The Annals of Applied Probability
2017, Vol. 27, No. 5, 2698–2752
DOI: 10.1214/16-AAP1256

A STOCHASTIC MCKEAN–VLASOV EQUATION FOR ABSORBING DIFFUSIONS ON THE HALF-LINE

BY BEN HAMBLY[1,*] AND SEAN LEDGER[2,†]

University of Oxford and Heilbronn Institute, University of Bristol†*

We study a finite system of diffusions on the half-line, absorbed when they hit zero, with a correlation effect that is controlled by the proportion of the processes that have been absorbed. As the number of processes in the system becomes large, the empirical measure of the population converges to the solution of a nonlinear stochastic heat equation with Dirichlet boundary condition. The diffusion coefficients are allowed to have finitely many discontinuities (piecewise Lipschitz) and we prove pathwise uniqueness of solutions to the limiting stochastic PDE. As a corollary, we obtain a representation of the limit as the unique solution to a stochastic McKean–Vlasov problem. Our techniques involve energy estimation in the dual of the first Sobolev space, which connects the regularity of solutions to their boundary behaviour, and tightness calculations in the Skorokhod M1 topology defined for distribution-valued processes, which exploits the monotonicity of the loss process L. The motivation for this model comes from the analysis of large portfolio credit problems in finance.

1. Introduction.

Motivation and framework. We prove the weak convergence of a system of interacting diffusions to the unique solution of a nonlinear stochastic PDE on the half-line. In our model, the diffusions are absorbed at the origin and the proportion of absorbed particles influences the diffusion coefficients, which leads to a description of the limiting system as the solution to a stochastic McKean–Vlasov problem. The motivation for studying the model in this paper is to extend the mathematical framework of [8] for the pricing of large portfolio credit derivatives to include processes whose dynamics are driven by statistics of the entire population. With more complicated interaction terms, the methods in [8] are no longer tractable and so we require new techniques. In particular, it is very difficult to analyse the correlation between pairs of particles in our model (an essential ingredient of [8]) and, from a practical perspective, it is desirable to allow the coefficients of the diffusions to be discontinuous, which presents a further complication.

Received April 2016; revised October 2016.

[1]Supported by the European Research Council under the European Union's Seventh Framework Programme (FP7-IDEAS-ERC, ERC grant agreement Nr. 291244).

[2]Supported by the Heilbronn Institute for Mathematical Research.

MSC2010 subject classifications. 60H15, 60H30, 60F15.

Key words and phrases. McKean–Vlasov problem, nonlinear SPDE, Skorokhod M1 topology.

Portfolio credit derivatives (such as the collateralised debt obligation—CDO) have a payoff structure which depends on the total notional value of the loss due to default of entities in the portfolio across the lifetime of the product, after a process of partial asset recovery takes place. We will not explore the financial details of these contracts (see [48]), but two important effects the modeller must capture are the intensity of defaults and the tendency for defaults to occur simultaneously. Common modelling approaches include copula-based models, in which the joint probability of default over a fixed time period is modelled directly, and reduced-form models, in which the default rates are modelled as correlated stochastic processes. The model we will consider is a *structural model*: default times are represented as the threshold hitting times of a collection of correlated stochastic processes. These models were introduced in the context of portfolio derivatives by [31] and [55], and their origins trace back to [5] and [44] for single-name derivatives.

Our general framework is as follows. Suppose we have a collection of $N \geq 1$ defaultable entities and a fixed finite time horizon $T > 0$. Assign the ith entity a risk process, $X^{i,N}$, called the *distance-to-default* process, with $\{X_0^{i,N}\}_{1\leq i\leq N}$ chosen to be positive independent random variables supported on $(0, \infty)$ with common law ν_0. Default of the ith entity is modelled as the first hitting time of zero of the distance-to-default process:

$$\tau^{i,N} := \inf\{t > 0 : X_t^{i,N} \leq 0\}. \tag{1.1}$$

The *empirical* and *loss processes* then track the spatial evolution of the surviving particles and the proportion of killed particles; defined respectively as

$$\nu_t^N := \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{t<\tau^{i,N}}\delta_{X_t^{i,N}}, \qquad L_t^N := \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\tau^{i,N}\leq t}. \tag{1.2}$$

Here, δ_x denotes the Dirac delta measure of the point $x \in \mathbb{R}$. The empirical process takes values in the sub-probability measures on $\mathbb{R}$ and the loss process takes values in $\mathbb{R}$. For $S \subseteq \mathbb{R}$, $\nu_t^N(S)$ is simply the proportion of the diffusions that take values in S at time t that have not yet hit the origin by time t:

$$\nu_t^N(S) = \frac{\#\{1 \leq i \leq N : X_t^{i,N} \in S \text{ and } t < \tau^{i,N}\}}{N},$$

hence we have the relationship

$$L_t^N = 1 - \nu_t^N(0, \infty).$$

In practice, once the dynamics of $X^{i,N}$ have been specified, the model could be used to generate realisations of L^N from which portfolio credit derivatives (options on L^N) could be priced using Monte Carlo routines. Instead, we will approximate L^N by its limit as $N \to \infty$. This is known as a large portfolio approximation, an idea first introduced in [51] and now found in several modern frameworks for copula-based models [13, 27, 43] and reduced-form models [22, 23, 45]. We will return to the question of how this approximation is generated in practice after a precise description of the limiting objects and mode of convergence.

Model specification. We will model the processes $\{X^{i,N}\}_{1\le i\le N}$ as correlated diffusions with parameters that are functions of the current proportional loss:

$$
\begin{aligned}
X_t^{i,N} &= X_0^i + \int_0^t \mu(s, X_s^{i,N}, L_s^N)\,ds + \int_0^t \sigma(s, X_s^{i,N})\rho(s, L_s^N)\,dW_s \\
&\quad + \int_0^t \sigma(s, X_s^{i,N})(1-\rho(s, L_s^N)^2)^{\frac{1}{2}}\,dW_s^i.
\end{aligned}
\tag{1.3}
$$

Here, $W, W^1, W^2, \ldots$ are independent standard Brownian motions and the precise conditions on the coefficients are given in Assumption 2.1. In particular, we assume ρ is piecewise Lipschitz with finitely many discontinuities in the loss variable $\ell \mapsto \rho(s, \ell)$. (It is easy, but perhaps not immediate, to show that this collection of processes exists; see Remark 2.2.)

In [8], this model is analysed for the case when the coefficients are constants and it is shown that the sequence of empirical process, $(\nu^N)_{N\ge 1}$, converges to a stochastic limit which can be characterised as the unique solution to a heat equation with constant coefficients and a random transport term driven by the systemic Brownian motion W [8], Theorem 1.1. However, numerical experiments show that the constant coefficient model is too simple to adequately capture the traded prices of CDOs across all tranches simultaneously [8], Section 5. This problem is common for Gaussian models—the tails of the risk processes are too light to produce large losses and so a large correlation parameter is required to generate scenarios in which many defaults occur over a given time horizon [26, 48]. Consequently, different products on the same underlying portfolio may produce different correlation parameters when calibrated to market prices. This phenomenon is known as *correlation skew* (see Figure 1).

There is a large literature addressing the drawbacks of Gaussian credit models. Examples include the addition of jump processes and heavy-tailed distributions [25, 41, 54], stochastic parameters and inhomogeneity [2, 7] and contagion effects [17, 28, 29]. Close relatives to our framework include [6], in which a jump process is added to the systemic factor, but in a discretised version of the system, and [32], in which the particles are taken to be general diffusions. In [1], the constant coefficient model is studied on the unit interval with absorbing boundaries at 0 and 1 and with an additional multiplicative killing rate as a model for mortgage pools.

Our present approach is inspired by Figure 1. Suppose μ and σ are fixed constants and ρ is only a function of ℓ. If $\ell \mapsto \rho(\ell)$ was piecewise constant across intervals corresponding to the CDO tranches in Figure 1, then an obvious strategy for calibrating ρ to the market prices is to calibrate the first level of ρ to the traded spread of the most junior tranche, fix this value, repeat the calibration procedure for the next most junior tranche spread and continue for all tranches. It is therefore a

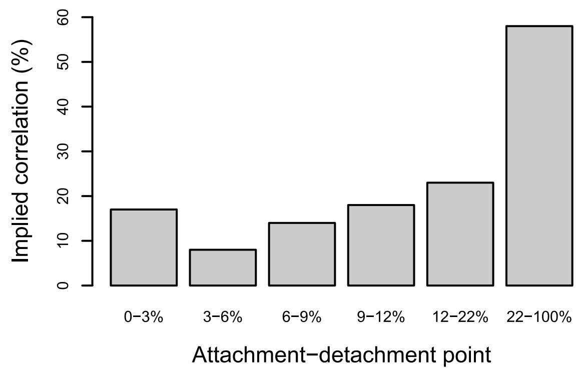


FIG. 1. *Implied correlation for each tranche for the data set from* [8], *Figure* 2, 7 *year maturity. With* ν_0, μ *and* σ *fixed in the constant coefficient model, the implied correlation for a given tranche is the value of the correlation parameter required to give a model spread equal to the market spread for that tranche. This is an example of correlation skew.*

natural assumption to allow the diffusion coefficients in (1.3) to have finitely many discontinuities. Piecewise Lipschitz coefficients encompass this class of models whilst giving an analytically tractable system.

Main results. The dynamics of an individual distance-to-default process, $X^{i,N}$, are controlled by the population behaviour, hence we have an example of a McKean–Vlasov system; see [50] for an overview. Some applications of these systems include the modelling of large collections of neurons and threshold hitting times for membrane potential levels in mathematical neuroscience [21, 42], the modelling of a large number of noncooperative agents in mean-field games [10, 12], filtering theory [3, 16] and mathematical genetics [18]. Examples in portfolio credit modelling include [17, 49] in which systems with contagion effects are analysed under their large population limits.

As $N \to \infty$, we will find that the influence of the idiosyncratic Brownian drivers, $W^1, W^2, \dots$, averages-out to a deterministic effect, but that the randomness due to the systemic Brownian motion, W, remains present. Hence, the system must be characterised as the pair (ν^N, W), and we will follow an established strategy to demonstrate the convergence in law of this pair and to characterise the limiting law:

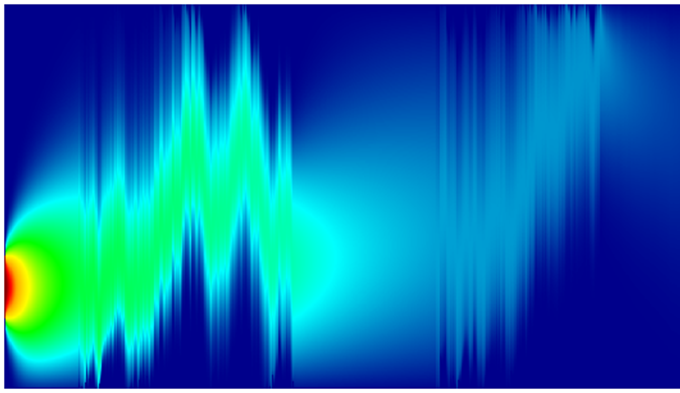
(i) Prove tightness of $(\nu^N, W)_{N\geq 1}$ (in a suitable topology).
(ii) Characterise the limit points as weak solutions of a nonlinear evolution equation.
(iii) Prove uniqueness of solutions for this equation.
(iv) Conclude all limiting laws agree, and hence that we have convergence in law.

The mathematical challenge comes from the interaction of the individuals through the boundary behaviour of the population and the discontinuities in the diffusion coefficients. A similar model has recently been studied where the particles interact through the quantiles of the empirical measure [15], however, there is no general uniqueness theory for this problem. For a model without systemic noise, there is a uniqueness theory in [35]. Discontinuous coefficients have been considered in [14], but only on the whole space and in the deterministic setting. In our model, parameter discontinuities are allowed because the limiting realisations of the loss process are strictly increasing (Proposition 4.6). This implies the infinite system spends a null set of time at points where the discontinuities in the coefficients prevent the application of the continuous mapping theorem (Corollary 5.7). Stochastic PDEs of McKean–Vlasov type are popular tools in the analysis of mean-field games with common noise [11, 36]. In [19, 20], a system of diffusions on the half-line is studied in which each particle undergoes a proportional jump towards zero whenever any of the particles hits the absorbing boundary at zero. The purpose of the model is to describe the self-excitatory behaviour of a large collection of neurons. For small values of the feedback parameter, existence and uniqueness theorems hold for the limiting system. It is shown in [9], however, that for large values of the feedback parameter the limiting system must blow-up (in the sense that no continuous solutions exist) and a complete existence and uniqueness theory in this case remains a challenge.

The topology we will use for establishing tightness of the sequence of laws of $(\nu^N, W)_{N\geq 1}$ is the product topology $(D_{\mathscr{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \mathrm{U})$, where $(D_{\mathscr{S}'}, \mathrm{M1})$ is the M1 topological space of distribution-valued càdlàgprocesses on $[0, T]$, introduced in [40], and $(C_{\mathbb{R}}, \mathrm{U})$ is the space of real-valued continuous functions on $[0, T]$ with the topology of uniform convergence. (Throughout, $\mathscr{S}$ denotes the space of rapidly decreasing functions and $\mathscr{S}'$ the space of tempered distributions.) It will not be necessary to explain the full details of the construction of $(D_{\mathscr{S}'}, \mathrm{M1})$, as the proof Theorem 1.1 uses only Theorem 3.2 and Proposition 2.7 of [40], together with facts about the classical M1 topology on $D_{\mathbb{R}}$. The M1 topology is helpful because monotone real-valued processes are automatically tight in $(D_{\mathbb{R}}, \mathrm{M1})$, a fact which has been exploited in many other applications (see [40] for references). In our infinite-dimensional setting, the decomposition trick in [40], Proposition 4.2, enables us to exploit the monotonicity of the loss process in proving tightness of the empirical process. Tightness on the product space implies the existence of subsequential limit points, whereby we recover the following.

THEOREM 1.1 (Existence). *Let (ν, W) realise a limiting law of the sequence $(\nu^N, W)_{N\geq 1}$. Then ν is a continuous process taking values in the sub-probability measures and satisfies the regularity conditions of Assumption 2.3 and the* limit SPDE:

$$\nu_t(\phi) = \nu_0(\phi) + \int_0^t \nu_s\big(\mu(s, \cdot, L_s)\partial_x\phi\big)\, ds + \frac{1}{2}\int_0^t \nu_s\big(\sigma^2(s, \cdot)\partial_{xx}\phi\big)\, ds$$
$$+ \int_0^t \nu_s\big(\sigma(s, \cdot)\rho(s, L_s)\partial_x\phi\big)\, dW_s \qquad \textit{with } L_t = 1 - \nu_t(\mathbf{1}_{(0,\infty)}),$$



$L^{-1}(1/5), \quad L^{-1}(2/5), \quad L^{-1}(3/5), \quad L^{-1}(4/5)$

$$\rho(t,\ell) = \begin{cases} 0, & \text{if } \ell \in [0, \frac{1}{5}) \cup [\frac{2}{5}, \frac{3}{5}) \cup [\frac{4}{5}, 1], \\ \frac{9}{10}, & \text{if } \ell \in [\frac{1}{5}, \frac{2}{5}) \cup [\frac{3}{5}, \frac{4}{5}). \end{cases}$$

FIG. 2. *Heat plot for the solution*, ν, *of the limit SPDE for a fixed sample path of* W. *Time is plotted on the horizontal axis*, *space on the vertical axis and the value of a pixel represents the* (*scaled*) *intensity of* ν *at that space-time point* (*blue for level zero increasing to dark red for maximal value*). *The initial condition is a step function*, $\mu = 0$, $\sigma = 1$ *and* ρ *is given above. Markers are added to show the times at which the loss process*, L, *reaches levels* $1/5$, $2/5$, $3/5$ *and* $4/5$. *Notice the corresponding three periods of smooth heat flow between the two periods of highly correlated motion.* (*Figure produced using the algorithm outlined in Section* 10.)

for every $t \in [0, T]$ *and* $\phi \in C^{\text{test}} := \{\phi \in \mathscr{S} : \phi(0) = 0\}$, *with probability* 1. *Furthermore, if the limit point is attained along the subsequence* $(\nu^{N_k}, W)_{k\geq 1}$, *then* $(L^{N_k}, W)_{k\geq 1}$ *converges in law to* (L, W) *on the product space* $(D_{\mathbb{R}}, \mathrm{M1}) \times (C_{\mathbb{R}}, \mathrm{U})$.

The limit SPDE is a nonlinear heat equation with stochastic transport term driven by the systemic Brownian motion (see Figure 2 for an example with an exaggerated correlation change), and the space of test functions, C^{test}, encodes the Dirichlet boundary conditions. In the limit, the idiosyncratic noise averages-out to produce the diffusive evolution equation. The intuition for this effect is explained easily in Section 3, however, a full proof of Theorem 1.1 requires more technical details and is given in Section 5. Several estimates involving purely probabilistic arguments are presented in Section 4, where a key result is Proposition 4.6 which shows (in an asymptotic sense) that over any nonzero time interval the system must lose a nonzero proportion of mass, and hence any limiting loss process is strictly increasing.

With Theorem 1.1 established, demonstrating the full weak convergence of $(\nu^N, W)_{N\geq 1}$ is a matter of proving uniqueness of solutions to the limit SPDE.

THEOREM 1.2 (Uniqueness/Law of large numbers). *Let ν_0 satisfy Assumption 2.1. Suppose that (ν, W) realises a limiting law of $(\nu^N, W)_{N\geq 1}$ and that $\tilde{\nu}$ satisfies Assumption 2.3. If ν and $\tilde{\nu}$ solve the limit SPDE in Theorem 1.1 with respect to W and ν_0, then with probability* 1

$$\nu_t(S) = \tilde{\nu}_t(S) \qquad \textit{for every } t \in [0, t] \textit{ and Borel measurable } S \subseteq \mathbb{R}.$$

Hence, there exists a unique law of a solution to the limit SPDE on $(D_{\mathcal{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \mathrm{U})$ and $(\nu^N, W)_{N\geq 1}$ converges weakly to this law. Furthermore, if (ν, W) realises the unique law, then $(L^N, W)_{N\geq 1}$ converges in law to (L, W) on $(D_{\mathbb{R}}, \mathrm{M1}) \times (C_{\mathbb{R}}, \mathrm{U})$, where $L_t = 1 - \nu_t(0, \infty)$.

REMARK 1.3 (Strong solutions). Theorem 1.2 shows that all weak solutions realise limiting laws, and amongst limiting laws we have pathwise uniqueness. Following [33], Corollary 5.3.23, we deduce that strong solutions exist on a sufficiently rich probability space, whereby ν (and hence L) is adapted to the filtration generated by W.

REMARK 1.4 (Density). In Corollary 7.4, we show that ν has a density process $V_t \in L^2(0, \infty)$ such that $\nu_t(\phi) = \int_0^\infty \phi(x) V_t(x)\, dx$ for all $\phi \in L^2(0, \infty)$ and $t \in [0, T]$. It is then instructive to write the limit SPDE formally as

$$\begin{aligned} V_t(x) = V_0(x) &- \int_0^t \partial_x\big(\mu(s, \cdot, L_s) V_s(\cdot)\big)\, ds + \frac{1}{2}\int_0^t \partial_{xx}\big(\sigma^2(s, \cdot) V_s(\cdot)\big)\, ds \\ &- \int_0^t \rho(s, L_s)\partial_x\big(\sigma(s, \cdot) V_s(\cdot)\big)\, dW_s \qquad \text{with } V_t(0) = 0. \end{aligned}$$

To prove Theorem 1.2 (Section 7), we use the kernel smoothing method from [8], which is a technique for mollifying potentially exotic solutions to the limit SPDE in order to work with smooth tractable objects, at the expense of a small approximation error. The technique was used on the whole space in [37, 38]. In [8], the approximation error is controlled in the space $L^2(0, \infty)$ and there the key quantity to control is the second moment of the mass near the origin: $\mathbf{E}\nu_t(0, \varepsilon)^2$, for a candidate solution ν. This approach succeeds because the quantity can be written in terms of the law of a two-dimensional Brownian motion in a wedge, for which explicit formulae are available. In that case, the kernel smoothing method can be used to give a precise description of the regularity of the solution [39]. As the particle interactions in our model are more complicated, however, these explicit formula are no longer available. Although we are able to show that the unique solution to the limit SPDE has a density in L^2 (Corollary 7.4), which is an auxiliary result towards Theorem 1.2, that method cannot be used to fully establish uniqueness as it relies on a crude upper bound for ν which neglects the effect of the absorbing boundary (Remark 7.5). Our solution to this problem is to adapt the kernel smoothing method to the dual of the first Sobolev space, which then only requires us to control the first moment $\mathbf{E}\nu_t(0,\varepsilon)$ (Section 6). This is an easier quantity to estimate as only individual particles need to be studied and not pairs of particles, hence we do not need to consider the complicated correlation between particles (see Propositions 4.4 and 5.6).

We must also deal with discontinuities in the coefficients of the limit SPDE and here the strict monotonicity of the limiting loss processes is again important. Our strategy is to prove uniqueness up to the first time the level of the loss reaches a discontinuity point of the coefficients, whereby continuity allows us to propagate the argument onto the next such time interval. With a strictly increasing loss process and only finitely many discontinuities, this argument terminates after finitely many iterations, whereby we have uniqueness on the whole time horizon $[0,T]$.

REMARK 1.5 (Pathological ρ). We cannot choose ρ arbitrarily and expect Theorem 1.2 to hold. As an example, let $\mu=0$, $\sigma=1$ and

$$\rho(t,\ell)=\begin{cases} q^{-1}, & \text{if } \ell = kq^{-n} \text{ for some prime } q, n\in\mathbb{N} \text{ and } 1\le k\le q^n-1,\\ 0, & \text{otherwise.}\end{cases}$$

For $N=q^n$, L^N is supported on $\{kq^{-n}\}_{0\le k\le q^n}$, hence ν^N behaves as the basic constant correlation system with $\rho=q^{-1}$, which we denote $\nu|_{\rho=q^{-1}}$. Therefore, $(\nu^{q^n})_{n\ge 1}$ converges weakly to $\nu|_{\rho=q^{-1}}$ as $n\to\infty$, hence there is a distinct limit point for every prime, so weak convergence fails for this example.

In Section 9, we recast our results as a stochastic McKean–Vlasov problem (with randomness from W) and this shows that ν can be written as the conditional law of a single tagged particle.

THEOREM 1.6 (Stochastic McKean–Vlasov problem). *Let* (ν, W) *be a strong solution to the limit SPDE* (*Remark* 1.3). *For any independent Brownian motion*, $W^{\perp}$, *there exists a continuous real-valued process*, X, *satisfying*

$$\begin{cases} X_t = X_0 + \int_0^t \mu(s,X_s,L_s)\,ds + \int_0^t \sigma(s,X_s)\rho(s,L_s)\,dW_s \\ \qquad + \int_0^t \sigma(s,X_s)\big(1-\rho(s,L_s)^2\big)^{\frac{1}{2}}\,dW_s^{\perp}, \\ \tau = \inf\{t>0 : X_t\le 0\}, \\ \nu_t(\phi) = \mathbf{E}\big[\phi(X_t)\mathbf{1}_{t<\tau}\,|\,W\big], \\ L_t = \mathbf{P}(\tau\le t\,|\,W). \end{cases}$$

(*Here*, X_0 *has law* ν_0 *and is independent of all other random variables*.) *Furthermore*, *the law of* (X, W) *is unique*.

Returning to the question of applying our model, regarding a portfolio credit derivative as an option on the loss process, L, with some payoff function, Ψ :

$D_{\mathbb{R}} \to \mathbb{R}$, the main practical question is how to accurately estimate $\mathbf{E}\Psi(L)$. This comes in two parts: we must first generate an approximation to L (through ν) to a given level of precision for a fixed Brownian trajectory and then we must combine such estimates into a random sample. In Section 10, we give an outline of a discrete-time algorithm for approximating the system and some potential variance reduction techniques. We leave the tasks of checking the benefits and correctness of these methods as open problems. A number of potential modifications to the model are also stated, along with their corresponding mathematical challenges.

Overview. In Section 2, we state the main technical assumptions on the model parameters and review their purpose. In Section 3, we derive the evolution equation satisfied by the empirical measure of the finite system, which gives a heuristic explanation for arriving at the limit SPDE in Theorem 1.1. In Section 4, several probabilistic estimates are derived for the finite system and these are applied in Section 5 to give a proof of Theorem 1.1. In Section 6, we describe the kernel smoothing method, which is the main tool for the proof of Theorem 1.2 in Section 7. In Section 8, several technical lemmas are presented which are used to in Section 7, but which are deferred for readability. In Section 9, we use our results to give a short proof of Theorem 1.6. In Section 10, we outline an algorithm for simulating the solution to the limit SPDE and discuss open problems relating to this and to potential model extensions.

2. Notation and assumptions.

The purpose of this section is to lay out the technical definitions omitted in the introduction and to explain their purpose.

ASSUMPTION 2.1 (Coefficient assumptions). Let $\mu : [0, T] \times \mathbb{R} \times [0, 1] \to \mathbb{R}$, $\sigma : [0, T] \times \mathbb{R} \to [0, \infty)$ and $\rho : [0, T] \times [0, 1] \to [0, 1)$ be the coefficients in (1.3) and ν_0 be the common law of the initial values of the distance-to-default processes introduced above (1.1). We assume that we have a sufficient large constant, $C \in (1, \infty)$, such that all the following hold:

(i) (Initial condition.) The probability measure ν_0 is supported on $(0, \infty)$, has a density $V_0 \in L^2(0, \infty)$ and satisfies

$$\nu_0(\lambda, \infty) = o\big(\exp\{-\alpha\lambda\}\big) \qquad \text{as } \lambda \to +\infty$$

for every $\alpha > 0$. [Note: $V_0 \in L^2(0, \infty)$ implies $\nu_0(0, \varepsilon) = O(\varepsilon^{1/2}) = o(1)$ as $\varepsilon \to 0$.]

(ii) (Spatial regularity.) For all fixed $t \in [0, T]$ and $\ell \in [0, 1]$, $\mu(t, \cdot, \ell), \sigma(t, \cdot) \in C^2(\mathbb{R})$ with

$$\big|\partial_x^n \mu(t, x, \ell)\big|, \big|\partial_x^n \sigma(t, x)\big| \le C$$

for all $t \in [0, T]$, $x \in \mathbb{R}$, $\ell \in [0, 1]$ and $n = 0, 1, 2$.

(iii) (Nondegeneracy.) For all $t \in [0, T]$, $x \in \mathbb{R}$, $\ell \in [0, 1]$

$$\sigma(t, x) \geq C^{-1} > 0, \qquad 0 \leq \rho(t, \ell) \leq 1 - C^{-1} < 1.$$

(iv) (Piecewise Lipschitz in loss.) There exists $0 = \theta_0 < \theta_1 < \cdots < \theta_k = 1$ such that

$$\left|\mu(t, x, \ell) - \mu(t, x, \bar{\ell})\right|, \left|\rho(t, \ell) - \rho(t, \bar{\ell})\right| \leq C|\ell - \bar{\ell}|,$$

whenever $t \in [0, T]$, $x \in \mathbb{R}$ and both $\ell, \bar{\ell} \in [\theta_{i-1}, \theta_i)$ for some $i \in \{1, 2, \dots, k\}$.

(v) (Integral constraint.) $\sup_{s\in[0,T]} \int_0^\infty |\partial_t \sigma(s, y)|\, dy < \infty$.

REMARK 2.2 ($X^{i,N}$ well-defined). To see that, we can find $\{X^{i,N}\}_{1\leq i\leq N}$ satisfying (1.3) notice that initially $L = 0$, so we can find N diffusions satisfying (1.3) up to the first time one of the diffusions hits the origin [i.e., with coefficients of the form $g(t, x, 0)$]—notice that the coefficients are globally Lipschitz by (ii) of Assumption 2.1, so standard diffusion theory applies. At this stopping time $L^N = 1/N$, and so the process can be restarted as a diffusion with coefficients $g(t, x, 1/N)$. This gives a solution up to the first time two particles have hit the origin. Repeating this argument gives the construction of $\{X^{i,N}\}_{1\leq i\leq N}$.

Condition (i) ensures that limiting realisations of the system satisfy the regularity conditions in Assumption 2.3, as required for Theorem 1.1. The tail assumption and boundary behaviour of ν_0 are used in Propositions 4.4 and 4.5 to show that ν^N inherits the corresponding properties at times $t > 0$, and this is transferred to limit points by Proposition 5.6.

The boundedness assumption on the coefficients, given by the case $n = 0$ in condition (ii), is used many times throughout this paper. The cases $n = 1$ and 2 are used in Lemmas 4.1 and 4.2 to relate the law of $X^{1,N}$ to that of a standard Brownian motion, and in Lemmas 8.1 and 8.2 to interchange coefficients and measures in the proof of Theorem 1.2.

Condition (iii) implies that there is always a diffusive effect acting on the system, and this ensures that the limiting system does not become degenerate. If $\sigma = 0$ or $\rho = 1$, then the particles are completely dependent and move according to a drift term given by μ and W. The assumption that ρ is bounded away from 1 is used directly in the proof of Theorem 1.2 in (7.2) and (7.7).

Condition (iv) is the main motivating assumption, which we have discussed at length in Section 1.

Condition (v) is purely a technical assumption to ensure that the drift coefficient, D, in Lemma 4.1 is uniformly bounded by a deterministic constant.

Finally, we will remark on the specific form of $\sigma = \sigma(t, x)$ and $\rho = \rho(t, \ell)$. From (1.3) we can write the dynamics of a single particle as

$$dX_t^{i,N} = \mu\big(t, X_t^{i,N}, L_t^N\big)\, dt + \sigma\big(t, X_t^{i,N}\big)\, dB_t^i,$$

where B^i is a Brownian motion. Although the $\{B^i\}_i$ are coupled through L^N, this representation allows us to relate the law of an *individual* particle to a standard Brownian motion as in Lemmas 4.1 and 4.2, since μ is bounded and σ is independent of L^N. A second advantage of the taking σ and ρ in this form is that the pairwise correlation between particles is purely a function of $\rho(t, L_t^N)$, and so is the same for all pairs. This is explicitly made use of in the construction of the time-change defined in (4.8), and there it is again important that the correlation function is bounded strictly away from 1, so that the system can be compared to a standard multi-dimensional Brownian motion.

Below are the constraints we place on solutions to the limit SPDE in Theorem 1.2 to ensure that we have uniqueness. As Theorem 1.1 indicates, these conditions are natural in the sense that all limit points of the finite system satisfy them.

ASSUMPTION 2.3 (Regularity conditions). Let ν be a càdlàg process taking values in the space of sub-probability measures on $\mathbb{R}$. The regularity conditions on ν are:

(i) (Loss function.) The process defined by $L_t := 1 - \nu_t(0, \infty)$ is nondecreasing at all times and is strictly increasing when $L_t < 1$.

(ii) (Support.) For every $t \in [0, T]$, ν_t is supported on $[0, \infty)$.

(iii) (Exponential tails.) For every $\alpha > 0$,

$$\mathbf{E} \int_0^T \nu_t(\lambda, +\infty)\, dt = o\big(e^{-\alpha\lambda}\big) \qquad \text{as } \lambda \to \infty.$$

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

$$\mathbf{E} \int_0^T \nu_t(0, \varepsilon)\, dt = O\big(\varepsilon^{1+\beta}\big) \qquad \text{as } \varepsilon \to 0.$$

(v) (Spatial concentration.) There exists $C > 0$ and $\delta > 0$ such that

$$\mathbf{E} \int_0^T \big|\nu_t(a, b)\big|^2\, dt \le C|b - a|^\delta \qquad \text{for all } a < b.$$

It is essential that limit points satisfy condition (i) in order to apply the continuous mapping theorem to recover the limit SPDE for limit points (Corollary 5.7). There, strict monotonicity ensures that there are only finitely many t such that $L_t = \theta_i$ for some i, and hence that this set of times is negligible in the limit. Knowing that L is monotone also allows us to split $[0, T]$ into consecutive intervals such that in the ith interval $L_t \in [\theta_i, \theta_{i+1})$, and this argument is used in the uniqueness proof in Section 7 (Case 2).

Condition (ii) is natural since ν^N is supported on $[0, \infty)$ by construction. However, it is also convenient to take our test functions, C^{test}, to be supported on $\mathbb{R}$, hence (ii) is needed to rule out pathological solutions that have support on the negative half-line and that would otherwise break the uniqueness claim.

Condition (iii) is used several times throughout Section 8 to check various integrability requirements. It is also used in Lemma 8.8 to relate ν and L via the H^{-1} norm.

Condition (iv) is the key boundary estimate discussed in Section 1. Its main use is in Lemma 7.6.

Condition (v) guarantees that solutions cannot become too concentrated in spatial locations. This is used to interchange coefficients and measures in Lemmas 8.1 and 8.2.

3. Dynamics of the finite particle system. This section introduces the empirical process approximation to the limit SPDE from Theorem 1.1 and explains the intuition behind the convergence of $(\nu^N)_{N\geq 1}$. Throughout, we will drop the dependence of the coefficients on the time, space and loss variables and use the following shorthand when it is safe to do so.

REMARK 3.1 (Shorthand notation). For fixed N, when there is no confusion, we may use the functional notation:

$$\mu_t = \mu(t, \cdot, L_t^N), \qquad \sigma_t = \sigma(t, \cdot), \qquad \rho_t = \rho(t, L_t^N).$$

PROPOSITION 3.2 (Finite evolution equation). *For every $N \geq 1$, $t \in [0, T]$ and $\phi \in C^{\text{test}}$*

$$\begin{aligned} \nu_t^N(\phi) = \nu_0^N(\phi) &+ \int_0^t \nu_s^N(\mu_s \partial_x \phi)\, ds + \frac{1}{2}\int_0^t \nu_s^N(\sigma_s^2 \partial_{xx}\phi)\, ds \\ &+ \int_0^t \nu_s^N(\sigma_s \rho_s \partial_x \phi)\, dW_s + I_t^N(\phi), \end{aligned}$$

where we have the idiosyncratic driver

$$I_t^N(\phi) := \frac{1}{N}\sum_{i=1}^N \int_0^t \sigma(s, X_s^{i,N})(1-\rho(s, L_s^N)^2)^{\frac{1}{2}} \partial_x \phi(X_s^{i,N}) \mathbf{1}_{s<\tau^{i,N}}\, dW_s^i.$$

PROOF. Apply Itô's formula to $\phi(X^{i,N})$ to obtain

$$\begin{aligned} \phi(X_{t\wedge\tau^{i,N}}^{i,N}) = \phi(X_0^{i,N}) &+ \int_0^t (\mu_s \partial_x \phi)(X_s^{i,N}) \mathbf{1}_{s<\tau^{i,N}}\, ds \\ &+ \frac{1}{2}\int_0^t (\sigma_s^2 \partial_{xx}\phi)(X_s^{i,N}) \mathbf{1}_{s<\tau^{i,N}}\, ds \\ &+ \int_0^t (\sigma_s \rho_s \partial_x \phi)(X_s^{i,N}) \mathbf{1}_{s<\tau^{i,N}}\, dW_s \\ &+ \int_0^t (\sigma_s (1-\rho_s^2)^{\frac{1}{2}} \partial_x \phi)(X_s^{i,N}) \mathbf{1}_{s<\tau^{i,N}}\, dW_s^i. \end{aligned}$$

If $\phi \in C^{\text{test}}$, then

$$\phi(X^{i,N}_{t\wedge\tau^{i,N}}) = \phi(X^{i,N}_t)\mathbf{1}_{t<\tau^{i,N}}. \tag{3.1}$$

Substituting this expression into the left-hand side above, summing over $i \in \{1, 2, \dots, N\}$ and multiplying by N^{-1} gives the result. $\square$

REMARK 3.3. We need to ensure that our test functions satisfy $\phi(0) = 0$ so that equation (3.1) is valid.

Since the idiosyncratic noise, I^N, is a sum of martingales with zero covariation, the process converges to zero in the limit as $N \to \infty$. This explains why we arrive at the limit SPDE in Theorem 1.1.

PROPOSITION 3.4 (Vanishing idiosyncratic noise). *For every* $\phi \in C^{\text{test}}$,

$$\mathbf{E} \sup_{t\in[0,T]} \left|I^N_t(\phi)\right|^2 = \|\partial_x\phi\|^2_\infty \cdot O(N^{-1}) \qquad \text{as } N \to \infty.$$

PROOF. Since σ and $\partial_x\phi$ are bounded, the result follows from Doob's martingale inequality and the fact that

$$\left[I^N_\cdot(\phi)\right]_t = \frac{1}{N^2}\sum_{i=1}^N \int_0^t \sigma(s, X^{i,N}_s)^2(1-\rho(s, L^N_s)^2)\partial_x\phi(X^{i,N}_s)^2\, ds. \qquad \square$$

The whole space process. In the proceeding sections, it will be useful to work with the process defined by

$$\bar{\nu}^N_t := \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}_t}, \tag{3.2}$$

which is a probability-measure valued processes on the whole of $\mathbb{R}$. Clearly, it is the case that

$$\nu^N_t(S) \le \bar{\nu}^N_t(S) \qquad \text{for all } N \ge 1, t \in [0, T] \text{ and } S \subseteq \mathbb{R}. \tag{3.3}$$

Since $\bar{\nu}^N$ is not affected by the absorbing boundary, from the work in Proposition 3.2 it follows that $\bar{\nu}^N$ satisfies the same evolution equation as ν^N, but on the whole space. This is encoded through the test functions.

PROPOSITION 3.5 (Evolution of $\bar{\nu}^N$). *For every* $N \ge 1$, $t \in [0, T]$ *and* $\phi \in \mathscr{S}$,

$$\begin{aligned}\bar{\nu}^N_t(\phi) = \nu^N_0(\phi) &+ \int_0^t \bar{\nu}^N_s(\mu_s\partial_x\phi)\, ds + \frac{1}{2}\int_0^t \bar{\nu}^N_s(\sigma^2_s\partial_{xx}\phi)\, ds \\ &+ \int_0^t \bar{\nu}^N_s(\sigma_s\rho_s\partial_x\phi)\, dW_s + \bar{I}^N_t(\phi),\end{aligned}$$

where

$$\bar{I}_t^N(\phi) := \frac{1}{N}\sum_{i=1}^{N}\int_0^t \sigma(s, X_s^{i,N})(1-\rho(s, L_s^N)^2)^{\frac{1}{2}}\partial_x\phi(X_s^{i,N})\,dW_s^i.$$

4. Probabilistic estimates. Here, we collect the main probabilistic estimates used in later proofs. The reader may wish to skip this section and use it only as a reference. We begin by noting the following simple result, which is just a consequence of the fact that $\{X^{i,N}\}_i$ are identically distributed: for any measurable $S \subseteq \mathbb{R}$, $N \geq 1$ and $t \in [0, T]$,

$$\mathbf{E}\nu_t^N(S) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{E}[\mathbf{1}_{X_t^{i,N}\in S; t<\tau^{i,N}}] = \mathbf{P}(X_t^{1,N} \in S; t < \tau^{1,N}). \tag{4.1}$$

Under $\mathbf{P}$, $X^{1,N}$ is a diffusion and with Lemmas 4.1 and 4.2 we are able to estimate (4.1) for relevant choices of S by relating the law of $X^{1,N}$ to that of standard Brownian motion. Specifically, in Corollary 4.3 and Propositions 4.4 and 4.5 we show that ν^N satisfies the corresponding estimates to those in Assumption 2.3(iii), (iv) and (v), which is of direct use in Proposition 5.6 when we take a limit as $N \to \infty$. In Propositions 4.6 and 4.7, we prove two estimates for which (4.1) is not helpful. These results require us to express the quantities of interest in terms of independent particles to show that certain events concerning the increments in the loss process are asymptotically negligible.

LEMMA 4.1 (Scale transformation). *Define* $\zeta : [0, T] \times \mathbb{R} \to \mathbb{R}$ *by*

$$\zeta(t, x) := \int_0^x \frac{dy}{\sigma(t, y)}$$

and $Z_t := \zeta(t, X_t^{1,N})$. *Then* $\operatorname{sgn}(Z_t) = \operatorname{sgn}(X_t^{1,N})$ *and* $dZ_t = D_t\,dt + dB_t$ *where* B *is the Brownian motion*

$$B_t = \int_0^t \rho(s, L_s^N)\,dW_s + \int_0^t (1-\rho(s, L_s^N)^2)^{\frac{1}{2}}\,dW_s^1$$

and the drift coefficient, D, *is given by*

$$D_t = \left(\frac{\mu}{\sigma} - \partial_x\sigma\right)(t, X_t^{1,N}, L_t^N) - \int_0^{X_t^{1,N}} \frac{\partial_t\sigma}{\sigma^2}(t, y)\,dy,$$

which is uniformly bounded (*in* N *and* t).

PROOF. Straightforward application of Itô's formula coupled with Assumption 2.1. $\square$

LEMMA 4.2 (Removing drift). *For every* $\delta \in (0,1)$, *there exists* $c_\delta > 0$ *such that*

$$\mathbf{P}\big(X_t^{1,N} \in S; t < \tau^{1,N}\big) \le c_\delta F_t\big(\zeta(t,S)\big)^\delta \qquad \textit{for every measurable } S \subseteq \mathbb{R},$$

where F_t *is the marginal law of a killed Brownian motion at time* t *with initial distribution* $\nu_0 \circ \zeta(0,\cdot)^{-1}$ *and* ζ *is as defined in Lemma* 4.1. *Likewise, if* $\bar{F}$ *is the marginal law of the Brownian motion without killing at the origin and with the same initial distribution*:

$$\mathbf{P}\big(X_t^{1,N} \in S\big) \le c_\delta \bar{F}_t\big(\zeta(t,S)\big)^\delta \qquad \textit{for every measurable } S \subseteq \mathbb{R}.$$

PROOF. Let Z be as in Lemma 4.1, then $\tau^{1,N}$ is also the first hitting time, τ^Z, of 0 by Z so

$$\mathbf{P}\big(X_t^{1,N} \in S; t < \tau^{1,N}\big) = \mathbf{P}\big(Z_t \in \zeta(t,S); t < \tau^Z\big). \tag{4.2}$$

Apply Girsanov's theorem with the change of measure

$$\frac{d\mathbf{Q}}{d\mathbf{P}}\bigg|_{\mathcal{F}_t} = \exp\bigg\{-\int_0^t D_s\,dB_s - \frac{1}{2}\int_0^t D_s^2\,ds\bigg\} =: \Xi_t,$$

then under $\mathbf{Q}$, Z is a standard Brownian motion with $Z_0 = \zeta(0, X_0^{1,N})$, and, for any $E \in \mathcal{F}_t$ and $p^{-1} + q^{-1} = 1$, Hölder's inequality gives

$$\mathbf{P}(E) = \mathbf{E}_{\mathbf{Q}}\big[\Xi_t^{-1}\mathbf{1}_E\big] \le \mathbf{E}_{\mathbf{Q}}\big[\Xi_t^{-p}\big]^{\frac{1}{p}}\mathbf{Q}(E)^{\frac{1}{q}} = \mathbf{E}_{\mathbf{P}}\big[\Xi_t^{1-p}\big]^{\frac{1}{p}}\mathbf{Q}(E)^{\frac{1}{q}} \le C_q\mathbf{Q}(E)^{\frac{1}{q}},$$

for some constant $C_q > 0$ as D is uniformly bounded. Applying this bound to (4.2) gives

$$\mathbf{P}\big(X_t^{1,N} \in S; t < \tau^{1,N}\big) \le C_q\mathbf{Q}\big(Z_t \in \zeta(t,S); t < \tau^Z\big)^{\frac{1}{q}} = C_q F_t\big(\zeta(t,S)\big)^{\frac{1}{q}}.$$

The result is then complete by taking $\delta = q^{-1}$. The case involving $\bar{F}$ follows by dropping the dependence on $\{t < \tau^{1,N}\}$. □

The following result is a simple consequence of Lemma 4.2 and controls the expected mass concentrated in an interval.

COROLLARY 4.3 (Spatial concentration). *For every* $\delta \in (0,1)$, *there exists* $c_\delta > 0$ *such that*

$$\mathbf{E}\int_0^T \nu_t^N(a,b)\,dt \le \mathbf{E}\int_0^T \bar{\nu}_t^N(a,b)\,dt \le c_\delta(b-a)^\delta,$$

for all $a < b$ *and* $N \ge 1$.

PROOF. Notice that $\zeta(t,(a,b)) \subseteq [\zeta(t,a),\zeta(t,b)]$, so with $\bar{F}$ as in Lemma 4.2

$$
\begin{aligned}
\bar{F}_t\big(\zeta\big(t,(a,b)\big)\big) &\le \int_0^\infty \int_{\zeta(t,a)}^{\zeta(t,b)} \frac{1}{\sqrt{2\pi t}} \exp\biggl\{-\frac{(x-\zeta(0,x_0))^2}{2t}\biggr\}\,dx\,\nu_0(dx_0) \\
&\le (2\pi t)^{-1/2}\big(\zeta(t,b)-\zeta(t,a)\big) \\
&= (2\pi t)^{-1/2}\int_a^b \frac{dy}{\sigma(t,y)} \le (2\pi t)^{-1/2}\cdot C\cdot(b-a),
\end{aligned}
$$

and then the result is immediate from Lemma 4.2 since $t \mapsto t^{-\delta/2}$ is integrable at the origin. □

Boundary estimate. A sharper application of Lemma 4.2 gives control of the concentration of mass near the origin. Notice the stronger rate of convergence due to the absorption at the boundary.

PROPOSITION 4.4 (Boundary estimate). *There exist $\beta > 0$ and $\delta \in (0,1)$ such that as $\varepsilon \to 0$*

$$
\mathbf{E}\nu_t^N(0,\varepsilon) = t^{-\frac{\delta}{2}}O\big(\varepsilon^{1+\beta}\big) \quad \textit{and} \quad \mathbf{E}\int_0^T \nu_t^N(0,\varepsilon)\,dt = O\big(\varepsilon^{1+\beta}\big),
$$

where the O's are uniform in $t \in [0,T]$ and $N \ge 1$.

PROOF. Let F be as in Lemma 4.2. The heat kernel for a Brownian motion absorbed at the origin is

$$
G_t(x_0,x) = (2\pi t)^{-\frac{1}{2}}\biggl[\exp\biggl\{-\frac{(x-x_0)^2}{2t}\biggr\} - \exp\biggl\{-\frac{(x+x_0)^2}{2t}\biggr\}\biggr], \tag{4.3}
$$

for $x_0, x, t > 0$. By using the bounds $G_t(x_0,x) \le (2\pi t)^{-1/2}$ and

$$
G_t(x_0,x) \le \frac{2x_0x}{\sqrt{2\pi t^3}} \exp\biggl\{-\frac{(x-x_0)^2}{2t}\biggr\},
$$

which follows from the simple estimate $1 - e^{-z} \le z$, for an arbitrary function $f = f(\varepsilon)$ we have, writing $\pi_0 := \nu_0 \circ \zeta(0,\cdot)^{-1}$, that

$$
\begin{aligned}
F_t\big((0,\varepsilon)\big) &\le c_1 t^{-\frac{1}{2}} \int_0^\varepsilon \int_0^{\varepsilon+f(\varepsilon)} \pi_0(dx_0)\,dx \\
&\qquad + c_1 t^{-\frac{3}{2}} \int_0^\varepsilon \int_{\varepsilon+f(\varepsilon)}^\infty x x_0 \exp\biggl\{-\frac{(x-x_0)^2}{2t}\biggr\}\pi_0(dx_0)\,dx \\
&\le c_1 t^{-\frac{1}{2}}\varepsilon\pi_0\big(0,\varepsilon+f(\varepsilon)\big) \\
&\qquad + c_1 t^{-\frac{3}{2}} \exp\biggl\{-\frac{f(\varepsilon)^2}{2t}\biggr\}\cdot\int_0^\varepsilon x\,dx\cdot\int_0^\infty x_0\pi_0(dx_0),
\end{aligned}
$$

where $c_1 > 0$ is a numerical constant. By Assumption 2.1(i), we have a constant $c_2 > 0$ such that

$$F_t\big((0,\varepsilon)\big) \le c_1 t^{-\frac{1}{2}} \varepsilon \nu_0\big(0, c_2\big(\varepsilon + f(\varepsilon)\big)\big) + c_2 t^{-\frac{3}{2}} \varepsilon^2 \exp\{-f(\varepsilon)^2/2t\}.$$

Since the function

$$u \mapsto u^{-\alpha} \exp\{-\beta/u\} \qquad \text{for } u > 0, \alpha, \beta > 0$$

is maximised at $u = \beta/\alpha$, we have the bound

$$F_t\big((0,\varepsilon)\big) \le c_3 t^{-\frac{1}{2}} \varepsilon \big\{\nu_0\big(0, c_2\big(\varepsilon + f(\varepsilon)\big)\big) + \varepsilon f(\varepsilon)^{-2}\big\}.$$

Taking $f(\varepsilon) = \varepsilon^{1/3}$ gives

$$F_t\big((0,\varepsilon)\big) = t^{-\frac{1}{2}} O\big(\varepsilon^{1+\frac{1}{6}}\big)$$

since $\nu_0(0,x) = O(x^{1/2})$ as $x \to 0$ [recall Assumption 2.1(i)]. The result is complete by applying Lemma 4.2 and noting that $\zeta(t,(0,\varepsilon)) \subseteq [0, \zeta(t,\varepsilon)] \subseteq [0, C\varepsilon]$. □

Tail estimate. A similar analysis applies for the decay of the mass that escapes to infinity.

PROPOSITION 4.5 (Tail estimate). *For every $\alpha > 0$, as $\lambda \to +\infty$*

$$\mathbf{E}\nu_t^N(\lambda,\infty) = o\big(\exp\{-\alpha\lambda\}\big) \qquad \textit{uniformly in } N \ge 1 \textit{ and } t \in [0,T].$$

PROOF. Working with $\bar{F}$ from Lemma 4.2 and splitting the range of integration at $\lambda/2$ gives

$$\begin{aligned}\bar{F}_t\big((\lambda,\infty)\big) &= \int_0^\infty \mathbf{P}(B_t > \lambda \mid B_0 = x)\pi_0(dx) \\ &\le c_1 t^{-\frac{1}{2}} \exp\left\{-\frac{\lambda^2}{8t}\right\} + \pi_0(\lambda/2,\infty),\end{aligned}$$

where $\pi_0 = \nu_0 \circ \zeta(0,\cdot)^{-1}$. By the conditions of Assumption 2.1, $\pi_0(\lambda/2,\infty) = o(e^{-\alpha\lambda})$, so

$$\bar{F}_t\big((\lambda,\infty)\big) \le c_1 t^{-\frac{1}{2}} e^{-2\lambda^2/t} + o\big(e^{-\alpha\lambda}\big) \le c_1\big\{t^{-\frac{1}{2}} e^{-\lambda^2/t}\big\} e^{-\lambda^2/T} + o\big(e^{-\alpha\lambda}\big),$$

as $\lambda \to \infty$, for every $\alpha > 0$. The result follows since $t \mapsto t^{-\frac{1}{2}} e^{-\lambda^2/t}$ is uniformly bounded for $\lambda \ge 1$, and using Lemma 4.2 with the fact that $\zeta(t,(\lambda,\infty)) \subseteq [\zeta(t,\lambda),\infty) \subseteq [C^{-1}\lambda,\infty)$. □

Loss increment estimate. So far the probabilistic estimates we have seen are consequences of the behaviour of the first moment of the diffusion processes. The next two estimates require knowledge of the correlation between particles and so are harder to prove. Heuristically, the first result shows that over any nonzero time interval a nonzero proportion of particles hit the absorbing boundary. Later in Proposition 5.6 this result will directly imply that limiting loss functions are strictly increasing whenever there is a nonzero proportion of mass remaining in the system.

PROPOSITION 4.6 (Asymptotic loss increment). *For all $t \in [0, T)$, $h > 0$ (such that $t + h \in [0, T]$) and $r < 1$*

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

PROOF. Begin by noticing that, for any $a, b > 0$, if $L^N_t < r$ and $\nu^N_t(a,\infty) \le b$, then $\nu^N_t(0, a) > 1 - r - b$. By applying Markov's inequality and Proposition 4.5, we get the bound

$$\begin{aligned}\mathbf{P}\big(L^N_{t+h} - L^N_t < \delta, L^N_t < r\big) &\le \mathbf{P}\big(L^N_{t+h} - L^N_t < \delta, \nu^N_t(0, a) > 1 - r - b\big)\\ &\quad + \mathbf{P}\big(\nu^N_t(a,\infty) > b\big)\\ &\le \mathbf{P}\big(L^N_{t+h} - L^N_t < \delta, \nu^N_t(0, a) > 1 - r - b\big)\\ &\quad + o\big(e^{-a}\big).\end{aligned}$$

Therefore, fix $b = 1 - r - c_0$, for $c_0 = \frac{1}{2}(1 - r)$, to arrive at

$$\begin{aligned}&\mathbf{P}\big(L^N_{t+h} - L^N_t < \delta, L^N_t < r\big)\\ &\quad\le \mathbf{P}\big(L^N_{t+h} - L^N_t < \delta, \nu^N_t(0, a) > c_0\big) + o\big(e^{-a}\big).\end{aligned} \tag{4.4}$$

We now concentrate on the first term in the right-hand side above with N, t and a fixed. Let $\mathcal{I}$ denote the random set of indices

$$\mathcal{I} := \big\{1 \le i \le N : X^{i,N}_t < a \text{ and } \tau^{i,N} > t\big\}.$$

If $\nu^N_t(0, a) > c_0$, then $\#\mathcal{I} \ge Nc_0$, so by conditioning on $\mathcal{I}$ (which is $\mathcal{F}_t$-measurable)

$$\begin{aligned}&\mathbf{P}\big(L^N_{t+h} - L^N_t \le \delta, \nu^N_t(0, a) > c_0\big)\\ &\quad\le \sum_{\mathcal{I}_0:\#\mathcal{I}_0\ge Nc_0} \mathbf{P}\big(L^N_{t+h} - L^N_t < \delta | \mathcal{I} = \mathcal{I}_0\big)\mathbf{P}(\mathcal{I} = \mathcal{I}_0)\end{aligned} \tag{4.5}$$

and

$$\mathbf{P}\big(L^N_{t+h} - L^N_t < \delta | \mathcal{I} = \mathcal{I}_0\big) \le \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{u\le h} X^{i,N}_{t+u} \le 0\Big\} < N\delta | \mathcal{I} = \mathcal{I}_0\Big). \tag{4.6}$$

To estimate the right-hand side of (4.6) take ζ as in Lemma 4.1 and define $Z_t^i := \zeta(t, X_t^{i,N})$ for $1 \le i \le N$. By Assumption 2.1, there exists a constant $c_1 > 0$ such that $|D_t^i| \le c_1$ for all t. Returning to (4.6), since $\zeta(t,x) \le 0$ if and only if $x \le 0$, we have

$$\mathbf{P}(L_{t+h}^N - L_t^N < \delta | \mathcal{I} = \mathcal{I}_0) \le \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{u \le h} Z_{t+u}^i \le 0\Big\} < N\delta | \mathcal{I} = \mathcal{I}_0\Big).$$

From the bound $Z_{t+u}^i \le Z_t^i + c_1 h + Y_u^i$, for $0 \le u \le h$, where

$$Y_u^i := I_u + J_u^i := \int_t^{t+u} \rho(s, L_s^N)\, dW_s + \int_t^{t+u} \sqrt{1 - \rho(s, L_s^N)^2}\, dW_s^i,$$

we obtain

$$\begin{aligned}
&\mathbf{P}(L_{t+h}^N - L_t^N < \delta | \mathcal{I} = \mathcal{I}_0) \\
&\quad \le \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{u \le h} Y_u^i \le -Z_t^i - c_1 h\Big\} < N\delta | \mathcal{I} = \mathcal{I}_0\Big).
\end{aligned}$$

From Assumption 2.1 $|Z_t^i| = O(|X_t^{i,N}|)$, so we have $c_2 > 0$ such that

$$\begin{aligned}
&\mathbf{P}(L_{t+h}^N - L_t^N < \delta | \mathcal{I} = \mathcal{I}_0) \\
&\quad \le \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{u \le h} Y_u^i \le -c_2 a - c_2\Big\} < N\delta | \mathcal{I} = \mathcal{I}_0\Big).
\end{aligned} \tag{4.7}$$

Our next step is to remove the dependence on the process I in (4.7). To do this, we split the probability on the event $\{\sup_{u \le h} |I_u| \ge c_2 a\}$ to get

$$\begin{aligned}
&\mathbf{P}(L_{t+h}^N - L_t^N < \delta | \mathcal{I} = \mathcal{I}_0) \\
&\quad \le \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{u \le h} J_u^i \le -2c_2 a - c_2\Big\} < N\delta | \mathcal{I} = \mathcal{I}_0\Big) \\
&\qquad + \mathbf{P}\Big(\sup_{u \le h} |I_u| \ge c_2 a | \mathcal{I} = \mathcal{I}_0\Big).
\end{aligned}$$

Since I is a martingale, this final probability is $o(1)$ as $a \to \infty$, by Doob's maximal inequality.

We have reduced the problem far enough to apply a time-change in order to extract the independence between the particles. To this end, conditioned on the event $\mathcal{I} = \mathcal{I}_0$, define

$$v(s) := \inf\Big\{u > 0 : \int_t^{t+u} (1 - \rho(u_0, L_{u_0}^N)^2)\, du_0 = s\Big\}, \tag{4.8}$$

then B, where $B^i := J_{v(\cdot)}^i$, is an $\mathbb{R}^{\#\mathcal{I}_0}$-valued standard Brownian motion, therefore,

$$\begin{aligned}
&\mathbf{P}(L_{t+h}^N - L_t^N < \delta | \mathcal{I} = \mathcal{I}_0) \\
&\quad \le \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{v(u) \in [0,h]} B_u^i \le -2c_2 a - c_2\Big\} < N\delta | \mathcal{I} = \mathcal{I}_0\Big) + o(1).
\end{aligned}$$

By Assumption 2.1, $c_3 u \le |v(u)|$, hence

$$\begin{aligned}
&\mathbf{P}\big(L^N_{t+h} - L^N_t < \delta | \mathcal{I} = \mathcal{I}_0\big) \\
&\quad \le \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{u\in[0,h/c_3]} B^i_u \le -2c_2 a - c_2\Big\} < N\delta | \mathcal{I} = \mathcal{I}_0\Big) + o(1) \\
&\quad \le \mathbf{P}\big(\#\{i \in \mathcal{I}_0 : B^i_{h/c_3} \le -2c_2 a - c_2\} < N\delta | \mathcal{I} = \mathcal{I}_0\big) + o(1) \\
&\quad \le \mathbf{P}\bigg(\frac{1}{N}\sum_{i\in\mathcal{I}_0} \mathbf{1}_{\xi^i \le -c_4(a+1)} < \delta\bigg) + o(1),
\end{aligned}$$

where $\{\xi^i\}_{1\le i\le N}$ is a collection of i.i.d. standard normal random variables and $c_3, c_4 > 0$ are further numerical constants. By symmetry, this final probability depends only on $\#\mathcal{I}_0$, hence

$$\mathbf{P}\big(L^N_{t+h} - L^N_t < \delta | \mathcal{I} = \mathcal{I}_0\big) \le \mathbf{P}\Bigg(\frac{1}{N}\sum_{i=1}^{\#\mathcal{I}_0} \mathbf{1}_{\xi^i \le -c_4(a+1)} < \delta\Bigg) + o(1).$$

Returning to (4.5), we now have

$$\begin{aligned}
&\mathbf{P}\big(L^N_{t+h} - L^N_t < \delta, \nu^N_t(0,a) > c_0\big) \\
&\quad \le \sum_{S_0:\#\mathcal{I}_0 \ge Nc_0} \mathbf{P}\Bigg(\frac{1}{N}\sum_{i=1}^{\#\mathcal{I}_0} \mathbf{1}_{\xi^i \le -c_4(a+1)} < \delta\Bigg)\mathbf{P}(\mathcal{I} = \mathcal{I}_0) + o(1) \\
&\quad \le \mathbf{P}\Bigg(\frac{1}{N}\sum_{i=1}^{Nc_0} \mathbf{1}_{\xi^i \le -c_4(a+1)} < \delta\Bigg) + o(1),
\end{aligned}$$

so the law of large numbers gives

$$\limsup_{n\to\infty} \mathbf{P}\big(L^N_{t+h} - L^N_t < \delta, L^N_t < r\big) \le \mathbf{1}_{c_0 p(a) \le \delta} + o(1), \tag{4.9}$$

where $p(a) := \mathbf{P}(\xi^1 \le -c_4(a+1))$ and where we have substituted back into (4.4). This inequality holds for all a and δ, with the $o(1)$ term denoting convergence as $a \to \infty$. We now choose the free parameter a to be a function of δ, specifically

$$a(\delta) := \big(2\log\log(1/\delta)\big)^{\frac{1}{2}}.$$

This guarantees that $a(\delta) \to \infty$ as $\delta \to 0$, but also

$$\begin{aligned}
\delta^{-1} p\big(a(\delta)\big) &\ge \frac{1}{2}\delta^{-1} a(\delta)^{-1} e^{-a(\delta)^2/2} \\
&= \frac{1}{\sqrt{2}}\delta^{-1}\big(\log(1/\delta)\big)^{-1}\big(\log\log(1/\delta)\big)^{1/2} \to \infty
\end{aligned}$$

as $\delta \to 0$, where we have used the well-known Gaussian estimate $\Phi(-x) \ge (x^{-1} - x^{-3})\phi(x) \ge \frac{1}{2}x^{-1}\phi(x)$, for Φ and ϕ the c.d.f. and p.d.f. of the standard normal distribution. Using this choice of $a(\delta)$ in (4.9) completes the result. □

The following is a partial converse of the previous result in that it shows that the system cannot lose a large amount of mass in a short period of time. It will be used in Proposition 5.1 to verify a sufficient condition for the tightness of $(\nu^N, W)_{N\geq 1}$.

PROPOSITION 4.7. *For every $t \in [0, T]$ and $\eta > 0$,*

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

PROOF. With $\varepsilon > 0$ fixed, we have

$$\begin{aligned}
&\mathbf{P}\big(L^N_{t+\delta} - L^N_t \geq \eta\big) \\
&\quad\leq \mathbf{P}\big(\nu^N_t(0,\varepsilon) \geq \eta/2\big) + \mathbf{P}\big(L^N_{t+\delta} - L^N_t \geq \eta, \nu^N_t(0,\varepsilon) < \eta/2\big) \\
&\quad\leq 2\eta^{-1}\mathbf{P}\big(X^{1,N}_t \in (0,\varepsilon)\big) + \mathbf{P}\big(L^N_{t+\delta} - L^N_t \geq \eta, \nu^N_t(0,\varepsilon) < \eta/2\big) \\
&\quad\leq \mathbf{P}\big(L^N_{t+\delta} - L^N_t \geq \eta, \nu^N_t(0,\varepsilon) < \eta/2\big) + o(1) \qquad \text{as } \varepsilon \to 0,
\end{aligned} \tag{4.10}$$

where the second line uses Markov's inequality and (4.1) and the third line uses Proposition 4.4 for $t > 0$ and Assumption 2.1 (i) for $t = 0$. Define $\mathcal{I}$ to be the random set of indices

$$\mathcal{I} := \big\{1 \leq i \leq N : X^{i,N}_t \geq \varepsilon\big\},$$

then conditioning on $\mathcal{I}$ gives

$$\begin{aligned}
&\mathbf{P}\big(L^N_{t+\delta} - L^N_t \geq \eta, \nu^N_t(0,\varepsilon) < \eta/2\big) \\
&\quad\leq \sum_{\mathcal{I}_0 : \#\mathcal{I}_0 \geq N(1-\eta/2)} \mathbf{P}\big(L^N_{t+\delta} - L^N_t \geq \eta | \mathcal{I} = \mathcal{I}_0\big)\mathbf{P}(\mathcal{I} = \mathcal{I}_0).
\end{aligned} \tag{4.11}$$

The conditional expectation in the summand can be bounded by

$$\begin{aligned}
&\mathbf{P}\big(L^N_{t+\delta} - L^N_t \geq \eta | \mathcal{I} = \mathcal{I}_0\big) \\
&\quad\leq \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{s\in[t,t+\delta]} X^{i,N}_s \leq 0\Big\} \geq \frac{N\eta}{2}\Big| \mathcal{I} = \mathcal{I}_0\Big) \\
&\quad\leq \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{s\in[t,t+\delta]} \big(X^{i,N}_s - X^{i,N}_t\big) \leq -\varepsilon\Big\} \geq \frac{N\eta}{2}\Big| \mathcal{I} = \mathcal{I}_0\Big).
\end{aligned}$$

With t fixed, define the process $U^i_s := \zeta(t+s, X^{i,N}_{t+s} - X^{i,N}_t)$, then

$$\mathbf{P}\big(L^N_{t+\delta} - L^N_t \geq \eta | \mathcal{I} = \mathcal{I}_0\big) \leq \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{s\in[0,\delta]} U^i_s \leq -c_5\varepsilon\Big\} \geq \frac{N\eta}{2}\Big| \mathcal{I} = \mathcal{I}_0\Big)$$

for $c_5 > 0$ a numerical constant. As for Z in Lemma 4.1, we have

$$\begin{aligned}
dU^i_s &= E^i_s\,ds + \rho\big(t+s, L^N_{t+s}\big)\,dW_{t+s} + \big(1 - \rho\big(t+s, L^N_{t+s}\big)^2\big)^{1/2}\,dW^i_{t+s} \\
&=: E^i_s\,ds + dI_s + dJ^i_s,
\end{aligned}$$

where E_s^i is uniformly bounded by Assumption 2.1, therefore, we can find $c_6 > 0$ such that

$$\begin{aligned}
&\mathbf{P}(L_{t+\delta}^N - L_t^N \geq \eta | \mathcal{I} = \mathcal{I}_0) \\
&\quad \leq \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{s\in[0,\delta]} J_s^i \leq -c_6(\varepsilon - \delta - a)\Big\} \geq \frac{N\eta}{2} \Big| \mathcal{I} = \mathcal{I}_0\Big) \\
&\qquad + \mathbf{P}\Big(\sup_{s\in[0,\delta]} |I_s| \geq a | \mathcal{I} = \mathcal{I}_0\Big).
\end{aligned}$$

By applying the time-change argument from (4.8) and using Markov's and Doob's maximal inequality, we have

$$\begin{aligned}
&\mathbf{P}(L_{t+\delta}^N - L_t^N \geq \eta | \mathcal{I} = \mathcal{I}_0) \\
&\quad \leq \mathbf{P}\Big(\#\Big\{i \in \mathcal{I}_0 : \inf_{s\in[0,\delta]} B_s^i \leq -c_7(\varepsilon - \delta - a)\Big\} \geq \frac{N\eta}{2}\Big) + O(\delta a^{-2}),
\end{aligned}$$

where B^i are independent standard Brownian motions, $a > 0$ and $c_7 > 0$ is a numerical constant.

Returning to (4.11) and noticing the the right-hand side above is maximised when $\mathcal{I}_0 = \{1, 2, \dots, N\}$

$$\begin{aligned}
&\mathbf{P}(L_{t+\delta}^N - L_t^N \geq \eta, \nu_t^N(0, \varepsilon) < \eta/2) \\
&\quad \leq \mathbf{P}\Big(\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\inf_{s\in[0,\delta]} B_s^i \leq -c_7(\varepsilon-\delta-a)\}} \geq \eta/2\Big) + O(\delta a^{-2}).
\end{aligned}$$

The law of large numbers and the distribution of the minimum of Brownian motion gives

$$\begin{aligned}
&\limsup_{N\to\infty} \mathbf{P}(L_{t+\delta}^N - L_t^N \geq \eta, \nu_t^N(0, \varepsilon) < \eta/2) \\
&\quad \leq \mathbf{1}_{\Phi(-c_7\delta^{-1/2}(\varepsilon-\delta-a))\geq\eta/2} + O(\delta a^{-2}),
\end{aligned} \tag{4.12}$$

provided $\varepsilon - \delta - a > 0$, where Φ is the normal c.d.f. We now make the choice

$$\varepsilon(\delta) = \delta^{1/2}\log(1/\delta) \quad \text{and} \quad a(\delta) = \delta^{1/2}\log\log(1/\delta),$$

which guarantees

$$\varepsilon(\delta) \to 0, \qquad \delta^{-1/2}(\varepsilon(\delta) - \delta - a(\delta)) \to \infty \quad \text{and} \quad \delta a(\delta)^{-2} \to 0,$$

as $\delta \to 0$. Hence, the result follows from (4.10), (4.11) and (4.12). $\square$

5. Tightness of the system and existence of solutions: Proof of Theorem 1.1. We will now use the results from Section 4 to prove Theorem 1.1, which follows directly from the combination of Propositions 5.5, 5.6 and 5.11. We first establish tightness of the sequence of the laws of $(\nu^N, W)_{N\geq 1}$ (Proposition 5.1) using the framework of [40]. The reader is referred to that article for the technical definitions of the topological spaces used in this section. Once we have tightness, we can then extract limit points of the sequence $(\nu^N, W)_{N\geq 1}$, and Propositions 5.3, 5.5 and 5.6 are devoted to recovering the properties of the limiting laws from the probabilistic properties of the finite system. Finally, the limit points are shown to satisfy the evolution equation in Theorem 1.1 via a martingale argument (Proposition 5.11) and care needs to be taken over the discontinuities in the coefficients of the limit SPDE (Corollary 5.7).

PROPOSITION 5.1 (Tightness). *The sequence $(\nu^N)_{N\geq 1}$ is tight on the space $(D_{\mathscr{S}'}, \mathrm{M1})$, hence $(\nu^N, W)_{N\geq 1}$ is tight on the space $(D_{\mathscr{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \mathrm{U})$, where $(C_{\mathbb{R}}, \mathrm{U})$ is the space of real-valued continuous paths with the topology of uniform convergence.*

REMARK 5.2. We note that a version of this result is given in [40], Theorem 4.3, for the case $\mu = 0$, $\sigma = 1$.

PROOF. The second statement follows from the first and the fact that joint tightness is implied by marginal tightness.

By [40], Theorem 3.2, it suffices to show that $(\nu^N(\phi))_{N\geq 1}$ is tight on $(D_{\mathbb{R}}, \mathrm{M1})$ for every $\phi \in \mathscr{S}$. To prove this we verify the conditions of [53], Theorem 12.12.2, the first of which is trivial because ν^N is a sub-probability measure so $|\nu_t^N(\phi)| \leq \|\phi\|_\infty$. Hence, we concentrate on condition (ii), which is implied by [40], Proposition 4.1, therefore, we are done if we can find $a, b, c > 0$ such that

$$\mathbf{P}\big(H_{\mathbb{R}}\big(\nu_{t_1}^N(\phi), \nu_{t_2}^N(\phi), \nu_{t_3}^N(\phi)\big) \geq \eta\big) \leq c\eta^{-a}|t_3 - t_1|^{1+b}, \tag{5.1}$$

for all $N \geq 1$, $\eta > 0$ and $0 \leq t_1 < t_2 < t_3 \leq T$, where

$$H_{\mathbb{R}}(x_1, x_2, x_3) := \inf_{\lambda\in(0,1)} \big|x_2 - (1-\lambda)x_1 - \lambda x_3\big| \qquad \text{for } x_1, x_2, x_3 \in \mathbb{R},$$

and if

$$\lim_{N\to\infty} \mathbf{P}\Big(\sup_{t\in(0,\delta)} \big|\nu_t^N(\phi) - \nu_0^N(\phi)\big| + \sup_{t\in(T-\delta,T)} \big|\nu_T^N(\phi) - \nu_t^N(\phi)\big| \geq \eta\Big) = 0, \tag{5.2}$$

for every $\eta > 0$.

With $\bar{\nu}^N$ as defined in (3.2), the decomposition in [40], Proposition 4.2, and Markov's inequality give

$$\begin{aligned}
&\mathbf{P}\big(H_{\mathbb{R}}\big(\nu_{t_1}^N(\phi), \nu_{t_2}^N(\phi), \nu_{t_3}^N(\phi)\big) \geq \eta\big) \\
&\quad \leq \eta^{-4}\mathbf{E}\big[\big(\big|\bar{\nu}_{t_1}^N(\phi) - \bar{\nu}_{t_2}^N(\phi)\big| + \big|\bar{\nu}_{t_2}^N(\phi) - \bar{\nu}_{t_3}^N(\phi)\big|\big)^4\big] \\
&\quad \leq 8\eta^{-4}\big(\mathbf{E}\big|\bar{\nu}_{t_1}^N(\phi) - \bar{\nu}_{t_2}^N(\phi)\big|^4 + \mathbf{E}\big|\bar{\nu}_{t_2}^N(\phi) - \bar{\nu}_{t_3}^N(\phi)\big|^4\big).
\end{aligned}$$

For any $t, s \in [0, T]$, from Hölder's inequality we obtain

$$\mathbf{E}\big|\bar{\nu}_t^N(\phi) - \bar{\nu}_s^N(\phi)\big|^4 \le \frac{1}{N}\sum_{i=1}^N \mathbf{E}\big|\phi\big(X^{i,N}_{t\wedge\tau^{i,N}}\big) - \phi\big(X^{i,N}_{s\wedge\tau^{i,N}}\big)\big|^4$$

$$\le \|\phi\|_{\mathrm{lip}}^4 \mathbf{E}\big|X^{i,N}_{t\wedge\tau^{i,N}} - X^{i,N}_{s\wedge\tau^{i,N}}\big|^4,$$

where $\|\phi\|_{\mathrm{lip}}$ is the Lipschitz constant of ϕ. By Assumption 2.1 and the Burkholder–Davis–Gundy inequality [46], Theorem IV.42.1, the final expectation above is $O(|t-s|^2)$ uniformly in N. Therefore, we have (5.1) with $a = 4$ and $b = 1$.

Now consider the first supremum in (5.2). By again using the decomposition from [40], Prop. 4.2, that is $\nu_t^N(\phi) = \bar{\nu}_t^N(\phi) - \phi(0)L_t^N$, we have

$$\mathbf{P}\Big(\sup_{t\in(0,\delta)} \big|\nu_t^N(\phi) - \nu_0^N(\phi)\big| \ge \eta\Big)$$

$$\le \mathbf{P}\Big(\sup_{t\in(0,\delta)} \big|\bar{\nu}_t^N(\phi) - \bar{\nu}_0^N(\phi)\big| \ge \eta/2\Big) + \mathbf{P}\big(|\phi(0)|L_\delta^N \ge \eta/2\big).$$

The first term on the right-hand side vanishes as $\delta \to 0$ by the same work as for (5.1) and the second term vanishes by Proposition 4.7. Therefore,

$$\mathbf{P}\Big(\sup_{t\in(0,\delta)} \big|\nu_t^N(\phi) - \nu_0^N(\phi)\big| \ge \eta\Big) \to 0, \qquad \text{as } \delta \to 0,$$

and likewise for $\mathbf{P}(\sup_{t\in(T-\delta,T)} |\nu_T^N(\phi) - \nu_t^N(\phi)| \ge \eta)$, so we have (5.2), which completes the proof. $\square$

Limit points. Tightness of $(\nu^N, W)_{N\ge1}$ ensures that the sequence is relatively compact [40], Theorem 3.2, hence every subsequence of $(\nu^N, W)_{N\ge1}$ has a further subsequence which converges in law. To avoid possible confusion about multiple distinct limit points, we will denote by (ν^*, W) any pair of processes that realises one of these limiting laws. Using $\Rightarrow$ to denote convergence in law, we have

$$\big(\nu^{N_k}, W\big) \Rightarrow \big(\nu^*, W\big), \qquad \text{on } (D_{\mathscr{S}'}, \mathrm{M1}) \times (C_{\mathbb{R}}, \mathrm{U}),$$

as $k \to \infty$, for some subsequence $(N_k)_{k\ge1}$. Establishing full weak convergence is equivalent to showing that there is exactly one limiting law.

So far, we have that any limiting empirical process, ν^*, is an element of $D_{\mathscr{S}'}$. The following result recovers ν^* as a probability-measure-valued process.

PROPOSITION 5.3. *Let (ν^*, W) realise a limiting law. Then ν_t^* is a sub-probability measure supported on* $[0, \infty)$ *for every* $t \in [0, T]$, *with probability* 1.

REMARK 5.4. Technically, what we will show is that, for every t, ν_t^* agrees with a sub-probability measure on $\mathscr{S}$ and from now on we associate ν_t^* with this measure.

PROOF OF PROPOSITION 5.3. Take $(\nu^{N_k}, W) \Rightarrow (\nu^*, W)$. Fix $\phi \in \mathscr{S}$, then by [40], Proposition 2.7(i), $\nu^{N_k}(\phi) \Rightarrow \nu^*(\phi)$ on $(D_{\mathbb{R}}, \mathrm{M1})$. Lemma 13.4.1 of [53] gives

$$\sup_{t\in[0,T]} \left|\nu_t^{N_k}(\phi)\right| \Rightarrow \sup_{t\in[0,T]} \left|\nu_t^*(\phi)\right|, \qquad \text{on } \mathbb{R},$$

therefore, the portmanteau theorem [4], Theorem 2.1, gives

$$\mathbf{P}\Big(\sup_{t\in[0,T]} \left|\nu_t^*(\phi)\right| > \|\phi\|_\infty\Big) \le \liminf_{k\to\infty} \mathbf{P}\Big(\sup_{t\in[0,T]} \left|\nu_t^{N_k}(\phi)\right| > \|\phi\|_\infty\Big) = 0,$$

with the final equality due to ν_t^N being a sub-probability measure. (The supremum over t ensures that the following argument holds for all t simultaneously.) By a similar analysis, we have that $\nu_t^*(\phi)$ is nonnegative when ϕ is nonnegative and $\nu_t^*(\phi) = 0$ when ϕ is supported on $(-\infty, 0)$. Hence, ν_t^* is a positive linear functional on $\mathscr{S}$, so extends to a positive linear functional, ξ_t, on the space, C_0, of continuous and compactly support function on $\mathbb{R}$ with the uniform topology. The Riesz representation theorem [47], Theorem 2.14, then implies that, for every t, there exists a regular Borel measure, ζ_t, such that

$$\xi_t(\phi) = \int_{\mathbb{R}} \phi(x)\zeta_t(dx) \qquad \text{for every } \phi \in C_0.$$

Associating ζ and ν^* gives the result. □

Now that it is safe to regard a limit point, $\nu^{N_k} \Rightarrow \nu^*$, as taking values in the sub-probability measures, it makes sense to introduce the *limit loss process* as $L_t^* := 1 - \nu_t^*(0, \infty)$. Of course, we would like to know that $L^{N_k} \Rightarrow L^*$ on $(D_{\mathbb{R}}, \mathrm{M1})$, however, the function $x \mapsto 1$ is not an element of $\mathscr{S}$, so [40], Proposition 2.7, does not allow us to deduce this fact from the continuous mapping theorem. To remedy this, we must work slightly harder.

PROPOSITION 5.5 (Convergence of the loss process). *Suppose that* $(\nu^{N_k}, W)_{k\ge1}$ *converges weakly to* (ν^*, W) *and that* $L_t^* := 1 - \nu_t^*(0, \infty)$. *Then* $(L^{N_k}, W)_{k\ge1}$ *converges weakly to* (L^*, W) *on* $(D_{\mathbb{R}}, \mathrm{M1}) \times (C_{\mathbb{R}}, \mathrm{U})$.

PROOF. For a contradiction, suppose that the weak convergence does not hold. Since $t \mapsto L_t^N$ is increasing, $L_t^N \in [0, 1]$ and we have Proposition 4.7, the conditions of [53] are satisfied and so $(L^N, \text{Theorem 12.12.2})_{N\ge1}$ is tight on $(D_{\mathbb{R}}, \mathrm{M1})$, and because marginal tightness implies joint tightness, $(L^N, W)_{N\ge1}$ is also tight. By taking a further subsequence if needed, assume that $(L^{N_k}, W)_{k\ge1} \Rightarrow (L^\dagger, W)_{k\ge1}$ for some $L^\dagger \in D_{\mathbb{R}}$.

Notice from [53], Theorem 12.4.1, that the canonical time projection from $(D_{\mathbb{R}}, \mathrm{M1})$ to $\mathbb{R}$ is only continuous at times for which its argument does not jump. That is, for every t, $\pi_t(x) := x_t$ is continuous at $x \in D_{\mathbb{R}}$ if and only if $x_{t-} = x_t$.

To this end, define $\mathrm{cont}(L^\dagger) = \{s \in [0,T] : \mathbf{P}(L^\dagger_{s-} = L^\dagger_s) = 1\}$, which we know by [4], Section 13, is cocountable in $[0,T]$. For $\lambda \in \mathbb{N}$ define $\phi_\lambda \in \mathscr{S}$ to be any function satisfying $\phi_\lambda = 1$ on $[-\lambda, \lambda]$, $\phi_\lambda = 0$ on $(-\infty, -2\lambda) \cup (2\lambda, \infty)$ and $\phi_\lambda \in (0,1)$ otherwise. By [40], Proposition 2.7(i), $\nu^{N_k}(\phi_\lambda) \Rightarrow \nu^*(\phi_\lambda)$, and define $\mathrm{cont}(\nu^*(\phi_\lambda)) = \{s \in [0,T] : \mathbf{P}(\nu^*_{s-}(\phi_\lambda) = \nu^*_s(\phi_\lambda)) = 1\}$. Take

$$\mathbb{T} := \mathrm{cont}(L^\dagger) \cap \bigcap_{\lambda=1}^{\infty} \mathrm{cont}(\nu^*(\phi_\lambda)),$$

which is cocountable (since it is the countable intersection of cocountable sets) and so is dense in $[0,T]$.

Since $(L^{N_k}, W) \Rightarrow (L^\dagger, W)$ and $\mathbb{T}$ is dense in $[0,T]$, if (L^*, W) and $(L^\dagger, W)$ are not equal in law on $(D_\mathbb{R}, \mathrm{M1})$, then it must be the case that not all of the finite-dimensional marginals of L^* and $L^\dagger$ on $\mathbb{T}$ are equal in law. It is no loss of generality to assume that there exists $\varepsilon > 0$, $m \in \mathbb{N}$, $f_i, g_i : \mathbb{R} \to \mathbb{R}$ bounded and Lipschitz and $t_1, \dots, t_m \in \mathbb{T}$ such that

$$\mathbf{E} \prod_{i=1}^{m} f_i(L^*_{t_i}) g_i(W_{t_i}) + \varepsilon \le \limsup_{k\to\infty} \mathbf{E} \prod_{i=1}^{m} f_i(L^{N_k}_{t_i}) g_i(W_{t_i}).$$

By Proposition 4.5,

$$\mathbf{E}\big|L^{N_k}_t - (1 - \nu^{N_k}_t(\phi_\lambda))\big| = O(e^{-\lambda}) \qquad \text{uniformly in } t \text{ and } N_k,$$

as $\lambda \to \infty$, therefore, the Lipschitz property of f_i gives

$$\mathbf{E} \prod_{i=1}^{m} f_i(L^*_{t_i}) g_i(W_{t_i}) + \varepsilon \le \limsup_{k\to\infty} \mathbf{E} \prod_{i=1}^{m} f_i(1 - \nu^{N_k}_{t_i}(\phi_\lambda)) g_i(W_{t_i}) + O(e^{-\lambda}),$$

but $t_i \in \mathrm{cont}(\nu^*(\phi_\lambda))$, so

$$\mathbf{E} \prod_{i=1}^{m} f_i(L^*_{t_i}) g_i(W_{t_i}) + \varepsilon \le \mathbf{E} \prod_{i=1}^{m} f_i(1 - \nu^*_{t_i}(\phi_\lambda)) g_i(W_{t_i}) + O(e^{-\lambda}).$$

Since ν^*_t is a probability measure $\nu^*_t(\phi_\lambda) \to \nu^*_t(\mathbb{R}) = 1 - L^*_t$ (recall from Proposition 5.3 that ν^*_t is supported on $[0,\infty)$), so taking $\lambda \to \infty$ gives the required contradiction. $\square$

We are now in a position to verify the first half of Theorem 1.1, which is that any limit point must satisfy the regularity conditions from Assumption 2.3.

PROPOSITION 5.6 (Regularity conditions). *If (ν^*, W) realises a limiting law of $(\nu^N, W)_{N\ge1}$, then ν^* satisfies Assumption 2.3.*

PROOF. First, ν^* takes values in the sub-probability measures by Proposition 5.3, and that result also gives Assumption 2.3(ii).

For conditions (iv) and (v) of Assumption 2.3, let $I = (x, y) \subseteq \mathbb{R}$ be any finite open interval. For $\delta > 0$, take any $\phi_\delta \in \mathscr{S}$ satisfying $\phi_\delta = 1$ on I, $\phi_\delta = 0$ on $(-\infty, x-\delta) \cup (y+\delta, \infty)$ and $\phi_\delta \in (0, 1)$ otherwise. Taking $(\nu^{N_k}, W) \Rightarrow (\nu^*, W)$ and noting that $\int_0^t \nu_s^{N_k}(\phi_\lambda)\,ds \Rightarrow \int_0^t \nu_s^*(\phi_\lambda)\,ds$ in $\mathbb{R}$ by [53], Theorem 11.5.1, and that these integrals are uniformly bounded (by $T\|\phi_\lambda\|_\infty = T$), we have

$$\mathbf{E}\int_0^T \nu_t^*(I)\,dt \le \mathbf{E}\int_0^T \nu_t^*(\phi_\delta)\,dt = \lim_{k\to\infty} \mathbf{E}\int_0^T \nu_t^{N_k}(\phi_\delta)\,dt.$$

For both conditions (iv) and (v), we have bounds on the right-hand side which are independent of N_k (Propositions 4.3 and 4.4), and then the conditions hold by sending $\delta \to 0$. For condition (iii), we have $y = \infty$, so $\phi_\delta \notin \mathscr{S}$. However, for $I = (\lambda, \eta)$ with $\eta > 0$, the above work gives

$$\begin{aligned}\mathbf{E}\int_0^T \nu_t^*(\lambda, \eta)\,dt &\le \lim_{k\to\infty} \mathbf{E}\int_0^T \nu_t^*(\phi_\lambda)\,dt \\ &\le \liminf_{k\to\infty} \mathbf{E}\int_0^T \nu_t^{N_k}(\lambda-\delta, \eta+\delta)\,dt \\ &= o\big(e^{-\alpha(\lambda-\delta)}\big),\end{aligned}$$

so sending $\delta \to 0$ and $\eta \to \infty$ (using the dominated convergence theorem) gives the result.

It remains to show (i) of Assumption 2.3. First, we prove that L^* is nondecreasing. By [4], Section 13, there is a (deterministic) cocountable set, $\mathbb{T}$, on which $(L_t^{N_k}, L_s^{N_k}) \Rightarrow (L_t^*, L_s^*)$ in $\mathbb{R}\times\mathbb{R}$. So for $s < t$ in $\mathbb{T}$ [4], Theorem 2.1, implies

$$\mathbf{P}\big(L_t^* - L_s^* < 0\big) \le \liminf_{k\to\infty} \mathbf{P}\big(L_t^{N_k} - L_s^{N_k} < 0\big) = 0,$$

and hence L^* is nondecreasing on $\mathbb{T}$. But $\mathbb{T}$ is dense in $[0, T]$ and L^* càdlàg, so we conclude L^* is nondecreasing on $[0, T]$. To deduce the strict monotonicity, Proposition 4.6 implies

$$\begin{aligned}\mathbf{P}\big(L_t^* - L_s^* = 0, L_s^* < r\big) &= \lim_\delta \mathbf{P}\big(L_t^* - L_s^* < \delta, L_s^* < r\big) \\ &\le \limsup_{\delta\to 0}\limsup_{k\to\infty} \mathbf{P}\big(L_t^{N_k} - L_s^{N_k} < \delta, L_s^{N_k} < r\big) = 0,\end{aligned}$$

whenever $r < 1$ and sending $r \uparrow 1$ gives the required result. □

So far, we have seen no reason why it is important L^* should be strictly increasing whenever the mass in the system is not completely depleted ($L^* < 1$). The following result is such an example and shows why this condition is needed to pass to a weak limit. The result will be applied directly in the next subsection.

COROLLARY 5.7 (Weak convergence of integrals). *Fix $t \in [0, T]$ and $\phi \in \mathscr{S}$. Let $g = g(t, x, \ell)$ be equal to either $\mu(t, x, \ell)$, $\sigma(t, x)^2$ or $\sigma(t, x, \ell)\rho(t, \ell)$. Define A to be all elements in $D_{\mathscr{S}'}$ that take values in the sub-probability measures and let $B = D_{[0,1]} \subseteq D_{\mathbb{R}}$. Then the map*

$$(\xi, \ell) \in A \times B \mapsto \int_0^t \xi_s\big(g(s, \cdot, \ell_s)\phi(\cdot)\big)\, ds \in \mathbb{R}$$

is continuous [*with respect to the product topology on* $(D_{\mathscr{S}'}, \mathrm{M1}) \times (D_{[0,1]}, \mathrm{M1})$] *at all point (ξ, ℓ) which satisfy the conditions of Assumption 2.3. Consequently, if $(\nu^{N_k}, W) \Rightarrow (\nu^*, W)$ then*

$$\int_0^t \nu^{N_k}\big(g\big(s, \cdot, L_s^{N_k}\big)\phi(\cdot)\big)\, ds \Rightarrow \int_0^t \nu^*\big(g\big(s, \cdot, L_s^*\big)\phi(\cdot)\big)\, ds \qquad \text{on } \mathbb{R}.$$

PROOF. For shorthand we will denote this map $\Psi : A \times B \to \mathbb{R}$. Suppose that $(\bar{\xi}, \bar{\ell}) \to (\xi, \ell)$ in $A \times B$, then

$$\begin{aligned} \big|\Psi(\bar{\xi}, \bar{\ell}) - \Psi(\xi, \ell)\big| &\le \left|\int_0^t \bar{\xi}_s\big(g(s, \cdot, \ell_s)\phi\big)\, ds - \int_0^t \xi_s\big(g(s, \cdot, \ell_s)\phi\big)\, ds\right| \\ &\quad + \int_0^t \big|\bar{\xi}_s\big(g(s, \cdot, \ell_s)\phi - g(s, \cdot, \bar{\ell}_s)\phi\big)\big|\, ds =: I + J. \end{aligned} \tag{5.3}$$

We will control I and J separately.

Begin by fixing $\varepsilon > 0$ and $\delta > 0$. Take $k = k(\delta) > 0$ sufficiently large so that $|g(s, x, \ell)\phi(x)| < \delta$ for all $s \in [0, T]$, $x \in \mathbb{R} \setminus [-k, k]$ and $\ell \in [0, 1]$, which is possible because g is bounded and ϕ is rapidly decreasing. Let ψ_ε be a mollifier and set $g^\varepsilon(s, x, \ell) := (g(s, \cdot, \ell) * \psi_\varepsilon)(x) \in C^\infty(\mathbb{R})$, then we have

$$\begin{aligned} I &\le \left|\int_0^t \bar{\xi}_s\big(g^\varepsilon(s, \cdot, \ell_s)\phi\big)\, ds - \int_0^t \xi_s\big(g^\varepsilon(s, \cdot, \ell_s)\phi\big)\, ds\right| \\ &\quad + 2\int_0^t \sup_{x \in \mathbb{R}}\big|\phi(x)\big|\big|g^\varepsilon(s, x, \ell_s) - g(s, x, \ell_s)\big|\, ds. \end{aligned}$$

Since $g^\varepsilon(s, \cdot, \ell) \in C^\infty(\mathbb{R})$ and $\phi \in \mathscr{S}$, $g^\varepsilon(s, \cdot, \ell)\phi(\cdot) \in \mathscr{S}$, so the first term vanishes as $\bar{\xi} \to \xi$. We can then split the second term as

$$\begin{aligned} \limsup_{\bar{\xi} \to \xi} I &\le 2\|\phi\|_\infty \int_0^t \sup_{x \in [-2k, 2k]} \big|g^\varepsilon(s, x, \ell_s) - g(s, x, \ell_s)\big|\, ds \\ &\quad + 2c\int_0^t \sup_{x \in \mathbb{R} \setminus [-2k, 2k]} \big|\phi(x)\big|\, ds, \end{aligned}$$

and here the first term vanishes as $\varepsilon \to 0$ by [24], App. C, Theorem 6, since $[-k, k]$ is compact and the second term can be guaranteed to be less than 2δ for k sufficiently large. Taking $\delta \to 0$ gives $\limsup I = 0$.

To deal with J in (5.3), first notice that since $\bar{\xi} \in A$

$$J \le \|\phi\|_\infty \int_0^t \sup_{x\in\mathbb{R}} \bigl|g(s,x,\ell_s) - g(s,x,\bar{\ell}_s)\bigr|\,ds.$$

Define $\mathbb{T}_0 := \{s \in [0,t] : \ell_s = \theta_i \text{ for some } i \in \{0,1,\dots,k\}\}$, where we recall Assumption 2.1 condition (v). For $\delta > 0$, let $\mathbb{T}_0^\delta := \{s \in [0,t] : \min_{0\le i\le k} |\theta_i - \ell_s| < \delta\}$. Define $\mathbb{T}_1$ to be all $s \in [0,t]$ such that $\ell_s = \ell_{s-}$, which we know is a cocountable set [53], Corollary 12.2.1. For $s \in \mathbb{T}_1$, $\bar{\ell}_s \to \ell_s$ in $\mathbb{R}$, so if $s \in \mathbb{T}_1 \setminus \mathbb{T}_0^\delta$ then eventually $\ell_s, \bar{\ell}_s \in [\theta_{i-1}, \theta_i)$ for some $i \in \{1,2,\dots,k\}$, whence $\sup_{x\in\mathbb{R}} |g(s,x,\ell_s) - g(s,x,\bar{\ell}_s)| \to 0$ by Assumption 2.1 condition (iv). We conclude

$$\limsup_{\bar{\xi}\to\xi} J \le c_1 \int_{([0,T]\setminus\mathbb{T}_1)\cup\mathbb{T}_0^\delta} ds \le c_1 k\delta \qquad \text{for every } \delta > 0,$$

where $c_1 > 0$ is a numerical constant due to Assumption 2.1. This completes the result. $\square$

Martingale approach. We complete this section and the proof of Theorem 1.1 by showing that the limit SPDE holds for a general limit point. For this, we will use a martingale argument and we introduce three processes.

DEFINITION 5.8 (Martingale components). For a fixed test function $\phi \in C^{\text{test}}$, define the maps:

(i) $M^\phi : D_{\mathscr{S}'} \times D_{[0,1]} \to D_{\mathbb{R}}$,

$$M^\phi(\xi,\ell)(t) := \xi_t(\phi) - \nu_0(\phi) - \int_0^t \xi_s\bigl(\mu(s,\cdot,\ell_s)\partial_x\phi\bigr)\,ds - \frac{1}{2}\int_0^t \xi_s\bigl(\sigma^2(s,\cdot)\partial_{xx}\phi\bigr)\,ds.$$

(ii) $S^\phi : D_{\mathscr{S}'} \times D_{[0,1]} \to D_{\mathbb{R}}$,

$$S^\phi(\xi,\ell)(t) := M^\phi(\xi,\ell)(t)^2 - \int_0^t \xi_s\bigl(\sigma(s,\cdot)\rho(s,\ell_s)\partial_x\phi\bigr)^2\,ds.$$

(iii) $C^\phi : D_{\mathscr{S}'} \times D_{\mathbb{R}} \times C_{\mathbb{R}} \to D_{\mathbb{R}}$,

$$C^\phi(\xi,\ell,w)(t) := M^\phi(\xi,\ell)(t)\cdot w(t) - \int_0^t \xi_s\bigl(\sigma(s,\cdot,\ell_s)\rho(s,\ell_s)\partial_x\phi\bigr)\,ds.$$

These processes capture the dynamics of the limit SPDE.

LEMMA 5.9 (Martingale approach). *Let W be a standard Brownian motion and let ξ and $L_t = 1 - \xi_t(0,\infty)$ be random processes satisfying the conditions of Assumption 2.3. If*

$$M^\phi(\xi,L), \qquad S^\phi(\xi,L) \quad \textit{and} \quad C^\phi(\xi,L,W)$$

are martingales for every $\phi \in C^{\text{test}}$, *then* ξ, L *and* W *satisfy the limit SPDE from Theorem* 1.1.

PROOF. The hypothesis gives

$$[M^{\phi}(\xi, L)]_t = \int_0^t \xi_s\big(\sigma(s, \cdot, L_s)\rho(s, L_s)\partial_x\phi\big)^2\, ds,$$

$$[M^{\phi}(\xi, L), W]_t = \int_0^t \xi_s\big(\sigma(s, \cdot, L_s)\rho(s, L_s)\partial_x\phi\big)\, ds,$$

hence

$$\left[M^{\phi}(\xi, L) - \int_0^{\cdot} \xi_s\big(\sigma(s, \cdot, L_s)\rho(s, L_s)\partial_x\phi\big)\, dW_s\right]_t = 0,$$

for every $t \in [0, T]$, which completes the proof. □

Our strategy is to take a limit in Proposition 3.2 and apply weak convergence. First, notice that we have the following.

LEMMA 5.10. *For every fixed* $\phi \in C^{\text{test}}$, *there exists a deterministic cocountable subset of* $[0, T]$ *on which*

$$M^{\phi}(\nu^{N_k}, L^{N_k})(t) \Rightarrow M^{\phi}(\nu^*, L^*)(t), \qquad S^{\phi}(\nu^{N_k}, L^{N_k})(t) \Rightarrow S^{\phi}(\nu^*, L^*)(t),$$
$$C^{\phi}(\nu^{N_k}, L^{N_k}, W)(t) \Rightarrow C^{\phi}(\nu^*, L^*, W)(t) \qquad \text{in } \mathbb{R}.$$

Furthermore, these sequences are uniformly bounded (for fixed ϕ*).*

PROOF. Note that all the above processes are uniformly bounded (for fixed ϕ) since ν^N is a probability measure. The result then follows by Corollary 5.7. □

PROPOSITION 5.11 (Evolution equation). *Suppose* $(\nu^{N_k}, W) \Rightarrow (\nu^*, W)$. *Then, for every* $\phi \in C^{\text{test}}$, *the processes* $M^{\phi}(\nu^*, L^*)$, $S^{\phi}(\nu^*, L^*)$ *and* $C^{\phi}(\nu^*, L^*, W)$ *from Definition* 5.8 *are martingales. Hence,* ν^* *and* W *satisfy the evolution equation from Theorem* 1.1. *Furthermore,* ν^* *is continuous.*

PROOF. Fix $\phi \in C^{\text{test}}$ and let $\mathbb{T}$ be the cocountable set of times on which we have the conclusion of Lemma 5.10. To show that $M^{\phi}(\nu^*, L^*)$ is a martingale, it is enough to show that, for any arbitrary $k \geq 1$, $s, t \in \mathbb{T}$, $s_1, \dots, s_k \in [0, s] \cap \mathbb{T}$ and $f_1, \dots, f_k : \mathbb{R} \to \mathbb{R}$ continuous and bounded, that the map defined by

$$F(\xi, \ell) := \big(M^{\phi}(\xi, \ell)(t) - M^{\phi}(\xi, \ell)(s)\big) \prod_{i=1}^{k} f_i\big(M^{\phi}(\xi, \ell)(s_i)\big)$$

satisfies $\mathbf{E}F(\nu^*, L^*) = 0$. By Lemma 5.10 and the boundedness and continuity of the f_i's

$$\mathbf{E}F(\nu^*, L^*) = \lim_{k\to\infty} \mathbf{E}F(\nu^{N_k}, L^{N_k}).$$

However, from Proposition 3.2, we have that $M^\phi(\nu^{N_k}, L^{N_k})$ is a martingale since

$$M^\phi(\nu^{N_k}, L^{N_k})(t) = \int_0^t \nu^{N_k}(\sigma(s, \cdot, L_s^{N_k})\rho(s, L^{N_k})\phi)\, dW_s + I_t^{N_k}(\phi), \tag{5.4}$$

therefore $\mathbf{E}F(\nu^{N_k}, L^{N_k}) = 0$ and so $M^\phi(\nu^*, L^*)$ is a martingale.

For S^ϕ, define the map

$$G(\xi, \ell) := (S^\phi(\xi, \ell)(t) - S^\phi(\xi, \ell)(s)) \prod_{i=1}^k f_i(S^\phi(\xi, \ell)(s_i)).$$

By applying Itô's formula to (5.4), we have

$$S^\phi(\nu^{N_k}, L^{N_k})(t) = S^\phi(\nu^{N_k}, L^{N_k})(0) + \text{martingale term} + 2[I^{N_k}(\phi)]_t.$$

So be the boundedness of the f_i and Proposition 3.4

$$\mathbf{E}G(\nu^{N_k}, L^{N_k}) = O(1/N_k),$$

so $\mathbf{E}G(\nu^*, L^*) = 0$ and $S^\phi(\nu^*, L^*)$ is a martingale. The work for C^ϕ follows similarly, so we omit it. The result is then complete by Lemma 5.9, and the continuity of $t \mapsto \nu_t^*$ follows by the fact that the right-hand side of the evolution equation in Theorem 1.1 is continuous. □

6. The kernel smoothing method. The kernel smoothing method converts a measure into an approximating family of functions and, by establishing uniform results on the functions, enables us to show the existence of a density for the measure. In the next section, we will use this to prove Theorem 1.2. Let ζ be a finite signed-measure and p_ε the Gaussian heat kernel

$$p_\varepsilon(x) := (2\pi\varepsilon)^{-1/2} \exp\{-x^2/2\varepsilon\}, \qquad x \in \mathbb{R}.$$

Begin by noting the familiar fact that ζ can be approximated by its convolution with p_ε: For every continuous and bounded $\phi : \mathbb{R} \to \mathbb{R}$,

$$\int_{\mathbb{R}} \phi(x)(\zeta * p_\varepsilon)(x)\, dx \to \zeta(\phi) = \int_{\mathbb{R}} \phi(x)\zeta(dx), \tag{6.1}$$

as $\varepsilon \to 0$, and

$$\bar{T}_\varepsilon \zeta(x) := (p_\varepsilon * \zeta)(x) = \int_{\mathbb{R}} p_\varepsilon(x - y)\zeta(dy) \tag{6.2}$$

is a $C^\infty(\mathbb{R})$ function. We will sometimes abuse notation and write $\bar{T}_\varepsilon\phi = p_\varepsilon * \phi$ when $\phi : \mathbb{R} \to \mathbb{R}$ is a function. With $(\cdot, \cdot)_2$ denoting the usual $L^2(\mathbb{R})$ inner product, we have

$$(\phi, \bar{T}_\varepsilon\zeta)_2 = \zeta(\bar{T}_\varepsilon\phi). \tag{6.3}$$

Our first observation is that $\bar{T}_\varepsilon$ is a contraction on $L^2(\mathbb{R})$:

PROPOSITION 6.1 (Contraction). *Let $f \in L^2(\mathbb{R})$. Then $\|\bar{T}_\varepsilon f\|_2 \le \|f\|_2$, where $\|\cdot\|_2$ is the L^2 norm on $\mathbb{R}$.*

PROOF. The Cauchy–Schwarz inequality gives

$$|\bar{T}_\varepsilon f(x)|^2 = \left|\int_{\mathbb{R}} p_\varepsilon(x-y) f(y)\, dy\right|^2 \le \int_{\mathbb{R}} p_\varepsilon(x-y)\, dy \cdot \int_{\mathbb{R}} p_\varepsilon(x-y) f(y)^2\, dy.$$

The first integral on the right-hand side integrates to one, then integrating over $x \in \mathbb{R}$ completes the proof. □

We now give a condition which shows how to recover the existence of a density via kernel smoothing.

PROPOSITION 6.2. *Suppose that ζ is a finite signed measure and*

$$\liminf_{\varepsilon \to 0} \|\bar{T}_\varepsilon \zeta\|_2 < \infty.$$

Then ζ has an $L^2(\mathbb{R})$ density, that is, there exists $f \in L^2(\mathbb{R})$ such that $\zeta(\phi) = (f, \phi)_2$, for every $\phi \in L^2(\mathbb{R})$. Furthermore, $\|\bar{T}_\varepsilon \zeta\|_2 \to \|f\|_2$ in $\mathbb{R}$.

PROOF. The hypothesis gives a bounded sequence $(\bar{T}_{\varepsilon_n}\zeta)_{n\ge1}$ in $L^2(\mathbb{R})$, with $\varepsilon_n \to 0$. By [24], Appendix D, Theorem 3, we can extract a weakly convergent subsequence

$$(\bar{T}_{\varepsilon_{n_k}}, \phi)_2 \to (f, \phi)_2 \qquad \text{for every } \phi \in L^2(\mathbb{R}),$$

for some $f \in L^2(\mathbb{R})$. But by (6.1) we conclude that $\zeta(\phi) = (f, \phi)_2$ for all $\phi \in \mathscr{S}$, and this gives the first result since $\mathscr{S}$ is dense in $L^2(\mathbb{R})$.

We now have that $\bar{T}_\varepsilon \zeta = \bar{T}_\varepsilon f$, therefore, by Proposition 6.1,

$$\limsup_{\varepsilon \to 0} \|\bar{T}_\varepsilon \zeta\|_2 \le \|f\|_2.$$

By (6.1), we also have

$$|(f, \phi)_2| = \lim_{\varepsilon \to 0} |(\bar{T}_\varepsilon \zeta, \phi)_2| \le \liminf_{\varepsilon \to 0} \|\bar{T}_\varepsilon \zeta\|_2 \|\phi\|_2 \qquad \text{for all } \phi \in \mathscr{S},$$

so $\|f\|_2 \le \liminf_{\varepsilon \to 0} \|\bar{T}_\varepsilon \zeta\|_2$, which completes the proof. □

Smoothing in H^{-1} and the antiderivative. The material above will be used to establish a preliminary regularity result (Proposition 7.1) in Section 7. However, for the main uniqueness proof we will work in a space of lower regularity and on the half-line. Recall that the first Sobolev space with Dirichlet boundary condition, $H_0^1(0,\infty)$, is defined to be the closure of $C_0^\infty(0,\infty)$ under the norm

$$\|f\|_{H^1(0,\infty)} := \left(\|f\|^2_{L^2(0,\infty)} + \|\partial_x f\|^2_{L^2(0,\infty)}\right)^{1/2}.$$

The dual of $H_0^1(0,\infty)$ will be denoted by H^{-1} and its norm by

$$\|\zeta\|_{-1} := \sup_{\|\phi\|_{H^1(0,\infty)}=1} \big|\zeta(\phi)\big|.$$

This is a natural space for us to work in due to the following.

PROPOSITION 6.3. *If ζ is a finite signed measure, then $\zeta \in H^{-1}$.*

PROOF. First, observe that $|\zeta(\phi)| \le |\zeta|\|\phi\|_\infty$, for every $\phi \in C_0^\infty(0,\infty)$. Morrey's inequality [24], Section 5.6, Theorem 4, gives a universal constant, $C > 0$, such that $\|\phi\|_\infty \le C\|\phi\|_{H^1}$, and this completes the proof. □

To work on the half-line, we will use the absorbing heat kernel defined, as in the proof of Proposition 4.4, by

$$G_\varepsilon(x,y) := p_\varepsilon(x-y) - p_\varepsilon(x+y) \qquad \text{for } x, y > 0 \tag{6.4}$$

and define

$$T_\varepsilon\zeta(x) := \int_0^\infty G_\varepsilon(x,y)\zeta(dy).$$

Notice that $G_\varepsilon(x,0) = 0$ for every x, so $y \mapsto G_\varepsilon(x,y)$ is an element of C^{test}, and also notice that $T_\varepsilon\zeta(0) = 0$. For $T_\varepsilon\zeta$ to approximate ζ, we need ζ to be supported on $[0,\infty)$.

PROPOSITION 6.4. *If ζ is supported on $[0,\infty)$, then*

$$(T_\varepsilon\zeta, \phi)_2 \to \zeta(\phi),$$

as $\varepsilon \to 0$, for every ϕ continuous, bounded and supported on $(0,\infty)$.

PROOF. Let $\tilde{\phi}(x) := \phi(-x)$, then from (6.1),

$$(T_\varepsilon\zeta, \phi)_2 = (\bar{T}_\varepsilon\zeta, \phi)_2 - (\bar{T}_\varepsilon\zeta, \tilde{\phi})_2 \to \zeta(\phi) - \zeta(\tilde{\phi}).$$

But by the hypotheses $\zeta(\tilde{\phi}) = 0$, as required. □

To access the H^{-1} norm, we will use the *antiderivative* defined by

$$\partial_x^{-1} f(x) := -\int_x^\infty f(y)\,dy \qquad \text{for } f : \mathbb{R} \to \mathbb{R} \text{ integrable}.$$

Notice that $\partial_x \partial_x^{-1} f = f$, and if $\partial_x f$ is also integrable, then $\partial_x^{-1}\partial_x f = f$, too. The result we will use in Section 7 is the following.

PROPOSITION 6.5. *If $\zeta \in H^{-1}$, then $\|\zeta\|_{-1} \le \liminf_{\varepsilon\to 0} \|\partial_x^{-1} T_\varepsilon\zeta\|_{L^2(0,\infty)}$.*

PROOF. First, notice that for fixed ε

$$\int_0^\infty \int_0^\infty \big(p_\varepsilon(x-y)+p_\varepsilon(x+y)\big)\,dx|\zeta|(dy)<\infty,$$

so $T_\varepsilon\zeta$ is integrable, and hence $\partial_x^{-1}T_\varepsilon\zeta$ is well-defined. Integration by parts gives

$$\big(\partial_x^{-1}T_\varepsilon\zeta,\partial_x\phi\big)_{L^2(0,\infty)}=(T_\varepsilon\zeta,\phi)_{L^2(0,\infty)}=\zeta(T_\varepsilon\phi),$$

for $\phi\in C^\infty(0,\infty)$. Therefore, by Proposition 6.4 we have

$$\begin{aligned}\big|\zeta(\phi)\big| &= \lim_{\varepsilon\to0}\big|\big(\partial_x^{-1}T_\varepsilon\zeta,\partial_x\phi\big)_2\big|\\ &\le \liminf_{\varepsilon\to0}\big\|\partial_x^{-1}T_\varepsilon\zeta\big\|_2\|\phi\|_2\\ &\le \liminf_{\varepsilon\to0}\big\|\partial_x^{-1}T_\varepsilon\zeta\big\|_2\|\phi\|_{H^1},\end{aligned}$$

which gives the result. □

7. Uniqueness of solutions; proof of Theorem 1.2. In this section, we will prove Theorem 1.2. Therefore, take ν, $\tilde\nu$ and W as in the statement with $(\nu^{N_k},W)_{k\ge1}\Rightarrow(\nu,W)$ along some subsequence. Let $L_t=1-\nu_t(0,\infty)$ and $\tilde L_t=1-\tilde\nu_t(0,\infty)$. The first step will be to show that ν has some L^2 regularity (Proposition 7.1), which is due to a comparison with $\bar\nu^{N_k}$ from (3.2) and from the dynamics of Proposition 3.5. We then use this fact, along with energy estimates in H^{-1}, to complete the proof. Several technical lemmas are used throughout this section, however, to aid readability, their full statements and proofs are deferred until Section 8.

L^2-Regularity. The result we will prove in this subsection is the following.

PROPOSITION 7.1 (L^2-regularity). *With ν as introduced at the start of Section 7,*

$$\sup_{s\in[0,T]}\sup_{\varepsilon>0}\|T_\varepsilon\nu_s\|_2^2<\infty \qquad \textit{with probability } 1.$$

We would like to work with some process $\bar\nu$ defined analogously to (3.2) that would satisfy the bound $\nu_t(S)\le\bar\nu_t(S)$, for every $t\in[0,T]$ and $S\subseteq\mathbb{R}$. At this stage, however, we are dealing only with weak limit points, so must recover the required process through a limiting procedure on $(\bar\nu^N)_{N\ge1}$.

LEMMA 7.2 (Whole space SPDE). *On a sufficiently rich probability space, there exists $(\nu^*,\bar\nu^*,W)$ such that (ν^*,W) is equal in law to (ν,W), $\nu_t^*(S)\le\bar\nu_t^*(S)$,* *for every* $t \in [0, T]$ *and* $S \subseteq \mathbb{R}$, *and* $\bar{\nu}^*$ *satisfies the limit SPDE on the whole space*:

$$\bar{\nu}_t^*(\phi) = \nu_0(\phi) + \int_0^t \bar{\nu}_s^*\big(\mu(s, \cdot, L_s)\partial_x\phi\big)\, ds + \frac{1}{2}\int_0^t \bar{\nu}_s^*\big(\sigma^2(s, \cdot)\partial_{xx}\phi\big)\, ds$$
$$+ \int_0^t \bar{\nu}_s^*\big(\sigma(s, \cdot)\rho(s, L_s)\partial_x\phi\big)\, dW_s \qquad \text{with } L_t^* = 1 - \nu_t^*(0, \infty),$$

for every $t \in [0, T]$ *and* $\phi \in \mathscr{S}$, *together with condition* (v) *of Assumption* 2.3 *and the two-sided tail bound*

$$\mathbf{E}\bar{\nu}_t^*\big((-\infty, -\lambda) \cup (\lambda, \infty)\big) = o\big(e^{-\alpha\lambda}\big) \qquad \text{as } \lambda \to +\infty,$$

for every $\alpha > 0$.

PROOF. Notice that in Proposition 5.1 we have carried out sufficient work to prove $(\bar{\nu}^N)_{N\geq 1}$ is tight on $(D_{\mathscr{S}'}, \mathrm{M1})$, hence $(\nu^N, \bar{\nu}^N, W)_{N\geq 1}$ is tight. We can therefore conclude that there is a subsequence $(N_{k_r})_{r\geq 1}$ for which $(\nu^{N_{k_r}}, \bar{\nu}^{N_{k_r}}, W)_{r\geq 1}$ converges in law. Any realisation of this limit must have a marginal law that agrees with the law of (ν, W). As the work in Propositions 5.3 and 5.11 is unchanged for $\bar{\nu}^N$ in place of ν^N, we conclude that $\bar{\nu}^*$ is probability-measure-valued and, due to Proposition 3.5, that $\bar{\nu}^*$ satisfies the limit SPDE on the whole space. Finally, we note that for every $\phi \in \mathscr{S}$ with $\phi \geq 0$ we have $\nu_t^{N_{k_r}}(\phi) \leq \bar{\nu}_t^{N_{k_r}}(\phi)$, therefore,

$$\mathbf{P}\big(\nu_t^*(\phi) > \bar{\nu}_t^*(\phi)\big) \leq \liminf_{r\to\infty} \mathbf{P}\big(\nu_t^{N_{k_r}}(\phi) > \bar{\nu}_t^{N_{k_r}}(\phi)\big) = 0,$$

for every $\phi \in \mathscr{S}$, $\phi \geq 0$, by [4], Theorem 2.1. This inequality holds for all t by the continuity of ν^* and $\bar{\nu}^*$ (which follows from being solutions to the limit SPDE) and suffices to give the required dominance. Condition (v) of Assumption 2.3 is satisfied by $\bar{\nu}^*$ because the proof of Corollary 4.3 uses only the behaviour of $\bar{\nu}^N$. Likewise, the two-sided tail estimate is satisfied due to the same work as in Proposition 4.5. $\square$

Our strategy is to use the kernel smoothing method with L^2-energy estimates on the SPDE satisfied by $\bar{\nu}^*$. This is possible because we do not have to take boundary effects into account, which is the main difficulty in the uniqueness proof that will follow. The following lemma relates $\bar{\nu}^*$ to Proposition 7.1.

LEMMA 7.3. *With* ν *and* $\bar{\nu}^*$ *as above and* $\bar{T}_\varepsilon$ *as in* (6.2), *if*

$$\liminf_{\varepsilon\to\infty} \mathbf{E}\Big[\sup_{s\in[0,T]} \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2\Big] < \infty$$

then Proposition 7.1 *holds*.

PROOF. Since $\nu^* \le \bar{\nu}^*$, $\liminf_{\varepsilon\to\infty} \mathbf{E}[\sup_{s\in[0,T]} \|\bar{T}_\varepsilon \nu_s^*\|_2^2] < \infty$. We would first like to deduce that this fact also holds for $\bar{T}_\varepsilon \nu$, but since the map $\nu_t \mapsto \|\bar{T}_\varepsilon \nu_t\|_2$ might not be continuous on $\mathscr{S}'$, more care must be taken.

By fixing $\{\phi_i\}_{i\ge 1}$ to be the Haar basis of $L^2(\mathbb{R})$, we have

$$\begin{aligned}\mathbf{E} \sup_{t\in[0,T]} \|\bar{T}_\varepsilon \nu_t\|_2^2 &= \mathbf{E} \sup_{t\in[0,T]} \lim_{k\to\infty} \sum_{i=1}^{k} (\bar{T}_\varepsilon \nu_t, \phi_i)_2^2 \\ &\le \liminf_{k\to\infty} \mathbf{E} \sup_{t\in[0,T]} \sum_{i=1}^{k} \nu_t(\bar{T}_\varepsilon \phi_i)^2,\end{aligned} \tag{7.1}$$

by (6.3) and Fatou's lemma. Since each ϕ_i is compactly supported, we have that $\bar{T}_\varepsilon \phi_i \in \mathscr{S}$, therefore, $\nu_t(\bar{T}_\varepsilon \phi_i)$ is equal in law to $\nu_t^*(\bar{T}_\varepsilon \phi_i)$, so by [53], Lemma 13.4.1,

$$\sup_{t\in[0,T]} \sum_{i=1}^{k} \nu_t(\bar{T}_\varepsilon \phi_i)^2 =_{\mathrm{law}} \sup_{t\in[0,T]} \sum_{i=1}^{k} \nu_t^*(\bar{T}_\varepsilon \phi_i)^2.$$

Returning to (7.1), we now have that

$$\mathbf{E} \sup_{t\in[0,T]} \|\bar{T}_\varepsilon \nu_t\|_2^2 \le \liminf_{k\to\infty} \mathbf{E} \sup_{t\in[0,T]} \sum_{i=1}^{k} \nu_t^*(\bar{T}_\varepsilon \phi_i)^2 \le \mathbf{E} \sup_{t\in[0,T]} \|\bar{T}_\varepsilon \nu_t^*\|_2^2.$$

By noting that $0 \le T_\varepsilon \nu_t \le \bar{T}_\varepsilon \nu_t$ and applying Fatou's lemma once more we arrive at

$$\begin{aligned}\mathbf{E}\Big[\liminf_{\varepsilon\to\infty} \sup_{s\in[0,T]} \|T_\varepsilon \nu_s\|_2^2\Big] &\le \mathbf{E}\Big[\liminf_{\varepsilon\to\infty} \sup_{s\in[0,T]} \|\bar{T}_\varepsilon \nu_s\|_2^2\Big] \\ &\le \liminf_{\varepsilon\to\infty} \mathbf{E} \sup_{t\in[0,T]} \|\bar{T}_\varepsilon \nu_t^*\|_2^2 < \infty.\end{aligned}$$

We now have that $\liminf_{\varepsilon\to\infty} \|T_\varepsilon \nu_s\|_2 < \infty$, for every $s \in [0,T]$, with probability 1. Proposition 6.2 implies that ν_t has an $L^2(\mathbb{R})$-density, V_t, for every t and that

$$\|V_s\|_2 \le \liminf_{\varepsilon\to 0} \|T_\varepsilon \nu_s\|_2 \le \liminf_{\varepsilon\to\infty} \sup_{s\in[0,T]} \|T_\varepsilon \nu_s\|_2,$$

therefore, $\sup_{s\in[0,T]} \|V_s\|_2 < \infty$, with probability 1. Then by Proposition 6.1,

$$\sup_{s\in[0,T]} \sup_{\varepsilon>0} \|T_\varepsilon \nu_s\|_2 \le \sup_{s\in[0,T]} \|V_s\|_2 < \infty,$$

almost surely, as required. □

As an immediate consequence of the final part of the previous proof and of the forthcoming proof of Proposition 7.1, we have the existence of a density process for ν.

COROLLARY 7.4 ($L^2(\mathbb{R})$-regularity). *With probability* 1, *for every* $t \in [0, T]$ *there exists* $V_t \in L^2(\mathbb{R})$ *such that* V_t *is supported on* $[0, \infty)$ *and is a density of* ν_t, *that is*,

$$\nu_t(\phi) = \int_0^\infty \phi(x) V_t(x)\, dx \qquad \textit{for every } \phi \in L^2(\mathbb{R}).$$

Furthermore, $\sup_{t\in[0,T]} \|V_t\|_2 < \infty$, *with probability* 1.

REMARK 7.5. We might hope that this argument could be used to prove uniqueness. However, notice that we have no control over $\nu - \tilde{\nu}$, as all we have are upper bounds on solutions.

PROOF OF PROPOSITION 7.1. Fix $x \in \mathbb{R}$ and set the function $y \mapsto p_\varepsilon(x - y) \in \mathscr{S}$ into the SPDE from Lemma 7.2 to get

$$\begin{aligned} d\bar{T}_\varepsilon \bar{\nu}_t^*(x) &= \bar{\nu}_t^*\big(\mu_t(y)\partial_y p_\varepsilon(x-y)\big)\, dt + \frac{1}{2}\bar{\nu}_t^*\big(\sigma_t(y)^2 \partial_{yy} p_\varepsilon(x-y)\big)\, dt \\ &\quad + \bar{\nu}_t^*\big(\sigma_t(y)\rho_t \partial_y p_\varepsilon(x-y)\big)\, dW_t \\ &= -\partial_x \bar{\nu}_t^*\big(\mu_t p_\varepsilon(x-\cdot)\big)\, dt + \frac{1}{2}\partial_{xx}\bar{\nu}_t^*\big(\sigma_t^2 p_\varepsilon(x-\cdot)\big)\, dt \\ &\quad - \rho_t \partial_x \bar{\nu}_t^*\big(\sigma_t p_\varepsilon(x-\cdot)\big)\, dW_t, \end{aligned}$$

with the shorthand from Remark 3.1. We would like to move the diffusion coefficients out of the integral against $\bar{\nu}^*$, and to do so we use Lemma 8.2:

$$\begin{aligned} d\bar{T}_\varepsilon \bar{\nu}_t^* &= -\big(\mu_t \partial_x \bar{T}_\varepsilon \bar{\nu}_t^* - \partial_x \mu_t \bar{\mathcal{H}}_{t,\varepsilon}^{\mu} + \bar{\mathcal{E}}_{t,\varepsilon}^{\mu}\big)\, dt \\ &\quad + \frac{1}{2}\partial_x\big(\sigma_t^2 \partial_x \bar{T}_\varepsilon \bar{\nu}_t^* - \partial_x \sigma_t^2 \bar{\mathcal{H}}_{t,\varepsilon}^{\sigma^2} + \bar{\mathcal{E}}_{t,\varepsilon}^{\sigma^2}\big)\, dt \\ &\quad - \rho_t\big(\sigma_t \partial_x \bar{T}_\varepsilon \bar{\nu}_t^* - \partial_x \sigma_t \bar{\mathcal{H}}_{t,\varepsilon}^{\sigma} + \bar{\mathcal{E}}_{t,\varepsilon}^{\sigma}\big)\, dW_t, \end{aligned}$$

where $\bar{\mathcal{H}}$ is as defined in Lemma 8.2 and the dependence on x is omitted for clarity. Applying Itô's formula to $(\bar{T}_\varepsilon \bar{\nu}_t^*(x))^2$ gives

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

Our strategy is to integrate over $x \in \mathbb{R}$, take a supremum over $t \in [0, T]$ and then take an expectation over the previous equation. For the first task, we appeal to Lemma 8.2, Lemma 8.3 and Young's inequality with free parameter $\eta > 0$ to obtain

$$\begin{aligned}\|\bar{T}_\varepsilon \bar{\nu}_t^*\|_2^2 \le \|\bar{T}_\varepsilon \nu_0\|_2^2 &+ c_\eta \int_0^t \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 \, ds + c_\eta \int_0^t \|\bar{T}_{2\varepsilon} \bar{\nu}_s^*\|_2^2 \, ds \\ &+ c_\eta \int_0^t \|\bar{\mathcal{E}}_{s,\varepsilon}^{\mu}\|_2^2 + \|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma^2}\|_2^2 + \|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}\|_2^2 \, ds \\ &- \int_0^t \int_{\mathbb{R}} [\sigma_s^2 \cdot (1 - (1+\eta)\rho_s^2) - \eta - \eta \mu_s^2] (\partial_x \bar{T}_\varepsilon \bar{\nu}_s^*)^2 \, dx \, ds \\ &- 2 \int_0^t \int_{\mathbb{R}} \rho_s \bar{T}_\varepsilon \bar{\nu}_s^* (\sigma_s \partial_x \bar{T}_\varepsilon \bar{\nu}_s^* + \partial_x \sigma_s \bar{\mathcal{H}}_{s,\varepsilon} + \bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}) \, dx \, dW_s,\end{aligned}$$

where $c_\eta > 0$ is a constant depending only on η. Considering the third line, by Assumption 2.1 it is possible to choose $\eta > 0$ small enough so that

$$\sigma_s^2(x)(1 - (1+\eta)\rho_s^2) - \eta - \eta \mu_s(x)^2 \ge 0 \qquad \text{for all } x \in \mathbb{R}, s \in [0, T], \tag{7.2}$$

therefore,

$$\begin{aligned}\|\bar{T}_\varepsilon \bar{\nu}_t^*\|_2^2 \le \|\bar{T}_\varepsilon \nu_0\|_2^2 &+ c_\eta \int_0^t \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 \, ds + c_\eta \int_0^t \|\bar{T}_{2\varepsilon} \bar{\nu}_s^*\|_2^2 \, ds \\ &+ c_\eta \int_0^t (\|\bar{\mathcal{E}}_{s,\varepsilon}^{\mu}\|_2^2 + \|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma^2}\|_2^2 + \|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}\|_2^2) \, ds \\ &- 2 \int_0^t \int_{\mathbb{R}} \rho_s \bar{T}_\varepsilon \bar{\nu}_s^* (\sigma_s \partial_x \bar{T}_\varepsilon \bar{\nu}_s^* + \partial_x \sigma_s \bar{\mathcal{H}}_{s,\varepsilon} + \bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}) \, dx \, dW_s.\end{aligned}$$

Using Lemma 8.5 to take a supremum over t and then expectation gives

$$\begin{aligned}\mathbf{E} \sup_{s \in [0,t]} \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 \le \|\bar{T}_\varepsilon \nu_0\|_2^2 &+ c_1 \mathbf{E} \int_0^t \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 \, ds + c_1 \mathbf{E} \int_0^t \|\bar{T}_{2\varepsilon} \bar{\nu}_s^*\|_2^2 \, ds \\ &+ c_1 \mathbf{E} \int_0^t (\|\bar{\mathcal{E}}_{s,\varepsilon}^{\mu}\|_2^2 + \|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma^2}\|_2^2 + \|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}\|_2^2) \, ds,\end{aligned}$$

where $c_1 > 0$ is a numerical constant.

Taking lim inf as $\varepsilon \to 0$ over the previous inequality and applying Proposition 6.1 (to $V_0 \in L^2$) and Lemma 8.2 yields

$$\begin{aligned}f(t) := \liminf_{\varepsilon \to 0} \mathbf{E} \sup_{s \in [0,t]} \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 &\le c_1 \|V_0\|_2^2 + 2c_1 \liminf_{\varepsilon \to 0} \mathbf{E} \int_0^t \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 \, ds \\ &\le c_1 \|V_0\|_2^2 + 2c_1 t f(t).\end{aligned}$$

Hence, for $t < 1/4c_1$ we have $f(t) \le 2c_1 \|V_0\|_2^2$. The proof is completed by propagating the argument onto $[1/4c_1, 2/4c_1]$ by the same work as above but started from $s = 1/4c_1$, rather than $s = 0$. This gives

$$\liminf_{\varepsilon \to 0} \mathbf{E}\Big[\sup_{s \in [(4c_1)^{-1}, 2(4c_1)^{-1}]} \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2\Big] \\ \leq 2c_1 \liminf_{\varepsilon \to 0} \mathbf{E}\Big[\sup_{s \in [0, (4c_1)^{-1}]} \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2\Big] \leq (2c_1)^2,$$

and so in general

$$\liminf_{\varepsilon \to 0} \mathbf{E}\Big[\sup_{s \in [k(4c_1)^{-1}, (k+1)(4c_1)^{-1}]} \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2\Big] \leq (2c_1)^{k+1} \qquad \text{for } k \geq 0.$$

Since the largest such k we need to take is $k_0 := 4c_1 T$, the simple bound

$$f(T) \leq \liminf_{\varepsilon \to 0} \mathbf{E} \sum_{k=0}^{k_0 - 1} \sup_{s \in [k(4c_1)^{-1}, (k+1)(4c_1)^{-1}]} \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 \leq \sum_{k=0}^{k_0 - 1} (2c_1)^{k+1} < \infty$$

completes the proof. $\square$

Resuming the uniqueness proof. Returning to proof of Theorem 1.2, notice that for a fixed $x > 0$, the function $y \mapsto G_\varepsilon(x, y)$ from (6.4) is an element of C^{test}. Setting into the SPDE for ν gives

$$d\nu_t\big(G_\varepsilon(x, \cdot)\big) = \nu_t\big(\mu_t \partial_y G_\varepsilon(x, \cdot)\big)\, dt + \frac{1}{2}\nu_t\big(\sigma_t^2 \partial_{yy} G_\varepsilon(x, \cdot)\big)\, dt \\ + \rho_t \nu_t\big(\sigma_t \partial_y G_\varepsilon(x, \cdot)\big)\, dW_t,$$

and by applying Lemma 8.6,

$$dT_\varepsilon \nu_t(x) = -\partial_x \nu_t\big(\mu_t G_\varepsilon(x, \cdot)\big)\, dt + \frac{1}{2}\partial_{xx}\nu_t\big(\sigma_t^2 G_\varepsilon(x, \cdot)\big)\, dt \\ - \rho_t \partial_x \nu_t\big(\sigma_t G_\varepsilon(x, \cdot)\big)\, dW_t - 2\partial_x \nu_t\big(\mu_t p_\varepsilon(x + \cdot)\big)\, dt \\ - 2\rho_t \partial_x \nu_t\big(\sigma_t p_\varepsilon(x + \cdot)\big)\, dW_t.$$

To introduce the antiderivative, we integrate the above equation over $x > 0$ and apply Lemma 8.3 to switch the time and space integrals. [Note: Lemma 8.3 is stated for $\bar{\nu}^*$, however the proof only relies on the tail bound from Assumption 2.3 condition (iii), which is satisfied by ν and $\tilde{\nu}$.] We arrive at

$$d\partial_x^{-1} T_\varepsilon \nu_t(x) = -\nu_t\big(\mu_t G_\varepsilon(x, \cdot)\big)\, dt + \frac{1}{2}\partial_x \nu_t\big(\sigma_t^2 G_\varepsilon(x, \cdot)\big)\, dt \\ - \rho_t \nu_t\big(\sigma_t G_\varepsilon(x, \cdot)\big)\, dW_t - 2\nu_t\big(\mu_t p_\varepsilon(x + \cdot)\big)\, dt \\ - 2\rho_t \nu_t\big(\sigma_t p_\varepsilon(x + \cdot)\big)\, dW_t,$$

which, after rewriting using the notation from Lemma 8.1, becomes

$$
\begin{aligned}
d\partial_x^{-1} T_\varepsilon \nu_t = {} & -\big(\mu_t T_\varepsilon \nu_t + \mathcal{E}^{\mu}_{t,\varepsilon}\big)\,dt + \frac{1}{2}\partial_x\big(\sigma_t^2 T_\varepsilon \nu_t + \mathcal{E}^{\sigma^2}_{t,\varepsilon}\big)\,dt \\
& - \rho_t\big(\sigma_t T_\varepsilon \nu_t + \mathcal{E}^{\sigma}_{t,\varepsilon}\big)\,dW_t - 2\nu_t\big(\mu_t p_\varepsilon(x+\cdot)\big)\,dt \\
& - 2\rho_t \nu_t\big(\sigma_t p_\varepsilon(x+\cdot)\big)\,dW_t.
\end{aligned}
\tag{7.3}
$$

We will now introduce the simplifying notation $o_{\mathrm{sq}}(1)$ to denote any family of $L^2(0,\infty)$-valued processes, $\{(f_{t,\varepsilon})_{t\in[0,T]}\}_{\varepsilon>0}$, satisfying

$$
\mathbf{E}\int_0^T \|f_{t,\varepsilon}\|^2_{L^2(0,\infty)}\,dt \to 0 \qquad \text{as } \varepsilon \to 0.
$$

Thus, a formal linear combination of $o_{\mathrm{sq}}(1)$ terms is of order $o_{\mathrm{sq}}(1)$. Therefore, (7.3) can be written (using Lemma 8.1) as

$$
\begin{aligned}
d\partial_x^{-1} T_\varepsilon \nu_t = {} & -\mu_t T_\varepsilon \nu_t\,dt + \frac{1}{2}\partial_x\big(\sigma_t^2 T_\varepsilon \nu_t + \mathcal{E}^{\sigma^2}_{t,\varepsilon}\big)\,dt - \sigma_t\rho_t T_\varepsilon \nu_t\,dW_t \\
& + o_{\mathrm{sq}}(1)\,dt + o_{\mathrm{sq}}(1)\,dW_t \\
& - 2\nu_t\big(\mu_t p_\varepsilon(x+\cdot)\big)\,dt - 2\rho_t\nu_t\big(\sigma_t p_\varepsilon(x+\cdot)\big)\,dW_t,
\end{aligned}
\tag{7.4}
$$

and we claim that the integrands in the final two terms are also of order $o_{\mathrm{sq}}(1)$. This claim is in fact the critical boundary result from [8], but here we only need first moment estimates.

LEMMA 7.6 (Boundary estimate). *We have*

$$
\mathbf{E}\int_0^T \int_0^\infty \Big(\int_0^\infty p_\varepsilon(x+y)\nu_t(dy)\Big)^2\,dx\,dt \to 0 \qquad \textit{as } \varepsilon \to 0,
$$

hence $\nu_t(\mu_t p_\varepsilon(x+\cdot)) = o_{\mathrm{sq}}(1)$ *and* $\nu_t(\sigma_t p_\varepsilon(x+\cdot)) = o_{\mathrm{sq}}(1)$.

PROOF. Begin by noting that

$$
\begin{aligned}
\big|\nu_t\big(p_\varepsilon(x+\cdot)\big)\big| &\le e^{-x^2/\varepsilon}\int_0^\infty p_\varepsilon(y)\nu_t(dy) \\
&\le c_1 e^{-x^2/\varepsilon}\varepsilon^{-1/2}\big[\nu_t\big(0,\varepsilon^\eta\big) + \exp\{-\varepsilon^{2\eta-1}/2\}\big],
\end{aligned}
$$

for $\eta \in (0,\frac{1}{2})$ a free parameter and $c_1 > 0$ a universal constant. Squaring and integrating over $x > 0$ gives

$$
\int_0^\infty \big|\nu_t\big(p_\varepsilon(x+\cdot)\big)\big|^2\,dx \le c_2\varepsilon^{-1/2}\big[\nu_t\big(0,\varepsilon^\eta\big)^2 + \exp\{-\varepsilon^{2\eta-1}\}\big],
$$

with $c_2 > 0$ another numerical constant. Condition (iv) of Assumption 2.3 and the fact that $\nu_t(S)^2 \le \nu_t(S)$, since ν_t is a sub-probability measure, allows us to write

$$
\mathbf{E}\int_0^T \int_0^\infty \big|\nu_t\big(p_\varepsilon(x+\cdot)\big)\big|^2\,dx = O\big(\varepsilon^{\eta(1+\beta)-1/2}\big) + O\big(\varepsilon^{-1/2}\exp\{-\varepsilon^{2\eta-1}\}\big),
$$

which vanishes if we choose η to satisfy

$$\frac{1}{2(1+\beta)} < \eta < \frac{1}{2},$$

and this completes the proof. $\square$

With Lemma 7.6, we can now reduce (7.4) to

$$\begin{aligned} d\partial_x^{-1} T_\varepsilon \nu_t = {} & -\mu_t T_\varepsilon \nu_t \, dt + \frac{1}{2}\partial_x\big(\sigma_t^2 T_\varepsilon \nu_t + \mathcal{E}_{t,\varepsilon}^{\sigma^2}\big)\, dt - \sigma_t \rho_t T_\varepsilon \nu_t \, dW_t \\ & + o_{\mathrm{sq}}(1)\, dt + o_{\mathrm{sq}}(1)\, dW_t, \end{aligned} \tag{7.5}$$

and this equation is also satisfied by $\tilde{\nu}$, as so far all we have used is Assumption 2.3. Writing $\Delta := \nu - \tilde{\nu}$ and $\delta_t^g(x) := g(t, x, L_t) - g(t, x, \tilde{L}_t)$, taking the difference of (7.5) for ν and $\tilde{\nu}$ yields

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

where $\tilde{\mathcal{E}}_{t,\varepsilon}^{\sigma^2}$ is as in Lemma 8.1, but with ν replaced by $\tilde{\nu}$. Applying Itô's formula to the square $(\partial_x^{-1} T_\varepsilon \Delta_t)^2$ gives

$$\begin{aligned} d\big(\partial_x^{-1} T_\varepsilon \Delta_t\big)^2 = {} & -2\partial_x^{-1} T_\varepsilon \Delta_t\big(\tilde{\mu}_t T_\varepsilon \Delta_t + \delta_t^\mu T_\varepsilon \nu_t\big)\, dt \\ & + \partial_x^{-1} T_\varepsilon \Delta_t \partial_x\big(\sigma_t^2 T_\varepsilon \Delta_t + \mathcal{E}_{t,\varepsilon}^{\sigma^2} - \tilde{\mathcal{E}}_{t,\varepsilon}^{\sigma^2}\big)\, dt \\ & - 2\partial_x^{-1} T_\varepsilon \Delta_t \sigma_t\big(\tilde{\rho}_t T_\varepsilon \Delta_t + \delta_t^\rho T_\varepsilon \nu_t\big)\, dW_t \\ & + \sigma_t^2\big(\tilde{\rho}_t T_\varepsilon \Delta_t + \delta_t^\rho T_\varepsilon \nu_t\big)^2\, dt \\ & + \partial_x^{-1} T_\varepsilon \Delta_t \cdot o_{\mathrm{sq}}(1)\, dt + \partial_x^{-1} T_\varepsilon \Delta_t \cdot o_{\mathrm{sq}}(1)\, dW_t + o_{\mathrm{sq}}(1)^2\, dt. \end{aligned} \tag{7.6}$$

Note that the initial condition for this equation is zero because ν and $\tilde{\nu}$ have the same initial condition.

Since the work in establishing the bounds in Lemma 8.3 only uses the tail estimate (iii) of Assumption 2.3, they remain valid and so, together with Lemma 8.7, the stochastic integrals in (7.6) are martingales for fixed x and ε. Therefore, first taking an expectation and then integrating over $x > 0$ and using Young's inequality with free parameter $\eta > 0$ produces a constant $c_\eta > 0$ such that

$$\begin{aligned} \mathbf{E}\big\|\partial_x^{-1} T_\varepsilon \Delta_t\big\|_2^2 \le {} & c_\eta \mathbf{E}\int_0^t \big\|\partial_x^{-1} T_\varepsilon \Delta_s\big\|_2^2\, ds + c_\eta \mathbf{E}\int_0^t \big\|\big(|\delta_s^\mu| + |\delta_s^\rho|\big)|T_\varepsilon \nu_s|\big\|_2^2\, ds \\ & - \mathbf{E}\int_0^t \int_0^\infty \big[\sigma_s^2\big(1 - (1+\eta)\tilde{\rho}_s^2\big) - \eta - \eta\tilde{\mu}_s^2\big]|T_\varepsilon \Delta_s|^2\, dx\, ds \\ & + o(1), \end{aligned} \tag{7.7}$$

where the terms involving $o_{\mathrm{sq}}(1)$ have collapsed to order $o(1)$. Also notice that (7.7) remains valid if t is a stopping time.

If it was the case that $\mathbf{E}\int_0^t \|T_\varepsilon \Delta_s\|_2^2\, ds = 0$, then by Proposition 6.2 we would have $\Delta = 0$ on $[0,t]$, and so would have completed the proof for this value of t. It is therefore no loss of generality to assume that this value is bounded away from zero for all $\varepsilon > 0$ sufficiently small. Then by taking $\eta > 0$ we can find a positive value $c_0 > 0$ such that

$$\begin{aligned}\mathbf{E}\|\partial_x^{-1} T_\varepsilon \Delta_t\|_2^2 &\le c\mathbf{E}\int_0^t \|\partial_x^{-1} T_\varepsilon \Delta_s\|_2^2\, ds \\ &\quad + c\mathbf{E}\int_0^t \big\|\big(|\delta_s^\mu| + |\delta_s^\rho|\big)|T_\varepsilon \nu_s|\big\|_2^2\, ds - c_0 + o(1),\end{aligned} \tag{7.8}$$

for $c > 0$ constant. We now want to introduce a comparison between solutions in the δ terms, and to do so we consider two cases.

Case 1: *Globally Lipschitz coefficients*. First, consider the simpler case where μ and ρ are Lipschitz in the loss variable, rather than piecewise Lipschitz. Therefore, we have $|\delta_t^g| \le C|L_t - \tilde{L}_t|$, so the inequality in (7.8) becomes

$$\begin{aligned}\mathbf{E}\|\partial_x^{-1} T_\varepsilon \Delta_t\|_2^2 &\le c_1\mathbf{E}\int_0^t \|\partial_x^{-1} T_\varepsilon \Delta_s\|_2^2\, ds \\ &\quad + c_1\mathbf{E}\int_0^t |L_s - \tilde{L}_s|^2 \|T_\varepsilon \nu_s\|_2^2\, ds - c_0 + o(1),\end{aligned}$$

with $c_1 > 0$ constant.

To bound the second term above, we introduce the stopping times

$$t_n := \inf\Big\{t > 0 : \sup_{s\in[0,T]} \sup_{\varepsilon>0} \|T_\varepsilon \nu_s\|_2^2 > n\Big\} \wedge T.$$

From Proposition 7.1 we know that $t_n \to T$ as $n \to \infty$, with probability 1. Since (7.7) is valid for stopping times, we have

$$\begin{aligned}\mathbf{E}\|\partial_x^{-1} T_\varepsilon \Delta_{t\wedge t_n}\|_2^2 &\le c_1\mathbf{E}\int_0^{t\wedge t_n} \|\partial_x^{-1} T_\varepsilon \Delta_s\|_2^2\, ds \\ &\quad + c_1 n\mathbf{E}\int_0^{t\wedge t_n} |L_s - \tilde{L}_s|^2\, ds - c_0 + o(1) \\ &\le c_1\mathbf{E}\int_0^t \|\partial_x^{-1} T_\varepsilon \Delta_{s\wedge t_n}\|_2^2\, ds \\ &\quad + c_1 n\mathbf{E}\int_0^t |L_{s\wedge t_n} - \tilde{L}_{s\wedge t_n}|^2\, ds - c_0 + o(1).\end{aligned}$$

By using the integrating factor $e^{-c_1 t}$, we obtain

$$\mathbf{E}\|\partial_x^{-1} T_\varepsilon \Delta_{t\wedge t_n}\|_2^2 \le c_1 n e^{c_1 T}\mathbf{E}\int_0^t |L_{s\wedge t_n} - \tilde{L}_{s\wedge t_n}|^2\, ds - c_0',$$

and applying Fatou's lemma and Propositions 6.3 and 6.5 gives

$$\mathbf{E}\|\Delta_{t\wedge t_n}\|_{-1}^2 \le c_1 n e^{c_1 T}\mathbf{E}\int_0^t |L_{s\wedge t_n} - \tilde{L}_{s\wedge t_n}|^2\,ds - c_0',$$

where $c_0' = c_0 e^{-c_1 T} > 0$.

Finally, we apply Lemma 8.8 to the above inequality to reintroduce Δ to the right-hand side. With fixed $\alpha > 0$, we have

$$\mathbf{E}\|\Delta_{t\wedge t_n}\|_{-1}^2 \le c_2(\delta^{-1} + \lambda)\mathbf{E}\int_0^t \|\Delta_{s\wedge t_n}\|_{-1}^2\,ds + c_2\delta + c_\alpha e^{-\alpha\lambda} - c_0',$$

where $c_2 > 0$ does not depend on α (but does depend on n). Now fix $\delta = c_0'/c_2$ so that we have

$$\mathbf{E}\|\Delta_{t\wedge t_n}\|_{-1}^2 \le c_3(1 + \lambda)\mathbf{E}\int_0^t \|\Delta_{s\wedge t_n}\|_{-1}^2\,ds + c_\alpha e^{-\alpha\lambda}$$

with $c_3 > 0$ independent of α. By using the integrating factor $e^{-c_3(1+\lambda)t}$, we deduce

$$\mathbf{E}\|\Delta_{t\wedge t_n}\|_{-1}^2 \le c_\alpha e^{c_3(1+\lambda)t-\alpha\lambda},$$

so setting $\alpha = 2c_3 t$ and sending $\lambda \to \infty$ gives $\mathbf{E}\|\Delta_{t\wedge t_n}\|_{-1}^2 = 0$. Therefore, $\nu = \tilde{\nu}$ on $[0, t_n]$, and since $t_n \to T$ we have Theorem 1.2 in Case 1.

Case 2: *Piecewise Lipschitz coefficients*. To extend the argument to the general case, we use a stopping argument and consider the system only on time intervals where the loss processes are in the same interval $[\theta_i, \theta_{i+1})$—recall Assumption 2.1.

Define the stopping times:

$$T_0 := \inf\{t > 0 : L_t \ge \theta_1\} \wedge T, \qquad \tilde{T}_0 := \inf\{t > 0 : \tilde{L}_t \ge \theta_1\} \wedge T$$

and $S_0 = T_0 \wedge \tilde{T}_0$. For the reason immediately proceeding (7.6), the argument in Case 1 can be replicated on $[0, S_0)$ by replacing t by $t \wedge S_0$, since before S_0, the coefficients can be compared using the Lipschitz property on $[\theta_0, \theta_1)$. Therefore, we conclude $\nu_t = \tilde{\nu}_t$ for $t \le S_0$, which forces $L_t = \tilde{L}_t$ for $t \le S_0$ and thus $T_0 = S_0 = \tilde{T}_0$.

We can then repeat the argument for the interval $[S_0, S_1)$, since $\Delta_{S_0} = 0$ (by continuity of ν and $\tilde{\nu}$), where

$$T_1 := \inf\{t > S_0 : L_t \ge \theta_2\} \wedge T, \qquad \tilde{T}_1 := \inf\{t > S_0 : \tilde{L}_t \ge \theta_2\} \wedge T$$

and $S_1 = T_1 \wedge \bar{T}_1$. Continuing up to S_k covers all the $[\theta_i, \theta_{i+1})$ intervals, and this completes the proof, since L and $\tilde{L}$ are increasing [Assumption 2.3, condition (i)] so $[0, T] \subseteq \bigcup_{i=0}^{k-1}[S_i, S_{i+1})$.

8. Technical lemmas. This section collects all the technical lemmas that were used in Section 7, and should be read only as a reference.

LEMMA 8.1. *Let $g_s(x) = g(s, x, L_s)$ where g is one of μ, σ or σ^2 and $L_s = 1 - \nu_s(0, \infty)$. Define the error term*

$$\mathcal{E}^g_{t,\varepsilon}(x) := \nu_t\big(g_t(\cdot)G_\varepsilon(x, \cdot)\big) - g_t(x)T_\varepsilon \nu_t(x).$$

Then

$$\mathbf{E}\int_0^T \|\mathcal{E}^g_{t,\varepsilon}\|^2_{L^2(0,\infty)}\,dt \to 0 \qquad \text{as } \varepsilon \to 0.$$

PROOF. Let $\lambda = \lambda(\varepsilon) \to \infty$, as $\varepsilon \to 0$, be a function that we will specify later. For any $x > 0$,

$$\begin{aligned}\big|\mathcal{E}^g_{t,\varepsilon}(x)\big| &\le \|\partial_x g\|_\infty \int_0^\infty |x-y|p_\varepsilon(x-y)\nu_t(dy)\\ &\le c_1\varepsilon^{\eta-\frac{1}{2}}\nu_t\big(x-\varepsilon^\eta, x+\varepsilon^\eta\big) + c_1\varepsilon^{-1/2}\exp\{-\varepsilon^{2\eta-1}/2\},\end{aligned}$$

with $c_1 > 0$ a universal constant, and where the second line follows by splitting the integral on $|y-x| < \varepsilon^\eta$ and its complement. By considering the range $x < \lambda$ and using condition (v) of Assumption 2.3,

$$\begin{aligned}\mathbf{E}\int_0^T \|\mathcal{E}^g_{t,\varepsilon}\|^2_{L^2(-\lambda,\lambda)}\,dt &= \lambda(\varepsilon)O\big(\varepsilon^{(2+\delta)\eta-1} + \varepsilon^{-1}\exp\{-\varepsilon^{2\eta-1}\}\big)\\ &= \lambda(\varepsilon)O\big(\varepsilon^\gamma\big),\end{aligned}\tag{8.1}$$

for some $\delta, \gamma > 0$, by fixing η in the range

$$\frac{1}{2+\delta} < \eta < \frac{1}{2}.$$

Now consider the range $x \ge \lambda$. Decomposing the y-integral on the range $y < x/2$ and its complement gives

$$\big|\mathcal{E}^g_{t,\varepsilon}(x)\big| \le 2\|g\|_\infty \int_0^\infty p_\varepsilon(x-y)\nu_t(dy) \le c_2 p_\varepsilon(x/2) + c_2\varepsilon^{-1/2}\nu_t\big(|x|/2, +\infty\big),$$

with $c_2 > 0$ another universal constant. Therefore,

$$\begin{aligned}&\mathbf{E}\int_0^T \|\mathcal{E}^g_{t,\varepsilon}\|^2_{L^2((-\lambda,\lambda)^c)}\,dt\\ &\quad= O\bigg(\varepsilon^{-1/2}e^{-\lambda(\varepsilon)^2/8\varepsilon}\int_{-\infty}^\infty p_\varepsilon(x/2)\,dx + \varepsilon^{-1}\int_{\lambda(\varepsilon)}^\infty e^{-x}\,dx\bigg)\\ &\quad= O\big(\varepsilon^{-1}e^{-\lambda(\varepsilon)}\big).\end{aligned}\tag{8.2}$$

Summing (8.1) and (8.2) and fixing $\lambda(\varepsilon) = \log(\varepsilon^{-2})$ completes the proof. $\square$

LEMMA 8.2. *Let $g_s(x) = g(s, x, L_s^*)$ where g is one of μ, σ or σ^2 and $L_s^* = 1 - \bar{\nu}_s^*(0, \infty)$. Define the error term*

$$\bar{\mathcal{E}}_{t,\varepsilon}^g(x) := \partial_x \bar{\nu}_t^* \big(g_t p_\varepsilon(x - \cdot)\big) - g_t(x) \partial_x \bar{T}_\varepsilon \bar{\nu}_t^*(x) + \partial_x g_t(x) \bar{\mathcal{H}}_{t,\varepsilon}^g(x)$$
$$\text{where } \bar{\mathcal{H}}_{t,\varepsilon}^g(x) := \bar{\nu}_t^*\big((x-y)\partial_x p_\varepsilon(x - \cdot)\big).$$

Then

$$\mathbf{E} \int_0^T \|\bar{\mathcal{E}}_{t,\varepsilon}^g\|_{L^2(\mathbb{R})}^2 \, dt \to 0 \qquad \text{as } \varepsilon \to 0$$

and there exists a numerical constant $c > 0$ such that

$$\big|\bar{\mathcal{H}}_{t,\varepsilon}^g(x)\big| \le c \bar{T}_{2\varepsilon} \bar{\nu}_t^*(x) \qquad \text{for all } t \in [0, T], x \in \mathbb{R} \text{ and } \varepsilon > 0.$$

PROOF. Interchanging differentiation and integration with respect to $\bar{\nu}_t^*$ gives

$$\bar{\mathcal{E}}_{t,\varepsilon}^g(x) = \int_{\mathbb{R}} \big[g_t(y) - g_t(x) + (y - x)\partial_x g_t(x)\big] \partial_x p_\varepsilon(x - y) \bar{\nu}_t^*(dy).$$

By bounding with the second-order derivative and using $\partial_x p_\varepsilon(x-y) = -2\varepsilon^{-1}(x - y) p_\varepsilon(x - y)$ gives

$$\big|\bar{\mathcal{E}}_{t,\varepsilon}^g(x)\big| \le \frac{1}{2} \int_{\mathbb{R}} \big|\partial_{xx} g_t(x)\big| |x - y|^3 \varepsilon^{-1} p_\varepsilon(x - y) \bar{\nu}_t^*(dy).$$

We therefore have the same order of ε as in Lemma 8.1, so the first result follows by the same work. For the second result, notice that

$$\big|z \partial_x p_\varepsilon(z)\big| = \frac{1}{\sqrt{2\pi\varepsilon}} \varepsilon^{-1} z^2 e^{-z^2/2\varepsilon} = \sqrt{2} \varepsilon^{-1} z^2 e^{-z^2/4\varepsilon} p_{2\varepsilon}(z),$$

and $\sup_{z \in \mathbb{R}} z^2 e^{-z^2/4\varepsilon} = \varepsilon$. $\square$

LEMMA 8.3 (Stochastic Fubini). *For all $n, m \ge 0$, $\varepsilon > 0$ and $t \in [0, T]$,*

$$\int_{\mathbb{R}} \left(\int_0^t \mathbf{E}\big[\big|\partial_x^n \bar{T}_\varepsilon \bar{\nu}_s^*(x) \cdot \partial_x^m \bar{T}_\varepsilon \bar{\nu}_s^*(x)\big|^2\big] \, ds \right)^{1/2} dx < \infty,$$

hence the stochastic Fubini theorem [52, (1.4)] *gives*

$$\int_{\mathbb{R}} \int_0^t g_t(x) \cdot \partial_x^n \bar{T}_\varepsilon \bar{\nu}_s^*(x) \cdot \partial_x^m \bar{T}_\varepsilon \bar{\nu}_s^*(x) \, dW_s \, dx$$
$$= \int_0^t \int_{\mathbb{R}} g_t(x) \cdot \partial_x^n \bar{T}_\varepsilon \bar{\nu}_s^*(x) \cdot \partial_x^m \bar{T}_\varepsilon \bar{\nu}_s^*(x) \, dx \, dW_s$$

whenever $\sup_{t \in [0,T], x \in \mathbb{R}} |g_t(x)| < \infty$.

PROOF. By applying Young's inequality and concavity of $z \mapsto \sqrt{z}$, it suffices to show that

$$\int_{\mathbb{R}} \Big(\int_0^t \mathbf{E}\big[|\partial_x^n \bar{T}_\varepsilon \bar{\nu}_s^*(x)|^4 \big] \, ds \Big)^{1/2} dx < \infty.$$

First, notice that

$$\partial_x^n \bar{T}_\varepsilon \bar{\nu}_s^*(x) = \bar{\nu}_s^*\big(\partial_x^n p_\varepsilon(x - \cdot)\big) = \bar{\nu}_s^*\big(P_n\big(\varepsilon^{-1}(x - \cdot)\big) p_\varepsilon(x - \cdot)\big),$$

where P_n is a polynomial of degree n. Since $\bar{\nu}_s^*$ is a probability measure, Hölder's inequality gives

$$\mathbf{E}\big[|\partial_x^n \bar{T}_\varepsilon \bar{\nu}_s^*(x)|^4 \big] \le \mathbf{E} \int_{\mathbb{R}} \big| P_n\big(\varepsilon^{-1}(x - y)\big)\big|^4 p_\varepsilon(x - y)^4 \bar{\nu}_s^*(dy). \tag{8.3}$$

For any value of x, the integrand above is bounded (recall that ε is fixed). Hence, it suffices to bound the right-hand side of (8.3) in terms of x only for large values of $|x|$. Splitting the y-integral on the region $|y| < x/2$ and its complement gives the bound

$$\begin{aligned} &\mathbf{E}\big[|\partial_x^n \bar{T}_\varepsilon \bar{\nu}_s^*(x)|^4 \big] \\ &\quad \le c_\varepsilon \mathbf{E} \bar{\nu}_s^*\big((x/2, +\infty) \cup (-\infty, -x/2)\big) + c_\varepsilon \exp\{-x^2/2\varepsilon\} = O\big(e^{-x}\big), \end{aligned}$$

where c_ε and the O depend only on ε and where we have used the tail estimate from Lemma 7.2. This suffices to complete the proof. □

LEMMA 8.4 (An integration-by-parts calculation). *Let $f, g \in C^1(\mathbb{R})$ be bounded with bounded first derivatives. Assume also that these functions and their first derivatives vanish at $\pm\infty$. Then*

$$\int_{\mathbb{R}} g(x) f(x) \partial_x f(x) \, dx = -\frac{1}{2} \int_{\mathbb{R}} \partial_x g(x) f(x)^2 \, dx.$$

PROOF. Integration by parts. □

LEMMA 8.5. *There exists a constant $c > 0$ such that*

$$\begin{aligned} &\mathbf{E} \sup_{u \in [0,t]} \bigg| 2 \int_0^u \int_{\mathbb{R}} \rho_s \bar{T}_\varepsilon \bar{\nu}_s^* \big(\sigma_s \partial_x \bar{T}_\varepsilon \bar{\nu}_s^* + \partial_x \sigma_s \bar{\mathcal{H}}_{s,\varepsilon} + \bar{\mathcal{E}}_{s,\varepsilon}^{\sigma} \big) \, dx \, dW_s \bigg| \\ &\qquad \le \frac{1}{2} \mathbf{E} \sup_{s \in [0,t]} \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 + c \mathbf{E} \int_0^t \|\bar{T}_\varepsilon \bar{\nu}_s^*\|_2^2 \, ds + c \mathbf{E} \int_0^t \|\bar{\mathcal{E}}_{s,\varepsilon}^{\sigma}\|_2^2 \, ds \end{aligned}$$

for all $t \in [0, T]$.

PROOF. By a similar analysis to (8.3), we know that, for every fixed ε, the integrand above is a rapidly decaying function of x, hence the stochastic integral is a martingale, so the Burkholder–Davis–Gundy inequality ([46], Theorem IV.42.1) gives a universal constant, $c_1 > 0$, for which the left-hand side above is bounded by

$$2c_1\mathbf{E}\biggl[\biggl(\int_0^t \biggl(\int_{\mathbb{R}} \rho_s \bar{T}_\varepsilon \bar{v}_s^* (\sigma_s \partial_x \bar{T}_\varepsilon \bar{v}_s^* + \partial_x \sigma_s \bar{\mathcal{H}}_{s,\varepsilon} + \bar{\mathcal{E}}_{s,\varepsilon}^\sigma)\, dx\biggr)^2 ds\biggr)^{1/2}\biggr].$$

By Lemma 8.4, this is equal to a constant multiple of

$$\mathbf{E}\biggl[\biggl(\int_0^t \biggl(\int_{\mathbb{R}} \rho_s \bar{T}_\varepsilon \bar{v}_s^* (-\partial_x \sigma_s \bar{T}_\varepsilon \bar{v}_s^* + \partial_x \sigma_s \bar{\mathcal{H}}_{s,\varepsilon} + \bar{\mathcal{E}}_{s,\varepsilon}^\sigma)\, dx\biggr)^2 ds\biggr)^{1/2}\biggr],$$

which, by Hölder's inequality, is bounded by a constant multiple of

$$\begin{aligned}
&\mathbf{E}\biggl[\biggl(\int_0^t \|\bar{T}_\varepsilon \bar{v}_s^*\|_2^2 \|-\partial_x \sigma_s \bar{T}_\varepsilon \bar{v}_s^* + \partial_x \sigma_s \bar{\mathcal{H}}_{s,\varepsilon}^\sigma + \bar{\mathcal{E}}_{s,\varepsilon}^\sigma\|_2^2\, ds\biggr)^{1/2}\biggr] \\
&\quad \le \mathbf{E}\biggl[\sup_{s\in[0,t]} \|\bar{T}_\varepsilon \bar{v}_s^*\|_2 \biggl(\int_0^t \|-\partial_x \sigma_s \bar{T}_\varepsilon \bar{v}_s^* + \partial_x \sigma_s \bar{\mathcal{H}}_{s,\varepsilon}^\sigma + \bar{\mathcal{E}}_{s,\varepsilon}^\sigma\|_2^2\, ds\biggr)^{1/2}\biggr].
\end{aligned}$$

The result then follows by applying Young's inequality with parameter $1/2$ and using the boundedness of the coefficients. $\square$

LEMMA 8.6 (Switching derivatives). *For all $x, y \in \mathbb{R}$ and $\varepsilon > 0$, we have*:

(i) $\partial_y G_\varepsilon(x, y) = -\partial_x G_\varepsilon(x, y) - 2\partial_x p_\varepsilon(x + y)$,
(ii) $\partial_{yy} G_\varepsilon(x, y) = \partial_{xx} G_\varepsilon(x, y)$.

PROOF. An easy calculation. $\square$

LEMMA 8.7. *For all $x > 0$, $t \in [0, T]$ and $\varepsilon > 0$,*

$$|\partial_x^{-1} T_\varepsilon \Delta_t(x)| \le \nu_t(x/2, +\infty) + \tilde{\nu}_t(x/2, +\infty) + e^{-x^2/8\varepsilon}.$$

PROOF. Split the integral

$$\partial_x^{-1} T_\varepsilon \nu_t(x) = -\int_x^\infty \int_0^\infty G_\varepsilon(y, z)\nu_t(dz)\, dy$$

at $z < x/2$ and its complement to obtain

$$\begin{aligned}
|\partial_x^{-1} T_\varepsilon \nu_t(x)| &\le \frac{1}{\sqrt{2\pi\varepsilon}} \int_x^\infty e^{-(y-x/2)^2/2\varepsilon}\, dy + \nu_t(x/2, +\infty) \\
&\le e^{-x^2/8\varepsilon} + \nu_t(x/2, +\infty).
\end{aligned}$$

The triangle inequality completes the result. $\square$

LEMMA 8.8. *Let* ν, $\tilde{\nu}$, L, $\tilde{L}$ *and* Δ *be as in Section 7. For every* $\alpha > 0$, *there exists a constant* $c_\alpha > 0$ *such that*

$$\mathbf{E}\int_0^t |L_s - \tilde{L}_s|^2\,ds \le c(\delta^{-1} + \lambda)\mathbf{E}\int_0^t \|\Delta_s\|_{-1}^2\,ds + c\delta + c_\alpha e^{-\alpha\lambda},$$

for all $t \in [0, 1]$, $0 < \delta < 1$ *and* $\lambda \ge 1$, *where* $c > 0$ *is a constant that does not depend on* α.

PROOF. For $0 < \delta < 1$ and $\lambda \ge 1$, let $\phi_{\delta,\lambda} \in H_0^1(0,\infty)$ be any cutoff function satisfying

$$\phi_{\delta,\lambda}(x)\begin{cases} = 0, & \text{if } x = 0,\\ \in (0,1), & \text{if } 0 < x < \delta,\\ = 1, & \text{if } \delta \le x \le \lambda,\\ \in (0,1), & \text{if } \lambda < x < \lambda + 1,\\ = 0, & \text{if } x \ge \lambda + 1,\end{cases}$$

$\|\partial_x\phi_{\delta,\lambda}\|_{L^\infty(0,\delta)} \le c_1\delta^{-1}$ and $\|\partial_x\phi_{\delta,\lambda}\|_{L^\infty(\lambda,\lambda+1)} \le c_1$, for some constant $c_1 > 0$. Then

$$\|\phi_{\delta,\lambda}\|_{H_0^1}^2 \le \int_0^{\lambda+1} dx + \int_0^\delta c_1^2\delta^{-2}\,dx + \int_\lambda^{\lambda+1} c_1^2\,dx = c_2(\delta^{-1} + \lambda),$$

for $c_2 > 0$ a constant. Therefore,

$$\begin{aligned} |L_t - \tilde{L}_t| &= |\nu_t(0,\infty) - \tilde{\nu}_t(0,\infty)|\\ &\le |\nu_t(\phi_{\delta,\lambda}) - \tilde{\nu}_t(\phi_{\delta,\lambda})| + |\nu_t(0,\delta)| + |\tilde{\nu}_t(0,\delta)|\\ &\quad + |\nu_t(\lambda,+\infty)| + |\tilde{\nu}_t(\lambda,+\infty)|\\ &\le c_2^{1/2}(\delta^{-1} + \lambda)^{1/2}\|\nu_t - \tilde{\nu}_t\|_{-1} + |\nu_t(0,\delta)| + |\tilde{\nu}_t(0,\delta)|\\ &\quad + |\nu_t(\lambda,+\infty)| + |\tilde{\nu}_t(\lambda,+\infty)|\end{aligned}$$

and so the result follows from conditions (iii) and (iv) of Assumption 2.3 [and that $|\nu_t(S)|^2 \le |\nu_t(S)|$ for all $S \subseteq \mathbb{R}$]. $\square$

The following result will be used in Section 9.

LEMMA 8.9 (Interchanging stochastic integration and conditional expectation). *Suppose we are working on a probability space with filtration* $\{\mathcal{F}_t\}$ *and* W *is a standard Brownian motion with natural filtration* $\{\mathcal{F}_t^W\}$. *Let* H *be a real-valued* $\{\mathcal{F}_t\}$-*adapted process with*

$$\mathbf{E}\int_0^T H_s^2\,ds < \infty.$$

Then, *with probability* 1,

$$\mathbf{E}\bigg[\int_0^t H_s\,dW_s\bigg|\mathcal{F}_t^W\bigg]=\int_0^t \mathbf{E}[H_s|\mathcal{F}_s^W]\,dW_s$$

and

$$\mathbf{E}\bigg[\int_0^t H_s\,dW_s^1\bigg|\mathcal{F}_t^W\bigg]=0$$

for every $t\in[0,T]$.

PROOF. As we can multiply H_s by $\mathbf{1}_{s<t}$, it suffices to take $t=T$. First, suppose that H is a basic process, that is,

$$H_u=Z\mathbf{1}_{s_1<u\le s_2},$$

where $s_1<s_2\le T$ are real numbers and Z is $\mathcal{F}_{s_1}$-measurable. Then

$$\begin{aligned}\mathbf{E}\bigg[\int_0^T H_s\,dW_s\bigg|\mathcal{F}_T^W\bigg]&=\mathbf{E}[Z(W_{s_2}-W_{s_1})|\mathcal{F}_T^W]\\&=\mathbf{E}[Z|\mathcal{F}_{s_1}^W](W_{s_2}-W_{s_1})\\&=\int_0^T\mathbf{E}[Z|\mathcal{F}_s^W]\mathbf{1}_{s_1<s\le s_2}\,dW_s\\&=\int_0^T\mathbf{E}[H_s|\mathcal{F}_s^W]\,dW_s\end{aligned}$$

and

$$\begin{aligned}\mathbf{E}\bigg[\int_0^T H_s\,dW_s^1\bigg|\mathcal{F}_t^W\bigg]&=\mathbf{E}[Z(W_{s_2}^1-W_{s_1}^1)|\mathcal{F}_t^W]\\&=\mathbf{E}[\mathbf{E}[Z(W_{s_2}^1-W_{s_1}^1)|\sigma(\mathcal{F}_T^W,\mathcal{F}_{s_1})]|\mathcal{F}_t^W]\\&=\mathbf{E}[Z\mathbf{E}[(W_{s_2}^1-W_{s_1}^1)|\sigma(\mathcal{F}_T^W,\mathcal{F}_{s_1})]|\mathcal{F}_t^W]\\&=\mathbf{E}[Z\mathbf{E}[W_{s_2}^1-W_{s_1}^1]|\mathcal{F}_T^W]=0,\end{aligned}$$

where we have used the fact that $W_{s_2}^1-W_{s_1}^1$ is independent of $\sigma(\mathcal{F}_T^W,\mathcal{F}_{s_1})$ since W^1 and W are independent and W^1 has independent increments. So the result holds in this case and immediately extends to linear combinations of basic processes. The usual density argument then allows us to extend the result to all required H. $\square$

9. Stochastic McKean–Vlasov problem; proof of Theorem 1.6. This section presents a short proof of Theorem 1.6. Take a strong solution (ν, W) to the limit SPDE (Remark 1.3), an independent Brownian motion $W^{\perp}$ and define X by

$$\begin{cases} X_t = X_0 + \int_0^t \mu(s, X_s, L_s)\, ds + \int_0^t \sigma(s, X_s)\rho(s, L_s)\, dW_s \\ \qquad + \int_0^t \sigma(s, X_s)\big(1 - \rho(s, L_s)^2\big)^{\frac{1}{2}}\, dW_s^{\perp} \\ \tau = \inf\{t > 0 : X_t \le 0\}. \end{cases}$$

[It is possible to find such an X by standard diffusion theory, since $t \to L_t = 1 - \nu_t(0, \infty)$ is given and fixed.] Let $\tilde{\nu}$ be the conditional law of X given W killed at zero, that is,

$$\tilde{\nu}_t(S) := \mathbf{P}(X_t \in S; t < \tau | W).$$

We will have the existence statement of Theorem 1.6 if we can prove $\nu = \tilde{\nu}$.

Applying Itô's formula to $\phi(X_t)$ as in the proof of Proposition 3.2 gives

$$\begin{aligned} \phi(X_t)\mathbf{1}_{t<\tau} = \phi(X_0) &+ \int_0^t (\mu_s \partial_x \phi)(X_s)\mathbf{1}_{s<\tau}\, ds + \frac{1}{2}\int_0^t (\sigma_s^2 \partial_{xx}\phi)(X_s)\mathbf{1}_{s<\tau}\, ds \\ &+ \int_0^t (\sigma_s \rho_s \partial_x \phi)(X_s)\mathbf{1}_{s<\tau}\, dW_s \\ &+ \int_0^t (\sigma_s (1 - \rho_s^2)^{\frac{1}{2}} \partial_x \phi)(X_s)\mathbf{1}_{s<\tau}\, dW_s^{\perp}. \end{aligned}$$

Take a conditional expectation with respect to W by applying Lemma 8.9 [and using that L is $\sigma(W)$-measurable] to get

$$\begin{aligned} \tilde{\nu}_t(\phi) = \nu_0(\phi) &+ \int_0^t \tilde{\nu}_s\big(\mu(s, \cdot, L_s)\partial_x \phi\big)\, ds + \frac{1}{2}\int_0^t \tilde{\nu}_s\big(\sigma^2(s, \cdot, L_s)\partial_{xx}\phi\big)\, ds \\ &+ \int_0^t \tilde{\nu}_s\big(\sigma(s, \cdot)\rho(s, L_s)\partial_x \phi\big)\, dW_s \qquad \text{with } L_t = 1 - \nu_t(0, \infty). \end{aligned}$$

Now, ν also satisfies this equation, however, in both cases the coefficients depend only on L. Therefore, we can regard L as fixed and ν and $\tilde{\nu}$ as solving the limit SPDE in the special case when coefficients do not depend on the loss-variable. This is a much easier linear problem and Theorem 1.2 is certainly sufficient to conclude $\nu = \tilde{\nu}$, as required.

We have also just shown that if (X, W) solves the McKean–Vlasov problem in Theorem 1.6, then its conditional law $\nu = \tilde{\nu}$ solves the limit SPDE. By Theorem 1.2, this fixes the law of ν, hence we have the uniqueness statement too.

10. Open problems. We end by giving some open problems arising from our model and its related extensions:

(i) As indicated at the end of Section 1, the most important practical question is how do we numerically approximate ν from a given realisation of W? This leads to the further questions of how do we combine these approximations to get an estimator for $\mathbf{E}\Psi(L)$, where $\Psi : D_{\mathbb{R}} \to \mathbb{R}$ is some payoff function, and how do we calibrate the model to any data on traded prices for options with payoff $\Psi(L)$?

Our proposed algorithm for the first problem is as follows. Here, we discretise the time variable and treat the outputs of the following subroutines as functions on $[0, \infty)$—in practise we would also need a discretisation scheme for the spatial variable, too, but we will not consider that problem here. Fix a precision level $\delta > 0$ and assume we are given a piecewise constant or piecewise linear approximation to a Brownian trajectory $t \mapsto w_t$ to precision at least δ (generating such a path contributes negligible computational cost in this algorithm) and an initial density $V^{(0)}$. Set $L^{(0)} = 0$. For $1 \le n \le T/\delta - 1$, form $V^{(n)}$ recursively by setting $V^{(n)} = u_\delta$ where u solves the deterministic linear PDE

$$\begin{aligned} du_t(x) &= -\mu\big(t, x, L^{(n-1)}\big)\partial_x u_t(x)\, dt + \frac{1}{2}\sigma(t, x)\rho\big(t, L^{(n-1)}\big)\partial_{xx} u_t(x)\, dt \\ &\quad - \sigma(t, x)\sqrt{1 - \rho\big(t, L^{(n-1)}\big)^2}\partial_x u_t(x)\, dw_t \qquad \text{with } u_t(0) = 0, \end{aligned} \tag{10.1}$$

for $t \in [0, \delta]$ and $x > 0$. Set $L^{(n)} = 1 - \int_0^\infty V^{(n)}(x)\, dx$ (calculated using some quadrature routine). Our approximation to the density process, V, of ν and the loss process, L, are given by piecewise interpolation of $\{V^{(n)}\}_n$ and $\{L^{(n)}\}_n$:

$$\begin{aligned} \tilde{V}_t &:= \big(1 - \operatorname{frac}\{s\}\big)V^{([s])} + \operatorname{frac}\{s\}V^{([s]+1)}, \\ \tilde{L}_t &:= \big(1 - \operatorname{frac}\{s\}\big)L^{([s])} + \operatorname{frac}\{s\}L^{([s]+1)}, \end{aligned}$$

where $s := t/\delta$, $[s]$ is the floor of s and $\operatorname{frac}\{s\} = s - [s]$.

In the case when σ and μ are constant and ρ depends only on the loss variable and w is given as a piecewise constant interpolation of W with precision δ, the solution to (10.1) can be written explicitly in terms of the Brownian transition kernel. A numerical solution can then be found by quadrature. (This instance of the algorithm was used to produce Figure 2.) If these assumption do not hold, then further approximations may be necessary. In [30], (10.1) is solved (for the constant coefficient case) by finite element methods and the scheme is proven to converge when the system is considered on the whole space. The authors conjecture and provide numerical evidence for a convergence rate for the scheme on the half-line with space-time discretisation. A first open problem is to verify that the piecewise-constant time-discretisation, $\tilde{V}$, above converges in law to the solution ν of limit SPDE as $\delta \to 0$. A harder problem is to establish the rate of convergence, in some appropriate norm, averaged over realisations of W.

Returning to the task of calculating the payoff $\mathbf{E}\Psi(L)$, we have the estimator

$$\mathcal{E}_{m,\delta} := \frac{1}{m}\sum_{i=1}^{m} \Psi(\tilde{L}_{w^i,\delta}),$$

where $\{w^i\}_{1\le i\le m}$ are independent standard Brownian motions and $\tilde{L}_{w,\delta}$ denotes the approximation to the loss function using the algorithm above with precision δ and Brownian trajectory w. As the Monte Carlo routine depends on δ, a natural variance reduction technique is to use multi-level Monte Carlo as in [30]. Another potentially useful technique is to alter the drift coefficient in (1.3) using Girsanov's theorem to produce a reweighted estimator. In the case when the payoff function, Ψ, is supported on large losses, and hence is sensitive only to rare events, changing the measure to one under which the particles have a large negative drift and multiplying by the appropriate Radon–Nikodym derivative is a form of importance sampling. A simpler observation in this scenario is that if the systemic Brownian motion has a realisation that has followed a largely increasing path on $[0, T]$, then although that realisation is likely to contribute little to $\mathcal{E}_{m,\delta}$, the negative of this realisation is likely to give a heavy contribution. Hence, the simple antithetic sampling routine in which we draw $2m$ samples of the common Brownian motion in pairs $(w, -w)$ is a candidate for variance reduction. An open problem is to verify the usefulness of these techniques either numerically or analytically.

(ii) Following on from the previous point, a natural extension to the model is to replace the systemic Brownian motion term in (1.3) with a Lévy process. This would allow the possibility of generating extreme losses. Mathematically, we expect to arrive at a nonlinear SPDE driven by a Lévy process on the half-line; see, for example, [34].

(iii) Another possibility for generating large systemic losses is to incorporate a contagion term in the particle dynamics along the lines of [19, 20]. For simplicity, consider the model where particles move according to the dynamics

$$\begin{aligned} X_t^{i,N} &= X_0^i + W_t^i - \alpha L_t^N, \\ \tau^i &= \inf\{t > 0 : X_t^{i,N} \le 0\}, \\ L_t^N &= \sum_{i=1}^N \mathbf{1}_{\tau^i \le t}, \end{aligned} \tag{10.2}$$

with $\alpha > 0$. Whenever a particle hits the origin, every other particle jumps by size α/N towards the boundary. This can begin an avalanche effect where a default causes many other entities to default. Convergence of a finite particle system to a limiting McKean–Vlasov equation is shown in [20], and it is known that for small values of α the solution is unique. For large values of α the limiting system undergoes a jump, whereby a macroscopic proportion of mass is lost in an infinitesimal period of time. It remains a challenge to prove uniqueness of solutions in this regime and to characterise a critical value of α. From our perspective, a natural extension is to consider the system with a common Brownian noise term between particles.

Acknowledgements. The authors thank the anonymous referees for their helpful corrections. We are grateful to Andreas Sojmark for his very thorough reading and suggestions for improvements. S. Ledger thanks Christoph Reisinger and Francois Delarue for discussions on this material.

REFERENCES

[1] AHMAD, F., HAMBLY, B. M. and LEDGER, S. (2016). A stochastic partial differential equation model for mortgage backed securities. Preprint.

[2] ANDERSEN, L. and SIDENIUS, J. (2005). Extensions to the Gaussian copula: Random recovery and random factor loadings. *J. Credit Risk* **1** 29–70.

[3] BAIN, A. and CRISAN, D. (2009). *Fundamentals of Stochastic Filtering*. *Stochastic Modelling and Applied Probability* **60**. Springer, New York. MR2454694

[4] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York.

[5] BLACK, F. and COX, J. (1976). Valuing corporate securities: Some effects of bond indenture provisions. *J. Finance* **31** 351–367.

[6] BUJOK, K. and REISINGER, C. (2012). Numerical valuation of basket credit derivatives in structural jump-diffusion models. *J. Comput. Finance* **15** 115–158.

[7] BURTSCHELL, X., GREGORY, J. and LAURENT, J.-P. (2007). Beyond the Gaussian copula: Stochastic and local correlation. *J. Credit Risk* **3** 31–62.

[8] BUSH, N., HAMBLY, B. M., HAWORTH, H., JIN, L. and REISINGER, C. (2011). Stochastic evolution equations in portfolio credit modelling. *SIAM J. Financial Math.* **2** 627–664.

[9] CÁCERES, M. J., CARRILLO, J. A. and PERTHAME, B. (2011). Analysis of nonlinear noisy integrate & fire neuron models: Blow-up and steady states. *J. Math. Neurosci.* **1** Art. 7, 33. MR2853216

[10] CARMONA, R. and DELARUE, F. (2013). Probabilistic analysis of mean-field games. *SIAM J. Control Optim.* **51** 2705–2734.

[11] CARMONA, R., DELARUE, F. and LACKER, D. (2015). Mean field games with common noise. Available at http://arxiv.org/abs/1407.6181.

[12] CHASSAGNEUX, J.-F., CRISAN, D. and DELARUE, F. (2015). A probabilistic approach to classical solutions of the master equation for large population equilibria. Available at http://arxiv.org/abs/1411.3009.

[13] CHERUBINI, U., LUCIANO, E. and VECCHIATO, W. (2004). *Copula Methods in Finance*. Wiley, Chichester. MR2250804

[14] CHIANG, T. S. (1994). McKean–Vlasov equations with discontinuous coefficients. *Soochow J. Math.* **20** 507–526. MR1309485

[15] CRISAN, D., KURTZ, T. G. and LEE, Y. (2014). Conditional distributions, exchangeable particle systems, and stochastic partial differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 946–974. MR3224295

[16] CRISAN, D. and XIONG, J. (2010). Approximate McKean–Vlasov representations for a class of SPDEs. *Stochastics* **82** 1–16.

[17] DAI PRA, P., RUNGGALDIER, W. J., SARTORI, E. and TOLOTTI, M. (2009). Large portfolio losses: A dynamic contagion model. *Ann. Appl. Probab.* **19** 347–394. MR2498681

[18] DAWSON, D. and GREVEN, A. (2014). *Spatial Fleming–Viot Models with Selection and Mutation*. Springer, Berlin.

[19] DELARUE, F., INGLIS, J., RUBENTHALER, S. and TANRÉ, E. (2015). Global solvability of a networked integrate-and-fire model of McKean–Vlasov type. *Ann. Appl. Probab.* **25** 2096–2133. MR3349003

[20] DELARUE, F., INGLIS, J., RUBENTHALER, S. and TANRÉ, E. (2015). Particle systems with a singular mean-field self-excitation. Application to neuronal networks. *Stochastic Process. Appl.* **125** 2451–2492. MR3322871

[21] DE MASI, A., GALVES, A., LÖCHERBACH, E. and PRESUTTI, E. (2015). Hydrodynamic limit for interacting neurons. *J. Stat. Phys.* **158** 866–902. MR3311484

[22] DING, X., GIESECKE, K. and TOMECEK, P. (2009). Time-changed birth processes and multiname credit derivatives. *Oper. Res.* **57** 990–1005.

[23] ERRAIS, E., GIESECKE, K. and GOLDBERG, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM J. Financial Math.* **1** 642–665.

[24] EVANS, L. C. (2010). *Partial Differential Equations*, 2nd ed. *Graduate Studies in Mathematics* **19**. Amer. Math. Soc., Providence, RI. MR2597943

[25] FANG, F., JÖNSSON, H., OOSTERLEE, C. and SCHOUTENS, W. (2010). Fast valuation and calibration of credit default swaps under Lévy processes. *J. Comput. Finance* **14** 1–30.

[26] FINGER, C. C. (2005). Issues in the pricing of synthetic CDOs. *J. Credit Risk* **1** 113–124.

[27] FREY, R. and MCNEIL, A. (2003). Dependent defaults in modes of portfolio credit risk. *J. Risk* **6** 59–92.

[28] GIESECKE, K., SPILIOPOULOS, K., SOWERS, R. B. and SIRIGNANO, J. A. (2015). Large portfolio asymptotics for loss from default. *Math. Finance* **25** 77–114.

[29] GIESECKE, K. and WEBER, S. (2006). Credit contagion and aggregate losses. *J. Econom. Dynam. Control* **30** 741–767. MR2224986

[30] GILES, M. and REISINGER, C. (2012). Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance. *SIAM J. Financial Math.* **3** 572–592.

[31] HULL, J. and WHITE, A. (2001). Valuing credit default swaps II: Modeling default correlations. *J. Derivatives* **8** 12–21.

[32] JIN, L. (2010). Particle systems and SPDEs with applications to credit modelling. Ph.D. thesis, Univ. Oxford.

[33] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer, New York.

[34] KIM, K.-H. (2014). A Sobolev space theory for parabolic stochastic PDEs driven by Lévy processes on C^1-domains. *Stochastic Process. Appl.* **124** 440–474. MR3131301

[35] KOLOKOLTSOV, V. N. (2013). Nonlinear diffusions and stable-like processes with coefficients depending on the median or VaR. *Appl. Math. Optim.* **68** 85–98. MR3072241

[36] KOLOKOLTSOV, V. N. and TROEVA, M. (2015). On the mean field games with common noise and the McKean–Vlasov SPDEs. Available at http://arxiv.org/abs/1506.04594.

[37] KOTELENEZ, P. (1995). A class of quasilinear stochastic partial differential equations of McKean–Vlasov type with mass conservation. *Probab. Theory Related Fields* **102** 159–188.

[38] KURTZ, T. G. and XIONG, J. (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.* **83** 103–126.

[39] LEDGER, S. (2014). 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. MR3249578

[40] LEDGER, S. (2016). Skorokhod's M1 topology for distribution-valued processes. *Electron. Commun. Probab.* **21** Paper No. 34, 11. MR3492929

[41] LINDSKOG, F. and MCNIEL, A. (2003). Common Poisson shock models: Applications to insurance and credit risk modelling. *Astin Bull.* **33** 209–238.

[42] LUÇON, E. and STANNAT, W. (2014). Mean field limit for disordered diffusions with singular interactions. *Ann. Appl. Probab.* **24** 1946–1993. MR3226169

[43] MERINO, S. and NYFELER, M. A. (2002). Calculating portfolio loss. *RISK* 82–86.

[44] MERTON, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *J. Finance* **29** 449–470.

[45] MORTENSEN, A. (2006). Semi-analytical valuation of basket credit derivatives in intensity-based models. *J. Derivatives* **13** 8–26.

[46] ROGERS, L. C. G. and WILLIAMS, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol. 2. Cambridge Mathematical Library*. Cambridge Univ. Press, Cambridge.

[47] RUDIN, W. (1987). *Real and Complex Analysis*, 3rd ed. McGraw-Hill, New York. MR0924157

[48] SCHÖNBUCHER, P. J. (2003). *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. Wiley, New York.

[49] SPILIOPOULOS, K., SIRIGNANO, J. A. and GIESECKE, K. (2014). Fluctuation analysis for the loss from default. *Stochastic Process. Appl.* **124** 2322–2362.

[50] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *Ecole D'Eté de Probabilités de Saint-Flour XIX*—1989 (P.-L. Hennequin, ed.). *Lecture Notes in Mathematics, Chapter* 3, **1464** 165–251. Springer, Berlin.

[51] VASICEK, O. (1991). Limiting loan loss probability distribution. Technical report, KMV Corporation.

[52] VERAAR, M. (2012). The stochastic Fubini theorem revisited. *Stochastics* **84** 543–551. MR2966093

[53] WHITT, W. (2002). *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer, New York. MR1876437

[54] WU, J.-L. and YANG, W. (2013). Valuation of synthetic CDOs with affine jump-diffusion processes involving Lévy stable distributions. *Math. Comput. Modelling* **57** 570–583. MR3011182

[55] ZHOU, C. (2001). An analysis of default correlations and multiple defaults. *Rev. Financ. Stud.* **14** 555–576.

MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD
WOODSTOCK ROAD
OXFORD
OX2 6GG
UNITED KINGDOM
E-MAIL: hambly@maths.ox.ac.uk

HEILBRONN INSTITUTE
FOR MATHEMATICAL RESEARCH
UNIVERSITY OF BRISTOL
HOWARD HOUSE
BRISTOL
BS8 1SN
UNITED KINGDOM
E-MAIL: sean.ledger@bristol.ac.uk