A large-population limit for a Markovian model of group-structured populations

A. Puhalskii, B. Simon

July 28, 2018

Abstract

A Markovian model of group-structured (two-level) population dynamics features births, deaths, and migrations of individuals, and fission and extinction of groups. These models are useful for studying group selection and other evolutionary processes that occur when individuals live in distinct groups. We show that the sample paths of a properly scaled sequence of these models converge in an appropriate Skorohod space to a deterministic trajectory that is a unique solution to a quasilinear evolution equation. The PDE model can therefore be justified as an approximation to the Markovian one.

1 Introduction

A group-structured population is one comprised of individuals that live in distinct groups. The individuals may be of different types, e.g., "cooperators" and "defectors", based on different heritable traits. Group-structured populations are very common in the natural world. Social animals often live in distinct groups, e.g., packs of wolves, ant colonies, tribes of hunter-gatherers; and group structure exist in other populations as well, e.g., the parasites that reside in a given host organism constitute a group. The individuals in a group-structured population are not necessarily multicellular organisms. For example, the individuals could be microbes, like Dictyostelium (a species of amoeba), in which case the corresponding groups would be the multicellular organisms made of them (slime molds).

Evolutionary theorists since Darwin's time have wondered how group structure may (or may not) affect evolution, and in particular the evolution of cooperative (altruistic) behaviors. Writing about prehistoric tribes of hunter-gatherers, Darwin [3, p. 166] famously said

> *There can be no doubt that a tribe including many members who, from possessing in high degree the spirit of patriotism, fidelity, obedience, courage, and*

sympathy, were always ready to aid one another, and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection.

In other words, although altruistic behavior may be reproductively disadvantageous for individuals (the hunter-gatherers) within the groups (the tribes), groups with more altruistic individuals fare better, and over time this could lead to the evolution of altruistic individuals in the population as a whole. Altruistic behavior cannot easily evolve in a well-mixed population[1] since altruists can be exploited by nonaltruists, thereby lowering their "fitness", but it may evolve more easily in group-structured populations, where more-cooperative groups have some survival advantage.

By the 1960's Darwin's thoughts on this topic had been rephrased in terms of the efficacy of group selection. The best arguments and mathematical models at that time, e.g., Maynard Smith [18], Williams [29], implied that group selection, while theoretically possible, was much too weak to have an effect in the natural world. Explanatory theories like "the selfish gene", Williams [29], Dawkins [4] and "kin selection", Hamilton [8], Maynard Smith [18], reinforced those arguments. Later, Wilson [30], Sober and Wilson [25], Wilson and Wilson [31], and others gradually reopened the debate, being more careful about the way group selection was defined. The new group selection proponents realized that group selection (properly defined) was a potent evolutionary force, however they lacked mathematical models to help make their cases. The problem with all the mathematical models of group selection up to that point was that they did not properly account for group-level birth and death events like fission and extinction. Furthermore, the models were typically designed for an equilibrium analysis only, so the time-dependent mechanisms that led to the evolution of cooperation/altruism remained mysterious. The later proponents of group selection were right about the efficacy of group selection, but they did not have a good mathematical explanation of how the phenomenon works.

In Simon [21] a Markovian model of group-structured population dynamics was proposed that featured individual-level events like births, deaths, and migrations, and also group-level events like fission and extinction. The paper also included a heuristic derivation of a PDE based on the Markovian model. The population dynamics from the Markovian model and PDE model are very similar to each other (Figure 1), leading one to suspect that there is a fundamental mathematical connection between the two models. The purpose of this paper is to make the connection precise. The PDE trajectory is proven to be the unique limit of a certain scaled sequence of sample paths from the Markovian model.

[1] If the individuals are sufficiently closely related then Hamilton's rule [8] shows that altruistic traits can evolve.

Numerical experiments with the Markovian and PDE models from Simon [21] showed that previous models that ignored or misrepresented group-level birth and death events vastly underestimated the strength of group selection, Simon, Fletcher, Doebeli [22], Simon and Pilosov [24]. There are other contemporary models of group-structured populations analyzed in the literature that also shed light on the process of group selection. Traulsen and Nowak [28] proposed a model based on nested Moran processes. They found conditions for a single mutant cooperator in a population of defectors to have a better chance of fixing in the population than a single mutant defector in a population of cooperators. Their results are valid in a limiting regime as the rates of group-level births and deaths approach zero, and as the fitness advantage for defectors approaches zero (weak selection). Another difference between their work and ours is that by looking only at fixation probabilities, their analysis did not involve population dynamics. Luo [16] and Luo and Mattingly [17] also considered nested Moran processes as models of group-structured populations. They obtained large–population limits for the group process. One limit is the solution of a deterministic PDE and the other one is a Fleming–Viot process. A wide variety of large–population limits for Markovian models of a similar kind, but restricted to a single biological level (no group structure), can be found in Champagnat, et.al. [2] and Puhalskii and Simon [20].

The present results are the same sort as the PDE limit in Luo and Mattingly [17], but the model of group-structured populations studied here is more (biologically) realistic. In particular, the birth and death events at each level are decoupled here, i.e., it is not necessary to assume that the number of groups and the sizes of the groups are constant, which is a defining feature of the Moran process. Furthermore, the group-level birth event in the present model is fission, which is more realistic than the "group cloning" birth event implicit in a model based on Moran processes. Migrations are introduced into the model and the number of types of individuals is not restricted to two, unlike in Luo and Mattingly [17]. On the technical side, abandoning the conventions of the Moran model forces one to grapple with convergence of processes assuming values in noncompact spaces of measures and accounting for migrations results in the limit equation being quasilinear whereas the PDE in Luo and Mattingly [17] is semilinear. In addition, rather than working with the generator of the Markov process of the number of groups which seems to be problematic in our setup, we work directly with the balance equations for the process trajectories. We prove convergence to the limit PDE for two different topologies on the state–space of the population process. One is the topology of weak convergence of nonnegative measures akin to that used in Luo and Mattingly [17] and the other is a stronger topology of weak convergence in an $\mathbb{L}^p$–space. The second mode of convergence stipulates a different scaling.

2 The Markovian model of group-structured populations

We consider a population of a finite number of groups, where each group consists of a finite number of individuals. There are ℓ types of individuals, e.g., cooperators and defectors. The state of a group is specified by the number of each type in the group. An $i = (i_1, \ldots, i_\ell)$–group is a group with i_1 type 1 individuals, i_2 type 2 individuals and so on. (We treat the i as $\mathbb{R}^\ell$–vectors, e.g., $i - i' = (i_1 - i'_1, \ldots, i_\ell - i'_\ell)$, and define $i \geq i'$ to mean that the entries of $i - i'$ are nonnegative. We denote $|i| = i_1 + \ldots + i_\ell$ and let e_k represent the kth element of the standard basis in $\mathbb{R}^\ell$, e.g., $e_1 = (1, 0, \ldots, 0)$.) Within each group, individuals independently give birth (asexually and without mutation) and die at stochastic rates which may depend on the individual type and the state of the group. The per capita birth rate of type k individuals in an i-group is $\beta^k(i)$, $k = 1, \ldots, \ell$. Likewise, the per capita death rate of type k individuals in an i-group is $\delta^k(i)$. (Naturally, we assume that $\beta^k(i) = \delta^k(i) = 0$ when $i_k = 0$.)

Let $X_t(i)$ be the number of i–groups in the population at time t. Then $\{X_t(i),\ i \in \mathbb{Z}_+^\ell \setminus \{0\}\}$ specifies the state of the population at time t. In the model, the groups independently die of extinction, the extinction rate for an i–group being $\epsilon(i) X_t^*$, where $X_t^* = \sum_i X_t(i)$ is the total number of groups at time t. An extinction event is the instantaneous death of all the individuals in a group, i.e., the death of the group. The other group–level event in the model is fission. Let us say that (unordered) set π_i of ℓ–vectors with nonnegative integer entries is a partition of i if the vectors are nonzero and $\sum_{j \in \pi_i} j = i$. An i–group fissions at rate $\phi(i)$ (independently of the states of the other groups) which means that i is split according to partition π_i, which is chosen at random, the elements of the partition determine the makeup of "offspring" groups and the "parent" group ceases to exist. Thus, there is a conservation of individuals under fission. (It is allowed for a partition to consist of the single vector i which constitutes "a nonproper fission".) The probability that an i–group is fissioned according to partition π_i is denoted by $\zeta_i(\pi_i)$. Evidently, $\sum_{\pi_i} \zeta_i(\pi_i) = 1$. Let $\pi_i(i')$ denote the number of i'–groups in partition π_i. We let $\eta(i, i') = \sum_{\pi_i} \pi_i(i') \zeta_i(\pi_i)$ denote the expected number of i'–groups produced by the fission. Finally, individuals can independently migrate from one group to another in the model. The per capita migration rate of type k individuals in an i–group is $\mu^k(i)$. It is assumed that a migrating individual chooses a group from the population to join (possibly, the one they are coming from), each with equal probability, considering themselves as a member of their group when deciding on the move.

Let

$B_t^k(i)$ represent the number of type k births in all i–groups in $[0,t]$, where $B_t^k(i)=0$ when $i_k=0$,

$D_t^k(i)$ represent the number of type k deaths in all i–groups in $[0,t]$, where $D_t^k(i)=0$ when $i_k=0$,

$M_t^k(i)$ represent the number of type k immigrations to all i–groups in $[0,t]$,

$\overline{M}_t^k(i)$ represent the number of type k emigrations from all i–groups in $[0,t]$, where $\overline{M}_t^k(i)=0$ when $i_k=0$,

$\overline{F}_t(i)$ represent the number of fissions of i–groups in $[0,t]$,

$F_t(i',i)$ represent the number of i–groups that are produced as a result of fissioning of i'–groups in $[0,t]$,

$E_t(i)$ represent the number of i–groups that get extinct in $[0,t]$.

It is assumed that all these processes take values in $\mathbb{Z}_+$, are equal to zero when $t=0$ and have nondecreasing piecewise constant rightcontinuous trajectories with lefthand limits. All the processes with the exception of $F_t(i',i)$ have unit jumps. The jump sizes of the process $F_t(i',i)$ are determined by the numbers of groups that the i'–groups may fission into. The birth, death, migration, and extinction processes are assumed to be independent Poisson processes. In order to be more specific, we introduce independent "primitives" as follows. Let, for $i\in\mathbb{Z}_+^\ell\setminus\{0\}$, $k\in\{1,\ldots,\ell\}$, $p\in\mathbb{N}$, $l\in\mathbb{N}$, $r\in\mathbb{N}$, and $X=(X(i))\in\mathbb{Z}_+^{\mathbb{Z}_+^\ell\setminus\{0\}}$ with $X^*=\sum_i X(i)\in(0,\infty)$,

$L_t^{B,k}(i,p,r)$ represent Poisson processes of rates $\beta^k(i)$,

$L_t^{D,k}(i,p,r)$ represent Poisson processes of rates $\delta^k(i)$,

$L_t^{\overline{M},k}(i,p,r)$ represent Poisson processes of rates $\mu^k(i)$,

$L_t^{\overline{F}}(i,p)$ represent Poisson processes of rates $\phi(i)$,

$L_t^E(i,p)$ represent Poisson processes of rates $\epsilon(i)$,

$\vartheta_i^k(p,r,X)$, with $i_k\ge 1$, represent random variables assuming values in $\mathbb{Z}_+^\ell\setminus\{0\}$ such that

$$\mathbf{P}\big(\vartheta_i^k(p,r,X)=i'\big)=\frac{X(i')}{X^*},$$

$\theta_i(p)$ represent random partitions of i distributed as ζ_i.

(Informally, p represents the index of a group and r represents the index of an individual in a group, $\vartheta_i^k(p,r,X)$ represents the makeup of the group the rth type k migrating individual of the pth i–group joins upon migration when the state of the population is X, and $\theta_i(p)$ is the set of groups that the pth fissioning i–group fissions into.) All these processes and random variables are assumed mutually independent for different i, or p, or r.

We let $\theta_i(i',p)$ represent the number of i'–groups in the partition $\theta_i(p)$ so that $\mathbf{P}(\theta_i(i',p)=j)=\sum_{\pi_i:\,\pi_i(i')=j}\zeta_i(\pi_i)$. It is noteworthy that

$$\mathbf{E}\theta_i(i',p)=\eta(i,i')\,. \tag{2.1}$$

We assume that the $(X_t(i)\,,t\geq 0)$ are $\mathbb{Z}_+$–valued processes which have piecewise constant rightcontinuous trajectories with limits on the left and that the following recursions are satisfied, with Δ representing the jump of a process, with $t-$ denoting the lefthand limit so that, e.g., $\Delta X_t(i)=X_t(i)-X_{t-}(i)$ and with $\mathbf{1}_\Gamma$ denoting the indicator function of event or element Γ:

$$\begin{aligned}
\Delta B_t^k(i) = \sum_{p=1}^{X_{t-}(i)}\sum_{r=1}^{i_k}\Delta L_t^{B,k}(i,p,r)\,,\ \Delta D_t^k(i) &= \sum_{p=1}^{X_{t-}(i)}\sum_{r=1}^{i_k}\Delta L_t^{D,k}(i,p,r)\,,\\
\Delta\overline{M}_t^k(i) = \sum_{p=1}^{X_{t-}(i)}\sum_{r=1}^{i_k}\Delta L_t^{\overline{M},k}(i,p,r)&(1-\mathbf{1}_{\vartheta_i^k(p,L_t^{\overline{M},k}(i,p,r),X_{t-})}(i)),\\
\Delta M_t^k(i) = \sum_{i'\neq i}\sum_{p=1}^{X_{t-}(i')}\sum_{r=1}^{i'_k}\Delta L_t^{\overline{M},k}(i',p,r)&\mathbf{1}_{\vartheta_{i'}^k(p,L_t^{\overline{M},k}(i',p,r),X_{t-})}(i)\,,\\
\Delta\overline{F}_t(i) = \sum_{p=1}^{X_{t-}(i)}\Delta L_t^{\overline{F}}(i,p)\,,\ \Delta F_t(i,i') &= \sum_{p=1}^{X_{t-}(i)}\theta_i(i',p)\Delta L_t^{\overline{F}}(i,p)\,,\\
\Delta E_t(i) &= \sum_{p=1}^{X_{t-}(i)X^*_{t-}}\Delta L_t^E(i,p)
\end{aligned} \tag{2.2}$$

and

$$\begin{aligned}
\Delta X_t(i) = &-\sum_{k=1}^{\ell}\Delta B_t^k(i)+\sum_{k=1}^{\ell}\Delta B_t^k(i-e_k)-\sum_{k=1}^{\ell}\Delta D_t^k(i)+\sum_{k=1}^{\ell}\Delta D_t^k(i+e_k)\\
&-\sum_{k=1}^{\ell}\Delta M_t^k(i)+\sum_{k=1}^{\ell}\Delta M_t^k(i-e_k)-\sum_{k=1}^{\ell}\Delta\overline{M}_t^k(i)+\sum_{k=1}^{\ell}\Delta\overline{M}_t^k(i+e_k)\\
&-\Delta\overline{F}_t(i)+\sum_{i'}\Delta F_t(i',i)-\Delta E_t(i)\,.
\end{aligned} \tag{2.3}$$

(We also assume that $B_t^k(i) = M_t^k(i) = 0$ when $i_k = -1$.) Induction on the jump epochs of the primitive processes shows that (2.2) and (2.3) admit a unique solution for given X_0 up to the time that $X_t^* = 0$, which may happen never. From that time on, we let the righthand sides of (2.2) vanish, so, $X_t(i) = 0$. If the $\beta^k(i)$ are bounded above uniformly in i by some β^k, then the population stays finite at all times, provided it is finite initially, because it does not exceed the value of the Yule process with birth rate $\sum_{k=1}^{\ell} \beta^k$ and the population at time zero being X_0^*.

Let $X_t = (X_t(i), i \in \mathbb{Z}_+^\ell \setminus \{0\})$. It is a Markov process with values in $\mathbb{Z}_+^{\mathbb{Z}_+^\ell \setminus \{0\}}$. One can view X_t as a density with respect to the counting measure on $\mathbb{Z}_+^\ell \setminus \{0\}$ so that X_t can be identified with the measure on $\mathbb{Z}_+^\ell \setminus \{0\}$ induced by the density and defined by $\Lambda_t(\Gamma) = \sum_i X_t(i)\mathbf{1}_\Gamma(i)$, where $\Gamma \subset \mathbb{Z}_+^\ell \setminus \{0\}$. Then, $(X_t, t \geq 0)$ can be referred to as a measure–valued process.

For the limit theorem, we consider a sequence of models as above, labelled with two parameters, m and n, which we let go to infinity. Accordingly, the variables we have introduced are supplemented with superscripts n and m, e.g., $B_t^{n,m,k}(i)$ stands for the number of type k births in all i–groups in $[0,t]$ for the (m,n)–model. Informally, m characterises the group number and n characterises the group sizes. It is assumed throughout that the functions $i_k(\beta^{n,m,k}(i) + \delta^{n,m,k}(i) + \mu^{n,m,k}(i))$, $\phi^{n,m}(i)$ and $m\epsilon^{n,m}(i)$ are bounded in $n, m,$ and i, that the $m\epsilon^{n,m}(i)$ are bounded away from zero, and that the number of groups that may be produced as a result of fissioning is bounded, i.e., the random variables $\theta_i^{n,m}(i',p)$ are bounded. It is convenient to extend the domain of i to all of $\mathbb{Z}_+^\ell$ so that $\beta^{n,m,k}(0)$ and similar quantities are defined, and define $X_t^{n,m}(0) = 0$.

Let

$$\Lambda_t^{n,m}(\Gamma) = \frac{1}{m} \sum_i X_t^{n,m}(i) \mathbf{1}_\Gamma\big(\frac{i}{n}\big), \tag{2.4}$$

where $\Gamma \subset \mathbb{R}_+^\ell$, and let $\Lambda^{n,m} = ((\Lambda_t^{n,m}(\Gamma), \Gamma \in \mathcal{B}(\mathbb{R}_+^\ell)), t \geq 0)$. Note that $\Lambda_t^{n,m}(\{0\}) = 0$. The process $\Lambda^{n,m}$ takes values in the space $\mathbb{M}_+(\mathbb{R}_+^\ell)$ of (nonnegative finite) Borel measures on $\mathbb{R}_+^\ell$, which is equipped with the weak topology and is, therefore, a complete separable metric space, see, Topsœ[27]. Accordingly, $\Lambda^{n,m}$ is a random element of the Skorohod space $\mathbb{D}(\mathbb{R}_+, \mathbb{M}_+(\mathbb{R}_+^\ell))$, see, e.g., Ethier and Kurtz [6] for the definition and properties. We introduce a number of other spaces. Let $\mathbb{C}(\mathbb{R}_+, \mathbb{M}_+(\mathbb{R}_+^\ell))$ denote the set of continuous $\mathbb{M}_+(\mathbb{R}_+^\ell)$–valued functions. Let $\mathbb{C}^1(\mathbb{R}_+^\ell)$ denote the set of real–valued functions on $\mathbb{R}_+^\ell$ that can be extended to functions with continuous derivatives defined on an open set containing $\mathbb{R}_+^\ell$. Let $\mathbb{C}_c^1(\mathbb{R}_+^\ell)$ denote the subset of $\mathbb{C}^1(\mathbb{R}_+^\ell)$ of functions of compact support.

Let, for $u = (u_1, \ldots, u_\ell) \in \mathbb{R}_+^\ell$, $\lfloor nu \rfloor = (\lfloor nu_1 \rfloor, \ldots, \lfloor nu_\ell \rfloor)$ and

$$\hat\beta^{n,m,k}(u) = \beta^{n,m,k}(\lfloor nu \rfloor),\ \hat\delta^{n,m,k}(u) = \delta^{n,m,k}(\lfloor nu \rfloor),\ \hat\mu^{n,m,k}(u) = \mu^{n,m,k}(\lfloor nu \rfloor),$$
$$\hat\phi^{n,m}(u) = \phi^{n,m}(\lfloor nu \rfloor),\ \hat\epsilon^{n,m}(u) = m\epsilon^{n,m}(\lfloor nu \rfloor)\,. \tag{2.5}$$

We also define, for $u \in \mathbb{R}_+^\ell$ and $\Gamma \subset \mathbb{R}_+^\ell$,

$$\hat\eta^{n,m}(u, \Gamma) = \sum_{i'} \eta^{n,m}(\lfloor nu \rfloor, i') \mathbf{1}_\Gamma\bigl(\frac{i'}{n}\bigr)\,. \tag{2.6}$$

(Note that $\hat\eta^{n,m}(u, \{i'/n\}) = \eta^{n,m}(\lfloor nu \rfloor, i')$.)

Let us assume that there exist functions $\hat\beta^k(u)$, $\hat\delta^k(u)$ and $\hat\mu^k(u)$, which belong to $\mathbb{C}^1(\mathbb{R}_+^\ell)$, such that the functions $u_k\hat\delta^k(u)$, $u_k\hat\beta^k(u)$, and $u_k\hat\mu^k(u)$ have bounded first order derivatives, continuous and bounded functions $\hat\phi(u)$ and $\hat\epsilon(u)$ and transition kernel $\hat\varphi(u, du')$, which is a finite measure on $\mathbb{R}_+^\ell$ for each u, with the total mass being uniformly bounded over $u \in \mathbb{R}_+^\ell$, such that $\int_{\mathbb{R}_+^\ell} f(u')\hat\varphi(u, du')$ is a continuous function of u and, as $n, m \to \infty$,

$$\sum_{k=1}^{\ell} \bigl(u_k|\hat\beta^{n,m,k}(u) - \hat\beta^k(u)| + u_k|\hat\delta^{n,m,k}(u) - \hat\delta^k(u)| + u_k|\hat\mu^{n,m,k}(u) - \hat\mu^k(u)|\bigr)$$
$$+ |\hat\phi^{n,m}(u) - \hat\phi(u)| + |\hat\epsilon^{n,m}(u) - \hat\epsilon(u)| + \Bigl|\int_{\mathbb{R}_+^\ell} f(u')\hat\phi^{n,m}(u)\hat\eta^{n,m}(u, du')$$
$$- \int_{\mathbb{R}_+^\ell} f(u')\hat\varphi(u, du')\Bigr| \to 0 \tag{2.7}$$

uniformly over compact sets of u, for all continuous bounded functions f of compact support. It is assumed further that the total population at time 0, which is $\sum_i \sum_{k=1}^\ell i_k X_0^{n,m}(i) = nm \int_{\mathbb{R}_+^\ell} |u| \Lambda_0^{n,m}(du)$, is finite. (Accordingly, $nm \int_{\mathbb{R}_+^\ell} |u| \Lambda_t^{n,m}(du)$ yields the total population at time t.) Let $|u| = u_1 + \ldots + u_\ell$.

Theorem 2.1. *1. Suppose that, for some $\hat\lambda_0 = (\hat\lambda_0(du)) \in \mathbb{M}_+(\mathbb{R}_+^\ell)$ such that $\hat\lambda_0(\mathbb{R}_+^\ell) > 0$ and $\int_{\mathbb{R}_+^\ell} |u|\hat\lambda_0(du) < \infty$, we have that $\Lambda_0^{n,m} \to \hat\lambda_0$ in probability in $\mathbb{M}_+(\mathbb{R}_+^\ell)$ and $\int_{\mathbb{R}_+^\ell} |u|\Lambda_0^{n,m}(du) \to \int_{\mathbb{R}_+^\ell} |u|\hat\lambda_0(du)$ in probability, as $n, m \to \infty$. In addition, suppose that*

$$\lim_{K\to\infty} \limsup_{n,m\to\infty} \mathbf{P}\bigl(\sum_i |X_0^{n,m}(i)|^2 > Km^2\bigr) = 0\,.$$

Then, the $\Lambda^{n,m}$ *converge in probability in* $\mathbb{D}(\mathbb{R}_+,\mathbb{M}_+(\mathbb{R}_+^\ell))$ *to* $(\hat\lambda_t\,,t\ge 0)\in \mathbb{C}(\mathbb{R}_+,\mathbb{M}_+(\mathbb{R}_+^\ell))$ *such that* $\hat\lambda_t(\mathbb{R}_+^\ell)>0$ *. It is uniquely specified by the requirement that, given* $f\in\mathbb{C}_c^1(\mathbb{R}_+^\ell)$ *,* $\int_{\mathbb{R}_+^\ell} f(u)\hat\lambda_t(du)$ *is differentiable and*

$$\frac{d}{dt}\int\limits_{\mathbb{R}_+^\ell} f(u)\hat\lambda_t(du)$$

$$= \int\limits_{\mathbb{R}_+^\ell}\Bigl(\sum_{k=1}^\ell\bigl(u_k\,(\hat\beta^k(u)-\hat\delta^k(u)-\hat\mu^k(u))+\frac{1}{\hat\lambda_t(\mathbb{R}_+^\ell)}\int\limits_{\mathbb{R}_+^\ell}u'_k\,\hat\mu^k(u')\,\hat\lambda_t(du')\bigr)\partial_{u_k}f(u)$$

$$+\int\limits_{\mathbb{R}_+^\ell} f(u')\,\hat\varphi(u,du')-(\hat\phi(u)+\hat\epsilon(u)\hat\lambda_t(\mathbb{R}_+^\ell))f(u)\Bigr)\hat\lambda_t(du)\,. \quad (2.8)$$

If $\hat\lambda_0$ *is absolutely continuous with respect to Lebesgue measure, then* $\hat\lambda_t$ *is absolutely continuous too.*

2. *In probability, locally uniformly in* t *,*

$$\int\limits_{\mathbb{R}_+^\ell}|u|\Lambda_t^{n,m}(du)\to\int\limits_{\mathbb{R}_+^\ell}|u|\hat\lambda_t(du)\,.$$

   If, in addition,

$$\int\limits_{\mathbb{R}_+^\ell}u\,\Lambda_0^{n,m}(du)\to\int\limits_{\mathbb{R}_+^\ell}u\,\hat\lambda_0(du)$$

   in probability, then

$$\int\limits_{\mathbb{R}_+^\ell}u\,\Lambda_t^{n,m}(du)\to\int\limits_{\mathbb{R}_+^\ell}u\,\hat\lambda_t(du)$$

   in probability locally uniformly in t *.*

3. *Suppose that, under the hypotheses of part 1,* $\hat\lambda_0$ *is absolutely continuous with respect to the Lebesgue measure, that its density* $\hat x_0=(\hat x_0(u)\,,u\in\mathbb{R}_+^\ell)$ *is a bounded and Lipschitz–continuous function, that* $\hat\varphi(u,du')=\overline{\varphi}(u,u')\,du'$ *with* $\overline{\varphi}(u,u')$ *having Sobolev derivative* $D_u\overline{\varphi}(u,u')$ *with respect to* u *for almost all* u' *such that* $\operatorname{ess\,sup}_{u\in\mathbb{R}_+^\ell}\int_{\mathbb{R}_+^\ell}(\overline{\varphi}(u,u')+|D_u\overline{\varphi}(u,u')|)\,du'<\infty$ *, and that the functions* $\hat\phi(u)$ *and* $\hat\epsilon(u)$ *are Lipschitz–continuous. Then the density* $\hat x_t=(\hat x_t(u)\,,u\in\mathbb{R}_+^\ell)$ *of* $\hat\lambda_t$ *is a bounded and Lipschitz–continuous function of* u*, is*

locally Lipschitz–continuous with respect to t, and, for almost all t and u with respect to the Lebesgue measure,

$$- \partial_t \hat{x}_t(u) = \sum_{k=1}^{\ell} \Big(\partial_{u_k} \big(\hat{x}_t(u) u_k \, (\hat{\beta}^k(u) - \hat{\delta}^k(u) - \hat{\mu}^k(u)) \big) + \frac{\int\limits_{\mathbb{R}_+^\ell} \hat{x}_t(u') u'_k \, \hat{\mu}^k(u') \, du'}{\int\limits_{\mathbb{R}_+^\ell} \hat{x}_t(u') \, du'} \, \partial_{u_k} \hat{x}_t(u) \Big) - \int\limits_{\mathbb{R}_+^\ell} \hat{x}_t(u') \overline{\varphi}(u, u') \, du'$$
$$+ \hat{x}_t(u) \hat{\phi}(u) + \hat{x}_t(u) \hat{\epsilon}(u) \int\limits_{\mathbb{R}_+^\ell} \hat{x}_t(u') \, du' . \quad (2.9)$$

Remark 2.1. If $\hat{\eta}^{n,m}(u, du') \to \hat{\eta}(u, du')$ weakly, as in the next example, then one may be able to take $\hat{\varphi}(u, du') = \hat{\phi}(u) \hat{\eta}(u, du')$.

Remark 2.2. As an example of the scaling, consider fissioning into one or two pieces that gives equal probability to every possible fission outcome:

$$\zeta_i(\{i', i - i'\}) = \frac{1}{\prod_{k=1}^{\ell} (i_k + 1)} .$$

Hence, the $\theta_i(i', p)$ assume values in $\{1, 2\}$ and

$$\eta(i, i') = \frac{2}{\prod_{k=1}^{\ell} (i_k + 1)} .$$

We define fission kernel $\hat{\eta}^{n,m}(u, du')$ by (2.6). Let $\phi^{n,m}(i) = \prod_{k=1}^{\ell} (i_k + 1) e^{-|i|/n} / n^\ell$. Consequently, $\hat{\phi}^{n,m}(u) = \prod_{k=1}^{\ell} (\lfloor n u_k \rfloor + 1) e^{-|\lfloor nu \rfloor|/n} / n^\ell$. We have that $\hat{\eta}^{n,m}(u, du') \to 2 / \big(\prod_{k=1}^{\ell} u_k \big) \, \mathbf{1}_{[0,u]}(u') \, du'$, $\hat{\phi}^{n,m}(u) \to \prod_{k=1}^{\ell} u_k e^{-|u|}$, and $\hat{\phi}^{n,m}(u) \hat{\eta}^{n,m}(u, du') \to 2 e^{-|u|} \mathbf{1}_{[0,u]}(u') \, du'$ weakly uniformly over u from bounded sets.

We now give a version of Theorem 2.1 for a stronger topology. Let

$$\check{\epsilon}^{n,m}(u) = m n^\ell \epsilon^{n,m}(\lfloor nu \rfloor) \, , \check{\eta}^{n,m}(u, u') = n^\ell \eta^{n,m}(\lfloor nu \rfloor, \lfloor nu' \rfloor) . \quad (2.10)$$

The functions $m n^\ell \epsilon^{n,m}(i)$ and $\phi^{n,m}(i) n^\ell \eta^{n,m}(i, i')$ are assumed to be bounded in n, m, i, and i'. Let us assume that there exist bounded Lipschitz–continuous functions $\hat{\beta}^k(u)$, $\hat{\delta}^k(u)$, and $\hat{\mu}^k(u)$, bounded continuous functions $\hat{\phi}(u)$ and $\check{\epsilon}(u)$, and function

$\check\varphi(u,u')$ such that the functions $u_k\hat\beta^k(u)$, $u_k\hat\delta^k(u)$, and $u_k\hat\mu^k(u)$ are bounded and Lipschitz–continuous, the function $\int_{\mathbb{R}_+^\ell} f(u')\check\varphi(u,u')\,du'$ is continuous with respect to u for any continuous function $f(u')$ of compact support, $\sup_{u\in\mathbb{R}_+^\ell}\int_{\mathbb{R}_+^\ell}\check\varphi(u,u')\,du'<\infty$, and for all bounded Borel measurable sets $\Theta\subset\mathbb{R}_+^\ell$ and continuous functions of compact support $f(u')$,

$$\int\limits_\Theta\Bigl(\sum_{k=1}^\ell\bigl(u_k|\hat\beta^{n,m,k}(u)-\hat\beta^k(u)|+u_k|\hat\delta^{n,m,k}(u)-\hat\delta^k(u)|+u_k|\hat\mu^{n,m,k}(u)-\hat\mu^k(u)|\bigr)$$
$$+|\hat\phi^{n,m}(u)-\hat\phi(u)|+|\check\epsilon^{n,m}(u)-\check\epsilon(u)|+|\int\limits_{\mathbb{R}_+^\ell}f(u')\hat\phi^{n,m}(u)\check\eta^{n,m}(u,u')\,du'$$
$$-\int\limits_{\mathbb{R}_+^\ell}f(u')\check\varphi(u,u')\,du'|\,du\Bigr)\to0\,,$$

as $n,m\to\infty$.

Let processes $\hat X^{n,m}=(\hat X^{n,m}_t\,,t\ge0)$ be defined by

$$\hat X^{n,m}_t=(\hat X^{n,m}_t(u),\,u\in\mathbb{R}_+^\ell)\text{ and }\hat X^{n,m}_t(u)=\frac{1}{m}\,X^{n,m}_t(\lfloor nu_1\rfloor,\ldots,\lfloor nu_\ell\rfloor)\,.$$

We assume that $\hat X^{n,m}_0\in\mathbb{L}^2(\mathbb{R}_+^\ell)\cap\mathbb{L}^1(\mathbb{R}_+^\ell)$. Both $\mathbb{L}^2(\mathbb{R}_+^\ell)$ and $\mathbb{L}^1(\mathbb{R}_+^\ell)$ are endowed with weak topologies which will be emphasised by denoting those spaces by $\mathbb{L}^2_w(\mathbb{R}_+^\ell)$ and $\mathbb{L}^1_w(\mathbb{R}_+^\ell)$, respectively. Specifically, $\mathbb{L}^2_w(\mathbb{R}_+^\ell)$ is endowed with the $\sigma\bigl(\mathbb{L}^2,\mathbb{L}^2\bigr)$–topology and $\mathbb{L}^1_w(\mathbb{R}_+^\ell)$ is endowed with the $\sigma\bigl(\mathbb{L}^1,\mathbb{L}^\infty\bigr)$–topology. Both spaces are completely regular topological spaces as topological groups. The definition and properties of the Skorohod spaces $\mathbb{D}(\mathbb{R}_+,\mathbb{L}^2_w(\mathbb{R}_+^\ell))$ and $\mathbb{D}(\mathbb{R}_+,\mathbb{L}^1_w(\mathbb{R}_+^\ell))$ follow from the analysis in Jakubowski [13].

Theorem 2.2. *1. If, for some $\check x_0=(\check x_0(u))\in\mathbb{L}^2(\mathbb{R}_+^\ell)\cap\mathbb{L}^1(\mathbb{R}_+^\ell)$ with $\check x_0(u)\ge0$, $\int_{\mathbb{R}_+^\ell}\check x_0(u)\,du>0$, and $\int_{\mathbb{R}_+^\ell}|u|\check x_0(u)\,du<\infty$, we have that $\hat X^{n,m}_0\to\check x_0$ in probability in $\mathbb{L}^1_w(\mathbb{R}_+^\ell)$ and $\int_{\mathbb{R}_+^\ell}|u|\hat X^{n,m}_0(u)\,du\to\int_{\mathbb{R}_+^\ell}|u|\check x_0(u)\,du$ in probability, as $n,m\to\infty$, and*

$$\lim_{K\to\infty}\limsup_{n,m\to\infty}\mathbf{P}\bigl(\sum_i|X^{n,m}_0(i)|^2>Km^2n^\ell\bigr)=0\,,$$

then the $\hat X^{n,m}$ converge in probability in $\mathbb{D}(\mathbb{R}_+,\mathbb{L}^1_w(\mathbb{R}_+^\ell))$ and in $\mathbb{D}(\mathbb{R}_+,\mathbb{L}^2_w(\mathbb{R}_+^\ell))$ to function $(\check x_t\,,t\ge0)$, where $\check x_t=(\check x_t(u),\,u\in\mathbb{R}_+^\ell)$, such that $\check x_t(u)\ge0$, $\int_{\mathbb{R}_+^\ell}\check x_t(u)\,du>0$, $\check x_t\in\mathbb{L}^1(\mathbb{R}_+^\ell)\cap\mathbb{L}^2(\mathbb{R}_+^\ell)$, and, for all t and almost all u, *the measure* $\hat{\lambda}_t(du) = \check{x}_t(u)\,du$ *satisfies* (2.8) *with* $\hat{\eta}(u)$ *and* $\hat{\epsilon}(u)$ *replaced with* $\check{\eta}(u)$ *and* $\check{\epsilon}(u)$ *, respectively.*

2. *In probability, locally uniformly in* t *,*

$$\int_{\mathbb{R}_+^\ell} |u|\hat{X}_t^{n,m}(u)\,du \to \int_{\mathbb{R}_+^\ell} |u|\check{x}_t(u)\,du\,.$$

If, in addition,

$$\int_{\mathbb{R}_+^\ell} u\,\hat{X}_0^{n,m}(u)\,du \to \int_{\mathbb{R}_+^\ell} u\,\check{x}_0(u)\,du$$

in probability, then

$$\int_{\mathbb{R}_+^\ell} u\,\hat{X}_t^{n,m}(u)\,du \to \int_{\mathbb{R}_+^\ell} u\,\check{x}_t(u)\,du$$

in probability locally uniformly in t *.*

3. *If, under the hypotheses of part 1,* $\check{x}_0$ *is a bounded and Lipschitz–continuous function,* $\check{\varphi}(u,u')$ *has Sobolev derivative* $D_u\check{\varphi}(u,u')$ *with respect to* u *for almost all* u' *such that* $\operatorname{ess\,sup}_{u\in\mathbb{R}_+^\ell}\int_{\mathbb{R}_+^\ell}(\check{\varphi}(u,u') + |D_u\check{\varphi}(u,u')|)\,du' < \infty$ *, and the functions* $\hat{\phi}(u)$ *and* $\check{\epsilon}(u)$ *are Lipschitz–continuous, then* $\check{x}_t(u)$ *is a bounded and Lipschitz–continuous function with respect to* u *, is locally Lipschitz–continuous with respect to* t *, and* (2.9) *is satisfied for almost all* t *and* u *.*

3 Proof of Theorem 2.1

The proof of Theorem 2.1 proceeds by establishing compactness of $\Lambda^{n,m}$ and ascertaining the limit point. Techniques of stochastic calculus are used extensively. Throughout the section, the hypotheses of part 1 of Theorem 2.1 are assumed to hold. We begin with a lemma on the properties of fission. Let b denote an upper bound on the number of offspring in a fission.

Lemma 3.1. *We have that*

$$\hat{\eta}^{n,m}(u,\mathbb{R}_+^\ell) \le b \tag{3.1}$$

and

$$\int_{\mathbb{R}_+^\ell} u'\hat{\eta}^{n,m}(u,du') = \frac{\lfloor nu\rfloor}{n}\,. \tag{3.2}$$

As a result,

$$\hat{\varphi}(u,\mathbb{R}_+^\ell) \le b\hat{\phi}(u) \tag{3.3}$$

and

$$\int_{\mathbb{R}_+^\ell} u'\hat{\varphi}(u,du') = u\hat{\phi}(u)\,. \tag{3.4}$$

Proof. By the analogue of (2.1),

$$\sum_{i'} \eta^{n,m}(i,i') = \sum_{i'} \mathbf{E}\theta_i^{n,m}(i') = \mathbf{E}\sum_{i'} \theta_i^{n,m}(i')\,.$$

The latter sum is the total number of pieces, so, it does not exceed b. Similarly,

$$\sum_{i'} i'\eta^{n,m}(i,i') = \mathbf{E}\sum_{i'} i'\theta_i^{n,m}(i')\,,$$

the latter sum being equal to i. Representations (3.1) and (3.2) now follow from (2.6). Since $\hat{\eta}^{n,m}(u,du') = 0$ when $|u'| > |u|$, (3.3) and (3.4) follow from (2.7), (3.1) and (3.2). ☐

Remark 3.1. Similarly, $\hat{\eta}^{n,m}(u,\mathbb{R}_+^\ell) = b$ and $\hat{\varphi}(u,\mathbb{R}_+^\ell) = b\hat{\phi}(u)$, provided every fission produces exactly b offspring. If, in addition, $\hat{\eta}^{n,m}(u,du') \to \hat{\eta}(u,du')$ weakly, then these relations carry over to $\hat{\eta}(u,du')$.

Let $\mathcal{F}_t^{n,m}$ represent the complete σ–algebra that is generated by the random variables $X_0^{n,m}(i)$, $L_s^{B,n,m,k}(i,p,r)$, $L_s^{D,n,m,k}(i,p,r)$, $L_s^{\overline{M},n,m,k}(i,p,r)$, $L_s^{\overline{F},n,m}(i,p)$, $L_s^{E,n,m}(i,p)$, $B_s^{n,m,k}(i)$, $D_s^{n,m,k}(i)$, $\overline{M}_s^{n,m,k}(i)$, $M_s^{n,m,k}(i)$, and $F_s^{n,m}(i,i')$, where $i \in \mathbb{Z}_+^\ell \setminus \{0\}$, $p \in \mathbb{N}$, $r \in \mathbb{N}$, $k \in \{1,\ldots,\ell\}$, and $0 \le s \le t$, and let $\mathbf{F}^{n,m} = (\mathcal{F}_t^{n,m}\,, t \ge 0)$ represent the associated filtration. Let us adopt the convention that $0/0 = 0$, that the analogues of the processes on the lefthand side of (2.2) are equal to zero when

$i = 0$ and define

$$
\begin{aligned}
N_t^{B,n,m,k}(i) &= B_t^{n,m,k}(i) - \int_0^t X_s^{n,m}(i) i_k \beta^{n,m,k}(i)\, ds\,,\\
N_t^{D,n,m,k}(i) &= D_t^{n,m,k}(i) - \int_0^t X_s^{n,m}(i) i_k \delta^{n,m,k}(i)\, ds\,,\\
N_t^{F,n,m}(i',i) &= F_t^{n,m}(i',i) - \int_0^t X_s^{n,m}(i')\phi^{n,m}(i')\eta^{n,m}(i',i)\, ds\,,\\
N_t^{\overline{F},n,m}(i) &= \overline{F}_t^{n,m}(i) - \int_0^t X_s^{n,m}(i)\phi^{n,m}(i)\, ds\,, \qquad (3.5)\\
N_t^{E,n,m}(i) &= E_t^{n,m}(i) - \int_0^t X_s^{n,m}(i) X_s^{n,m,*}\epsilon^{n,m}(i)\, ds\,,\\
N_t^{M,n,m,k}(i) &= M_t^{n,m,k}(i) - \int_0^t \sum_{i'\neq i} X_s^{n,m}(i') i'_k \mu^{n,m,k}(i')\frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\, ds\,,\\
N_t^{\overline{M},n,m,k}(i) &= \overline{M}_t^{n,m,k}(i) - \int_0^t X_s^{n,m}(i) i_k \mu^{n,m,k}(i)\big(1-\frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\big)\, ds\,.
\end{aligned}
$$

We note that the righthand sides are equal to zero after the time when $X_t^{n,m,*}$ hits zero. Let

$$\alpha_i^{n,m}(i') = \mathbf{E}\theta_i^{n,m}(i',1)^2\,.$$

We note that

$$\alpha_i^{n,m}(i') \le b\eta^{n,m}(i,i')\,. \qquad (3.6)$$

By the analogues of (2.1) and (2.2), and by Lemma A.1 in the appendix, the processes on the righthand sides of (3.5) are locally square integrable martingales, whose predictable quadratic variation processes are as follows, see, e.g., Liptser and

Shiryayev [15] for the corresponding definitions,

$$\langle N^{B,n,m,k}(i)\rangle_t = \int_0^t X_s^{n,m}(i) i_k \beta^{n,m,k}(i)\,ds\,,\ \langle N^{D,n,m,k}(i)\rangle_t = \int_0^t X_s^{n,m}(i) i_k \delta^{n,m,k}(i)\,ds\,,$$
$$\langle N^{E,n,m}(i)\rangle_t = \int_0^t X_s^{n,m}(i) X_s^{n,m,*}\epsilon^{n,m}(i)\,ds\,,\ \langle N^{\overline{F},n,m}(i)\rangle_t = \int_0^t X_s^{n,m}(i)\phi^{n,m}(i)\,ds\,,$$
$$\langle N^{F,n,m}(i',i)\rangle_t = \int_0^t X_s^{n,m}(i')\phi^{n,m}(i')\alpha_{i'}^{n,m}(i)\,ds\,,$$
$$\langle N^{M,n,m,k}(i)\rangle_t = \int_0^t \sum_{i'\neq i} X_s^{n,m}(i') i'_k \mu^{n,m,k}(i') \frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\,ds\,,$$
$$\langle N^{\overline{M},n,m,k}(i)\rangle_t = \int_0^t X_s^{n,m}(i) i_k \mu^{n,m,k}(i)\big(1 - \frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\big)\,ds\,. \tag{3.7}$$

The nonzero predictable covariance processes are

$$\langle N^{M,n,m,k}(i), N^{\overline{M},n,m,k}(i')\rangle_t = (1 - \mathbf{1}_i(i'))\int_0^t X_s^{n,m}(i') i'_k \mu^{n,m,k}(i') \frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\,ds\,,$$
$$\langle N^{F,n,m}(i,i'), N^{\overline{F},n,m}(i)\rangle_t = \int_0^t X_s^{n,m}(i)\phi^{n,m}(i)\eta^{n,m}(i,i')\,ds\,, \tag{3.8}$$
$$\langle N^{F,n,m}(i,i'), N^{F,n,m}(i,j')\rangle_t = \int_0^t X_s^{n,m}(i)\phi^{n,m}(i)\mathbf{E}\theta_i^{n,m}(i',1)\theta_i^{n,m}(j',1)\,ds\,.$$

Let

$$R_t^{n,m} = \frac{1}{m}\,X_t^{n,m,*} = \Lambda_t^{n,m}(\mathbb{R}_+^\ell)\,, \tag{3.9}$$

so, $R_0^{n,m} \to R_0 = \hat{\lambda}_0(\mathbb{R}_+^\ell) > 0$ in probability, as $n,m \to \infty$. The processes $(R_t^{n,m}, t \geq 0)$ are random elements of $\mathbb{D}(\mathbb{R}_+,\mathbb{R})$. Let us recall that a sequence of stochastic processes with trajectories in a Skorohod space is said to be C–tight if it is tight for convergence in distribution in the Skorohod space and the limit points are laws of continuous path processes, see, e.g., Jacod and Shiryaev [12].

Lemma 3.2. *The sequence of processes* $(R_t^{n,m}\,, t\ge 0)$ *is* C*–tight and, given* $t>0$*, there exists* $\rho>0$ *such that* $\mathbf{P}(\inf_{s\le t} R_s^{n,m}>\rho)\to 1$*, as* $n,m\to\infty$*.*

Proof. By the analogue of (2.3),

$$\Delta X_t^{n,m,*}=\sum_i\sum_{i'}\Delta F_t^{n,m}(i',i)-\sum_{k=1}^{\ell}\Delta D_t^{n,m,k}(e_k)-\sum_i \Delta\overline{F}_t^{n,m}(i)-\sum_i\Delta E_t^{n,m}(i)\,. \tag{3.10}$$

Since $\sum_i \Delta F_t^{n,m}(i',i)\le b\Delta\overline{F}_t^{n,m}(i')$,

$$\Delta X_t^{n,m,*}\le b\sum_i \Delta\overline{F}_t^{n,m}(i)-\sum_i\Delta E_t^{n,m}(i)\,.$$

Hence,

$$\Delta R_t^{n,m}\le\Delta\overline{R}_t^{n,m}\,, \tag{3.11}$$

where

$$\overline{R}_t^{n,m}=R_0^{n,m}+\frac{b}{m}\sum_i\overline{F}_t^{n,m}(i)-\frac{1}{m}\sum_i E_t^{n,m}(i)\,. \tag{3.12}$$

Let $N_t^{\overline{R},n,m}(i)=(bN_t^{\overline{F},n,m}(i)-N_t^{E,n,m}(i))/m$ and $N_t^{\overline{R},n,m}=\sum_i N_t^{\overline{R},n,m}(i)$ so that by (3.5),

$$\overline{R}_t^{n,m}=R_0^{n,m}+\int_0^t\sum_i\frac{1}{m}X_s^{n,m}(i)\big(b\phi^{n,m}(i)-X_s^{n,m,*}\epsilon^{n,m}(i)\big)\,ds+N_t^{\overline{R},n,m}\,. \tag{3.13}$$

Since the processes $N^{\overline{F},n,m}(i)=(N_t^{\overline{F},n,m}(i)\,,t\ge 0)$ and $N^{E,n,m}(i)=(N_t^{E,n,m}(i)\,,t\ge 0)$ are locally square integrable martingales with disjoint jumps, it follows by (3.7) that the process $N^{\overline{R},n,m}=(N_t^{\overline{R},n,m}\,,t\ge 0)$ is a locally square integrable martingale with the predictable quadratic variation process

$$\langle N^{\overline{R},n,m}\rangle_t=\int_0^t\sum_i\frac{1}{m^2}X_s^{n,m}(i)(b^2\phi^{n,m}(i)+X_s^{n,m,*}\epsilon^{n,m}(i))\,ds\,. \tag{3.14}$$

By (3.5), (3.12), (3.13), and the Itô formula for semimartingales, see, e.g., Theorem 1 on p.118 in Liptser and Shiryayev [15], on taking into account that the processes

$(\overline{F}_t^{n,m}\,, t \geq 0)$ and $(E_t^{n,m}\,, t \geq 0)$ have unit jumps,

$$(\overline{R}_t^{n,m})^2 = (R_0^{n,m})^2 + \int_0^t 2\overline{R}_{s-}^{n,m}\, d\overline{R}_s^{n,m} + \frac{1}{m^2}\sum_{s\leq t}(b^2\sum_i(\Delta\overline{F}_s^{n,m}(i))^2 + \sum_i(\Delta E_s^{n,m}(i))^2)$$
$$= (R_0^{n,m})^2 + \int_0^t 2\overline{R}_s^{n,m}\sum_i \frac{1}{m}\, X_s^{n,m}(i)\big(b\phi^{n,m}(i) - R_s^{n,m} m\epsilon^{n,m}(i)\big)\, ds$$
$$+ \int_0^t 2\overline{R}_{s-}^{n,m}\, dN_s^{R,n,m} + \frac{1}{m^2}\sum_i(b^2\overline{F}_t^{n,m}(i) + E_t^{n,m}(i))$$
$$\leq (R_0^{n,m})^2 + \int_0^t 2\overline{R}_s^{n,m}\sum_i \frac{b}{m}\, X_s^{n,m}(i)\phi^{n,m}(i)\, ds$$
$$+ \int_0^t 2\overline{R}_{s-}^{n,m}\, dN_s^{R,n,m} + \frac{1}{m^2}\sum_i(b^2\overline{F}_t^{n,m}(i) + E_t^{n,m}(i))\,.$$

Hence, on recalling (3.5), (3.9), and (3.11), by $m\epsilon^{n,m}(i)$ and $\phi^{n,m}(i)$ being bounded, there exists $K_0 > 0$ such that, for all $t > 0$,

$$(\overline{R}_t^{n,m})^2 \leq (R_0^{n,m})^2 + \frac{K_0}{m}\, t + K_0\int_0^t (\overline{R}_s^{n,m})^2\, ds + \int_0^t 2\overline{R}_{s-}^{n,m}\, dN_s^{\overline{R},n,m}$$
$$+ \frac{1}{m^2}\sum_i(N_t^{\overline{F},n,m}(i) + N_t^{E,n,m}(i))\,. \quad (3.15)$$

Let, for $K_1 > 0$,

$$\tau_{K_1} = \inf\{s \geq 0 :\, \overline{R}_s^{n,m} > K_1\}\,.$$

Then, by $R_0^{n,m}$ being $\mathcal{F}_0^{n,m}$–measurable and by $N^{\overline{R},n,m}$, $N^{\overline{F},n,m}(i)$ and $N^{E,n,m}(i)$ being locally square integrable martingales, whose predictable quadratic variation processes are bounded for $t \leq \tau_{K_1}$ by (3.7) and (3.14), so that the local martingale on the righthand side of (3.15) stopped at τ_{K_1} is a martingale, we have that

$$\mathbf{E}\big(\int_0^{t\wedge\tau_{K_1}} 2\overline{R}_{s-}^{n,m}\, dN_s^{\overline{R},n,m} + \frac{1}{m^2}\sum_i(N_{t\wedge\tau_{K_1}}^{\overline{F},n,m}(i) + N_{t\wedge\tau_{K_1}}^{E,n,m}(i))\big) = 0\,.$$

By (3.15) and Gronwall's inequality, for $K_2 > 0$,

$$\mathbf{E}(\overline{R}^{n,m}_{t\wedge\tau_{K_1}})^2\mathbf{1}_{[0,K_2]}(R^{n,m}_0) \le (K_2^2 + \frac{K_0}{m}\,t)e^{K_0t}\,.$$

Letting $K_1 \to \infty$ implies, by Fatou's lemma, that

$$\mathbf{E}(\overline{R}^{n,m}_t)^2\mathbf{1}_{[0,K_2]}(R^{n,m}_0) \le (K_2^2 + \frac{K_0}{m}\,t)e^{K_0t}\,. \tag{3.16}$$

By (3.14), (3.16), the $\phi^{n,m}(i)$ and $m\epsilon^{n,m}(i)$ being bounded, for $\gamma > 0$,

$$\lim_{n,m\to\infty} \mathbf{P}(\langle N^{\overline{R},n,m}\rangle_t > \gamma) = 0\,, \tag{3.17}$$

so, by the Lenglart–Rebolledo inequality, see, e.g., Theorem 3 on p.66 in Liptser and Shiryayev [15], in probability,

$$\lim_{m,n\to\infty} \sup_{s\le t} |N^{\overline{R},n,m}_s| = 0\,.$$

By (3.9), (3.13), and Gronwall's inequality, for some $K' > 0$,

$$\overline{R}^{n,m}_t \le (R^{n,m}_0 + \sup_{s\le t}|N^{\overline{R},n,m}_s|)e^{K't}\,.$$

It follows that

$$\lim_{K\to\infty}\limsup_{n,m\to\infty} \mathbf{P}(\sup_{s\le t} \overline{R}^{n,m}_s > K) = 0\,. \tag{3.18}$$

By (3.5) and (3.10),

$$\begin{aligned} R^{n,m}_t = R^{n,m}_0 + \int_0^t \sum_i \frac{X^{n,m}_s(i)}{m}\big(\sum_{i'} \eta^{n,m}(i,i')\phi^{n,m}(i) - \phi^{n,m}(i) - X^{n,m,*}_s\epsilon^{n,m}(i)\big)\,ds \\ - \sum_{k=1}^{\ell}\int_0^t \frac{X^{n,m}_s(e_k)}{m}\,\delta^{n,m,k}(e_k)\,ds - \sum_{k=1}^{\ell}\frac{N^{D,n,m,k}_t(e_k)}{m} + \frac{N^{R,n,m}_t}{m}\,, \end{aligned} \tag{3.19}$$

where $(N^{R,n,m}_t\,, t \ge 0)$ is a locally square integrable martingale with the predictable quadratic variation process

$$\langle N^{R,n,m}\rangle_t = \sum_i \int_0^t X^{n,m}_s(i)\big(\mathbf{E}\big(\sum_{i'}\theta^{n,m}_i(i',1) - 1\big)^2\phi^{n,m}(i) + X^{n,m,*}_s\epsilon^{n,m}(i)\big)\,ds\,.$$

Since $\sum_{i'} \theta_i^{n,m}(i',1) \le b$, by (3.14) and (3.17), we have that $\langle N^{R,n,m}\rangle_t/m^2 \to$ in probability, as $n,m\to\infty$, so, in probability,

$$\lim_{n,m\to\infty} \frac{1}{m} \sup_{s\le t} |N^{R,n,m}| = 0\,. \tag{3.20}$$

Hence, for arbitrary $K_3 > 0$,

$$\begin{aligned} R_t^{n,m} - \frac{N_t^{R,n,m}}{m} + \sum_{k=1}^{\ell} \frac{N_t^{D,n,m,k}(e_k)}{m} &= R_0^{n,m} - K_3 \int_0^t (R_s^{n,m} - \frac{N_s^{R,n,m}}{m} \\ &+ \sum_{k=1}^{\ell} \frac{N_s^{D,n,m,k}(e_k)}{m})\, ds + \int_0^t \big(K_3 R_s^{n,m} - \sum_{k=1}^{\ell} \frac{X_s^{n,m}(e_k)}{m}\, \delta^{n,m,k}(e_k) \\ &+ \sum_i \frac{X_s^{n,m}(i)}{m} (\sum_{i'} \eta^{n,m}(i,i')\phi^{n,m}(i) - \phi^{n,m}(i) - X_s^{n,m,*}\epsilon^{n,m}(i)) - K_3 \frac{N_s^{R,n,m}}{m} \\ &+ K_3 \sum_{k=1}^{\ell} \frac{N_s^{D,n,m,k}(e_k)}{m}\big)\, ds\,. \end{aligned} \tag{3.21}$$

By (3.21) and the fact that $\sum_{i'} \eta^{n,m}(i,i') \le b$ according to Lemma 3.1, if $\sup_{s\le t} R_s^{n,m} \le K$ then, solving for the lefthand side of (3.21) and picking K_3 great enough so that

$$\begin{aligned} K_3 R_s^{n,m} - \sum_{k=1}^{\ell} \frac{X_s^{n,m}(e_k)}{m}\, \delta^{n,m,k}(e_k) + \sum_i \frac{X_s^{n,m}(i)}{m} (\sum_{i'} \eta^{n,m}(i,i')\phi^{n,m}(i) \\ - \phi^{n,m}(i) - X_s^{n,m,*}\epsilon^{n,m}(i)) \ge 0 \end{aligned}$$

when $0 \le s \le t$,

$$\begin{aligned} R_t^{n,m} - \frac{N_t^{R,n,m}}{m} + \sum_{k=1}^{\ell} \frac{N_t^{D,n,m,k}(e_k)}{m} &= e^{-K_3 t} R_0^{n,m} + e^{-K_3 t} \int_0^t e^{K_3 s} \big(K_3 R_s^{n,m} \\ &- \sum_{k=1}^{\ell} \frac{X_s^{n,m}(e_k)}{m}\, \delta^{n,m,k}(e_k) + \sum_i \frac{X_s^{n,m}(i)}{m} (\sum_{i'} \eta^{n,m}(i,i')\phi^{n,m}(i) \\ &- \phi^{n,m}(i) - X_s^{n,m,*}\epsilon^{n,m}(i)) - K_3 \frac{N_s^{R,n,m}}{m} + K_3 \sum_{k=1}^{\ell} \frac{N_s^{D,n,m,k}(e_k)}{m}\big)\, ds \\ &\ge e^{-K_3 t} R_0^{n,m} - \sup_{s\le t} \frac{|N_s^{R,n,m}|}{m} - \sum_{k=1}^{\ell} \sup_{s\le t} \frac{|N_s^{D,n,m,k}(e_k)|}{m}\,. \end{aligned} \tag{3.22}$$

By (3.7), (3.9), (3.18), the $\delta^{n,m,k}(e_k)$ being bounded and the Lenglart–Rebolledo inequality, in probability, for $k = 1, \ldots, \ell$,

$$\lim_{m,n\to\infty} \sup_{s\le t} \frac{1}{m} |N_t^{D,n,m,k}(e_k)| = 0\,,$$

which implies, by (3.20), (3.22), and $\hat{\lambda}_0(\mathbb{R}_+^\ell)$ being positive, that the $R_t^{n,m}$ are locally uniformly asymptotically separated away from zero in probability. In addition, by (3.19) and the $m\epsilon^{n,m}(i)$ being bounded,

$$\lim_{\sigma\to 0} \limsup_{n,m\to\infty} \mathbf{P}\big(\sup_{s,s'\in[0,t]:\,|s-s'|\le\sigma} |R_s^{n,m} - R_{s'}^{n,m}| > \gamma \big) = 0\,,$$

so, $R^{n,m}$ is C–tight.

□

Let us introduce $\hat{\beta}^{n,m}(u) = (\hat{\beta}^{n,m,1}(u), \ldots, \hat{\beta}^{n,m,\ell}(u))$ and $\hat{\delta}^{n,m}(u) = (\hat{\delta}^{n,m,1}(u), \ldots, \hat{\delta}^{n,m,\ell}(u))$, and let $\cdot$ denote the inner product in $\mathbb{R}^\ell$.

Lemma 3.3. *The sequence of processes* $\big(\int_{\mathbb{R}_+^\ell} |u| \Lambda_t^{n,m}(du)\,, t \ge 0\big)$ *is* C*–tight.*

Proof. By the analogue of (2.3),

$$\sum_i i_k \Delta X_t^{n,m}(i) = \sum_i \Delta B_t^{n,m,k}(i) - \sum_i \Delta D_t^{n,m,k}(i) - \sum_i i_k \Delta E_t^{n,m}(i)\,,$$

so,

$$\sum_i \sum_{k=1}^{\ell} i_k X_t^{n,m}(i) = \sum_i \sum_{k=1}^{\ell} i_k X_0^{n,m}(i) + \sum_i \sum_{k=1}^{\ell} B_t^{n,m,k}(i) - \sum_i \sum_{k=1}^{\ell} D_t^{n,m,k}(i) - \sum_i |i| E_t^{n,m}(i)\,. \quad (3.23)$$

Therefore, on recalling (2.4) and (3.5),

$$\int_{\mathbb{R}_+^\ell} |u| \Lambda_t^{n,m}(du) \le \int_{\mathbb{R}_+^\ell} |u| \Lambda_0^{n,m}(du) + \int_0^t \int_{\mathbb{R}_+^\ell} u \cdot \hat{\beta}^{n,m}(u) \Lambda_s^{n,m}(du)\, ds + \frac{1}{mn} N_t^{B,n,m}\,, \quad (3.24)$$

where

$$N_t^{B,n,m} = \sum_i \sum_{k=1}^{\ell} N_t^{B,n,m,k}(i)\,.$$

By (3.7), the process $N^{B,n,m} = (N_t^{B,n,m}\,, t \geq 0)$ is a locally square integrable martingale with the predictable quadratic variation process

$$\langle N^{B,n,m}\rangle_t = \int_0^t \sum_i \sum_{k=1}^{\ell} i_k \beta^{n,m,k}(i) X_s^{n,m}(i)\, ds = nm \int_0^t \int_{\mathbb{R}_+^\ell} u\cdot\hat{\beta}^{n,m}(u)\Lambda_s^{n,m}(du)\, ds\,. \tag{3.25}$$

Let $\Theta_K^{n,m}$ represent the event that $\int_{\mathbb{R}_+^\ell} |u|\Lambda_0^{n,m}(du) \leq K$, where $K > 0$, and let, for $K_1 > 0$,

$$\tau_{K_1} = \inf\{t \geq 0 : \int_0^t \int_{\mathbb{R}_+^\ell} u \cdot \hat{\beta}^{n,m}(u)\Lambda_s^{n,m}(du)\, ds > K_1\}\,.$$

The process $(N_{t\wedge\tau_{K_1}}^{B,n,m}\,, t \geq 0)$ being a martingale implies by (3.24) that $\mathbf{E}\mathbf{1}_{\Theta^{n,m}(K)} \int_{\mathbb{R}_+^\ell} |u|\Lambda_{t\wedge\tau_{K_1}}^{n,m}(du)$ is finite, so by the $\hat{\beta}^{n,m,k}(u)$ being bounded, provided K is great enough,

$$\mathbf{E}\mathbf{1}_{\Theta^{n,m}(K)} \int_{\mathbb{R}_+^\ell} |u|\Lambda_{t\wedge\tau_{K_1}}^{n,m}(du) \leq K + K\int_0^t \mathbf{E}\mathbf{1}_{\Theta^{n,m}(K)} \int_{\mathbb{R}_+^\ell} |u|\Lambda_{s\wedge\tau_{K_1}}^{n,m}(du)\, ds\,.$$

By Gronwall's inequality and Fatou's lemma,

$$\mathbf{E}\mathbf{1}_{\Theta^{n,m}(K)} \int_{\mathbb{R}_+^\ell} |u|\Lambda_t^{n,m}(du) \leq Ke^{Kt}\,.$$

Since, for $K_2 > 0$,

$$\mathbf{P}(\int_{\mathbb{R}_+^\ell} |u|\Lambda_t^{n,m}(du) > K_2) \leq 1 - \mathbf{P}(\Theta_K^{n,m}) + \frac{Ke^{Kt}}{K_2}\,,$$

we have that

$$\lim_{K_2\to\infty} \limsup_{n,m\to\infty} \mathbf{P}(\int_{\mathbb{R}_+^\ell} |u|\Lambda_t^{n,m}(du) > K_2) \leq \limsup_{n,m\to\infty}(1 - \mathbf{P}(\Theta_K^{n,m}))\,,$$

which implies that the lefthand side equals zero by K being arbitrary and by the fact that $\int_{\mathbb{R}_+^\ell}|u|\Lambda_0^{n,m}(du) \to \int_{\mathbb{R}_+^\ell}|u|\hat{\lambda}_0(du)$ in probability. By (3.25), for $\gamma > 0$,

$$\lim_{n,m\to\infty}\mathbf{P}\big(\frac{1}{m^2n^2}\langle N^{B,n,m}\rangle_t > \gamma\big) = 0\,,$$

so, by the Lenglart–Rebolledo inequality,

$$\lim_{n,m\to\infty}\mathbf{P}\big(\frac{1}{mn}\sup_{s\le t}|N_s^{B,n,m}| > \gamma\big) = 0\,. \tag{3.26}$$

By (3.24) and Gronwall's inequality, for K great enough,

$$\sup_{s\le t}\int\limits_{\mathbb{R}_+^\ell}|u|\Lambda_s^{n,m}(du) \le \big(\int\limits_{\mathbb{R}_+^\ell}|u|\Lambda_0^{n,m}(du) + \sup_{s\le t}\frac{1}{mn}\,|N_t^{B,n,m}|\big)e^{Kt}\,.$$

By (3.26),

$$\lim_{K_2\to\infty}\limsup_{n,m\to\infty}\mathbf{P}\big(\sup_{s\le t}\int\limits_{\mathbb{R}_+^\ell}|u|\Lambda_s^{n,m}(du) > K_2\big) = 0\,. \tag{3.27}$$

By (3.5) and (3.23), in analogy with (3.24), for $s \le t$,

$$\begin{aligned}
&|\int\limits_{\mathbb{R}_+^\ell}|u|\Lambda_t^{n,m}(du) - \int\limits_{\mathbb{R}_+^\ell}|u|\Lambda_s^{n,m}(du)|\\
&\qquad\le \int\limits_s^t\int\limits_{\mathbb{R}_+^\ell}\big(u\cdot(\hat{\beta}^{n,m}(u) + \hat{\delta}^{n,m}(u)) + |u|\hat{\epsilon}^{n,m}(u)R_{\tilde{s}}^{n,m}\big)\Lambda_{\tilde{s}}^{n,m}(du)\,d\tilde{s}\\
&\qquad+ \frac{1}{mn}\,|(N_t^{B,n,m} - N_s^{B,n,m}) + (N_t^{D,n,m} - N_s^{D,n,m}) + (N_t^{E,n,m} - N_s^{E,n,m})|\\
&\qquad\le (t-s)\sup_{\tilde{s}\in[s,t]}\int\limits_{\mathbb{R}_+^\ell}\big(u\cdot(\hat{\beta}^{n,m}(u) + \hat{\delta}^{n,m}(u)) + |u|\hat{\epsilon}^{n,m}(u)R_{\tilde{s}}^{n,m}\big)\Lambda_{\tilde{s}}^{n,m}(du)\\
&\qquad+ \frac{1}{mn}\,|(N_t^{B,n,m} - N_s^{B,n,m}) + (N_t^{D,n,m} - N_s^{D,n,m}) + (N_t^{E,n,m} - N_s^{E,n,m})|\,,
\end{aligned}$$

where

$$N_{\tilde{t}}^{D,n,m} = \sum_i\sum_{k=1}^{\ell}N_{\tilde{t}}^{D,n,m,k}(i)\,,\quad N_{\tilde{t}}^{E,n,m} = \sum_i N_{\tilde{t}}^{E,n,m}(i)\,.$$

Similarly to (3.26), on recalling (3.7),

$$\lim_{n,m\to\infty} \mathbf{P}\big(\sup_{t\le L} \frac{|N_t^{D,n,m}|+|N_t^{E,n,m}|}{mn} > \gamma\big) = 0\,.$$

Hence, invoking (3.27) once again and Lemma 3.2,

$$\lim_{\chi\to 0}\limsup_{n,m\to\infty} \mathbf{P}\big(\sup_{s,t\le L,\,|t-s|<\chi} \big|\int_{\mathbb{R}_+^\ell} |u|(\Lambda_t^{n,m}(du)-\Lambda_s^{n,m}(du))\big| > \gamma\big) = 0\,.$$

□

We develop more semimartingale decompositions. Rearranging in the analogue of (2.3), accounting for (2.2) and (3.5) and assuming that $X_s^{n,m}(i-e_k)=0$ when $i_k=0$, yields, for $i\in\mathbb{Z}_+^\ell$,

$$\begin{aligned} X_t^{n,m}(i)-X_0^{n,m}(i) = \int_0^t \Big(\sum_{k=1}^\ell \big(&-X_s^{n,m}(i)i_k(\beta^{n,m,k}(i)+\delta^{n,m,k}(i)) \\ &+ X_s^{n,m}(i-e_k)(i_k-1)\beta^{n,m,k}(i-e_k) + X_s^{n,m}(i+e_k)(i_k+1)\delta^{n,m,k}(i+e_k) \\ &- X_s^{n,m}(i)i_k\mu^{n,m,k}(i) + X_s^{n,m}(i+e_k)(i_k+1)\mu^{n,m,k}(i+e_k) \\ &- \sum_{i'\ne i} X_s^{n,m}(i')i'_k\mu^{n,m,k}(i')\frac{X_s^{n,m}(i)}{X_s^{n,m,*}} + \sum_{i'\ne i-e_k} X_s^{n,m}(i')i'_k\mu^{n,m,k}(i')\frac{X_s^{n,m}(i-e_k)}{X_s^{n,m,*}}\big) \\ &- X_s^{n,m}(i)\phi^{n,m}(i) + \sum_{i'} X_s^{n,m}(i')\phi^{n,m}(i')\eta^{n,m}(i',i) \\ &- X_s^{n,m}(i)X_s^{n,m,*}\epsilon^{n,m}(i)\Big)\,ds + N_t^{n,m}(i)\,, \end{aligned} \quad (3.28)$$

where

$$\begin{aligned} N_t^{n,m}(i) = \sum_{k=1}^\ell \big(&-N_t^{B,n,m,k}(i) - N_t^{D,n,m,k}(i) + N_t^{B,n,m,k}(i-e_k) + N_t^{D,n,m,k}(i+e_k) \\ &-N_t^{\overline{M},n,m,k}(i) - N_t^{M,n,m,k}(i) + N_t^{\overline{M},n,m,k}(i+e_k) + N_t^{M,n,m,k}(i-e_k)\big) \\ &-N_t^{\overline{F},n,m}(i) + \sum_{i'} N_t^{F,n,m}(i',i) - N_t^{E,n,m}(i)\,. \end{aligned} \quad (3.29)$$

Hence, the predictable quadratic variation process of $N^{n,m}(i) = (N_t^{n,m}(i)\,, t \geq 0)$ is as follows

$$\langle N^{n,m}(i)\rangle_t = \sum_{k=1}^{\ell}\big(\langle N^{B,n,m,k}(i)\rangle_t + \langle N^{D,n,m,k}(i)\rangle_t + \langle N^{B,n,m,k}(i-e_k)\rangle_t$$
$$+ \langle N^{D,n,m,k}(i+e_k)\rangle_t + \langle N^{\overline{M},n,m,k}(i)\rangle_t + \langle N^{M,n,m,k}(i)\rangle_t + \langle N^{\overline{M},n,m,k}(i+e_k)\rangle_t$$
$$+ \langle N^{M,n,m,k}(i-e_k)\rangle_t + \langle N^{\overline{F},n,m}(i)\rangle_t + \sum_{i'}\langle N^{F,n,m}(i',i)\rangle_t + \langle N^{E,n,m}(i)\rangle_t$$
$$- \langle N^{\overline{M},n,m,k}(i), N^{M,n,m,k}(i-e_k)\rangle_t - \langle N^{M,n,m,k}(i), N^{\overline{M},n,m,k}(i+e_k)\rangle_t$$
$$+ \langle N^{\overline{M},n,m,k}(i+e_k), N^{M,n,m,k}(i-e_k)\rangle_t\big)\,, \quad (3.30)$$

where $\langle N^{\overline{M},n,m,k}(i), N^{M,n,m,k}(i-e_k)\rangle_t = \langle N^{\overline{M},n,m,k}(i+e_k), N^{M,n,m,k}(i-e_k)\rangle_t = 0$ when $i_k = 0$.

Lemma 3.4. *For all $L > 0$,*

$$\lim_{K\to\infty}\limsup_{n,m\to\infty}\mathbf{P}\big(\sup_{t\in[0,L]}\sum_i |X_t^{n,m}(i)|^2 > Km^2\big) = 0\,.$$

Proof. On writing (3.28) as

$$X_t^{n,m}(i) = X_0^{n,m}(i) + A_t^{n,m}(i) + N_t^{n,m}(i)\,,$$

we have that

$$X_t^{n,m}(i)^2 = X_0^{n,m}(i)^2 + 2\sum_{0<s\leq t} X_{s-}^{n,m}(i)\Delta X_s^{n,m}(i) + \sum_{0<s\leq t}(\Delta X_s^{n,m}(i))^2$$

$$= X_0^{n,m}(i)^2 + 2\int_0^t X_s^{n,m}(i)\,dA_s^{n,m}(i) + 2\int_0^t X_{s-}^{n,m}(i)dN_s^{n,m}(i) + \sum_{0<s\leq t}(\Delta X_s^{n,m}(i))^2\,. \quad (3.31)$$

As a consequence, on recalling that $|\Delta X_s^{n,m}(i)| \leq b$,

$$X_t^{n,m}(i)^2 \leq X_0^{n,m}(i)^2 + 2\int_0^t X_s^{n,m}(i)\,dA_s^{n,m}(i) + 2\int_0^t X_{s-}^{n,m}(i)dN_s^{n,m}(i)$$
$$+ b\sum_{0<s\leq t}|\Delta X_s^{n,m}(i)|\,. \quad (3.32)$$

By (3.28),

$$\int_0^t X_s^{n,m}(i)\,dA_s^{n,m}(i) = \int_0^t X_s^{n,m}(i)\,\Big(\sum_{k=1}^{\ell}\big(-X_s^{n,m}(i)i_k(\beta^{n,m,k}(i)+\delta^{n,m,k}(i))$$
$$+X_s^{n,m}(i-e_k)(i_k-1)\beta^{n,m,k}(i-e_k)+X_s^{n,m}(i+e_k)(i_k+1)\delta^{n,m,k}(i+e_k)$$
$$-X_s^{n,m}(i)i_k\mu^{n,m,k}(i)+X_s^{n,m}(i+e_k)(i_k+1)\mu^{n,m,k}(i+e_k)$$
$$+\sum_{i'\neq i-e_k} X_s^{n,m}(i')i'_k\mu^{n,m,k}(i')\frac{X_s^{n,m}(i-e_k)}{X_s^{n,m,*}}-\sum_{i'\neq i} X_s^{n,m}(i')i'_k\mu^{n,m,k}(i')\frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\big)$$
$$+\sum_{i'} X_s^{n,m}(i')\phi^{n,m}(i')\eta^{n,m}(i',i)-X_s^{n,m}(i)\phi^{n,m}(i)-X_s^{n,m}(i)X_s^{n,m,*}\epsilon^{n,m}(i)\Big)\,ds\,.$$

Since the functions $i_k(\beta^{n,m,k}(i)+\delta^{n,m,k}(i)+\mu^{n,m,k}(i))$, $\phi^{n,m}(i)$ and $\eta^{n,m}(i',i)$ are bounded, there exist $K_1>0$ and $K_2>0$ such that

$$|\int_0^t X_s^{n,m}(i)\,dA_s^{n,m}(i)| \le K_1\int_0^t X_s^{n,m}(i)\big(X_s^{n,m}(i)+\sum_{k=1}^{\ell}\big(X_s^{n,m}(i-e_k)+X_s^{n,m}(i+e_k)\big)$$
$$+\sum_{i'} X_s^{n,m}(i')\phi^{n,m}(i')\eta^{n,m}(i',i)+X_s^{n,m}(i)X_s^{n,m,*}\epsilon^{n,m}(i)\big)\,ds$$
$$\le K_2\int_0^t \big(X_s^{n,m}(i)^2+\sum_{k=1}^{\ell}\big(X_s^{n,m}(i-e_k)^2+X_s^{n,m}(i+e_k)^2\big)$$
$$+X_s^{n,m}(i)X_s^{n,m,*}+X_s^{n,m}(i)^2X_s^{n,m,*}\epsilon^{n,m}(i)\big)\,ds\,.$$

On recalling that $m\epsilon^{n,m}(i)$ is bounded in n,m,i, for some $K_3>0$,

$$\sum_i|\int_0^t X_s^{n,m}(i)\,dA_s^{n,m}(i)| \le K_3\int_0^t\sum_i X_s^{n,m}(i)^2(1+R_s^{n,m})\,ds+K_3m^2\int_0^t(R_s^{n,m})^2\,ds\,. \tag{3.33}$$

By (2.5), (2.6), (3.7), (3.8), (3.9), and (3.30), for some $K_4>0$,

$$\sum_i d\langle N^{n,m}(i)\rangle_t \le K_4m(1+(R_t^{n,m})^2)\,dt\,. \tag{3.34}$$

Let, for $\gamma>0$,

$$\tau_\gamma^{n,m}=\inf\{t\ge 0:\ R_t^{n,m}>\gamma\}\,. \tag{3.35}$$

Let $X^{n,m}(i)\circ N^{n,m}(i)_t=\int_0^t X^{n,m}_{s-}(i)dN^{n,m}_s(i)$ and $X^{n,m}(i)\circ N^{n,m}(i)=\bigl(X^{n,m}(i)\circ N^{n,m}(i)_t\,,t\ge 0\bigr)$. The process $X^{n,m}(i)\circ N^{n,m}(i)$ is a locally square integrable martingale with the predictable quadratic variation process $(\int_0^t X^{n,m}_s(i)^2 d\langle N^{n,m}(i)\rangle_s\,,t\ge 0\bigr)$. By (3.34), $\int_0^{t\wedge\tau^{n,m}_\gamma} X^{n,m}_s(i)^2 d\langle N^{n,m}(i)\rangle_s$ is bounded for given n,m. Therefore, $\bigl(X^{n,m}(i)\circ N^{n,m}_{t\wedge\tau^{n,m}_\gamma}(i)\,,t\ge 0\bigr)$ is a martingale, so $\mathbf{E}(X^{n,m}(i)\circ N^{n,m}(i)_{t\wedge\tau^{n,m}_\gamma}|\mathcal{F}^{n,m}_0)=0$. By (3.32) and (3.33), for $R>0$, introducing the event $\Gamma^{n,m}=\{\sum_{i'}X^{n,m}_0(i')^2\le Rm^2\}$,

$$\mathbf{E}\sum_i X^{n,m}_{t\wedge\tau^{n,m}_\gamma}(i)^2\mathbf{1}_{\Gamma^{n,m}}\le Rm^2+2K_3(1+\gamma)\int_0^t\mathbf{E}\sum_i X^{n,m}_{s\wedge\tau^{n,m}_\gamma}(i)^2\mathbf{1}_{\Gamma^{n,m}}\,ds$$
$$+2K_3m^2\gamma^2t+b\mathbf{E}\sum_{0<s\le t}|\Delta X^{n,m}_{s\wedge\tau^{n,m}_\gamma}(i)|\,.$$

By the analogue of (2.3),

$$\mathbf{E}\sum_{0<s\le t}|\Delta X^{n,m}_{s\wedge\tau^{n,m}_\gamma}(i)|\le\sum_i\sum_{k=1}^{\ell}\mathbf{E}\bigl(B^{n,m,k}_{t\wedge\tau^{n,m}_\gamma}(i)+D^{n,m,k}_{t\wedge\tau^{n,m}_\gamma}(i)+M^{n,m,k}_{t\wedge\tau^{n,m}_\gamma}(i)+\overline{M}^{n,m,k}_{t\wedge\tau^{n,m}_\gamma}(i)\bigr)$$
$$+\sum_i\mathbf{E}\overline{F}^{n,m}_{t\wedge\tau^{n,m}_\gamma}(i)+\sum_i\sum_{i'}\mathbf{E}F^{n,m}_{t\wedge\tau^{n,m}_\gamma}(i',i)+\sum_i\mathbf{E}E^{n,m}_{t\wedge\tau^{n,m}_\gamma}(i)\,.$$

Since the processes on the righthand sides of (3.5) are local martingales and $X^{n,m,*}_s\le m\gamma$ when $s<\tau^{n,m}_\gamma$, for some $K_5>0$,

$$\mathbf{E}\sum_{0<s\le t}|\Delta X^{n,m}_{s\wedge\tau^{n,m}_\gamma}(i)|\le K_5m\gamma(1+\gamma)t\,. \tag{3.36}$$

It follows that

$$\mathbf{E}\sum_i X^{n,m}_{t\wedge\tau^{n,m}_\gamma}(i)^2\mathbf{1}_{\Gamma^{n,m}}\le Rm^2+2K_3(1+\gamma)\int_0^t\mathbf{E}\sum_i X^{n,m}_{s\wedge\tau^{n,m}_\gamma}(i)^2\mathbf{1}_{\Gamma^{n,m}}\,ds$$
$$+2K_3m^2\gamma^2t+bK_5m\gamma(1+\gamma)t\,.$$

By Gronwall's inequality,

$$\mathbf{E}\sum_i X^{n,m}_{t\wedge\tau^{n,m}_\gamma}(i)^2\mathbf{1}_{\Gamma^{n,m}}\le(Rm^2+2K_3\gamma^2tm^2+bK_5m\gamma(1+\gamma)t)e^{2K_3(1+\gamma)t}\,. \tag{3.37}$$

By (3.31) and (3.33), recalling (2.10), for some $K_6 > 0$,

$$2\sum_i \Big|\int_0^t X_{s-}^{n,m}(i)dN_s^{n,m}(i)\Big| \le \sum_i X_t^{n,m}(i)^2 + \sum_i X_0^{n,m}(i)^2 +$$
$$2\sum_i \Big|\int_0^t X_s^{n,m}(i)\, dA_s^{n,m}(i)\Big| + b\sum_{0<s\le t}\sum_i |\Delta X_s^{n,m}(i)|$$
$$\le \sum_i X_t^{n,m}(i)^2 + \sum_i X_0^{n,m}(i)^2 + K_6\int_0^t \sum_i X_s^{n,m}(i)^2\, ds + K_6 m^2 \int_0^t (R_s^{n,m})^2\, ds$$
$$+ K_6 \int_0^t \sum_i X_s^{n,m}(i)^2 R_s^{n,m}\, ds + b\sum_{0<s\le t}\sum_i |\Delta X_s^{n,m}(i)|\,.$$

Therefore, by (3.36) and (3.37),

$$\mathbf{2E}\sum_i \Big|\int_0^{t\wedge\tau_\gamma^{n,m}} X_{s-}^{n,m}(i)dN_s^{n,m}(i)\Big|\mathbf{1}_{\Gamma^{n,m}} \le (1+K_6 t + K_6 t\gamma)(Rm^2 + 2K_3\gamma^2 t m^2$$
$$+ bK_5 m\gamma(1+\gamma)t)e^{2K_3(1+\gamma)t} + Rm^2 + K_6 t\gamma^2 m^2 + bK_5 m\gamma(1+\gamma)t\,.$$

On applying Doob's inequality, for $L > 0$ and $K > 0$,

$$\mathbf{P}\Big(\mathbf{1}_{\Gamma^{n,m}} \sup_{t\le L\wedge\tau_\gamma^{n,m}} \Big|\sum_i \int_0^t X_{s-}^{n,m}(i)\, dN_s^{n,m}(i)\Big| > Km^2\Big)$$
$$\le \frac{1}{2Km^2}\Big((1+K_6 L + K_6 L\gamma)(Rm^2 + 2K_3\gamma^2 L m^2$$
$$+ bK_5 m\gamma(1+\gamma)L)e^{2K_3(1+\gamma)L} + Rm^2 + K_6 L\gamma^2 m^2 + bK_5 m\gamma(1+\gamma)L\Big)\,.$$

By (3.31), (3.33), (3.36), and Gronwall's inequality,

$$\sup_{t\le L\wedge\tau_\gamma^{n,m}} \sum_i X_t^{n,m}(i)^2 \le \Big(\sum_i X_0^{n,m}(i)^2 + 2K_3 m^2\gamma^2 L$$
$$+ 2\sup_{t\le L\wedge\tau_\gamma^{n,m}} \Big|\sum_i \int_0^t X_{s-}^{n,m}(i)dN_s^{n,m}(i)\Big| + bK_5 m\gamma(1+\gamma)L\Big)e^{2K_3(1+\gamma)L}\,.$$

It follows that, for arbitrary $K' > 0$,

$$\mathbf{P}\big(\sup_{t\le L}\sum_i X_t^{n,m}(i)^2 > K'm^2\big) \le \mathbf{P}(\sum_i X_0^{n,m}(i)^2 > Rm^2) + \mathbf{P}(\tau_\gamma^{n,m} \le L)$$

$$+ \mathbf{P}\big(\mathbf{1}_{\Gamma^{n,m}} \sup_{t\le L\wedge\tau_\gamma^{n,m}} |\sum_i \int_0^t X_{s-}^{n,m}(i)dN_s^{n,m}(i)| > \frac{m^2}{2}\,\big(K'e^{-2K_3(1+\gamma)L} - R$$

$$- 2K_3\gamma^2 L - \frac{bK_5\gamma(1+\gamma)L}{m}\big)\big) \le \mathbf{P}(\sum_i X_0^{n,m}(i)^2 > Rm^2) + \mathbf{P}(\tau_\gamma^{n,m} \le L)$$

$$+ \frac{1}{m^2\big(K'e^{-2K_3(1+\gamma)L} - R - 2K_3\gamma^2L - bK_5\gamma(1+\gamma)L/m\big)}$$

$$\Big((1+K_6L+K_6L\gamma)(Rm^2 + 2K_3\gamma^2Lm^2 + bK_5m\gamma(1+\gamma)L)e^{2K_3(1+\gamma)L}$$

$$+ Rm^2 + K_6L\gamma^2m^2 + bK_5m\gamma(1+\gamma)L\Big)\,.$$

On recalling (3.35),

$$\lim_{K'\to\infty}\limsup_{n,m\to\infty}\mathbf{P}(\sup_{t\le L}\sum_i X_t^{n,m}(i)^2 > K'm^2)$$

$$\le \limsup_{n,m\to\infty}\mathbf{P}(\sum_i X_0^{n,m}(i)^2 > Rm^2) + \limsup_{n,m\to\infty}\mathbf{P}\big(\sup_{t\le L} R_t^{n,m} \ge \gamma)\,.$$

Letting $R\to\infty$, $\gamma\to\infty$, and accounting for Lemma 4.1 and the hypotheses of the theorem yield the convergence

$$\lim_{K'\to\infty}\limsup_{n,m\to\infty}\mathbf{P}\big(\sup_{t\le L}\sum_i X_t^{n,m}(i)^2 > K'm^2\big) = 0\,.$$

□

Let $a_t(u)$, where $t\in\mathbb{R}$, $u\in\mathbb{R}^\ell$, represent a bounded and continuously differentiable function compactly supported in u uniformly over t from bounded intervals.

By (3.9) and (3.28), for $t \geq 0$,

$$\int_0^t \sum_i a_s\big(\frac{i}{n}\big)\frac{dX_s^{n,m}(i)}{m} = \int_0^t \sum_i a_s\big(\frac{i}{n}\big)\Big(\sum_{k=1}^{\ell}\Big(-\frac{X_s^{n,m}(i)}{m} i_k(\beta^{n,m,k}(i)+\delta^{n,m,k}(i))$$
$$+\frac{X_s^{n,m}(i-e_k)}{m}(i_k-1)\beta^{n,m,k}(i-e_k)+\frac{X_s^{n,m}(i+e_k)}{m}(i_k+1)\delta^{n,m,k}(i+e_k)$$
$$+\sum_{i'\neq i-e_k}\frac{X_s^{n,m}(i')}{m}\, i'_k\mu^{n,m,k}(i')\frac{X_s^{n,m}(i-e_k)}{mR_s^{n,m}}-\sum_{i'\neq i}\frac{X_s^{n,m}(i')}{m}\, i'_k\mu^{n,m,k}(i')\frac{X_s^{n,m}(i)}{mR_s^{n,m}}$$
$$-\frac{X_s^{n,m}(i)}{m}\, i_k\mu^{n,m,k}(i)+\frac{X_s^{n,m}(i+e_k)}{m}(i_k+1)\mu^{n,m,k}(i+e_k)\Big)-\frac{X_s^{n,m}(i)}{m}\phi^{n,m}(i)$$
$$+\sum_{i'}\frac{X_s^{n,m}(i')}{m}\phi^{n,m}(i')\eta^{n,m}(i',i)-\frac{X_s^{n,m}(i)}{m}\sum_{i'}X_s^{n,m}(i')\epsilon^{n,m}(i)\Big)\, ds+N_t'^{n,m}, \tag{3.38}$$

where

$$N_t'^{n,m}=\frac{1}{m}\sum_i a_t\big(\frac{i}{n}\big)N_t^{n,m}(i)\,. \tag{3.39}$$

Taking into account that by integration by parts

$$a_t\big(\frac{i}{n}\big)\frac{X_t^{n,m}(i)}{m}=a_0\big(\frac{i}{n}\big)\frac{X_0^{n,m}(i)}{m}+\int_0^t\frac{X_s^{n,m}(i)}{m}\,\partial_s a_s\big(\frac{i}{n}\big)\,ds+\int_0^t a_s\big(\frac{i}{n}\big)\frac{dX_s^{n,m}(i)}{m}\,,$$

changing summation indices and regrouping in (3.38) yield

$$\sum_i a_t\big(\frac{i}{n}\big)\frac{X_t^{n,m}(i)}{m}=\sum_i a_0\big(\frac{i}{n}\big)\frac{X_0^{n,m}(i)}{m}$$
$$+\int_0^t\sum_i\Big(\partial_s a_s\big(\frac{i}{n}\big)\frac{X_s^{n,m}(i)}{m}+\frac{X_s^{n,m}(i)}{m}\sum_{k=1}^{\ell}\Big(n\big(a_s\big(\frac{i+e_k}{n}\big)-a_s\big(\frac{i}{n}\big)\big)\ \frac{i_k}{n}\ \beta^{n,m,k}(i)$$
$$+n\big(a_s\big(\frac{i-e_k}{n}\big)-a_s\big(\frac{i}{n}\big)\big)\frac{i_k}{n}\,\delta^{n,m,k}(i)+n\big(a_s\big(\frac{i-e_k}{n}\big)-a_s\big(\frac{i}{n}\big)\big)\ \frac{i_k}{n}\,\mu^{n,m,k}(i)$$
$$+n\big(a_s\big(\frac{i+e_k}{n}\big)-a_s\big(\frac{i}{n}\big)\big)\sum_{i'\neq i}\frac{i'_k}{n}\ \frac{\mu^{n,m,k}(i')X_s^{n,m}(i')}{mR_s^{n,m}}\Big)$$
$$+a_s\big(\frac{i}{n}\big)\sum_{i'}\frac{X_s^{n,m}(i')}{m}\,\phi^{n,m}(i')\eta^{n,m}(i',i)-a_s\big(\frac{i}{n}\big)\phi^{n,m}(i)\,\frac{X_s^{n,m}(i)}{m}$$

$$- a_s\big(\frac{i}{n}\big)\,\frac{X^{n,m}_s(i)}{m}\sum_{i'} X^{n,m}_s(i')\epsilon^{n,m}(i)\Big)\,ds + N'^{n,m}_t\,. \quad (3.40)$$

Owing to (3.29) and (3.39),

$$\begin{aligned} &N'^{n,m}_t \\ &= \frac{1}{m}\sum_i\Big(\sum_{k=1}^{\ell}\Big(\big(a_t\big(\frac{i+e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)N^{B,n,m,k}_t(i)+\big(a_t\big(\frac{i-e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)N^{D,n,m,k}_t(i) \\ &\quad + \big(a_t\big(\frac{i+e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)N^{M,n,m,k}_t(i)+\big(a_t\big(\frac{i-e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)N^{\overline{M},n,m,k}_t(i)\Big) \\ &\quad + \sum_{i'} a_t\big(\frac{i}{n}\big)N^{F,n,m}_t(i',i) - a_t\big(\frac{i}{n}\big)\big(N^{\overline{F},n,m}_t(i)+N^{E,n,m}_t(i)\big)\Big)\,. \end{aligned}$$

Thus, $N'^{n,m} = (N'^{n,m}_t\,, t \ge 0)$ is a locally square integrable martingale with the predictable quadratic variation process

$$\begin{aligned} \langle N'^{n,m}\rangle_t = \frac{1}{m^2}\sum_i\Big(\sum_{k=1}^{\ell}\Big(&\big(a_t\big(\frac{i+e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)^2\langle N^{B,n,m,k}(i)\rangle_t \\ &+\big(a_t\big(\frac{i-e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)^2\langle N^{D,n,m,k}(i)\rangle_t \\ &+\big(a_t\big(\frac{i+e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)^2\langle N^{M,n,m,k}(i)\rangle_t \\ &+\big(a_t\big(\frac{i-e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)^2\langle N^{\overline{M},n,m,k}(i)\rangle_t \\ +\sum_{i'}\big(a_t\big(\frac{i+e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)&\big(a_t\big(\frac{i'-e_k}{n}\big)-a_t\big(\frac{i'}{n}\big)\big)\langle N^{M,n,m,k}(i),N^{\overline{M},n,m,k}(i')\rangle_t\Big) \\ &+a_t\big(\frac{i}{n}\big)^2\Big(\sum_{i'}\langle N^{F,n,m}(i',i)\rangle_t+\langle N^{E,n,m}(i)\rangle_t+\langle N^{\overline{F},n,m}(i)\rangle_t\Big) \\ &+a_t\big(\frac{i}{n}\big)\sum_{i'}\sum_j a_t\big(\frac{j}{n}\big)\langle N^{F,n,m}(i',i),N^{F,n,m}(i',j)\rangle_t \\ &-a_t\big(\frac{i}{n}\big)\sum_{i'} a_t\big(\frac{i'}{n}\big)\langle N^{F,n,m}(i,i'),N^{\overline{F},n,m}(i)\rangle_t\Big)\,. \end{aligned}$$

Substitutions from (3.7) and (3.8) with the account of (3.1) yield

$$\langle N'^{n,m}\rangle_t = \frac{1}{m^2}\sum_i\Big(\sum_{k=1}^{\ell}\Big(\big(a_t\big(\frac{i+e_k}{n}\big)-a_t\big(\frac{i}{n}\big)\big)^2\int_0^t X^{n,m}_s(i)i_k\beta^{n,m,k}(i)\,ds$$

$$+ \big(a_t\big(\frac{i-e_k}{n}\big) - a_t\big(\frac{i}{n}\big)\big)^2 \int_0^t X_s^{n,m}(i) i_k \delta^{n,m,k}(i)\, ds$$

$$+ \big(a_t\big(\frac{i+e_k}{n}\big) - a_t\big(\frac{i}{n}\big)\big)^2 \int_0^t \sum_{i' \neq i} X_s^{n,m}(i') i'_k \mu^{n,m,k}(i') \frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\, ds$$

$$+ \big(a_t\big(\frac{i-e_k}{n}\big) - a_t\big(\frac{i}{n}\big)\big)^2 \int_0^t X_s^{n,m}(i) i_k \mu^{n,m,k}(i) \big(1 - \frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\big)\, ds$$

$$+ \sum_{i' \neq i} \big(a_t\big(\frac{i+e_k}{n}\big) - a_t\big(\frac{i}{n}\big)\big)\big(a_t\big(\frac{i'-e_k}{n}\big) - a_t\big(\frac{i'}{n}\big)\big) \int_0^t X_s^{n,m}(i') i'_k \mu^{n,m,k}(i') \frac{X_s^{n,m}(i)}{X_s^{n,m,*}}\, ds\Big)$$

$$+ \sum_{i'} a_t\big(\frac{i}{n}\big)^2 \int_0^t X_s^{n,m}(i') \phi^{n,m}(i') \alpha^{n,m}(i', i)\, ds$$

$$+ a_t\big(\frac{i}{n}\big)^2 \int_0^t X_s^{n,m}(i) \big(X_s^{n,m,*} \epsilon^{n,m}(i) + \phi^{n,m}(i)\big)\, ds$$

$$+ \int_0^t X_s^{n,m}(i) \phi^{n,m}(i) \mathbf{E}\big(\sum_{i'} a_t\big(\frac{i'}{n}\big) \theta_i^{n,m}(i', 1)\big)^2\, ds$$

$$- a_t\big(\frac{i}{n}\big) \sum_{i'} a_t\big(\frac{i'}{n}\big) \int_0^t X_s^{n,m}(i) \eta^{n,m}(i, i') \phi^{n,m}(i)\, ds\Big). \quad (3.41)$$

Let

$$a_t^n(u) = a_t\big(\frac{\lfloor nu \rfloor}{n}\big) \quad (3.42)$$

and

$$Y_t^{n,m} = \int_{\mathbb{R}_+^\ell} a_t^n(u) \Lambda_t^{n,m}(du)\,. \quad (3.43)$$

By (2.4), (2.5), (2.6), (3.40), and (3.41),

$$Y_t^{n,m} = Y_0^{n,m} + \int_0^t Z_s^{n,m}\, ds + N_t'^{n,m}\,, \quad (3.44)$$

where

$$Z_s^{n,m} = \int\limits_{\mathbb{R}_+^\ell} \Big(\partial_s a_s^n(u) + \sum_{k=1}^\ell \Big(n\big(a_s^n\big(u+\frac{e_k}{n}\big) - a_s^n(u)\big)\,\frac{\lfloor nu_k\rfloor}{n}\,\hat\beta^{n,m,k}(u)$$

$$+ n\big(a_s^n\big(u-\frac{e_k}{n}\big) - a_s^n(u)\big)\,\frac{\lfloor nu_k\rfloor}{n}\,(\hat\delta^{n,m,k}(u)+\hat\mu^{n,m,k}(u))$$

$$+ n\big(a_s^n\big(u+\frac{e_k}{n}\big) - a_s^n(u)\big)\int\limits_{\mathbb{R}_+^\ell}\frac{1}{R_s^{n,m}}\,\frac{\lfloor nu'_k\rfloor}{n}\,\hat\mu^{n,m,k}(u')\,\Lambda_s^{n,m}(du')\Big)$$

$$+ \int\limits_{\mathbb{R}_+^\ell} a_s^n(u')\hat\phi^{n,m}(u)\hat\eta^{n,m}(u,du') - a_s^n(u)\big(\hat\phi^{n,m}(u) + R_s^{n,m}\hat\epsilon^{n,m}(u)\big)\Big)\Lambda_s^{n,m}(du)$$

$$- \int_0^t \sum_i \sum_{k=1}^\ell \big(a_s\big(\frac{i+e_k}{n}\big) - a_s\big(\frac{i}{n}\big)\big)\frac{1}{R_s^{n,m}}\,i_k\mu^{n,m,k}(i)\frac{X_s^{n,m}(i)^2}{m^2}\,ds \quad (3.45)$$

and, by (3.41), on letting $B_{1/n}(u) = \{u' \in \mathbb{R}_+^\ell : \max_{k=1,\dots,\ell}|u'_k - u_k| < 1/n\}$,

$$\langle N'^{n,m}\rangle_t = \frac{1}{m}\int_0^t \int\limits_{\mathbb{R}_+^\ell}\Big(\sum_{k=1}^\ell \Big(n\big(a_t^n\big(u+\frac{e_k}{n}\big) - a_t^n(u)\big)^2\frac{\lfloor nu_k\rfloor}{n}\,\hat\beta^{n,m,k}(u)$$

$$+ n\big(a_t^n\big(u-\frac{e_k}{n}\big) - a_t^n(u)\big)^2\frac{\lfloor nu_k\rfloor}{n}\,\hat\delta^{n,m,k}(u)$$

$$+ n\big(a_t^n\big(u+\frac{e_k}{n}\big) - a_t^n(u)\big)^2 \int\limits_{\mathbb{R}_+^\ell\setminus B_{1/n}(u)}\frac{1}{R_s^{n,m}}\,\frac{\lfloor nu'_k\rfloor}{n}\,\hat\mu^{n,m,k}(u')\,\Lambda_s^{n,m}(du')$$

$$+ n\big(a_t^n\big(u-\frac{e_k}{n}\big) - a_t^n(u)\big)^2\frac{\lfloor nu_k\rfloor}{n}\,\hat\mu^{n,m,k}(u)\big(1-\frac{\hat X_s^{n,m}(u)}{R_s^{n,m}}\big)$$

$$+ n\big(a_t^n\big(u+\frac{e_k}{n}\big) - a_t^n(u)\big)\frac{1}{R_s^{n,m}}$$

$$\int\limits_{\mathbb{R}_+^\ell\setminus B_{1/n}(u)}\big(a_t^n\big(u'-\frac{e_k}{n}\big) - a_t^n(u')\big)\frac{\lfloor nu'_k\rfloor}{n}\,\hat\mu^{n,m,k}(u')\,\Lambda_s^{n,m}(du')\Big)$$

$$+ \hat\phi^{n,m}(u)\int\limits_{\mathbb{R}_+^\ell} a_t^n(u')^2\hat\alpha^{n,m}(u,du') + a_t^n(u)^2\big(R_s^{n,m}\hat\epsilon^{n,m}(u) + \hat\phi^{n,m}(u)\big)$$

$$+ \hat\phi^{n,m}(u)\mathbf{E}\big(\sum_{i'} a_t\big(\frac{i'}{n}\big)\theta^{n,m}_{\lfloor nu\rfloor}(i',1)\big)^2$$

$$- a_t^n(u)\hat{\phi}^{n,m}(u) \int_{\mathbb{R}_+^\ell} a_t^n(u')\hat{\eta}^{n,m}(u,du')\Big)\Lambda_s^{n,m}(du)\,ds\,,$$

where, in analogy with (2.6), $\hat{\alpha}^{n,m}(u,\Gamma) = \sum_{i'} \alpha^{n,m}(\lfloor nu \rfloor, i'/n)\mathbf{1}_\Gamma(i'/n)$. We note that $\mathbf{E}\big(\sum_{i'} a_t(i'/n)\theta^{n,m}_{\lfloor nu \rfloor}(i',1)\big)^2 \le b^2 \sup_{t,u} a_t(u)^2$ and, by (3.6),

$$\hat{\alpha}^{n,m}(u,\mathbb{R}_+^\ell) = \sum_{i'} \alpha^{n,m}\big(\lfloor nu \rfloor, i'\big) \le b\eta^{n,m}(\lfloor nu \rfloor, \mathbb{R}_+^\ell) \le b^2\,.$$

Therefore, on recalling (3.9), (3.42), the fact that the $a_t(u)$ are differentiable of bounded support and the boundedness hypotheses of the theorem, we have that, given $L > 0$, for some $\hat{K} > 0$, which may depend on L, for $t \le L$,

$$\langle N'^{n,m} \rangle_t \le \frac{\hat{K}}{m} \int_0^t (1 + (R_s^{n,m})^2)\,ds\,.$$

By Lemma 3.2, Lemma 3.3 and the Lenglart–Rebolledo inequality, for $\chi > 0$,

$$\lim_{n,m\to\infty} \mathbf{P}(\sup_{t\le L} |N_t'^{n,m}| > \chi) = 0\,. \tag{3.46}$$

Lemma 3.5. *The sequence $\Lambda^{n,m}$ is C-tight for convergence in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{M}_+(\mathbb{R}_+^\ell))$.*

Proof. By Theorem 4.6 in Jakubowski [13] and Topsœ [26] (or Topsœ [27]), it is sufficient to prove that, for all $L > 0$ and $\gamma > 0$,

$$\lim_{K\to\infty} \limsup_{n,m\to\infty} \mathbf{P}\big(\sup_{t\in[0,L]} \Lambda_t^{n,m}(\mathbb{R}_+^\ell) > K\big) = 0\,,$$

$$\lim_{K\to\infty} \limsup_{n,m\to\infty} \mathbf{P}\big(\sup_{t\in[0,L]} \Lambda_t^{n,m}(u:\,|u| > K) > \gamma\big) = 0$$

and, for all continuous functions g of compact support and all $\gamma > 0$,

$$\lim_{\chi\to 0} \limsup_{n,m\to\infty} \mathbf{P}(\sup_{\substack{s,t\in[0,L]:\\ |s-t|\le\chi}} |\int_{\mathbb{R}_+^\ell} g(u)(\Lambda_s^{n,m}(du) - \Lambda_t^{n,m}(du))| > \gamma) = 0\,.$$

The first and second requirements are fulfilled by Lemma 3.3 (see (3.18) and (3.27)). Let us note that, by (3.45), (3.9), Lemma 3.2 and Lemma 3.3,

$$\lim_{K\to\infty} \limsup_{n,m\to\infty} \mathbf{P}(\sup_{s\le t} |Z_s^{n,m}| > K) = 0\,.$$

Therefore, the third limit follows from (3.43), (3.44), (3.45), and (3.46). ∎

We now identify limit points of the $\Lambda^{n,m}$. Let $(\tilde{\lambda}_t\,, t \geq 0)$ represent a limit point in distribution of $\Lambda^{n,m}$ along a subsequence in $\mathbb{D}(\mathbb{R}_+, \mathbb{M}_+(\mathbb{R}_+^\ell))$. We keep the notation (n,m) for the subsequence. Since the functions $a_t^n(u)$ are bounded uniformly in t, u and n and converge to $a_t(u)$ uniformly in u locally uniformly in t, see (3.42), by (3.43) and the continuous mapping theorem, in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$,

$$Y_t^{n,m} \to Y_t = \int\limits_{\mathbb{R}_+^\ell} a_t(u)\, \tilde{\lambda}_t(du)\,. \tag{3.47}$$

On recalling (3.9) and Lemma 3.2, we obtain that, in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$,

$$R_s^{n,m} \to R_s = \tilde{\lambda}_s(\mathbb{R}_+^\ell)\,. \tag{3.48}$$

By Lemma 3.2, the latter quantity is bounded away from zero locally uniformly in s with probability 1. Since the function $a_t(u)$ is continuously differentiable and is of compact support in u locally uniformly in t, $a_s^n(u) \to a_s(u)$, $\partial_s a_s^n(u) \to \partial_s a_s(u)$, and $n\big(a_s^n\big(u \pm e_k/n\big) - a_s^n(u)\big) \to \pm\partial_{u_k} a_s(u)$ uniformly in u locally uniformly in t, as $n \to \infty$. Therefore, in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$, for $k = 1, \ldots, \ell$,

$$\int\limits_{\mathbb{R}_+^\ell} n\big(a_s^n\big(u + \frac{e_k}{n}\big) - a_s^n(u)\big)\Lambda_s^{n,m}(du) \to \int\limits_{\mathbb{R}_+^\ell} \partial_{u_k} a_s(u)\tilde{\lambda}_s(du)\,, \tag{3.49}$$

and, since the $u_k\hat{\mu}^{n,m,k}(u)$ are bounded and the convergences in (2.7) hold,

$$\int\limits_{\mathbb{R}_+^\ell} \frac{\lfloor nu_k \rfloor}{n}\, \hat{\mu}^{n,m,k}(u)\Lambda_s^{n,m}(du) \to \int\limits_{\mathbb{R}_+^\ell} u_k\hat{\mu}^k(u)\, \tilde{\lambda}_s(du)\,. \tag{3.50}$$

Since the convergences in (3.48), (3.49), and (3.50) hold jointly, in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$,

$$\begin{aligned}\int\limits_{\mathbb{R}_+^\ell} n\big(a_s^n\big(u + \frac{e_k}{n}\big) - a_s^n(u)\big)\,\Lambda_s^{n,m}(du) \int\limits_{\mathbb{R}_+^\ell} \frac{1}{R_s^{n,m}}\, \frac{\lfloor nu_k' \rfloor}{n}\, \hat{\mu}^{n,m,k}(u')\,\Lambda_s^{n,m}(du') \\ \to \frac{\int\limits_{\mathbb{R}_+^\ell} u_k\hat{\mu}^k(u)\,\tilde{\lambda}_s(du)}{\tilde{\lambda}_s(\mathbb{R}_+^\ell)} \int\limits_{\mathbb{R}_+^\ell} \partial_{u_k} a_s(u)\tilde{\lambda}_s(du)\,.\end{aligned} \tag{3.51}$$

Similar lines of reasoning show that, jointly in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^4)$, and jointly with the convergence in (3.51), for $k = 1, \ldots, \ell$,

$$\int_{\mathbb{R}_+^\ell} n\big(a_s^n\big(u + \frac{e_k}{n}\big) - a_s^n(u)\big) \, \frac{\lfloor nu_k \rfloor}{n} \, \hat{\beta}^{n,m,k}(u) \, \Lambda_s^{n,m}(du) \to \int_{\mathbb{R}_+^\ell} \partial_{u_k} a_s(u) \, u_k \hat{\beta}^k(u) \, \tilde{\lambda}_s(du),$$

$$\int_{\mathbb{R}_+^\ell} n\big(a_s^n(u) - a_s^n\big(u - \frac{e_k}{n}\big)\big) \, \frac{\lfloor nu_k \rfloor}{n} \, \hat{\delta}^{n,m,k}(u) \, \Lambda_s^{n,m}(du) \to \int_{\mathbb{R}_+^\ell} \partial_{u_k} a_s(u) u_k \hat{\delta}^k(u) \tilde{\lambda}_s(du) \, ,$$

$$\int_{\mathbb{R}_+^\ell} n\big(a_s^n\big(u + \frac{e_k}{n}\big) - a_s^n(u)\big) \Lambda_s^{n,m}(du) \int_{\mathbb{R}_+^\ell} \frac{1}{R_s^{n,m}} \, \frac{\lfloor nu'_k \rfloor}{n} \, \hat{\mu}^{n,m,k}(u') \, \Lambda_s^{n,m}(du') \to \frac{\int_{\mathbb{R}_+^\ell} u_k \hat{\mu}^k(u) \, \tilde{\lambda}_s(du)}{\tilde{\lambda}_s(\mathbb{R}_+^\ell)} \int_{\mathbb{R}_+^\ell} \partial_{u_k} a_s(u) \tilde{\lambda}_s(du) \, ,$$

$$\int_{\mathbb{R}_+^\ell} n\big(a_s^n(u) - a_s^n\big(u - \frac{e_k}{n}\big)\big) \, \frac{\lfloor nu_k \rfloor}{n} \, \hat{\mu}^{n,m,k}(u) \, \Lambda_s^{n,m}(du) \to \int_{\mathbb{R}_+^\ell} \partial_{u_k} a_s(u) u_k \hat{\mu}^k(u) \, \tilde{\lambda}_s(du) \, ,$$

and

$$\int_{\mathbb{R}_+^\ell} \big(\hat{\phi}^{n,m}(u) \int_{\mathbb{R}_+^\ell} a_s^n(u') \hat{\eta}^{n,m}(u, du') - a_s^n(u) \big(\hat{\phi}^{n,m}(u) + R_s^{n,m} \hat{\epsilon}^{n,m}(u)\big)\big) \, \Lambda_s^{n,m}(du) \to \int_{\mathbb{R}_+^\ell} \big(\int_{\mathbb{R}_+^\ell} a_s(u') \hat{\varphi}(u, du') - a_s(u) \big(\hat{\phi}(u) + R_s \hat{\epsilon}(u)\big)\big) \, \tilde{\lambda}_s(du) \, .$$

Let us note also that, by Lemma 3.4, for arbitrary $\gamma > 0$,

$$\lim_{n,m\to\infty} \mathbf{P}\big(\sup_{s\le t} \sum_i \sum_{k=1}^{\ell} |a_s\big(\frac{i+e_k}{n}\big) - a_s\big(\frac{i}{n}\big)| i_k\, \mu^{n,m,k}(i) \hat{X}^{n,m}_s(i)^2 > \gamma\big) = 0\,.$$

Therefore, by (3.45), in distribution in $\mathbb{D}(\mathbb{R}_+,\mathbb{R})$,

$$Z^{n,m}_s \to Z_s\,, \tag{3.52}$$

where

$$\begin{aligned} Z_s = \int\limits_{\mathbb{R}^\ell_+} \partial_s a_s(u)\tilde{\lambda}_s(du) + \sum_{k=1}^{\ell}\Big(\int\limits_{\mathbb{R}^\ell_+} \partial_{u_k} a_s(u)\, u_k(\hat{\beta}^k(u) - \hat{\delta}^k(u) - \hat{\mu}^k(u))\,\tilde{\lambda}_s(du) \\ + \frac{\int\limits_{\mathbb{R}^\ell_+} u_k\,\hat{\mu}^k(u)\,\tilde{\lambda}_s(du)}{\tilde{\lambda}_s(\mathbb{R}^\ell_+)} \int\limits_{\mathbb{R}^\ell_+} \partial_{u_k} a_s(u)\,\tilde{\lambda}_s(du)\Big) \\ + \int\limits_{\mathbb{R}^\ell_+}\big(\int\limits_{\mathbb{R}^\ell_+} a_s(u')\hat{\varphi}(u,du') - a_s(u)\hat{\phi}(u) - \tilde{\lambda}_s(\mathbb{R}^\ell_+) a_s(u)\,\hat{\epsilon}(u)\big)\tilde{\lambda}_s(du)\,. \end{aligned}$$

Since Z_s has continuous trajectories owing to Lemmas 3.2 and 3.3, by (3.44), (3.46), (3.47), and (3.52),

$$Y_t = Y_0 + \int\limits_0^t Z_s\, ds\,.$$

By continuity of Z_s, Y_t is continuously differentiable with respect to t and

$$\begin{aligned} \frac{d}{dt}\int\limits_{\mathbb{R}^\ell_+} a_t(u)\,\tilde{\lambda}_t(du) = \int\limits_{\mathbb{R}^\ell_+} \partial_t a_t(u)\tilde{\lambda}_t(du) \\ + \sum_{k=1}^{\ell}\Big(\int\limits_{\mathbb{R}^\ell_+} \partial_{u_k} a_t(u) u_k(\hat{\beta}^k(u) - \hat{\delta}^k(u) - \hat{\mu}^k(u))\tilde{\lambda}_t(du) \\ + \frac{\int\limits_{\mathbb{R}^\ell_+} u_k\,\hat{\mu}^k(u)\,\tilde{\lambda}_t(du)}{\tilde{\lambda}_t(\mathbb{R}^\ell_+)} \int\limits_{\mathbb{R}^\ell_+} \partial_{u_k} a_t(u)\tilde{\lambda}_t(du)\Big) \end{aligned}$$

$$+\int_{\mathbb{R}_+^\ell}\Big(\int_{\mathbb{R}_+^\ell} a_t(u')\hat\varphi(u,du') - a_t(u)\hat\phi(u) - \tilde\lambda_t(\mathbb{R}_+^\ell)a_t(u)\,\hat\epsilon(u)\Big)\tilde\lambda_t(du)\,. \quad (3.53)$$

We now prove that $\tilde\lambda_t$ is specified uniquely. Given $\nu\in\mathbb{M}_+(\mathbb{R}_+^\ell)$ such that $\nu(\mathbb{R}_+^\ell)>0$, we define, for $y\in\mathbb{C}_c^1(\mathbb{R}_+^\ell)$,

$$A(\nu)y(u) \;=\; \sum_{k=1}^{\ell}\Big(u_k(\hat\beta^k(u)\,-\,\hat\delta^k(u)\,-\,\hat\mu^k(u))\,+\,\frac{\int_{\mathbb{R}_+^\ell} u_k'\,\hat\mu^k(u')\nu(du')}{\nu(\mathbb{R}_+^\ell)}\Big)\partial_{u_k}y(u)\,.$$

We also let, for bounded functions y,

$$B(\nu)y(u) = \int_{\mathbb{R}_+^\ell} y(u')\hat\varphi(u,du') - \big(\hat\phi(u)+\hat\epsilon(u)\nu(\mathbb{R}_+^\ell)\big)y(u)\,. \quad (3.54)$$

We write (3.53) as

$$\frac{d}{ds}\,\langle a_s,\tilde\lambda_s\rangle = \langle\partial_s a_s + (A(\tilde\lambda_s)+B(\tilde\lambda_s))a_s,\tilde\lambda_s\rangle\,. \quad (3.55)$$

(In the rest of the section, $\langle\cdot,\cdot\rangle$ represents the pairing between $\mathbb{L}^\infty(\mathbb{R}_+^\ell)$ and $\mathbb{M}(\mathbb{R}_+^\ell)$, with $\mathbb{M}(\mathbb{R}_+^\ell)$ denoting the set of signed Borel measures on $\mathbb{R}_+^\ell$ with the total variation norm.) Let

$$F_t^k(u) = u_k(\hat\beta^k(u)-\hat\delta^k(u)-\hat\mu^k(u)) + \frac{1}{\tilde\lambda_t(\mathbb{R}_+^\ell)}\int_{\mathbb{R}_+^\ell} u_k'\,\hat\mu^k(u')\,\tilde\lambda_t(du')\,, \quad (3.56)$$

$F_t(u) = (F_t^1(u),\ldots,F_t^\ell(u))$ and let $\psi_{s,t}(u) = \big(\psi_{s,t}^1(u),\ldots,\psi_{s,t}^\ell(u)\big)$, where $s\in\mathbb{R}$ and $t\in\mathbb{R}$, be defined by $\psi_{s,s}(u)=u$ and by

$$\partial_t\psi_{s,t}(u) = F_t(\psi_{s,t}(u))\,. \quad (3.57)$$

Since $u_k\hat\beta^k(u)$, $u_k\hat\delta^k(u)$, and $u_k\hat\mu^k(u)$ are $\mathbb{C}^1$–functions with bounded derivatives and the ratio on the righthand side of (3.56) is a continuous function of t, the function $\psi_{s,t}(u)$ is continuously differentiable in (s,t,u) with the u–derivatives being uniformly bounded locally uniformly in (s,t), see, e.g., Theorem 3.1 on p.95 in Hartman [9]. Let, for $f\in\mathbb{R}^{\mathbb{R}_+^\ell}$,

$$U_{s,t}f(u) = f(\psi_{s,t}(u))\,, \quad (3.58)$$

where $u \in \mathbb{R}_+^\ell$. For $f \in \mathbb{C}_c^1(\mathbb{R}_+^\ell)$, $U_{s,t}f(u)$ is continuously differentiable in (s,t,u) and

$$\partial_s U_{s,t} f = -A(\tilde{\lambda}_s) U_{s,t} f\,. \tag{3.59}$$

the temporal derivatives on the lefthand side being for the sup–norm. (One way to ascertain the equation is to use the flow property that $\psi_{s,t} = \psi_{r,t} \circ \psi_{s,r}$.) We now draw on Luo and Mattingly [17] by letting in (3.55), for t fixed, $a_s = U_{s,t} f$, where $f \in \mathbb{C}_c^1(\mathbb{R}_+^\ell)$. By (3.58) and (3.59), $\partial_s a_s = -A(\tilde{\lambda}_s) a_s$, $a_t = f$, and $a_0 = U_{0,t} f$. By (3.55) and the fact that $\tilde{\lambda}_0 = \hat{\lambda}_0$,

$$\langle f, \tilde{\lambda}_t \rangle = \langle U_{0,t} f, \hat{\lambda}_0 \rangle + \int_0^t \langle B(\tilde{\lambda}_s) U_{s,t} f, \tilde{\lambda}_s \rangle\, ds\,. \tag{3.60}$$

Via limits akin to the Daniell integral, this equality extends to bounded Borel functions f. Let $\breve{\lambda}_t$ represent another limit point of $\Lambda^{n,m}$, so, it solves (3.60) as well. Since $\tilde{\lambda}_0 = \breve{\lambda}_0$, we have that

$$\langle f, \tilde{\lambda}_t - \breve{\lambda}_t \rangle = \int_0^t \langle B(\tilde{\lambda}_s) U_{s,t} f, \tilde{\lambda}_s - \breve{\lambda}_s \rangle ds + \int_0^t \langle (B(\tilde{\lambda}_s) - B(\breve{\lambda}_s)) U_{s,t} f, \breve{\lambda}_s \rangle ds\,.$$

By (3.54), by $U_{s,t}$ being a contraction for the sup–norm and by (3.60), on recalling that $\sup_{u \in \mathbb{R}_+^\ell} \hat{\varphi}(u, \mathbb{R}_+^\ell) < \infty$, given $T > 0$, there exists $K > 0$ such that, for all $t \le T$,

$$\|\tilde{\lambda}_t - \breve{\lambda}_t\| \le K \int_0^t \|\tilde{\lambda}_s - \breve{\lambda}_s\|\, ds\,,$$

with $\|\cdot\|$ representing the total variation norm on $\mathbb{M}(\mathbb{R}_+^\ell)$. By Gronwall's inequality, $\tilde{\lambda}_t = \breve{\lambda}_t$. Thus, (3.53) has a unique solution, which concludes the convergence proof. We have also proved that $\hat{\lambda}_t$ is specified uniquely by (3.53) which is the same equation as (2.8).

Let us prove that if $\hat{\lambda}_0$ admits a density with respect to Lebesgue measure, then $\hat{\lambda}_t$ does too. Let $\mathcal{N}$ denote the set of Lebesgue measurable functions on $\mathbb{R}_+^\ell$ that are not greater than one in absolute value and are equal to zero a.e. with respect to the Lebesgue measure. Let $f \in \mathcal{N}$. By the uniqueness of solutions to (3.57), $\psi_{s,t}^{-1}(u)$ is well defined, satisfies the initial condition $\psi_{s,s}^{-1}(u) = u$ and the version of (3.57) in reverse time

$$\partial_t \psi_{s,t}^{-1}(u) = -F_t(\psi_{s,t}^{-1}(u))\,.$$

Therefore, $\psi_{s,t}^{-1}(u)$ is of class C^1 in u. Since $\{u:\, U_{s,t}f(u) \ne 0\} = \psi_{s,t}^{-1}\{u:\, f(u) \ne 0\}$ and sets of Lebesgue measure zero are preserved under C^1–maps, see, e.g., Lemma

1.1 on p.68 in Hirsch [11], $U_{s,t}f = 0$ a.e. As $\hat{\lambda}_0$ admits a density, $\langle U_{0,t}f, \hat{\lambda}_0\rangle = 0$. By (3.60), for some $K > 0$,

$$\operatorname*{ess\,sup}_{f\in\mathcal{N}} |\langle f, \hat{\lambda}_t\rangle| \le K \int_0^t \operatorname*{ess\,sup}_{f\in\mathcal{N}} |\langle U_{s,t}f, \hat{\lambda}_s\rangle|\, ds \le K \int_0^t \operatorname*{ess\,sup}_{f\in\mathcal{N}} |\langle f, \hat{\lambda}_s\rangle|\, ds\,.$$

By Gronwall's inequality, $|\langle f, \hat{\lambda}_s\rangle| = 0$ when $f \in \mathcal{N}$, so, $\hat{\lambda}_s$ has a density which we denote by $\hat{x}_s(u)$. Part 1 has been proved.

We prove part 2. Since the $\beta^k(u)$ and $\delta^k(u)$ are bounded and $\sup_{s\le t}\int_{\mathbb{R}_+^\ell} |u|\hat{\lambda}_s(du) < \infty$, for all $t > 0$, by approximation, (2.8) holds for $f(u) = |u|$ so that

$$\frac{d}{dt}\int_{\mathbb{R}_+^\ell} |u|\hat{\lambda}_t(du) = \int_{\mathbb{R}_+^\ell}\Bigl(u\cdot\bigl(\hat{\beta}(u) - \hat{\delta}(u)\bigr) + \int_{\mathbb{R}_+^\ell} |u'|\hat{\varphi}(u, du') \\ - |u|\bigl(\hat{\phi}(u) + \hat{\epsilon}(u)\hat{\lambda}_t(\mathbb{R}_+^\ell)\bigr)\Bigr)\hat{\lambda}_t(du)\,. \quad (3.61)$$

By (3.4), $\int_{\mathbb{R}_+^\ell} |u'|\hat{\varphi}(u, du') = |u|\hat{\phi}(u)$. Substitution in (3.61) yields

$$\frac{d}{dt}\int_{\mathbb{R}_+^\ell} |u|\,\hat{\lambda}_t(du) = \int_{\mathbb{R}_+^\ell} u\cdot(\hat{\beta}(u) - \hat{\delta}(u))\hat{\lambda}_t(du) - \hat{\lambda}_t(\mathbb{R}_+^\ell)\int_{\mathbb{R}_+^\ell} |u|\,\hat{\epsilon}(u)\hat{\lambda}_t(du)\,.$$

On the other hand, (3.23) can be written as

$$\int_{\mathbb{R}_+^\ell} |u|\Lambda_t^{n,m}(du) = \int_{\mathbb{R}_+^\ell} |u|\Lambda_0^{n,m}(du) + \int_0^t\int_{\mathbb{R}_+^\ell} u\cdot\bigl(\hat{\beta}^{n,m}(u) - \hat{\delta}^{n,m}(u)\bigr)\Lambda_s^{n,m}(du)\,ds \\ - \int_0^t \Lambda_s^{n,m}(\mathbb{R}_+^\ell)\int_{\mathbb{R}_+^\ell} |u|\hat{\epsilon}^{n,m}(u)\Lambda_s^{n,m}(du)\,ds + \frac{1}{mn}\,\overline{N}_t^{n,m}\,,$$

where $\sup_{t\le L}|\overline{N}_t^{n,m}|/(mn) \to 0$ in probability, as $m, n \to \infty$, the latter convergence being proved in analogy with (3.26). By Lemma 3.2 and Lemma 3.3, in probability,

$$\int_{\mathbb{R}_+^\ell} |u|\Lambda_t^{n,m}(du) \to \int_{\mathbb{R}_+^\ell} |u|\hat{\lambda}_0(du) + \int_0^t\int_{\mathbb{R}_+^\ell} u\cdot(\hat{\beta}(u) - \hat{\delta}(u))\hat{\lambda}_s(du)\,ds \\ - \int_0^t \hat{\lambda}_s(\mathbb{R}_+^\ell)\int_{\mathbb{R}_+^\ell} |u|\hat{\epsilon}(u)\hat{\lambda}_s(du)\,ds\,.$$

Therefore, in probability in $\mathbb{D}(\mathbb{R}_+,\mathbb{R})$,

$$\int_{\mathbb{R}_+^\ell} |u|\Lambda_t^{n,m}(du) \to \int_{\mathbb{R}_+^\ell} |u|\hat{\lambda}_t(du)\,.$$

By Lemma 3.3, the convergence holds locally uniformly in t . The other assertion of part 2 is proved similarly. Part 2 has been proved.

We address now the regularity properties of $\hat{x}_s$, so, we assume the hypotheses of part 3 of the theorem to hold. Since $u \to \psi_{s,t}(u)$ is a diffeomorphism, given $f \in \mathbb{L}^\infty(\mathbb{R}_+^\ell)$ and $z \in \mathbb{L}^1(\mathbb{R}_+^\ell)$, by a change of variables, see, e.g., Theorem 2.6 on p.505 in Lang [14], $\langle U_{s,t}f, z\rangle_1 = \langle f, z\circ\psi_{s,t}^{-1}\, J(\psi_{s,t}^{-1})\rangle_1$, where $\langle\cdot,\cdot\rangle_1$ represents the pairing between $\mathbb{L}^\infty$ and $\mathbb{L}^1$ and $J(\psi_{s,t}^{-1})$ denotes the absolute value of the Jacobian determinant of $\psi_{s,t}^{-1}$. Let us denote

$$U_{s,t}^* z = z\circ\psi_{s,t}^{-1}\, J(\psi_{s,t}^{-1})\,. \tag{3.62}$$

It is a bounded operator on $\mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$, with norms being bounded locally uniformly, cf., Proposition 9.6 on p.270 in Brezis [1]. (The u–derivatives of $\psi_{s,t}^{-1}(u)$ are bounded locally uniformly in s and t .) By (3.54), the $B(\hat{\lambda}_s)$ are bounded operators on $\mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$ too, with locally uniformly bounded norms.

By (3.60), for $f \in \mathbb{L}^\infty(\mathbb{R}_+^\ell)$,

$$\langle f, \hat{x}_t\rangle_1 = \langle f, U_{0,t}^*\hat{x}_0\rangle_1 + \int_0^t \langle f, U_{s,t}^* B(\hat{\lambda}_s)\hat{x}_s\rangle_1\, ds\,,$$

so, for almost all u ,

$$\hat{x}_t(u) = U_{0,t}^*\hat{x}_0(u) + \int_0^t U_{s,t}^* B(\hat{\lambda}_s)\hat{x}_s(u)\, ds\,. \tag{3.63}$$

We can therefore redefine $\hat{x}_t(u)$ as the latter righthand side, which makes it a continuous function of t for all $u \in \mathbb{R}_+^\ell$. Continuity of $\hat{\lambda}_t(\mathbb{R}_+)$ implies, as in the proof of Scheffe's theorem, that $\hat{x}_t$ is continuous in t in $\mathbb{L}^1(\mathbb{R}_+^\ell)$. In particular, $\int_0^t U_{s,t}^* B(\hat{\lambda}_s)\hat{x}_s\, ds$ is well defined as a Riemann integral in $\mathbb{L}^1(\mathbb{R}^\ell)$ and

$$\hat{x}_t = U_{0,t}^*\hat{x}_0 + \int_0^t U_{s,t}^* B(\hat{\lambda}_s)\hat{x}_s\, ds\,. \tag{3.64}$$

As in the proof of Lemma 4.5 on p.142 in Pazy [19], see also Theorem 9.19 on p.488 in Engel and Nagel [5], there exists family V_t of bounded linear operators on $\mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$ such that, for $z \in \mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$,

$$V_t z = U_{0,t}^* z + \int_0^t U_{s,t}^* B(\hat{\lambda}_s) V_s z\, ds\,. \tag{3.65}$$

Specifically, one defines

$$V_t^{(0)} z = U_{0,t}^* z\,, \quad V_t^{(m)} z = \int_0^t U_{s,t}^* B(\hat{\lambda}_s) V_s^{(m-1)} z\, ds$$

and

$$V_t z = \sum_{m=0}^{\infty} V_t^{(m)} z\,,$$

the convergence holding in $\mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$ because, as induction shows, $\|V_t^{(m)}\| \le K_1 K_2^m t^m/m!$, where K_1 is an upper bound for $\|U_{0,t}^*\|$ and K_2 is an upper bound for $\|U_{s,t}^*\| \|B(\hat{\lambda}_s)\|$. Let $\overline{x}_t = V_t \hat{x}_0$. By (3.64), (3.65), and the uniqueness of $\hat{\lambda}_t$, $\overline{x}_t = \hat{x}_t$ as elements of $\mathbb{L}^1(\mathbb{R}_+^\ell)$, so $\hat{x}_t \in \mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$. By (3.62), $U_{s,t}^* z(u)$ is Lipschitz–continuous with respect to t, provided $z \in \mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$. By (3.63), $\hat{x}_t(u)$ is Lipschitz–continuous with respect to t. Now (2.9) is obtained from (2.8) via integration by parts. Part 3 has been proved.

4 Proof of Theorem 2.2

The proof proceeds along similar lines as the one of Theorem 2.1. We give the main points. Let, in analogy with (3.9),

$$\check{R}_t^{n,m} = \frac{1}{mn^\ell} \sum_i X_t^{n,m}(i) = \int_{\mathbb{R}_+^\ell} \hat{X}_t^{n,m}(u)\, du\,.$$

The next three lemmas are proved similarly to Lemma 3.2, Lemma 3.3, and Lemma 3.4, respectively.

Lemma 4.1. *The sequence $(\check{R}_t^{n,m}, t \ge 0)$ is C–tight and, given $t > 0$, there exists $\rho > 0$ such that $\mathbf{P}(\inf_{s \le t} \check{R}_s^{n,m} > \rho) \to 1$, as $n, m \to \infty$.*

Lemma 4.2. *The sequence* $\left(\int_{\mathbb{R}_+^\ell}|u|\hat{X}_t^{n,m}(u)\,du\,,t\geq 0\right)$ *is* C*–tight.*

Lemma 4.3. *For all* $L>0$*,*

$$\lim_{K\to\infty}\limsup_{n,m\to\infty}\mathbf{P}\big(\sup_{t\in[0,L]}\sum_i|X_t^{n,m}(i)|^2>Km^2n^\ell\big)=0\,.$$

In analogy with (3.44),

$$\check{Y}_t^{n,m}=\check{Y}_0^{n,m}+\int_0^t\check{Z}_s^{n,m}\,ds+\check{N}_t'^{n,m}\,, \tag{4.1}$$

where

$$\check{Y}_t^{n,m}=\int_{\mathbb{R}_+^\ell}a_t^n(u)\hat{X}_t^{n,m}(u)\,du\,, \tag{4.2}$$

$$\begin{aligned}\check{Z}_s^{n,m}=\int_{\mathbb{R}_+^\ell}\Big(\partial_sa_s^n(u)+\sum_{k=1}^\ell\Big(n\big(a_s^n\big(u+\frac{e_k}{n}\big)-a_s^n(u)\big)\,\hat{X}_s^{n,m}(u)\,\frac{\lfloor nu_k\rfloor}{n}\,\hat{\beta}^{n,m,k}(u)\\+n\big(a_s^n\big(u-\frac{e_k}{n}\big)-a_s^n(u)\big)\,\hat{X}_s^{n,m}(u)\,\frac{\lfloor nu_k\rfloor}{n}\,(\hat{\delta}^{n,m,k}(u)+\hat{\mu}^{n,m,k}(u))\\+n\big(a_s^n\big(u+\frac{e_k}{n}\big)-a_s^n(u)\big)\hat{X}_s^{n,m}(u)\int_{\mathbb{R}_+^\ell}\frac{\hat{X}_s^{n,m}(u')}{\check{R}_s^{n,m}}\,\frac{\lfloor nu_k'\rfloor}{n}\,\hat{\mu}^{n,m,k}(u')\,du'\Big)\\+\hat{X}_s^{n,m}(u)\int_{\mathbb{R}_+^\ell}a_s^n(u')\hat{\phi}^{n,m}(u)\check{\eta}^{n,m}(u,u')\,du'-a_s^n(u)\,\hat{X}_s^{n,m}(u)\,\big(\hat{\phi}^{n,m}(u)+\check{R}_s^{n,m}\check{\epsilon}^{n,m}(u)\big)\Big)\,du\\+\int_0^t\int_{\mathbb{R}_+^\ell}\hat{X}_s^{n,m}(u)\int_{B_{1/n}(\lfloor nu\rfloor/(2n))}a_s^n(u')\hat{\phi}^{n,m}(u)\check{\eta}^{n,m}(u,u')\,du'\,du\\-\sum_i\sum_{k=1}^\ell\big(a_s\big(\frac{i+e_k}{n}\big)-a_s\big(\frac{i}{n}\big)\big)\frac{1}{R_s^{n,m}}\,i_k\mu^{n,m,k}(i)\frac{X_s^{n,m}(i)^2}{m^2n^\ell}\end{aligned}$$

and $(\check{N}_t'^{n,m}\,,t\geq 0)$ is a locally square integrable martingale with the predictable quadratic variation process, cf. (3.41),

$$\langle \check{N}'^{n,m}\rangle_t = \frac{1}{m}\int\limits_0^t \int\limits_{\mathbb{R}_+^\ell} \Big(\sum_{k=1}^{\ell}\Big(\frac{1}{n^{\ell-1}}\,\big(a_t^n\big(u+\frac{e_k}{n}\big)-a_t^n(u)\big)^2\hat{X}_s^{n,m}(u)\frac{\lfloor nu_k\rfloor}{n}\,\hat{\beta}^{n,m,k}(u)$$

$$+\frac{1}{n^{\ell-1}}\,\big(a_t^n\big(u-\frac{e_k}{n}\big)-a_t^n(u)\big)^2\hat{X}_s^{n,m}(u)\frac{\lfloor nu_k\rfloor}{n}\,\hat{\delta}^{n,m,k}(u)$$

$$+\,n\big(a_t^n\big(u+\frac{e_k}{n}\big)-a_t^n(u)\big)^2\hat{X}_s^{n,m}(u)\int\limits_{\mathbb{R}_+^\ell\setminus B_{1/n}(\lfloor nu\rfloor/n)}\frac{\hat{X}_s^{n,m}(u')}{\check{R}_s^{n,m}}\,\frac{\lfloor nu'_k\rfloor}{n}\,\hat{\mu}^{n,m,k}(u')\,du'$$

$$+\frac{1}{n^{\ell-1}}\,\big(a_t^n\big(u-\frac{e_k}{n}\big)-a_t^n(u)\big)^2\hat{X}_s^{n,m}(u)\frac{\lfloor nu_k\rfloor}{n}\,\hat{\mu}^{n,m,k}(u)\big(1-\frac{\hat{X}_s^{n,m}(u)}{n^\ell\check{R}_s^{n,m}}\big)$$

$$+\,n\big(a_t^n\big(u+\frac{e_k}{n}\big)-a_t^n(u)\big)\frac{\hat{X}_s^{n,m}(u)}{n^\ell\check{R}_s^{n,m}}$$

$$\int\limits_{\mathbb{R}_+^\ell\setminus B_{1/n}(\lfloor nu\rfloor/n)}\big(a_t^n\big(u'-\frac{e_k}{n}\big)-a_t^n(u')\big)\hat{X}_s^{n,m}(u')\frac{\lfloor nu'_k\rfloor}{n}\,\hat{\mu}^{n,m,k}(u')\,du'\Big)$$

$$+\frac{1}{n^\ell}\,\hat{\phi}^{n,m}(u)\hat{X}_s^{n,m}(u)\int\limits_{\mathbb{R}_+^\ell}a_t^n(u')^2\check{\alpha}^{n,m}(u,u')\,du'$$

$$+\frac{1}{n^\ell}\,a_t^n(u)^2\hat{X}_s^{n,m}(u)\big(\check{R}_s^{n,m}\check{\epsilon}^{n,m}(u)+\hat{\phi}^{n,m}(u)\big)$$

$$+\frac{1}{n^\ell}\,\hat{X}_s^{n,m}(u)\hat{\phi}^{n,m}(u)\mathbf{E}\big(\sum_{i'}a_t\big(\frac{i'}{n}\big)\theta_{\lfloor nu\rfloor}^{n,m}(i',1)\big)^2$$

$$-\frac{1}{n^\ell}\,a_t^n(u)\hat{X}_s^{n,m}(u)\hat{\phi}^{n,m}(u)\int\limits_{\mathbb{R}_+^\ell}a_t^n(u')\check{\eta}^{n,m}(u,u')\,du'\Big)du\,ds\,,$$

where $\check{\alpha}^{n,m}(u,u') = n^\ell\alpha_{\lfloor nu\rfloor}^{n,m}(\lfloor nu'\rfloor)$. By (3.6), $\check{\alpha}^{n,m}(u,u') \le b\check{\eta}^{n,m}(u,u')$. Besides, $\mathbf{E}\big(\sum_{i'}a_t(i'/n)\theta_{\lfloor nu\rfloor}^{n,m}(i',1)\big)^2 \le b\sup_{t,u}a_t(u)^2\int_{\mathbb{R}_+^\ell}\check{\eta}^{n,m}(u,u')\,du'$. By Lemma 4.1, Lemma 4.2 and (3.1), $\langle\check{N}'^{n,m}\rangle_t \to 0$ in probability, as $n,m\to\infty$. Therefore, for $\gamma > 0$,

$$\lim_{n,m\to\infty}\mathbf{P}(\sup_{t\le L}|\check{N}_t'^{n,m}| > \gamma) = 0\,. \tag{4.3}$$

The following analogue of Lemma 3.5 holds.

Lemma 4.4. *The sequence $\hat{X}^{n,m}$ is C–tight in $\mathbb{D}(\mathbb{R}_+,\mathbb{L}_w^1(\mathbb{R}_+^\ell))$ and in $\mathbb{D}(\mathbb{R}_+,\mathbb{L}_w^2(\mathbb{R}_+^\ell))$.*

Proof. We prove, first, the C–tightness in $\mathbb{D}(\mathbb{R}_+, \mathbb{L}^2_w(\mathbb{R}^\ell_+))$. Taking as F in Theorem 4.6 in Jakubowski [13] the set of functions $f \to \int_0^L g(u)f(u)\,du$, where $L > 0$, $f \in \mathbb{L}^2(\mathbb{R}^\ell_+)$ and $g(u)$ is continuously differentiable of compact support, it is sufficient to prove that, for all $L > 0$,

$$\lim_{K\to\infty} \limsup_{n,m\to\infty} \mathbf{P}\Big(\sup_{t\in[0,L]} \int_{\mathbb{R}^\ell_+} |\hat{X}^{n,m}_t(u)|^2\,du > K\Big) = 0 \tag{4.4}$$

and, for all differentiable g of compact support and all $\gamma > 0$,

$$\lim_{\chi\to 0} \limsup_{n,m\to\infty} \mathbf{P}\Big(\sup_{\substack{s,t\in[0,L]:\\ |s-t|\le\chi}} \Big| \int_{\mathbb{R}^\ell_+} g(u)(\hat{X}^n_s(u) - \hat{X}^n_t(u))\,du\Big| > \gamma\Big) = 0\,. \tag{4.5}$$

The convergence in (4.4) is the statement of Lemma 4.3. Similarly to the proof of Lemma 3.5, (4.5) follows from (4.1), (4.2), (2.5), (2.10), and (4.3).

Since $\hat{X}^{n,m}$ is relatively compact for convergence in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{L}^2_w(\mathbb{R}^\ell_+))$ the relative compactness of $\hat{X}^{n,m}$ for convergence in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{L}^1_w(\mathbb{R}^\ell_+))$ holds if, for any $\gamma > 0$,

$$\lim_{K\to\infty} \limsup_{n,m\to\infty} \mathbf{P}\Big(\sup_{t\le L} \int_{\mathbb{R}^\ell_+} \mathbf{1}_{[K,\infty)}(|u|)\hat{X}^{n,m}_t(u)\,du > \gamma\Big) = 0\,,$$

which follows from Lemma 4.2. □

Let $(\check{x}_s\,, s \ge 0)$ represent a limit point of $\hat{X}^{n,m}$ for both topologies. We identify it in a similar fashion to that in the proof of Theorem 2.1. As $n, m \to \infty$, along a subnet, in distribution in $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$,

$$(\check{Y}^{n,m}_s, \check{Z}^{n,m}_s) \to (\check{Y}_s, \check{Z}_s)\,,$$

where $\check{Y}_s = \int_{\mathbb{R}^\ell_+} a_s(u)\check{x}_s(u)\,du$ and

$$\check{Z}_s = \int_{\mathbb{R}^\ell_+} \partial_s a_s(u)\check{x}_s(u)\,du + \sum_{k=1}^{\ell}\Big(\int_{\mathbb{R}^\ell_+} \partial_{u_k} a_s(u)\,\check{x}_s(u)u_k(\hat{\beta}^k(u) - \hat{\delta}^k(u) - \hat{\mu}^k(u))\,du + \frac{\int_{\mathbb{R}^\ell_+} \check{x}_s(u)u_k\,\hat{\mu}^k(u)\,du}{\int_{\mathbb{R}^\ell_+} \check{x}_s(u)\,du} \int_{\mathbb{R}^\ell_+} \partial_{u_k} a_s(u)\check{x}_s(u)\,du\Big)$$

$$+\int_{\mathbb{R}_+^\ell} a_s(u) \int_{\mathbb{R}_+^\ell} \check{x}_s(u')\check{\varphi}(u',u)\,du'\,du - \int_{\mathbb{R}_+^\ell} a_s(u)\,\check{x}_s(u)\hat{\phi}(u)\,du$$
$$- \int_{\mathbb{R}_+^\ell} \check{x}_s(u)\,du \int_{\mathbb{R}_+^\ell} a_s(u)\,\check{x}_s(u)\check{\epsilon}(u)\,du\,.$$

Since $\check{Z}_s$ has continuous trajectories, by (3.44), (3.46), (3.47), and (3.52),

$$\frac{d}{dt}\int_{\mathbb{R}_+^\ell} a_t(u)\check{x}_t(u)\,du = \check{Z}_t\,. \tag{4.6}$$

Since $(\check{x}_t,\,t \geq 0) \in \mathbb{C}(\mathbb{R}_+, \mathbb{L}^2(\mathbb{R}_+^\ell) \cap \mathbb{L}^1(\mathbb{R}_+^\ell))$ and the functions $u_k\hat{\beta}^k(u)$, $u_k\hat{\delta}^k(u)$, and $u_k\hat{\mu}^k(u)$ are bounded, (4.6) holds for $(a_t\,, t \geq 0) \in \mathbb{C}(\mathbb{R}_+, \mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)+\mathbb{W}^{1,2}(\mathbb{R}_+^\ell))$. One next introduces $U_{s,t}$ by (3.57) and (3.58). It is an operator on $\mathbb{L}^\infty(\mathbb{R}_+^\ell)$. Since the functions $u_k\hat{\beta}^k(u)$, $u_k\hat{\delta}^k(u)$, and $u_k\hat{\mu}^k(u)$ are bounded and Lipschitz–continuous, by (3.57) and Theorem 3.1.1 on p.76 in Hille [10], the functions $\psi_{s,t}(u)$ are Lipschitz–continuous with respect to u. Similarly, $\psi_{s,t}^{-1}(u)$ is Lipschitz–continuous as well. By the argument of the proof of Proposition 9.6 on p.270 in Brezis [1], see also Problem 7.5 on p.174 in Gilbarg and Trudinger [7], $\mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$ is invariant under $U_{t,s}$. Furthermore, $U_{t,s}$ is a bounded operator on $\mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$ and, given $f \in \mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$, $U_{t,s}f(u)$ is an absolutely continuous function of s and $\partial_s U_{s,t}f(u) = -A(\tilde{\lambda}_s)U_{s,t}f(u)$ a.e. As in the proof of Theorem 2.1, one lets $a_s = U_{s,t}f$ in (4.6) and obtains (3.60) with $\tilde{\lambda}_s(du) = \check{x}_s(u)\,du$. As before, that implies that $\tilde{\lambda}_s$ is specified uniquely, so, $\check{x}_s(u)$ is specified uniquely for almost all u with respect to the Lebesgue measure.

Let $\hat{x}_0 \in \mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$, $\check{\varphi}(u,u')$ have Sobolev derivative $D_u\check{\varphi}(u,u')$ with respect to u for almost all u' such that $\operatorname{ess\,sup}_{u\in\mathbb{R}_+^\ell} \int_{\mathbb{R}_+^\ell} (\check{\varphi}(u,u') + |D_u\check{\varphi}(u,u')|)\,du' < \infty$, and $\hat{\phi}(u)$ and $\check{\epsilon}(u)$ be elements of $\mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$. The u_k–derivatives of $\psi_{t,s}^{-1}(u)$ being bounded implies that $U_{s,t}^*$ is well defined and is a bounded operator on $\mathbb{W}^{1,\infty}(\mathbb{R}_+^\ell)$. In analogy with (3.63), one obtains that

$$\check{x}_t = U_{0,t}^*\hat{x}_0 + \int_0^t U_{s,t}^* B(\hat{\lambda}_s)\check{x}_s\,ds\,.$$

The rest of the proof mimics the proof of Theorem 2.1.

A Appendix

Lemma A.1. *Let $N = (N_t\,, t \geq 0)$ be a Poisson process of rate λ adapted to filtration $\mathbf{F} = (\mathcal{F}_t\,, t \geq 0)$. Let process $X = (X_t\,, t \geq 0)$ be adapted to $\mathbf{F}$ as well. Let $\xi(k,x)$*

be bounded random variables that are independent of $\mathcal{F}_{\tau_k-}$, with τ_k denoting the kth jump of N_t. Let $Y_t = \int_0^t \xi(N_s, X_{s-})\, dN_s$ and let $\mathbf{G} = (\mathcal{G}_t\,, t \geq 0)$ be the smallest filtration that contains $\mathbf{F}$ such that $Y = (Y_t\,, t \geq 0)$ is $\mathbf{G}$–adapted. Then the process Y has $\mathbf{G}$–compensator $\tilde{Y}_t = \int_0^t \mathbf{E}\xi(N_s+1, x)|_{x=X_s}\, \lambda\, ds$. The locally square integrable martingale $(Y_t - \tilde{Y}_t\,, t \geq 0)$ has the process $\big(\int_0^t \mathbf{E}\xi(N_s + 1, x)^2|_{x=X_s}\, \lambda\, ds\,, t \geq 0\big)$ as the predictable quadratic variation process.

Proof. We prove the first claim. It suffices to show that, for arbitrary $\mathbf{G}$–stopping time τ, $\mathbf{E}\int_0^\tau \xi(N_s, X_{s-})\, dN_s = \mathbf{E}\int_0^\tau \mathbf{E}\xi(N_s + 1, x)|_{x=X_s}\, \lambda\, ds$. Since $\{\tau_k \leq \tau\} \in \mathcal{G}_{\tau_k-}$, the latter σ–algebra and the $\xi(k, x)$ are independent and X_{τ_k-} is $\mathcal{G}_{\tau_k-}$–measurable,

$$\mathbf{E}\int_0^\tau \xi(N_s, X_{s-})\, dN_s = \sum_k \mathbf{E}\, \mathbf{1}_{\{\tau_k \leq \tau\}}\, \xi(k, X_{\tau_k-}) = \sum_k \mathbf{E}\big(\, \mathbf{1}_{\{\tau_k \leq \tau\}}\, \mathbf{E}(\xi(k, X_{\tau_k-})|\mathcal{G}_{\tau_k-})\big)$$

$$= \sum_k \mathbf{E}\big(\, \mathbf{1}_{\{\tau_k \leq \tau\}}\, \mathbf{E}(\xi(k, x)|\mathcal{G}_{\tau_k-})|_{x=X_{\tau_k-}}\big) = \sum_k \mathbf{E}\big(\, \mathbf{1}_{\{\tau_k \leq \tau\}}\, \mathbf{E}(\xi(k, x)|_{x=X_{\tau_k-}})\big)$$

$$= \mathbf{E}\int_0^\tau \mathbf{E}\xi(N_{s-} + 1, x)|_{x=X_{s-}}\, dN_s = \mathbf{E}\int_0^\tau \mathbf{E}\xi(N_s + 1, x)|_{x=X_s}\, \lambda ds\,.$$

In order to prove the second claim, we write by the Itô formula for the locally square integrable martingale $M_t = Y_t - \tilde{Y}_t$,

$$M_t^2 = 2\int_0^t M_{s-}\, dM_s + \sum_{s \leq t} (\Delta M_s)^2 = 2\int_0^t M_{s-}\, dM_s + \int_0^t \xi(N_s, X_{s-})^2\, dN_s\,.$$

By the first part of the proof, the compensator of $\int_0^t \xi(N_s, X_{s-})^2\, dN_s$ is given by $\int_0^t \mathbf{E}\xi(N_s + 1, x)^2|_{x=X_s}\, \lambda ds$. □

References

[1] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, 2011.

[2] Champagnat, N., Ferriere, R., and Meleard, S. Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. *Theor. Popul. Biol.* 69, 297-321, 2006.

[3] Darwin, C. The Descent of Man and Selection in Relation to Sex. London: Murray, 1871.

[4] Dawkins, R. The Selfish Gene. Oxford University Press, New York, 1976

[5] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[6] S.N. Ethier and T.G. Kurtz. *Markov processes. Characterization and convergence.* Wiley, 1986.

[7] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order.* Springer-Verlag, Berlin, 2nd edition, 1998. Revised 3rd printing.

[8] Hamilton, W.D. The Genetical Evolution of Social Behavior. *Journal of Theoretical Biology.* 7, 1-16, 1964.

[9] P. Hartman. *Ordinary differential equations.* John Wiley & Sons, Inc., New York-London-Sydney, 1964.

[10] E. Hille. *Lectures on ordinary differential equations.* Addison-Wesley Publ. Co., Reading, Mass.-London-Don Mills, Ont., 1969.

[11] M.W. Hirsch. *Differential topology*, volume 33 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1994. Corrected reprint of the 1976 original.

[12] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, 1987.

[13] A. Jakubowski. On the Skorokhod topology. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(3):263–285, 1986.

[14] S. Lang. *Real and functional analysis*, volume 142 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, third edition, 1993.

[15] R.Sh. Liptser and A.N. Shiryayev. *Theory of martingales.* Kluwer, 1989.

[16] Luo, S. A Unifying Framework Reveals Key Properties of Multilevel Selection. *Journal of Theoretical Biology.* 341, 41-52, 2013.

[17] Luo, S., and Mattingly, J.C. Scaling limits of a model for selection at two scales. *Nonlinearity* 30(4): 1682–1707, 2017.

[18] Maynard-Smith, J. Group Selection and Kin Selection. *Nature.* 201, 1145-1146, 1964.

[19] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[20] Puhalskii, A. and Simon, B. Discrete evolutionary birth-death processes and their large-population limits. *Stochastic Models.* 28, 399-412, 2012.

[21] Simon, B. A Dynamical Model of Two-Level Selection. *Evolutionary Ecology Research.* 12, 555-588, 2010.

[22] Simon, B., Fletcher, J.A., and Doebeli, M. Towards a general theory of group selection. *Evolution.* 67, 1561-1572, 2013.

[23] Simon, B. and Nielsen, A. Numerical solutions and animations of group selection dynamics. *Evolutionary Ecology Research.* 14: 757-768, 2012.

[24] Simon, B. and Pilosov, M. Group-level events are catalysts in the evolution of cooperation. *Journal of Theoretical Biology.* 410, 125-136, 2016.

[25] Sober, E. and Wilson, D.S. Unto Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, Cambridge, MA., 1998

[26] F. Topsøe. Compactness in spaces of measures. *Studia Mathematica*, 36:195–221, 1970.

[27] F. Topsøe. *Topology and Measure*, volume 133 of *Lecture Notes in Mathematics*. Springer, 1970.

[28] Traulsen, A. and Nowak, M.A. Evolution of cooperation by multilevel selection. *Proc. National Academy of Sciences.* 103(29): 10952-10955, 2006.

[29] Williams, G.C. Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought. Princeton: Princeton University Press, 1966.

[30] Wilson, D.S. A Theory of Group Selection. *Proceedings of the National Academy of Sciences.* 72, 143-146, 1975.

[31] Wilson, D.S. and Wilson, E.O. Rethinking the theoretical foundations of sociobiology. *Q. Rev. Biol.*, 82 (4), 327-348, 2007.

[32] Wynne-Edwards, V.C. Animal Dispersion in Relation to Social Behavior. Edinburgh: Oliver and Boyd, 1962.