Contagious McKean–Vlasov Systems with Positive Feedback

Macro-Modelling of Systemic Risk in Large Financial Systems



Andreas Søjmark

St. Hilda's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

To my parents Else & Per

Acknowledgements

Firstly, I am grateful to my advisor Ben Hambly for supervising the work that went into this thesis and for pointing me towards McKean–Vlasov type problems. Secondly, I wish to thank Sean Ledger for playing a huge role in my DPhil studies and for inviting me to Bristol on several occasions. Without a doubt, Ben and Sean's involvement have made this thesis far more meaningful and far more likely to be of—at least somewhat—interest. Mistakes and blunders are, of course, all mine.

Outside of academia, and most of all, I am thankful to my parents for being such a warm part of my life and for always supporting me in my endeavours. They have played a huge role in who I am and where I am today. Similarly, I wish to thank my brothers and my closest friends for all the fun we have together. As for those who knew me before the days of maths, they deserve an extra thanks for only ever so rarely finding the courage to question what the devil went wrong, when I skipped a turn and ran off a cliff into uncharted mathematical waters.

Concerning life at the maths institute, it has been a true privilege to be part of the brilliant bunch of people that makes up the CDT PDE as well as the Mathematical and Computational Finance Group. Above all, it has been immensely good fun to share an office with Andy and Jono! Moreover, I would like to thank Christoph Reisinger and Vadim Kaushansky for several discussions centred around their shared interest in some of the problems considered in this thesis.

Finally, I am thankful to the EPSRC for paying my most basic fees; and I wholeheartedly thank the various Danish foundations who helped stave off the ever imminent threat of economic ruin, funding my necessities. Special thanks are also due to my private donors Christian and Frederik for providing me with constant refuge in the unholy city of London, and to Klaus—my personal Maecenas—for hosting me in the holiest City of New York whenever my studies took me there.

Abstract

We introduce a family of contagious McKean–Vlasov systems with positive feedback which are shown to arise as the mean field (or large population) limits of an associated family of finite particle systems. In the finite systems, each particle performs some variant of Brownian motion until it hits a lower barrier—at which point it is absorbed and the surviving particles are shifted down in the direction of the barrier. The latter may lead to further absorptions and hence a positive feedback loop is created that can model contagious feedback effects from hitting the barrier.

More specifically, we propose to use this class of McKean–Vlasov systems as a simple framework for macroscopic modelling of systemic risk in large financial systems. In this case, each particle represents the distance-to-default of a financial institution, which is simply a measure of its financial 'healthiness' living on the positive half-line, and absorption at the origin then corresponds to default. The positive feedback emerges naturally by imposing that, whenever there is a default, all the other particles must be shifted closer to the origin—and thus closer to default—through an endogenous notion of default contagion.

In view of the above, the main contributions of this thesis are twofold. Firstly, we establish the convergence of the aforementioned finite dimensional particle systems to well-defined characterizations of their mean field limits. Secondly, we provide novel results on uniqueness and regularity for the appertaining class of limiting McKean–Vlasov problems.

Contents

1 Introduction **1**

1.1 Systemic Risk 1

1.1.1 Two streams of literature on systemic risk 2

1.2 Main objectives of the thesis 3

1.2.1 Macro-modelling: a system-wide perspective risk 3

1.2.2 Remarks on related models of default contagion 4

1.2.3 Large portfolio credit risk with default clustering 6

1.3 Overview of mathematical contributions 6

1.3.1 The SPDE problem with gradual feedback 6

1.3.2 A McKean–Vlasov problem with instantaneous feedback and blow-ups 7

1.4 Outline of the thesis 9

1.4.1 Remarks on related mathematical literature 11

2 Upper Dirichlet Heat Kernel Type Estimates for Drifts of at most Linear Growth **13**

2.1 Main results 13

2.1.1 Outline of the chapter 15

2.2 The Dirichlet Heat Kernel and Boundary Estimates for Absorbed Brownian Motion 15

2.2.1 Absorbed Brownian Motion 16

2.2.2 The heat equation on a half-line 17

2.2.3 Dirichlet Heat Kernel Estimate 18

2.3 Sub-Gaussianity 19

2.3.1 Change of measure 22

2.4 The density estimates 25

2.4.1 Estimates on the half-line with absorption at zero 27

2.4.2 Estimates on the whole space 31

2.4.3 Proof of Theorem 2.1.1 32

3 An SPDE Model for Systemic Risk with Endogenous Contagion 34
3.1 Introduction 34
3.1.1 A simple model of default contagion 35
3.1.2 A simple model for systemic risk: contagion, common exposures, and herding 38
3.1.3 Outline of the chapter 40
3.2 Main results 41
3.2.1 Structural conditions 41
3.2.2 The limit SPDE 43
3.2.3 Financial contagion and default clustering 46
3.3 The finite particle system 50
3.3.1 Evolution equation for the finite system 50
3.3.2 Regularity properties of the particles 51
3.3.3 Regularity properties of the loss process 52
3.4 Tightness and convergence 53
3.4.1 Convergence of the mean process 56
3.4.2 The limit SPDE: Proof of Theorem 3.2.4 58
3.5 Uniqueness arguments 62
3.5.1 Energy estimates and smoothing 62
3.5.2 L^2 regularity and containment of boundary effects 66
3.5.3 The smoothed H^{-1} estimate: Proof of Proposition 3.5.1 71

4 Instantaneous Feedback ⁎ Part I: Local Well-Posedness and Regularity up to Explosion 75
4.1 Introduction 75
4.1.1 Antecedent literature motivated by neuroscience 77
4.2 Main results: On local well-posedness and the occurence of blow-ups . 77
4.2.1 Uniqueness up to explosion 82
4.2.2 Overview of the chapter 85
4.3 Minimal jumps and α-fragility: proof of Proposition 4.2.2 85
4.4 Contractivity of Γ and minimality of differentiable solutions: proof of Theorem 4.2.6 90
4.5 Stability of Γ and the fixed-point argument: proof of Theorem 4.2.7 . 92
4.6 Bootstrapping and full uniqueness up to explosion time: proof of Theorem 4.2.8 100

4.6.1 A duality lemma . . . 101
4.6.2 Bootstrapping of solutions . . . 102
4.6.3 Monotonicity and trapping . . . 104
4.7 Extensions to more general coefficients . . . 108
4.7.1 Stability of the fixed point map . . . 109
4.7.2 Density bounds for bootstrapping . . . 111
4.7.3 Contractivity of the fixed point map . . . 112
4.7.4 Minimality, monotonicity and trapping . . . 113

5 Instantaneous Feedback ⋇ Part II: Convergence of the Particle System with Common Noise 114
5.1 Introduction . . . 114
5.1.1 Overview of the chapter . . . 116
5.2 The possibility and timing of blow-ups . . . 116
5.2.1 The idiosyncratic model with $\rho = 0$ . . . 118
5.2.2 The common noise model with $\rho > 0$ . . . 121
5.3 Existence and particle approximation . . . 124
5.3.1 Preliminary estimates . . . 131
5.3.2 Compactness, Continuity, and Convergence . . . 135
5.3.3 Construction of filtrations . . . 144
5.3.4 Remark on numerics . . . 146

6 Instantaneous Feedback ⋇ Part III: Global Results on Uniqueness and Regularity 147
6.1 Introduction . . . 147
6.1.1 Recent history of the problem . . . 150
6.1.2 Overview of the chapter . . . 151
6.2 Uniqueness in the weak feedback regime . . . 151
6.3 Short-time uniqueness after a blow-up . . . 154
6.3.1 Analyticity of the density in the interior . . . 158
6.4 Global behaviour of the loss with blow-ups . . . 161
6.4.1 Locally weak feedback after a blow-up . . . 163
6.4.2 1/2-Hölder continuity of the loss . . . 165

A Auxiliary Results for Chapter 3 168
A.1 Some technical lemmas . . . 168
A.2 Proofs of Propositions 3.3.5 and 3.3.6 . . . 170

Bibliography **173**

List of Figures

3.2.1 The figure shows two heat plots of the density $(t, x) \mapsto V_t(x)$ from Remark 3.2.5, based on a numerical simulation of the SPDE for a fixed realization of W^0. On the left $\alpha = 1.5$, while on the right $\alpha = 0$. The other parameters are: $\rho = 0.1$, $\sigma = 1$, $b = 0$, and $\mathfrak{K}$ is an isosceles triangle on $[0, 0.015]$ with height $2/0.015$. The realization of W^0 starts on a slightly negative trend, but then moves back up again. 47

3.2.2 In these two heat plots of $(t, x) \mapsto V_t(x)$, we have $\alpha = 2$ on both sides and otherwise the parameters are the same as in Fig. 3.2.1. However, the simulations are done for two different realizations of the common noise W^0: on the left it declines steadily, while on the right it increases correspondingly. 48

4.2.1 Example of a solution to (4.2, 4.3) showing two blow-up times. The pixel intensity represents the value of the solution density at that space-time coordinate. The initial condition is a linear combination of indicator functions of three disjoint sets. 79

4.2.2 On the left, V_{t-} is the density just before a jump of size ΔL_t. This density is then translated by $\alpha \Delta L_t$ and the mass falling below the boundary at zero equals the change in the loss, which gives (4.6). After the jump, the system is restarted from the density on the right. Notice that, in general, this new initial condition will not vanish at the origin. 80

4.3.1 The function from Example 4.3.2 is on the left. The candidate jump sizes—that is, the solutions to (4.6)—are the points on the right where the graphs intersect. The point $(1, 1)$ gives the minimal allowed jump size of 1, since $x = 1$ is the value given by (PJC) 86

4.6.1 Given an initial density, V_0, the ε-deleted initial condition constructed in (4.20) is obtained by killing the mass on $(0, \varepsilon)$ and shifting the density towards the origin by the amount $\alpha \nu_0(0, \varepsilon) + \frac{1}{4}\varepsilon$. The proof of Lemma 4.6.6 shows us that V_0^ε vanishes in a neighbourhood of zero. 105

4.6.2 On a small time interval, the unique differentiable solution, L, from Corollary 4.6.4 and the ε-deleted solutions from Definition 4.6.5 trap any candidate càdlàg solution, $\bar{L}$, by Lemma 4.6.7. Here $\eta > \varepsilon$. Lemma 4.6.8 shows that $L^\varepsilon \to L$ uniformly on a small time interval as $\varepsilon \to 0$, and so we see that $L = \bar{L}$ is forced since the area between the curves above shrinks to zero. 106

5.2.1 The plots show different realisations of the solution to (MV) with $\rho = 0.5$ and the same initial condition. Each pixel represents the value of the density at that space-time point, with space on the vertical and time on the horizontal. On the right, the common noise, B^0, decreases sufficiently quickly to cause a blow-up to occur. 117

6.2.1 A visual proof of Lemma 6.2.1 with $f(x) = x$: The red curve is X^1 and the blue curve is X^2, coupled through the same initial value and driving noise. Writing $L^1_t - L^2_t = \mathbb{P}(\tau_1 \leq t < \tau_2)$, we see that, as x_0 decreases, we only get contributions to this probability from when the red curve first hits the x-axis until the blue curve first hits it. Hence the probability equals $\nu_0(0.59, 0.72)$. Note that the difference between the curves is bounded by $\|L^1 - L^2\|_t$. 153

6.3.1 Two pathological initial densities that satisfy (6.4) with $\alpha = 5/3$. The blue curve takes infinitely many values above the critical value of $3/5$ near zero. The red curve remains strictly below $3/5$, but all derivatives vanish at the origin, so control of the form in (6.5) is not possible. . 156

Chapter 1

Introduction

"How did you go bankrupt?" Bill asked. "Two ways," Mike said...
"Gradually and then suddenly!"
– Bill Gorton and Mike Campbell in The Sun Also Rises

1.1 Systemic Risk

One of the most important lessons of the 2007–2009 financial crisis is the imperative nature of system-wide perspectives on risk. That is, financial models need to take into account the interconnectedness of the financial system, and they must incorporate sensible notions of financial contagion whereby the distress of one institution can lead to losses for the other members of the system. In the financial literature, these issues are studied under the umbrella term of *systemic risk*, which a recent survey paper by Benoit et al. [8] defines as the risk that several market participants become simultaneously affected by severe losses that spread throughout the financial system.

While this notion of systemic risk has already had important implications for macroprudential policies (Benoit et al. [8], Duffie [49]) and central bank stress testing practices (Dees et al. [44]), there is still a pressing need for a better understanding of the dynamic feedback effects and amplification mechanisms that seem to have been the real drivers of the financial crisis. Indeed, it is widely accepted that—up to an order of magnitude—the extent of the crisis cannot be explained by simple references to (linear) exogenous shocks such as the devaluation of mortgage-backed securities (Cochrane [32], Brunnermeier [14], Hellwig [76]). Instead, smaller scale shocks are understood to have unfolded into a spiral of events rooted within the financial system itself, and amplified by a myriad of interactions between the individual institutions.

Accordingly, there have been judicious calls for a better understanding of the *endogenous* (nonlinear) nature of systemic risk (Pedersen [97], Danielsson et al. [39]),

and it has been emphasized that systemic risk is inherently *dynamic* with a gradual build-up typically taking place in the background before it materializes in a crisis (Brunnermeier et al. [15]).

1.1.1 Two streams of literature on systemic risk

A decade after the global financial crisis, there is by now a significant body of work dealing with the complex web of interactions in the financial system and the dual role of interconnectedness as a source of diversification or a channel for contagion. In terms of mathematical modelling, it is possible to identify two overarching approaches to the challenges of systemic risk.

Firstly, there is the large literature on *network-based* models for clearing and contagion, which extend the early frameworks of Eisenberg and Noe [51] and Allen and Gale [2]. For a comprehensive overview of this literature and its applications, we refer to the review by Glasserman and Young [67], the book by Hurd [78], and the empirical study of Cont et al. [34] (see also [3, 24, 99] and the references therein for various different points of view). While these network models are principally static, we note that a dynamic continuous time Eisenberg–Noe framework has recently been proposed in Banerjee et al. [5].

Next, there is the much smaller literature on dynamic continuous-time *mean field* models in the spirit of Carmona et al. [28] and Garnier et al. [62]. Important contributions to this literature can be found in [6, 23, 25, 58, 59, 63]. In general, these models benefit from a richer dynamic and stochastic structure, but they tend to focus on quite simple interbank interactions that neglect defaults and fail to incorporate financial contagion. A related line of literature is focused on dynamic intensity-based models for large portfolio credit risk [38, 40, 64, 65, 66, 102], where implicit notions of default contagion are incorporated through interacting intensities.

The models proposed in this thesis can be seen as a natural continuation of the mean field literature. However, in contrast to the existing models, we develop a more flexible framework that incorporates default contagion endogenously via a structural mechanism for defaults. This form of contagion has not previously been considered in the mean field literature. Moreover, our structural approach to contagion differs markedly from the related reduced form literature (for large portfolio credit risk), where contagion comes from self-exciting point processes, and it has the additional benefit of being conceptually close to the network based approaches.

1.2 Main objectives of the thesis

As emphasized in the abstract, the main contributions of this thesis consist of a series of novel mathematical results on the well-posedness of contagious McKean–Vlasov systems with positive feedback. The nature of these contributions is briefly explained in Section 1.3 and we then give an outline of the contents of the thesis in Section 1.4.

In addition to the mathematical results, we believe it is both interesting and important to study systemic risk in large financial markets from a 'macroscopic' (or 'system-wide') point of view—in the sense of working with large population limits via McKean–Vlasov or mean field theory. In short, this provides a low parameter approach to modelling the financial system as a whole, whereby one can hope to capture (or at least highlight) the most important mechanisms at work, while avoiding the need to calibrate complex high-dimensional systems. Naturally, much work is needed to produce more realistic mean field models of contagion in large financial systems, however, we believe it is another key contribution of this thesis to have started off this analysis (which we stress has occurred in tandem with work by S. Nadtochiy and M. Shkolnikov as mentioned in Section 1.4.1, where we discuss related developments in the mathematical literature).

1.2.1 Macro-modelling: a system-wide perspective risk

Consider a 'representative' member of a large financial system and let $X_t \in [0, \infty)$ denote a (stochastic) measure of its financial health at time t, as summarised by a non-negative number which we call its *distance-to-default*. Neglecting system-wide effects, it is a classical approach in structural credit risk theory to simply model X_t in isolation by a Brownian motion with drift—where we say that default occurs at the first hitting time of zero. If 'isolation' is taken to mean that all the actors are assumed independent, then this means that the overall health of the financial system is effectively described by a linear (deterministic) heat equation.

The motivation for this thesis is to introduce *systemic risk* into this picture by instead proposing a mean field model derived from an interacting particle system that incorporates simple notions of (i) common exposures, (ii) herd behaviour, and (iii) endogenous contagion. As we will see in Chapter 3, this implies that the health of the financial system will now be described by a mean field type stochastic partial differential equation (SPDE) on the positive half-line, which is driven by the noise of the common exposures and has nonlocal nonlinearities coming from the system-wide interactions (see Theorems 3.2.4 and 3.2.6). Furthermore, in a suitable sense, we can

translate this as saying that the dynamics of a 'representative' financial institution is no longer just a Brownian motion with drift but rather a conditional McKean–Vlasov type diffusion with dependence on the conditional law of its paths given the noise of the common exposures (see Theorem 3.2.7).

The model we propose in Section 3 incorporates a *gradual* notion of default contagion, modulated by an *impact kernel*, which controls how fast the feedback from defaults is felt by other members of the system. Mathematically, this can be seen as introducing a dampening effect on the positive feedback arising from the nonlinearity of the SPDE, which in turn allows for a global description of its behaviour. If, on the other hand, we allow the feedback from defaults to become instantaneous (meaning that the aforementioned impact kernel degenerates to a dirac mass at 0), then the nonlinearity can give rise to singularities whereby the global SPDE description of the mean field limit ceases to exist. For this reason, it becomes necessary to focus instead on the McKean–Vlasov description of the limit, which is the subject of Chapters 4 through 6. In Section 1.3 below, we provide a brief overview of the formulation of the main mathematical problems, and in Section 1.3.2 we discuss the model with singularities in the simpler—but still rather intricate—case where there is no common source of noise.

1.2.2 Remarks on related models of default contagion

As already mentioned, there has been a surge of academic interest in systemic risk during the ten years that have passed since the global financial crisis—both in the economics literature and more recently within the mathematical finance community.

Building on the original Eisenberg–Noe framework [51], Rogers and Veraart [99] provides a rigorous framework for studying the ways in which costly defaults (with recovery rates being less than unity) can lead to default cascades in interbank networks. A very influential and closely related—but ultimately heuristic—model for default cascades (with vanishing recovery rates) was offered by Gai and Kapadia [61] already in the immediate aftermath of the global financial crisis. Guided by a different random graph approach, another influential framework for default contagion is that of Amini, Cont, and Minca [3], which gives simple analytic expressions for the asymptotic size of default cascades in interbank networks as the number of banks tends to infinity. Finally, it is worth emphasising that there is a much longer history of seeking to model contagion in credit risk theory, as for example evidenced by the influential one-period model of Davis and Lo [41]. For an overview of this earlier credit risk literature, we refer to the review by Davis [42].

All of the above frameworks differ substantially from the one we propose in this thesis, although it is worth pointing out that we are interested in large population limits of interacting systems just as [3] focuses on asymptotic results for random graphs. However, unlike [3], our limiting equations have intricate nonlinearities that do not allow for simple analytic conclusions, so their application to systemic risk must rely on numerical algorithms (one such algorithm is proposed in Section 5.3.4). Moreover, a key difference is that the models we propose here are dynamically evolving in continuous time—with the defaults and appertaining contagion arising endogenously from these dynamics.

Out of the existing literature on systemic risk, two contributions can be singled out as being particularly close in spirit to our framework. First of all, Battiston et al. [7] provides a heuristic analysis of an interacting system of stochastic differential equations living in $[0, 1]$, which model the robustness of a finite system of banks and closely resembles the finite dimensional system from Chapter 4 of this thesis. Similarly, Lipton [90] provides a rigorous analysis—also in the finite dimensional setting—of a system of interacting distances-to-default that is essentially analogous to the notion of contagion derived in Chapter 4. For a simple illustration of the connection between Lipton's model and the contents of this thesis, we refer to the recent paper by Lipton et al. [82], which considers a numerical algorithm for a mean field version of the model in [90] based in part on the results from Chapter 4. In fact, we only became aware of the papers [7] and [82] after the completion of Chapters 3 and 4, thanks to M. Shkolnikov and V. Kaushansky, respectively, and there are certainly significant differences between the frameworks (including our impact kernel and our focus on various mean field type interactions along with the corresponding McKean–Vlasov analysis).

In light of the above, the framework developed in this thesis provides a complimentary perspective on the models from [7, 82] and, most importantly, it introduces a rigorous analysis of the mean field limits for these families of systemic risk models. Interestingly, J.-P. Bouchaud made us aware that Gualdi et al. [69] has applied a variant of the mean field limits we obtain in order to analyse endogenous macroeconomic shocks (directly at the 'macro' or 'mean field' level). This analysis, however, never addresses the well-posedness of the model nor the emergence as a mean field limit of a finite system. Therefore, an important contribution of this thesis is to provide rigorous justification for this type of modelling within the realm of contagious McKean–Vlasov systems with positive feedback.

1.2.3 Large portfolio credit risk with default clustering

Although the focus of this thesis is on systems of financial institutions, our framework can also be applied to the study of default clustering in large portfolios of more general defaultable entities. That is, the models we propose in Chapters 3–6 could serve as *structural* large portfolio models for the pricing of credit derivatives, extending the earlier frameworks of [16, 73].

Recently, there has been a great deal of interest in disentangling the roles of contagion and common risk factors as drivers of corporate default clustering (Azizpour et al. [4], Lando and Nielsen [84], Duffie et al. [50]). However, the literature remains inconclusive and has focused almost exclusively on self-exciting point processes, so the model we present here could serve as a first step towards a structural platform from which to approach these matters.

1.3 Overview of mathematical contributions

In this section we briefly introduce the precise formulation of the main mathematical problems studied in this thesis. Our starting point is the SPDE model for systemic risk discussed in Section 1.2.1 above (which is the subject of Chapters 2 and 3). Next, we consider a limiting case of this problem that arises when the gradual impact of contagion becomes instantaneous. In this latter situation, the SPDE point of view is not always tractable and one is therefore led to focus more on the associated McKean–Vlasov formulation of the problem (which is the subject of Chapters 4 through 6).

1.3.1 The SPDE problem with gradual feedback

Recall the model described in Section 1.2.1 above. In Chapter 3 we show that this model leads to a 'macro-level' (or 'systemic') description of the financial system through a nonlinear SPDE on the positive half-line that models the density $V_t(x)$ of banks, at time t, with a given financial healthiness x (or, more precisely, a given distance-to-default $x > 0$, where $x = 0$ corresponds to default).

In the most general setting, the SPDE problem for the (sub-probability) density V_t is of the form

$$\begin{aligned} dV_t(x) = {} & \tfrac{1}{2}\partial_{xx}\big(\sigma(t,x)^2 V_t(x)\big)dt - \partial_x\big(b(t,x,\nu_t)V_t(x)\big)dt \\ & - \rho(t,\nu_t)\partial_x\big(\sigma(t,x)V_t(x)\big)dB_t^0 + \partial_x\big(\alpha(t,x,\nu_t)V_t(x)\big)d\mathfrak{L}_t, \quad V_t(0) = 0, \end{aligned}$$

with $d\nu_t(x) = V_t(x)dx$ and $\mathfrak{L}_t = (\mathfrak{K} \star L)_t$, where $L_t = 1 - \int_0^\infty V_t(x)dx$. Here L captures the loss of mass in the system (i.e., the proportion of defaults), which feeds back into the equation for V_t, and B^0 is a Brownian motion modelling the common exposures. Moreover, $\mathfrak{K}$ is the earlier mentioned *impact kernel* that modulates the gradual impact of contagion. Notice that the contagion emerges as a nonlinear transport term proportional to the derivative of $\mathfrak{L}_t$, which—as an integration by parts shows—is formally given by the flux across the Dirichlet boundary at zero, smoothed in time according to $\mathfrak{K}$, that is,

$$\frac{d}{dt}\mathfrak{L}_t = \frac{1}{2}\int_0^t \mathfrak{K}(t-s)\sigma(s,0)\partial_x V_s(0)ds.$$

Our first main contribution of this thesis is to establish the well-posedness of this nonstandard SPDE problem, and to show that it arises as the unique mean field limit of a finite model for the financial system (as briefly described in Section 1.2.1 above). These results are presented in Chapter 3, building on the work in Chapter 2. The details of the mathematical innovations and the connections to antecedent literature are discussed at length in Section 3.2.

1.3.2 A McKean–Vlasov problem with instantaneous feedback and blow-ups

Similarly to the model outlined in Section 1.2.1, suppose the distances-to-default of each financial institution (in the finite system) are given by independent Brownian motions with a positive feedback effect from defaults that *instantaneously* shifts all the other institutions closer to the default barrier—according to some parameter $\alpha > 0$ capturing the strength of the feedback. In the large population limit, this leads to a McKean–Vlasov problem for a 'representative' distance-to-default, X, given by

$$\begin{cases} X_t = X_0 + B_t - \alpha L_t \\ \tau = \inf\{t \geq 0 : X_t \leq 0\} \\ L_t = \mathbb{P}(\tau \leq t), \end{cases} \tag{MV}$$

where B is standard Brownian motion and $X_0 > 0$ is a random start point independent of B. We analyse this problem (and variants) in Chapters 4 through 6, where we provide novel results on well-posedness along with important regularity results for the loss process L and a suggested numerical algorithm for simulating the mean-field limit directly. For now, however, we simply wish to give the reader a feel for the problem (MV), by briefly showing what it looks like as a PDE problem—although

it turns out that one needs to be quite careful about how to interpret this point of view. Either way, it allows to briefly highlight how the problem connects to earlier contributions in the PDE literature for closely related problems.

1.3.2.1 PDE viewpoint: the supercooled Stefan problem

Given a solution X to (MV), we define, for each time $t > 0$, the sub-probability measure $\nu_t = \mathbb{P}(X_t \in \cdot, t < \tau)$, and note that this gives the law of X_t with an absorbing boundary at the origin. Denoting by V_t the density of ν_t, we have $\int_0^\infty \phi(x)V_t(x)dx = \mathbb{E}[\phi(X_t)\mathbf{1}_{t<\tau}]$. From here, it can be deduced from an application of Itô's formula and integration by parts (for suitable test functions ϕ) that V_t is formally a solution of the Cauchy–Dirichlet problem

$$\partial_t V_t(x) = \tfrac{1}{2}\partial_{xx}V_t(x) + \alpha L_t'\partial_x V_t(x), \qquad L_t' = \tfrac{1}{2}\partial_x V_t(0), \qquad V_t(0) = 0, \tag{1.1}$$

for x in the positive half-line, with initial condition V_0. In this PDE formulation, we can see that the contagious feedback from defaults emerges as a nonlinear (and nonlocal) transport term that pushes mass towards the origin at a rate proportional to the current flux of defaulting banks crossing the origin.

Setting $v(t,x) := -\alpha V_t(x - \alpha L_t)$, for $\alpha > 0$, the equations for V and L become

$$\partial_t v = \tfrac{1}{2}\partial_{xx}v \text{ on } (\alpha L_t, \infty), \qquad \partial_x v(t, \alpha L_t) = -\tfrac{1}{2}\alpha L_t', \qquad v(t, \alpha L_t) = 0. \tag{1.2}$$

This is a Stefan problem modelling the freezing of a supercooled liquid occupying the semi-infinite strip $(\alpha L_t, \infty)$: the liquid is initially supercooled to a temperature $v(0,x) = -\alpha V_0(x)$ that is below the freezing point $v = 0$, and the evolving 'freezing front' is given by $x = \alpha L_t$. Forgetting for a moment that V_0 is a probability density, if $v(0,\cdot) = -c$, then the well-posedness of this system displays a clear dichotomy: for $c < 1$ it admits an explicit similarity solution, while no solution can exist for $c \geq 1$. This situation motivates the recent analysis of Dembo and Tsai [48], which investigates the critical case $c = 1$ via the scaling behaviour of a discretised particle approximation.

A related line of study comes from the work of Calvez et al. [18, 19], which considers the PDE

$$\partial_t u(t,x) = \tfrac{1}{2}\partial_{xx}u(t,x) + u(t,0)\partial_x u(t,x), \qquad \partial_x u(t,0) = -u(t,0)^2, \tag{1.3}$$

for $x \in (0,\infty)$, as a model for cell polarisation: u is the density of molecular markers on a cell, identified with the positive half-line, and the time evolution of the density

is coupled with the concentration of markers on the cell membrane at $x = 0$. If we set $\tilde{V}_t(x) := \alpha^{-1}\int_0^x u(t,y)dy$, then we get back (1.1) with initial condition $\tilde{V}_0(x) = \alpha^{-1}\int_0^x u_0(y)dy$. Calvez et al. [18, 19] show that (1.3) admits a global weak solution if $\int_0^\infty u_0(y)dy \leq 1$, while u explodes to infinity in finite time if $\int_0^\infty u_0(y)dy > 1$ with u_0 non-increasing (see also Lepoutrea et al. [89] for some refinements).

Note that in both of the above examples, the solvability depends critically on how the initial condition compares to α^{-1}. This same relationship will play an important role in the results of the next sections. Returning for now to the supercooled Stefan problem, there is an early literature on its well-posedness for slight variations of (1.2) on a finite strip. This goes back to Fasano and Primicerio [54, 55], who give conditions on the initial datum under which the system is uniquely solvable in the class of classical solutions—either for all times or up to some finite time where L' explodes and beyond which the solution is understood to have ceased to exist.

If a source term $L_t'\delta_0(x - c)$ is added to (1.1), for some $c > 0$, where δ_0 is the Dirac delta function, then the mass of the system is preserved. Recalling the particle system briefly discussed in Section 1.2.1 above, this corresponds to the case where the particles are instantly reset to a predetermined value c upon reaching the absorbing boundary instead of being taken out of the system. This form of the particle system is motivated by mathematical neuroscience, where it arises from networked integrate-and-fire models for the 'firing' of electrically coupled neurons (see e.g. Brunel et al. for a heuristic derivation [96]). By reducing the analysis to that of a Stefan-like problem, Carillo et al. [29] show that, for $\alpha \leq 0$, a unique classical solution exists for all times, while, for $\alpha > 0$, classical solutions exist up to a possibly finite explosion time (for initial conditions that are C^1 up to the boundary and vanish there). For further background, see also [17, 30].

At this point, it is important to emphasise that—in contrast with the PDE solutions—the limiting solutions to (MV), which we will construct in Chapter 5 (see also [46, 95]), are global in time: they do not cease to exist at some explosion time, even despite the fact that there must be blow-ups in the form of jump discontinuities for large enough $\alpha > 0$, as we show in Theorem 4.2.1.

1.4 Outline of the thesis

The contents of this thesis form two connected components, deriving from the papers [74, 75, 87, 88] written during my DPhil studies in Oxford. Taken together, Chapters 2 and 3 form the first component, containing the results from [75] which is a joint

work with my advisor Ben Hambly. The second component consists of the remaining Chapters 4–6: Chapter 4 corresponds to [74], which is a joint work with Ben Hambly and Sean Ledger, while Chapter 5 is based on the paper [87] and Chapter 6 is based on the paper [88] both of which are co-authored with Sean Ledger.

Throughout the thesis, our main motivation—and reference point—is the mean field model for systemic risk that we propose and analyse in Chapter 3. The model is derived as the large population limit of a system of interacting particles on the positive half-line, where each particle represents the distance-to-default of a given financial institution, and absorption at the origin corresponds to default. The two key features of the model are intended to capture the interplay between common exposures and default contagion through (i) direct correlation via a common source of noise, and (ii) nonlinear interaction via an endogenous notion of contagion whereby defaults shift the distances-to-default closer to zero.

The purpose of Chapter 3 is to characterise, and then analyse, the resulting large population limit, which is given by a stochastic partial differential equation (SPDE) on the positive half-line with an absorbing boundary at the origin. As we will see, there is a delicate balance between the realisations of the common noise driving this SPDE and the nonlinearity from the contagion mechanism, which can cause the loss of mass to accelerate rapidly, corresponding to a period of systemic default clustering.

As part of our analysis of the model in Chapter 3, we rely on a family of Dirichlet heat kernel type estimates for the densities of absorbed diffusions on the positive half-line. The derivation of these estimates is the subject of Chapter 2. The main novelty is that the estimates allow for drifts of up to linear growth and, in addition, our probabilistic approach is robust to drifts which may depend on randomness outside the realm of just the current state of the diffusion.

In Chapters 4, 5, and 6, we study a limiting case of the model from Chapter 3, where defaults have an instantaneous feedback effect on the rest of the system. This makes the analysis of the mean field limit quite intricate, as contagious blow-ups can now arise from the aforementioned rapid accelerations in the loss of mass. Nevertheless, despite the theoretical challenges, such blow-ups are very interesting to understand, since they can be seen as providing a means of characterizing breakdowns of the financial system.

In Chapter 4 we analyse the mean field limit when there is no common noise. Using a fixed-point argument, we construct a differentiable solution of the resulting McKean–Vlasov problem up to a first explosion time. Our main contribution is a proof of uniqueness (up to a unique explosion time) in the class of càdlàg functions, for which

a novel minimality result and trapping technique are introduced. Furthermore, we allow initial conditions to include densities with any power law decay at the boundary, and we connect the exponent of this decay with the regularity of the solution in small time in a precise way.

In Chapter 5 we extend the model with instantaneous feedback to include a common source of noise driven by Brownian motion. In this case, the mean field limit is now given by a *conditional* McKean–Vlasov problem. We prove existence of solutions to this problem, by realising them as suitable limit points of the finite-dimensional particle system, and we show that the pathwise realisation of the common noise can both trigger and prevent blow-ups.

Finally, in Chapter 6 we present a global uniqueness argument for the conditional McKean–Vlasov problem in the weak feedback regime. By weak feedback we mean the case where the contagion parameters are small enough to prevent blow-ups in solutions. Next, we show how the same techniques can be extended to give short-time uniqueness after blow-ups, regardless of the feedback strength, in the model without a common noise. The heart of our approach is a surprisingly simple probabilistic comparison argument that is robust in the sense that it does not ask for any regularity of the solutions. By combining this with the results from Chapter 4, we end the thesis by giving a fairly complete description of the global behaviour of solutions with blow-ups in the case where there is no common noise—alas, we are only able to prove this under certain restrictions on the evolution of the system.

1.4.1 Remarks on related mathematical literature

Many important ideas in Chapter 3 can be traced back to developments in the recent work of Hambly and Ledger [73], and the analysis in Chapters 4–6 follows on from the foundational work of Delarue et al. [45, 46]. Our approach to well-posedness and regularity in Chapters 4 and 6 is entirely separate from that of [45, 46], however, we exploit the notion of a 'physical' solution which was first introduced in [46]. Also, the weak convergence analysis of Chapter 5 (for the large population mean field limit) builds on several important ideas first developed in [46]. We discuss in detail how our contributions relate to these earlier works in the respective chapters.

Moreover, we stress that there has been a sequence of exciting new developments by Nadtochiy and Shkolnikov [94, 95] and, even more recently, by Delarue, Nadtochiy, and Shkolnikov [47]. These contributions have appeared concurrently with the work that has gone into this thesis. The financial models we propose and the appertaining mathematical results were all developed independently of the efforts of Nadtochiy,

Shkolnikov, and Delarue. As we detail in the respective chapters, the combined set of results arising from these papers along with this thesis complement each other nicely. Furthermore, we note that Campi and Fischer [20] and Campi et al. [21] have recently started a related line of literature on finite stochastic differential games and the associated mean field games in the presence of absorbing boundaries.

Finally, we recall that—as briefly mentioned in Section 1.3.2.1 above—the models we study are closely linked to the 'supercooled' Stefan problem. In that sense, our results can be seen as generalizing the forerunning PDE literature on this problem through a probabilistic study of the true nature of the blow-ups (with the results culminating in the complete picture provided by [47]). In addition, we note that our models with a common noise can also be seen as generalising the corresponding stochastic version of the 'supercooled' Stefan problem (with a Brownian transport term or, equivalently, Brownian noise in the moving boundary). Recently, there has been a surge of interest in stochastic Stefan type problems motivated largely by the modelling of limit order books in financial markets, see e.g. Cont and Müller [33], Müller [93], and Hambly and Kalsi [71, 72].

Chapter 2

Upper Dirichlet Heat Kernel Type Estimates for Drifts of at most Linear Growth

This chapter studies the boundary and tail behaviour of the individual transition densities for a quite general system of interacting real-valued SDEs, where each individual SDE (sometimes referred to as a 'particle') is started on the positive half-line and absorbed when it first hits the origin. The resulting estimates will form a key part of the analysis in Chapter 3 and, in particular, they lead to the absorbing density bounds for conditional McKean–Vlasov diffusions presented in Theorem 3.2.7.

2.1 Main results

The main result of this chapter is Theorem 2.1.1 below. The density bounds contained therein are akin to the classical Dirichlet heat kernel estimates which are available under stronger growth assumptions on the drift (see e.g. [70, 31]). As far as we aware, there are no results in the literature on Dirichlet heat kernel estimates when the drift is allowed to have linear growth on an unbounded domain. Consequently, we believe the results of this chapter could be of independent interest—aside from their use in Chapters 3 and 4 of this thesis.

$$d\mathbf{X}_t^N = \boldsymbol{b}(t, \mathbf{X}_t^N)dt + \boldsymbol{\sigma}(t, \mathbf{X}_t^N)\big(\rho(t, \mathbf{X}_t^N)d\mathbf{B}_t^0 + \sqrt{1-\rho(t, \mathbf{X}_t^N)^2}d\mathbf{B}_t\big), \tag{2.1}$$

where $\mathbf{B}^0 = (B^0, \dots, B^0)^\top$ and $\mathbf{B} = (B^1, \dots, B^N)^\top$, for a family of independent Brownian motions $B^0, B^1, \dots, B^N$. Writing $\boldsymbol{b}_t = \boldsymbol{b}(t, \mathbf{X}_t^N)$ and $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}(t, \mathbf{X}_t^N)$, we also

assume that

$$\boldsymbol{b}_t = \begin{bmatrix} b^1\big(t, X_t^1, \sum_{i=1}^N a_i^N X_t^i, \omega\big) \\ \vdots \\ b^N\big(t, X_t^N, \sum_{i=1}^N a_i^N X_t^i, \omega\big) \end{bmatrix} \quad \text{and} \quad \boldsymbol{\sigma}_t = \begin{bmatrix} \sigma^1\big(t, X_t^1\big) & & 0 \\ & \ddots & \\ 0 & & \sigma^N\big(t, X_t^N\big) \end{bmatrix},$$

with uniform ellipticity of $\boldsymbol{\sigma}$ and a uniform linear growth bound

$$\big|b^i\big(t, X_t^i, M_t^N, \omega\big)\big| \le C\big(1 + |X_t^i| + |M_t^N|\big), \quad M_t^N := \sum_{i=1}^N a_j^N X_t^j, \quad \sum_{j=1}^N a_j^N = 1.$$

The particular structure of the coefficients is motivated by the model we study in Chapter 3. The precise form of $\boldsymbol{b}$ is not essential (we just need some linear growth bound), but we do make use of the fact that the matrix $\boldsymbol{\sigma}$ is diagonal, as it allows us to conveniently scale it to unity for each individual diffusion. The specific growth bound—in terms of only one individual particle and a weighted average of the other particles—is needed to ensure that the density bounds are uniform in the size, N, of the system, which is crucial to our analysis in Chapter 3.

Given a solution to the system (2.1) on an interval $[0, T]$ we denote by X^i the i'th coordinate process. The first hitting time of the origin for this individual particle is defined by $\tau_i = \inf\{t > 0 : X_t^i \le 0\}$. We have the following result.

Theorem 2.1.1. *Suppose the system (2.1) has a strong solution and that the structural conditions listed in Assumption 2.1.2 below are all satisfied. Then the absorbed process $(X_t^i, t < \tau_i)$ has a density $\mathfrak{p}_t^{i,N}(x, y)$ conditional on $X_0^i = y$, and this density satisfies the following uniform bounds. For every $\epsilon > 0$ there exist $\kappa \in (0, 1)$ and $C, c > 0$, uniformly in $N \ge 1$ and $t \in (0, T]$, such that, for all $x, y \in (0, \infty)$,*

$$\mathfrak{p}_t^{i,N}(x, y) \le C\bigg(\frac{1}{\sqrt{t}}\Big(\frac{x}{\sqrt{t}} \wedge 1\Big)\Big(\frac{y}{\sqrt{t}} \wedge 1\Big) + \Big(\frac{x^\kappa y^\kappa}{t^\kappa} \wedge 1\Big)e^{\epsilon y^2}\bigg)\big(1 \wedge e^{-\frac{(x-y)^2}{ct} + c_{x,y}}\big) \tag{2.2}$$

and

$$\mathfrak{p}_t^{i,N}(x, y) \le C\big(t^{-\frac{1}{2}} + e^{\epsilon y^2}\big)e^{-\frac{(x-y)^2}{ct}}, \tag{2.3}$$

where $c_{x,y} \lesssim |x - y|(x \wedge y)$. If $\sigma(t, x) = \sigma_1(t)\sigma_2(x)$, then we can take $c_{x,y} \equiv 0$.

The proof of the above theorem is the subject of Sections 2.2 through 2.4, culminating in Section 2.4.3. In addition to the assumption that a strong solution exists, the following structural conditions are critical to the proof.

Assumption 2.1.2. First of all, we assume that the coefficients $\boldsymbol{b}$ and $\boldsymbol{\sigma}$ take the form outlined immediately after (2.1). Furthermore, we assume the following:

i. *Linear growth of the drift:* Uniformly in $i = 1, \dots, N$, $t \in [0, T]$, and $\omega \in \Omega$, there exists $C > 0$ such that $|b(t, x, m, \omega)| \leq C(1 + |x| + |m|)$ for all $x, m \in \mathbb{R}$.

ii. *Regularity of the volatility*: Uniformly in $i = 1, \dots, N$ and $t \in [0, T]$, there exist $C, c > 0$ such that $c \leq \sigma^i(t, x) < C$ for all $x \in \mathbb{R}$. Furthermore, $(t, x) \mapsto \sigma(t, x)$ is $\mathcal{C}^{1,1}([0, T] \times \mathbb{R})$ with uniform bounds $|\partial_t \sigma^i(t, x)| \leq C$ and $|\partial_x \sigma^i(t, x)| \leq C$ for all $x \in \mathbb{R}$. If $\sigma(t, x) = \sigma_1(t)\sigma_2(x)$, then no smoothness in time is needed.

iii. *Sub-Gaussian initial law:* The random start points $X_0^1, \dots, X_0^N$ are i.i.d and independent of the Brownian motions. Moreover, there exists $\epsilon > 0$ such that

$$\mathbb{P}\big(X_0^1 \in (\lambda, \infty)\big) = O\big(\exp(-\epsilon\lambda^2)\big) \quad \text{as} \quad \lambda \to \infty.$$

2.1.1 Outline of the chapter

The purpose of this section is to prove the density estimates stated in Theorem 2.1.1. Our approach will rely on techniques that are entirely probabilistic in nature and we follow a simple intuitive procedure.

First of all, we provide a sub-Gaussian estimate in Section 2.3, which shows how the sub-Gaussianity of the initial law propagates nicely for all positive times. This, in particular, ensures sufficiently fast decay of the tails of each SDE. In Section 2.4, we then transform the SDEs into Brownian motions with drift that have the same hitting times of the origin, and then we use the sub-Gaussianity from the previous section to introduce a change of measure related to the drift.

Finally, we derive the density estimates in Section 2.5, by comparing any given transformed particle at time $t > 0$ with an independent absorbed Brownian motion (under the original measure) started from the transformed particle's position at earlier times $s < t$ and run for the remaining time $t - s$.

Before embarking on this journey, we devote Section 2.2 to a brief discussion of some classical results in the benchmark case of just one absorbed Brownian motion.

2.2 The Dirichlet Heat Kernel and Boundary Estimates for Absorbed Brownian Motion

To set the stage for the rest of the thesis, this section explores the boundary behaviour of an absorbed Brownian motion started on the positive half-line and killed when it first hits zero. In particular, we are interested in the decay of its density near the origin and at infinity. Moreover, we will give a brief exposition of its connection to the Cauchy–Dirichlet problem for the heat equation $\partial_t v = \frac{1}{2}\partial_{xx} v$ on $[0, \infty) \times \mathbb{R}_+$.

2.2.1 Absorbed Brownian Motion

Let $X_t := X_0 + B_t$ denote a one dimensional Brownian motion with random start point $X_0 > 0$ (independent of B), and define its first hitting time of zero as

$$\tau := \inf\{t > 0 : X_t \leq 0\}.$$

This gives rise to an *absorbed* Brownian motion $(X_t, t < \tau)$ whose marginal laws are understood to be given by the family of sub-probability measures ν_t, where

$$\nu_t(A) := \mathbb{P}(X_t \in A\,,\, t < \tau) \quad \forall A \in \mathcal{B}(\mathbb{R})$$

with $\nu_0 = \text{Law}(X_0)$. By means of the strong Markov property of Brownian motion, we can derive an explicit expression for ν_t from first principles. Conditioning on the random start point, we have

$$\begin{aligned}\nu_t(a,b) &= \int_0^\infty \mathbb{P}\big(y + B_t \in (a,b)\,,\, t < \tau_y\big) d\nu_0(y) \\ &= \int_0^\infty \mathbb{P}\big(y + B_t \in (a,b))d\nu_0(y) - \int_0^\infty \mathbb{P}\big(y + B_t \in (a,b)\,,\, t \geq \tau_y\big) d\nu_0(y)\end{aligned}$$

where $\tau_y = \inf\{t > 0 : y + B_t \leq 0\}$. In words, we need to subtract of the probability that the Brownian motion belongs to (a,b) at time t but crossed the origin on its journey. Stopped at the point where it hits zero, the restarted process is again a Brownian motion with equal probability of travelling to (a,b) or $(-b,-a)$. Thus, the latter of the above probabilities should simply equal the probability of a Brownian motion started at y travelling to $(-b,-a)$. To make this precise, notice that we obviously have $B_{\tau_y} = -y$, and hence

$$\begin{aligned}\mathbb{P}\big(y + B_t \in (a,b)\,,\, t \geq \tau_y\big) &= \mathbb{P}\big(B_t - B_{\tau_y} \in (a,b)\,,\, t \geq \tau_y\big) \\ &= \mathbb{E}\big[\mathbf{1}_{t \geq \tau_y} \mathbb{P}\big(B_t - B_{\tau_y} \in (a,b) \mid \mathcal{F}^B_{\tau_y}\big)\big].\end{aligned}$$

By the strong Markov property, $B_t - B_{\tau_y}$ is Gaussian with mean 0 and variance $t - \tau_y$ conditional on τ_y. Since this distribution is symmetric around zero, we deduce that

$$\begin{aligned}\mathbb{P}\big(y + B_t \in (a,b)\,,\, t \geq \tau_y\big) &= \mathbb{E}\big[\mathbf{1}_{t \geq \tau_y} \mathbb{P}\big(B_t - B_{\tau_y} \in (-b,-a) \mid \mathcal{F}^B_{\tau_y}\big)\big] \\ &= \mathbb{P}\big(B_t - B_{\tau_y} \in (-b,-a)\,,\, t \geq \tau_y\big) \\ &= \mathbb{P}\big(y + B_t \in (-b,-a)\big),\end{aligned}$$

where we have noted that $t \geq \tau_y$ is implied by $y + B_t \in (-b, -a)$. Exploiting again the symmetry of B, it follows that

$$\begin{aligned}\nu_t(a,b) &= \int_0^\infty \mathbb{P}\big(y + B_t \in (a,b)\big) d\nu_0(y) - \int_0^\infty \mathbb{P}\big(-y + B_t \in (a,b)\big) d\nu_0(y) \\ &= \int_a^b \int_0^\infty G_t(x,y) d\nu_0(y) dx,\end{aligned}$$

where G is the *Dirichlet heat kernel* defined by

$$G_t(x,y) := \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t} - \frac{1}{\sqrt{2\pi t}} e^{-(x+y)^2/2t}. \tag{2.4}$$

In particular, we have shown that, for every $t > 0$, the (sub-probability) measure ν_t has a density V_t, which takes the form

$$V_t(x) = \int_0^\infty G_t(x, x_0) d\nu_0(x_0).$$

2.2.2 The heat equation on a half-line

Let G denote the Dirichlet heat kernel (2.4) and let V be as above. Suppose ν_0 has a density $V_0 \in L^1(\mathbb{R}_+)$ with $|V_0(x)| \leq Ce^{cx^2}$ for some $C, c > 0$. Then it is easily verified [22, Section 3.3] that the function

$$V_t(x) = \int_0^\infty G_t(x, x_0) V_0(x_0) dx_0, \tag{2.5}$$

is infinitely differentiable in the quarter-plane $(0, \infty) \times \mathbb{R}_+$ and satisfies

$$\partial_t V_t(x) = \frac{1}{2} \partial_{xx} V(t,x) \quad \text{for all} \quad (t,x) \in (0,\infty) \times \mathbb{R}_+.$$

Clearly, we also have the Dirichlet boundary condition $V_t(0) = 0$, for all $t > 0$, since $G_t(0, x_0) = 0$, and it is fairly straightforward to show (see e.g. [22, Section 3.4]) that V_t respects the initial condition in the sense of being continuous at $t = 0$. More specifically, we have $\lim_{t \downarrow 0} V_t(x) = V_0(x)$ at all continuity points of V_0, with the convergence being uniform on compact subsets of intervals of continuity. In summary, the function $(t,x) \mapsto V_t(x)$ is a classical (smooth) solution of the Cauchy-Dirichlet problem

$$\begin{cases} \partial_t v(t,x) = \frac{1}{2} \partial_{xx} v(t,x) & \text{on} \quad (0,\infty) \times (0,\infty) \\ v(0,x) = V_0(x) & \text{on} \quad \{0\} \times (0,\infty) \\ v(t,0) = 0 & \text{on} \quad [0,\infty) \times \{0\}. \end{cases} \tag{2.6}$$

Under very mild conditions on V_0 (see e.g. [22, Section 4.4]), the solutions to this problem are unique and hence they coincide with (2.5).

As an alternative way of seeing that the absorbed Brownian motion corresponds to (2.6), one could use Itô's formula to deduce that ν_t satisfies a weak formulation of the PDE. We will see many examples of such reasoning later in this thesis.

2.2.3 Dirichlet Heat Kernel Estimate

When deriving the absorbing density estimate (2.2), our strategy essentially amounts to expressing a general (absorbing) density as a suitable perturbation of the classical Dirichlet heat kernel G_t. Indeed, the first part of (2.2) will be familiar from the following bound on the Dirichlet heat kernel, which we make use of in Section 2.4.

Lemma 2.2.1 (Dirichlet heat kernel estimate)**.** *There exists a constant $C > 0$ such that*

$$G_t(x,y) \leq \frac{C}{\sqrt{t}}\Big(\frac{x}{\sqrt{t}} \wedge 1\Big)\Big(\frac{y}{\sqrt{t}} \wedge 1\Big)e^{-\frac{(x-y)^2}{4t}} \tag{2.7}$$

for all $x, y \in \mathbb{R}_+$ at all positive times $t > 0$.

Proof. The properties of the exponential function gives that

$$G_t(x,y) = \frac{1}{\sqrt{2\pi t}}\big(1 - e^{\frac{2xy}{t}}\big)e^{-\frac{(x-y)^2}{2t}} \leq \frac{1}{\sqrt{2\pi t}}\Big(1 \wedge \frac{2xy}{t}\Big)e^{-\frac{(x-y)^2}{2t}}.$$

If $x, y \leq \sqrt{t}$ or $x, y \geq \sqrt{t}$, then it is immediate that

$$\Big(1 \wedge \frac{xy}{t}\Big) \leq \Big(\frac{x}{\sqrt{t}} \wedge 1\Big)\Big(\frac{y}{\sqrt{t}} \wedge 1\Big)$$

and hence the claim is true in these cases. By the symmetry in x and y, it only remains to consider the case where $x \leq \sqrt{t}$ and $y \geq \sqrt{t}$. To this end, we can observe that

$$\frac{y}{\sqrt{t}} = \frac{x}{\sqrt{t}} + \frac{y-x}{\sqrt{t}} \leq 1 + \Big(1 + \frac{(x-y)^2}{4t}\Big) \leq 1 + e^{\frac{(x-y)^2}{4t}},$$

for all $x, y \geq 0$, and hence

$$\Big(1 \wedge \frac{xy}{t}\Big)e^{-\frac{(x-y)^2}{2t}} = \frac{x}{\sqrt{t}}\frac{y}{\sqrt{t}}e^{-\frac{(x-y)^2}{4t}}e^{-\frac{(x-y)^2}{4t}} \leq \frac{x}{\sqrt{t}}\Big(e^{-\frac{(x-y)^2}{4t}} + 1\Big)e^{-\frac{(x-y)^2}{4t}}.$$

Consequently, if $x \leq \sqrt{t}$ and $y \geq \sqrt{t}$, then

$$\Big(1 \wedge \frac{xy}{t}\Big)e^{-\frac{(x-y)^2}{2t}} \leq \Big(1 + e^{-\frac{(x-y)^2}{4t}}\Big)\Big(\frac{x}{\sqrt{t}} \wedge 1\Big)\Big(\frac{y}{\sqrt{t}} \wedge 1\Big)e^{-\frac{(x-y)^2}{4t}},$$

which proves the claim. □

2.3 Sub-Gaussianity

For notational convenience, we define the dominating processes

$$\Lambda_t^{i,N} := |X_t^i| + \sum_{j=1}^N a_j^N |X_t^j|, \quad \Gamma_t^{i,N} := |X_t^i|^2 + \sum_{j=1}^N a_j^N |X_t^j|^2. \tag{2.8}$$

Notice that, by Assumption 2.1.2, we have

$$|X_t^i|, |M_t^N|, |b^i(t, X_t^i, M_t^N, \omega)| \lesssim 1 + \Lambda_t^{i,N},$$

where '$\lesssim$' denotes than an equality holds up to some fixed numerical constant.

Proposition 2.3.1. *Let Assumption 2.1.2 be satisfied. For every $\epsilon > 0$, there is a (smooth) decreasing function $t \mapsto \eta_t$ with $\eta_0 = \epsilon/2$ and $\eta_t > 0$ such that, for all $t > 0$ and $N \geq 1$,*

$$\mathbb{E}\left[e^{\eta_t \Gamma_t^{i,N}}\right] \leq e^{c\int_0^t \eta_s ds} \mathbb{E}\left[e^{\epsilon |X_0^i|^2}\right]^{\frac{1}{2}} \mathbb{E}\left[e^{\epsilon \sum_{j=1}^N a_j^N |X_0^j|^2}\right]^{\frac{1}{2}}. \tag{2.9}$$

In particular, X_t^i, M_t^N, and $\Lambda_t^{i,N}$ are all sub-Gaussian. Moreover, we stress that the estimate (2.9) is uniform in $N \geq 1$ as long as the last expectation (involving the weighted average of the initial conditions) is bounded uniformly.

Proof. Fix a strictly positive (deterministic) function $\eta \in \mathcal{C}^1(\mathbb{R}_+)$ with $\eta_0 = \epsilon/2$ and define $\Psi \in \mathcal{C}^{1,2}(\mathbb{R}_+ \times \mathbb{R}^N)$ by

$$\Psi(t, x) := e^{\eta_t (x_i^2 + \sum_j a_j^N x_j^2)}.$$

For simplicity of notation, we write Γ_t in place of $\Gamma_t^{i,N}$. Then $\Psi(t, X_t) = e^{\eta_t \Gamma_t}$, where we have defined $X_t := (X_t^1, \ldots, X_t^N)$. Introducing the stopping times

$$\tau_n := \inf\{t \geq 0 : |X_t| > n\} \quad \text{for} \quad n \geq 1,$$

and applying Itô's formula, we get

$$e^{\eta_{t\wedge\tau_n} \Gamma_{t\wedge\tau_n}} = e^{\eta_0 \Gamma_0} + \int_0^{t\wedge\tau_n} (\partial_t + \mathcal{L}) e^{\eta_s \Gamma_s} ds + \sum_{j=1}^N \int_0^{t\wedge\tau_n} \sigma_s^j \partial_{x_j} e^{\eta_s \Gamma_s} dB_s^j, \tag{2.10}$$

where $dB_s^j = \rho_t dW_t^0 + \sqrt{1-\rho_t^2} dW_t^j$ and

$$\mathcal{L}\Psi = \sum_{j=1}^N b_t^j \partial_{x_j} \Psi + \frac{1}{2} \sum_{j=1}^N (\sigma_t^j)^2 \partial_{x_j}^2 \Psi + \frac{1}{2} \sum_{k \neq j} \sigma_t^j \sigma_t^k \rho_t^2 \partial_{x_j x_k} \Psi.$$

Here we have used the short-hand notation $\rho_t = \rho(t, X_t)$, $\sigma_t^i = \sigma(t, X_t^i)$, and likewise for b_t^i. Computing the derivatives, we see that

$$\begin{aligned}(\partial_t + \mathcal{L})e^{\eta_t\Gamma_t} = e^{\eta_t\Gamma_t}\Big(&\dot\eta_t\Gamma_t + 2\eta_t\sum_{j=1}^N b_t^j(\delta_{ij}+a_j^N)X_t^j + \\ &+\ \eta_t\sum_{j=1}^N(\sigma_t^j)^2(\delta_{ij}+a_j) + 2\eta_t^2\sum_{j=1}^N(\sigma_t^j)^2(\delta_{ij}+a_j^N)^2(X_t^j)^2 \\ &+\ 2\eta_t^2\rho_t^2\sum_{k\neq j}\sigma_t^j\sigma_t^k(\delta_{ij}+a_j^N)(\delta_{ik}+a_k^N)X_t^jX_t^k\Big).\end{aligned}$$

Using the bounds $|\sigma^j| \le C_\sigma$ and $|b^j| \le C_b(1 + |X_t^j| + \sum_\ell a_\ell^N|X_t^\ell|)$ as well as the basic inequalities

$$x \le 1 + x^2, \quad (x+y)^2 \le 2(x^2+y^2), \quad \text{and} \quad \big(\textstyle\sum_{j=1}^N a_j^N x_j\big)^2 \le \sum_{j=1}^N a_j^N x_j^2,$$

one then easily verifies that

$$(\partial_t + \mathcal{L})e^{\eta_t\Gamma_t} \le \Big(2\eta_t(C_b + C_\sigma^2) + (\dot\eta_t + 10C_b\eta_t + 8C_\sigma^2\eta_t^2)\Gamma_t\Big)e^{\eta_t\Gamma_t}.$$

Consequently, if only we can choose $\eta \in \mathcal{C}^1(\mathbb{R}_+)$ such that

$$\dot\eta_t + 10C_b\eta_t + 8C_\sigma^2\eta_t^2 = 0, \quad \eta_0 = \epsilon/2,$$

then we will have

$$(\partial_t + \mathcal{L})e^{\eta_t\Gamma_t} \le 2\eta_t(C_\sigma^2 + C_b)e^{\eta_t\Gamma_t}. \tag{2.11}$$

This is achieved by taking

$$\eta_t := \frac{5C_b\epsilon}{10C_be^{10C_bt} + 4C_\sigma^2\epsilon(e^{10C_bt}-1)}. \tag{2.12}$$

With this choice for η, the inequality (2.11) is satisfied and thus (2.10) gives

$$e^{\eta_{t\wedge\tau_n}\Gamma_{t\wedge\tau_n}} \le e^{\eta_0\Gamma_0} + 2(C_\sigma^2 + C_b)\int_0^{t\wedge\tau_n}\eta_s e^{\eta_s\Gamma_s}ds + \sum_{j=1}^N\int_0^{t\wedge\tau_n}\sigma_s^j\partial_{x_j}e^{\eta_s\Gamma_s}dB_s^j. \tag{2.13}$$

The stochastic integral is a true martingale, so taking expectations gives

$$\mathbb{E}\big[e^{\eta_{t\wedge\tau_n}\Gamma_{t\wedge\tau_n}}\big] \le \mathbb{E}\big[e^{\eta_0\Gamma_0}\big] + 2(C_\sigma^2 + C_b)\int_0^t\eta_s\mathbb{E}\big[e^{\eta_{s\wedge\tau_n}\Gamma_{s\wedge\tau_n}}\big]ds.$$

By Cauchy–Schwarz (recall $\eta_0 = \epsilon/2$), we have

$$\mathbb{E}\big[e^{\eta_0\Gamma_0}\big] \le \mathbb{E}\Big[e^{\epsilon|X_0^i|^2}\Big]^{\frac12}\mathbb{E}\Big[e^{\epsilon\sum_j a_j^N|X_0^j|^2}\Big]^{\frac12}$$

and hence it follows from Grönwall's inequality that

$$\mathbb{E}\left[e^{\eta_{t\wedge\tau_n}\Gamma_{t\wedge\tau_n}}\right] \leq e^{2(C_\sigma^2+C_b)\int_0^t \eta_s ds}\mathbb{E}\left[e^{\epsilon|X_0^i|^2}\right]^{\frac{1}{2}}\mathbb{E}\left[e^{\epsilon\sum_j a_j^N|X_0^j|^2}\right]^{\frac{1}{2}}, \tag{2.14}$$

Noting that $\tau_n \uparrow \infty$ as $n \to \infty$, we have $e^{\eta_{t\wedge\tau_n}\Gamma_{t\wedge\tau_n}} \to e^{\eta_t\Gamma_t}$ as $n \to \infty$, by the continuity of $t \mapsto \eta_t$ and $t \mapsto X_t^j$, for $j = 1, \ldots, N$. Consequently, Fatou's lemma yields

$$\mathbb{E}\left[e^{\eta_t\Gamma_t}\right] \leq \liminf_{n\to\infty}\mathbb{E}\left[e^{\eta_{t\wedge\tau_n}\Gamma_{t\wedge\tau_n}}\right] \leq e^{2(C_\sigma^2+C_b)\int_0^t \eta_s ds}\mathbb{E}\left[e^{\epsilon|X_0^i|^2}\right]^{\frac{1}{2}}\mathbb{E}\left[e^{\epsilon\sum_j a_j^N|X_0^j|^2}\right]^{\frac{1}{2}}.$$

Recall that the function $t \mapsto \eta_t$, as defined in (2.12), does not depend on N. Moreover, we can observe that

$$\mathbb{E}\left[e^{\epsilon\sum_j a_j^N|X_0^j|^2}\right]^{\frac{1}{2}} \leq \sum_{j=1}^N \mathbb{E}\left[a_j^N e^{\epsilon|X_0^j|^2}\right]^{\frac{1}{2}}$$

so if we, for example, have $0 \leq a_j^N \leq CN^{-1}$, for all j, then it is clear that the estimate is uniform in $N \geq 1$. □

The above result also yields control over the running max of $\Gamma^{i,N}$.

Corollary 2.3.2. *Given η_T from Proposition 2.3.1, fix any $\epsilon < \eta_T/2$. Then*

$$\mathbb{E}\left[e^{\epsilon\sup_{t\leq T}\Gamma_t^{i,N}}\right] \leq C_{T,\epsilon},$$

for some constant $C_{T,\epsilon} > 0$. Furthermore, the estimate is uniform in $N \geq 1$ under the same condition as in Proposition 2.3.1.

Proof. Setting $\xi_t := p\eta_t/2$ for $p < 1$, we can derive the same estimate as (2.13) only for ξ. As $p < 1$, we now have that $|\partial_{x_j}e^{\xi_s\Gamma_s}|^2$ is integrable by Proposition 2.3.1, so we can apply Burkholder–Davis–Gundy to control the running max of the stochastic integral. Thus, taking expectation and using monotone convergence, we get

$$\mathbb{E}\left[e^{\frac{p\eta_T}{2}\sup_{t\leq T}\Gamma_t}\right] \leq \mathbb{E}\left[\sup_{t\leq T} e^{\xi_t\Gamma_t}\right] \leq C_{T,p}.$$

As $p < 1$ was arbitrary, the result follows. □

2.3.1 Change of measure

In this section we first transform each particle into a Brownian motion with drift in a way that preserves the hitting times of the origin. Next, we then use the sub-Gaussianity to introduce a change-of-measure that can remove this drift and finally we obtain an important estimate for the associated Radon-Nikodym derivative.

Lemma 2.3.3. *Define the transformation* $\Upsilon \in C^{1,2}([0,T]\times\mathbb{R})$ *by*

$$(t,x) \mapsto \Upsilon_t(x) := \int_0^x \frac{1}{\sigma(t,y)} dy.$$

Fixing an arbitrary index $i \in \{1,\dots,N\}$*, we let* $Z_t := \Upsilon_t(X_t^i)$*. Then*

$$dZ_t = \hat{b}_t^i dt + dB_t^i \quad \textit{with} \quad Z_0 = \Upsilon_0(X_0^i),$$

where B^i *is a Brownian motion and the (random) drift* $\hat{b}^i$ *obeys the growth condition*

$$|\hat{b}_t^i| \lesssim 1 + |X_t^i| + |M_t^N|. \tag{2.15}$$

Furthermore, the transformed process Z *satisfies*

$$\operatorname{sgn}(Z_t) = \operatorname{sgn}(X_t^i) \quad \text{and} \quad |Z_t| \lesssim |X_t^i|. \tag{2.16}$$

Proof. Note that

$$\partial_t \Upsilon_t(x) = -\int_0^x \frac{\partial_t \sigma(t,y)}{\sigma(t,y)^2} dy, \;\; \partial_x \Upsilon_t(x) = \frac{1}{\sigma(t,x)}, \;\; \partial_{xx}^2 \Upsilon_t(x) = -\frac{\partial_x \sigma(t,x)}{\sigma(t,x)^2}.$$

Defining $B_t^i := \int_0^t \rho_s dW_s^0 + \int_0^t \sqrt{1-\rho_s^2} dW_s^i$, we have $dX_t^i = b_t^i dt + \sigma_t dB_t$ with $d\langle X^i\rangle_t = (\sigma_t^i)^2 dt$. Hence Itô's formula yields $dZ_t = \hat{b}_t^i dt + dB_t^i$, where

$$\hat{b}_t^i := \frac{b_t^i}{\sigma(t,X_t^i)} - \frac{1}{2}\partial_x \sigma(t,X_t^i) - \int_0^{X_t^i} \frac{\partial_t \sigma(t,y)}{\sigma(t,y)^2} dy.$$

Now, the bound on Z_t and the statement about its sign in (2.16) follow directly from the definition of Υ, since σ is strictly positive and bounded away from zero. Similarly, the growth condition in (2.15) follows from the properties of the coefficients as stated in Assumption 2.1.2. □

Lemma 2.3.4. *Fix* $i \in \{1,\dots,N\}$ *and define the stochastic exponential*

$$\mathcal{E}_t := \exp\Big(-\int_0^t \hat{b}_s^i dB_s^i - \frac{1}{2}\int_0^t (\hat{b}_s^i)^2 ds\Big),$$

where $\hat{b}^i$ is the drift of Z as defined in Lemma 2.3.3. Then Z is a Brownian motion under the probability measure $\mathbb{Q}$ given by the Radon-Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = \mathcal{E}_t$$

with initial value Z_0 distributed according to $\varpi_0 := \nu_0 \circ \Upsilon_0^{-1}$, where $\nu_0 := \mathrm{Law}(X_0^i)$.

Proof. For notational convenience we drop the superscript i and define

$$\mathcal{E}_{s,t} := \exp\Big(-\int_s^t \hat{b}_u dB_u - \frac{1}{2}\int_s^t \hat{b}_u^2 du\Big), \quad \mathcal{E}_{0,t} = \mathcal{E}_t.$$

By standard arguments, $\mathcal{E}_t$ is a positive continuous local martingale and hence also a supermartingale with $\mathcal{E}_0 = 1$. The claim is now that $\mathcal{E}_t$ is in fact a true martingale on $[0,T]$, which amounts to showing that $\mathbb{E}\mathcal{E}_t = 1$ for all $t \in [0,T]$.

Recall that $|\hat{b}_s| \leq C(1 + \Lambda_s^{i,N})$. While we cannot appeal to Novikov's condition directly, the sub-Gaussianity of $\Lambda_s^{i,N}$ will allows us to apply it on every interval of a fine enough partition of $[0,T]$. To see this, we fix an arbitrary $n \geq 1$ and partition the interval $[0,T]$ by $t_0 < \cdots < t_n$ where $t_k := kT/n$. An application of Jensen's inequality then yields

$$\begin{aligned}\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_{t_{k-1}}^{t_k} \hat{b}_u^2 du\Big)\Big] &\leq \frac{n}{T}\int_{t_{k-1}}^{t_k} \mathbb{E}\Big[\exp\Big(\frac{T}{2n}\hat{b}_u^2\Big)du\Big] \\ &\leq \sup_{t\in[t_{k-1},t_k]} \mathbb{E}\Big[\exp\Big(\frac{C^2T}{n}(1+\Lambda_t^{i,N})^2\Big)\Big].\end{aligned}$$

By choosing $n \geq 1$ sufficiently large, so that $2C^2T/n \leq \eta_T$, we can deduce from Proposition 2.3.1 that

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_{t_{k-1}}^{t_k} \hat{b}_u^2 du\Big)\Big] < \infty, \quad \text{for each} \quad k = 1, \ldots, n.$$

In particular, Novikov's condition now implies that $(\mathcal{E}_t)_{t\in[0,t_1]}$ is a true martingale and hence

$$\mathbb{E}[\mathcal{E}_t] = 1, \quad \text{for all} \quad t \in [0,t_1].$$

Noting that $(\mathcal{E}_{t_1,t})_{t\in[t_1,t_2]}$ is again a stochastic exponential with $\mathcal{E}_{t_1,t_1} = 1$, another application of Novikov's condition shows that $(\mathcal{E}_{t_1,t})_{t\in[t_1,t_2]}$ is a martingale. Consequently, we have

$$\mathbb{E}\left[\mathcal{E}_{t_1,t} \mid \mathcal{F}_{t_1}\right] = \mathcal{E}_{t_1,t_1} = 1$$

and hence

$$\mathbb{E}\mathcal{E}_t = \mathbb{E}[\mathcal{E}_{t_1}\mathcal{E}_{t_1,t}] = \mathbb{E}\big[\mathcal{E}_{t_1}\mathbb{E}[\mathcal{E}_{t_1,t} \mid \mathcal{F}_{t_1}]\big] = \mathbb{E}[\mathcal{E}_{t_1}] = 1$$

for all $t \in [t_1, t_2]$. Considering inductively the intervals $[t_{k-1}, t_k]$ for $k = 3, \dots, n$, it follows by the same reasoning that $\mathbb{E}[\mathcal{E}_t] = 1$ for all $t \in [0, T]$. Thus, $\mathcal{E}$ is a true martingale on $[0, T]$ and hence Girsanov's theorem implies that the process $Z = (Z_t)_{t\in[0,T]}$ given by

$$dZ_t = \hat{b}_t dt + dB_t, \quad Z_0 = \Upsilon_0(X_0^1),$$

is a Brownian motion under $\mathbb{Q}$ with initial distribution $\varpi_0 = \nu_0 \circ \Upsilon_0^{-1}$. □

Other than Z being a Brownian motion under $\mathbb{Q}$, we shall also need a specific estimate on the Radon–Nikodym derivative.

Lemma 2.3.5. *Fix $i \in \{1, \dots, N\}$ and let $\mathcal{E}_t$ be the stochastic exponential defined in Lemma 2.3.4. Given any $\epsilon > 0$ such that $\mathbb{E}[\exp(\epsilon|X_0^j|^2)] < \infty$, for $j = 1, \dots, N$, there exists $p > 1$ close enough to 1 such that*

$$\mathbb{E}\big[\mathcal{E}_t^{1-p} \mid X_0^i\big] \le C \exp\big(\epsilon (X_0^i)^2\big), \quad \textit{where} \quad C = C(p, \epsilon, T).$$

Moreover, for any $q > 1$, we have

$$\mathbb{E}\big[|\hat{b}_t^i|^q \mid X_0^i\big] \le C\big(1 + (X_0^i)^q\big), \quad \textit{where} \quad C = C(q, T).$$

Proof. In the notation of Lemma 2.3.4, we begin by defining

$$Y_t^i := -\int_0^t \hat{b}_s^i dB_s^i.$$

Using the conditional Hölder's inequality, with $p > 1$, we then have

$$\begin{aligned}\mathbb{E}\big[\mathcal{E}_t^{1-p} \mid X_0^i\big] &= \mathbb{E}\Big[\big(e^{pY_t^i - \frac{1}{2}\langle pY^i\rangle_t}\big)^{\frac{p-1}{p}} \big(e^{\frac{(p+1)}{2}(p-1)p\langle Y^i\rangle_t}\big)^{\frac{1}{p}} \mid X_0^i\Big] \\ &\le \mathbb{E}\Big[e^{pY_t^i - \frac{1}{2}\langle pY^i\rangle_t} \mid X_0^i\Big]^{\frac{p-1}{p}} \mathbb{E}\Big[e^{\frac{(p+1)}{2}(p-1)p\langle Y^i\rangle_t} \mid X_0^i\Big]^{\frac{1}{p}}\end{aligned}$$

Noting that the first term on the right-hand side is bounded by 1, since X_0^i is independent of B^i, we conclude that

$$\mathbb{E}\big[\mathcal{E}_t^{1-p} \mid X_0^i\big] \le \mathbb{E}\big[e^{C_p \int_0^t (\hat{b}_s^i)^2 ds} \mid X_0^i\big]^{\frac{1}{p}} \text{ with } C_p := \frac{(p+1)}{2}(p-1)p.$$

The key observation here is that $C_p \downarrow 0$ as $p \downarrow 1$. Recalling the linear growth bound $|\hat{b}_s^i| \lesssim 1 + \Lambda_s^{i,N}$, we can apply Jensen's inequality to see that

$$\mathbb{E}\big[\mathcal{E}_t^{1-p} \mid X_0^i\big] \lesssim \left(\frac{1}{T}\int_0^T \mathbb{E}\big[e^{TCC_p(\Lambda_s^{i,N})^2} \mid X_0^i\big] ds\right)^{\frac{1}{p}}.$$

By the above observation concerning C_p, we can fix a power p_0 close enough to 1 such that $TCC_{p_0} \leq \eta_T/2$, where $t \mapsto \eta_t$ is given by Proposition 2.3.1. In turn,

$$\mathbb{E}\big[\mathcal{E}_t^{1-p_0} \mid X_0^i\big] \lesssim \sup_{s\in[0,T]} \mathbb{E}\big[e^{\eta_s(\Lambda_s^{i,N})^2/2} \mid X_0^i\big]^{\frac{1}{p_0}}. \tag{2.17}$$

Given that X_0^i is independent of all the Brownian motions and the initial conditions X_0^j, for $j \neq i$, we still have the estimate (2.14) from the proof of Proposition 2.3.1, so

$$\sup_{s\in[0,T]} \mathbb{E}\big[e^{\eta_T(\Lambda_s^{i,N})^2/2} \mid X_0^i\big] \leq Ce^{\epsilon(X_0^i)^2/p_0}\mathbb{E}\Big[e^{\epsilon\sum_{j\neq i} a_j^N(X_0^j)^2}\Big]^{\frac{1}{2p_0}}.$$

Combining this with the estimate (2.17), the first claim of the proposition follows.

Finally, using the bound $|\hat{b}_t^i| \lesssim 1 + \Lambda_t^{i,N}$, the second claim follows by a standard Grönwall argument, so we leave out the proof. □

2.4 The density estimates

In this section, we derive the desired density estimates for each X^i absorbed at zero. We will do this by controlling the probability $\mathbb{P}[X_t^i \in S,\, t < \tau_i]$ for every $S \in \mathcal{B}(0,\infty)$, where we recall that $\tau_i = \inf\{t > 0 : X_t^i \leq 0\}$.

Using the scale transformation Υ from Lemma 2.3.3, and noting that $x \mapsto \Upsilon_t(x)$ is bijective with $\Upsilon_t(x) \leq 0$ if and only if $x \leq 0$, we have

$$\mathbb{P}_{X_0^i=x_0}[X_t^i \in S,\, t < \tau_i] = \mathbb{P}_{Z_0=z}[Z_t \in S_t,\, t < \tau],$$

where $\tau := \inf\{t > 0 : Z_t \leq 0\}$, $Z_t := \Upsilon(X_t^i)$, $S_t := \Upsilon_t(S)$, and $z := \Upsilon_0(x_0)$. From this point onwards, we fix (arbitrarily) a time horizon $t > 0$, a Borel set $S \in \mathcal{B}(0,\infty)$, and a starting point $x_0 > 0$, meaning that t, S_t, and z are fixed in what follows.

As above, we will use the notation $\mathbb{P}_{Z_0=z}$ to denote the regular conditional probability given $Z_0 = z$, and we will use the notation $\mathbb{P}_z$ to denote that a Brownian motion is started at a fixed initial point z.

In short, the idea is now to approximate the transformed particle $Z_{t\wedge\tau}$ by running it up to time s, for some $s < t$, and then running an independent absorbed Brownian motion W, started from $Z_{s\wedge\tau}$, for the remaining time $t - s$. More precisely, given our fixed inputs (t, S_t, and z), we are interested in the map

$$s \mapsto \mathbb{E}_{Z_0=z}\big[\mathbb{P}_{Z_{s\wedge\tau}}[W_{t-s} \in S_t,\; t - s < \tau_W]\big], \tag{2.18}$$

for all $s \in [0,t]$, where $\tau_W := \inf\{t > 0 : W_t \leq 0\}$. This leads us to define, for every $s < t$, the function

$$u(s,x) := \mathbb{P}_x\big[W_{t-s} \in S_t,\ t - s < \tau_W\big] = \int_{S_t} G_{t-s}(y,x)dy,$$

where $G_t(y,x) = p_t(x-y) - p_t(x+y)$ with $p_t(x) = (2\pi t)^{-\frac{1}{2}} \exp(-x^2/2t)$. Note that the resulting u is a classical solution of the terminal-boundary value problem

$$\begin{cases} \partial_s u(s,x) + \frac{1}{2}\partial_{xx} u(s,x) = 0 & \text{on} \quad [0,t) \times (0,\infty) \\ u(t,x) = \mathbf{1}_{S_t}(x) & \text{on} \quad \{t\} \times (0,\infty) \\ u(s,0) = 0 & \text{on} \quad [0,t) \times \{0\} \end{cases} \tag{2.19}$$

We can write (2.18) more succinctly as

$$s \mapsto v(s,z) := \mathbb{E}_{Z_0=z}[u(s, Z_{s\wedge\tau})], \tag{2.20}$$

and note that

$$v(0,z) = \mathbb{P}_z[W_{t\wedge\tau_W} \in S_t] \quad \text{and} \quad v(t,z) = \mathbb{P}_{Z_0=z}[Z_{t\wedge\tau} \in S_t].$$

By bounded convergence, it is immediate that $s \mapsto v(s,z)$ is continuous. Additionally, we show in Lemma 2.4.1 below that it is in fact absolutely continuous on $[0,t_0]$ for any $t_0 < t$. Consequently, if only we can show that the (a.e.) derivative $\partial_s v$ extends to $L^1(0,t)$, then we will have absolute continuity on all of $[0,t]$ with

$$\mathbb{P}_{Z_0=z}[Z_{t\wedge\tau} \in S_t] = \mathbb{P}_z[W_{t\wedge\tau_W} \in S_t] + \int_0^t \partial_s v(s,z)ds. \tag{2.21}$$

Therefore, the key is simply to establish the right control over $s \mapsto \partial_s v(s,z)$. We embark on this in the next section, but first we prove the previous claim about the absolute continuity.

Lemma 2.4.1. *For fixed $z > 0$ and $t > 0$, the mapping $s \mapsto v(s,z)$ from (2.20) is in $AC[0,t_0]$, for every $t_0 < t$, with (a.e.) derivative*

$$\partial_s v(s,z) = \mathbb{E}_{Z_0=z}\big[\mathbf{1}_{\{s<\tau\}} \hat{b}_s \partial_x u(s, Z_s)\big].$$

Proof. Fix $s \in [0,t)$ and $h > 0$ such that $h < t - s$. Recalling Lemma 2.3.3, we have $dZ_{t\wedge\tau} = \mathbf{1}_{\{t<\tau\}}\hat{b}_t dt + \mathbf{1}_{\{t<\tau\}} dB_t$, so an application of Itô's formula yields

$$\begin{aligned} v(s+h,z) - v(s,z) &= \mathbb{E}_{Z_0=z}\Big[u(s+h, Z_{(s+h)\wedge\tau}) - u(s, Z_{s\wedge\tau})\Big] \\ &= \mathbb{E}_{Z_0=z}\bigg[\int_s^{(s+h)\wedge\tau} (\partial_s + \frac{1}{2}\Delta) u_r dr + \int_s^{s+h} \mathbf{1}_{\{r<\tau\}} \hat{b}_r \partial_x u_r dr + \int_s^{(s+h)\wedge\tau} \partial_x u_r dB_r\bigg]. \end{aligned}$$

The first term on the right-hand side vanishes by (2.19) and, since $\partial_x u(r,x)$ is bounded on the interval $[s, s+h] \subseteq [0,t)$, the stochastic integral is a true martingale. Noting also that $\mathbb{E}[|\sup_{r\le T} \hat{b}_r|] < \infty$ (see e.g. Corollary 2.3.2), Fubini's theorem thus implies that

$$v(s+h,z) - v(s,z) = \int_s^{s+h} \mathbb{E}_{Z_0=z}\big[\mathbf{1}_{\{r<\tau\}} \hat{b}_r \partial_x u(r, Z_r)\big] dr.$$

This proves the claim. □

2.4.1 Estimates on the half-line with absorption at zero

Given the expression for $\partial_s v(s,z)$ in Lemma 2.4.1, an application of Hölder's inequality with $q > 1$ yields

$$\begin{aligned}
\big|\partial_s v(s,z)\big| &\le \mathbb{E}_{Z_0=z}\big[\mathbf{1}_{\{s<\tau\}} |\hat{b}_s| |\partial_x u(s, Z_s)|\big] \\
&\le \mathbb{E}_{Z_0=z}\Big[\mathbf{1}_{\{s<\tau\}} |\hat{b}_s| \int_{S_t} |\partial_x G_{t-s}(y, Z_s)| dy\Big]. \\
&\le \mathbb{E}_{X_0^i=x_0}\big[|\hat{b}_s|^{\frac{q}{q-1}}\big]^{\frac{q-1}{q}} \int_{S_t} \mathbb{E}_{Z_0=z}\big[\mathbf{1}_{\{s<\tau\}} |\partial_x G_{t-s}(y, Z_s)|^q\big]^{\frac{1}{q}} dy \\
&\le C(1+x_0) \int_{S_t} \mathbb{E}_{Z_0=z}\big[\mathbf{1}_{\{s<\tau\}} |\partial_x G_{t-s}(y, Z_s)|^q\big]^{\frac{1}{q}} dy.
\end{aligned}$$

Here the last inequality follows from the second claim in Lemma 2.3.5 and we emphasize that q can be taken arbitrarily close to 1. By introducing the Radon–Nikodym derivative $\mathcal{E}$ from Lemma 2.3.4, Hölder's inequality with $p > 1$ yields

$$\begin{aligned}
\int_{S_t} \mathbb{E}_{Z_0=z}\big[\mathbf{1}_{\{s<\tau\}} \big|\partial_x G_{t-s}(y, Z_s)\big|^q\big]^{\frac{1}{q}} dy &= \int_{S_t} \mathbb{E}^{\mathbb{Q}}_{Z_0=z}\big[\mathcal{E}_s^{-1} \mathbf{1}_{\{s<\tau\}} \big|\partial_x G_{t-s}(y, Z_s)\big|^q\big]^{\frac{1}{q}} dy \\
&\le \mathbb{E}_{X_0^i=x_0}\big[\mathcal{E}_s^{1-p}\big]^{\frac{1}{pq}} \int_{S_t} \mathbb{E}^{\mathbb{Q}}_{Z_0=z}\big[\mathbf{1}_{\{s<\tau\}} \big|\partial_x G_{t-s}(y, Z_s)\big|^a\big]^{\frac{1}{a}} dy,
\end{aligned}$$

where $a = a(p,q) := q\frac{p}{p-1} > 1$. For any $\delta > 0$, we can take p close enough to 1 so that the first estimate from Lemma 2.3.5 applies. Consequently, there exists $a > 1$ large enough such that

$$\begin{aligned}
\big|\partial_s v(s,z)\big| &\lesssim e^{\delta x_0^2} \int_{S_t} \mathbb{E}^{\mathbb{Q}}_z\big[\mathbf{1}_{s<\tau} \big|\partial_x G_{t-s}(y, Z_{s\wedge\tau})\big|^a\big]^{\frac{1}{a}} dy \\
&= e^{\delta x_0^2} \int_{S_t} \left(\int_0^\infty \big|\partial_x G_{t-s}(y,x)\big|^a G_s(x,z) dx\right)^{\frac{1}{a}} dy, \qquad (2.22)
\end{aligned}$$

where we have used that $Z_{s\wedge\tau}$ is an absorbing Brownian motion under $\mathbb{Q}$, which follows from Lemma 2.3.4.

Before proceeding, we state two elementary bounds, which we will become useful later.

Lemma 2.4.2. *Fix any $x, y \in \mathbb{R}$ and $t > s$. Then, for all powers $a \geq 1$,*

$$e^{-\frac{a(y-x)^2}{2(t-s)}} e^{-\frac{(x-z)^2}{2s}} \leq e^{-\frac{(y-z)^2}{2t}} e^{-\frac{t}{2s}\left(\frac{(x-y)}{\sqrt{t-s}} + \frac{\sqrt{t-s}(y-z)}{t-s+sa}\right)^2}, \tag{2.23}$$

$$e^{-\frac{a(y+x)^2}{2(t-s)}} e^{-\frac{(x-z)^2}{2s}} \leq e^{-\frac{(y+z)^2}{2t}} e^{-\frac{t}{2s}\left(\frac{(y+x)}{\sqrt{t-s}} + \frac{\sqrt{t-s}(y+z)}{t-s+sa}\right)^2}. \tag{2.24}$$

Proof. Writing $(x-z)^2 = ((y-z)+(x-y))^2$ and expanding the square, we have

$$\begin{aligned} e^{-\frac{(y-x)^2}{2(t-s)}} e^{-\frac{(x-z)^2}{2s}} &= e^{-\frac{(t-s+as)(x-y)^2}{2s(t-s)}} e^{-\frac{(y-z)^2}{2s}} e^{-\frac{2(y-z)(x-y)}{2s}} \\ &= e^{-\frac{as(y-z)^2}{2s(t-s+as)}} e^{-\frac{(t-s+as)(x-y)^2}{2s(t-s)}} e^{-\frac{(t-s)(y-z)^2}{2s(t-s+as)}} e^{-\frac{2(y-z)(x-y)}{2s}} \\ &= e^{-\frac{a(y-z)^2}{2(t-s+as)}} e^{-\frac{t-s+as}{2s}\left(\frac{(x-y)}{\sqrt{t-s}} + \frac{\sqrt{t-s}(y-z)}{t-s+sa}\right)^2}. \end{aligned}$$

Hence the first inequality follows once we notice that $t-s+as \geq t$ and $t-s+as \leq at$, by the assumptions that $t > s$ and $a \geq 1$. The argument is analogous for the second inequality. □

Recalling (2.21), the desired density estimates will follow if we can obtain suitable bounds on the right-hand side of (2.22).

Proposition 2.4.3. *For any $\delta > 0$ there exists $a > 1$ such that, for every $S \in \mathcal{B}(0,\infty)$, it holds uniformly in $i = 1, \dots, N$ and $N \geq 1$ that*

$$\begin{aligned} \mathbb{P}_{X_0^i = x_0}\big[X^i_{t\wedge\tau_i} \in S\big] \leq \int_S G_t\big(\Upsilon_t(x), \Upsilon_0(x_0)\big)\partial_x \Upsilon_t(x) dx \\ + C_a e^{\delta x_0^2} \int_S \Big(t^{-\frac{1}{a}} \Upsilon_0(x_0)^{\frac{1}{a}} \Upsilon_t(x)^{\frac{1}{a}} \wedge 1\Big) e^{-\frac{(\Upsilon_t(x) - \Upsilon_0(x_0))^2}{4at}} \partial_x \Upsilon_t(x) dx. \end{aligned}$$

Proof. Given $\delta > 0$, we can choose $a > 1$ such that (2.22) holds. Writing out the expressions for $\partial_x G_{t-s}$ and G_s, and noting that

$$e^{-(x-z)^2/2s} - e^{-(x+z)^2/2s} \leq \Big(\frac{2xz}{s} \wedge 1\Big) e^{-(x-z)^2/2s}, \tag{2.25}$$

we get $|\partial_s v(s,z)| \lesssim e^{\delta x_0^2} I(s)$, where $I(s) := s^{-\frac{1}{2a}}(t-s)^{-\frac{1}{2}} \tilde{I}(s)$ with

$$\tilde{I}(s) := \int_{S_t} \left(\int_0^\infty \Big| \frac{(y-x)}{t-s} e^{-\frac{(y-x)^2}{2(t-s)}} + \frac{(y+x)}{t-s} e^{-\frac{(y+x)^2}{2(t-s)}} \Big|^a \frac{xz}{s} e^{-\frac{(x-z)^2}{2s}} dx \right)^{\frac{1}{a}} dy.$$

Recalling (2.21), we have

$$\mathbb{P}_{X_0^i = x_0}\big[X^i_{t\wedge\tau_i} \in S\big] = \mathbb{P}_{Z_0 = z}\big[Z_{t\wedge\tau} \in S_t\big] \leq \mathbb{P}_z\big[W_{t\wedge\tau_W} \in S\big] + C e^{\delta x_0^2} \int_0^t I(s) ds \tag{2.26}$$

for some $C > 0$, so the claim amounts to controlling I. In the next six steps, we show that the integral of I is indeed finite and that this gives rise to the desired estimate.

Step 1. To estimate $I(s)$, we begin by observing that

$$\Big|\frac{(y-x)}{t-s}e^{-\frac{(y-x)^2}{2(t-s)}}+\frac{(y+x)}{t-s}e^{-\frac{(y+x)^2}{2(t-s)}}\Big|\le\frac{|y-x|}{t-s}\Big(e^{-\frac{(y-x)^2}{2(t-s)}}-e^{-\frac{(y+x)^2}{2(t-s)}}\Big)+\frac{2y}{t-s}e^{-\frac{(y+x)^2}{2(t-s)}}.$$

For the first term on the right-hand side, we can use (2.25) to see that

$$\frac{|y-x|}{t-s}\Big(e^{-\frac{(y-x)^2}{2(t-s)}}-e^{-\frac{(y+x)^2}{2(t-s)}}\Big)\le\frac{|y-x|}{t-s}\Big(\frac{2xy}{t-s}\wedge 1\Big)e^{-\frac{(y-x)^2}{2(t-s)}}=:f_1(s). \tag{2.27}$$

For the second term, it will prove useful to observe that, for $x,y>0$,

$$\frac{y}{t-s}e^{-\frac{(y+x)^2}{2(t-s)}}\le\Big(\frac{x+y}{x}\Big)^{\frac{1}{a}}\frac{y}{t-s}e^{-\frac{(y+x)^2}{2(t-s)}}\le\frac{y^{\frac{1}{a}}}{x^{\frac{1}{a}}}\frac{(x+y)}{t-s}e^{-\frac{(y+x)^2}{2(t-s)}}=:f_2(s), \tag{2.28}$$

where we have used that $y^{1-\frac{1}{a}}\le(x+y)^{1-\frac{1}{a}}$ since $a>1$.

Based on (2.27) and (2.28), we have $I(s)\le I_1(s)+2I_2(s)$, where

$$I_k(s):=s^{-\frac{3}{2a}}(t-s)^{-\frac{1}{2}}z^{\frac{1}{a}}\int_{S_t}\Big(\int_0^\infty f_k(s)^a x e^{-\frac{(x-z)^2}{2s}}dx\Big)^{\frac{1}{a}}dy,\quad k=1,2.$$

Step 2. We begin with the second term, I_2. Using (2.24) of Lemma 2.4.2, we have

$$\begin{aligned}I_2(s)=\ & s^{-\frac{3}{2a}}\frac{1}{t-s}\int_{S_t}z^{\frac{1}{a}}y^{\frac{1}{a}}\Big(\int_0^\infty\Big(\frac{x+y}{\sqrt{t-s}}\Big)^a e^{-\frac{a(y+x)^2}{2(t-s)}}e^{-\frac{(x-z)^2}{2s}}dx\Big)^{\frac{1}{a}}dy\\ &\le s^{-\frac{3}{2a}}\frac{1}{t-s}\int_{S_t}z^{\frac{1}{a}}y^{\frac{1}{a}}e^{\frac{-(y+z)^2}{2t}}\Big(\int_0^\infty\Big(\frac{x+y}{\sqrt{t-s}}\Big)^a e^{-\frac{t}{2s}\left(\frac{(y+x)}{\sqrt{t-s}}+\frac{\sqrt{t-s}(y+z)}{t-s+sa}\right)^2}dx\Big)^{\frac{1}{a}}dy\end{aligned}$$

To evaluate the inner integral, we perform a change of variables with

$$w=t^{\frac{1}{2}}s^{-\frac{1}{2}}\Big(\frac{(x+y)}{\sqrt{t-s}}+\frac{\sqrt{t-s}(y+z)}{t-s+\sigma s}\Big),\quad dx=(t-s)^{\frac{1}{2}}t^{-\frac{1}{2}}s^{\frac{1}{2}}dw,$$

where we note that

$$0\le\frac{(x+y)}{\sqrt{t-s}}=t^{-\frac{1}{2}}s^{\frac{1}{2}}w-\frac{\sqrt{t-s}(y+z)}{t-s+\sigma s}\le w$$

and hence

$$I_2(s)\le s^{-\frac{1}{a}}(t-s)^{\frac{1}{2a}-1}t^{-\frac{1}{2a}}\Big(\int_0^\infty w^a e^{-\frac{w^2}{2}}dw\Big)^{\frac{1}{a}}\int_{S_t}z^{\frac{1}{a}}y^{\frac{1}{a}}e^{\frac{-(y+z)^2}{2at}}dy.$$

Since $a>1$, we have $\int_0^t s^{-\frac{1}{a}}(t-s)^{\frac{1}{2a}-1}ds=c_a t^{-\frac{1}{2a}}$ for some $c_a>0$, so I_2 is in $L^1(0,t)$ with

$$\int_0^t I_2(s)ds\le C_a t^{-\frac{1}{a}}\int_{S_t}z^{\frac{1}{a}}y^{\frac{1}{a}}e^{\frac{-(y+z)^2}{2at}}dy, \tag{2.29}$$

where $C_a > 0$ is some numerical constant depending only on a.

Step 3. In order to estimate the first term, I_1, we rely on the inequality

$$x\Big(\frac{2xy}{t-s} \wedge 1\Big)^a \leq 4y\frac{|x-y|^2}{t-s} + 2y \quad \text{for any } a > 1. \tag{2.30}$$

To see that this inequality is true, simply note that, when $x \leq y$, we trivially have

$$x\Big(\frac{2xy}{t-s} \wedge 1\Big)^a \leq x \leq y$$

while, for $y \leq x$, we can write $x = (x-y) + y$ to obtain

$$x\Big(\frac{2xy}{t-s} \wedge 1\Big)^a \leq 2(|x-y|^2 + y^2)\Big(\frac{2y}{t-s} \wedge \frac{1}{y}\Big) \leq 2|x-y|^2\frac{2y}{t-s} + 2y.$$

Now, using the inequality (2.30) as well as (2.23) from Lemma 2.4.2, it follows that

$$\begin{aligned}
I_1(s) =\ & s^{-\frac{1}{2a}}(t-s)^{-\frac{1}{2}}\int_{S_t}\bigg(\int_0^\infty \Big|\Big(\frac{2xy}{t-s} \wedge 1\Big)\frac{|y-x|}{t-s}e^{-\frac{(y-x)^2}{2(t-s)}}\Big|^a \frac{xz}{s}e^{-\frac{(x-z)^2}{2s}}dx\bigg)^{\frac{1}{a}}dy \\
&= s^{-\frac{3}{2a}}(t-s)^{-\frac{1}{2}}\int_{S_t}\bigg(\int_0^\infty xz\Big(\frac{2xy}{t-s} \wedge 1\Big)^a \Big|\frac{y-x}{\sqrt{t-s}}\Big|^a e^{-\frac{a(y-x)^2}{2(t-s)}}e^{-\frac{(x-z)^2}{2s}}dx\bigg)^{\frac{1}{a}}dy \\
&\lesssim s^{-\frac{3}{2a}}(t-s)^{-\frac{1}{2}}\int_{S_t} z^{\frac{1}{a}}y^{\frac{1}{a}}e^{-\frac{(y-z)^2}{2at}}\left(\int_0^\infty F(s,x,y,z)dx\right)^{\frac{1}{a}}dy,
\end{aligned}$$

where

$$F(s,x,y,z) := \left(\Big|\frac{y-x}{\sqrt{t-s}}\Big|^a + \Big|\frac{y-x}{\sqrt{t-s}}\Big|^{a+2}\right)e^{-\frac{t}{2s}\left(\frac{(y-x)}{\sqrt{t-s}} + \frac{\sqrt{t-s}(y-z)}{t-s+sa}\right)^2}.$$

For the inner integral, we make the change of variables

$$w = t^{\frac{1}{2}}s^{-\frac{1}{2}}\left(\frac{x-y}{\sqrt{t-s}} + \frac{\sqrt{t-s}(y-z)}{t-s+as}\right), \quad dx = (t-s)^{\frac{1}{2}}t^{-\frac{1}{2}}s^{\frac{1}{2}}dw,$$

and set

$$f(w) = f(s,w,y,z) := \Big|t^{-\frac{1}{2}}s^{\frac{1}{2}}w - \frac{\sqrt{t-s}(y-z)}{t-s+as}\Big|.$$

This yields

$$I_1(s) \lesssim s^{-\frac{1}{a}}(t-s)^{\frac{1}{2a}-1}t^{-\frac{1}{2a}}\int_{S_t} z^{\frac{1}{a}}y^{\frac{1}{a}}e^{-\frac{(y-z)^2}{2at}}\bigg(\int_{\mathbb{R}}\big(f(w)^a + f(w)^{a+2}\big)e^{-\frac{w^2}{2}}dw\bigg)^{\frac{1}{a}}dy.$$

Noting that $f(w) \leq |w| + t^{-1}(t-s)^{\frac{1}{2}}|y-z|$, we split up I_1 accordingly. Since

$$t^{-\frac{1}{2a}}\int_0^t s^{-\frac{1}{a}}(t-s)^{\frac{1}{2a}-1}ds = c_a t^{-\frac{1}{a}}, \quad t^{-1-\frac{1}{2a}}\int_0^t s^{-\frac{1}{a}}(t-s)^{\frac{1}{2a}-\frac{1}{2}}ds = c_a' t^{-\frac{1}{2}-\frac{1}{a}},$$

$$\text{and} \quad t^{-1-\frac{5}{2a}}\int_0^t s^{-\frac{1}{a}}(t-s)^{\frac{3}{2a}-\frac{1}{2}}ds = c_a'' t^{-\frac{1}{2}-\frac{2}{a}},$$

we thus obtain

$$\int_0^t I_1(s)ds \lesssim C_a t^{-\frac{1}{a}} \int_{S_t} z^{\frac{1}{a}} y^{\frac{1}{a}} e^{-\frac{(y-z)^2}{2at}} dy + C'_a t^{-\frac{1}{2}-\frac{1}{a}} \int_{S_t} z^{\frac{1}{a}} y^{\frac{1}{a}} e^{-\frac{(y-z)^2}{2at}} |y-z| \, dy$$
$$+ C''_a t^{-\frac{1}{2}-\frac{2}{a}} \int_{S_t} z^{\frac{1}{a}} y^{\frac{1}{a}} e^{-\frac{(y-z)^2}{2at}} |y-z|^{1+\frac{2}{a}} \, dy, \tag{2.31}$$

where

$$C'_a = c'_a \left(\int_{\mathbb{R}} e^{-\frac{w^2}{2}} dw \right)^{\frac{1}{a}} = c'_a 2^{\frac{1}{2a}} \pi^{\frac{1}{2a}}, \quad C''_a = c''_a \left(\int_{\mathbb{R}} e^{-\frac{w^2}{2}} dw \right)^{\frac{1}{a}} = c''_a 2^{\frac{1}{2a}} \pi^{\frac{1}{2a}},$$
$$\text{and} \quad C_a = c_a \left(\int_{\mathbb{R}} \big(|w|^a + |w|^{a+2} \big) e^{-\frac{w^2}{2}} dw \right)^{\frac{1}{a}}.$$

Step 4. We can now combine the estimates for I_1 and I_2 from Steps 2 and 3. To simplify the resulting expressions, it is useful to recall the elementary inequalities

$$|y-z| \, e^{-\frac{(y-z)^2}{2at}} \leq C_a t^{\frac{1}{2}} e^{-\frac{(y-z)^2}{4at}}, \quad (y+z) e^{\frac{-(y+z)^2}{2at}} \leq C_a t^{\frac{1}{2}} e^{-\frac{(y+z)^2}{4at}},$$
$$\text{and} \quad t^{-\frac{1}{a}} |y-z|^{1+\frac{2}{a}} \, e^{-\frac{(y-z)^2}{2at}} \leq C_a t^{\frac{1}{2}} e^{-\frac{(y-z)^2}{4at}}.$$

Given (2.26), and using the above inequalities, it follows from (2.29) and (2.31) that there exists a constant $C_a > 0$ such that

$$\mathbb{P}_{X_0^i = x_0} \big[X^i_{t \wedge \tau_i} \in S \big] \leq \int_{S_t} G_t(y,z) dy + C_a e^{\delta x_0^2} \int_{S_t} t^{-\frac{1}{a}} z^{\frac{1}{a}} y^{\frac{1}{a}} e^{-\frac{(y-z)^2}{4at}} dy. \tag{2.32}$$

Step 5. It remains to observe that (2.32) also holds with 1 in place of $t^{-\frac{1}{a}} z^{\frac{1}{a}} y^{\frac{1}{a}}$. To see this, note that in Step 1 we also have $|\partial_s v(s,z)| \lesssim e^{\delta x_0^2} (J_1(s) + J_2(s))$, where

$$\begin{cases} J_1(s) := s^{-\frac{1}{2a}} (t-s)^{-1} \displaystyle\int_{S_t} \Big(\int_0^\infty \Big| \frac{y-x}{\sqrt{t-s}} \Big|^a e^{-\frac{a(y-x)^2}{2(t-s)}} e^{-\frac{(x-z)^2}{2s}} dx \Big)^{\frac{1}{a}} dy \\ J_2(s) := s^{-\frac{1}{2a}} (t-s)^{-1} \displaystyle\int_{S_t} \Big(\int_0^\infty \Big| \frac{y+x}{\sqrt{t-s}} \Big|^a e^{-\frac{a(y+x)^2}{2(t-s)}} e^{-\frac{(x-z)^2}{2s}} dx \Big)^{\frac{1}{a}} dy. \end{cases}$$

Using Lemma 2.4.2 and performing the same changes of variables as in Steps 2 and 3, the bound follows by similar arguments only with substantially simplified computations.

Step 6. Recall that $S_t = \Upsilon_t(S)$ and $z = \Upsilon_0(x_0)$. In view of Step 4 and Step 5, the proof is therefore finished by performing the change of variables $y = \Upsilon_t(x)$. □

2.4.2 Estimates on the whole space

If we ignore the absorption at the boundary, and instead work on the whole real line, then the estimates simplify significantly. This leads to the following result.

Proposition 2.4.4. *For every $\delta > 0$ there exists $a > 1$ and $C_a > 0$ such that, uniformly in $i = 1, \dots, N$ and $N \geq 1$, it holds for every $S \in \mathcal{B}(\mathbb{R})$ that*

$$\mathbb{P}_{X_0^i = x_0}\big[X_t^i \in S\big] \leq \int_S p_t\big(\bar{\Upsilon}_t(x, x_0)\big)\partial_x \bar{\Upsilon}_t(x, x_0)dx + C_a \int_S e^{-\frac{(\bar{\Upsilon}_t(x,x_0))^2}{4at}} \partial_x \bar{\Upsilon}_t(x, x_0)dx,$$

where $p_t(x) = (2\pi t)^{-\frac{1}{2}} \exp(x^2/2t)$ and $\bar{\Upsilon}_t(x, x_0) = \int_{x_0}^x \sigma(t, y)^{-1} dy$.

Proof. Consider Lemma 2.3.3 with $\bar{Z}_t := \bar{\Upsilon}_t(X_t^i, x_0)$ in place of $Z_t = \Upsilon_t(X_t^i) = \bar{\Upsilon}_t(X_t^i, 0)$ and let $\bar{\mathbb{Q}}$ denote the corresponding measure from Lemma 2.3.4 such that $\bar{Z}_t$ is a standard Brownian motion under $\bar{\mathbb{Q}}$, started at 0 when $X_0^i = x_0$. Setting $\bar{S}_t := \bar{\Upsilon}_t(S, x_0)$, we have

$$\mathbb{P}_{X_0^i = x_0}\big[X_t^i \in S\big] = \mathbb{P}_{Z_0 = 0}\big[\bar{Z}_t \in \bar{S}_t\big].$$

From here, the bound follows by analogy with Proposition 2.4.3, only for the new measure $\bar{\mathbb{Q}}$ and the corresponding functions

$$\bar{v}(s) := \mathbb{E}_{Z_0=0}\big[\bar{u}(s, \bar{Z}_s)\big], \quad \bar{u}(s, x) := \mathbb{P}_x\big[W_{t-s} \in \bar{S}_t\big] = \int_{\bar{S}_t} p_{t-s}(y - x)dy.$$

Since there are no boundary effects, the estimates simplify. In fact, the arguments are the same as for J_1 in Step 5 of the proof of Proposition 2.4.3 with $z = 0$. □

2.4.3 Proof of Theorem 2.1.1

In view of Propositions 2.4.3 and 2.4.4, only a few observations remain before we can deduce the density estimates (2.2) and (2.3), thus proving Theorem 2.1.1. Given Assumption 2.1.2, it holds by construction of $\bar{\Upsilon}$ that there are constants $C_1, C_2 > 0$ such that $C_1|x - x_0| \leq |\bar{\Upsilon}_t(x, x_0)| \leq C_2|x - x_0|$ and $|\partial_x \Upsilon_t(x)| \leq C_2$ for all $x \in \mathbb{R}$ and $t \in [0, T]$. Consequently, the whole-space estimate (2.3) is an immediate consequence of Proposition 2.4.4.

For the estimate with boundary decay, we begin by recalling the classical Dirichlet heat kernel estimate from Lemma 2.7, namely

$$G_t(x, x_0) \leq C\frac{1}{\sqrt{t}}\Big(\frac{x}{\sqrt{t}} \wedge 1\Big)\Big(\frac{x_0}{\sqrt{t}} \wedge 1\Big)\exp\Big(\frac{-(x - x_0)^2}{4t}\Big),$$

for all $x, x_0 \geq 0$ and $t > 0$. Next, we recall that, by definition of Υ, there exists $C > 0$ such that $|\Upsilon_t(x)| \leq Cx$ and $|\partial_x \Upsilon_t(x)| \leq C$ for all $x \geq 0$ and $t \in [0, T]$. Finally, we can observe that

$$\begin{aligned}\big(\Upsilon_t(x) - \Upsilon_0(x_0)\big)^2 &\geq \left(\int_{x \wedge x_0}^{x \vee x_0} \frac{dy}{\sigma(t, y)}\right)^2 + \int_{x \wedge x_0}^{x \vee x_0} \frac{2dy}{\sigma(t, y)} \int_0^{x \wedge x_0} \frac{1}{\sigma(t, y)} - \frac{1}{\sigma(0, y)} dy \\ &\geq C(x - x_0)^2 - C' t|x - x_0|(x \wedge x_0),\end{aligned}$$

by simply relying on the properties of σ and $\partial_t\sigma$ from Assumption 2.1.2. Combining these observations, the density estimate (3.12) follows from Proposition 2.4.3 by taking $c_{x,y} := 4aC'|x-y|(x \wedge y)$ and $\kappa := 1/a$.

It remains to observe that we can take $c_{x,y} \equiv 0$ when $\sigma(t,x) = \sigma_1(t)\sigma_2(x)$ and that this case does not rely on any smoothness of $t \mapsto \sigma(t,x)$, as mentioned in part (ii) of Assumption 2.1.2. To see this, the point is that it suffices to scale away the spatial component of the volatility. Specifically, we can consider the analogue of Proposition 2.4.3 with

$$\tilde{\Upsilon}(x) := \int_0^x \frac{dy}{\sigma_2(y)} \quad \text{and} \quad \tilde{Z}^i_t := \tilde{\Upsilon}(X^i_t) = \tilde{\Upsilon}(X^i_0) + \int_0^t \tilde{b}^i_s ds + \int_0^t \sigma_1(s) dB_s.$$

Fix $S \in \mathcal{B}(0,\infty)$ and set $\tilde{S} := \tilde{\Upsilon}(S)$. Fixing also $t \in (0,T]$, we define the time change

$$s \mapsto A^{\sigma_1}_{s,t} := \int_s^t \sigma_1(r)^2 dr.$$

Then we can replicate the arguments from the start of Section 2.3 and subsection 2.3.1 with $\tilde{v}(s,z) := \mathbb{E}^z[\tilde{u}(s, \tilde{Z}_{s\wedge\tilde{\tau}})]$ and

$$\tilde{u}(s,x) := \mathbb{P}_x\big[W_{A^{\sigma_1}_{s,t}} \in \tilde{S},\ A^{\sigma_1}_{s,t} < \tau_W\big] = \int_{\tilde{S}} G^{\sigma_1}_{s,t}(y,x)dy,$$

where $(s,x,y) \mapsto G^{\sigma_1}_{s,t}(y,x)$ is the Green's function for $\partial_s f + \frac{1}{2}\sigma_1(s)^2\partial_{xx} f = 0$ as a terminal-boundary value problem on $[0,t) \times (0,\infty)$ with $f(s,0) = 0$. That is,

$$G^{\sigma_1}_{s,t}(y,x) = \big(2\pi A^{\sigma_1}_{s,t}\big)^{-\frac{1}{2}}\bigg(\exp\Big(\frac{-(x-y)^2}{2A^{\sigma_1}_{s,t}}\Big) - \exp\Big(\frac{-(x+y)^2}{2A^{\sigma_1}_{s,t}}\Big)\bigg).$$

Since there are constants $c_1, c_2 > 0$ such that $c_1(t-s) \le A^{\sigma_1}_{s,t} \le c_2(t-s)$, the same estimates as in the proof of Proposition 2.4.3 yield

$$\mathbb{P}_{X^i_0=x_0}\big[X^i_{t\wedge\tau_i} \in S\big] = \int_{\tilde{S}} G^{\sigma_1}_{s,t}(y, \tilde{\Upsilon}(x_0))dy + \int_0^t \partial_s\tilde{v}(s, \tilde{\Upsilon}(x_0))ds$$

and, using also the bounds $|\tilde{\Upsilon}(x)| \le cx$ and $|\partial_x\tilde{\Upsilon}| \le c$, we thus have

$$\begin{aligned}\mathbb{P}_{X^i_0=x_0}\big[X^i_{t\wedge\tau_i} \in S\big] \lesssim \int_S & G^{\sigma_1}_{0,t}\big(\tilde{\Upsilon}(x), \tilde{\Upsilon}(x_0)\big)dx \\ & + e^{\delta x_0^2}\int_S \big(t^{-\frac{1}{a}}x^{\frac{1}{a}}x_0^{\frac{1}{a}} \wedge 1\big)e^{-\frac{(\tilde{\Upsilon}(x)-\tilde{\Upsilon}(x_0))^2}{4at}}dx.\end{aligned}$$

As $\tilde{\Upsilon}$ does not depend on time, we have $|\tilde{\Upsilon}(x) - \tilde{\Upsilon}(x_0)| = |\int_{x_0}^x \sigma_2(y)^{-1}dy|$, so there exist $c_1', c_2' > 0$ such that $c_1'|x-x_0| \le |\tilde{\Upsilon}(x) - \tilde{\Upsilon}(x_0)| \le c_2'|x-x_0|$. Furthermore, $G^{\sigma_1}_{0,t}$ is easily seen to satisfy an analogous bound to that of $G_t(x,y)$ from Lemma 2.7. Therefore, we can conclude that the density estimate (2.2) holds with $c_{x,y} \equiv 0$, as desired. This finishes the proof of Proposition 2.1.1.

Chapter 3

An SPDE Model for Systemic Risk with Endogenous Contagion

3.1 Introduction

In this chapter we propose a dynamic mean-field model for 'systemic risk' in large financial systems, derived from a system of interacting diffusions on the positive half-line with an absorbing boundary at the origin. These diffusions represent the distances-to-default of financial institutions and absorption at zero corresponds to default. As a way of modelling correlated exposures and herd behaviour, we consider a common source of noise and a form of mean-reversion in the drift. Moreover, we introduce an endogenous contagion mechanism whereby the default of one institution causes a drop in the distances-to-default of the other institutions. In this way, we aim to capture key 'system-wide' effects on risk, through what we have decided to call a contagious McKean–Vlasov system with positive feedback (in the sense of absorptions pulling the system closer to the absorbing boundary).

The mean-field (or large population) limit of the finite system turns out to be characterized uniquely by a nonlinear and non-local SPDE on the half-line with a Dirichlet boundary condition. The density of this SPDE gives the conditional law of a non-standard 'conditional' McKean–Vlasov diffusion, for which we provide a novel upper Dirichlet heat kernel type estimate based on the results from Chapter 2.

Depending on the realizations of the common noise and the rate of mean reversion, we will see that the SPDE can exhibit rapid accelerations in the loss of mass at the boundary. In other words, the contagion mechanism can give rise to periods of significant systemic default clustering.

3.1.1 A simple model of default contagion

Our starting point is inspired by recent dynamic frameworks for the *structural* modelling of large portfolio credit risk as proposed in [16, 73]. In these frameworks, a portfolio of defaultable entities is modelled as a system of interacting diffusion processes, where defaults occur at the first hitting time of a lower barrier in line with the original first-passage model of Black and Cox [12] (extending the seminal work of Merton [91], where default can only happen at the terminal time).

To be precise, we identify each financial institution (henceforth: *bank*) with a notion of its *distance-to-default* given by

$$Y_t^i = \log A_t^i - \log D_t^i, \quad \text{for } i = 1, \dots, N,$$

where A_t^i is the market value of bank i's assets and D_t^i denotes its default barrier. These distances-to-default will be modelled by suitable stochastic processes on $(0, \infty)$ with absorption at the origin corresponding to default. The precise dynamics will be specified below, but first we show how to incorporate the contagion mechanism.

For simplicity, we assume that the system can be described by assigning a weight a_i^N to each of the banks, where

$$a_i^N = \frac{a_i(X_0^i)}{\sum_{j=1}^N a_j(X_0^j)} \quad \text{with} \quad c \le a_i(\cdot) \le C, \qquad i = 1, \dots, N, \tag{3.1}$$

for fixed constants $c, C > 0$. Here the idea is that the realisation of the initial values X_0^i fix how important bank i is (relative to the other banks) and that this is then reflected in the weight a_i^N. Notice that $\sum_{i=1}^N a_i^N = 1$ in agreement with the fact that the weights describe the *relative* importance of each bank. Moreover, we stress that the above bounds on a_i imply $a_i^N = O(1/N)$ as $N \to \infty$. In particular, no single bank can have a macroscopic effect on the system in the large population limit.

Remark 3.1.1 (Common noise). Since our model will contain a *common* source of noise, there is a simple way of introducing a group of particularly influential 'systemically important' banks. Indeed, we could model a group of such banks by a separate diffusion and then treat this as a common input in the dynamics of the smaller banks.

3.1.1.1 The endogenous contagion mechanism

Suppose bank j is the first to default. We propose that this will have a contagious impact on any other bank $i \neq j$ by means of a gradual 'discounting' of its asset values,

materializing over some short period of time following the default. The size of the 'discount' will be determined by bank j's relative weight a_j^N (defined in (3.1)) and a parameter α_t^i capturing how costly defaults are to bank i at time t (see Remark 3.1.2 below for a precise interpretation of the contagion parameter α). Specifically, we model the resulting contagion by discounting the asset values of the other banks according to the rule

$$A_\cdot^i \longmapsto \hat{A}_\cdot^i := \exp\Big(-a_j^N \int_0^\cdot \alpha_s^i d\mathfrak{L}_s^{j,N}\Big) A_\cdot^i, \quad \text{for each } \ i \neq j, \tag{3.2}$$

with

$$\mathfrak{L}_t^{j,N} := \int_0^t \mathfrak{K}(t-r)\mathbf{1}_{\{r \geq \tau_j\}} dr,$$

where the *impact kernel* $\mathfrak{K}$ models the gradual realisation of the losses spurred by the default.

Observe that we have $\mathfrak{L}_t^{j,N} = 0$ for all $t \leq \tau_j$, so the discounting in (3.2) only takes effect at the time of bank j's default. Furthermore, we require that the kernel satisfies $\|\mathfrak{K}\|_{L^1} = 1$, so $\mathfrak{L}_t^{j,N} = 1$ for all $t \geq \tau_j + \varepsilon$, whenever the support of $\mathfrak{K}$ equals $[0, \varepsilon]$ for some $\varepsilon > 0$, which is the setup we have in mind (see Remark 3.1.2 below). We stress that the total negative impact from contagion is not restricted to losses from direct counterparty exposures, but may also have arisen from more indirect sources such as emerging liquidity shortages, fire sales and drops in confidence.

Remark 3.1.2 (Interpretation of α). The parameter α can be interpreted as follows. Consider the case where $\alpha \geq 0$ is a fixed constant and $\mathrm{supp}\mathfrak{K} \subseteq [0, \varepsilon]$. At time $t = \tau_j + \varepsilon$, the discounting in (3.2) is then of the form

$$\hat{A}_t^i = \exp(-\alpha a_j^N) A_t^i \simeq (1 - \alpha a_j^N) A_t^i,$$

for large N, since $a_j^N = O(1/N)$. In other words: *by time $\tau_j + \varepsilon$, the default of bank j has caused each bank $i \neq j$ to lose a proportion αa_j^N of their asset values, relative to what they would have been worth without the contagion.*

As more banks default, we continue to apply the rule from (3.2). Therefore, the actual (updated) asset values, $\hat{A}$, are given by

$$\hat{A}_t^i := \prod_{j \neq i} \exp\Big(-a_j^N \int_0^t \alpha_s^i d\mathfrak{L}_s^{j,N}\Big) A_t^i = \exp\Big(-\sum_{j \neq i} a_j^N \int_0^t \alpha_s^i d\mathfrak{L}_s^{j,N}\Big) A_t^i,$$

for $i = 1, \ldots, N$, where $\mathfrak{L}_t^{j,N} := \int_0^t \mathfrak{K}(t-r)\mathbf{1}_{\{r \geq \hat{\tau}_j\}} dr$ are defined in terms of the actual default times $\hat{\tau}_j := \inf\{t > 0 : \hat{Y}_t^j \leq 0\}$ for the actual (updated) distances-to-default,

namely $\hat{Y}_t^j := \log \hat{A}_t^j - \log D_t^j$. Summing over the terms $a_j^N \mathfrak{L}_s^{j,N}$, we get

$$\hat{A}_t^i = \exp\Big(-\int_0^t \alpha_s^i d\mathfrak{L}_s^N\Big) A_t^i, \quad \text{for} \quad t < \hat{\tau}_i, \quad i = 1, \ldots, N, \tag{3.3}$$

where

$$\mathfrak{L}_t^N := \int_0^t \mathfrak{K}(t-s) L_s^N ds \quad \text{and} \quad L_t^N := \sum_{j=1}^N a_j^N \mathbf{1}_{\{t \geq \hat{\tau}_j\}}. \tag{3.4}$$

Taking logarithms in (3.3), it follows that the actual (updated) distances-to-default, $\hat{Y}$, have dynamics of the form

$$d\hat{Y}_t^i = \, dY_t^i - \alpha_t^i d\mathfrak{L}_t^N \quad \text{for} \quad t < \hat{\tau}_i, \quad i = 1, \ldots, N.$$

Here Y^i is simply the original distance-to-default without contagion, while the latter part is a new contagion term driven by the *contagion process*, $\mathfrak{L}^N$, from (3.4).

Remark 3.1.3 (The impact kernel). Concerning the impact kernel, $\mathfrak{K}$, the only assumption we will make is that $\mathfrak{K} \in \mathcal{W}_0^{1,1}(\mathbb{R}_+)$ with $\|\mathfrak{K}\|_1 = 1$, where $\mathcal{W}_0^{1,p}(\mathbb{R}_+)$ denotes the Sobolev space with one weak derivative in L^p and zero trace. The benefits of this construction are that $\mathfrak{L}^N = \mathfrak{K} \star L^N$ remains adapted, it inherits the monotonicity of L^N, and $(\mathfrak{L}^N)' = \mathfrak{K}' \star L^N$ is in $L^\infty(0,T)$. The role of the kernel is to model a continuous notion of latency whereby the full impact of contagion is realised gradually (depending on the kernel's support) as counterparty exposures are sorted out and indirect effects start to kick in.

Remark 3.1.4 (Capital Structure). As would be expected, the default contagion alone cannot deplete the entire asset base. However, due in large part to the low volatility of banking assets in normal times, financial institutions tend to have leverage ratios (here defined as liabilities over assets) as high as 85–95% (see e.g. [10, 68]). Thus, there is ample room for the contagion to be detrimental—in the sense that relatively small adjustments to the asset values can result in distress.

Remark 3.1.5 (Connectivity of the system). It is worth noting that one can introduce a simple notion of the connectivity of the system through the contagion parameter α. Suppose, for example, that each default only affects a randomly sampled proportion $\hat{p}$ of the banks, each losing αa_j^N times their asset values upon bank j's default. As N gets large, this has a similar effect on the system as all the banks incurring the smaller loss of $\hat{\alpha} a_j^N$ times their asset values, where $\hat{\alpha} := \hat{p}\alpha$.

3.1.1.2 A concrete example: contagion and common exposures

Consider a system of N banks. Suppose that, in the absence of contagion, the assets of all these banks can be modelled by a system of geometric Brownian motions, which are correlated through a common Brownian motion and otherwise have the same drift and volatility. That is, suppose

$$\frac{dA_t^i}{A_t^i} = \mu dt + \sigma\big(\sqrt{1-\rho^2}dW_t^i + \rho dW_t^0\big), \qquad \text{for} \quad i = 1, \dots, N,$$

where $W^0, W^1, \dots$ is a sequence of independent Brownian motions. In addition, we assume for simplicity that each default barrier D_t^i is constant in time and given by a random variable D^i (of course, we could also model the default barrier by a diffusion).

In the absence of contagion, the distances-to-default are then given by

$$Y_t^i = Y_0^i + \mu t + \sigma\big(\sqrt{1-\rho^2}W_t^i + \rho W_t^0\big), \qquad Y_0^i = \log A_0^i - \log D^i.$$

Depending on the realisations of the Y_0^i's, each bank is assigned a relative weight a_i^N given by (3.1). As we saw in Remark 3.1.2, our contagion mechanism amounts to imposing that, upon the default of bank j, each bank $i \neq j$ loses a proportion αa_j^N of its a priori asset value over an ε amount of time (modulated by $\mathfrak{K}$).

Writing X^i for the updated distances-to-default, $\hat{Y}^i$, it follows that, when accounting for contagion, the actual distances-to-default have dynamics of the form

$$\begin{cases} dX_t^i = \mu dt + \sigma\big(\sqrt{1-\rho^2}dW_t^i + \rho dW_t^0\big) - \alpha d\mathfrak{L}_t^N, \\ \mathfrak{L}_t^N = \int_0^t \mathfrak{K}(t-s)L_s^N ds, \quad L_t^N = \sum_{j=1}^N a_j^N \mathbf{1}_{t \geq \tau_j}, \\ \tau_i = \inf\{t > 0 : X_t^i \leq 0\}. \end{cases}$$

In the next section, we introduce a general version of this model which allows for much more general coefficients as opposed to just constants. This larger generality can, for example, be used to include effects of herding behaviour in the drift, and it can allow the various coefficients to depend on time as well as adjust to the current state of the system.

3.1.2 A simple model for systemic risk: contagion, common exposures, and herding

In addition to contagion, we want our model for systemic risk to include common exposures and a notion of herding. The two latter effects have already been considered in Carmona et al. [28] by means of a (constant-coefficient) particle system with a common Brownian motion and mean reversion in the drift.

There are several reasons why it is sensible to include herding via a form of mean reversion in the drift of the distances-to-default. First of all, it is unreasonable if the distance-to-default of a single bank becomes disproportionately large and this is taken care of by mean reversion. More generally, banks tend to engage in similar business models and they make similar investment decisions, so it is natural that, if the healthiness of some banks start to decline, then this should also pull down the healthiness of the other banks (but in a smooth way that is separate from the contagion which only occurs at times of default). Therefore, we will include herding as a form of mean-reversion that adjusts the drifts of each distance-to-default depending on its value vis-à-vis a weighted average of the surviving banks' distances-to-default.

We are now in a position to present a precise formulation of our 'base case' model for systemic risk. Letting X^j denote the actual distances-to-default with default times $\tau_j := \inf\{t > 0 : X_t^j \leq 0\}$, for $j = 1, \ldots, N$, we propose to model a large financial system by an interacting particle system of the form

$$\begin{aligned} dX_t^j = \mu(t, X_t^j)dt + \pi(t, \nu_t^N) \sum_{i=1}^N a_i^N \Big(\big(X_t^i \mathbf{1}_{\{t<\tau_i\}} + \gamma(t, \nu_t^N)\mathbf{1}_{\{t\geq\tau_i\}} \big) - X_t^j \Big) dt \\ + \sigma(t, X_t^j)\big(\sqrt{1-\rho(t,\nu_t^N)^2} dW_t^j + \rho(t, \nu_t^N) dW_t^0 \big) - \alpha(t, X_t^j, \nu_t^N) d\mathfrak{L}_t^N, \end{aligned}$$

where W^0 and $W^1, \ldots, W^N$ are independent Brownian motions with

$$\mathfrak{L}_t^N = \int_0^t \mathfrak{K}(t-s) L_s^N ds, \quad L_t^N = \sum_{i=1}^N a_i^N \mathbf{1}_{\{t\geq\tau_i\}}, \quad \text{and} \quad \nu_t^N = \sum_{i=1}^N a_i^N \mathbf{1}_{\{t<\tau_i\}} \delta_{X_t^i}.$$

Notice that, after each default, we do not renormalize the mean reverting interaction in the drift as it is intended to reflect herding in investment decisions: if banks are defaulting, this suggests the investments are not performing and thus defaults should not suddenly adjust the drift upwards by a renormalization.

We emphasize that our model is intended for the study of short-run market imperfections rather than the long-run behaviour of the financial system. The parameters can be summarized as follows:

- σ and μ model, respectively, the volatility of the banks' assets and the return of the assets in the absence of any contagion or herding.
- ρ is the correlation parameter reflecting the extent of common exposures, that is, how correlated the assets are in the absence of any other interactions.
- α is the contagion parameter capturing that, upon bank i's default, the other banks lose a proportion αa_i^N of their assets (see Remark 3.1.2).

- π determines the rate of mean reversion which adjusts the core return μ due to herding in investment decisions or other interbank interactions.

- γ captures the effect a defaulted bank has on the drift, and could be influenced by actions of the central bank or government seeking to stabilize the system.

In order to capture system-wide influences on these variables, it is natural to allow them to depend on the empirical measure ν^N. In particular, the correlation can then act as an indirect source of contagion—in line with the observation that correlations tend to increase in times of financial distress (see e.g. Cont and Wagalath [37, 36]). Similarly, the rate of herding and the costliness of contagion can vary with the health of the system, which can capture potentially self-reinforcing amplification mechanisms. Furthermore, we will incorporate discontinuities in the dependence on the losses, L^N, which can allow for more abrupt adjustments of the herding, the correlation, or the costliness of defaults (see also Section 1.2.3). In this regard, we note that the idea of having a loss-dependent correlation with finitely many discontinuities (see part iii. of Assumption 3.2.1) was introduced in [73] as a possible way to accommodate for the implied correlation skew across the tranches of a CDO.

3.1.3 Outline of the chapter

In Section 3.2 we state our main results concerning the existence of a unique mean field limit for a general version of the model from Section 3.1.2. Furthermore, we discuss some qualitative insights for systemic risk. In Section 3.3 we study the regularity of the particle system, which is centred around the boundary and tail behaviour of the densities of the particles. The backbone of this is a novel family of upper Dirichlet heat kernel type estimates, which can be derived from the results of Chapter 2. In Section 3.4 we proceed to establish tightness of the system and we show that the resulting limit points are solutions to a nonlinear SPDE on the positive half-line. In Section 3.5 we rely on energy estimates to prove uniqueness of the SPDE and thus we deduce the full convergence in law to this limit. We present the uniqueness proof in Section 3.5.1.1, postponing the technical estimates to Sections 3.5.2 and 3.5.3. Therefore, the reader can get a complete picture of existence and uniqueness by only reading up to and including Section 3.5.1.

3.2 Main results

To make our framework as flexible as possible, we will consider a more general version of the model introduced in Section 3.1.2. Specifically, we will focus on particle systems of the form

$$\begin{cases} dX_t^i = b(t, X_t^i, \nu_t^N)dt + \sigma(t, X_t^i)\sqrt{1-\rho(t,\nu_t^N)}dW_t^i \\ \qquad\quad + \sigma(t, X_t^i)\rho(t,\nu_t^N)dW_t^0 - \alpha(t, X_t^i, \nu_t^N)d\mathfrak{L}_t^N, \\ \mathfrak{L}_t^N = (\mathfrak{K} \star L^N)_t, \quad L_t^N = 1 - \nu_t^N(0,\infty), \\ \nu_t^N = \sum_{i=1}^N a_i^N \mathbf{1}_{\{t<\tau_i\}}\delta_{X_t^i}, \quad \tau_i = \inf\{t \geq 0 : X_t^i \leq 0\}, \end{cases} \tag{3.5}$$

where $\mathfrak{K} \in \mathcal{W}_0^{1,1}(\mathbb{R}_+)$ and $W^0, \dots, W^N$ are independent Brownian motions. Concerning the weights, we assume that there exist $C, c > 0$ such that

$$a_i^N = \frac{a_i(X_0^i)}{\sum_{j=1}^N a_j(X_0^j)} \quad \text{with} \quad c \leq a_i(\cdot) \leq C \quad \text{for each} \quad i = 1, \dots, N. \tag{3.6}$$

For the well-posedness of (3.5) we refer to the beginning remarks of Section 3. As regards the model from Section 3.1.2, we stress that the drift b depends explicitly on the *mean process* $M^N := \langle \nu_\cdot^N, \mathrm{Id}\rangle$ and similarly the *loss process* L^N plays a vital rôle. More generally, these two mean field statistics can serve as useful indicators of systemic risk. However, their convergence as $N \to \infty$ does not follow directly from our notion of convergence for the empirical measures ν^N, so they will require special attention.

3.2.1 Structural conditions

In view of Section 3.1.2, we need local notions of Lipschitz continuity in X_t^i and ν_t^N, and we must allow the drift to have linear growth in both X_t^i and M_t^N. Thus, we need suitable notions of distance on the space of sub-probability measures $\mathbf{M}_{\leq 1}(\mathbb{R})$, which leads us to introduce the Kantorovich type distances

$$d_0(\mu,\tilde{\mu}) := \sup\{|\langle \mu - \tilde{\mu}, \psi\rangle| : \|\psi\|_{\mathrm{Lip}} \leq 1, |\psi(0)| \leq 1\}, \text{ and}$$
$$d_1(\mu,\tilde{\mu}) := \sup\{|\langle \mu - \tilde{\mu}, \psi\rangle| : \|\psi\|_{\mathrm{Lip}} \leq 1, \|\psi\|_\infty \leq 1\}.$$

Assumption 3.2.1. Let the coefficients of (3.5) be of the form $\sigma(t,x)$, $b(t,x,\mu,\ell^\mu)$, $\alpha(t,x,\mu,\ell^\mu)$, and $\rho(t,\mu,\ell^\mu)$, where $\ell^\mu := 1 - \mu(0,\infty)$. We assume:

i. Linear growth and space/time regularity: The map $x \mapsto g(t,x,\mu,\ell)$, for $g = b, \alpha$, is in $\mathcal{C}^2(\mathbb{R})$ and $(t,x) \mapsto \sigma(t,x)$ is in $\mathcal{C}^{1,2}([0,T]\times\mathbb{R})$. There exists $C > 0$ s.t.

$$|g(t,x,\mu,\ell)| \leq C(1 + |x| + \langle \mu, |\cdot|\rangle), \quad |\partial_x^{(n)} g(t,x,\mu,\ell)| \leq C, \quad n = 1,2,$$
$$|\sigma(t,x)| \leq C, \quad |\partial_t \sigma(t,x)| \leq C, \quad \textit{and} \quad |\partial_x^{(n)}\sigma(t,x)| \leq C, \quad n = 1,2.$$

ii. Local d_0/d_1-Lipschitz continuity in μ: Let $g = b, \alpha$. There exists $C > 0$ s.t.

$$|g(t,x,\mu,\ell) - g(t,x,\tilde{\mu},\ell)| \le C(1 + |x| + \langle \mu, |\cdot| \rangle) d_0(\mu,\tilde{\mu})$$
$$|\rho(t,\mu,\ell) - \rho(t,\tilde{\mu},\ell)| \le C(1 + \langle \mu, |\cdot| \rangle) d_1(\mu,\tilde{\mu})$$

iii. Piecewise local Lipschitz continuity in ℓ: There exist $0 = \theta_0 < \cdots < \theta_k = 1$ s.t.

$$\big|g(t,x,\mu,\ell) - g(t,x,\mu,\tilde{\ell})\big| \le C(1 + |x| + \langle \mu, |\cdot| \rangle)|\ell - \tilde{\ell}|$$

whenever $\ell, \tilde{\ell} \in [\theta_{i-1}, \theta_i)$ for some $1 \le i \le k$, where g is any of b, α, or ρ.

iv. Non-degeneracy: There exists $\epsilon > 0$ s.t. $0 < \epsilon \le \sigma$ and $0 \le \rho \le 1 - \epsilon$.

v. Sub-Gaussian initial law: $X_0^1, \ldots, X_0^N$ are i.i.d and independent of the driving Brownian motions. X_0^1 has law μ_0 with a density in $L^2(0,\infty)$ and

$$\exists \epsilon > 0 \quad \text{s.t.} \quad \mu_0(\lambda,\infty) = O(\exp\{-\epsilon\lambda^2\}) \quad \text{as} \quad \lambda \to \infty.$$

Moreover, $\nu_0^N := \sum_{i=1}^N a_i^N \delta_{X_0^i}$ converges weakly to a probability measure ν_0 with a density $V_0 \in L^2(0,\infty)$ s.t. $\|xV_0\|_{L^2} = \int_0^\infty |xV_0(x)|^2 dx < \infty$.

Remark 3.2.2. In (i) it suffices for $\partial_{xx} b$, $\partial_{xx}\sigma$ and $\partial_t \sigma$ to exist as weak derivatives in L^∞, and if $\sigma(t,x) = \sigma_1(t)\sigma_2(x)$ then no regularity in time is needed.

In order to establish the uniqueness of the limit SPDE, we impose some natural regularity conditions. As we will show in Section 3.4, these conditions are weaker than those that hold for the limit points of ν^N.

Definition 3.2.3. Let $\mathcal{D} = \mathcal{D}[0,T]$ denote the class of processes $\tilde{\nu} = (\tilde{\nu}_t)_{t\in[0,T]}$, whose marginals $\tilde{\nu}_t$ take values in $\mathbf{M}_{\le 1}(\mathbb{R})$ and satisfy:

i. Support on $\mathbb{R}_+$: For every $t \in [0,T]$, $\tilde{\nu}_t$ is supported on $\mathbb{R}_+ = [0,\infty)$.

ii. Increasing loss: $\tilde{L}_t := 1 - \tilde{\nu}_t(0,\infty)$ is strictly increasing when $\tilde{L} < 1$.

iii. Exponential tails: $\tilde{M}_t := \langle \tilde{\nu}_t, \mathrm{Id} \rangle < \infty$ and, for every $\epsilon > 0$, we have

$$\mathbb{E}\big[\textstyle\int_0^T \tilde{\nu}_t(a,\infty)dt\big] = o\big(\exp(-\epsilon a)\big) \quad \text{as} \quad a \to \infty.$$

iv. Boundary decay: There exists $\beta > 0$ such that

$$\mathbb{E}\big[\textstyle\int_0^T \tilde{\nu}_t(0,\varepsilon)dt\big] = O(\varepsilon^{1+\beta}) \quad \text{as} \quad \varepsilon \to 0.$$

v. *Spatial concentration*: There exist $C > 0$ and $\delta > 0$ such that

$$\mathbb{E}\big[\textstyle\int_0^T |\tilde{\nu}_t(a,b)|^2 dt\big] \le C|b-a|^\delta \quad \forall a, b \in \mathbb{R}.$$

3.2.2 The limit SPDE

Let $\mathscr{S}$ denote the space of Schwartz functions on $\mathbb{R}$ and let $\mathscr{S}'$ denote its dual, the space of tempered distributions. Then we can think of the empirical measures ν^N as living in the space $D_{\mathscr{S}'} = D_{\mathscr{S}'}[0,T]$ consisting of $\mathscr{S}'$-valued càdlàg processes on $[0,T]$. With a view to exploiting the monotonicity of the loss process, L^N, it becomes natural to use weak convergence with respect to Skorohod's M1 topology on $D_{\mathscr{S}'}$. The extension of the M1 topology to this setting was introduced in [86] and further details can be found there.

It remains to mention that we consider the limit SPDE in its weak formulation with respect to the space of test functions $\mathscr{C}_0$ as defined by

$$\mathscr{C}_0 := \{\phi \in \mathscr{S} : \phi(0) = 0\}.$$

This encodes the idea that the SPDE is posed as a Dirichlet problem on $\mathbb{R}_+$.

Theorem 3.2.4 (Existence of the limiting SPDE)**.** *If Assumption 3.2.1 is satisfied, then (ν^N, W^0) has limits points (ν, W^0) on $(D_{\mathscr{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$, which are continuous $\mathbf{M}_{\leq 1}(\mathbb{R}_+)$-valued processes in the class $\mathcal{D}$ from Definition 3.2.3. Moreover, almost surely, each (ν, W^0) obeys the limit SPDE*

$$\begin{aligned} d\langle \nu_t, \phi\rangle &= \langle \nu_t, b(t,\cdot,\nu_t)\partial_x\phi\rangle dt + \tfrac{1}{2}\langle \nu_t, \sigma(t,\cdot)^2\partial_{xx}\phi\rangle dt \\ &\quad + \langle \nu_t, \sigma(t,\cdot)\rho(t,\nu_t)\partial_x\phi\rangle dW_t^0 - \langle \nu_t, \alpha(t,\cdot,\nu_t)\partial_x\phi\rangle d\mathfrak{L}_t, \end{aligned} \tag{3.7}$$

with $\mathfrak{L}_t := \int_0^t \mathfrak{K}(t-s)L_s ds$ and $L_t := 1 - \nu_t(0,\infty)$, for all $t \in [0,T]$ and $\phi \in \mathscr{C}_0$. Finally, if ν is attained along a particular subsequence $(\nu^{N_k}, W^0)_{k\geq 1}$, then (L^{N_k}, W^0) and (M^{N_k}, W^0) converge weakly to (L, W^0) and (M, W^0) on $(D_{\mathbb{R}}, \mathrm{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$, where $M_t := \langle \nu_t, \mathrm{Id}\rangle$.

It is instructive to think of the measure-valued SPDE (3.7) in terms of the equation for its density process, V_t, which exists in L^2 by Corollary 3.5.4. Based on the boundary decay from Theorem 3.2.7 below, a simple Borel–Cantelli argument implies $V_t(0) = 0$ in the sense that $\lim_{\varepsilon \downarrow 0} \varepsilon^{-1}\nu_t(0,\varepsilon) = 0$, see Lemma A.1.4. Therefore, integrating by parts in (3.7) leads to the following observation.

Remark 3.2.5 (SPDE for the density). If ν is a limit point from Theorem 3.2.4, then it has a density process V_t in $L^2(\mathbb{R}_+)$ which, in the weak sense, satisfies the Dirichlet problem

$$\begin{aligned} dV_t(x) &= \tfrac{1}{2}\partial_{xx}\big(\sigma(t,x)^2 V_t(x)\big)dt - \partial_x\big(b(t,x,\nu_t)V_t(x)\big)dt \\ &\quad - \rho(t,\nu_t)\partial_x\big(\sigma(t,x)V_t(x)\big)dW_t^0 + \partial_x\big(\alpha(t,x,\nu_t)V_t(x)\big)d\mathfrak{L}_t, \quad V_t(0) = 0, \end{aligned}$$

with $\mathfrak{L}_t = (\mathfrak{K} \star L)_t$ and $L_t = 1 - \int_0^\infty V_t(x)dx$. Pretending that the derivative at zero exists, we can use this SPDE for V and integrate by parts to obtain the formal expression

$$\frac{d}{dt}\mathfrak{L}_t = \frac{1}{2}\int_0^t \mathfrak{K}(t-s)\sigma(s,0)\partial_x V_s(0)ds.$$

That is, at least formally, the contagion term driven by $\mathfrak{L}$ acts like an extra transport term proportional to the flux across the boundary—smoothed in time according to $\mathfrak{K}$. (In the case of constant coefficients and $\alpha \equiv 0$, sharp regularity results can be found in [85]).

Since every limit point of the finite particle system obeys the SPDE (3.7), we can deduce the full weak convergence to this SPDE once we have uniqueness in the class of solutions satisfying the conditions from Definition 3.2.3.

Theorem 3.2.6 (Uniqueness of the SPDE and functional LLN)**.** *Let (ν, W^0) be as in Theorem 3.2.4 and suppose $\tilde{\nu}$ is another solution to the SPDE (3.7) belonging to the class $\mathcal{D}$ from Definition 3.2.3. Then, almost surely,*

$$\nu_t(A) = \tilde{\nu}_t(A) \quad \forall t \in [0,T] \quad \forall A \in \mathcal{B}(\mathbb{R}).$$

In particular, any solution to the SPDE (3.7) has a unique law on the product space $(D_{\mathscr{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$ and (ν^N, W^0) converges weakly to this law. Furthermore, jointly with W^0, the loss process L^N and the mean process M^N converge weakly to L and M as defined in Theorem 3.2.4.

In view of the pathwise uniqueness, a Yamada–Watanabe argument ensures that the unique solution to the SPDE can be taken to be strong on a suitable probability space. Our final result shows that the unique solution to the SPDE can thus be recast as the conditional law, given the common Brownian motion W^0, of a 'conditional' McKean–Vlasov type diffusion with absorption at the origin. Furthermore, we establish an Aronson type upper bound on the density of this absorbed diffusion and show that it has power law decay near the origin.

Theorem 3.2.7 (Conditional McKean–Vlasov formulation)**.** *Let (ν, W^0) be the unique strong solution to the SPDE (3.7). Then, for any Brownian motion $W \perp (X_0, W^0)$, we have*

$$\nu_t = \mathbb{P}\big(X_t \in \cdot\,,\, t < \tau \mid W^0\big) \quad \textit{for} \quad \tau := \inf\{t > 0 : X_t \leq 0\},$$

where X_t is the unique solution to the conditional McKean–Vlasov diffusion

$$\begin{cases} dX_t = b(t, X_t, \nu_t)dt + \sigma(t, X_t)\sqrt{1-\rho(t,\nu_t)^2}dW_t \\ \qquad\quad + \sigma(t, X_t)\rho(t, \nu_t)dW_t^0 - \alpha(t, X_t, \nu_t)d\mathfrak{L}_t \\ \mathfrak{L}_t = (\mathfrak{K} \star L)_t, \quad L_t = \mathbb{P}(\tau \le t \mid W^0), \quad X_0 \sim \nu_0. \end{cases}$$

Moreover, each marginal of the process absorbed at zero has a (sub-probability) density $\mathfrak{p}(t,x) = \mathbb{E}[V_t(x)]$, which satisfies the following bounds: Defining

$$\varphi_t^{\kappa,\varepsilon}(x,y) := \frac{1}{\sqrt{t}}\Big(\frac{x}{\sqrt{t}} \wedge 1\Big)\Big(\frac{y}{\sqrt{t}} \wedge 1\Big) + \Big(\frac{x^\kappa y^\kappa}{t^\kappa} \wedge 1\Big)e^{\epsilon y^2},$$

we have that, for any $\epsilon > 0$, there exists $\kappa \in (0,1)$ and $C, c > 0$ such that

$$\mathfrak{p}(t,x) \le C\int_0^\infty \varphi_t^{\kappa,\varepsilon}(x,y)\big(1 \wedge e^{-\frac{(x-y)^2}{ct} + c_{x,y}}\big)d\nu_0(y) \tag{3.8}$$

and

$$\mathfrak{p}(t,x) \le C\int_0^\infty \big(t^{-\frac{1}{2}} + e^{\epsilon y^2}\big)e^{-\frac{(x-y)^2}{ct}}d\nu_0(y), \tag{3.9}$$

where $c_{x,y} \le c_0|x-y|(x \wedge y)$ for some $c_0 > 0$. If $\sigma(t,x) = \sigma_1(t)\sigma_2(x)$, then $c_{x,y} \equiv 0$.

Proof. Given Theorems 3.2.4 and 3.2.6, the first claim is straightforward. Indeed, treating ν_t as given, the SDE has a strong solution, $\tilde{X}$, by the standard theory, so we can define $\tilde{\nu}_t := \mathbb{P}[\tilde{X}_t \in \cdot,\ t < \tau \mid W^0]$. Arguing as in Section 9 of [73], with the obvious changes, we can apply Itô's formula to $\phi(\tilde{X}_t)$, for $\phi \in \mathscr{C}_0$, to show that $\tilde{\nu}$ solves a linear version of the SPDE (3.7) with the fixed ν. Hence $\tilde{\nu} = \nu$ by uniqueness (for the linear SPDE) which proves the claim. The density estimates (3.8) and (3.9) are more involved, but they will follow from Proposition 3.4.2 based on the work in in Chapter 2. □

Remark 3.2.8 (Gaussian tails). The bound (3.9) also holds on the whole space. Note that the correction term $c_{x,y}$ in (3.8) is only relevant when $x, y \to \infty$ jointly (inside a cone determined by $c_{x,y}$) and in this case the tail is controlled by (3.9). Note also that the factor $e^{\epsilon y^2}$ requires the initial law to be sub-Gaussian. This is a natural consequence of the linear growth. If e.g. the drift is bounded, then $e^{\epsilon y^2}$ can be dropped and κ can be taken arbitrarily close to 1. Similarly, the factor $e^{\epsilon y^2}$ drops if the linear growth is in terms of $|M_t - M_0|$ and $|X_t - X_0|$.

For the proofs of Theorems 3.2.4 and 3.2.6, our techniques build upon and extend the methods of [73], which dealt with a similar SPDE problem, albeit without the contagion term (i.e. $\alpha \equiv 0$) and with a bounded drift as well as with coefficients

that only depend on ν via the losses L. The main insight from [73] is that it can be fruitful to approach uniqueness via suitable energy estimates in H^{-1} (the dual of the first Sobolev space) when combined with careful control over $\mathbb{E}[\nu_t(0,\varepsilon)]$ as $\varepsilon \downarrow 0$ and $\mathbb{E}[\nu_t(\lambda,\infty)]$ as $\lambda \uparrow \infty$. However, in order for this to work in our setting, several extensions of the arguments in [73] are needed and, crucially, we rely on novel upper bounds for the densities of the absorbed particles (Proposition 3.3.3). Ultimately, we thus arrive at the essential H^{-1} energy estimate (Proposition 3.5.1 and Lemma 3.5.2), by establishing power law decay of $x \mapsto \mathbb{E}[V_t(x)]$ near the Dirichlet boundary at zero and Gaussian tails towards infinity (Corollary 3.3.4).

3.2.3 Financial contagion and default clustering

Recall that the nonlinearities of the limit SPDE are of a non-local nature tied to the flux across the origin. This differentiates our setting from the existing theory for Zakai type SPDEs and it has striking consequences for the qualitative behaviour of the solution.

In particular, the health of the financial system—governed by the limit SPDE—depends critically on the interplay between the common noise W^0 and the nonlinear effects of the contagion process $\mathfrak{L}$ (see Figure 3.2.1). At least conceptually, this captures the main forces at work in the 2007–2009 financial crisis, where contagion ensued from the correlated corrections across the U.S. housing markets as *"financial institutions had levered up on similar large portfolios of securities and loans that faced little idiosyncratic risk, but large amounts of systematic risk"* (Acharya et al. [1]).

The left-hand plot in Figure 3.2.1 illustrates how contagion can cause a period of significant default clustering (from $t = 0.2$ to $t = 0.4$), with the right-hand plot confirming that the system would have done just fine in the absence of contagion. Indeed, the entire lower group of 'unhealthy' banks default in the left plot, while the right plot shows a comfortable recovery from the slight initial deterioration (caused by the common exposures).

Notice also that the demise of the lower 'unhealthy' part of the system causes a substantial drop in the distance-to-default of the 'healthier' upper part. However, these problems do not result in further defaults, so the contagious effects die out. As illustrated by Figure 3.2.2, a sharper decline of the common exposures can prompt much more severe contagion, which in turn can result in a default cascade that also wipes out the healthier upper part (over a very short period of time).

Finally, we emphasise the critical rôle of common exposures as an instigator of such periods of high contagion. This is made clear by the rightmost plot of Figure 3.2.2,

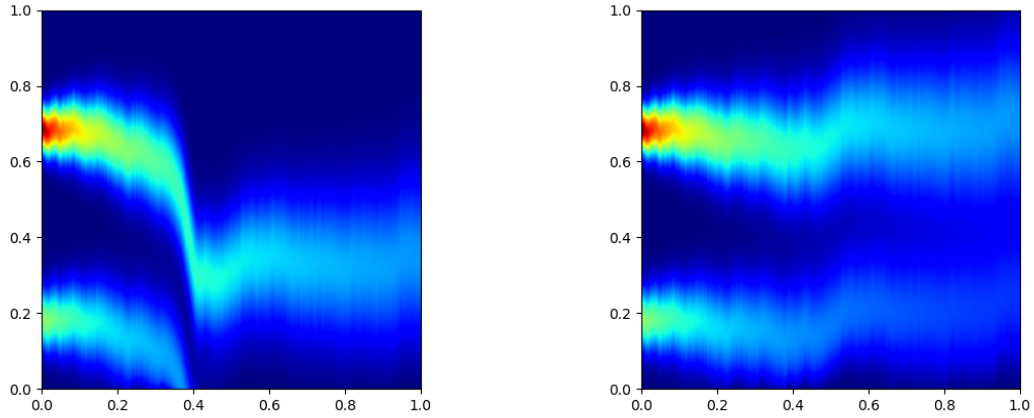


Figure 3.2.1: The figure shows two heat plots of the density $(t, x) \mapsto V_t(x)$ from Remark 3.2.5, based on a numerical simulation of the SPDE for a fixed realization of W^0. On the left $\alpha = 1.5$, while on the right $\alpha = 0$. The other parameters are: $\rho = 0.1$, $\sigma = 1$, $b = 0$, and $\mathfrak{K}$ is an isosceles triangle on $[0, 0.015]$ with height $2/0.015$. The realization of W^0 starts on a slightly negative trend, but then moves back up again.

which shows the system trending solidly upwards in a contrasting scenario where the common exposures are doing well. The importance of the common noise (as portrayed by Fig. 3.2.2) agrees with observations from the network-based literature, which suggest that idiosyncratic shocks are unlikely to significantly impact a large network, whereas the addition of common shocks can generate substantial losses from contagion (see e.g. the discussion in Cont et al. [34]). Also, recalling the model from Section 3.1.2, we note that a high rate of herding can have the dual effect of producing a healthier system in 'normal' times, while serving as an amplifier of potential crises by causing even more default clustering if the common exposures decline significantly. Similarly, the correlation function presents another endogenous channel that can amplify the effect of declining common exposures. We intend to return to a closer study of the interplay between the different parameters in future work.

3.2.3.1 Default cascades and the limiting case of instantaneous contagion

In view of the steep decline in Figure 3.2.2, it is interesting to consider what happens as the contagious impact of defaults becomes instantaneous. That is, as a sequence of impact kernels $(\mathfrak{K}_\varepsilon)$ approximates the dirac mass at zero.

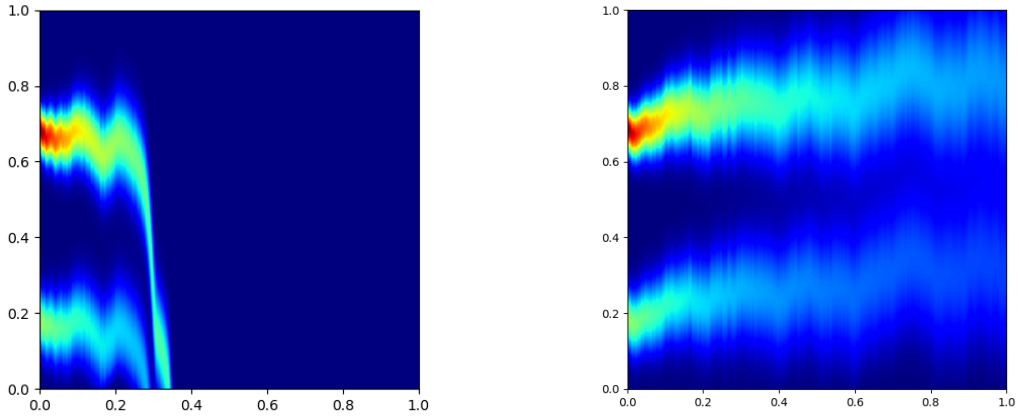


Figure 3.2.2: In these two heat plots of $(t, x) \mapsto V_t(x)$, we have $\alpha = 2$ on both sides and otherwise the parameters are the same as in Fig. 3.2.1. However, the simulations are done for two different realizations of the common noise W^0: on the left it declines steadily, while on the right it increases correspondingly.

In the case of constant coefficients and no common noise, the resulting limiting problem has recently been studied independently by Nadtochiy and Shkolnikov [94] and the author of this thesis together with Ledger and Hambly [74], building on earlier work by Delarue et al. [45, 46]. To be specific, the limiting object must be understood as a McKean–Vlasov problem of the form

$$X_t = X_0 + bt + \sigma W_t - \alpha L_t, \quad L_t = \mathbb{P}(\tau \leq t), \tag{3.10}$$

where $\tau = \inf\{t > 0 : X_t \leq 0\}$. As it is, (3.10) is ill-posed, however, as we discuss in Chapter 4, it has been conjectured that it is well-posed in the class of 'physical' solutions as first introduced in [46]. In fact, at the time of finishing this thesis, the conjecture has been resolved in the case of constant coefficients by the very recent preprint [47]. We discuss these development in depth in the final part of Chapter 6.

At the time when the results of this chapter were developed (i.e., at the time of preparing the preprint [75]), the following was known about solutions to (3.10). Global existence of a physical solution could be readily deduced from the results in [46] (see also Chapter 5), and, if $\alpha > 0$ is sufficiently small, then the earlier result [45, Thm. 2.4] gives that there is a unique global solution such that $t \mapsto L_t$ is in $\mathcal{C}^1[0, T]$. However, uniqueness for a general $\alpha > 0$ remained unsettled: If α is sufficiently large (given X_0), then it was noted in [74] that $t \mapsto L_t$ cannot be continuous, and uniqueness in the class of physical solutions was then only known up to the first

(finite, but strictly positive) time when the $\mathcal{W}^{1,2}$ norm of L explodes. These results are the subject of Chapter 4 of this thesis, see in particular Theorem 4.2.8.

Mathematically, the phase transition in α is very interesting: it means that the steep decline in Figure 3.2.2 may degenerate into a jump discontinuity whereby a macroscopic part of the system is lost at the blink of an eye. Financially, such a jump could offer an idealized definition of a true 'systemic default cascade', in line with analysis of [94] for a variant of (3.10). However, beyond the benefits of a precise definition, this may be too categorical, as instantaneous default cascades are not observed in practice and, from a systemic risk point of view, it is the default clustering and steep declines in distances-to-default that matter, not whether they materialized over short periods of time or as jumps. Therefore, we believe the framework proposed in this chapter can serve as the reference model, with the instantaneous problem (3.10) arising as an important limiting case. We analyse this limiting case—and its variations—in Chapters 4, 5, and 6.

For the purposes of this chapter, however, we emphasize that there are two theoretical advantages of the model presented in this chapter. Firstly, it needs no extra notion of 'physical' solutions and it makes sense as an SPDE globally (whereas, in the instantaneous case, one can only hope to have a notion of an SPDE in the intervals between blow-ups). Secondly, the SPDE fully characterizes the unique mean field limit of the finite system, which is currently not guaranteed for (3.10) except for the case of constant coefficients where the aforementioned conjecture has now been resolved by [47] (see Chapter 6). Consequently, this issue remains pronounced with a common source of noise and with the full generality of the coefficients, as nothing is then known about uniqueness of the corresponding McKean–Vlasov problem (aside from the result for the weak feedback regime discussed in Chapter 6).

3.2.3.2 Connections to mathematical neuroscience

Interestingly, our setup is closely related to nonlinear *leaky integrate-and-fire* models for electrically coupled neurons with noisy input. These models can be phrased as particle systems, where each SDE corresponds to the electrical potential of a neuron, and when this potential reaches a threshold voltage, the neuron is then said to *spike*, thus causing it to fire an electrical signal to the other neurons exciting them to higher voltage levels.

As suggested by Inglis and Talay [79], the transmission of this signal can be modelled by a cable equation, which translates to a gradual impact of the signal in analogy with our contagion mechanism. However, instead of being absorbed at the

boundary, the neurons that spike are instantly reset to a predetermined value (called the resting potential) and they then continue to evolve according to this rule ad infinitum. In the mean field limit, this yields a McKean–Vlasov problem analogous to that of Theorem 2.7 except for the resetting. So far, this has only been studied without common noise (and with simpler interactions) in which case the global well-posedness was proved in [79], building on ideas from [45, 46]. See [13, 92, 96, 104] for further background on the neuroscience literature and the relevance of common noise.

3.3 The finite particle system

We begin by observing that the finite system (3.5) is well-posed for each $N \geq 1$. To see this, we write $\mathbf{X}^N = (X^1, \dots, X^N)$ and express the system as a vector-valued SDE

$$d\mathbf{X}_t^N = \boldsymbol{b}(t, \mathbf{X}_t^N)dt + \boldsymbol{\sigma}(t, \mathbf{X}_t^N)\big(\rho_t d\mathbf{W}_t^0 + (1-\rho_t^2)^{\frac{1}{2}} d\mathbf{W}_t\big) - \boldsymbol{\alpha}(t, \mathbf{X}_t^N)d\mathfrak{L}_t^N,$$

where the dependence on ν_t^N is captured as a function of $\mathbf{X}_t^N$.

Remark 3.3.1. In the above, $\rho_t = \rho(t, \nu_t^N)$. Similarly, we will sometimes use the shorthand notation $\alpha_t^i = \alpha(t, X_t^i, \nu_t^N)$ and $b_t^i = b(t, X_t^i, \nu_t^N)$ as well as $\sigma_t^i = \sigma(t, X_t^i)$.

Recall from (iii) of Assumption 3.2.1 that the Lipschitz continuity in L_t^N only holds piecewise. However, on each of the intervals between defaults, L_t^N is simply equal to a fixed $\mathcal{F}_0$-measurable random variable with $L_t^N = 0$ on $[0, \varsigma_1)$ and then $L_t^N = \sum_{k=1}^n a_{i_k}^N$ on $[\varsigma_n, \varsigma_{n+1})$, where ς_n is the time of the n'th default (and $i_1, \dots, i_n$ have defaulted). Thus, on each of these intervals we can treat the coefficients as functions of just t and the surviving members of $\mathbf{X}_t^N$, with (i)-(ii) of Assumption 3.2.1 giving local Lipschitz continuity in the euclidean norm. That is, we can solve the system inductively on each of the intervals by running it up to the next default and then restarting it with the new (fixed) value for L_t^N and with the defaulted particle removed from ν_t^N. Hence the well-posedness follows by the standard theory for SDEs with locally Lipschitz coefficients of at most linear growth.

3.3.1 Evolution equation for the finite system

Since we have sufficient symmetry in the coefficients of the particles, we can obtain a single evolution equation for the dynamics of the empirical measures. Furthermore, the assumptions on the weights a_i^N will ensure that enough averaging is taking place in order for the idiosyncratic noise in this equation to vanish in the large population

limit. These observations are made precise in the proposition below and the work that follows. In the proof of the next result, and in the rest of the chapter, we will use the notation '$\lesssim$' to denote that an inequality holds up to a fixed numerical constant.

Proposition 3.3.2 (Evolution of the empirical measures). *For all $N \geq 1$, we have*

$$\begin{aligned} d\langle \nu_t^N, \phi\rangle &= \langle \nu_t^N, b(t,\cdot,\nu_t^N)\partial_x\phi\rangle dt + \tfrac{1}{2}\langle \nu_t^N, \sigma(t,\cdot)^2\partial_{xx}\phi\rangle dt \\ &\quad + \langle \nu_t^N, \sigma(t,\cdot)\rho_t\partial_x\phi\rangle dW_t^0 - \langle \nu_t^N, \alpha(t,\cdot)\partial_x\phi\rangle d\mathfrak{L}_t^N + dI_t^{\phi,N} \end{aligned} \tag{3.11}$$

for every $\phi \in \mathscr{C}_0$, where the idiosyncratic part, $I^{\phi,N}$, satisfies

$$\mathbb{E}\Big[\sup_{t\leq T}|I_t^{\phi,N}|^2\Big] = O(N^{-1}) \quad as \quad N\to\infty.$$

Proof. Notice that, since $\phi(0)=0$ for $\phi\in\mathscr{C}_0$, we have

$$\langle \nu_t^N,\phi\rangle = \sum_{i=1}^N a_i^N \mathbf{1}_{\{s<\tau_i\}}\phi(X_t^i) = \sum_{i=1}^N a_i^N\phi(X_{t\wedge\tau_i}^i), \quad \text{for} \quad \phi\in\mathscr{C}_0.$$

Applying Itô's formula to $\phi(X^i_{t\wedge\tau_i})$, we see that (3.11) holds with

$$I_t^{\phi,N} := \sum_{i=1}^N \int_0^t a_i^N\sigma(s,X_s^i)(1-\rho(s,\nu_s^N)^2)^{\frac{1}{2}}\partial_x\phi(X^i_{s\wedge\tau_i})dW_s^i.$$

By independence of the W^i's and the boundedness of σ and $\partial_x\phi$, we get

$$\mathbb{E}\big[\langle I_\cdot^{\phi,N}\rangle_t\big] = \sum_{i=1}^N \mathbb{E}\Big[\int_0^{t\wedge\tau_i}(a_i^N)^2(\sigma_s^i)^2(1-\rho_s^2)\partial_x\phi(X_s)^2ds\Big] \lesssim \sum_{i=1}^N \mathbb{E}\big[(a_i^N)^2\big].$$

Thus, the claim follows by Doob's martingale inequality, since $a_i^N \lesssim N^{-1}$. □

3.3.2 Regularity properties of the particles

The next result plays a central role in our analysis. It consists of a uniform family of upper Dirichlet heat kernel type estimates for the densities of the particles.

Proposition 3.3.3 (Absorbing density estimates). *Let X_t^i be given by (3.5) under Assumption 3.2.1. Then the absorbed process $(X_t^i, t<\tau_i)$ has a transition density $\mathfrak{p}_t^{i,N}$, which satisfies the following bounds: For every $\epsilon>0$ there exist $\kappa\in(0,1)$ and constants $C,c>0$, uniformly in $N\geq 1$ and $t\in(0,T]$, such that*

$$\mathfrak{p}_t^{i,N}(x,y) \leq C\Big(\frac{1}{\sqrt{t}}\Big(\frac{x}{\sqrt{t}}\wedge 1\Big)\Big(\frac{y}{\sqrt{t}}\wedge 1\Big) + \Big(\frac{x^\kappa y^\kappa}{t^\kappa}\wedge 1\Big)e^{\epsilon y^2}\Big)\big(1\wedge e^{-\frac{(x-y)^2}{ct}+c_{x,y}}\big) \tag{3.12}$$

and

$$\mathfrak{p}_t^{i,N}(x,y) \leq C\big(t^{-\frac{1}{2}} + e^{\epsilon y^2}\big)e^{-\frac{(x-y)^2}{ct}}, \tag{3.13}$$

where $c_{x,y}\lesssim |x-y|(x\wedge y)$. If $\sigma(t,x)=\sigma_1(t)\sigma_2(x)$, then $c_{x,y}\equiv 0$.

Proof. The proof is given in Chapter 2, see Section 2.4.3. □

The above estimates provide critical control over (the first moment of) the mass of ν^N near zero as well as Gaussian decay towards infinity (Corollary 3.4). As we verify in Proposition 3.4.3, these features carry over to the limit points of the particle system, which will be essential to the proof of the energy estimates in Section 5.

Corollary 3.3.4 (Decay of mass at the boundaries)**.** *For all $a, b \geq 0$ we have $\mathbb{E}[\nu_t^N(a,b)] \lesssim t^{-\frac{1}{2}}|b-a|$. Furthermore, it holds uniformly in $N \geq 1$ and $t \in (0,T]$ that*

$$\begin{cases} \exists \epsilon > 0: & \mathbb{E}[\nu_t^N(a,\infty)] = O\big(\exp(-\epsilon a^2)\big) \quad as \quad a \to \infty \\ \exists \delta \in (0,1],\, \beta > 0: & \mathbb{E}[\nu_t^N(0,\varepsilon)] = t^{-\frac{\delta}{2}} O(\varepsilon^{1+\beta}) \quad as \quad \varepsilon \to 0 \end{cases}$$

Proof. Recalling the definition of ν_t^N and using that $a_i^N \leq C/N$, it follows from Proposition 3.3.3 that, for any $(a,b) \subseteq \mathbb{R}_+$, we have

$$\mathbb{E}[\nu_t^N(a,b)] \leq \frac{C}{N} \sum_{i=1}^{N} \mathbb{P}\big(X^i_{t\wedge\tau_i} \in (a,b)\big) \leq C' \int_0^\infty \int_a^b f_t(x,y) dx d\mu_0(y),$$

where $f_t(x,y)$ can take the value of either of the right-hand sides of (3.12) or (3.13). Given (3.13) and the sub-Gaussianity of μ_0, the first two claims are immediate. Similarly, the final claim follows from (3.12) by exploiting the power law decay at the boundary for $t > 0$. □

3.3.3 Regularity properties of the loss process

Below we present two important results about the limiting behaviour of the loss process. The first result assures that, in the large population limit, it is strictly increasing when there is mass left in the system. This is crucial for the convergence to the limit SPDE (Section 3.4.2) and for uniqueness (Section 3.5.1.1) as it ensures the loss process cannot get stuck at one of the coefficients' finitely many discontinuity points (Assumption 3.2.1(iii)).

Proposition 3.3.5. *For any $t \in [0,T)$ and $h > 0$, it holds for all $r < 1$ that*

$$\lim_{\delta \to 0} \limsup_{N \to \infty} \mathbb{P}(L^N_{t+h} - L^N_t < \delta,\, L^N_t < r) = 0.$$

Proof. See Section A.2 in the Appendix. □

While the system is progressively losing mass by the previous proposition, the next result ensures that there cannot be too large losses in arbitrarily small amounts of time. This is used in the tightness arguments below (Section 3.4).

Proposition 3.3.6. *For every $t \in [0,T]$ and $\eta > 0$, we have*

$$\lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P}(L^N_{t+\delta} - L^N_t \geq \eta) = 0.$$

Proof. See Section A.2 in the Appendix. □

The above relies on the impact kernel $\mathfrak{K}$ being in $\mathcal{W}^{1,1}$. As discussed in Section 3.2.3.1, instantaneous contagion can result in the limit loss process having jumps.

3.4 Tightness and convergence

The aim of this section is to recover the SPDE (3.7) by passing to the limit in the finite evolution equation (3.3.2). To this end, we need to prove tightness of (ν^N) and then we require some continuity results for the integrals in (3.3.2).

Lemma 3.4.1. *For all $s,t \in [0,T]$, we have*

$$\mathbb{E}\big[|X^i_{t\wedge\tau_i} - X^i_{s\wedge\tau_i}|^4\big] = O(|t-s|^2) \quad as \quad |t-s| \to 0,$$

uniformly in $N \geq 1$ and $i = 1,\dots,N$, where $\tau_i = \inf\{t > 0 : X^i_t \leq 0\}$.

Proof. Using the equation for X^i_t with $dB^i_t = \sqrt{1-\rho_t^2}dW^i_t + \rho_t dW^0_t$, we have

$$\mathbb{E}\Big[\big|X^i_{t\wedge\tau_i} - X^i_{s\wedge\tau_i}\big|^4\Big] \lesssim \mathbb{E}\Big[\Big(\int_s^t \sigma_r \mathbf{1}_{r<\tau_i} dB^i_r\Big)^4\Big] + \mathbb{E}\Big[\Big(\int_s^t \big|b^i_r - \alpha^i_r (\mathfrak{L}^N)'_t\big| dr\Big)^4\Big].$$

Since σ_u is bounded, the Burkholder–Davis–Gundy inequality yields

$$\mathbb{E}\Big[\Big(\int_s^t \sigma_r \mathbf{1}_{\{r<\tau_i\}} dB^i_r\Big)^4\Big] \lesssim \mathbb{E}\big[\langle B_\cdot - B_s\rangle^2_t\big] = (t-s)^2.$$

Letting $\Lambda^{i,N}_r := |X^i_t| + \sum_{j=1}^N a^N_j |X^j_t|$, the conditions in Assumption 3.2.1 imply the bound $|b^i_r| + |\alpha^i_r| \lesssim 1 + \Lambda^{i,N}_r$. Noting also that $(\mathfrak{L}^N)' \in L^\infty$, Jensen's inequality gives

$$\mathbb{E}\Big[\Big(\int_s^t \big|b^i_r - \alpha^i_r (\mathfrak{L}^N)'_r\big| dr\Big)\Big] \leq C(t-s)^4 \sup_{r \leq T} \mathbb{E}\big[(1+\Lambda^{i,N}_r)^4\big],$$

where last term is bounded uniformly in $N \geq 1$, by (3.13) of Proposition 3.3.3. □

Armed with this lemma and the control on the loss process from Section 3, the tightness now follows fairly easily from our choice of topology.

Proposition 3.4.2 (Tightness). *The sequence* (ν^N, W^0) *is tight on* $(D_{\mathscr{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$ *and any limit point* ν^* *satisfies* $\nu_t^* \in \mathbf{M}_{\leq 1}(\mathbb{R}_+)$ *for all* $t \in [0,T]$*, almost surely. Moreover, if we define* $L_t^* := 1 - \nu_t^*(0,\infty)$*, then* L^* *is strictly increasing when* $L^* < 1$*, and* (L^N, W^0) *converges weakly to* (L^*, W^0) *on* $(D_{\mathbb{R}}, \mathrm{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$ *whenever* (ν^N, W^0) *converges weakly to* (ν^*, W^0) *on* $(D_{\mathscr{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$.

Proof. For the tightness of (ν^N, W^0), it suffices to show the M1 tightness of $\langle \nu^N, \phi \rangle$ on $D_{\mathbb{R}}$ for every $\phi \in \mathscr{S}$, by Theorem 3.2 in [86]. To this end, we verify the sufficient conditions (i) and (ii) from Theorem 12.12.3 in [105]. The first condition is trivial since $\langle \nu^N, \phi \rangle$ is uniformly bounded by $|\langle \nu^N, \phi \rangle| \leq \|\phi\|_\infty$ for all $N \geq 1$. For the second condition, we can consider the decomposition

$$\langle \nu_t^N, \phi \rangle = \langle \hat{\nu}_t^N, \phi \rangle - \phi(0) L_t^N \quad \text{with} \quad \hat{\nu}_t^N := \textstyle\sum_{i=1}^N a_i^N \delta_{X^i_{t \wedge \tau_i}}. \tag{3.14}$$

The advantage of this is that the monotone part, $\phi(0)L_t^N$, is immaterial to the M1 modulus of continuity. Indeed, by Propositions 4.1 and 4.2 of [86], it is sufficient to verify that

$$\mathbb{E}\Big[\big|\langle \hat{\nu}_t^N, \phi \rangle - \langle \hat{\nu}_s^N, \phi \rangle\big|^4\Big] = O(|t-s|^2) \quad \forall s,t \in [0,T] \tag{3.15}$$

and that, for every $\varepsilon > 0$,

$$\lim_{\delta \to 0} \lim_{N \to \infty} \mathbb{P}\Big[\sup_{t \leq \delta} |\langle \nu_t^N - \nu_0^N, \phi \rangle| + \sup_{t \in (T-\delta, T)} |\langle \nu_T^N - \nu_t^N, \phi \rangle| > \varepsilon\Big] = 0. \tag{3.16}$$

Since $a_i^N \leq C/N$, Jensen's inequality and the Lipschitz continuity of $\phi \in \mathscr{S}$ imply

$$\mathbb{E}\Big[\big|\langle \hat{\nu}_t^N, \phi \rangle - \langle \hat{\nu}_s^N, \phi \rangle\big|^4\Big] \leq \frac{C}{N} \|\phi\|_{\mathrm{Lip}} \sum_{i=1}^N \mathbb{E}\Big[\big|X^i_{t \wedge \tau_i} - X^i_{s \wedge \tau_i}\big|^4\Big].$$

Thus, we can conclude from Lemma 3.4.1 that (3.15) is satisfied. With regard to (3.16), the decomposition (3.14) yields

$$\mathbb{P}\Big[\sup_{t \leq \delta} |\langle \nu_t^N - \nu_0^N, \phi \rangle| > \varepsilon\Big] \leq \mathbb{P}\Big[\sup_{t \leq \delta} |\langle \hat{\nu}_t^N - \hat{\nu}_0^N, \phi \rangle| \geq \frac{\varepsilon}{2}\Big] + \mathbb{P}\Big[|\phi(0)| L_\delta^N \geq \frac{\varepsilon}{2}\Big]$$

and likewise for the supremum over $t \in (T-\delta, T)$. By Markov's inequality and the same arguments as above, the first term vanishes uniformly in $N \geq 1$ as $\delta \to 0$. Combining this with Proposition 3.3.6, we deduce (3.16), which proves the tightness of (ν^N, W^0). For the final claims, Proposition 5.3 of [73] ensures that, for any limit point ν^*, each marginal ν_t^* can be recovered as an element of $\mathbf{M}_{\leq 1}(\mathbb{R}_+)$. Moreover, using Propositions 3.3.5 and 3.3.6, the claims about L^* and L^N follow by arguing as in Propositions 5.5 and 5.6 of [73]. □

As in the statement of Theorem 3.2.7, the decay of the mass near zero, for the limit points ν^*, will be captured by the function

$$\varphi_t^{\kappa,\varepsilon}(x,y) := \frac{1}{\sqrt{t}}\Big(\frac{x}{\sqrt{t}} \wedge 1\Big)\Big(\frac{y}{\sqrt{t}} \wedge 1\Big) + \Big(\frac{x^\kappa y^\kappa}{t^\kappa} \wedge 1\Big)e^{\epsilon y^2}.$$

Specifically, the uniform density estimates from Section 3 imply the following.

Proposition 3.4.3. *For every $\epsilon > 0$ there exist $\kappa \in (0,1)$ and $C, c > 0$ such that any continuous limit point ν^* of (ν^N) satisfies*

$$\begin{cases} \mathbb{E}[\nu_t^*(a,b)] \le C \displaystyle\int_0^\infty \int_a^b \varphi_t^{\kappa,\epsilon}(x,y)\big(1 \wedge e^{-\frac{(x-y)^2}{ct} + c_{x,y}}\big)dxd\nu_0(y) \\ \mathbb{E}[\nu_t^*(a,b)] \le C \displaystyle\int_0^\infty \int_a^b \big(t^{-\frac{1}{2}} + e^{\epsilon y^2}\big)e^{-\frac{(x-y)^2}{ct}}dxd\nu_0(y), \end{cases}$$

for any $(a,b) \subseteq \mathbb{R}_+$. Here $c_{x,y} \lesssim |x-y|(x \wedge y)$, and if $\sigma(t,x) = \sigma_1(t)\sigma_2(x)$ then $c_{x,y} \equiv 0$. In particular, ν^* satisfies the conditions of Definition 3.2.3 with the pointwise properties

$$\begin{cases} \exists \epsilon > 0 : & \mathbb{E}[\nu_t^*(\lambda,\infty)] = O\big(\exp(-\epsilon\lambda^2)\big) \quad as \quad \lambda \to \infty \\ \exists \delta \in (0,1],\, \beta > 0 : & \mathbb{E}[\nu_t^*(0,\varepsilon)] = t^{-\frac{\delta}{2}}O(\varepsilon^{1+\beta}) \quad as \quad \varepsilon \to 0 \end{cases}$$

Proof. Fix an interval $(a,b) \subseteq \mathbb{R}_+$ and fix a small $\varepsilon > 0$ s.t. $a - \varepsilon > 0$. Let φ_ε be a smooth cut-off function in $\mathcal{C}_c^\infty(\mathbb{R},[0,1])$ equal to 1 on (a,b) and vanishing outside $(a-\varepsilon, b+\varepsilon)$. Then, as in the proof of Corollary 3.3.4, we have

$$\mathbb{E}[\langle \nu_t^N, \varphi_\varepsilon\rangle] \le \mathbb{E}[\nu_t^N(a-\varepsilon, b+\varepsilon)] \le C\int_0^\infty \int_{a-\varepsilon}^{b+\varepsilon} f_t(x,y)dxd\nu_0(y), \tag{3.17}$$

where $f_t(x,y)$ can take the value of either of the right-hand sides of (3.12) and (3.13). Here we have also used that μ_0 is dominated by a constant times $\nu_0 = \lim_N \sum_{i=1}^N a_i^N \delta_{X_0^i}$, in light of (3.6). Since $\varphi_\varepsilon \in \mathscr{S}$, after passing to a convergent subsequence, $\langle \nu^{N_k}, \varphi_\varepsilon\rangle$ converges weakly to $\langle \nu^*, \varphi_\varepsilon\rangle$ on $D_\mathbb{R}$ by [86, Prop. 2.7]. Thus, by the assumption that ν^* is continuous, [105, Thm. 12.4.1] gives that $\langle \nu_t^{N_k}, \varphi_\varepsilon\rangle$ converges weakly to $\langle \nu_t^*, \varphi_\varepsilon\rangle$ on $\mathbb{R}$ for all $t \in [0,T]$. Using bounded convergence, we get convergence of the means and hence, for all $\varepsilon > 0$,

$$\mathbb{E}[\nu_t^*(a,b)] \le \mathbb{E}[\langle \nu_t^*, \varphi_\varepsilon\rangle] = \lim_{k\to\infty} \mathbb{E}[\langle \nu_t^{N_k}, \varphi_\varepsilon\rangle]$$

Recalling (3.17), the result follows by dominated convergence. □

Remark 3.4.4. If ν^* is not continuous, $\int_0^T \langle \nu_t^{N_k}, \varphi_\varepsilon\rangle dt \Rightarrow \int_0^T \langle \nu_t^*, \varphi_\varepsilon\rangle dt$ is still true by [105, Thm. 11.5.1], where '$\Rightarrow$' denotes weak convergence. Hence Assumption 3.2.3 holds for limit points of (ν^N) without a priori knowledge of continuity.

3.4.1 Convergence of the mean process

Let $M_t^N := \langle \nu_t^N, \psi \rangle$ for any given Lipschitz function $\psi \in \mathrm{Lip}(\mathbb{R})$. While we are mainly interested in the case $\psi = \mathrm{Id}$, other applications may call for a general Lipschitz function, so we allow for this in the ensuing arguments.

Given $\nu^N \Rightarrow \nu^*$ on $(D_{\mathscr{S}'}, \mathrm{M1})$, where '$\Rightarrow$' denotes weak convergence, we would like to know that $M^N \Rightarrow M^* := \langle \nu^*, \psi \rangle$ on $(D_{\mathbb{R}}, \mathrm{M1})$. However, ψ need not be in $\mathscr{S}$, so we cannot simply appeal to the continuity of the projection map from $(D_{\mathscr{S}'}, \mathrm{M1})$ to $(D_{\mathbb{R}}, \mathrm{M1})$. Nevertheless, we can work around this obstacle by exploiting the uniformly exponential tails of the empirical measures in expectation, cf. Proposition 3.3.3.

First of all, we can observe that $(M^N)_{N \geq 1}$ is tight on $(D_{\mathbb{R}}, \mathrm{M1})$. To see this, we verify the necessary conditions (i) and (ii) from Theorem 12.12.3 in [105]. The first condition amounts to showing that, for all $\varepsilon > 0$, there exists $c > 0$ such that

$$\mathbb{P}\Big[\sup_{t \leq T} |M_t^N| > c \Big] \leq \varepsilon \quad \forall N \geq 1.$$

This property is immediate, since we even have uniformly sub-Gaussian tails of the running max of M_t^N, by Corrollary 2.3.2. For the second condition, we can rely on the decomposition (3.14), which amounts to

$$M_t^N = \hat{M}_t^N - \psi(0) L_t^N, \quad \hat{M}_t^N := \langle \hat{\nu}_t^N, \psi \rangle.$$

Since ψ is Lipschitz, the tightness then follows by repeating the same arguments as in the proof of Proposition 3.4.2. Given this, we can now demonstrate the weak convergence of the mean process.

Proposition 3.4.5 (Functional LLN for the empirical means)**.** *Suppose* $(\nu^N, W^0) \Rightarrow (\nu^*, W^0)$ *on* $(D_{\mathscr{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$*. Then we have* $(M^N, W^0) \Rightarrow (M^*, W^0)$ *on the product space* $(D_{\mathbb{R}}, \mathrm{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$*.*

Proof. Let $(M^{N_k}, W^0)_{k \geq 1}$ be an arbitrary subsequence. By the tightness of (M^{N_k}), we can pass to a further convergent subsequence, also indexed by $k \geq 1$. Let $(M^\dagger, W^0)$ denote the weak limit of this subsequence and note that we still have $\nu^{N_k} \Rightarrow \nu^*$. We need to show that the laws of (M^*, W^0) and $(M^\dagger, W^0)$ agree. To this end, let $\phi_\lambda := \varphi_\lambda \psi$, where $\varphi_\lambda \in \mathcal{C}_c^\infty(\mathbb{R}; [0,1])$ is a standard cutoff function equal to 1 on $[-\lambda, \lambda]$. Then $\phi_\lambda \in \mathscr{S}(\mathbb{R})$ and $\phi_\lambda \to \psi$ pointwise as $\lambda \to \infty$. Since $\phi_\lambda \in \mathscr{S}$, the projection map $\pi^{\phi_\lambda} : (D_{\mathscr{S}'}, \mathrm{M1}) \to (D_{\mathbb{R}}, \mathrm{M1})$ is continuous (see [86, Prop. 2.7]), so we get

$$\langle \nu^{N_k}, \phi_\lambda \rangle \Rightarrow \langle \nu^*, \phi_\lambda \rangle \quad \text{on} \quad (D_{\mathbb{R}}, \mathrm{M1}), \tag{3.18}$$

by the continuous mapping theorem. Moreover, for all $\lambda > 0$, we have

$$\mathbb{E}\big[|\langle \nu_t^{N_k}, \psi - \phi_\lambda\rangle|\big] \le \mathbb{E}\big[\langle \nu_t^{N_k}, |\psi|\mathbf{1}_{[\lambda,\infty)}\rangle\big] \lesssim \mathbb{E}\big[\langle \nu_t^{N_k}, (1+|x|)\mathbf{1}_{[\lambda,\infty)}\rangle\big],$$

and hence the exponential tails (in expectation) imply

$$\mathbb{E}\big[|\langle \nu_t^{N_k}, \psi - \phi_\lambda\rangle|\big] = o(1) \text{ as } \lambda \to \infty, \text{ uniformly in } k \ge 1,\ t \le T, \tag{3.19}$$

see Lemma A.1.3. Next, we define

$$\mathbb{T} := \mathrm{cont}(M^\dagger) \textstyle\bigcap_{\lambda\in\mathbb{N}} \mathrm{cont}(\langle \nu^*, \phi_\lambda\rangle), \ \ \mathrm{cont}(\xi) = \{t \in [0,T] : \mathbb{P}[\xi_{t-} = \xi_t] = 1\},$$

and note that this set is co-countable since $\langle \nu^*, \phi_\lambda\rangle$ and $M^\dagger$ are in $D_{\mathbb{R}}$ (see e.g. [105, Cor. 12.2.1]). Let $\widetilde{\mathbb{T}} := \mathbb{T} \cap \{t \in [0,T] : \mathbb{P}[M_t^* < \infty] = 1\}$, where the latter set has full Lebesgue measure, so $\widetilde{\mathbb{T}}$ is dense in $[0,T]$. We will show that all the finite dimensional distributions of (M^*, W^0) and $(M^\dagger, W^0)$ agree for indices $\{t_1, \dots, t_l\} \in \widetilde{\mathbb{T}}$. By a monotone class argument, it suffices to show that

$$\mathbb{E}\bigg[\prod_{i=1}^l f_i(M_{t_i}^*) g_i(W_{t_i}^0)\bigg] = \lim_{k\to\infty} \mathbb{E}\bigg[\prod_{i=1}^l f_i(M_{t_i}^{N_k}) g_i(W_{t_i}^0)\bigg] \tag{3.20}$$

for bounded $f_i, g_i \in \mathrm{Lip}(\mathbb{R}; \mathbb{R}_+)$, $i = 1, \dots, l$, where we have used convergence of the marginals $(M_{t_i}^{N_k}, W_{t_i}^0)_{i=1,\dots,l} \Rightarrow (M_{t_i}^\dagger, W_{t_i}^0)_{i=1,\dots,l}$, see [105, Thm. 11.6.6]. By definition of M^N, we can observe that

$$\big|f_i(M_{t_i}^{N_k}) - f_i\big(\langle \nu_{t_i}^{N_k}, \phi_\lambda\rangle\big)\big| \le \|f_i\|_{\mathrm{Lip}} |\langle \nu_{t_i}^{N_k}, \psi - \phi_\lambda\rangle|, \quad i = 1, \dots, l. \tag{3.21}$$

Using the upper bounds on each $f_i(M_{t_i}^{N_k})$, as implied by (3.21), it follows from (3.19) and the boundedness of $f_i, g_i \ge 0$, that

$$\mathbb{E}\bigg[\prod_{i=1}^l f_i(M_{t_i}^{N_k}) g_i(W_{t_i}^0)\bigg] \le \mathbb{E}\bigg[\prod_{i=1}^l f_i(\langle \nu_{t_i}^{N_k}, \phi_\lambda\rangle) g_i(W_{t_i}^0)\bigg] + \{o(1) \text{ terms}\} \tag{3.22}$$

as $\lambda \to \infty$, where the $o(1)$ terms are uniform in $k \ge 1$. By construction of $\widetilde{\mathbb{T}}$, [105, Thm. 12.4.1] implies that the canonical projections $\pi_{t_i} : (D_{\mathbb{R}}, \mathrm{M1}) \to \mathbb{R}$ are continuous at each $\langle \nu^{N_k}, \phi_\lambda\rangle$ for $\lambda \in \mathbb{N}$. Thus, with $\langle \nu_{t_i}^{N_k}, \phi_\lambda\rangle = \pi_{t_i}\langle \nu^{N_k}, \phi_\lambda\rangle$, the weak convergence (3.18) implies that, by taking a lim sup over $k \ge 1$ on both sides of (3.22),

$$\lim_{k\to\infty} \mathbb{E}\bigg[\prod_{i=1}^l f_i(M_{t_i}^{N_k}) g_i(W_{t_i}^0)\bigg] \le \mathbb{E}\bigg[\prod_{i=1}^l f_i(\langle \nu_{t_i}^*, \phi_\lambda\rangle) g_i(W_t^0)\bigg] + o(1),$$

as $\lambda \to \infty$. Finally, we note that $\langle \nu_t^*, \phi_\lambda\rangle \to \langle \nu_t^*, \psi\rangle = M_t^*$ as $\lambda \to \infty$, by dominated convergence, since $|\phi_\lambda| \le |\psi|$ and ψ is integrable with respect to ν_t^* for all $t \in \widetilde{\mathbb{T}}$.

Consequently, by the boundedness and continuity of each (f_i, g_i), we can send $\lambda \to \infty$ to find that

$$\lim_{k\to\infty} \mathbb{E}\bigg[\prod_{i=1}^{l} f_i(M_{t_i}^{N_k}) g_i(W_{t_i}^0)\bigg] \leq \mathbb{E}\bigg[\prod_{i=1}^{l} f_i(M_{t_i}^{*}) g_i(W_{t_i}^0)\bigg].$$

Now, by relying instead on the lower bounds for each $f_i(M_{t_i}^{N_k})$ as implied by (3.21), analogous arguments yield the reverse inequality, thus proving (3.20). □

3.4.2 The limit SPDE: Proof of Theorem 3.2.4

By [86, Thm. 3.2], the tightness from Proposition 3.4.2 implies relative compactness in the sense that every subsequence of (ν^N, W^0) has a weakly convergent subsequence. Using this, we can show that the corresponding integrals in the finite evolution equation (3.11) converge weakly (Proposition 3.4.7) and then we employ a martingale argument (Proposition 3.4.8) to see that this gives rise to the desired limit SPDE.

Remark 3.4.6 (A Skorokhod Representation result). For the convergence results, we shall need a version of Skorokhod's representation theorem. While $(D_{\mathscr{S}'}, \mathrm{M1})$ fails to be Polish, it is, nevertheless, a completely regular Suslin space. In particular, condition (10) of [80, Thm. 2] holds and hence any weakly convergent subsequence of (ν^N) has a further subsequence with the usual almost sure Skorokhod representation property. Firstly, $(D_{\mathscr{S}'}, \mathrm{M1})$ is completely regular and Hausdorff by [86, Prop. 2.7] and [81, Thm. 2.1.1]. Secondly, [86, Prop. 2.7] and its proof gives $D_{\mathscr{S}'} = \bigcup_{n=1}^{\infty} D_{S_{\text{-}}n}$ where each $(D_{S_{\text{-}}n}, \mathrm{M1})$ is Polish and the inclusions $D_{S_{\text{-}}n} \hookrightarrow D_{\mathscr{S}'}$ are M1-continuous. Thus, each $D_{S_{\text{-}}n}$ is Suslin in the Hausdorff space $(D_{\mathscr{S}'}, \mathrm{M1})$ and hence their union is Suslin under the finest topology such that the inclusions are continuous [100, Thm. I.II.3]. As the M1-topology is coarser, we conclude that $(D_{\mathscr{S}'}, \mathrm{M1})$ is Suslin. In turn, [80, Thm. 2] indeed applies.

Proposition 3.4.7 (Functional continuity). *Fix $t \leq T$ and $\phi \in \mathscr{S}$, and define*

$$\Psi, \Phi : \{\zeta \in D_{\mathscr{S}'} : \zeta_s \in \mathbf{M}_{\leq 1}(\mathbb{R})\} \times D_{\mathbb{R}} \to \mathbb{R} \quad \textit{by}$$

$$\Psi(\zeta, \ell) := \int_0^t \langle \zeta_s, g(s, \cdot, \zeta_s, \ell_s)\phi \rangle ds, \;\; \Phi(\zeta, \ell) := \int_0^t \langle \zeta_s, \alpha(s, \cdot, \zeta_s, \ell_s)\phi \rangle d(\mathfrak{K} \star \ell)_s,$$

where g is a placeholder for any of b, σ^2, or $\sigma\rho$. Suppose $(\nu^N, L^N) \Rightarrow (\nu^*, L^*)$ on $(D_{\mathscr{S}'}, \mathrm{M1}) \times (D_{\mathbb{R}}, \mathrm{M1})$, then $\Psi^N := \Psi(\nu^N, L^N)$ converges weakly to $\Psi^* := \Psi(\nu^*, L^*)$ on $\mathbb{R}$, and likewise for Φ.

Proof. Fix a bounded $f \in \mathrm{Lip}(\mathbb{R})$. By Remark 3.4.6, any subsequence of (ν^N, L^N) has a further subsequence, also indexed by N, for which we can assume almost sure convergence. By applying the triangle inequality, we have

$$\begin{aligned}\big|\mathbb{E}[f(\Psi^N)]-\mathbb{E}[f(\Psi^*)]\big| &\lesssim \mathbb{E}\bigg[\Big|\int_0^t \big\langle \nu_s^* - \nu_s^N, g(s,\cdot,\nu_s^*,L_s^*)\phi\big\rangle ds\Big|\bigg] \\ &\quad + \mathbb{E}\bigg[\Big|\int_0^t \big\langle \nu_s^N, g(s,\cdot,\nu_s^*,L_s^N)\phi - g(s,\cdot,\nu_s^N,L_s^N)\phi\big\rangle ds\Big|\bigg] \\ &\quad + \mathbb{E}\bigg[\Big|\int_0^t \big\langle \nu_s^N, g(s,\cdot,\nu_s^*,L_s^*)\phi - g(s,\cdot,\nu_s^*,L_s^N)\phi\big\rangle ds\Big|\bigg] \\ &=: \mathbb{E}[I_1^N] + \mathbb{E}[I_2^N] + \mathbb{E}[I_3^N].\end{aligned}$$

Starting with I_1, fix $\delta > 0$ and recall $|g| \le |x| + C(\nu^*)$ with $C(\nu^*) \lesssim 1 + \sup_{s\le t}|M_s^*|$. Hence, using that $\phi \in \mathscr{S}$, we can take $\lambda = \lambda(\delta)$ sufficiently large so that

$$\sup_{x\in\mathbb{R}\setminus[-\lambda,\lambda]} |g(s,x,\nu_s^*,L_s^*)\phi(x)| < (1 + C(\nu^*))\delta/2 \quad \forall s \in [0,T]. \tag{3.23}$$

Now take a family of mollifiers $\psi_\varepsilon \in \mathcal{C}_c^\infty(\mathbb{R})$ and consider the (random) mollifications

$$g_s^\varepsilon(x) := \big(g(s,\cdot,\nu_s^*,L_s^*) \star \psi_\varepsilon\big)(x), \quad \varepsilon > 0.$$

As ν^* and ν^N are sub-probability measures, we have

$$I_1 \le \Big|\int_0^t \big\langle \nu_s^* - \nu_s^N, g_s^\varepsilon\phi\big\rangle ds\Big| + 2\|\phi\|_\infty \int_0^t \sup_{|x|\le\lambda} |g_s^\varepsilon - g_s| ds + (1 + C(\nu^*))\delta t.$$

Since $g_s^\varepsilon \to g_s$ uniformly in $x \in [\lambda,\lambda]$, dominated convergence implies that, for ε sufficiently small, the second term is less than δ in expectation. Likewise, it follows from $g_s^\varepsilon\phi \in \mathscr{S}$ that, for N large, the first term is less than δ in expectation. Hence $\mathbb{E}[I_1^N] \le C(\delta + \delta + \delta t)$ for large enough N, where C only depends on $\mathbb{E}[C(\nu^*)]$, so we can conclude that $\mathbb{E}[I_1^N] \to 0$ as $N \to \infty$.

For I_2^N, the local Lipschitz continuity of g, as well as $\phi \in \mathscr{S}$ and $\nu_s^N \in \mathbf{M}_{\le 1}$, gives

$$I_2^N \le C(\nu^*) \int_0^t \sup\{|\langle \nu_s^* - \nu_s^N, \psi\rangle| : \psi \in \mathcal{C}_{d_0}\} ds,$$

where $\mathcal{C}_{d_0} := \{\psi \in \mathcal{C}(\mathbb{R}) : \|\psi\|_{\mathrm{Lip}} \le 1, |\psi(x)| \le 1 + |x|\}$, and $C(\nu^*) \lesssim 1 + \sup_{s\le t}|M_s^*|$. Fix $\delta > 0$ and take $\lambda = \lambda(\delta)$ large (to be determined later). By the Arzelà–Ascoli theorem, there is a finite family $\psi_1, \dots, \psi_k \in \mathcal{C}_{d_0}$ supported in $[-\lambda,\lambda]$ so that, for each $\psi \in \mathcal{C}_{d_0}$,

$$\sup\{|\psi(x) - \psi_i(x)| : x \in [-\lambda,\lambda]\} < \delta/2$$

for some $i \in \{1, \dots, k\}$. Fixing any $\psi \in \mathcal{C}_{d_0}(\mathbb{R})$ and the corresponding ψ_i,

$$|\langle \nu^* - \nu^N, \psi \rangle| \le \Big| \int_0^\lambda (\psi - \psi_i) d(\nu^* - \nu^N) \Big| + \Big| \int_\lambda^\infty (\psi - \psi_i) d(\nu^* - \nu^N) \Big| + |\langle \nu^* - \nu^N, \psi_i \rangle|.$$

By construction, the first term is bounded by δ uniformly in $N \ge 1$. Moreover, by Cauchy–Schwartz and Jensen's inequality, we have

$$\mathbb{E}\Big[C(\nu^*) \int_0^t \Big| \int_\lambda^\infty (\psi - \psi_i) d(\nu_s^* - \nu_s^N) \Big| ds \Big] \le C \int_0^t \mathbb{E}\big[\langle \nu_s^* + \nu_s^N, |\cdot|^2 \mathbf{1}_{[\lambda,\infty)} \rangle \big] ds,$$

and hence $\mathbb{E}[I_2^N]$ is bounded by a constant multiple C of

$$\delta t + \int_0^t \mathbb{E}\big[\langle \nu_s^* + \nu_s^N, |\cdot|^2 \mathbf{1}_{[\lambda,\infty)} \rangle \big] ds + \sup_{i \le k} \int_0^t \mathbb{E}\big[|\langle \nu_s^* - \nu_s^N, \psi_i \rangle| \big] ds.$$

By Corollary 3.3.4 and Proposition 3.4.3, Lemma A.1.3 gives that the middle term vanishes uniformly in $N \ge 1$ as $\lambda \to \infty$, so we can take λ large enough so that it is uniformly bounded by δ. For the final term, recall that ψ_i has compact support in $[-\lambda, \lambda]$, so we can use the same mollification argument as for I_1^N. Since there are only finitely many ψ_i's to consider, it follows that we can take N sufficiently large to ensure that $\mathbb{E}[I_2^N] \le C(\delta t + \delta + 2\delta)$, where C is a fixed numerical constant. This proves that $\mathbb{E}[I_2^N]$ vanishes as $N \to \infty$.

Finally, we consider the last integral I_3^N. By Assumption 3.2.1(iii), we have

$$\big| \big\langle \nu_s^N, g(s, \cdot, \nu_s^*, \ell_s^*)\phi - g(s, \cdot, \nu_s^*, \ell_s^N)\phi \big\rangle \big| \lesssim C(\nu^*) |\ell_s^* - \ell_s^N| \tag{3.24}$$

whenever $\ell_s^*, \ell_s^N \in [\theta_{i-1}, \theta_i)$ for some $i = 1, \dots, k$. Let $\{\ell^*, (\ell^N)_{N \ge 1}\}$ represent a fixed realization of $\{L^*, (L^N)_{N \ge 1}\}$. Then $\ell^N \to \ell^*$ in $(D_{\mathbb{R}}, \mathrm{M1})$ and hence $\ell_s^N \to \ell_s^*$ for any fixed $s \in \{r \in [0,t] : \ell_{r-}^* = \ell_r^*\}$. Fix $\varepsilon > 0$. Since ℓ^* is strictly increasing (by Assumption 3.2.3), we can take $\delta = \delta(\varepsilon)$ small enough so that

$$\mathrm{Leb}\big(\{ r \in [0,T] : |\theta_i - \ell_r^*| < \delta \text{ for some } i \} \big) \le \varepsilon.$$

On the other hand, if $|\theta_i - \ell_s^*| \ge \delta$ for all i, then we eventually have $\ell_s^*, \ell_s^N \in [\theta_{i-1}, \theta_i)$ for some i, so (3.24) applies. Thus, for the given realization of the randomness,

$$\limsup_{N \to \infty} I_3^N \le C'(\nu^*) \int_0^t \mathbf{1}_{\{r \in [0,T] \,:\, |\theta_i - \ell_r^*| < \delta \text{ for some } i\}}(s) ds \le C'(\nu^*) \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, we deduce that $\lim_N I_3^N = 0$ almost surely. Noting the uniformity of the bound (3.24), dominated convergence gives $\mathbb{E}[I_3^N] \to 0$ as $N \to \infty$.

It remains to prove $\mathbb{E}[f(\Phi^N)] \to \mathbb{E}[f(\Phi^*)]$ as $N \to \infty$. Once we observe that $|(\mathfrak{K}' \star L^*)_s| \le \|\mathfrak{K}'\|_1$ and $|\int_0^t (\mathfrak{K}' \star (L^* - L^N))_s ds| \le \|\mathfrak{K}'\|_1 \int_0^t |L_s^* - L_s^N| ds$, this follows by arguments analogous to those for Ψ. $\square$

Based on our weak convergence analysis, we are finally in a position to pass to the limit in the finite evolution equation (3.11).

Proposition 3.4.8 (Martingale argument)**.** *Fix an arbitrary $\phi \in \mathscr{C}_0$. For all $(\zeta, \ell, w) \in D_{\mathscr{S}'} \times D_{\mathbb{R}} \times C_{\mathbb{R}}$, we can define elements of $D_{\mathbb{R}}$ by*

$$\begin{aligned}
\mathcal{M}_t(\zeta,\ell) &:= \langle \zeta_t, \phi\rangle - \langle \nu_0, \phi\rangle - \textstyle\int_0^t \langle \zeta_s, b(s,\cdot,\zeta_s,\ell_s)\partial_x\phi\rangle ds \\
&\quad + \tfrac{1}{2}\textstyle\int_0^t \langle \zeta_s, \sigma(s,\cdot)^2\partial_{xx}\phi\rangle ds - \textstyle\int_0^t \langle \zeta_s, \alpha(s,\cdot,\zeta_s,\ell_s)\partial_x\phi\rangle d(\mathfrak{K}\star\ell)_s, \\
\mathcal{N}_t(\zeta,\ell) &:= \mathcal{M}_t(\zeta,\ell)^2 - \textstyle\int_0^t \langle \zeta_s, \sigma(s,\cdot)\rho(s,\zeta_s,\ell_s)\partial_x\phi\rangle^2 ds, \\
\mathcal{K}_t(\zeta,\ell,w) &:= \mathcal{M}_t(\zeta,\ell)\cdot w_t - \textstyle\int_0^t \langle \zeta_s, \sigma(s,\cdot)\rho(s,\zeta_s,\ell_s)\partial_x\phi\rangle ds.
\end{aligned}$$

If $(\nu^N, L^N, W^0) \Rightarrow (\nu^*, L^*, W^0)$, then $\mathcal{M}(\nu^*, L^*)$, $\mathcal{N}(\nu^*, L^*)$, and $\mathcal{K}(\nu^*, L^*, W^0)$ are all continuous martingales.

Proof. Let $\mathcal{M}_t^N := \mathcal{M}_t(\nu^N, L^N)$ and $\mathcal{M}_t^* := \mathcal{M}_t(\nu^*, L^*)$. Fix $s, t \in [0,T]$ with $s < t$ and define, for any $s_1, \dots, s_n \in [0,s]$,

$$F_{\mathcal{M}}^N := \big(\mathcal{M}_t^N - \mathcal{M}_s^N\big)\prod_{i=1}^n f_i(\mathcal{M}_{s_i}^N) \quad \text{and} \quad F_{\mathcal{M}}^* := \big(\mathcal{M}_t^* - \mathcal{M}_s^*\big)\prod_{i=1}^n f_i(\mathcal{M}_{s_i}^*),$$

where $f_1, \dots, f_n \in C_b(\mathbb{R})$ are arbitrary. Proceeding analogously for $\mathcal{N}$ and $\mathcal{K}$, it follows from Proposition 3.4.7 and the continuous mapping theorem that

$$F_{\mathcal{M}}^N \Rightarrow F_{\mathcal{M}}^*, \quad F_{\mathcal{N}}^N \Rightarrow F_{\mathcal{N}}^*, \quad \text{and} \quad F_{\mathcal{K}}^N \Rightarrow F_{\mathcal{K}}^*.$$

Using this, and appealing to the finite dimensional evolution equation from Proposition 3.3.2, the goal is now to show that, for $\mathcal{H} = \mathcal{M}, \mathcal{N}, \mathcal{K}$, we have

$$\mathbb{E}\big[F_{\mathcal{H}}^*\big] = \lim_{N\to\infty} \mathbb{E}\big[F_{\mathcal{H}}^N\big] = 0, \tag{3.25}$$

thus proving that $\mathcal{M}^*$, $\mathcal{N}^*$, and $\mathcal{K}^*$ are true martingales, by a standard monotone class argument. Relying on uniform integrability to deduce the convergence of the means in (3.25), the proof of this convergence follows by a straightforward modification to the arguments in [73, Prop. 5.11]. □

By [86, Thm. 3.2] and the tightness from Propositions 3.4.2 and 3.4.5, we can extract a weakly convergent subsequence $(\nu^N, L^N, W^0) \Rightarrow (\nu^*, L^*, W^0)$. Consequently, Proposition 3.4.8 and the Doob–Meyer decomposition theorem allows us to conclude that, for each test function $\phi \in \mathscr{C}_0$, we have

$$\begin{aligned}
\langle \mathcal{M}(\nu^*, L^*)\rangle_t &= \textstyle\int_0^t \langle \nu_s^*, \sigma(s,\cdot)\rho(s,\nu_s^*, L_s^*)\partial_x\phi\rangle^2 ds \ \text{ and} \\
\big\langle \mathcal{M}(\nu^*, L^*), W^0\big\rangle_t &= \textstyle\int_0^t \langle \nu_s^*, \sigma(s,\cdot)\rho(s,\nu_s^*, L_s^*)\partial_x\phi\rangle\, ds,
\end{aligned}$$

so it holds for all $t \in [0,T]$ that

$$\left\langle \mathcal{M}(\nu^*, L^*) - \int_0^\cdot \langle \nu_s^*, \sigma(s,\cdot)\rho(s,\nu_s^*,L_s^*)\partial_x\phi\rangle \, dW_s^0 \right\rangle_t = 0.$$

This is precisely saying that (ν^*, W^0) satisfies the SPDE (3.7) and so the proof of Theorem 3.2.4 is complete.

3.5 Uniqueness arguments

The purpose of this section is to present a detailed proof of Theorem 3.2.6. In view of Section 3.4.2, we fix a limit point ν of the sequence of empirical measures (ν^N) and let $\tilde{\nu}$ denote another candidate solution to the SPDE (3.7). Then the strategy is to establish an energy estimate in H^{-1} for the difference $\Delta_t := \nu_t - \tilde{\nu}_t$, where H^{-1} is the usual dual space of $H_0^1 = \mathcal{W}_0^{1,2}(\mathbb{R}_+)$. More precisely, we will rely on a 'smoothed' H^{-1} estimate given in Proposition 3.5.1. Based on this estimate, we show how to derive the uniqueness of the SPDE in Section 3.5.1.1, and then Sections 3.5.2 and 3.5.3 are devoted to the proof of Proposition 3.5.1.

3.5.1 Energy estimates and smoothing

Rather than estimating the H^{-1} norm of Δ_t directly, our approach relies on smoothing the solutions ν and $\tilde{\nu}$ by means of convolution (with a family of kernels approximating the identity). In this way, we can manipulate the resulting equations classically. Since our problem is phrased on the positive half-line with absorption at the boundary, it is natural to consider the family of Dirichlet heat kernels G_ε given by

$$G_\varepsilon(x,y) := p_\varepsilon(x-y) - p_\varepsilon(x+y), \quad p_\varepsilon(x) = (2\pi\varepsilon)^{-\frac{1}{2}} \exp(-x^2/2\varepsilon). \tag{3.26}$$

We denote the action of G_ε on a measure μ by $\mathcal{T}_\varepsilon\mu$, that is

$$(\mathcal{T}_\varepsilon\mu)(x) := \int_0^\infty G_\varepsilon(x,y)d\mu(y).$$

For simplicity of presentation, we introduce the notation ∂_x^{-1} for the anti-derivative, which amounts to

$$(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)(x) := -\int_x^\infty (\mathcal{T}_\varepsilon\Delta_t)(y)dy. \tag{3.27}$$

Recall the embedding $\mathbf{M}_\pm \hookrightarrow H^{-1}$, where $\mathbf{M}_\pm$ is the space of finite signed measures on $\mathbb{R}$ (with total variation norm). As in [73, Prop. 6.5], we then have

$$\|\Delta_t\|_{H^{-1}} \leq \liminf_{\varepsilon\to 0} \|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t\|_2, \tag{3.28}$$

where $\|\cdot\|_2$ denotes the L^2-norm on $\mathbb{R}_+$. Therefore, we can estimate the H^{-1} norm of the difference Δ_t via the anti-derivatives of the smoothed solutions.

Let us now briefly outline the key ideas behind our approach. The first observation is that $y \mapsto G_\varepsilon(x,y)$ is an admissible test function in $\mathscr{C}_0$, so we can plug it into the SPDE and thus obtain expressions for the smoothed solutions, $\mathcal{T}_\varepsilon\nu_t$ and $\mathcal{T}_\varepsilon\tilde{\nu}_t$. Integrating these to introduce the anti-derivatives, and looking at their difference, this then allows us to obtain an equation for $\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t$ in terms of $\mathcal{T}_\varepsilon\Delta_t$ along with the occurrence of $\mathcal{T}_\varepsilon\nu_t$ and some critical 'boundary effects' as well as a collection of simpler error terms.

In order to control the L^2-norm of $\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t$, we thus need a uniform estimate on $\mathcal{T}_\varepsilon\nu_t$, and we need to contain the boundary effects (and error terms) as $\varepsilon \to 0$. These two tasks are the subject of Section 3.5.2, which then allows us to derive the aforementioned L^2 estimate for $\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t$ in Section 3.5.3. However, before any of this, we begin by showing how to derive the uniqueness of the SPDE given that the desired energy estimate holds.

3.5.1.1 Uniqueness of the SPDE: Proof of Theorem 3.2.6

As above, we let ν be a limit point of the particle system and suppose $\tilde{\nu}$ is another solution to the SPDE (3.7), satisfying Assumption 3.2.3.

Recall that the local Lipschitz continuity in the loss variable only holds in a piecewise fashion on the intervals $[\theta_{i-1},\theta_i)$ for $i = 1,\dots,k$. This hurdle is easily overcome by the following piecewise stopping argument: Suppose we can prove uniqueness on $[0,t_0]$ if $L_t, \tilde{L}_t \in [0,\theta_1)$ for $t < t_0$ (see Prop. 3.5.1 and the ensuing arguments) and introduce the stopping times

$$\tau_1 := \inf\{t > 0 : L_t \geq \theta_1\} \wedge T \quad \text{and} \quad \tilde{\tau}_1 := \inf\{t > 0 : \tilde{L}_t \geq \theta_1\} \wedge T.$$

Then $L_t, \tilde{L}_t \in [0,\theta_1)$ for $t < \varsigma_1 := \tau_1 \wedge \tilde{\tau}_1$, so we get uniqueness up to ς_1. Note that, by the uniqueness, $L = \tilde{L}$ on $[0,\varsigma_1]$ and hence $\varsigma_1 = \tau_1 = \tilde{\tau}_1$. Thus, we can repeat the uniqueness arguments on $[\varsigma_1,\varsigma_2)$, where $L_t, \tilde{L}_t \in [\theta_1,\theta_2)$, by defining $\varsigma_2 := \tau_2 \wedge \tilde{\tau}_2$ for

$$\tau_2 := \inf\{t > \varsigma_1 : L_t \geq \theta_2\} \wedge T \quad \text{and} \quad \tilde{\tau}_2 := \inf\{t > \varsigma_1 : \tilde{L}_t \geq \theta_2\} \wedge T.$$

Continuing in this way for $\varsigma_3,\dots,\varsigma_{k-1}$, we get uniqueness on all of $[0,T]$, since L_t and $\tilde{L}_t$ are strictly increasing (recall part (ii) of Assumption 3.2.3).

Below we prove uniqueness when the local Lipschitz continuity in L holds everywhere, noting that the arguments imply uniqueness if L and $\tilde{L}$ are confined to a particular piece $[\theta_{i-1},\theta_i)$. Thus, in view of the above stopping argument, the next result will

suffice to complete the proof of Theorem 3.2.6. This result is what we referred to as the 'smoothed' H^{-1} estimate at the beginning of Section 5.

Proposition 3.5.1 (Smoothed H^{-1} Estimate)**.** *Suppose the local Lipschitz continuity in Assumption 3.2.1(iii) holds everywhere, as opposed to piecewise. Then there is a sequence of stopping times (t_n), with $t_n \uparrow T$ as $n \uparrow \infty$, such that, for all $t \in [0,T]$,*

$$\begin{aligned}\mathbb{E}\big[\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_{t\wedge t_n}\|_2^2\big] + c_0\mathbb{E}\Big[\int_0^{t\wedge t_n}\|\mathcal{T}_\varepsilon\Delta_s\|_2^2 ds\Big] \le c_n\mathbb{E}\Big[\int_0^{t\wedge t_n}|L_s-\tilde{L}_s|^2 ds\Big]\\ + c_n\mathbb{E}\Big[\int_0^{t\wedge t_n} d_0(\nu_s,\tilde{\nu}_s)\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\|_2 ds\Big] + c_n\mathbb{E}\Big[\int_0^{t\wedge t_m} d_1(\nu_s,\tilde{\nu}_s)^2 ds\Big] + o(1),\end{aligned}$$

as $\varepsilon \downarrow 0$, for a fixed constant c_0 and constants c_n only depending on n.

Proof. The proof is postponed to Sections 3.5.2 and 3.5.3. □

Note that, at this point, Grönwall already gives that $\mathbb{E}[\int_0^{t\wedge t_n}\|\frac{d\Delta_s}{dx}\|_2 ds]$ is finite and, in particular, $\Delta_{s\wedge t_n}$ has a density in L^2, which we make use of below. Observe also that $|L-\tilde{L}|$ is dominated by $d_1(\nu,\tilde{\nu})$, however, it is included in the estimate as it shows where the above 'piecewise stopping argument' comes into play. The next lemma relates the left- and right-hand sides of the smoothed H^{-1} estimate, thus opening the door for a stronger Grönwall argument that will allow us to finish the proof of Theorem 3.2.6.

Lemma 3.5.2. *There exists $c>0$ so that, for all $s \le T$, $\delta \in (0,1)$ and $\lambda > 1$,*

$$\begin{cases} d_0(\nu_s,\tilde{\nu}_s) \le c\lambda(1+\delta^{-1})\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2 + c\delta^{\frac{1}{2}}\big\|\frac{d\Delta_s}{dx}\big\|_2 + x_s^\lambda + z_s^\varepsilon \\ d_1(\nu_s,\tilde{\nu}_s) \le c\sqrt{\lambda+\delta^{-1}}\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2 + c\delta^{\frac{1}{2}}\big\|\frac{d\Delta_s}{dx}\big\|_2 + x_s^\lambda + z_s^\varepsilon, \end{cases}$$

as $\varepsilon \downarrow 0$, where $(x_t^\lambda)_{t\le T}$ is a process such that $\mathbb{E}[\int_0^T |x_s^\lambda|^2 ds] \le c_a e^{-a\lambda}$, for all $a>0$, and $(z_s^\varepsilon)_{s\le T}$ is a process such that $\mathbb{E}[\int_0^T |z_s^\varepsilon|^2 ds] = o(1)$ as $\varepsilon \downarrow 0$.

Inserting the bound for d_0 in Proposition 3.5.1, the leading term that we get is $\lambda(1+\delta^{-1})\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\|_2^2$, for which it is crucial that λ is not squared. For the rest of the terms coming from d_0, however, we can safely apply Young's inequality to bound them. After also inserting the bound for d_1, it follows that

$$\begin{aligned}\mathbb{E}\big[\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_{t\wedge t_n}\|_2^2\big] \le\ \delta c_n'\mathbb{E}\Big[\int_0^{t\wedge t_n}\Big\|\frac{d\Delta_s}{dx}\Big\|_2^2 ds\Big] - c_0\mathbb{E}\Big[\int_0^{t\wedge t_n}\|\mathcal{T}_\varepsilon\Delta_s\|_2^2 ds\Big]\\ + c_n'(1+\lambda+\lambda\delta^{-1})\int_0^t \mathbb{E}\big[\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_{s\wedge t_n}\|_2^2 ds\big] + c_a' e^{-a\lambda} + o(1), \qquad (3.29)\end{aligned}$$

as $\varepsilon \downarrow 0$. By construction of $\mathcal{T}_\varepsilon$, $\|\mathcal{T}_\varepsilon \Delta_s\|_2 \leq \|\frac{d\Delta_s}{dx}\|_2$ and $\|\mathcal{T}_\varepsilon \Delta_s\|_2 \to \|\frac{d\Delta_s}{dx}\|_2$, so dominated convergence gives

$$\mathbb{E}\Big[\int_0^{t\wedge t_n} \big\|\mathcal{T}_\varepsilon \Delta_s\big\|_2 ds\Big] \to \mathbb{E}\Big[\int_0^{t\wedge t_n} \Big\|\frac{d\Delta_s}{dx}\Big\|_2 ds\Big] \quad \text{as} \quad \varepsilon \downarrow 0.$$

Therefore, by taking $\delta := c_0/2c_n'$, the sum of the first two terms on the right-hand side of (3.29) is eventually non-positive for small $\varepsilon > 0$. In turn, we can apply the integrating factor $\exp\{c_n'(1+\lambda+\lambda\delta^{-1})t\}$ to the anti-derivative term in (3.29) and deduce that

$$\mathbb{E}\big[\|\partial_x^{-1}\mathcal{T}_\varepsilon \Delta_{t\wedge t_n}\|_2^2\big] \leq c_a' c_n' (1+\lambda+\lambda\delta^{-1}) e^{c_n'(1+\lambda+\lambda\delta^{-1})T} e^{-a\lambda} + o(1)$$

as $\varepsilon \downarrow 0$. Recalling (3.28) and appealing to Fatou's lemma, it follows that

$$\mathbb{E}\big[\|\Delta_{t\wedge t_n}\|_{H^{-1}}^2\big] \leq c_{a,n,T}'(1+\lambda+\lambda\delta^{-1}) \exp\{\lambda(1+\delta^{-1})T - a\lambda\}.$$

Consequently, we can simply take $a := 2(1+\delta^{-1})T$ and send $\lambda \to \infty$ to arrive at

$$\mathbb{E}\big[\|\Delta_{t\wedge t_n}\|_{H^{-1}}^2\big] = 0 \quad \forall t \in [0,T].$$

Since n was arbitrary and $t_n \uparrow T$ as $n \uparrow \infty$, we conclude that $\nu_t = \tilde{\nu}_t$ for all $t \in [0,T]$. This completes the proof of Theorem 3.2.6.

Proof of Lemma 3.5.2. Fix $\psi \in \mathcal{C}_{d_0}$, where $\mathcal{C}_{d_0} := \{f : \|f\|_{\mathrm{Lip}} \leq 1, |f(x)| \leq 1 + |x|\}$. Fixing also $\lambda > 1$ and $\delta > 0$, we can take a cut-off function $\chi \in \mathcal{C}_c^\infty(\mathbb{R})$ equal to 1 on $[\delta, \lambda - 1]$ and supported in $(\delta/2, \lambda)$ such that

$$|\chi| \leq 1, \quad |\partial\chi| \leq C_\chi/\delta \ \text{ on } \ [\delta/2, \delta], \quad \text{and} \quad |\partial\chi| \leq C_\chi \ \text{ on } \ [\lambda-1, \lambda].$$

Observe that $\chi\psi \in \mathcal{W}_0^{1,\infty}(\delta/2, \lambda)$ with

$$\begin{aligned}\|\partial(\chi\psi)\|_2^2 &= \int_0^\delta |\partial\chi|^2 |\psi|^2 \, dx + \int_{\lambda-1}^\lambda |\partial\chi|^2 |\psi|^2 \, dx + \int_0^\lambda |\chi|^2 |\partial\psi|^2 \, dx \\ &\leq C_\chi^2(\delta^{-1}+1)\|\psi\|_{L^\infty(0,\lambda)}^2 + \lambda.\end{aligned}$$

Next, we can observe that

$$\begin{aligned}|\langle \Delta_s, \psi\rangle| &\leq |\langle \Delta_s, \chi\psi\rangle| + |\langle \Delta_s, (\chi-1)\psi \mathbf{1}_{[0,\delta]}\rangle| + \langle \nu_s + \tilde{\nu}_s, |\psi| \mathbf{1}_{[\lambda-1,\infty)}\rangle \\ &\leq |\langle \Delta_s, \chi\psi\rangle| + (1+\delta)\delta^{1/2}\Big\|\tfrac{d\Delta_s}{dx}\Big\|_2 + \langle \nu_s + \tilde{\nu}_s, (1+|\cdot|)\mathbf{1}_{[\lambda-1,\infty)}\rangle\end{aligned}$$

Now let $(\cdot,\cdot)$ be the inner product on L^2. Then integration by parts and Cauchy–Schwarz gives

$$\begin{aligned}|\langle \Delta_s, \chi\psi\rangle| &\le \big|(\mathcal{T}_\varepsilon \Delta_s, \chi\psi)\big| + \big|(\mathcal{T}_\varepsilon\Delta_s,\chi\psi) - \langle \Delta_s,\chi\psi\rangle\big| \\ &\le \big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2 \big\|\partial_x(\chi\psi)\big\|_2 + \big|(\mathcal{T}_\varepsilon\Delta_s,\chi\psi) - \langle\Delta_s,\chi\psi\rangle\big|.\end{aligned}$$

By the Arzelà–Ascoli theorem applied to $\{\chi f : f \in \mathcal{C}_{d_0}\}$, we can find a finite family $\{\varphi_i \in \mathrm{Lip}(\mathbb{R}) : i = 1,\dots,k(\delta,\lambda)\}$ with support in $[\delta/2,\lambda]$ so that, for any $f \in \mathcal{C}_{d_0}$,

$$\sup_{x\in\mathbb{R}} |\varphi_i(x) - \chi f(x)| \le \delta^{\frac{1}{2}}\lambda^{-\frac{1}{2}}, \quad \text{for some } \ i = 1,...,k(\delta,\lambda).$$

Consequently, there is a φ_i such that

$$\big|(\mathcal{T}_\varepsilon\Delta_s,\chi\psi) - \langle\Delta_s,\chi\psi\rangle\big| \le \big|(\mathcal{T}_\varepsilon\Delta_s,\varphi_i) - \langle\Delta_s,\varphi_i\rangle\big| + 2\sqrt{\delta}\big\|\tfrac{d\Delta_s}{dx}\big\|_2,$$

where we have used Cauchy–Schwarz with $\|\mathcal{T}_\varepsilon\Delta_s\|_2 \le \|\frac{d\Delta_s}{dx}\|_2$ and $\|\varphi_i - \chi\psi\|_2 \le \sqrt{\delta}$. Note that $(\mathcal{T}_\varepsilon\Delta_s,\varphi_i) \to \langle\Delta_s,\varphi_i\rangle$ as $\varepsilon\to 0$ (see e.g. [73, Prop. 6.4]), so defining

$$z_s^\varepsilon = z_s^\varepsilon(\delta,\lambda) := \sup\{|(\mathcal{T}_\varepsilon\Delta_s,\varphi_i) - \langle\Delta_s,\varphi_i\rangle| : i = 1,\dots,k(\delta,\lambda)\},$$

we have $\mathbb{E}[\int_0^T |z_s^\varepsilon|^2]ds \to 0$ as $\varepsilon\to 0$ by bounded convergence. Consequently,

$$|\langle\Delta_s,\chi\psi\rangle| \le \big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2\big\|\partial_x(\chi\psi)\big\|_2 + 2\sqrt{\delta}\big\|\tfrac{d\Delta_s}{dx}\big\|_2 + z_s^\varepsilon,$$

where z^ε is as required by the lemma. Finally, combining the above, we can take suprema over ψ in the function classes that define d_0 and d_1 to find that

$$d_0(\nu_s,\tilde\nu_s) \le c\lambda\sqrt{1+\delta^{-1}}\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2 + c\delta^{\frac{1}{2}}\big\|\tfrac{d\Delta_s}{dx}\big\|_2 + x_s^\lambda + z_s^\varepsilon,$$

and

$$d_1(\nu_s,\tilde\nu_s) \le c\sqrt{\lambda+\delta^{-1}}\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2 + c\delta^{\frac{1}{2}}\big\|\tfrac{d\Delta_s}{dx}\big\|_2 + x_s^\lambda + z_s^\varepsilon,$$

where $x_s^\lambda := \langle \nu_s + \tilde\nu_s, (1+|\cdot|)\mathbf{1}_{[\lambda-1,\infty)}\rangle$. By the exponential decay properties of x_s^λ in λ, as guaranteed by Lemma A.1.3, this completes the proof. □

3.5.2 L^2 regularity and containment of boundary effects

Our first task in this section is to establish a weighted L^2 estimate for $\mathcal{T}_\varepsilon\nu_t$ that is uniform in $\varepsilon > 0$ and $t \in [0,T]$. This can be achieved by exploiting the extra control that we have over ν because it arises as a limit point of the particle system.

Specifically, we can use that the empirical measures (ν^N) are dominated by their whole-space counterparts so that, in the limit, ν is dominated by a solution to the

whole-space version of the SPDE (3.7) with M and L still defined in terms of ν. This is crucial as we can perform the L^2 estimates for the whole-space SPDE without any boundary effects. On the other hand, the estimates for the half-line only succeed because we work in the weaker space H^{-1}, where we can control the boundary effects solely by relying on the boundary decay from Proposition 3.4.3 and Assumption 3.2.3 (see Lemma 5.5 and Section 5.3.2).

Proposition 3.5.3. *Let ν^* be a limit point of (ν^N). Then*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\sup_{\varepsilon>0}\int_0^\infty (1+x^2)(\mathcal{T}_\varepsilon\nu_t^*)^2dx\right]<\infty. \tag{3.30}$$

Proof. The rough ideas are the same as [73, Prop. 7.1], so we will refer back to this in order to avoid duplication. First, let $\bar{\nu}^N := \sum_{i=1}^N a_i^N \delta_{X_t^i}$ and note $\nu^N(A) \le \bar{\nu}^N(A)$ for all $A \in \mathcal{B}(\mathbb{R})$. Now, the same work as for (ν^N) ensures $(\nu^N, \bar{\nu}^N)$ is tight with limit points $\nu^* \le \bar{\nu}^*$, where $\bar{\nu}^*$ satisfies the whole-space analogue of the SPDE (3.7) in the sense that the space of test functions is all of $\mathscr{S}$, but with M and L still defined in terms of ν^*. Arguing as in Lemmas 7.2 and 7.3 of [73], to prove (3.30) it suffices to bound

$$J_1 := \liminf_{\varepsilon\to 0}\mathbb{E}\left[\sup_{t\le T}\int_{\mathbb{R}}(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*)^2dx\right] \text{ and } J_2 := \liminf_{\varepsilon\to 0}\mathbb{E}\left[\sup_{t\le T}\int_{\mathbb{R}}x^2(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*)^2dx\right],$$

where the action of $\bar{\mathcal{T}}_\varepsilon$ is given by

$$(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*)(x) := \int_{\mathbb{R}} p_\varepsilon(x-y)d\bar{\nu}_t^*(y).$$

Note that $y \mapsto p_\varepsilon(x-y)$ is certainly not in $\mathscr{C}_0$, but it has rapid decay at $\pm\infty$, so it is an admissible test function for the whole-space SPDE satisfied by $\bar{\nu}^*$. We split the remaining parts of the proof into three steps.

Step 1. We begin by showing that $J_1 \lesssim \int_0^\infty V_0(x)^2dx$. To simplify things, let

$$\mathfrak{b}_t(x) := b(t,x,\nu_t,L_t) - \alpha(t,x,\nu_t,L_t)\mathfrak{L}_t'.$$

Then $\mathfrak{b}_t$ will play the rôle of the drift μ_t in the proof of Proposition 7.1 from [73]. Proceeding as in [73], we can test the SPDE with $p_\varepsilon(\cdot - y)$ and introduce appropriate error terms in order to get a tractable expression for $\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*$. Next, we can then apply Itô's formula and thus convert this into an expression for the square $(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*)^2$. In this

way, we arrive at

$$\begin{aligned} d(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*)^2 =& -2(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*)\big(\mathfrak{b}_t\partial_x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*) - \partial_x\mathfrak{b}_t\bar{\mathcal{H}}^{\mathfrak{b}}_{t,\varepsilon} + \bar{\mathcal{E}}^{\mathfrak{b}}_{t,\varepsilon}\big)dt \\ &+ (\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*)\partial_x\big(\sigma_t^2\partial_x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*) - \partial_x\sigma_t^2\bar{\mathcal{H}}^{\sigma^2}_{t,\varepsilon} + \bar{\mathcal{E}}^{\sigma^2}_{t,\varepsilon}\big)dt \\ &- 2\rho_t(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*)\big(\sigma_t\partial_x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*) - \partial_x\sigma_t\bar{\mathcal{H}}^{\sigma}_{t,\varepsilon} + \bar{\mathcal{E}}^{\sigma}_{t,\varepsilon}\big)dW_t \\ &+ \rho_t^2\big(\sigma_t\partial_x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*) - \partial_x\sigma_t\bar{\mathcal{H}}^{\sigma}_{t,\varepsilon} + \bar{\mathcal{E}}^{\sigma}_{t,\varepsilon}\big)^2dt, \end{aligned} \tag{3.31}$$

where the error terms $\bar{\mathcal{E}}$ and $\bar{\mathcal{H}}$ are defined in Lemma A.1.2 of the Appendix. Compared to [73], we must be careful with the linear growth of $\mathfrak{b}_t$, however, this is taken care of via integration by parts: Since $\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*$ vanishes at $\pm\infty$ (by the whole-space analogue of Lemma A.1.4 for the tails of $\bar{\nu}_t^*$), we thus get

$$\begin{aligned} -\int_{\mathbb{R}}\int_0^t 2\mathfrak{b}_s(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)\partial_x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)dsdx &= -\int_0^t\int_{\mathbb{R}}\mathfrak{b}_s\partial_x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)^2dxds \\ &= \int_0^t\int_{\mathbb{R}}\partial_x\mathfrak{b}_s(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)^2dxds \le C\int_0^t\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*\big\|_2^2ds. \end{aligned} \tag{3.32}$$

Given this, the plan is to integrate over $x\in\mathbb{R}$ in the equation (3.31) in order to obtain an estimate for the L^2-norm of $\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*$. After integrating over x, we can then appeal to the previous estimate (3.32) and, similarly, we can perform another integration by parts in the second line on the right-hand side of (3.31). Using this, and the fact that $|\bar{\mathcal{H}}^g_{t,\varepsilon}|\lesssim\bar{\mathcal{T}}_{2\varepsilon}\bar{\nu}_t^*$ by Lemma A.1.2, we can thus apply Young's inequality (with free parameter $\theta>0$) to see that

$$\begin{aligned} \big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*\big\|_2^2 \le& \big\|\bar{\mathcal{T}}_\varepsilon\nu_0\big\|_2^2 + C_\theta\int_0^t\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*\big\|_2^2ds + C_\theta\int_0^t\big\|\bar{\mathcal{T}}_{2\varepsilon}\bar{\nu}_s^*\big\|_2^2ds \\ &+ C_\theta\int_0^t\big(\big\|\bar{\mathcal{E}}^{\mathfrak{b}}_{s,\varepsilon}\big\|_2^2 + \big\|\bar{\mathcal{E}}^{\sigma^2}_{s,\varepsilon}\big\|_2^2 + \big\|\bar{\mathcal{E}}^{\sigma}_{s,\varepsilon}\big\|_2^2\big)ds \\ &+ \int_0^t\int_{\mathbb{R}}\big(\rho_s^2\sigma_s^2 + \theta\rho_s^2\sigma_s^2 - \sigma_s^2 + \theta\big)\big|\partial_x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)\big|^2dxds \\ &+ \int_0^t\int_{\mathbb{R}}\rho_s\partial_x\sigma_s(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)^2 + 2\rho_s(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)(\partial_x\sigma_s^2\bar{\mathcal{H}}^{\sigma}_{s,\varepsilon} - \bar{\mathcal{E}}^{\sigma}_{s,\varepsilon})dxdW_s. \end{aligned} \tag{3.33}$$

Here we have used the stochastic Fubini theorem to switch the order of integration in the stochastic integral (allowed due to the exponential tails as in [73, Lem. 8.3]) and we have also integrated by parts in the dx-integral inside the stochastic integral.

Step 2. Since ρ^2 is bounded away from 1 (and σ^2 is bounded away from 0), we can choose θ sufficiently small so that the third line of (3.33) is negative and hence it

can be discarded. Raising both sides of (3.33) to a power $k \geq 1$, we thus have

$$\begin{aligned}\sup_{r\leq t}\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_r^*\big\|_2^{2k} \leq\ & C\big\|\bar{\mathcal{T}}_\varepsilon\nu_0\big\|_2^{2k} + C\int_0^t \big(\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*\big\|_2^{2k} + \big\|\bar{\mathcal{T}}_{2\varepsilon}\bar{\nu}_s^*\big\|_2^{2k}\big)ds \\ &+ C\int_0^t \big(\big\|\bar{\mathcal{E}}_{s,\varepsilon}^{\mathfrak{b}}\big\|_2^{2k} + \big\|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma^2}\big\|_2^{2k} + \big\|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}\big\|_2^{2k}\big)ds \\ &+ C\Big(\sup_{r\leq t}\int_0^r\int_{\mathbb{R}} \rho_s\partial_x\sigma_s(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)^2 + 2\rho_s(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)(\partial_x\sigma_s^2\bar{\mathcal{H}}_{s,\varepsilon}^{\sigma} - \bar{\mathcal{E}}_{s,\varepsilon}^{\sigma})dxdW_s\Big)^k, \qquad (3.34)\end{aligned}$$

with $C = C(k)$. By the Burkholder–Davis–Gundy, Hölder and Young's inequalities, the expected supremum of the stochastic integral is bounded by a constant times

$$\begin{aligned}&\mathbb{E}\Big[\Big(\int_0^t\Big(\int_{\mathbb{R}}(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)\big(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^* + |\partial_x\sigma_s^2\bar{\mathcal{H}}_{s,\varepsilon}^{\sigma}| + |\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}|\big)dx\Big)^2 ds\Big)^{k/2}\Big] \\ &\qquad \leq C'\mathbb{E}\Big[\sup_{r\leq t}\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_r^*\big\|_2^{2k}\Big] + C''\mathbb{E}\Big[\int_0^t \big(\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*\big\|_2^{2k} + \big\|\bar{\mathcal{T}}_{2\varepsilon}\bar{\nu}_t^*\big\|_2^{2k} + \big\|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}\big\|_2^{2k}\big)ds\Big],\end{aligned}$$

Consequently, taking expectations in (3.34) gives

$$\begin{aligned}\inf_{\varepsilon\leq\varepsilon'}\mathbb{E}\Big[\sup_{r\leq t}\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_r^*\big\|_2^{2k}\Big] \leq\ & C_0\inf_{\varepsilon\leq\varepsilon'}\big\|\bar{\mathcal{T}}_\varepsilon\nu_0\big\|_2^{2k} + tC_0\inf_{\varepsilon\leq\varepsilon'}\mathbb{E}\Big[\sup_{r\leq t}\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_r^*\big\|_2^{2k}\Big] \\ &+ C_0\inf_{\varepsilon\leq\varepsilon'}\mathbb{E}\Big[\int_0^t \big(\big\|\bar{\mathcal{E}}_{s,\varepsilon}^{\mathfrak{b}}\big\|_2^{2k} + \big\|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma^2}\big\|_2^{2k} + \big\|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}\big\|_2^{2k}\big)ds\Big].\end{aligned}$$

If we now restrict to $t \leq T_0 := 1/2C_0$ and send $\varepsilon' \to 0$, then we get

$$\liminf_{\varepsilon\to 0}\mathbb{E}\Big[\sup_{t\leq T_0}\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*\big\|_2^{2k}\Big] \leq 2C_0\big\|V_0\big\|_2^{2k},$$

where we have used that $\|\bar{\mathcal{T}}_\varepsilon\nu_0\|_2 \to \|V_0\|_2$ as $\varepsilon \to 0$ and that the error terms vanish by Lemma A.1.2. For $k = 1$, this proves the bound on J_1 in small time, that is, for $t \in [0, T_0]$ with $T_0 = 1/2C_0$. The extension of this bound to all of $[0, T]$ follows by propagating the argument onto the finitely many intervals $[T_k, T_{k+1} \wedge T]$ with $T_{k+1} = T_k + 1/2C_0$ for $k = 1, \ldots, \lceil 2C_0T\rceil$ (as in the proof of [73, Prop. 7.1]).

Step 3. We now show how to extend the previous work to prove that J_2 is finite. In order to succeed at this, we will need to control the 4th moments of $\sup_{t\leq T}\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_t^*\|_2$, which is the reason for introducing the power k in Step 2 above. The idea is simply to multiply by x^2 in (3.31) before integrating over $x \in \mathbb{R}$ and then proceed as in Steps 1 and 2 above. Beginning with the first term, an integration by part yields

$$\begin{aligned}-\int_{\mathbb{R}}\int_0^t 2x^2\mathfrak{b}_s(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)\partial_x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)dsdx =\ & \int_0^t\int_{\mathbb{R}}\big(x^2\partial_x\mathfrak{b}_s + 2x\mathfrak{b}_s\big)(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)^2dxds \\ \leq\ & C\int_0^t\big\|x(\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*)\big\|_2^2ds + C\int_0^t(1+|M_s|)\big\|\bar{\mathcal{T}}_\varepsilon\bar{\nu}_s^*\big\|_2^2ds.\end{aligned}$$

Crucially, the second term on the right-hand side can be controlled by

$$\mathbb{E}\Big[\int_0^t (1+|M_s|)\big\|\bar{\mathcal{T}}_\varepsilon \bar{\nu}_s^*\big\|_2^2 ds\Big] \le \int_0^t \mathbb{E}[(1+|M_s|)^2]ds + \int_0^t \mathbb{E}\big[\big\|\bar{\mathcal{T}}_\varepsilon \bar{\nu}_s^*\big\|_2^4\big] ds \tag{3.35}$$

which is finite because of the result in Step 2 with $k = 2$. Arguments completely analogous to those that led to (3.33)–(3.34) then yield

$$\begin{aligned}\big\|x\bar{\mathcal{T}}_\varepsilon \bar{\nu}_t^*\big\|_2^2 \le \big\|x\bar{\mathcal{T}}_\varepsilon \nu_0\big\|_2^2 &+ C_\theta \int_0^t \big\|x\bar{\mathcal{T}}_\varepsilon \bar{\nu}_s^*\big\|_2^2 ds + C_\theta \int_0^t \big\|x\bar{\mathcal{T}}_{2\varepsilon} \bar{\nu}_s^*\big\|_2^2 ds + C' \\ &+ C_\theta \int_0^t \big(\big\|x\bar{\mathcal{E}}_{s,\varepsilon}^{\mathfrak{b}}\big\|_2^2 + \big\|x\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma^2}\big\|_2^2 + \big\|x\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}\big\|_2^2\big) ds \\ &- 2\int_0^t \int_{\mathbb{R}} x^2 \rho_s \bar{\mathcal{T}}_\varepsilon \bar{\nu}_s^* (\sigma_s \partial_x \bar{\mathcal{T}}_\varepsilon \bar{\nu}_t^* + \partial_x \sigma_s \bar{\mathcal{H}}_{s,\varepsilon}^{\sigma} + \bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}) dx dW_s,\end{aligned} \tag{3.36}$$

where the extra constant C' comes from (3.35). Given this, we can argue as in Step 2 (with $k = 1$) to bound J_2 by a multiple of $C' + \|xV_0\|_2^2$. □

Corollary 3.5.4. *Any limit point ν^* of (ν^N) has an L^2-valued density process $(V_t^*)_{t\ge 0}$ with $\|xV_t^*\|_2 < \infty$, where x denotes the identity map $x \mapsto x$.*

Proof. Given Proposition 3.5.3, this follows by a standard weak compactness argument for the L^2-bounded sequence $(\mathcal{T}_\varepsilon \nu_t^*)_{\epsilon>0}$. □

The next lemma will allow us to handle the critical 'boundary effects' in the smoothed estimate for $\partial_x^{-1}\mathcal{T}_\varepsilon \Delta_t$, by ensuring that these terms vanish as $\varepsilon \to 0$. This depends delicately on the behaviour of the mass of the solutions near the boundary. Thus, the essential ingredient is the boundary decay from Definition 3.2.3(iv), which is satisfied by the limit points because of Proposition 3.4.3.

Lemma 3.5.5. *Let μ satisfy the conditions in Definition 3.2.3 and let $g_t(y)$ be a (stochastic) function with $|g_t(y)| \lesssim 1 + |y| + M_t$, where $M_t := \langle \mu_t, \psi\rangle$ for some $\psi \in$ Lip$(\mathbb{R})$. Then*

$$\mathbb{E}\Big[\int_0^T \int_0^\infty \big|\big\langle \mu_t, g_t(\cdot) p_\varepsilon(x+\cdot)\big\rangle\big|^2 dx dt\Big] \to 0 \quad \text{as} \quad \varepsilon \to 0.$$

Proof. Notice first that, by Jensen's inequality,

$$\big|\big\langle \mu_t, g_t(\cdot) p_\varepsilon(x+\cdot)\big\rangle\big|^2 \le C\varepsilon^{-1} e^{-\frac{x^2}{\varepsilon}} \int_0^\infty (1 + y^2 + M_t^2) e^{-\frac{y^2}{\varepsilon}} d\mu_t(y),$$

and hence

$$\int_0^\infty \big|\big\langle \mu_t, g_t(\cdot) p_\varepsilon(x+\cdot)\big\rangle\big|^2 dx \le C \int_0^\infty (1 + y^2 + M_t^2)\varepsilon^{-\frac{1}{2}} e^{-\frac{y^2}{\varepsilon}} d\mu_t(y).$$

We divide the rest of the proof into three cases, where either 1, y^2 or M_t^2 appears in the integrand. The first case follows as in [73, Lem. 7.6]. For the second case, fix $0 < \varepsilon < 1$, and let $\eta \in (0,1)$ be a free parameter. Splitting the integral on $y \leq \varepsilon^\eta$ and its complement gives

$$\varepsilon^{-\frac{1}{2}} \int_0^\infty y^2 e^{-\frac{y^2}{\varepsilon}} d\mu_t(y) \leq \varepsilon^{-\frac{1}{2}} \Big(\mu_t(0, \varepsilon^\eta) + \langle \mu_t, y^2 \rangle \exp(-\varepsilon^{2\eta-1}) \Big). \tag{3.37}$$

By conditions (iii)-(iv) of Definition 3.2.3, and Lemma A.1.3, it follows that

$$\mathbb{E}\Big[\int_0^T \int_0^\infty y^2 \varepsilon^{-\frac{1}{2}} e^{-\frac{y^2}{\varepsilon}} d\mu_t(y) dt \Big] \leq C \varepsilon^{-\frac{1}{2}} \Big(\varepsilon^{\eta(1+\beta)} + \exp(-\varepsilon^{2\eta-1}) \Big) \tag{3.38}$$

for a constant $\beta > 0$, where the right-hand side of (3.38) converges to zero as long as

$$(2+2\beta)^{-1} < \eta < 2^{-1}.$$

For the final case, we can rely on Hölder's inequality to see that

$$\mathbb{E}\Big[\int_0^T M_t^2 \int_0^\infty \varepsilon^{-\frac{1}{2}} e^{-\frac{y^2}{\varepsilon}} d\mu_t dt \Big] \leq \mathbb{E}\Big[\int_0^T M_t^{2q} dt \Big]^{\frac{1}{q}} \mathbb{E}\Big[\int_0^T \Big| \int_0^\infty \varepsilon^{-\frac{1}{2}} e^{-\frac{y^2}{\varepsilon}} d\mu_t \Big|^p dt \Big]^{\frac{1}{p}}.$$

Since $M_t^{2q} \lesssim \mu_t(0,\infty) + \langle \mu_t, y^{2q} \rangle$ by the Lipschitz continuity of ψ, we can deduce that $\mathbb{E}\int_0^T M_t^{2q} dt < \infty$ for all $q > 1$ by Definition 3.2.3(iii) and Lemma A.1.3. Moreover, arguing precisely as we did in (3.37)–(3.38),

$$\mathbb{E}\Big[\int_0^T \Big| \int_0^\infty \varepsilon^{-\frac{1}{2}} e^{-\frac{y^2}{\varepsilon}} d\mu_t(y) \Big|^p dt \Big] \leq C \varepsilon^{-\frac{p}{2}} \Big(\varepsilon^{\eta(1+\beta)} + \exp(-p\varepsilon^{2\eta-1}) \Big),$$

so the claim follows by taking $p := 1 + \beta/2$ and choosing η in the range

$$(1+\beta/2)(2+2\beta)^{-1} < \eta < 2^{-1}.$$

This finishes the proof. □

3.5.3 The smoothed H^{-1} estimate: Proof of Proposition 3.5.1

In this section, we complete the proof of the 'smoothed' H^{-1} estimate as presented in Proposition 3.5.1.

Remark 3.5.6. As in [73], we denote by $o_{\mathrm{sq}}(1)$ any family of L^2 valued processes $\{(\zeta_t^\varepsilon)_{t \leq T}\}_{\varepsilon > 0}$ that satisfy $\mathbb{E}\int_0^T \|\zeta_t^\varepsilon\|_2^2 dt \to 0$ as $\varepsilon \downarrow 0$. For example, Lemma 3.5.5 and Lemma A.1.1 ensure that the 'boundary effects' and error terms will be $o_{\mathrm{sq}}(1)$.

As outlined at the start of Section 3.5.1, we will derive an equation for $\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t$ by testing the SPDE for ν and $\tilde{\nu}$ with the kernel $y \mapsto G_\varepsilon(\cdot, y)$ and then integrating the difference of the resulting expressions over (x, ∞) for $x > 0$. As in the proof of Proposition 3.5.3, we set $\mathfrak{b}_t := b_t - \alpha_t\mathfrak{L}'_t$ and $\tilde{\mathfrak{b}}_t := \tilde{b}_t - \tilde{\alpha}_t\tilde{\mathfrak{L}}'_t$. Then we can argue as in [73, Sec. 7], using Lemma 3.5.5 and Lemma A.1.1 in place of their counterparts in [73], to conclude that

$$\begin{aligned} d\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t = &- \big(\tilde{\mathfrak{b}}_t\mathcal{T}_\varepsilon\Delta_t + \delta_t^{\mathfrak{b}}\mathcal{T}_\varepsilon\nu_t\big)dt + \tfrac{1}{2}\partial_x\big(\sigma_t^2\mathcal{T}_\varepsilon\Delta_t + \mathcal{E}_{t,\varepsilon}^{\sigma^2} - \tilde{\mathcal{E}}_{t,\varepsilon}^{\sigma^2}\big)dt \\ &- \sigma_t(\tilde{\rho}\mathcal{T}_\varepsilon\Delta_t + \delta_t^{\rho}\mathcal{T}_\varepsilon\nu_t)dW_t^0 + o_{\mathrm{sq}}(1)dt + o_{\mathrm{sq}}(1)dW_t^0, \end{aligned}$$

where $\mathcal{E}^g$ is an error term defined in Lemma A.1.1, with $\tilde{\mathcal{E}}^g$ defined analogously only for $\tilde{\nu}$, and

$$\delta_t^g := g(t, x, \nu_t, L_t) - g(t, x, \tilde{\nu}_t, \tilde{L}_t).$$

Applying Itô's formula to the square, it follows that

$$\begin{aligned} d(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)^2 = &-2(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)\big(\tilde{\mathfrak{b}}_t\mathcal{T}_\varepsilon\Delta_t + \delta_t^{\mathfrak{b}}\mathcal{T}_\varepsilon\nu_t\big)dt \\ &+ (\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)\partial_x\big(\sigma_t^2\mathcal{T}_\varepsilon\Delta_t + \mathcal{E}_{t,\varepsilon}^{\sigma^2} - \tilde{\mathcal{E}}_{t,\varepsilon}^{\sigma^2}\big)dt \\ &+ \sigma_t^2\big(\tilde{\rho}\mathcal{T}_\varepsilon\Delta_t + \delta_t^{\rho}\mathcal{T}_\varepsilon\nu_t\big)^2dt \\ &- 2\alpha(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)\sigma_t\big(\tilde{\rho}\mathcal{T}_\varepsilon\Delta_t + \delta_t^{\rho}\mathcal{T}_\varepsilon\nu_t\big)dW_t \\ &+ (\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)o_{\mathrm{sq}}(1)dt + (\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)o_{\mathrm{sq}}(1)dW_t + o_{\mathrm{sq}}(1)^2dt. \end{aligned} \tag{3.39}$$

We can now integrate the above in $x \in \mathbb{R}_+$ to arrive at the L^2 norm of $\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t$. Then it remains to estimate the resulting integrals on the right-hand side—a procedure we split into five short steps (deviating substantially from [73]). For convenience, we define

$$|M|_{t,\star} := \sup_{s\le t}\{1 + |M_t| + |\tilde{M}_t|\} \text{ and } \|\mathcal{T}\nu\|_{t,\star} := \sup_{s\le t}\sup_{\varepsilon>0}\|(1+x)\mathcal{T}_\varepsilon\nu_s\|_2.$$

Step 1. We start by considering the first line on the right-hand side of (3.39). Note that we can write $2(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)(\mathcal{T}_\varepsilon\Delta_t) = \partial_x(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)^2$ in the first term. Hence, we can integrate by parts in x and use Young's inequality with free parameter θ to get

$$\begin{aligned} -\int_0^\infty 2\tilde{\mathfrak{b}}_t(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)(\mathcal{T}_\varepsilon\Delta_t)dx &= -\int_0^\infty(\partial_x\tilde{\mathfrak{b}}_t)(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)^2dx \\ &\quad + 2\tilde{\mathfrak{b}}_t(0)\int_0^\infty(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)(\mathcal{T}_\varepsilon\Delta_t)dx \\ &\le C_{\mathfrak{b}}\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t\|_2^2 + C_\theta C_{\mathfrak{b}}|M|_{t,\star}^2\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t\|_2^2 + \theta\|\mathcal{T}_\varepsilon\Delta_t\|_2^2. \end{aligned}$$

For the second integral in the first line of (3.39), we recall that

$$|\delta_t^{\mathfrak{b}}| \lesssim (|x| + |M|_{t,\star})\big(d_0(\nu_t, \tilde{\nu}_t) + |L_t - \tilde{L}_t| + |\mathfrak{L}'_t - \tilde{\mathfrak{L}}'_t|\big).$$

Therefore, using Cauchy–Schwarz on the d_0 term and Young's inequality on the other terms, we get

$$\begin{aligned}\Big|\int_0^\infty \delta_t^{\mathfrak{b}}(\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)\mathcal{T}_\varepsilon\nu_t dx\Big| &\lesssim |M|_{t,\star}\|\mathcal{T}\nu\|_{t,\star}\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t\|_2 d_0(\nu_t,\tilde{\nu}_t)\\ &\quad + |M|_{t,\star}^2\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t\|_2^2 + |M|_{t,\star}^2\|\mathcal{T}\nu\|_{t,\star}^2(|L_t - \tilde{L}_t|^2 + |\mathfrak{L}'_t - \tilde{\mathfrak{L}}'_t|^2),\end{aligned}$$

where we can note that

$$\begin{aligned}\int_0^t |M|_{s,\star}^2\|\mathcal{T}\nu\|_{s,\star}^2|\mathfrak{L}'_s - \tilde{\mathfrak{L}}'_s|^2 ds &\le |M|_{t,\star}^2\|\mathcal{T}\nu\|_{t,\star}^2\int_0^t\Big|\int_0^s \mathfrak{K}'(s-r)(L_r - \tilde{L}_r)dr\Big|^2 ds\\ &\le |M|_{t,\star}^2\|\mathcal{T}\nu\|_{t,\star}^2\|\mathfrak{K}'\|_1^2\int_0^t \big|L_r - \tilde{L}_r\big|^2 dr.\end{aligned}$$

Step 2. Now consider the second line on the right-hand-side of (3.39). By performing an integration by parts and using Young's inequality with free parameter θ, we get

$$\begin{aligned}&\int_0^\infty (\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t)\partial_x\big(\sigma_t^2\mathcal{T}_\varepsilon\Delta_t + \mathcal{E}_{t,\varepsilon}^{\sigma^2} - \tilde{\mathcal{E}}_{t,\varepsilon}^{\sigma^2}\big)dx\\ &\qquad = -\big\|\sigma_t\mathcal{T}_\varepsilon\Delta_t\big\|_2^2 - \int_0^\infty (\mathcal{T}_\varepsilon\Delta_t)(\mathcal{E}_{t,\varepsilon}^{\sigma^2} - \tilde{\mathcal{E}}_{t,\varepsilon}^{\sigma^2})dx\\ &\qquad \le -\big\|\sigma_t\mathcal{T}_\varepsilon\Delta_t\big\|_2^2 + C_\theta\big\|\mathcal{E}_{t,\varepsilon}^{\sigma^2} - \tilde{\mathcal{E}}_{t,\varepsilon}^{\sigma^2}\big\|_2^2 + \theta\big\|\mathcal{T}_\varepsilon\Delta_t\big\|_2^2,\end{aligned}$$

where we have used that $\mathcal{T}_\varepsilon\Delta_t$, $\mathcal{E}_{t,\varepsilon}^g$, and $\tilde{\mathcal{E}}_{t,\varepsilon}^g$ are zero at zero and vanish at infinity.

Step 3. In the third line on the right-hand side of (3.39), we expand the square and apply Young's inequality with free parameter θ to the lower order term. This yields

$$\begin{aligned}\int_0^\infty \sigma_t^2\big(\tilde{\rho}\mathcal{T}_\varepsilon\Delta_t + \delta_t^\rho\mathcal{T}_\varepsilon\nu_t\big)^2 dx &\le \big\|\sigma_t\tilde{\rho}_t\mathcal{T}_\varepsilon\Delta_t\big\|_2^2 + C_\sigma|\delta_t^\rho|^2\big\|\mathcal{T}_\varepsilon\nu_t\big\|_2^2\\ &\quad + C_\theta|\delta^\rho|^2\big\|\mathcal{T}_\varepsilon\nu_t\big\|_2^2 + \theta\big\|\mathcal{T}_\varepsilon\Delta_t\big\|_2^2.\end{aligned}$$

Also, we recall here that $|\delta_t^\rho|^2 \lesssim |M|_{t,\star}^2(|L_t - \tilde{L}_t|^2 + d_1(\nu_t,\tilde{\nu}_t)^2)$.

Step 4. When taking expectation in (3.39), the stochastic integrals vanish. Thus, by taking expectation and integrating over $x > 0$, it follows from Steps 1–3 and

Young's inequality with free parameter $\theta > 0$ that

$$\begin{aligned}\mathbb{E}\Big[\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_t\big\|_2^2\Big] &\leq C_\theta\mathbb{E}\Big[\int_0^t \|\mathcal{T}\nu\|_{s,\star}^2|M|_{s,\star}^2\big(|L_s-\tilde{L}_s|^2+d_1(\nu_s,\tilde{\nu}_s)^2\big)ds\Big]\\ &\quad + C_\theta\mathbb{E}\Big[\int_0^t |M|_{s,\star}^2\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2^2\,ds\Big]\\ &\quad + C_\theta\mathbb{E}\Big[\int_0^t |M|_{s,\star}\|\mathcal{T}\nu\|_{s,\star}\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2 d_0(\nu_s,\tilde{\nu}_s)ds\Big]\\ &\quad + \mathbb{E}\Big[\int_0^t\int_0^\infty(\sigma_s^2\tilde{\rho}_s^2-\sigma_s^2+2\theta)|\mathcal{T}_\varepsilon\Delta_s|^2dxds\Big]+o(1),\end{aligned}\tag{3.40}$$

as $\varepsilon\downarrow 0$, for a constant C_θ which only depends on the free parameter θ.

Step 5. Since ρ is bounded away from 1 and σ is bounded away from 0, we can take the free parameter θ sufficiently small so that (for all x and t)

$$\sigma(s,x)^2\rho(t,\tilde{\nu}_t)^2-\sigma(s,x)^2+2\theta\leq -c_0\tag{3.41}$$

for a fixed constant $c_0>0$. Next, we can consider the stopping times

$$t_n:=\inf\big\{t>0:\|\mathcal{T}\nu\|_{t,\star}^2>n\ \text{ or }\ |M|_{t,\star}^2>n\big\}\wedge T,$$

for $n\geq 1$, and notice that, by Proposition 3.5.3, we have $t_n\uparrow T$ as $n\to\infty$. Evaluating the estimate (3.40) at $t\wedge t_n$ and using (3.41), we get

$$\begin{aligned}&\mathbb{E}\Big[\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_{t\wedge t_n}\big\|_2^2\Big]+c_0\mathbb{E}\Big[\int_0^{t\wedge t_n}\big\|\mathcal{T}_\varepsilon\Delta_s\big\|_2^2ds\Big]\\ &\leq Cn^2\mathbb{E}\Big[\int_0^{t\wedge t_n}d_0(\nu_s,\tilde{\nu}_s)\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_s\big\|_2ds\Big]+Cn\int_0^t\Big[\mathbb{E}\big\|\partial_x^{-1}\mathcal{T}_\varepsilon\Delta_{s\wedge t_n}\big\|_2^2ds\Big]\\ &\quad + Cn^2\mathbb{E}\Big[\int_0^{t\wedge t_n}|L_s-\tilde{L}_s|^2ds\Big]+Cn^2\mathbb{E}\Big[\int_0^{t\wedge t_n}d_1(\nu_s,\tilde{\nu}_s)^2ds\Big]+o(1).\end{aligned}$$

Finally, by applying the integrating factor $\exp(-Cnt)$ to the first term on the right-hand side, we obtain the estimate from Proposition 3.5.1 with $c_n:=n^2(e^{CT}-1)$. This completes the proof.

Chapter 4

Instantaneous Feedback ⁂ Part I: Local Well-Posedness and Regularity up to Explosion

In this chapter we analyse the local well-posedness of a contagious McKean–Vlasov system where the positive feedback from absorptions affects the particles instantaneously. As discussed in Section 3.2.3.1 of the previous chapter, such a system appears naturally as the limiting case of the model presented there, if we consider a sequence of impact kernels $\mathfrak{K}_\epsilon$ converging to a Dirac mass at 0. Alternatively, one can develop a model for instantaneous contagion directly—without reference to the setting of Chapter 3—which is the route we take here. Interestingly, even in the large population limit, the positive feedback effect can now become singular and thus generate blow-ups (which we already hinted at in Section 3.2.3.1).

4.1 Introduction

As our motivating example, we consider a large financial system consisting of N banks. Following the structural approach to credit risk, we say that the i'th bank defaults when its asset value, A^i_t, hits a default barrier, D^i_t. This gives rise to the notion of *distance-to-default*, for which a simple model could be of the form

$$Y^i_t := \log(A^i_t) - \log(D^i_t) = X^i_0 + \int_0^t b(s)ds + \int_0^t \sigma(s)dB^i_s, \qquad i = 1, \dots, N,$$

where $X^1_0, \dots, X^N_0$ are i.i.d. copies of X_0, and $B^1, \dots, B^N$ are independent Brownian motions (independent of the initial conditions). Next, we can introduce an element of contagion with a model in which the default of one bank causes the other banks to

lose a proportion α/N of their assets. For large N, we have $1-\alpha/N \simeq \exp\{-\alpha/N\}$, so the new asset values, $\hat{A}$, can then be defined by

$$\hat{A}^{i,N}_t := \prod_{j=1}^{N} \exp\Big\{-\frac{\alpha}{N}\mathbf{1}_{t\geq\tau^{j,N}}\Big\}A^i_t = \exp\Big\{-\frac{\alpha}{N}\sum_{j=1}^{N}\mathbf{1}_{t\geq\tau^{j,N}}\Big\}A^i_t, \quad t<\tau^{i,N},$$

where $\tau^{i,N} := \inf\{t>0 : X^{i,N}_t \leq 0\}$ for $X^{i,N}_t := \log(\hat{A}^{i,N}_t) - \log(D^i_t)$. After taking logarithms, it follows that the new distances-to-default, $X^{i,N}$, have dynamics given by the interacting particle system

$$dX^{i,N}_t = b(t)dt + \sigma(t)dB^i_t - \alpha dL^N_t \quad \text{with} \quad L^N_t = \frac{1}{N}\sum_{j=1}^{N}\mathbf{1}_{t\geq\tau^{j,N}}, \tag{4.1}$$

for $i=1,\dots,N$, with the same initial conditions as above.

We note that one must be a little careful with the rigorous interpretation of the loss process, L^N, as it is part of the dynamics in (4.1). Nevertheless, its meaning should be entirely transparent, so we leave the details of this to Chapter 5 (see in particular Section 5.3) where we analyse the convergence of the particle system. For now, we shall be satisfied with knowing that the system does indeed converge along subsequences—in a suitable sense—and that the limit points are solutions to the McKean–Vlasov problem (MV) presented below.

As opposed to considering the convergence of the particle system, this chapter studies the mean-field version of the problem directly, namely (MV), by way of a fixed point approach. As regards the particular form of the coefficients in (4.1), we will concentrate the analysis in this chapter on the simplest case (MV) and then we devote the final Section 4.7 to a discussion of more general coefficients including the addition of a spatially dependent drift. Without further ado, the subject of this chapter is the McKean–Vlasov problem

$$\begin{cases} X_t = X_0 + B_t - \alpha L_t \\ \tau = \inf\{t\geq 0 : X_t \leq 0\} \\ L_t = \mathbb{P}(\tau \leq t), \end{cases} \tag{MV}$$

where $\alpha\in\mathbb{R}$ is a constant, B is a standard Brownian motion and X_0 is an independent random variable distributed on the positive half-line. We denote the law of the initial condition X_0 by ν_0. A solution to this problem is a deterministic and initially zero càdlàg function $t\mapsto L_t$ that is increasing and for which (MV) holds for any Brownian motion B. Viewing (MV) as an SDE in X, notice that there is no distinction to be made between strong and weak notions of solution: knowing L fixes the law of X and, together with any Brownian motion, L fixes a pathwise construction of X.

When $\alpha > 0$, the equations have a positive feedback effect and this is the case we consider here — the situation for $\alpha \leq 0$ is classical, and existence and uniqueness of smooth solutions for the corresponding Fokker–Planck equation is known already from [17, 29, 30].

4.1.1 Antecedent literature motivated by neuroscience

We emphasise that the first version of the problem (MV) appeared in the mathematical neuroscience literature as a mean-field limit of a large network of electrically coupled neurons [17, 29, 45, 46], similarly to the discussion in Section 3.2.3.2 of the previous chapter. In particular, the (subsequential) convergence of (4.1) has already been addressed in [46], albeit for a slight variant of the system with constant volatility.

To be precise, the neuroscience setting can be described as follows: Each neuron is identified with an electrical potential (given by an SDE) and when it reaches a threshold voltage, the neuron fires an electrical signal to the other neurons, which then become excited to higher voltage levels. After reaching the threshold, the neuron is instantaneously reset to a predetermined value and it then continues to evolve according to this rule indefinitely. Therefore, the model is different to our setting as we do not reset the mean-field particle in (MV), however, the essential mathematical difficulties from the positive feedback remain common to both models.

4.2 Main results: On local well-posedness and the occurence of blow-ups

Mathematically, the McKean–Vlasov problem (MV) can be recast in a number of ways. We denote the law of X_t killed at the origin by ν_t and set $\nu_t(\phi) := \mathbb{E}[\phi(X_t)\mathbf{1}_{t<\tau}]$, for suitable test functions ϕ. Then, provided X is continuous, a simple application of Itô's formula gives the nonlinear PDE

$$\begin{cases} \langle \nu_t, \phi \rangle = \langle \nu_0, \phi \rangle + \frac{1}{2} \int_0^t \langle \nu_s, \partial_{xx}\phi \rangle ds - \alpha \int_0^t \langle \nu_s, \partial_x \phi \rangle dL_s \\ L_t = 1 - \int_0^\infty d\nu_s(x), \end{cases} \tag{4.2}$$

for $\phi \in C^2$ with $\phi(0) = 0$ and $t \geq 0$. Writing V_t for the density of ν_t (which exists by Proposition 4.3.1), we can formally integrate by parts in (4.2) to find the Dirichlet problem

$$\partial_t V_t(x) = \tfrac{1}{2}\partial_{xx} V_t(x) + \alpha L_t' \partial_x V_t(x), \qquad L_t' = \tfrac{1}{2}\partial_x V_t(0), \qquad V_t(0) = 0, \tag{4.3}$$

on the positive half-line. In other words, the law of X (absorbed at the origin) solves the heat equation on $(0,\infty)$ with a drift term proportional to the flux across the Dirichlet boundary at zero — the latter being a highly singular nonlocal nonlinearity. Setting $v(t,x) := -V_t(x-\alpha L_t)$ in (4.3), the equations for v and L can be viewed as a Stefan problem with supercooling on the semi-infinite strip $(\alpha L_t, \infty)$. From this point of view, it is known that L_t' may explode in finite time as has been analysed (on a finite strip) in the series of papers [55, 56, 57, 77].

A third characterisation of (MV) is as the solution to the integral equation

$$\int_0^\infty \Phi\Big(-\frac{x-\alpha L_t}{t^{1/2}}\Big)\nu_0(dx) = \int_0^t \Phi\Big(\alpha\frac{L_t-L_s}{(t-s)^{1/2}}\Big)dL_s, \tag{4.4}$$

where Φ is the c.d.f. of the Normal distribution. This is a Volterra integral equation of the first kind, and a derivation can be obtained by following [98]. The formulation that is most helpful here is to view the problem as a fixed point of the map Γ defined by

$$\begin{cases} X_t^\ell = X_0 + B_t - \alpha\ell_t \\ \tau^\ell = \inf\{t\geq 0 : X_t^\ell \leq 0\} \\ \Gamma[\ell]_t = \mathbb{P}(\tau^\ell \leq t), \end{cases} \tag{4.5}$$

that is, a solution to $\Gamma[L] = L$. We will use this map in stating and proving our main theorems. The key is to find a suitable space on which Γ stabilises and to show that it is contractive (Theorems 4.2.6 & 4.2.7). Note that [45, 94] also take this approach to the problem, but our techniques differ in that they are entirely probabilistic and do not rely on PDE estimates.

A very interesting feature of the problem (MV) is that it exhibits a phase transition in the continuity of solutions as the feedback strength, α, increases. The methods of [45], for the neuroscience version of the problem, show that for α sufficiently small (and $\nu_0 = \delta_x$ with $x > 0$) there is a unique solution to (MV) in the class of continuously differentiable functions. On the other hand, the extremely simple proof below (modified from [17]) shows that (MV) cannot have a continuous solution for values of α that are sufficiently large. When this is the case, the positive feedback becomes too great and at some point in time the loss process, L, undergoes a jump discontinuity, which we call a *blow-up*. In other words, in an infinitesimal period of time, a macroscopic proportion of the mass in the PDE (4.2) is lost at the boundary — i.e. the system blows up (see Figure 4.2.1). It is intriguing that this result can be proved so simply and with no technical estimates, although of course the threshold obtained below is not sharp and the proof does not reveal anything about the nature of the blow-ups.

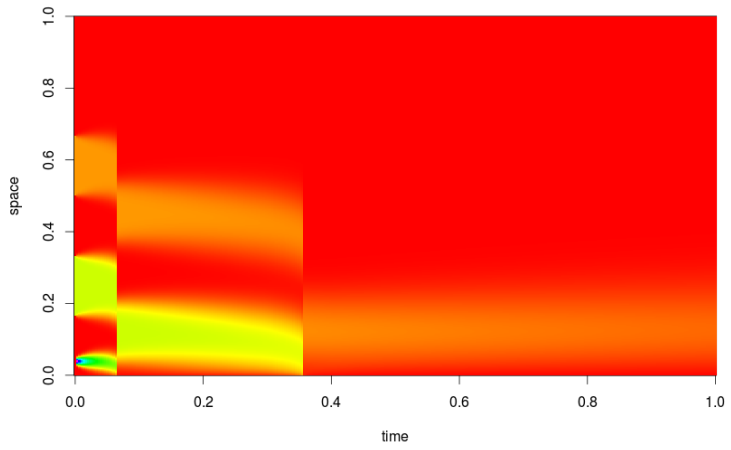


Figure 4.2.1: Example of a solution to (4.2, 4.3) showing two blow-up times. The pixel intensity represents the value of the solution density at that space-time coordinate. The initial condition is a linear combination of indicator functions of three disjoint sets.

Theorem 4.2.1 (Blow-up for large α)**.** *Let* $m_0 := \int_0^\infty x\nu_0(dx)$*. If* $\alpha > 2m_0$*, then any solution to (MV) cannot be continuous for all times.*

Proof. For a contradiction, suppose L solves (MV) and is continuous. Stopping X at the hitting time τ, we then get

$$0 \leq X_{t\wedge\tau} = X_0 + B_{t\wedge\tau} - \alpha L_{t\wedge\tau}.$$

Taking expectations and rearranging,

$$m_0 \geq \alpha\mathbb{E}[L_{t\wedge\tau}] \to \alpha\mathbb{E}[L_\tau] = \alpha\int_0^\infty L_s dL_s \qquad \text{as } t \to \infty,$$

where we note that $\tau < \infty$ a.s. and $L_\infty = 1$, since Brownian motion hits every level with probability 1. As L is continuous and increasing, the integral can be computed exactly, so we get

$$m_0 \geq \tfrac{1}{2}\alpha(L_\infty^2 - L_0^2) = \tfrac{1}{2}\alpha,$$

which is the required contradiction. □

From Theorem 4.2.1 it is clear that, in general, we cannot restrict our search to solutions of (MV) that are continuous and so we must allow càdlàg solutions. If a

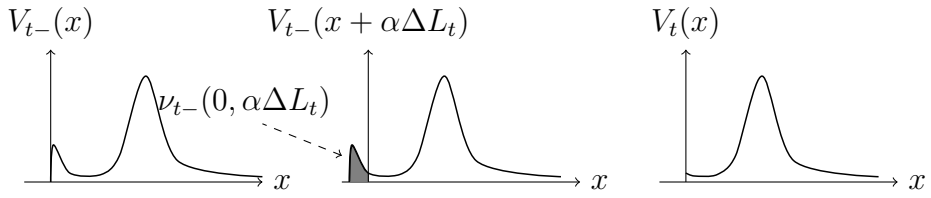


Figure 4.2.2: On the left, V_{t-} is the density just before a jump of size ΔL_t. This density is then translated by $\alpha\Delta L_t$ and the mass falling below the boundary at zero equals the change in the loss, which gives (4.6). After the jump, the system is restarted from the density on the right. Notice that, in general, this new initial condition will not vanish at the origin.

solution has a jump of size ΔL_t at time t, then the instantaneous loss must equal the mass of ν_{t-} absorbed at the boundary after a translation by $-\alpha\Delta L_t$, that is:

$$\nu_{t-}(0, \alpha\Delta L_t) = \Delta L_t, \tag{4.6}$$

see Figure 4.2.2. Unfortunately, this equation alone is not sufficient to determine the jump sizes of a discontinuous solution. In particular, $\Delta L_t = 0$ is always a solution, but in general this is an invalid jump size by Theorem 4.2.1. To continue the solution after a blow-up, we must decide how to choose the jump size from the solution set of (4.6). In [46] the authors introduce the term *physical solution* for a solution, L, that satisfies

$$\Delta L_t = \inf\{x \geq 0 : \nu_{t-}(0, \alpha x) < x\} \tag{PJC}$$

for all $t \geq 0$. The next result justifies that this is a natural condition, as it is the smallest possible choice of jump size that admits càdlàg solutions (see Section 4.3 for a proof and further discussion).

Proposition 4.2.2 (Physical solutions have minimal jumps)**.** *Suppose L is any càdlàg process satisfying (MV). Then*

$$\Delta L_t \geq \inf\{x \geq 0 : \nu_{t-}(0, \alpha x) < x\},$$

for every $t \geq 0$. In particular, if the right-hand side is non-zero for some $t \geq 0$, then L has a blow-up at time t.

It is helpful to consider the density function, V_{t-}, of ν_{t-} in light of Proposition 4.2.2. If, at some time t, V_{t-} is greater than or equal to the critical value of α^{-1} on a non-zero interval about the origin, then a blow-up in L is forced to occur at that time.

In [46] it is shown that there exist (global) physical solutions to the neuroscience version of equation (MV), for any $\alpha > 0$, albeit with an initial measure ν_0 that vanishes in a neighbourhood of zero. Those solutions arise as topological limit points of a corresponding finite particle system analogous to (4.1). Consequently, the authors are unable to establish regularity results on the solutions. The main advantage of a fixed point argument is that solutions are guaranteed to have known regularity, at least on a small time interval.

Our motivation in this chapter is to make progress towards the following conjecture:

Conjecture 4.2.3 (Global uniqueness)**.** *We conjecture that solutions to (4.2.1) which satisfy the physical jump condition (PJC) are unique. Furthermore, we conjecture that the solution is continuously differentiable between jump times, and we predict $\sqrt{t}$-singularities immediately before and after jumps, which is to say: For all $t_0 > 0$ that satisfy $\Delta L_{t_0} > 0$, we have $L'_{t_0+h} = O(|h|^{-1/2})$, as $h \to 0$.*

Currently, we are far from proving Conjecture 4.2.3, and the major obstruction concerns the initial conditions. To see why, notice that after a jump has taken place, we must, in general, restart the system from an initial law that has a density which does not vanish at the origin (see Figure 4.2.2, as well as Proposition 4.3.1 concerning the existence of a density for ν at all times). In fact, without further analysis, all we can say about the measure ν_t after a blow-up at time t is that

$$\inf\{x \geq 0 : \nu_t(0, \alpha x) < x\} = 0.$$

Therefore, to attack Conjecture 4.2.3 it is necessary to make progress towards the following simpler goal:

Conjecture 4.2.4 (Uniqueness for non-vanishing initial laws)**.** *Suppose ν_0 has a density V_0 and satisfies $\inf\{x \geq 0 : \nu_0(0, \alpha x) < x\} = 0$. Then we conjecture that the solution, L, to (MV) is unique and continuously differentiable up to a small time with a singularity at $t = 0$. Moreover, provided $V_0(0+) < \alpha^{-1}$, we predict that the singularity is of the form $L'_t = O(t^{-1/2})$ as $t \downarrow 0$.*

At the time when we wrote the paper [74], which underlies this chapter, initial conditions that do not vanish at the origin were outside the scope of known results. (See Chapter 6 for some new developments that—at least partially—answer the above conjectures building on the contributions of this chapter).

The results that existed before [74] only tackled uniqueness in a too restrictive class of candidate solutions. To be precise, in the forerunning literature, the closest

to Conjecture 4.2.4 is [94, Thm. 2.6] which is established for a slight variant of our problem: Given an initial density $V_0 \in H^1(0,\infty)$ with $V_0(0) = 0$, it gives existence and uniqueness up to the first time the L^2-norm of L' explodes in the class of candidate solutions for which $L \in H^1(0,t_0)$ for some $t_0 > 0$. Here $H^1(a,b)$ denotes the usual Sobolev space with one weak derivative in $L^2(a,b)$. We note that these conditions on V_0 imply that it decays like $o(x^{1/2})$ near the origin, so it is far from what is needed in order to proceed to a global uniqueness theory. The aforementioned was preceded by [45, Thm. 4.1], which gives existence and uniqueness on a small time interval in the class of C^1 solutions, starting from an initial density that decays like $O(x)$ as $x \to 0$. The connection between the boundary decay of the initial condition and the short-time regularity of solutions will feature prominently in our results below.

Moreover, the results in the literature predating this chapter left open the question of whether there is (short-time) uniqueness in the wider class of càdlàg solutions. This is more than just a technical curiosity: Indeed, the physical and financial motivations for (MV) derive from the corresponding particle system, as presented in (4.1), but the requisite regularity of its limit points is not known and could be difficult to verify a priori — all we know is that they are càdlàg. Thus, it obstructs the full convergence in law of the finite system (i.e. the propagation of chaos).

4.2.1 Uniqueness up to explosion

Our contribution here is to construct an H^1 solution to (MV) up to the first time its H^1 norm explodes, and to show that it is the unique solution on this time interval amongst all possible càdlàg solutions (Theorem 4.2.8). We allow initial densities with any power law decay at the origin, that is, decay of order $O(x^\beta)$ as $x \to 0$ for some $0 < \beta < 1$. To be precise, we consider initial conditions, ν_0, that have a bounded initial density, V_0, for which we can find constants $C, D, x_\star > 0$ such that

$$V_0(x) \leq Cx^\beta \mathbf{1}_{x \leq x_\star} + D\mathbf{1}_{x > x_\star}, \qquad \text{for every } x > 0. \tag{4.7}$$

A key observation is that it is only the values of C, $x_\star$ and β that determine the behaviour of solutions in short time (Theorem 4.2.7). The reason for assuming an initial density is that this state is reached after any non-zero period of time anyway (Proposition 4.3.1). Below, we state our main results chronologically to make it clear how they are connected.

First we introduce the following subsets of H^1:

Definition 4.2.5. For $\gamma \in (0, \frac{1}{2})$, $A > 0$ and $t_0 > 0$, let $\mathcal{S}(\gamma, A, t_0)$ denote the subset of $H^1(0, t_0)$ given by

$$\mathcal{S}(\gamma, A, t_0) := \{\ell \in H^1(0, t_0) : \ell'_t \leq At^{-\gamma} \text{ for a.e. } t \in [0, t_0]\}.$$

Our first argument is to show that, on these sets, Γ is an L^∞-contraction. The proof follows by comparing the first hitting times of a single Brownian motion driven by two different drift functions (Proposition 4.4.1), which is a coupling of the X^ℓ processes in (4.5) to the same Brownian motion B. As a by-product of this method we can deduce that differentiable solutions are minimal in the class of potential càdlàg solutions, which is an important ingredient in the main uniqueness result (Theorem 4.2.8). The power of the technique lies in the ability to estimate the positive part of the difference $(L_t - \bar{L}_t)^+$, for two solutions L and $\bar{L}$, assuming only the regularity properties of L and not $\bar{L}$.

Theorem 4.2.6 (Contraction and minimality)**.** *For any $\gamma \in (0, \frac{1}{2})$ and $A > 0$, there exists $t_0 > 0$ such that for all $\ell, \bar{\ell} \in \mathcal{S}(\gamma, A, t_0)$*

$$\|\Gamma[\ell] - \Gamma[\bar{\ell}]\|_{L^\infty(0,t_0)} \leq \tfrac{1}{2}\|\ell - \bar{\ell}\|_{L^\infty(0,t_0)}.$$

Moreover, if there exists a solution $L \in \mathcal{S}(\gamma, A, t_0)$ to (MV) for some $\gamma \in (0, \frac{1}{2})$, $A > 0$ and $t_0 > 0$ and another solution $\bar{L}$ that is càdlàg, then $\bar{L}_t \geq L_t$ for all $t < t_0$.

To construct a fixed point solution, we find choices of the parameters γ, A and t_0 such that the map Γ stabilises on the set $\mathcal{S}(\gamma, A, t_0)$. An interesting product of our technique is that we are able to recover the exact regularity of the solutions at time zero from the decay of the initial condition near the Dirichlet boundary. Thus, the main factor in the short-time growth of solutions is the behaviour of the heat equation with absorbing boundary conditions and not the feedback effect in the model. This should be an indication that something new is required to tackle Conjecture 4.2.3. Crucially, we rely on Girsanov's Theorem to control $t \mapsto B_t - \alpha\ell_t$, for which we require $\ell \in H^1(0, t_0)$. Notice, however, that if $V_0(+0) > 0$ as in the hypothesis of Conjecture 4.2.3, then we must have $\ell_t \geq \text{const.} \times \sqrt{t}$ (by comparison with the case $\alpha = 0$), and so we can no longer expect this approach to be viable.

Theorem 4.2.7 (Stability and fixed point)**.** *There exists a constant $K > 0$ depending only on β, C and $x_\star$ such that for all $\varepsilon > 0$ there exists t_0 for which*

$$\Gamma : \mathcal{S}(\tfrac{1-\beta}{2}, K + \varepsilon, t_0) \to \mathcal{S}(\tfrac{1-\beta}{2}, K + \varepsilon, t_0).$$

Hence Γ has a fixed point, $L = \Gamma[L]$, in $\mathcal{S}(\frac{1-\beta}{2}, K + \varepsilon, t_0)$ and this solution is unique in the class of candidate solutions in $\bigcup_{0<\gamma<1/2,\, A>0} \mathcal{S}(\gamma, A, t_0)$.

This short-time fixed point construction can be extended by observing that if L is in the space $H^1(0,t_0)$, then, by appealing to Girsanov's Theorem, we can show that the density V_{t_0-} must have decay of order $O(x^\beta)$ near the Dirichlet boundary. In other words, we can recover the regularity of the initial condition at time t_0, and so exactly the same argument can be applied to the system restarted from time t_0. Consequently, we can extend the solution constructed in Theorem 4.2.7 onto another small non-zero time interval and iterate this procedure, so long as the solution we obtain is in H^1. The catch that prevents this procedure from giving a differentiable solution for all times (recall that this would contradict Theorem 4.2.1) is that we lose control of the constants in the relevant $\mathcal{S}$ subsets. Hence the construction only applies up to the first time the H^1-norm of the solution explodes. Trivially, this occurs no later than the first jump, however, the mathematical challenge in attempting to restart the problem after an explosion time seems no less difficult than Conjecture 4.2.4.

Theorem 4.2.8 (Uniqueness up to explosion)**.** *There exists a solution, L, to (MV) up to time*

$$t_{\text{explode}} := \sup\{t > 0 : \|L\|_{H^1(0,t)} < \infty\} \in (0, \infty]$$

such that, for every $t_0 < t_{\text{explode}}$, we have $L \in \mathcal{S}(\frac{1-\beta}{2}, K, t_0)$ for some constant $K > 0$. Furthermore, L is the unique solution on $[0, t_{\text{explode}})$: If $\bar{L}$ is any other generic càdlàg solution to (MV) on $[0,t]$ for some $t < t_{\text{explode}}$, then $L_s = \bar{L}_s$ for all $s \in [0,t]$.

The analogous result for the McKean–Vlasov problem corresponding to (4.1) is presented in Section 4.7. The first step towards uniqueness is given by Theorem 4.2.6: The differentiable solution constructed in the first part of Theorem 4.2.8 is minimal, so it must be a lower bound of any generic candidate càdlàg solution, at least in small time. To obtain an upper bound, we introduce a family of processes in which we kill a small proportion, $\varepsilon > 0$, of the initial condition at time zero. In Section 4.6 we show that these modified solutions cannot overlap and so they bound the generic solutions from above. By returning to the contraction argument in Theorem 4.2.6, we show that these modified solutions converge to the differentiable solution as the amount of deleted mass, ε, tends to zero. Thus, the envelope of solutions shrinks to zero size and this forces uniqueness of solutions. The power of this method is that it circumvents the need to have quantitative information about the generic candidate càdlàg solutions.

Remark 4.2.9 (Propagation of chaos). Theorem 4.2.8 resolves ambiguity about the validity of the propagation of chaos. By following the methods in [46], it can be shown that the particle system in (4.1) has limit points that converge in law (with respect to a suitable topology) to a càdlàg solution of (MV). Now that we have uniqueness amongst general càdlàg solutions, we can conclude that this is in fact full convergence, up to the first explosion time, and not just subsequential convergence.

4.2.2 Overview of the chapter

In Section 4.3 we motivate and explain the physical condition (PJC) and prove that it is necessary in order to have a càdlàg solution. In Section 4.4 we prove Theorem 4.2.6 via a comparison argument. In Section 4.5 we prove Theorem 4.2.7 by finding a space on which the map Γ stabilises. In Section 4.6 we prove Theorem 4.2.8 by extending the fixed point argument up to the explosion time and introducing the ε-deleted solutions used to bound candidate solutions. In Section 4.7 we show how our arguments can be extended to incorporate more general drift and diffusion coefficients.

4.3 Minimal jumps and α-fragility: proof of Proposition 4.2.2

Our aims in this section are twofold: Firstly, we prove that the physical jump condition in (PJC) yields the solution with the smallest possible jump size (at any given instance) and, secondly, we provide some intuition for this behaviour.

We begin by observing that regardless of the initial condition, at any time $t > 0$, the measure ν_t will have a density.

Proposition 4.3.1 (Existence of a density process)**.** *Let $(\nu_t)_{t\geq 0}$ be the law of a solution, X, to (MV) absorbed at the origin. For all $t > 0$, ν_t is absolutely continuous with respect to the Lebesgue measure and therefore has a density function, $V_t : (0, \infty) \to (0, \infty)$. Moreover, if ν_0 has a density $V_0 \in L^\infty$, then $\|V_t\|_\infty \leq \|V_0\|_\infty$.*

Proof. By omitting the killing at zero, we have the upper bound

$$\nu_t(S) \leq \mathbb{P}(X_0 + B_t - \alpha L_t \in S) = \int_S \int_0^\infty p_t(x_0 + x - \alpha L_t)\nu_0(dx_0)dx,$$

where p_t is the Brownian transition kernel. Hence $\nu_t(S) \leq (2\pi t)^{-1/2}\mathrm{Leb}(S)$, so the first claim follows from the Radon–Nikodym Theorem. If ν_0 has a density V_0 in L^∞, then we get

$$V_t(x) \leq \int_0^\infty p_t(x_0 + x - \alpha L_t)V_0(x_0)dx_0 \leq \|V_0\|_\infty ,$$

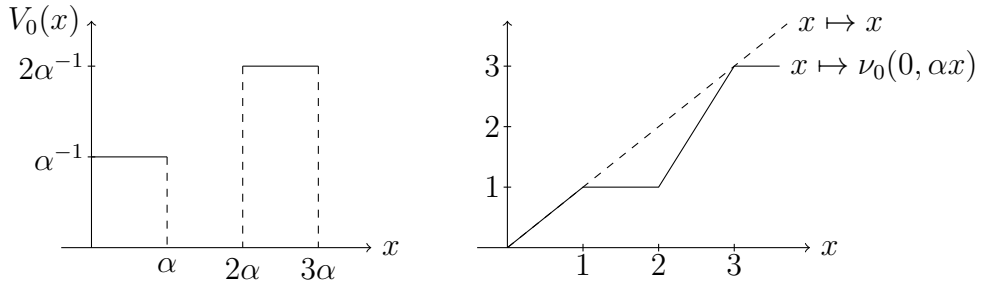


Figure 4.3.1: The function from Example 4.3.2 is on the left. The candidate jump sizes—that is, the solutions to (4.6)—are the points on the right where the graphs intersect. The point $(1, 1)$ gives the minimal allowed jump size of 1, since $x = 1$ is the value given by (PJC)

since $\int_{\mathbb{R}} p_t(y)dy = 1$. This proves the second claim. □

As remarked in Section 6.4, equation (4.6) alone is insufficient to determine the jump size at a given time. Not only is zero always a solution of that equation, but there may indeed be many other solutions that differ from the physical solution given by (PJC).

Example 4.3.2. Suppose we have an initial law ν_0 with density

$$V_0(x) = \alpha^{-1}\mathbf{1}_{0<x<\alpha} + 2\alpha^{-1}\mathbf{1}_{2\alpha<x<3\alpha}.$$

Then any jump size $\Delta L_0 \in [0, 1] \cup \{3\}$ solves (4.6), however, the physical solution is given by the condition $\Delta L_0 = 1$, see Figure 4.3.1. In particular, this example conveys very clearly the extent of the a priori non-uniqueness of (MV) if we do not insist on the physical jump condition. By at least insisting on right-continuity, Proposition 4.2.2 already rules out the continuum of jump sizes below 1, however, the jump size of 3 is only ruled out by declaring it to be non-physical.

An alternative way to view the physical jump condition (PJC), which better motivates Proposition 4.2.2, is the notion of *fragility*. To state a proper definition, consider first the sequence

$$\begin{aligned} f_0(\mu, \alpha, \varepsilon) &:= \mu(0, \alpha\varepsilon) \\ f_{n+1}(\mu, \alpha, \varepsilon) &:= \mu(0, \alpha\varepsilon + \alpha f_n(\mu, \alpha, \varepsilon)), \qquad \text{for } n \geq 0, \end{aligned} \tag{4.8}$$

for a given measure μ and constants $\varepsilon > 0$ and $\alpha > 0$. Notice that $f_n(\mu, \alpha, \varepsilon)$ is increasing in n and in ε. Hence we can deduce that the limit $f_\infty(\mu, \alpha, \varepsilon)$ exists and is an increasing function of ε. Therefore, we can sensibly define:

Definition 4.3.3 (α-fragility)**.** The measure, μ, is said to be *α-fragile* if

$$\lim_{\varepsilon \to 0} f_\infty(\mu, \alpha, \varepsilon) \neq 0.$$

To see why α-fragility is related to physical solutions, consider starting the heat equation from an initial law ν_0. In small time, we will immediately lose mass at the boundary, say an ε amount. To approximate the contagious system in (MV), we must then shift the measure down towards the origin by an amount $\alpha\varepsilon$. If we apply this to ν_0, we obtain a loss of $f_0(\nu_0, \varepsilon, \alpha)$, which we can then further shift our initial condition by, and so on, hence obtaining $f_\infty(\nu_0, \alpha, \varepsilon)$ in the limit. If this terminal quantity does not shrink to zero with ε, we should not expect the solution to (MV) to be right-continuous at $t = 0$: An infinitesimal loss of mass starts a cascade of losses summing to a non-zero amount. Indeed, this is a heuristic version of the argument presented in the proof of Proposition 4.2.2 below. For now, we notice that there is, in fact, an exact correspondence between $f_\infty(\mu, \alpha, 0+)$ and the physical jump condition (PJC).

Proposition 4.3.4. *For any atomless measure μ on the positive half-line and $\alpha > 0$ we have*

$$\lim_{\varepsilon \to 0} f_\infty(\mu, \alpha, \varepsilon) = \inf\{x \geq 0 : \mu(0, \alpha x) < x\}.$$

Proof. Write $f := \lim_{\varepsilon \to 0} f_\infty(\mu, \alpha, \varepsilon)$ and

$$x_0 := \inf\{x \geq 0 : \mu(0, \alpha x) < x\}.$$

Suppose $f < x_0$. Then we can find $\varepsilon > 0$ such that $f_\infty(\mu, \alpha, \varepsilon) + \varepsilon < x_0$. By taking the limit as $n \to \infty$ in (4.8) we get

$$f_\infty(\mu, \alpha, \varepsilon) = \mu(0, \alpha\varepsilon + \alpha f_\infty(\mu, \alpha, \varepsilon)) \geq \varepsilon + f_\infty(\mu, \alpha, \varepsilon),$$

where the inequality is due to the definition of x_0, but this is clearly a contradiction.

Suppose $f > x_0$. By definition of x_0, we can find a sequence $(x_\delta) \downarrow x_0$ as $\delta \downarrow 0$ such that

$$\mu(0, \alpha x_\delta) \leq x_\delta - \delta.$$

Fix any such $\delta > 0$ and take $\varepsilon \in (0, x_\delta \wedge \delta)$, then we have

$$f_0(\mu, \alpha, \varepsilon) = \mu(0, \alpha\varepsilon) \leq \mu(0, \alpha x_\delta) \leq x_\delta - \delta \leq x_\delta - \varepsilon.$$

Since $f_0(\mu, \alpha, \varepsilon) + \varepsilon \leq x_\delta$ we may now repeat the argument to get

$$f_1(\mu, \alpha, \varepsilon) = \mu(0, \alpha\varepsilon + \alpha f_0(\mu, \alpha, \varepsilon)) \leq \mu(0, \alpha x_\delta) \leq x_\delta - \delta \leq x_\delta - \varepsilon,$$

and so on. Proceeding inductively, we conclude that $f \leq x_\delta - \delta$, so taking $\delta \downarrow 0$ gives $f \leq x_0$, which is a contradiction. □

Example 4.3.5. Returning to the density V_0 from Example 4.3.2, we define $N_\varepsilon := \lfloor \varepsilon^{-1} - 1 \rfloor$ for $\varepsilon < 1$. Then $\alpha(n+1)\varepsilon \leq \alpha$ for all $n \leq N_\varepsilon$, so we get

$$f_n(V_0, \alpha, \varepsilon) = (n+1)\varepsilon \leq 1 \qquad \forall n \leq N_\varepsilon.$$

Thereafter, we have $\alpha < \alpha(N_\varepsilon + 2)\varepsilon < 2\alpha$, so $f_{N_\varepsilon+1}(V_0, \alpha, \varepsilon) = 1$ and, since $\alpha < \alpha\varepsilon + \alpha < 2\alpha$, it follows that $f_n(V_0, \alpha, \varepsilon)$ remains equal to 1 for all $n > N_\varepsilon$. Consequently, we have $f_\infty(V_0, \alpha, \varepsilon) = 1$ for all $\varepsilon < 1$, and hence $f_\infty(V_0, \alpha, 0+) = 1$, which indeed agrees with the physical jump condition.

Remark 4.3.6 (Restarting solutions at a non-zero time). To prove Proposition 4.2.2 it will be sufficient to take $t = 0$ in the statement of the result, since we can always restart the system at a time $t > 0$, taking t as our new time origin. This might be confusing at first glance as we specified $L_0 = 0$ in our definition of a solution to (MV). To be clear, if we want to solve (MV) from a time $t > 0$ onwards, and we have a solution L up to time t, then we can solve the problem

$$\begin{cases} \widetilde{X}^x_s = x + B_s - \alpha \widetilde{L}_s \\ \widetilde{\tau}^x = \inf\{s \geq 0 : \widetilde{X}^x_s \leq 0\} \\ \widetilde{L}_s = \int_0^\infty \mathbb{P}(\widetilde{\tau}^x \leq s) V_t(x) dx \\ \widetilde{L}_0 = 0 \text{ and } \widetilde{L} \text{ càdlàg}, \end{cases}$$

for $s > 0$, where V_t is the density for the problem at time t. Indeed, this is exactly the same formulation as (MV), except the initial condition is a sub-probability density. We then have that

$$L_u := L_t + \widetilde{L}_{u-t}$$

is an extension of L for $u \geq t$, which solves (MV). Therefore, restarting the system at a non-zero time is just a matter of normalising the initial condition.

Proof of Proposition 4.2.2. For a contradiction, suppose that L is a càdlàg process satisfying (MV) for which the inequality in Proposition 4.2.2 is violated. As noted in Remark 4.3.6, it suffices to assume that this occurs at $t = 0$, and this implies that $x_0 := \inf\{x > 0 : \nu_0(0, \alpha x) < x\} > 0$ and $L_h \to 0$ as $h \downarrow 0$. By definition of x_0 we have

$$\nu_0(0, x) \geq \alpha^{-1} x \quad \text{for} \quad x < \alpha x_0.$$

Thus, if F is a decreasing differentiable function, then

$$\begin{aligned}\int_0^{\alpha x_0} F(x)\nu_0(dx) &= F(\alpha x_0)\nu_0(0,\alpha x_0) - \int_0^{\alpha x_0} F'(x)\nu_0(0,x)dx \\ &\geq \alpha^{-1}\Big(\alpha x_0 F(\alpha x_0) - \int_0^{\alpha x_0} xF'(x)dx\Big) \\ &= \alpha^{-1}\int_0^{\alpha x_0} F(x)dx.\end{aligned}$$

Using this lower bound, it holds for any $h > 0$ that

$$\begin{aligned}L_h = \mathbb{P}(\tau \leq h) &\geq \int_0^{\infty} \mathbb{P}(B_h \leq \alpha L_h - x)\nu_0(dx) \\ &\geq \alpha^{-1}\int_0^{\alpha x_0} \mathbb{P}(B_h - \alpha L_h \leq -x)dx \\ &= \alpha^{-1}h^{1/2}\int_{-\alpha h^{-1/2}L_h}^{\alpha h^{-1/2}(x_0 - L_h)} \Phi(-y)dy.\end{aligned}$$

Here Φ is the Normal c.d.f. and we shall also need the Normal p.d.f., ϕ. Observe that the function

$$\Psi(x) := \phi(x) - x\Phi(-x), \qquad x \in \mathbb{R}$$

satisfies

$$\Psi'(x) = -\Phi(-x), \qquad \text{for every } x \in \mathbb{R}. \tag{4.9}$$

Now let $c_0 := \alpha x_0/2$. Since we are assuming $L_h \to 0$, we have $\alpha(x_0 - L_h) > c_0$ for all $h > 0$ sufficiently small, and hence

$$\alpha h^{-1/2}L_h \geq \int_{-\alpha h^{-1/2}L_h}^{c_0 h^{-1/2}} \Phi(-y)dy = \big[x\Phi(-x) - \phi(x)\big]_{-\alpha h^{-1/2}L_h}^{c_0 h^{-1/2}}.$$

Using $\Phi(\alpha h^{-1/2}L_h) = 1 - \Phi(-\alpha h^{-1/2}L_h)$ and $\phi(-\alpha h^{-1/2}L_h) = \phi(\alpha h^{-1/2}L_h)$, we can rearrange this inequality to find that

$$\begin{aligned}&\phi(\alpha h^{-1/2}L_h) - \alpha h^{-1/2}L_h\Phi(-\alpha h^{-1/2}L_h) \\ &\qquad \leq \phi(c_0 h^{-1/2}) - c_0 h^{-1/2}\Phi(-c_0 h^{-1/2}).\end{aligned}$$

In other words, for all $h > 0$ sufficiently small, we have

$$\Psi(\alpha h^{-1/2}L_h) \leq \Psi(c_0 h^{-1/2}),$$

but this then gives us our required contradiction: By (4.9), Ψ is strictly decreasing, and if $L_h \to 0$ as $h \downarrow 0$ then we can certainly find h sufficiently small so that $\alpha h^{-1/2}L_h < c_0 h^{-1/2}$. □

4.4 Contractivity of Γ and minimality of differentiable solutions: proof of Theorem 4.2.6

The main objective in this section is to prove Theorem 4.2.6, which states that differentiable solutions are minimal and that Γ is contractive on L^∞. The key technique is the following comparison result, which is obtained by coupling two outputs, $\Gamma[\ell]$ and $\Gamma[\bar{\ell}]$, of the map (4.5) to the same driving Brownian motion.

Proposition 4.4.1 (Comparison)**.** *Let ℓ and $\bar{\ell}$ be two increasing and initially zero càdlàg functions and Γ be defined as in (4.5) for model parameters α and ν_0. Suppose that ℓ is continuous on $[0, t_0)$. Then*

$$(\Gamma[\ell]_t - \Gamma[\bar{\ell}]_t)^+ \leq \int_0^t \Big\{ 2\Phi\Big(\alpha \frac{(\ell_s - \bar{\ell}_s)^+}{\sqrt{t-s}}\Big) - 1 \Big\} d\Gamma[\ell]_s, \quad \text{for every } t < t_0,$$

where $x^+ := \max\{x, 0\}$ and Φ is the Normal c.d.f.

Proof. It is no loss of generality to take a Brownian motion, B, and initial value, X_0, such that (4.5) holds with this B and this X_0 for both ℓ and $\bar{\ell}$, that is

$$X_t^\ell = X_0 + B_t - \alpha \ell_t \qquad \text{and} \qquad X_t^{\bar{\ell}} = X_0 + B_t - \alpha \bar{\ell}_t,$$

where we denote the respective hitting times of zero by τ and $\bar{\tau}$.

By conditioning on the value of τ we have

$$\begin{aligned} \Gamma[\ell]_t - \Gamma[\bar{\ell}]_t &\leq \mathbb{P}(\tau \leq t, \bar{\tau} > t) \\ &= \int_0^t \mathbb{P}(\bar{\tau} > t | \tau = s) d\Gamma[\ell]_s \\ &= \int_0^t \mathbb{P}(\inf_{u \in [0,t]} X_u^{\bar{\ell}} > 0 | \tau = s) d\Gamma[\ell]_s \\ &\leq \int_0^t \mathbb{P}(\inf_{u \in [s,t]} \{X_s^{\bar{\ell}} + B_u - B_s - \alpha(\bar{\ell}_u - \bar{\ell}_s)\} > 0 | \tau = s) d\Gamma[\ell]_s, \end{aligned} \tag{4.10}$$

where in the final line we have discarded the contribution on $u \in [0, s]$. Since $\bar{\ell}$ is increasing and $\alpha \geq 0$ we can further bound (4.10) by

$$\Gamma[\ell]_t - \Gamma[\bar{\ell}]_t \leq \int_0^t \mathbb{P}(\inf_{u \in [s,t]} \{X_s^{\bar{\ell}} + B_u - B_s\} > 0 | \tau = s) d\Gamma[\ell]_s.$$

On the event $\{\tau = s\}$ we have $X_s^\ell = 0$, so

$$X_s^{\bar{\ell}} = X_s^{\bar{\ell}} - X_s^\ell = \alpha(\ell_s - \bar{\ell}_s)$$

and thus we have

$$\begin{aligned}\Gamma[\ell]_t - \Gamma[\bar{\ell}]_t &\le \int_0^t \mathbb{P}(\inf_{u\in[s,t]}\{B_u - B_s\} > -\alpha(\ell_s - \bar{\ell}_s))d\Gamma[\ell]_s \\ &\le \int_0^t \mathbb{P}(\inf_{u\in[s,t]}\{B_u - B_s\} > -\alpha(\ell_s - \bar{\ell}_s)^+)d\Gamma[\ell]_s \\ &= \int_0^t \Big\{2\Phi\Big(\alpha\frac{(\ell_s - \bar{\ell}_s)^+}{\sqrt{t-s}}\Big) - 1\Big\}d\Gamma[\ell]_s.\end{aligned}$$

Since the right-hand side is positive, we can replace the left-hand side by its maximum with zero, so the proof is complete. □

When $\Gamma[\ell]$ is differentiable with power law control on its derivative near zero, then we are able to use this derivative to get a more direct bound.

Corollary 4.4.2. *Suppose that $\Gamma[\ell] \in \mathcal{S}(\gamma, A, t_0)$ for some $\gamma \in (0, \frac{1}{2})$, $A > 0$ and $t_0 > 0$ (recall Definition 4.2.5). Then there exists $c_0 > 0$, independent of t_0 and ℓ, such that*

$$(\Gamma[\ell]_t - \Gamma[\bar{\ell}]_t)^+ \le c_0 \int_0^t \frac{(\ell_s - \bar{\ell}_s)^+}{(t-s)^{\frac{1}{2}} s^\gamma} ds, \qquad \text{for every } t \le t_0,$$

where $\bar{\ell}$ is any càdlàg function that is increasing and initially zero.

Proof. We begin by noticing that $x \mapsto 2\Phi(\alpha x) - 1$ is bounded above on $x \ge 0$ by the linear function

$$x \mapsto c_1 x := \alpha\sqrt{2/\pi}\, x.$$

Therefore, by Proposition 4.4.1,

$$(\Gamma[\ell]_t - \Gamma[\bar{\ell}]_t)^+ \le c_1 \int_0^t \frac{(\ell_s - \bar{\ell}_s)^+}{(t-s)^{\frac{1}{2}}} \Gamma[\ell]'_s ds.$$

The result follows since $\Gamma[\ell]'_s \le As^{-\gamma}$, by definition of $\mathcal{S}(\gamma, A, t_0)$. □

It is now a straightforward task to prove Theorem 4.2.6.

Proof of Theorem 4.2.6. With A and γ fixed as in the statement of the result, take $\ell, \bar{\ell} \in \mathcal{S}(\gamma, A, t_0)$, where we will show how to take t_0 sufficiently small (and independent of ℓ and $\bar{\ell}$) so that we have the result. With $c_0 > 0$ as in Corollary 4.4.2 and the symmetry in $\ell \leftrightarrow \bar{\ell}$, we have

$$\begin{aligned}|\Gamma[\ell]_t - \Gamma[\bar{\ell}]_t| \le c_0 \int_0^t \frac{|\ell_s - \bar{\ell}_s|}{(t-s)^{\frac{1}{2}} s^\gamma} ds &\le c_0 \|\ell - \bar{\ell}\|_{L^\infty(0,t_0)} \int_0^t \frac{ds}{(t-s)^{\frac{1}{2}} s^\gamma} \\ &\le c_1 t_0^{\frac{1}{2}-\gamma} \|\ell - \bar{\ell}\|_{L^\infty(0,t_0)},\end{aligned}$$

for $t < t_0$, where $c_1 > 0$ is a constant independent of t_0, ℓ and $\bar{\ell}$. Consequently, it suffices to take t_0 such that $c_1 t_0^{1/2-\gamma} \leq \frac{1}{2}$.

For the second half of the result, take t_0 now fixed as in the statement. By the same estimate as above applied to Corollary 4.4.2, we have

$$(L_t - \bar{L}_t)^+ \leq c_1 t_1^{\frac{1}{2}-\gamma} \|(L - \bar{L})^+\|_{L^\infty(0,t_1)}, \qquad \text{for every } t \leq t_1,$$

where t_1 is any time with $t_1 < t_0$. Therefore, taking $t_1 > 0$ such that $c_1 t_1^{\frac{1}{2}-\gamma} < 1$ forces

$$\|(L - \bar{L})^+\|_{L^\infty(0,t_1)} = 0,$$

that is $\bar{L}_t \geq L_t$ for $t \leq t_1$.

By returning to Corollary 4.4.2, and repeating the same estimate again, we can deduce that

$$(L_t - \bar{L}_t)^+ \leq \|(L - \bar{L})^+\|_{L^\infty(t_1,t_2)} \int_{t_1}^{t} \frac{ds}{(t-s)^{\frac{1}{2}} s^\gamma} ds, \qquad \text{for every } t \leq t_2,$$

for any $t_2 < t_0$. Thus, by taking $t_2 > t_1$ sufficiently close to t_1, we can again force

$$\|(L - \bar{L})^+\|_{L^\infty(0,t_2)} = 0.$$

Continuing to repeat this argument, we obtain a sequence of times $t_1 < t_2 < \cdots < t_n < \cdots < t_0$. If $t_n \to t_\infty < t_0$, then the argument also applies at time t_∞ and so can be restarted. Furthermore, if this procedure ever terminates at a time strictly less than t_0, then the argument can be restarted (by left continuity) for a non-zero time, thus contradicting the termination. Hence we conclude that we have $\bar{L}_t \geq L_t$ up to the time t_0. □

4.5 Stability of Γ and the fixed-point argument: proof of Theorem 4.2.7

Here we show that we can choose parameters such that Γ maps $\mathcal{S}(\gamma, A, t_0)$ into itself. From this and Theorem 4.2.6, we are then able to conclude the existence of a solution in short time by the Banach fixed point theorem. We will also carefully track the contribution to the regularity of the solution near time zero due to the decay of the initial density near the Dirichlet boundary. To this end, we will see that the exponent $\frac{1-\beta}{2}$ in Theorem 4.2.7 comes directly from Lemma 4.5.2 below.

Lemma 4.5.1. *With ℓ and Γ above*

$$\Gamma[\ell]_t = 1 - \int_0^\infty \int_0^\infty G_t(x_0, x)dx\,\nu_0(dx_0) + 2\alpha \int_0^t \mathbb{E}[p_{t-s}(X_s)\mathbf{1}_{s<\tau}]d\ell_s,$$

where

$$p_t(x) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}, \qquad G_t(x, y) := p_t(x - y) - p_t(x + y).$$

Proof. Fix t and take u to be the solution to the backward heat equation on the positive half-line with Dirichlet boundary condition:

$$\partial_t u(s, x) + \tfrac{1}{2}\partial_{xx} u(s, x) = 0, \qquad u(\cdot, 0) = 0,\ u(t, x) = \mathbf{1}_{x>0}.$$

That is, let

$$u(s, x) := \int_0^\infty G_{t-s}(x, y)dy.$$

By applying Itô's formula to $s \mapsto u(s, X_{s\wedge\tau})$ we obtain

$$du(s, X_{s\wedge\tau}) = \mathbf{1}_{s<\tau}\{\partial_s + \tfrac{1}{2}\partial_{xx}\}u(s, X_s)ds + \mathbf{1}_{s<\tau}\partial_x u(s, X_s)d(B_s - \alpha\ell_s).$$

Since $\partial_s u + \frac{1}{2}\partial_{xx} u = 0$, integrating in time and taking expectation gives

$$\mathbb{E}[u(t, X_{t\wedge\tau})] = \mathbb{E}[u(0, X_0)] - \alpha \int_0^t \mathbb{E}[\mathbf{1}_{s<\tau}\partial_x u(s, X_s)]d\ell_s,$$

where the endpoints take the values

$$\mathbb{E}[u(t, X_{t\wedge\tau})] = \mathbb{P}(\tau > t) = 1 - \Gamma[\ell]_t$$
$$\mathbb{E}[u(0, X_0)] = \mathbb{E}\int_0^\infty G_t(X_0, x)dx = \int_0^\infty \int_0^\infty G_t(x_0, x)dx\,\nu_0(dx_0).$$

It remains to notice that

$$\begin{aligned}
\partial_x u(s, x) &= \int_0^\infty \frac{1}{\sqrt{2\pi(t-s)}} \partial_x \Big\{ e^{-\frac{(x-y)^2}{2(t-s)}} - e^{-\frac{(x+y)^2}{2(t-s)}} \Big\} dy \\
&= -\int_0^\infty \frac{1}{\sqrt{2\pi(t-s)}} \partial_y \Big\{ e^{-\frac{(x-y)^2}{2(t-s)}} + e^{-\frac{(x+y)^2}{2(t-s)}} \Big\} dy \\
&= \frac{2}{\sqrt{2\pi(t-s)}} e^{-\frac{x^2}{2(t-s)}} = 2p_{t-s}(x).
\end{aligned}$$

□

Our strategy is to apply the method of difference quotients [53, Sect. 5.8.2, Thm. 3] using Lemma 4.5.1. With $t, \delta > 0$ fixed, the starting point is to write

$$\begin{aligned}\Gamma[\ell]_{t+\delta} - \Gamma[\ell]_t &= -\int_0^\infty \int_0^\infty (G_{t+\delta}(x_0, x) - G_t(x_0, x)) dx \nu_0(dx_0) \\ &\quad + 2\alpha \int_t^{t+\delta} \mathbb{E}[p_{t+\delta-s}(X_s)\mathbf{1}_{s<\tau}] d\ell_s \\ &\quad + 2\alpha \int_0^t \mathbb{E}[(p_{t+\delta-s}(X_s) - p_{t-s}(X_s))\mathbf{1}_{s<\tau}] d\ell_s \\ &=: I_1 + I_2 + I_3. \end{aligned} \tag{4.11}$$

In order to estimate these three integrals we must make use of the assumption (4.7) on the behaviour of ν_0 near the origin. Recall that we have constants $C, D, x_\star, \beta > 0$ such that

$$V_0(x) \leq Cx^\beta \mathbf{1}_{x \leq x_\star} + D\mathbf{1}_{x > x_\star}.$$

Lemma 4.5.2 (I_1)**.** *There exists $t_0 = t_0(C, D, x_0)$ and $K = K(C, x_0)$ such that*

$$0 \leq I_1 \leq K\delta t^{-\frac{1-\beta}{2}},$$

for all $t < t_0$ and $\delta > 0$.

Remark 4.5.3 (Constants). In the proof below we will allow the constants to increase as necessary, but we will pay close attention to ensure that K does not depend on D (whereas t_0 does).

Proof of Lemma 4.5.2. By the fundamental theorem of calculus, and properties of G, we have

$$\begin{aligned} I_1 &= -\delta \int_0^\infty \int_0^\infty \int_0^1 \partial_t G_{t+\theta\delta}(z, x) d\theta dx \nu_0(dz) \\ &= -\frac{\delta}{2} \int_0^\infty \int_0^\infty \int_0^1 \partial_{xx} G_{t+\theta\delta}(z, x) d\theta dx \nu_0(dz) \\ &= \frac{\delta}{2} \int_0^1 \int_0^\infty \partial_x G_{t+\theta\delta}(0, z) \nu_0(dz) d\theta \\ &= \frac{\delta}{\sqrt{2\pi}} \int_0^1 \int_0^\infty \frac{z}{(t+\theta\delta)^{3/2}} e^{-\frac{z^2}{2(t+\theta\delta)}} V_0(z) dz d\theta \\ &= \frac{\delta}{\sqrt{2\pi}} \int_0^1 \int_0^\infty \frac{z}{(t+\theta\delta)^{1/2}} e^{-\frac{z^2}{2}} V_0((t+\theta\delta)^{\frac{1}{2}} z) dz d\theta. \end{aligned}$$

Splitting the z-integral and using the bound on V_0 gives

$$\begin{aligned} I_1 &\leq \frac{\delta}{\sqrt{2\pi}} \int_0^1 \Big(C \int_0^{x_\star (t+\theta\delta)^{-\frac{1}{2}}} (t+\theta\delta)^{-\frac{1-\beta}{2}} z^{1+\beta} e^{-\frac{z^2}{2}} dz \\ &\qquad\qquad + D \int_{x_\star (t+\theta\delta)^{-\frac{1}{2}}}^{\infty} (t+\theta\delta)^{-\frac{1}{2}} z e^{-\frac{z^2}{2}} dz \Big) d\theta \\ &\leq \frac{\delta}{\sqrt{2\pi}} \Big(K(t+\delta)^{-\frac{1-\beta}{2}} + D(t+\delta)^{-\frac{1}{2}} e^{-\frac{x_\star^2}{2(t+\delta)}} \Big). \end{aligned}$$

Note that the final term is a bounded function of D and $x_\star$, so by taking t_0 sufficiently small we can conclude the result. □

In order to control I_2 from equation (4.11) we need to make some assumptions on the regularity of ℓ. Provided ℓ is in H^1, we have that $t \mapsto B_t - \alpha \ell_t$ is absolutely continuous with respect to Brownian motion, and so we can proceed as below with a change of measure.

Lemma 4.5.4 (A Girsanov argument)**.** *Let $F : \mathbb{R} \to \mathbb{R}$ be measurable and bounded and assume that $\ell \in H^1(0, t_0)$ for some $t_0 > 0$. Then for all $q > 1$ and $t \in [0, t_0]$*

$$\mathbb{E}[F(X_t)\mathbf{1}_{t<\tau}] \leq \exp\Big\{ \frac{\alpha^2 \|\ell'\|^2_{L^2(0,t)}}{2(q-1)^2} \Big\} \mathbb{E}[F(W_t)^q \mathbf{1}_{t<\tau^W}]^{\frac{1}{q}}$$

where W is a standard Brownian motion, started from ν_0, and τ^W is its first hitting time of the origin, i.e. $\tau^W = \inf\{t \geq 0 : W_t = 0\}$.

Proof. Let Z_t denote the Radon–Nikodym derivative

$$Z_t = \frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = \exp\Big\{ \alpha \int_0^t \ell'_s dB_s - \frac{\alpha^2}{2} \int_0^t (\ell'_s)^2 ds \Big\}$$

for $t \in [0, t_0]$. Under $\mathbb{Q}$, $t \mapsto B_t - \alpha \ell_t$ is a standard Brownian motion by Girsanov's Theorem. Therefore, applying Hölder's inequality gives

$$\begin{aligned} \mathbb{E}[F(X_t)\mathbf{1}_{t<\tau}] &= \mathbb{E}^{\mathbb{Q}}[Z_t^{-1} F(X_t)\mathbf{1}_{t<\tau}] \\ &\leq \mathbb{E}^{\mathbb{Q}}[Z_t^{-r}]^{\frac{1}{r}} \mathbb{E}^{\mathbb{Q}}[F(X_t)^q \mathbf{1}_{t<\tau}]^{\frac{1}{q}} \\ &\leq \mathbb{E}[Z_t^{1-r}]^{\frac{1}{r}} \mathbb{E}[F(W_t)^q \mathbf{1}_{t<\tau^W}]^{\frac{1}{q}}, \end{aligned}$$

where $q^{-1} + r^{-1} = 1$. Finally, it is a simple calculation to see that

$$\mathbb{E}[Z_t^{1-r}]^{\frac{1}{r}} = \exp\Big\{ \frac{\alpha^2}{2}(r-1)^2 \|\ell'\|^2_{L^2(0,t)} \Big\},$$

and this completes the proof. □

Lemma 4.5.5. *Suppose ℓ satisfies the conditions of Lemma 4.5.4 and fix $q > 1$. Then there is a constant $c_0 > 0$ depending only on V_0 and q such that, for all $u, s > 0$ and $\zeta, \eta \geq 0$,*

$$\begin{aligned}&\mathbb{E}[X_s^\zeta p_u(X_s)\mathbf{1}_{s<\tau}]\\ &\qquad \leq c_0 \exp\Big\{\frac{\alpha^2\|\ell'\|^2_{L^2(0,s)}}{2(q-1)^2}\Big\}(u^{\frac{\beta}{2}+\eta+\frac{q\zeta}{2}-\frac{q-1}{2}}s^{-\eta}+u^{\eta+\frac{q\zeta}{2}-\frac{q-1}{2}}s^{\frac{\beta}{2}-\eta})^{\frac{1}{q}}.\end{aligned}$$

Proof. We first prove the bound for Brownian motion, and then we appeal to the previous lemma for the full result. From the assumption (4.7) we can certainly find a constant $c_0 > 0$ such that $V_0(x) \leq c_0 x^\beta$ for all $x > 0$, and so

$$\mathbb{E}[W_s^{q\zeta}p_u(W_s)^q\mathbf{1}_{s<\tau^W}] \leq \frac{c_0}{u^{\frac{q}{2}}}\int_0^\infty\int_0^\infty x^{q\zeta}e^{-\frac{qx^2}{2u}}G_s(x_0,x)x_0^\beta dx_0dx,$$

where from here on we absorb the relevant numerical constants into c_0. Using the estimate

$$G_s(x_0,x) \leq \frac{2}{(2\pi s)^{1/2}}\Big(\frac{xx_0}{s}\wedge 1\Big)e^{-\frac{(x-x_0)^2}{2s}}$$

together with the substitution $x_0 \mapsto x_0 - x$ gives

$$\begin{aligned}&\mathbb{E}[W_s^{q\zeta}p_u(W_s)^q\mathbf{1}_{s<\tau^W}] \qquad (4.12)\\ &\leq \frac{c_0}{u^{\frac{q}{2}}s^{\frac{1}{2}}}\int_0^\infty x^{q\zeta}e^{-\frac{qx^2}{2u}}\int_{-x}^\infty (x+x_0)^\beta\Big(\frac{x(x+x_0)}{s}\wedge 1\Big)e^{-\frac{x_0^2}{2s}}dx_0dx.\end{aligned}$$

Now notice that for generic $a, b \geq 0$ we have the inequality

$$(a+b)\wedge 1 \leq (a\wedge 1)+(b\wedge 1).$$

Hence it holds for $x \geq 0$, $x_0 \geq -x$ and $\eta \in [0,1]$ that

$$\frac{x(x+x_0)}{s}\wedge 1 \leq \Big(\frac{x^2}{s}\wedge 1\Big)+\Big(\frac{|x_0x|}{s}\wedge 1\Big) \leq \Big(\frac{x^2}{s}\Big)^\eta+\Big(\frac{|x_0x|}{s}\Big)^{2\eta}.$$

Combining this inequality with $(x+x_0)^\beta \leq c_1(|x|^\beta+|x_0|^\beta)$, for $c_1 > 0$ a constant depending only on β, allows us to bound the inner integral in (4.12) by

$$\begin{aligned}&\int_{-x}^\infty (|x|^\beta+|x_0|^\beta)\Big(\Big(\frac{x^2}{s}\Big)^\eta+\Big(\frac{|x_0x|}{s}\Big)^{2\eta}\Big)e^{-\frac{x_0^2}{2s}}dx_0\\ &\qquad \leq c_0 s^{\frac{1}{2}}\int_{-\infty}^\infty (|x|^\beta+s^{\beta/2}|x_0|^\beta)\Big(\Big(\frac{x^2}{s}\Big)^\eta+\Big(\frac{|x_0x|}{s^{1/2}}\Big)^{2\eta}\Big)e^{-\frac{x_0^2}{2}}dx_0\\ &\qquad \leq c_0(|x|^{\beta+2\eta}s^{\frac{1}{2}-\eta}+|x|^{2\eta}s^{\frac{1}{2}+\frac{\beta}{2}-\eta}).\end{aligned}$$

Setting this into (4.12) and making the change of variables $x \mapsto u^{1/2}x$ we obtain

$$\mathbb{E}[W_s^{q\zeta}p_u(W_s)^q\mathbf{1}_{s<\tau^W}] \leq c_0(u^{\frac{\beta}{2}+\eta+\frac{q\zeta}{2}-\frac{q-1}{2}}s^{-\eta}+u^{\eta+\frac{q\zeta}{2}-\frac{q-1}{2}}s^{\frac{\beta}{2}-\eta}).$$

Finally, the result for X in place of W follows by invoking Lemma 4.5.4. □

Lemma 4.5.6 (I_2)**.** *If $\ell \in \mathcal{S}(\gamma, A, t_0)$, then there exist constants $c_1 > 0$, $\theta_1 > 0$ and $\theta_2 \in [0, 1)$ (depending on ℓ) such that*

$$0 \le I_2 \le c_1 \delta^{1+\theta_1} t^{-\frac{1}{2}\theta_2}, \qquad \text{for every } t \in (0, t_0).$$

Remark 4.5.7. Formally, the integrand in I_2 converges to a Dirac mass at zero, evaluated at X_s and multiplied by $\mathbf{1}_{s<\tau}$. We expect such an expression to vanish since $s < \tau$ implies $X_s > 0$. While this argument is not rigorous, it suggests why $I_2 = o(\delta)$ and hence why this term will not contribute to the bound on $\Gamma[\ell]'$ in the proof of Theorem 4.2.7.

Proof of Lemma 4.5.6. We begin by applying the bound $\ell'_s \le As^{-\gamma}$ to see that

$$\begin{aligned} I_2 &= 2\alpha \int_t^{t+\delta} \mathbb{E}[p_{t+\delta-s}(X_s)\mathbf{1}_{s<\tau}] d\ell_s = 2\alpha \int_t^{t+\delta} \mathbb{E}[p_{t+\delta-s}(X_s)\mathbf{1}_{s<\tau}] \ell'_s ds \\ &\le 2\alpha A t^{-\gamma} \int_t^{t+\delta} \mathbb{E}[p_{t+\delta-s}(X_s)\mathbf{1}_{s<\tau}] ds. \end{aligned}$$

Taking $u = t + \delta - s$ and $\zeta = 0$ in Lemma 4.5.5, and bounding with the worst-case exponents, gives

$$I_2 \le c_2 t^{-\gamma-\frac{\eta}{q}} \int_t^{t+\delta} (t+\delta-s)^{\frac{\eta}{q}-\frac{1}{2}(1-\frac{1}{q})} ds = c_3 \delta^{1+\frac{\eta}{q}-\frac{1}{2}(1-\frac{1}{q})} t^{-\gamma-\frac{\eta}{q}},$$

where c_2 and c_3 are constants depending on ℓ. The result is now complete by choosing η and q so that $\gamma + \frac{\eta}{q} < \frac{1}{2}$ (possible since $\gamma < \frac{1}{2}$) and $\frac{\eta}{q} - \frac{1}{2}(1-\frac{1}{q}) > 0$ (take q sufficiently close to 1 and maintain constant ratio η/q). □

The third term, I_3, is not so simple to control.

Lemma 4.5.8 (I_3)**.** *If $\gamma < \frac{1}{2}$ and $\ell \in \mathcal{S}(\gamma, A, t_0)$, then there exist constants $c_0 > 0$, $c_1 > 0$ and $\theta_3 > 0$ (depending only on V_0 and γ) such that*

$$|I_3| \le c_0 \alpha A \delta \exp\{c_1 \alpha^2 A^2 t_0^{1-2\gamma}\} t^{-\frac{1-\beta}{2}+\theta_3}, \quad \text{for every } t \in (0, t_0).$$

Proof. We begin by applying the fundamental theorem of calculus to see that

$$\begin{aligned} p_{t+\delta-s}(X_s) - p_{t-s}(X_s) &= \int_0^\delta \partial_u p_{t-s+u}(X_s) du \\ &= \frac{1}{2} \int_0^\delta (X_s^2 (t-s+u)^{-2} - (t-s+u)^{-1}) p_{t-s+u}(X_s) du \end{aligned}$$

and therefore

$$|p_{t+\delta-s}(X_s) - p_{t-s}(X_s)| \le \frac{1}{2} \int_0^\delta (X_s^2 (t-s)^{-2} + (t-s)^{-1}) p_{t-s+u}(X_s) du.$$

Taking an expectation and using Proposition 4.5.5 gives

$$\begin{aligned}&\mathbb{E}[|p_{t+\delta-s}(X_s) - p_{t-s}(X_s)|\mathbf{1}_{s<\tau}] \\ &\leq c_0\delta \exp\left\{\frac{\alpha^2\|\ell'\|^2_{L^2(0,t_0)}}{2(q-1)^2}\right\}\left((t-s)^{\frac{\beta}{2}+\eta-\frac{q-1}{2}-1}s^{-\eta} + (t-s)^{\eta-\frac{q-1}{2}-1}s^{\frac{\beta}{2}-\eta}\right)^{\frac{1}{q}},\end{aligned} \tag{4.13}$$

where c_0 depends only on V_0 and the choice of q and η.

By applying the bound (4.13) to the expression for I_3, we get

$$\begin{aligned}|I_3| \leq 2Ac_0\alpha\delta \exp\left\{\frac{\alpha^2\|\ell'\|^2_{L^2(0,t_0)}}{2(q-1)^2}\right\}\Big(&J(\tfrac{1}{q}(\tfrac{\beta}{2}+\eta-\tfrac{q-1}{2}-1), -\tfrac{\eta}{q}-\gamma) \\ &+ J(\tfrac{1}{q}(\eta-\tfrac{q-1}{2}-1), \tfrac{\beta}{2q}-\tfrac{\eta}{q}-\gamma)\Big),\end{aligned}$$

where J is defined to be

$$J(a,b) := \int_0^t (t-s)^a s^b ds = Ct^{1+a+b}, \tag{4.14}$$

for a constant $C = C(a,b) > 0$, provided $a > -1$ and $b > -1$. To keep the above exponents bigger than -1, we need to select η and q so that $\frac{\eta}{q} < 1-\gamma$, whereby we obtain

$$\begin{aligned}|I_3| &\leq Ac_0\alpha\delta \exp\left\{\frac{\alpha^2\|\ell'\|^2_{L^2(0,t_0)}}{2(q-1)^2}\right\}t^{\frac{\beta}{2q}-\frac{1}{2}(\frac{1}{q}-1)-\gamma} \\ &= Ac_0\alpha\delta \exp\left\{\frac{\alpha^2\|\ell'\|^2_{L^2(0,t_0)}}{2(q-1)^2}\right\}t^{-\frac{1-\beta}{2}+\frac{1-\beta}{2}(1-\frac{1}{q})+\frac{1}{2}-\gamma},\end{aligned}$$

where we have absorbed numerical constants into c_0. Since $\gamma < \frac{1}{2}$, we can take q sufficiently close to 1 so that we have the required exponent. The proof is then complete by noting that

$$\|\ell'\|^2_{L^2(0,t_0)} \leq \int_0^{t_0} A^2 s^{-2\gamma} ds = \frac{A^2}{1-2\gamma} t_0^{1-2\gamma}.$$

□

Proposition 4.5.9 (Stability of Γ)**.** *There exists a constant $K > 0$ depending only on C and $x_\star$ such that for every $\varepsilon > 0$ there exists $t_0 > 0$ for which*

$$\Gamma : \mathcal{S}(\tfrac{1-\beta}{2}, K+\varepsilon, t_0) \to \mathcal{S}(\tfrac{1-\beta}{2}, K+\varepsilon, t_0),$$

where t_0 also depends on the model parameters.

Proof. Take K and t_0 as in the conclusion of Lemma 4.5.2. We will decrease the value of t_0 throughout the proof, but this is the K in the statement of the result.

First we check that $\Gamma : \mathcal{S}(\frac{1-\beta}{2}, K+\varepsilon, t_0) \to H^1(0,t_0)$. Combine Lemmas 4.5.2, 4.5.6 and 4.5.8 to get

$$\begin{aligned}\left|\frac{\Gamma[\ell]_{t+\delta} - \Gamma[\ell]_t}{\delta}\right| &\leq Kt^{-\frac{1-\beta}{2}} + c_2\delta^{\theta_1}t^{-\frac{1}{2}\theta_2} \\ &\quad + c_0\alpha(K+\varepsilon)\exp\{c_1\alpha^2(K+\varepsilon)^2t_0^\beta\}t^{-\frac{1-\beta}{2}+\theta_3},\end{aligned} \tag{4.15}$$

and so we have

$$\limsup_{\delta\to 0}\int_0^{t_0}\left|\frac{\Gamma[\ell]_{t+\delta} - \Gamma[\ell]_t}{\delta}\right|^2 dt < \infty.$$

By the method of difference quotients [53, Sect. 5.8.2, Thm. 3] we conclude that $\Gamma[\ell]$ is in $H^1(0,t_0)$, as required.

Taking a pointwise limit in (4.15) gives

$$\begin{aligned}|\Gamma[\ell]'_t| &\leq Kt^{-\frac{1-\beta}{2}} + c_0\alpha(K+\varepsilon)\exp\{c_1\alpha^2(K+\varepsilon)^2t_0^\beta\}t^{-\frac{1-\beta}{2}+\theta_3} \\ &\leq Kt^{-\frac{1-\beta}{2}} + c_0\alpha(K+\varepsilon)\exp\{c_1\alpha^2(K+\varepsilon)^2t_0^\beta\}t_0^{\theta_3}t^{-\frac{1-\beta}{2}}.\end{aligned}$$

With $\varepsilon > 0$ fixed, we can now take $t_0 > 0$ sufficiently small (since the constants in the bound in Lemma 4.5.8 do not depend on the value of t_0) so that

$$c_0\alpha(K+\varepsilon)\exp\{c_1\alpha^2(K+\varepsilon)^2t_0^\beta\}t_0^{\theta_3} < \varepsilon,$$

which completes the proof. □

With Proposition 4.5.9 now in place, it remains to show that we can find a fixed point and deduce that it lives in one of the sets $\mathcal{S}(\frac{1-\beta}{2}, K+\varepsilon, t_0)$.

Proof of Theorem 4.2.7. Take K, ε and t_0 as in the conclusion of Proposition 4.5.9. Take t_0 sufficiently small so that the first half of Theorem 4.2.6 holds. Define the sequence

$$\ell^{(0)} := 0, \qquad \ell^{(n)} := \Gamma[\ell^{(n-1)}], \qquad \text{for every } n \geq 1.$$

By the Banach fixed point theorem, we know that there exists a limit point $\ell^{(n)} \to L$ in $L^\infty(0,t_0)$, as $n\to\infty$, and that $\Gamma[L] = L$. So L solves (MV) on $[0,t_0)$.

To see $L \in \mathcal{S}(\frac{1-\beta}{2}, K+\varepsilon, t_0)$, notice that, since $\Gamma[\ell]$ is bounded by 1, dominated convergence gives

$$\int_0^{t_0-\delta}\left|\frac{L_{t+\delta}-L_t}{\delta}\right|^2 dt \leq \limsup_{n\to\infty}\int_0^{t_0-\delta}\left|\frac{\ell^{(n)}_{t+\delta}-\ell^{(n)}_t}{\delta}\right|^2 dt, \quad \text{for every } \delta > 0. \tag{4.16}$$

Proposition 4.5.9 ensures $\ell^{(n)} \in \mathcal{S}(\frac{1-\beta}{2}, K+\varepsilon, t_0)$, so we have the estimate

$$|\ell^{(n)}_{t+\delta} - \ell^{(n)}_t| = \Big| \int_t^{t+\delta} (\ell^{(n)}_s)' ds \Big| \leq (K+\varepsilon) t^{-\frac{1-\beta}{2}} \delta.$$

Using this in (4.16) yields

$$\int_0^{t_0-\delta} \Big| \frac{L_{t+\delta} - L_t}{\delta} \Big|^2 dt \leq (K+\varepsilon) \int_0^{t_0} t^{-\frac{1-\beta}{2}} dt < \infty,$$

and hence the method of difference quotients [53, Sect. 5.8.2, Thm. 3] gives that L is in $H^1(0,t_0)$. Moreover, it holds pointwise almost everywhere in $(0,t_0)$ that

$$0 \leq \frac{L_{t+\delta} - L_t}{\delta} = \lim_{n\to\infty} \frac{\ell^{(n)}_{t+\delta} - \ell^{(n)}_t}{\delta} \leq (K+\varepsilon) t^{-\frac{1-\beta}{2}},$$

so sending $\delta \to 0$ gives that $L \in \mathcal{S}(\frac{1-\beta}{2}, K+\varepsilon, t_0)$, as required.

The uniqueness statement follows immediately by applying the second half of Theorem 4.2.6. □

4.6 Bootstrapping and full uniqueness up to explosion time: proof of Theorem 4.2.8

Here we show how the solutions from the fixed point argument in the previous section can be extended up to the first time their H^1 norm explodes. (Trivially, this time occurs before or at the first jump time.) The key to this bootstrapping method is to notice that if $\ell \in H^1(0,t_0)$, for some t_0, then necessarily $V_{t_0-}(x) = O(x^\beta)$ (Lemma 4.6.2). Not only does this show that t_0 cannot be a jump time, it also allows us to apply the fixed point argument once more, thus extending the solution to $H^1(0,t_1)$ for some $t_1 > t_0$ (Corollary 4.6.3). The first half of Theorem 4.2.8 then follows by iterating this argument (Corollary 4.6.4).

After the above, we proceed to prove the second half of Theorem 4.2.8. The idea is to consider modified initial conditions for which a fixed portion of the initial density is erased and added to the initial value of the loss process (Definition 4.6.5). An argument that shows solutions cannot overlap (Lemma 4.6.7) then allows us to trap any general càdlàg solution of (MV) between the modified solutions and the minimal differentiable solution. The proof concludes by showing that the size of this trapping envelope shrinks to zero as the size of the initial modification is taken to zero (Lemma 4.6.8), thus forcing the minimal solution and the general càdlàg solution to be equal (see Figure 4.6.2).

4.6.1 A duality lemma

Note that the increments $X_s - X_0$ are independent of the random start point X_0 for all $s \geq 0$. On a fixed time interval $[0,t]$, we define the *dual* process $\widehat{X}$ by

$$\widehat{X}^x_s := x + X_{t-s} - X_t, \quad \text{for} \quad s \in [0,t].$$

This process is dual to X on $[0,t]$ in the sense that

$$\int_{\mathbb{R}} \mathbb{E}_{X_0=x}[\phi(X_t)]\psi(x)dx = \int_{\mathbb{R}} \phi(x)\mathbb{E}[\psi(\widehat{X}^x_t)]dx,$$

for any two non-negative measurable functions ϕ and ψ (the proof follows as in the next lemma). Similarly, there is also a duality relation for the processes X and $\widehat{X}$ absorbed when they first hit zero, which is a version of *Hunt's swithing identity* as presented for Levy processes in [9, Thm. II.1.5].

Lemma 4.6.1 (Hunt's switching identity)**.** *For any two non-negative measurable functions $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ and $\psi : \mathbb{R}_+ \to \mathbb{R}_+$, we have*

$$\int_0^\infty \mathbb{E}_{X_0=x}[\phi(X_t)\mathbf{1}_{t<\tau}]\psi(x)dx = \int_0^\infty \phi(x)\mathbb{E}[\psi(\widehat{X}^x_t)\mathbf{1}_{t<\widehat{\tau}_x}]dx,$$

where

$$\tau := \inf\{s > 0 : X_s \leq 0\} \quad \textit{and} \quad \widehat{\tau}_x := \inf\{s > 0 : \widehat{X}^x_s \leq 0\}.$$

Proof. First of all, we can observe that

$$\begin{aligned}\{t < \tau\} &= \big\{X_0 + \inf_{s\in[0,t]}(X_s - X_0) < 0\big\} = \big\{X_0 + \inf_{s\in[0,t]}(X_{t-s} - X_0) < 0\big\} \\ &= \big\{X_0 + (X_t - X_0) + \inf_{s\in[0,t]}(X_{t-s} - X_t) < 0\big\}\end{aligned}$$

and hence, conditional on $X_0 = x$, we have

$$\{t < \tau\} = \big\{t < \widehat{\tau}_{(x+X_t-X_0)}\big\}.$$

By Tonelli's theorem and the change of variables $x \mapsto x + X_t - X_0$, we can thus conclude that

$$\begin{aligned}\int_{\mathbb{R}} \mathbb{E}_{X_0=x}[\phi(X_t)\mathbf{1}_{t<\tau}]\psi(x)dx &= \int_{\mathbb{R}} \mathbb{E}[\phi(x + X_t - X_0)\mathbf{1}_{t<\widehat{\tau}_{(x+X_t-X_0)}}]\psi(x)dx \\ &= \mathbb{E}\Big[\int_{\mathbb{R}} \phi(x)\mathbf{1}_{t<\widehat{\tau}_x}\psi(x + X_0 - X_t)dx\Big] \\ &= \int_{\mathbb{R}} \phi(x)\mathbb{E}[\psi(\widehat{X}^x_t)\mathbf{1}_{t<\widehat{\tau}_x}]dx.\end{aligned}$$

This proves the claim. □

4.6.2 Bootstrapping of solutions

Lemma 4.6.2 (Recovery of initial exponents)**.** *Suppose that $L \in H^1(0,t_0)$ solves (MV) for some $t_0 > 0$, with the usual assumption (4.7) on V_0. Then there exists further constants $C', D', x'_\star > 0$ such that*

$$V_{t_0-}(x) \leq C' x^\beta \mathbf{1}_{x<x'_\star} + D' \mathbf{1}_{x \geq x'_\star}, \qquad \text{for every } x > 0.$$

Proof. Note that there exists a limit $L_{t_0-} = \lim_{s \uparrow t_0} L_s$, since L must be increasing. Therefore, there exists a left limit density V_{t_0-} (recall Proposition 4.3.1). Now fix $t < t_0$. Our strategy is to apply the above mentioned version of Hunt's switching identity, so define the dual process

$$d\widehat{X}_s = dB_s - \alpha d\widehat{L}_s, \qquad \widehat{L}_s = L_{t-s}, \qquad \widehat{\tau} = \inf\{s > 0 : \widehat{X}_s \leq 0\},$$

for $s \in [0,t]$. Then, by Lemma 4.6.1, we have

$$\int_{x \in \mathbb{R}} \phi(x) \mathbb{E}_{X_0=x}[\psi(X_t)\mathbf{1}_{t<\tau}]dx = \int_{x \in \mathbb{R}} \psi(x) \mathbb{E}_{\widehat{X}_0=x}[\phi(\widehat{X}_t)\mathbf{1}_{t<\widehat{\tau}}]dx,$$

for all non-negative measurable functions ϕ and ψ. (Note that, since $L \in H^1(0,t_0)$ and $\Delta L_{t_0} = 0$ by the hypotheses, we have continuity of L). By taking $\phi = V_0$ and ψ an arbitrary non-negative measurable function, we can conclude that

$$V_t(x) = \mathbb{E}_{X_0=x}[V_0(\widehat{X}_t)\mathbf{1}_{t<\widehat{\tau}}] \tag{4.17}$$

for almost every $x > 0$.

Since $L \in H^1(0,t_0)$, $\widehat{X}$ is just a Brownian motion with a H^1 drift. Therefore we can apply Lemma 4.5.4 with $\widehat{X}$ in place of X and $F = V_0$ to get

$$\mathbb{E}_{X_0=x}[V_0(\widehat{X}_t)\mathbf{1}_{t<\widehat{\tau}}] \leq C_q \mathbb{E}_{W_0=x}[(V_0(W_t)\mathbf{1}_{t<\tau^W})^q]^{\frac{1}{q}}$$

for any $q > 1$, where W is a Brownian motion and $C_q < \infty$ depends on L, t_0 and q and is independent of t. (We have absorbed the exponential factor into C_q.) By the assumption that $V_0(x) = O(x^\beta)$, and increasing C_q as needed, we deduce that

$$\mathbb{E}_{X_0=x}[V_0(\widehat{X}_t)\mathbf{1}_{t<\widehat{\tau}}] \leq C_q \mathbb{E}_{W_0=x}[(W_t\mathbf{1}_{t<\tau^W})^{q\beta}]^{\frac{1}{q}}.$$

Provided we take $q > 1$ such that $q\beta < 1$, we can apply Jensen's inequality to get

$$\mathbb{E}_{X_0=x}[V_0(\widehat{X}_t)\mathbf{1}_{t<\widehat{\tau}}] \leq C_q \mathbb{E}_{W_0=x}[W_t\mathbf{1}_{t<\tau^W}]^\beta = C_q x^\beta.$$

Putting this into (4.17) gives the required bound near zero at time t. Proposition 4.3.1 gives $V_t(x) \leq \|V_0\|_\infty$, so we have the required boundedness away from the origin too. Since the constants obtained above are independent of t, we can find C', D' and $x'_\star$ such that

$$V_t(x) \leq C' x^\beta \mathbf{1}_{x<x'_\star} + D' \mathbf{1}_{x \geq x'_\star}$$

for all $x > 0$ and $t < t_0$. By sending $t \uparrow t_0$, we have the result. □

The implication of Lemma 4.6.2 is that at the end of the fixed-point argument from Section 4.5 we can restart the argument with new initial conditions that have the same power law decay. As a result we can push our construction of solutions by a further non-zero amount of time. Notice, however, that we lose control of the exact constant that we had in Theorem 4.2.7 and hence the proceeding results are qualitative, not quantitative.

Corollary 4.6.3 (Extending solutions)**.** *Suppose we have a solution $L \in H^1(0, t_0)$ to (MV). Then there exists $t_1 > t_0$ such that L can be extended to a solution of (MV) on $(0, t_1)$ for which $L \in H^1(0, t_1)$. Furthermore, if $L \in \mathcal{S}(\frac{1-\beta}{2}, K_0, t_0)$ for some $K_0 > 0$, then we can find $K_1 > 0$ such that the extension satisfies $L \in \mathcal{S}(\frac{1-\beta}{2}, K_1, t_1)$.*

Proof. Since we have $L \in H^1(0, t_0)$, Lemma 4.6.2 implies that V_{t_0-} satisfies the condition (4.7), for some constants C, D and $x_\star$ (that are possibly different to those for V_0). Therefore, by Theorem 4.2.7 there exists $K_1 > 0$ and $u_0 > 0$ such that we can find $F \in \mathcal{S}(\frac{1-\beta}{2}, K_1, u_0)$ solving

$$\begin{cases} X_{t_0+u} = X_{t_0} + B_u - \alpha F_u \\ \tau^{(t_0)} = \inf\{u \geq 0 : X_{t_0+u} \leq 0\} \\ F_u = \mathbb{P}(\tau^{(t_0)} \leq u), \end{cases} \tag{4.18}$$

for $u < u_0$ (recall Remark 4.3.6). It follows that, if we define $t_1 := t_0 + u_0$ and

$$\widetilde{L}_t := \begin{cases} L_t, & \text{if } 0 < t < t_0 \\ F_{t-t_0}, & \text{if } t_0 \leq t < t_1, \end{cases}$$

then $\widetilde{L}$ extends L and solves (MV) for all $t \in (0, t_1)$ with the required derivative control. □

Naturally, we can iterate Corollary 4.6.3 to prove the first half of Theorem 4.2.8.

Corollary 4.6.4 (Bootstrap to explosion time)**.** *There exists a solution, L, to (MV) up to time*

$$t_{\text{explode}} := \sup\{t > 0 : \|L\|_{H^1(0,t)} < \infty\} \in (0, \infty]$$

such that, for every $t_0 < t_{\text{explode}}$, we have $L \in \mathcal{S}(\frac{1-\beta}{2}, K, t_0)$ for some $K > 0$.

Proof. By repeating Corollary 4.6.3, we can find an infinite sequence of times $t_0 < \cdots < t_n < \cdots$ over which we can successively extend L. If $t_\infty := \lim_{n\to\infty} t_n$ is such that $\|L\|_{H^1(0,t_\infty)} = \infty$, then we are done. Otherwise, by the left continuity of L, we can restart the argument from t_∞ by applying Corollary 4.6.3. This procedure cannot terminate at a time for which $\|L\|_{H^1(0,t)} < \infty$, or else it can be restarted, hence we conclude the result. □

4.6.3 Monotonicity and trapping

Our main technical construction will be a solution to (MV) for which we delete a portion of the initial condition near the boundary and add that mass to the loss at time zero, before finally shifting the density towards zero accordingly.

Definition 4.6.5 (ε-deleted solutions)**.** Suppose X_0 has a density V_0 and let $\varepsilon > 0$. Define the McKean–Vlasov problem

$$\begin{cases} X_t^\varepsilon = X_0 \mathbf{1}_{X_0 \geq \varepsilon} - \frac{1}{4}\varepsilon + B_t - \alpha L_t^\varepsilon \\ \tau^\varepsilon = \inf\{t \geq 0 : X_t^\varepsilon \leq 0\} \\ L_t^\varepsilon = \nu_0(0,\varepsilon) + \int_\varepsilon^\infty \mathbb{P}_{X_0 = x}(\tau^\varepsilon \leq t)\nu_0(dx), \end{cases} \tag{4.19}$$

where solutions, L^ε, are taken to be càdlàg.

It is not immediately clear that we can solve the above problem for $\varepsilon > 0$, however, careful inspection of the initial loss reveals that we can.

Lemma 4.6.6 (ε-deleted solutions exist)**.** *Assume ν_0 has a density satisfying (4.7). Then there exist $K > 0$, $t_0 > 0$ and $\varepsilon_0 > 0$ such that, for every $\varepsilon \in (0, \varepsilon_0)$, there is a solution $L^\varepsilon \in \mathcal{S}(\frac{1-\beta}{2}, K, t_0)$ to (4.19).*

Proof. We begin by noting that, for $\varepsilon < \varepsilon_0 < x_\star$, we have

$$\nu_0(0,\varepsilon) \leq \int_0^\varepsilon C x^\beta dx \leq \frac{C}{1+\beta}\varepsilon^{1+\beta} \leq \frac{C\varepsilon_0^\beta}{1+\beta}\varepsilon.$$

Hence we can certainly take ε_0 sufficiently small such that

$$\nu_0(0,\varepsilon) \leq \tfrac{1}{4}\alpha^{-1}\varepsilon, \quad \text{for every } \varepsilon < \varepsilon_0.$$

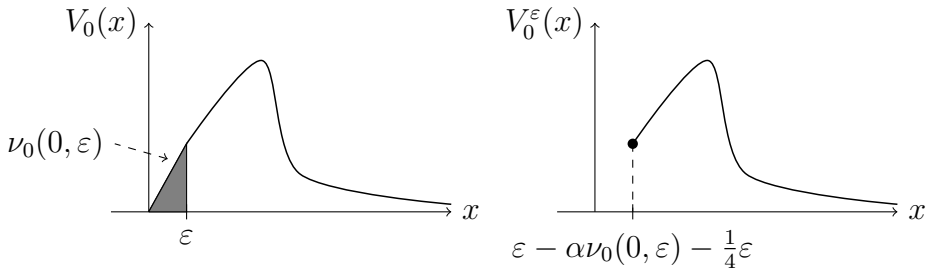


Figure 4.6.1: Given an initial density, V_0, the ε-deleted initial condition constructed in (4.20) is obtained by killing the mass on $(0,\varepsilon)$ and shifting the density towards the origin by the amount $\alpha\nu_0(0,\varepsilon) + \frac{1}{4}\varepsilon$. The proof of Lemma 4.6.6 shows us that V_0^ε vanishes in a neighbourhood of zero.

This guarantees that $\alpha\nu_0(0,\varepsilon) + \frac{1}{4}\varepsilon \leq \frac{1}{2}\varepsilon < \varepsilon$, so we can rewrite (4.19) as

$$L_t^\varepsilon = \nu_0(0,\varepsilon) + F_t^\varepsilon, \tag{4.20}$$

where F_t^ε is given by the McKean–Vlasov problem

$$\begin{cases} X_t^{\varepsilon,x} = x + B_t - \alpha F_t^\varepsilon \\ \tau^{\varepsilon,x} = \inf\{t \geq 0 : X_t^{\varepsilon,x} \leq 0\} \\ F_t^\varepsilon = \int_0^\infty \mathbb{P}(\tau^{\varepsilon,x} \leq t) V_0^\varepsilon(x)dx \\ V_0^\varepsilon(x) = V_0(x + \alpha\nu_0(0,\varepsilon) + \frac{1}{4}\varepsilon)\mathbf{1}_{x+\alpha\nu_0(0,\varepsilon)+\varepsilon/4\geq\varepsilon}, \end{cases}$$

see Figure 4.6.1. As noted in Remark 4.3.6, we can solve to find F^ε, since V_0^ε vanishes on $x < \frac{1}{2}\varepsilon$, so we certainly have the control in (4.7) for some choice of constants. Furthermore, we have

$$\begin{aligned} V_0^\varepsilon(x) &\leq C(x + \alpha\nu_0(0,\varepsilon) + \tfrac{1}{4}\varepsilon)^\beta \mathbf{1}_{x+\alpha\nu_0(0,\varepsilon)+\varepsilon/4\geq\varepsilon} \\ &\leq C(x + \tfrac{1}{2}\varepsilon)^\beta \mathbf{1}_{x\geq\varepsilon/2} \leq 2^\beta C x^\beta, \end{aligned}$$

so we can find constants such that (4.7) holds for V_0^ε uniformly in $\varepsilon \in (0,\varepsilon_0)$. Since the arguments in Section 4.5 for proving Theorem 4.2.7 depend only on the constants in (4.7), we can conclude that the parameters of $\mathcal{S}$ obtained in Theorem 4.2.7 are constant across the class of initial densities $\{V_0^\varepsilon\}_\varepsilon$, for $\varepsilon < \varepsilon_0$. Hence we have the result. □

The solutions in Lemma 4.6.6 are useful because we have the following two comparison results, which say that the ε-deleted solutions dominate and converge to the traditional solutions of (MV) — see Figure 4.6.2.

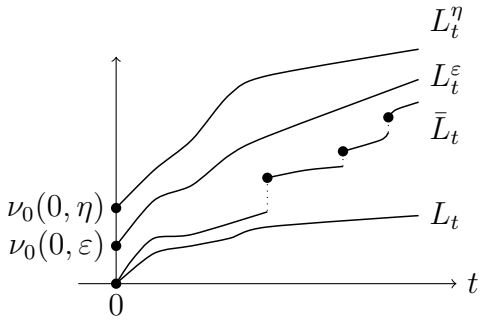


Figure 4.6.2: On a small time interval, the unique differentiable solution, L, from Corollary 4.6.4 and the ε-deleted solutions from Definition 4.6.5 trap any candidate càdlàg solution, $\bar{L}$, by Lemma 4.6.7. Here $\eta > \varepsilon$. Lemma 4.6.8 shows that $L^\varepsilon \to L$ uniformly on a small time interval as $\varepsilon \to 0$, and so we see that $L = \bar{L}$ is forced since the area between the curves above shrinks to zero.

Lemma 4.6.7 (Monotonicity)**.** *Let $\varepsilon > 0$ and $t_0 > 0$ be fixed. Assume that L is a generic solution to (MV) and that L^ε solves (4.19) and that both are continuous on $[0, t_0)$. Then*

$$L_t^\varepsilon > L_t, \qquad \text{for every } t \in [0, t_0).$$

Proof. By construction, the result is true at time $t = 0$. For a contradiction, let t be the first time at which $L_t^\varepsilon = L_t$. Couple both solutions to the same Brownian motion:

$$\begin{aligned} X_s &= X_0 + B_s - \alpha L_s \\ X_s^\varepsilon &= X_0 \mathbf{1}_{X_0 \geq \varepsilon} - \tfrac{1}{4}\varepsilon + B_s - \alpha L_s^\varepsilon, \end{aligned}$$

for $s \in [0, t)$, so that taking the difference gives

$$X_s - X_s^\varepsilon = X_0 \mathbf{1}_{X_0 < \varepsilon} + \tfrac{1}{4}\varepsilon - \alpha(L_s - L_s^\varepsilon) \geq \tfrac{1}{4}\varepsilon.$$

Therefore, from the definition of L_t^ε, we get

$$\begin{aligned} L_t^\varepsilon = \mathbb{P}(\inf_{0 \leq u \leq t} X_u^\varepsilon \leq 0) &= \lim_{s \uparrow t} \mathbb{P}(\inf_{0 \leq u \leq s} X_u^\varepsilon \leq 0) \\ &\geq \lim_{s \uparrow t} \mathbb{P}(\inf_{0 \leq u \leq s} X_u \leq \tfrac{1}{4}\varepsilon) \\ &= L_t + \mathbb{P}(\inf_{0 \leq u \leq t} X_u \in (0, \tfrac{1}{4}\varepsilon]) \end{aligned} \tag{4.21}$$

Now fix any $f \in H^1([0,t])$ such that $\|\alpha L - f\|_{L^\infty([0,t])} < \frac{1}{16}\varepsilon$. Then we have

$$L_t^\varepsilon \geq L_t + \mathbb{P}(\inf_{0 \leq u \leq t} \{X_0 + B_u - f_u\} \in (\tfrac{1}{16}\varepsilon, \tfrac{3}{16}\varepsilon)). \tag{4.22}$$

By Girsanov's Theorem, $u \mapsto B_u - f_u$ is absolutely continuous with respect to Brownian motion, and the infimum of Brownian motion has a density, therefore the probability on the right-hand side of (4.22) is non-zero. Hence $L^\varepsilon_t > L_t$, which is the required contradiction. □

Lemma 4.6.8 (Convergence)**.** *Suppose there exist constants $\gamma \in (0, \frac{1}{2})$, $K > 0$, $t_0 > 0$ and $\varepsilon_0 > 0$ such that $L, L^\varepsilon \in \mathcal{S}(\gamma, K, t_0)$ for all $\varepsilon \in (0, \varepsilon_0)$. Then there exists $t_1 \in (0, t_0]$ such that*

$$\|L^\varepsilon - L\|_{L^\infty(0,t_1)} \to 0, \qquad \text{as } \varepsilon \to 0.$$

Proof. We know from Lemma 4.6.7 that $L^\varepsilon > L$ on $[0, t_0)$. By following the argument in Section 4.4 for the proof of Theorem 4.2.6, and coupling L and L^ε to the same Brownian motion, we have

$$\begin{aligned}
0 \le L^\varepsilon_t - L_t &= \mathbb{P}(\tau^\varepsilon \le t < \tau) \\
&= \int_0^t \mathbb{P}(t < \tau | \tau^\varepsilon = s) dL^\varepsilon_s + \nu_0(0, \varepsilon) \\
&\le \int_0^t \Phi\Big(\frac{\alpha(L^\varepsilon_s - L_s) + \frac{1}{4}\varepsilon}{(t-s)^{1/2}}\Big)(L^\varepsilon)'_s ds + \nu_0(0,\varepsilon) \\
&\le c_0(\|L^\varepsilon - L\|_{L^\infty(0,t)} + \tfrac{1}{4}\varepsilon) \int_0^t \frac{ds}{(t-s)^{1/2} s^\gamma} + \nu_0(0,\varepsilon) \\
&\le c_0 t^{1/2-\gamma}(\|L^\varepsilon - L\|_{L^\infty(0,t)} + \tfrac{1}{4}\varepsilon) + \nu_0(0,\varepsilon),
\end{aligned}$$

where $c_0 > 0$ is a constant independent of t, increasing from line to line as necessary. Therefore, taking a supremum over $t \le t_1$ and taking t_1 sufficiently small, we have

$$\|L^\varepsilon - L\|_{L^\infty(0,t_1)} \le \tfrac{1}{2}\|L^\varepsilon - L\|_{L^\infty(0,t_1)} + o(1),$$

which completes the proof. □

We are now in a position to complete the proof of the main uniqueness theorem. As already indicated, the idea is to trap a general candidate solution from below by the differentiable solution from Corollary 4.6.4 and from above by the ε-deleted solutions, and then to shrink the resulting envelope to zero using the uniform control in Lemma 4.6.8

Proof of Theorem 4.2.8. Take the bootstrapped solution, L, from Corollary 4.6.4. We will be done if we can show that there is not a solution to (MV) on $[0, t_\star)$ that is distinct from L, for any $t_\star \le t_{\text{explode}}$. So for a contradiction, suppose that $\bar{L}$ is such a solution and let t_{jump} be the first discontinuity of $t \mapsto \bar{L}_t$.

By Lemma 4.6.6 we can find $K > 0$, $t_0 > 0$ and a family L^ε satisfying the hypotheses of Lemma 4.6.8. By decreasing t_0 so that $t_0 < t_{\text{jump}}$, Lemma 4.6.7 and Theorem 4.2.6 guarantee

$$L_t \leq \bar{L}_t < L^\varepsilon_t, \qquad \text{for every } t < t_0.$$

By taking $t_1 \in [0, t_0)$ from the conclusion of Lemma 4.6.8 and sending $\varepsilon \to 0$, we conclude that $L = \bar{L}$ on $[0, t_1]$.

This argument can now be restarted from time t_1 and iterated as in the proof of Corollary 4.6.4, hence we conclude that $L = \bar{L}$ up to the minimum of t_{jump} and $t_\star$. If $t_\star \leq t_{\text{jump}}$, then we are done, so suppose $t_{\text{jump}} < t_\star$. By left-continuity, we have that $\nu_{t_{\text{jump}}-} = \bar{\nu}_{t_{\text{jump}}-}$, and, since L does not have a jump at time t_{jump}, the physical jump condition (4.6) gives

$$\begin{aligned}\Delta \bar{L}_{t_{\text{jump}}-} &= \inf\{x \geq 0 : \bar{\nu}_{t_{\text{jump}}-}(0, \alpha x) < x\} \\ &= \inf\{x \geq 0 : \nu_{t_{\text{jump}}-}(0, \alpha x) < x\} = 0.\end{aligned}$$

This contradicts t_{jump} being a jump time of $\bar{L}$, so the proof is complete. □

4.7 Extensions to more general coefficients

In this section we consider some simple extensions of (MV) that incorporate more general drift and diffusion coefficients. The aim is to outline how the analysis can be reduced to that of the previous sections and thus we will only provide sketches of the proofs.

To be specific, we consider the McKean–Vlasov problem

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t)dB_t - \alpha dL_t \\ \tau = \inf\{t \geq 0 : X_t \leq 0\} \\ L_t = \mathbb{P}(\tau \leq t), \end{cases} \tag{4.23}$$

where B is a Brownian motion and the independent initial condition, X_0, is given by ν_0, which is assumed to satisfy (4.7) and taken to be sub-Gaussian. In terms of the coefficients, we assume that $b(t, x)$ is Lipschitz in space with $|b(t, x)| \leq C(1 + |x|)$ and we impose the non-degeneracy condition $c \leq \sigma(t) \leq C$ for all $t \in [0, T]$. Furthermore, we impose an upper bound $b(t, x) \leq M$ for positive $x \geq 0$ (this is only used in Section 4.7.3 and it could be omitted if we could ensure $\kappa = 1$ in Section 4.7.1).

The above setup includes the financial model (4.1). Moreover, it e.g. allows for a Brownian motion with an Ornstein–Uhlenbeck type drift modelling the attraction to

a 'resting state' — as in the original neuroscience model [45]. This could also be of interest in the financial framework, for example to include a target leverage ratio or to model some notion of 'flocking to default' as in [59].

Theorem 4.7.1 (Existence and uniqueness up to explosion)**.** *Let the above assumptions be in place. Then there exists a solution to (4.23) up to time*

$$t_{\text{explode}} := \sup\{t > 0 : \|L\|_{H^1(0,t)} < \infty\} \in (0, \infty]$$

such that, for some $\bar{\beta} > 0$, it holds for every $t_0 < t_{\text{explode}}$ that L is in $\mathcal{S}(\frac{1-\bar{\beta}}{2}, K, t_0)$ for some $K > 0$. Furthermore, this solution is unique in the class of candidate solutions with $L \in \bigcup_{0<\gamma<1/2,\, A>0} \mathcal{S}(\gamma, A, t_0)$.

As for (MV), the notions of weak and strong uniqueness for (4.23), regarded as an SDE in X, are equivalent: taking the difference of two solutions with the same B and L cancels all terms apart from the drift, whereby strong uniqueness is immediate from the Lipschitz property of $x \mapsto b(t, x)$. In essence, the main point in the proof of Theorem 4.7.1 is to have good control on the boundary decay of the relevant transition densities akin to the classical Dirichlet heat kernel estimates.

When there is no spatial dependence in the drift — as in (4.1) — we can do better with essentially no extra work. Specifically, we can directly replicate the monotonicity and trapping arguments from Section 4.6.3, which gives uniqueness amongst generic càdlàg solutions.

Theorem 4.7.2 (Generic uniqueness)**.** *Suppose the above assumptions are in place and that, in addition, $b(t, x) = b(t)$. Then we have uniqueness on $[0, t_{\text{explode}})$ amongst all generic càdlàg solutions to (4.23) in the same sense as Theorem 4.2.8.*

The remainder of this section concerns the proofs of Theorems 4.7.1 and 4.7.2, however, we will only show how to reduce the analysis to that of Sections 4.4, 4.5 and 4.6. The proof of Theorem 4.7.1 is split into three succinct parts: first we discuss the stability of the fixed point map (Section 4.7.1), then we revisit the bootstrapping of short-time existence (Section 4.7.2), and finally we extend the contractivity arguments (Section 4.7.3). Last, the proof of Theorem 4.7.2 is outlined in Section 4.7.4.

4.7.1 Stability of the fixed point map

Let Γ denote the fixed point map for (4.7.1) defined by analogy with (4.5). Then we need to ensure that Proposition 4.5.9, concerning the stability of Γ, also holds true in this more general setting.

The starting point is still the expression for $\Gamma[\ell]$ from Lemma 4.5.1, which remains valid with X given by (4.23). Next, we can observe that the Girsanov argument from Lemma 4.5.4 will now take the form

$$\mathbb{E}[F(X_t)\mathbf{1}_{t<\tau}] \leq \exp\Big\{\frac{\alpha^2\|\ell'\|^2_{L^2(0,t)}}{2c^2(q-1)^2}\Big\}\mathbb{E}[F(\widetilde{X}_t)^q\mathbf{1}_{t<\widetilde{\tau}}]^{\frac{1}{q}},$$

with

$$d\widetilde{X}_t = b(t,\widetilde{X}_t)dt + \sigma(t)dW_t \quad \text{and} \quad \widetilde{\tau} = \inf\big\{t>0 : \widetilde{X}_t \leq 0\big\},$$

where W is a standard Brownian motion and $\widetilde{X}_0$ is an independent random variable distributed according to ν_0. Indeed, we can simply consider the Radon–Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = \exp\Big\{\alpha\int_0^t \frac{\ell'_s}{\sigma(s)}dB_s - \frac{\alpha^2}{2}\int_0^t \frac{(\ell'_s)^2}{\sigma(s)^2}ds\Big\}$$

and otherwise repeat the arguments from the proof of Lemma 4.5.4. From here, we can apply [75, Thm. 2.7] to see that the absorbed process $\widetilde{X}$ has a transition density $\widetilde{V}_t(x,x_0)$ which satisfies the Dirichlet heat kernel type estimate

$$\widetilde{V}_t(x,x_0) \leq C\Big(\big(t^{-1}xx_0\wedge 1\big)t^{-\frac{1}{2}} + \big(t^{-\kappa}x^\kappa x_0^\kappa \wedge 1\big)e^{\delta x_0^2}\Big)e^{-c(x-x_0)^2/t}, \tag{4.24}$$

for some constants $C>0$ and $\kappa\in(0,1]$ only depending on $\delta>0$, where the latter can be taken sufficiently small such that $\int_0^\infty e^{\delta x^2}\nu_0(dx)<\infty$.

Remark 4.7.3. The estimate (4.24) is derived in [75, Sec. 6] by first showing that

$$\widetilde{V}_t(x,x_0) = G^\sigma_{0,t}(x,x_0) + \int_0^t \mathbb{E}_{\widetilde{X}_0=x_0}[b(s,\widetilde{X}_s)\partial_y G^\sigma_{s,t}(x,\widetilde{X}_s)\mathbf{1}_{s<\widetilde{\tau}}]ds,$$

where $G^\sigma_{\cdot,t}$ is the Green's function for $\partial_s u(s,x) = -\frac{1}{2}\sigma(s)^2\partial_{xx}u(s,x)$ as a terminal-boundary value problem on $[0,t)\times\mathbb{R}_+$ with $u(\cdot,0)=0$. This follows from ideas closely related to those in the proof of Lemma 4.5.1 and the estimate then comes from a careful change of measure in the second term to remove the drift of $\widetilde{X}$.

With a view towards Section 4.7.2, fix any $\bar\kappa<\kappa$ and set $\bar\beta := \beta\wedge\bar\kappa$, where β is the decay exponent of the initial condition. Then the power law decay of $\widetilde{V}_t(x,x_0)$ near the boundary, cf. (4.24), implies that we can replicate the arguments from Section 4.5 with $\widetilde{V}_t$ in place of the Dirichlet heat kernel, G_t. In this way Proposition 4.5.9 remains true only with $\bar\beta$ in place of β.

4.7.2 Density bounds for bootstrapping

Let ν_t denote the law of X_t killed at the origin. In order to allow bootstrapping of our short-time results, we need to establish analogoues of Lemma 4.17 and Corollary 4.6.3.

First, referring to (4.24) as above, we can argue by analogy with Proposition 4.3.1 to see that ν_t has a density, V_t. Moreover, these same arguments show that V_t is bounded with $\|V_t\|_\infty \leq \|V_0\|_\infty + C\int_0^\infty e^{\delta x^2}\nu_0(dx)$.

Now suppose $L \in H^1(0,t_0)$ for some $t_0 > 0$ and fix $t < t_0$. As in Section 4.6, we can then define, for all $s \leq t$, the dual process

$$d\widehat{X}_s = -b(t-s, X_{t-s})ds + \sigma(t-s)dB_s - \alpha d\widehat{L}_s,$$

where $\widehat{L}_s = L_{t-s}$ and B is a Brownian motion. As in the proof of Lemma 4.17, it follows from Hunt's switching identity that

$$V_t(x) = \mathbb{E}_{\widehat{X}_0=x}[V_0(\widehat{X}_t)\mathbf{1}_{t<\widehat{\tau}}], \quad \widehat{\tau} = \inf\big\{s \geq 0 : \widehat{X}_s \leq 0\big\}.$$

Thus, by the same Girsanov argument as in the extension of Lemma 4.5.4 above, it holds for any $q > 1$ that

$$V_t(x) \leq C_q \mathbb{E}_{Y_0=x}\left[V_0(Y_t)^q \mathbf{1}_{t<\tau_Y}\right]^{\frac{1}{q}}, \tag{4.25}$$

where

$$dY_s = -b(t-s, X_{t-s})ds + \sigma(t-s)dB_s,$$

and τ_Y is the first hitting time of zero by Y. As for X, the absorbed process Y has a transition density $U_t(y, y_0)$, so we get

$$\mathbb{E}_{Y_0=x}\left[V_0(Y_t)^q \mathbf{1}_{t<\tau_Y}\right] = \int_0^\infty V_0(y)^q U_t(y,x)dy. \tag{4.26}$$

Appealing again to [75], we see that U satisfies precisely the same bound as (4.24). Now observe that $(t^{-1}xy \wedge 1) \leq t^{-\beta q}x^{\beta q}y^{\beta q}$ for $\beta q \leq 1$ and recall that $V_0(y) \leq Cy^\beta$. Using this and the order x^κ boundary decay of the second term in (4.24), as applied to $U_t(y,x)$, it follows easily from (4.25) and (4.26) that

$$V_t(x) \leq C_q'\big(x^\beta + x^{\kappa/q}\big), \quad \text{for all } t < t_0,$$

for any $q > 1$ such that $\beta q \leq 1$. Taking q close enough to 1 so that $\kappa/q \leq \bar{\beta}$ (with $\bar{\beta}$ fixed in Section 4.7.1) we can conclude that V_t satisfies condition (4.7) for every $t \in [0, t_0)$ with $\bar{\beta}$ as the decay exponent. Given this, we immediately obtain the desired analogues of Lemma 4.17 and Corollary 4.6.3.

4.7.3 Contractivity of the fixed point map

In this subsection we finally show that Γ remains an L^∞ contraction, by extending the first part of Theorem 1.6, and then we complete the proof of Theorem 4.7.1.

With $\ell, \bar{\ell} \in \mathcal{S}(\gamma, A, t_0)$ as in the statement of Theorem 4.2.6, consider for $t \in [0, t_0)$ the corresponding processes X^ℓ and $X^{\bar{\ell}}$ given by

$$dX_t^\ell = b(t, X_t^\ell)dt + \sigma(t)dB_t - \alpha d\ell_t, \quad dX_t^{\bar{\ell}} = b(t, X_t^{\bar{\ell}})dt + \sigma(t)dB_t - \alpha d\bar{\ell}_t,$$

coupled through the same Brownian motion and initial condition. Let τ_ℓ and $\tau_{\bar{\ell}}$ denote the respective hitting times of zero.

Since $\bar{\ell}$ is increasing, and recalling also the upper bound on b for $x \geq 0$, it holds for all $s \leq u < \tau_{\bar{\ell}}$ that

$$X_u^{\bar{\ell}} - X_s^{\bar{\ell}} \leq \bar{C} \int_s^u \sigma(r)^2 dr + \int_s^u \sigma(r) dB_r =: \bar{X}_{s,u}. \tag{4.27}$$

Moreover, by the Lipschitz continuity of the drift we have

$$|X_t^{\bar{\ell}} - X_t^\ell| \leq C \int_0^t |X_r^{\bar{\ell}} - X_r^\ell| dr + \alpha \|\ell - \bar{\ell}\|_{L^\infty(0,t)},$$

so Grönwall's lemma yields

$$|X_t^{\bar{\ell}} - X_t^\ell| \leq \alpha e^{Ct} \|\ell - \bar{\ell}\|_{L^\infty(0,t)}.$$

Therefore, on the event $\{\tau_\ell = s\}$ it holds for all $s < t_0$ that

$$X_s^{\bar{\ell}} = X_s^{\bar{\ell}} - X_s^\ell \leq \alpha C_0 \|\ell - \bar{\ell}\|_{L^\infty(0,s)}. \tag{4.28}$$

Using (4.27) and (4.28), and arguing as in (4.10), we deduce that

$$\begin{aligned} \Gamma[\ell]_t - \Gamma[\bar{\ell}]_t &\leq \int_0^t \mathbb{P}\big(\inf_{u \in [s,t]} \bar{X}_{s,u} + \alpha C_0 \|\ell - \bar{\ell}\|_{L^\infty(0,s)} > 0 \big) d\Gamma[\ell]_s \\ &= \int_0^t \int_0^\infty \bar{p}_{t-s}\big(x, \alpha C_0 \|\ell - \bar{\ell}\|_{L^\infty(0,s)}\big) dx d\Gamma[\ell]_s, \end{aligned}$$

for $t < t_0$, where a simple time-change shows that $\bar{p}_r(x, x_0)$ is given explicitly by

$$\bar{p}_r(x, x_0) = \frac{1}{\sqrt{2\pi\varsigma(r)}} \Big(1 - \exp\Big\{-\frac{2xx_0}{\varsigma(r)}\Big\}\Big) \exp\Big\{-\frac{(x - x_0 - \bar{C}\varsigma(r))^2}{2\varsigma(r)}\Big\}$$

with $\varsigma(r) = \int_s^{r+s} \sigma(u)^2 du$. Noting that $c(t-s) \leq \varsigma(t-s) \leq C(t-s)$ and using the bound $1 - e^{-z} \leq z \wedge 1$, it is straightforward to verify that

$$\int_0^\infty \bar{p}_{t-s}(x, \alpha C_0 y) dx \leq C_0'(t-s)^{-\frac{1}{2}} y, \qquad \text{for } \ y \geq 0.$$

Thus, it holds for all $t < t_0$ that

$$\Gamma[\ell]_t - \Gamma[\bar{\ell}]_t \leq C_0' \|\ell - \bar{\ell}\|_{L^\infty(0,t)} \int_0^t (t-s)^{-\frac{1}{2}} d\Gamma[\ell]_s.$$

From here we can repeat the proof of Theorem 4.2.6 to get contractivity of Γ and by Section 4.7.1 we have stability. Thus, the short-time version of Theorem 4.7.1 holds by the same arguments as in the proof of Theorem 4.2.7. Finally, Section 4.7.2 allows us to bootstrap this result up to the H^1 explosion time and so the full Theorem 4.7.1 follows.

4.7.4 Minimality, monotonicity and trapping

In this final subsection we consider the special case $b(t,x) = b(t)$ and sketch the remaining steps towards the proof of Theorem 4.7.2.

Proceeding as in Section 4.7.3, the Grönwall argument is now no longer needed and instead we immediately get $X_s^{\bar{L}} = X_s^{\bar{L}} - X_s^L \leq \alpha(L_s - \bar{L}_s)^+$ on $\{\tau_L = s\}$. Therefore, the ensuing arguments yield

$$(L_t - \bar{L}_t)^+ \leq C' \|(L - \bar{L})^+\|_{L^\infty(0,t)} \int_0^t (t-s)^{-\frac{1}{2}} dL_s$$

and hence the exact same reasoning as in the second part of Theorem 4.2.6 ensures that any differentiable solution $L \in \mathcal{S}(\gamma, A, t_0)$ is minimal.

In terms of the monotonicity and trapping procedure, the only change in Lemma 4.6.7 is that equation (4.22) now reads as

$$L_t^\varepsilon \geq L_t + \mathbb{P}(\inf_{0 \leq u \leq t} \{X_0 + \textstyle\int_0^u b(r)dr + \int_0^u \sigma(r) dB_r - f_u\} \in (\frac{1}{16}\varepsilon, \frac{3}{16}\varepsilon)). \tag{4.29}$$

Consequently, the proof of Lemma 4.6.7 goes through by absolute continuity with respect to the time-changed Brownian motion $u \mapsto \int_0^u \sigma(r) dB_r$. Furthermore, the proof of Lemma 4.6.8 follows immediately by applying the same reasoning as in Section 4.7.3. Given this, the proof of Theorem 4.7.2 can be finished in precisely the same way as the second part of Theorem 4.2.8.

Chapter 5

Instantaneous Feedback ⁜ Part II: Convergence of the Particle System with Common Noise

This chapter continues our study of the contagious McKean–Vlasov system from Chapter 4 with instantaneous feedback. Specifically, we study the case where all the particles are correlated through a common Brownian motion.

5.1 Introduction

This chapter studies the convergence of a contagious McKean–Vlasov system with common noise and illuminates 'blow-up' phenomena in the large population limit. As we will see in detail later, this large population limit gives rise to a conditional McKean–Vlasov problem of the form

$$\begin{cases} X_t = X_0 + \sqrt{1-\rho^2}B_t + \rho B_t^0 - \alpha L_t \\ \tau = \inf\{t \geq 0 : X_t \leq 0\} \\ L_t = \mathbb{P}(\tau \leq t \,|\, B^0). \end{cases} \tag{CMV}$$

Here B and B^0 are independent Brownian motions, X_0 is an independent random variable with law ν_0 supported on $(0, \infty)$, and $\alpha > 0$ and $\rho \in [0, 1)$ are fixed constants. The *common noise*, B^0, plays a pivotal role: in certain cases it has the power to provoke or prevent a *blow-up*, which we define as a jump discontinuity of $t \mapsto L_t$. A key property of (MV) is that these blow-ups occur endogenously: all the random drivers are continuous, yet jump discontinuities can occur through the feedback effect alone. Our main technical results in this chapter are twofold. Firstly, we analyse the blow-ups of (MV) and derive conditions that determine if and when such jump

discontinuities occur. Secondly, we provide an existence result for (MV) with more general drift and diffusion coefficients, and we show that it arises as the limit of the corresponding finite particle system, for a large class of initial conditions.

To get a feel for the workings of (MV), consider the conditional law of X absorbed at the origin given the common noise B^0. This defines a flow of random sub-probability measures $\nu_t := \mathbb{P}(X_t \in \cdot\,,\, t < \tau \,|\, B^0)$ for all $t \geq 0$. A priori, the jump sizes of L are not uniquely specified by (MV), however, at any time $t \geq 0$, a canonical choice is fixed in terms of the left limit $\nu_{t-} = \mathbb{P}(X_{t-} \in \cdot\,,\, t \leq \tau \,|\, B^0)$ according to the jump conditions (5.1) and (PJC) introduced in Section 2.

Remark 5.1.1 (Stochastic evolution equation). Lemma 5.2.1 below shows that the flow ν_t has a density process V_t. In section 3, we will see that—between jump times and equipped with the jump condition (PJC)—this density process is governed by a generalized notion of a stochastic evolution equation (see Remark 5.13 for details of this). The equation comprises a Laplacian from the diffusive effect of the full Brownian noise and a stochastic transport term driven by the common noise B^0. Moreover, the positive feedback emerges as a nonlinear (and nonlocal) transport term that pushes mass towards the boundary in proportion to the flux across the origin—with the capacity to create blow-ups. Two heat plots of V_t, for different realisations of B^0, are presented in Figure 5.2.1—with one of the plots exhibiting a blow-up. See Section 5.3.4 for a brief description of the numerical algorithm.

Our main interest in (MV) stems from its potential to serve as a mean-field model for the interplay between common exposures and contagion in large financial systems, as discussed in [75]. In this case $X_t \in [0, \infty)$ is the *distance-to-default* of a 'typical' bank or credit-risky asset, ρ captures the extent of common exposures, and $\alpha > 0$ imposes a contagious feedback effect from defaults. In this setting, a jump discontinuity of L_t corresponds to a 'default cascade' where the feedback in the accumulation of loss explodes and a macroscopic proportion of the financial system is lost instantaneously.

The above model without the common noise motivates the analysis in [74] and numerical schemes for this model have been explored in [83]. A similar financial framework also provides the motivation for [94, 95]—the latest paper extending the idiosyncratic ($\rho = 0$) model to more generic random drivers (playing the role of the B here) and replacing the constant feedback parameter with a graph of interactions. At a more heuristic level, an analogous model for macroeconomic crises has been proposed in [69] based on a PDE ($\rho = 0$) version of the stochastic evolution equation mentioned in Remark 5.1.1, albeit without addressing the possibility of blow-ups.

Another motivation comes from mathematical neuroscience, where (MV) models the voltage level across a typical neuron in a large mean-field network. This application (with $\rho = 0$) is the focus of [17, 29, 30, 45, 46], where the feedback term describes the jump in voltage level experienced when a neighbouring neuron spikes (i.e. hits a threshold, thus emitting an action potential and then being reset). The addition of a common noise ($\rho > 0$) is novel and partly captures the effect of some systemic random stimulus influencing all neurons simultaneously, perhaps due to some external influence in the environment. While such common noise mean-field models are used in the neuroscience literature, see e.g. [13, 104], they have not been treated rigorously. Finally, we stress that common noise McKean–Vlasov systems have recently appeared in the context of mean-field games—see [26, 27] for a start.

Questions of existence and uniqueness for (MV) are delicate, not least because of the possibility of blow-ups. At present, results are only available for the case $\rho = 0$ [45, 46, 74, 94], with uniqueness only known before the first blow-up [74, 94]. In the present chapter we establish global existence of a 'relaxed' version of (MV) via limit points of the finite particle system, which provides the justification for the aforementioned mean-field models. We mention here that the relaxation of (MV) concerns the adaptedness of L (or ν) to B^0, but further details are left to Section 3.

5.1.1 Overview of the chapter

In Section 5.2, we derive a number of results about the possibility and timing of blow-ups, for a given solution of (MV). In doing so, our main objective is a simple juxtaposition of the behaviour with and without the common noise (Theorem 5.2.7). In Section 5.3, we start by formulating the finite particle system that motivates our analysis of (MV). We then show that, for any subsequence, there is a further subsequence along which the system converges to a 'relaxed' solution of the desired conditional McKean–Vlasov problem (Theorem 5.3.2). Since we do not address uniqueness, the question of whether there is full propagation of chaos is left open.

5.2 The possibility and timing of blow-ups

We begin by recalling the notion of a *physical solution* for (MV, $\rho = 0$) [45, 74, 94], which stipulates that X and L must be càdlàg with jump sizes satisfying the equation

$$\nu_{t-}([0, \alpha \Delta L_t]) = \Delta L_t, \tag{5.1}$$

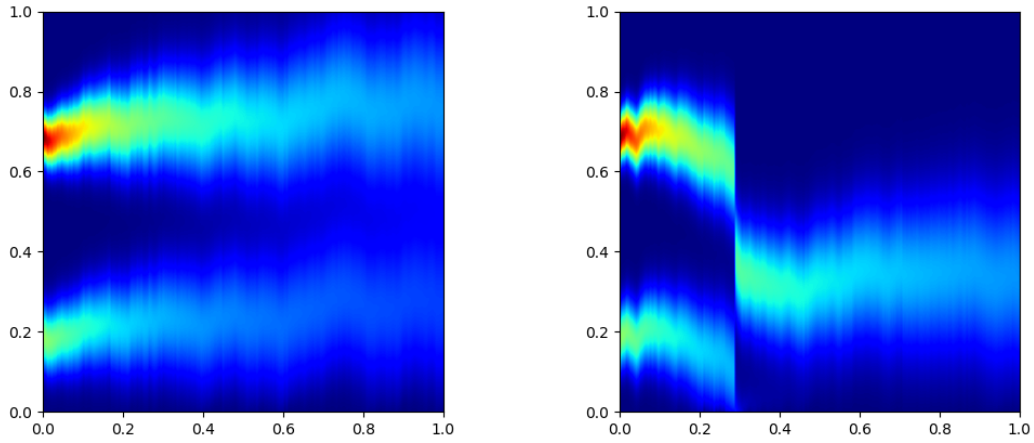


Figure 5.2.1: The plots show different realisations of the solution to (MV) with $\rho = 0.5$ and the same initial condition. Each pixel represents the value of the density at that space-time point, with space on the vertical and time on the horizontal. On the right, the common noise, B^0, decreases sufficiently quickly to cause a blow-up to occur.

together with the minimality constraint

$$\Delta L_t = \inf\{x > 0 : \nu_{t-}([0, \alpha x]) < x\}, \tag{PJC}$$

where we note that $\nu_{t-} = \mathbb{P}(X_{t-} \in \cdot, t \leq \tau)$ in the idiosyncratic ($\rho = 0$) model. We refer to (PJC) as the *physical jump condition*, and we recall that illustrations of this condition can be found in Figures 1.2 and 2.1 from Chapter 4. As we will verify later (see Proposition 5.3.5), these constraints (5.1) and (PJC) remain valid when $\rho > 0$, in which case they are of course equalities of random variables with $\nu_{t-} = \mathbb{P}(X_{t-} \in \cdot, t \leq \tau \,|\, B^0)$.

When proving results that exclude a jump, the key ingredient will be the minimality constraint (PJC). Conversely, when proving results that enforce a jump, our main tool will be the moment method from [74, Thm. 1.1], which states that a blow-up must occur if $\alpha/2 > m_0 := \mathbb{E}X_0$ (see also Proposition 5.2.4 below).

As in the case $\rho = 0$ [74, Prop. 2.1], it is easy to check that ν_t has a bounded density V_t, at all times $t > 0$, regardless of the initial condition.

Lemma 5.2.1 (Density process)**.** *Given a solution to (MV), the associated flow of random sub-probability measures ν_t has a density process $(V_t)_{t>0}$ with $V_t \in L^\infty(0, \infty)$.*

Proof. Fix any $t > 0$ and let p_t denote the density of B_t. Omitting the absorption at the origin, and using Tonelli to change the order of integration, we have

$$\begin{aligned}\nu_t(A) &\leq \mathbb{P}\big(X_0 + \sqrt{1-\rho^2}B_t + \rho B_t^0 - \alpha L_t \in A \mid B^0\big) \\ &= \int_A \int_0^\infty p_{(1-\rho^2)t}(x - x_0 - \rho B_t^0 + \alpha L_t)\nu_0(dx_0)dx,\end{aligned}$$

for any Borel set $A \in \mathcal{B}(0,\infty)$. Consequently, $\nu_t(A) \leq (2\pi(1-\rho^2)t)^{-1/2}\mathrm{Leb}(A)$, for all $A \in \mathcal{B}(0,\infty)$, so the result follows from the Radon–Nikodym theorem. □

5.2.1 The idiosyncratic model with $\rho = 0$

Without the common noise, $t \mapsto \nu_t$ and $t \mapsto L_t$ are deterministic. Hence the blow-up times of (MV) with $\rho = 0$ are completely deterministic. Since we know from Lemma 5.2.1 that each ν_t has a density, a simple observation is that, whenever we have $V_{t_0-}(x) < \alpha^{-1}$ in a neighbourhood of 0, then $\Delta L_{t_0} = 0$ is fixed by (PJC), and so a blow-up cannot occur at that time. In the opposite direction, if $V_{t_0-}(x) \geq \alpha^{-1}$ in a neighbourhood of 0, then a blow-up must occur at time t_0.

Furthermore, from the proof of Lemma 5.2.1, we can see that

$$\|V_t\|_\infty \leq \min\{\|V_0\|_\infty, (2\pi t)^{-1/2}\}, \qquad \text{for every } t \geq 0, \tag{5.2}$$

with $\|V_0\|_\infty$ interpreted as $+\infty$ if ν_0 does not have a density. Consequently, we have:

Corollary 5.2.2. *For any initial condition, ν_0, no blow-up can occur strictly after time $\alpha^2/2\pi$. If ν_0 has a density, V_0, satisfying $\|V_0\|_\infty < \alpha^{-1}$, then a blow-up never occurs.*

By working harder, we can extend this approach to show that if the initial density is supported far from the boundary, then there is insufficient time for the mass to reach the boundary before the time decay in (5.2) prevents a blow-up. The next result gives a linear range in α, which is clearly not sharp, but will be useful in Section 5.2.2.

Proposition 5.2.3 (Linear spatial bounds for blow-up)**.** *If ν_0 is supported on $(0, \frac{1}{2}\alpha)$ then a blow-up must occur. If ν_0 is supported on $(\frac{5}{4}\alpha, +\infty)$ then a blow-up will never occur.*

Proof. If ν_0 is supported on $(0, \frac{1}{2}\alpha)$, then $\mathbb{E}X_0 < \frac{1}{2}\alpha$, so the first part is immediate from [74, Thm. 1.1]. For the second part notice that, by disregarding the absorption

at the origin, we have the upper bound

$$\begin{aligned}\nu_t([0,\alpha x]) &\le \mathbb{P}(X_0 + B_t - \alpha L_t \in [0,\alpha x]) \\ &= \int_0^\infty \int_0^{\alpha x} \frac{1}{\sqrt{2\pi t}} \exp\Big\{ - \frac{(z - x_0 + \alpha L_t)^2}{2t} \Big\} V_0(x_0) dz dx_0.\end{aligned}$$

Suppose $\operatorname{supp}(\nu_0) \subseteq (\alpha b, +\infty)$ for $b > 1$. Since $L_t \le 1$

$$\begin{aligned}\nu_t([0,\alpha x]) &\le \int_0^\infty \int_0^{\alpha x} \frac{1}{\sqrt{2\pi t}} \exp\Big\{ - \frac{(z - \alpha(b-1))^2}{2t} \Big\} V_0(x_0) dz dx_0 \\ &\le \frac{\alpha x}{\sqrt{2\pi t}} \exp\Big\{ - \frac{\alpha^2 (b - 1 - x)^2}{2t} \Big\},\end{aligned}$$

for every $x < b - 1$. Using that $t \mapsto (2\pi t)^{-1/2} e^{-c^2/2t}$ has maximum value $(2\pi e c^2)^{-1/2}$ gives

$$\nu_t([0,\alpha x]) \le \frac{\alpha x}{\sqrt{2\pi e \alpha^2 (b-1-x)^2}} < x$$

for all $x < b - 1 - (2\pi e)^{-1/2}$. Since $2\pi e > 16$, taking $b = \frac{5}{4}$ is sufficient. □

Transformed loss processes

A question arising from [94] is what happens when we replace the linear loss term in (MV, $\rho = 0$) with a general function of the loss process, namely

$$\begin{cases} X_t = X_0 + B_t - \alpha f(L_t) \\ \tau = \inf\{t \ge 0 : X_t \le 0\} \\ L_t = \mathbb{P}(\tau \le t), \end{cases} \tag{5.3}$$

with $f : [0,1) \to \mathbb{R}$. The case $f(x) = -\log(1-x)$ is considered in [94]. To avoid mass escaping to infinity, we can restrict to f bounded below—then we know that $L_t \to 1$ as $t \to \infty$, which in particular is useful in applications of Proposition 5.2.4 below. While we focus on $f(x) = x$ in this as chapter, as in the motivating papers [45, 46, 74, 75], we note that our results extend immediately to (5.3) for any Lipschitz continuous f. Also, it takes only minor adaptations (see [94, Sec. 7] for the kind of changes that would be needed) to cover functions whose Lipschitz constant explodes when $x \uparrow 1$ such as $f(x) = -\log(1-x)$ or $f(x) = (1-x)^{-1}$. For a discontinuous f, one would have to be more careful. A first simple observation is that if f has a positive jump at $x_0 < 1$, then L will have a blow-up at $t_0 = L^{-1}(x_0)$, however, if f has a negative jump, then L might still be continuous. In any case, for whatever function allows for a solution, the following result is the natural extension of [74, Thm. 1.1].

Proposition 5.2.4 (Blow-up for transformed loss)**.** *Suppose L solves (5.3) for some measurable $f : [0, 1) \to \mathbb{R}_+$. If*

$$m_0 := \mathbb{E}X_0 < \alpha\Big(\int_0^{L_t} f(x)dx + (1 - L_t)f(L_t)\Big),$$

then a blow-up must occur before or at time t.

Proof. For a contradiction, suppose that X is continuous up to and including time t. Then, by stopping X at the first time it reaches zero,

$$0 \leq X_{t\wedge\tau} = X_0 + B_{t\wedge\tau} - \alpha f(L_{t\wedge\tau}).$$

Taking expectation, and observing that $s \mapsto L_s$ gives the law of τ,

$$0 \leq m_0 - \alpha\int_0^{\infty} f(L_{t\wedge s})dL_s = m_0 - \alpha\Big(\int_0^{t} f(L_s)dL_s + (1 - L_t)f(L_t)\Big).$$

The proof is now complete by noting that L is of finite variation, since it is increasing, and continuous up to and including time t, by assumption, so the integral on the right-hand side can be re-written as in the statement of the result. □

By taking $f(x) = x$ and sending $t \to \infty$ in Proposition 5.2.4, we recover [74, Thm. 1.1]. Notice that, as soon as the integral of f over $(0, 1)$ is positive, then for any initial condition we can take α large enough to cause a blow-up. In addition, if f is chosen so that its integral over $(0, 1)$ is positively infinite, e.g. $f(x) = (1 - x)^{-1}$, then a blow-up is guaranteed for any initial conditions and any $\alpha > 0$. Note, however, that the case $f(x) = -\log(1 - x)$ from [94] has finite integral, so in this case it is unclear from Proposition 5.2.4 whether a blow-up must occur for every fixed $\alpha > 0$.

Remark 5.2.5. In the preprint [95]—which appeared almost simultaneously with the paper [87] underlying this chapter—solutions to (5.3) are constructed directly (without reference to a particle system) via an extension of Schauder's fixed point theorem. The main result [95, Thm. 2.2] holds for more general stochastic processes Z in place of the Brownian motion B (as in the setting of Chapter 6). If one defines by L^w a solution to (5.3) with driving noise $\sqrt{1 - \rho^2}B_t + \rho w_t$, for a fixed Brownian path w, then this yields a different way of constructing solutions to (MV) with $L_t^w = \mathbb{P}(t \geq \tau \mid B^0 = w)$, although they need not satisfy the jump condition (PJC) and, of course, they are not linked to the motivating particle system (cf. Section 5.3).

5.2.2 The common noise model with $\rho > 0$

For the idiosyncratic ($\rho = 0$) model, the only parameters controlling blow-ups are the feedback strength, α, and initial condition, ν_0. The novelty in the common noise model ($\rho > 0$) is that there are choices of initial condition for which the realisations of the common noise determine whether a blow-up occurs or not. Note, however, that the conclusions of Corollary 5.2.2 still hold, as can be seen from the proof of Lemma 5.2.1.

Since the solution L is no longer deterministic, the occurrence of a blow-up is now a random event. Theorem 5.2.7 shows that we can find initial conditions that do not blow up for the idiosyncratic model, but have a non-zero probability of blow-up when there is common noise ($\rho > 0$). The proof uses an explicit construction to show that a blow-up is forced if the sample path of B^0 is sufficiently negative so as to quickly transport the initial mass towards the boundary. Similarly, we identify analogous situations where the realisation of B^0 can prevent a blow-up. See Figure 5.2.1 for an illustration of the former scenario.

Lemma 5.2.6 (No-crossing lemma)**.** *Suppose we have ν_0 such that* $\operatorname{supp}(\nu_0) \subseteq [x_\star, +\infty)$*, for some $x_\star > 0$, and $f, \bar{f} : [0, \infty) \to \mathbb{R}$ deterministic and continuous that satisfy $f(0), \bar{f}(0) > -x_\star$. Let L and $\bar{L}$ be solutions to*

$$\begin{cases} X_t = Z_0 + \sqrt{1-\rho^2}B_t + f(t) - \alpha L_t \\ \tau = \inf\{t \geq 0 : X_t \leq 0\} \\ L_t = \mathbb{P}(\tau \leq t), \end{cases} \qquad \begin{cases} \bar{X}_t = Z_0 + \sqrt{1-\rho^2}B_t + \bar{f}(t) - \alpha \bar{L}_t \\ \bar{\tau} = \inf\{t \geq 0 : \bar{X}_t \leq 0\} \\ \bar{L}_t = \mathbb{P}(\bar{\tau} \leq t), \end{cases}$$

where $Z_0 \sim \nu_0$. If $f(t) > \bar{f}(t)$ for all $t \in [0, t_0)$ and L is continuous up to time t_0, then $L_t \leq \bar{L}_t$ for all $t \in [0, t_0]$.

Proof. First observe that if $X_s \geq \bar{X}_s$ for all $s < t$, then

$$L_t = L_{t-} = \mathbb{P}(\inf_{s<t} X_s < 0) \leq \mathbb{P}(\inf_{s<t} \bar{X}_s < 0) = \bar{L}_{t-} \leq \bar{L}_t, \tag{5.4}$$

so we are done if we can show $X_t \geq \bar{X}_t$ for all $t \in [0, t_0)$. Let $t_\star$ be the first time $X_t \leq \bar{X}_t$. Since $X_0 > \bar{X}_0$, by right-continuity we have $t_\star > 0$. By continuity of X and the fact that $\bar{X}$ only has downwards jumps ($\bar{L}$ has only upwards jumps), $X_{t_\star} = \bar{X}_{t_\star}$, and this forces

$$\alpha(\bar{L}_{t_\star} - L_{t_\star}) = \bar{f}(t_\star) - f(t_\star) < 0.$$

But this is a contradiction, because the left-hand side is non-negative by (5.4), since $X_t > \bar{X}_t$ for $t < t_\star$. □

Theorem 5.2.7 (Blow-up and common noise)**.** *(i). For every $\alpha > 0$, there exists an initial density V_0 such that, when $\rho = 0$, any solution L to (MV) does not blow up, while, for every $\rho > 0$,*

$$\mathbb{P}(L \text{ has a blow-up}) > 0.$$

(ii). Likewise, there are initial densities such that L has a blow-up when $\rho = 0$, but, for every $\rho > 0$,

$$\mathbb{P}(L \text{ has a blow-up}) < 1.$$

Proof. (i). Take the step-function $V_0(x) = \delta^{-1}\mathbf{1}\{2\alpha \le x \le 2\alpha + \delta\}$. If $\rho = 0$, then a blow-up cannot occur by Proposition 5.2.3, regardless of $\delta > 0$. For $\rho > 0$, our strategy is to consider sample paths of B^0 that move the initial mass close to the boundary in very short time so that only a small amount of mass is lost, and then we apply a modified moment argument to force a blow-up. To this end, define

$$\mathcal{A}_{t_0,\delta,m} := \{\, f : [0,t_0] \to \mathbb{R} \text{ s.t. } |f(t) - mt| < \delta \text{ for all } t \in [0,t_0] \,\},$$

and fix the parameter $m = -2(\alpha - \delta)/t_0$ with $\delta < \alpha$. By Lemma 5.2.6, whenever $\rho B^0_\cdot \in \mathcal{A}_{t_0,\delta,m}$, either L has a blow-up on $[0,t_0]$ or $L \le \bar{L}$ on $[0,t_0]$, where $\bar{L}$ is a solution to the idiosyncratic problem

$$\begin{cases} \bar{X}_t = X_0 - 2\alpha + \delta + \sqrt{1-\rho^2}B_t - \alpha \bar{L}_t \\ \bar{\tau} = \inf\{t \ge 0 : \bar{X}_t \le 0\} \\ \bar{L}_t = \mathbb{P}(\bar{\tau} \le t). \end{cases}$$

Restrict to the case where L does not blow-up on $[0,t_0]$. We are now free to tune t_0. Since $\bar{L}_t \downarrow 0$ as $t \downarrow 0$, take $t_0 \le \delta^2$ such that $\bar{L}_{t_0} \le (1 \wedge \alpha)\delta$.

Next we modify the moment argument from [74, Thm. 1.1]. For $t > t_0$

$$0 \le X_{t\wedge\tau} = X_{t_0\wedge\tau} + \sqrt{1-\rho^2}\Delta_{t_0,t}B_{\cdot\wedge\tau} + \rho\Delta_{t_0,t}B^0_{\cdot\wedge\tau} - \alpha\Delta_{t_0,t}L_{\cdot\wedge\tau}$$

where we use the notation $\Delta_{u,v}f := f(v) - f(u)$ for increments. Taking a conditional expectation and assuming L is continuous upto and including t (which will eventually lead to a contradiction)

$$\tfrac{1}{2}\alpha(L_t - L_{t_0})(2 - L_t - L_{t_0}) \le \mathbb{E}[X_{t_0\wedge\tau} \,|\, B^0] + \rho\mathbb{E}[\Delta_{t_0,t}B^0_{\cdot\wedge\tau} \,|\, B^0].$$

Using that $L_{t_0} \le \delta$ and $L_t \le 1$ gives

$$\tfrac{1}{2}\alpha(L_t - \delta)(1 - \delta) \le \mathbb{E}[X_{t_0\wedge\tau} \,|\, B^0] + \rho\mathbb{E}[\Delta_{t_0,t}B^0_{\cdot\wedge\tau} \,|\, B^0]. \tag{5.5}$$

To estimate the first moment, notice that

$$X_{t_0\wedge\tau} \le |X_{t_0}| \le |X_0 + \rho B^0_{t_0}| + \sqrt{1-\rho^2}\,|B_{t_0}| + \alpha |L_{t_0}|.$$

Taking expectation and using Cauchy–Schwarz and $t_0 \le \delta^2$ gives

$$\mathbb{E}[X_{t_0\wedge\tau} \,|\, B^0] \le \mathbb{E}[|X_0 + \rho B^0_{t_0}| \,|\, B^0] + 2\delta,$$

and since $\rho B^0 \in \mathcal{A}_{t_0,\delta,m}$ we conclude $\mathbb{E}[X_{t_0\wedge\tau} \,|\, B^0] \le 5\delta$. Define

$$\mathcal{B}_{t_0,t_1,\delta} := \{\, f : [0,t_1] \to \mathbb{R} \text{ s.t. } |f(t) - f(t_0)| < \delta \text{ for all } t \in [t_0,t_1]\,\},$$

then if we also have $\rho B^0_\cdot \in \mathcal{B}_{t_0,t_1,\delta}$, (5.5) becomes

$$\tfrac{1}{2}\alpha(L_t - \delta)(1-\delta) \le 7\delta, \qquad \text{i.e.} \quad L_t \le \delta + \frac{14\delta}{\alpha(1-\delta)}. \tag{5.6}$$

Now fix $\delta > 0$ sufficiently small so that the right-hand side is less than $1/2$. But by comparison with Brownian motion, $L_t \to 1$ as $t \to \infty$, so we can certainly find t_1 large enough to contradict (5.6). Therefore

$$0 < \mathbb{P}(\rho B^0_\cdot \in \mathcal{A}_{t_0,m,\delta} \cap \mathcal{B}_{t_0,t_1,\delta}) \le \mathbb{P}(L \text{ has a blow-up}).$$

(ii). Take $V_0(x) = 4\alpha^{-1}\mathbf{1}\{\frac{1}{4}\alpha \le x \le \frac{1}{2}\alpha\}$. If $\rho = 0$ then a blow-up occurs by [74, Thm. 1.1], since $\mathbb{E}X_0 = \frac{3}{8}\alpha < \frac{1}{2}\alpha$. For $\rho > 0$, define $\bar{L}$ to be the solution to the idiosyncratic problem with initial density $x \mapsto 4\alpha^{-1}\mathbf{1}\{\frac{1}{8}\alpha \le x \le \frac{3}{8}\alpha\}$. Since $\bar{L}$ is right-continuous, $\bar{L}_t \to 0$ as $t \to 0$, therefore by adapting the argument in Proposition 5.2.3 we can show that there exists $t_0 > 0$ such that L is continuous on $[0,t_0]$. Furthermore it is no loss of generality to take $t_0 > 0$ small enough so that $\bar{L}_{t_0} \le \frac{1}{16}$. Fix $\delta = \frac{1}{16}\alpha$ and $m > 0$, then if $\rho B^0_\cdot \in \mathcal{A}_{t_0,m,\delta}$, Lemma 5.2.6 implies $L \le \bar{L} \le \frac{1}{16}$ on $[0,t_0]$.

By disregarding absorption at zero, for $t \le t_0$ we have

$$\begin{aligned}
\nu_{t-}([0,\alpha x]) &\le \mathbb{P}(X_{t-} \in [0,\alpha x] \mid B^0) \\
&= \mathbb{P}(X_0 + \sqrt{1-\rho^2}B_t + \rho B^0_t - \alpha L_{t-} \in [0,\alpha x] \mid B^0) \\
&= \int_0^\infty \int_0^{\alpha x} \frac{1}{\sqrt{2\pi(1-\rho^2)t}} \exp\Big\{ - \frac{(z - x_0 - \rho B^0_t + \alpha L_{t-})^2}{2(1-\rho^2)t} \Big\} V_0(x_0)\,dz\,dx_0,
\end{aligned} \tag{5.7}$$

since B is independent of B^0 and L, which are both B^0-measurable. Therefore if $x \le \frac{1}{16}$ then we can bound the integrand to get

$$\nu_{t-}([0,\alpha x]) \le \frac{\alpha}{\sqrt{2\pi(1-\rho^2)t}} \exp\Big\{ - \frac{\alpha^2}{512(1-\rho^2)t} \Big\} \cdot x.$$

Hence we can further restrict $t_0 > 0$ so that for all $t \leq t_0$, $\nu_{t-}(0, \alpha x) < x$, and so if $\rho B^0_\cdot \in \mathcal{A}_{t_0,m,\delta}$, then L does not have a blow-up on $[0, t_0]$, by (PJC).

The work so far has been independent of m, which we now fix as $m = t_0^{-1}K$. By Corollary 5.2.2, we know that a blow-up cannot occur after time $t_1 := \alpha^2/2\pi + t_0$. If $\rho B^0_\cdot \in \mathcal{B}_{t_0,t_1,\alpha}$ then the bound in (5.7) gives

$$\nu_{t-}([0, \alpha x]) \leq \frac{\alpha}{\sqrt{2\pi(1-\rho^2)t_0}} \exp\Big\{ - \frac{(K - \frac{7}{4}\alpha)^2}{2(1-\rho^2)t_1} \Big\} \cdot x,$$

for all $x \leq 1$. Therefore, take K large enough so that $\nu_{t-}([0, \alpha x]) < x$, which ensures no blow-up can occur on $[t_0, t_1]$ either, and hence

$$0 < \mathbb{P}(\rho B^0_\cdot \in \mathcal{A}_{t_0,m,\delta} \cap \mathcal{B}_{t_0,t_1,\alpha}) \leq \mathbb{P}(L \text{ does not have a blow-up}),$$

as required. □

5.3 Existence and particle approximation

In this section, we show how solutions to a 'relaxed' version of (MV) emerge as the support for the limit points of a finite particle system with positive feedback. It is this particle system that motivates the applications described in the introduction and hence it is a salient point that 'relaxed' solutions to (MV) indeed describe the mean-field limit. Nevertheless, we emphasise that uniqueness is not addressed in this chapter and hence there may not be full propagation of chaos (see Chapter 6 for a uniqueness results with common noise in the weak feedback regime). The finite system can be formulated as a systems of SDEs

$$\begin{cases} dX^{i,N}_t = b(t, X^{i,N}_t)dt + \sigma(t)\big(\sqrt{1-\rho(t)^2}dB^i_t + \rho(t)dB^0_t\big) - \alpha dL^N_t \\ L^N_t = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \mathbf{1}_{t \geq \tau^{i,N}} \quad \text{with} \quad \tau^{i,N} = \inf\{t > 0 : X^{i,N}_t \leq 0\}, \end{cases} \tag{5.8}$$

for $i = 1, \ldots, N$, absorbed when they drop below the origin. Here, and it what follows, $B^0, \ldots, B^N$ are independent Brownian motions and $X^1_0, \ldots, X^N_0$ are i.i.d. copies of X_0 (independent of the Brownian motions).

Some care has to be taken in specifying what is meant by L^N in (5.8). Clearly, it only jumps when $X^{j,N}_{t-} = 0$ for some (previously unabsorbed) particle j, but if taking $\Delta L^N_t = 1/N$ means that $X^{i,N}_{t-} - \alpha \Delta L^N_t \leq 0$ for more (previously unabsorbed) particles $i \neq j$, then these are also absorbed at time t, and so forth. Thus, a cascade is initiated, which is resolved by demanding that ΔL^N_t equals the smallest k/N such

that the downward shifts from taking $\Delta L^N_t = k/N$ in (5.8) result in no more than a total of k particles being absorbed at time t, based on their positions at 'time $t-$'. In other words, we require L^N to be the unique piecewise constant càdlàg process satisfying

$$\Delta L^N_t = \inf\big\{\tfrac{k}{N} \geq 0 : \boldsymbol{\nu}^N_{t-}\big([0, \alpha\tfrac{k}{N}]\big) \leq \tfrac{k}{N}\big\}, \quad \text{where} \quad \boldsymbol{\nu}^N_t := \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{t<\tau^{i,N}}\delta_{X^{i,N}_t},$$

for every $t \geq 0$. This, of course, is nothing but the discrete analogue of the physical jump condition (PJC). From here, existence and uniqueness of (5.8) is immediate, under Assumption 5.3.1, as a unique strong solution can be constructed recursively on the intervals of constancy for L^N. We have included a spatial drift $b(t,x)$ with up to linear growth, to e.g. allow for Ornstein–Uhlenbeck type dynamics, as this is desirable for the motivating financial [75] and neuroscience [13] applications (our arguments for such a drift extend immediately to a drift also depending on the empirical mean in a Lipschitz way, but we leave this out for simplicity of presentation).

The fact that the finite system is of mean-field type—but interacting through the paths rather than just marginals—can be summarised by writing

$$L^N_t = \mathbf{P}^N(\hat{\tau} \leq t) \;\text{ and }\; \boldsymbol{\nu}^N_t = \mathbf{P}^N(\hat{\eta} \in \cdot, t < \hat{\tau}), \quad \text{where} \quad \mathbf{P}^N := \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}.$$

Here the empirical measures, $\mathbf{P}^N$, are understood as random probability measures on the space of càdlàg paths, $D_{\mathbb{R}}$, and $\hat{\eta}$ denotes the canonical process on $D_{\mathbb{R}}$ with $\hat{\tau}$ denoting its first hitting time of zero, that is, $\hat{\tau} = \inf\{t > 0 : \hat{\eta}_t < 0\}$. We will use this notation throughout the remaining part of the chapter. Moreover, we will consistently use the boldface font (e.g. $\mathbf{P}$ and $\boldsymbol{\nu}$) for *random* measures, whereas the blackboard font (e.g. $\mathbb{P}$ and its expectation operator $\mathbb{E}$) is reserved for fixed background spaces.

Heuristically, one could now hope that—along a subsequence and in a suitable topology—the pair $(\mathbf{P}^N, B^0)$ will converge in law to some limit $(\mathbf{P}^*, B^0)$ which satisfies $\mathbf{P}^* = \mathrm{Law}(X \,|\, B^0)$, where X solves (MV) with $L_t = \mathbf{P}^*(\hat{\tau} \leq t)$. However, the idea that $\mathbf{P}^*$, and hence L, should turn out to be B^0-measurable is asking too much of the weak limit point $(\mathbf{P}^*, B^0)$. Fortunately, this property can be relaxed vis-à-vis the original formulation of (MV) without altering any essential features of the problem. This relaxation, explained below, is crucial in order to recover (MV) as a large population limit of the finite system (5.8). Specifically, we define a 'relaxed' solution to (MV) as comprising a family $(X, B, B^0, \mathbf{P})$, defined on a probability space $(\Omega, \mathbb{P}, \mathcal{F})$ of choice,

such that

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t)\big(\sqrt{1-\rho(t)^2}dB_t + \rho(t)dB_t^0\big) - \alpha dL_t \\ \tau = \inf\{t > 0 : X_t \leq 0\} \text{ and } \mathbf{P} = \mathrm{Law}(X \,|\, B^0, \mathbf{P}) \text{ under } \mathbb{P} \\ L_t = \mathbb{P}(\tau \leq t \,|\, B^0, \mathbf{P}) \text{ and } (B^0, \mathbf{P}) \perp B \text{ under } \mathbb{P}, \end{cases} \tag{rCMV}$$

where $\mathbf{P}$ is random probability measure, X is càdlàg, and (B, B^0) is 2-dim Brownian motion under $\mathbb{P}$.

We stress that the fixed point type constraint $\mathbf{P} = \mathrm{Law}(X \,|\, B^0, \mathbf{P})$ is an important part of the solution, which characterizes the extra information in the conditioning and encodes the fact that $L_t = \mathbf{P}(\hat{\tau} \leq t)$, as should be required for a limit $\mathbf{P}$ of $\mathbf{P}^N$ with $L^N = \mathbf{P}^N(\hat{\tau} \leq t)$. Observe also that, if the random measure $\mathbf{P}$ is B^0-measurable, then the above simplifies to the exact form of (MV). In the absence of B^0-measurability, however, it is the independence of $(\mathbf{P}, B^0)$ and B that constitutes the most important point. It allows (rCMV) to retain all the essential aspects of (MV) and, in particular, it ensures that the results from Section 2 are equally valid for (rCMV)—by simply conditioning on the pair $(\mathbf{P}, B^0)$ whenever we conditioned on B^0. Furthermore, this independence will feature prominently in the discussion after Theorem 5.3.2 below, where a slightly different point of view on (MV) and (rCMV) is presented. We begin our analysis by outlining the conditions under which our results hold.

Assumption 5.3.1 (Structural conditions)**.** The drift, $b(t, x)$, is assumed to be Lipschitz in space with linear growth bound $|b(t, x)| \leq C(1 + |x|)$. The volatility, $\sigma(t)$, and the correlation, $\rho(t)$, are taken to be deterministic functions in $\mathcal{C}^\kappa(0, T)$, for some $\kappa > 1/2$, satisfying the non-degeneracy conditions $0 \leq \rho(t) \leq 1 - \epsilon$ and $\epsilon \leq \sigma(t) \leq \epsilon^{-1}$, for some $\epsilon \in (0, 1)$. Furthermore, we assume that $\int x^8 d\nu_0 < \infty$, and that $\nu_0([0, x]) \leq \frac{1-\epsilon}{2}\alpha^{-1}x$ for all $x > 0$ sufficiently small, for some $\epsilon \in (0, 1)$, where ν_0 is the distribution of the initial condition X_0 which takes values in $(0, \infty)$.

Notice that, by a standard Grönwall argument, the moment assumption on X_0 guarantees $\sup_{N\geq 1} \mathbb{E}[\sup_{t\leq T} |X_t^{1,N}|^8] < \infty$, which we shall need later. We can now state the main result of this section, whose proof is the subject of the remaining part of the chapter (see Sections 3.1 and 3.2).

Theorem 5.3.2 (Existence and convergence)**.** *Suppose Assumption 5.3.1 is satisfied. Then any subsequence of the pair $(\mathbf{P}^N, B^0)$ has a further subsequence, which converges in law to some $(\mathbf{P}, B^0)$, where B^0 is a Brownian motion also denoted B^0 and $\mathbf{P}$ is a random probability measure $\mathbf{P} : \Omega \to \mathcal{P}(D_\mathbb{R})$. Given this limit, there is a background space $(\Omega, \mathcal{F}, \mathbb{P})$, which carries the Brownian motion B^0, another Brownian motion*

$B \perp_{\mathbb{P}} B^0$, and a càdlàg process $(X_t)_{t\in[0,T]}$, together with the random probability measure $\mathbf{P}$, such that (rCMV) is satisfied by $(X, B, B^0, \mathbf{P})$. Moreover, if we define

$$\boldsymbol{\nu}_t := \mathbf{P}(\hat{\eta}_t \in \cdot\,,\, t < \hat{\tau}) = \mathbb{P}(X_t \in \cdot\,,\, t < \tau \,|\, B^0, \mathbf{P}),$$

then we have the minimality constraint

$$\Delta L_t = \inf\{x \geq 0 : \boldsymbol{\nu}_{t-}([0, \alpha x]) < x\}, \tag{5.9}$$

which determines the jump sizes of L.

Note that $L_t = 1 - \boldsymbol{\nu}_t(0,\infty)$ with the jump sizes determined by $\boldsymbol{\nu}_{t-}$. Hence it is natural to consider $(X, B, B^0, \boldsymbol{\nu})$ as the leanest way of writing the solution, where $\boldsymbol{\nu} = (\boldsymbol{\nu}_t)_{t\in[0,T]}$ is viewed as a measure-valued stochastic process. In line with this point of view, we have the following result, which is proved in Section 3.3.

Proposition 5.3.3 (Filtration)**.** *Define $\boldsymbol{\nu} := (\boldsymbol{\nu}_t)_{t\in[0,T]}$ as a stochastic process whose marginals are the random measures $\boldsymbol{\nu}_t : \Omega \to \mathrm{M}(\mathbb{R})$, where $\mathrm{M}(\mathbb{R})$ is the space of finite measures on $\mathbb{R}$ (equipped with its Borel σ-algebra under the total variation norm). Then there is a natural choice of filtration $(\mathcal{F}_t)_{t\in[0,T]}$, forming a filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, for which $(X, B, B^0, \boldsymbol{\nu})$ is adapted and (B, B^0) is a 2d Brownian motion.*

The filtration alluded to above is natural in the sense that it is generated by $(X, B, B^0, \boldsymbol{\nu})$, but we reserve the details for Section 3.3. The main point of the proof is to ensure that B and B^0 remain Brownian motions in the filtration, which essentially amounts to verifying that $\boldsymbol{\nu}_t$ is in fact independent of the future increments of B^0 (by transferring this property from the particle system).

Following on from Proposition 5.3.3, there is a natural sub-filtration $\mathcal{F}_t^{B^0,\boldsymbol{\nu}} \subset \mathcal{F}_t$ generated by B^0 and $\boldsymbol{\nu}$, which is independent of B due to $(B^0, \mathbf{P}) \perp B$ (again, we refer to Section 3.3 for the details of this construction). The adaptedness to this sub-filtration implies that $\boldsymbol{\nu}$ and L are in fact given by $\langle \boldsymbol{\nu}_t, \phi\rangle = \mathbb{E}[\phi(X_t)\mathbf{1}_{t<\tau} \,|\, \mathcal{F}_t^{B^0,\boldsymbol{\nu}}]$ and $L_t = \mathbb{P}(\tau \leq t \,|\, \mathcal{F}_t^{B^0,\boldsymbol{\nu}})$, where we underline once more that $\mathcal{F}_t^{B^0,\boldsymbol{\nu}}$ is independent of B, while the state process, X, is adapted to the larger filtration $\mathcal{F}_t$. As we will now demonstrate, the time dynamics of $\boldsymbol{\nu}_t$ satisfies—at least formally—a certain stochastic evolution equation driven by B^0, and the fact that L_t is conditional on $\mathcal{F}_t^{B^0,\boldsymbol{\nu}}$, rather than just $\mathcal{F}_t^{B^0}$, simply amounts to $\boldsymbol{\nu}_t$ not necessarily being adapted to its driving noise.

To derive the dynamics of $\boldsymbol{\nu}$, we will consider how it acts on smooth test functions vanishing at the origin, say, $\phi \in C_b^2(\mathbb{R})$ with $\phi(0) = 0$. Let $t_0 \equiv 0$, and let $t_k < t_{k+1}$

denote two successive (random) jump times of L. For $t \in [t_k, t_{k+1})$, Itô's formula then yields

$$\begin{aligned}\phi(X_t)\mathbf{1}_{t<\tau} = \phi(X_{t_k}) &+ \frac{1}{2}\int_{t_k}^t \mathbf{1}_{s<\tau}\partial_{xx}\phi(X_s)\sigma(s)ds + \int_{t_k}^t \mathbf{1}_{s<\tau}\partial_x\phi(X_s)b(s,X_s)ds \\ &- \alpha\int_{t_k}^t \mathbf{1}_{s<\tau}\partial_x\phi(X_s)dL_s + \int_{t_k}^t \mathbf{1}_{s<\tau}\partial_x\phi(X_s)\sigma(s)\rho(s)dB_s^0 \\ &+ \int_{t_k}^t \mathbf{1}_{s<\tau}\partial_x\phi(X_s)\sigma(s)\sqrt{1-\rho(s)^2}dB_s, \end{aligned} \tag{5.10}$$

where we have used that $\phi(X_t)\mathbf{1}_{t<\tau} = \phi(X_{t\wedge\tau})$, for all $t \geq 0$, by virtue of $\phi(0) = 0$. Notice that $\{t_k \leq t < t_{k+1}\}$ is in $\mathcal{F}_t^{B^0,\boldsymbol{\nu}}$, so the event $t \in [t_k, t_{k+1})$ is fixed conditional on this filtration. Recalling that $\langle \boldsymbol{\nu}_t, \phi\rangle = \mathbb{E}[\phi(X_t)\mathbf{1}_{t<\tau} \,|\, \mathcal{F}_t^{B^0,\boldsymbol{\nu}}]$, we can thus take a conditional expectation in (5.10) to find that, on $[t_k, t_{k+1})$,

$$\begin{cases} d\langle \boldsymbol{\nu}_t, \phi\rangle = \langle \boldsymbol{\nu}_t, \mathcal{A}_t\phi\rangle dt - \langle \boldsymbol{\nu}_t, \alpha\partial_x\phi\rangle dL_t + \langle \boldsymbol{\nu}_t, \sigma(t)\rho(t)\partial_x\phi\rangle dB_t^0 \\ L_t = 1 - \boldsymbol{\nu}_t(0,\infty), \end{cases} \tag{5.11}$$

for all $\phi \in C_b^2(\mathbb{R})$ with $\phi(0) = 0$, where $\mathcal{A}_t := \frac{1}{2}\sigma(t)^2\partial_{xx} + b(t,\cdot)\partial_x$, and the initial condition is given by the translation

$$\boldsymbol{\nu}_{t_k} = \boldsymbol{\nu}_{t_k-}(\cdot + \alpha\Delta L_{t_k}), \quad \text{where} \quad \Delta L_{t_k} = \inf\{x \geq 0 : \boldsymbol{\nu}_{t_k-}([0,\alpha x]) < x\}, \tag{5.12}$$

which is specified entirely in terms of the evolution on the previous interval $[t_{k-1}, t_k)$.

To deduce (5.11) from (5.10) we have used that the final integral in (5.10) is killed by the conditional expectation, since $\mathcal{F}_t^{B^0,\boldsymbol{\nu}}$ is independent of B, and that the two middle integrals (with integrators L and B^0) commute with the conditional expectation by virtue of L and B^0 both being $\mathcal{F}_t^{B^0,\boldsymbol{\nu}}$-adapted (see e.g. [73, Lemma 8.9]).

Remark 5.3.4 (Generalized SPDE). We can view (5.11)-(5.12) as giving rise to a generalized notion of a nonlinear (and nonlocal) SPDE for $\boldsymbol{\nu}$ driven by B^0. Indeed, recalling that $\boldsymbol{\nu}_t$ has a density V_t, by Lemma 5.2.1, we can interpret (5.11) as a weak solution of

$$\begin{cases} dV_t(x) = \mathcal{A}_t^* V_t(x)dt + \alpha\partial_x V_t(x)dL_t + \sigma(t)\rho(t)\partial_x V_t(x)dB_t^0 \\ L_t = 1 - \int_0^\infty V_t(x)dx \ \text{ and } \ V_t(0) = 0, \end{cases} \tag{5.13}$$

on $[t_k, t_{k+1})$, with initial condition $V_{\tau_k}(x) = V_{\tau_k-}(x + \alpha\Delta L_{\tau_k})$. Here $\mathcal{A}_t^*$ is the adjoint of $\mathcal{A}_t$ defined by $\mathcal{A}_t^*\phi(x) = \frac{1}{2}\sigma(t)^2\partial_{xx}\phi(x) + \partial_x[b(t,x)\phi(x)]$. Aside from being weak in the sense of generalized functions, our solution to (5.13) is also probabilistically weak

in the sense that the filtration is allowed to be strictly bigger than that generated by the driving noise B^0 alone (see e.g. [43] for other examples of such a situation).

In terms of calling (5.13) an SPDE, one should of course be careful that the equation must be understood recursively on the intervals between jump times (which could in principle accumulate), and we stress that, even on these intervals, L cannot be expected to be absolutely continuous when $\rho \not\equiv 0$ (while it is naturally of finite variation, as it is increasing). The regularity of (5.13)—and indeed the proper framework in which to understand it rigorously—are interesting questions, which we do not pursue here.

Regularity issues aside, we find the 'recursive' SPDE point of view very instructive, and it pinpoints precisely how (rCMV) and (MV) represent the same problem. The fact that $\mathbf{P}$ can fail to be the true conditional law given B^0 corresponds to L_t being conditional on $\mathcal{F}_t^{B^0,\boldsymbol{\nu}}$ in place of $\mathcal{F}_t^{B^0}$, which is simply a consequence of $\boldsymbol{\nu}$ only being a probabilistically weak solution to the generalized SPDE. Without a result on uniqueness, this is as much as one can hope for, when passing to a limit point in our particle system. In particular, we note that, without insisting on the physical jump condition (5.9), one is immediately faced with non-uniqueness of solutions to (rCMV) and hence (5.13). If, on the other hand, one could prove pathwise uniqueness with (5.9) in place, then a Yamada–Watanabe argument would yield a setup in which $\boldsymbol{\nu}$ is adapted to B^0. The latter is the approach taken in [73, 75] for closely related classes of (globally defined) SPDEs on the half-line, however, those uniqueness arguments do not extend to the present model due to the irregularities in L.

Before embarking on the proofs of Theorem 5.3.2 and Proposition 5.3.3 (in the three coming subsections) we provide a final result. It says that the condition (5.9) constrains the smallest jump size a solution to (rCMV)—or (MV)—can have whilst also being càdlàg. The proof is based on an adjustment to the argument in [74, Prop. 1.2] for the model without a common source of noise.

Proposition 5.3.5 (Minimality of jumps)**.** *If L is a càdlàg process satisfying (rCMV), then*

$$\Delta L_t \geq \inf\{x > 0 : \boldsymbol{\nu}_{t-}([0, \alpha x]) < x\}$$

with probability one.

Proof. We will proceed by contradiction, so suppose L is a càdlàg solution for which the inequality is violated at some time $t \geq 0$. By following the strategy in the proof

of [74, Prop. 1.2] it is no loss of generality to have $t = 0$, so take $x_0 > 0$ such that $\nu_0([0, x]) \geq \alpha^{-1} x$ for $x < \alpha x_0$. Then we have

$$L_h \geq \alpha^{-1} \int_0^{\alpha x_0} \mathbb{P}_{X_0 = x}(\tau \leq h \,|\, B^0, \mathbf{P}) dx. \tag{5.14}$$

We must account for the common noise and drift term in the dynamics of X. To this end, we introduce the notation

$$Y_t^0 := \int_0^t \sigma(s)\rho(s) dB_s^0, \qquad Y_t := \int_0^t \sigma(s)\sqrt{1 - \rho(s)^2} dB_s.$$

Using the bounds on b from Assumption 5.3.1, we then have

$$X_t \leq X_0 + Ct + C \int_0^t X_s ds + Y_t^0 + Y_t - \alpha L_t,$$

for all $t \leq \tau$. Applying Grönwall's lemma gives

$$X_t \leq e^{Ct} X_0 + C e^{Ct} t + \bar{Y}_t^0 + \bar{Y}_t - \alpha \bar{L}_t \leq e^{Ct} X_0 + C e^{Ct} t + \bar{Y}_t^0 + \bar{Y}_t - \alpha L_t$$

where we have used the short-hand $\bar{f}_t := \int_0^t e^{C(t-s)} df_s$ and noted that $\bar{L} \geq L$. Since $e^{Ch} = 1 + O(h)$ as $h \downarrow 0$, taking a (random) subsequence of $h \downarrow 0$ for which $\bar{Y}_h^0 \leq -h^{1/2}$ (recall $\bar{Y}^0$ is time-changed Brownian motion) we have

$$X_h \leq X_0 + \bar{Y}_h - \alpha L_h,$$

along that subsequence for $h > 0$ sufficiently small.

Returning to (5.14),

$$L_h \geq \alpha^{-1} \int_0^{\alpha x_0} \mathbb{P}(\bar{Y}_h \leq -x + \alpha L_h \,|\, B^0, \mathbf{P}) dx.$$

where $\bar{Y}_h \sim \text{Normal}(0, v(h))$ and independent of $(B^0, \mathbf{P})$ with $v(h) := \int_0^h \sigma(t)(1 - \rho(t)^2) e^{C(h-t)} dt \asymp h$. Therefore, after a change of variables, we get

$$\alpha v(h)^{-\frac{1}{2}} L_h \geq \int_{-\alpha v(h)^{-1/2} L_h}^{\alpha v(h)^{-1/2} x_0 - \alpha v(h)^{-1/2} L_h} \Phi(-y) dy,$$

with Φ the normal cdf. At this point the remainder of the proof is identical to that of [74, Prop. 1.2]. □

5.3.1 Preliminary estimates

Due to our particular setup and the choice of topology in Section 3.2 below, the key prerequisite for tightness is sufficient control over the loss process, L^N, near the initial time $t = 0$. To achieve this, we need to exploit some control on the initial condition near the origin. The assumption $\nu_0([0,x]) \leq \frac{1-\epsilon}{2}\alpha^{-1}x$ is by no means optimal, but it is already quite general (for $\rho = 0$, [46, 94] assumes ν_0 is supported away from the boundary) and it allows for an intuitive argument to control the loss of mass. The point of the constant $c = \frac{1-\epsilon}{2} < 1$ is that we can utilise $c < 1 - c$. Moreover, we rely on the fact that ρ is bounded away from 1, as we need to single out the contribution of the common noise and utilise the averaging of the independent Brownian motions in a law of large numbers argument.

Proposition 5.3.6 (Initial control of the loss)**.** *The loss process, L^N, satisfies*

$$\limsup_{N\to\infty} \mathbb{P}\big(L^N_\delta \geq \alpha^{-1}\log(\delta^{-1})\delta^{\frac{1}{2}}\big) = 0.$$

for all $\delta \in (0,\delta_0)$, for some small enough $\delta_0 > 0$.

Proof. **Step 1.** Fix $1-\epsilon < \lambda' < \lambda < 1$. Setting $\varepsilon = \varepsilon(\delta) := \delta^{\frac{1}{2}}\log(\delta^{-1})$, we have

$$\mathbb{P}(L^N_\delta \geq \alpha^{-1}\varepsilon) \leq \mathbb{P}\big(L^N_\delta \geq \alpha^{-1}\varepsilon,\ \nu^N_0([0,\varepsilon]) \leq \tfrac{\lambda'}{2}\alpha^{-1}\varepsilon\big) + \mathbb{P}\big(\nu^N_0([0,\varepsilon]) \geq \tfrac{\lambda'}{2}\alpha^{-1}\varepsilon\big),$$

where $\nu^N_0 := \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_0}$. Observe that, for all ε sufficiently small,

$$\lim_{N\to\infty} \mathbb{P}\big(\nu^N_0([0,\varepsilon]) \geq \tfrac{\lambda'}{2}\alpha^{-1}\varepsilon\big) = \mathbf{1}_{\{\nu_0([0,\varepsilon]) \geq \frac{\lambda'}{2}\alpha^{-1}\varepsilon\}} = 0,$$

by the law of large numbers and Assumption 5.3.1. Note also that $\nu^N_0([0,\varepsilon]) \leq \frac{\lambda'}{2}\alpha^{-1}\varepsilon$ if and only if $\#\{i : X^{i,N}_0 \in [0,\varepsilon]\} \leq N\frac{\lambda'}{2}\alpha^{-1}\varepsilon$. Thus, letting $\mathcal{I}$ denote the random index set $\mathcal{I} = \{i : X^{i,N}_0 \geq \varepsilon\}$, we get

$$\mathbb{P}\big(L^N_\delta \geq \alpha^{-1}\varepsilon,\, \nu^N_0([0,\varepsilon]) \leq \tfrac{\lambda'}{2}\alpha^{-1}\varepsilon\big) \leq \sum_{|\mathcal{I}_0| \geq N(1-\frac{\lambda'\varepsilon}{2\alpha})} \mathbb{P}(L^N_\delta \geq \alpha^{-1}\varepsilon \mid \mathcal{I} = \mathcal{I}_0)\mathbb{P}(\mathcal{I} = \mathcal{I}_0).$$

Step 2. Next we define

$$\tau := \inf\{t \geq 0 : L^N_t \geq \tfrac{\lambda}{2}\alpha^{-1}\varepsilon\} \quad \text{and} \quad \mathcal{J} := \big\{i : \inf_{s<\tau\wedge\delta}(X^{i,N}_s - X^{i,N}_0) \leq -\tfrac{1}{2}\varepsilon\big\}.$$

Then we claim that, for every $\mathcal{I}_0$ with $|\mathcal{I}_0| \geq N(1-\frac{\lambda'}{2}\alpha^{-1}\varepsilon)$, we have

$$\mathbb{P}(L^N_\delta \geq \alpha^{-1}\varepsilon \mid \mathcal{I} = \mathcal{I}_0) \leq \mathbb{P}\big(|\mathcal{I}_0 \cap \mathcal{J}| \geq N\tfrac{\lambda-\lambda'}{2}\alpha^{-1}\varepsilon \mid \mathcal{I} = \mathcal{I}_0\big). \tag{5.15}$$

Suppose, for a contradiction, that $L^N_\delta \geq \alpha^{-1}\varepsilon$ but $|\mathcal{I}_0 \cap \mathcal{J}| < N\frac{\lambda-\lambda'}{2}\alpha^{-1}\varepsilon$. Then

$$|\mathcal{I}_0 \cap \mathcal{J}| + |\mathcal{I}_0^{\complement}| < N\tfrac{\lambda-\lambda'}{2}\alpha^{-1}\varepsilon + N\tfrac{\lambda'}{2}\alpha^{-1}\varepsilon = N\tfrac{\lambda}{2}\alpha^{-1}\varepsilon,$$

so even if all these particles are killed in the interval $[0, \tau\wedge\delta]$, their total contribution to $L^N_{\tau\wedge\delta}$ is less than $\frac{\lambda}{2}\alpha^{-1}\varepsilon$. In particular, this means that, at time $\tau\wedge\delta$, the largest possible downward jump for the remaining particles (indexed by $\mathcal{I}_0 \cap \mathcal{J}^{\complement}$) is $-\frac{\lambda}{2}\varepsilon$. Notice, however, that these remaining particles all satisfy $X^{j,N}_0 \geq \varepsilon$ and $\inf_{s<\tau\wedge\delta}(X^{j,N}_s - X^{j,N}_0) > -\frac{1}{2}\varepsilon$, so we must have $\inf_{s<\tau\wedge\delta} X^{j,N}_s > \frac{1}{2}\varepsilon$. Consequently, a translation by $-\frac{\lambda}{2}\varepsilon$ cannot cause any of them to drop below the origin and hence $L^N_{\tau\wedge\delta} < \frac{\lambda}{2}\alpha^{-1}\varepsilon$, so $\delta < \tau$, which certainly implies $L^N_\delta < \alpha^{-1}\varepsilon$. This yields the desired contradiction.

Step 3. Note that $\alpha L^N_s < \frac{\lambda}{2}\varepsilon$ for $s < \tau\wedge\delta$. Therefore, using the linear growth assumption on the drift, we deduce that $\inf_{s<\tau\wedge\delta}(X^{i,N}_s - X^{i,N}_0) \leq -\frac{1}{2}\varepsilon$ implies

$$-\delta C(1 + \sup_{s<\delta}|X^{i,N}_s|) + \inf_{s<\delta}\{Y^i_s + Y^0_s\} - \tfrac{\lambda}{2}\varepsilon \leq -\tfrac{1}{2}\varepsilon, \tag{5.16}$$

where

$$Y^0_s := \int_0^s \sigma(r)\rho(r)dB^0_r \quad \text{and} \quad Y^i_s := \int_0^s \sigma(r)\sqrt{1-\rho(r)^2}dB^i_r.$$

From here, we split the analysis on the intersection of the events

$$\Big\{\sup_{s<\delta}|Y^0_s| \leq \delta^{\frac{1}{2}}\log\log(\delta^{-1})\Big\} \quad \text{and} \quad \Big\{\sup_{s<\delta}|X^{i,N}_s| \leq \delta^{-\frac{1}{2}}\log\log(\delta^{-1})\Big\},$$

and their complements. It is easily verified that both complements occur with probability $o(1)$ as $\delta \downarrow 0$, uniformly in $N \geq 1$. On the intersection of the two events, it follows from the previous observation (5.16) that $\inf_{s<\delta} Y^i_s$ must be smaller than $-\frac{1}{2}\varepsilon + \frac{\lambda}{2}\varepsilon + \delta C + 2\delta^{\frac{1}{2}}\log\log(\delta^{-1})$ for all $i \in \mathcal{J}$. Hence we can conclude from the final estimate in Step 1 and the claim (5.15) in Step 2 that

$$\mathbb{P}(L^N_\delta \geq \alpha^{-1}\varepsilon) \leq \mathbb{P}\Big(\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\left\{\inf_{s<\delta} Y^i_s \leq -\frac{1-\lambda}{2}\varepsilon + \delta C + 2\delta^{\frac{1}{2}}\log\log(\delta^{-1})\right\}} \geq \tfrac{\lambda-\lambda'}{2}\alpha^{-1}\varepsilon\Big) + o(1)$$

as $N \uparrow \infty$ and $\delta \downarrow 0$, where the $o(1)$ terms in δ are uniform in $N \geq 1$. Recall that $\varepsilon = \delta^{\frac{1}{2}}\log(\delta^{-1})$. Since ρ is bounded away from 1 (by Assumption 5.3.1), we can perform a time-change in each Y^i, and then the law of large numbers yields

$$\limsup_{N\to\infty}\mathbb{P}(L^N_\delta \geq \alpha^{-1}\varepsilon) \leq \mathbf{1}_{\left\{\Phi\left(-c[\frac{1-\lambda}{2}\log(\delta^{-1}) - \delta^{\frac{1}{2}}C - 2\log\log(\delta^{-1})]\right) \geq \frac{\lambda-\lambda'}{2}\alpha^{-1}\log(\delta^{-1})\delta^{\frac{1}{2}}\right\}} + o(1)$$

as $\delta \downarrow 0$, for some $c > 0$, where Φ is the standard normal cdf. Noting that Φ evaluated at the given point is of order $O(\exp\{-\text{const.} \times (\log\delta^{-1})^2\})$, the indicator eventually becomes zero as $\delta \downarrow 0$ and thus the proof is complete. $\square$

The minimality constraint (PJC) will be recovered from the following lemma, together with Proposition 5.3.5. The proof follows by the same arguments as in Proposition 5.3 of [46], but we provide the full details for completeness. The point is that, unlike the previous lemma, there is no need to isolate the effects of the common and idiosyncratic noise.

Lemma 5.3.7. *Fix $t < T$. There is a constant $C > 0$ such that, for every $\delta > 0$ sufficiently small,*

$$\mathbb{P}\big(\nu^N_{t-}([0, \alpha x + C\delta^{\frac{1}{3}}]) \geq x \quad \forall x \leq L^N_{t+\delta} - L^N_{t-} - C\delta^{\frac{1}{3}}\big) \geq 1 - C\delta,$$

whenever $N \geq \delta^{-\frac{1}{3}}$.

Proof. Note that $N(L^N_{t+\delta} - L^N_{t-})$ equals the number of particles that are absorbed in the interval $[t, t+\delta]$. Let $Y^i_{t,s} := \int_t^s \sigma(r)\rho(r)dB^0_r + \int_t^s \sigma(r)\sqrt{1-\rho(r)^2}dB^i_r$ and define

$$E^{i,k}_1 := \Big\{ X^{i,N}_{t-} - \alpha\frac{k}{N} - C\delta(1 + \sup_{s\leq t+\delta}|X^i_s|) - \sup_{s\in[t,t+\delta]}|Y^i_{t,s}| \leq 0,\ t \leq \tau^i \Big\}.$$

Then it must be the case that

$$\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{E^{i,k}_1} \geq \frac{k}{N} \quad \text{for} \quad k = 0, 1, \ldots, N(L^N_{t+\delta} - L^N_{t-}). \tag{5.17}$$

Fix an arbitrary $\lambda_0 \leq L^N_{t+\delta} - L^N_{t-} - 2\delta^{\frac{1}{3}}$ and set $k_0 := \big\lfloor N(\lambda_0 + 2\delta^{\frac{1}{3}})\big\rfloor \leq N(L^N_{t+\delta} - L^N_{t-})$. This way, (5.17) applies to k_0 and we have $\lambda_0 \geq \frac{k_0}{N} - 2\delta^{\frac{1}{3}}$ as well as $\frac{k_0}{N} \geq \lambda_0 + 2\delta^{\frac{1}{3}} - \frac{1}{N}$. Introducing the events

$$E^i_2 := \big\{ C\delta(1 + \sup_{s\leq t+\delta}|X^{i,N}_s|) + \sup_{s\in[t,t+\delta]}|Y^i_{t,s}| \geq \delta^{\frac{1}{3}} \big\},$$

it follows that, on $E^{i,k_0}_1 \cap (E^i_2)^{\complement}$, we have $X^i_{t-} - \alpha\frac{k_0}{N} \leq \delta^{\frac{1}{3}}$, and hence

$$X^i_{t-} - \alpha\lambda_0 \leq X^i_{t-} - \alpha\big(\frac{k_0}{N} - 2\delta^{\frac{1}{3}}\big) \leq (1 + 2\alpha)\delta^{\frac{1}{3}}.$$

Consequently,

$$\nu^N_{t-}([0, \alpha\lambda_0 + (1+2\alpha)\delta^{\frac{1}{3}}]) = \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{X^i_{t-} \leq \alpha\lambda_0 + (1+2\alpha)\delta^{\frac{1}{3}},\ t\leq\tau^i\}} \geq \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{E^{i,k_0}_1}\mathbf{1}_{(E^i_2)^{\complement}}.$$

Defining the event $E := \{\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{E_2^i} \le \delta^{\frac{1}{3}}\}$, we deduce that, on E,

$$\begin{aligned}\nu_{t-}^N([0,\alpha\lambda_0+(1+2\alpha)\delta^{\frac{1}{3}}]) &\ge \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{E_1^{i,k_0}} - \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{E_2^i} \ge \frac{k_0}{N} - \delta^{\frac{1}{3}} \\ &\ge \big(\lambda_0 + 2\delta^{\frac{1}{3}} - \frac{1}{N}\big) - \delta^{\frac{1}{3}} = \lambda_0 + \delta^{\frac{1}{3}} - \frac{1}{N}.\end{aligned}$$

Taking $N \ge \delta^{-\frac{1}{3}}$ implies $\delta^{\frac{1}{3}} - N^{-1} \ge 0$, so we can finally conclude that

$$\nu_{t-}^N([0,\alpha\lambda_0+(1+2\alpha)\delta^{\frac{1}{3}}]) \ge \lambda_0 \qquad \text{on } E, \tag{5.18}$$

for our arbitrary choice of $\lambda_0 \le L_{t+\delta}^N - L_{t-}^N - 2\delta^{\frac{1}{3}}$. It remains to observe that there exists $C > 0$, independent of δ, such that $\mathbb{P}(E^{\complement}) \le C\delta$. To see this, we can apply Markov's inequality twice, Cauchy–Schwarz and Burkholder–Davis–Gundy (to Y^i) to get

$$\begin{aligned}\mathbb{P}(E^{\complement}) &= \mathbb{P}\Big(\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{E_2^i} > \delta^{\frac{1}{3}}\Big) \le \delta^{-\frac{1}{3}}\frac{1}{N}\sum_{i=1}^N \mathbb{P}(E_2^i) \\ &\le c\delta^5\mathbb{E}\Big[(1+\sup_{s\le T}|X_s^{1,N}|)^8\Big] + c\delta^{-3}\mathbb{E}\Big[\sup_{s\in[t,t+\delta]}|Y_{t,s}^i|^8\Big] \le c_1\delta^5 + c_1\delta.\end{aligned}$$

Recalling (5.18) and taking $C := \max\{2c_1, 1+2\alpha\}$ finishes the proof. □

Finally, for reasons of continuity, we will need the following lemma, which is conveniently phrased at the level of a tagged particle. It captures the simple—yet essential—fact that the Brownian noise causes the particles to dip below the origin immediately after first hitting it.

Lemma 5.3.8. *For any $T > 0$ and $\delta > 0$, we have*

$$\lim_{\varepsilon\downarrow 0}\limsup_{N\to\infty}\mathbb{P}\Big(\inf_{s\le\delta\wedge(T-\tau^{1,N})}\big\{X_{s+\tau^{1,N}}^{1,N} - X_{\tau^{1,N}}^{1,N}\big\} > -\varepsilon,\ \tau^{1,N} < T\Big) = 0.$$

Proof. Since L^N is increasing, and using the linear bound on the drift, the probability in question is controlled by

$$\mathbb{P}\Big(\inf_{s\le\delta\wedge(T-\tau^{1,N})}\big\{sC\lambda + Y_{s+\tau^{1,N}}^1 - Y_{\tau^{1,N}}^1\big\} > -\varepsilon,\ \tau^{1,N} < T\Big) + \mathbb{P}\big(\sup_{s\le T}|X_s^{1,N}| > \lambda\big)$$

where $Y_s^1 := \int_0^s \sigma(r)\sqrt{1-\rho(r)^2}dB_r^1 + \int_0^s \sigma(r)\rho(r)dB_r^0$. By the strong Markov property, the first term is $o(1)$ as $\varepsilon \downarrow 0$ uniformly in $N \ge 1$ (for any fixed λ). Sending $\lambda \to \infty$, the second term vanishes (uniformly in $N \ge 1$) and hence the result follows. □

5.3.2 Compactness, Continuity, and Convergence

To establish the convergence of the finite particle system, we follow the ideas of [46]. In particular, we extend the particles from (5.8) to $[0, \bar{T}]$, for a fixed $\bar{T} > T$, by adding purely Brownian noise on $(T, \bar{T}]$. This amounts to replacing the empirical measures $\mathbf{P}^N$ by

$$\tilde{\mathbf{P}}^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{\tilde{X}^{i,N}} \quad \text{where} \quad \tilde{X}_t^{i,N} := \begin{cases} X_t^{i,N}, & t \in [0,T] \\ X_T^{i,N} + B_t^i - B_T^i, & t \in (T, \bar{T}]. \end{cases} \tag{5.19}$$

For simplicity of notation, we will drop the '$\sim$' and simply write $X_t^{i,N}$ and $\mathbf{P}^N$, with the understanding that $X_t^{i,N}$ is given by $\tilde{X}_t^{i,N}$ when we are working on the full interval $[0, \bar{T}]$. Notice that this construction automatically extends Proposition 5.3.6 to hold at the endpoint $\bar{T}$, and the choice of Brownian noise on $(T, \bar{T}]$ ensures the validity of Lemma 5.3.8 with $\bar{T}$ in place of T (of course, we are only actually interested in the dynamics of the system up to time T).

As in [46, Sec. 4.1], we endow $D_{\mathbb{R}} = D_{\mathbb{R}}[0, \bar{T}]$ with Skorokhod's M1 topology because it will allow us to circumvent irregularities in the loss, L^N, by virtue of its monotonicity. For properties of this topology we will be referring to [86, 105] (in [105] it is called the strong M1 topology and denoted SM_1). Importantly, these properties rely on the members of $D_{\mathbb{R}}$ being left-continuous at the terminal time, which is the reason for the continuous extension of the particle system to $\bar{T}$. Furthermore, we emphasise that $(D_{\mathbb{R}}, \mathrm{M1})$ is a Polish space [105, Thm. 12.8.1] and that its Borel sigma algebra is generated by the marginal projections [105, Thm. 11.5.2].

Similarly to [46, Lemmas 5.4 and 5.5], once we have a result such as Proposition 5.3.6 at both endpoints $t = 0$ and $t = \bar{T}$, the tightness of the empirical measures is an easy consequence of the properties of the M1 topology (Proposition 5.3.9). From Prokhorov's therorem [11, Thm. 5.1], we thus obtain a weakly convergent subsequence of the pair $(\mathbf{P}^N, B^0)$, and the goal is then to relate the resulting limit point to a solution of (rCMV).

Before seeking to characterize the limit point $(\mathbf{P}^*, B^0)$, the first task is to ensure that $L^N = \mathbf{P}^N(t \geq \hat{\tau})$ converges to $\mathbf{P}^*(t \geq \hat{\tau})$ whenever $\mathbf{P}^N$ converges $\mathbf{P}^*$. This is achieved through a continuity result (Lemma 5.3.10), which is an analogue of [46, Lemma 5.6, Prop. 5.8, Lemma 5.9] and its proof follows similarly, once we observe that it suffices to rely on Lemma 5.3.8, which is better suited to the common noise.

Finally, we need to characterize $\mathbf{P}^*$ as the conditional law of a process X satisfying the desired conditional McKean–Vlasov equation such that $(\mathbf{P}^*, B^0)$ is independent of the other Brownian motion B (which is constructed as part of the solution). For

this part, we rely on a martingale argument (Lemma 3.11 and Proposition 3.12) that differs from the approach in [46]—although still close in spirit—and which is better suited to a common noise. Another benefit of this argument is that it does not care about the form of the volatility σ, whereas the approach in [46] is tailored to a constant volatility. Aside from the independence result for $\mathbf{P}^*$ and B (Lemma 3.13), all our results extend to a (bounded and non-degenerate) Lipschitz continuous $x \mapsto \sigma(t,x)$, so without the common noise we can generalize the results of [46] in that sense, but we leave out the details of this. Instead, we push ahead and implement the plan outlined above, starting with the tightness of the empirical measures in the M1 topology.

Proposition 5.3.9 (Tightness of the empirical measures)**.** *Let $\mathfrak{T}^{wk}_{\mathrm{M1}}$ denote the topology of weak convergence on $\mathcal{P}(D_{\mathbb{R}})$ induced by the* M1 *topology on $D_{\mathbb{R}}$. Then the empirical measures $(\mathbf{P}^N)_{N\geq 1}$ are tight on $(\mathcal{P}(D_{\mathbb{R}}), \mathfrak{T}^{wk}_{\mathrm{M1}})$.*

Proof. Since $(D_{\mathbb{R}}, \mathrm{M1})$ is a Polish space, so is $(\mathcal{P}(D_{\mathbb{R}}), \mathfrak{T}^{wk}_{\mathrm{M1}})$. Therefore, by a classical result [103, Ch.I, Prop. 2.2], it suffices to verify that $(X^{1,N})_{N\geq 1}$ is tight on $(D_{\mathbb{R}}, \mathrm{M1})$. Let

$$H_{\mathbb{R}}(x_1, x_2, x_3) := \inf_{\lambda\in[0,1]} |x_2 - (1-\lambda)x_1 + \lambda x_3|.$$

Following [86, Sec. 4], we can deduce the tightness of $(X^{1,N})_{N\geq 1}$ by showing that

$$\mathbb{P}\big(H_{\mathbb{R}}(X^{1,N}_{t_1}, X^{1,N}_{t_2}, X^{1,N}_{t_3}) \geq \delta\big) \leq C\delta^{-4}|t_3 - t_1|^2 \tag{5.20}$$

for all $N \geq 1$, $\delta > 0$ and $0 \leq t_1 < t_2 < t_3 \leq \bar{T}$, along with

$$\lim_{\delta\downarrow 0} \lim_{N\to\infty} \mathbb{P}\Big(\sup_{t\in(0,\delta)} \big|X^{1,N}_t - X^{1,N}_0\big| + \sup_{t\in(\bar{T}-\delta,\bar{T})} \big|X^{1,N}_{\bar{T}} - X^{1,N}_t\big| \geq \varepsilon \Big) = 0 \tag{5.21}$$

for every $\varepsilon > 0$. For the first condition, observe that

$$H_{\mathbb{R}}(X^{1,N}_{t_1}, X^{1,N}_{t_2}, X^{1,N}_{t_3}) \leq |Z_{t_1} - Z_{t_2}| + |Z_{t_2} - Z_{t_3}| + \inf_{\lambda\in[0,1]} |L^N_{t_2} - (1-\lambda)L^N_{t_1} - \lambda L^N_{t_3}|,$$

where Z is given by

$$dZ_t = b(t, X^{1,N}_t)dt + \sigma(t)\sqrt{1-\rho(t)^2}dB^1_t + \sigma(t)\rho(t)dB^0_t.$$

Since L^N is increasing, the final term on the right-hand side is zero and hence (5.20) follows easily from Markov's inequality by controlling the increments of Z. Finally, (5.21) holds by virtue of Proposition 5.3.6 and the fact that it also applies at the artificial endpoint $\bar{T}$, as remarked above. □

Recall that each $\mathbf{P}^N$ is a random probability measure on $(D_{\mathbb{R}}, \mathcal{B}(D_{\mathbb{R}}))$, and that $L_t^N = \mathbf{P}^N(t \geq \hat{\tau})$, where $\hat{\tau}$ is the first hitting time of zero for the canonical process on $D_{\mathbb{R}}$. That is, written as a random variable on $D_{\mathbb{R}}$, we have $\hat{\tau}(\zeta) = \inf\{t > 0 : \zeta_t \leq 0\}$ for every $\zeta \in D_{\mathbb{R}}$. The next lemma provides a continuity result for L^N with respect to limit points of $\mathbf{P}^N$. It will play a central role in our further weak convergence analysis.

Lemma 5.3.10 (Continuity result for the loss)**.** *Suppose $\mathbf{P}^N$ converges weakly to $\mathbf{P}^*$ on $(\mathcal{P}(D_{\mathbb{R}}), \mathfrak{T}_{\mathrm{M1}}^{wk})$ and let $\mathbb{Q}^* := \mathrm{Law}(\mathbf{P}^*)$. At $\mathbb{Q}^*$-a.e. μ in $\mathcal{P}(D_{\mathbb{R}})$, the mapping $\mu \mapsto \mu(t \geq \hat{\tau})$ is continuous with respect to $\mathfrak{T}_{\mathrm{M1}}^{wk}$, for all $t \in \mathbb{T}^\mu \cap [0, \bar{T})$, where $\mathbb{T}^\mu$ is the set of continuity points of $t \mapsto \mu(t \geq \hat{\tau})$.*

Proof. First of all, we claim that, for any $\delta > 0$, we have

$$\mathbb{E}\big[\mathbf{P}^*\big(\hat{\tau} < \bar{T},\ \inf_{s \leq \delta \wedge (\bar{T} - \hat{\tau})} \{\eta_{s+\hat{\tau}} - \eta_{\hat{\tau}}\} = 0\big)\big] = 0. \tag{5.22}$$

To see this, notice that

$$\begin{aligned} &\mathbb{E}\big[\mathbf{P}^*\big(\hat{\tau} < \bar{T},\ \inf_{s \leq \delta \wedge (\bar{T} - \hat{\tau})} \{\eta_{s+\hat{\tau}} - \eta_{\hat{\tau}}\} = 0\big)\big] \\ &\quad \leq \lim_{\varepsilon \to 0} \liminf_{N \to \infty} \mathbb{E}\big[\mathbf{P}^N\big(\hat{\tau} < \bar{T},\ \inf_{s \leq \delta \wedge (\bar{T} - \hat{\tau})} \big\{\eta_{s+\hat{\tau}} - \eta_{\hat{\tau}}\big\} > -\varepsilon\big)\big] \\ &\quad = \lim_{\varepsilon \to 0} \liminf_{N \to \infty} \mathbb{P}\big(\tau^{1,N} < \bar{T},\ \inf_{s \leq \delta \wedge (\bar{T} - \tau^{1,N})} \big\{X^{1,N}_{s+\tau^{1,N}} - X^{1,N}_{\tau^{1,N}}\big\} > -\varepsilon\big). \end{aligned} \tag{5.23}$$

By Lemma 5.3.8, the right-hand side of (5.23) vanishes and hence the claim is indeed true. Consequently, almost every realization of $\mathbf{P}^*$ is supported on the subset of ζ's in $D_{\mathbb{R}}$ with $\zeta_0 > 0$ such that $\hat{\tau}(\zeta) < \bar{T}$ implies $\inf_{s \leq \delta \wedge (\bar{T} - \hat{\tau})}\{\zeta_{s+\hat{\tau}} - \zeta_{\hat{\tau}}\} < 0$ for any $\delta > 0$.

Consider now a sequence $\mu^N \to \mu^*$ in $(\mathcal{P}(D_{\mathbb{R}}), \mathfrak{T}_{\mathrm{M1}}^{wk})$, where μ^* is supported on the aforementioned subset. By the previous claim, the set of such μ^*'s has full measure under $\mathbb{Q}^* = \mathrm{Law}(\mathbf{P}^*)$. Applying Skorokhod's representation theorem [11, Thm. 6.7], there are càdlàg processes Z^N and Z (defined on the same background space) such that Z^N converges almost surely to Z in $(D_{\mathbb{R}}, \mathrm{M1})$ with

$$\mu^N(t \geq \hat{\tau}) = \mathbb{E}\big[\mathbf{1}_{(-\infty,0]}(\inf_{s \leq t} Z_s^N)\big] \quad \text{and} \quad \mu^*(t \geq \hat{\tau}) = \mathbb{E}\big[\mathbf{1}_{(-\infty,0]}(\inf_{s \leq t} Z_s)\big].$$

By [105, Thm. 12.4.1], we have $Z_0^N \to Z_0 > 0$ almost surely, so we immediately deduce from dominated convergence that $\mu^N(0 \geq \hat{\tau}) \to \mu^*(0 \geq \hat{\tau}) = 0$.

For $t > 0$, more work is needed. Let Ω_0 be a set of full probability on which $Z^N \to Z$ in the M1 topology. Then, by [105, Thm. 13.4.1], it holds for each $\omega \in \Omega_0$ that

there is a set of times $\mathbb{T}_\omega$ of full Lebesgue measure in $[0,\bar{T}]$ such that $\inf_{s\le t} Z^N_s(\omega) \to \inf_{s\le t} Z_s(\omega)$ for every $t \in \mathbb{T}_\omega$. Moreover, by the support of μ^*, it holds almost surely that Z becomes strictly negative immediately after first touching zero, so, for almost all $\omega \in \Omega_0$, we have $\mathbf{1}_{(-\infty,0]}(\inf_{s\le t} Z^N_s(\omega)) \to \mathbf{1}_{(-\infty,0]}(\inf_{s\le t} Z_s(\omega))$ for all $t \in \mathbb{T}_\omega$ except the first t for which $\inf_{s\le t} Z_s(\omega) = 0$. Consequently, dominated convergence implies the almost sure convergence

$$\int_I \mathbf{1}_{(-\infty,0]}(\inf_{s\le t} Z^N_s)dt \to \int_I \mathbf{1}_{(-\infty,0]}(\inf_{s\le t} Z_s)dt \quad \forall I \in \mathcal{B}(0,\bar{T}).$$

After taking expectations and using Tonelli, another application of dominated convergence yields

$$\int_I \mu^N(t \ge \hat{\tau})dt \to \int_I \mu^*(t \ge \hat{\tau})dt \quad \forall I \in \mathcal{B}(0,\bar{T}).$$

Let $\ell^N_t := \mu^N(t \ge \hat{\tau})$ and consider a truncation of $t \mapsto \ell^N_t$ given by $\ell^{N,\varepsilon}_t := \ell^N_{(t\vee\varepsilon)\wedge(\bar{T}-\varepsilon)}$. Then $\ell^{N,\varepsilon}_t = \mu^N(t \ge \hat{\tau})$ for $t \in [\varepsilon, \bar{T}-\varepsilon]$ and $(\ell^{N,\varepsilon})_{N\ge1}$ is automatically relatively compact in the M1 topology [105, Thm. 12.12.2], so we can pass to a convergent subsequence $\ell^{N_k,\varepsilon} \to \ell^\varepsilon$. By [105, Thm. 12.4.1], $\ell^{N_k,\varepsilon}$ converges pointwise to ℓ^ε on the co-countable set of continuity points for $t \mapsto \ell^\varepsilon_t$. In particular, dominated convergence yields

$$\int_I \mu^{N_k}(t \ge \hat{\tau})dt = \int_I \ell^{N_k,\varepsilon}_t dt \to \int_I \ell^\varepsilon_t dt \quad \forall I \in \mathcal{B}(\varepsilon, \bar{T}-\varepsilon).$$

Using the right-continuity, we thus deduce that $\ell^\varepsilon_t = \mu^*(t \ge \hat{\tau})$ everywhere on $(\varepsilon, \bar{T}-\varepsilon)$. Consequently, $\mu^{N_k}(t \ge \hat{\tau}) \to \mu^*(t \ge \hat{\tau})$ for every $t \in \mathbb{T}^{\mu^*} \cap (\varepsilon, \bar{T}-\varepsilon)$. Sending $\varepsilon \to 0$ (and noting that we would reach the same conclusion for any subsequence), we conclude that $\mu^N(t \ge \hat{\tau}) \to \mu^*(t \ge \hat{\tau})$ for all $t \in \mathbb{T}^{\mu^*} \cap [0,\bar{T})$. □

Using the tightness of $\mathbf{P}^N$ from Proposition 5.3.9, we can fix a limit point $\mathbb{P}^*$ of $\mathbb{P}^N := \mathrm{Law}(\mathbf{P}^N, B^0)$ on $\mathcal{P}(D_\mathbb{R}) \times C_\mathbb{R}$. From now on, we shall work with this particular limit point (keeping in mind that the arguments apply to any limit point) and, despite having passed to a subsequence, we shall simply write $\mathbb{P}^N \Rightarrow \mathbb{P}^*$. For concreteness, we define $\Omega^* := \mathcal{P}(D_\mathbb{R}) \times C_\mathbb{R}$ and introduce the random variables $\mathbf{P}^*(\mu, w) := \mu$ and $B^0(\mu, w) := w$ on the background space $(\Omega^*, \mathbb{P}^*, \mathcal{B}(\Omega^*))$. Note that the joint law of $(\mathbf{P}^*, B^0)$ is $\mathbb{P}^*$ with $\mathcal{B}(\Omega^*) = \sigma(\mathbf{P}^*, B^0)$. Given this, we define $L^* := \mathbf{P}^*(\,\cdot \ge \hat{\tau})$ along with the co-countable set of times

$$\mathbb{T} := \big\{t \in [0,\bar{T}] : \mathbb{P}^*(L^*_t = L^*_{t-}) = 1,\ \mathbb{E}^*[\mathbf{P}^*(\hat{\eta}_t = \hat{\eta}_{t-})] = 1\big\}. \tag{5.24}$$

By Proposition 5.3.6 and the proof of Proposition 5.3.9, we knew from the outset that L^N is tight in $D_{\mathbb{R}}$ on $[0,\bar{T}]$. By Lemma 5.3.10 and the continuous mapping theorem [11, Thm. 2.7], we can deduce that the finite dimensional distributions of its limit (along our fixed subsequence) agree with those of L^*, and hence we identify L^* as the limit. Of course, we are actually interested in the dynamics on $[0,T]$, so we note that we also have convergence at the process level in $D_{\mathbb{R}}$ on $[0,T_0]$, for any $T_0 \in [0,T] \cap \mathbb{T}$, since the restriction is continuous so long as T_0 is $\mathbb{P}^*$-almost surely a continuity point of L^*.

We now proceed to derive some further properties of the limiting law $\mathbb{P}^*$, which will ultimately enable us to construct a probability space that supports a solution to (rCMV). First of all, we define the map $\mathcal{M} : \mathcal{P}(D_{\mathbb{R}}) \times D_{\mathbb{R}} \to D_{\mathbb{R}}$ by

$$\mathcal{M}(\mu,\eta) := \eta - \eta_0 - \int_0^{\cdot} b(s,\eta_s)ds - \alpha\mu(\,\cdot \geq \hat{\tau}). \tag{5.25}$$

Fix an arbitrary choice of times $s_0, t_0 \in \mathbb{T} \cap [0,T)$ with $s_0 < t_0$ and $s_1,\dots,s_k \in [0,s_0] \cap \mathbb{T}$, for $\mathbb{T}$ as defined in (5.24), and let $F : D_{\mathbb{R}} \to \mathbb{R}$ be given by

$$F(\zeta) := (\zeta_{t_0} - \zeta_{s_0}) \prod_{i=1}^{k} f_i(\zeta_{s_i})$$

for arbitrary $f_1,\dots,f_k \in C_b(\mathbb{R})$. Based on this, we define the functionals

$$\begin{cases} \Psi(\mu) := \big\langle \mu, F(\mathcal{M}(\mu,\cdot))\big\rangle, \\ \Upsilon(\mu) := \big\langle \mu, F\big(\mathcal{M}(\mu,\cdot)^2 - \int_0^{\cdot}\sigma(s)^2 ds\big)\big\rangle, \ \text{ and} \\ \Theta(\mu,w) := \big\langle \mu, F\big(\mathcal{M}(\mu,\cdot)\times w - \int_0^{\cdot}\sigma(s)\rho(s)ds\big)\big\rangle. \end{cases}$$

As an application of Lemma 5.3.10, we obtain the following continuity result.

Lemma 5.3.11 (Functional continuity)**.** *For $\mathbb{P}^*$-almost every μ, we have $\Psi(\mu^n)$, $\Upsilon(\mu^n)$ and $\Theta(\mu^n,w^n)$ converging to $\Psi(\mu)$, $\Upsilon(\mu)$ and $\Theta(\mu,w)$ whenever $(\mu^n,w^n) \to (\mu,w)$ in $(\mathcal{P}(D_{\mathbb{R}}),\mathfrak{T}^{wk}_{\mathrm{M1}}) \times (C_{\mathbb{R}},\|\cdot\|_\infty)$ along a sequence for which $\langle \mu^n, |\mathcal{M}_t(\mu^n,\cdot)|^p\rangle$ is bounded uniformly in $n \geq 1$, for some $p > 2$, at any $t \geq 0$.*

Proof. By definition of $\mathbb{T}$, it holds for $\mathbb{P}^*$-almost every μ that $\mu(t \geq \hat{\tau}) = \mu(t- \geq \hat{\tau})$ and $\mu(\eta_t = \eta_{t-}) = 1$ for all $t \in \mathbb{T}$. Appealing to Lemma 5.3.10, we can thus restrict to a set of μ's with probability one under $\mathbb{P}^*$ such that $\mu^n(t \geq \hat{\tau})$ converges to $\mu(t \geq \hat{\tau})$ for $t = t_0$ and $t = s_0,\dots,s_k$. Fix any one of these μ's and suppose $(\mu^n,w^n) \to (\mu,w)$ in $(\mathcal{P}(D_{\mathbb{R}}),\mathfrak{T}^{wk}_{\mathrm{M1}}) \times (C_{\mathbb{R}},\|\cdot\|_\infty)$ with μ^n satisfying the above integrability assumption. We

start by showing that $\Psi(\mu^n) \to \Psi(\mu)$. Invoking Skorokhod's representation theorem [11, Thm. 6.7], we can write

$$\Psi(\mu^n) = \mathbb{E}[F(\mathcal{M}(\mu^n, Z^n))] \quad \text{and} \quad \Psi(\mu) = \mathbb{E}[F(\mathcal{M}(\mu, Z))],$$

where $Z^n \to Z$ almost surely in $(D_{\mathbb{R}}, \mathrm{M1})$ with $\mathrm{Law}(Z^n) = \mu^n$, $\mathrm{Law}(Z) = \mu$, and $\mathbb{P}(Z_t = Z_{t-})$ for $t = t_0, s_0, \ldots, s_k$. By standard properties of M1 convergence [105, Thm. 12.4.1], we have pointwise convergence at the continuity points of Z and, in particular, $Z^n_s \to Z_s$ for almost every $s \in [0, T]$. As Z^n is a convergent and hence bounded sequence, we deduce from the continuity of $b(s, \cdot)$ and its linear growth bound that, almost surely,

$$Z^n_t - Z^n_0 - \int_0^t b(s, Z^n_s)ds \;\to\; Z_t - Z_0 - \int_0^t b(s, Z_s)ds, \quad \text{for } t = t_0, s_0, \ldots, s_k, \tag{5.26}$$

by dominated convergence. In turn, we conclude that $F(\mathcal{M}(\mu^n, Z^n))$ converges almost surely to $F(\mathcal{M}(\mu, Z))$ in $\mathbb{R}$. It remains to observe that $F(\mathcal{M}(\mu^n, Z^n))$ is uniformly integrable, since $|F(\mathcal{M}(\mu^n, Z^n)| \leq |\mathcal{M}_{t_0}(\mu^n, Z^n)| + |\mathcal{M}_{s_0}(\mu^n, Z^n)|$, and $\mathbb{E}[|\mathcal{M}_t(\mu^n, Z^n)|^p]$ equals $\langle \mu^n, |\mathcal{M}_t(\mu^n, \cdot)|^p \rangle$, which is uniformly bounded for a power $p > 1$, by assumption. Hence the convergence of $\Psi(\mu^n)$ to $\Psi(\mu^n)$ follows from Vitali's convergence theorem. The proof is analogous for Υ and Θ. □

Define a new background space $\bar{\Omega} := \Omega^* \times D_{\mathbb{R}} = \mathcal{P}(D_{\mathbb{R}}) \times C_{\mathbb{R}} \times D_{\mathbb{R}}$ equipped with its Borel sigma algebra $\mathcal{B}(\bar{\Omega})$ and the probability measure $\bar{\mathbb{P}}$ given by

$$\bar{\mathbb{P}}(A) := \int_{\mathcal{P}(D_{\mathbb{R}}) \times C_{\mathbb{R}}} \mu(\{\eta : (\mu, w, \eta) \in A\}) d\mathbb{P}^*(\mu, w), \qquad \forall A \in \mathcal{B}(\bar{\Omega}). \tag{5.27}$$

For simplicity of presentation, we shall not distinguish notationally between random variables defined on Ω^* and their canonical extensions to $\bar{\Omega}$.

Proposition 5.3.12. *Let $\mathcal{M}$ be given by (5.25). Then $\mathcal{M}_\cdot$, $\mathcal{M}^2_\cdot - \int_0^\cdot \sigma(s)^2 ds$ and $\mathcal{M}_\cdot \times B^0_\cdot - \int_0^\cdot \sigma(s)\rho(s)ds$ are all continuous martingales on $(\bar{\Omega}, \bar{\mathbb{P}}, \mathcal{B}(\bar{\Omega}))$.*

Proof. Recall that $\mathbb{P}^N \Rightarrow \mathbb{P}^*$, where $\mathbb{P}^N$ is the law of $(\mathbf{P}^N, B^0)$. Using Skorokhod's representation theorem, we can find an almost surely convergent sequence $(\mathbf{Q}^N, B^N) \to (\mathbf{Q}^*, B^*)$ in $(\mathcal{P}(D_{\mathbb{R}}), \mathfrak{T}^{wk}_{M1}) \times (C_{\mathbb{R}}, \|\cdot\|_\infty)$ with $\mathbb{P}^N = \mathrm{Law}(\mathbf{Q}^N, B^N)$ and $\mathbb{P}^* = \mathrm{Law}(\mathbf{Q}^*, B^*)$. Notice that

$$\langle \mathbf{P}^N, |\mathcal{M}_t(\mathbf{P}^N, \cdot)|^p \rangle = \frac{1}{N} \sum_{i=1}^N \Big| \int_0^t \sigma(s)\rho(s) dB^0_s + \int_0^t \sigma(s)\sqrt{1 - \rho(s)^2} dB^i_s \Big|^p,$$

for any given $p > 2$, which converges with probability one (e.g. by the Birkhoff–Khinchin ergodic theorem) to a nice functional $G(B^0)$, for the fixed B^0 from the particle system. Since the laws of $(\mathbf{P}^N, B^0)$ and $(\mathbf{Q}^N, B^N)$ agree, for each $N \geq 1$, we deduce that $\langle \mathbf{Q}^N, |\mathcal{M}_t(\mathbf{Q}^N, \cdot)|^p \rangle - G(B^N)$ converges in probability, and using also that $B^N \to B^*$ almost surely, we can thus check that there is a subsequence for which $\langle \mathbf{Q}^N, |\mathcal{M}_t(\mathbf{Q}^N, \cdot)|^p \rangle$ is bounded with probability one. After passing to this subsequence, $(\mathbf{Q}^N, B^N)$ satisfies the assumptions of Lemma 5.3.11 with probability one, so we can deduce from the functional continuity that $\Psi(\mathbf{P}^N)$ converges in law to $\Psi(\mathbf{P}^*)$—along a subsequence still indexed by $N \geq 1$. The same argument applies to $\Upsilon(\mathbf{P}^N)$ and $\Theta(\mathbf{P}^N, B^0)$.

Next, we can observe that, for all $N \geq 1$,

$$\mathbb{E}\big[\Psi(\mathbf{P}^N)\big] = \mathbb{E}\Big[F\Big(\int_0^{\cdot} \sigma(s)\rho(s)dB_s^0 + \int_0^{\cdot} \sigma(s)\sqrt{1-\rho(s)^2}dB_s^1\Big)\Big] = 0.$$

Furthermore, $\mathbb{E}[|\Psi(\mathbf{P}^N)|^p]$ is clearly uniformly bounded in $N \geq 1$, for any given $p > 1$, so we have uniform integrability. Since $\Psi(\mathbf{P}^N)$ converges weakly to $\Psi(\mathbf{P}^*)$, as we argued above, the uniform integrability gives convergence of the means [11, Thm. 3.5] and so it holds by construction of $\bar{\mathbb{P}}$ that

$$\bar{\mathbb{E}}[F(\mathcal{M})] = \mathbb{E}^*[\Psi(\mathbf{P}^*)] = \lim_{N\to\infty} \mathbb{E}\big[\Psi(\mathbf{P}^N)\big] = 0.$$

By definition of F we can thus deduce that $\mathcal{M}$ is indeed a martingale under $\bar{\mathbb{P}}$. For the pathwise continuity of $t \mapsto \mathcal{M}_t$, note that

$$\bar{\mathbb{E}}[|\mathcal{M}_t - \mathcal{M}_s|^4] \leq \limsup_{N\to\infty} \mathbb{E}[\langle \mathbf{P}^N, |\mathcal{M}_t - \mathcal{M}_s|^4 \rangle] = \mathbb{E}\Big[\Big|\int_s^t \sigma(r)dW_r\Big|^4\Big] = O(|t-s|^2),$$

where W is a standard Brownian motion and the last equality follows from Burkholder–Davis–Gundy. By Kolmogorov's continuity criterion [52, Chp. 3, Thm. 8.8], we conclude that $\mathcal{M}$ has a continuous version. The proof is similar for the two other processes, using the convergence in law (along a subsequence) of $\Upsilon(\mathbf{P}^N)$ to $\Upsilon(\mathbf{P}^*)$ and of $\Theta(\mathbf{P}^N, B^0)$ to $\Theta(\mathbf{P}^*, B^0)$, respectively. □

Based on the previous proposition, we can now finalise the proof of Theorem 5.3.2.

Proof of Theorem 5.3.2. Fix the probability space $(\bar{\Omega}, \bar{\mathbb{P}}, \mathcal{B}(\bar{\Omega}))$ as introduced in (5.27) and recall that this includes fixing a limit point $(\mathbf{P}^*, B^0)$ of $(\mathbf{P}^N, B^0)$. Now define a càdlàg process X on $\bar{\Omega}$ by $X(\mu, w, \eta) := \eta$. Then it holds by construction of $\bar{\mathbb{P}}$ that

$$\bar{\mathbb{P}}\big(X \in A,\, (\mathbf{P}^*, B^0) \in S\big) = \int_S \mu(A)d\mathbb{P}^*(\mu, w),$$

where we recall that $\mathbb{P}^*$ is the joint law of $(\mathbf{P}^*, B^0)$. Consequently, we have

$$\bar{\mathbb{P}}(X \in A \,|\, \mathbf{P}^*, B^0) = \mathbf{P}^*(A), \qquad \forall A \in \mathcal{B}(D_{\mathbb{R}}),$$

as desired. Next, Proposition 5.3.12 gives that

$$\mathcal{M}_t = X_t - X_0 - \int_0^t b(s, X_s)ds - \alpha \mathbf{P}^*(\hat{\tau} \leq t)$$

is a continuous martingale with

$$\langle \mathcal{M}, \mathcal{M} \rangle_t = \int_0^t \sigma(s)^2 ds \qquad \text{and} \qquad \langle \mathcal{M}, B^0 \rangle_t = \int_0^t \sigma(s)\rho(s)ds.$$

By Lévy's characterisation theorem, we deduce that there exists a Brownian motion, B, that is independent of B^0 and for which

$$\mathcal{M}_t = \int_0^t \sigma(s)\rho(s)dB_s^0 + \int_0^t \sigma(s)\sqrt{1-\rho(s)^2}dB_s.$$

Furthermore, Lemma 5.3.13 below ensures B is independent of $(\mathbf{P}^*, B^0)$, so it follows that $(X, \mathbf{P}^*, B^0, B)$ is a solution to (rCMV).

Finally, by passing to a limit as $N \to \infty$ in Lemma 5.3.7 and then sending $\delta \downarrow 0$, we obtain ($\bar{\mathbb{P}}$-a.s.)

$$\Delta \hat{L}_t \leq \inf\{x > 0 : \nu^*_{t-}([0, \alpha x]) < x\}. \tag{5.28}$$

Since $\hat{L}$ is càdlàg by construction, Proposition 5.3.5 ensures that we have equality in (5.28) and thus the existence proof is complete. □

It remains to verify that $\mathbf{P}^*$ really is independent of the idiosyncratic Brownian motion, B, as constructed above.

Lemma 5.3.13. *In the proof of Theorem 5.3.2, we have* $(\mathbf{P}^*, B^0) \perp B$ *under* $\bar{\mathbb{P}}$.

Proof. By Assumption 5.3.1, $t \mapsto \sigma(t)\rho(t)$ is in $\mathcal{C}^\kappa(0,T)$ for some $\kappa > 1/2$. Hence the map

$$w \mapsto \int_0^\cdot \sigma(s)\rho(s)dw_s$$

makes sense as a Young integral for any Brownian path w [60, Thm. 6.8] and it agrees almost surely with the corresponding Itô integral against B^0 under $\mathbb{P}^*$. Moreover, if $(w^n) \subset \cap_{\kappa>1/2} \mathcal{C}^\kappa(0,T)$ converges uniformly to $w \in \cap_{\kappa>1/2} \mathcal{C}^\kappa(0,T)$, then the pathwise integrals also converge [60, Prop. 6.11-6.12], so by analogy with Lemma 5.3.11 we get $\Lambda(\mu^n, w^n) \to \Lambda(\mu, w)$ where

$$\Lambda(\mu, w) := \Big\langle \mu, F\Big(\mathcal{M}(\mu, \cdot) - \int_0^\cdot \sigma(s)\rho(s)dw_s\Big)\Big\rangle^2,$$

whenever $(\mu^n, w^n) \to (\mu, w)$, for $\mathbb{P}^*$-almost every (μ, w), provided μ^n satisfies the integrability assumption of Lemma 5.3.11 and w^n is a sequence of Brownian paths. By the same reasoning as in the proof of Proposition 5.3.12, a Skorokhod representation argument gives that it suffices to only have continuity of Λ along such sequences, in order to deduce that $\Lambda(\mathbf{P}^N, B^0)$ converges in law to $\Lambda(\mathbf{P}^*, B^0)$.

Next, we define the process

$$Y_t := \mathcal{M}_t - \int_0^t \sigma(s)\rho(s)dB_s^0 = \int_0^t \sigma(s)\sqrt{1-\rho(s)^2}dB_s.$$

By definition, we then have

$$\int_{\Omega^*} \big\langle \mathbf{P}^*(\omega), F\big(Y(\omega, \cdot)\big)\big\rangle^2 d\mathbb{P}^*(\omega) = \mathbb{E}^*[\Lambda(\mathbf{P}^*, B^0)].$$

It is immediate that $\Lambda(\mathbf{P}^N, B^0)$ is uniformly integrable, so the above mentioned convergence in law yields convergence of the means, and hence

$$\int_{\Omega^*} \big\langle \mathbf{P}^*(\omega), F\big(Y(\omega, \cdot)\big)\big\rangle^2 d\mathbb{P}^*(\omega) = \lim_{N\to\infty} \mathbb{E}\big[\Lambda(\mathbf{P}^N, B^0)\big] = 0, \tag{5.29}$$

where we have used that

$$\mathbb{E}[\Lambda(\mathbf{P}^N, B^0)] \le \frac{C}{N^2}\sum_{i=1}^N \mathbb{E}\bigg[\Big(\int_{s_0}^{t_0} \sigma(s)\sqrt{1-\rho(s)^2}dB_s^i\Big)^2\bigg] \to 0.$$

By (5.29), it holds for $\mathbb{P}^*$-a.e. $\omega \in \Omega^*$ that $\eta \mapsto Y(\omega, \eta)$ is a martingale under $\mathbf{P}^*(\omega)$. A similar argument, for

$$\Gamma(\mu, w) := \bigg\langle \mu, F\bigg(\Big(\mathcal{M}(\mu, \cdot) - \int_0^{\cdot} \sigma(s)\rho(s)dw_s\Big)^2 - \int_0^{\cdot} \sigma(s)^2(1-\rho(s)^2)ds\bigg)\bigg\rangle^2,$$

shows that the quadratic variation of $\eta \mapsto Y(\omega, \eta)$ is $\int_0^{\cdot} \sigma(s)^2(1-\rho(s)^2)ds$ and, as in the proof of Theorem 5.3.2, Levy's characterisation theorem then allows us to deduce that, for $\mathbb{P}^*$-a.e. $\omega \in \Omega^*$, the process $\eta \mapsto Y(\omega, \eta)$ is a time-changed Brownian motion under $\mathbf{P}^*(\omega)$ agreeing with the law of Y under $\bar{\mathbb{P}}$. In particular,

$$\begin{aligned}\bar{\mathbb{P}}\big(Y \in A,\, (\mathbf{P}^*, B^0) \in S\big) &= \int_S (\mathbf{P}^*(\omega))(\{\eta : Y(\omega, \eta) \in A\})d\mathbb{P}^*(\omega) \\ &= \int_S \bar{\mathbb{P}}(Y \in A)d\mathbb{P}^*(\omega) = \bar{\mathbb{P}}(Y \in A)\bar{\mathbb{P}}\big((\mathbf{P}^*, B^0) \in S\big).\end{aligned}$$

Finally, we can observe that $B = (B_t)_{t\in[0,T]}$ is given as the uniform limit in probability (under $\bar{\mathbb{P}}$) of Riemann sums of $\int_0^{\cdot}(\sigma(s)\sqrt{1-\rho(s)^2})^{-1}dY_s$. By the above, each of these finite sums are independent of $(\mathbf{P}^*, B^0)$ and hence so is B. This finishes the proof. $\square$

5.3.3 Construction of filtrations

This final section is concerned mainly with the construction of the filtrations discussed at the start of Section 3. In particular, we prove Proposition 5.3.3. As in the proof of Theorem 5.3.2 above, we continue to work with the background space $(\bar{\Omega}, \mathcal{F}, \mathbb{P})$ from (5.27), and the co-countable set of times $\mathbb{T}$ from (5.24), for a fixed limit point $(\mathbf{P}^*, B^0)$. Associated to this limit point, we have the absorbing marginal flow $t \mapsto \boldsymbol{\nu}_t^* := \mathbf{P}^*(\hat{\eta}_t \in \cdot\,,\, t < \hat{\tau})$, where, as usual, $\hat{\eta}$ denotes the canonical process on $D_{\mathbb{R}}$ and $\hat{\tau}$ is its first hitting time of zero.

Recall that $\bar{\Omega} = \mathcal{P}(D_{\mathbb{R}}) \times C_{\mathbb{R}} \times D_{\mathbb{R}}$ with its product Borel σ-algebra $\mathcal{F}$. Recall also that $X_t(\mu, w, \zeta) = \zeta_t$ and $B_t^0(\mu, w, \zeta) = w_t$, and note that $\boldsymbol{\nu}_t^*(\mu, w, \zeta) = \mu(\hat{\eta}_t \in \cdot\,,\, t < \hat{\tau})$. Finally, we can observe that, by virtue of these processes satisfying (rCMV), B_t will be adapted to any filtration that makes X_t, B_t^0, and $\langle \boldsymbol{\nu}_t^*, 1\rangle$ adapted.

Let us begin by constructing the desired filtration for just B^0 and $\boldsymbol{\nu}^* = (\boldsymbol{\nu}_t^*)_{t\geq 0}$. As formulated in Proposition 5.3.3, we want each marginal $\boldsymbol{\nu}_t^*$ to be measurable as an $\mathrm{M}(\mathbb{R})$-valued random variable, where the Borel-sigma algebra for $\mathrm{M}(\mathbb{R})$ comes from the duality with (a closed subspace of) the bounded continuous functions $C_b(\mathbb{R})$ in the uniform topology. Therefore, we will work with the filtration which, at time t, is generated by the projections $B^0 \mapsto B_s^0$ and $\boldsymbol{\nu}^* \mapsto \langle \boldsymbol{\nu}_s^*, \phi\rangle$ for all $s \leq t$ and all $\phi \in C_b(\mathbb{R})$. That is, we define

$$\mathcal{F}_t^{B^0,\boldsymbol{\nu}^*} := \sigma\Big((\mu, w) \mapsto \big(\langle \mu, \phi(\hat{\eta}_s)\mathbf{1}_{s<\hat{\tau}}\rangle, w_s\big) : \phi \in C_b(\mathbb{R}),\, s \leq t\Big)$$

The critical point is now to ensure that B^0 remains a Brownian motion in this filtration. The next lemma will allow us to deduce just that.

Proposition 5.3.14 (Convergence of the absorbing marginal flow)**.** *For any family of bounded continuous functions $\phi_1, \dots, \phi_k \in C_b(\mathbb{R})$ we have*

$$\big(\langle \boldsymbol{\nu}_{t_1}^N, \phi_1\rangle, \dots, \langle \boldsymbol{\nu}_{t_k}^N, \phi_k\rangle\big) \Rightarrow \big(\langle \boldsymbol{\nu}_{t_1}^*, \phi_1\rangle, \dots, \langle \boldsymbol{\nu}_{t_k}^*, \phi_k\rangle\big),$$

as $N \to \infty$, for any set of times $t_1 \dots, t_k \in \mathbb{T}$.

Proof. Fix an arbitrary $t \in \mathbb{T}$ and any $\phi \in C_b(\mathbb{R})$. We can then observe that

$$\langle \boldsymbol{\nu}_t^N, \phi\rangle = \langle \boldsymbol{\nu}_t^N, \phi_0\rangle, \quad \text{where} \quad \phi_0(x) := \begin{cases} \phi(x) & x \geq 0 \\ \phi(0) & x < 0 \end{cases},$$

and likewise for $\boldsymbol{\nu}_t^*$. Noting that $\phi_0(\hat{\eta}_{t\wedge\hat{\tau}}) = \phi_0(\hat{\eta}_t)\mathbf{1}_{t<\hat{\tau}} + \phi_0(\hat{\eta}_{\hat{\tau}})\mathbf{1}_{t\geq\hat{\tau}}$ with $\phi_0(\hat{\eta}_{\hat{\tau}}) = \phi(0)$, we therefore have the decomposition

$$\langle \boldsymbol{\nu}_t^N, \phi\rangle = \langle \mathbf{P}^N, \phi_0(\hat{\eta}_{t\wedge\hat{\tau}})\rangle - \phi(0)\mathbf{P}^N(t \geq \hat{\tau}),$$

and, again, the same is true for the limit $\boldsymbol{\nu}^*_t$. Hence the proposition will follow from the continuous mapping theorem if we can show that, given any $t \in \mathbb{T}$, for $\mathbb{P}^*$-almost every μ,

$$\langle \mu^n, \phi_0(\hat{\eta}_{t\wedge\hat{\tau}})\rangle \to \langle \mu, \phi_0(\hat{\eta}_{t\wedge\hat{\tau}})\rangle \quad \text{and} \quad \mu^n(t \geq \hat{\tau}) \to \mu(t \geq \hat{\tau}) \tag{5.30}$$

whenever $\mu^n \to \mu$ in $(\mathcal{P}(D_{\mathbb{R}}), \mathfrak{T}^{wk}_{\mathrm{M1}})$. Since $t \in \mathbb{T}$, we have $\mu(t \geq \hat{\tau}) = \mu(t- \geq \hat{\tau})$ for $\mathbb{P}^*$-almost every μ, so the claim is true for the second part of (5.30) by Lemma 5.3.10.

Noting that $t \in \mathbb{T}$ also implies $\mu(\hat{\eta}_t = \hat{\eta}_{t-})$ for $\mathbb{P}^*$-almost every μ, Skorokhod's representation theorem allows us to rewrite the first part of (5.30) as $\mathbb{E}[\phi(Z^n_{t\wedge\tau^n})] \to \mathbb{E}[\phi(Z_{t\wedge\tau})]$ whenever $Z^n \to Z$ a.s. in $(D_{\mathbb{R}}, \mathrm{M1})$ with $\mathbb{P}(Z_t = Z_{t-}) = 1$.

Since t is almost surely a continuity point of Z, there is an event Ω_t of probability one, for which $Z^n_t \to Z_t$ and $\inf_{s\leq t} Z^n_s \to \inf_{s\leq t} Z_s$, by [105, Thms. 12.4.1 & 13.4.1]. On the event $\Omega_t \cap \{\inf_{s\leq t} Z_s > 0\}$, we eventually have $\phi_0(Z^n_{t\wedge\tau^n}) = \phi_0(Z^n_t)$, which converges to $\phi_0(Z_t) = \phi_0(Z_{t\wedge\tau})$. Similarly, on the event $\Omega_t \cap \{\inf_{s\leq t} Z_s < 0\}$, we eventually have $\phi_0(Z^n_{t\wedge\tau^n}) = \phi(0)$, which agrees with $\phi_0(Z_{t\wedge\tau}) = \phi(0)$. For the remaining event $\Omega_t \cap \{\inf_{s\leq t} Z_s = 0\}$, it follows as in the proof of Theorem 5.3.10 that, up to a null set, $Z_t = 0$ and hence $\phi_0(Z_{t\wedge\tau}) = \phi_0(Z_t) = \phi(0)$. In principle, $\phi_0(Z^n_{t\wedge\tau^n})$ may oscillate between the values $\phi(0)$ and $\phi_0(Z^n_t)$, but the latter converges to $\phi_0(Z_t) = \phi(0)$. Combining the above, we have $\phi_0(Z^n_{t\wedge\tau^n}) \to \phi_0(Z_{t\wedge\tau})$ almost surely, so $\mathbb{E}[\phi(Z_{t\wedge\tau})]$ is indeed the limit of $\mathbb{E}[\phi(Z^n_{t\wedge\tau^n})]$, by dominated convergence. □

Let G be an arbitrary bounded continuous function $G : C_{\mathbb{R}} \to \mathbb{R}$. Fixing any $t_1, \ldots, t_n \in \mathbb{T} \cap [0, t]$ and $f_1, \ldots, f_n \in C_b(\mathbb{R})$, the previous proposition gives that

$$\begin{aligned}\bar{\mathbb{E}}\Big[G(B^0_{t+\cdot} - B^0_t) \prod_{i=1}^n f_i(\langle \boldsymbol{\nu}^*_{t_i}, \phi_i\rangle)\Big] &= \lim_{N\to\infty} \mathbb{E}\Big[G(B^0_{t+\cdot} - B^0_t) \prod_{i=1}^n f_i(\langle \boldsymbol{\nu}^N_{t_i}, \phi_i\rangle)\Big] \\ &= \bar{\mathbb{E}}\Big[G(B^0_{t+\cdot} - B^0_t)\Big]\bar{\mathbb{E}}\Big[\prod_{i=1}^n f_i(\langle \boldsymbol{\nu}^*_{t_i}, \phi_i\rangle)\Big]. \end{aligned} \tag{5.31}$$

Here the second equality follows from the simple observation that $(\boldsymbol{\nu}^N_s)_{s\leq t}$ is independent of $(B^0_{s+t} - B^0_t)_{s\geq 0}$, since the finite particle system has a unique strong solution with $(B^1, \ldots, B^N)$ independent of B^0. We deduce from (5.31) that the future increments of B^0 are independent of $\mathcal{F}^{B^0,\boldsymbol{\nu}^*}_t$ and so B^0 remains a Brownian motion in this filtration. Furthermore, the filtration is independent of B, since both B^0 and $\boldsymbol{\nu}^*$ are independent of B, as ensured by Lemma 5.3.13.

Finally, we need a larger filtration $\mathcal{F}_t$ for which the full solution $(X, B, B^0, \boldsymbol{\nu}^*)$ is adapted. This can be achieved by also including the pre-images of the projections

$X \mapsto X_s$. Therefore, we define

$$\mathcal{F}_t := \sigma\Big((\mu, w, \zeta) \mapsto \big(\langle \mu, \phi(\hat{\eta}_s)\mathbf{1}_{s<\hat{\tau}}\rangle, w_s, \zeta_s\big) : \phi \in C_b(\mathbb{R}),\, s \leq t\Big).$$

From our construction of the solution $(X, B, B^0, \boldsymbol{\nu}^*)$ on $(\bar{\Omega}, \bar{\mathbb{P}}, \mathcal{F})$, we note that we can write $B_t(\mu, w, \zeta)$ in terms of ζ_s, w_s, and $\langle \mu, \phi(\hat{\eta}_s)\mathbf{1}_{s<\hat{\tau}}\rangle$ with $\phi \equiv 1$, for $s \leq t$, so B_t is $\mathcal{F}_t$-measurable. By analogy with (5.31), at any time $t \geq 0$, the future increments of B and B^0 are independent of $\mathcal{F}_t$. Thus, they are both Brownian motions with respect to this filtration. This completes the proof of Proposition 5.3.3.

5.3.4 Remark on numerics

The density functions shown in Figure 5.2.1 were generated according to the following quadrature-based algorithm (and similarly for Figure 4.2.1 in the previous chapter). Begin with an initial probability density function, V_0, and fix step sizes δt and δx for creating a mesh on the time-space grid $[0, T] \times [0, u]$. Here, u an upper bound chosen large enough so that truncating the solution above level u in the spatial component is an acceptable approximation to the process on the half-line.

At each point on the spatial slice of the grid at time δt, approximate the density at that point by evolving the solution according heat equation with drift term given by the common noise increment $B^0_{\delta t} - B^0_0$ (i.e. numerically evalutate the integral corresponding to this transformation at all points on the slice). This gives a candidate density $U_{\delta t}$. Now estimate the candidate loss $L^{(0)}_{\delta t} = 1 - \int_0^\infty U_{\delta t}(x)dx$ numerically, and then iterate the process $L^{(n+1)}_{\delta t} = L^{(0)}_{\delta t} + \int_0^{\alpha L^{(n)}_{\delta t}} U_{\delta t}(x)dx$ until some stopping rule is reached. In our case, the stopping rule we use is to stop when $L^{(n+1)}_{\delta t} - L^{(n)}_{\delta t} < \varepsilon$ for the first time, for a given threshold $\varepsilon > 0$.

The final value in the sequence $L^{(n)}_{\delta t}$ is our approximation, $L_{\delta t}$, to the loss over the first time step. We then translate the solution by $\alpha L_{\delta t}$ in the direction of the boundary (by re-interpolating $U_{\delta t}$ onto the shifted grid) to give our approximation $V_{\delta t}$ of the density after the first time-step. This process is then repeated across all time-steps. We conjecture that the numerical method described here is convergent, but this remains an open problem.

Chapter 6

Instantaneous Feedback ⁎ Part III: Global Results on Uniqueness and Regularity

In this chapter we study uniqueness of the McKean–Vlasov problem

$$\begin{cases} X_t = X_0 + Z_t - \alpha f(L_t) \\ \tau = \inf\{t \geq 0 : X_t \leq 0\} \\ L_t = \mathbb{P}(\tau \leq t), \end{cases} \tag{MV}$$

where $X_0 \in (0, \infty)$ is a random start point, Z is an independent continuous stochastic process, $\alpha \in \mathbb{R}$ is the feedback parameter, and $f : [0, 1] \to \mathbb{R}$ is a continuous function. By a solution to this problem we mean an increasing càdlàg function $L : [0, \infty) \to [0, 1]$ that satisfies (MV) and is initially zero. As usual, we denote by ν_0 the law of X_0 and the corresponding density is denoted by V_0.

6.1 Introduction

Recently, variants of (MV) have been studied in [74, 75, 87, 94, 95] motivated by the study of contagion in large financial markets, and in [45, 46, 79] motivated by nonlinear integrate-and-fire models for large networks of electrically coupled neurons. In this respect, our main interest is the case $Z_t = \rho\beta_t + \sqrt{1-\rho^2}B_t$, where B_t is a standard Brownian motion and β_t is a fixed Brownian path: uniqueness of the deterministic L in (MV) would then give pathwise uniqueness of the stochastic L in the conditional McKean–Vlasov problem

$$\begin{cases} X_t = X_0 + \sqrt{1-\rho^2}B_t + \rho B_t^0 - \alpha L_t \\ \tau = \inf\{t \geq 0 : X_t \leq 0\} \\ L_t = \mathbb{P}(\tau \leq t \mid B^0), \end{cases} \tag{CMV}$$

where B and B^0 are independent Brownian motions. In Chapter 5, 'relaxed' solutions to (CMV), for which the adaptedness of L to B^0 is relaxed, are obtained as limit points of the following finite particle system: N particles move according to Brownian motions correlated through B^0, except that when a particle hits the origin, it is absorbed, and this then has a contagious feedback effect that causes all the other particles to jump down by α/N with $\alpha > 0$ — possibly leading to more particles being absorbed and hence further rounds of downward jumps.

If, as in the previous chapters, each particle measures the 'distance-to-default' of a financial entity, and absorption at zero corresponds to default, then such a feedback loop could model a *cascade* of bankruptcies caused by the interplay between default contagion ($\alpha > 0$) and common exposures ($\rho > 0$).

Remark 6.1.1 (A possible model of indirect contagion). To illustrate some flexibility in the interpretation of the above mentioned particle system, we note that a variant could serve as a possible model for indirect contagion caused by the forced selling of a common asset in the spirit of [35]. In this case, each of the N particles would model the log leverage ratio of a bank—defined as the log of capital over assets—for which a minimum value is enforced by regulation. When reaching some lower threshold above this minimum, the distressed bank must sell assets to increase their leverage ratios, and if part of these sales pertain to common illiquid assets ($\rho > 0$), then it will depress the price of these, hence causing a drop in the leverage ratios of all the other banks by α/N (with $\alpha > 0$). Note, however, that a bank reaching the threshold should now be reset to some higher pre-specified leverage ratio (after the selling of assets) and then continue to evolve instead of defaulting. Thus, there is an analogous feedback effect from absorptions (when hitting the barrier), which pushes down all the particles, however—instead of actually being absorbed—the particles are immediately restarted at a pre-specified value. Naturally, the selling of the common asset should lower the extent of common exposures, which could be achieved by letting ρ depend on the analogue of L (as was considered in earlier chapters), but we leave this out here as it is not the focus of this chapter.

Blow-ups and the physical jump condition

As a result of the positive feedback loops, for large enough $\alpha > 0$, the solutions to (CMV) can develop jump discontinuities—which we call *blow-ups* as in the previous chapters (following [74, 87]). In order to study uniqueness in these cases, we recall that it is necessary to resolve ambiguity at the blow-ups (see e.g. Figure 4.3.1 and

the accompanying discussion in Chapter 4). As usual, this is achieved by specifying that the appertaining jump sizes must satisfy the physical jump condition

$$\Delta L_t = \inf\{x > 0 : \nu_{t-}([0, \alpha x]) < x\}, \tag{PJC}$$

with probability 1, where $\nu_t(S) := \mathbb{P}(X_t \in S,\, t < \tau \mid B^0)$. We reiterate that (PJC) is the correct specification of jump sizes for two reasons: First, in the case of (CMV), the condition gives the minimal jump sizes that are necessary for L to be càdlàg, by Proposition 4.2.2 in Chapter 4. Secondly, the solutions constructed in Chapter 5 (and [46]) are obtained as limit points of finite particle systems as described above, for which the corresponding discrete version of (PJC) gives the only sensible physical description of the jump behaviour, and this discrete condition then yields (PJC) in the limit. It should also be noted that generic (possibly non-physical) solutions to (MV) can be constructed directly from a generalised Schauder fixed point theorem as proved by Nadtochyi and Shkolnikov [95] under suitable conditions on Z and f. For (MV), the condition (PJC) should be understood in terms of $\nu_t = \mathbb{P}(X_t \in \cdot,\, t < \tau)$, corresponding to a pathwise realisation of B^0 in (CMV).

With the physical jump condition in place, it is then natural to only consider initial densities that do not want to jump immediately (in alignment with L being càdlàg and initially zero). Indeed, it is natural to restrict solely to states that could be reached by the evolving system and, recalling that (PJC) is a rule for the left limit ν_{t-}, this translates as those initial conditions ν_0 satisfying

$$\inf\{x > 0 : \nu_0([0, \alpha x]) < x\} = 0, \tag{6.1}$$

in the case $f(x) = x$ that is. While this condition is necessary for càdlàg solutions under Brownian drivers as in (CMV), it is possible to construct càdlàg solutions to (MV) violating this condition for more general Z (see [95, (3.15)]). Specifically, one can take $Z_t := B_t + A_t$ with $A_t := \alpha\mathbb{P}(X_0 + \inf_{s<t} B_s \leq 0)$ and $f(x) = x$. If X_0 is such that (6.1) is positive, and it is not imposed that positivity in (6.1) should cause L to jump immediately, then $L := A$ gives a continuous solution to (MV) despite violating (6.1).

In this chapter, the only mathematical contributions concerning jumps will be for the case when Z is a standard Brownian motion and $f(x) = x$ (see Section 6.3). If considering a generic feedback function f, the condition corresponding to (PJC) is simply

$$\Delta L_t = \inf\{x > 0 : \nu_{t-}([0, \alpha \cdot (f(x + L_{t-}) - f(L_{t-}))]) < x\}.$$

6.1.1 Recent history of the problem

Aside from the earlier mentioned existence results [87, 95], the literature on (MV) is centred around the case where Z is a standard Brownian motion, B, up to various forms of drift and volatility. For clarity, we thus focus on $Z_t = B_t$ and $f(x) = x$ in this subsection.

Over the last few years, the study of uniqueness and regularity for (MV) has been continued through independent efforts of Nadtochiy and Shkolnikov [94] and the author of this thesis together with Ledger and Hambly [74] (these are the results presented in Chapter 4). Here is a brief overview of the developments:

- Let $X_0 = x_0 > 0$. By [45] there is an $\alpha_0 \in (0, 1]$ such that, for any $\alpha \in (0, \alpha_0)$, (MV) has a C^1 solution L on any time-interval, and the solution is unique in this class. The result is formulated for the (neuronal) version where particles are reset upon hitting the boundary.

- In the same setting as [45], for any $\alpha > 0$, [46] obtains global solutions to (MV) as limit points of the (neuronal) particle system described earlier. Moreover, there is propagation of chaos provided there is uniqueness among the limit points.

- Let $V_0 \in H^1_0(0, \infty)$. For any $\alpha > 0$, [94] gives a solution L to (MV) up to an explosion time $t_\star > 0$ such that $L' \in L^2(0, t)$ for $t < t_\star$ and $\|L'\|_{L^2(0,t)}$ explodes as $t \uparrow t_\star$. Moreover, there is uniqueness up to $t_\star$ in this class of solutions.

- Let $V_0 \in L^\infty$ and $V_0(x) \leq Cx^\beta$ for some $C, \beta > 0$. For any $\alpha > 0$, [74] gives a solution L to (MV) up to an explosion time $t_\star > 0$ such that $L' \in L^2(0, t)$ for $t < t_\star$ and $\|L'\|_{L^2(0,t)}$ explodes as $t \uparrow t_\star$. Moreover, there is uniqueness up to $t_\star$ among all candidate solutions, and $|L'_t| \leq K_0 t^{-\frac{1-\beta}{2}}$ on $[0, t_0]$ for any $t_0 < t_\star$.

Note that the final result can be combined with that of the second bullet point to give propagation of chaos for (CMV, $\rho = 0$) up to the explosion time (or globally for small $\alpha > 0$). Based on the above, numerical schemes for (CMV, $\rho = 0$) have been proposed and analysed by Kaushansky and Reisinger [83] and Kaushansky, Lipton and Reisinger [82], up to the explosion time.

6.1.2 Overview of the chapter

In Section 6.2 we prove global uniqueness of (MV) under a smallness condition on the contagion parameter $\alpha > 0$, i.e. (6.2), which defines the weak feedback regime. The result applies with any continuous stochastic process, Z, as the driver. In Section 6.3 we specialise to (MV) with $Z_t = B_t$ and $f(x) = x$, but for a general $\alpha > 0$, and we show that the comparison argument from Section 6.2 can be extended to give short-time uniqueness after a blow-up. Finally, we devote Section 6.4 to a discussion of what can be said about the global behaviour of the idiosyncratic model—that is, with $Z_t = B_t$ and $f(x) = x$—in the presence of blow-ups.

Concerning the latter, we mention the very recent preprint [47] by Delarue, Nadtochiy and Shkolnikov, which—building on the analysis in [45, 94, 74]—provided the first results on global uniqueness and regularity for this problem. Unless we explicitly indicate otherwise, the arguments in Section 6.4 were developed independently of [47], however, we stress that the contents of Section 6.4 remain unpublished at the time of writing and, in particular, they did not form part of the earlier preprint [88] which contained the results from Sections 6.1–6.3 of this chapter.

6.2 Uniqueness in the weak feedback regime

In this section we prove uniqueness of solutions to (MV) in the *weak feedback regime*, that is, when the feedback parameter, $\alpha > 0$, is sufficiently small relative to the initial condition and the feedback function f (Theorem 6.2.2).

Our main method of proof is the simple comparison argument in Lemma 6.2.1. It is important to recognise that this method is of 'zeroth order', in the sense that it makes no use of differential or analytic properties of the solutions, L, as a function of time or of the initial density, V_0, as a function of space. This is very natural probabilistically, as the McKean–Vlasov formulation (MV) does not involve any derivatives and always makes sense globally, in contrast to the PDE point of view discussed in Section 1.1.

The value of a zeroth order approach becomes clear when there is a rough drift in the driving noise, Z, as the analytical tools developed in [45, 74, 94] cannot be applied to such a setting. For example, path by path, the solutions to (CMV) will be non-differentiable when $\rho > 0$, yet the uniqueness argument presented below is robust enough to include (CMV) in the weak feedback regime (Theorem 6.2.3). In fact, this is the first uniqueness result for any version of (MV) with a rough drift. Moreover, the robustness of our approach will turn out to be very useful even without a rough drift, when attacking the question of uniqueness after a blow-up in Section 3.

With a view to what follows, it is convenient to introduce the notation

$$\|\ell\|_t := \sup_{s\le t}|\ell(s)| \quad \text{and} \quad \|f\|_{\mathrm{Lip}(x)} := \sup_{y\neq z\in[0,x]}\frac{|f(y)-f(z)|}{|y-z|}.$$

For now, our first task is the essential comparison argument alluded to above. A visualisation of its proof is provided by Figure 6.2.1

Lemma 6.2.1 (Comparison)**.** *Suppose L and $\bar{L}$ are two solutions to (MV) with the same initial condition ν_0. Then*

$$|L_t-\bar{L}_t| \le \mathbb{E}[\nu_0(\sup_{s\le t}\{\alpha f(L_s)-Z_s\}, \sup_{s\le t}\{\alpha f(\bar{L}_s)-Z_s\})] \\ \vee \mathbb{E}[\nu_0(\sup_{s\le t}\{\alpha f(\bar{L}_s)-Z_s\}, \sup_{s\le t}\{\alpha f(L_s)-Z_s\})].$$

Proof. Couple the solutions via the Brownian motion and random start point:

$$\begin{cases} X_t = X_0 + Z_t - \alpha f(L_t) \\ \tau = \inf\{t\ge 0 : X_t \le 0\} \\ L_t = \mathbb{P}(\tau \le t), \end{cases} \qquad \begin{cases} \bar{X}_t = X_0 + Z_t - \alpha f(\bar{L}_t) \\ \bar{\tau} = \inf\{t\ge 0 : \bar{X}_t \le 0\} \\ \bar{L}_t = \mathbb{P}(\bar{\tau} \le t), \end{cases}$$

then we have

$$\begin{aligned} L_t - \bar{L}_t &\le \mathbb{P}(\inf_{s\le t} X_s \le 0, \inf_{s\le t}\bar{X}_s > 0) \\ &= \mathbb{P}(\inf_{s\le t}\{X_0 + Z_s - \alpha f(L_s)\} \le 0, \inf_{s\le t}\{X_0 + Z_s - \alpha f(\bar{L}_s)\} > 0) \\ &= \mathbb{P}(\sup_{s\le t}\{\alpha f(L_s) - Z_s\} \le X_0 < \sup_{s\le t}\{\alpha f(\bar{L}_s) - Z_s\}). \end{aligned}$$

Conditioning on X_0 gives

$$\begin{aligned} L_t - \bar{L}_t &\le \mathbb{E}\int_0^\infty \mathbf{1}\{\sup_{s\le t}\{\alpha f(L_s)-Z_s\} \le x_0 < \sup_{s\le t}\{\alpha f(\bar{L}_s)-Z_s\}\}\nu_0(dx_0) \\ &= \mathbb{E}[\nu_0(\sup_{s\le t}\{\alpha f(L_s)-Z_s\}, \sup_{s\le t}\{\alpha f(\bar{L}_s)-Z_s\})]. \end{aligned}$$

Notice that the same inequality with L and $\bar{L}$ interchanged also holds, by symmetry, therefore we have the result. □

From here the main uniqueness result is almost immediate. Note that we refer to the *weak feedback regime*, as the uniqueness is guaranteed through a smallness condition on $\alpha \ge 0$, namely (6.2) below.

Theorem 6.2.2 (Uniqueness in the weak feedback regime)**.** *Suppose that ν_0 has a density satisfying $V_0 : (0,\infty) \to (0,\infty)$ and let L and $\bar{L}$ be two solutions to (MV). If*

$$\alpha \cdot \|V_0\|_\infty \cdot \|f\|_{\mathrm{Lip}(L_t \vee \bar{L}_t)} < 1, \tag{6.2}$$

then $L = \bar{L}$ on $[0,t]$.

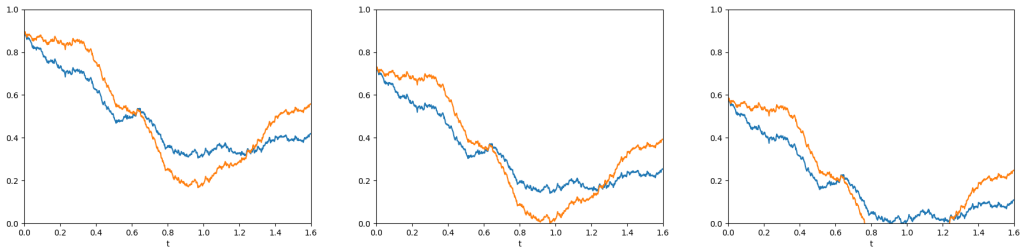


Figure 6.2.1: A visual proof of Lemma 6.2.1 with $f(x) = x$: The red curve is X^1 and the blue curve is X^2, coupled through the same initial value and driving noise. Writing $L_t^1 - L_t^2 = \mathbb{P}(\tau_1 \leq t < \tau_2)$, we see that, as x_0 decreases, we only get contributions to this probability from when the red curve first hits the x-axis until the blue curve first hits it. Hence the probability equals $\nu_0(0.59, 0.72)$. Note that the difference between the curves is bounded by $\|L^1 - L^2\|_t$.

Proof. Let L and $\bar{L}$ be two solutions to (MV). Applying the inequality $\bar{L}_s = L_s + \bar{L}_s - L_s \geq L_s - \|L - \bar{L}\|_r$, for $s \leq r$, to Lemma 6.2.1 gives

$$\begin{aligned}|L_r - \bar{L}_r| \leq \mathbb{E}[\nu_0(-\alpha\|f(L) - f(\bar{L})\|_r + \sup_{s\leq r}\{\alpha f(L_s) - Z_s\}, \sup_{s\leq r}\{\alpha f(L_s) - Z_s\})] \\ \vee \mathbb{E}[\nu_0(-\alpha\|f(L) - f(\bar{L})\|_r + \sup_{s\leq r}\{\alpha f(\bar{L}_s) - Z_s\}, \sup_{s\leq r}\{\alpha f(\bar{L}_s) - Z_s\})].\end{aligned}$$

Writing $Z_r^* := \sup_{s\leq r}\{\alpha f(L_s) - Z_s\}$ and $\bar{Z}^*$ for the same supremum corresponding to $\bar{L}$, the above becomes

$$|L_r - \bar{L}_r| \leq \mathbb{E}\Big[\int_{Z_r^* - \alpha\|f(L)-f(\bar{L})\|_r}^{Z_r^*} V_0(x)dx\Big] \vee \mathbb{E}\Big[\int_{\bar{Z}_r^* - \alpha\|f(L)-f(\bar{L})\|_r}^{\bar{Z}_r^*} V_0(x)dx\Big],$$

and so it is immediate that, for all $r \geq 0$,

$$|L_r - \bar{L}_r| \leq \alpha\|V_0\|_\infty\|f(L) - f(\bar{L})\|_r.$$

Using the Lipschitz property of f and taking a supremum over $r \in [0, t]$ gives

$$\|L - \bar{L}\|_t \leq \alpha\|V_0\|_\infty\|f\|_{\mathrm{Lip}(L_t \vee \bar{L}_t)}\|L - \bar{L}\|_t.$$

Therefore the hypothesis in the statement forces $\|L - \bar{L}\|_t = 0$, as required. □

Recall the notation $\nu_t = \mathbb{P}(X_t \in \cdot\,, t < \tau)$. If the driver Z_t has a density p_t, then ν_t has a density V_t with $\|V_t\|_\infty \leq \|V_0\|_\infty$. Indeed, ignoring the absorption at the origin,

$$\nu_t(S) \leq \int_0^\infty \int_S p_t(x_0 + z - \alpha f(L_t))V_0(x_0)dzdx_0 \leq \|V_0\|_\infty \cdot |S|. \tag{6.3}$$

Hence the smallness condition (6.2) gives $\Delta L_t = 0$ by virtue of the physical jump condition, so the solution remains continuous for all time in the weak feedback regime.

As immediate corollaries of Theorem 6.2.2 we obtain the following two results which are the first contributions to the question of uniqueness for the conditional McKean–Vlasov problem (CMV) from [87] and for a generic continuous driving noise in (MV) as considered in [95].

Theorem 6.2.3 (Uniqueness of CMV in the weak feedback regime)**.** *The solutions to the conditional McKean–Vlasov problem (CMV) are pathwise unique when* $\|V_0\|_\infty < \alpha^{-1}$*. Furthermore, this ensures that the 'relaxed' solutions constructed in [87, Thm. 3.2] are in fact adapted to the systemic driver* B^0*, and thus it implies full convergence in law of the finite particle system from [87] to the unique solution of (CMV).*

Compared with Section 1.1, (CMV) gives rise to a stochastic PDE, namely

$$dV_t(x) = \tfrac{1}{2}\partial_{xx}V_t(x)dt + \alpha\partial_x V_t(x)dL_t - \rho\partial_x V_t(x)dB_t^0$$

with $L_t = 1 - \int_0^\infty V_t(x)dx$, understood in the weak sense for test functions $\phi \in C^2(0,\infty)$ with $\phi(0) = 0$ [87, Rmk. 3.13]. In the weak feedback regime this SPDE makes sense globally and, by virtue of Theorem 6.2.3, the solution is adapted to the driving noise.

Theorem 6.2.4 (Small-time uniqueness in [95])**.** *Suppose* $\|V_0\|_\infty < \alpha^{-1}$*. The McKean–Vlasov system (MV) with* $f(x) = -\log(1-x)$*, as defined in [95, Thm. 2.2], has a unique solution for all* $t \geq 0$ *such that* $L_t < 1 - \alpha\|V_0\|_\infty$*.*

We should note that Theorem 6.2.2 is the most straightforward application of Lemma 6.2.1. Indeed, it is possible to obtain stronger results with harder work and a more specific setting. The next section gives one such example.

6.3 Short-time uniqueness after a blow-up

For the remaining part of this paper we return to the setting $Z_t = B_t$ and $f(x) = x$. In this case, there is full uniqueness of (MV) up to an explosion time $t_\star > 0$ [74, Thm. 1.8], but while solutions continue to exist there are currently no results on whether the system restarts in a unique way.

More specifically, the current approaches to uniqueness break down, as the system may be restarting from a density $V_{t_\star}$ with mass at the boundary where, in principle, all we know is that (PJC) imposes

$$\inf\{x > 0 : \textstyle\int_0^{\alpha x} V_{t_\star}(y)dy < x\} = 0. \tag{6.4}$$

The problem here is that (6.4) implies little about the regularity of $V_{t_\star}$ near zero. Without further information, we cannot rule out pathological cases like those in Figure 6.3.1, so it seems difficult to gain sufficient control to prove that uniqueness can be propagated after a blow-up. In practise, however, we do not expect these edge cases to arise, and indeed we are able to show here that we have at least polynomial control on the density after a jump time of L (Proposition 6.3.2 and Theorem 6.3.3). As the next result shows, this will be sufficient to give small time uniqueness after the blow-up. The idea is to use the methods from the previous section, but to ensure that insufficient time has passed for a large amount of mass to reach the boundary.

Theorem 6.3.1 (Short-time uniqueness)**.** *Suppose ν_0 has a density, V_0, for which there exists $c > 0$, $x_0 > 0$ and $n \in \mathbb{N}$ such that*

$$V_0(x) \leq \alpha^{-1} - cx^n, \qquad \text{for all } x < x_0. \tag{6.5}$$

Let L and $\bar{L}$ be two solution to (MV) with $Z_t = B_t$ and $f(x) = x$. Then there exists $t_0 > 0$ such that $L_t = \bar{L}_t$ for all $t \in [0, t_0]$.

Proof. For shorthand let $Z_t^* := \sup_{s \leq t}\{\alpha L_s - B_s\}$ and

$$F(t) := \mathbb{E}\nu_0([-\alpha\|L - \bar{L}\|_t + Z_t^*, Z_t^*)) \leq E(t) \cdot \alpha\|L - \bar{L}\|_t,$$

where $E(t) := \mathbb{E}[\sup_{x \in [-\alpha\|L-\bar{L}\|_t + Z_t^*, Z_t^*)} V_0(x)]$. We decompose this expectation into the three regions:

$$Z_t^* \in [0, \alpha\|L - \bar{L}\|_t + t), \qquad Z_t^* \in [\alpha\|L - \bar{L}\|_t + t, x_0), \qquad Z_t^* \in [x_0, \infty).$$

Since L and $\bar{L}$ are càdlàg, we can take $t_1 > 0$ sufficiently small so that for all $t \leq t_1$ we have $\alpha\|L - \bar{L}\|_t + t \leq x_0/2$ and $\alpha L_t, \alpha \bar{L}_t \leq x_0/4$. Therefore

$$\begin{aligned}
E(t) &\leq \alpha^{-1}\mathbb{P}(Z_t^* \in [0, \alpha\|L - \bar{L}\|_t + t]) + (\alpha^{-1} - ct^n)\mathbb{P}(Z_t^* \in [\alpha\|L - \bar{L}\|_t + t, x_0]) \\
&\qquad + \|V_0\|_\infty \mathbb{P}(Z_t^* \in [x_0, \infty)) \\
&\leq \alpha^{-1} - ct^n\mathbb{P}(Z_t \in [x_0/2, x_0]) + \|V_0\|_\infty \mathbb{P}(Z_t^* \in [x_0, \infty)) \\
&\leq \alpha^{-1} - ct^n\mathbb{P}(\sup_{s \leq t} B_s \in [\tfrac{1}{2}x_0, \tfrac{3}{4}x_0]) + \|V_0\|_\infty \mathbb{P}(\sup_{s \leq t} B_s \in [\tfrac{3}{4}x_0, \infty)) \\
&= \alpha^{-1} - ct^n(\Phi(-x_0/2t^{1/2}) - \Phi(-3x_0/4t^{1/2})) + \|V_0\|_\infty \Phi(-3x_0/4t^{1/2}),
\end{aligned}$$

where Φ is the normal cdf. Using the asymptotic bounds

$$\Phi(-cx) \asymp x^{-1}e^{-c^2x^2/2}, \qquad \text{as } x \to \infty,$$

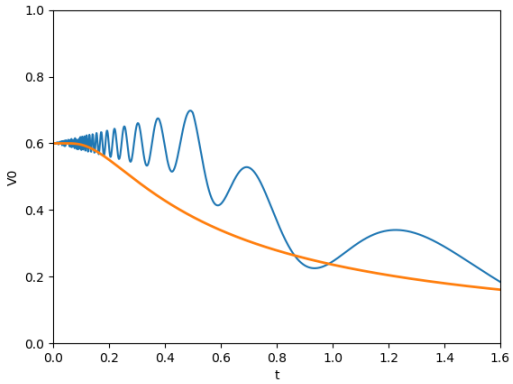


Figure 6.3.1: Two pathological initial densities that satisfy (6.4) with $\alpha = 5/3$. The blue curve takes infinitely many values above the critical value of $3/5$ near zero. The red curve remains strictly below $3/5$, but all derivatives vanish at the origin, so control of the form in (6.5) is not possible.

we can find $t_2 \in (0, t_1]$ sufficiently small so that $E(t) < \alpha^{-1}$ for all $t \leq t_2$, which guarantees that $F(t) < \|L - \bar{L}\|_t$. By symmetry and Lemma 6.2.1, we conclude that

$$|L_t - \bar{L}_t| < \|L - \bar{L}\|_t, \qquad \text{for every } t \in (0, t_2]. \tag{6.6}$$

Let $t_0 > 0$ be the smaller of the first jump times of L and $\bar{L}$ and t_2. By continuity there exists $t_\star \in [0, t_0]$ for which the supremum is attained:

$$|L_{t_\star} - \bar{L}_{t_\star}| = \|L - \bar{L}\|_{t_\star}$$

If $t_\star = 0$, then $\|L - \bar{L}\|_{t_0} = 0$ and we are done. Otherwise $t_\star \in (0, t_0]$, but this contradicts (6.6) at $t = t_\star$. □

The next result allows us to propagate Theorem 6.3.1 to blow-up times. The main idea is that, although we cannot control the solution density near zero at arbitrary times, we can prove that the density is analytic in the interior of the half-line. At a blow-up this is then sufficient to give control at zero, since the new point at the origin was in the interior of the density before the jump discontinuity. Our proof of analyticity rely on kernel smoothing and energy estimate techniques from [73, 75], and is presented in the next section.

Proposition 6.3.2 (Analyticity)**.** *Suppose ν_0 has a density $V_0 \in L^2(0, \infty)$ and V_t is the corresponding density process for a solution of (MV). Then for all $t > 0$ and $x > 0$, $y \mapsto V_t(y)$ is analytic at the point x.*

Theorem 6.3.3 (Short-time uniqueness after a blow-up)**.** *Suppose ν_0 has a density $V_0 \in L^2(0,\infty)$. Let $t_\star$ be the first blow-up time of the solution. Then the system can be restarted at time $t_\star$ and the restarted solution is the unique solution on an additional non-zero time interval, amongst all possible càdlàg solutions on this period.*

Proof. Note that, at the first blow-up time $t_\star$, we have $V_{t_\star}(x) = V_{t_\star -}(x + \alpha \Delta L_{t_\star})$, for $x \geq 0$. So although $V_{t_\star -}$ is not analytic at $x = 0$, because $\Delta L_{t_\star} > 0$ and $V_{t_\star -}$ is analytic in the interior, $V_{t_\star}$ is analytic at $x = 0$. Therefore we have a series expansion

$$V_{t_\star}(x) = V_{t_\star}(0) + \sum_{n \geq 1} c_n x^n, \qquad \text{for every } x \in [0, x_0],$$

for some $x_0 > 0$. If $V_{t_\star}(0) < \alpha^{-1}$ we have the required condition on $V_{t_\star}$ by taking x_0 sufficiently small.

If $V_{t_\star}(0) > \alpha^{-1}$, then, since the physical jump condition (PJC) on ΔL ensures that $V_{t_\star}$ satisfies (6.4), by taking x sufficiently small we have a contradiction. Therefore suppose $V_{t_\star}(0) = \alpha^{-1}$, then by the last case we cannot have $c_n = 0$ for all $n \geq 1$, so let $n_0 := \min\{n : c_n \neq 0\}$. For $x_1 > 0$ sufficiently small we have

$$V_{t_\star}(x) \geq \alpha^{-1} + (c_{n_0} + \varepsilon) x^{n_0}, \qquad \text{for every } x \in [0, x_1],$$

where $|\varepsilon| \leq \frac{1}{2}|c_{n_0}|$. Again if $c_{n_0} > 0$ then we contradict (6.4), hence $c_{n_0} < 0$ and so we have that $V_{t_\star}$ satisfies the condition (6.5) in Theorem 6.3.1. Consequently, Theorem 6.3.1 can be applied up to a small time after the first blow-up time $t_\star$. □

Although we get small-time uniqueness after a blow-up, it is important to note that global uniqueness is out of reach of the techniques presented in this section, as they do not immediately allow for bootstrapping at continuity times. Simply put, we pay a price for the robustness of our comparison arguments, by way of not obtaining any regularity. This translates into a lack of control over the boundary behaviour of the densities, and hence we cannot rule out the appearance of pathological states like those in Figure 6.3.1 at continuity times. However, combined with the regularity results from Chapter 4 we can indeed give a fairly precise characterization of the global behaviour, although we must appeal to [47] to complete the picture (see Section 6.4).

Let us briefly remark here that, starting from an initial density satisfying (6.5), and simply ignoring the absorption at the origin, it is easy to establish a zeroth order estimate $\nu_t(0,x) \leq \alpha^{-1}x - Cx^k$ for some $C > 0$ and $k \in \mathbb{N}$ near $x = 0$ on a small time interval following the initial time. In particular, there can be no jump discontinuities on this interval, which is essentially saying that—after a blow-up—we are back in the weak feedback regime, locally, for a small enough period of time. This is the starting point for our analysis in Section 6.4.

6.3.1 Analyticity of the density in the interior

In this final subsection we present a proof of Proposition 6.3.2. As already mentioned, it goes via kernel smoothing, so for any $\delta > 0$ and any measure μ on $(0, \infty)$, we define

$$T_\delta \mu(x) := \int_0^\infty G_\delta(x_0, x)\mu(dx_0) \quad \text{and} \quad T_\delta^r \mu(x) := \int_0^\infty G_\delta^r(x_0, x)\mu(dx_0),$$

for $x \geq 0$, where G_δ and G_δ^r are, respectively, the absorbing and reflecting Gaussian kernels on the positive half-line, given by

$$G_\delta(x, y) := \frac{1}{\sqrt{2\pi\delta}}\Big\{e^{-\frac{(x-y)^2}{2\delta}} - e^{-\frac{(x+y)^2}{2\delta}}\Big\}, \quad G_\delta^r(x, y) := \frac{1}{\sqrt{2\pi\delta}}\Big\{e^{-\frac{(x-y)^2}{2\delta}} + e^{-\frac{(x+y)^2}{2\delta}}\Big\}.$$

We will sometimes abuse notation and apply T_δ and T_δ^r to $\phi \in C^\infty(0, \infty)$.

Lemma 6.3.4 (Existence of weak derivatives)**.** *Suppose* $\liminf_{\delta\to 0} \|\partial_x^n T_\delta \mu\|_2 < \infty$*, then* μ *has an* n^{th} *weak derivative* $\partial_x^n \mu \in L^2(0, \infty)$ *and* $\|\partial_x^n T_\delta \mu\|_2 \to \|\partial_x^n \mu\|_2$ *as* $\delta \to 0$*.*

From [74, Prop. 2.1], we know that regardless of the initial condition, for all times $t > 0$, ν_t has a density function $V_t : (0, \infty) \to (0, \infty)$. A similar argument shows that:

Lemma 6.3.5. *If* ν_0 *has a density* $V_0 \in L^2(0, \infty)$*, then* $\|V_t\|_2 \leq \|V_0\|_2$ *for every* $t \geq 0$*.*

Let $\langle \nu_t, \phi \rangle := \int \phi d\nu_t = \mathbb{E}[\phi(X_t)\mathbf{1}_{t<\tau}]$ and let $t_\star$ be the first jump time of L. Strictly before $t_\star > 0$, Itô's formula gives the weak PDE

$$\langle \nu_t, \phi \rangle = \langle \nu_0, \phi \rangle + \frac{1}{2}\int_0^t \langle \nu_s, \partial_{xx}\phi \rangle ds - \alpha \int_0^t \langle \nu_s, \partial_x \phi \rangle dL_s,$$

for $\phi \in C^2(0, \infty)$ with $\phi(0) = 0$. Now take $\phi = T_\delta \psi$ for an arbitrary $\psi \in C_0^\infty(0, \infty)$. Integrating by parts, and differentiating n times, we can deduce that

$$d\partial_x^n T_\delta \nu_t(x) = \tfrac{1}{2}\partial_x^{n+2} T_\delta \nu_t(x)dt + \alpha \partial_x^{n+1} T_\delta^r \nu_t(x) dL_t.$$

for $x \geq 0$ (a.e.). Using $d(\partial_x^n T_\delta \nu_t)^2 = 2\partial_x^n T_\delta \nu_t(x) d\partial_x^n T_\delta \nu_t(x)$, and rearranging, we get

$$\begin{aligned} d\partial_x^n T_\delta \nu_t(x)^2 = \partial_x^n T_\delta \nu_t(x)\partial_x^{n+2} T_\delta \nu_t(x)dt + 2\alpha \partial_x^n T_\delta \nu_t(x)\partial_x^{n+1} T_\delta \nu_t(x) dL_t \\ + 4\alpha \partial_x^n T_\delta \nu_t(x)\partial_x^{n+1} R_\delta \nu_t(x) dL_t, \end{aligned} \tag{6.7}$$

where we have introduced the remainder term

$$R_\delta \nu_t(x) := \int_0^\infty \frac{1}{\sqrt{2\pi\delta}} e^{-\frac{(x+x_0)^2}{2\delta}} \nu_t(dx_0).$$

Fix any two open sets $U \Subset W \Subset (0, \infty)$, where '$\Subset$' denotes compact containment. Let ζ be a smooth cut-off function with $\zeta = 1$ on U, $\zeta \in (0, 1)$ on $W \setminus U$, and $\zeta = 0$ otherwise. Note that $|\partial_x \zeta| + |\partial_{xx}\zeta| \leq C\mathbf{1}_{W\setminus U}$, where C only depends on W and U.

Lemma 6.3.6 (Smoothed energy estimate)**.** *For all integers $a \geq 2$ and $b \geq 1$ we have*

$$t_\star^b \|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_{t_\star -}\|_2^2 + \int_0^{t_\star} t^b \|\zeta^{\frac{a}{2}} \partial_x^{n+1} T_\delta \nu_t\|_2^2 dt \\ \leq c_1 a(a-1) \int_0^{t_\star} t^b \|\zeta^{\frac{a-2}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dt + b \int_0^{t_\star} t^{b-1} \|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dt + o(1),$$

as $\delta \to 0$, where $t_\star > 0$ the first jump time of L.

Proof. Multiplying (6.7) by ζ^a and using the integration-by-parts formulae

$$\int g \cdot f \cdot f' = -\int g|f|^2 \quad \text{and} \quad \int g \cdot f \cdot f'' = -\int g|f'|^2 + \frac{1}{2}\int g''|f|^2,$$

we get

$$d\|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_t\|_2^2 + \|\zeta^{\frac{a}{2}} \partial_x^{n+1} T_\delta \nu_t\|_2^2 dt = \tfrac{1}{2}\||\partial_{xx}\zeta^a|^{\frac{1}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dt - 2\alpha \|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dL_t \\ + 4\alpha \Big(\int_0^\infty \zeta^a(x) \partial_x^n T_\delta \nu_t(x) \partial_x^{n+1} R_\delta \nu_t(x) dx \Big) dL_t.$$

Since $|\partial_x \zeta|, |\partial_{xx}\zeta| \leq C$, Young's inequality with small enough parameter $\eta > 0$ gives

$$d\|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_t\|_2^2 + \|\zeta^{\frac{a}{2}} \partial_x^{n+1} T_\delta \nu_t\|_2^2 dt \\ \leq c_1 a(a-1) \|\zeta^{\frac{a-2}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dt + 4\alpha\eta^{-1} \|\zeta^{\frac{a}{2}} \partial_x^{n+1} R_\delta \nu_t\|_2^2 dL_t,$$

where we have used $a\zeta^{(a-1)/2} \leq a(a-1)\zeta^{(a-2)/2}$. Differentiating $t \mapsto t^b \|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_t\|_2^2$ with respect to t and taking the range of integration up to $t_\star$ gives

$$t_\star^b \|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_{t_\star -}\|_2^2 + \int_0^{t_\star} t^b \|\zeta^{\frac{a}{2}} \partial_x^{n+1} T_\delta \nu_t\|_2^2 dt + b \int_0^{t_\star} t^{b-1} \|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dt \\ \leq c_1 a(a-1) \int_0^{t_\star} t^b \|\zeta^{\frac{a-2}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dt + 4\alpha\eta^{-1} \int_0^{t_\star} t^b \|\zeta^{\frac{a}{2}} \partial_x^{n+1} R_\delta \nu_t\|_2^2 dL_t.$$

It remains to show the final term above vanishes as $\delta \to 0$. Writing $p_\delta(\cdot)$ for the standard Gaussian transition kernel, it follows from the definition of R_δ that

$$|\partial_x^{n+1} R_\delta \nu_t(x)| \leq \Big| \partial_x^{n+1} \int_0^\infty p_\delta(x+y) \nu_t(dy) \Big| \leq \int_0^\infty |Q(\delta^{-\frac{1}{2}} y, \delta^{-\frac{1}{2}} x)| p_\delta(x+y) \nu_t(dy),$$

for a two-variable polynomial Q. Note $p_\delta(x+y) \leq p_\delta(y) \cdot e^{-x^2/2\delta}$. Applying Lemma 6.3.4 and Cauchy–Schwarz, we conclude $|\partial_x^{n+1} R_\delta \nu_t(x)| = O(e^{-x^2/4\delta})$ uniformly in t. Since ζ is supported on W,

$$\|\zeta^{\frac{a}{2}} \partial_x^{n+1} R_\delta \nu_t\|_2 = O(e^{-w^2/2\delta})$$

uniformly in t, where $w := \inf W > 0$. This is sufficient to complete the proof. □

As Lemma 6.3.6 gives control over the $(n+1)^{\text{th}}$ spatial derivative in terms of the n^{th} derivative, we can use it to inductively prove that $V_{t_\star -}$ has weak derivatives of all orders, and is therefore smooth in the interior of $(0,\infty)$.

Corollary 6.3.7 (Smoothness)**.** *If ν_0 has a density $V_0 \in L^2(0,\infty)$, then $V_{t_\star -} \in C^\infty(0,\infty)$. Furthermore $V_t \in C^\infty(0,\infty)$ for almost all $t \in (0,t_\star)$ and*

$$t_\star^b \|\zeta^{\frac{a}{2}} \partial_x^n V_{t_\star -}\|_2^2 + \int_0^{t_\star} t^b \|\zeta^{\frac{a}{2}} \partial_x^{n+1} V_t\|_2^2 dt \\ \leq c_1 a(a-1) \int_0^{t_\star} t^b \|\zeta^{\frac{a-2}{2}} \partial_x^n V_t\|_2^2 dt + b \int_0^{t_\star} t^{b-1} \|\zeta^{\frac{a}{2}} \partial_x^n V_t\|_2^2 dt.$$

Proof. Fix $n \geq 0$. Suppose that for all $U \Subset W \Subset (0,\infty)$, $a \geq 2n$, and $b \geq n$ we have

$$\liminf_{\delta \to 0} \int_0^{t_\star} t^b \|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dt < \infty. \tag{6.8}$$

Then Fatou's Lemma and Lemma 6.3.4 imply that $\partial_x^n V_t$ exists and is in $L^2(U)$ for every $U \Subset (0,\infty)$ and almost all $t \in (0,t_\star)$, with

$$\int_0^{t_\star} t^b \|\zeta^{\frac{a}{2}} \partial_x^n T_\delta \nu_t\|_2^2 dt \to \int_0^{t_\star} t^b \|\zeta^{\frac{a}{2}} \partial_x^n V_t\|_2^2 dt < \infty.$$

Therefore taking a $\liminf$ over δ in Lemma 6.3.6 gives that (6.8) holds for $n+1$, for any $a \geq 2(n+1)$, $b \geq n+1$. Since (6.8) holds for $n = 0$ and all $a, b \geq 0$, by Lemma 6.3.4, using induction we conclude that the statement holds for all $n \geq 0$. Returning to Lemma 6.3.6 once more, we deduce that

$$\liminf_{\delta \to 0} \|\partial_x^n T_\delta \nu_{t_\star -}\|_{L^2(U)}^2 \leq \liminf_{\delta \to 0} \|\zeta^n \partial_x^n T_\delta \nu_{t_\star -}\|_2^2 < \infty,$$

for every $n \geq 0$. Therefore $V_{t_\star -}$ has weak derivatives of all orders, so is smooth. Finally, sending $\delta \to 0$ in Lemma 6.3.6 gives the inequality. □

Finally we are in a position to complete the proof of Proposition 6.3.2 by using the inequality in Corollary 6.3.7 to show that $x \mapsto V_{t_\star -}(x)$ is analytic for all $x \in (0,\infty)$.

Proof of Proposition 6.3.2. Introduce the short-hand notation

$$I(n,a,b) := \int_0^{t_\star} t^b \|\zeta^{\frac{a}{2}} \partial_x^n V_t\|_2^2 dt,$$

where we recall the cut-off ζ is defined for fixed $U \Subset W \Subset (0,\infty)$. Corollary 6.3.7 implies

$$I(n+1,a,b) \leq c_1 a(a-1) I(n,a-2,b) + b I(n,a,b-1) \leq (c_1 t_\star a(a-1) + b) I(n,a-2,b-1).$$

Iterating the argument gives

$$I(n, 2n, n) \leq I(0, 0, 0) \prod_{1 \leq i \leq n} (c_1 t_\star 2i(2i-1) + i) \leq C^n \cdot (2n)!,$$

for $C > 0$ a constant depending on V_0 and ζ. Returning to Corollary 6.3.7 we have

$$t_\star^b \|\zeta^{\frac{a}{2}} \partial_x^n V_{t_\star -}\|_2^2 \leq (c_1 t_\star a(a-1) + b) I(n, a-2, b-1).$$

Therefore setting $a = 2n + 2$ and $b = n + 1$ gives

$$\|\partial_x^n V_{t_\star -}\|_{L^2(U)}^2 \leq t_\star^{-(n+1)} (c_1 t_\star (2n+2)(2n+1) + n + 1) \cdot C^n \cdot (2n)! = (C')^n \cdot (2n)!,$$

for a further constant $C' > 0$ also depending on $t_\star$. Since $(2n)! \leq 4^n (n!)^2$, we conclude that

$$\|\partial_x^n V_{t_\star -}\|_{L^2(U)} \leq (C'')^n \cdot n!, \qquad \text{for every } n \geq 0. \tag{6.9}$$

Morrey's inequality [53, Sect. 5.6, Thm. 4] gives a constant $c_2 > 0$ such that

$$\|\partial_x^n V_{t_\star -}\|_{L^\infty(U)} \leq c_2 \|\partial_x^n V_{t_\star -}\|_{H^1(U)} \leq (C''')^n \cdot n!,$$

where we have applied (6.9), for $C''' > 0$ a further constant independent of n. This inequality guarantees that $V_{t_\star -}$ is analytic in the interior of U, and therefore, since $U \Subset (0, \infty)$ was arbitrary, $V_{t_\star -}$ is analytic at all strictly positive spatial points. □

6.4 Global behaviour of the loss with blow-ups

We begin by reiterating that this section, like the one before, only covers the special case of (MV) with $Z_t = B_t$ and $f(x) = x$. Concerning this particular problem, we state two theorems that provide (i) a characterization of the global regularity of loss, and (ii) a result on global uniqueness. These results are part of ongoing work, and, unfortunately, they still only provide partial answers to the conjectures from Chapter 4. Be that as it may, the purpose of these theorems is not to be as complete as possible, but rather to highlight how the analysis in this thesis compares with that of the recent preprint [47]. In particular, we only provide brief sketches of the proofs based on previous results and the two propositions in Sections and below.

The next theorem characterizes the global behaviour of a given global—possibly non-unique—physical solution to the idiosyncratic McKean–Vlasov problem (MV) driven by Brownian motion. Before stating the result, we recall that the continuity times $t > 0$—of a given solution with density V_t—are precisely those times for which

$$\inf\Big\{x > 0 : \int_0^{\alpha x} V_{t-}(y)dy < x\Big\} = 0,$$

since the physical jump condition then gives $\Delta L_t = 0$ and hence $V_t = V_{t-}$.

Theorem 6.4.1. *Suppose that, at all times of continuity $t \geq 0$, we have $V_t(x) \leq \alpha^{-1}$ in a small neighbourhood $(0, x_t)$. (Note that this can only be violated if $x \mapsto V_t(x)$ oscillates around α^{-1} with an accumulation point at $V_t(0+) = \alpha^{-1}$). Then, at any given time $t_0 \geq 0$, the loss $t \mapsto L_t$ is continuously differentiable on $(t_0, t_{explode})$, where $t_{explode} > t_0$ is the next time its H^1 norm explodes. Moreover, L is guaranteed to be $\frac{1}{2}$-Hölder continuous at the potentially singular time t_0 provided (i) $\Delta L_{t_0} = 0$ with $V_{t_0-}(0+) < \alpha^{-1}$, or (ii) $\Delta L_{t_0} > 0$ with $V_{t_0}(0+) < \alpha^{-1}$ after resolving the jump.*

Proof. Fix an arbitrary $t_0 \geq 0$. Suppose first that the physical jump condition gives $\Delta L_{t_0} = 0$ and hence $V_{t_0} = V_{t_0-}$. Then the assumption of the theorem gives that $V_{t_0}(x) \leq \alpha^{-1}$ on $(0, x_0)$, for some $x_0 > 0$. Applying Proposition 6.4.3, we thus have that $V_t(x) < \alpha^{-1}$ on $(0, x_1)$, for all $t \in (t_0, t_1]$, for some $x_1 > 0$ and $t_1 > t_0$. Now fix any $t \in (t_0, t_1]$. Then, by the previous observation, Proposition 6.4.4 gives that $s \mapsto L_s$ is $\frac{1}{2}$-Hölder on $[t, t + \epsilon]$ for some ϵ. Based on this, one can replicate the Krylov–Safanov argument from [47, Prop. 2.5] to show that $V_s(x) \leq Cx^\beta$ near zero for $s \in [t, t + \epsilon']$, for suitable constants. Finally, as in [47, Prop. 3.2], we can now invoke the results from Chapter 4 (Theorems 4.2.7 and 4.2.8) to see that there is a continuously differentiable solution up to the next explosion of the H^1 norm after the arbitrary time $t \in (t_0, t_1]$, since we are restarting from precisely the kind of initial density that is considered in Chapter 4. Concerning point (i) of the second statement, notice that, if $V_{t_0} < \alpha^{-1}$, then Proposition 6.4.4 gives $\frac{1}{2}$-Hölder continuity on a closed interval $[t_0, t_0 + \epsilon]$ and hence the claim follows.

It remains to consider the situations where the physical jump condition gives $\Delta L_{t_0} > 0$, meaning that there is a jump discontinuity at time t_0. However, by the proof of Theorem 6.3.3, we are guaranteed to be back in the setting of the previous case after having resolved the jump. This completes proof. □

In addition to the above description of the solutions, we have the following uniqueness result for (MV), again in the Brownian case.

Theorem 6.4.2. *Suppose that, at any time of continuity $t \geq 0$, we have a bound $V_t(x) \leq \alpha^{-1} - C_t x^{n_t}$ on some interval $(0, x_t)$, for some constants $C_t, n_t \geq 0$. Then there is global uniqueness in the class of physical solutions.*

Proof. This follows from Theorem 6.3.3, as the assumption ensures it can also be applied at any explosion time that does not involve a jump. □

Needless to say, the above results are only partial, since they place constraints on the kind of states that V_t can reach at continuity times (in the possible cases where the H^1 norm of L explodes at time t without forcing a blow-up in the form of a jump discontinuity). In short, we are disallowing the pathological states depicted in Figure 6.3.1. The key contribution of [47] is to verify that, if the initial density does not change monotonicity infinitely often on compacts, then the same is true for V_{t-} at all positive times [47, Thm. 1.1]. Hence Theorem 6.4.1 applies unconditionally for such initial densities. Moreover, for the same class of initial densities, [47, Thm. 1.4] provides a new uniqueness argument that confirms the global uniqueness amongst all physical solutions.

6.4.1 Locally weak feedback after a blow-up

Consider an arbitrary time $t_0 \geq 0$ and, if t_0 is a blow-up time, assume that we have resolved the jump. Then the physical jump condition (with V_{t_0} in place of V_{t_0-}) gives

$$\inf\Big\{x > 0 : \int_0^{\alpha x} V_{t-}(y)dy < x\Big\} = 0.$$

The next result shows that if we restart the system from the density V_{t_0}, and if V_{t_0} is no greater than α^{-1} in a small neighbourhood of zero, then, for a short amount time, we are back in the weak feedback regime locally near zero. That is, the future densities are guaranteed to be strictly below α^{-1} in a small enough (but uniform) neighbourhood of the origin, for a sufficiently small amount of time.

For simplicity of presentation (and without loss of generality), we take $t_0 = 0$ in the statement of the result. The reasoning is similar in flavour to the probabilistic ideas in Theorem 6.2.2 and Theorem 6.3.1. Also, we stress that an analogous—if slightly different—result has recently appeared in [47, Section 2.1], however, our analysis was developed independently.

Proposition 6.4.3. *Suppose $V_0(x) \leq C_\star \mathbf{1}_{(0,x_\star)} + D_\star \mathbf{1}_{[x_\star,\infty)}$ and assume further that V_0 is such that $\Delta L_0 = 0$. Then there exists $t_0, x_0 > 0$ such that*

$$V_t(x) \leq (1 - \delta(t))C_\star \mathbf{1}_{(0,x_0)} + D_\star \mathbf{1}_{[x_0,\infty)} \quad \textit{for all} \quad x > 0, \tag{6.10}$$

at all times $t \in [0, t_0]$, where $\delta(t) > 0$ for $t > 0$ and δ is increasing with $\lim_{t\downarrow 0} \delta(t) = 0$.

Proof. Let p_t denote the density of the Brownian motion B_t. Recall that L is increasing and right-continuous with $L_0 = 0$, by assumption. Set $x_0 := x_\star/4$ and

consider an arbitrary interval (a,b) in $[0,x_0)$. By ignoring the absorption at the origin, we have

$$\begin{aligned}
\nu_t([a,b]) &\le \mathbb{P}\left(a \le X_0 + B_t - \alpha L_t \le b\right) \\
&= \mathbb{P}\left(a - B_t + \alpha L_t \le X_0 \le b - B_t + \alpha L_t\right) \\
&= \mathbb{E}\Big[\int_{(a-B_t+\alpha L_t)\vee 0}^{b-B_t+\alpha L_t} V_0(x)dx\Big] \\
&= \int_{-\infty}^{b+\alpha L_t} \int_{(a-y+\alpha L_t)\vee 0}^{b-y+\alpha L_t} V_0(x)p_t(y)dxdy. \\
&= \int_{-b-\alpha L_t}^{b+\alpha L_t} \int_{a-y+\alpha L_t}^{b-y+\alpha L_t} V_0(x)p_t(y)dxdy + \int_{-\infty}^{-b-\alpha L_t} \int_{a-y+\alpha L_t}^{b-y+\alpha L_t} V_0(x)p_t(y)dxdy \\
&=: I_1 + I_2
\end{aligned}$$

Now consider any $t_0 > 0$ such that $\alpha L_{t_0} < x_0$. In the following, we restrict to $t < t_0$. Recalling that $b < \varepsilon_0$, we notice that, if $y \ge -b - \alpha L_t$, then $b - y + \alpha L_t < 4x_0 = x_\star$. Hence we have

$$I_1 = \int_{-b-\alpha L_t}^{b+\alpha L_t} \int_a^b V_0(z + \alpha L_t - y)p_t(y)dzdy \le C_\star(b-a)\int_{-b-\alpha L_t}^{b+\alpha L_t} p_t(y)dy,$$

since $z + \alpha L_t - y < x_\star$ on the domain of integration.

Let $c_0 := x_0 - \alpha L_{t_0} > 0$. Since $L_t \le L_{t_0}$, we have

$$2b + 2\alpha L_t + c_0 \le 2b + \alpha L_t + x_0 < 4x_0 = x_\star,$$

so we can conclude that

$$\begin{aligned}
I_2 &= \int_{-\infty}^{-b-\alpha L_t-c_0} \int_{a-y+\alpha L_t}^{b-y+\alpha L_t} V_0(x)dx p_t(y)dy + \int_{-b-\alpha L_t-c_0}^{-b-\alpha L_t} \int_a^b V_0(z+\alpha L_t - y)dz p_t(y)dy \\
&\le D_\star(b-a)\int_{-\infty}^{-b-\alpha L_t-c_0} p_t(y)dy + C_\star(b-a)\int_{-b-\alpha L_t-c_0}^{-b-\alpha L_t} p_t(y)dy
\end{aligned}$$

Changing variables $y \mapsto y + c_0$, we note that

$$\begin{aligned}
\int_{-\infty}^{-b-\alpha L_t-c_0} p_t(y)dy &= \int_{-\infty}^{-b-\alpha L_t} \frac{1}{\sqrt{2\pi t}} e^{-(y-c_0)^2/2t}dy \\
&= e^{-c_0^2/2t}\int_{-\infty}^{-b-\alpha L_t} \frac{1}{\sqrt{2\pi t}} e^{c_0 y/t} e^{-y^2/2t}dy \\
&\le e^{-c_0^2/2t}\int_{-\infty}^{-b-\alpha L_t} \frac{1}{\sqrt{2\pi t}} e^{-y^2/2t}dy \\
&= e^{-c_0^2/2t}\int_{b+\alpha L_t}^{\infty} p_t(y)dz
\end{aligned}$$

By taking t_0 smaller, we only make c_0 bigger, so we can ensure that $D_\star e^{-c_0^2/2t} \leq C_\star$ for $t \leq t_0$. Thus, we conclude that, for small enough t_0, it holds for all $t < t_0$ that

$$I_2 \leq (b-a)C_\star \left(\int_{b+\alpha L_t}^{\infty} p_t(y)dz + \int_{-b-\alpha L_t-c_0}^{-b-\alpha L_t} p_t(y)dy \right),$$

for any interval (a, b) in $[0, x_0)$. Combining the estimates for I_1 and I_2, we have

$$\begin{aligned}\nu_t([a,b]) &\leq C_\star(b-a)\left(\int_{b+\alpha L_t}^{\infty} p_t(y)dz + \int_{-b-\alpha L_t-c_0}^{-b-\alpha L_t} p_t(y)dy + \int_{-b-\alpha L_t}^{b+\alpha L_t} p_t(y)dy \right) \\ &\leq C_\star(b-a)\left(1 - \int_{-\infty}^{-x_0-\alpha L_t-c_0} p_t(y)dy \right).\end{aligned}$$

Consequently, for any $t < t_0$, we have

$$V_t(x) \leq (1-\delta(t))C_\star \quad \text{with} \quad \delta(t) := \int_{-\infty}^{-x_0-\alpha L_t-c_0} p_t(y)dy$$

for all $x \in (0, x_0)$. This completes the proof. □

The estimate (6.10) is certainly not optimal, and indeed the proof is very crude—by way of having ignored the absorbing boundary—but this is precisely the point. Without any assumptions on the regularity of the solution, it guarantees that the solution is well-behaved in a small time interval after a blow-up. In particular, it ensures that the uniqueness argument from Section 6.3 can be propagated for another small time interval after a blow-up. Moreover, it can serve as the basis for improving the known regularity of the solution, which we exploit in the next subsection.

6.4.2 1/2-Hölder continuity of the loss

Suppose t_0 is a blow-up time and let V_{t_0} be the new initial density after having resolved the jump. Then we know from the proof of Theorem 6.3.3 that we must have $V_{t_0}(x) \leq \alpha^{-1}$ on a small neighbourhood of zero (in fact, we know even more). Therefore, Proposition 6.4.3 gives that $V_{t_0+\epsilon}$ is strictly below α^{-1} on a small neighbourhood of the origin (for small $\epsilon > 0$). Exploiting this, we can show that the loss $t \mapsto L_t$ is $\frac{1}{2}$-Hölder continuous to the right of t_0, and the Hölder constant is uniform near t_0 if $V_{t_0}(0+) < \alpha^{-1}$. We note here that our arguments were developed independently of [47, Prop. 2.2] which is an essentially identical result, yet it is obtained by quite different methods.

Proposition 6.4.4. *Fix an arbitrary time $t > 0$ and $\delta_0 > 0$. Suppose we have $V_s(x) \leq C_\star \mathbf{1}_{(0,x_\star)} + D_\star \mathbf{1}_{[x_\star,\infty)}$, for all $x > 0$, for every $s \in [t, t+\delta_0]$, with $C_\star < \alpha^{-1}$. Then $s \mapsto L_s$ is $\frac{1}{2}$-Hölder continuous on $[t, t_\star]$, for a small enough $t_\star = t_\star(C_\star, D_\star) > t$, with uniform Hölder constant $K_\star = K_\star(C_\star, D_\star)$.*

Proof. Recall that we can write $\tilde{L}_\delta = L_{t+\delta} - L_t$ for the restarted loss process $\tilde{L}$, as defined in terms of the restarted particle $\tilde{X}^x_s$ from Remark 4.3.6.

Let $\tilde{V}_s(y) := \int_0^\infty \tilde{V}^x_s(y) V_t(x)dx$, where $\tilde{V}^x_s$ is the density of the absorbed particle $(\tilde{X}^x_s, s < \tilde{\tau}^x)$. By construction of $\tilde{X}^x_s$, we have $\tilde{V}_\delta = V_{t+\delta}$, at least in the almost everywhere sense. In particular, the assumption of the proposition implies that $\tilde{V}_\delta(x) \leq (1-\epsilon)\alpha^{-1}$ on $(0, x_\star)$, for all $\delta \in [0, \delta_0]$, for some $\epsilon > 0$ only depending on the constant $C_\star < \alpha^{-1}$.

By repeating the arguments from the proof of Lemma 4.5.1 for $\tilde{X}^{x_0}_s$ in place of X, we arrive at the expression

$$\tilde{L}_\delta = \int_0^\infty \Big(1 - \int_0^\infty G_\delta(x, x_0)dx\Big) V_t(x_0)dx_0 + 2\alpha \int_0^\delta \int_0^\infty \mathbb{E}[p_{\delta - s}(X^{x_0}_s)] V_t(x_0)dx_0 d\tilde{L}_s. \tag{6.11}$$

By changing variables $x \mapsto \delta^{\frac{1}{2}} x$ in the first inner integral, we have

$$1 - \int_0^\infty G_\delta(x, x_0)dx = 1 - \frac{1}{\sqrt{\pi}} \int_{-x_0\delta^{-\frac{1}{2}}}^{x_0\delta^{-\frac{1}{2}}} e^{-\frac{x^2}{2}} dx.$$

Next, in order to isolate the δ, we perform the change of variables $x_0 \mapsto \delta^{\frac{1}{2}} x_0$ in the outer integral. Using also that V_t is bounded by $D_\star$, this yields

$$\begin{aligned}\int_0^\infty \Big(1 - \int_0^\infty G_\delta(x, x_0)dx\Big) V_t(x_0)dy &\leq \delta^{\frac{1}{2}} D_\star \int_0^\infty \Big(1 - \frac{1}{\sqrt{\pi}} \int_{-x_0}^{x_0} e^{-\frac{y^2}{2}} dy\Big) dx_0 \\ &= \frac{D_\star}{\pi} \delta^{\frac{1}{2}}.\end{aligned}$$

Next, we can observe that, for $s < \delta$,

$$\begin{aligned}\int_0^\infty \mathbb{E}[p_{\delta - s}(X^{x_0}_s)] V_t(x_0)dx_0 &= \int_0^{x_\star} \frac{1}{\sqrt{2\pi(\delta - s)}} e^{-\frac{x^2}{2(\delta - s)}} \tilde{V}_\delta(x)dx \\ &\quad + \int_{x_\star}^\infty \frac{1}{\sqrt{2\pi(\delta - s)}} e^{-\frac{x^2}{2(\delta - s)}} \tilde{V}_\delta(x)dx \\ &\leq \frac{1}{2}(1-\epsilon)\alpha^{-1} + \frac{1}{4}\epsilon\alpha^{-1} = \frac{1}{2}(1 - \frac{\epsilon}{2})\alpha^{-1},\end{aligned}$$

for all $\delta \in [0, \delta_\star]$, for a fixed $\delta_\star \in (0, \delta_0)$, since we can make the second integral as small as we like by ensuring that $\delta - s$ is small. Inserting the two estimates into (6.11) we finally obtain

$$\tilde{L}_\delta \leq (1 - \frac{\epsilon}{2})\tilde{L}_\delta + \delta^{\frac{1}{2}} \frac{D_\star}{\pi} \delta^{\frac{1}{2}}$$

and hence there exists $K_\star := K_\star(C_\star, D_\star) > 0$ such that

$$L_{t+\delta} - L_t = \tilde{L}_\delta \leq K_\star \delta^{\frac{1}{2}},$$

for all $\delta \in [0, \delta_\star]$. This proves the Hölder continuity at t. By repeating the same arguments for $s \in [t, t+\delta_\star)$, we obtain $\frac{1}{2}$-Hölder continuity at every $s \in [t, t+\delta_\star)$ with the same Hölder constant. This completes the proof. □

We end this section by observing that the previous result is sharp in regard to the level of Hölder continuity.

Corollary 6.4.5. *Suppose* $\lim_{x \downarrow 0} V_t(x) \geq \epsilon > 0$. *Then the* $\frac{1}{2}$*-Hölder continuity at time* t *is sharp. That is, there exists* $\delta_0 > 0$ *and* $C_0 > 0$ *such that*

$$L_{t+\delta} - L_t \geq C\delta^{\frac{1}{2}}$$

for all $\delta < \delta_0$.

Proof. Recall the expression for $\tilde{L}_\delta = L_{t+\delta} - L_t$ from the proof of Proposition 6.4.4. By continuity, we have $V_t(x) \geq \frac{\epsilon}{2}$ in a small neighborhood $(0, x_\epsilon)$, and hence we can immediately deduce that

$$\begin{aligned}\tilde{L}_\delta &\geq \delta^{\frac{1}{2}} \int_0^\infty \Big(1 - \frac{1}{\sqrt{\pi}} \int_{-x_0}^{x_0} e^{-\frac{y^2}{2}} dy\Big) V_t(\delta^{\frac{1}{2}} x_0) dx_0 \\ &\geq \delta^{\frac{1}{2}} \frac{\epsilon}{2} \int_0^{x_\epsilon \delta^{-\frac{1}{2}}} \Big(1 - \frac{1}{\sqrt{\pi}} \int_{-x_0}^{x_0} e^{-\frac{y^2}{2}} dy\Big) dx_0.\end{aligned}$$

This proves the claim. □

Appendix A

Auxiliary Results for Chapter 3

A.1 Some technical lemmas

Lemma A.1.1. *Suppose ν is a measure valued process satisfying the conditions in Definition 3.2.3 and let $g_s(x) = g(s, x, \nu_s, L_s)$, where g is any of b, $\mathfrak{b}$, σ, or σ^2, with $L_s = 1 - \nu_s(0, \infty)$. Define the error term*

$$\mathcal{E}^g_{t,\varepsilon}(x) := \langle \nu_t, g_t(\cdot) G_\varepsilon(x, \cdot) \rangle - g_t(x)(\mathcal{T}_\varepsilon \nu_t)(x).$$

Then we have

$$\mathbb{E}\Big[\int_0^T \big\|\mathcal{E}^g_{t,\varepsilon}\big\|^2_{L^2(\mathbb{R})}\, dt\Big] \to 0 \quad as \quad \varepsilon \downarrow 0.$$

Proof. This follows by a simple modification of Lemma 8.1 in [73]. The only thing to note is that the crude bound $|g_t(x) - g_t(y)| \leq 2\|g_t\|_\infty$ no longer holds, as g_t need not be bounded. However, the arguments from [73] still apply if we instead rely on $|g_t(x) - g_t(y)| \leq \|\partial_x g_t\|_\infty |x - y|$. □

Lemma A.1.2. *Suppose $\bar{\nu}$ satisfies the whole-space analogues of conditions (iii)-(iv) in Definition 3.2.3. Let $g_s = g(s, \cdot, \nu_s, L_s)$, where g is any of b, $\mathfrak{b}$, σ, or σ^2, and define the error terms*

$$\begin{aligned}
\bar{\mathcal{E}}^g_{t,\varepsilon}(x) &:= \langle \bar{\nu}_t, g_t(\cdot) p_\varepsilon(x, \cdot) \rangle - g_t(x) \partial_x (\bar{\mathcal{T}}_\varepsilon \bar{\nu}_t)(x) + \partial_x g_t(x) \bar{\mathcal{H}}^g_{t,\varepsilon}(x), \\
\bar{\mathcal{H}}^g_{t,\varepsilon}(x) &:= \langle \bar{\nu}_t, (x - \cdot) \partial_x p_\varepsilon(x - \cdot) \rangle .
\end{aligned}$$

Then, for $k = 1, 2$, we have

$$\mathbb{E}\Big[\int_0^T \big\|\bar{\mathcal{E}}^g_{t,\varepsilon}\big\|^{2k}_{L^2(\mathbb{R})} dt\Big] \to 0, \quad \mathbb{E}\Big[\int_0^T \big\|x\bar{\mathcal{E}}^g_{t,\varepsilon}\big\|^2_{L^2(\mathbb{R})} dt\Big] \to 0, \quad as \quad \varepsilon \downarrow 0.$$

Proof. This is immediate from the obvious modifications of Lemma 8.2 of [73]. □

Lemma A.1.3. *Let ν_t^N be as defined in (3.5) and let $\tilde{\nu}_t$ be any measure valued process satisfying Definition 3.2.3. Then, for every $a > 0$, we have*

$$\mathbb{E}\Big[\int_0^T \big\langle \tilde{\nu}_t, x^k \mathbf{1}_{[\lambda,\infty)}(x)\big\rangle dt\Big] = o(\lambda^k e^{-a\lambda}) \quad as \quad \lambda \to \infty, \tag{A.1}$$

and, likewise, it holds uniformly in $N \geq 1$ and $t \in [0,T]$ that

$$\mathbb{E}\Big[\big\langle \nu_t^N, x^k \mathbf{1}_{[\lambda,\infty)}(x)\big\rangle\Big] = o(\lambda^k e^{-a\lambda}) \quad as \quad \lambda \to \infty \tag{A.2}$$

Proof. Fix an arbitrary $a > 0$ and let $\varepsilon > 0$ be given. By Corollary 3.3.4, there exists $\lambda_0 \geq 0$ such that

$$\mathbb{E}\left[\nu_t^N(\lambda,\infty)\right] \leq \varepsilon e^{-a\lambda} \quad \forall \lambda \geq \lambda_0, \tag{A.3}$$

uniformly in $N \geq 1$ and $t \in [0,T]$. Given any $\lambda \geq \lambda_0$ we let $\{s_0, s_1, \ldots\}$ denote the partition of $[\lambda,\infty)$ with $s_i - s_{i-1} = 1/a$. By (A.3) and monotone convergence, we get

$$\mathbb{E}\big[\langle \nu_t^N, x\mathbf{1}_{[\lambda,\infty)}\rangle\big] \leq \sum_{i=1}^{\infty} s_i \mathbb{E}\big[\nu_t^N(s_{i-1},\infty)\big] \leq \varepsilon \sum_{i=1}^{\infty} s_i e^{-as_{i-1}} = \varepsilon e \sum_{i=1}^{\infty} s_i e^{-as_i}$$

Now, $x \mapsto xe^{-ax}$ is decreasing for $x \geq 1/a$, so taking $\lambda_0' := \max\{1/a, \lambda_0\}$ and noting that $s_0 = \lambda$, it holds for all $\lambda > \lambda_0'$ that

$$\mathbb{E}\big[\langle \nu_t^N, x\mathbf{1}_{[\lambda,\infty)}\rangle\big] \leq \varepsilon e \int_{\lambda}^{\infty} xe^{-ax}dx = \varepsilon e\big(\frac{\lambda}{a} + \frac{1}{a^2}\big)e^{-a\lambda}.$$

This proves (A.2) for $k = 1$ and the work for $k \geq 2$ is analogous. The claim in (A.1) follows similarly, by relying on the exponential tail property from Definition 3.2.3. □

Lemma A.1.4. *Let ν be a limit point from Theorem 3.2.4 and let $\tilde{\nu}$ be any measure valued process satisfying Definition 3.2.3. Then, almost surely (as $\lambda \to \infty$),*

$$\lim_{\varepsilon \downarrow 0} \frac{\nu_t(0,\varepsilon)}{\varepsilon} = 0, \quad \nu_t(\lambda,\infty) = O(e^{-\lambda}), \quad and \quad \int_0^T \tilde{\nu}_t(\lambda,\infty)dt = O(e^{-\lambda}).$$

Proof. For the first claim, recall from Proposition 3.4.3 that there exists $\delta \in (0,1]$ and $\beta > 0$ such that $\mathbb{E}[\nu_t(0,\varepsilon)] = t^{-\frac{\delta}{2}} O(\varepsilon^{1+\beta})$ as $\varepsilon \to 0$. Using Markov's inequality, we thus deduce that, for any $\theta > 0$ and n sufficiently large,

$$\mathbb{P}\Big[n^{2/\beta}\nu_t(0,n^{-2/\beta}) > \theta\Big] \leq \theta^{-1} n^{2/\beta} \mathbb{E}\big[\nu_t(0,n^{-2/\beta}\big] \leq Ct^{-\frac{\delta}{2}}\theta^{-1}n^{-2}.$$

Hence the Borel–Cantelli lemma gives $\limsup_n n^{2/\beta}\nu_t(0,n^{-2/\beta}) = 0$ almost surely. Now, given $\varepsilon > 0$, we have $(n+1)^{-2/\beta} < \varepsilon \leq n^{-2/\beta}$ for some $n \geq 1$, so we deduce that

$$\limsup_{\varepsilon \downarrow 0} \frac{\nu_t(0,\varepsilon)}{\varepsilon} \leq \limsup_{n \geq 1} \frac{\nu_t(0,n^{-2/\beta})}{(n+1)^{-2/\beta}} \leq \limsup_{n \geq 1} \frac{\nu_t(0,n^{-2/\beta})}{n^{-2/\beta}} \Big(\frac{n+1}{n}\Big)^{\frac{2}{\beta}}.$$

Since the latter is zero with probability 1, this proves the first claim. The two remaining results follow by analogous considerations for the tail probabilities, using the exponential tail properties from Proposition 3.4.3 and Definition 3.2.3. □

A.2 Proofs of Propositions 3.3.5 and 3.3.6

As an immediate consequence of Corollary 2.3.2, we have the following.

Lemma A.2.1. *Let* $\Lambda_t^{i,N} = |X_t^i| + \sum_{j=1}^N a_j^N |X_t^j|$, *as in (2.8). Then*

$$\mathbb{P}\Big[\sup_{t\le T} \Lambda_t^{i,N} \ge \lambda\Big] = o(1) \quad as \quad \lambda \to \infty.$$

uniformly in $N \ge 1$ *and* $i = 1, \dots, N$.

Based on this lemma, we can adapt the arguments from Section 4 of [73] to prove Proposition 3.3.5 and Proposition 3.3.6.

Proof of Proposition 3.3.5. Arguing as in the proof of Proposition 4.6 in [73], for all $a > 0$ and $\theta := \frac{1}{2}(1-r)$, we have

$$\mathbb{P}\big[L_{t+h}^N - L_t^N < \delta, L_t^N < r\big] \le \mathbb{P}\big[L_{t+h}^N - L_t^N < \delta,\ \nu_t^N(0,a) > \theta\big] + o(e^{-a}).$$

Let $E := \{L_{t+h}^N - L_t^N < \delta,\ \nu_t^N(0,a) > \theta\}$ and define a random index set $\mathcal{I}$ by

$$\mathcal{I} := \big\{1 \le i \le N : X_t^i < a,\ t < \tau_i,\ a_i^N > \theta/2N\big\}.$$

Note that the particles with index in $\{1 \le i \le N : a_i^N \le \frac{1}{2}\theta N^{-1}\}$ can contribute at most $\frac{1}{2}\theta$ towards $\nu_t^N(0,a)$. Recalling also that $a_i^N \le m/N$ for some $m > 0$, see (3.6), it follows that on the event E we must have $|\mathcal{I}| > N\theta/2m$. Hence

$$\mathbb{P}[E] \le \sum_{\mathcal{I}_0 : |\mathcal{I}_0| > \frac{\theta N}{2m}} \mathbb{P}[E \,|\, \mathcal{I} = \mathcal{I}_0]\mathbb{P}[\mathcal{I} = \mathcal{I}_0]. \tag{A.4}$$

Furthermore, since $L_{t+h}^N - L_t^N < \delta$ on E while $a_i^N > \theta/2N$ for $i \in \mathcal{I}$, we deduce that

$$\mathbb{P}[E \,|\, \mathcal{I} = \mathcal{I}_0] \le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I}_0 : \inf_{u\le h} X_{t+u}^i \le 0\big\}\big| < 2\delta N/\theta \mid \mathcal{I} = \mathcal{I}_0\Big]. \tag{A.5}$$

In order to estimate (A.5), we let $Z_t^i = \Upsilon_t(X_t^i)$ as in Lemma 2.3.3 and recall that Z_t^i then satisfies

$$dZ_t^i = \hat{b}_t^i dt + dB_t^i, \quad |\hat{b}_t^i| \le c(1 + \Lambda_t^{i,N}), \quad |Z_t^i| \le c'(1 + |X_t^i|). \tag{A.6}$$

Using that $Z_t^i \le 0$ if and only if $X_t^i \le 0$, the estimate in (A.5) implies

$$\mathbb{P}[E \,|\, \mathcal{I} = \mathcal{I}_0] \le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I} : \inf_{u\le h} Z_{t+u}^i \le 0\big\}\big| < 2\delta N/\theta \mid \mathcal{I} = \mathcal{I}_0\Big].$$

Moreover, it is immediate from (A.6) that

$$Z^i_{t+u} \le Z^i_t + c_1 h(1+\tilde{\Lambda}^{i,N}_h) + \tilde{B}^i_u, \quad \tilde{B}^i_u := B^i_{t+u} - B^i_t, \quad \tilde{\Lambda}^{i,N}_h := \sup_{u\le h} \Lambda^{i,N}_{t+u}.$$

In particular, if $\inf_{u\le h} \tilde{B}^i_u \le -Z^i_t - ch(1+\tilde{\Lambda}^{i,N}_h)$, then $\inf_{u\le h} Z^i_{t+u} \le 0$, and hence

$$\mathbb{P}[E \,|\, \mathcal{I} = \mathcal{I}_0] \le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I} : \inf_{u\le h} \tilde{B}^i_u \le -Z^i_t - ch(1+\tilde{\Lambda}^{i,N}_h)\big\}\big| < \tfrac{2\delta}{\theta} N \mid \mathcal{I} = \mathcal{I}_0\Big].$$

Using the bound $|Z^i_t| \le c'|X^i_t|$ from (A.6), the definition of $\mathcal{I}$ implies that there is a constant $\tilde{c} > 0$ such that

$$\mathbb{P}[E \,|\, \mathcal{I} = \mathcal{I}_0] \le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I} : \inf_{u\le h} \tilde{B}^i_u \le -\tilde{c}a - \tilde{c}(1+\tilde{\Lambda}^{i,N}_h)\big\}\big| < \tfrac{2\delta}{\theta} N \mid \mathcal{I} = \mathcal{I}_0\Big].$$

Recalling the fact that

$$\tilde{B}^i_u = \int_t^{t+u} \rho_s dW^0_s + \int_t^{t+u} \sqrt{1-\rho_s^2} dW^i_s =: I_u + J^i_u,$$

we split the above probability on $\{\sup_{u\le h} |I_u| < \tilde{c}a\} \cap \{\sup_{u\le h} \Lambda^{i,N}_{t+u} < a\}$ and its complement. In this way, we get

$$\begin{aligned}\mathbb{P}[E \,|\, \mathcal{I} = \mathcal{I}_0] \le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I} : \inf_{u\le h} J^i_u \le -3\tilde{c}a - \tilde{c}\big\}\big| < \tfrac{2\delta}{\theta} N \mid \mathcal{I} = \mathcal{I}_0\Big] \\ + \mathbb{P}\Big[\sup_{u\le h} |I_u| \ge \tilde{c}a\Big] + \mathbb{P}\Big[\sup_{u\le h} \Lambda^{i,N}_{t+u} \ge a\Big].\end{aligned}$$

From Lemma A.2.1 we know that the last term is $o(1)$ as $a \to \infty$ uniformly in $N \ge 1$. Similarly, I_u is a martingale, and hence the second term is also $o(1)$ as $a \to \infty$ by Doob's Maximal Inequality. Concerning the first term, notice that ρ_s does not depend on i and recall that it is bounded away from 1, so by performing a single random time-change we can assume the J^i's are independent standard Brownian motions (up to a constant multiplying h). Recalling (A.4), the proof can then be completed via a law of large numbers argument as in Proposition 4.6 of [73], by choosing the free parameter a as a suitable function of δ. □

Proof of Proposition 3.3.6. Arguing as in Proposition 4.7 of [73], it suffices to show that

$$\lim_{\delta\to 0} \lim_{N\to\infty} \mathbb{P}\big[L^N_{t+\delta} - L^N_t \ge \eta,\ \nu^N_t(0,\varepsilon) < \eta/2\big] = 0.$$

Let $E := \{L^N_{t+\delta} - L^N_t \ge \eta,\ \nu^N_t(0,\varepsilon) < \eta/2\}$ and define the random index set

$$\mathcal{I} := \big\{1 \le i \le N : X^i_t \ge \varepsilon \ \text{ or } \ t \ge \tau_i\big\}.$$

By (3.6), there exists m such that $a_i^N \le m/N$, so noting that $\{\nu_t^N(0,\varepsilon) < \eta/2\}$ is contained in $\{|\mathcal{I}| \ge \frac{N}{m}(1-\frac{\eta}{2})\}$, we have

$$\mathbb{P}[E] \le \sum_{\mathcal{I}_0 : |\mathcal{I}_0| \ge \frac{N}{m}(1-\frac{\eta}{2})} \mathbb{P}[E \mid \mathcal{I} = \mathcal{I}_0]\mathbb{P}[\mathcal{I} = \mathcal{I}_0].$$

The conditional probabilities can be estimated by

$$\begin{aligned}\mathbb{P}[E \mid \mathcal{I} = \mathcal{I}_0] &\le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I}_0 : \inf_{s\in[t,t+\delta]} X_s^i \le 0, \;\; X_t^i \ge \varepsilon\big\}\big| \ge \frac{N\eta}{2m} \mid \mathcal{I} = \mathcal{I}_0\Big] \\ &\le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I}_0 : \inf_{s\in[t,t+\delta]} (X_s^i - X_t^i) \le -\varepsilon\big\}\big| \ge \frac{N\eta}{2m} \mid \mathcal{I} = \mathcal{I}_0\Big] \qquad \text{(A.7)}\end{aligned}$$

Using the scale transform Υ from Lemma 2.3.3, we let $U_s^i := \Upsilon_{t+s}(X_{t+s}^i) - \Upsilon_{t+s}(X_t^i)$, which is the integral of $y \mapsto \sigma(t+s,y)^{-1}$ over the interval $[X_t^i, X_{t+s}^i]$. By analogy with Lemma 2.3.3, we get

$$dU_s^i = u_s^i ds + \rho_{t+s} dW_{t+s}^0 + (1-\rho_{t+s}^2)^{\frac{1}{2}} dW_{t+s}^i =: u_s^i ds + dI_s + dJ_s^i \qquad \text{(A.8)}$$

where the drift satisfies $|u_s^i| \le c_1(1+\Lambda_{t+s}^{i,N})$. By construction of Υ and the boundedness from below of $1/\sigma$, say $1/\sigma \ge c_2 > 0$, it follows from (A.7) that

$$\mathbb{P}[E \mid \mathcal{I} = \mathcal{I}_0] \le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I}_0 : \inf_{s\le\delta} U_s^i \le -c_2\varepsilon\big\}\big| \ge \frac{N\eta}{2m} \mid \mathcal{I} = \mathcal{I}_0\Big].$$

Given $\delta > 0$, fix a constant $a = a(\delta) > 0$ to be specified later. Using the decomposition (A.8) of U_s^i and the growth estimate for its drift u_s^i, we see that, on the event

$$\Big\{\sup_{s\le\delta} \Lambda_{t+s}^{i,N} < a/\delta\Big\} \cap \Big\{\sup_{s\le\delta} |I_s| < a\Big\}, \qquad \text{(A.9)}$$

if $i \in \mathcal{I}_0$ is such that $\inf_{s\le\delta} U_s^i \le -c_2\varepsilon$, then

$$\begin{aligned}\inf_{s\le\delta} J_s^i &\le -c_2\varepsilon + \delta c_1(1 + \sup_{s\le\delta} \Lambda_{t+s}^{i,N}) + \sup_{s\le\delta} |I_s| \\ &\le -c_2\varepsilon + c_1\delta + (1+c_1)a \;\le -c_3(\varepsilon - \delta - a)\end{aligned}$$

Consequently, splitting up the desired probability on the event (A.9) and its complement,

$$\begin{aligned}\mathbb{P}[E \mid \mathcal{I} = \mathcal{I}_0] \le \mathbb{P}\Big[\big|\big\{i \in \mathcal{I}_0 : \inf_{s\le\delta} J_s^i \le -c_3(\varepsilon-\delta-a)\big\}\big| \ge \frac{N\eta}{2m} \mid \mathcal{I} = \mathcal{I}_0\Big] \\ + \mathbb{P}\Big[\sup_{s\le\delta} \Lambda_{t+s}^{i,N} \ge a/\delta\Big] + \mathbb{P}\Big[\sup_{s\le\delta} |I_s| \ge a\Big].\end{aligned}$$

By Doob's maximal inequality, the last term is bounded by δa^{-2}, so choosing $a = a(\delta)$ s.t. $\delta a^{-2} \to 0$ as $\delta \to 0$, we can ensure that it vanishes as $\delta \to 0$. Moreover, requiring also that $\delta/a \to 0$ as $\delta \to 0$, it follows from Lemma A.2.1 that the second term vanishes (uniformly in N) as $\delta \to 0$. These requirements are satisfied by $a = a(\delta) := \delta^{1/2}\log\log(1/\delta)$, and with this choice for a, the proof can now be finished by the same arguments as in Proposition 4.7 of [73]. □

Bibliography

[1] Acharya, V.V., Pedersen, L.H., Philippon, T., Richardson, M.: Measuring Systemic Risk. *Rev. Finan. Stud.* **30**, 2–47 (2017)

[2] Allen, F., Gale, D.: Financial Contagion. *J. Polit. Economy* **108**, 1–33 (2000)

[3] Amini, H., Cont,R., Minca, A.: Resilience to contagion in financial networks *Mathematical Finance* **26**, 329–365 (2016)

[4] Azizpour, S., Giesecke, K., Schwenkler, G.: Exploring the Sources of Default Clustering. *J. Finan. Econ.* **129**, 154–183 (2018)

[5] Banerjee, T., Bernstein, A., Feinstein, Z.: Dynamic clearing and contagion in financial networks. Preprint. Available at `ArXiv:1801.02091` (2018)

[6] Benazzoli, C., Campi, L., Di Persio, L.: Mean-field games with controlled jump-diffusion dynamics: existence results and an interbank illiquid market model. Preprint. Available at `arXiv:1703.01919` (2017)

[7] Battiston, S., Delli Gatti, D., Gallegati, M., Greenwald, B., Stiglitz, J. E.: Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control* **36**, 1121–1141

[8] Benoit, S., Colliard, J.-E., Hurlin, C., Pérignon, C.: Where the Risks Lie: A Survey on Systemic Risk. *Review of Finance* **21**, 109–152 (2017)

[9] Bertoin, J.: *Lévy processes*. Cambridge Tracts in Mathematics. Cambridge University Press (1996)

[10] Berg, T., Gider, J.: What Explains the Difference in Leverage between Banks and Non-Banks? *J. Finan. Quant. Anal.* **52**, 2677–2702 (2017)

[11] Billingsley, P.: *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition (1999)

[12] Black, F., Cox, J. C.: Valuing corporate securities: some effects of bond indenture provisions. *J. Finance* **31**, 351–367 (1976).

[13] N. Brunel. Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. *Journal of Computational Neuroscience* **8**, 183–208 (2000).

[14] Brunnermeier, M.K.: Deciphering the Liquidity and Credit Crunch 2007-2008. *J. Econ. Perspect.* **23**, 77–100 (2009)

[15] Brunnermeier, M.K., Gorton, G., Krishnamurthy, A.: Risk Topography. *NBER Macroeconomics Annual* **26**, 149–176 (2012)

[16] Bush, N., Hambly B.M., Haworth, H., Jin, L., Reisinger, C.: Stochastic evolution equations in portfolio credit modelling. *SIAM J. Fin. Math.* **2**, 627–664 (2011)

[17] Cáceres, M.J., Carrillo, J.A., Perthame, B.: Analysis of non-linear noisy integrate & fire neuron models: blow-up and steady states. *The Journal of Mathematical Neuroscience*, **1**, 1–33 (2011)

[18] Calvez, V., Meunier, N., Voituriez, R.: A one-dimensional Keller-Segel equation with a drift issued from the boundary. *C. R. Math. Acad. Sci. Paris*, **348**, 629–634 (2010).

[19] Calvez, V., Hawkins, R., Meunier, N., Voituriez, R.: Analysis of a nonlocal model for spontaneous cell polarization. *SIAM J. Appl. Math.*, **72**, 594–622 (2012).

[20] Campi, L., Fischer, M.: N-player games and mean-field games with absorption. *Ann. Appl. Probab.* **28**, 2188–2242 (2018)

[21] Campi, L., Ghio, M., Livieri, G.: N-player games and mean-field games with smooth dependence on past absorptions. Preprint. Available at `arXiv:1902.02670` (2019)

[22] Cannon, J.R.: *The One-Dimensional Heat Equation.* Encyclopedia Math. Appl. **23**, Addison–Wesley, Reading, MA (1984)

[23] Capponi, A., Bo, L.: Systemic Risk in Interbanking Networks. *SIAM J. Fin. Math.* **6**, 386–424 (2015)

[24] Capponi, A., Chen, P.-C.: Systemic risk mitigation in financial networks. *J. Econom. Dynam. Control* **58**, 152–166 (2015)

[25] Capponi, A., Sun, X., Yao, D.: A Dynamic Network Model of Interbank Lending: Systemic Risk and Liquidity Provisioning. Preprint. Available at SSRN: 3028417 (2017)

[26] Carmona, R., Delarue, F.: *Probabilistic Theory of Mean Field Games with Applications II. Mean Field Games with Common Noise and Master Equations.* Probability Theory and Stochastic Modelling 84, Springer (2018)

[27] Carmona, R., Delarue, F., Lacker, D.: Mean field games with common noise. *Annals of Probability*, **44**, 3740–3803 (2016)

[28] Carmona, R., Fouque, J.-P., Sun, L.-H.: Mean Field Games and Systemic Risk. *Comm. Math. Sci.* **13**, 911–933 (2015)

[29] Carrillo, J.A., González, M.D.M.., Gualdani, M.P., Schonbek, M.E.: Classical solutions for a nonlinear Fokker–Planck equation arising in computational neuroscience. *Comm. Partial Differential Equations* **38**, 385–409 (2013)

[30] Carrillo, J.A., Perthame, B., Salort, D., Smets, D.: Qualitative properties of solutions for the noisy integrate and fire model in computational neuroscience. *Nonlinearity* **28**, 3365–3388 (2015)

[31] Cho, S., Kim, P., and Park, H.: Two-sided estimates on Dirichlet heat kernels for time-dependent parabolic operators with singular drifts in $C^{1,\alpha}$-domains. *J. Differential Equations* **252**, 1101–1145 (2012)

[32] Cochrane, J.H.: Toward a Run-Free Financial System. In: Baily, M.N., Taylor, J.B. (eds.) *Across the Great Divide*, 197–249. Hoover Press (2014)

[33] Cont, R., Müller, M.S.: A stochastic partial differential equation model for limit order book dynamics. Preprint. Available at SSRN:3366536 (2019)

[34] Cont, R, Moussa, A., Santos, E.B.: Network structure and systemic risk in banking systems. In: Fouque, J.-P., Langsam, J. (eds.) *Handbook of Systemic Risk*, 327–368. CUP (2013)

[35] Cont, R., Schaaning, E.: Monitoring Indirect Contagion. *J. Banking Finance* **104**, 85–102 (2019)

[36] Cont, R., Wagalath, L.: Running for the exit: distressed selling and endogenous correlation in financial markets. *Math. Finance* **23**, 718–741 (2013)

[37] Cont, R., Wagalath, L.: Fire sale forensics: measuring endogenous risk. *Math. Finance* **26**, 835–866 (2016)

[38] Cvitanic, J., Ma, J., Zhang, J.: The law of large numbers for self-exciting correlated defaults. *Stoch. Proc. Appl.* **122**, 2781–2810 (2012)

[39] Danielsson, J., Shin, H.S., Zigrand, J.-P.: Endogenous Extreme Events and the Dual Role of Prices. *Annu. Rev. Econ.* **4**, 111–129 (2012)

[40] Dai Pra, P., Wolfgang, J. R., Sartori, E., Tolotti, M.: Large portfolio losses: A dynamic contagion model. *Ann. Appl. Probab.* **19**, 347–394 (2009)

[41] Davis, M., Lo, V.: Infectious defaults. *Quant. Finance* **1**, 382–387 (2001)

[42] Davis, M.: Contagion models in credit risk. In: Lipton, A., Rennie, A. (eds.) *Oxford Handbook of Credit Risk Derivatives*. Oxford University Press, (2011).

[43] Dawson, D., Vaillancourt, J.: Stochastic McKean–Vlasov equations. *NoDEA* **2**, 199–229 (1995)

[44] Dees, S., Henry, J., Martin, R.: *STAMPE: Stress-Test Analytics for Macroprudential Purposes in the Euro Area*, ECB (2017)

[45] Delarue, F., Inglis, J., Rubenthaler, S., Tanré, E.: Global solvability of a networked integrate-and-fire model of McKean–Vlasov type. *Ann. Appl. Probab.* **25**, 2096–2133 (2015)

[46] Delarue, F., Inglis, J., Rubenthaler, S., Tanré, E.: Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stoch. Proc. Appl.* **125**, 2451–2492 (2015)

[47] Delarue, F., Nadtochiy, S., Shkolnikov, M.: Global solutions to the supercooled Stefan problem with blow-ups: regularity and uniqueness. Preprint. Available at `arXiv:1902.05174` (2019)

[48] Dembo, A., Tsai, L.C.: The Criticality of a Randomly-Driven Front. To appear in *Arch. Ration. Mech. Anal.* Available at `arXiv:1705.10017` (2017)

[49] Duffie, D.: Financial Regulatory Reform After the Crisis: An Assessment. *Management Science*, **64**, 4471–4965 (2017)

[50] Duffie, D., Eckner, A., Horel, G., Saita, L.: Frailty correlated default. *J. Finance* **64**, 2089–2123 (2009)

[51] Eisenberg, L., Noe, T.H.: Systemic risk in financial systems. *Management Science* **47**, 236–249 (2001)

[52] Ethier, S., Kurtz, T.: *Markov Processes: Characterization and Convergence.* Wiley (1986)

[53] Evans, L.C.: *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society (2010)

[54] Fasano, A., Primicerio, M.: New results on some classical parabolic free-boundary problems. *Quart. Appl. Math.*, **38**, 439–460 (1981)

[55] Fasano, A., Primicerio, M.: A critical case for the solvability of Stefan-like problems, *Math. Methods Appl. Sci.*, **5**, 84–96 (1983)

[56] Fasano, A., Primicerio, M., Howison, S.D., Ockendon, J.R.: On the singularities of one dimensional Stefan problems with supercooling. In: *Mathematical Models for Phase Change Problems*, Ed. J.F. Rodrigues, International Series of Numerical Mathematics 88, Birkhäuser Verlag, Basel (1989)

[57] Fasano, A., Primicerio, M., Howison, S.D., Ockendon, J.R.: Some remarks on the regularization of supercooled one-phase Stefan problems in one dimension, *Quart. Appl. Math.*, **47**, 153–168 (1990)

[58] Fouque, J.-P., Ichiba, T.: Stability in a model of interbank lending. *SIAM J. Fin. Math.* **4**, 784–803 (2013)

[59] Fouque, J.-P., Sun, L.-H. Systemic risk illustrated. In: Fouque, J.-P., Langsam, J. (eds.) *Handbook on Systemic Risk*, 444–452. Cambridge University Press (2013)

[60] Friz, P., Victoir, N.: *Multidimensional Stochastic Processes as Rough Paths.* Cambridge Studies in Advanced Mathematics, Cambridge University Press (2010)

[61] Gai, P., Kapadia, S.: Contagion in financial networks. *Proc. R. Soc. A* **466**, 2401–2423 (2010)

[62] Garnier, J., Papanicolaou, G., Yang, T.-W.: Large deviations for a mean field model of systemic risk. *SIAM J. Fin. Math.* **4**, 151–184 (2013)

[63] Garnier, J., Papanicolaou, G., Yang, T.-W.: A risk analysis for a system stabilized by a central agent. *Risk and Decision Analysis* **6**, 97–120 (2017)

[64] Giesecke, K., Schwenkler, G., Sirignano, J.: Inference for Large Financial Systems. To appear in *Mathematical Finance*. Available at `SSRN:3012751` (2017).

[65] Giesecke, K., Spiliopoulos, K., Sowers, R.: Default clustering in large portfolios: Typical events. *Ann. Appl. Probab.* **23**, 348–385 (2013)

[66] Giesecke, K., Spiliopoulos, K., Sowers, R., Sirignano, J.: Large portfolio asymptotics for loss from default. *Math. Finance* **25**, 77–114 (2015)

[67] Glasserman, P., Young, H.P.: Contagion in Financial Networks. *J. Econ. Literature* **54**, 779–831 (2016)

[68] Gornall, W., Strebulaev, I.A.: Financing as a Supply Chain: The Capital Structure of Banks and Borrowers. *J. Finan. Econ.* **129**, 510–530 (2018)

[69] Gualdi, S., Bouchaud, J.-P., Cencetti, G., Tarzia, M., Zamponi, F.:. Endogenous crisis waves: a stochastic model with synchronized collective behavior. *Physical Review Letters*, **114**, 088701, 1–5 (2015)

[70] Gyrya, P., Saloff-Coste L.: *Neumann and Dirichlet Heat Kernels in Inner Uniform Domains*. Astérisque **336**, Société Mathématique de France (2011)

[71] Hambly, B., Kalsi, J.: A reflected moving boundary problem driven by space-time white noise. Preprint, to appear in *Stoch. Partial Differ. Equ. Anal. Comput.* Available at `arXiv:1805.10166` (2018)

[72] Hambly, B., Kalsi, J.: Stefan problems for reflected SPDES driven by space-time white noise. Preprint, to appear in *Stoc. Proc. Applic.* Available at `arXiv:1806.04739` (2018)

[73] Hambly, B., Ledger, S.: A Stochastic McKean–Vlasov Equation for Absorbing Diffusions on the Half-Line. *Ann. Appl. Probab.*, **27**, 2698–2752 (2017)

[74] Hambly, B., Ledger, S., Søjmark, A.: A McKean–Vlasov Equation with Positive Feedback and Blow-ups. Preprint. To appear in *Ann. Appl. Probab.* Available at arXiv:1801.07703 (2018).

[75] Hambly, B., Søjmark, B.: An SPDE model for systemic risk with endogenous contagion. *Finance and Stochastics* **157**, 535–594 (2019)

[76] Hellwig, M. F.: Systemic Risk in the Financial Sector: An Analysis of the Subprime-Mortgage Financial Crisis. *DE Economist*, **157**, 129–207 (2009)

[77] Herrero, M.A., Velázquez, J.J.L.: Singularity formation in the one-dimensional supercooled Stefan problem, *European J. Appl. Math.*, **7**, 119–150 (1996)

[78] Hurd, T.: *Contagion! The Spread of Systemic Risk in Financial Networks*, Springer, New York (2016)

[79] Inglis, J., Talay, D.: Mean-Field Limit of a Stochastic Particle System Smoothly Interacting Through Threshold Hitting-Times and Applications to Neural Networks with Dendritic Component. *SIAM J. Math. Anal.* **47**, 3884–3916 (2015)

[80] Jakubowski, A.: The Almost Sure Skorokhod Representation for Subsequences in Nonmetric Spaces. *Theory Probab. Appl.* **42**, 167–175 (1998)

[81] Kallianpur, G. and Xiong, J.: *Stochastic Differential Equations in Infinite Dimensional Spaces*. Institute of Mathematical Statistics: Lecture notes, monograph series (1995)

[82] Kaushansky, V., Lipton, A., Reisinger, C.: Semi-analytical solution of a McKean-Vlasov equation with feedback through hitting a boundary. Preprint. Available at `arXiv:1808.05311` (2018)

[83] Kaushansky, V., Reisinger, C.: Simulation of particle systems interacting through hitting times. To appear in *Discrete and Continuous Dynamical Systems*. Available at `arXiv:1805.11678` (2018)

[84] Lando, D., Nielsen, M.: Correlation in corporate defaults: Contagion or conditional independence? *J. of Finan. Intermediation* **19**, 355–372 (2010)

[85] Ledger, S.: Sharp regularity near an absorbing boundary for solutions to second order SPDEs in a half-line with constant coefficients. *Stoch. Partial Differ. Equ. Anal. Comput.* **2**, 1–26 (2014)

[86] Ledger, S.: Skorokhod's M1 topology for distribution-valued processes. *Electron. Commun. Probab.* **21**, 1–11 (2016)

[87] Ledger, S., Søjmark, A.: At the mercy of the common noise: blow-ups in a conditional McKean–Vlasov problem. Preprint. Available at `arXiv:1807.05126` (2018)

[88] Ledger, S., Søjmark, A.: Uniqueness for contagious McKean–Vlasov systems in the weak feedback regime. Preprint. Available at `arXiv:1811.12356` (2018)

[89] Lepoutrea, T., Meunier, N., Muller, N.: Cell polarisation model: The 1D case. *J. Maths. Pures Appl.*, **101**, 152–171 (2014)

[90] Lipton, A: Modern monetary circuit theory, stability of interconnected banking network, and balance sheet optimization for individual banks. *Int. J. Theor. Appl. Finance* **19**, 1–57 (2016)

[91] Merton, R. C.: On the pricing of corporate debt: the risk structure of interest rates. *J. Finance* **29**, 449–470 (1974)

[92] Moreno-Bote, R., Parga, N.: Response of integrate-and-fire neurons to noisy inputs filtered by synapses with arbitrary time scales: firing rate and correlations. *Neural Comput.* **22**, 1528–1572 (2010)

[93] Müller, M.S.: A stochastic Stefan-type problem under first-order boundary conditions. *Ann. Appl. Probab.* **28**, 2335–2369 (2018)

[94] Nadtochiy, S., Shkolnikov, M.: Particle systems with singular interaction through hitting times: application in systemic risk modeling. *Ann. Appl. Probab.* **29**, 89–129 (2019)

[95] Nadtochiy, S., Shkolnikov, M.: Mean field systems on networks, with singular interaction through hitting times. Preprint. Available at `arXiv:1807.02015` (2018)

[96] Ostojic, S., Brunel, N., Hakim, V.: Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *J. Comput. Neurosci.* **26**, 369–392 (2009)

[97] Pedersen, L.H.: When Everyone Runs for the Exit. *Int. J. Central Banking* **5**, 177–199 (2009)

[98] Peskir, G.: On Integral Equations Arising in the First-Passage Problem for Brownian Motion, *J. Integral Equations Appl.* **14**, 397–423 (2002)

[99] Rogers, L.C.G., Veraart, L.A.M.: Failure and rescue in an interbank network, *Management Science* **59**, 882–898 (2013)

[100] Schwartz, L.: *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*. Tata Institute of Fundamental Research, Oxford University Press, London (1973)

[101] Sowers, R., Spiliopoulos, K.: Default clustering in large pools: Large deviations. *SIAM J. Fin. Math.* **6**, 86–116 (2015)

[102] Spiliopoulos, K.: Systemic Risk and Default Clustering for Large Financial Systems. In: Friz, P., Gatheral, J., Gulisashvili, A., Jacqier, A., Teichmann, J. (eds.) *Large Deviations and Asymptotic Methods in Finance*, 529–557. Springer (2015)

[103] Sznitman, A.-S.: Topics in propagation of chaos. In Paul-Louis Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour XIX – 1989*, volume 1464 of *Lecture Notes in Mathematics*, Chapter 3, 165–251, Springer Berlin Heidelberg (1991)

[104] Ullner, E., Politi, A., Torcini, A.: Ubiquity of collective irregular dynamics in balanced networks of spiking neurons *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **28**, 1–6 (2018)

[105] Whitt, W.: *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer Series in Operations Research and Financial Engineering. Springer (2002)