mathrm{d}B_t \qquad (7.1.5)\\
&\qquad + U^i_x(\widehat{X}^i_t,t)\widehat{X}^i_t\Big(\nu_i\pi^i_t\mathrm{d}W^i_t - \theta_i\big(\frac{1}{n-1}\sum_{k\neq i}^{n}\nu_k\pi^k_t\mathrm{d}W^k_t\big)\Big) \\
&\qquad + \frac{1}{2}U^i_{xx}(\widehat{X}^i_t,t)\Big((\nu_i\pi^i_t)^2 + \frac{\theta_i^2}{n-1}\overline{(\nu\pi_t)^2}^{(-i)} + \Big(\sigma_i\pi^i_t - \theta_i\overline{\sigma\pi_t}^{(-i)}\Big)^2\Big)\big(\widehat{X}^i_t\big)^2\mathrm{d}t \\
&\qquad - U^i_x(\widehat{X}^i_t,t)\big(c^i_t - \theta_i\overline{c}^{(-i)}_t\big)\widehat{X}^i_t\mathrm{d}t + V^i(\hat{c}^i_t\widehat{X}^i_t,t)\mathrm{d}t\Big],
\end{aligned}
$$

with $Q^i(\widehat{X}^i_0,0) = Q^i\big(x^i_0(\overline{x}^{(-i)}_0)^{-\theta_i},0\big) = u^i_0\big(x^i_0(\overline{x}^{(-i)}_0)^{-\theta_i}\big)$ and that the B, W^j are all i.i.d.

By Definition 5.5.2, the process $Q^i(\widehat{X}^i_t,t)$ becomes a martingale at the optimal $(\pi^{i,*},c^{i,*})$, hence, direct computations using first order conditions (∂_{π^i}"drift" $=\partial_{c^i}$"drift" $=0$) yield

$$
\begin{cases}
\pi^i_t &= \dfrac{1}{\nu_i^2+\sigma_i^2}\Big(\theta_i\sigma_i\overline{\sigma\pi_t}^{(-i)} - \big(\mu_i - \theta_i\sigma_i\overline{\sigma\pi_t}^{(-i)}\big)\dfrac{U^i_x(\widehat{X}^i_t,t)}{U^i_{xx}(\widehat{X}^i_t,t)\widehat{X}^i_t}\Big), \\
c^i_t &= \dfrac{(V^i_x)^{-1}\Big(U^i_x(\widehat{X}^i_t,t)\big(\tilde{c}^{(-i)}_t\big)^{\theta_i},t\Big)\big(\tilde{c}^{(-i)}_t\big)^{\theta_i}}{\widehat{X}^i_t}.
\end{cases} \qquad (7.1.6)
$$

Injecting π^i_t in the drift term of (7.1.5) and simplifying we arrive at the consistency condition (7.1.2), we do not carry out this step explicitly, nonetheless, using that the pair (U^i,V^i) solves (7.1.2), equation (7.1.5) simplifies to

$$
\begin{aligned}
\mathrm{d}Q^i(\widehat{X}^i_t,t) &= e^{-\rho_i t}\Big[U^i_x(\widehat{X}^i_t,t)\widehat{X}^i_t\Big(\nu_i\pi^i_t\mathrm{d}W^i_t - \theta_i\big(\frac{1}{n-1}\sum_{k\neq i}^{n}\pi^k_t\nu_k\mathrm{d}W^k_t\big)\Big) \\
&\quad + U^i_x(\widehat{X}^i_t,t)\Big(\sigma_i\pi^i_t - \theta_i\overline{\sigma\pi_t}^{(-i)}\Big)\widehat{X}^i_t\mathrm{d}B_t \\
&\quad + \frac{1}{2}U^i_{xx}(\widehat{X}^i_t,t)\frac{1}{\nu_i^2+\sigma_i^2}\left|\pi^i_t(\nu_i^2+\sigma_i^2) - \Big(\theta_i\sigma_i\overline{\sigma\pi_t}^{(-i)} - \mu_i\frac{U^i_x(\widehat{X}^i_t,t)}{U^i_{xx}(\widehat{X}^i_t,t)\widehat{X}^i_t}\Big)\right|^2(\widehat{X}^i_t)^2\mathrm{d}t \qquad (7.1.7)\\
&\quad + V^i(\hat{c}^i_t\widehat{X}^i_t,t)\mathrm{d}t - U^i_x(\widehat{X}^i_t,t)c^i_t\widehat{X}^i_t\mathrm{d}t - \widetilde{V}^i(U^i_x(\widehat{X}^i_t,t),t)\mathrm{d}t\Big].
\end{aligned}
$$

The concavity assumption of $U^i(x,t)$ implies that the third term of the expression above is non-positive and vanishes when (7.1.6) holds. At the same time by the definition of $\widetilde{V}^i$, the Fenchel-Legendre transform of V^i, we notice that

$$
\widetilde{V}^i(U^i_x(\widehat{X}^i_t,t),t) = V^i(\hat{c}^{i,*}_t\widehat{X}^i_t,t) - U^i_x(\widehat{X}^i_t,t)c^{i,*}_t\widehat{X}^i_t,
$$

that yields non-positivity and extremality of the last line of (7.1.7), when $c^i = c^{i,*}$. We can conclude that, if $(\pi^{i,*}_t,c^{i,*}_t) = (\pi^i_t,c^i_t) \in \mathcal{A}^i$ and the associated process $\widehat{X}^{i,*}$ is well-defined (solution to (5.4.5) with $(\pi^{i,*},c^{i,*})$), the process $U^i(\widehat{X}^{i,*}_t,t)$ is a local-martingale, otherwise it is a local supermartingale. The result concludes. □

In contrast with [72] we do not know the explicit form of the utility as it is part of the solution. Hence we exploit convex duality properties to prove the extremality argument in combination with HJB-type methodology. In [52] or [36] the authors only explored the investment problem and this issue does not appear.

7.2 Forward performance with initial CRRA power preference

For $i \in \{1, \ldots, n\}$, the dynamics of Q^i is given by (7.1.1). Then, the solution to the PDE given in Equation (7.1.2) has the following form

$$U^i(x,t) = \begin{cases} \frac{1}{1-\frac{1}{\delta_i}} x^{1-\frac{1}{\delta_i}} f_i(t), & \delta_i \neq 1,\ \delta_i > 0, \\ \log(x) f_i(t) + h_i(t), & \delta_i = 1, \end{cases} \tag{7.2.1}$$

where $x > 0$ and $(f_i(t))_{t \geqslant 0}, (h_i(t))_{t \geqslant 0}$ are the maps independent of x satisfying $f_i(0) = 1$ and $h_i(0) = 0$, are sufficiently integrable and $t \mapsto f_i(t),\ t \mapsto h_i(t)$ are differentiable.

Remark 7.2.1. *For the sake of simplicity we omit the presentation for the logarithmic utility and emphasise that the optimal policy for this case coincides with the general power case with $\delta_i = 1$ plugged inside.*

The structure of the consistency condition (7.1.2) implies V^i (see [41, Proposition 4.5]) as

$$V^i(x,t) = \frac{1}{\epsilon_i} \frac{1}{1 - \frac{1}{\delta_i}} x^{1-\frac{1}{\delta_i}} g_i(t), \tag{7.2.2}$$

where $x > 0$ and $(g_i(t))_{t \geqslant 0}$ is a map independent of x satisfying $g_i(0) = 1$ and sufficiently integrable. We refer to the constants $\delta_i > 0,\ \delta_i \neq 1$, and $\theta_i \in [0,1]$ as *personal risk tolerance* and *competition weight*, respectively, whereas $\epsilon_i > 0$ captures the *relative importance the agent assigns to the wealth compared to consumption*. Finally, ρ_i stands for *personal discount rate* of the agent.

In this case of initial power preferences we have that the *local risk tolerance function* for U^i satisfies $r^i = -U^i_x / U^i_{xx} = \delta_i x$, and hence

$$\pi_t^{i,*} = \frac{1}{\nu_i^2 + \sigma_i^2} \Big(\theta_i \sigma_i \overline{\sigma\pi_t}^{(-i)} + \Big(\mu_i - \theta_i \sigma_i \overline{\sigma\pi_t}^{(-i)} \Big) \delta_i \Big).$$

We now proceed to find this case's optimal consumption map c^i and utility fields. Injecting the expressions above for $U^i(x,t),\ V^i(x,t)$ in (7.1.2) together with the optimal consumption policy given by (7.1.6) yields an ODE system for (f_i, g_i, c^i). We have

$$\begin{cases} c_t^i = \big(\tilde{c}_t^{(-i)}\big)^{\theta_i(1-\delta_i)} \epsilon_i^{-\delta_i} \left(\dfrac{g_i(t)}{f_i(t)} \right)^{\delta_i}, \\ f_i'(t) + \Big(\eta_i + \big(1 - \frac{1}{\delta_i}\big) \theta_i \bar{c}_t^{(-i)} \Big) f_i(t) + \frac{\epsilon_i^{-\delta_i}}{\delta_i} \big(\tilde{c}_t^{(-i)}\big)^{\theta_i(1-\delta_i)} f_i(t) \left(\dfrac{g_i(t)}{f_i(t)} \right)^{\delta_i} = 0 \ , \end{cases} \tag{7.2.3}$$

where

$$\begin{aligned} \eta_t^i = \Big(1 - \frac{1}{\delta_i}\Big) \bigg(& \frac{\delta\big(\mu_i - \theta_i \sigma_i \overline{\sigma\pi_t}^{(-i)} (1 - \frac{1}{\delta_i})\big)^2}{2(\nu_i^2 + \sigma_i^2)} - r(1 - \theta_i) - \theta_i \overline{\mu\pi_t}^{(-i)} + \frac{\theta_i}{2} \overline{\Sigma\pi_t^2}^{(-i)} \\ & + \frac{\theta_i^2}{2} \Big(\big(\overline{\sigma\pi_t}^{(-i)}\big)^2 + \frac{1}{n-1} \overline{(\nu\pi_t)^2}^{(-i)} \Big) \big(1 - \frac{1}{\delta_i}\big) \bigg) - \rho_i. \end{aligned} \tag{7.2.4}$$

We have two equations for three unknowns, now we need one further assumption for the nature of f_i and g_i in order to solve the system.

Assumption 7.2.2. *Let $g_i(t) = f_i(t)^{1-\kappa_i}$, where $\kappa_i \in \mathbb{R}$.*

Remark 7.2.3. *Here, as opposed to* [72] *or* [64] *where one is naturally led to $\kappa_i = 1$, we find a non-trivial time-dependence structure of the consumption utility. We highlight that the FPP approach naturally supports the time dynamics of the utilities in contrast to the HJB approach that does not additionally weight the integrand with a function of time.*

We have two distinct cases.

Case 1: Let $\kappa_i \neq 0$. We obtain and solve a classical ODE of *Bernoulli equation* type (we omit the dependence in t for simplification)

$$f_i' + a_i(t)f_i + b_i(t)f_i^{1-\kappa_i\delta_i} = 0 \;\text{ for }\; a_i(t) = \eta_t^i - \theta_i\bar{c}_t^{(-i)} \;\text{ and }\; b_i(t) = \frac{\epsilon_i^{-\delta_i}}{\delta_i}\big(\tilde{c}_t^{(-i)}\big)^{\theta_i(1-\delta_i)}.$$

Substituting $k_i(t) := f_i(t)^{-\kappa_i\delta_i}$ we find the ODE for $k_i(t)$,

$$\frac{1}{\kappa_i\delta_i}k_i' - a_i(t)k_i - b_i(t) = 0.$$

Hence, solving the ODE we have

$$k_i(t) = e^{\kappa_i\delta_i\int_0^t a_i(s)\mathrm{d}s} + \kappa_i\delta_i\int_0^t e^{\kappa_i\delta_i\int_s^t a(r)dr}b_i(s)\mathrm{d}s,$$

and $f_i(t) = k_i(t)^{\frac{1}{\kappa_i\delta_i}}$, $g_i(t) = f_i(t)^{1-\kappa_i} = k_i(t)^{\frac{1-\kappa_i}{\kappa_i\delta_i}}$. Plugging the f_i, g_i expressions in the first equation of (7.2.3) along with the general form of the consumption strategy we can extract the form of optimal consumption map.

Case 2. Let $\kappa_i = 0$. In this case, the optimal consumption policy of agent i simplifies considerably. More precisely, from (7.2.3) we obtain

$$c_t^i = \big(\tilde{c}_t^{(-i)}\big)^{\theta_i(1-\delta_i)}\epsilon_i^{\delta_i}.$$

7.2.1 The Forward Nash equilibrium

In view of the *best responses* of Proposition 7.1.3 we now solve for the *simultaneous best responses* as to establish the existence of a Nash equilibrium.

Definition 7.2.4 (Forward Nash equilibrium)**.** *Let for any $i \in \{1,\dots,n\}$, $(\pi^{i,*}, c^{i,*}) \in \mathcal{A}^i$ and $(\pi^{i,*}, c^{i,*})$ is the optimal strategy in the sense of Proposition 7.1.3. Let $\big(\pi^i, c^i\big) \in \mathcal{A}^i$. Let Q^i be the $\mathbb{F}$-progressively measurable random field satisfying*

$$Q^i(x,t) := e^{-\rho_i t}U^i(x,t) + \int_0^t e^{-\rho_i s}V^i(\hat{c}_s^i x, s)ds,$$

where $\hat{c}^i$ is given by (5.4.4) *with $c^j = c^{j,*},\ j \neq i$ and $U^i, V^i : \Omega \times (0,\infty) \times [0,\infty) \to \mathbb{R}$ are two other $\mathbb{F}$-progressively measurable random fields.*

A forward Nash equilibrium consists of n-triples of $\mathcal{F}_t$-adapted maps $(Q^i, \pi^{i,*}, c^{i,*})$ with $i = 1,\dots,n$ such that for any $t \geqslant 0$ the following conditions hold.

i) The mappings $x \mapsto U^i(x,t)$ and $x \mapsto V^i(x,t)$ are $\mathbb{P}$-a.s. strictly increasing and strictly concave;

ii) Let managers $j \neq i$ acting according to $(\pi^{j,*}, c^{j,*})$, manager i acts according to $(\pi^i, c^i) \in \mathcal{A}^i$, and take $\widehat{X}^i$ as the associated relative performance wealth process (5.4.5) *and* relative consumption metric *$\hat{c}^i$ given by* (5.4.4). *Then $Q^i(\widehat{X}_t^i, t)$ is a (local) supermartingale.*

iii) Let all *managers acting according to $(\pi^{j,*}, c^{j,*})$, and take $\widehat{X}^{i,*}$ as the associated* relative performance wealth process (5.4.5) *and* relative consumption metric *$\hat{c}^{i,*}$ given by* (5.4.4). *Then $Q^i(\widehat{X}_t^{i,*}, t)$ is a (local) martingale.*

Equilibrium with FPPs of separable power factor form

In order to obtain explicit results we focus on the *separable power factor form* case of (7.2.1)-(7.2.2) for which $U^i_x/U^i_{xx} = V^i_x/V^i_{xx} = -\delta_i x$. More notably, at the level at which we have formulated our problem, we recover the results of [72, Theorem 3] for which one has $U^i_x/U^i_{xx} = -\delta_i x$, for all t and $x > 0$ (note their Remark 5).

Assumption 7.2.5. *Set* $g_i(t) := f_i(t)^{1-\kappa}$ *in* (7.2.1)-(7.2.2) *for some* $\kappa \in \mathbb{R}$ *for any* $i = 1, \dots, n$.

One can see that the PDE (7.1.2) depends on the two unknown functions U and V, and thus makes for an under-determined system. Hence, Assumption 7.2.5 is essential to ensure that the equation is solvable for the unique Nash equilibrium. At the same time we have to assume κ to be constant across the whole population in order to get a solution for the Nash equilibria.

Theorem 7.2.6. *Let the conditions of Proposition 7.1.3 hold for all agents* $i \in \{1, \dots, n\}$. *Assume furthermore that agents have separable power factor form forward maps* U^i, V^i *with initial conditions.*

$$U^i(x,0) = \epsilon_i V^i(x,0) = \frac{1}{1-\frac{1}{\delta_i}} x^{1-\frac{1}{\delta_i}}, \quad \epsilon_i > 0,\ \delta_i > 0,\ \delta_i \neq 1,$$

as in (7.2.1)-(7.2.2). *Define the quantities*

$$\varphi^\sigma_n = \frac{1}{n}\sum_{k=1}^n \delta_k \frac{\sigma_k \mu_k}{\nu_k^2 + \sigma_k^2\big(1 + \frac{\theta_k(1-\delta_k)}{n-1}\big)}, \qquad \psi^\sigma_n = \frac{1}{n-1}\sum_{k=1}^n \theta_k(1-\delta_k)\frac{\sigma_k^2}{\nu_i^2 + \sigma_i^2\big(1+\frac{\theta_k(1-\delta_k)}{n-1}\big)},$$

$$\lambda_i = \epsilon_i^{-\frac{\delta_i}{1+\frac{\theta_i}{n-1}(1-\delta_i)}} \Big(\widetilde{\epsilon^\delta}\Big)^{\frac{\theta_i(1-\delta_i)}{(\overline{\theta(1-\delta)}-1)\left(1+\frac{\theta_i}{n-1}(1-\delta_i)\right)}}, \quad \beta_i = \frac{1}{1+\frac{\theta_i}{n-1}(1-\delta_i)}\Big(\frac{\theta_i(1-\delta_i)}{\overline{\theta(1-\delta)}-1}\overline{\eta\delta} - \eta_i\delta_i\Big),$$

$$\widetilde{\epsilon^\delta} = \Bigg(\prod_{k=1}^n \epsilon_k^{\frac{\delta_k}{1+\frac{\theta_k}{n-1}(1-\delta_k)}}\Bigg)^{\frac{1}{n}}, \quad \overline{\eta\delta} = \frac{1}{n}\sum_{k=1}^n \frac{\eta_k\delta_k}{1+\frac{\theta_k}{n-1}(1-\delta_k)},$$

and $$\overline{\theta(1-\delta)} = \frac{1}{n-1}\sum_{k=1}^n \frac{\theta_k(1-\delta_k)}{1+\frac{\theta_k}{n-1}(1-\delta_k)}.$$

Let as well, for $i = 1, \dots, n$, *the map* η^i_t *be defined as*

$$\begin{aligned}\eta^i_t = \eta_i := \Big(1 - \frac{1}{\delta_i}\Big)\Bigg(&\frac{\delta\big(\mu_i - \theta_i\sigma_i\overline{\sigma\pi}^{(-i)}(1-\frac{1}{\delta_i})\big)^2}{2(\nu_i^2+\sigma_i^2)} - r(1-\theta_i) - \theta_i\overline{\mu\pi}^{(-i)} + \frac{\theta_i}{2}\overline{\Sigma\pi^2}^{(-i)} \\ &+ \frac{\theta_i^2}{2}\Big(\big(\overline{\sigma\pi}^{(-i)}\big)^2 + \frac{1}{n-1}\overline{(\nu\pi)^2}^{(-i)}\Big)\Big(1-\frac{1}{\delta_i}\Big)\Bigg) - \rho_i,\end{aligned} \tag{7.2.5}$$

where the explicit expressions for $\overline{\sigma\pi}^{(-i)}$, $\overline{\mu\pi}^{(-i)}$, $\overline{(\nu\pi)^2}^{(-i)}$ *and* $\overline{\Sigma\pi^2}^{(-i)}$ *are found in the supplementary material in* (A.2.1), (A.2.2), (A.2.3), (A.2.4) *respectively.*

If $\psi^\sigma_n \neq 1$ *and Hypothesis 7.2.5 holds* ($\kappa_i = \kappa$ *for all* i) *then a unique optimal candidate strategy exists with the optimal* $\pi^{i,*}$ *and* $c^{i,*}$ *given by*

$$\pi^{i,*} = \frac{1}{\nu_i^2 + \sigma_i^2\big(1+\frac{\theta_i(1-\delta_i)}{n-1}\big)}\Big(\theta_i\sigma_i(1-\delta_i)\Big(1+\frac{1}{n-1}\Big)\frac{\varphi^\sigma_n}{1-\psi^\sigma_n} + \mu_i\delta_i\Big), \tag{7.2.6}$$

$$c^{i,*}_t = \begin{cases}\Big(\frac{1}{\beta_i} + \Big(\frac{1}{\lambda_i} - \frac{1}{\beta_i}\Big)e^{\kappa\beta_i t}\Big)^{-1}, & \beta_i \neq 0,\\ \big(-\kappa t + \frac{1}{\lambda_i}\big)^{-1}, & \beta_i = 0,\end{cases} \tag{7.2.7}$$

and η^i_t *in* (7.2.5) *is constant in time. Moreover, optimal candidate strategy is a* forward Nash equilibrium strategy *if either* $\kappa \leqslant 0$ *or* $\kappa > 0$, $\beta_i > \lambda_i$.

The forward Nash equilibria is given by the n*-triples* $\{(Q^{i,*}, \pi^{i,*}, c^{i,*})\}_{i=1,\dots,n}$ *where the* $Q^{i,*}(x,t) = e^{-\rho_i t}U^{i,*}(x,t) + \int_0^t e^{-\rho_i s}V^{i,*}(\hat{c}^i_s x, s)ds$, *with* $U^{i,*}, V^{i,*}$ *the solution of* (7.1.2) *(of the form* (7.2.1)-

(7.2.2)*) with the optimal strategies* $\pi^{\cdot,*}, c^{\cdot,*}$ *plugged-in. The maps* f_i, g_i *can be determined up to a closed formula given below in* (A.2.12).

If $\psi_n^\sigma = 1$ *or* $\kappa > 0,\ \beta_i \leqslant \lambda_i$ *then there exists no Nash equilibrium.*

Proof. The complete proof can be found in Section A.2 of the supplementary material. □

The parameter κ is interpreted as the *market-risk relative consumption preference* and in Section 7.4 below we discuss at length its economic features and comparative interpretation. The investment strategy (7.2.6) is a classical expression matching that of [36, 72, 73]. The expression for the consumption strategy (7.2.7) has similarities to the results found in [72, Equation (5)] with the crucial difference of κ. When $\kappa = 1$ then the results coincide with [72].

For some combination of parameters λ_i, β_i and κ, $(c_t^{i,*})_{t\geqslant 0}$ is not admissible due to the finite-time blow-up. We comment further on the admissibility of c^* in Section 7.4.

Remark 7.2.7 (Open question)**.** *To obtain a Nash equilibria result, we are restricted to* κ *being constant across the agents, and the same will happen for the MFG case in the next Section. The constraint is technical and stems from the difficulty in expressing the consistency equation for the averages of all* c^i *(i.e.,* (A.2.8) *is not longer reachable) and without such an equation the system cannot be solved. At the same time, for a "best response" all* κ_i*'s may be distinct. The open question is then if one allows for different* κ_i*'s, how to carry out the aggregation procedure as to express the consistency equation for the averages of all* c^i*?*

In [11] *An example of how this problem was overcome can be found . There, aggregation was achieved through the so-called weighted-dilated inf-convolution, and the authors found the dynamic equation of the Representative agent of their economy, which allowed them to solve their Nash-equilibrium problem. The representative agent methodology remains unexplored in the context of the FPP. For instance, it is unclear how inf-convolution can be carried out in this framework as the FPP is not known in advance. It is endogenously found as part of the problem's solution in opposition to the classic utility game-type problems.*

Lastly, if one sets $\kappa = 1$ then our FPP solution can be rewritten in respect to a time horizon T to recover exactly the results of [72] (and [73] if $V^i = 0$) and those of [52] for two players and no consumption. Such calculations are straightforward hence omitted. We close the section with a corollary for the single stock case.

Corollary 7.2.8. *Let* $\mu_i = \mu,\ \sigma_i = \sigma,\ \nu_i = 0$*, for all* $i = 1, \ldots, n$ *and* $\mu, \sigma > 0$*. Then* $\lambda_i,\ \beta_i,\ \theta_{crit},\ \overline{\rho\delta},\ \overline{\delta}$ *are expressed as*

$$\lambda_i = \epsilon_i^{-\frac{\delta_i}{1+\frac{\theta_i}{n-1}(1-\delta_i)}} \left(\widetilde{\epsilon^\delta}\right)^{\frac{\theta_i(1-\delta_i)}{(\overline{\theta(1-\delta)}-1)(1+\frac{\theta_i}{n-1}(1-\delta_i))}}, \quad \theta_{crit} = \frac{1 - \frac{1}{n-1}\sum_{k=1}^n \frac{\theta_k(1-\delta_k)}{1+\frac{\theta_k}{n-1}(1-\delta_k)}}{\overline{\delta}},$$

$$\overline{\delta} = \frac{1}{n-1}\sum_{k=1}^n \frac{\delta_k}{1+\frac{\theta_k}{n-1}(1-\delta_k)}, \quad \overline{\rho\delta} = \frac{1}{n}\sum_{k=1}^n \frac{\rho_k\delta_k}{1+\frac{\theta_k}{n-1}(1-\delta_k)}$$

$$\begin{aligned}\beta_i = &\frac{\mu^2}{2\sigma^2\big(1+\frac{\theta_i}{n-1}(1-\delta_i)\big)}\Big(1 - \frac{\delta_i}{1+\frac{\theta_i}{n-1}(1-\delta_i)}\Big)\Big(1 - \frac{\theta_i}{\theta_{crit}\big(1+\frac{\theta_i}{n-1}\big)}\Big)\Big(\delta_i + \frac{\theta_i}{\theta_{crit}}(1-\delta_i)\Big)\\ &+ \frac{r}{\big(1+\frac{\theta_i}{n-1}(1-\delta_i)\big)}(1-\delta_i)\left(1 - \frac{\theta_i}{\theta_{crit}}\right) + \frac{1}{\big(1+\frac{\theta_i}{n-1}(1-\delta_i)\big)}\frac{\theta_i}{\theta_{crit}}\frac{\overline{\rho\delta}}{\overline{\delta}}(1-\delta_i) + \rho_i\delta_i.\end{aligned}$$

Then, if $\theta_{crit} \neq 1$ *the optimal candidate strategy exists with the optimal* $(\pi^{i,*}, c^{i,*})$

$$\pi^{i,*} = \frac{\mu}{\sigma^2\big(1+\frac{\theta_i}{n-1}(1-\delta_i)\big)}\Big(\delta_i + \frac{\theta_i}{\theta_{crit}}(1-\delta_i)\Big),$$

$$c_t^{i,*} = \begin{cases}\left(\frac{1}{\beta_i} + \Big(\frac{1}{\lambda_i} - \frac{1}{\beta_i}\Big)e^{\kappa\beta_i t}\right)^{-1}, & \beta_i \neq 0,\\ \big(-\kappa t + \frac{1}{\lambda_i}\big)^{-1}, & \beta_i = 0.\end{cases}$$

Furthermore, the optimal candidate strategy is a Nash equilibrium, if either $\kappa < 0$*, or* $\kappa > 0,\ \lambda_i < \beta_i$*.*

7.3 Forward mean-field game

The following section mimics one of 6.2, for completeness and clarity of presentation we repeat these arguments and emphasise the notable discrepancies.

As in the CARA result, one sees that the optimal strategy and forward utility map of 7.2.6 for some agent depend on that agent's specific parameters (model parameters, initial wealth, risk tolerance and performance concern, etc.) and on certain averages of the parameters of all agents. This makes a case for a MFG approach to the game. In this section, and inspired by [36], we formalise the concept of CRRA forward mean field Nash game with consumption. We use the concept of *type distributions* introduced in [73].

We focus on forward maps that at time $t = 0$ are of power type,

$$U^i(x, m, 0) = \epsilon_i V^i(x, m, 0) = \begin{cases} \frac{1}{1-\frac{1}{\delta_i}} \left(x m^{-\theta_i} \right)^{1-\frac{1}{\delta_i}}, & \delta > 0,\ \delta_i \neq 1, \\ \log(x m^{-\theta_i}), & \delta_i = 1, \end{cases}$$

where $x > 0, m > 0$ denote the wealth of agent and the average wealth of other agents respectively (for U^i) or the consumption of the agent and the average consumption of the other agents respectively (for V^i). We refer to $\delta_i > 0$ and $\theta_i \in [0, 1]$ as *personal risk tolerance*, *competition weight* parameters, respectively, whereas $\epsilon_i > 0$ captures the *relative importance the agent assigns to the wealth compared to consumption*. Finally, for admissible set of strategy pairs $(\pi^{i,*}, c^{i,*}),\ i = 1, \dots, n$, we recall the Forward relative performance process $Q^i(x, t) := e^{-\rho_i t} U^i(x, t) + \int_0^t e^{-\rho_i s} V^i(\hat{c}^i_s x, s) \mathrm{d}s$, where c^i_s is given by (5.4.4). We refer to ρ_i as *personal discount rate*. Here, as in the corresponding n-player game (see Remark 7.2.1), we skip the logarithmic case.

For the n-agent game, we define for each agent $i = 1, \dots, n$ the *type vector*

$$\zeta_i := (x^i_0, \delta_i, \theta_i, \epsilon_i, \rho_i, \breve{\mu}_i, \nu_i, \sigma_i),$$

which characterises uniquely each agent i. These *type vectors* induce an empirical measure, called the *type distribution*, which is probability measure on the *type space*

$$\mathcal{Z}^e := (0, \infty) \times (0, \infty) \times [0, 1] \times (0, \infty) \times [0, \infty) \times (0, \infty) \times [0, \infty) \times [0, \infty), \tag{7.3.1}$$

given by

$$m_n(A) = \frac{1}{n} \sum_{i=1}^{n} 1_A(\zeta_i), \quad \text{for Borel sets } A \subset \mathcal{Z}^e.$$

Recalling Section 6.2 we assume that as the number of agents becomes large, $n \to \infty$, the above empirical measure m_n has a weak limit m in the sense that $\int_{\mathcal{Z}^e} f\, dm_n \to \int_{\mathcal{Z}^e} f\, dm$ for every bounded continuous function f on $\mathcal{Z}^e$. For example, this holds almost surely if the ζ_i's are i.i.d. samples from m. Let $\zeta = (\xi, \delta, \theta, \epsilon, \rho, \breve{\mu}, \nu, \sigma)$ denote a $\mathcal{Z}^e$-valued random variable with limiting distribution m.

As in Section 6.2 The *mean field game* (MFG) we define next allows us to derive the limiting strategy as the outcome of a self-contained equilibrium problem, which intuitively represents a game with a continuum of agents with type distribution m. Rather than directly modelling a continuum of agents we follow the MFG paradigm of modelling a single generic agent who we view as randomly selected from the population. The probability measure m represents the distribution of type parameters among the continuum of agents, equivalently, the generic agent's type vector is a random variable with law m. Heuristically, each agent in the continuum trades in a single stock driven by two Brownian motions, one of which is unique to this agent and one of which is common to all agents. We extend the Forward Nash equilibrium of Definition 7.2.4 to the MFG setting below.

7.3.1 Agents as type-distribution samples and the market

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F})_{t \geqslant 0}, \mathbb{P})$ be a stochastic basis supporting two independent Brownian motions $W = (W_t)_{t \geqslant 0}$ and $B = (B_t)_{t \geqslant 0}$ together with a random vector ζ having distribution m

and given by

$$\zeta = (\xi, \delta, \theta, \epsilon, \rho, \breve{\mu}, \nu, \sigma), \tag{7.3.2}$$

with values in the space $\mathcal{Z}^e$ defined in (7.3.1), and independent of W and B. Let $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ denote the smallest filtration satisfying the usual assumptions for which ζ is $\mathcal{F}_0^{\mathrm{MF}}$-measurable and both W and B are adapted. Let also $\mathbb{F}^B = (\mathcal{F}_t^B)_{t\in[0,T]}$ denote the natural filtration generated by the Brownian motion B.

The *generic agent's* wealth process is

$$dX_t = rX_t + \pi_t X_t\big((\breve{\mu} - r)\mathrm{d}t + \nu \mathrm{d}W_t + \sigma \mathrm{d}B_t\big) - c_t X_t \mathrm{d}t, \quad X_0 = \xi, \tag{7.3.3}$$

for a self-financing strategy, $(\pi_t)_{t\geqslant 0}$, standing for the fraction of wealth invested in the risky asset and consumption policy $(c_t^i)_{t\geqslant 0}$, representing the instantaneous rate of consumption per unit of wealth. Together they must belong to the admissible set

$$\mathcal{A}_{\mathrm{MF}} = \Big\{(\pi, c) : \mathbb{F}\text{-progressively measurable } \mathbb{R}\times(0,\infty)\text{-valued process } (\pi_t, c_t)_{t\geqslant 0},$$
$$\mathbb{E}\Big[\int_0^t (|\pi_s|^2 + |c_s|^2)\mathrm{d}s\Big] < \infty, \text{ for any } t > 0\Big\}.$$

The risk-free interest rate r is deterministic and fixed for the entire population. The random variable ξ is the initial wealth of the generic agent, whereas $(\breve{\mu}, \nu, \sigma)$ are the market parameters. As in previous sections we denote $\mu := \breve{\mu} - r$ as a "excess return". In the sequel, the parameters $\delta,\ \theta,\ \epsilon$ and ρ will affect the risk and consumption preferences of the generic agent. Each agent among the continuum will have different preference parameters and hence these eight parameters are $\mathcal{F}_0$-random, and each has the exact same interpretation an in the n-player game of the earlier section.

7.3.2 The mean field equilibrium

The formulation of the forward Nash game of Section 7.2 drives the formulation of the mean field game we discuss here, see also [36, 72]. Recall that in the MFG-formulation the *generic agent* has no influence on the average wealth of the continuum of agents, as but one agent amid a continuum.

We next introduce the concept of *mean field (MF)-forward relative performance*, π^*, c^* is the *MF-equilibrium* and, the main object of interest the *MF-Forward relative performance equilibrium*.

Assumption 7.3.1. *Assume the derivatives $U_t(x,t)$, $U_x(x,t)$, $U_{xx}(x,t)$ and $V_x(y,t)$, $V_{xx}(x,t)$ exist for all $t \geqslant 0,\ x > 0$, $\mathbb{P}$-a.s. and furthermore, $x \mapsto U(x,t)$ and $x \mapsto V(x,t)$ are strictly increasing ($U_x, V_x > 0$) and strictly concave ($U_{xx}, V_{xx} < 0$) for any $t \geqslant 0,\ x > 0$, $\mathbb{P}$-a.s.*

Given this setup we next define our concept of equilibrium (see also [36]).

Definition 7.3.2 (MF-CRRA-Forward relative performance equilibrium (for the generic manager)). *Let $(\overline{X}_t)_{t\geqslant 0}$ and $(\overline{\Gamma}_t)_{t\geqslant 0}$ be $\mathbb{F}^B$-adapted positive square integrable stochastic processes representing the geometric average wealth and geometric average consumption of the continuum of agents respectively. Let $(\pi, c) \in \mathcal{A}^{\mathrm{MF}}$ and $X^{\pi,c}$ solve* (7.3.3) *with π and c.*

Set a $\mathbb{F}^{\mathrm{MF}}$-progressively measurable random field $(Q(x,t))_{t\geqslant 0}$ having, for some $\rho \geqslant 0$, the dynamics

$$Q(x,t) = e^{-\rho t}U(x,t) + \int_0^t e^{-\rho s}V(\hat{c}_s x, s)ds, \tag{7.3.4}$$

where $\hat{c}_t = c_t \overline{\Gamma_t}^{-\theta}$ and $U, V : \Omega\times(0,\infty)\times[0,\infty) \to \mathbb{R}$ are two other $\mathbb{F}$-progressively measurable random fields.

The field Q is an MF-forward relative performance *for the* generic manager *if, for all $t \geqslant 0$, the following conditions hold:*

i) The mappings $x \mapsto U(x,t),\ x \mapsto V(x,t)$, are $\mathbb{P}$-a.s. strictly increasing and strictly concave;

ii) For any $(\pi, c) \in \mathcal{A}^{MF}$*,* $Q(X_t^{\pi,c}\overline{X}_t^{-\theta}, t)$ *is a (local) supermartingale and* $X^{\pi,c}$ *is the* generic agent's *wealth process solving* (7.3.3) *for the strategy* (π, c)*;*

iii) There exists $(\pi^*, c^*) \in \mathcal{A}^{MF}$ *such that* $Q(X_t^*\overline{X}_t^{-\theta}, t)$ *is a (local) martingale where* X^* *solves* (7.3.3) *with* (π^*, c^*) *plugged in as the strategy;*

iv) We call (π^*, c^*) *of iii) a* MF-equilibrium *if*

$$\overline{X}_t = \exp \mathbb{E}[\log X_t^* | \mathcal{F}_t^B] \quad \textit{and} \quad \overline{\Gamma}_t = \exp \mathbb{E}[\log c_t^* | \mathcal{F}_t^B] \quad \textit{for all } t \geqslant 0, \ \mathbb{P}\textit{-a.s.,}$$

where X^* *solves* (7.3.3) *with* (π^*, c^*) *plugged in as the strategy.*

We denote the sextuple $(U, V, \pi^*, c^*, \overline{X}, \overline{\Gamma})$ *satisfying i)-iv) the* MF-Forward relative performance equilibrium.

The last point can be understood as a fixed point argument which creates a compatibility condition between the generic agent within the continuum of agents. In fact, conditionally on the Brownian motion B each agent faces an independent noise W and an independent type vector ζ. As argued in [36, 72, 73], conditionally on B, all agents faces i.i.d. copies of the same optimisation problem. The law of large numbers suggests that the geometric average terminal wealth of the whole population should be $\exp \mathbb{E}[\log X_t^* | \mathcal{F}_t^B]$.

Our construction allows us to identify $\exp \mathbb{E}[\log X_t^* | \mathcal{F}_t^B]$ with a certain dynamics and, in turn, treat this component as an additional uncontrolled state process. This avoids altogether the conceptualisation of the master equation for models with different types of agents.

It is unclear if the definition for forward MFG we propose is sufficiently tractable (concretely point iv)) to address the case of random coefficients or the general setting of a forward utility map with a full Itô-field representation (with volatility). This same issue is already present in the earlier works [72, 73] if one wanted to generalise to random coefficients there. Nonetheless, the definition works very well within the Merton market model setup, yielding a particularly tractable framework of study to understand the inherent difficulties of the general case. Additionally, our goal is to further interpret the results as this is one of the first works combining MFGs with forward performance criteria.

7.3.3 Solving the optimisation problem

We now present the main result of this section which is the existence of a *MF-Forward relative performance equilibrium* for the generic manager according to Definition 7.3.2 within the context of a structural assumption similar to (7.2.1).

From the methodological point of view, the problem is solved as before. Apply Itô's formula to $Q(Z_t^{\pi,c}, t)$, determine the optimal strategy (π^*, c^*) and the consistency condition (the PDE) for Q such that the first three conditions of Definition 7.3.2 hold. The last condition, to show that (π^*, c^*) is indeed the MFG Forward equilibrium will follow by construction as the whole formulation is built with the fixed-point condition embedded from the start.

We next present the *separable power factor form* assumption for U and V.

Assumption 7.3.3. *Let* $U(x,t), V(x,t)$ *admit the following form, for* $x > 0,\ t \geqslant 0,$

$$U(x,t) = \frac{1}{1-\frac{1}{\delta}} x^{1-\frac{1}{\delta}} f(t), \qquad V(x,t) = \frac{1}{\epsilon} \frac{1}{1-\frac{1}{\delta}} x^{1-\frac{1}{\delta}} g(t),$$

for $t \mapsto f(t)$ *differentiable for any* $t \geqslant 0$ *and* $f(0) = 1$*, for some random variable* $\epsilon > 0$ *and* $g(t) = f(t)^{1-\kappa}$ *for some* $\kappa \in \mathbb{R}$ *(where we have* $g(0) = 1^{1-\kappa} = 1$*). Moreover, let for any* $t \geqslant 0,\ \int_0^t |f(s)|^2 ds < \infty,\ \mathbb{P}$*-a.s.*

The parameter $\delta > 0, \delta \neq 1$ represents the agent's risk tolerance while $\epsilon > 0$, associated to the consumption element, represents the relative importance the agent assigns to the consumption when compared to the wealth. The parameter κ represent the *market-risk relative consumption preference* (in Section 7.4 we discuss its economic features).

As in the previous section, this assumption is a natural one to make in order to solve (7.3.7) below (see additionally [41], [4, Definition 2.9] or [28, Section 4]).

Theorem 7.3.4. *Assume that $\delta > 0,\ \theta \in [0,1],\ \mu > 0,\ \sigma \geqslant 0,\ \nu \geqslant 0$ such that $\sigma^2 + \nu^2 > 0$. Define constants $\psi^\sigma, \varphi^\sigma, \psi^\mu, \varphi^\mu$, assume they are finite and have explicit form given by*

$$\psi^\sigma = \mathbb{E}\Big[\theta(1-\delta)\frac{\sigma^2}{\nu^2+\sigma^2}\Big], \varphi^\sigma = \mathbb{E}\Big[\delta\frac{\mu\sigma}{\nu^2+\sigma^2}\Big], \psi^\mu = \mathbb{E}\Big[\theta(1-\delta)\frac{\mu\sigma}{\nu^2+\sigma^2}\Big], \varphi^\mu = \mathbb{E}\Big[\delta\frac{\mu^2}{\nu^2+\sigma^2}\Big].$$

Assume further $\psi^\sigma \neq 1$ and define for some $\kappa \in \mathbb{R}$,

$$\pi^* = \frac{1}{\nu^2+\sigma^2}\Big(\theta(1-\delta)\sigma\frac{\varphi^\sigma}{1-\psi^\sigma} + \mu\delta\Big), \quad c_t^* = \begin{cases} \Big(\frac{1}{\beta} + \Big(\frac{1}{\lambda} - \frac{1}{\beta}\Big)e^{-\kappa\beta t}\Big)^{-1}, & \beta \neq 0, \\ \big(\kappa t + \frac{1}{\lambda}\big)^{-1}, & \beta = 0, \end{cases} \tag{7.3.5}$$

where

$$\lambda = \epsilon^{-\delta}\big(e^{\mathbb{E}[\delta\log\epsilon]}\big)^{\frac{\theta(1-\delta)}{\mathbb{E}[\theta(1-\delta)]-1}}, \qquad \beta = \frac{\theta(1-\delta)}{\mathbb{E}[\theta(1-\delta)]-1}\mathbb{E}[\eta\delta] - \eta\delta,$$

and

$$\begin{aligned}\eta = \Big(1-\frac{1}{\delta}\Big)\Big(&\frac{\delta\big(\mu - \theta\sigma\overline{\sigma\pi^*}(1-\frac{1}{\delta})\big)^2}{2(\nu^2+\sigma^2)} - r(1-\theta) \\ &- \theta_i\overline{\mu\pi^*} + \frac{\theta}{2}\overline{\Sigma(\pi^*)^2} + \frac{\theta^2}{2}\big(\overline{\sigma\pi^*}\big)^2\big(1-\frac{1}{\delta}\big)\Big) - \rho,\end{aligned} \tag{7.3.6}$$

where $\overline{\sigma\pi^*}$, $\overline{\mu\pi^*}$ and $\overline{\Sigma(\pi^*)^2}$ are given by (A.3.4)*,* (A.3.7) *and* (A.3.8) *respectively.*

Take the field Q of (7.3.4) *with U, V satisfying Hypothesis 7.3.1 and 7.3.3 (for the κ above), and the consistency PDE*

$$\begin{aligned}U_t = \theta\Big(&\frac{\varphi^\sigma}{1-\psi^\sigma}\psi^\mu + \varphi^\mu - \frac{\sigma\frac{\varphi^\sigma}{1-\psi^\sigma}(\mu - \theta\sigma\frac{\varphi^\sigma}{1-\psi^\sigma})}{\nu^2+\sigma^2} - r(1-\theta) - \frac{\theta}{2}\overline{\Sigma(\pi^*)^2} - \frac{\theta}{2}\Big(\frac{\varphi^\sigma}{1-\psi^\sigma}\Big)^2\Big)xU_x \\ &+ \rho U + \frac{\big(\mu - \theta\sigma\frac{\varphi^\sigma}{1-\psi^\sigma}\big)^2}{2(\nu^2+\sigma^2)}\frac{(U_x)^2}{U_{xx}} + \frac{1}{2}\theta^2\Big(\frac{\varphi^\sigma}{1-\psi^\sigma}\Big)^2\Big(\frac{\sigma^2}{\nu^2+\sigma^2} - 1\Big)x^2U_{xx} \\ &+ \theta U_x\bar{c}_t - \widetilde{V}(U_x, t),\end{aligned} \tag{7.3.7}$$

where $\overline{\Sigma(\pi^*)^2}$ is given by (A.3.8)*, $\bar{c}_t = \mathbb{E}[c_t^*]$, where c^* is from* (7.3.5) *and $\widetilde{V}$ is Fenchel-Legendre transform of V (with initial condition of power type according to Hypothesis 7.3.3 with $t = 0$).*

Then, if $\kappa < 0$ or $\kappa > 0,\ \beta > \lambda$ there exists a unique (parameterised by κ) MF-Forward CRRA relative performance equilibrium in the sense of Definition 7.3.2. The MF-equilibrium strategy is given by (π^*, c^*) from (7.3.5) *and MF-forward CRRA relative performance utility is given by Q, for U, V satisfying* (7.3.7) *and $\overline{X}_t = \exp\mathbb{E}\big[\log(X_t^{\pi^*,c^*})\big],\ \overline{\Gamma}_t = \exp\mathbb{E}[\log(c_t^*)]$. Finally, utility time dynamics f and g satisfy*

$$f(t) = \begin{cases} \Big(\big(c_t\big)^{\frac{1}{\delta}}\big(\tilde{c}_t\big)^{\theta(\frac{1}{\delta}-1)}\epsilon\Big)^{-\frac{1}{\kappa}}, & \kappa \neq 0, \\ \exp\Big\{\ \big(\theta(\frac{1}{\delta}-1)\mathbb{E}[\lambda] - \eta - \frac{\lambda}{\delta}\big)\,t\Big\}, & \kappa = 0, \end{cases} \quad \text{and} \quad g(t) = f(t)^{1-\kappa}. \tag{7.3.8}$$

If $\psi^\sigma = 1$ or $\kappa > 0, \beta \leqslant \lambda$, then there exists no MF-equilibrium.

Proof. The complete proof can be found in Section A.3 of the supplementary material. ☐

We address the admissibility of c^* in Section 7.4.

We now provide the single stock case result.

Corollary 7.3.5 (Single stock)**.** *Let μ, σ, ν be deterministic with $\nu = 0$ and $\mu, \sigma > 0$. Then λ, β*

are expressed as

$$\lambda = \epsilon^{-\delta}\big(e^{\mathbb{E}[\delta \log \epsilon]}\big)^{\frac{\theta(1-\delta)}{\mathbb{E}[\theta(1-\delta)]-1}}, \tag{7.3.9}$$
$$\beta = \bigg(\frac{\mu^2}{2\sigma^2}\Big(\delta + \frac{\theta}{\theta_{crit}}(1-\delta)\Big) + r\bigg)(1-\delta)\big(1 - \frac{\theta}{\theta_{crit}}\big) + \frac{\theta}{\theta_{crit}}\frac{\mathbb{E}[\rho\delta]}{\mathbb{E}[\delta]}(1-\delta) + \rho\delta,$$

where $\theta_{crit} = \frac{1-\mathbb{E}[\theta(1-\delta)]}{\mathbb{E}[\delta]}$.

If $\mathbb{E}[\theta(1-\delta)] \neq 1$ *then a optimal candidate strategy exists, with the optimal strategy* (π^*, c^*) *given by*

$$\pi^* = \frac{\mu}{\sigma^2}\Big(\delta + \frac{\theta}{\theta_{crit}}(1-\delta)\Big) \text{ and } c^*_t = \begin{cases} \Big(\frac{1}{\beta} + \big(\frac{1}{\lambda} - \frac{1}{\beta}\big)e^{\kappa\beta t}\Big)^{-1}, & \beta \neq 0, \\ \big(-\kappa t + \frac{1}{\lambda}\big)^{-1}, & \beta = 0. \end{cases} \tag{7.3.10}$$

Furthermore, the optimal candidate strategy is a MF-equilibrium, if either $\kappa \leqslant 0$, *or* $\kappa > 0,\ \lambda < \beta$.

Remark 7.3.6 (On convergence of the Nash equilibria of n-player game)**.** *We can see that the MF-equilibrium agrees with the limit of the n-player game equilibrium strategies, as $n \to +\infty$, $\mathbb{P}$-a.s.. The respective parameters and their functions converge to the averages of type distributions by the strong law of large numbers.*

7.4 A discussion of the results

In this section, we focus on the interpretation of the results obtained by way of analyzing the single stock case of the mean field game given as by Corollary 7.3.5 (compare with Corollary 7.2.8). This approach provides already valuable insights on the model, capturing the relations and their interpretation whilst avoiding the complex, intricate dependencies of the general case. This case also allows for easier comparison with the results from [72] who used standard utility maps. For the sake of simplicity, if not specified, we write π, c to refer to (π^*, c^*) the mean field single stock optimal candidate strategy given by (7.3.10). For a full overview, we keep mention of the inadmissible equilibria cases, even though our discussion focuses on the MF-equilibrium strategy. We do not discuss the optimal investment strategy π and omit some details on the optimal consumption strategy c since these overlap with preceding work [37,72,73]. We restrict our attention to the *market-risk relative consumption preference* κ and its interplay with the other parameters in the scope of the optimal consumption c. The case $\kappa = 1$ is just that of [72, Section 4] and we point the reader to the discussion there. Before starting, we emphasize that the behaviours found here, when $\kappa \neq 1$, are not a symmetrised version of those of [72] – this is easily seen in Figure 7.4.3 and Table 7.4.1.

Classical parameters

Recovering classical results. Under no performance concerns and no competition (e.g., $\theta = 0$) we recover the classical results by Merton [78] and Samuelson [85] in (7.3.10): agents invest a constant fraction of wealth in the stock, the consumption strategy is time-dependent, and both investment and consumption strategies are independent – this holds for the classic utility results in [73]. Both here and in [37] the results differ from the mentioned ones as the classical results encapsulate a dependence on the time horizon. When $\kappa = 1$ we recover the results in [72].

The market parameters β, λ *in* (7.3.9). As in [72] and Corollary 7.3.5 we rewrite β in (7.3.9) as

$$\beta = \frac{\mu^2}{2\sigma^2}(1-\delta_{\text{eff}})\delta_{\text{eff}} + r(1-\delta_{\text{eff}}) + \rho\delta'_{\text{eff}},$$

$$\text{where} \quad \delta_{\text{eff}} = (1-\frac{\theta}{\theta_{\text{crit}}})\delta + \frac{\theta}{\theta_{\text{crit}}} \quad \text{and} \quad \delta'_{\text{eff}} = (1-\frac{\theta}{\theta'_{\text{crit}}})\delta + \frac{\theta}{\theta'_{\text{crit}}}$$

$$\text{for} \quad \theta_{\text{crit}} = \frac{1-\mathbb{E}[\theta(1-\delta)]}{\mathbb{E}[\delta]} \quad \text{and} \quad \theta'_{\text{crit}} = \rho\frac{1-\mathbb{E}[\theta(1-\delta)]}{\mathbb{E}[\delta\rho]}.$$

When ρ is deterministic across the population we have $\delta_{\text{eff}} = \delta'_{\text{eff}}$ and $\theta_{\text{crit}} = \theta'_{\text{crit}}$. Furthermore, β reduces to the following expression $\beta = \frac{\mu^2}{2\sigma^2}(1-\delta)\delta + r(1-\delta) + \rho\delta$, which is reminiscent of the classical Merton result. Additionally, setting $r = \rho = 0$, we can rewrite β as

$$\beta = \frac{1}{2}\pi\mu(\delta-1)\Big(\frac{\theta}{\theta_{\text{crit}}} - 1\Big),$$

which shows that β is, in fact, the expected return times some additional risk (on top of the log risk-tolerant investor) and a different competition proportion (on top of θ_{crit}-competitive investor). We hence interpret β as the *expected effective portfolio return*.

When it comes to λ (see (7.3.9)),

$$\lambda = \frac{1}{\epsilon^\delta} \times (H_{\text{population}})^{\theta(1-\delta)} \quad \text{where} \quad H_{\text{population}} = \big(e^{\mathbb{E}[\delta \log \epsilon]}\big)^{\frac{1}{\mathbb{E}[\theta(1-\delta)]-1}},$$

the main relation to notice is its inverse dependence on the *agent's relative perception of wealth relative to consumption* ϵ and scaled with the risk tolerance δ, and then adjusted by the agent's specific weighting (associated to the agent's risk aversion and competition preferences) of the population parameter $H_{\text{population}}$.

Hence, we treat λ as an "effective" ϵ. As originally ϵ captures the linear relation between utility from wealth and utility from consumption in the general FPP process (see Hypothesis 7.3.3), we characterise λ as the *effective relative perception of wealth with respect to consumption*.

The market-risk relative consumption preference κ

Before fully diving into the analysis of κ, we differentiate between two settings and their interpretations. An agent in the absence of competition ($\theta = 0$) or purely looking for a "best response" to the actions undertaken by the other agents (see Remark 7.2.7) can determine its own idiosyncratic κ. At the same time, for the Nash forward equilibrium, all players must interact through a common κ. In either case, we interpret κ as a *market-risk relative consumption preference* observing that when the agent is free to choose κ for herself ($\theta = 0$ case or just best response), the agent looks for the most suitable market investment-consumption environment via the choice of κ.

The concept of Elasticity of intertemporal substitution (EIS)

We start by calculating one of the main parameters in models of dynamic choice of consumption in macroeconomics and finance [88]. The *elasticity of intertemporal substitution* (EIS) measures the agent's willingness to substitute future consumption for present consumption in response to changes in investment opportunities and is given by [5,56]

$$\text{EIS} := -\frac{d\big(\partial_t c_t/c_t\big)}{d\big(\partial_t V_x(c_t z,t)/V_x(c_t z,t)\big)},$$

for some[1] $z > 0$. EIS measures the response of consumption growth to the deviation of *real interest rate*, the latter captured by the denominator of the above expression.

If EIS < 0, the agent is subjected to *income effect*. On the contrary, EIS > 0 is known as *substitution effect* (see introduction). Finally, EIS $= 0$ corresponds to a case of constant consumption rate where the agent is indifferent about the real-time growth and would like to spend the constant fraction of income all over the time.

We introduce the reader to the new concept of *Elasticity of Conformity*.

Definition 7.4.1 (Elasticity of Conformity (EC)). *Let c_t and $\tilde{c}_t$ be the consumption of the generic agent and geometric average* (5.4.4) *respectively. We define* Elasticity of Conformity γ_t *as*

$$\gamma_t := \frac{d\left(\partial_t c_t / c_t\right)}{d\left(\partial_t \tilde{c}_t / \tilde{c}_t\right)}. \tag{7.4.1}$$

The *Elasticity of Conformity (EC)* captures how the change in the agent's consumption is affected by the deviation in the consumption of the typical agent (as compared against the geometric average) across time. The EC's negative/positive sign reflects that the agent's decision aligns with the primary trend of investors, i.e., to increase/decrease consumption in response to the decaying geometric average consumption. Moreover, the choice of the agent to maintain a non-zero level of consumption is captured by their EIS asymptotically going to 0 as $t \to \infty$ (see Figure 7.4.1 a) and c)).

We now calculate the EIS of the generic agent

$$\begin{aligned}
\text{EIS} := -\frac{d\big(\partial_t c_t / c_t\big)}{d\big(\partial_t V_x(c_t z, t)/V_x(c_t z, t)\big)} &= -\frac{d\big(\partial_t c_t / c_t\big)}{d\left(\dfrac{V_{xx}(c_t z, t)\partial_t c_t z + V_x(c_t z, t)\frac{\partial_t g(t)}{g(t)}}{V_x(c_t z, t)}\right)} \\
&= \frac{d\big(\partial_t c_t / c_t\big)}{d\Big(\frac{1}{\delta}\big(\partial_t c_t / c_t\big) - \big(\partial_t g(t)/g(t)\big)\Big)},
\end{aligned}$$

given the CRRA property $V_x/V_{xx} = -\delta x$. Injecting (7.3.8), $g(t) = \Big(\big(c_t\big)^{\frac{1}{\delta}}\big(\tilde{c}_t\big)^{\theta(\frac{1}{\delta}-1)}\epsilon\Big)^{\frac{\kappa-1}{\kappa}}$, yields

$$\frac{\partial_t g(t)}{g(t)} = \frac{\partial_t c_t}{c_t}\frac{\kappa-1}{\kappa}\frac{1}{\delta} + \frac{\partial_t \tilde{c}_t}{\tilde{c}_t}\frac{\kappa-1}{\kappa}\theta\left(\frac{1}{\delta}-1\right) = \frac{\partial_t c_t}{c_t}\frac{\kappa-1}{\kappa}\frac{1}{\delta}\left(1+\theta\big(1-\delta\big)\frac{1}{G_t}\right),$$

where $G_t := \dfrac{\partial_t c_t / c_t}{\partial_t \tilde{c}_t / \tilde{c}_t}$. Returning to the EIS expression from before, we find

$$\text{EIS} = \frac{d\big(\partial_t c_t / c_t\big)}{d\Big(\frac{1}{\delta}\big(\partial_t c_t / c_t\big) - \frac{\kappa-1}{\kappa}\frac{1}{\delta}\Big(1+\theta\big(1-\delta\big)\frac{1}{G_t}\Big)\big(\partial_t c_t / c_t\big)\Big)} = \frac{\delta\kappa}{1-\theta(1-\kappa)(\delta-1)\frac{1}{\gamma_t}}.$$

In contrast with the classic utility optimisation approach where one has $\text{EIS}^{\text{classic}} = \delta$ the FPP approach allows an intrinsic time-dependence of V and thus of the EIS. Namely

$$\text{EIS}_t^{(\theta\neq 0)} = \frac{\text{EIS}^{(\theta=0)}}{1-\theta(1-\kappa)(\delta-1)\frac{1}{\gamma_t}}, \quad \text{where} \quad \text{EIS}^{(\theta=0)} = \kappa\delta = \kappa \times \text{EIS}^{\text{classic}}, \tag{7.4.2}$$

with γ_t standing for the time-dependent *Elasticity of Conformity (EC)*.

The FPP here allows one to disentangle risk tolerance δ and Elasticity of intertemporal substitution (EIS) – standard utility theory does not allow this. This same feature has been reported in [12, Slide 22] through the use of recursive utilities (in a more restricted Merton market framework than ours but within MFGs) and which is close to our result (7.4.2). In

[1] The formula presented has stated the agent's rate of consumption equal to $c_t z$, where c_t is a rate of consumption for unit of wealth and $z > 0$ some fixed amount of wealth. On the one hand, fixing the amount of wealth helps us to purely investigate the relative effect of consumption, on the other, it allows to assess its time differential to calculate EIS, in contrast with $c_t X_t$ having an Itô-differential form and thus not computable.

contrast with standard utility theory, the FPP setting captures the different dimensions of risk, represented by κ, that purely comes from the consumption environment and further scales the EIS. When $\kappa = 1$, the case of standard CRRA utility framework and analysed in [72], immediately yields $\text{EIS} = \delta = \text{EIS}^{\text{classic}}$.

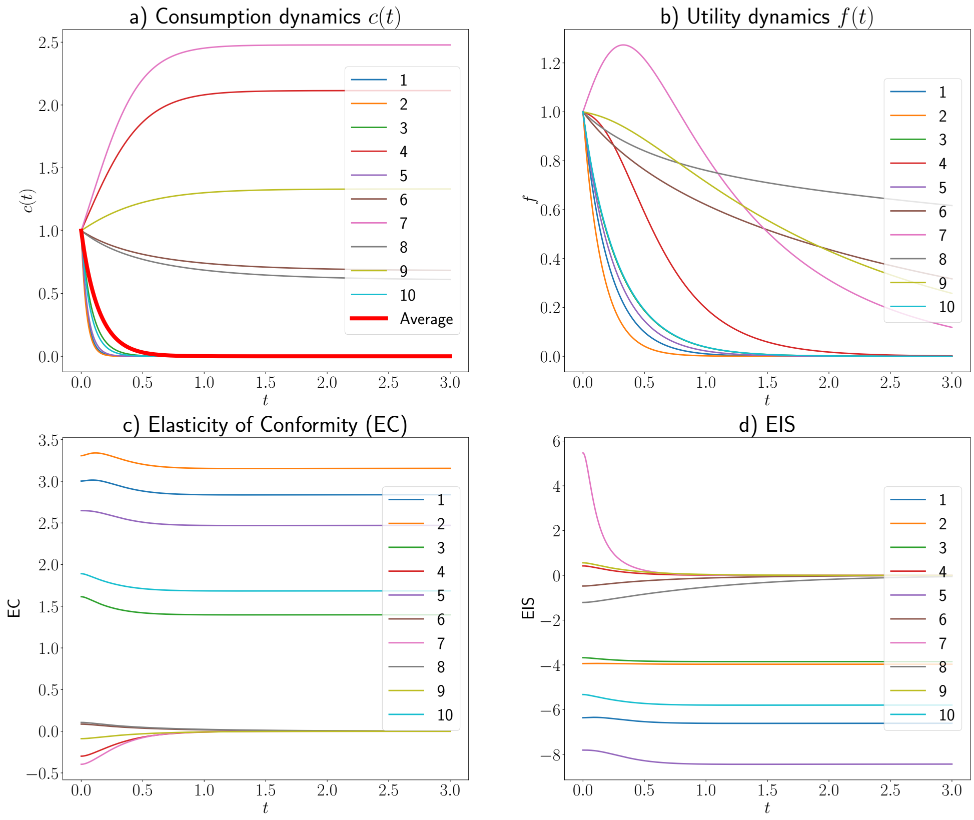

Figure 7.4.1: Simulated agents' types as a function of time: a) Optimal consumption (7.3.10); b) Utility dynamics f (7.3.8); c) Elasticity of Conformity (EC) (7.4.1); d) EIS (7.4.2). All the parameters are independent as a components of joint type's distribution and simulated as follows: $\delta/2 \sim Beta(1.5, 1.5)$, $\theta \sim Beta(3, 5)$, $\rho \sim Beta(1, 19)$, $r = 0.05$, $\lambda = 1$, $\mu = 5$, $\sigma = 1$, $\kappa = -2$, $\theta_{\text{crit}} = 0.375$, where $Beta(\cdot, \cdot)$ refers to the Beta distribution.

Figure 7.4.1 displays simulations of consumption, utility, EC and EIS for a collection of agents with a certain fixed distribution of the type vector. Looking into Figure 7.4.1a) it is worth pointing out that, within our simulations, the average (geometric) consumption asymptotically decays to 0 due to the behaviour of the majority of the agents being under the income effect. As β is bounded from above but not from below, and given the symmetric distribution of type component for δ and θ around critical values of 1 and θ_{crit} respectively, the simulated β is much more likely to be negative and hence driving the consumption to 0 on the long run. Agents that do not reduce their consumption to a zero constant level over time are indifferent to the main herd of investors, and one sees their Elasticity of Conformity (Figure 7.4.1c)) converging to zero. Moreover, their consumption behaviour is becoming perfectly inelastic in the long run ($\text{EIS} \to 0$ as $t \to \infty$ in Figure 7.4.1d)).

In contrast, agents that choose to gradually cut their consumption to 0 (following the market trend) have a constant positive EC and a constant negative EIS in the long term.

The market-risk relative consumption preference κ under no performance concerns: $\theta = 0$ (no competition)

Choosing different EIS and Risk Tolerance. In the absence of competition, i.e., when $\theta = 0$, (7.4.2) already yields a non-standard $\text{EIS}^{(\theta=0)} = \kappa\delta$, whereas standard utility theory with CRRA type utilities yields $\text{EIS}^{\text{classic}} = \delta$. Without competition, as we no longer require κ to be uniform and deterministic across the whole population, the agent can choose in which consumption environment he is willing to invest, hence, as opposed to [72], the agent can determine her own risk tolerance δ and $\text{EIS}^{(\theta=0)}$ independently.

For $\kappa > 0$ then $\text{EIS}^{(\theta=0)} > 0$ and hence the agent is subjected to *income* effect, otherwise, when $\kappa < 0$ then $\text{EIS}^{(\theta=0)} < 0$ and the agent allows for a *substitution* effect. Finally, $\kappa = 0$ reflects a constant consumption rate as $\text{EIS}^{(\theta=0)} = 0$. This last scenario is particularly important as empirical research finds EIS in general to be around 0 for the average household (see [56], [17], [31]). In our framework, $\text{EIS} = 0$ may simply mean $\kappa = 0$ instead of a low risk tolerance $\delta = 0$.

We do not present any formal proof but argue that κ cannot be made to disappear from the problem at the expense of redefining the remaining model parameters (namely, $r, \rho, \nu, \epsilon, \beta, \lambda$) and then referring to a variant of the model (with the redefined parameters). In effect, κ represents a new degree of freedom in modelling. Nonetheless, a scaling effect between κ, β, λ at the level of consumption strategy is present. Concretely, given β and λ, define $\hat{\beta} := \kappa\beta$, $\hat{\lambda} := \kappa\lambda$, and $c_t^{(\beta,\lambda,\kappa)} := c_t^*$ with c_t^* the expression given in (7.3.10) for $\kappa \neq 0$ (and β, λ). Then $c^{(\beta,\lambda,\kappa)}$ can be rewritten as follows

$$c_t^{(\beta,\lambda,\kappa)} = \kappa^{-1} c_t^{(\hat{\beta},\hat{\lambda},\kappa=1)}.$$

It is important to remark that for $\kappa = 1$ the consumption $c^{(\hat{\beta},\hat{\lambda},\kappa=1)}$ is that of [72]. By construction $\lambda > 0$ but $\hat{\lambda} \in \mathbb{R}$, on the other hand, for $\kappa \geqslant 0$ the parameters $\hat{\beta}$, $\hat{\lambda}$ preserve the original sign of β and λ.

The market-risk relative consumption preference κ under performance concerns: $\theta \neq 0$

The agents are now synchronised on their consumption preference as they compete (κ is the same for all agents, see Remark 7.2.7). From (7.4.2) the agent sees her consumption dynamics being re-scaled by competition ($\theta \neq 0$) and the market environment ($\kappa \neq 0$). The agents can access their EIS, (7.4.2), by adjusting the risk-competition preferences (the EC depends on β and λ as implied by the agents' parameters).

The "market-risk relative consumption preference" parameter κ. This parameter appears in Hypothesis 7.2.2 and 7.3.3 to weigh the utility of wealth versus consumption through $g(t) = f(t)^{1-\kappa}$. It reflects the weight in time the environment assigns to the consumption process compared to the accumulation of wealth. In parallel with the risk tolerance δ of the standard utility $u_0(x) = \frac{1}{1-\frac{1}{\delta}} x^{1-\frac{1}{\delta}}$, we interpret κ as a parameter of market risk, reflecting how the market is encouraging its agents to prefer consumption over wealth and vice versa. Recalling for simplicity the no-competition EIS, $\text{EIS}^{(\theta=0)} = \kappa\delta$, we can see that κ appears as another dimension of risk, further adjusting each agent's consumption response to the deviations in the real interest rate.

Figure 7.4.2 displays the effect of κ on consumption and utility dynamics for the generic agent. When $\kappa > 0$, the agent's utility is more sensitive to deviations of wealth than to consumption (as $g(t) = f(t)^{1-\kappa}$) and two cases need to be distinguished: $\kappa \geqslant 1$ and $0 < \kappa < 1$. The market environment for large κ makes the agents' utility from wealth and consumption inversely related. In this scenario, the percent increase in utility from wealth will correspond to a decrease in utility from wealth and vice versa, making the agent *consumption-to-wealth averse*. On the other hand, when $0 < \kappa < 1$ the environment pushes the agent to derive a moderate utility from consumption relative to the utility from wealth. From a percent increase or decrease in utility from consumption, the agent will obtain a much larger respective change in her utility from wealth.

If $\kappa < 0$ the general trend is reversed from that of $0 < \kappa < 1$. The changes in consumption utility have a larger weight relative to those of the wealth utility dynamics. We refer to all subcases of $\kappa < 1$ as *consumption-to-wealth tolerant*, indicating that the agents' utilities from the accumulation of wealth and consumption are aligned. Still, we emphasise that the quality of their mutual relation is subjected to the magnitude of κ. Finally, $\kappa = 0$ prescribes $f = g$, and the environment tells the agent to prefer wealth and consumption equally. Here, $c_t = \lambda$ for all

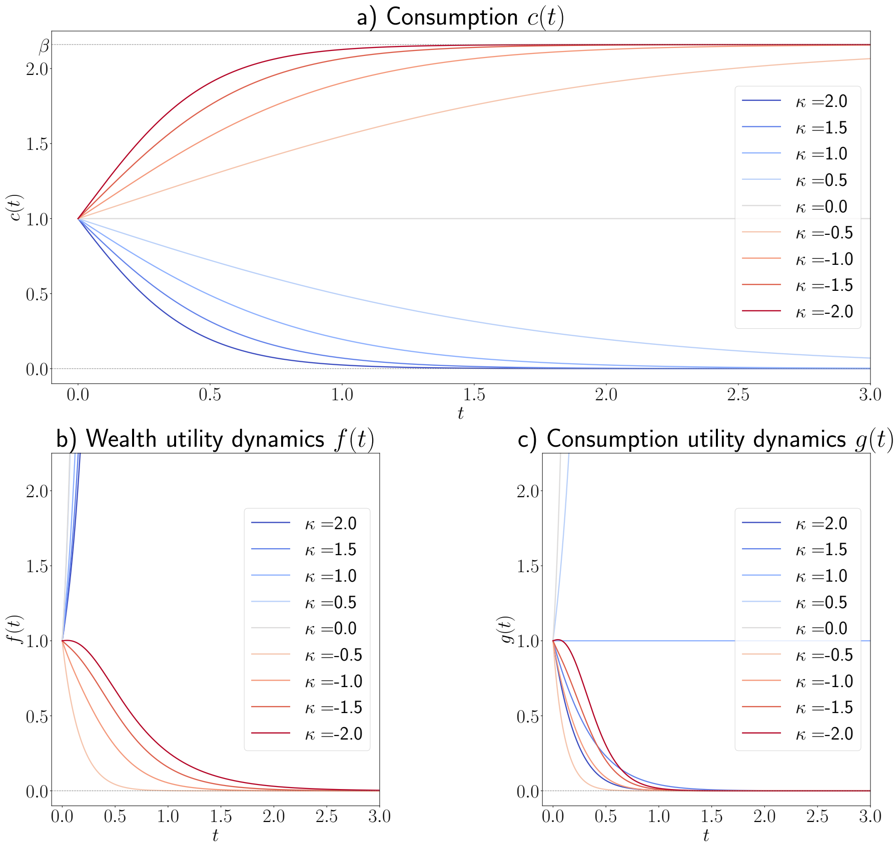

Figure 7.4.2: Plots of different elements from a single agent's optimization problem for κ from -2 to 2: a) Consumption dynamics $c(t)$ with lower asymptote at Zero and upper asymptote at β; b) Wealth utility dynamics $f(t)$; c) Consumption utility dynamics $g(t)$. The parameters are $\delta = 0.5,\ \theta = 0.6,\ \rho = 0.04,\ r = 0.05,\ \beta = 2.16,\ \lambda = 1,\ \mu = 5,\ \sigma = 1$.

$t \geqslant 0$ (see (7.3.10)), in other words, the agent has no dynamic preference to deviate from the effective relative perception of wealth with respect to consumption λ. Looking back at (7.3.10), one can rewrite c as

$$c_t = \left(\frac{1}{\beta} + \left(\frac{1}{\lambda} - \frac{1}{\beta}\right)e^{\kappa\beta t}\right)^{-1} = \left(\frac{1}{\beta}\left(1 - e^{\kappa\beta t}\right) + \frac{1}{\lambda}e^{\kappa\beta t}\right)^{-1}.$$

Roughly, with $\kappa = 0$, the market does not have the agent competing their λ against their expected effective portfolio return β.

Figure 7.4.3 displays the effect of κ on the consumption map for different combinations of λ, β. When $\kappa \in (-1, 1)$ the steady-state of consumption (the asymptote) is reached much later.

In this model, a unique macro-economics cause-effect phenomenon of κ is noteworthy. Assume a central planner that can control κ then, given market parameters, by manipulating κ the central planner can encourage or discourage consumption across the whole market. This is a well-known (Keynesian economics) policy for *economic stimulus* in recession times.

We briefly discuss the effect of the discounting factor ρ and the riskless rate r. By looking into (7.3.9), one can notice the additive effect of r and ρ on the expected effective portfolio return. The sign of β affects the direction and rate of the agent's consumption (7.3.10). Thus, by choosing the personal discount factor appropriately the agent can adjust for β, making it larger/smaller, as to exceed the important thresholds of λ and 0 (that take place in c^* (7.3.10) – see Table 7.4.1).

The finer interplay of κ with β and λ on the consumption policy c. We refer to [72, Section 4] comprising the analysis for $\rho = r = 0$ and $\kappa = 1$. This reduction, upon inspection, extends to our case when $\kappa > 0$, thus we focus on the case where $\kappa < 0$, i.e., the market encourages the

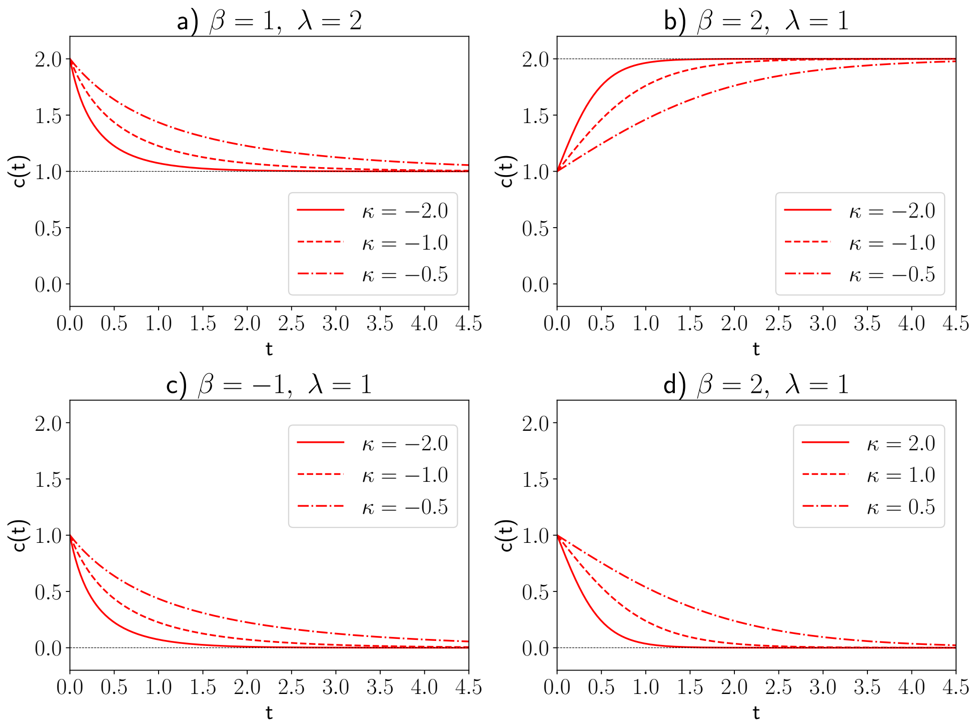

Figure 7.4.3: Plots of optimal c of (7.3.10) for various combinations of parameters with a highlight on the monotonic behaviour of $\kappa \mapsto (c_t)_{t\geqslant 0}$ by way in which the asymptote is reached.

substitution effect.

In terms of market behaviour, there are 4 cases to be distinguished: $\beta > \lambda$, $0 \leqslant \beta < \lambda$, $\beta < 0 < \lambda$ and $\beta = \lambda$ and we summarise them in Table 7.4.1 (by monotonic behaviour, associated asymptotes, and admissibility). If $\beta = \lambda$ then κ plays no role in c, but it does affect the utility map. We emphasise that the expected rate of return of wealth, $\partial_t \mathbb{E}[\log X_t|\mathcal{F}_0]$, follows the opposite direction of the monotonicity of c.

First, c_t is increasing in time, when $\beta > \lambda$. Indeed, when the relative perception of wealth is big, i.e., if $\epsilon \nearrow \infty$ then $\lambda \searrow 0^+$, the agent's consumption does not vanish and monotonically increases to β; this is in stark contrast with the case $\beta > \lambda$ under $\kappa > 0$, where $c_t \searrow 0^+$ monotonically as $t \to \infty$ (compare Figure 7.4.3 b) and d). From the beginning, the agent has a focus on wealth thus intends to consume less in the short-term (Figure 7.4.3 b). However, in the long run, the agent follows the market preference towards consumption (see Figure 7.4.2, b) and c) for negative κ's), and raises her proportion of consumption up to some asymptotically stable level ($= \beta$). As the agent's wealth increases, she cannot resist the environment's push to spend more despite her preference for wealth. This particular situation mimics a “keeping up with the Joneses” [51] behaviour given a population of ultra-rich agents: the effective utility from getting richer comes from increasing the level of consumption and is accelerated by competition and the intensive consumption environment.

The trend is similar to when $\beta < \lambda$, which is monotonically decreasing. Indeed, when $\epsilon \searrow 0^+$, $\lambda \nearrow +\infty$, the agent prefers consumption over wealth (while the market pushes agents to consume $\kappa < 0$), so the agent consumes at a higher rate from the onset. However, we distinguish two sub-cases here. The agent with an expected effective portfolio return $\beta \geqslant 0$, who seeks to stabilise her consumption and thus decrease it asymptotically up to the safe rate β (Figure 7.4.3 a). When the agent chooses $\beta < 0$, she monotonically reduces her consumption rate to 0, even if she had started from a higher level. That is the case when the agent is very risk-tolerant ($\delta > 1$) or very competitive ($\theta > \theta_{\text{crit}}$) or both, so the agent accumulates wealth even if she is deriving more utility-value from consumption over time (Figure 7.4.3 c).

Figure 7.4.3 d) features the *income* effect. The agent is pushed to asymptotically decrease her consumption to 0, due to the environment promoting utility from wealth over utility from consumption – see the positive κ cases of Figure 7.4.2 b) and c).

Finally, when $\kappa = 0$ or $\beta = \lambda$ the consumption c_t is constant equal to λ. In the former scenario, $\kappa = 0$, the dynamics of the agent's preferences towards consumption and wealth is identical (captured by $f(t) = g(t)$ and thus EIS $= 0$). Hence, the agent's optimal consumption policy is the constant effective relative perception of wealth λ. Nonetheless, the agent's utility is affected by κ. All cases are neatly summarised in Table 7.4.1. See also Figure C.1 in the supplementary material in [35, Appendix C] (see arXiv:2012.01235) for the regions of monotonicity for c_t as function of κ, δ, θ.

Table 7.4.1: Behaviour of $(c_t)_{t\geqslant 0}$ of (7.3.10) as a map of its parameters. Here $\searrow, \nearrow$ denote convergence from above/below respectively (as $t \to \infty$). Apparent is the asymmetry of cases for $\kappa < 0$ against $\kappa > 0$.

Parameters			Consumption c of (7.3.10)		
κ	β	$\lambda - \beta$	optim. candidate	admissibility	c_t as $t \to \infty$
< 0	> 0	> 0	✔	✔	$\searrow \beta$
		< 0			$\nearrow \beta$
		$= 0$			constant throughout
	$= 0$	> 0			$\searrow 0$
	< 0	> 0			$\searrow 0$
> 0	> 0	> 0	✔	✘	finite-time blow-up
		< 0	✔	✔	$\searrow 0$
		$= 0$	✔	✔	constant throughout
	$= 0$	> 0	✔	✘	finite-time blow-up
	< 0	> 0	✔	✔	finite-time blow-up
$= 0$	any	any	✔	✔	constant throughout

The environment's influence on the agents

One may wonder about the environment's impact on the agent in the particular case of EIS being close to zero. When the environment hardly distinguishes the long-run preference for consumption from the one for wealth ($\kappa \to 0$), the agent who wants to faster reach a constant level of consumption needs to adapt by changing her risk and competition parameters in order to increase the absolute value of β. For example, assume further no-discounting and zero risk-free interest rate, e.g., $\rho = r = 0$, a highly risk-tolerant market with $\mathbb{E}[\delta] \to \infty$ and mutually independent types θ and δ, hence from (7.3.9) we have $\theta_{\text{crit}} = \mathbb{E}[\theta]$. Then, the agent changes in the following ways: the agent avoids the competition but accepts more risk ($\theta \to 0,\ \delta > 1$); or competes at the average market level and accepts even more risk ($\theta \to \mathbb{E}[\theta],\ \delta \gg 1$); or accepts being highly competitive, but risk-averse ($\theta > \mathbb{E}[\theta],\ \delta < 1$).

Lastly, we note that β, as a quadratic function of δ_{eff}, is bounded from above by $\frac{\mu^2}{8\sigma^2}$ but is not bounded from below. That means that under a mild *substitution* regime ($\kappa \to 0^-$), the rational agent cannot outperform the upper-bound effective rate of portfolio return β (shown in Fig.7.4.2a and Fig.7.4.3b). Nevertheless, by adjusting the risk-competition parameters (δ, θ), the agent can decrease her consumption to zero as fast as he wants to. Notably, the choice of ρ helps the agent adjust the magnitude of β and hence the consumption preferences.

At this point, the market highly influences the choices of the player. Suppose the agent wants to maintain a desired level of consumption. In that case, she can do so by re-setting her risk-competition preferences (and in accordance to remain optimal "as a martingale").

7.5 Open questions

From the construction we provided, here and in the much simpler case [36], it is still open how to tackle the full MFG generalisation to market models featuring random coefficients. Here the mean field aggregation approach created in [73] is not possible anymore and a new tool is needed. We refer to Remark 7.2.7 regarding the open issue of allowing agent to choose different κ parameters.

Open is the so-called mean-field aggregation problem where different agents use utility maps from different families, e.g. CRRA and CARA: [42] would be a starting point for the finite-player case while the mean-field case would requires the multi-class approach of [8, Section 8] with the parameterisation technique of from our Section 6.2. Many other questions can be posed in this context of mean-field forward utilities, ranging from possible non-solvability [48], to risk-sharing [11], ergodic problems [50] and associated numerics [54].

Generalising the dynamics of the forward performance utility map (7.1.1) or (7.3.4) outside Assumption 7.3.1 to a fully Itô-dynamics and stochastic strategies is also open. A crucial tool for such would be a general Itô-Wentzell-Lions chain rule as developed in [38] and would likely build along [43] or [82]; or alternatively, an approach similar to [28] can be taken where the volatility of the FPP is exogenously postulated as the agent's preferences (Forward-Backward SDEs is the tool of choice there). It is also open exploring game competition in the setting of [28] either for the finite-player game or the MFG as in [3,36,52]. As pointed in [36, Section 5], many other questions can be posed in this context of mean field FPP, ranging from possible non-solvability [48], to risk-sharing pricing [11], indifference-pricing [27, 74], ergodic problems [50]. The cases and analysis of [3] can be extrapolated to the MFG case as well.

A particular case of interest playing to the strengths of the forward performance process framework, and which we do not explore here, is the *dynamic model selection* problem. FPPs allow for a dynamic update of market parameters to accommodate a switch in the market environment or the agent's perception of risk which the standard utility theory does not allow. In fact, [36, Section 4.4] carries over to the construction we have provided here – such would allow the agent to update the market-risk relative consumption preference κ. An alternative view to this problem exploiting FPPs and time dependent risk parameters can be found in [86].

Lastly, a comparative study between the capabilities of forward utility maps versus Epstein-Zin preferences within the Kreps-Porteus recursive utilities is to the best of our knowledge an open question. As shown, the forward performance criteria features the crucial property of the recursive utility and hence this work stands to span new economics studies on EIS [1,31,88] .

Appendix A

Proofs of Part III

A.1 Supplementary calculations

Proof of Proposition 6.1.2. We recall the optimal strategy is given by (6.1.4), where we define

$$\widehat{\sigma} := \overline{(\pi\sigma)}_t^{(-i)}, \;\; B_t^\nu := \theta_i^2 \frac{1}{(n-1)^2}\sum_{k\neq i}(\pi_t^k \nu_k)^2, \;\; M_t^\mu := \theta_i \overline{(\pi\mu)}_t^{(-i)} = \frac{\theta_i}{n-1}\sum_{k\neq i}\pi_t^k \mu_k.$$

The drift of (6.1.3) becomes (we omit the argument in U_t, U_x, U_{xx} and use $\widehat{\sigma} := \overline{(\pi\sigma)}_t^{(-i)}$)

$$\begin{aligned}
&U_t^i + U_x^i\big(\pi_t^i\mu_i - M_t^\mu\big) + \frac{1}{2}U_{xx}^i\Big[(\pi_t^i\nu_i)^2 + B_t^\nu + \big(\pi_t^i\sigma_i - \theta_i\overline{(\pi\sigma)}_t^{(-i)}\big)^2\Big] \\
=&\Big(U_t^i - M_t^\mu U_x^i + \frac{1}{2}U_{xx}^i B_t^\nu\Big) + \frac{1}{2}U_{xx}^i\Big[\big(\theta_i\widehat{\sigma}\big)^2 - (\pi_t^i)^2(\nu_i^2+\sigma_i^2)\Big] \\
=&U_t^i + U_x^i\Big[\theta_i\sigma_i\widehat{\sigma}\mu_i\frac{1}{\nu_i^2+\sigma_i^2} - M_t^\mu\Big] - \frac{\mu_i^2}{2}\frac{1}{\nu_i^2+\sigma_i^2}\frac{(U_x^i)^2}{U_{xx}^i} \\
&+\frac{1}{2}U_{xx}^i\Big\{B_t^\nu + \big(\theta_i\widehat{\sigma}\big)^2 - \frac{1}{\nu_i^2+\sigma_i^2}\big(\theta_i\sigma_i\widehat{\sigma}\big)^2\Big\} \\
=&U_t^i + U_x^i\Big[\frac{\mu_i\theta_i\sigma_i\widehat{\sigma}}{\nu_i^2+\sigma_i^2} - \theta_i\overline{(\pi\mu)}_t^{(-i)}\Big] - \frac{\mu_i^2}{2(\nu_i^2+\sigma_i^2)}\frac{(U_x^i)^2}{U_{xx}^i} \\
&+\frac{1}{2}U_{xx}^i\Big\{\theta_i^2\frac{1}{(n-1)^2}\sum_{k\neq i}(\pi_t^k\nu_k)^2 + \big(\theta_i\widehat{\sigma}\big)^2\Big[1 - \frac{\sigma_i^2}{\nu_i^2+\sigma_i^2}\Big]\Big\}.
\end{aligned}$$

Equation (6.1.2) now follows as U_t^i needs to be chosen such that the equation is zero. We inject in the drift of (6.1.3) the expression (6.1.2) and obtain a simplified version

$$\begin{aligned}
&-\Big\{U_x^i\Big[\theta_i\sigma_i\widehat{\sigma}\mu_i\frac{1}{\nu_i^2+\sigma_i^2} - M_t^\mu\Big] - \frac{\mu_i^2}{2}\frac{1}{\nu_i^2+\sigma_i^2}\frac{(U_x^i)^2}{U_{xx}^i} + \frac{1}{2}U_{xx}^i\Big\{B_t^\nu + \big(\theta_i\widehat{\sigma}\big)^2 \\
&-\frac{1}{\nu_i^2+\sigma_i^2}\big(\theta_i\sigma_i\widehat{\sigma}\big)^2\Big\}\Big\} + U_x^i\big(\pi_t^i\mu_i - M_t^\mu\big) \\
&+\frac{1}{2}U_{xx}^i\Big[(\pi_t^i\nu_i)^2 + B_t^\nu + \big(\pi_t^i\sigma_i\big)^2 - 2\pi_t^i\sigma_i\theta_i\widehat{\sigma} + \big(\theta_i\widehat{\sigma}\big)^2\Big] \\
=&\frac{U_{xx}^i}{2}\frac{1}{\nu_i^2+\sigma_i^2}\Big(\big((\pi_t^i)^2(\nu_i^2+\sigma_i^2)^2 - 2\big(\pi_t^i(\nu_i^2+\sigma_i^2)\big)\Big(\sigma_i\theta_i\widehat{\sigma} - \mu_i\frac{U_x^i}{U_{xx}^i}\Big)\Big) \\
&+-\frac{1}{\nu_i^2+\sigma_i^2}\frac{U_{xx}^i}{2}\frac{2}{U_{xx}^i}\Big\{U_x^i\Big[\theta_i\sigma_i\widehat{\sigma}\mu_i\Big] - \frac{\mu_i^2}{2}\frac{(U_x^i)^2}{U_{xx}^i} + \frac{1}{2}U_{xx}^i\Big\{-\big(\theta_i\sigma_i\widehat{\sigma}\big)^2\Big\}\Big\} \\
=&\frac{U_{xx}^i}{2}\frac{1}{\nu_i^2+\sigma_i^2}\Big|\pi_t^i(\nu_i^2+\sigma_i^2) - \Big(\sigma_i\theta_i\widehat{\sigma} - \mu_i\frac{U_x^i}{U_{xx}^i}\Big)\Big|^2,
\end{aligned}$$

which results in (6.1.5). $\square$

Proof of Equation (6.2.7). We take up the drift of (6.2.7) and we have just by re-organizing the terms

$$\begin{aligned} 0 &= U_t(Z_t^\pi,t) + U_x(Z_t^\pi,t)\big(\mu\pi_t - \theta\overline{\mu\pi_t}\big) + \frac{1}{2}U_{xx}(Z_t^\pi,t)\Big((\nu\pi_t)^2 + \big(\sigma\pi_t - \theta\overline{\sigma\pi_t}\big)^2\Big) \\ &= \Big(U_t - U_x\theta\overline{\mu\pi_t} + \frac{1}{2}U_{xx}\theta^2(\overline{\sigma\pi_t})^2\Big) \\ &\qquad + \frac{1}{2}\frac{U_{xx}}{(\nu^2+\sigma^2)}\Big(\pi_t^2(\nu^2+\sigma^2)^2 - 2\pi_t(\nu^2+\sigma^2)\Big\{\theta\sigma\overline{\sigma\pi_t} - \mu\frac{U_x}{U_{xx}}\Big\}\Big) \end{aligned}$$

We recall the optimal strategy given by (6.2.8), where we complete the square inside the U_{xx} term in the SPDE above we have

$$\begin{aligned} 0 = \Bigg\{ & U_t + U_x \cdot \Big(\mu\frac{\theta\sigma\overline{\sigma\pi_t}}{(\nu^2+\sigma^2)} - \theta\overline{\mu\pi_t}\Big) + \frac{1}{2}U_{xx}\cdot\theta^2(\overline{\sigma\pi_t})^2\Big(1 - \frac{\sigma^2}{\nu^2+\sigma^2}\Big) \\ & - \frac{1}{2}\frac{\mu^2}{(\nu^2+\sigma^2)}\frac{(U_x)^2}{U_{xx}}\Bigg\} + \frac{1}{2}\frac{U_{xx}}{(\nu^2+\sigma^2)}\Big|\pi_t(\nu^2+\sigma^2) - \Big(\theta\sigma\overline{\sigma\pi_t} - \mu\frac{U_x}{U_{xx}}\Big)\Big|^2 \end{aligned}$$

Under the CARA condition $U_x/U_{xx} = -\delta$ and the choice of the optimal strategy, the remaining drift must zero-out. We then have

$$U_t = -\frac{U_x}{2(\nu^2+\sigma^2)}\cdot\Big(\mu\theta\sigma\overline{\sigma\pi_t} + \delta\mu^2\Big) + U_{xx}\frac{(\theta\,\sigma\,\overline{\sigma\pi_t})^2}{2(\nu^2+\sigma^2)} - \frac{1}{2}U_{xx}\cdot(\theta\,\overline{\sigma\pi_t})^2 + U_x\Big(\theta\overline{\mu\pi_t}\Big).$$

□

A.2 Proof of Theorem 7.2.6

We prove Theorem 7.2.6 in full detail.

Proof. Step 1. Finding the investment strategy.

First, we deal with the investment policy. Injecting the condition $U_x/U_{xx} = -\delta_i x$ in (7.1.6) leads to the system ($i \in \{1,\dots,n\}$)

$$\pi_t^{i,*} = \frac{1}{\nu_i^2 + \sigma_i^2\big(1 + \frac{\theta_i(1-\delta_i)}{n-1}\big)}\Big(\theta_i\sigma_i(1-\delta_i)\frac{n}{n-1}\overline{\sigma\pi_t} + \mu_i\delta_i\Big), \quad \text{where} \quad \overline{\sigma\pi_t} = \frac{1}{n}\sum_{k=1}^{n}\sigma_k\pi_t^{k,*}.$$

The last identity expresses $\pi^{i,*}$ as a function of the unknown $\overline{\sigma\pi_t}$. To determine it we multiply both sides of the $\pi^{i,*}$ expression by σ_i and average over $i \in \{1,\dots,n\}$. This yields the solvability condition

$$\overline{\sigma\pi_t} = \overline{\sigma\pi_t}\psi_n^\sigma + \varphi_n^\sigma \quad \Leftrightarrow \quad \overline{\sigma\pi} = \frac{\varphi_n^\sigma}{1-\psi_n^\sigma} \quad \text{as long as} \quad \psi_n^\sigma \neq 1. \tag{A.2.1}$$

Plugging the expression $\overline{\sigma\pi}$ into that for $\pi^{i,*}$ yields the result. That the optimal strategies are constant is now obvious and one finds that the corresponding η (see (7.2.5)) is time-independent with $\eta_t^i = \eta_i$.

If $\psi_n^\sigma = 1$, then there exists no Nash equilibrium.

Step 2. Finding the explicit form of η_i.

Just like for $\overline{\sigma\pi_t}$, we obtain an expression for $\overline{\mu\pi_t} = \frac{1}{n}\sum_{i=1}^{k}\mu_k\pi_t^{k,*}$ by multiplying $\pi^{i,*}$ by μ_i on both sides and averaging over i. We have

$$\overline{\mu\pi_t} = \frac{n}{n-1}\frac{\varphi_n^\sigma}{1-\psi_n^\sigma}\psi_n^\mu + \phi_n^\mu, \qquad \overline{\mu\pi_t}^{(-i)} = \frac{n}{n-1}\overline{\mu\pi_t} - \frac{1}{n-1}\mu_i\pi_t^i, \tag{A.2.2}$$

where φ_n^μ, ψ_n^μ are defined as

$$\varphi_n^\mu = \frac{1}{n}\sum_{k=1}^n \delta_k \frac{\mu_k^2}{\nu_k^2+\sigma_k^2(1+\frac{\theta_k(1-\delta_k)}{n-1})}, \qquad \psi_n^\mu = \frac{1}{n-1}\sum_{k=1}^n \theta_k(1-\delta_k)\frac{\mu_k\sigma_k}{\nu_k^2+\sigma_k^2(1+\frac{\theta_k(1-\delta_k)}{n-1})}.$$

Similarly, defining

$$\overline{(\nu\pi_t)^2} = \frac{1}{n}\sum_{k=1}^n (\nu_k\pi_t^k)^2 \quad \Rightarrow \quad \overline{(\nu\pi_t)^2} = \frac{1}{n}\sum_{k=1}^n \Big(\frac{\nu_k\theta_k\sigma_k\cdot\frac{n}{n-1}\cdot\frac{\varphi_n^\sigma}{1-\psi_n^\sigma}+\nu_k\mu_k\delta_k}{\nu_k^2+\sigma_k^2\big(1+\frac{\theta_k(1-\delta_k)}{n-1}\big)}\Big)^2, \tag{A.2.3}$$

together with $\overline{(\nu\pi_t)^2}^{(-i)} = \frac{n}{n-1}\overline{(\nu\pi_t)^2} - \frac{1}{n-1}(\nu_i\pi_t^i)^2$. Finally $\overline{\Sigma\pi_t^2} = \overline{(\nu\pi_t)^2} + \overline{(\sigma\pi_t)^2}$, and like for (A.2.3) we have

$$\overline{(\sigma\pi_t)^2} = \frac{1}{n}\sum_{k=1}^n \Big(\frac{\sigma_k\theta_k\sigma_k\frac{n}{n-1}\frac{\varphi_n^\sigma}{1-\psi_n^\sigma}+\sigma_k\mu_k\delta_k}{\nu_k^2+\sigma_k^2\big(1+\frac{(1-\delta_k)\theta_k}{n-1}\big)}\Big)^2 \tag{A.2.4}$$

$$\text{with } \widehat{\Sigma\pi_t^2} = \frac{n}{n-1}\overline{\Sigma\pi_t^2} - \frac{1}{n-1}(\nu_i^2+\sigma_i^2)(\pi_t^i)^2.$$

Replacing these expressions in that of the constant $\eta_t^i = \eta_i$ in Equation (7.2.4) of Section 7.2 one obtains the sought expression (in the Theorem's statement).

Step 3. Finding the consumption strategy and the relative performance utilities.

The system of Equations (7.2.3) under Hypothesis 7.2.5, i.e., $g_i(t) = f_i(t)^{1-\kappa}$, becomes

$$\begin{cases} c_t^i = \epsilon_i^{-\delta_i}\big(\tilde{c}_t^{(-i)}\big)^{\theta_i(1-\delta_i)} f_i(t)^{-\kappa\delta_i}, \\ f_i'(t) + \Big(\eta_i + \theta_i\big(1-\frac{1}{\delta_i}\big)\bar{c}_t^{(-i)}\Big) f_i(t) + \frac{\epsilon_i^{-\delta_i}}{\delta_i}\big(\tilde{c}_t^{(-i)}\big)^{\theta_i(1-\delta_i)} f_i(t)^{1-\kappa\delta_i} = 0, \end{cases} \tag{A.2.5}$$

where η is given by (7.2.5). We mainly follow the machinery of the proof [72, Theorem 2.2] to obtain the closed form solution. We repeat these arguments and emphasise the notable discrepancies. Substituting the LHS of the first equation of (A.2.5) into the second, we obtain the linear ODE which solves as

$$f_i(t) = \exp\left(-\int_0^t \left(\eta_i + \theta_i\left(1-\frac{1}{\delta_i}\right)\bar{c}_s^{(-i)} + \frac{1}{\delta_i}c_s^i\right)\mathrm{d}s\right).$$

Now plugging it back into the first equation of (A.2.5) we obtain

$$c_t^i \exp\Big(-\kappa\int_0^t c_s^i \mathrm{d}s\Big) = \epsilon_i^{-\delta_i}\big(\tilde{c}_t^{(-i)}\big)^{\theta_i(1-\delta_i)} e^{-\kappa\delta_i\eta_i t}\exp\Big(-\kappa(1-\delta_i)\theta_i\int_0^t \big(\bar{c}_s^{(-i)}\big)\mathrm{d}s\Big).$$

After rewriting it with respect to $\tilde{c}_t = \Big(\prod_{k=1}^n c_t^k\Big)^{\frac{1}{n}}$ and $\bar{c}_t = \frac{1}{n}\sum_{k=1}^n c_t^k$, we get

$$c_t^i \exp\Big(-\kappa\int_0^t c_s^i \mathrm{d}s\Big) = \epsilon_i^{-\frac{\delta_i}{1+\frac{\theta_i}{n-1}(1-\delta_i)}}\,\tilde{c}_t^{\frac{n}{n-1}\frac{\theta_i(1-\delta_i)}{1+\frac{\theta_i}{n-1}(1-\delta_i)}}\,e^{\frac{-\kappa\eta_i\delta_i}{1+\frac{\theta_i}{n-1}(1-\delta_i)}t} \tag{A.2.6}$$
$$\times\exp\Big(-\kappa\frac{n}{n-1}\frac{\theta_i(1-\delta_i)}{1+\frac{\theta_i}{n-1}(1-\delta_i)}\int_0^t \bar{c}_s\mathrm{d}s\Big).$$

We take the geometric average of Equation (A.2.6) over $i = 1,\dots,n$ to obtain

$$\tilde{c}_t\exp\Big(-\kappa\int_0^t \bar{c}_s\mathrm{d}s\Big) = \Big(\widetilde{\epsilon^\delta}\Big)^{-1} e^{\kappa\overline{\eta\delta}t}\Big(\tilde{c}_t\exp\Big(-\kappa\int_0^t \bar{c}_s\mathrm{d}s\Big)\Big)^{\overline{\theta(1-\delta)}}, \tag{A.2.7}$$

where

$$\widetilde{\epsilon^{\delta}} = \left(\prod_{k=1}^{n} \epsilon_k^{\frac{\delta_k}{1+\frac{\theta_k}{n-1}(1-\delta_k)}}\right)^{\frac{1}{n}}, \quad \overline{\eta\delta} = \frac{1}{n}\sum_{k=1}^{n} \frac{\eta_i\delta_i}{1+\frac{\theta_k}{n-1}(1-\delta_k)},$$
$$\text{and} \quad \overline{\theta(1-\delta)} = \frac{1}{n-1}\sum_{k=1}^{n} \frac{\theta_k(1-\delta_k)}{1+\frac{\theta_k}{n-1}(1-\delta_k)}.$$

Hence we obtain that

$$\tilde{c}_t \exp\Big(-\kappa\int_0^t \bar{c}_s \mathrm{d}s\Big) = \Big(\widetilde{\epsilon^{\delta}}\Big)^{\frac{1}{\overline{\theta(1-\delta)}-1}} e^{\frac{-\kappa\overline{\eta\delta}}{\overline{\theta(1-\delta)}-1}t}. \tag{A.2.8}$$

Using the previous equation we rewrite (A.2.6) as

$$c_t^i \exp\Big(-\kappa\int_0^t c_s^i \mathrm{d}s\Big) = \lambda_i e^{-\kappa\beta_i t}, \tag{A.2.9}$$

where

$$\lambda_i = \epsilon_i^{-\frac{\delta_i}{1+\frac{\theta_i}{n-1}(1-\delta_i)}} \Big(\widetilde{\epsilon^{\delta}}\Big)^{\frac{n}{n-1}\frac{\theta_i(1-\delta_i)}{(\overline{\theta(1-\delta)}-1)(1+\frac{\theta_i}{n-1}(1-\delta_i))}},$$
$$\beta_i = \frac{1}{1+\frac{\theta_i}{n-1}(1-\delta_i)}\bigg(\frac{n}{n-1}\frac{\theta_i(1-\delta_i)}{\overline{\theta(1-\delta)}-1}\overline{\eta\delta} - \eta_i\delta_i\bigg).$$

Now consider two distinct cases $\kappa \neq 0$ and $\kappa = 0$.

Case 1: Let $\kappa \neq 0$. Integrating (A.2.9) from 0 to t and taking logarithms we get

$$\kappa\int_0^t c_s^i \mathrm{d}s = \begin{cases} -\log\Big(1+\frac{\lambda_i}{\beta_i}\big(e^{-\kappa\beta_i t}-1\big)\Big), & \beta_i \neq 0, \\ -\log(1-\lambda_i\kappa t), & \beta_i = 0. \end{cases} \tag{A.2.10}$$

Finally differentiating (A.2.10) with respect to t, we obtain

$$c_t^i = \begin{cases} \bigg(\frac{1}{\beta_i} + \Big(\frac{1}{\lambda_i}-\frac{1}{\beta_i}\Big)e^{\kappa\beta_i t}\bigg)^{-1}, & \beta_i \neq 0, \\ \big(-\kappa t + \frac{1}{\lambda_i}\big)^{-1}, & \beta_i = 0. \end{cases} \tag{A.2.11}$$

By direct inspection it is obvious that for certain combinations of parameters, namely $\kappa > 0,\ \beta_i < \lambda_i$ and $\kappa > 0,\ \beta_i = 0$, the optimal consumption c in not continuous and can even be negative. These cases are not admissible. Admissibility against parameter combinations is summarised in Table 7.4.1.

Now, from the first line of (A.2.5), we get that

$$f_i(t) = \Big(\big(c_t^i\big)^{\frac{1}{\delta_i}}\big(\tilde{c}_t^{(-i)}\big)^{\theta_i(\frac{1}{\delta_i}-1)}\epsilon_i\Big)^{-\frac{1}{\kappa}} \quad \text{and} \quad g_i(t) = f_i(t)^{1-\kappa}. \tag{A.2.12}$$

Case 2: Let $\kappa = 0$. Then (A.2.9) immediately yields $c_t^i = \lambda_i$ which is the same result if one sets $\kappa = 0$ in (A.2.11). □

A.3 Proof of Theorem 7.3.4

We prove Theorem 7.3.4 in full detail.

Proof. We proceed stepwise in order to construct the constant mean field-equilibrium. To that end we must solve ii)-iii) in Definition 7.3.2 for given processes $\overline{X}, \overline{\Gamma}$ associated to some $(\pi, c) \in \mathcal{A}_{\mathrm{MF}}$. Condition iv) of the MF-equilibrium allows us to focus only on processes of the form

$\overline{X}_t = \exp \mathbb{E}[\log X_t|\mathcal{F}_t^B]$ and $\overline{\Gamma}_t = \exp \mathbb{E}[\log c_t|\mathcal{F}_t^B]$ where X solves (7.3.3) for some strategy $(\pi, c) \in \mathcal{A}_{MF}$.

Step 0. The average wealth process. To solve the above problem given $(\overline{X}_t)_{t\geqslant 0}$ it suffices to restrict ourselves to processes $(\overline{X}_t)_{t\geqslant 0}$ satisfying $\overline{X}_t = \exp \mathbb{E}[\log X_t^\pi|\mathcal{F}_t^B]$, $\mathbb{P}$-a.s. We then have via Itô's formula and the arguments from [73] $\mathbb{P}$-a.s.

$$\begin{aligned}
\overline{X}_t &= \exp \mathbb{E}[\log X_t|\mathcal{F}_t^B] \\
&= \exp \mathbb{E}\Big[\log \xi + \int_0^t \big(\mu\pi_s - \frac{1}{2}\pi_s^2(\nu^2+\sigma^2)\big)\mathrm{d}s \\
&\qquad\qquad\qquad + \int_0^t \nu\pi_s \mathrm{d}W_s + \int_0^t \sigma\pi_s \mathrm{d}B_s - \int_0^t c_s \mathrm{d}s \Big| \mathcal{F}_t^B\Big] \\
&= \exp\Big[\overline{\log\xi} + \int_0^t \big(\overline{\mu\pi_s} - \frac{1}{2}\overline{\Sigma\pi_s^2}\big)\mathrm{d}s + \int_0^t \overline{\sigma\pi_s}\mathrm{d}B_s - \int_0^t \overline{c}_s \mathrm{d}s\Big] \\
&= \overline{\xi} + \Big(\int_0^t \eta\overline{X}_s \mathrm{d}s + \int_0^t \overline{\sigma\pi_s}\,\overline{X}_s \mathrm{d}B_s - \int_0^t \overline{c}_s\overline{X}_s\mathrm{d}s\Big),
\end{aligned}$$

where, for consistency of notation with respect to the previous section, we denote

$$\eta := \overline{\mu\pi_s} - \frac{1}{2}(\overline{\Sigma\pi_s^2} - \overline{\sigma\pi_s}^2),\;\; \overline{\xi} := \exp\mathbb{E}[\log\xi],\;\; \overline{\mu\pi_s} := \mathbb{E}[\mu\pi_s],\;\; \overline{\sigma\pi_s} := \mathbb{E}[\sigma\pi_s], \quad \overline{c} = \mathbb{E}[c].$$

Hence, for $(\pi, c) \in \mathcal{A}^{\mathrm{MF}}$ we can define the process $Z^{\pi,c} = X^{\pi,c}\overline{X^{\pi,c}}^{-\theta}$. By Itô's formula we derive its SDE dynamics as

$$\begin{aligned}
\frac{\mathrm{d}Z_t^{\pi,c}}{Z_t^{\pi,c}} &= \big(\mu\pi_t - \theta\overline{\mu\pi_t} + \frac{\theta}{2}\overline{\Sigma\pi_t^2} + \frac{\theta^2}{2}\overline{\sigma\pi_t}^2 - \theta\sigma\pi_t\overline{\sigma\pi_t}\big)\mathrm{d}t + \nu\pi_t\mathrm{d}W_t + \big(\sigma\pi_t - \theta\overline{\sigma\pi_t}\big)\mathrm{d}B_t, \\
&\quad - (c_t - \theta\overline{c}_t)\mathrm{d}t, \quad Z_0^{\pi,c} = \xi(\overline{\xi})^{-\theta}.
\end{aligned}$$

We proceed to solve the MFG Forward performance problem of Definition 7.3.2 with its help.

Applying Itô's formula to $U(Z_t^{\pi,c}, t)$ yields

$$\begin{aligned}
\mathrm{d}Q(Z_t^{\pi,c}, t) &= U_t(Z_t^{\pi,c}, t)\mathrm{d}t + U_x(Z_t^{\pi,c}, t)\mathrm{d}Z_t^{\pi,c} + \frac{1}{2}U_{xx}(Z_t^{\pi,c}, t)\mathrm{d}\langle Z_t^{\pi,c}\rangle + V(Z_t^{\pi,c}, t)\mathrm{d}t \\
&= \Big[U_t(Z_t^{\pi,c}, t) + U_x(Z_t^{\pi,c}, t)\big(\mu\pi_t - \theta\overline{\mu\pi} + \frac{\theta}{2}\overline{\Sigma\pi^2} + \frac{\theta^2}{2}\overline{\sigma\pi}^2 - \theta\sigma\pi_t\overline{\sigma\pi}\big)Z_t^{\pi,c} \\
&\qquad + \frac{1}{2}U_{xx}(Z_t^{\pi,c}, t)\Big((\nu\pi_t)^2 + \big(\sigma\pi_t - \theta\overline{\sigma\pi}\big)^2\Big)(Z_t^{\pi,c})^2\Big]\mathrm{d}t \\
&\qquad + U_x(Z_t^{\pi,c}, t)\nu\pi_t Z_t^{\pi,c}\mathrm{d}W_t + U_x(Z_t^{\pi,c}, t)\big(\sigma\pi_t - \theta\overline{\sigma\pi}\big)Z_t^{\pi,c}\mathrm{d}B_t \\
&\qquad + U_x(Z_t^{\pi,c}, t)(c_t - \theta\overline{c}_t)\mathrm{d}t + V(\hat{c}_t Z_t^{\pi,c}, t)\mathrm{d}t,
\end{aligned} \tag{A.3.1}$$

with $U(Z_0^{\pi,c}, 0) = U(\xi(\overline{\xi})^{-\theta}, 0) = \frac{1}{1-\frac{1}{\delta}}\big(\xi(\overline{\xi})^{-\theta}\big)^{1-\frac{1}{\delta}}$ and using that B, W are i.i.d.

Step 1. The candidate best responses strategies π^*, c^*. As before, the process $U(Z_t^{\pi,c}, t)$ becomes a Martingale at the optimum π. Direct computations using first order conditions (∂_π"drift" $=$ ∂_c"drift" $= 0$) yield

$$\begin{cases}
0 + U_x\cdot\big(\mu - 0 - \theta\sigma\overline{\sigma\pi_t}\big)Z_t^{\pi,c} + \frac{1}{2}U_{xx}\Big(2\pi\nu^2 + 2\big(\sigma\pi_t - \theta\overline{\sigma\pi_t}\big)\sigma\Big)\big(Z_t^{\pi,c}\big)^2 = 0, \\
-U_x(Z_t^{\pi,c}, t)Z_t^{\pi,c} + V_x(\hat{c}_t Z_t^{\pi,c}, t)\frac{Z_t^{\pi,c}}{\left(\overline{c}_t^{(-i)}\right)^{\theta_i}} = 0,
\end{cases}$$

$$\Rightarrow \quad \begin{cases}
\pi_t &= \frac{1}{\nu^2+\sigma^2}\Big(\theta\sigma\overline{\sigma\pi_t} + (\mu - \theta\sigma\overline{\sigma\pi_t})\dfrac{U_x}{U_{xx}Z_t^{\pi,c}}\Big), \\
c_t &= \dfrac{(V_x)^{-1}\Big(U_x(Z_t^{\pi,c}, t)\tilde{c}_t^\theta, t\Big)\tilde{c}_t^\theta}{Z_t^{\pi,c}}.
\end{cases} \tag{A.3.2}$$

Now we inject into the first equation the CRRA constraint $U_x/U_{xx} = -\delta x$ and use Hypothesis

7.3.3 to obtain

$$\pi_t = \frac{1}{\nu^2+\sigma^2}\Big(\theta\sigma\overline{\sigma\pi_t} + (\mu - \theta\sigma\overline{\sigma\pi_t})\delta\Big) \quad \text{and} \quad c_t = \tilde{c}_t^{\theta(1-\delta)} f(t)^{-\delta\kappa}.$$

By inspection, it is clear that π^* is a $\mathcal{F}_0^{\mathrm{MF}}$-measurable RV which is independent of time and is well defined as long as $\overline{\sigma\pi}$ is finite. The derivation of the closed form of the optimal consumption needs further work and is carried out further below.

Step 2. The optimality of the strategy. In contrast to the n-player optimisation the mean field game is defined with reference to a pair of average processes $\overline{X}_t$ and $\overline{\Gamma}_t$ against which the equilibrium is defined through a fixed-point stationarity identity. We provide a verification procedure similar to that in [36, Proof of Theo. 2]. The original constant strategy π is a MF-equilibrium if and only if for all $t \geqslant 0,\ \mathbb{P}$-a.s.

$$\begin{cases} \mathbb{E}[\log X_t^{\pi,c}|\mathcal{F}_t^B] = \mathbb{E}[\log X_t^{\pi^*,c^*}|\mathcal{F}_t^B], \\ \mathbb{E}[\log c_t|\mathcal{F}_t^B] = \mathbb{E}[\log c_t^*|\mathcal{F}_t^B], \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} \overline{\log\xi} + \int_0^t \big(\overline{\mu\pi_s} - \frac{1}{2}\overline{\Sigma\pi_s^2}\big)\mathrm{d}s + \int_0^t \overline{\sigma\pi_s}\mathrm{d}B_s - \int_0^t \bar{c}_s\mathrm{d}s \\ \qquad\qquad = \overline{\log\xi} + \int_0^s \big(\overline{\mu\pi_s^*} - \frac{1}{2}\overline{\Sigma(\pi_s^*)^2}\big)\mathrm{d}s + \int_0^t \overline{\sigma\pi_s^*}\mathrm{d}B_s - \int_0^t \overline{c_s^*}\mathrm{d}s, \\ \tilde{c}_t = \tilde{c}_t^*, \end{cases}$$

where we denote $\tilde{c}_t := \exp\mathbb{E}[\log c_t|\mathcal{F}_t^B] = \exp\mathbb{E}[\log c_t^*|\mathcal{F}_t^B] =: \tilde{c}_t^*$. After taking expectations in the first equation it follows that π is a MF-equilibrium if and only if the following three conditions hold $\mathbb{P}$-a.s.

$$\begin{cases} \overline{\sigma\pi_t} = \overline{\sigma\pi_t^*}, \\ \int_0^t \big(\overline{\mu\pi_s} - \frac{1}{2}\overline{\Sigma\pi_s^2}\big)\mathrm{d}s \ - \int_0^t \overline{c_s}\mathrm{d}s = \int_0^t \big(\overline{\mu\pi_s^*} - \frac{1}{2}\overline{\Sigma(\pi_s^*)^2}\big)\mathrm{d}s - \int_0^t \overline{c_s^*}\mathrm{d}s, \\ \tilde{c}_t = \tilde{c}_t^*. \end{cases} \tag{A.3.3}$$

Using (A.3.2) (with $U_x/U_{xx} = -\delta x$ replaced in) one derives (using the expressions $\varphi^\sigma, \psi^\sigma$)

$$\sigma\pi_t^* = \theta(1-\delta)\frac{\sigma^2}{\nu^2+\sigma^2}\overline{\sigma\pi_t} + \delta\frac{\mu\sigma}{\nu^2+\sigma^2} \qquad \Rightarrow \quad \overline{\sigma\pi^*} = \overline{\sigma\pi_t}\psi^\sigma + \varphi^\sigma.$$

Using that $\overline{\sigma\pi_t} = \overline{\sigma\pi_t^*}$ yields solvability if $\psi^\sigma = \mathbb{E}\big[\theta(1-\delta)\frac{\sigma^2}{\nu^2+\sigma^2}\big] \neq 1$. Thus

$$\overline{\sigma\pi^*} = \overline{\sigma\pi} = \frac{\varphi^\sigma}{1-\psi^\sigma} = \text{constant}, \tag{A.3.4}$$

and the π^* expression of (7.3.5) follows. To exploit the next condition, we solve PDE (7.3.7) under Hypothesis 7.3.3. Together with the optimal candidate consumption we have

$$\begin{cases} c_t^* = \epsilon^{-\delta}(\tilde{c}_t)^{\theta(1-\delta)} f_i(t)^{-\kappa\delta}, \\ f'(t) + \Big(\chi - \theta\big(1-\frac{1}{\delta}\big)\big(\overline{\mu\pi_t} - \frac{1}{2}\overline{\Sigma\pi_t^2} - \bar{c}_t\big)\Big) f(t) + \frac{\epsilon^{-\delta}}{\delta}(\tilde{c}_t)^{\theta(1-\delta)} f(t)^{1-\kappa\delta} = 0, \end{cases} \tag{A.3.5}$$

with

$$\chi = \Big(1-\frac{1}{\delta}\Big)\Bigg(\frac{\delta\big(\mu - \theta\sigma\overline{\sigma\pi^*}(1-\frac{1}{\delta})\big)^2}{2(\nu^2+\sigma^2)} + \frac{\theta^2}{2}\big(\overline{\sigma\pi^*}\big)^2\big(1-\frac{1}{\delta}\big) - r(1-\theta)\Bigg) - \rho.$$

Plugging the first equation of (A.3.5) into the second one, we solve for f to obtain

$$f(t) = \exp\Big(-\int_0^t \Big(\chi - \theta\big(1-\frac{1}{\delta}\big)\big(\overline{\mu\pi_s} - \frac{1}{2}\overline{\Sigma\pi_s^2} - \bar{c}_s\big) + \frac{1}{\delta}c_s^*\Big)\mathrm{d}s\Big).$$

Now we substitute it back into the first equation of (A.3.5) to get

$$c_t^* \exp\big(-\kappa \int_0^t c_s^* \mathrm{d}s\big) = \epsilon^{-\delta}(\tilde{c}_t)^{\theta(1-\delta)} e^{\chi\delta\kappa t} \\ \times \exp\Big(\theta(1-\delta)(\kappa-1)\int_0^t \big(\overline{\mu\pi_s} - \frac{1}{2}\overline{\Sigma\pi_s^2} - \bar{c}_s\big)\mathrm{d}s\Big).$$

Now we substitute according to second equilibrium identity of (A.3.3) to obtain

$$c_t^* \exp\big(-\kappa \int_0^t c_s^* \mathrm{d}s\big) = \epsilon^{-\delta}(\tilde{c}_t)^{\theta(1-\delta)} e^{\eta\delta\kappa t} \exp\big(-\theta(1-\delta)\kappa \int_0^t \overline{c_s^*}\mathrm{d}s\big), \tag{A.3.6}$$

where

$$\eta_t = \chi - \theta\big(1-\frac{1}{\delta}\big)\Big(\overline{\mu\pi_t^*} - \frac{1}{2}\overline{\Sigma(\pi_t^*)^2}\Big).$$

Taking the logarithm, expectation and exponent on both sides of (A.3.6) we get

$$\tilde{c}_t^* \exp\big(-\kappa \int_0^t \overline{c_s^*}\mathrm{d}s\big) = \big(\widetilde{\epsilon^\delta}\big)^{-1} e^{\kappa\overline{\eta\delta}t}\Big(\tilde{c}_t \exp\big(-\kappa\int_0^t \overline{c_s^*}\mathrm{d}s\big)\Big)^{\overline{\theta(1-\delta)}},$$

where $\widetilde{\epsilon^\delta} = \exp\mathbb{E}[\delta\log\epsilon]$, $\overline{\eta\delta} = \mathbb{E}[\eta\delta]$ and $\overline{\theta(1-\delta)} = \mathbb{E}[\theta(1-\delta)]$. Using the last equilibrium identity, $\tilde{c}_t = \tilde{c}_t^*$, we rewrite the expression as

$$\tilde{c}_t^* \exp\big(-\kappa \int_0^t \overline{c_s^*}\mathrm{d}s\big) = \widetilde{\epsilon^\delta}^{\frac{1}{\overline{\theta(1-\delta)}-1}} e^{-\frac{\overline{\eta\delta}}{\overline{\theta(1-\delta)}-1}\kappa t}.$$

Plugging it back into (A.3.6) we obtain

$$c_t^* \exp\big(-\kappa \int_0^t c_s^* \mathrm{d}s\big) = \lambda e^{-\kappa\beta t},$$

where

$$\lambda = \epsilon^{-\delta}\big(e^{\mathbb{E}[\delta\log\epsilon]}\big)^{\frac{\theta(1-\delta)}{\mathbb{E}[\theta(1-\delta)]-1}}, \qquad \beta = \frac{\theta(1-\delta)}{\mathbb{E}[\theta(1-\delta)]-1}\mathbb{E}[\eta\delta] - \eta\delta.$$

The same arguments used in the proof of Theorem 7.2.6 yield

$$c_t^* = \begin{cases} \Big(\frac{1}{\beta} + \big(\frac{1}{\lambda} - \frac{1}{\beta}\big)e^{\kappa\beta t}\Big)^{-1}, & \beta \neq 0, \\ \big(-\kappa t + \frac{1}{\lambda}\big)^{-1}, & \beta = 0. \end{cases}$$

As in Theorem 7.2.6, one finds that for certain combinations of parameters, namely $\kappa > 0,\ \beta < \lambda$ and $\kappa > 0,\ \beta = 0$, the optimal consumption c in not continuous and even can be negative. These cases are not admissible and admissibility against parameter combinations is summarised in Table 7.4.1.

Finally, from the first line of (A.3.5), we get that

$$f(t) = \Big(\big(c_t\big)^{\frac{1}{\delta}}\big(\tilde{c}_t\big)^{\theta(\frac{1}{\delta}-1)}\epsilon\Big)^{-\frac{1}{\kappa}} \quad \text{and thus} \quad g(t) = \Big(\big(c_t\big)^{\frac{1}{\delta}}\big(\tilde{c}_t\big)^{\theta(\frac{1}{\delta}-1)}\epsilon\Big)^{\frac{\kappa-1}{\kappa}}.$$

Now we are left to conclude using the closed forms of $\overline{\mu\pi_t^*}$ and $\overline{\Sigma(\pi_t^*)^2}$. First, we have

$$\overline{\mu\pi_t^*} = \frac{\varphi^\sigma}{1-\psi^\sigma}\psi^\mu + \varphi^\mu = \text{constant}, \tag{A.3.7}$$

using ψ^μ, φ^μ. Finally, we find the expression for $\overline{\Sigma\pi^2}$. Multiplying separately the π^* of (7.3.5) by σ and ν, squaring, taking expectation and summing the results, we have

$$\overline{\Sigma\pi^2} = \mathbb{E}\Big[\frac{1}{\nu^2+\sigma^2}\Big(\theta(1-\delta)\sigma\frac{\varphi^\sigma}{1-\psi^\sigma}+\mu\delta\Big)^2\Big]. \tag{A.3.8}$$

The explicit form of the expression η follows by injecting these identities in (7.3.6).

Step 3. The MFG forward performance process dynamics. Injecting the consistency PDE (7.3.7) in the expression for $\mathrm{d}U(Z_t^{\pi,c},t)$ given into (A.3.1) yields

$$\begin{aligned}
&\mathrm{d}Q(Z_t^{\pi,c},t)\\
&= U_x(Z_t^{\pi,c},t)\Big(\nu\pi_t\mathrm{d}W_t + \big(\sigma\pi_t - \theta\frac{\varphi^\sigma}{1-\psi^\sigma}\big)\mathrm{d}B_t\Big)Z_t^{\pi,c}\\
&\quad+\frac{1}{2}U_{xx}(Z_t^{\pi,c},t)\frac{1}{(\nu^2+\sigma^2)}\Big|\pi_t(\nu^2+\sigma^2)-\Big(\theta(1-\delta)\sigma\frac{\varphi^\sigma}{1-\psi^\sigma}+\mu\delta\Big)\Big|^2(Z_t^{\pi,c})^2\mathrm{d}t\\
&\quad+V(\hat{c}_tZ_t^{\pi,c},t)-U_x(Z_t^{\pi,c},t)c_tZ_t^{\pi,c}-\widetilde{V}(U_x(Z_t^{\pi,c},t),t).
\end{aligned}$$

$\square$

A.4 Regions of monotonicity of optimal consumption

We enhance Fig. 3 (case $\kappa = 1$) of the earlier contribution [72] to the context of our work and comment further on the finer interplay of κ with β and λ on the consumption policy c (see end of Section 7.4).

In Figure A.4.1, we have two pictures of the regions of monotonicity of c_t for $\kappa > 0$ and $\kappa < 0$. We can see that having κ across the region given by $\delta = 1$ symmetrically reverses the direction of monotonicity for c. The region of consumption in the plot having a constant consumption regime has a constant color (we do not mark such level curves apart from the region boundaries).

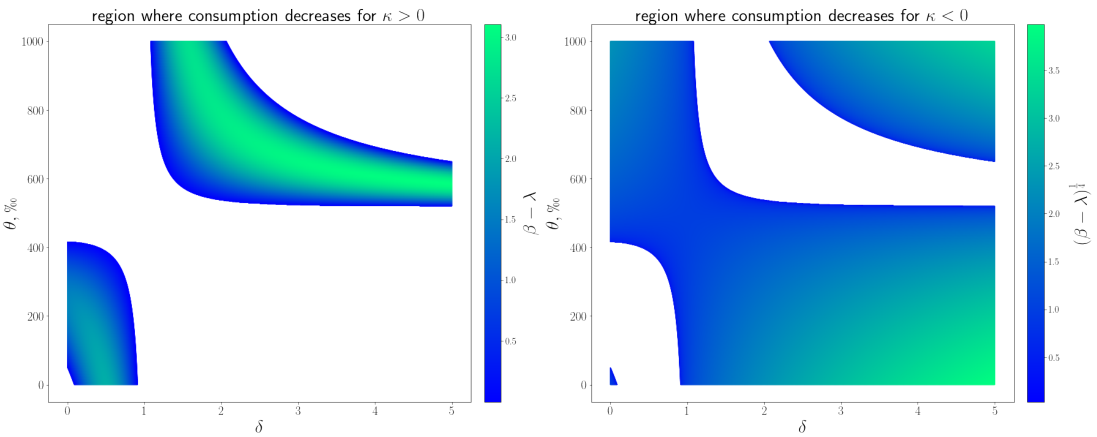

Figure A.4.1: The regions of monotonicity for c_t for $\kappa > 0$ and $\kappa < 0$ as function of δ and θ (in ‰), the agent lying inside (outside) the coloured region decreases (increases) consumption rate over time. The agents on the border consume at a constant rate. The colour gradient relates to the speed of monotonicity characterised by $\beta - \lambda$ or a function of it. The set of parameters is taken from [72, Fig. 3 (case $\kappa = 1$)], namely $\mu = 5,\ \sigma = 1,\ \epsilon = 1,\ \mathbb{E}[\log\epsilon] = 0,\ \mathbb{E}[\theta(1-\delta)] = -1.6,\ \mathbb{E}[\delta] = 5,\ \theta_{\text{crit}} = 0.52,\ \rho = r = 0.$

Bibliography

[1] K. K Aase, *Recursive utility using the stochastic maximum principle*, Quantitative Economics **7** (2016), no. 3, 859–887.

[2] F. Ahmad, B. M. Hambly, and S. Ledger, *A stochastic partial differential equation model for the pricing of mortgage-backed securities*, Stochastic Process. Appl. **128** (2018), no. 11, 3778–3806. MR3860010

[3] M. Anthropelos, T. Geng, and T. Zariphopoulou, *Competition in fund management and forward relative performance criteria*, 2020. ResearchGate preprint, DOI:10.13140/RG.2.2.20712.90889.

[4] L. Avanesyan, M. Shkolnikov, and R. Sircar, *Construction of a class of forward performance processes in stochastic factor models, and an extension of Widder's theorem*, Finance Stoch. **24** (2020), no. 4, 981–1011. MR4152279

[5] S. A. Azar, *The Nexus Between the Elasticity of Intertemporal Substitution and the Coefficient of Relative Risk Aversion*, International Journal of Financial Research **9** (2018July), no. 3, 98–102.

[6] P. Bank and F. Riedel, *Existence and structure of stochastic equilibria with intertemporal substitution*, Finance Stoch. **5** (2001), no. 4, 487–509. MR1861998

[7] E. Bayraktar, A. Cecchin, A. Cohen, and F. Delarue, *Finite state mean field games with wright–fisher common noise*, Journal de Mathématiques Pures et Appliquées **147** (2021), 98–162.

[8] A. Bensoussan, J. Frehse, and P. Yam, *Mean field games and mean field type control theory*, SpringerBriefs in Mathematics, Springer, New York, 2013. MR3134900

[9] F. E. Benth, K. H. Karlsen, and K. Reikvam, *Optimal portfolio management rules in a non-Gaussian market with durability and intertemporal substitution*, Finance Stoch. **5** (2001), no. 4, 447–467. MR1861996

[10] F. Berrier and M. Tehranchi, *Forward utility of investment and consumption*, preprint **2** (2011), no. 9, 12.

[11] J. Bielagk, A. Lionnet, and G. Dos Reis, *Equilibrium pricing under relative performance concerns*, SIAM J. Financial Math. **8** (2017), no. 1, 435–482. MR3666773

[12] A. Bismuth and O. Guéant, *It's all relative: mean field game extensions of Merton's problem*, 2019. Presentation at DEA Krakow, work forthcoming.

[13] M. Bossy, J.-F. Jabir, and D. Talay, *On conditional mckean lagrangian stochastic models*, Probability Theory and Related Fields **151** (2011Oct), no. 1, 319–351.

[14] M. J Brennan, *The optimal number of securities in a risky asset portfolio when there are fixed costs of transacting: Theory and some empirical results*, Journal of Financial and Quantitative analysis (1975), 483–496.

[15] R. Buckdahn, J. Li, S. Peng, and C. Rainer, *Mean-field stochastic differential equations and associated PDEs*, Ann. Probab. **45** (2017), no. 2, 824–878. MR3630288

[16] C. H. Campani and R. Garcia, *Approximate analytical solutions for consumption/investment problems under recursive utility and finite horizon*, The North American Journal of Economics and Finance **48** (2019), 364–384.

[17] J. Y Campbell and N G. Mankiw, *The response of consumption to income: a cross-country investigation*, European economic review **35** (1991), no. 4, 723–756.

[18] P. Cardaliaguet, *Notes on mean field games*, Technical report, 2013.

[19] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, *The master equation and the convergence problem in mean field games:(ams-201)*, Vol. 381, Princeton University Press, 2019.

[20] R. Carmona and F. Delarue, *The master equation for large population equilibriums*, Stochastic analysis and applications 2014, 2014, pp. 77–128.

[21] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications I*, 1st ed., Probability Theory and Stochastic Modelling, vol. 84, Springer International Publishing, 2017.

[22] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications II*, 1st ed., Probability Theory and Stochastic Modelling, vol. 84, Springer International Publishing, 2017.

[23] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications i-ii*, Springer, 2018.

[24] R. A. Carmona and B. Rozovskii (eds.), *Stochastic partial differential equations: six perspectives*, Mathematical Surveys and Monographs, vol. 64, American Mathematical Society, Providence, RI, 1999. MR1661761

[25] T. Cavallazzi, *Itô-Krylov's formula for a flow of measures*, arXiv preprint arXiv:2110.05251 (2021).

[26] J.-F. Chassagneux, D. Crisan, and F. Delarue, *A probabilistic approach to classical solutions of the master equation for large population equilibria*, arXiv preprint arXiv:1411.3009 (2014).

[27] W. F. Chong, *Pricing and hedging equity-linked life insurance contracts beyond the classical paradigm: The principle of equivalent forward preferences*, Insurance: Mathematics and Economics **88** (2019), 93–107.

[28] W. F. Chong and G. Liang, *Optimal investment and consumption with forward preferences and uncertain parameters*, arXiv preprint arXiv:1807.01186 (2018).

[29] R. Cont and D.-A. Fournié, *Functional Itô calculus and stochastic integral representation of martingales*, Ann. Probab. **41** (2013), no. 1, 109–133. MR3059194

[30] D. Crisan and E. McMurray, *Smoothing properties of McKean-Vlasov SDEs*, Probab. Theory Related Fields **171** (2018), no. 1-2, 97–148. MR3800831

[31] L. D. Cundy, *Essays on the elasticity of intertemporal substitution*, University of Iowa, 2018. PhD dissertation.

[32] A. B. de Leon, D. Holm, E. Luesink, and S. Takao, *Implications of Kunita-Itô-Wentzell formula for k-forms in stochastic fluid dynamics*, arXiv preprint arXiv:1903.07201 (2019).

[33] C. Deng, X. Su, and C. Zhou, *Relative wealth concerns with partial information and heterogeneous priors*, arXiv preprint arXiv:2007.11781 (2020).

[34] G. dos Reis and V. Platonov, *On the relation between Stratonovich and Itô integrals with functional integrands of conditional measure flows*, arXiv preprint arXiv:2111.03523 (2021).

[35] G. dos Reis and V. Platonov, *Forward utility and market adjustments in relative investment-consumption games of many players*, arXiv preprint arXiv:2012.01235 (2020).

[36] G. dos Reis and V. Platonov, *Forward utilities and mean-field games under relative performance concerns*, From Particle Systems to Partial Differential Equations. ICPS 2019, ICPS 2018, ICPS 2017, 2021, pp. 227–251.

[37] G. dos Reis and V. Platonov, *Forward utilities and mean-field games under relative performance concerns*, From particle systems to partial differential equations, 2021, pp. 227–251.

[38] G. dos Reis and V. Platonov, *Itô-Wentzell-Lions Formula for Measure Dependent Random Fields under Full and Conditional Measure Flows*, Potential Analysis (2022). 1572-929X.

[39] R. Duboscq and A. Réveillac, *Stochastic regularization effects of semi-martingales on random functions*, J. Math. Pures Appl. (9) **106** (2016), no. 6, 1141–1173. MR3565418

[40] D. Duffie and L. G Epstein, *Stochastic differential utility*, Econometrica: Journal of the Econometric Society (1992), 353–394.

[41] N. El Karoui, C. Hillairet, and M. Mrad, *Consistent utility of investment and consumption: a forward/backward SPDE viewpoint*, Stochastics **90** (2018), no. 6, 927–954. MR3814804

[42] N. El Karoui, C. Hillairet, and M. Mrad, *Construction of an aggregate consistent utility, without Pareto optimality. Application to long-term yield curve modeling*, Frontiers in stochastic analysis—BSDEs, SPDEs and their applications, 2019, pp. 169–199. MR4008344

[43] N. El Karoui and M. Mrad, *An exact connection between two solvable SDEs and a nonlinear utility stochastic PDE*, SIAM J. Financial Math. **4** (2013), no. 1, 697–736. MR3106475

[44] L. G. Epstein and S. E. Zin, *Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework*, Econometrica **57** (1989), no. 4, 937–969.

[45] X. Erny, E. Löcherbach, and D. Loukianova, *Conditional propagation of chaos for mean field systems of interacting neurons*, Electron. J. Probab. **26** (2021), Paper No. 20, 25. MR4235471

[46] G.-E. Espinosa and N. Touzi, *Optimal Investment under Relative Performance Concerns*, Mathematical Finance **25** (2015), no. 2, 221–257.

[47] F. Flandoli and F. Russo, *Generalized integration and stochastic ODEs*, Ann. Probab. **30** (2002), no. 1, 270–292. MR1894108

[48] C. Frei and G. Dos Reis, *A financial market with interacting investors: does an equilibrium exist?*, Mathematics and financial economics **4** (2011), no. 3, 161–182. MR2796281

[49] G. Fu, X. Su, and C. Zhou, *Mean field exponential utility game: A probabilistic approach*, arXiv preprint arXiv:2006.07684 (2020).

[50] W. Fung Chong, Y. Hu, G. Liang, and T. Zariphopoulou, *An ergodic BSDE approach to forward entropic risk measures: representation and large-maturity behavior*, arXiv:1607.02289 (2016).

[51] J. Gali, *Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices*, Journal of Money, Credit and Banking (1994), 1–8.

[52] T. Geng and T. Zariphopoulou, *Passive and competitive investment strategies under relative forward performance criteria*, Available at SSRN 2870040 (2017).

[53] M. Gerencsér and I. Gyöngy, *Localization errors in solving stochastic partial differential equations in the whole space*, Math. Comp. **86** (2017), no. 307, 2373–2397. MR3647962

[54] E. Gobet and M. Mrad, *Convergence rate of strong approximations of compound random maps, application to SPDEs*, Discrete Contin. Dyn. Syst. Ser. B **23** (2018), no. 10, 4455–4476. MR3927828

[55] X. Guo, H. Pham, and X. Wei, *Itô's formula for flow of measures on semimartingales*, arXiv preprint arXiv:2010.05288 (2020).

[56] R. E. Hall, *Intertemporal substitution in consumption*, Journal of Political Economy **96** (1988), no. 2, 339–357.

[57] V. Henderson and D. Hobson, *Horizon-unbiased utility functions*, Stochastic processes and their applications **117** (2007), no. 11, 1621–1641.

[58] A. Hindy, C.-F. Huang, and D. Kreps, *On intertemporal preferences in continuous time: The case of certainty*, Journal of Mathematical Economics **21** (1992), no. 5, 401–440.

[59] D. Holm and R Hu, *Stochastic effects of waves on currents in the ocean mixed layer* (2020).

[60] M. Huang, R. P. Malhamé, and P. E. Caines, *Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst. **6** (2006), no. 3, 221–251. MR2346927

[61] M. Kacperczyk, C. Sialm, and L. Zheng, *On the industry concentration of actively managed equity mutual funds*, The Journal of Finance **60** (2005), no. 4, 1983–2011.

[62] S. Källblad, *Black's inverse investment problem and forward criteria with consumption*, SIAM Journal on Financial Mathematics **11** (2020), no. 2, 494–525.

[63] O. Kammar, *A note on Fréchet diffrentiation under Lebesgue integrals*, 2016. Preprint.

[64] I. Karatzas, J. P Lehoczky, S. P Sethi, and S. E Shreve, *Explicit solution of a general consumption/investment problem*, Optimal consumption and investment with bankruptcy, 1997, pp. 21–56.

[65] V. N. Kolokoltsov and M. Troeva, *Regularity and sensitivity for mckean-vlasov spdes*, AIP Conference Proceedings **1907** (2017), no. 1, 030046.

[66] D. M. Kreps and E. L. Porteus, *Temporal resolution of uncertainty and dynamic choice theory*, Econometrica **46** (1978), no. 1, 185–200.

[67] N. V. Krylov, *On the Itô-Wentzell formula for distribution-valued processes and related topics*, Probab. Theory Related Fields **150** (2011), no. 1-2, 295–319. MR2800911

[68] N. V. Krylov and T. Wang, *Filtering partially observable diffusions up to the exit time from a domain*, Stochastic Process. Appl. **121** (2011), no. 8, 1785–1815. MR2811024

[69] H. Kunita, *Some extensions of Itô's formula*, Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French), 1981, pp. 118–141. MR622557

[70] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1997. Reprint of the 1990 original. MR1472487

[71] D. Lacker, M. Shkolnikov, and J. Zhang, *Inverting the Markovian projection, with an application to local stochastic volatility models*, The Annals of Probability **48** (2020), no. 5, 2189 –2211.

[72] D. Lacker and A. Soret, *Many-player games of optimal consumption and investment under relative performance criteria*, Math. Financ. Econ. **14** (2020), no. 2, 263–281. MR4076185

[73] D. Lacker and T. Zariphopoulou, *Mean field and n-agent games for optimal investment under relative performance criteria*, to appear in Mathematical Finance (2017).

[74] T. Leung, R. Sircar, and T. Zariphopoulou, *Forward indifference valuation of american options*, Stochastics An International Journal of Probability and Stochastic Processes **84** (2012), no. 5-6, 741–770.

[75] J. Li, W. Li, and G. Liang, *A game theoretical approach to homothetic robust forward investment performance processes in stochastic factor models*, Available at SSRN 3601075 (2020).

[76] P.-L. Lions, *Cours au College de France: : Théorie des jeux a champs moyens*, Available at www.college-de-france.fr (2007).

[77] A. Matoussi and M. Mrad, *Dynamic utility and related nonlinear SPDE driven by lévy noise.*, HaL preprint hal-03025475 (2020Nov).

[78] R. C Merton, *Lifetime portfolio selection under uncertainty: The continuous-time case*, The review of Economics and Statistics (1969), 247–257.

[79] M. Musiela and T. Zariphopoulou, *Portfolio choice under dynamic investment performance criteria*, Quant. Finance **9** (2009), no. 2, 161–170. MR2512986

[80] M. Musiela and T. Zariphopoulou, *Portfolio choice under space-time monotone performance criteria*, SIAM J. Financial Math. **1** (2010), no. 1, 326–365. MR2652069

[81] M. Musiela and T. Zariphopoulou, *Investments and forward utilities*, preprint (2006).

[82] M. Musiela and T. Zariphopoulou, *Stochastic partial differential equations and portfolio choice*, Contemporary quantitative finance, 2010, pp. 195–216. MR2732847

[83] D. Ocone and É. Pardoux, *A generalized Itô-Ventzell formula. Application to a class of anticipating stochastic differential equations*, 1989, pp. 39–71. MR995291

[84] B. L. Rozovskiĭ, *Stochastic evolution systems*, Mathematics and its Applications (Soviet Series), vol. 35, Kluwer Academic Publishers Group, Dordrecht, 1990. Linear theory and applications to nonlinear filtering, Translated from the Russian by A. Yarkho. MR1135324

[85] P. A Samuelson, *Lifetime portfolio selection by dynamic stochastic programming*, The review of economics and statistics (1969), 239–246.

[86] M. S. Strub and X. Y. Zhou, *Evolution of the Arrow-Pratt measure of risk-tolerance for predictable forward utility processes*, Available at SSRN 3276638 (2018).

[87] M. Talbi, N. Touzi, and J. Zhang, *Dynamic programming equation for the mean field optimal stopping problem*, arXiv preprint arXiv:2103.05736 (2021).

[88] J. Thimme, *Intertemporal substitution in consumption: A literature review*, Journal of Economic Surveys **31** (2017), no. 1, 226–257.

[89] S. Van Nieuwerburgh and L. Veldkamp, *Information immobility and the home bias puzzle*, The Journal of Finance **64** (2009), no. 3, 1187–1215.

[90] A. D. Ventzel, *On equations of theory of conditional Markov processes*, Theory of probability and its applications, ussr, 1965, pp. 357–361.

[91] M. Veraar, *The stochastic Fubini theorem revisited*, Stochastics **84** (2012), no. 4, 543–551. MR2966093

[92] P. Weil, *The equity premium puzzle and the risk-free rate puzzle*, Journal of Monetary Economics **24** (1989), no. 3, 401–421.

[93] G. Zitković, *A dual characterization of self-generation and exponential forward performances*, Ann. Appl. Probab. **19** (2009), no. 6, 2176–2210. MR2588243