The Annals of Applied Probability
2020, Vol. 30, No. 5, 2311–2354
https://doi.org/10.1214/20-AAP1559

PROPAGATION OF CHAOS AND THE MANY-DEMES LIMIT FOR WEAKLY INTERACTING DIFFUSIONS IN THE SPARSE REGIME

BY MARTIN HUTZENTHALER[*] AND DANIEL PIEPER[†]

Fakultät für Mathematik, Universität Duisburg-Essen, [*]*martin.hutzenthaler@uni-due.de;* [†]*daniel.pieper@uni-due.de*

Propagation of chaos is a well-studied phenomenon and shows that weakly interacting diffusions may become independent as the system size converges to infinity. Most of the literature focuses on the case of exchangeable systems where all involved diffusions have the same distribution and are "of the same size". In this paper, we analyze the case where only a few diffusions start outside of an accessible trap. Our main result shows that in this "sparse regime" the system of weakly interacting diffusions converges in distribution to a forest of excursions from the trap. In particular, initial independence propagates in the limit and results in a forest of independent trees.

1. Introduction. The notion of "propagation of chaos" was originally termed by Mark Kac [15] and refers to a relation between microscopic and macroscopic models. Microscopic descriptions, on the one hand, are based on molecules (or particles, individuals, subpopulations, etc.) and model their interactions and driving forces. Macroscopic descriptions, on the other hand, are based on macroscopic observables such as the density and model the dynamics of these quantities. To connect microscopic and macroscopic descriptions, the density in the D-molecule microscopic model should converge as $D \to \infty$ to the density in the macroscopic model. Now Kac's idea behind the terminology "propagation of chaos" is that if the initial distribution is "chaotic" (e.g., positions and velocities of molecules are purely random and independent), then the dynamics of the microscopic model destroys this independence, but finitely many fixed molecules should in the limit as $D \to \infty$ evolve independently (depending on all other molecules only through deterministic macroscopic observables such as the density). In this sense, independence of finitely many fixed molecules "propagates".

Next we give a formal statement of "propagation of chaos" for weakly interacting diffusions. Let $I \subseteq \mathbb{R}$ be a closed interval (we focus on one-dimensional cases), let the set $\mathcal{M}_1(I)$ of probability measures on I be equipped with the 1-Wasserstein metric, let $b, \tilde{\sigma} : I \times \mathcal{M}_1(I) \to \mathbb{R}$ be measurable functions, let $W(i)$, $i \in \mathbb{N}$, be independent standard Brownian motions, for every $D \in \mathbb{N}$ let $X^D = \{(X_t^D(i))_{t\in[0,\infty)} : i \in \{1, \dots, D\}\}$ have state space I^D and be a solution of the stochastic differential equation (SDE)

$$
\begin{aligned}
dX_t^D(i) = {}& b\left(X_t^D(i), \frac{1}{D}\sum_{j=1}^{D} \delta_{X_t^D(j)}\right) dt \\
&+ \tilde{\sigma}\left(X_t^D(i), \frac{1}{D}\sum_{j=1}^{D} \delta_{X_t^D(j)}\right) dW_t(i), \qquad (1.1)\\
&t \in (0,\infty), i \in \{1, \dots, D\},
\end{aligned}
$$

Received April 2018; revised May 2019.
MSC2020 subject classifications. Primary 60K35; secondary 60J70, 92D25.
Key words and phrases. Propagation of chaos, interacting diffusions, mean-field approximation, many-demes limit, measure-valued processes, tree of excursions, excursion measure, altruistic defense.

and let $M(i)$, $i \in \mathbb{N}$, be independent and identically distributed (i.i.d.) and be a solution of the SDE

$$\begin{aligned} dM_t(i) &= b\big(M_t(i), \mathbb{P}(M_t(i) \in \cdot)\big)\,dt \\ &\quad + \tilde{\sigma}\big(M_t(i), \mathbb{P}(M_t(i) \in \cdot)\big)\,dW_t(i), \quad t \in (0,\infty), i \in \mathbb{N}. \end{aligned} \tag{1.2}$$

Then under the additional assumptions that $I = \mathbb{R}$, that $b, \tilde{\sigma}$ are globally Lipschitz continuous, that $\tilde{\sigma}$ is bounded, that b satisfies a linear growth condition, and that for all $D \in \mathbb{N}$ it holds that X_0^D and $(M_0(i))_{i\in\{1,\dots,D\}}$ have the same distribution (in particular, the components of X_0^D are i.i.d.), Theorem 1 in [22] implies for all $k \in \mathbb{N}$ that

$$\big(X_t^D(1), \dots, X_t^D(k)\big)_{t\in[0,\infty)} \overset{D\to\infty}{\Longrightarrow} \big(M_t(1), \dots, M_t(k)\big)_{t\in[0,\infty)} \tag{1.3}$$

in the sense of convergence in distribution on $C([0,\infty), I^k)$ and that

$$\left(\frac{1}{D}\sum_{j=1}^{D} \delta_{X_t^D(j)}\right)_{t\in[0,\infty)} \overset{D\to\infty}{\Longrightarrow} \big(\mathbb{E}[\delta_{M_t(1)}]\big)_{t\in[0,\infty)} \tag{1.4}$$

in the sense of convergence in distribution on $C([0,\infty), \mathcal{M}_1(I))$. So although the components of X^D depend on each other through the empirical distribution process $(\frac{1}{D}\sum_{j=1}^{D} \delta_{X_t^D(j)})_{t\in[0,\infty)}$ for every finite $D \in \mathbb{N}$, in the limit as $D \to \infty$ a finite number of fixed components become independent since they only "depend" on each other through the deterministic process $(\mathbb{E}[\delta_{M_t(1)}])_{t\in[0,\infty)}$. Theorem 4.1 in [8] implies (1.3) and (1.4) under more general assumptions including strict positivity of $\tilde{\sigma}$. Moreover, Proposition 4.29 in [11] and Proposition 3.1 in [12] imply (1.3) and (1.4) for certain cases where $\tilde{\sigma}$ is locally Hölder-$\frac{1}{2}$-continuous in the first argument and does not depend on the second argument. For further results on propagation of chaos see, for example, [1, 18, 20, 21, 23, 27]. The limit (1.4) is also referred to as mean-field approximation. The SDE (1.2) is referred to as mean-field SDE or SDE of McKean–Vlasov type. An essential observation for all of these results is that $X^D(i)$, $i \in \{1,\dots,D\}$, are exchangeable for every $D \in \mathbb{N}$ so that all components have the same distribution and are—informally speaking—of the "same size".

In this paper, we focus on the case $I = [0,1]$ and interpret elements of $[0,1]$ as frequencies (e.g., of a certain property within a subpopulation) and think of a population which is spatially separated into finitely many subpopulations (also denoted as "demes") which are labeled by the elements of $\{1,\dots,D\}$, where $D \in \mathbb{N}$. We assume that a subpopulation stays in frequency 0 as long as there is no immigration into this subpopulation. Our question is: What is the limit of X^D as $D \to \infty$ if only one entry in the vector $(X_0^D(i))_{i\in\mathbb{N}}$ is nonzero? We will assume that X_0 is a $[0,1]^{\mathbb{N}}$-valued random variable which is almost surely summable and that for all $D \in \mathbb{N}$ and all $i \in \{1,\dots,D\}$ it holds almost surely that $X_0^D(i) = X_0(i)$. We will refer to this case as *sparse regime*. In particular, in the sparse regime X_0^D cannot be exchangeable (and nontrivial) for every $D \in \mathbb{N}$. The puzzling question is now how does independence of the initial frequencies propagate in the *many-demes limit* (cf., e.g., [28]) as $D \to \infty$?

We will study this nontrivial question under the simplifying assumption that b is affine-linear in the second argument and that $\tilde{\sigma}$ is constant in the second argument. More precisely, let $f : [0,1]^2 \to \mathbb{R}$, $h : [0,1] \to \mathbb{R}$, $\sigma : [0,1] \to [0,\infty)$, and $h_D : [0,1] \to \mathbb{R}$, $D \in \mathbb{N}$, be functions which satisfy Setting 1.1 below. In the special case where $I = [0,1]$ and where for all $(x,\nu) \in I \times \mathcal{M}_1(I)$ it holds that $b(x,\nu) = \int_I y f(y,x)\,\nu(dy) + h_D(x)$ and $\tilde{\sigma}(x,\nu) = \sigma(x)$, for every $D \in \mathbb{N}$ the solution X^D of (1.1) solves the SDE

$$\begin{aligned} dX_t^D(i) &= \frac{1}{D}\sum_{j=1}^{D} X_t^D(j) f\big(X_t^D(j), X_t^D(i)\big)\,dt + h_D\big(X_t^D(i)\big)\,dt \\ &\quad + \sqrt{\sigma^2\big(X_t^D(i)\big)}\,dW_t(i), \quad t \in (0,\infty), i \in \{1,\dots,D\}. \end{aligned} \tag{1.5}$$

We allow the function h_D to depend on $D \in \mathbb{N}$ in order to include weak immigration (one could think of $h_D(x) = h(x) + \frac{\mu}{D}$ where $\mu \in [0, \infty)$ is the immigration rate into the total population and where $h(0) = 0$).

Now we describe heuristically the propagation of initial independence in the many-demes limit. For this, we assume for simplicity for all $D \in \mathbb{N}$ that $h_D = h$ (no immigration) and that $X_0(i) = 0$ for all $i \in \mathbb{N} \cap [3, \infty)$. The total mass is bounded in D for every time point. As a consequence, the first summand on the right-hand side of (1.5) converges to zero and the first deme $X^D(1)$ converges to the solution of the SDE

$$dY_t = h(Y_t)\,dt + \sqrt{\sigma^2(Y_t)}\,dW_t(1), \quad t \in (0, \infty), \tag{1.6}$$

as $D \to \infty$. Mass emigrates from this first deme. This mass will not migrate to deme 2 (or deme 1) since the immigration rate $\frac{1}{D} X_t^D(1) f(X_t^D(1), X_t^D(2))$ at time $t \in [0, \infty)$ from deme 1 to deme 2 vanishes as $D \to \infty$. Thus, this mass migrates to a deme with index in $\{3, 4, \ldots, D\}$ and there will be a finite number of demes where this mass immigrates and founds a nonvanishing subpopulation. From these subpopulations again mass emigrates. This mass again will not migrate to deme 1, 2, 3, or any other deme with fixed index $i \in \mathbb{N}$ since the total migration rate into a deme with fixed index vanishes in the many-demes limit. Instead, this mass migrates again to randomly chosen demes (which are "empty" with asymptotic probability one) and founds nonvanishing subpopulations. Consequently, since every migrating mass populates "empty" demes (with asymptotic probability one), the subpopulations which originate from descendants of migrants from deme 1 constitute a tree of independent subpopulations. Analogously, the subpopulations which originate from descendants of migrants from deme 2 constitute a tree of independent subpopulations. In addition, these two trees are disjoint (and thus driven by independent families of Brownian motions) and therefore independent if $X_0(1)$ and $X_0(2)$ are independent random variables. In other words, independence of the family $\{X_0(i) : i \in \mathbb{N}\}$ propagates in the many-demes limit and results in a forest of independent trees of independent subpopulations. A formal statement of this "propagation of chaos" result in the sparse regime will be proved in Theorem 1.4 below. We note that, as opposed to the exchangeable regime, "propagation of independence" does not mean that fixed demes (e.g., deme 3 and 4) become independent in the limit as $D \to \infty$ (which is rather trivial) but that the full progenies of individuals starting on deme 3 and on deme 4 do not interfere in the limit as $D \to \infty$.

In the literature, this type of "propagation of chaos" has already been established in two special cases. Theorem 3.3 in [11] proves the analog of Theorem 1.4 below in the special case where the infinitesimal variance σ^2 is additive (and where $I = [0, \infty)$ and for all $x, y \in [0, \infty)$ it holds that $f(y, x) = 1$) and this additivity of infinitesimal variances is a strong tool for decomposing the total population into "loop-free" processes. Moreover, Proposition 2.9 in [5] proves an analog of Theorem 1.4 below in the special case where for all $x, y \in [0, 1]$ and all $D \in \mathbb{N}$ it holds that $\sigma^2(x) = dx(1-x)$, $f(y, x) = c$, $h_D(x) = -cx + sx(1-x) + \frac{m}{D}(1-x)$ where $c, d, m, s \in (0, \infty)$ are positive constants and where the forest of excursions is replaced by a dynamic description hereof which is a continuous atomic-valued Markov process and where independence of disjoint trees is not obvious. In this special case of Wright–Fisher diffusions with selection and rare mutation, there exists a duality with a particle jump process and this duality is a very strong tool. Our more general setup allows for new applications, one of which is carried out in Section 1.5 below. Moreover, there are many related results for interacting particle systems or systems of interacting diffusions where branching processes appear in "sparse" regimes. For example, it is a classical result that the number of alleles of one type in a Wright–Fisher model (Moran model) converges to a branching process in discrete (continuous) time as the population size converges to infinity if the initial numbers of alleles of this type are bounded. For results with SuperBrownian motion appearing in suitable rescalings see, for example, [2–4, 6].

The structure of this paper is as follows. In Section 1.2 we introduce the forest of excursions, in Section 1.3 we state our main result Theorem 1.4, and in Section 1.5 we specify an application to altruistic defense traits. The proof of Theorem 1.4 consists essentially of two major steps. In Section 2 we prove that if ancestral lineages of individuals never come back to a deme, then the resulting "loop-free" processes (see the SDE (2.4) below) converge in the many-demes limit (see Lemma 2.17 below). Moreover, in Section 3 we show that the distance between the D-demes process (1.5) and the corresponding "loop-free" process converges suitably to zero as $D \to \infty$ (see Lemma 3.8 below). The principal idea of reducing the problem to loop-free processes stems from [11]. Throughout this paper, we use the notation from Section 1.1 below without further mentioning.

1.1. *Notation.* For all $x, y \in \mathbb{R}$ we define $x \wedge y := \min\{x, y\}$, $x \vee y := \max\{x, y\}$, $x^+ := x \vee 0$, $x^- := -(x \wedge 0)$, and $\operatorname{sgn}(x) := \mathbf{1}_{(0,\infty)}(x) - \mathbf{1}_{(-\infty,0)}(x)$. We define $\sup \varnothing := -\infty$ and $\inf \varnothing := \infty$. We write $\mathbb{N} := \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For all $N \in \mathbb{N}$ we write $[N] := \{1, \dots, N\}$ and $[N]_0 := [N] \cup \{0\}$.

For the remainder of this subsection, let (E, d_E) and (F, d_F) be metric spaces and let $d \in \mathbb{N}$ and $m \in \mathbb{N}_0$. We denote by $\mathcal{B}(E)$ the Borel σ-algebra on (E, d_E) and by $\mathcal{M}_{\mathrm{f}}(E)$ the set of finite measures on $(E, \mathcal{B}(E))$ endowed with the weak topology. For every $s \in [0, \infty)$ we denote by $D([s, \infty), E)$ the set of all càdlàg functions $f : [s, \infty) \to E$ endowed with the Skorokhod topology. We denote by $C(E, F)$ the set of all continuous functions $f : E \to F$ and by $\operatorname{Lip}(E, F)$ the set of all Lipschitz continuous functions $f : E \to F$. We denote by $C_b^2(\mathbb{R}, \mathbb{R})$ the set of twice continuously differentiable bounded functions $\psi : \mathbb{R} \to \mathbb{R}$ with bounded first and second derivative. For every $\psi : \mathbb{R} \to \mathbb{R}$ we write $\|\psi\|_\infty := \sup_{x \in \mathbb{R}} |\psi(x)| \in [0, \infty]$. We denote by $C^m([0, 1]^d, \mathbb{R})$ the set of functions $\psi : [0, 1]^d \to \mathbb{R}$ whose partial derivatives of order 0 through m exist and are continuous on $[0, 1]^d$. For every $\psi : [0, 1]^d \to \mathbb{R}$ we define $\|\psi\|_\infty := \sup_{x \in [0,1]^d} |\psi(x)| \in [0, \infty]$. For every multiindex $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ of length $|\alpha| := \sum_{k=1}^d \alpha_k$ we write $\partial^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$. For every $\psi \in C^m([0, 1]^d, \mathbb{R})$ we set $\|\psi\|_{C^m} := \max_{\alpha \in \mathbb{N}_0^d, |\alpha| \le m} \|\partial^\alpha \psi\|_\infty$.

By a solution of an SDE driven by Brownian motions we mean a stochastic process with continuous sample paths which is adapted to the filtration generated by the Brownian motions and the initial value and which satisfies the integrated SDE for every time point almost surely.

1.2. *Setting and forest of excursions.* In this subsection, we gather the assumptions that we impose in our main result, Theorem 1.4 below, and we introduce the forest of excursions which plays the role of a limiting object in our main result.

In the following Setting 1.1, we collect our assumptions on the coefficients of the SDE (1.5). Under these assumptions, for every $D \in \mathbb{N}$ the SDE (1.5) has a unique strong solution with continuous sample paths in $[0, 1]^D$; see Theorem 3.2 in [26]. Moreover, under these assumptions, the SDE (1.6) has a unique strong solution $Y = (Y_t)_{t \in [0,\infty)}$ with continuous sample paths in $[0, 1]$ for which 0 is a trap, that is, for all $t, s \in [0, \infty)$ it holds that ($Y_t = 0$ implies $Y_{t+s} = 0$).

SETTING 1.1 (Coefficient functions). Let $\mu \in [0, \infty)$, $f \in C^3([0, 1]^2, \mathbb{R})$, $h \in C^3([0, 1], \mathbb{R})$, and $h_D \in C^3([0, 1], \mathbb{R})$, $D \in \mathbb{N}$, have the properties that $\sup_{D \in \mathbb{N}} \|h_D\|_{C^2} < \infty$, that $\lim_{D \to \infty} D h_D(0) = \mu$, for all $x \in [0, 1]$ that $\lim_{D \to \infty} h_D(x) = h(x)$, and for all $D \in \mathbb{N}$ and all $y \in (0, 1]$ that $y f(y, 1) + h_D(1) \le 0$, that $f(y, 0) > 0$, that $2\mu \ge D h_D(0) \ge 0 = h(0)$, and that $h(1) < 0$.

Let $L_f, L_h \in [0,\infty)$ be such that for all $D \in \mathbb{N}$ and all $x, y, u, v \in [0,1]$ it holds that $|f(y,x) - f(v,u)| \le L_f|y - v| + L_f|x - u|$, that $|f(y,x)| \le L_f$, that $|h(x) - h(y)| \le L_h|x - y|$, and that $|h_D(x) - h_D(y)| \le L_h|x - y|$.

Let $\sigma^2 \in C^3([0,1],\mathbb{R})$ satisfy that $\sigma^2(0) = 0 = \sigma^2(1)$ and for all $x \in (0,1)$ that $\sigma^2(x) > 0$. Let $L_\sigma \in [0,\infty)$ be such that for all $x, y \in [0,1]$ it holds that $|\sigma^2(x) - \sigma^2(y)| \le L_\sigma|x - y|$.

For every $D \in \mathbb{N}$, we denote by $\tilde{h}_D : [0,1] \to \mathbb{R}$ the function that satisfies for all $x \in [0,1]$ that $\tilde{h}_D(x) = h_D(x) - h_D(0)$.

In addition to Setting 1.1, we impose the following assumptions involving the scale function S of Y, which imply that Y hits zero in finite time almost surely (a straightforward adaptation of Lemma 9.5 and Lemma 9.6 in [10] to the state space $[0,1]$ shows that this is ensured by (1.10) below) and that there exists an excursion measure for Y.

SETTING 1.2 (Scale function). Assume that Setting 1.1 holds and that

$$\lim_{(0,\frac{1}{2})\ni\varepsilon\to 0} \int_\varepsilon^{\frac{1}{2}} \frac{h(x)}{\sigma^2(x)}\,dx \in \mathbb{R}.$$

We define the functions $s, S : [0,1) \to [0,\infty)$ and $\tilde{a} : [0,1] \to [0,\infty)$ by

$$[0,1) \ni z \mapsto s(z) := \exp\left(-\int_0^z \frac{h(x)}{\frac{1}{2}\sigma^2(x)}\,dx\right) \in [0,\infty), \tag{1.7}$$

$$[0,1) \ni y \mapsto S(y) := \int_0^y s(z)\,dz \in [0,\infty), \tag{1.8}$$

$$[0,1] \ni y \mapsto \tilde{a}(y) := yf(y,0) \in [0,\infty). \tag{1.9}$$

We further assume that

$$\int_0^{\frac{1}{2}} \frac{S(y)}{\sigma^2(y)s(y)}\,dy + \int_{\frac{1}{2}}^1 \frac{\tilde{a}(y)}{\sigma^2(y)s(y)}\,dy < \infty. \tag{1.10}$$

We define the *set of excursions from zero* by

$$U := \left\{ \begin{array}{c} \eta \in C(\mathbb{R},[0,1]) : \text{there exists } t_0 \in (0,\infty) \text{ such that} \\ \text{it holds for all } t \in (0,t_0) \text{ that } \eta_t > 0 \text{ and} \\ \text{it holds for all } t \in (-\infty,0] \cup [t_0,\infty) \text{ that } \eta_t = 0 \end{array} \right\}.$$

Moreover, we denote by $D(\mathbb{R},[0,1])$ the set of all càdlàg functions $f : \mathbb{R} \to [0,1]$ and we define

$$V := \{\eta \in D(\mathbb{R},[0,1]) : \eta_t = 0 \text{ for all } t \in (-\infty,0)\} \supseteq U.$$

In the situation of Setting 1.2, Theorem 1 in [10] adapted to the state space $[0,1]$ shows that there exists a unique σ-finite measure Q on U satisfying the following property: For every bounded and continuous function $F : C([0,\infty),[0,1]) \to \mathbb{R}$ with the property that there exists a $\delta > 0$ such that for all $\chi \in C([0,\infty),[0,1])$ with $\sup_{t\in[0,\infty)} \chi_t < \delta$ it holds that $F(\chi) = 0$, it holds that

$$\lim_{(0,1)\ni\varepsilon\to 0} \frac{1}{S(\varepsilon)} \mathbb{E}[F(Y) \mid Y_0 = \varepsilon] = \int F(\eta)Q(d\eta).$$

The measure Q is called the *excursion measure* associated with Y; see also [24]. A straightforward adaptation of Lemma 9.8 in [10] to the state space $[0,1]$ and assumption (1.10) imply that

$$\int\int_0^\infty \tilde{a}(\chi_t)\,dt\,Q(d\chi) = \int_0^1 \frac{\tilde{a}(y)}{\frac{1}{2}\sigma^2(y)s(y)}\,dy < \infty. \tag{1.11}$$

For the convergence result, we further assume the following Setting 1.3 for the initial distributions.

SETTING 1.3 (Sparse initial condition). Assume that Setting 1.2 holds. For every $i \in \mathbb{N}$ let $X_0(i)$ be a $[0,1]$-valued random variable and for every $D \in \mathbb{N}$ let $\{(X_t^D(i))_{t\in[0,\infty)} : i \in [D]\}$ be a solution of (1.5) such that a.s. $\sum_{i=1}^{\infty} X_0(i) < \infty$ and such that for all $D \in \mathbb{N}$ and all $i \in [D]$ it holds a.s. that $X_0^D(i) = X_0(i)$.

Under the assumption of Setting 1.3, we now construct the associated forest of excursions. For that, let $Y(i) = (Y_t(i))_{t\in[0,\infty)}$, $i \in \mathbb{N}$, be solutions of (1.6) driven by independent Brownian motions which are independent of $(X_0(i))_{i\in\mathbb{N}}$ such that for all $i \in \mathbb{N}$ it holds a.s. that $Y_0(i) = X_0(i)$. These trajectories describe the demes of the initial population. Let $\{\Pi^{\varnothing}\} \cup \{\Pi^{(n,s,\chi)} : (n,s,\chi) \in \mathbb{N}_0 \times [0,\infty) \times V\}$ be an independent family of Poisson point processes on $[0,\infty) \times U$ with intensity measures

$$\mathbb{E}\big[\Pi^{\varnothing}(dt \otimes d\eta)\big] = \mu\, dt \otimes Q(d\eta)$$

and

$$\mathbb{E}\big[\Pi^{(n,s,\chi)}(dt \otimes d\eta)\big] = \tilde{a}(\chi_{t-s})\, dt \otimes Q(d\eta), \quad (n,s,\chi) \in \mathbb{N}_0 \times [0,\infty) \times V.$$

The points of $\Pi^{\varnothing}$ and $\Pi^{(n,s,\chi)}$ are interpreted as tuples of times and paths providing the population times of new demes and the evolution of the population inside these demes. Here, $\Pi^{\varnothing}$ describes the demes whose founders immigrated into the system, while $\Pi^{(n,s,\chi)}$ describes the demes which descend from a deme with population size trajectory $(\chi_{t-s})_{t\in[0,\infty)}$ (which is zero before its population time s) and where the ancestral lineages of individuals living on these demes have exactly $n \in \mathbb{N}$ migration events (only counting migration events within the system). The 0th generation is the random σ-finite measure on $[0,\infty) \times V$ defined through $\mathcal{T}^{(0)} := \sum_{i=1}^{\infty} \delta_{(0,(\mathbf{1}_{t\in[0,\infty)} Y_t(i))_{t\in\mathbb{R}})} + \Pi^{\varnothing}$. For every $n \in \mathbb{N}_0$ the $(n+1)$th generation is the random σ-finite measure on $[0,\infty) \times U$ representing all the demes which have been colonized from demes of the nth generation, that is, $\mathcal{T}^{(n+1)} := \int \Pi^{(n,s,\chi)} \mathcal{T}^{(n)}(ds \otimes d\chi)$. The *forest of excursions* $\mathcal{T}$ is then the sum of all of these measures $\mathcal{T} := \sum_{n\in\mathbb{N}_0} \mathcal{T}^{(n)}$.

A straightforward adaptation of Lemma 5.2, Lemma 9.9, and Lemma 9.10 in [10] to the state space $[0,1]$ shows for every $t \in [0,\infty)$ that the total mass $\int \chi_{t-s} \mathcal{T}(ds \otimes d\chi)$ has finite expectation and is thus finite almost surely. Moreover, in the case where $\mu = 0$ (no immigration) and where there exists an $x \in (0,1]$ such that for all $i \in \mathbb{N}$ it holds that $X_0(i) = x\mathbf{1}_{i=1}$, Theorem 5 in [10] yields that the total mass process dies out (i.e., $\int \chi_{t-s} \mathcal{T}(ds \otimes d\chi)$ converges to zero in probability as $t \to \infty$) if and only if

$$\int \int_0^{\infty} \tilde{a}(\chi_t)\, dt\, Q(d\chi) \le 1. \tag{1.12}$$

1.3. *Main result*: *Propagation of chaos in the sparse regime*. In this subsection, we state our main theorem.

THEOREM 1.4 (Convergence to a forest of excursions). *Assume that* Setting 1.3 *holds and let $\mathcal{T}$ be the forest of excursions constructed in* Section 1.2. *Then it holds that*

$$\left(\sum_{i=1}^{D} X_t^D(i) \delta_{X_t^D(i)}\right)_{t\in[0,\infty)} \xRightarrow{D\to\infty} \left(\int \eta_{t-s} \delta_{\eta_{t-s}} \mathcal{T}(ds \otimes d\eta)\right)_{t\in[0,\infty)} \tag{1.13}$$

in the sense of convergence in distribution on $D([0,\infty), \mathcal{M}_{\mathrm{f}}([0,1]))$.

The form of the left-hand side in (1.13) might look unfamiliar at first glance. We note that the sequence $\{(\sum_{i=1}^{D} \delta_{X_t^D(i)})_{t\in[0,\infty)} : D \in \mathbb{N}\}$ has no chance of being relatively compact as stochastic processes with values in $\mathcal{M}_f([0,1])$, as the total masses diverge to infinity as $D \to \infty$. By weighting the point masses as in (1.13), we avoid this problem and retain a well-understood state space. Alternatively, one can change the (topology of the) state space. This is done in [11], where σ-finite measures on $(0, 1]$ with the vague topology are used to prove convergence of $\{(\sum_{i=1}^{D} \delta_{X_t^D(i)})_{t\in[0,\infty)} : D \in \mathbb{N}\}$ instead.

We emphasize that the limiting object $\mathcal{T}$ is easier to analyze than the solution of the SDE (1.5) due to its tree structure and since general branching processes are very well understood, resulting, for example, in the criterion (1.12). Further properties of the limiting process in the case $\mu = 0$ without immigration are investigated in [10].

1.4. *Main ideas and structure of the proof of Theorem* 1.4. In Section 2.1 we decompose (1.5) into processes with migration levels in the sense that the next higher migration level is driven by successful migrations essentially only from the migration level directly below, see (2.1) below. We then couple these migration level processes to "loop-free" processes which we obtain by pretending that on a fixed deme all individuals have the same migration level, which turns (2.1) into (2.4) below. We show in Section 2.4 that these loop-free processes converge in the limit as $D \to \infty$ to the forest of excursions. To prove this result, we apply induction on the number of migration steps, which is useful since, conditionally on the lower migration levels, the processes in the next higher migration level in the loop-free processes evolve as independent diffusions. The convergence of independent diffusions of this kind is obtained in Section 2.3 in Lemma 2.15 below. This result can be seen as a functional Poisson limit theorem and follows essentially from Lemma 2.11 below which is proved in [11] by reversing time.

The principal idea of a decomposition into loop-free processes is taken from [11]. There, however, the considerations were restricted to the "classical" migration term of the form $\frac{1}{D}\sum_{j=1}^{D}(X_t^D(j) - X_t^D(i))$, while here we allow for "nonlinear" migrations of the form $\frac{1}{D}\sum_{j=1}^{D} X_t^D(j) f(X_t^D(j), X_t^D(i))$. By considering the case where for all $x, y \in [0, 1]$ it holds that $f(y, x) = 1$ and where for all $x \in [0, 1]$ and all $D \in \mathbb{N}$ we replace $h_D(x)$ by $h_D(x) - x$ in (1.5), we recover the classical migration term in our framework.

To complete the proof of Theorem 1.4, it remains to show that the migration level processes and the loop-free processes have the same limit as $D \to \infty$. This is carried out in Section 3.2. This has already been proved in [11] for the case where σ^2 is of the form $\sigma^2(x) = \beta x$ for a constant $\beta \in (0,\infty)$. This assumption crucially simplified the situation, as this completely decouples the infinitesimal variances of the migration level processes and allows to prove that the L^1-distance between the migration level processes and the loop-free processes converges to zero as $D \to \infty$ using Gronwall's inequality. In the more general situation of this paper, we instead show that a certain form of a weak distance between the migration level processes and the loop-free processes converges to zero as $D \to \infty$, see Lemma 3.8 below. The proof relies on applying Itô's formula to suitable evaluations of the semigroup of the loop-free processes, see (3.26) below.

1.5. *Application*: *Altruistic defense traits*. Let $\alpha, \beta, \kappa \in (0,\infty)$, let $\mu_\infty \in [0,\infty)$, let $a \in (1,\infty)$, let $(\mu_D)_{D\in\mathbb{N}} \subseteq [0, 1]$ be such that $\lim_{D\to\infty} D\mu_D = \mu_\infty$ and such that for all $D \in \mathbb{N}$ it holds that $D\mu_D \le 2\mu_\infty$, and let $b \in C^3([0, 1], \mathbb{R})$ be such that $b(1) = 0 \le b(0)$. Let $f : [0, 1]^2 \to \mathbb{R}$, $h : [0, 1] \to \mathbb{R}$, $h_D : [0, 1] \to \mathbb{R}$, $D \in \mathbb{N}$, and $\sigma^2 : [0, 1] \to [0,\infty)$ be the functions satisfying for all $D \in \mathbb{N}$ and all $x, y \in [0, 1]$ that $f(y, x) = \kappa(a - x)\frac{a-x}{a}\frac{1}{a-y}$, that $h(x) = -\kappa x(a-x)\frac{1}{a} - \alpha x(1-x)$, that $h_D(x) = h(x) + \mu_D b(x)$, and that $\sigma^2(x) = \beta(a-x)x(1-x)$. Then it holds for all $x, y \in [0,1]$ that $yf(y,x) + h(x) = \kappa\frac{a-x}{a-y}(y-x) - \alpha x(1-x)$. Thus, for every $D \in \mathbb{N}$, (1.5) specializes to the SDE

$$
\begin{aligned}
dX_t^D(i) = {} & \frac{\kappa}{D}\sum_{j=1}^{D}\frac{a - X_t^D(i)}{a - X_t^D(j)}\big(X_t^D(j) - X_t^D(i)\big)\,dt \\
& - \alpha X_t^D(i)\big(1 - X_t^D(i)\big)\,dt + \mu_D b\big(X_t^D(i)\big)\,dt \\
& + \sqrt{\beta\big(a - X_t^D(i)\big)X_t^D(i)\big(1 - X_t^D(i)\big)}\,dW_t(i), \\
& t \in (0,\infty), i \in [D].
\end{aligned}
\tag{1.14}
$$

Theorem 1.3 in [12] shows (in the case where for all $D \in \mathbb{N}$ it holds that $\mu_D = 0$, that is, the case of no mutation) that the solution process of the SDE (1.14) arises as a diffusion limit of the relative frequency of altruistic individuals in the host population in a Lotka–Volterra type host-parasite model. The reason for the functional form of the coefficient functions is as follows. Since individuals migrate at fixed rate, altruistic individuals in the host population (let $H_t^{D,N}(i)$ denote the total number of hosts on deme $i \in [D]$ at time $t \in [0,\infty)$) migrate at rate $\frac{\kappa}{N}m(i,j)\frac{H_t^{D,N}(j)}{H_t^{D,N}(i)}$ from deme j to deme i. The total host population sizes evolve much faster (a separation of time scales occurs) than the relative frequencies and they stabilize locally in time at $\frac{1}{\beta(a-x)}$ if x is the current frequency of altruists, where a, β are suitable parameters. Thus, if time is measured in multiples of N, the migration rates converge to

$$
w\text{-}\lim_{N\to\infty} N\frac{\kappa}{N}m(i,j)\frac{H_{tN}^{D,N}(j)}{H_{tN}^{D,N}(i)} = \kappa m(i,j)\frac{\beta(a - X_t^D(i))}{\beta(a - X_t^D(j))}.
$$

The resampling rate in the Wright–Fisher diffusion is inverse-proportional to the total mass and this explains why the squared diffusion term is

$$
w\text{-}\lim_{N\to\infty}\frac{1}{H_{tN}^{D,N}(i)}\frac{A_{tN}^{D,N}(i)}{H_{tN}^{D,N}(i)}\left(1 - \frac{A_{tN}^{D,N}(i)}{H_{tN}^{D,N}(i)}\right) = \beta\big(a - X_t^D(i)\big)X_t^D(i)\big(1 - X_t^D(i)\big).
$$

Define $\mu := \mu_\infty b(0)$. We check the assumptions of Theorem 1.4: Indeed, $\mu \in [0,\infty)$, $f \in C^3([0,1]^2,\mathbb{R})$, $h, \sigma^2 \in C^3([0,1],\mathbb{R})$, for all $D \in \mathbb{N}$ it holds that $h_D \in C^3([0,1],\mathbb{R})$, it holds that $\sup_{D\in\mathbb{N}}\|h_D\|_{C^2} \le \|h\|_{C^2} + \|b\|_{C^2} < \infty$ and that $\lim_{D\to\infty} Dh_D(0) = \lim_{D\to\infty} D\mu_D b(0) = \mu_\infty b(0) = \mu$, for all $x \in [0,1]$ it holds that $\lim_{D\to\infty} h_D(x) = h(x)$, for all $D \in \mathbb{N}$ and all $y \in (0,1]$ it holds that $yf(y,1) + h_D(1) = yf(y,1) + h(1) = \frac{\kappa y(a-1)^2}{a(a-y)} - \frac{\kappa(a-1)}{a} \le 0$, that $f(y,0) = \frac{\kappa a}{a-y} > 0$, that $Dh_D(0) = D\mu_D b(0) \le 2\mu_\infty b(0) = 2\mu$, that $Dh_D(0) = D\mu_D b(0) \ge 0$, that $h(0) = 0$, and that $h(1) = -\frac{\kappa(a-1)}{a} < 0$. Moreover, it holds that $\sigma^2(0) = 0 = \sigma^2(1)$ and for all $x \in (0,1)$ that $\sigma^2(x) > 0$. Thus, Setting 1.1 is satisfied. Furthermore, it holds that

$$
\begin{aligned}
\lim_{(0,\frac{1}{2})\ni\varepsilon\to 0}\int_\varepsilon^{\frac{1}{2}}\frac{h(x)}{\sigma^2(x)}\,dx &= \lim_{(0,\frac{1}{2})\ni\varepsilon\to 0}\int_\varepsilon^{\frac{1}{2}}\frac{-\kappa x(a-x)\frac{1}{a} - \alpha x(1-x)}{\beta(a-x)x(1-x)}\,dx \\
&= \lim_{(0,\frac{1}{2})\ni\varepsilon\to 0}\int_\varepsilon^{\frac{1}{2}}\frac{-\kappa}{a\beta(1-x)} - \frac{\alpha}{\beta(a-x)}\,dx \\
&= \lim_{(0,\frac{1}{2})\ni\varepsilon\to 0}\left(\frac{\kappa}{a\beta}\left(\ln\left(1 - \frac{1}{2}\right) - \ln(1-\varepsilon)\right)\right.
\end{aligned}
\tag{1.15}
$$

$$+\frac{\alpha}{\beta}\left(\ln\left(a-\frac{1}{2}\right)-\ln(a-\varepsilon)\right)\Bigg)$$
$$=\frac{\kappa}{a\beta}\ln\left(1-\frac{1}{2}\right)+\frac{\alpha}{\beta}\left(\ln\left(a-\frac{1}{2}\right)-\ln(a)\right)\in\mathbb{R}.$$

Let $s, S : [0, 1) \to [0, \infty)$ and $\tilde{a} : [0, 1] \to [0, \infty)$ be given by (1.7), (1.8), and (1.9), respectively. Then it holds for all $z \in [0, 1)$ that

$$s(z)=\exp\left(\int_0^z \frac{2\kappa}{a\beta(1-x)}+\frac{2\alpha}{\beta(a-x)}\,dx\right)=(1-z)^{-\frac{2\kappa}{a\beta}}\left(\frac{a-z}{a}\right)^{-\frac{2\alpha}{\beta}} \tag{1.16}$$

and

$$S(z)=\int_0^z s(x)\,dx\le zs(z). \tag{1.17}$$

We obtain from (1.17) that

$$\int_0^{\frac{1}{2}}\frac{S(y)}{\sigma^2(y)s(y)}\,dy\le\int_0^{\frac{1}{2}}\frac{1}{\beta(a-y)(1-y)}\,dy\le\frac{1}{2\beta(a-\frac{1}{2})(1-\frac{1}{2})}<\infty \tag{1.18}$$

and it follows from (1.16) and from the fact that $\frac{2\kappa}{a\beta}-1\in(-1,\infty)$ that

$$\begin{aligned}\int_{\frac{1}{2}}^1\frac{\tilde{a}(y)}{\sigma^2(y)s(y)}\,dy&=\int_{\frac{1}{2}}^1\frac{\frac{\kappa a y}{a-y}}{\beta(a-y)y(1-y)}(1-y)^{\frac{2\kappa}{a\beta}}\left(\frac{a-y}{a}\right)^{\frac{2\alpha}{\beta}}\,dy\\&=\frac{\kappa a}{\beta}a^{-\frac{2\alpha}{\beta}}\int_{\frac{1}{2}}^1(1-y)^{\frac{2\kappa}{a\beta}-1}(a-y)^{\frac{2\alpha}{\beta}-2}\,dy\\&\le\frac{\kappa a}{\beta}a^{-\frac{2\alpha}{\beta}}\left(\left(a-\frac{1}{2}\right)^{\frac{2\alpha}{\beta}-2}+(a-1)^{\frac{2\alpha}{\beta}-2}\right)\int_{\frac{1}{2}}^1(1-y)^{\frac{2\kappa}{a\beta}-1}\,dy\\&<\infty.\end{aligned} \tag{1.19}$$

Hence, Setting 1.2 is satisfied. Therefore, Theorem 1.4 is applicable to the SDE (1.14) for any initial configuration satisfying Setting 1.3.

For the remainder of this subsection we consider the case where $\mu_\infty = 0$ and where there exists an $x \in (0, 1]$ such that for all $D \in \mathbb{N}$ and all $i \in [D]$ it holds that $X_0^D(i) = x\mathbf{1}_{i=1}$. We obtain from (1.16), (1.18), and (1.19) together with a straightforward adaptation of Lemma 9.6, Lemma 9.9, and Lemma 9.10 in [10] to the state space [0, 1] that the assumptions of Theorem 5 in [10] are satisfied. An application of the latter theorem shows that the total mass process $(\int \chi_{t-s}\mathcal{T}(ds\otimes d\chi))_{t\in[0,\infty)}$ dies out (i.e., it converges to zero in probability as $t\to\infty$) if and only if (1.12) holds. Equations (1.11) and (1.16) yield that

$$\begin{aligned}\int\int_0^\infty\tilde{a}(\chi_t)\,dt\,Q(d\chi)&=\int_0^1\frac{\tilde{a}(y)}{\frac{1}{2}\sigma^2(y)s(y)}\,dy\\&=\int_0^1\frac{\frac{\kappa a y}{a-y}}{\frac{1}{2}\beta(a-y)y(1-y)}(1-y)^{\frac{2\kappa}{a\beta}}\left(\frac{a-y}{a}\right)^{\frac{2\alpha}{\beta}}\,dy.\end{aligned}$$

This together with (1.12) and the fact that $\frac{2\kappa}{a\beta}\int_0^1 (1-y)^{\frac{2\kappa}{a\beta}-1}\,dy = 1$ proves that the total mass process dies out if and only if

$$\begin{aligned}
0 &\geq \int_0^1 \frac{\frac{\kappa a y}{a-y}}{\frac{1}{2}\beta(a-y)y(1-y)}(1-y)^{\frac{2\kappa}{a\beta}}\left(\frac{a-y}{a}\right)^{\frac{2\alpha}{\beta}}\,dy - 1 \\
&= \frac{2\kappa}{a\beta}\int_0^1 (1-y)^{\frac{2\kappa}{a\beta}-1}\left(\frac{a-y}{a}\right)^{\frac{2\alpha}{\beta}-2}\,dy - 1 \\
&= \frac{2\kappa}{a\beta}\int_0^1 (1-y)^{\frac{2\kappa}{a\beta}-1}\left(\left(\frac{a-y}{a}\right)^{\frac{2\alpha}{\beta}-2} - 1\right)dy.
\end{aligned}$$

Consequently, the total mass process dies out if and only if $\alpha \geq \beta$. Our interpretation is that the altruistic defense trait can spread in the host population in the many-demes limit if the cost of defense α is smaller than the benefit of defense β.

2. Convergence of the loop-free processes.

2.1. *Migration level processes and loop-free processes.* Throughout this subsection, assume that Setting 1.1 holds. To prove Theorem 1.4, we use a decomposition into migration levels. We say that an individual has *migration level $k \in \mathbb{N}_0$* at time $t \in [0,\infty)$ if its ancestral lineage up to time t contains exactly k migration steps (within the system). To formalize this, we define for all $D \in \mathbb{N}$ that $X^{D,-1} := 0$ and consider for every $D \in \mathbb{N}$ the SDE

$$\begin{aligned}
dX_t^{D,k}(i) &= \frac{1}{D}\sum_{j=1}^{D} X_t^{D,k-1}(j) f\biggl(\sum_{m\in\mathbb{N}_0} X_t^{D,m}(j), \sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)\biggr)\,dt \\
&\quad + \frac{X_t^{D,k}(i)}{\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)}\tilde{h}_D\biggl(\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)\biggr)\,dt + \mathbf{1}_{k=0} h_D(0)\,dt \\
&\quad + \sqrt{\frac{X_t^{D,k}(i)}{\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)}\sigma^2\biggl(\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)\biggr)}\,dW_t^k(i), \\
&\quad t \in (0,\infty), (i,k) \in [D]\times\mathbb{N}_0,
\end{aligned} \tag{2.1}$$

where $\{W^k(i) : (i,k) \in \mathbb{N}\times\mathbb{N}_0\}$ is a set of independent standard Brownian motions. Throughout this paper, we consider weak solutions of (2.1) with initial distribution and values in $\{(x_{i,k})_{(i,k)\in[D]\times\mathbb{N}_0} \in [0,1]^{[D]\times\mathbb{N}_0} : \sum_{k\in\mathbb{N}_0} x_{i,k} \in [0,1] \text{ for all } i \in [D]\}$. The existence of such solutions can be shown as in Lemma 4.3 in [11]. These processes will be referred to as *migration level processes*.

The following lemma shows that (1.5) can be recovered from (2.1) by summing over all migration levels.

LEMMA 2.1 (Decomposition into migration levels). *Assume that* Setting 1.1 *holds, let* $D \in \mathbb{N}$, *let* $\{(X_t^{D,k}(i), W_t^k(i))_{t\in[0,\infty)} : (i,k) \in [D]\times\mathbb{N}_0\}$ *be a weak solution of* (2.1), *and let* $W(i) = (W_t(i))_{t\in[0,\infty)}$, $i \in [D]$, *be continuous adapted processes satisfying for all* $i \in [D]$ *and all* $t \in [0,\infty)$ *that a.s.*

$$\begin{aligned}
W_t(i) &= \int_0^t \mathbf{1}_{\{\sum_{m\in\mathbb{N}_0} X_s^{D,m}(i)>0\}} \sum_{k\in\mathbb{N}_0}\sqrt{\frac{X_s^{D,k}(i)}{\sum_{m\in\mathbb{N}_0} X_s^{D,m}(i)}}\,dW_s^k(i) \\
&\quad + \int_0^t \mathbf{1}_{\{\sum_{m\in\mathbb{N}_0} X_s^{D,m}(i)=0\}}\,dW_s^0(i).
\end{aligned} \tag{2.2}$$

Then $W = \{W(i) : i \in [D]\}$ is a D-dimensional standard Brownian motion and the process $\tilde{X}^D = \{(\tilde{X}^D_t(i))_{t\in[0,\infty)} : i \in [D]\}$ defined for all $i \in [D]$ and all $t \in [0,\infty)$ by

$$\tilde{X}^D_t(i) := \sum_{k\in\mathbb{N}_0} X^{D,k}_t(i) \tag{2.3}$$

is the unique solution of (1.5) driven by the Brownian motion W.

PROOF. The processes $W(i)$, $i \in [D]$, are continuous local martingales whose cross-variation processes satisfy for all $i, j \in [D]$ and all $t \in [0,\infty)$ that $\langle W(i), W(j)\rangle_t = \delta_{ij} t$. Lévy's characterization of Brownian motion (see, e.g., Theorem 3.3.16 in [16]) implies that W is a D-dimensional standard Brownian motion. Moreover, it follows from summing (2.1) over $k \in \mathbb{N}_0$ that $\tilde{X}^D$ satisfies (1.5) with Brownian motion given by (2.2). Pathwise uniqueness of the SDE (1.5) in the situation of Setting 1.1 follows from Theorem 3.2 in [26]. This finishes the proof of Lemma 2.1. $\square$

In the limit as $D \to \infty$, the migration level processes are essentially loop-free in the following sense. We define for all $D \in \mathbb{N}$ that $Z^{D,-1} := 0$ and consider for every $D \in \mathbb{N}$ the SDE

$$\begin{aligned} dZ^{D,k}_t(i) &= \frac{1}{D}\sum_{j=1}^{D} Z^{D,k-1}_t(j) f\big(Z^{D,k-1}_t(j), Z^{D,k}_t(i)\big)\,dt \\ &\quad + \tilde{h}_D\big(Z^{D,k}_t(i)\big)\,dt + \mathbf{1}_{k=0} h_D(0)\,dt \\ &\quad + \sqrt{\sigma^2\big(Z^{D,k}_t(i)\big)}\,dW^k_t(i), \quad t \in (0,\infty), (i,k) \in [D]\times\mathbb{N}_0, \end{aligned} \tag{2.4}$$

where $\{W^k(i) : (i,k) \in \mathbb{N}\times\mathbb{N}_0\}$ is a set of independent standard Brownian motions. Existence and uniqueness of strong solutions of (2.4) follow from Theorem 3.2 in [26]. These processes will be referred to as *loop-free processes*.

SETTING 2.2 (Coupling of migration level and loop-free processes). Assume that Setting 1.1 holds. For every $D \in \mathbb{N}$ let

$$\big\{\big(X^{D,k}_t(i), W^k_t(i)\big)_{t\in[0,\infty)} : (i,k) \in [D]\times\mathbb{N}_0\big\}$$

be a weak solution of (2.1) with initial distribution and values in

$$\left\{(x_{i,k})_{(i,k)\in[D]\times\mathbb{N}_0} \in [0,1]^{[D]\times\mathbb{N}_0} : \sum_{k\in\mathbb{N}_0} x_{i,k} \in [0,1] \text{ for all } i \in [D]\right\}.$$

For every $D \in \mathbb{N}$ and every $x \in [0,1]^{[D]\times\mathbb{N}_0}$ we denote by

$$\big\{\big(Z^{D,k,x}_t(i)\big)_{t\in[0,\infty)} : (i,k) \in [D]\times\mathbb{N}_0\big\}$$

continuous adapted processes that are defined on the stochastic basis given by the weak solution of (2.1), satisfy (2.4) with Brownian motion given by the Brownian motion of the weak solution of (2.1), and further satisfy for all $(i,k) \in [D]\times\mathbb{N}_0$ that a.s. $Z^{D,k,x}_0(i) = x_{i,k}$. Whenever we omit the index x, we consider the solution of (2.4) satisfying for all $(i,k) \in [D]\times\mathbb{N}_0$ that a.s. $Z^{D,k}_0(i) = X^{D,k}_0(i)$. For notational simplicity, we do not distinguish notationally between the possibly different stochastic bases and Brownian motions for different $D \in \mathbb{N}$.

2.2. *Moment and regularity estimates.* In this subsection, we collect some preparatory results. We start with the following lemma which provides an estimate for the first moment of the total mass process.

LEMMA 2.3 (First moment). *Assume that* Setting 2.2 *holds and let* $T \in [0, \infty)$. *Then we have for all* $D \in \mathbb{N}$ *that*

$$\sup_{t\in[0,T]} \mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_t^{D,k}(i)\right] \leq e^{(L_f+L_h)T}\left(\mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)\right] + 2\mu T\right) \tag{2.5}$$

and (2.5) *holds with* $X_t^{D,k}(i)$ *replaced by* $Z_t^{D,k}(i)$.

PROOF. For all $D \in \mathbb{N}$ let $\{W(i) : i \in [D]\}$ and $\tilde{X}^D$ be as in Lemma 2.1. Setting 1.1 implies for all $D \in \mathbb{N}$ and all $x, y \in [0, 1]$ that $f(y, x) \leq L_f$ and $h_D(x) \leq L_h x + 2\mu/D$. Together with Lemma 2.1, this shows for all $D \in \mathbb{N}$ and all $t \in [0, \infty)$ that a.s.

$$\begin{aligned}\sum_{i=1}^{D}\tilde{X}_t^D(i) \leq \sum_{i=1}^{D}\tilde{X}_0^D(i) + 2\mu t + (L_f + L_h)\int_0^t \sum_{i=1}^{D}\tilde{X}_s^D(i)\,ds \\ + \sum_{i=1}^{D}\int_0^t \sqrt{\sigma^2(\tilde{X}_s^D(i))}\,dW_s(i).\end{aligned} \tag{2.6}$$

The stochastic integrals on the right-hand side of (2.6) are martingales since the integrands are globally bounded. Hence, (2.6) and Tonelli's theorem imply for all $D \in \mathbb{N}$ and all $t \in [0, \infty)$ that

$$\mathbb{E}\left[\sum_{i=1}^{D}\tilde{X}_t^D(i)\right] \leq \mathbb{E}\left[\sum_{i=1}^{D}\tilde{X}_0^D(i)\right] + 2\mu t + (L_f + L_h)\int_0^t \mathbb{E}\left[\sum_{i=1}^{D}\tilde{X}_s^D(i)\right]ds.$$

Gronwall's inequality then yields for all $D \in \mathbb{N}$ and all $t \in [0, \infty)$ that

$$\mathbb{E}\left[\sum_{i=1}^{D}\tilde{X}_t^D(i)\right] \leq e^{(L_f+L_h)t}\left(\mathbb{E}\left[\sum_{i=1}^{D}\tilde{X}_0^D(i)\right] + 2\mu t\right).$$

Taking the supremum over $t \in [0, T]$ and using (2.3) proves (2.5). The proof for the loop-free processes is similar. The proof of Lemma 2.3 is thus completed. □

The following lemma is a variant of Lemma 2.3 and, heuristically speaking, shows that uniformly in the number of demes essentially only finitely many migration levels contribute to the total mass.

LEMMA 2.4 (Essentially only finitely many levels). *Assume that* Setting 2.2 *holds and that*

$$\sum_{k\in\mathbb{N}_0}\sup_{D\in\mathbb{N}} \mathbb{E}\left[\sum_{i=1}^{D} X_0^{D,k}(i)\right] < \infty. \tag{2.7}$$

Then we have for all $T \in [0, \infty)$ *that*

$$\sum_{k\in\mathbb{N}_0}\sup_{D\in\mathbb{N}}\sup_{t\in[0,T]} \mathbb{E}\left[\sum_{i=1}^{D} X_t^{D,k}(i)\right] \leq e^{(L_f+L_h)T}\left(\sum_{k\in\mathbb{N}_0}\sup_{D\in\mathbb{N}} \mathbb{E}\left[\sum_{i=1}^{D} X_0^{D,k}(i)\right] + 2\mu T\right) \tag{2.8}$$

and (2.8) *holds with* $X_t^{D,k}(i)$ *replaced by* $Z_t^{D,k}(i)$.

PROOF. In the situation of Setting 1.1, it holds for all $D \in \mathbb{N}$ and all $x, y \in [0,1]$ that $f(y,x) \le L_f$, that $\tilde{h}_D(x) \le L_h x$, and that $D h_D(0) \le 2\mu$. Moreover, the stochastic integral part of (2.1) yields a martingale. These facts, (2.1), and Tonelli's theorem show for all $k \in \mathbb{N}_0$, all $D \in \mathbb{N}$, and all $t \in [0,\infty)$ that

$$\mathbb{E}\Bigg[\sum_{i=1}^{D} X_t^{D,k}(i)\Bigg] \le \mathbb{E}\Bigg[\sum_{i=1}^{D} X_0^{D,k}(i)\Bigg] + \mathbf{1}_{k=0} 2\mu t + \int_0^t L_f \mathbb{E}\Bigg[\sum_{i=1}^{D} X_s^{D,k-1}(i)\Bigg] + L_h \mathbb{E}\Bigg[\sum_{i=1}^{D} X_s^{D,k}(i)\Bigg]\, ds.$$

This implies for all $T \in [0,\infty)$ and all $k \in \mathbb{N}_0$ that

$$\begin{aligned}\sup_{D\in\mathbb{N}} \sup_{t\in[0,T]} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_t^{D,k}(i)\Bigg] \le{}& \sup_{D\in\mathbb{N}} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_0^{D,k}(i)\Bigg] + \mathbf{1}_{k=0} 2\mu T \\ &+ L_f \int_0^T \sup_{D\in\mathbb{N}} \sup_{u\in[0,s]} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_u^{D,k-1}(i)\Bigg] ds \\ &+ L_h \int_0^T \sup_{D\in\mathbb{N}} \sup_{u\in[0,s]} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_u^{D,k}(i)\Bigg] ds.\end{aligned} \tag{2.9}$$

Lemma 2.3 and (2.7) show that the right-hand side of (2.9) is finite. For every $K \in \mathbb{N}$ a summation of (2.9) over $k \in [K]_0$ and Gronwall's inequality yield for all $T \in [0,\infty)$ that

$$\begin{aligned}&\sum_{k=0}^{K} \sup_{D\in\mathbb{N}} \sup_{t\in[0,T]} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_t^{D,k}(i)\Bigg] \\ &\quad \le e^{(L_f+L_h)T} \Bigg(\sum_{k=0}^{K} \sup_{D\in\mathbb{N}} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_0^{D,k}(i)\Bigg] + 2\mu T\Bigg).\end{aligned}$$

Taking the limit as $K \to \infty$ proves (2.8). The proof for the loop-free processes is similar. This completes the proof of Lemma 2.4. □

The following lemma gives an estimate for the second moment of the total mass process.

LEMMA 2.5 (Second moment). *Assume that* Setting 2.2 *holds. Then we have for all* $D \in \mathbb{N}$ *and all* $T \in [0,\infty)$ *that*

$$\begin{aligned}&\mathbb{E}\Bigg[\sup_{t\in[0,T]} \Bigg(\sum_{i=1}^{D} \sum_{k\in\mathbb{N}_0} X_t^{D,k}(i)\Bigg)^2\Bigg] \\ &\quad \le e^{(8L_\sigma + 4(L_f+L_h)^2 T)T} \Bigg(4\mathbb{E}\Bigg[\Bigg(\sum_{i=1}^{D} \sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)\Bigg)^2\Bigg] + 8T\big(L_\sigma + 2\mu^2 T\big)\Bigg)\end{aligned} \tag{2.10}$$

and (2.10) *holds with* $X_t^{D,k}(i)$ *replaced by* $Z_t^{D,k}(i)$.

PROOF. For all $D \in \mathbb{N}$ let $\{W(i) : i \in [D]\}$ and $\tilde{X}^D$ be as in Lemma 2.1. Setting 1.1 implies for all $D \in \mathbb{N}$ and all $x, y \in [0,1]$ that $|f(y,x)| \le L_f$ and $|h_D(x)| \le L_h x + 2\mu/D$.

This and Lemma 2.1 yield for all $D \in \mathbb{N}$ and all $t \in [0, \infty)$ that a.s.

$$\left|\sum_{i=1}^{D} \tilde{X}_t^D(i)\right| \le \left|\sum_{i=1}^{D} \tilde{X}_0^D(i)\right| + (L_f + L_h) \int_0^t \left|\sum_{i=1}^{D} \tilde{X}_s^D(i)\right| ds + 2\mu t + \left|\sum_{i=1}^{D} \int_0^t \sqrt{\sigma^2(\tilde{X}_s^D(i))}\, dW_s(i)\right|.$$

The Minkowski inequality then implies for all $D \in \mathbb{N}$ and all $T \in [0, \infty)$ that

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} \left(\sum_{i=1}^{D} \tilde{X}_t^D(i)\right)^2\right]^{\frac{1}{2}} &\le \mathbb{E}\left[\left(\sum_{i=1}^{D} \tilde{X}_0^D(i)\right)^2\right]^{\frac{1}{2}} + 2\mu T \\ &\quad + (L_f + L_h) \int_0^T \mathbb{E}\left[\left(\sum_{i=1}^{D} \tilde{X}_s^D(i)\right)^2\right]^{\frac{1}{2}} ds \\ &\quad + \mathbb{E}\left[\sup_{t\in[0,T]} \left|\sum_{i=1}^{D} \int_0^t \sqrt{\sigma^2(\tilde{X}_s^D(i))}\, dW_s(i)\right|^2\right]^{\frac{1}{2}}.\end{aligned} \tag{2.11}$$

Using Doob's L^2-inequality (see, e.g., Corollary 2.2.17 in [7]), the Itô isometry, Setting 1.1, and the fact that for all $x \in \mathbb{R}$ it holds that $2x \le 1 + x^2$, we obtain for all $D \in \mathbb{N}$ and all $T \in [0, \infty)$ that

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\in[0,T]} \left|\sum_{i=1}^{D} \int_0^t \sqrt{\sigma^2(\tilde{X}_s^D(i))}\, dW_s(i)\right|^2\right] \\ &\le 4\int_0^T L_\sigma \mathbb{E}\left[\sum_{i=1}^{D} \tilde{X}_s^D(i)\right] ds \le 2L_\sigma T + 2L_\sigma \int_0^T \mathbb{E}\left[\left(\sum_{i=1}^{D} \tilde{X}_s^D(i)\right)^2\right] ds.\end{aligned} \tag{2.12}$$

Equations (2.11) and (2.12), the fact that it holds for all $x_1, \dots, x_4 \in \mathbb{R}$ that $(\sum_{i=1}^4 x_i)^2 \le 4\sum_{i=1}^4 x_i^2$, and Hölder's inequality yield for all $D \in \mathbb{N}$ and all $T \in [0, \infty)$ that

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\in[0,T]} \left(\sum_{i=1}^{D} \tilde{X}_t^D(i)\right)^2\right] \\ &\le 4\mathbb{E}\left[\left(\sum_{i=1}^{D} \tilde{X}_0^D(i)\right)^2\right] + 16\mu^2 T^2 \\ &\quad + 4(L_f + L_h)^2 \left(\int_0^T \mathbb{E}\left[\left(\sum_{i=1}^{D} \tilde{X}_s^D(i)\right)^2\right]^{\frac{1}{2}} ds\right)^2 \\ &\quad + 8L_\sigma T + 8L_\sigma \int_0^T \mathbb{E}\left[\left(\sum_{i=1}^{D} \tilde{X}_s^D(i)\right)^2\right] ds \\ &\le 4\mathbb{E}\left[\left(\sum_{i=1}^{D} \tilde{X}_0^D(i)\right)^2\right] + 8T(L_\sigma + 2\mu^2 T) \\ &\quad + (8L_\sigma + 4(L_f + L_h)^2 T) \int_0^T \mathbb{E}\left[\sup_{u\in[0,s]} \left(\sum_{i=1}^{D} \tilde{X}_u^D(i)\right)^2\right] ds.\end{aligned}$$

This, Gronwall's inequality, and (2.3) prove (2.10). The proof for the loop-free processes is similar. This completes the proof of Lemma 2.5. $\square$

The following lemma is a consequence of Lemma 2.5 and allows for a localization procedure in the total mass uniformly in the number of demes.

LEMMA 2.6 (Localization argument). *Assume that* Setting 2.2 *and*

$$\sup_{D\in\mathbb{N}} \mathbb{E}\left[\left(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)\right)^2\right] < \infty \tag{2.13}$$

hold and for every $D, M \in \mathbb{N}$ define the stopping time

$$\tau_M^D := \inf\left\{t \in [0,\infty) : \sum_{i=1}^{D}\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i) \geq M\right\}. \tag{2.14}$$

Then it holds for all $T \in [0,\infty)$ that

$$\lim_{M\to\infty} \sup_{D\in\mathbb{N}} \mathbb{E}\left[\sup_{t\in[0,T]} \sum_{i=1}^{D}\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)\mathbf{1}_{\{\tau_M^D\leq T\}}\right] = 0. \tag{2.15}$$

PROOF. For all $D, M \in \mathbb{N}$ and $T \in [0,\infty)$ we have that

$$\{\tau_M^D \leq T\} = \left\{\sup_{t\in[0,T]} \sum_{i=1}^{D}\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i) \geq M\right\}.$$

This implies for all $D, M \in \mathbb{N}$ and all $T \in [0,\infty)$ that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]} \sum_{i=1}^{D}\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)\mathbf{1}_{\{\tau_M^D\leq T\}}\right] \\
&\quad\leq \frac{1}{M}\mathbb{E}\left[\sup_{t\in[0,T]}\left(\sum_{i=1}^{D}\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)\right)^2\right].
\end{aligned}$$

This, Lemma 2.5, and (2.13) show (2.15). The proof of Lemma 2.6 is thus completed. $\square$

Throughout the rest of this subsection and in Section 2.3 below, the following Setting 2.7 will frequently be referred to. In the situation of Setting 2.7, for every $D \in \mathbb{N}$ the SDE (2.17) below with $g = g_D$ has a unique strong solution with continuous sample paths in $[0, 1]$; see, for example, Theorem 5.4.22, Proposition 5.2.13, and Corollary 5.3.23 in [16]. This fact will tacitly be used in the remainder of this paper.

SETTING 2.7 (Time-dependent immigration). Assume that Setting 1.1 holds and that $g_D : [0,\infty) \times [0,1] \to \mathbb{R}$, $D \in \mathbb{N}$, are measurable functions that satisfy for all $D \in \mathbb{N}$ and all $t \in [0,\infty)$ that $g_D(t,0) \geq 0$, that $\frac{1}{D}g_D(t,1) + \tilde{h}_D(1) \leq 0$, that

$$\sup_{u\in[0,\infty)} \sup_{\substack{x,y\in[0,1]\\ x\neq y}} \frac{|g_D(u,x) - g_D(u,y)|}{|x-y|} < \infty,$$

and that

$$\sup_{M\in\mathbb{N}} \int_0^t \sup_{x\in[0,1]} \left|g_M(u,x)\right|^2 du < \infty. \tag{2.16}$$

For all $s \in [0, \infty)$, all $D \in \mathbb{N}$, and all measurable functions $g : [0, \infty) \times [0, 1] \to \mathbb{R}$ we consider the one-dimensional SDE

$$dY_{t,s}^{D,g} = \frac{1}{D} g\big(t, Y_{t,s}^{D,g}\big)\,dt + \tilde{h}_D\big(Y_{t,s}^{D,g}\big)\,dt + \sqrt{\sigma^2\big(Y_{t,s}^{D,g}\big)}\,dW_t, \quad t \in [s, \infty), \tag{2.17}$$

where W is a standard Brownian motion. We adopt the same notation and write, for example, $Y_{t,s}^{D,\zeta}$ (or $Y_{t,s}^{D,c}$) when the function $g : [0, \infty) \times [0, 1] \to \mathbb{R}$ in (2.17) is replaced by a function $\zeta : [0, \infty) \to \mathbb{R}$ (or by a constant $c \in \mathbb{R}$).

The following lemma estimates the L^1-distance between certain solutions of (2.17).

LEMMA 2.8 (L^1-regularity). *Assume* Setting 1.1, *let* $g_D : [0, \infty) \times [0, 1] \to \mathbb{R}$, $D \in \mathbb{N}$, *and* $\tilde{g}_D : [0, \infty) \times [0, 1] \to \mathbb{R}$, $D \in \mathbb{N}$, *be two sequences of functions satisfying* Setting 2.7, *let* $s \in [0, \infty)$, *and for every* $D \in \mathbb{N}$ *let* $(Y_{t,s}^{D,g_D})_{t \in [s,\infty)}$ *and* $(Y_{t,s}^{D,\tilde{g}_D})_{t \in [s,\infty)}$ *be solutions of* (2.17) *with respect to the same Brownian motion. Then it holds for all* $D \in \mathbb{N}$ *and all* $t \in [s, \infty)$ *that*

$$\begin{aligned}
&\mathbb{E}\big[\big|Y_{t,s}^{D,g_D} - Y_{t,s}^{D,\tilde{g}_D}\big|\big] \\
&\quad \le e^{L_h(t-s)}\bigg(\mathbb{E}\big[\big|Y_{s,s}^{D,g_D} - Y_{s,s}^{D,\tilde{g}_D}\big|\big] \\
&\qquad + \frac{1}{D}\int_s^t \mathbb{E}\big[\big|g_D\big(u, Y_{u,s}^{D,g_D}\big) - \tilde{g}_D\big(u, Y_{u,s}^{D,\tilde{g}_D}\big)\big|\big]\,du\bigg).
\end{aligned}$$

PROOF. As in Theorem 1 in [29] (see also, e.g., the proof of Lemma 3.3 in [14]) an approximation of $\mathbb{R} \ni x \mapsto |x| \in \mathbb{R}$ with C^2-functions and exploiting that

$$\sup_{x,y \in [0,1], x \ne y} \frac{|\sqrt{\sigma^2(x)} - \sqrt{\sigma^2(y)}|^2}{|x - y|} < \infty$$

shows for all $D \in \mathbb{N}$ and all $t \in [s, \infty)$ that a.s.

$$\begin{aligned}
\big|Y_{t,s}^{D,g_D} - Y_{t,s}^{D,\tilde{g}_D}\big| &= \big|Y_{s,s}^{D,g_D} - Y_{s,s}^{D,\tilde{g}_D}\big| \\
&\quad + \int_s^t \operatorname{sgn}\big(Y_{u,s}^{D,g_D} - Y_{u,s}^{D,\tilde{g}_D}\big)\,d\big(Y_{u,s}^{D,g_D} - Y_{u,s}^{D,\tilde{g}_D}\big).
\end{aligned} \tag{2.18}$$

For every $D \in \mathbb{N}$ and every $t \in [s, \infty)$ let M_t^D be a real-valued random variable satisfying a.s. that

$$M_t^D = \int_s^t \operatorname{sgn}\big(Y_{u,s}^{D,g_D} - Y_{u,s}^{D,\tilde{g}_D}\big)\Big(\sqrt{\sigma^2\big(Y_{u,s}^{D,g_D}\big)} - \sqrt{\sigma^2\big(Y_{u,s}^{D,\tilde{g}_D}\big)}\Big)\,dW_u. \tag{2.19}$$

Then (2.18) and Setting 1.1 imply for all $D \in \mathbb{N}$ and all $t \in [s, \infty)$ that a.s.

$$\begin{aligned}
\big|Y_{t,s}^{D,g_D} - Y_{t,s}^{D,\tilde{g}_D}\big| &\le \big|Y_{s,s}^{D,g_D} - Y_{s,s}^{D,\tilde{g}_D}\big| \\
&\quad + \frac{1}{D}\int_s^t \big|g_D\big(u, Y_{u,s}^{D,g_D}\big) - \tilde{g}_D\big(u, Y_{u,s}^{D,\tilde{g}_D}\big)\big|\,du \\
&\quad + L_h \int_s^t \big|Y_{u,s}^{D,g_D} - Y_{u,s}^{D,\tilde{g}_D}\big|\,du + M_t^D.
\end{aligned} \tag{2.20}$$

Since the integrand of the stochastic integral in (2.19) is globally bounded, it holds for all $D \in \mathbb{N}$ and all $t \in [s, \infty)$ that $\mathbb{E}[M_t^D] = 0$. Therefore, (2.20) and Tonelli's theorem imply for all $D \in \mathbb{N}$ and all $t \in [s, \infty)$ that

$$\begin{aligned}\mathbb{E}[|Y_{t,s}^{D,g_D} - Y_{t,s}^{D,\tilde{g}_D}|] \leq \mathbb{E}[|Y_{s,s}^{D,g_D} - Y_{s,s}^{D,\tilde{g}_D}|] \\ + \frac{1}{D}\int_s^t \mathbb{E}[|g_D(u, Y_{u,s}^{D,g_D}) - \tilde{g}_D(u, Y_{u,s}^{D,\tilde{g}_D})|]\, du \\ + L_h \int_s^t \mathbb{E}[|Y_{u,s}^{D,g_D} - Y_{u,s}^{D,\tilde{g}_D}|]\, du.\end{aligned}$$

Gronwall's inequality then finishes the proof of Lemma 2.8. □

The following lemma provides us with a second moment estimate for solutions of (2.17) and is similar to Lemma 2.5. Its proof is completely analogous to that of Lemma 2.5 and thus omitted here.

LEMMA 2.9 (Second moment). *Assume that* Setting 2.7 *holds*, *let* $s \in [0, \infty)$, *let* $\{W(i) : i \in \mathbb{N}\}$ *be a set of independent standard Brownian motions*, *and for every* $D \in \mathbb{N}$ *and every* $i \in [D]$ *let* $(Y_{t,s}^{D,g_D}(i))_{t\in[s,\infty)}$ *be a solution of* (2.17) *driven by* $W(i)$ *such that it holds a.s. that* $Y_{s,s}^{D,g_D}(i) = 0$. *Then it holds for all* $D \in \mathbb{N}$ *and all* $T \in [s, \infty)$ *that*

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\in[s,T]}\left(\sum_{i=1}^{D} Y_{t,s}^{D,g_D}(i)\right)^2\right] \\ &\quad \leq 3(T-s)e^{(6L_\sigma+3L_h^2(T-s))(T-s)}\left(\int_s^T \sup_{x\in[0,1]}|g_D(u,x)|^2\, du + 2L_\sigma\right).\end{aligned}$$

The following lemma follows from a straightforward adaptation of Lemma 9.8 in [10] to the state space [0, 1].

LEMMA 2.10 (Finite excursion area). *Assume that* Setting 1.2 *holds*. *Then it holds that*

$$\int\int_0^\infty \eta_t\, dt\, Q(d\eta) = \int_0^1 \frac{y}{\frac{1}{2}\sigma^2(y)s(y)}\, dy < \infty.$$

2.3. *Poisson limit of independent diffusions with vanishing immigration*. To show the convergence of the loop-free processes in Section 2.4 below, we first prove a Poisson limit for independent diffusions with vanishing immigration, see Lemma 2.15 below, based on the following lemma which is essentially Lemma 4.19 in [11] and proved there utilizing a time reversion argument.

LEMMA 2.11 (Poisson limit, constant case). *Assume that* Setting 1.2 *holds*, *let* $c, s \in [0, \infty)$, *let* $D_0 \in \mathbb{N}$ *be such that it holds for all* $D \in \mathbb{N} \cap [D_0, \infty)$ *that* $c/D + \tilde{h}_D(1) \leq 0$, *for every* $D \in \mathbb{N} \cap [D_0, \infty)$ *let* $(Y_{t,s}^{D,c})_{t\in[s,\infty)}$ *be a solution of* (2.17) *satisfying a.s. that* $Y_{s,s}^{D,c} = 0$, *and let* $\phi_D : [0, 1] \to \mathbb{R}$, $D \in \mathbb{N}_0$, *be functions with the property that*

$$\sup_{\substack{x,y\in[0,1]\\ x\neq y}} \sup_{D\in\mathbb{N}} \frac{|\phi_D(x) - \phi_D(y)|}{|x-y|} < \infty,$$

that $\lim_{D\to\infty} D|\phi_D(0)| = 0$, *and for all* $y \in [0, 1]$ *that* $\lim_{D\to\infty} \phi_D(y) = \phi_0(y)$. *Then it holds for all* $t \in [s, \infty)$ *that*

$$\lim_{D\to\infty} D\mathbb{E}[\phi_D(Y_{t,s}^{D,c})] = c\int_s^t \int \phi_0(\eta_{t-u})Q(d\eta)\, du.$$

For every $T \in (0,\infty)$ and every $s \in [0,T)$ we define

$$\mathcal{E}_{s,T} := \left\{ \begin{array}{c} C([s,T],[0,1]) \ni \eta \mapsto \prod_{i=1}^{n} \psi_i(\eta_{t_i}) \in \mathbb{R} : \\ n \in \mathbb{N}, \psi_1,\ldots,\psi_n \in \mathrm{Lip}([0,1],\mathbb{R}), \\ \text{and } t_1,\ldots,t_n \in [s,T] \text{ with } t_1 \leq \cdots \leq t_n \end{array} \right\}. \tag{2.21}$$

From the Lipschitz continuity and boundedness of the involved functions in (2.21), it follows for all $T \in (0,\infty)$ and all $s \in [0,T)$ that the elements of $\mathcal{E}_{s,T}$ are globally Lipschitz continuous in the sense of Lemma 2.12. The proof of Lemma 2.12 is straightforward and therefore omitted.

LEMMA 2.12 (Lipschitz continuity). *Let $n \in \mathbb{N}$ and let $\psi_1,\ldots,\psi_n \in \mathrm{Lip}([0,1],\mathbb{R})$. Then there exists a constant $L \in [0,\infty)$ such that it holds for all $x_1,\ldots,x_n \in [0,1]$ and all $y_1,\ldots,y_n \in [0,1]$ that*

$$\left| \prod_{i=1}^{n} \psi_i(x_i) - \prod_{i=1}^{n} \psi_i(y_i) \right| \leq L \sum_{i=1}^{n} |x_i - y_i|.$$

The following two lemmas generalize Lemma 2.11 in a suitable way. The proof of Lemma 2.13 below is analogous to the proof of Lemma 4.20 in [11] and therefore omitted here.

LEMMA 2.13 (Poisson limit, piecewise constant case). *Assume that* Setting 1.2 *holds and let $T \in (0,\infty)$. Then for all $n \in \mathbb{N}$, all $s \in [0,T)$, all $c_1,\ldots,c_n \in [0,\infty)$, all $\psi_1,\ldots,\psi_n \in \mathrm{Lip}([0,1],\mathbb{R})$, all $t_0,\ldots,t_n \in [s,T]$ with $s = t_0 \leq t_1 \leq \cdots \leq t_n \leq T$, all $\zeta : [0,\infty) \to [0,\infty)$ satisfying for all $t \in [0,\infty)$ that $\zeta(t) = \sum_{i=1}^{n} c_i \mathbf{1}_{[t_{i-1},t_i)}(t)$, all $D_0 \in \mathbb{N}$ such that it holds for all $D \in \mathbb{N} \cap [D_0,\infty)$ that $\max\{c_1,\ldots,c_n\}/D + \tilde{h}_D(1) \leq 0$, all solutions $(Y_{t,s}^{D,\zeta})_{t\in[s,\infty)}$, $D \in \mathbb{N} \cap [D_0,\infty)$, of (2.17) satisfying for all $D \in \mathbb{N} \cap [D_0,\infty)$ a.s. that $Y_{s,s}^{D,\zeta} = 0$, and all $F \in \mathcal{E}_{s,T}$ satisfying for all $\eta \in C([s,T],[0,1])$ that $F(\eta) = \prod_{i=1}^{n} \psi_i(\eta_{t_i})$ and $F(0) = 0$ it holds that*

$$\lim_{D\to\infty} D\mathbb{E}\big[F\big((Y_{t,s}^{D,\zeta})_{t\in[s,T]}\big)\big] = \int_s^T \zeta(u) \int F\big((\eta_{t-u})_{t\in[s,T]}\big) Q(d\eta)\, du.$$

The following lemma generalizes Lemma 2.13 and is a crucial ingredient in the proof of Lemma 2.15 below.

LEMMA 2.14 (Poisson limit, general case). *Assume* Setting 1.2 *and* Setting 2.7, *let $T \in (0,\infty)$, let $s \in [0,T)$, for every $D \in \mathbb{N}$ let $(Y_{t,s}^{D,g_D})_{t\in[s,\infty)}$ be a solution of (2.17) satisfying a.s. that $Y_{s,s}^{D,g_D} = 0$, and let $g : [0,\infty) \times [0,1] \to \mathbb{R}$ be a measurable function satisfying for all $t \in [0,\infty)$ that $g(t,0) \geq 0$, that $[0,1] \ni x \mapsto g(t,x) \in \mathbb{R}$ is continuous, and that*

$$\lim_{D\to\infty} \int_s^T \sup_{x\in[0,1]} \big| g_D(u,x) - g(u,x) \big|\, du = 0. \tag{2.22}$$

Then it holds for all $F \in \mathcal{E}_{s,T}$ with $F(0) = 0$ that

$$\begin{aligned} &\lim_{D\to\infty} D\mathbb{E}\big[F\big((Y_{t,s}^{D,g_D})_{t\in[s,T]}\big)\big] \\ &\quad = \int_s^\infty g(u,0) \int F\big((\eta_{t-u})_{t\in[s,T]}\big) Q(d\eta)\, du \in \mathbb{R}. \end{aligned} \tag{2.23}$$

PROOF. In a first step, we assume that $(g_D)_{D\in\mathbb{N}}$ and g are uniformly bounded by $K\in\mathbb{N}$. Fix $F\in\mathcal{E}_{s,T}$ with $F(0)=0$ for the rest of the proof and let $m\in\mathbb{N}$, $\psi_1,\dots,\psi_m\in\mathrm{Lip}([0,1],\mathbb{R})$, and $t_1,\dots,t_m\in[s,T]$ with $t_1\le\cdots\le t_m$ be such that it holds for all $\eta\in C([s,T],[0,1])$ that $F(\eta)=\prod_{i=1}^m\psi_i(\eta_{t_i})$. We choose step functions $\zeta^{(n)}:[s,T]\to[0,\infty)$, $n\in\mathbb{N}$, with the property that $\zeta^{(n)}(\cdot)\to g(\cdot,0)$ almost everywhere as $n\to\infty$ and such that it holds for all $n\in\mathbb{N}$ that $\zeta^{(n)}\le K$. For every $n\in\mathbb{N}$ we extend $\zeta^{(n)}$ to $[0,\infty)$ by setting it to zero outside of $[s,T]$. Setting 1.1 implies the existence of $D_0\in\mathbb{N}$ such that we have for all $D\in\mathbb{N}\cap[D_0,\infty)$ that $K/D+\tilde{h}_D(1)\le 0$. For every $n\in\mathbb{N}$ and every $D\in\mathbb{N}\cap[D_0,\infty)$ let $(Y_{t,s}^{D,\zeta^{(n)}})_{t\in[s,\infty)}$ be a solution of (2.17) satisfying a.s. that $Y_{s,s}^{D,\zeta^{(n)}}=0$. Since we may let F depend trivially on further time points, Lemma 2.13 yields for every $n\in\mathbb{N}$ that

$$\lim_{D\to\infty}D\mathbb{E}\big[F\big((Y_{t,s}^{D,\zeta^{(n)}})_{t\in[s,T]}\big)\big]=\int_s^T\zeta^{(n)}(u)\int F\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du. \tag{2.24}$$

We are going to show that (2.24) converges to (2.23) as $n\to\infty$. Let $L_F\in[0,\infty)$ be a Lipschitz constant of F in the sense of Lemma 2.12. Then Lemma 2.12 and Lemma 2.8 applied with $\tilde{g}_D=\zeta^{(n)}$ imply for all $n\in\mathbb{N}$ and all $D\in\mathbb{N}\cap[D_0,\infty)$ that

$$\begin{aligned}
&\big|D\mathbb{E}\big[F\big((Y_{t,s}^{D,g_D})_{t\in[s,T]}\big)\big]-D\mathbb{E}\big[F\big((Y_{t,s}^{D,\zeta^{(n)}})_{t\in[s,T]}\big)\big]\big|\\
&\quad\le DL_F\sum_{i=1}^m\mathbb{E}\big[\big|Y_{t_i,s}^{D,g_D}-Y_{t_i,s}^{D,\zeta^{(n)}}\big|\big]\\
&\quad\le mL_Fe^{L_h(T-s)}\int_s^T\mathbb{E}\big[\big|g_D\big(u,Y_{u,s}^{D,g_D}\big)-\zeta^{(n)}(u)\big|\big]\,du\\
&\quad\le mL_Fe^{L_h(T-s)}\bigg(\int_s^T\sup_{x\in[0,1]}\big|g_D(u,x)-g(u,x)\big|\,du\\
&\qquad+\int_s^T\mathbb{E}\big[\big|g\big(u,Y_{u,s}^{D,g_D}\big)-g(u,0)\big|\big]\,du\\
&\qquad+\int_s^T\big|g(u,0)-\zeta^{(n)}(u)\big|\,du\bigg).
\end{aligned}\tag{2.25}$$

The first summand on the right-hand side of (2.25) converges to zero as $D\to\infty$ by (2.22). The dominated convergence theorem and the fact that $Y_{\cdot,s}^{D,g_D}$ converges to zero in distribution as $D\to\infty$ yield that the second summand on the right-hand side of (2.25) converges to zero as $D\to\infty$. Finally, the dominated convergence theorem ensures that the third summand on the right-hand side of (2.25) converges to zero as $n\to\infty$. Altogether, it follows that

$$\lim_{n\to\infty}\lim_{D\to\infty}\big|D\mathbb{E}\big[F\big((Y_{t,s}^{D,g_D})_{t\in[s,T]}\big)\big]-D\mathbb{E}\big[F\big((Y_{t,s}^{D,\zeta^{(n)}})_{t\in[s,T]}\big)\big]\big|=0.$$

This proves convergence of the left-hand side of (2.24) to the left-hand side of (2.23) as $n\to\infty$. Lemma 2.12, $F(0)=0$, and Lemma 2.10 ensure that $\int_s^T\int|F((\eta_{t-u})_{t\in[s,T]})|Q(d\eta)\,du<\infty$. This, the fact that we have for all $n\in\mathbb{N}$ that $\zeta^{(n)}\le K$, and the dominated convergence theorem show that

$$\begin{aligned}
&\lim_{n\to\infty}\int_s^T\zeta^{(n)}(u)\int F\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du\\
&\quad=\int_s^T g(u,0)\int F\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du.
\end{aligned}\tag{2.26}$$

It remains to note that $F(0)=0$ ensures

$$\int_s^T g(u,0)\int F\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du = \int_s^\infty g(u,0)\int F\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du. \tag{2.27}$$

Hence, (2.26) and (2.27) show that the right-hand sides of (2.24) and (2.23) are equal in the limit $n\to\infty$.

For the rest of the proof, we return to the case of general $(g_D)_{D\in\mathbb{N}}$ and g. For all $D, K\in\mathbb{N}$ let $(Y_{t,s}^{D,g_D\wedge K})_{t\in[s,\infty)}$ be a solution of (2.17) satisfying a.s. that $Y_{s,s}^{D,g_D\wedge K}=0$. It holds for all $K\in\mathbb{N}$ that

$$\begin{aligned}&\lim_{D\to\infty}\int_s^T \sup_{x\in[0,1]}\big|g_D(u,x)\wedge K-g(u,x)\wedge K\big|\,du\\&\quad\le \lim_{D\to\infty}\int_s^T \sup_{x\in[0,1]}\big|g_D(u,x)-g(u,x)\big|\,du=0.\end{aligned}\tag{2.28}$$

We note that (2.16) and (2.22) imply that

$$\int_s^T \sup_{x\in[0,1]}\big|g(u,x)\big|\,du<\infty. \tag{2.29}$$

Lemma 2.8 applied with $(\tilde g_D)_{D\in\mathbb{N}}=(g_D\wedge K)_{D\in\mathbb{N}}$, arguments as in (2.25), the dominated convergence theorem, and (2.29) then show that

$$\begin{aligned}&\varlimsup_{K\to\infty}\varlimsup_{D\to\infty}\big|D\mathbb{E}\big[F\big((Y_{t,s}^{D,g_D})_{t\in[s,T]}\big)\big]-D\mathbb{E}\big[F\big((Y_{t,s}^{D,g_D\wedge K})_{t\in[s,T]}\big)\big]\big|\\&\quad\le mL_Fe^{L_h(T-s)}\varlimsup_{K\to\infty}\int_s^T\big|g(u,0)-g(u,0)\wedge K\big|\,du=0.\end{aligned}\tag{2.30}$$

For all $i\in[m]$ and $x\in[0,1]$ we write $\psi_i(x)=\psi_i(x)^+-\psi_i(x)^-$ to obtain a decomposition of F of the form $F=F^+-F^-$, where F^+ and F^- are finite sums of nonnegative functions in $\mathcal{E}_{s,T}$ and satisfy $F^+(0)=0=F^-(0)$. Due to this and (2.28), the first part of the proof yields for all $K\in\mathbb{N}$ that

$$\begin{aligned}&\lim_{D\to\infty}D\mathbb{E}\big[F^+\big((Y_{t,s}^{D,g_D\wedge K})_{t\in[s,T]}\big)\big]\\&\quad=\int_s^\infty\big(g(u,0)\wedge K\big)\int F^+\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du.\end{aligned}\tag{2.31}$$

The monotone convergence theorem ensures that

$$\begin{aligned}&\lim_{K\to\infty}\int_s^\infty\big(g(u,0)\wedge K\big)\int F^+\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du\\&\quad=\int_s^\infty g(u,0)\int F^+\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du.\end{aligned}\tag{2.32}$$

Moreover, Lemma 2.12, $F(0)=0$, Lemma 2.8, (2.28), and (2.29) yield for all $K\in\mathbb{N}$ that

$$\begin{aligned}&\lim_{D\to\infty}D\mathbb{E}\big[F^+\big((Y_{t,s}^{D,g_D\wedge K})_{t\in[s,T]}\big)\big]\\&\quad\le L_Fme^{L_h(T-s)}\int_s^T\sup_{x\in[0,1]}\big|g(u,x)\wedge K\big|\,du\\&\quad\le L_Fme^{L_h(T-s)}\int_s^T\sup_{x\in[0,1]}\big|g(u,x)\big|\,du<\infty,\end{aligned}$$

which together with (2.31) yields that

$$\sup_{K\in\mathbb{N}}\int_s^\infty (g(u,0)\wedge K)\int F^+\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du<\infty.$$

The same is true when we replace F^+ by F^-. This implies that

$$\int_s^\infty g(u,0)\int F\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du$$

is well defined as a real number. Hence, combining (2.30), (2.31), and (2.32) for F^+ and F^- yields that

$$\begin{aligned}
&\lim_{D\to\infty} D\mathbb{E}\big[F\big((Y_{t,s}^{D,g_D})_{t\in[s,T]}\big)\big]\\
&\quad=\lim_{K\to\infty}\lim_{D\to\infty} D\mathbb{E}\big[F\big((Y_{t,s}^{D,g_D\wedge K})_{t\in[s,T]}\big)\big]\\
&\quad=\int_s^\infty g(u,0)\int F\big((\eta_{t-u})_{t\in[s,T]}\big)Q(d\eta)\,du\in\mathbb{R}.
\end{aligned}$$

This finishes the proof of Lemma 2.14. $\square$

With Lemma 2.14 in hand, we are ready to prove the following Poisson limit lemma for independent diffusions with vanishing immigration, which generalizes Lemma 4.21 in [11] to state-dependent g.

LEMMA 2.15 (Poisson limit for independent diffusions with vanishing immigration). *Assume that* Setting 1.2 *and* Setting 2.7 *hold*, *let* $s\in[0,\infty)$, *let* $\{W(i):i\in\mathbb{N}\}$ *be a set of independent standard Brownian motions*, *for every* $D\in\mathbb{N}$ *and every* $i\in[D]$ *let* $(Y_{t,s}^{D,g_D}(i))_{t\in[s,\infty)}$ *be a solution of* (2.17) *driven by* $W(i)$ *such that it holds a.s. that* $Y_{s,s}^{D,g_D}(i)=0$, *let* $g:[0,\infty)\times[0,1]\to\mathbb{R}$ *be a measurable function satisfying for all* $t\in[0,\infty)$ *that* $g(t,0)\geq 0$, *that* $[0,1]\ni x\mapsto g(t,x)\in\mathbb{R}$ *is continuous*, *and that*

$$\lim_{D\to\infty}\int_s^{t\vee s}\sup_{x\in[0,1]}\big|g_D(u,x)-g(u,x)\big|\,du=0,$$

and let Π *be a Poisson point process on* $[s,\infty)\times U$ *with intensity measure* $\mathbb{E}[\Pi(du\otimes d\eta)]=g(u,0)\,du\otimes Q(d\eta)$. *Then it holds that*

$$\left(\sum_{i=1}^D Y_{t,s}^{D,g_D}(i)\delta_{Y_{t,s}^{D,g_D}(i)}\right)_{t\in[s,\infty)}\xRightarrow{D\to\infty}\left(\int \eta_{t-u}\delta_{\eta_{t-u}}\Pi(du\otimes d\eta)\right)_{t\in[s,\infty)}$$

in the sense of convergence in distribution on $D([s,\infty),\mathcal{M}_{\mathrm{f}}([0,1]))$.

PROOF. Fix $\varphi\in C^2([0,1],\mathbb{R})$ for the rest of this paragraph. We define the function $\phi:[0,1]\to\mathbb{R}$ by $[0,1]\ni x\mapsto\phi(x):=x\varphi(x)\in\mathbb{R}$ and for all $D\in\mathbb{N}$ and all $t\in[s,\infty)$ we define $S_{t,s}^D:=\sum_{i=1}^D\phi(Y_{t,s}^{D,g_D}(i))$. The fact that there exists a constant $L_\phi\in[0,\infty)$ such that for all $x\in[0,1]$ it holds that $|\phi(x)|\leq L_\phi x$ and Markov's inequality yield for all $D,K\in\mathbb{N}$ and all $t\in[s,\infty)$ that

$$\begin{aligned}
\mathbb{P}\big(|S_{t,s}^D|\geq K\big)&\leq\frac{1}{K}\mathbb{E}\left[\sum_{i=1}^D\big|\phi\big(Y_{t,s}^{D,g_D}(i)\big)\big|\right]\\
&\leq\frac{L_\phi}{K}\mathbb{E}\left[\sum_{i=1}^D Y_{t,s}^{D,g_D}(i)\right]=\frac{L_\phi}{K}D\mathbb{E}\big[Y_{t,s}^{D,g_D}(1)\big].
\end{aligned}$$

This, Lemma 2.8, and Setting 2.7 imply for all $t \in [s, \infty)$ that

$$\lim_{K\to\infty} \sup_{D\in\mathbb{N}} \mathbb{P}(|S^D_{t,s}| \geq K) \leq \lim_{K\to\infty} \frac{L_\phi}{K} e^{L_h(t-s)} \sup_{D\in\mathbb{N}} \int_s^t \sup_{x\in[0,1]} |g_D(u,x)|\, du = 0. \tag{2.33}$$

For every $T \in (s, \infty)$ and every $D \in \mathbb{N}$ let $\mathcal{S}^D_T$ be the set of all stopping times with respect to the natural filtration of $S^D_{\cdot,s}$ that are bounded by T. For all $y \in [0, 1]$, all $D \in \mathbb{N}$, and all $t \in [s, \infty)$ we define

$$(G^D_t \phi)(y) := \left(\frac{1}{D} g_D(t, y) + \tilde{h}_D(y)\right)\phi'(y) + \frac{1}{2}\sigma^2(y)\phi''(y).$$

Setting 1.1 and $\phi \in C^2([0, 1], \mathbb{R})$ imply that there exists a constant $C_\phi \in [1, \infty)$ such that it holds for all $y \in [0, 1]$, all $D \in \mathbb{N}$, and all $t \in [s, \infty)$ that

$$|(G^D_t \phi)(y)| \leq C_\phi\left(\frac{1}{D} \sup_{x\in[0,1]} |g_D(t, x)| + y\right) \quad \text{and} \quad \sigma^2(y)\phi'^2(y) \leq C^2_\phi y,$$

Jensen's inequality implies for all $x_1, x_2, x_3 \in \mathbb{R}$ that $(\sum_{i=1}^3 x_i)^2 \leq 3\sum_{i=1}^3 x_i^2$, Hölder's inequality shows for every $\delta \in [0, \infty)$ and every integrable function $\alpha : [0, \delta] \to \mathbb{R}$ that $(\int_0^\delta \alpha(u)\, du)^2 \leq \delta \int_0^\delta (\alpha(u))^2\, du$, and it holds for all $x \in \mathbb{R}$ that $x \leq 1 + x^2$. Itô's formula, the Itô isometry, and the preceding estimates show for all $T \in (s, \infty)$, all $\bar{\delta} \in [0, 1]$, all $D \in \mathbb{N}$, all $\tau \in \mathcal{S}^D_T$, and all $\delta \in [0, \bar{\delta}]$ that

$$\begin{aligned}
&\mathbb{E}[(S^D_{\tau+\delta,s} - S^D_{\tau,s})^2] \\
&\quad = \mathbb{E}\left[\left(\sum_{i=1}^D \int_\tau^{\tau+\delta} (G^D_u\phi)(Y^{D,g_D}_{u,s}(i))\, du + \sum_{i=1}^D \int_\tau^{\tau+\delta} (\sqrt{\sigma^2}\phi')(Y^{D,g_D}_{u,s}(i))\, dW_u(i)\right)^2\right] \\
&\quad \leq 3C^2_\phi \mathbb{E}\left[\left(\int_\tau^{\tau+\delta} \sup_{x\in[0,1]} |g_D(u,x)|\, du\right)^2\right] + 3C^2_\phi \mathbb{E}\left[\left(\sum_{i=1}^D \int_0^\delta Y^{D,g_D}_{\tau+u,s}(i)\, du\right)^2\right] \\
&\qquad + 3\mathbb{E}\left[\sum_{i=1}^D \int_0^\delta \sigma^2(Y^{D,g_D}_{\tau+u,s}(i))\phi'^2(Y^{D,g_D}_{\tau+u,s}(i))\, du\right] \\
&\quad \leq 3C^2_\phi \delta \mathbb{E}\left[\int_\tau^{\tau+\delta} \sup_{x\in[0,1]} |g_D(u,x)|^2\, du\right] + 3C^2_\phi \delta \mathbb{E}\left[\int_0^\delta \left(\sum_{i=1}^D Y^{D,g_D}_{\tau+u,s}(i)\right)^2 du\right] \\
&\qquad + 3C^2_\phi \mathbb{E}\left[\int_0^\delta \sum_{i=1}^D Y^{D,g_D}_{\tau+u,s}(i)\, du\right] \\
&\quad \leq 3C^2_\phi \bar{\delta} \int_s^{T+1} \sup_{x\in[0,1]} |g_D(u,x)|^2\, du + 6C^2_\phi \bar{\delta} \mathbb{E}\left[\sup_{u\in[s,T+1]} \left(\sum_{i=1}^D Y^{D,g_D}_{u,s}(i)\right)^2\right] \\
&\qquad + 3C^2_\phi \bar{\delta}.
\end{aligned}$$

This, Lemma 2.9, and Setting 2.7 imply for all $T \in (s, \infty)$ that

$$\lim_{\bar{\delta}\to 0} \sup_{D\in\mathbb{N}} \sup_{\tau\in\mathcal{S}^D_T} \sup_{\delta\in[0,\bar{\delta}]} \mathbb{E}[(S^D_{\tau+\delta,s} - S^D_{\tau,s})^2] = 0. \tag{2.34}$$

By Aldous' tightness criterion (see, e.g., Theorem 3.8.6 in [7]), (2.33) and (2.34) ensure that

$$\left\{\left(\sum_{i=1}^D Y^{D,g_D}_{t,s}(i)\varphi(Y^{D,g_D}_{t,s}(i))\right)_{t\in[s,\infty)} : D \in \mathbb{N}\right\} \tag{2.35}$$

is relatively compact. Since $\varphi \in C^2([0,1],\mathbb{R})$ was arbitrary, it follows from (2.35), from Theorem 2.1 in [25], and from Prokhorov's theorem (see, e.g., Theorem 3.2.2 in [7]) that

$$\left\{\left(\sum_{i=1}^{D} Y_{t,s}^{D,g_D}(i)\delta_{Y_{t,s}^{D,g_D}(i)}\right)_{t\in[s,\infty)} : D \in \mathbb{N}\right\} \tag{2.36}$$

is relatively compact.

In the following, we identify the limit points of (2.36) by showing convergence of finite-dimensional distributions. For that, fix $n \in \mathbb{N}$, fix $\varphi_1,\dots,\varphi_n \in \mathrm{Lip}([0,1],[0,\infty))$, and fix $t_1,\dots,t_n \in [s,\infty)$ with $t_1 \le \cdots \le t_n$. For every $j \in [n]$ we define the function $\phi_j : [0,1] \to [0,\infty)$ by $[0,1] \ni x \mapsto \phi_j(x) := x\varphi_j(x) \in [0,\infty)$. The fact that $Y_{\cdot,s}^{D,g_D}(i)$, $i \in [D]$, are i.i.d. for all $D \in \mathbb{N}$ yields for all $D \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(-\sum_{j=1}^{n}\sum_{i=1}^{D}\phi_j\left(Y_{t_j,s}^{D,g_D}(i)\right)\right)\right] \\
&\quad= \prod_{i=1}^{D}\mathbb{E}\left[\exp\left(-\sum_{j=1}^{n}\phi_j\left(Y_{t_j,s}^{D,g_D}(i)\right)\right)\right] \\
&\quad= \left(1 - \frac{D\mathbb{E}[1-\exp(-\sum_{j=1}^{n}\phi_j(Y_{t_j,s}^{D,g_D}(1)))]}{D}\right)^D.
\end{aligned} \tag{2.37}$$

For all $x_1,\dots,x_n \in [0,1]$ it holds that

$$1-\exp\left(-\sum_{j=1}^{n}\phi_j(x_j)\right) = \sum_{j=1}^{n}\left(1-\exp(-\phi_j(x_j))\right)\exp\left(-\sum_{i=1}^{j-1}\phi_i(x_i)\right). \tag{2.38}$$

This shows that (2.37) involves the expectation of a sum. Each summand has the form of a functional $F \in \mathcal{E}_{s,t_n}$ with $F(0)=0$. On compact subintervals of $[0,\infty)$, the sequence of functions $x \mapsto (1-\frac{x}{D})^D$, $D \in \mathbb{N}$, converges uniformly to the function $x \mapsto e^{-x}$ as $D \to \infty$. This, Lemma 2.14 applied to each summand of the sum obtained from (2.37) and (2.38), and Campbell's formula (see, e.g., Theorem 24.14 in [17]) show that

$$\begin{aligned}
&\lim_{D\to\infty}\mathbb{E}\left[\exp\left(-\sum_{j=1}^{n}\sum_{i=1}^{D}\phi_j\left(Y_{t_j,s}^{D,g_D}(i)\right)\right)\right] \\
&\quad= \exp\left(-\lim_{D\to\infty} D\mathbb{E}\left[1-\exp\left(-\sum_{j=1}^{n}\phi_j\left(Y_{t_j,s}^{D,g_D}(1)\right)\right)\right]\right) \\
&\quad= \exp\left(-\int\int_s^{\infty}\left(1-\exp\left(-\sum_{j=1}^{n}\phi_j(\eta_{t_j-u})\right)\right)g(u,0)\,du\,Q(d\eta)\right) \\
&\quad= \mathbb{E}\left[\exp\left(-\sum_{j=1}^{n}\int\phi_j(\eta_{t_j-u})\Pi(du\otimes d\eta)\right)\right].
\end{aligned}$$

This implies the convergence of finite-dimensional distributions of (2.36) and completes the proof of Lemma 2.15. □

2.4. *Convergence of the loop-free processes*. In this subsection, we show convergence of the loop-free processes using Lemma 2.15. For that, we make the following assumption, which implies that the initial population has migration level zero and that its total mass has finite second moment.

SETTING 2.16 (Initial distribution). Assume that Setting 1.3 and Setting 2.2 hold, that

$$\mathbb{E}\left[\left(\sum_{i=1}^{\infty} X_0(i)\right)^2\right] < \infty, \tag{2.39}$$

and that it holds for all $D \in \mathbb{N}$ and all $i \in [D]$ that $\mathcal{L}(X_0^{D,0}(i)) = \mathcal{L}(X_0(i))$ and for all $D \in \mathbb{N}$ and all $(i,k) \in [D] \times \mathbb{N}$ that $\mathcal{L}(X_0^{D,k}(i)) = \delta_0$.

The following lemma establishes the convergence of the loop-free processes and is analogous to Lemma 4.22 in [11]. Recall that by Setting 2.2, which is satisfied by assumption in the following lemma, the loop-free processes fulfill for all $D \in \mathbb{N}$ and all $(i,k) \in [D] \times \mathbb{N}_0$ that a.s. $Z_0^{D,k}(i) = X_0^{D,k}(i)$.

LEMMA 2.17 (Convergence of the loop-free processes). *Assume that* Setting 2.16 *holds and let $\mathcal{T}$ be the forest of excursions constructed in* Section 1.2. *Then it holds that*

$$\left(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} Z_t^{D,k}(i)\delta_{Z_t^{D,k}(i)}\right)_{t\in[0,\infty)} \xRightarrow{D\to\infty} \left(\int \eta_{t-s}\delta_{\eta_{t-s}}\mathcal{T}(ds\otimes d\eta)\right)_{t\in[0,\infty)}$$

in the sense of convergence in distribution on $D([0,\infty),\mathcal{M}_{\mathrm{f}}([0,1]))$.

PROOF. Fix $\varphi \in C^2([0,1],\mathbb{R})$ for the rest of this paragraph. We define the function $\phi : [0,1] \to \mathbb{R}$ by $[0,1] \ni x \mapsto \phi(x) := x\varphi(x) \in \mathbb{R}$ and for all $D \in \mathbb{N}$ and all $t \in [0,\infty)$ we define $S_t^D := \sum_{i=1}^D \sum_{k\in\mathbb{N}_0} \phi(Z_t^{D,k}(i))$. The fact that there exists a constant $L_\phi \in [0,\infty)$ such that for all $x \in [0,1]$ it holds that $|\phi(x)| \le L_\phi x$ and Markov's inequality yield for all $D, K \in \mathbb{N}$ and all $t \in [0,\infty)$ that

$$\mathbb{P}(|S_t^D| \ge K) \le \frac{1}{K}\mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} |\phi(Z_t^{D,k}(i))|\right] \le \frac{L_\phi}{K}\mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} Z_t^{D,k}(i)\right].$$

This, Lemma 2.3, and Setting 2.16 imply for all $t \in [0,\infty)$ that

$$\lim_{K\to\infty}\sup_{D\in\mathbb{N}} \mathbb{P}(|S_t^D| \ge K) \le \lim_{K\to\infty}\sup_{D\in\mathbb{N}} \frac{L_\phi}{K}\mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} Z_t^{D,k}(i)\right] = 0. \tag{2.40}$$

For every $T \in (0,\infty)$ and every $D \in \mathbb{N}$ let $\mathcal{S}_T^D$ be the set of stopping times with respect to the natural filtration of S^D that are bounded by T. For all $D \in \mathbb{N}$ and all $x = (x_{i,k})_{(i,k)\in[D]\times\mathbb{N}_0} \in [0,1]^{[D]\times\mathbb{N}_0}$ we define $\psi^D(x) := \sum_{i=1}^D \sum_{k\in\mathbb{N}_0} \phi(x_{i,k})$ and

$$\begin{aligned}(G^D\psi^D)(x) := \sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0}\Bigg(&\frac{\mathbf{1}_{k>0}}{D}\sum_{j=1}^{D} x_{j,|k-1|} f(x_{j,|k-1|}, x_{i,k}) \\ &+ \tilde{h}_D(x_{i,k}) + \mathbf{1}_{k=0} h_D(0)\Bigg)\phi'(x_{i,k}) \\ &+ \frac{1}{2}\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \sigma^2(x_{i,k})\phi''(x_{i,k}).\end{aligned}$$

Setting 1.1 and $\phi \in C^2([0,1],\mathbb{R})$ imply that there exists a constant $C_\psi \in [1,\infty)$ such that it holds for all $D \in \mathbb{N}$, all $x = (x_{i,k})_{(i,k)\in[D]\times\mathbb{N}_0} \in [0,1]^{[D]\times\mathbb{N}_0}$, and all $y \in [0,1]$ that

$|(G^D\psi^D)(x)| \le C_\psi(2\mu + \sum_{i=1}^D \sum_{k\in\mathbb{N}_0} x_{i,k})$ and that $\sigma^2(y)\phi'^2(y) \le C_\psi^2 y$, Jensen's inequality ensures for all $x_1, x_2 \in \mathbb{R}$ that $(x_1+x_2)^2 \le 2(x_1^2+x_2^2)$, Hölder's inequality yields for every $\delta \in [0,\infty)$ and every integrable function $\alpha : [0,\delta] \to \mathbb{R}$ that $(\int_0^\delta \alpha(u)\,du)^2 \le \delta \int_0^\delta (\alpha(u))^2\,du$, and it holds for all $x \in \mathbb{R}$ that $2x \le 1 + x^2$. Itô's formula, the Itô isometry, and the preceding estimates show for all $T \in (0,\infty)$, all $\bar\delta \in [0,1]$, all $D \in \mathbb{N}$, all $\tau \in \mathcal{S}_T^D$, and all $\delta \in [0,\bar\delta]$ that

$$\begin{aligned}
&\mathbb{E}\big[(S^D_{\tau+\delta} - S^D_\tau)^2\big] \\
&= \mathbb{E}\Bigg[\Bigg(\int_\tau^{\tau+\delta} (G^D\psi^D)(Z_u^{D,\cdot}(\cdot))\,du + \sum_{i=1}^D \sum_{k\in\mathbb{N}_0} \int_\tau^{\tau+\delta} (\sqrt{\sigma^2}\phi')(Z_u^{D,k}(i))\,dW_u^k(i)\Bigg)^2\Bigg] \\
&\le 2C_\psi^2 \mathbb{E}\Bigg[\Bigg(\int_0^\delta 2\mu + \sum_{i=1}^D \sum_{k\in\mathbb{N}_0} Z^{D,k}_{\tau+u}(i)\,du\Bigg)^2\Bigg] \\
&\quad + 2\mathbb{E}\Bigg[\sum_{i=1}^D \sum_{k\in\mathbb{N}_0} \int_0^\delta (\sqrt{\sigma^2}\phi')^2(Z^{D,k}_{\tau+u}(i))\,du\Bigg] \\
&\le 2C_\psi^2 \delta \mathbb{E}\Bigg[\int_0^\delta \Bigg(2\mu + \sum_{i=1}^D \sum_{k\in\mathbb{N}_0} Z^{D,k}_{\tau+u}(i)\Bigg)^2 du\Bigg] + 2C_\psi^2 \mathbb{E}\Bigg[\int_0^\delta \sum_{i=1}^D \sum_{k\in\mathbb{N}_0} Z^{D,k}_{\tau+u}(i)\,du\Bigg] \\
&\le 3C_\psi^2 \bar\delta \mathbb{E}\Bigg[\sup_{t\in[0,T+1]} \Bigg(2\mu + \sum_{i=1}^D \sum_{k\in\mathbb{N}_0} Z_t^{D,k}(i)\Bigg)^2\Bigg] + C_\psi^2 \bar\delta.
\end{aligned}$$

This, Lemma 2.5, and Setting 2.16 imply for all $T \in (0,\infty)$ that

$$\lim_{\bar\delta\to 0} \sup_{D\in\mathbb{N}} \sup_{\tau\in\mathcal{S}_T^D} \sup_{\delta\in[0,\bar\delta]} \mathbb{E}\big[(S^D_{\tau+\delta} - S^D_\tau)^2\big] = 0. \tag{2.41}$$

By Aldous' tightness criterion (see, e.g., Theorem 3.8.6 in [7]), (2.40) and (2.41) ensure that

$$\Bigg\{\Bigg(\sum_{i=1}^D \sum_{k\in\mathbb{N}_0} Z_t^{D,k}(i)\varphi(Z_t^{D,k}(i))\Bigg)_{t\in[0,\infty)} : D \in \mathbb{N}\Bigg\} \tag{2.42}$$

is relatively compact. Since $\varphi \in C^2([0,1],\mathbb{R})$ was arbitrary, it follows from (2.42), from Theorem 2.1 in [25], and from Prokhorov's theorem (see, e.g., Theorem 3.2.2 in [7]) that

$$\Bigg\{\Bigg(\sum_{i=1}^D \sum_{k\in\mathbb{N}_0} Z_t^{D,k}(i)\delta_{Z_t^{D,k}(i)}\Bigg)_{t\in[0,\infty)} : D \in \mathbb{N}\Bigg\} \tag{2.43}$$

is relatively compact.

In the following, we identify the limit points of (2.43) by showing convergence of finite-dimensional distributions. For that, fix $n \in \mathbb{N}$, fix $\varphi_1, \dots, \varphi_n \in \mathrm{Lip}([0,1],[0,\infty))$, and fix $t_1, \dots, t_n \in [0,\infty)$ with $t_1 \le \cdots \le t_n$. For every $j \in [n]$ we define the function $\phi_j : [0,1] \to [0,\infty)$ by $[0,1] \ni x \mapsto \phi_j(x) := x\varphi_j(x) \in [0,\infty)$. Next we show that it holds for all $m \in \mathbb{N}_0$ that

$$\begin{aligned}
&\lim_{D\to\infty} \mathbb{E}\Bigg[\exp\Bigg(-\sum_{j=1}^n \sum_{i=1}^D \sum_{k=0}^m \phi_j(Z_{t_j}^{D,k}(i))\Bigg)\Bigg] \\
&\qquad = \mathbb{E}\Bigg[\exp\Bigg(-\sum_{j=1}^n \sum_{k=0}^m \int \phi_j(\eta_{t_j-s})\mathcal{T}^{(k)}(ds\otimes d\eta)\Bigg)\Bigg].
\end{aligned} \tag{2.44}$$

We prove (2.44) by induction on $m \in \mathbb{N}_0$. For all $\eta \in C([0,\infty),[0,1])$ we define $F(\eta) := \sum_{j=1}^n \phi_j(\eta_{t_j}) \in \mathbb{R}$. For every $D \in \mathbb{N}$ let $(Y_{t,0}^{D,Dh_D(0)}(i))_{t\in[0,\infty)}$, $i \in [D]$, be solutions of (2.17) driven by independent Brownian motions such that for all $i \in [D]$ it holds a.s. that $Y_{0,0}^{D,Dh_D(0)}(i) = X_0(i)$ and let $(\bar{Y}_{t,0}^{D,Dh_D(0)}(i))_{t\in[0,\infty)}$, $i \in [D]$, be independent solutions of (2.17) satisfying for all $i \in [D]$ a.s. that $\bar{Y}_{0,0}^{D,Dh_D(0)}(i) = 0$. Then $Y_{\cdot,0}^{D,Dh_D(0)}$ is equal in distribution to $Z^{D,0}$. Note that it holds for all $x, y, z \in [0,\infty)$ that $|e^{-x} - e^{-(y+z)}| \le |e^{-x}(1 - e^{-y})| + |e^{-y}(e^{-x} - e^{-z})| \le 1 - e^{-y} + |x - z|$. Moreover, there exists a constant $L_F \in [0,\infty)$ such that it holds for all $\eta, \bar{\eta} \in C([0,\infty),[0,1])$ that $|F(\eta) - F(\bar{\eta})| \le L_F \sum_{j=1}^n |\eta_{t_j} - \bar{\eta}_{t_j}|$. These facts and Lemma 2.8 imply that

$$
\begin{aligned}
&\lim_{K\to\infty}\lim_{D\to\infty}\left|\mathbb{E}\left[\exp\left(-\sum_{i=K+1}^{D} F\bigl(Y_{\cdot,0}^{D,Dh_D(0)}(i)\bigr)\right)\right]\right. \\
&\qquad \left. - \mathbb{E}\left[\exp\left(-\sum_{i=1}^{D} F\bigl(\bar{Y}_{\cdot,0}^{D,Dh_D(0)}(i)\bigr)\right)\right]\right| \\
&\quad \le \lim_{K\to\infty}\lim_{D\to\infty}\mathbb{E}\left[1-\exp\left(-\sum_{i=1}^{K} F\bigl(\bar{Y}_{\cdot,0}^{D,Dh_D(0)}(i)\bigr)\right)\right] \\
&\qquad + \lim_{K\to\infty} L_F n e^{L_h t_n} \sum_{i=K+1}^{\infty} \mathbb{E}[X_0(i)].
\end{aligned}
\tag{2.45}
$$

The second summand on the right-hand side of (2.45) is zero due to Setting 2.16. For every $i \in \mathbb{N}$ the process $\bar{Y}_{\cdot,0}^{D,Dh_D(0)}(i)$ converges weakly to zero as $D \to \infty$, so the first summand on the right-hand side of (2.45) is also zero. On the other hand, for every $i \in \mathbb{N}$ the process $Y_{\cdot,0}^{D,Dh_D(0)}(i)$ converges weakly to $Y(i)$ as $D \to \infty$ (see, e.g., Theorem 4.8.10 in [7]). These observations and Lemma 2.15 with $s = 0$, $(g_D)_{D\in\mathbb{N}} = (Dh_D(0))_{D\in\mathbb{N}}$, and $g = \mu$ imply that

$$
\begin{aligned}
&\lim_{D\to\infty}\mathbb{E}\left[\exp\left(-\sum_{i=1}^{D} F\bigl(Z^{D,0}(i)\bigr)\right)\right] \\
&\quad = \lim_{K\to\infty}\lim_{D\to\infty}\mathbb{E}\left[\exp\left(-\sum_{i=1}^{K} F\bigl(Y_{\cdot,0}^{D,Dh_D(0)}(i)\bigr)\right)\right] \\
&\qquad \times \mathbb{E}\left[\exp\left(-\sum_{i=K+1}^{D} F\bigl(Y_{\cdot,0}^{D,Dh_D(0)}(i)\bigr)\right)\right] \\
&\quad = \lim_{K\to\infty}\mathbb{E}\left[\exp\left(-\sum_{i=1}^{K} F\bigl(Y(i)\bigr)\right)\right]\lim_{D\to\infty}\mathbb{E}\left[\exp\left(-\sum_{i=1}^{D} F\bigl(\bar{Y}_{\cdot,0}^{D,Dh_D(0)}(i)\bigr)\right)\right] \\
&\quad = \mathbb{E}\left[\exp\left(-\sum_{i=1}^{\infty} F\bigl(Y(i)\bigr)\right)\right]\mathbb{E}\left[\exp\left(-\int F(\eta_{\cdot-s})\Pi^{\varnothing}(ds\otimes d\eta)\right)\right] \\
&\quad = \mathbb{E}\left[\exp\left(-\int F(\eta_{\cdot-s})\mathcal{T}^{(0)}(ds\otimes d\eta)\right)\right].
\end{aligned}
$$

This establishes (2.44) in the base case $m = 0$. For the induction step $\mathbb{N}_0 \ni m \to m+1$ the induction hypothesis and relative compactness for all $\tilde{m} \in [m]_0$ of

$$
\left\{\left(\sum_{i=1}^{D}\sum_{k=0}^{\tilde{m}} Z_t^{D,k}(i)\delta_{Z_t^{D,k}(i)}\right)_{t\in[0,\infty)} : D \in \mathbb{N}\right\}
$$

imply for all $\tilde{m} \in [m]_0$ and all $\varphi \in C([0,1],\mathbb{R})$ that

$$\left(\sum_{i=1}^{D}\sum_{k=0}^{\tilde{m}} Z_t^{D,k}(i)\varphi\big(Z_t^{D,k}(i)\big)\right)_{t\in[0,\infty)} \xRightarrow{D\to\infty} \left(\sum_{k=0}^{\tilde{m}}\int \eta_{t-s}\varphi(\eta_{t-s})\mathcal{T}^{(k)}(ds\otimes d\eta)\right)_{t\in[0,\infty)}. \tag{2.46}$$

By the Skorokhod representation theorem (see, e.g., Theorem 3.1.8 in [7]), we may assume almost sure convergence in (2.46). Consequently, we may assume for all $\varphi \in C([0,1],\mathbb{R})$ that

$$\left(\sum_{i=1}^{D} Z_t^{D,m}(i)\varphi\big(Z_t^{D,m}(i)\big)\right)_{t\in[0,\infty)} \xrightarrow[\text{a.s.}]{D\to\infty} \left(\int \eta_{t-s}\varphi(\eta_{t-s})\mathcal{T}^{(m)}(ds\otimes d\eta)\right)_{t\in[0,\infty)}. \tag{2.47}$$

For every $D \in \mathbb{N}$ let $g_D, g : [0,\infty)\times[0,1]\times\Omega \to \mathbb{R}$ be functions which satisfy for all $(t,x) \in [0,\infty)\times[0,1]$ that $g_D(t,x) = \sum_{j=1}^{D} Z_t^{D,m}(j) f(Z_t^{D,m}(j),x)$ and $g(t,x) = \int \eta_{t-s} f(\eta_{t-s},x)\mathcal{T}^{(m)}(ds\otimes d\eta)$. The sequence of functions $(g_D)_{D\in\mathbb{N}}$ satisfies Setting 2.7 almost surely. Moreover, the function g satisfies almost surely for all $t\in[0,\infty)$ that $g(t,0)\geq 0$. Furthermore, the fact that it holds for all $t\in[0,\infty)$ that $\int \sup_{x\in[0,1]}|\eta_{t-s}f(\eta_{t-s},x)| \times \mathcal{T}^{(m)}(ds\otimes d\eta) \leq L_f \int \eta_{t-s}\mathcal{T}(ds\otimes d\eta)$, the fact that it holds for all $t\in[0,\infty)$ that $\mathbb{E}[\int \eta_{t-s}\mathcal{T}(ds\otimes d\eta)] < \infty$, and the dominated convergence theorem yield that it holds almost surely for all $t\in[0,\infty)$ that $[0,1]\ni x\mapsto g(t,x)\in\mathbb{R}$ is continuous. Equation (2.47) and the assumptions on f imply almost surely for all $t\in[0,\infty)$ that $(g_D(t,\cdot))_{D\in\mathbb{N}}$ is an equicontinuous sequence and this together with (2.47) yields almost surely for all $t\in[0,\infty)$ that

$$\lim_{D\to\infty}\sup_{x\in[0,1]}\big|g_D(t,x)-g(t,x)\big| = 0. \tag{2.48}$$

It follows almost surely for all $t\in[0,\infty)$ from (2.47) that $[0,t]\ni u\mapsto \int \eta_{u-s}\mathcal{T}^{(m)}(ds\otimes d\eta)\in\mathbb{R}$ is càdlàg and therefore square-integrable and thus that

$$\sup_{D\in\mathbb{N}}\int_0^t\left(\sum_{i=1}^{D} Z_u^{D,m}(i)\right)^2 du < \infty.$$

This, the fact that for all $D\in\mathbb{N}$ and all $u\in[0,\infty)$ it holds that $\sup_{x\in[0,1]}|g_D(u,x) - g(u,x)| \leq L_f(\sum_{i=1}^{D} Z_u^{D,m}(i) + \int \eta_{u-s}\mathcal{T}^{(m)}(ds\otimes d\eta))$, and Theorem 6.18 and Corollary 6.21 in [17] imply almost surely for all $t\in[0,\infty)$ that the family

$$\left\{[0,t]\ni u\mapsto \sup_{x\in[0,1]}\big|g_D(u,x)-g(u,x)\big|\in\mathbb{R} : D\in\mathbb{N}\right\}$$

is uniformly integrable. This, Theorem 6.25 in [17], and (2.48) show almost surely for all $t\in[0,\infty)$ that

$$\lim_{D\to\infty}\int_0^t \sup_{x\in[0,1]}\big|g_D(u,x)-g(u,x)\big|\,du = 0.$$

Conditionally on $(Z^{M,m})_{M\in\mathbb{N}}$, for every $D\in\mathbb{N}$ a version of $Z^{D,m+1}$ is given by $Y^{D,g_D}_{\cdot,0}$ satisfying for all $i\in[D]$ and all $t\in[0,\infty)$ that a.s.

$$\begin{aligned} Y^{D,g_D}_{t,0}(i) = &\int_0^t \frac{1}{D} g_D\big(u, Y^{D,g_D}_{u,0}(i)\big) + \tilde{h}_D\big(Y^{D,g_D}_{u,0}(i)\big)\,du \\ &+ \int_0^t \sqrt{\sigma^2\big(Y^{D,g_D}_{u,0}(i)\big)}\,dW^{m+1}_u(i). \end{aligned}$$

Therefore, Lemma 2.15 yields that a.s.

$$\begin{aligned} &\lim_{D\to\infty} \mathbb{E}\bigg[\exp\bigg(-\sum_{i=1}^D F\big(Z^{D,m+1}(i)\big)\bigg)\,\bigg|\,(Z^{M,m})_{M\in\mathbb{N}}, \mathcal{T}^{(m)}\bigg] \\ &= \lim_{D\to\infty} \mathbb{E}\bigg[\exp\bigg(-\sum_{i=1}^D F\big(Y^{D,g_D}_{\cdot,0}(i)\big)\bigg)\,\bigg|\,(Z^{M,m})_{M\in\mathbb{N}}, \mathcal{T}^{(m)}\bigg] \qquad (2.49) \\ &= \mathbb{E}\bigg[\exp\bigg(-\int F(\eta_{\cdot-s})\Pi^{(m+1)}(ds\otimes d\eta)\bigg)\,\bigg|\,\mathcal{T}^{(m)}\bigg], \end{aligned}$$

where $\Pi^{(m+1)}$ conditioned on $\mathcal{T}^{(m)}$ is a Poisson point process on $[0,\infty)\times U$ with the property that for all bounded measurable $\Phi:[0,\infty)\times U\to[0,\infty)$ it holds almost surely that

$$\begin{aligned} &\int \Phi(s,\eta)\mathbb{E}\big[\Pi^{(m+1)}(ds\otimes d\eta)\mid \mathcal{T}^{(m)}\big] \\ &= \int \Phi(s,\eta)\bigg(\int \tilde{a}(\chi_{s-r})\mathcal{T}^{(m)}(dr\otimes d\chi)\bigg)\,ds\otimes Q(d\eta) \\ &= \int\int \Phi(s,\eta)\tilde{a}(\chi_{s-r})\,ds\otimes Q(d\eta)\mathcal{T}^{(m)}(dr\otimes d\chi) \\ &= \int\int \Phi(s,\eta)\mathbb{E}\big[\Pi^{(m,r,\chi)}(ds\otimes d\eta)\big]\mathcal{T}^{(m)}(dr\otimes d\chi). \end{aligned}$$

This shows that $\Pi^{(m+1)}$ conditioned on $\mathcal{T}^{(m)}$ is equal in distribution to $\int \Pi^{(m,r,\chi)}\mathcal{T}^{(m)}(dr\otimes d\chi) = \mathcal{T}^{(m+1)}$. This, (2.49), and the induction hypothesis show that

$$\begin{aligned} &\lim_{D\to\infty} \mathbb{E}\bigg[\exp\bigg(-\sum_{k=0}^{m+1}\sum_{i=1}^D F\big(Z^{D,k}(i)\big)\bigg)\bigg] \\ &= \mathbb{E}\bigg[\lim_{D\to\infty} \exp\bigg(-\sum_{k=0}^{m}\sum_{i=1}^D F\big(Z^{D,k}(i)\big)\bigg) \\ &\quad \times \mathbb{E}\bigg[\exp\bigg(-\sum_{i=1}^D F\big(Z^{D,m+1}(i)\big)\bigg)\,\bigg|\,(Z^{M,m})_{M\in\mathbb{N}}, \mathcal{T}^{(m)}\bigg]\bigg] \\ &= \mathbb{E}\bigg[\exp\bigg(-\sum_{k=0}^{m}\int F(\eta_{\cdot-s})\mathcal{T}^{(k)}(ds\otimes d\eta)\bigg) \\ &\quad \times \mathbb{E}\bigg[\exp\bigg(-\int F(\eta_{\cdot-s})\mathcal{T}^{(m+1)}(ds\otimes d\eta)\bigg)\,\bigg|\,\mathcal{T}^{(m)}\bigg]\bigg] \\ &= \mathbb{E}\bigg[\exp\bigg(-\sum_{k=0}^{m+1}\int F(\eta_{\cdot-s})\mathcal{T}^{(k)}(ds\otimes d\eta)\bigg)\bigg]. \end{aligned}$$

This finishes the induction step $\mathbb{N}_0 \ni m \to m+1$ and hence proves (2.44). Lemma 2.4, the fact that

$$\lim_{m\to\infty} \mathbb{E}\left[\exp\left(-\sum_{k=0}^{m}\sum_{j=1}^{n}\int \phi_j(\eta_{t_j-s})\mathcal{T}^{(k)}(ds\otimes d\eta)\right)\right] = \mathbb{E}\left[\exp\left(-\sum_{k=0}^{\infty}\sum_{j=1}^{n}\int \phi_j(\eta_{t_j-s})\mathcal{T}^{(k)}(ds\otimes d\eta)\right)\right],$$

and (2.44) imply the convergence of finite-dimensional distributions of (2.43). Therefore, this finishes the proof of Lemma 2.17. □

3. Convergence to a forest of excursions. To prove Theorem 1.4 in Section 3.3 below, we first show that the migration level processes and the loop-free processes have the same limit as $D \to \infty$; see Lemma 3.8 below. Our method of proof is the integration by parts formula for semigroups; see (3.22), (3.23), and (3.26) below. For this, we first derive moment estimates in Section 3.1 and uniform bounds on the derivatives of the semigroups of the loop-free processes in Lemma 3.3.

3.1. *Results for the migration level processes.* The following lemma implies that individuals on the same deme have essentially the same migration level in the limit $D \to \infty$ and is analogous to Lemma 4.24 in [11].

LEMMA 3.1 (Essentially one level per deme). *Assume that* Setting 2.2 *holds*, *that*

$$\sup_{D\in\mathbb{N}} \mathbb{E}\left[\left(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)\right)^2\right] < \infty, \tag{3.1}$$

and that

$$\lim_{D\to\infty} \mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_0^{D,m}(i)\right] = 0. \tag{3.2}$$

Then it holds for all $T \in (0,\infty)$ *that*

$$\lim_{D\to\infty} \sup_{t\in[0,T]} \mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_t^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_t^{D,m}(i)\right] = 0. \tag{3.3}$$

PROOF. Fix $T \in (0,\infty)$ for the rest of the proof. For every $D, M \in \mathbb{N}$ we consider the stopping time τ_M^D defined in (2.14). Since it holds for all $D \in \mathbb{N}$, all $i \in [D]$, and all $t \in [0,T]$ that $\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i) \in [0,1]$, we obtain for all $D, M \in \mathbb{N}$ that

$$\begin{aligned} &\sup_{t\in[0,T]} \mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_t^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_t^{D,m}(i)\right] \\ &\quad\le \sup_{t\in[0,T]} \mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_t^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_t^{D,m}(i)\mathbf{1}_{\{\tau_M^D > T\}}\right] \\ &\quad\quad + \mathbb{E}\left[\sup_{t\in[0,T]} \sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_t^{D,k}(i)\mathbf{1}_{\{\tau_M^D \le T\}}\right]. \end{aligned} \tag{3.4}$$

Lemma 2.6 ensures that the second summand on the right-hand side of (3.4) converges to zero uniformly in $D \in \mathbb{N}$ as $M \to \infty$. To prove (3.3) it therefore suffices to show that the first summand on the right-hand side of (3.4) converges to zero as $D \to \infty$ for all $M \in \mathbb{N}$. We fix $M \in \mathbb{N}$ for the rest of the proof. For every $D, K \in \mathbb{N}$ and every $t \in [0, \infty)$ let $M_t^{D,K}$ be a real-valued random variable satisfying that a.s.

$$M_t^{D,K} = \sum_{i=1}^{D} \sum_{\substack{k,m=0\\ m\neq k}}^{K} \int_0^t X_s^{D,k}(i) \sqrt{\frac{X_s^{D,m}(i)}{\sum_{l\in\mathbb{N}_0} X_s^{D,l}(i)} \sigma^2\Big(\sum_{l\in\mathbb{N}_0} X_s^{D,l}(i)\Big)}\, dW_s^m(i). \tag{3.5}$$

Itô's formula, (2.1), (3.5), and Setting 1.1 yield for all $D, K \in \mathbb{N}$ and all $t \in [0, \infty)$ that a.s.

$$\begin{aligned}
&\sum_{i=1}^{D}\sum_{k=0}^{K} X_t^{D,k}(i) \sum_{\substack{m=0\\ m\neq k}}^{K} X_t^{D,m}(i) \\
&\quad \le \sum_{i=1}^{D}\sum_{k=0}^{K} X_0^{D,k}(i) \sum_{\substack{m=0\\ m\neq k}}^{K} X_0^{D,m}(i) + 2\int_0^t \sum_{i=1}^{D}\sum_{k=0}^{K} X_s^{D,k}(i) h_D(0)\, ds \\
&\qquad + 2\int_0^t \sum_{i=1}^{D} \sum_{\substack{k,m=0\\ m\neq k}}^{K} X_s^{D,k}(i)\bigg(\frac{1}{D}\sum_{j=1}^{D} L_f X_s^{D,m-1}(j) + L_h X_s^{D,m}(i)\bigg)\, ds \\
&\qquad + 2M_t^{D,K}.
\end{aligned} \tag{3.6}$$

For every $D, K \in \mathbb{N}$ the fact that

$$\sum_{i=1}^{D} \sum_{\substack{k,m=0\\ m\neq k}}^{K} \int_0^T (X_s^{D,k}(i))^2 \frac{X_s^{D,m}(i)}{\sum_{l\in\mathbb{N}_0} X_s^{D,l}(i)} \sigma^2\Big(\sum_{l\in\mathbb{N}_0} X_s^{D,l}(i)\Big)\, ds \le DTL_\sigma$$

implies that $(M_t^{D,K})_{t\in[0,T]}$ is a martingale. Using this, using for all $D \in \mathbb{N}$ and all $s \in [0, T]$ that $\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_{s\wedge\tau_M^D}^{D,k}(i) \le M + \sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)$, and applying the optional sampling theorem (see, e.g., Theorem 2.2.13 in [7]) and Tonelli's theorem, we obtain from (3.6) for all $D, K \in \mathbb{N}$ and all $t \in [0, T]$ that

$$\begin{aligned}
&\mathbb{E}\bigg[\sum_{i=1}^{D}\sum_{k=0}^{K} X_{t\wedge\tau_M^D}^{D,k}(i) \sum_{\substack{m=0\\ m\neq k}}^{K} X_{t\wedge\tau_M^D}^{D,m}(i)\bigg] \\
&\quad \le \mathbb{E}\bigg[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_0^{D,m}(i)\bigg] \\
&\qquad + 2Th_D(0)\bigg(M + \mathbb{E}\bigg[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)\bigg]\bigg) \\
&\qquad + \frac{1}{D} 2TL_f \mathbb{E}\bigg[\bigg(M + \sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)\bigg)^2\bigg] \\
&\qquad + 2L_h \int_0^t \mathbb{E}\bigg[\sum_{i=1}^{D}\sum_{k=0}^{K} X_{s\wedge\tau_M^D}^{D,k}(i) \sum_{\substack{m=0\\ m\neq k}}^{K} X_{s\wedge\tau_M^D}^{D,m}(i)\bigg]\, ds.
\end{aligned}$$

This, the fact that we have for all $D \in \mathbb{N}$ that $h_D(0) \le 2\mu/D$, Gronwall's inequality, and the monotone convergence theorem ensure for all $D \in \mathbb{N}$ that

$$
\begin{aligned}
&\sup_{t\in[0,T]} \mathbb{E}\Bigg[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_t^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_t^{D,m}(i)\mathbf{1}_{\{\tau_M^D>T\}}\Bigg] \\
&\quad\le \sup_{t\in[0,T]} \mathbb{E}\Bigg[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_{t\wedge\tau_M^D}^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_{t\wedge\tau_M^D}^{D,m}(i)\Bigg] \\
&\quad\le e^{2L_hT}\Bigg(\mathbb{E}\Bigg[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_0^{D,m}(i)\Bigg] \\
&\qquad + \frac{1}{D}2T(2\mu+L_f)\sup_{N\in\mathbb{N}} \mathbb{E}\Bigg[\Bigg(M+\sum_{i=1}^{N}\sum_{k\in\mathbb{N}_0} X_0^{N,k}(i)\Bigg)^2\Bigg]\Bigg).
\end{aligned}
$$

Letting $D \to \infty$ and applying (3.2) and (3.1) finishes the proof of Lemma 3.1. □

The following lemma implies that the total mass is not evenly distributed over all demes and is analogous to Lemma 4.23 in [11].

LEMMA 3.2 (Concentration of mass). *Assume that* Setting 1.2 *and* Setting 2.2 *hold*, *that*

$$
\sum_{k\in\mathbb{N}_0}\sup_{D\in\mathbb{N}} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_0^{D,k}(i)\Bigg] < \infty, \tag{3.7}
$$

that

$$
\sup_{D\in\mathbb{N}} \mathbb{E}\Bigg[\Bigg(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)\Bigg)^2\Bigg] < \infty, \tag{3.8}
$$

and that

$$
\lim_{\delta\to 0}\sup_{D\in\mathbb{N}} \mathbb{E}\Bigg[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \big(X_0^{D,k}(i)\wedge\delta\big)\Bigg] = 0. \tag{3.9}
$$

Then it holds for all $T \in (0,\infty)$ that

$$
\lim_{\delta\to 0}\sum_{k\in\mathbb{N}_0}\sup_{D\in\mathbb{N}}\sup_{t\in[0,T]} \mathbb{E}\Bigg[\sum_{i=1}^{D}\big(X_t^{D,k}(i)\wedge\delta\big)\Bigg] = 0. \tag{3.10}
$$

PROOF. Fix $T \in (0,\infty)$ for the rest of the proof. For every $D, M \in \mathbb{N}$ we consider the stopping time τ_M^D defined in (2.14). Then it holds for all $\delta \in (0,\infty)$, all $K \in \mathbb{N}_0$, and all $M \in \mathbb{N}$ that

$$
\begin{aligned}
&\sum_{k\in\mathbb{N}_0}\sup_{D\in\mathbb{N}}\sup_{t\in[0,T]} \mathbb{E}\Bigg[\sum_{i=1}^{D}\big(X_t^{D,k}(i)\wedge\delta\big)\Bigg] \\
&\quad\le \sum_{k=0}^{K}\sup_{D\in\mathbb{N}}\sup_{t\in[0,T]} \mathbb{E}\Bigg[\sum_{i=1}^{D}\big(X_t^{D,k}(i)\wedge\delta\big)\mathbf{1}_{\{\tau_M^D>T\}}\Bigg] + \sum_{k=K}^{\infty}\sup_{D\in\mathbb{N}}\sup_{t\in[0,T]} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_t^{D,k}(i)\Bigg] \\
&\qquad + \sum_{k=0}^{K}\sup_{D\in\mathbb{N}} \mathbb{E}\Bigg[\sup_{t\in[0,T]}\sum_{i=1}^{D}\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)\mathbf{1}_{\{\tau_M^D\le T\}}\Bigg].
\end{aligned} \tag{3.11}
$$

Lemma 2.4 and (3.7) imply that the second summand on the right-hand side of (3.11) converges to zero as $K \to \infty$, while Lemma 2.6 and (3.8) ensure for all $K \in \mathbb{N}_0$ that the third summand on the right-hand side of (3.11) converges to zero as $M \to \infty$. To prove (3.10) it therefore suffices to show for all $K \in \mathbb{N}_0$ and all $M \in \mathbb{N}$ that the first summand on the right-hand side of (3.11) converges to zero as $\delta \to 0$. We fix $k \in \mathbb{N}_0$ and $M \in \mathbb{N}$ for the rest of the proof. Setting 1.1 implies the existence of $D_0 \in \mathbb{N}$ such that for all $D \in \mathbb{N} \cap [D_0, \infty)$ we have $L_f M/D + h_D(1) \le 0$. For every $D \in \mathbb{N} \cap [D_0, \infty)$ let $\tilde{X}^D$ be as in Lemma 2.1. Moreover, for every $D \in \mathbb{N} \cap [D_0, \infty)$ and every $i \in [D]$ let $(\tilde{Y}_{t,0}^{D,L_f M+Dh_D(0)}(i))_{t\in[0,\infty)}$ and $(Y_{t,0}^{D,L_f M+Dh_D(0)}(i))_{t\in[0,\infty)}$ be two solutions of (2.17) with respect to the same Brownian motion satisfying a.s. that $\tilde{Y}_{0,0}^{D,L_f M+Dh_D(0)}(i) = \tilde{X}_0^D(i)$ and $Y_{0,0}^{D,L_f M+Dh_D(0)}(i) = 0$. Lemma 2.1 above, Setting 1.1 above, and Lemma 3.3 in [14] show for all $D \in \mathbb{N} \cap [D_0, \infty)$, all $i \in [D]$, and all $t \in [0, T]$ that $\tilde{X}_t^D(i)$ is stochastically bounded from above by $\tilde{Y}_{t,0}^{D,L_f M+Dh_D(0)}(i)$ on the event $\{\tau_M^D > T\}$. This, (2.3), the fact that for all $a, b, \delta \in [0, \infty)$ it holds that $|a \wedge \delta - b \wedge \delta| = |a \wedge \delta - b \wedge \delta| \wedge \delta \le |a - b| \wedge \delta$, and Jensen's inequality for the conditional expectation applied for all $\delta \in (0, \infty)$ with the concave function $[0, 1] \ni x \mapsto x \wedge \delta \in \mathbb{R}$ yield for all $\delta \in (0, \infty)$, all $t \in [0, T]$, and all $D \in \mathbb{N} \cap [D_0, \infty)$ that

$$
\begin{aligned}
&\mathbb{E}\Bigg[\sum_{i=1}^{D}(X_t^{D,k}(i) \wedge \delta)\mathbf{1}_{\{\tau_M^D>T\}}\Bigg] \\
&\quad\le \sum_{i=1}^{D}\mathbb{E}\big[(\tilde{X}_t^D(i) \wedge \delta)\mathbf{1}_{\{\tau_M^D>T\}}\big] \le \sum_{i=1}^{D}\mathbb{E}\big[\tilde{Y}_{t,0}^{D,L_f M+Dh_D(0)}(i) \wedge \delta\big] \\
&\quad\le \sum_{i=1}^{D}\mathbb{E}\big[Y_{t,0}^{D,L_f M+Dh_D(0)}(i) \wedge \delta\big] \\
&\qquad+ \sum_{i=1}^{D}\mathbb{E}\big[\mathbb{E}\big[\big|\tilde{Y}_{t,0}^{D,L_f M+Dh_D(0)}(i) - Y_{t,0}^{D,L_f M+Dh_D(0)}(i)\big| \mid \tilde{X}_0^D(i)\big] \wedge \delta\big].
\end{aligned}
$$

This, Lemma 2.8, and the fact that it holds for all $t \in [0, T]$ that $e^{L_h t} \ge 1$ ensure for all $\delta \in (0, \infty)$, all $t \in [0, T]$, and all $D \in \mathbb{N} \cap [D_0, \infty)$ that

$$
\begin{aligned}
\mathbb{E}\Bigg[\sum_{i=1}^{D}(X_t^{D,k}(i) \wedge \delta)\mathbf{1}_{\{\tau_M^D>T\}}\Bigg] &\le D\mathbb{E}\big[Y_{t,0}^{D,L_f M+Dh_D(0)}(1) \wedge \delta\big] \\
&\qquad+ e^{L_h t}\sum_{i=1}^{D}\mathbb{E}\big[\tilde{X}_0^D(i) \wedge \delta\big].
\end{aligned}
$$

This, Lemma 2.14 with $(g_D)_{D\ge D_0} = (L_f M + Dh_D(0))_{D\ge D_0}$ and $g = L_f M + \mu$, equation (2.3), and subadditivity for all $\delta \in (0, \infty)$ of $[0, 1] \ni x \mapsto x \wedge \delta \in \mathbb{R}$ imply for all $\delta \in (0, \infty)$ that

$$
\begin{aligned}
&\sup_{t\in[0,T]} \varlimsup_{D\to\infty} \mathbb{E}\Bigg[\sum_{i=1}^{D}(X_t^{D,k}(i) \wedge \delta)\mathbf{1}_{\{\tau_M^D>T\}}\Bigg] \\
&\quad\le (L_f M + \mu)\int\int_0^T (\chi_{T-r} \wedge \delta)\,dr\,Q(d\chi) \\
&\qquad+ e^{L_h T}\sup_{D\in\mathbb{N}}\mathbb{E}\Bigg[\sum_{i=1}^{D}\sum_{m\in\mathbb{N}_0}(X_0^{D,m}(i) \wedge \delta)\Bigg].
\end{aligned}
\tag{3.12}
$$

The first summand on the right-hand side of (3.12) converges to zero as $\delta \to 0$ by the dominated convergence theorem and Lemma 2.10. The second summand on the right-hand side of (3.12) converges to zero as $\delta \to 0$ due to (3.9). This completes the proof of Lemma 3.2. $\square$

3.2. *The migration level processes and the loop-free processes have the same limit.* For the rest of this paragraph, we fix $K \in \mathbb{N}_0$ and assume that Setting 2.2 holds. For all $D \in \mathbb{N}$ we denote by $\{S_t^D : t \in [0,\infty)\}$ the strongly continuous contraction semigroup on $C([0,1]^{[D]\times[K]_0}, \mathbb{R})$ associated with $\{(Z_t^{D,k}(i))_{t\in[0,\infty)} : (i,k) \in [D] \times [K]_0\}$; see Remark 3.2 in [26]. Then for all $D \in \mathbb{N}$, all $t \in [0,\infty)$, all $\psi \in C([0,1]^{[D]\times[K]_0}, \mathbb{R})$, and all $x \in [0,1]^{[D]\times[K]_0}$ it holds that

$$(S_t^D\psi)(x) = \mathbb{E}\big[\psi\big((Z_t^{D,k,x}(i))_{(i,k)\in[D]\times[K]_0}\big)\big]. \tag{3.13}$$

For every $D \in \mathbb{N}$ the semigroup $\{S_t^D : t \in [0,\infty)\}$ has as its generator the closure of the operator G^D acting on $C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$, given for all $\psi \in C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$ and all $x = (x_{i,k})_{(i,k)\in[D]\times[K]_0} \in [0,1]^{[D]\times[K]_0}$ by

$$\begin{aligned}(G^D\psi)(x) = {} & \sum_{i=1}^{D}\sum_{k=0}^{K}\Bigg(\frac{\mathbf{1}_{k>0}}{D}\sum_{j=1}^{D} x_{j,|k-1|} f(x_{j,|k-1|}, x_{i,k}) \\ & + \tilde{h}_D(x_{i,k}) + \mathbf{1}_{k=0}h_D(0)\Bigg)\frac{\partial\psi}{\partial x_{i,k}}(x) \\ & + \frac{1}{2}\sum_{i=1}^{D}\sum_{k=0}^{K}\sigma^2(x_{i,k})\frac{\partial^2\psi}{\partial x_{i,k}^2}(x).\end{aligned} \tag{3.14}$$

The following lemma establishes uniform bounds on the derivatives of the semigroups of the loop-free processes.

LEMMA 3.3 (Uniform C^2-bound). *Assume that* Setting 2.2 *holds, let $K \in \mathbb{N}_0$, and for every $D \in \mathbb{N}$ let $\{S_t^D : t \in [0,\infty)\}$ be as in* (3.13). *Then there exists $c \in [0,\infty)$ such that it holds for all $D \in \mathbb{N}$, all $t \in [0,\infty)$, and all $\psi^D \in C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$ that $S_t^D\psi^D \in C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$ and*

$$\|S_t^D\psi^D\|_{C^2} \le e^{ct}\|\psi^D\|_{C^2}. \tag{3.15}$$

PROOF. For every $D \in \mathbb{N}$ and every $(i,k) \in [D] \times [K]_0$ let the function $b_{i,k} : [0,1]^{[D]\times[K]_0} \to \mathbb{R}$ satisfy for all $x = (x_{j,l})_{(j,l)\in[D]\times[K]_0} \in [0,1]^{[D]\times[K]_0}$ that

$$b_{i,k}(x) = \frac{\mathbf{1}_{k>0}}{D}\sum_{j=1}^{D} x_{j,|k-1|} f(x_{j,|k-1|}, x_{i,k}) + \tilde{h}_D(x_{i,k}) + \mathbf{1}_{k=0}h_D(0).$$

Then it holds for all $D \in \mathbb{N}$ and all $\alpha \in \mathbb{N}_0^{[D]\times[K]_0}$ with $|\alpha| = 1$ that

$$\begin{aligned}&\sum_{i=1}^{D}\sum_{k=0}^{K}\|\partial^\alpha b_{i,k}\|_\infty \\ &\quad \le \|f\|_\infty + \left\|\frac{\partial f}{\partial x_1}\right\|_\infty + \left\|\frac{\partial f}{\partial x_2}\right\|_\infty + \left\|\frac{d\tilde{h}_D}{dx}\right\|_\infty \le 3\|f\|_{C^1} + \|\tilde{h}_D\|_{C^1}\end{aligned}$$

and for all $D \in \mathbb{N}$ and all $\alpha \in \mathbb{N}_0^{[D]\times[K]_0}$ with $|\alpha| = 2$ that

$$\begin{aligned}\sum_{i=1}^{D}\sum_{k=0}^{K}\|\partial^\alpha b_{i,k}\|_\infty &\le 2\left(\left\|\frac{\partial f}{\partial x_1}\right\|_\infty + \left\|\frac{\partial f}{\partial x_2}\right\|_\infty\right) \\ &\quad + \left\|\frac{\partial^2 f}{\partial x_1^2}\right\|_\infty + 2\left\|\frac{\partial^2 f}{\partial x_1\,\partial x_2}\right\|_\infty + \left\|\frac{\partial^2 f}{\partial x_2^2}\right\|_\infty + \left\|\frac{d^2\tilde{h}_D}{dx^2}\right\|_\infty \\ &\le 8\|f\|_{C^2} + \|\tilde{h}_D\|_{C^2}.\end{aligned}$$

We define

$$c := 4\Bigl(8\|f\|_{C^2} + \sup_{D\in\mathbb{N}}\|\tilde{h}_D\|_{C^2}\Bigr) + \frac{1}{2}\|\sigma^2\|_{C^2},$$

which is finite due to Setting 1.1. Then Theorem 4.1 in [13] shows for all $D \in \mathbb{N}$, all $t \in [0,\infty)$, and all $\psi^D \in C^2([0,1]^{[D]\times[K]_0},\mathbb{R})$ that $S_t^D\psi^D \in C^2([0,1]^{[D]\times[K]_0},\mathbb{R})$ and that (3.15) holds. The proof of Lemma 3.3 is thus completed. □

The following lemma follows immediately from Theorem 3.16 in [19] and Lemma 3.3 above.

LEMMA 3.4 (Kolmogorov backward equation). *Assume that* Setting 2.2 *holds*, *let $T \in (0,\infty)$, let $D \in \mathbb{N}$, let $K \in \mathbb{N}_0$, let $\psi \in C^2([0,1]^{[D]\times[K]_0},\mathbb{R})$, let $\{S_t^D : t \in [0,\infty)\}$ be as in* (3.13), *let G^D be as in* (3.14), *and define the function $u : [0,T]\times[0,1]^{[D]\times[K]_0} \to \mathbb{R}$ by*

$$[0,T]\times[0,1]^{[D]\times[K]_0} \ni (t,x) \mapsto u(t,x) := \bigl(S_{T-t}^D\psi\bigr)(x) \in \mathbb{R}.$$

Then it holds that $u \in C^{1,2}([0,T]\times[0,1]^{[D]\times[K]_0},\mathbb{R})$ and it holds for all $t \in [0,T]$ and all $x \in [0,1]^{[D]\times[K]_0}$ that $u(T,x) = \psi(x)$ and

$$\frac{\partial u}{\partial t}(t,x) + \bigl(G^D u\bigr)(t,x) = 0.$$

The following lemma shows that finitely many levels of the migration level processes and of the loop-free processes have the same limit as $D \to \infty$ at every fixed time point.

LEMMA 3.5 (Asymptotic equality for one-dimensional distributions). *Assume that* Setting 1.2 *and* Setting 2.2 *hold*, *that*

$$\sum_{k\in\mathbb{N}_0}\sup_{D\in\mathbb{N}}\mathbb{E}\left[\sum_{i=1}^{D}X_0^{D,k}(i)\right] < \infty, \tag{3.16}$$

that

$$\sup_{D\in\mathbb{N}}\mathbb{E}\left[\left(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0}X_0^{D,k}(i)\right)^2\right] < \infty, \tag{3.17}$$

that

$$\lim_{D\to\infty}\mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0}X_0^{D,k}(i)\sum_{m\in\mathbb{N}_0\setminus\{k\}}X_0^{D,m}(i)\right] = 0, \tag{3.18}$$

and that

$$\lim_{\delta\to 0}\sup_{D\in\mathbb{N}}\mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0}\bigl(X_0^{D,k}(i)\wedge\delta\bigr)\right] = 0, \tag{3.19}$$

let $T \in (0, \infty)$ *and* $K \in \mathbb{N}_0$, *for every* $D \in \mathbb{N}$ *let* $\psi^D \in C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$, *and suppose that* $\sup_{D\in\mathbb{N}} \|\psi^D\|_{C^2} < \infty$. *Then it holds that*

$$\begin{aligned}\lim_{D\to\infty} \big|\mathbb{E}\big[\psi^D\big(\big(X_T^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big] \\ - \mathbb{E}\big[\psi^D\big(\big(Z_T^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big]\big| = 0.\end{aligned} \tag{3.20}$$

PROOF. For every $D \in \mathbb{N}$ let $\{S_t^D : t \in [0,\infty)\}$ be as in (3.13), let G^D be as in (3.14), and define the function $u^D : [0,T] \times [0,1]^{[D]\times[K]_0} \to \mathbb{R}$ by

$$[0,T]\times[0,1]^{[D]\times[K]_0} \ni (t,x) \mapsto u^D(t,x) := \big(S_{T-t}^D \psi^D\big)(x) \in \mathbb{R}. \tag{3.21}$$

Equations (3.21) and (3.13) yield for all $D \in \mathbb{N}$ that

$$\mathbb{E}\big[u^D\big(T, \big(X_T^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big] = \mathbb{E}\big[\psi^D\big(\big(X_T^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big] \tag{3.22}$$

and

$$\mathbb{E}\big[u^D\big(0, \big(X_0^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big] = \mathbb{E}\big[\psi^D\big(\big(Z_T^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big]. \tag{3.23}$$

This shows that (3.20) is implied by

$$\begin{aligned}\lim_{D\to\infty} \big|\mathbb{E}\big[u^D\big(T, \big(X_T^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big] \\ - \mathbb{E}\big[u^D\big(0, \big(X_0^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big]\big| = 0.\end{aligned} \tag{3.24}$$

Lemma 3.4 implies for all $D \in \mathbb{N}$ that $u^D \in C^{1,2}([0,T]\times[0,1]^{[D]\times[K]_0}, \mathbb{R})$ and for all $D \in \mathbb{N}$, all $t \in [0,T]$, and all $x \in [0,1]^{[D]\times[K]_0}$ that

$$\frac{\partial u^D}{\partial t}(t,x) + \big(G^D u^D\big)(t,x) = 0. \tag{3.25}$$

For every $D \in \mathbb{N}$ (a small variation with different orders of differentiability of) Whitney's extension theorem (see, e.g., Theorem 2.3.6 in [9]) ensures that u^D can be extended to a function in $C^{1,2}([0,\infty)\times\mathbb{R}^{[D]\times[K]_0}, \mathbb{R})$. Then Itô's formula, (2.1), (3.25), (3.14), and Tonelli's theorem yield for all $D \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\big[u^D\big(T, \big(X_T^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big] - \mathbb{E}\big[u^D\big(0, \big(X_0^{D,k}(i)\big)_{(i,k)\in[D]\times[K]_0}\big)\big] \\
&= \int_0^T \mathbb{E}\Bigg[\sum_{i=1}^D \sum_{k=0}^K \frac{\partial u^D}{\partial x_{i,k}}\big(s, \big(X_s^{D,m}(j)\big)_{(j,m)\in[D]\times[K]_0}\big) \\
&\quad \times \Bigg\{\frac{1}{D}\sum_{j=1}^D X_s^{D,k-1}(j)\Bigg(f\Bigg(\sum_{m\in\mathbb{N}_0} X_s^{D,m}(j), \sum_{m\in\mathbb{N}_0} X_s^{D,m}(i)\Bigg) \\
&\quad - f\big(X_s^{D,k-1}(j), X_s^{D,k}(i)\big)\Bigg) \\
&\quad + \frac{X_s^{D,k}(i)}{\sum_{m\in\mathbb{N}_0} X_s^{D,m}(i)}\tilde{h}_D\Bigg(\sum_{m\in\mathbb{N}_0} X_s^{D,m}(i)\Bigg) - \tilde{h}_D\big(X_s^{D,k}(i)\big)\Bigg\} \\
&\quad + \frac{1}{2}\sum_{i=1}^D \sum_{k=0}^K \frac{\partial^2 u^D}{\partial x_{i,k}^2}\big(s, \big(X_s^{D,m}(j)\big)_{(j,m)\in[D]\times[K]_0}\big) \\
&\quad \times \Bigg\{\frac{X_s^{D,k}(i)}{\sum_{m\in\mathbb{N}_0} X_s^{D,m}(i)}\sigma^2\Bigg(\sum_{m\in\mathbb{N}_0} X_s^{D,m}(i)\Bigg) - \sigma^2\big(X_s^{D,k}(i)\big)\Bigg\}\Bigg]\,ds.
\end{aligned} \tag{3.26}$$

Setting 1.1 implies for all $D \in \mathbb{N}$, all $(x, y) \in \{(x_1, x_2) \in [0, 1]^2 : x_1 + x_2 \leq 1\}$, and all $\delta \in (0, 1)$ that

$$\begin{aligned} \left|\frac{x}{x+y}\tilde{h}_D(x+y) - \tilde{h}_D(x)\right| &\leq \frac{x}{x+y}\left|\tilde{h}_D(x+y) - \tilde{h}_D(x)\right| + \frac{y}{x+y}\left|\tilde{h}_D(x)\right| \\ &\leq 2L_h \frac{xy}{x+y} \\ &\leq 2L_h(x \wedge y) \\ &\leq \mathbf{1}_{x \leq \delta} 2L_h(x \wedge \delta) + \mathbf{1}_{x > \delta} 2L_h y \\ &\leq 2L_h(x \wedge \delta) + \frac{2L_h}{\delta} xy. \end{aligned} \tag{3.27}$$

Analogously, Setting 1.1 implies for all $(x, y) \in \{(x_1, x_2) \in [0, 1]^2 : x_1 + x_2 \leq 1\}$ and all $\delta \in (0, 1)$ that

$$\left|\frac{x}{x+y}\sigma^2(x+y) - \sigma^2(x)\right| \leq 2L_\sigma(x \wedge \delta) + \frac{2L_\sigma}{\delta} xy. \tag{3.28}$$

Equations (3.27) and (3.28) with $x = X_s^{D,k}(i)$ and $y = \sum_{m \in \mathbb{N}_0 \setminus \{k\}} X_s^{D,m}(i)$, Setting 1.1, and (3.26) show for all $\delta \in (0, 1)$ and all $D \in \mathbb{N}$ that

$$\begin{aligned} &\left|\mathbb{E}[u^D(T, (X_T^{D,k}(i))_{(i,k) \in [D] \times [K]_0})] - \mathbb{E}[u^D(0, (X_0^{D,k}(i))_{(i,k) \in [D] \times [K]_0})]\right| \\ &\quad \leq \sup_{t \in [0,T]} \|u^D(t, \cdot)\|_{C^2} T \\ &\qquad \times \Biggl(L_f \sup_{t \in [0,T]} \mathbb{E}\Biggl[\sum_{j=1}^D \sum_{k \in \mathbb{N}_0} X_t^{D,k-1}(j) \sum_{m \in \mathbb{N}_0 \setminus \{k-1\}} X_t^{D,m}(j)\Biggr] \\ &\qquad + \frac{L_f}{D} \mathbb{E}\Biggl[\sup_{t \in [0,T]} \Biggl(\sum_{i=1}^D \sum_{k \in \mathbb{N}_0} X_t^{D,k}(i)\Biggr)^2\Biggr] \\ &\qquad + (2L_h + L_\sigma) \sup_{t \in [0,T]} \mathbb{E}\Biggl[\sum_{i=1}^D \sum_{k \in \mathbb{N}_0} (X_t^{D,k}(i) \wedge \delta)\Biggr] \\ &\qquad + \frac{2L_h + L_\sigma}{\delta} \sup_{t \in [0,T]} \mathbb{E}\Biggl[\sum_{i=1}^D \sum_{k \in \mathbb{N}_0} X_t^{D,k}(i) \sum_{m \in \mathbb{N}_0 \setminus \{k\}} X_t^{D,m}(i)\Biggr]\Biggr). \end{aligned} \tag{3.29}$$

Lemma 3.3 and $\sup_{D \in \mathbb{N}} \|\psi^D\|_{C^2} < \infty$ imply that

$$\sup_{D \in \mathbb{N}} \sup_{t \in [0,T]} \|u^D(t, \cdot)\|_{C^2} < \infty.$$

The first and the fourth summand on the right-hand side of (3.29) converge to zero as $D \to \infty$ by Lemma 3.1 and assumptions (3.17) and (3.18). The second summand on the right-hand side of (3.29) converges to zero as $D \to \infty$ by Lemma 2.5 and assumption (3.17). The third summand on the right-hand side of (3.29) converges to zero uniformly in $D \in \mathbb{N}$ as $\delta \to 0$ by Lemma 3.2 and assumptions (3.16), (3.17), and (3.19). By letting first $D \to \infty$ and then $\delta \to 0$, (3.24) follows. This finishes the proof of Lemma 3.5. □

The following lemma shows that in the situation of Setting 2.16, the assumptions of Lemma 3.1 and Lemma 3.2 and some of the assumptions in Lemma 3.5 are satisfied.

LEMMA 3.6 (Well-behaved initial distribution). *Assume* Setting 2.16. *Then it holds that*

$$\sum_{k\in\mathbb{N}_0} \sup_{D\in\mathbb{N}} \mathbb{E}\left[\sum_{i=1}^{D} X_0^{D,k}(i)\right] < \infty, \tag{3.30}$$

that

$$\sup_{D\in\mathbb{N}} \mathbb{E}\left[\left(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i)\right)^2\right] < \infty, \tag{3.31}$$

that

$$\lim_{D\to\infty} \mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_0^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_0^{D,m}(i)\right] = 0, \tag{3.32}$$

and that

$$\lim_{\delta\to 0} \sup_{D\in\mathbb{N}} \mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} (X_0^{D,k}(i) \wedge \delta)\right] = 0.$$

PROOF. Equations (3.30), (3.31), and (3.32) follow immediately from the structure of the initial distribution given in Setting 2.16 and from (2.39). Moreover, Setting 2.16 and the dominated convergence theorem show that

$$\lim_{\delta\to 0} \sup_{D\in\mathbb{N}} \mathbb{E}\left[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} (X_0^{D,k}(i) \wedge \delta)\right] = \lim_{\delta\to 0} \mathbb{E}\left[\sum_{i=1}^{\infty}(X_0(i) \wedge \delta)\right] = 0.$$

This completes the proof of Lemma 3.6. □

The following lemma uses the Markov property in order to generalize Lemma 3.5 to finitely many time points.

LEMMA 3.7 (Asymptotic equality for f.d.d.s). *Assume* Setting 2.16, *let* $K \in \mathbb{N}_0$, *and let* $t_1, t_2, \ldots \in [0,\infty)$ *with* $t_1 < t_2 < \cdots$. *Then for all* $n \in \mathbb{N}$ *and all* $(\psi_j^D)_{D,j\in\mathbb{N}} \subseteq C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$ *which satisfy for all* $j \in \mathbb{N}$ *that* $\sup_{D\in\mathbb{N}} \|\psi_j^D\|_{C^2} < \infty$ *it holds that*

$$\begin{aligned}\lim_{D\to\infty}\Bigg|&\mathbb{E}\left[\prod_{j=1}^{n} \psi_j^D\big((X_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\right] \\ &- \mathbb{E}\left[\prod_{j=1}^{n} \psi_j^D\big((Z_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\right]\Bigg| = 0.\end{aligned} \tag{3.33}$$

PROOF. We prove (3.33) by induction on $n \in \mathbb{N}$. The base case $n = 1$ has been settled in Lemma 3.5, where the conditions (3.16), (3.17), (3.18), and (3.19) are satisfied due to Lemma 3.6. To show the induction step $\mathbb{N} \ni n \to n+1$, we fix $(\psi_j^D)_{D,j\in\mathbb{N}} \subseteq C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$ which satisfy for all $j \in \mathbb{N}$ that $\sup_{D\in\mathbb{N}} \|\psi_j^D\|_{C^2} < \infty$. For every $D \in \mathbb{N}$ we define the function $\psi^D : [0,1]^{[D]\times[K]_0} \to \mathbb{R}$ by

$$[0,1]^{[D]\times[K]_0} \ni x \mapsto \psi^D(x) := \psi_n^D(x)\mathbb{E}\big[\psi_{n+1}^D\big((Z_{t_{n+1}-t_n}^{D,k,x}(i))_{(i,k)\in[D]\times[K]_0}\big)\big].$$

Then Lemma 3.3 proves for every $D \in \mathbb{N}$ that $\psi^D \in C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$. Moreover, it follows from Lemma 3.3 that $\sup_{D\in\mathbb{N}} \|\psi^D\|_{C^2} < \infty$. Therefore, the induction hypothesis (applied to $\psi_1^D, \dots, \psi_{n-1}^D, \psi^D$) yields that

$$\begin{aligned}\lim_{D\to\infty}\Bigg|&\mathbb{E}\Bigg[\prod_{j=1}^{n-1}\psi_j^D\big((X_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\psi^D\big((X_{t_n}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\Bigg]\\ &-\mathbb{E}\Bigg[\prod_{j=1}^{n-1}\psi_j^D\big((Z_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\psi^D\big((Z_{t_n}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\Bigg]\Bigg|=0.\end{aligned}\tag{3.34}$$

By the Markov property it holds for all $D \in \mathbb{N}$ that

$$\begin{aligned}&\mathbb{E}\Bigg[\prod_{j=1}^{n-1}\psi_j^D\big((Z_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\psi^D\big((Z_{t_n}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\Bigg]\\ &\quad=\mathbb{E}\Bigg[\prod_{j=1}^{n+1}\psi_j^D\big((Z_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\Bigg].\end{aligned}$$

Moreover, we observe for all $D \in \mathbb{N}$ that

$$\begin{aligned}&\mathbb{E}\Bigg[\prod_{j=1}^{n-1}\psi_j^D\big((X_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\psi^D\big((X_{t_n}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\Bigg]\\ &\quad=\mathbb{E}\Bigg[\prod_{j=1}^{n}\psi_j^D\big((X_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\\ &\qquad\times\mathbb{E}\big[\psi_{n+1}^D\big((Z_{t_{n+1}-t_n}^{D,k,x}(i))_{(i,k)\in[D]\times[K]_0}\big)\big]\big|_{x=(X_{t_n}^{D,k}(i))_{(i,k)\in[D]\times\mathbb{N}_0}}\Bigg].\end{aligned}$$

When the initial distribution is given by $(X_{t_n}^{D,k}(i))_{(i,k)\in[D]\times\mathbb{N}_0}$, the conditions (3.16), (3.17), (3.18), and (3.19) are fulfilled due to Lemma 3.6, Lemma 2.4, Lemma 2.5, Lemma 3.1, and Lemma 3.2. Therefore, Lemma 3.5 implies that

$$\begin{aligned}\lim_{D\to\infty}\Bigg|&\mathbb{E}\Bigg[\prod_{j=1}^{n}\psi_j^D\big((X_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\\ &\quad\times\mathbb{E}\big[\psi_{n+1}^D\big((Z_{t_{n+1}-t_n}^{D,k,x}(i))_{(i,k)\in[D]\times[K]_0}\big)\big]\big|_{x=(X_{t_n}^{D,k}(i))_{(i,k)\in[D]\times\mathbb{N}_0}}\Bigg]\\ &-\mathbb{E}\Bigg[\prod_{j=1}^{n}\psi_j^D\big((X_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\\ &\quad\times\mathbb{E}\big[\psi_{n+1}^D\big((X_{t_{n+1}-t_n}^{D,k,x}(i))_{(i,k)\in[D]\times[K]_0}\big)\big]\big|_{x=(X_{t_n}^{D,k}(i))_{(i,k)\in[D]\times\mathbb{N}_0}}\Bigg]\Bigg|\\ =0.&\end{aligned}$$

The Markov property yields for all $D \in \mathbb{N}$ that

$$\begin{aligned}
&\mathbb{E}\Bigg[\prod_{j=1}^{n} \psi_j^D\big((X_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big) \\
&\qquad \times \mathbb{E}\big[\psi_{n+1}^D\big((X_{t_{n+1}-t_n}^{D,k,x}(i))_{(i,k)\in[D]\times[K]_0}\big)\big]\big|_{x=(X_{t_n}^{D,k}(i))_{(i,k)\in[D]\times\mathbb{N}_0}}\Bigg] \qquad (3.35)\\
&= \mathbb{E}\Bigg[\prod_{j=1}^{n+1} \psi_j^D\big((X_{t_j}^{D,k}(i))_{(i,k)\in[D]\times[K]_0}\big)\Bigg].
\end{aligned}$$

Combining (3.34) through (3.35) proves the induction step $\mathbb{N} \ni n \to n+1$ and hence finishes the proof of Lemma 3.7. □

The following lemma is the main result of Section 3.2 and shows that the migration level processes and the loop-free processes have the same limit as $D \to \infty$.

LEMMA 3.8 (Migration level and loop-free processes have the same limit). *Assume that* Setting 2.16 *holds*, *let* $n \in \mathbb{N}$, *let* $\phi_1, \dots, \phi_n \in C^2([0,1], \mathbb{R})$ *with the property that for all* $j \in [n]$ *it holds that* $\phi_j(0) = 0$, *let* $\psi \in C_b^2(\mathbb{R}, \mathbb{R})$, *and let* $t_1, \dots, t_n \in [0, \infty)$ *with* $t_1 < \cdots < t_n$. *Then it holds that*

$$\begin{aligned}
\lim_{D\to\infty}\Bigg|&\mathbb{E}\Bigg[\prod_{j=1}^{n} \psi\Bigg(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \phi_j(X_{t_j}^{D,k}(i))\Bigg)\Bigg] \\
&- \mathbb{E}\Bigg[\prod_{j=1}^{n} \psi\Bigg(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \phi_j(Z_{t_j}^{D,k}(i))\Bigg)\Bigg]\Bigg| = 0. \qquad (3.36)
\end{aligned}$$

PROOF. In a first step, we are going to reduce the considerations to $k \in [K]_0$ for finite $K \in \mathbb{N}_0$. The assumptions on $\phi_1, \dots, \phi_n$, and ψ imply the existence of constants $L_\phi, L_\psi \in [0, \infty)$ such that it holds for all $j \in [n]$ and all $x \in [0,1]$ that $|\phi_j(x)| \le L_\phi x$ and for all $x_1, \dots, x_n \in \mathbb{R}$ and all $y_1, \dots, y_n \in \mathbb{R}$ that $|\prod_{j=1}^n \psi(x_j) - \prod_{j=1}^n \psi(y_j)| \le L_\psi \sum_{j=1}^n |x_j - y_j|$. It follows for all $K \in \mathbb{N}_0$ that

$$\begin{aligned}
\sup_{D\in\mathbb{N}}\Bigg|&\mathbb{E}\Bigg[\prod_{j=1}^{n} \psi\Bigg(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \phi_j(X_{t_j}^{D,k}(i))\Bigg)\Bigg] \\
&- \mathbb{E}\Bigg[\prod_{j=1}^{n} \psi\Bigg(\sum_{i=1}^{D}\sum_{k=0}^{K} \phi_j(X_{t_j}^{D,k}(i))\Bigg)\Bigg]\Bigg| \qquad (3.37)\\
&\le L_\psi L_\phi \sum_{j=1}^{n}\sum_{k=K+1}^{\infty} \sup_{D\in\mathbb{N}} \mathbb{E}\Bigg[\sum_{i=1}^{D} X_{t_j}^{D,k}(i)\Bigg].
\end{aligned}$$

The right-hand side of (3.37) converges to zero as $K \to \infty$ by Lemma 3.6 and Lemma 2.4. The analogous statement holds when $X_{t_j}^{D,k}(i)$ is replaced by $Z_{t_j}^{D,k}(i)$ in (3.37). To prove (3.36) it therefore suffices to show for all $K \in \mathbb{N}_0$ that

$$\begin{aligned}
\lim_{D\to\infty}\Bigg|&\mathbb{E}\Bigg[\prod_{j=1}^{n} \psi\Bigg(\sum_{i=1}^{D}\sum_{k=0}^{K} \phi_j(X_{t_j}^{D,k}(i))\Bigg)\Bigg] \\
&- \mathbb{E}\Bigg[\prod_{j=1}^{n} \psi\Bigg(\sum_{i=1}^{D}\sum_{k=0}^{K} \phi_j(Z_{t_j}^{D,k}(i))\Bigg)\Bigg]\Bigg| = 0. \qquad (3.38)
\end{aligned}$$

We fix $K \in \mathbb{N}_0$ for the rest of the proof. For every $j \in [n]$ and $D \in \mathbb{N}$ we define the function $\psi_j^D : [0,1]^{[D]\times[K]_0} \to \mathbb{R}$ by

$$[0,1]^{[D]\times[K]_0} \ni (x_{i,k})_{(i,k)\in[D]\times[K]_0} = x \mapsto \psi_j^D(x) := \psi\left(\sum_{i=1}^{D}\sum_{k=0}^{K}\phi_j(x_{i,k})\right).$$

It follows for all $j \in [n]$ that $\sup_{D\in\mathbb{N}} \|\psi_j^D\|_\infty \le \|\psi\|_\infty$. Since $\phi_1, \ldots, \phi_n$, and ψ are twice continuously differentiable, it holds for all $j \in [n]$ and all $D \in \mathbb{N}$ that $\psi_j^D \in C^2([0,1]^{[D]\times[K]_0}, \mathbb{R})$. Furthermore, the chain rule and the product rule imply for all $j \in [n]$, all $D \in \mathbb{N}$, all $(\tilde{i}, \tilde{k}), (\tilde{j}, \tilde{l}) \in [D] \times [K]_0$, and all $x \in [0,1]^{[D]\times[K]_0}$ that

$$\left|\frac{\partial \psi_j^D}{\partial x_{\tilde{i},\tilde{k}}}(x)\right| = \left|\psi'\left(\sum_{i=1}^{D}\sum_{k=0}^{K}\phi_j(x_{i,k})\right)\phi_j'(x_{\tilde{i},\tilde{k}})\right| \le \|\psi'\|_\infty \|\phi_j'\|_\infty$$

and

$$\begin{aligned}\left|\frac{\partial^2 \psi_j^D}{\partial x_{\tilde{j},\tilde{l}}\partial x_{\tilde{i},\tilde{k}}}(x)\right| &= \left|\psi''\left(\sum_{i=1}^{D}\sum_{k=0}^{K}\phi_j(x_{i,k})\right)\phi_j'(x_{\tilde{j},\tilde{l}})\phi_j'(x_{\tilde{i},\tilde{k}})\right.\\ &\qquad\left. + \mathbf{1}_{(\tilde{i},\tilde{k})=(\tilde{j},\tilde{l})}\psi'\left(\sum_{i=1}^{D}\sum_{k=0}^{K}\phi_j(x_{i,k})\right)\phi_j''(x_{\tilde{i},\tilde{k}})\right| \\ &\le \|\psi''\|_\infty \|\phi_j'\|_\infty^2 + \|\psi'\|_\infty \|\phi_j''\|_\infty.\end{aligned}$$

It follows for all $j \in [n]$ that $\sup_{D\in\mathbb{N}} \|\psi_j^D\|_{C^2} < \infty$. Then Lemma 3.7 shows (3.38) which in turn finishes the proof of Lemma 3.8. $\square$

3.3. *Proof of Theorem* 1.4. PROOF OF THEOREM 1.4. In a first step, we prove Theorem 1.4 under the additional assumption that

$$\mathbb{E}\left[\left(\sum_{i=1}^{\infty} X_0(i)\right)^2\right] < \infty. \tag{3.39}$$

Analogously to the proofs of Lemma 2.15 and Lemma 2.17, one shows that

$$\left\{\left(\sum_{i=1}^{D} X_t^D(i)\delta_{X_t^D(i)}\right)_{t\in[0,\infty)} : D \in \mathbb{N}\right\} \tag{3.40}$$

is relatively compact. In the following, we identify the limit points of (3.40) by proving convergence of finite-dimensional distributions. For that, fix $n \in \mathbb{N}$, fix $\varphi_1, \ldots, \varphi_n \in C^2([0,1], \mathbb{R})$, fix $\psi \in C_b^2(\mathbb{R}, \mathbb{R})$, and fix $t_1, \ldots, t_n \in [0, \infty)$ with $t_1 < \cdots < t_n$. For every $j \in [n]$ we define the function $\phi_j : [0,1] \to \mathbb{R}$ by $[0,1] \ni x \mapsto \phi_j(x) := x\varphi_j(x) \in \mathbb{R}$. For every $D \in \mathbb{N}$ let

$$\{(X_t^{D,k}(i), W_t^k(i))_{t\in[0,\infty)} : (i,k) \in [D] \times \mathbb{N}_0\}$$

be a weak solution of (2.1) with initial distribution satisfying for all $i \in [D]$ that $\mathcal{L}(X_0^{D,0}(i)) = \mathcal{L}(X_0(i))$ and for all $(i,k) \in [D] \times \mathbb{N}$ that $\mathcal{L}(X_0^{D,k}(i)) = \delta_0$ and let $\{(Z_t^{D,k}(i))_{t\in[0,\infty)} : (i,k) \in [D] \times \mathbb{N}_0\}$ be a solution of (2.4) on the same probability space with Brownian motion given by the Brownian motion of the weak solution of (2.1) and started in $(X_0^{D,k}(i))_{(i,k)\in[D]\times\mathbb{N}_0}$.

Due to this and assumption (3.39), Setting 2.2 and Setting 2.16 are satisfied. First, Lemma 2.1 shows for all $D \in \mathbb{N}$ that

$$\mathbb{E}\Big[\prod_{j=1}^{n} \psi\Big(\sum_{i=1}^{D} \phi_j\big(X_{t_j}^{D}(i)\big)\Big)\Big] = \mathbb{E}\Big[\prod_{j=1}^{n} \psi\Big(\sum_{i=1}^{D} \phi_j\Big(\sum_{k\in\mathbb{N}_0} X_{t_j}^{D,k}(i)\Big)\Big)\Big]. \tag{3.41}$$

The assumptions on $\varphi_1, \dots, \varphi_n$ imply the existence of a constant $L_\phi \in [0,\infty)$ such that it holds for all $j \in [n]$ and all $x, y \in [0,1]$ that

$$\big|\phi_j(x) - \phi_j(y)\big| \le L_\phi |x-y|. \tag{3.42}$$

From this we obtain for all $D \in \mathbb{N}$, all $j \in [n]$, all $t \in [0,\infty)$, and all $\delta \in (0,\infty)$ that

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\sum_{i=1}^{D} \phi_j\Big(\sum_{m\in\mathbb{N}_0} X_t^{D,m}(i)\Big)\Big(1 - \sum_{k\in\mathbb{N}_0} \mathbf{1}_{\{X_t^{D,k}(i)\ge\delta\}}\Big)\Big|\Big] \\
&\quad\le L_\phi \mathbb{E}\Big[\sum_{i=1}^{D}\sum_{m\in\mathbb{N}_0} \big(X_t^{D,m}(i)\wedge\delta\big)\Big] \\
&\qquad + \frac{L_\phi}{\delta^2}\mathbb{E}\Big[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_t^{D,k}(i) \sum_{l\in\mathbb{N}_0\setminus\{k\}} X_t^{D,l}(i)\Big];
\end{aligned} \tag{3.43}$$

cf. (4.114) and (4.115) in [11]. The fact that there exists a constant $L_\psi \in [0,\infty)$ such that it holds for all $x_1, \dots, x_n \in \mathbb{R}$ and all $y_1, \dots, y_n \in \mathbb{R}$ that $|\prod_{j=1}^{n}\psi(x_j) - \prod_{j=1}^{n}\psi(y_j)| \le L_\psi \sum_{j=1}^{n}|x_j - y_j|$ together with (3.42) and (3.43) proves for all $D \in \mathbb{N}$ and all $\delta \in (0,1)$ that

$$\begin{aligned}
&\Big|\mathbb{E}\Big[\prod_{j=1}^{n} \psi\Big(\sum_{i=1}^{D} \phi_j\Big(\sum_{m\in\mathbb{N}_0} X_{t_j}^{D,m}(i)\Big)\Big)\Big] \\
&\qquad - \mathbb{E}\Big[\prod_{j=1}^{n} \psi\Big(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \phi_j\big(X_{t_j}^{D,k}(i)\big)\Big)\Big]\Big| \\
&\quad\le L_\psi \sum_{j=1}^{n}\mathbb{E}\Big[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \mathbf{1}_{\{X_{t_j}^{D,k}(i)\ge\delta\}}\Big|\phi_j\Big(\sum_{m\in\mathbb{N}_0} X_{t_j}^{D,m}(i)\Big) - \phi_j\big(X_{t_j}^{D,k}(i)\big)\Big|\Big] \\
&\qquad + L_\psi \sum_{j=1}^{n}\mathbb{E}\Big[\Big|\sum_{i=1}^{D} \phi_j\Big(\sum_{m\in\mathbb{N}_0} X_{t_j}^{D,m}(i)\Big)\Big(1 - \sum_{k\in\mathbb{N}_0} \mathbf{1}_{\{X_t^{D,k}(i)\ge\delta\}}\Big)\Big|\Big] \\
&\qquad + L_\psi \sum_{j=1}^{n}\mathbb{E}\Big[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \mathbf{1}_{\{X_{t_j}^{D,k}(i)<\delta\}}\big|\phi_j\big(X_{t_j}^{D,k}(i)\big)\big|\Big] \\
&\quad\le \frac{2L_\psi L_\phi}{\delta^2}\sum_{j=1}^{n}\mathbb{E}\Big[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} X_{t_j}^{D,k}(i) \sum_{m\in\mathbb{N}_0\setminus\{k\}} X_{t_j}^{D,m}(i)\Big] \\
&\qquad + 2L_\psi L_\phi \sum_{j=1}^{n}\mathbb{E}\Big[\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0} \big(X_{t_j}^{D,k}(i)\wedge\delta\big)\Big].
\end{aligned} \tag{3.44}$$

Lemma 3.6 and Lemma 3.1 ensure that the first summand on the right-hand side of (3.44) converges to zero as $D \to \infty$, while Lemma 3.6 and Lemma 3.2 show that the second summand on the right-hand side of (3.44) converges to zero uniformly in $D \in \mathbb{N}$ as $\delta \to 0$. By letting first $D \to \infty$ and then $\delta \to 0$, we therefore obtain from (3.44) that

$$\lim_{D\to\infty}\left|\mathbb{E}\left[\prod_{j=1}^{n}\psi\left(\sum_{i=1}^{D}\phi_j\left(\sum_{k\in\mathbb{N}_0}X_{t_j}^{D,k}(i)\right)\right)\right]\right.$$
$$\left.-\mathbb{E}\left[\prod_{j=1}^{n}\psi\left(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0}\phi_j\left(X_{t_j}^{D,k}(i)\right)\right)\right]\right|=0. \tag{3.45}$$

Lemma 2.17 shows that

$$\lim_{D\to\infty}\mathbb{E}\left[\prod_{j=1}^{n}\psi\left(\sum_{i=1}^{D}\sum_{k\in\mathbb{N}_0}\phi_j\left(Z_{t_j}^{D,k}(i)\right)\right)\right]$$
$$=\mathbb{E}\left[\prod_{j=1}^{n}\psi\left(\int\phi_j(\eta_{t_j-s})\mathcal{T}(ds\otimes d\eta)\right)\right]. \tag{3.46}$$

Combining (3.41), (3.45), Lemma 3.8, and (3.46) shows that

$$\lim_{D\to\infty}\mathbb{E}\left[\prod_{j=1}^{n}\psi\left(\sum_{i=1}^{D}\phi_j\left(X_{t_j}^{D}(i)\right)\right)\right]=\mathbb{E}\left[\prod_{j=1}^{n}\psi\left(\int\phi_j(\eta_{t_j-s})\mathcal{T}(ds\otimes d\eta)\right)\right].$$

This implies the convergence of finite-dimensional distributions of (3.40) and proves Theorem 1.4 under the additional assumption (3.39).

It remains to prove Theorem 1.4 in the case when (3.39) fails to hold. Fix a bounded continuous function $F : D([0,\infty), \mathcal{M}_{\mathrm{f}}([0,1])) \to \mathbb{R}$ for the rest of the proof. Then Setting 1.3 and the previous step imply that a.s.

$$\lim_{D\to\infty}\mathbb{E}\left[F\left(\left(\sum_{i=1}^{D}X_t^D(i)\delta_{X_t^D(i)}\right)_{t\in[0,\infty)}\right)\,\Big|\,\left(X_0(i)\right)_{i\in\mathbb{N}}\right]$$
$$=\mathbb{E}\left[F\left(\left(\int\eta_{t-s}\delta_{\eta_{t-s}}\mathcal{T}(ds\otimes d\eta)\right)_{t\in[0,\infty)}\right)\,\Big|\,\left(X_0(i)\right)_{i\in\mathbb{N}}\right].$$

Then it follows from taking expectations and from the dominated convergence theorem that

$$\lim_{D\to\infty}\mathbb{E}\left[F\left(\left(\sum_{i=1}^{D}X_t^D(i)\delta_{X_t^D(i)}\right)_{t\in[0,\infty)}\right)\right]$$
$$=\mathbb{E}\left[F\left(\left(\int\eta_{t-s}\delta_{\eta_{t-s}}\mathcal{T}(ds\otimes d\eta)\right)_{t\in[0,\infty)}\right)\right].$$

This finishes the proof of Theorem 1.4. □

Acknowledgment. The authors thank an anonymous referee for many valuable comments.

Both authors are supported by the German Research Foundation (DFG) through Priority Program 1590 "Probabilistic Structures in Evolution", Grant number HU 1889/3-2.

REFERENCES

[1] BUCKDAHN, R., DJEHICHE, B., LI, J. and PENG, S. (2009). Mean-field backward stochastic differential equations: A limit approach. *Ann. Probab.* **37** 1524–1565. MR2546754 https://doi.org/10.1214/08-AOP442

[2] CHETWYND-DIGGLE, J. A. and ETHERIDGE, A. M. (2018). SuperBrownian motion and the spatial Lambda-Fleming–Viot process. *Electron. J. Probab.* **23** Art. ID 71. MR3835477 https://doi.org/10.1214/18-EJP191

[3] CHETWYND-DIGGLE, J. A. and KLIMEK, A. (2019). Rare mutations in the spatial Lambda-Fleming–Viot model in a fluctuating environment and SuperBrownian motion. Available at arXiv:1901.04374.

[4] COX, J. T. and PERKINS, E. A. (2005). Rescaled Lotka–Volterra models converge to super-Brownian motion. *Ann. Probab.* **33** 904–947. MR2135308 https://doi.org/10.1214/009117904000000973

[5] DAWSON, D. A. and GREVEN, A. (2014). *Spatial Fleming–Viot Models with Selection and Mutation. Lecture Notes in Math.* **2092**. Springer, Cham. MR3155790 https://doi.org/10.1007/978-3-319-02153-9

[6] DURRETT, R. and PERKINS, E. A. (1999). Rescaled contact processes converge to super-Brownian motion in two or more dimensions. *Probab. Theory Related Fields* **114** 309–399. MR1705115 https://doi.org/10.1007/s004400050228

[7] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Wiley, New York. MR0838085 https://doi.org/10.1002/9780470316658

[8] GÄRTNER, J. (1988). On the McKean–Vlasov limit for interacting diffusions. *Math. Nachr.* **137** 197–248. MR0968996 https://doi.org/10.1002/mana.19881370116

[9] HÖRMANDER, L. (1990). *The Analysis of Linear Partial Differential Operators. I: Distribution Theory and Fourier Analysis*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **256**. Springer, Berlin. MR1065993 https://doi.org/10.1007/978-3-642-61497-2

[10] HUTZENTHALER, M. (2009). The Virgin Island model. *Electron. J. Probab.* **14** 1117–1161. MR2511279 https://doi.org/10.1214/EJP.v14-646

[11] HUTZENTHALER, M. (2012). Interacting diffusions and trees of excursions: Convergence and comparison. *Electron. J. Probab.* **17** Art. ID 71. MR2968678 https://doi.org/10.1214/EJP.v17-2278

[12] HUTZENTHALER, M., JORDAN, F. and METZLER, D. (2015). Altruistic defense traits in structured populations. Available at arXiv:1505.02154.

[13] HUTZENTHALER, M. and PIEPER, D. (2018). Differentiability of semigroups of stochastic differential equations with Hölder-continuous diffusion coefficients. Available at arXiv:1803.10608.

[14] HUTZENTHALER, M. and WAKOLBINGER, A. (2007). Ergodic behavior of locally regulated branching populations. *Ann. Appl. Probab.* **17** 474–501. MR2308333 https://doi.org/10.1214/105051606000000745

[15] KAC, M. (1956). Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954–1955, *Vol. III* 171–197. Univ. California Press, Berkeley, CA. MR0084985

[16] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

[17] KLENKE, A. (2014). *Probability Theory: A Comprehensive Course*, 2nd ed. *Universitext*. Springer, London. MR3112259 https://doi.org/10.1007/978-1-4471-5361-0

[18] LASRY, J.-M. and LIONS, P.-L. (2007). Mean field games. *Jpn. J. Math.* **2** 229–260. MR2295621 https://doi.org/10.1007/s11537-007-0657-8

[19] LIGGETT, T. M. (2010). *Continuous Time Markov Processes: An Introduction. Graduate Studies in Mathematics* **113**. Amer. Math. Soc., Providence, RI. MR2574430 https://doi.org/10.1090/gsm/113

[20] MCKEAN, H. P. JR. (1967). Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations* (*Lecture Series in Differential Equations*, *Session* 7, *Catholic Univ.*, 1967) 41–57. Air Force Office Sci. Res., Arlington, VA. MR0233437

[21] MÉLÉARD, S. and ROELLY-COPPOLETTA, S. (1987). A propagation of chaos result for a system of particles with moderate interaction. *Stochastic Process. Appl.* **26** 317–332. MR0923112 https://doi.org/10.1016/0304-4149(87)90184-0

[22] OELSCHLÄGER, K. (1984). A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.* **12** 458–479. MR0735849

[23] OELSCHLÄGER, K. (1985). A law of large numbers for moderately interacting diffusion processes. *Z. Wahrsch. Verw. Gebiete* **69** 279–322. MR0779460 https://doi.org/10.1007/BF02450284

[24] PITMAN, J. and YOR, M. (1982). A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete* **59** 425–457. MR0656509 https://doi.org/10.1007/BF00532802

[25] ROELLY-COPPOLETTA, S. (1986). A criterion of convergence of measure-valued processes: Application to measure branching processes. *Stochastics* **17** 43–65. MR0878553 https://doi.org/10.1080/17442508608833382

[26] SHIGA, T. and SHIMIZU, A. (1980). Infinite-dimensional stochastic differential equations and their applications. *J. Math. Kyoto Univ.* **20** 395–416. MR0591802 https://doi.org/10.1215/kjm/1250522207

[27] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX*—1989. *Lecture Notes in Math*. **1464** 165–251. Springer, Berlin. MR1108185 https://doi.org/10.1007/BFb0085169

[28] WAKELEY, J. and TAKAHASHI, T. (2004). The many-demes limit for selection and drift in a subdivided population. *Theor. Popul. Biol*. **66** 83–91. https://doi.org/10.1016/j.tpb.2004.04.005

[29] YAMADA, T. and WATANABE, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ*. **11** 155–167. MR0278420 https://doi.org/10.1215/kjm/1250523691