Electronic Journal of Probability

Electron. J. Probab. **17** (2012), no. 71, 1–49.
ISSN: 1083-6489 DOI: 10.1214/EJP.v17-2278

Interacting diffusions and trees of excursions: convergence and comparison

Martin Hutzenthaler*

Abstract

We consider systems of interacting diffusions with local population regulation representing populations on countably many islands. Our main result shows that the total mass process of such a system is bounded above by the total mass process of a tree of excursions with appropriate drift and diffusion coefficients. As a corollary, this entails a sufficient, explicit condition for extinction of the total mass as time tends to infinity. On the way to our comparison result, we establish that systems of interacting diffusions with uniform migration between finitely many islands converge to a tree of excursions as the number of islands tends to infinity. In the special case of logistic branching, this leads to a duality between a tree of excursions and the solution of a McKean-Vlasov equation.

Keywords: Island model; virgin island model; mean field model; McKean-Vlasov limit; extinction; excursion measure; many-demes-limit; propagation of chaos.
AMS MSC 2010: Primary 60K35, Secondary 60E15; 92D25.
Submitted to EJP on April 6, 2011, final version accepted on August 17, 2012.
Supersedes arXiv:1104.1060v1.

1 Introduction

In population dynamics and population genetics a prominent role is played by diffusion processes on I^G with $I = [0,\infty)$ or $I = [0,1]$ driven by stochastic differential equations (SDEs) of the form

$$dX_t(i) = \Big[\sum_{j\in G} X_t(j)m(j,i) - X_t(i) + \mu\big(X_t(i)\big)\Big]dt + \sqrt{\sigma^2\big(X_t(i)\big)}\,dB_t(i), \quad i \in G, \tag{1.1}$$

for $t \geq 0$ where G is a finite or countable set and $(B_t(i))_{t\geq 0}$, $i \in G$, are independent standard Brownian motions and where $(m(j,i))_{j,i\in G}$ is a stochastic matrix. We will refer to the solution $(X_t)_{t\geq 0}$ of (1.1) as (G,m,μ,σ^2)-process. In appropriate timescales and for suitable choices of μ and σ^2, the component $X_t(i)$ describes the (rescaled) *population size* on island $i \in G$ at time $t \geq 0$, or the relative frequency of a genetic type that is present on island $i \in G$ at time $t \geq 0$. The linear interaction term on the right-hand side of (1.1) models a *mass flow* between the islands, which might be caused by migration of individuals or a flow of genes. Here we will use a picture from population dynamics. The coefficient $\sigma^2(x)$ then is the infinitesimal variance of the local population size given its current value $x \in [0,\infty)$. A classical case are Feller's branching diffusions where $\sigma^2(x) = \mathrm{const}\cdot x$ for $x \in [0,\infty)$. Moreover the drift term μ describes the mean growth rate of a local population apart from immigration from other islands and emigration. A prototypical example is $\mu(x) = x(K-x)$, i.e. logistic growth.

*University of Munich, Germany. E-mail: hutzenthaler@bio.lmu.de

Here we address the question of the maximal effect of the migration matrix $m(i,j)_{i,j\in G}$ for fixed μ and σ^2. We will show in Theorem 3.8 that the total mass process $(\sum_{i\in G} X_t(i))_{t\geq 0}$ of $(X_t)_{t\geq 0}$ is dominated by a tree of excursions, which has been constructed in [21]. This dominating process does not depend on G or on the migration matrix. The intuition which leads to this comparison result is as follows. Consider a model with supercritical population-size independent branching and additional deaths due to competition within each island of G, i.e., μ is a concave function with $\mu(0)=0$ and σ^2 is a linear function. Now compare different distributions of individuals over space. If there is at most one individual on each island, then there are no deaths due to competition until the next birth or migration event. If, however, all individuals are on the same island, then there are deaths due to competition. The effect of the population-size independent branching is the same in both situations. We infer that more individuals survive if the distribution of individuals is more uniform over space. Now the migration dynamics which distributes mass uniformly over space would be uniform migration on G. As there is no uniform migration on an infinite set G, we approximate G with larger and larger finite subsets and consider uniform migration on the finite subsets. This intuition leads to considering an N-island model $(X_t^N)_{t\geq 0}$ for $N\in\mathbb{N}$ which is the solution of (1.1) with $G:=\{1,\ldots,N\}$ and $m(i,j):=\frac{1}{N}$, $i,j\in G$, that is, the solution of

$$dX_t^N(i)=\left[\frac{1}{N}\sum_{j=1}^{N}X_t^N(j)-X_t^N(i)+\mu\big(X_t^N(i)\big)\right]dt+\sqrt{\sigma^2\big(X_t^N(i)\big)}\,dB_t(i),\quad i\in\{1,2,\ldots,N\},\tag{1.2}$$

for $t\in[0,\infty)$ for every $N\in\mathbb{N}$. We will refer to this N-island model as (N,μ,σ^2)-process. Now if the initial configuration X_0^N converges in distribution to X_0 as $N\to\infty$, then the above intuition leads to the assertion that the total mass of the (G,m,μ,σ^2)-process is dominated by the limit of the total mass process of an N-island model $(X_t^N)_{t\geq 0}$ as $N\to\infty$

$$\sum_{i\in G}X_t(i)\leq\lim_{N\to\infty}\sum_{i=1}^{N}X_t^N(i),\quad t\geq 0,\tag{1.3}$$

where the stochastic order being used here will be specified in (3.14) below. We confirm this intuition in Theorem 3.8. Please note that a comparison result analogous to (1.3) of a (G,m,μ,σ^2)-process with an N-island process cannot be expected in general for finite $N\in\mathbb{N}$.

Next we review part of the literature on the comparison and on the limit in (1.3) and begin with the limits of the N-island model as $N\to\infty$. Starting with Kac (1957) and McKean (1967), the case of exchangeable initial configurations X_0^N, $N\in\mathbb{N}$, (so that $\sum_{j=1}^N X_0^N(j)$ is $O(N)$ as $N\to\infty$) has been studied intensively (e.g. [18, 30, 35] and the references therein). The most general result assumes the drift and the diffusion coefficient to depend continuously on x_i and on the measure $\frac{1}{N}\sum_{j=1}^N\delta_{x_j}$ for $x\in I^N$. If, in addition, certain Lyapunov conditions hold and if $\sigma(x)\neq 0$ for all $x\in I$, then $(X_t^N(i))_{t\geq 0}$ converges in distribution as $N\to\infty$ to a limiting process $(M_t)_{t\geq 0}$ for every $i\in\mathbb{N}$ which in our case solves the McKean-Vlasov equation

$$dM_t=\Big[\mathbb{E}M_t-M_t+\mu(M_t)\Big]\,dt+\sqrt{\sigma^2\big(M_t\big)}\,dB_t\tag{1.4}$$

for $t\in[0,\infty)$, see Theorem 4.1 of [18]. In particular, the Lyapunov assumptions of this theorem are satisfied if the coefficients are locally Lipschitz continuous and satisfy a linear growth condition, see Proposition 5.1 of [18]. To the best of our knowledge, our case of Hölder-$\frac{1}{2}$-continuous diffusion coefficients which are not strictly positive has not been considered so far. Proposition 4.29 below fills this gap. The idea of comparing a (G,m,μ,σ^2)-process with a limit of N-island processes is due to Hutzenthaler and Wakolbinger (2007). Their Proposition 2.2 establishes for the case of an associated initial configuration X_0 with identically distributed marginals and concave μ that $X_t(i)\leq M_t$ in a suitable stochastic order for every $t\in[0,\infty)$ and every $i\in G$. As a corollary hereof and of an extinction result for the McKean-Vlasov process (1.4), Hutzenthaler and Wakolbinger (2007) obtain a sufficient condition for local extinction of the (G,m,μ,σ^2)-process, see Theorem 1 in [23].

In this paper we focus on initial configurations with bounded total mass (so that $\sum_{j=1}^N X_0^N(j)$ is $O(1)$ as $N\to\infty$). The main difficulty here is that the limit of $(X_t^N)_{t\geq 0}$ as $N\to\infty$ was unknown so far. We will establish convergence of $(X_t^N)_{t\geq 0}$ as $N\to\infty$ in Theorem 3.3 below for the case of additive σ^2. This is the first convergence result of the N-island process in case of bounded total mass (together with the concurrent preprint [13] of Dawson and Greven who consider interacting Wright-Fisher diffusions with selection and rare mutations). The limiting process turns out to be a *forest of (mass) excursions* started in x_i, $i \in \mathbb{N}$. This forest of excursions has been constructed and analysed in [21] and has been denoted as *virgin island model*, see also Section 2 for a formal definition. In this model, every migrant populates a new island, similar to the infinite-alleles model in which every mutation produces a new allele. Jean Bertoin turned this idea into a *tree of alleles* [7] and a *partition into colonies* [6] (see also Section 7 of Pardoux and Wakolbinger (2010) for a connection with the virgin island model). In particular, Bertoin [7] shows in case of state-independent branching that the virgin island model is the diffusion approximation of a discrete mass branching process.

Here is a brief heuristics how a forest of (mass) excursions emerges from the (N, μ, σ^2)-processes as $N \to \infty$. Due to bounded initial mass, the total mass up to a finite time is also bounded in $N \in \mathbb{N}$, see Lemma 4.7. So the total mass that immigrates into a fixed island is of order $O(1/N)$ as $N \to \infty$. Among this immigrated mass might be a "successful" founder producing a substantial family. The probability of this event is of order $O(1/N)$ (this follows from (2.4) since $S(\varepsilon) \sim S^{'}(0)\varepsilon$ as $\varepsilon \to \infty$). Among N islands, there is then a Poisson number of islands having a founder whose progeny reaches a fixed level $\delta > 0$, say. Moreover all of these founders are on distinct islands. The reason for this is that the probability to have two successful founders on the same island is of order $O(1/N^2)$. Consequently, this does not appear within N islands in the limit as $N \to \infty$, see Lemma 4.21 for the details. Thus, in the limit as $N \to \infty$, every "successful" emigrant populates a previously unpopulated island. The evolution of the population size on every freshly populated island is described by a random (mass) excursion. These mass excursions are born (densely in time) in a Poissonian manner on ever new islands with an intensity proportional to the currently extant mass, and, once born, follow the SDE

$$dY_t = -Y_t dt + \mu(Y_t)dt + \sqrt{\sigma^2(Y_t)}\, dB_t\,. \tag{1.5}$$

Formally, this is described by means of the *excursion measure* Q associated with (1.5) in the sense of Pitman and Yor (1982) (see also [21]). The intensity measure with which a path $(\eta_t)_{t\geq 0}$ spawns a "daughter" excursion born at time $t \geq 0$ is $\eta_t\, dt \otimes Q$. The roots of the forest are random paths which are independent solutions of (1.5). The virgin island model is then a countable family $\mathcal{V} := \{(s, \chi)\}$ of islands where island (s, χ) is populated at time $s \geq 0$ and carries mass χ_{t-s} at time $t \geq 0$.

One tree of excursions in the forest of excursions is illustrated in Figure 1. Note that Figure 1 does not contain the whole tree. In fact islands are populated by emigrants densely in time but only finitely many excursions started by these emigrants reach a given strictly positive height. Now a noteworthy observation is that the tree structure provides us with independence of disjoint subtrees. Putting this differently, the virgin island model is a branching process in discrete time in the sense of Jiřina (1958) except that there are now infinitely many types, one for each excursion path. Due to this branching structure, the virgin island model is easier to study than the N-island process. Several authors considered analogue of the virgin island model in the case of state-independent branching; see [1, 5, 6, 7, 14, 20, 29, 42, 46] for a selection of articles.

To state our convergence result of Theorem 3.3 more formally, denote the *population size spectrum* of the (N, μ, σ^2)-process as $\zeta_t^N := \sum_{i=1}^{N} \delta_{X_t^N(i)}$ for all $t \in [0, \infty)$ and for every $N \in \mathbb{N}$. Furthermore define the population size spectrum of the virgin island model as $\zeta_t := \sum_{(s,\chi)\in\mathcal{V}} \delta_{\chi_{t-s}}$ for all $t \in [0, \infty)$. Our convergence result asserts that the population size spectra converge in distribution, i.e., $\zeta^N \to \zeta$ in distribution as $N \to \infty$; see Theorem 3.3 for the precise statement.

Now we state the comparison result (1.3) more precisely. Recall that μ is the infinitesimal mean in a non-spatial situation. We assume μ to be *subadditive*. Then a population of size x that is separated into two islands experiences (in sum) a larger growth rate than a population of the same size that is concentrated on one island. Thus the virgin island model should offer in expectation a more prolific evolution of the total mass than a model (1.1) with the same coefficients μ and σ^2. The infinitesimal variance σ^2 has an impact on a comparison in distribution. More precisely the stochastic order in (1.3) depends on whether σ^2 is superadditive, additive or subadditive. In our prototype example of additive σ^2, the stochastic order

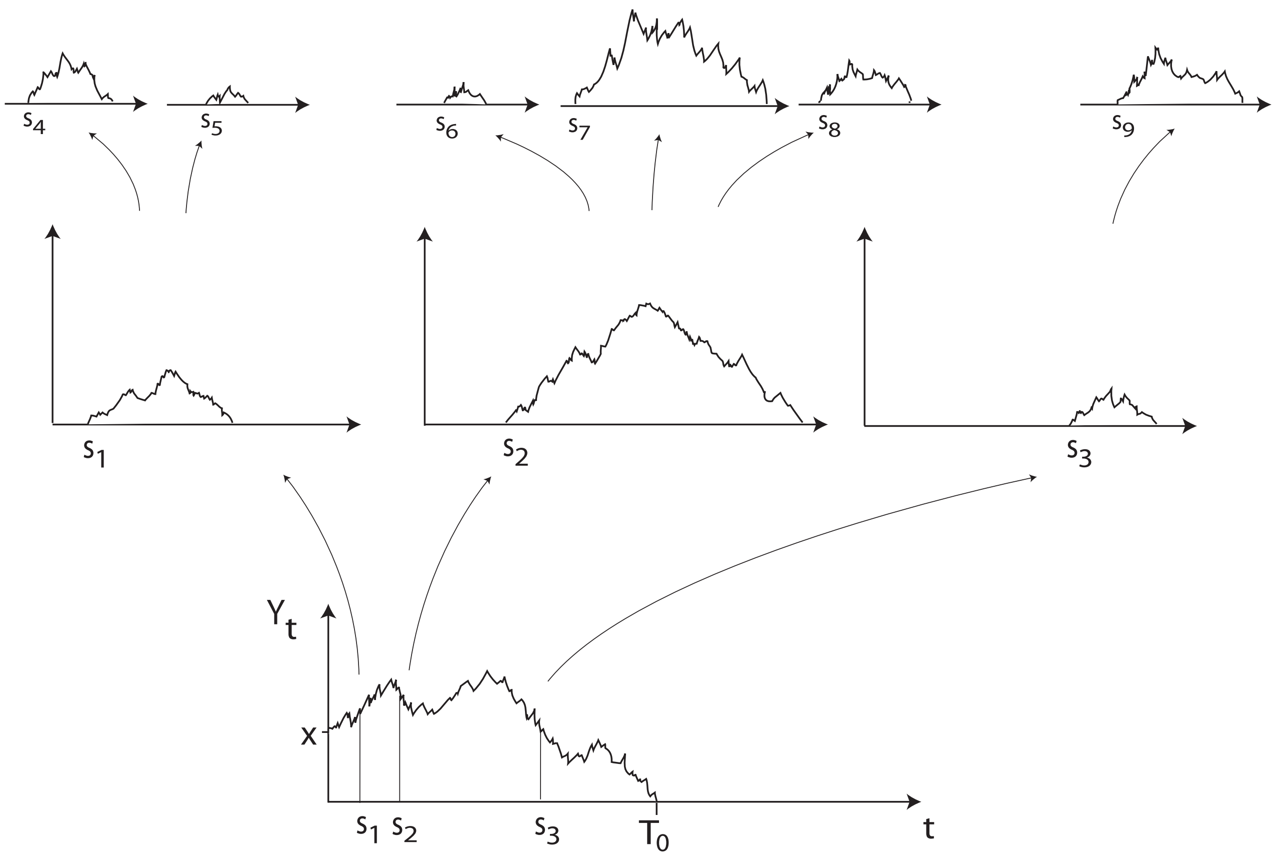

Figure 1: Subtree of the Virgin Island Model. Only offspring islands with a certain excursion height are drawn. Note that infinitely many islands are colonized e.g. between times s_1 and s_2.

is the usual increasing order. In that case we have that

$$\mathbb{E}\Big[f\Big(\sum_{i\in G} X_t(i)\Big)\Big] \leq \mathbb{E}\Big[f\Big(\sum_{(s,\chi)\in\mathcal{V}} \chi_{t-s}\Big)\Big], \quad t \geq 0, \tag{1.6}$$

for every non-decreasing function $f\colon [0,\infty) \to [0,\infty)$. If σ^2 is superadditive or subadditive, then we will use concave, non-decreasing functions and convex, non-decreasing functions, respectively. In fact Theorem 3.8 is a comparison result not only on the one-dimensional, but on the finite dimensional distributions. Thereby results on the distribution of the total mass of the virgin island model have an immediate impact on the distribution of interacting diffusions with local population regulation.

As a very special application of our comparison result, we obtain a sufficient condition for global extinction for interacting diffusions with local population regulation. Here we speak of global extinction if the total mass $\sum_{i\in G} X_t(i)$ converges to zero in distribution as $t \to \infty$ whenever $\sum_{i\in G} X_0(i) < \infty$. Theorem 2 of [21] shows under certain assumptions that global extinction of the virgin island model with coefficients μ and σ is equivalent to

$$\int_0^\infty \frac{y}{\sigma^2(y)/2} \exp\left(\int_0^y \frac{-x+\mu(x)}{\sigma^2(x)/2}\,dx\right) dy \leq 1. \tag{1.7}$$

As a consequence of (1.6), if condition (1.7) is satisfied, then the (G, m, μ, σ^2)-process dies out globally no matter what the migration matrix $m(\cdot,\cdot)$ is. Here are further implications of our comparison result. Theorem 2 of [21] together with (1.6) implies an upper bound for the survival probability of $(X_t)_{t\geq 0}$. Theorem 3 of [21] yields an upper bound for $\mathbb{E}[\int_0^\infty \sum_{i\in G} X_s(i)ds]$ if the left-hand side of (1.7) is strictly smaller than one and an upper bound for the growth rate of $\int_0^t \mathbb{E}[\sum_{i\in G} X_s(i)ds]$ as $t \to \infty$ if (1.7) fails to hold. Moreover Theorem 4 of [21] implies an upper bound for the growth rate of $\sum_{i\in G} X_t(i)$ as $t \to \infty$ if (1.7) fails to hold.

Here is a selection of models for which one could think of a similar comparison result. Mueller and Tribe (1994) investigate a one-dimensional SPDE analog of interacting Feller branching diffusions with logistic growth. Bolker and Pacala (1997) propose a branching random walk in which the individual mortality rate is increased by a weighted sum of the entire population. Etheridge (2004) studies two diffusion limits hereof. The “stepping stone version of the Bolker-Pacala model” is a system of interacting Feller branching diffusions with non-local logistic growth. The “superprocess version of the Bolker-Pacala model” is an analog of this in continuous space. Blath, Etheridge and Meredith (2007) study a two-type version hereof, which is a spatial extension of the classical Lotka-Volterra model. Fournier and Méléard (2004) generalize the model of Bolker and Pacala (1997) by allowing spatial dependence of all rates. A model in discrete time and discrete space is constructed in Birkner and Depperschmidt (2007). In that paper an individual has a Poisson number of offspring with mean depending on the current configuration and, once created, offspring take an independent random walk step from the location of their mother.

2 The virgin island model

The virgin island model without immigration has been introduced in [21]. Here we slightly generalize this model by adding independent immigration of mass.

The virgin island model is an analog of (1.1) in which every emigrant populates a new island. Islands with positive mass at time zero evolve as the one-dimensional diffusion $(Y_t)_{t\geq 0}$ solving (1.5). The following assumption guarantees existence and uniqueness of a strong $[0,\infty)$-valued solution of equation (1.5), see e.g. Theorem IV.3.1 in [24].

Assumption 2.1. *The set I is an interval of length $|I|\in(0,\infty]$ which is either of the form $[0,|I|]$ if $|I|<\infty$ or of the form $[0,\infty)$ if $|I|=\infty$. The functions $\mu\colon I\to\mathbb{R}$ and $\sigma^2\colon I\to[0,\infty)$ are locally Lipschitz continuous in I and satisfy $\mu(0)=0=\sigma^2(0)$ and if $|I|<\infty$, then $\mu(|I|)\leq 0=\sigma^2(|I|)$. The function $\sigma^2(\cdot)$ is strictly positive on $(0,|I|)$ and the function $\mu(\cdot)$ is globally upward Lipschitz continuous, that is, $\mu(x)-\mu(y)\leq L_\mu|x-y|$ whenever $x>y\in I$ where $L_\mu\in[0,\infty)$ is a finite constant. Furthermore σ^2 satisfies the growth condition $\sigma^2(y)\leq L_\sigma(y+y^2)$ for all $y\in I$ where $L_\sigma\in[0,\infty)$ is a finite constant.*

Note that zero is a trap for $(Y_t)_{t\geq 0}$, that is, $Y_t=0$ implies $Y_{t+s}=0$ for all $s\geq 0$.

Mass emigrates from each island at rate one and colonizes new islands. A new population should evolve as the process $(Y_t)_{t\geq 0}$ and should start from a single individual which has mass zero due to the diffusion approximation. Thus we need the law of excursions of $(Y_t)_{t\geq 0}$ from the trap zero. For this, define the set of excursions from zero by

$$U:=\big\{\chi\in\mathbf{C}\big((-\infty,\infty),[0,\infty)\big)\colon T_0\in(0,\infty],\,\chi_t=0\ \ \forall\, t\in(-\infty,0]\cup[T_0,\infty)\big\} \tag{2.1}$$

where $T_y=T_y(\chi):=\inf\{t>0\colon \chi_t=y\}$ is the first hitting time of $y\in[0,\infty)$. The set U is furnished with locally uniform convergence. For existence of the excursion measure Q and in order to apply the results of [21], we need to assume additional properties of $\mu(\cdot)$ and of $\sigma^2(\cdot)$. For the motivation of these assumptions, we refer the reader to [21]. Assume $\int_0^\varepsilon \frac{y}{\sigma^2(y)}dy<\infty$ for some $0<\varepsilon<|I|$. Then the scale function $S\colon I\to[0,\infty)$ defined through

$$s(z):=\exp\Big(-\int_0^z\frac{-x+\mu(x)}{\sigma^2(x)/2}\,dx\Big),\quad S(y):=\int_0^y s(z)\,dz,\quad z,y\in I. \tag{2.2}$$

is well-defined.

Assumption 2.2. *The functions $\mu(\cdot)$ and $\sigma^2(\cdot)$ satisfy*

$$\int_0^\varepsilon\frac{y}{\sigma^2(y)}\,dy<\infty \text{ and } \int_\varepsilon^{|I|}\frac{y}{\sigma^2(y)s(y)}\,dy<\infty \tag{2.3}$$

for some $0<\varepsilon<|I|$.

Under Assumption 2.2, the process $(Y_t)_{t\geq 0}$ hits zero in finite time almost surely and the expected total emigration intensity of the virgin island model is finite, see Lemma 9.5 and Lemma 9.6 in [21]. Moreover the scale function $S(\cdot)$ is well-defined and satisfies $S^{'}(0) = 1$. A generic example which satisfies Assumption 2.2 is $\mu(y) = c_1 y^{\kappa_1} - c_2 y^{\kappa_2}$, $\sigma^2(y) = c_3 y^{\kappa_3}$ where $c_1, c_2, c_3 > 0$, $\kappa_2 > \kappa_1 \geq 1$ and $\kappa_3 \in [1, 2)$. Assumption 2.2 is not met by $\sigma^2(y) = y^2$.

The excursion measure Q of the SDE (1.5) has first been described in Pitman and Yor (1982). Here we use the description of Q as limit of the law of $(Y_t)_{t\geq 0}$ started in $\varepsilon > 0$ and rescaled with $S(\varepsilon)$ as $\varepsilon \to 0$. More formally, under Assuming 2.1 and 2.2, Theorem 1 in [21] implies that there exists a unique σ-finite measure Q on U such that

$$\lim_{\varepsilon\to 0} \frac{1}{S(\varepsilon)} \mathbb{E}^{\varepsilon}\big[F(Y)\big] = \int F(\chi) Q(d\chi) \tag{2.4}$$

for every bounded continuous function $F\colon \mathbf{C}\big([0,\infty),[0,\infty)\big) \to \mathbb{R}$ for which there exists a $\delta > 0$ such that $F(\chi) = 0$ whenever $\sup_{t\geq 0} \chi_t < \delta$. The reader might want to think of Q as describing the evolution of a population founded by a single individual. In the special case $\sigma^2(y) = 2\beta y$, $\mu(y) = cy$ with $\beta > 0$ and $c \in \mathbb{R}$, the process $(Y_t)_{t\geq 0}$ is Feller's branching diffusion whose law is infinitely divisible. In that case the excursion measure coincides with the canonical measure.

Having introduced the excursion measure, we now construct the virgin island model with constant immigration rate $\theta \in [0,\infty)$ and started in $(x_k)_{k\in\mathbb{N}} \subset I$. Let $\{(Y_t^{k,x_k})_{t\geq 0}\colon k \in \mathbb{N}\}$ be independent solutions of (1.5) such that $Y_0^{k,x_k} = x_k$ almost surely. Moreover let Π^θ be a Poisson point process on $[0,\infty) \times U$ with intensity measure

$$\mathbb{E}\big[\Pi^\theta(dt \otimes d\psi)\big] = \theta\, dt \otimes Q(d\psi). \tag{2.5}$$

The elements of the Poisson point process Π^θ are the islands whose founders immigrated into the system. Next we construct all islands which are colonized from a given mother island. Let $\{\Pi^{(n,s,\chi)}\colon (n,s,\chi) \in \mathbb{N}_0 \times [0,\infty) \times \mathrm{C}\big([0,\infty), I\big)\}$ be a set of independent Poisson point processes on $[0,\infty) \times U$ with intensity measure

$$\mathbb{E}\big[\Pi^{(n,s,\chi)}(dt \otimes d\psi)\big] = \chi_{t-s}\, dt \otimes Q(d\psi) \quad (n,s,\chi) \in \mathbb{N}_0 \times [0,\infty) \times \mathrm{C}\big([0,\infty), I\big). \tag{2.6}$$

All ingredients are assumed to be independent. The elements of the Poisson point process $\Pi^{(n,s,\chi)}$ are the islands which descend from an island with population size trajectory $(\chi_{t-s})_{t\geq 0}$ and where the ancestral lineages of individuals living on these islands have exactly $n \in \mathbb{N}_0$ migration events. Now the 0-th generation is a random σ-finite measure on $[0,\infty) \times \mathrm{C}\big([0,\infty), I\big)$ defined through $\mathcal{V}^{(0)} := \sum_{k\in\mathbb{N}} \delta_{(0,Y^{k,x_k})} + \Pi^\theta$. The $(n+1)$-st generation, $n \geq 0$, is the random σ-finite measure on all islands which have been colonized from islands of the n-th generation, that is, $\mathcal{V}^{(n+1)} := \sum_{(s,\xi)\in\mathcal{V}^{(n)}} \Pi^{(n,s,\xi)}$ for all $n \in \mathbb{N}_0$. The virgin island model $\mathcal{V}$ is then the sum of all of these measures

$$\mathcal{V} := \sum_{n\in\mathbb{N}_{\geq 0}} \mathcal{V}^{(n)}. \tag{2.7}$$

We call $\mathcal{V}$ the virgin island model with immigration rate θ and initial configuration $(x_k)_{k\in\mathbb{N}}$.

3 Main results

We begin with convergence of the N-island process. In this convergence, we allow the drift function μ_N and the diffusion function σ_N to depend on N in order to include the case of weak immigration. For example, one could be interested in an N-island model with logistic branching and weak immigration at rate $\frac{\theta}{N}$ on each island. In that case, one would set $\mu_N(x) = \frac{\theta}{N} + \gamma x(K-x)$ and $\sigma_N^2(x) := x$ for $x \in I$. The equation of the N-island process now reads as

$$dX_t^N(i) = \left[\frac{1}{N}\sum_{j=1}^N X_t^N(j) - X_t^N(i) + \mu_N\big(X_t^N(i)\big)\right] dt + \sqrt{\sigma_N^2\big(X_t^N(i)\big)}\, dB_t(i) \tag{3.1}$$

where $i = 1, \ldots, N$ and where $(B_t(i))_{t\geq 0}$, $i \in \mathbb{N}$, are independent standard Brownian motions. The idea to include weak immigration into a convergence result is due to [12] who independently obtain convergence of an N-island model using different methods.

Define $\tilde{\mu}_N(x) := \mu_N(x) - \mu_N(0)$ for $x \in I$. For the (N, μ_N, σ_N^2)-process to converge, we need assumptions on μ_N, σ_N^2 and on the initial distribution.

Assumption 3.1. *Define $I := [0,\infty)$. The functions $\mu_N, \mu\colon I \to \mathbb{R}$ are locally Lipschitz continuous on I. The sequence $(\mu_N)_{N\in\mathbb{N}}$ converges pointwise to μ as $N \to \infty$. In addition, $N\cdot\mu_N(0) \to \theta \in [0,\infty)$ as $N \to \infty$ and $0 \le N\mu_N(0) \le 2\theta$ for all $N \in \mathbb{N}$. The diffusion functions σ_N^2 and σ^2 are linear, that is, $\sigma_N^2(x) = \beta_N x$ and $\sigma^2(x) = \beta x$ for some constants $\beta_N, \beta > 0$ and all $N \in \mathbb{N}$. Furthermore β_N converges to β as $N \to \infty$. Assumptions 2.1 and 2.2 hold for μ and σ^2. Moreover $(\mu_N)_{N\in\mathbb{N}}$ is uniformly upward Lipschitz continuous, that is, $\mu_N(x) - \mu_N(y) \le L_\mu|x-y|$ for all $x \ge y \in I$, $N \in \mathbb{N}$ and some constant $L_\mu \in [0,\infty)$.*

Here is an example. If $\mu_N(x) = \frac{\theta}{N} + C_1 x^{\kappa_1} - C_2 x^{\kappa_2}$ and $\sigma_N^2(x) = x$, then Assumption 3.1 is satisfied if $1 \le \kappa_1 < \kappa_2$ and $C_2(\kappa_1 - 1) > 0$.

Assumption 3.2. *The random variables $(X_0(i))_{i\in\mathbb{N}}$ and $(X_0^N(i))_{i\le N}$ are defined on the same probability space for each $N \in \mathbb{N}$. There exists a random permutation $\pi_1^N, \ldots, \pi_N^N$ of $\{1,\ldots,N\}$ for each $N \in \mathbb{N}$ such that*

$$\lim_{N\to\infty} \mathbb{E}\Big[\sum_{i=1}^{N} |X_0(i) - X_0^N(\pi_i^N)|\Big] = 0. \tag{3.2}$$

Furthermore the total mass of $X_0(\cdot)$ has finite expectation $\mathbb{E}\big[\sum_{i\in\mathbb{N}} X_0(i)\big] < \infty$.

If $(x_i)_{i\in\mathbb{N}} \subset I$ is a summable sequence, then Assumption 3.2 is satisfied for $X_0(i) = x_i$ and $X_0^N(i) = x_i$, $i \le N \in \mathbb{N}$.

Next we introduce the topology for the weak convergence of the N-island process. What will be relevant here is not any specific numbering of the islands but the statistics (or “spectrum”) of their population sizes, described by the sum of Dirac measures at each time point, that is,

$$\Big(\sum_{i=1}^{N} \delta_{X_t^N(i)}\Big)_{t\le T} \tag{3.3}$$

where δ_x is the Dirac measure on $x \in I$. The state space of the measure-valued process (3.3) is the set $\mathcal{M}_\sigma\big(I\big)$ of σ-finite measures on I. We equip the state space $\mathcal{M}_\sigma\big(I\big)$ with the vague topology on $I \setminus \{0\}$. For weak convergence of $\mathcal{M}_\sigma\big(I\big)$-valued processes, we equip the space of càdlàg-functions from $[0,\infty)$ to $\mathcal{M}_\sigma\big(I\big)$ with the Skorokhod topology (e.g. [16]).

Now we formulate the convergence of the (N, μ_N, σ_N^2)-process defined in (3.1).

Theorem 3.3. *Suppose that $\big(\mu_N\big)_{N\in\mathbb{N}}$ and $\big(\sigma_N^2\big)_{N\in\mathbb{N}}$ satisfy Assumption 3.1 and that the initial configurations $(X_0^N(i))_{i\le N}$ and $(X_0(i))_{i\in\mathbb{N}}$ satisfy Assumption 3.2. Then, for every $T \in [0,\infty)$, we have that*

$$\Big(\sum_{i=1}^{N} \delta_{X_t^N(i)}\Big)_{t\le T} \xrightarrow{w} \Big(\sum_{(s,\eta)\in\mathcal{V}} \delta_{\eta_{t-s}}\Big)_{t\le T} \qquad \text{as } N \to \infty \tag{3.4}$$

in distribution where $\mathcal{V}$ is the virgin island model with immigration rate $\theta = \lim_{N\to\infty} N\mu_N(0)$ and initial configuration $(X_0(i))_{i\in\mathbb{N}}$.

The proof is deferred to Section 4.

Remark 3.4. *For readability we rewrite the convergence in terms of test functions. The weak convergence (3.4) is equivalent to*

$$\lim_{N\to\infty} \mathbb{E}\bigg[F\bigg(\Big(\sum_{i=1}^{N} f\big(X_t^N(i)\big)\Big)_{t\le T}\bigg)\bigg] = \mathbb{E}\bigg[F\bigg(\Big(\sum_{(s,\eta)\in\mathcal{V}} f\big(\eta_{t-s}\big)\Big)_{t\le T}\bigg)\bigg]. \tag{3.5}$$

for every bounded continuous function F on $\mathrm{C}\big([0,T],\mathbb{R}\big)$ and every continuous function $f\colon I \to \mathbb{R}$ with compact support in $(0,|I|)$. This equivalence follows from Theorem 2.2 of [38] if $\mathcal{M}_\sigma\big(I\big)$ is equipped with the weak topology, the case of the vague topology follows analogously. Applications often require functions f with non-compact support. The following condition might be useful in that case. Let $\bar{F}$ be a continuous function on $\mathrm{C}\big([0,T],\mathbb{R}\big)$ satisfying the Lipschitz condition

$$\big|\bar{F}(\eta) - \bar{F}(\bar{\eta})\big| \le L_{\bar{F}} \sum_{j=1}^{n} \big|\eta_{t_j} - \bar{\eta}_{t_j}\big| \qquad \forall\, \eta, \bar{\eta} \in \mathrm{C}\big([0,T],\mathbb{R}\big) \tag{3.6}$$

for some $0 \leq t_1 \leq \cdots \leq t_n \leq T$. In addition let $\bar{f}\colon I \to \mathbb{R}$ be a continuous function satisfying $|\bar{f}(x)| \leq L_{\bar{f}} x$ for all $x \in I$. Then following the arguments in the proof of Lemma 4.21 below, one can show that (3.5) *holds with F and f replaced by $\bar{F}$ and $\bar{f}$, respectively.* $\diamond$

The assumptions of Theorem 3.3 are satisfied for branching diffusions with local population regulation. A prominent example is the N-island model with logistic drift $\mu(x) = \gamma x(K - x)$ and with $\sigma^2(x) = 2\beta x$ for $x \in [0, \infty)$ and some constants $\gamma, K, \beta > 0$. More generally, Theorem 3.3 can be applied if $\mu(x) = \gamma x - c(x)$ and $\sigma^2(x) = 2\beta x$ for $x \in [0, \infty)$ where $c(\cdot)$ is a concave function with $c'(0) \in \mathbb{R}$. We believe that Theorem 3.3 also holds for non-linear infinitesimal variances such as $\sigma^2(x) = x(1 - x)$ in case of the Wright-Fisher diffusion. Our proof requires linearity only for one argument which is the step from equation (4.105) to equation (4.106).

In case of logistic branching, we obtain a noteworthy duality of the total mass process

$$V_t := \sum_{(s,\chi)\in\mathcal{V}} \chi_{t-s}, \quad t \geq 0. \tag{3.7}$$

of the virgin island model with the mean field model $(M_t)_{t\geq 0}$ defined in (1.4). By Theorem 3 of [23], systems of interacting Feller branching diffusions with logistic drift satisfy a duality relation which for the $(N, \gamma y(K - y), 2\beta y)$-process reads as

$$\mathbb{E}^{y\delta_1}\left[e^{-\frac{\gamma}{\beta}x\sum_{i=1}^N X_t^N(i)}\right] = \mathbb{E}^{x\underline{1}}\left[e^{-\frac{\gamma}{\beta}X_t^N(1)y}\right] \qquad \forall\, x, y, t \geq 0 \tag{3.8}$$

where the notation $\mathbb{E}^{y\delta_1}$ refers to the initial configuration $X_0^N = (y, 0, \ldots, 0)$ and $\mathbb{E}^{x\underline{1}}$ refers to $X_0^N = (x, \ldots, x)$. This duality is established in [23] via a generator calculation, in Swart (2006) via dualities between Lloyd-Sudbury particle models and in [3] by following ancestral lineages of forward and backward processes in a graphical representation. Now let $N \to \infty$ in (3.8). Then the left-hand side converges to the Laplace transform of the total mass process of the virgin island model (without immigration) according to Theorem 3.3 and the right-hand side converges to the Laplace transform of the mean field model (1.4) according to Proposition 4.29 below. This proves the following corollary.

Corollary 3.5. *Let $(V_t)_{t\geq 0}$ be the total mass process of the virgin island model without immigration starting on only one island. Furthermore let $(M_t)_{t\geq 0}$ be the solution of* (1.4)*, both with coefficients $\mu(y) = \gamma y(K - y)$ and $\sigma^2(y) = 2\beta y$ for $y \in [0, \infty)$ where $\gamma, K, \beta > 0$. Then*

$$\mathbb{E}^y\left[e^{-\frac{\gamma}{\beta}xV_t}\right] = \mathbb{E}^x\left[e^{-\frac{\gamma}{\beta}M_t y}\right] \qquad \forall\, x, y, t \geq 0 \tag{3.9}$$

where $\mathbb{E}^y$ and $\mathbb{E}^x$ refer to $V_0 = y$ and $M_0 = x$, respectively.

Together with known results on the mean field model (1.4), this corollary leads to a computable expression for the extinction probability of the virgin island model.

Corollary 3.6. *Let $(V_t)_{t\geq 0}$ be as in Corollary 3.5. Then V_t converges to a random variable V_∞ in distribution as $t \to \infty$. If*

$$\int_0^\infty \exp\left(K\gamma x - \frac{\gamma\beta}{2}x^2\right)\cdot \exp(-x)\, dx \leq 1, \tag{3.10}$$

then $\mathbb{P}[V_\infty = 0] = 1$. If condition (3.10) *fails to hold, then*

$$\mathbb{P}^y[V_\infty = 0] = 1 - \mathbb{P}^y[V_\infty = \infty] = \int_0^\infty e^{-\frac{\gamma}{\beta}yx}\,\Gamma_\rho(dx) \in (0, 1) \tag{3.11}$$

where $\mathbb{P}^y$ refers to $V_0 = y \in (0, \infty)$. The parameter $\rho \in (0, \infty)$ is the unique solution of

$$\int_0^\infty y^{\frac{\rho}{\beta}}(K - y)\exp\Big(\frac{\gamma K - 1}{\beta}y - \frac{\gamma}{2\beta}y^2\Big)\, dy = 0 \tag{3.12}$$

and the probability distribution Γ_ρ is defined by

$$\Gamma_\rho(dx) = \frac{C_\rho}{\beta x}\exp\Big(\int_K^x \frac{(\rho - z) + \gamma z(K - z)}{\beta z} dz\Big)\, dx \tag{3.13}$$

on $(0, \infty)$ where C_ρ is a normalizing constant.

Proof. Theorem 2 of [21] shows convergence in distribution of V_t to V_∞ as $t \to \infty$ and $\mathbb{P}(V_\infty = 0) = 1$ if (3.10) holds. If (3.10) fails to hold, then Corollary 3.5 together with convergence of V_t implies convergence in distribution of M_t to a variable M_∞ as $t \to \infty$. The distribution of M_∞ is necessarily an invariant distribution of the mean field model (1.4) and is nontrivial. Lemma 5.1 of [23] shows that there is exactly one nontrivial invariant distribution for (1.4) and this distribution is given by (3.13). □

The second main result is a comparison of systems of locally regulated diffusions with the virgin island model. For its formulation, we introduce three stochastic orders which are inspired by Cox et al. (1996). Let $Z = (Z_t)_{t\geq 0}$ and $\tilde{Z} = (\tilde{Z}_t)_{t\geq 0}$ be two stochastic processes with state space I. We say that Z is dominated by $\tilde{Z}$ with respect to a set $\mathbb{F}$ of test functions on path space if

$$Z \leq_{\mathbb{F}} \tilde{Z} \;:\Longleftrightarrow\; \mathbb{E}\big[F(Z)\big] \leq \mathbb{E}\big[F(\tilde{Z})\big] \quad \forall\, F \in \mathbb{F}. \tag{3.14}$$

The first order is 'the usual stochastic order' $\leq_{\mathrm{st}}$ in which Z is dominated by $\tilde{Z}$ if there is a coupling of Z and $\tilde{Z}$ in which Z_t is dominated by $\tilde{Z}_t$ for all $t \geq 0$ almost surely. Assuming path continuity, an equivalent condition is as follows. Denote the set of non-decreasing test functions of $n \in \mathbb{N}_{\geq 0}$ arguments by

$$\mathcal{F}_{+\pm}^{(n)} := \mathcal{F}_{+\pm}^{(n)}(S) := \Big\{ f\colon S^n \to \mathbb{R} \,|\, f \text{ is non-decreasing, } f \text{ is bounded or } f \geq 0 \Big\} \tag{3.15}$$

for a set $S \subseteq [0,\infty)$. Furthermore let $\mathbb{F}_{+\pm}$ be the set of non-decreasing functions which depend on finitely many time-space points

$$\begin{aligned}\mathbb{F}_{+\pm} := \mathbb{F}_{+\pm}(G,S) := \Big\{ F\colon \mathrm{C}\big([0,\infty)\times G, S\big) \to \mathbb{R} \,|\, \exists n \in \mathbb{N}_0\, \exists\, (t_1,i_1),\ldots,(t_n,i_n) \in [0,\infty)\times G \\ \exists f \in \mathcal{F}_{+\pm}^{(n)}(S) \text{ such that } F(\eta) = f\big(\eta_{t_1}(i_1),\ldots,\eta_{t_n}(i_n)\big) \Big\}.\end{aligned} \tag{3.16}$$

If there is no space component, then we simply write $\mathbb{F}_{+\pm}(S)$. In this notation, $Z \leq_{\mathrm{st}} \tilde{Z}$ is equivalent to $Z \leq_{\mathbb{F}_{+\pm}} \tilde{Z}$, see Subsection 4.B.1 in Shaked and Shanthikumar (1994).

We will use two more stochastic orders. In the literature, the set of non-decreasing, convex functions is often used. Here an adequate set is the collection of non-decreasing functions whose second order partial derivatives are non-negative. As we do not want to assume smoothness, we slightly weaken the latter assumption. We say for $1 \leq i,j \leq n$ that a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is (i,j)-*convex* if

$$f(z + h_1 e_i + h_2 e_j) - f(z + h_1 e_i) - f(z + h_2 e_j) + f(z) \geq 0 \quad \forall\, z \in \mathbb{R}^n, h_1, h_2 \geq 0. \tag{3.17}$$

Note that if f is smooth, then this is equivalent to $\frac{\partial^2}{\partial x_i \partial x_j} f \geq 0$. In addition note that f is (i,i)-convex if and only if f is convex in the i-th component. Moreover we say that f is (i,j)-concave if $-f$ is (i,j)-convex. A function is called directionally convex (e.g. Shaked and Shanthikumar 1990) if it is (i,j)-convex for all $1 \leq i,j \leq n$. Such functions are also referred to as L-superadditive functions (e.g. Rüschendorf 1983). Define the set of increasing, directionally convex functions as

$$\mathcal{F}_{++}^{(n)} := \Big\{ f \in \mathcal{F}_{+\pm}^{(n)} \colon\ f \text{ is } (i,j)\text{-convex for all } 1 \leq i,j \leq n \Big\} \tag{3.18}$$

and similarly $\mathcal{F}_{+-}$ with '(i,j)-convex' replaced by '(i,j)-concave'. Furthermore define $\mathbb{F}_{++}$ and $\mathbb{F}_{+-}$ as in (3.16) with $\mathcal{F}_{+\pm}^{(n)}$ replaced by $\mathcal{F}_{++}^{(n)}$ and $\mathcal{F}_{+-}^{(n)}$, respectively. Now we have introduced three stochastic orders $\leq_{\mathbb{F}_{+\pm}}$, $\leq_{\mathbb{F}_{+-}}$ and $\leq_{\mathbb{F}_{++}}$. Note that $\mathcal{F}_{++}^{(n)}$ contains all mixed monomials and that $\mathcal{F}_{+-}^{(n)}$ contains all functions $1 - \exp\big(-\sum_{i=1}^n \lambda_i x_i\big)$ with $\lambda_1,\ldots,\lambda_n \geq 0$.

For the solution $(X_t)_{t\geq 0}$ of (1.1) to be well-defined, we additionally assume the migration matrix to be substochastic.

Assumption 3.7. *The set G is (at most) countable and the matrix $(m(j,i))_{j,i\in G}$ is non-negative and substochastic, i.e., $m(j,k) \geq 0$ and $\sum_{i\in G} m(j,i) \leq 1$ for all $j,k \in G$.*

Note that Assumption 2.1 together with Assumption 3.7 guarantees existence and uniqueness of a strong solution of (1.1) with values in $\{x \in I^G \colon |x| < \infty\}$. This follows from Proposition 2.1 and inequality (48) of [23] by letting the weight function $\sigma_i \nearrow 1$ for $i \in G$ and using monotone convergence.

Theorem 3.8. *Assume 2.1, 2.2 and 3.7. If μ is concave and if σ^2 is superadditive, then we have that*

$$\Big(\sum_{i\in G} X_t(i)\Big)_{t\geq 0} \leq_{\mathbb{F}_{+-}([0,\infty))} \Big(V_t\Big)_{t\geq 0}. \tag{3.19}$$

If μ is concave and σ^2 is subadditive, then inequality (3.19) *holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{++}$. If μ is subadditive and σ^2 is additive, then inequality* (3.19) *holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{+\pm}$.*

The proof is deferred to Section 5. Comparisons of diffusions at fixed time points are well-known. Cox et al. (1996) establish a comparison between the finite-dimensional distributions of two diffusions where the test functions have product structure. To the best of our knowledge, Theorem 3.8 is the first comparison result with general test functions on the finite-dimensional distributions. The techniques we develop for this in Subsection 5.1 might allow to generalize the comparison results of Cox et al. (1996) on interacting diffusions and the comparison results of Bergenthum and Rüschendorf (2007) on semimartingales.

The assumption of μ being subadditive is natural in the following sense. Let us assume that letting two 1-island processes with initial masses x and y, respectively, evolve independently is better in expectation for the total mass than letting one 1-island process with initial mass $x+y$ evolve. This assumption implies that

$$\mu(x+y) = \lim_{t\to 0}\frac{\mathbb{E}X_t^{x+y}-x-y}{t} \leq \lim_{t\to 0}\frac{\mathbb{E}X_t^{x}+\mathbb{E}X_t^{y}-x-y}{t} = \mu(x)+\mu(y) \tag{3.20}$$

for all $x,y,x+y\in I$ and thus subadditivity of the infinitesimal mean μ. If σ^2 is not additive, then we need the stronger assumption of μ being concave for Lemma 5.4.

From Theorem 3.8 and a global extinction result for the virgin island model, we obtain a condition for global extinction of systems of locally regulated diffusions. According to Theorem 2 of [21], the total mass of the virgin island model converges in distribution to zero as $t\to\infty$ if and only if condition (3.21) below is satisfied. Together with Theorem 2, this proves the following corollary.

Corollary 3.9. *Assume 2.1, 3.7 and 2.2. Suppose that μ is subadditive and σ^2 is additive, or that μ is concave and σ^2 is superadditive. Then*

$$\int_0^{|I|} \frac{y}{\sigma^2(y)/2}\exp\left(\int_0^y \frac{-x+\mu(x)}{\sigma^2(x)/2}\,dx\right)dy \leq 1, \tag{3.21}$$

implies global extinction of the solution $(X_t)_{t\geq 0}$ of (1.1)*, that is, $\sum_{i\in G} X_t(i) \xrightarrow{w} 0$ as $t\to\infty$ whenever $\sum_{i\in G} X_0(i)<\infty$ almost surely.*

Proof. The function $[0,\infty)\ni x\mapsto 1-e^{-\lambda x}\in\mathcal{F}_{+\pm}\cap\mathcal{F}_{+-}$ for every $\lambda\in[0,\infty)$. If σ^2 is superadditive (or even additive), then Theorem 2 above and Theorem 2 of [21] imply that

$$\mathbb{E}\left[1-e^{-\lambda\sum_{i\in G}X_t(i)}\right] \leq \mathbb{E}\left[1-e^{-\lambda V_t}\right] \xrightarrow{t\to\infty} 0 \tag{3.22}$$

for all $\lambda\in[0,\infty)$. Convergence of the Laplace transform then implies weak convergence. □

In case of logistic branching ($\mu(y)=\gamma y(K-y)$, $\sigma^2(y)=2\beta y$), condition (3.21) simplifies to condition (3.10).

4 Convergence to the virgin island model

4.1 Outline

First we outline the intuition behind the proof. The virgin island process is a tree of excursions whereas the N-island process has no tree structure. It happens in the latter process that different emigrants colonize the same island. In addition, the N-island process is not loop-free. An individual could migrate from island 1 to island 3 and then back to island 1. That these two effects vanish in the limit as the number of islands tends to infinity will be established in two separate steps.

The first step ensures that the limit of the N-island process as $N\to\infty$ is loop-free. For this purpose, we decompose the N-island process according to the number of migration steps. Throughout the paper, we say that an individual has *migration level* $k \in \mathbb{N}_0$ at time $t \in [0,\infty)$ if its ancestral lineage contains exactly k migration steps. For example, an individual starting on island 1 at time 0, moving to island 3 and then back to island 1 has migration level 2. Let $N \in \mathbb{N}$. We define a system $\{(X_t^{N,k}(i))_{t\geq 0} \colon k \in \mathbb{N}_0, i \leq N\}$ of diffusions such that $X_t^{N,k}(i)$ consists of the mass on island i at time t with migration level k. Recall $\tilde{\mu}_N(x) = \mu_N(x) - \mu_N(0)$ for all $x \in I$. Define $X_0^{N,k}(i) = \mathbb{1}_{k=0} X_0^N(i)$ for all $i \in \{1,\dots,N\}$ and $k \in \mathbb{N}_0$. Let $\big\{\big(X_t^{N,k}(i), B_t^k(i)\big)_{t\geq 0} \colon i \leq N, k \in \mathbb{N}_0\big\}$ be a solution of

$$\begin{aligned} dX_t^{N,k}(i) = & \left(\frac{1}{N}\sum_{j=1}^N X_t^{N,k-1}(j) - X_t^{N,k}(i)\right) dt + \mathbb{1}_{k=0}\mu_N(0)\,dt \\ & + \frac{X_t^{N,k}(i)}{\sum_{m\geq 0} X_t^{N,m}(i)} \tilde{\mu}_N\bigg(\sum_{m\geq 0} X_t^{N,m}(i)\bigg) dt \\ & + \sqrt{\frac{X_t^{N,k}(i)}{\sum_{m\geq 0} X_t^{N,m}(i)} \sigma_N^2\Big(\sum_{m\geq 0} X_t^{N,m}(i)\Big)}\, dB_t^k(i) \end{aligned} \tag{4.1}$$

for all $t \in [0,\infty)$, $i = 1,\dots,N$ and $k \in \mathbb{N}_0$ where $X_t^{N,-1} := 0$ for $t \geq 0$ and $i \leq N$ and where the family $\{(B_t^k(i))_{t\geq 0} \colon i,k \in \mathbb{N}_0\}$ is a system of independent standard Brownian motions. Here we implicitly used the continuous extension of $\frac{x}{x+y}\tilde{\mu}_N(x+y)$ and of $\frac{x}{x+y}\sigma_N^2(x+y)$ as functions of $(x,y) \in [0,\infty)^2 \setminus \{(0,0)\}$ into the point $(0,0)$, where $N \in \mathbb{N}$. Any weak solution of (4.1) will be denoted as a (N,μ_N,σ_N^2)-process with migration levels. See Lemma 4.3 for existence of a weak solution of (4.1).

Lemma 4.23 below indicates that the individuals with migration level k at a fixed time are concentrated on essentially finitely many islands in the limit $N \to \infty$. A later migration event will not hit these essentially finitely many islands because hitting a fixed island has probability $\frac{1}{N}$. Therefore we expect that all individuals on an island have the same migration level. Inserting this into (4.1) suggests to consider the solution $\big\{(Z_t^{N,k}(i))_{t\geq 0} \colon i \leq N, k \in \mathbb{N}_0\big\}$ of

$$\begin{aligned} dZ_t^{N,k}(i) = & \bigg(\frac{1}{N}\sum_{j=1}^N Z_t^{N,k-1}(j) - Z_t^{N,k}(i) + \mathbb{1}_{k=0}\mu_N(0) + \tilde{\mu}_N\big(Z_t^{N,k}(i)\big)\bigg)\, dt \\ & + \sqrt{\sigma_N^2\big(Z_t^{N,k}(i)\big)}\, dB_t^k(i), \quad Z_0^{N,k}(i) = X_0^{N,k}(i), \qquad i = 1,\dots,N, \end{aligned} \tag{4.2}$$

where $Z_t^{N,-1} := 0$ for all $t \geq 0$ and $i \leq N$. We will refer to this solution as the loop-free (N,μ_N,σ_N^2)-process or as loop-free N-island model. Note that this is a $(\bar{G},\bar{m},\mu,\sigma^2)$-process with $\bar{G} := \{1,2,\dots,N\} \times \mathbb{N}_0$ and migration matrix $\bar{m}\big((i,k),(j,l)\big) = \frac{1}{N}\mathbb{1}_{l=k+1}$ for $(i,k),(j,l) \in \bar{G}$. In particularly, we may and will choose $\big\{(Z_t^{N,k}(i))_{t\geq 0} \colon i \leq N, k \in \mathbb{N}_0\big\}$ to be the solution of (4.2) with respect to the Brownian motion of a weak solution of (4.1) for every $N \in \mathbb{N}$. Consequently there exists a unique strong solution under Assumption 2.1. Lemma 4.25 below establishes the assertion that the distance between the (N,μ_N,σ_N^2)-process with migration levels and the loop-free (N,μ_N,σ_N^2)-process converges to zero in a suitable sense as $N \to \infty$. It turns out that some difficulties arise from the different forms of the diffusion coefficients in the (N,μ_N,σ_N^2)-process with migration levels and in the loop-free (N,μ_N,σ_N^2)-process. As we could not resolve these difficulties, we additionally assume for Lemma 4.25 that σ_N^2 is linear. Then we have that $x\sigma_N^2(y)/y = \sigma_N^2(x)$ for $x = X_t^{N,k}(i)$ and $y = \sum_{m=0}^\infty X_t^{N,m}(i)$ and the diffusion coefficients in (4.1) and in (4.2) are similar. Our proof of Lemma 4.25 is a moment estimate in the spirit of Yamada and Watanabe (1971).

In Subsection 4.3 we show that two emigrants colonize different islands in the limit $N \to \infty$. Let us rephrase this more formally. Recall that $\{(B_t^k(i))_{t\geq 0} \colon i \leq N\}$ is independent of $\{(B_t^l(i))_{t\geq 0} \colon l < k, i \leq N\}$. Thus, conditioned on $\{\sum_{j=1}^N Z_t^{N,k-1}(j) \colon t \geq 0\}$, the loop-free (N,μ_N,σ_N^2)-process $(Z_t^{N,k}(i))_{t\geq 0}$ on island i with migration level $k \geq 1$ evolves as the solution of

$$dY_{t,s}^{N,\zeta} = \frac{\zeta_N(t)}{N}\,dt - Y_{t,s}^{N,\zeta}dt + \tilde{\mu}_N\big(Y_{t,s}^{N,\zeta}\big)\,dt + \sqrt{\sigma_N^2\big(Y_{t,s}^{N,\zeta}\big)}dB_t,\ t \geq s, \tag{4.3}$$

starting at time $s = 0$ in $Y_{s,s}^{N,\zeta} = 0$ driven by the Brownian motion $(B_t)_{t\geq 0} = (B_t^k(i))_{t\geq 0}$ for each $i \leq N$ where $\zeta_N(t) := \sum_{j=1}^N Z_t^{N,k-1}(j)$. Note that $(Y_{t,s}^{N,\zeta}(i))_{t\geq s}$, $i \leq N$, are independent and identically distributed. Now let $\zeta_N\colon [0,\infty) \to N\cdot I := \{N\cdot x\colon x\in I\}$ be a fixed path and let $(\bar{Y}^{N,\zeta}_{t,s}(i))_{t\geq s}$, $i\leq N$, be independent solutions of (4.3). We are interested in the total mass $\big(\sum_{i=1}^N \bar{Y}^{N,\zeta}_{t,s}(i)\big)_{t\geq s}$ as $N\to\infty$. As $\bar{Y}^{N,\zeta}_{s,s}(i)=0$ and as the immigration rate on island i tends to zero, the process $(\bar{Y}^{N,\zeta}_{t,s}(i))_{t\geq s}$ converges to zero as $N\to\infty$ for every $i\in\mathbb{N}$. However, as mass of order $O(\frac{\zeta_N}{N})$ immigrates on a fixed island, the probability that the excursion started by these immigrants reaches a certain level $\delta>0$, say, is of order $O(\frac{\zeta_N}{N})$ as the convergence in (2.4) indicates. Now as there are N independent trials, the Poisson limit theorem should imply that

$$\Big(\sum_{i=1}^N \delta_{\bar{Y}^{N,\zeta}_{t,s}(i)}\Big)_{s\leq t\leq T} \xrightarrow{w} \Big(\int \delta_{\eta_{t-u}}\Pi(du,d\eta)\Big)_{s\leq t\leq T} \qquad \text{as } N\to\infty \tag{4.4}$$

where Π is a Poisson point process with intensity measure $\lim_{N\to\infty}\zeta_N(u)du\otimes Q(d\eta)$ if this limit exists. We will prove (4.4) in Lemma 4.21 by reversing time.

For convergence of the loop-free (N,μ_N,σ_N^2)-process, we do not need to assume linearity of the diffusion function. Here we may replace Assumption 3.1 with the following weaker assumption.

Assumption 4.1. *The functions $\mu_N,\mu\colon I\to\mathbb{R}$ and $\sigma_N^2,\sigma^2\colon I\to[0,\infty)$ are locally Lipschitz continuous on I. The sequence $(\mu_N,\sigma_N^2)_{N\in\mathbb{N}}$ converges pointwise to (μ,σ^2) as $N\to\infty$. In addition, $N\cdot\mu_N(0)\to\theta\in[0,\infty)$ as $N\to\infty$ and $N\mu_N(0)\leq 2\theta$ for all $N\in\mathbb{N}$. The functions μ_N and σ_N^2 satisfy $\mu_N(0)\geq 0=\sigma_N^2(0)$ and if $|I|<\infty$, then $\mu_N(|I|)\leq 0=\sigma_N^2(|I|)$. Assumption 2.1 and 2.2 hold for μ and σ^2. Moreover $(\mu_N)_{N\in\mathbb{N}}$ is uniformly upward Lipschitz continuous in zero, that is $\mu_N(x)-\mu_N(y)\leq L_\mu|x-y|$ for all $x\geq y\in I$, $N\in\mathbb{N}$ and some constant $L_\mu\in[0,\infty)$. The sequence $(\sigma_N^2)_{N\in\mathbb{N}}$ satisfies the uniform growth condition $\sigma_N^2(y)\leq L_\sigma(y+y^2)$ for all $y\in I$, $N\in\mathbb{N}$ where $L_\sigma\in[0,\infty)$ is a finite constant and satisfies that $\liminf_{0<y\to 0}\inf_{N\in\mathbb{N}}\sigma_N^2(y)/\sigma^2(y)>0$.*

Note that if σ_N^2 is linear, then Assumption 4.1 implies Assumption 3.1.

some steps of our proof are based on second-moment estimates and require the following assumption of uniformly finite second moments of the initial distribution. This assumption is then relaxed in further steps.

Assumption 4.2. *The initial distribution satisfies that*

$$\sup_{N\in\mathbb{N}}\mathbb{E}\Big[\Big(\sum_{i=1}^N X_0^N(i)\Big)^2\Big]<\infty. \tag{4.5}$$

4.2 Preliminaries

In this subsection we establish preliminary results such as moment estimates and existence of the processes. The quick reader might want to skip this subsection. We begin with weak existence of the N-island process with migration levels.

Lemma 4.3. *Assume 4.1. The (N,μ_N,σ_N^2)-process with migration levels exists in the weak sense, that is, equation* (4.1) *has a weak solution for every $N\in\mathbb{N}$.*

Proof. As the proof is fairly standard, we only give an outline. Approximate (4.1) with stochastic differential equations for which weak solutions exist. For example, approximate μ_N and σ_N^2 locally uniformly with bounded continuous functions $\mu_{N,n}$ and $\sigma_{N,n}^2$, respectively. Consider the solution $(X_t^{N,k,n})_{t\geq 0}$ of (4.1) with μ_N and σ_N^2 replaced by $\mu_{N,n}$ and $\sigma_{N,n}^2$, respectively, and which only depends on the migration levels $k\leq n$. Then this solution has a weak solution according to Theorem V.23.5 and Theorem V.20.1 of [40] as the coefficients are bounded and continuous and the stochastic differential equation is finite-dimensional. Show that the formal generator hereof converges to the formal generator associated with (4.1). In addition establish tightness of $(X^{N,k,n})_{n\in\mathbb{N}}$ using moment estimates as in Lemma 4.7 for fixed $N\in\mathbb{N}$ but uniformly in $n\in\mathbb{N}$. Then apply the tightness criterion of Aldous (1978). Then there exists a converging subsequence and its limit solves the martingale problem associated with (4.1), see Lemma 4.5.1 in [16]. From this solution of the martingale problem, construct a weak solution of (4.1) as in Theorem V.20.1 of [40]. □

Next we prove that the N-island model with migration levels is indeed a decomposition of the N-island model (3.1).

Lemma 4.4. *Assume 4.1. Fix $N \in \mathbb{N}$ and let $\big\{\big(X_t^{N,k}(i), B_t^k(i)\big)_{t\geq 0}\colon k \in \mathbb{N}_0, i \leq N\big\}$ be a solution of* (4.1)*. Then $\big\{(\tilde{X}_t^N(i))_{t\geq 0}\colon i \leq N\big\}$ defined through*

$$\tilde{X}_t^N(i) := \sum_{k\geq 0} X_t^{N,k}(i), \quad t \geq 0, i \leq N, \tag{4.6}$$

is the unique solution of the N-island model (3.1) *corresponding to standard Brownian motions defined through*

$$dB_t(i) = \mathbb{1}_{\sum_{m\geq 0} X_t^{N,m}(i)>0} \sum_{k\geq 0} \sqrt{\frac{X_t^{N,k}(i)}{\sum_{m\geq 0} X_t^{N,m}(i)}} dB_t^k(i) + \mathbb{1}_{\sum_{m\geq 0} X_t^{N,m}(i)=0} dB_t^0(i) \quad t \geq 0 \tag{4.7}$$

for every $i \leq N$.

Proof. The process $(B_t(i))_{t\geq 0}$ is a continuous martingale with quadratic variation process $[B(i), B(j)]_t = \delta_{ij} t$. Therefore Lévy's characterization (e.g. Theorem IV.33.1 in [40]) implies that (4.7) defines a standard Brownian motion. Moreover it follows from summing (4.1) over $k \in \mathbb{N}_0$ that $\big(\tilde{X}_t^N(i)\big)_{t\geq 0}$ solves (3.1). Pathwise uniqueness of (3.1) has been established in Proposition 2.1 of [23]. □

In the following lemmas, let the process $\{(X_t^{N,k}(i), B_t^k(i))_{t\geq 0}\colon k \in \mathbb{N}_0, i \leq N\}$ be a solution of (4.1) and let the process $\{(Z_t^{N,k}(i), B_t^k(i))_{t\geq 0}\colon k \in \mathbb{N}_0, i \leq N\}$ be the solution of (4.2). Define a stopping time $\tilde{\tau}_K^N \in [0, \infty)$ through

$$\tilde{\tau}_K^N := \inf\Big\{t \geq 0\colon \sum_{i=1}^N \sum_{m\geq 0} \big(X_t^{N,m}(i) + Z_t^{N,m}(i)\big) \geq K\Big\} \tag{4.8}$$

for every $K \in [0, \infty)$ and every $N \in \mathbb{N}$.

Lemma 4.5. *Assume 4.1. Then we have that*

$$\sup_{t\leq T} \mathbb{E}\Big[\sum_{i=1}^N \sum_{k=0}^\infty X_t^{N,k}(i) \Big| (X_0^{N,l})_{l\in\mathbb{N}_0} = (x^N \mathbb{1}_{l=0})_{l\in\mathbb{N}_0}\Big] \leq \Big(2\theta T + \sum_{i=1}^N x_i^N\Big) e^{L_\mu T} \tag{4.9}$$

for every configuration $x^N \in I^N$, every $T \in [0, \infty)$ and every $N \in \mathbb{N}$. The analogous assertion holds for the loop-free (N, μ_N, σ_N^2)-processes, $N \in \mathbb{N}$.

Proof. Fix $N \in \mathbb{N}$ and $T < \infty$. By Assumption 4.1 we have that $\mu_N(x) \leq L_\mu x + \frac{2\theta}{N}$ for all $x \in I$. According to Lemma 4.4, $\big\{(\tilde{X}_t^N(i))_{t\geq 0}\colon i \leq N\big\}$ defined through (4.6) is a solution of (3.1). Sum (3.1) over $i \leq N$, stop at time $\tilde{\tau}_K^N$ and take expectations to obtain that

$$\mathbb{E}\Big[\sum_{i=1}^N \tilde{X}_{t\wedge\tilde{\tau}_K^N}^N(i)\Big] \leq \sum_{i=1}^N x_i^N + \int_0^t L_\mu \mathbb{E}\Big[\sum_{i=1}^N \tilde{X}_{s\wedge\tilde{\tau}_K^N}^N(i)\Big] + 2\theta \, ds \tag{4.10}$$

for every $t \leq T$ and $K \in [0, \infty)$. Note that the right-hand side is finite. Now Gronwall's inequality implies that

$$\mathbb{E}\Big[\sum_{i=1}^N \tilde{X}_{t\wedge\tilde{\tau}_K^N}^N(i)\Big] \leq \Big(\sum_{i=1}^N x_i^N + 2\theta T\Big) e^{L_\mu T} \tag{4.11}$$

for all $t \leq T$ and $K \in [0, \infty)$. Letting $K \to \infty$, path continuity and Fatou's lemma yield that

$$\begin{aligned}\sup_{t\leq T} \mathbb{E}\Big[\sum_{i=1}^N \tilde{X}_t^N(i)\Big] &= \sup_{t\leq T} \mathbb{E}\Big[\sum_{i=1}^N \liminf_{K\to\infty} \tilde{X}_{t\wedge\tilde{\tau}_K^N}^N(i)\Big] \\ &\leq \sup_{t\leq T} \liminf_{K\to\infty} \mathbb{E}\Big[\sum_{i=1}^N \tilde{X}_{t\wedge\tilde{\tau}_K^N}^N(i)\Big] \leq \Big(\sum_{i=1}^N x_i^N + 2\theta T\Big) e^{L_\mu T}.\end{aligned} \tag{4.12}$$

This proves inequality (4.9). The inequality for the loop-free N-island process follows similarly. □

Lemma 4.6. *Assume 3.2 and 4.1. Then we have that*

$$\sum_{k\geq 0}\sup_{N\in\mathbb{N}}\sup_{t\leq T}\sum_{i=1}^{N}\Big(\mathbb{E}X_t^{N,k}(i)+\mathbb{E}Z_t^{N,k}(i)\Big)<\infty \tag{4.13}$$

for every $T\in[0,\infty)$.

Proof. We prove inequality (4.13) for the solution of (4.1). The estimate for the solution of (4.2) is analogous.

Recall $\tilde{\mu}_N(x)=\mu_N(x)-\mu_N(0)$ for all $x\in I$. Apply Itô's formula to $\sum_{i=1}^N X_t^{N,k}(i)$, take expectations, estimate $\tilde{\mu}_N(x)\leq L_\mu x$ for all $x\in I$ and take suprema to obtain that

$$\begin{aligned}\sup_{N\in\mathbb{N}}\sup_{t\leq T}\mathbb{E}\sum_{i=1}^{N}X_t^{N,k}(i)&\leq \mathbb{1}_{k=0}\sup_{N\in\mathbb{N}}\mathbb{E}\sum_{i=1}^{N}X_0^N(i)\\ &\quad+\int_0^T 2\theta\mathbb{1}_{k=0}+\sup_{N\in\mathbb{N}}\sup_{s\leq t}\mathbb{E}\sum_{i=1}^{N}\Big(X_s^{N,k-1}(i)+L_\mu X_s^{N,k}(i)\Big)\,dt\end{aligned} \tag{4.14}$$

for all $T\geq 0$ and all $k\in\mathbb{N}_0$. Note that the right-hand side is finite due to Lemma 4.5 and Assumption 3.2. Summing over $k\leq K\in\mathbb{N}$ and applying Gronwall's inequality implies that

$$\sum_{k=0}^{K}\sup_{N\in\mathbb{N}}\sup_{t\leq T}\mathbb{E}\sum_{i=1}^{N}X_t^{N,k}(i)\leq\Big(\sup_{M\in\mathbb{N}}\mathbb{E}\sum_{i=1}^{M}X_0^M(i)+2\theta T\Big)\cdot e^{(1+L_\mu)T} \tag{4.15}$$

for every $K\in\mathbb{N}$. Letting $K\to\infty$ proves (4.13). ☐

Lemma 4.7. *Assume 4.1. Then we have that*

$$\begin{aligned}&\mathbb{E}\bigg[\Big(\sup_{t\leq T}\sum_{i=1}^{N}\sum_{m\geq 0}X_t^{N,m}(i)\Big)^2\bigg|(X_0^{N,l})_{l\in\mathbb{N}_0}=(x^N\mathbb{1}_{l=0})_{l\in\mathbb{N}_0}\bigg]\\ &\leq 4\bigg[\Big(\sum_{i=1}^{N}x_i^N\Big)^2+2\theta T+1\bigg]\Big(1+T(4\theta+L_\sigma)e^{L_\mu T}\Big)e^{(2L_\mu+L_\sigma)T}\end{aligned} \tag{4.16}$$

for every configuration $x^N\in I^N$*, every* $T\in[0,\infty)$ *and every* $N\in\mathbb{N}$*. The analogous assertion holds for the loop-free* (N,μ_N,σ_N^2)*-process.*

Proof. Fix $N\in\mathbb{N}$, $T\in[0,\infty)$ and a configuration $x^N\in I^N$. According to Lemma 4.4, $(\sum_{m=0}^\infty X_t^{N,m})_{t\geq 0}$ is an N-island model. Recall from Assumption 4.1 that $\mu_N(x)\leq L_\mu x+\frac{2\theta}{N}=:\bar{\mu}_N(x)$ for all $x\in I$. Thus Lemma 3.3 of [23] implies that the (N,μ_N,σ_N^2)-process is dominated by the $(N,\bar{\mu}_N,\sigma_N^2)$-process $(\bar{X}_t^N)_{t\geq 0}$ starting in $\bar{X}_0^N=x^N$. Using Itô's formula and $\sigma_N^2(x)\leq L_\sigma(x+x^2)$ for all $x\in I$, we get that

$$\begin{aligned}&\mathbb{E}\bigg[\Big(\sum_{i=1}^{N}\bar{X}^N_{t\wedge\tilde{\tau}_K^N}(i)\Big)^2\bigg]\\ &=\Big(\sum_{i=1}^{N}x_i^N\Big)^2+\mathbb{E}\bigg[\int_0^{t\wedge\tilde{\tau}_K^N}2L_\mu\Big(\sum_{i=1}^{N}\bar{X}_s^N(i)\Big)^2+4\theta\sum_{i=1}^{N}\bar{X}_s^N(i)+\sum_{i=1}^{N}\sigma_N^2\big(\bar{X}_s^N(i)\big)\,ds\bigg]\\ &\leq\Big(\sum_{i=1}^{N}x_i^N\Big)^2+\int_0^t(2L_\mu+L_\sigma)\mathbb{E}\bigg[\Big(\sum_{i=1}^{N}\bar{X}^N_{s\wedge\tilde{\tau}_K^N}(i)\Big)^2\bigg]+(4\theta+L_\sigma)\Big(2\theta T+\sum_{i=1}^{N}x_i^N\Big)e^{L_\mu T}\,ds\\ &\leq\Big(\sum_{i=1}^{N}x_i^N\Big)^2\big(1+T(4\theta+L_\sigma)e^{L_\mu T}\big)+\int_0^t(2L_\mu+L_\sigma)\mathbb{E}\bigg[\Big(\sum_{i=1}^{N}\bar{X}^N_{s\wedge\tilde{\tau}_K^N}(i)\Big)^2\bigg]ds+T\big(4\theta+L_\sigma\big)\big(2\theta T+1\big)e^{L_\mu T}\end{aligned} \tag{4.17}$$

for every $t\leq T$ and every $K\in\mathbb{N}$. We used Lemma 4.5 for the last but one inequality and the estimate $a\leq 1+a^2$ for $a\in\mathbb{R}$ for the last inequality. Note that the right-hand side is finite. Applying Doob's L^2

submartingale inequality (e.g. Theorem II.70.2 in [39]) to the submartingale $\sum_{i=1}^N \bar{X}_t^N(i)$, using Fatou's lemma and applying Gronwall's inequality to (4.17), we conclude that

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\le T}\Big(\sum_{i=1}^N\sum_{m\ge 0}X_t^{N,m}(i)\Big)^2\Big]\le 4\sup_{t\le T}\mathbb{E}\Big[\Big(\sum_{i=1}^N\bar{X}_t^N(i)\Big)^2\Big]\le 4\sup_{t\le T}\mathbb{E}\Big[\liminf_{K\to\infty}\Big(\sum_{i=1}^N\bar{X}^N_{t\wedge\tilde{\tau}_K^N}(i)\Big)^2\Big]\\&\le 4\sup_{t\le T}\liminf_{K\to\infty}\mathbb{E}\Big[\Big(\sum_{i=1}^N\bar{X}^N_{t\wedge\tilde{\tau}_K^N}(i)\Big)^2\Big]\le 4\Big[\Big(\sum_{i=1}^N x_i^N\Big)^2+2\theta T+1\Big]\Big(1+T(4\theta+L_\sigma)e^{L_\mu T}\Big)e^{(2L_\mu+L_\sigma)T}.\end{aligned}\tag{4.18}$$

The proof in the case of the loop-free N-island model is analogous. □

Recall $\tilde{\tau}_K^N$ from (4.8). Next we show that stopping at the time $\tilde{\tau}_K^N$ has no impact within a finite time interval in the limit $K\to\infty$.

Lemma 4.8. *Assume 4.1 and 4.2. Then any solution of* (4.1) *satisfies that*

$$\limsup_{K\to\infty}\sup_{N\in\mathbb{N}}\mathbb{E}\Big[\sup_{t\le T}\sum_{i=1}^N\sum_{m\ge 0}X_t^{N,m}(i)\mathbb{1}_{\tilde{\tau}_K^N\le T}\Big]=0\tag{4.19}$$

for every $T\in[0,\infty)$. The analogous assertion holds for the loop-free (N,μ_N,σ_N^2)-process.

Proof. Rewriting $\{\tilde{\tau}_K^N\le T\}=\{\sup_{t\le T}\sum_{i=1}^N\sum_{m\ge0}(X_t^{N,m}(i)+Z_t^{N,m}(i))\ge K\}$, the assertion follows from the Markov inequality and from the second-moment estimate of Lemma 4.7. □

Lemma 4.9. *Assume 4.1. Then the N-island process $(X_t^N)_{t\ge0}$ solving the SDE* (3.1) *satisfies that*

$$\begin{aligned}&\mathbb{E}\left[\tfrac{1}{N}\sum_{i=1}^N\sup_{t\in[0,T]}\big(X_t^N(i)\big)^2\Big|X_0^N=x^N\right]\\&\le 2\left(\tfrac{1}{N}\sum_{i=1}^N\big(x_i^N\big)^2+\tfrac{24T\theta^2}{N^2}+8L_\sigma T\Big(\tfrac{2\theta T}{N}+\tfrac{1}{N}\sum_{i=1}^N x_i^N\Big)e^{L_\mu T}\right)\exp\big(40(1+T)(1+L_\mu+L_\sigma)^2T\big)\end{aligned}\tag{4.20}$$

for every configuration $x^N\in I^N$, every $T\in[0,\infty)$ and every $N\in\mathbb{N}$.

Proof. Fix $N\in\mathbb{N}$ and $x^N\in I^N$ throughout the proof. Assumption 4.1 implies that $\mu_N(x)\le L_\mu x+\frac{2\theta}{N}=:\bar{\mu}_N(x)$ for all $x\in I$. Thus Lemma 3.3 of [23] implies that the (N,μ_N,σ_N^2)-process $(X_t^N)_{t\ge0}$ is dominated by the $(N,\bar{\mu}_N,\sigma_N^2)$-process $(\bar{X}_t^N)_{t\ge0}$ starting in $\bar{X}_0^N=x^N$. Applying Doob's L^2 submartingale inequality (e.g. Theorem II.70.2 in [39]), Jensen's inequality and $\sigma_N^2(x)\le L_\sigma(x+x^2)$ for all $x\in I$, we get that

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\in[0,T]}\big(\bar{X}_t^N(i)-x_i^N\big)^2\right]\\&=\mathbb{E}\left[\sup_{t\in[0,T]}\Big(\int_0^t\tfrac{1}{N}\sum_{j=1}^N\bar{X}_s^N(j)-\bar{X}_s^N(i)+L_\mu\bar{X}_s^N(i)+\tfrac{2\theta}{N}ds+\int_0^t\sigma_N(\bar{X}_s^N(i))dB_s(i)\Big)^2\right]\\&\le 2\mathbb{E}\Big[\Big(\int_0^T\tfrac{1}{N}\sum_{j=1}^N\bar{X}_s^N(j)+(L_\mu+1)\bar{X}_s^N(i)+\tfrac{2\theta}{N}\,ds\Big)^2\Big]+2\mathbb{E}\left[\sup_{t\in[0,T]}\Big(\int_0^t\sigma_N\big(\bar{X}_s^N(i)\big)dB_s(i)\Big)^2\right]\\&\le 6T\int_0^T\mathbb{E}\Big[\Big(\tfrac{1}{N}\sum_{j=1}^N\bar{X}_s^N(j)\Big)^2+(L_\mu+1)^2\Big(\bar{X}_s^N(i)\Big)^2+\tfrac{4\theta^2}{N^2}\Big]ds+8\int_0^T\mathbb{E}\left[\sigma_N^2\big(\bar{X}_s^N(i)\big)\right]ds\\&\le\int_0^T\mathbb{E}\Big[6T\tfrac{1}{N}\sum_{j=1}^N\Big(\bar{X}_s^N(j)\Big)^2+\big(6T(L_\mu+1)^2+8L_\sigma\big)\Big(\bar{X}_s^N(i)\Big)^2+\tfrac{24T\theta^2}{N^2}+8L_\sigma\bar{X}_s^N(i)\Big]ds\end{aligned}$$

for all $i \in \{1,2,\ldots,N\}$ and, using Lemma 4.5,

$$\begin{aligned}
&\mathbb{E}\left[\tfrac{1}{N}\sum_{i=1}^{N}\sup_{t\in[0,T]}\left(\bar{X}_t^N(i)\right)^2\right] - 2\tfrac{1}{N}\sum_{i=1}^{N}\left(x_i^N\right)^2 \le 2\mathbb{E}\left[\tfrac{1}{N}\sum_{i=1}^{N}\sup_{t\in[0,T]}\left(\bar{X}_t^N(i)-x_i^N\right)^2\right] \\
&\le 40(1+T)(1+L_\mu+L_\sigma)^2\int_0^T \mathbb{E}\left[\tfrac{1}{N}\sum_{i=1}^{N}\sup_{r\in[0,s]}\left(\bar{X}_r^N(i)\right)^2\right]ds + \tfrac{48T\theta^2}{N^2} + 16L_\sigma T\tfrac{1}{N}\left(2\theta T+\sum_{i=1}^{N}x_i^N\right)e^{L_\mu T}
\end{aligned}$$

for all $T\in[0,\infty)$. The right-hand side is finite due to Lemma 4.7. Therefore, Gronwall's lemma implies that

$$\begin{aligned}
&\mathbb{E}\left[\tfrac{1}{N}\sum_{i=1}^{N}\sup_{t\in[0,T]}\left(X_t^N(i)\right)^2\right] \le \mathbb{E}\left[\tfrac{1}{N}\sum_{i=1}^{N}\sup_{t\in[0,T]}\left(\bar{X}_t^N(i)\right)^2\right] \\
&\le \left(2\tfrac{1}{N}\sum_{i=1}^{N}\left(x_i^N\right)^2 + \tfrac{48T\theta^2}{N^2} + 16L_\sigma T\tfrac{1}{N}\left(2\theta T+\sum_{i=1}^{N}x_i^N\right)e^{L_\mu T}\right)\exp\left(40(1+T)(1+L_\mu+L_\sigma)^2T\right)
\end{aligned} \tag{4.21}$$

for all $T\in[0,\infty)$ and this finishes the proof. □

Next we prove some preliminary results for the solution $(Y_{t,s}^{N,\zeta})_{t\ge 0}$ of (4.3).

Lemma 4.10. *Assume 4.1. Let $\zeta_N\colon[0,\infty)\to N\cdot I=\{N\cdot x\colon x\in I\}$ be a locally square Lebesgue integrable function for every $N\in\mathbb{N}$. Then we have that*

$$\mathbb{E}^x\Big[\big(\sup_{s\le t\le T}Y_{t,s}^{N,\zeta}\big)^2\Big] \le C_T\Big[x+x^2+\int_s^T\frac{\zeta_N(r)}{N}+\Big(\frac{\zeta_N(r)}{N}\Big)^2\,dr\Big] \tag{4.22}$$

for all $x\in I$, $0\le s\le T$, $N\in\mathbb{N}$ and some constant $C_T<\infty$ which does not depend on x, N or on ζ_N.

Proof. The proof is similar to the proof of Lemma 4.7, so we omit it. □

Lemma 4.11. *Assume 4.1. Let $\zeta_N\colon[0,\infty)\to N\cdot I$ be a locally square Lebesgue integrable function for every $N\in\mathbb{N}$. Furthermore let $\big(Y_{t,s}^{N,\zeta}(i)\big)_{t\ge s}$, $i\le N$, be independent solutions of* (4.3) *starting in $Y_{s,s}^{N,\zeta}(i)=0$, $i\le N$, for every $N\in\mathbb{N}$. Then we have that*

$$\mathbb{E}\Big[\sup_{s\le t\le T}\Big(\sum_{i=1}^{N}Y_{t,s}^{N,\zeta}(i)\Big)^2\Big] \le C_T\Big[\int_s^T\zeta_N(r)+\frac{\big(\zeta_N(r)\big)^2}{N}\,dr\Big] \tag{4.23}$$

for all $0\le s\le T$, $N\in\mathbb{N}$ and some constant $C_T<\infty$ which does not depend on N or ζ_N.

Proof. The proof is similar to the proof of Lemma 4.7, so we omit it. □

Lemma 4.12. *Assume 4.1 and fix $T\in[0,\infty)$. Let $(Y_{t,s}^{N,\zeta})_{t\ge s}$ and $(Y_{t,s}^{N,\tilde{\zeta}})_{t\ge s}$ be two solutions of* (4.3) *with respect to the same Brownian motion such that $Y_{s,s}^{N,\zeta}=x$ and $Y_{s,s}^{N,\tilde{\zeta}}=y$. If $\zeta_N,\tilde{\zeta}_N\colon[0,T]\to N\cdot I$ are square Lebesgue integrable, then we have that*

$$\mathbb{E}\big[(Y_{t,s}^{N,\zeta}-Y_{t,s}^{N,\tilde{\zeta}})^+\big] \le e^{L_\mu(t-s)}\left(\frac{1}{N}\int_s^t(\zeta_N(r)-\tilde{\zeta}_N(r))^+\,dr+(x-y)^+\right) \tag{4.24}$$

for all $N\in\mathbb{N}$, for all $0\le s\le t\le T$ and all $x,y\in I$ where $z^+=\max(z,0)$ for all $z\in\mathbb{R}$.

Proof. As in Theorem 1 of Yamada and Watanabe (1971) [47], an approximation of $x\to x^+$ with C^2-functions (see also the proof of Lemma 4.25 for this approximation) results in

$$d\big(Y_{t,s}^{N,\zeta}-Y_{t,s}^{N,\tilde{\zeta}}\big)^+ = \mathbb{1}_{Y_{t,s}^{N,\zeta}-Y_{t,s}^{N,\tilde{\zeta}}\ge 0}\,d\big(Y_{t,s}^{N,\zeta}-Y_{t,s}^{N,\tilde{\zeta}}\big). \tag{4.25}$$

Taking expectations, the upward Lipschitz continuity of $\tilde{\mu}_N$ implies that

$$\begin{aligned}
&\mathbb{E}\big[\big(Y_{t,s}^{N,\zeta}-Y_{t,s}^{N,\tilde{\zeta}}\big)^+\big] \\
&\le (x-y)^+ + \frac{1}{N}\int_s^t\big(\zeta_N(r)-\tilde{\zeta}_N(r)\big)^+\,dr + L_\mu\int_s^t\mathbb{E}\big[\big(Y_{r,s}^{N,\zeta}-Y_{r,s}^{N,\tilde{\zeta}}\big)^+\big]\,dr
\end{aligned} \tag{4.26}$$

for all $t \geq s$. The right-hand side is finite due to Lemma 4.10. Therefore, Gronwall's inequality implies (4.24). $\square$

Now we study the solution $(Y_{t,s}^{N,c})_{t\geq s}$ of (4.3) in which $\zeta_N \equiv c$ is a constant $c \in [0,\infty)$ for every $N \in \mathbb{N}$ with $N \geq c/|I|$. Let a point $\alpha \in (0,|I|)$ be fixed. Recall the scale function S from (2.2). Define the scale function S_N of $(Y_{t,s}^{N,c})_{t\geq s}$ through

$$s_N(z) := \exp\Big(-\int_\alpha^z \frac{c/N}{\sigma_N^2(x)/2}\,dx - \int_0^z \frac{-x+\tilde{\mu}_N(x)}{\sigma_N^2(x)/2}\,dx\Big),\ S_N(y) := \int_0^y s_N(z)\,dz \tag{4.27}$$

for all $y,z \in I$. We point out that using two reference points (α and 0) in the definition of S_N is unusual but this definition differs from the standard definition with a single reference point α only by a constant factor. The next two lemmas involve the speed measures

$$m(dy) := \frac{2}{\sigma^2(y)s(y)}\,dy, \quad m_N(dy) := \frac{2}{\sigma_N^2(y)s_N(y)}\,dy \tag{4.28}$$

as measures on $(0,|I|)$ for every $N \in \mathbb{N}$. Note that under Assumption 4.1, $m_N(\cdot)$ converges as a measure on $(0,|I|)$ vaguely to $m(\cdot)$ as $N \to \infty$.

Lemma 4.13. *Assume 4.1 and 2.2. If $\zeta_N\colon [0,\infty) \to N{\cdot}I$ is locally square Lebesgue integrable for every $N \in \mathbb{N}$, then*

$$\lim_{\delta\to 0}\sup_{N\in\mathbb{N}}\int_0^\delta \Big|\mathbb{E}^x\big[f_N(Y_{t,s}^{N,\zeta})\big] - \mathbb{E}^0\big[f_N(Y_{t,s}^{N,\zeta})\big]\Big|\,m_N(dx) = 0 \tag{4.29}$$

for all $s \leq t < \infty$ and all functions $f_N\colon I \to \mathbb{R}$, $N \in \mathbb{N}$, with $\sup_{x\neq y\in I}\sup_{N\in\mathbb{N}}\frac{|f_N(x)-f_N(y)|}{|x-y|} < \infty$.

Proof. Fix $s \leq t < \infty$ and define $C := \sup_{x\neq y\in I}\sup_{N\in\mathbb{N}}\frac{|f_N(x)-f_N(y)|}{|x-y|} \in [0,\infty)$. Let $(Y_{u,s}^{N,\zeta,x})_{u\geq s}$ and $(Y_{u,s}^{N,\zeta,0})_{u\geq s}$ be solutions of (4.3) with respect to the same Brownian motion satisfying $Y_{s,s}^{N,\zeta,x} = x$ and $Y_{s,s}^{N,\zeta,0} = 0$. According to Assumptions 4.1 and 2.2, there exist real numbers $\varepsilon,\delta_0 \in (0,|I|\wedge\alpha)$ such that $\sigma_N^2(y) \geq \varepsilon\sigma^2(y)$ for all $y \in [0,\delta_0]$ and all $N \in \mathbb{N}$ and such that $\int_0^{\delta_0} y/\sigma^2(y)\,dy < \infty$. The first-moment estimate of Lemma 4.12 provides us with the inequality

$$\begin{aligned}
&\sup_{N\in\mathbb{N}}\int_0^\delta \Big|\mathbb{E}^x\big[f_N(Y_{t,s}^{N,\zeta})\big] - \mathbb{E}^0\big[f_N(Y_{t,s}^{N,\zeta})\big]\Big|\,m_N(dx)\\
&\leq \sup_{N\in\mathbb{N}}\int_0^\delta \mathbb{E}\Big[C\big|Y_{t,s}^{N,\zeta,x} - Y_{t,s}^{N,\zeta,0}\big|\Big]\,m_N(dx) \leq \sup_{N\in\mathbb{N}}\int_0^\delta C\cdot e^{L_\mu t}\cdot x\,m_N(dx)\\
&\leq Ce^{L_\mu t}\int_0^\delta x\frac{2}{\varepsilon\sigma^2(x)}\,dx\cdot\exp\Big(\int_0^{\delta_0}\frac{2L_\mu z}{\varepsilon\sigma^2(z)}\,dz\Big)
\end{aligned} \tag{4.30}$$

for every $\delta \in (0,\delta_0)$. The right-hand side of (4.30) converges to zero as $\delta \to 0$ by the dominated convergence theorem. $\square$

Lemma 4.14. *Assume 4.1. If $c > 0$, then*

$$\frac{c}{N}m_N((0,\delta)) \longrightarrow 1 \qquad \text{as } N \to \infty \tag{4.31}$$

for every $\delta \in (0,|I|\wedge\alpha)$ such that $\int_0^\delta y/\sigma^2(y)\,dy < \infty$ and $\inf_{y\in(0,\delta)}\inf_{N\in\mathbb{N}}\sigma_N^2(y)/\sigma^2(y) > 0$.

Proof. Fix $\delta \in (0,|I|\wedge\alpha)$ such that $\int_0^\delta y/\sigma^2(y)\,dy < \infty$ and $\inf_{y\in(0,\delta)}\inf_{N\in\mathbb{N}}\sigma_N^2(y)/\sigma^2(y) > 0$ and fix $c \in (0,\infty)$. Integration by parts yields that

$$\begin{aligned}
\frac{c}{N}m_N\big((0,\delta)\big) &= \int_0^\delta \frac{c/N}{\sigma_N^2(y)/2}\exp\Big(\int_\alpha^y \frac{c/N}{\sigma_N^2(x)/2}\,dx\Big)\exp\Big(\int_0^y \frac{-x+\tilde{\mu}_N(x)}{\sigma_N^2(x)/2}\,dx\Big)\,dy\\
&= \Big[\exp\Big(\int_\alpha^y \frac{c/N}{\sigma_N^2(x)/2}\,dx + \int_0^y \frac{-x+\tilde{\mu}_N(x)}{\sigma_N^2(x)/2}\,dx\Big)\Big]_0^\delta\\
&\quad - \int_0^\delta \frac{-y+\tilde{\mu}_N(y)}{\sigma_N^2(y)/2}\exp\Big(\int_\alpha^y \frac{c/N}{\sigma_N^2(x)/2}\,dx + \int_0^y \frac{-x+\tilde{\mu}_N(x)}{\sigma_N^2(x)/2}\,dx\Big)\,dy
\end{aligned} \tag{4.32}$$

for every $N \in \mathbb{N}$. As σ_N^2 is Lipschitz continuous in $[0,\alpha]$ and $\sigma_N(0)=0$, $\int_\alpha^0 \frac{1}{\sigma_N^2(x)/2}dx = -\infty$ for every $N \in \mathbb{N}$. Letting $N \to \infty$ in (4.32) and applying the dominated convergence theorem shows that

$$\begin{aligned} &\lim_{N\to\infty} \frac{c}{N} m_N\big((0,\delta)\big) \\ &= \exp\Big(\int_0^\delta \frac{-x+\mu(x)}{\sigma^2(x)/2}\,dx\Big) - \int_0^\delta \frac{-y+\mu(y)}{\sigma^2(y)/2} \exp\Big(\int_0^y \frac{-x+\mu(x)}{\sigma^2(x)/2}\,dx\Big)\,dy \\ &= \exp\Big(\int_0^\delta \frac{-x+\mu(x)}{\sigma^2(x)/2}\,dx\Big) - \Big[\exp\Big(\int_0^y \frac{-x+\mu(x)}{\sigma^2(x)/2}\,dx\Big)\Big]_0^\delta \end{aligned} \tag{4.33}$$

which is equal to one. □

We recall the following lemma from [21], see Lemma 9.8 there.

Lemma 4.15. *Assume 2.1 and 2.2. Let Q be the excursion measure defined through* (2.4)*. Then*

$$\int \Big(\int_0^\infty \chi_t\,dt\Big)\,Q(d\chi) = \int_0^{|I|} \frac{y}{\sigma^2(y)/2} \exp\Big(\int_0^y \frac{-x+\mu(x)}{\sigma^2(x)/2}\,dx\Big)\,dy < \infty. \tag{4.34}$$

The last result of this subsection is a variation of the second moment estimate of Lemma 4.7. Define a stopping time $\tau_K^N \in [0,\infty]$ through

$$\tau_K^N := \inf\Big(\Big\{t \ge 0\colon\ \sum_{i=1}^N X_t^N(i) \ge K\Big\} \cup \{\infty\}\Big) \tag{4.35}$$

for every $K \in [0,\infty)$ and every $N \in \mathbb{N}$.

Lemma 4.16. *Assume 4.1, 2.2 and 4.2. Then we have that*

$$\sup_{N\in\mathbb{N}} \sum_{i=1}^N \mathbb{E}\Big[\Big(\sup_{t\in[0,T]} X_{t\wedge\tau_K^N}^N(i)\Big)^2\Big] < \infty \tag{4.36}$$

for all $T \in [0,\infty)$ and all $K \in \mathbb{N}$.

Proof. Fix $T \in [0,\infty)$ and $K \in \mathbb{N}$. Lemma 3.3 in [23] shows that, on the event $\{\tau_K^N \ge t\}$, $X_t^N(i)$ is bounded from above by $Y_{t,0}^{N,K+N\mu_N(0)}$ for all $t \in [0,T]$ almost surely for every $N \in \mathbb{N}$. By Assumption 4.1 we have that $N\mu_N(0) \le 2\theta$ for all $N \in \mathbb{N}$. Together with the second-moment estimate of Lemma 4.10, this implies that

$$\begin{aligned} \sum_{i=1}^N \mathbb{E}\Big[\Big(\sup_{t\in[0,T]} X_{t\wedge\tau_K^N}^N(i)\Big)^2\Big] &\le \sum_{i=1}^N \mathbb{E}\Big[\mathbb{E}^{X_0^N(i)}\Big[\Big(\sup_{t\in[0,T]} Y_{t,0}^{N,K+2\theta}\Big)^2\Big]\Big] \\ &\le C_T \sum_{i=1}^N \Big(\mathbb{E}\big[X_0^N(i)\big] + \mathbb{E}\Big[\Big(X_0^N(i)\Big)^2\Big] + T\frac{K+2\theta}{N} + T\frac{(K+2\theta)^2}{N^2}\Big) \\ &\le C_T\Big(\sup_{M\in\mathbb{N}} \mathbb{E}\Big[\sum_{i=1}^M X_0^M(i)\Big] + \sup_{M\in\mathbb{N}} \mathbb{E}\Big[\Big|\sum_{i=1}^M X_0^M(i)\Big|^2\Big] + T(K+2\theta) + T(K+2\theta)^2\Big) \end{aligned} \tag{4.37}$$

for every $N \in \mathbb{N}$ and some constant $C_T < \infty$. The right-hand side is finite due to Assumption 4.2. □

4.3 Poisson limit of independent diffusions with vanishing immigration

In this subsection, we prove (4.4) which is the central step in the proof of Theorem 3.3. Our proof is based on reversing time in the stationary process. For the time reversal, we consider the following stationary situation. Excursions from zero of the solution process $(Y_t)_{t\ge0}$ of the SDE (1.5) start at times given by the points of an homogeneous Poisson point process on $\mathbb{R}$ with rate 1. This process of immigrating excursions is invariant for the dynamics of $(Y_t)_{t\ge0}$ restricted to non-extinction, see (4.38). Now the time reversal of an excursion is again governed by the excursion measure, see Lemma 4.17. As a consequence, reversing time in the process of immigrating excursions does not change the distribution.

Let us retell this argument more formally. Consider a Poisson point process Π on $(-\infty,\infty)\times U$ with intensity measure $ds\otimes Q$. Then $\sum_{(s,\eta)\in\Pi}\delta_{(\eta_{t-s})_{t\geq 0}}$ is the process of immigrating excursions. Note that at a fixed time t, $\sum_{(s,\eta)\in\Pi}\delta_{\eta_{t-s}}$ is a Poisson point process on $(0,\infty)$ with intensity measure

$$\int_{-\infty}^{\infty} Q(\eta_{t-s}\in dy)\,ds = \int_{-\infty}^{0} Q(\eta_{-s}\in dy)\,ds = m(dy) \tag{4.38}$$

where m is the speed measure defined in (4.28). Here we used that $\eta_t=0$ for $t\in(-\infty,0]$ for all $\eta\in U$. The relation (4.38) between the speed measure and the excursion measure has been established in Lemma 9.8 of [21] by exploiting a well-known explicit formula for $\mathbb{E}^y\int_0^\infty f(Y_s)\,ds$. It is also well-known (e.g. (15.5.34) in [27]) that the speed measure m is an invariant measure for the sub-Markov semigroup $\mathbb{E}^{\cdot} f(Y_t)\mathbb{1}_{Y_t>0}$. This can also be seen from (4.38) by noting that $Q(\eta_s\in dy)$ is an entrance measure for this sub-Markov semigroup. Thus the process of immigrating excursions is indeed invariant for the dynamics of $(Y_t)_{t\geq 0}$ restricted to non-extinction.

Now we show that reversing time in the process of immigrating excursions does not change the distribution of the process. The process $(Y_t)_{t\geq 0}$ restricted to non-extinction is time-reversible when started in the invariant measure m, that is,

$$\int_I \mathbb{E}^x F\big((Y_t)_{t\leq T}\big)\mathbb{1}_{Y_T>0} m(dx) = \int_I \mathbb{E}^x F\big((Y_{T-t})_{t\leq T}\big)\mathbb{1}_{Y_T>0} m(dx) \tag{4.39}$$

for every $T\in[0,\infty)$ and every non-negative measurable function on $\mathrm{C}\big([0,T]\big)$, see Section 13 of Chapter 15 in [27]. The next lemma shows that if the speed measure m is replaced by the left-hand side of (4.38), then (4.39) can be extended to allow for extinction. First we state the Markov property of the excursion measure. Definition (2.4) of Q as rescaled law of $(Y_t)_{t\geq 0}$ together with the Markov property of $(Y_t)_{t\geq 0}$ implies that

$$\int F\Big(\big(\eta_t\big)_{0\leq t\leq T}\Big)\tilde F\Big(\big(\eta_{T+t}\big)_{t\geq 0}\Big)Q(d\eta) = \int F\Big(\big(\eta_t\big)_{0\leq t\leq T}\Big)\mathbb{E}^{\eta_T}\tilde F\Big(\big(Y_t\big)_{t\geq 0}\Big)Q(d\eta) \tag{4.40}$$

for all measurable functions $F,\tilde F\colon \mathrm{C}\big([0,\infty),[0,\infty)\big)\to[0,\infty)$ satisfying $F(\underline{0})=0=\tilde F(\underline{0})$ and every $T\in[0,\infty)$. Here and below, $\underline{0}$ denotes the function which is $\equiv 0$.

Lemma 4.17. *Assume 2.1 and 2.2. Then*

$$\int\int_{-\infty}^{\infty} F\big((\eta_{t-s})_{t\in\mathbb{R}}\big)ds\,Q(d\eta) = \int\int_{-\infty}^{\infty} F\big((\eta_{T-t-s})_{t\in\mathbb{R}}\big)ds\,Q(d\eta) \tag{4.41}$$

for all $T\in\mathbb{R}$ and all measurable functions $F\colon \mathrm{C}\big([0,\infty)\big)\to[0,\infty)$.

Proof. It suffices (see e.g. Theorem 14.12 in [28]) to establish (4.41) for $F_n(\eta):=\prod_{i=1}^n f_i(\eta_{t_i})$ where $t_1<\ldots<t_n\in\mathbb{R}$ and $f_1,\ldots,f_n\in \mathrm{C}_b\big([0,\infty),[0,\infty)\big)$. If $F_n(\underline{0})>0$, then both sides of (4.41) are infinite. For the rest of the proof, we assume $F_n(\underline{0})=0$, that is, $f_i(0)=0$ for at least one $i\in\{1,\ldots,n\}$. We may even assume $f_i\in \mathrm{C}_c\big((0,\infty),[0,\infty)\big)$ for at least one $i\in\{1,\ldots,n\}$. Otherwise approximate f_i monotonically from below with test functions which have compact support. In addition, we may without loss of generality assume $t_1=0=T$. Otherwise use a time translation.

If F_n vanishes on $\{\eta\colon \eta_0=0 \text{ or } \eta_{t_n}=0\}$, then (4.41) is essentially (4.39). To see this, consider

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int \mathbb{1}_{\eta_{-s}>0}F_n\big((\eta_{t-s})_{t\in\mathbb{R}}\big)\mathbb{1}_{\eta_{t_n-s}>0}\,Q(d\eta)ds\\
\underset{(4.40)}{=}&\int_{-\infty}^{\infty}\int \mathbb{1}_{\eta_{-s}>0}\mathbb{E}^{\eta_{-s}}\Big[F_n\big((Y_t)_{t\in[0,t_n]}\big)\mathbb{1}_{Y_{t_n}>0}\Big]\,Q(d\eta)ds\\
\underset{(4.38)}{=}&\int_I \mathbb{E}^x\Big[F_n\big((Y_t)_{t\in[0,t_n]}\big)\mathbb{1}_{Y_{t_n}>0}\Big]m(dx).
\end{aligned} \tag{4.42}$$

Applying (4.39) with $T = t_n$, reversing the calculation in (4.42) and substituting $s - t_n \mapsto s$ shows that

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int \mathbb{1}_{\eta_{-s}>0} F_n\big((\eta_{t-s})_{t\in[0,t_n]}\big)\mathbb{1}_{\eta_{t_n-s}>0}\, Q(d\eta)ds\\
&= \int_{-\infty}^{\infty}\int \mathbb{1}_{\eta_{-s}>0} F_n\big((\eta_{t_n-t-s})_{t\in[0,t_n]}\big)\mathbb{1}_{\eta_{t_n-s}>0}\, Q(d\eta)ds.\\
&= \int_{-\infty}^{\infty}\int \mathbb{1}_{\eta_{-t_n-s}>0} F_n\big((\eta_{-t-s})_{t\in[0,t_n]}\big)\mathbb{1}_{\eta_{-s}>0}\, Q(d\eta)ds.
\end{aligned} \tag{4.43}$$

We prove (4.41) with F replaced by F_n by induction on $n \in \mathbb{N}$. The base case $n = 1$ follows from a time translation. The induction step $n-1 \to n$ follows directly from (4.43) if $f_1(0) = 0 = f_n(0)$. If $f_n(0) > 0$, then

$$\begin{aligned}
&\int\int_{-\infty}^{\infty} F_n\big((\eta_{t-s})_{t\in[0,t_n]}\big)\mathbb{1}_{\eta_{-s}>0}\mathbb{1}_{\eta_{t_n-s}=0}\, ds\, Q(d\eta)\\
&= f_n(0)\Big(\int\int_{-\infty}^{\infty} F_{n-1}\big((\eta_{t-s})_{t\in[0,t_n]}\big)\mathbb{1}_{\eta_{-s}>0}\big(1-\mathbb{1}_{\eta_{t_n-s}>0}\big)\, ds\, Q(d\eta)\Big)\\
&= f_n(0)\Big(\int\int_{-\infty}^{\infty} F_{n-1}\big((\eta_{-t-s})_{t\in[0,t_n]}\big)\mathbb{1}_{\eta_{-s}>0}\big(1-\mathbb{1}_{\eta_{-t_n-s}>0}\big)\, ds\, Q(d\eta)\Big)\\
&= \int\int_{-\infty}^{\infty} F_n\big((\eta_{-t-s})_{t\in[0,t_n]}\big)\mathbb{1}_{\eta_{-s}>0}\mathbb{1}_{\eta_{-t_n-s}=0}\, ds\, Q(d\eta).
\end{aligned} \tag{4.44}$$

For the second step we used linearity, applied the induction hypothesis and equation (4.43) and again used linearity. Adding (4.43) and (4.44) proves the induction step in case of $f_1(0) = 0$. The remaining case $f_1(0) > 0$ follows from a similar calculation as in (4.44). This completes the proof of Lemma 4.17. □

Lemma 4.18. *Assume 2.1 and 2.2. Then the solution process $(Y_t)_{t\geq 0}$ of the SDE (1.5) satisfies that*

$$\int_I \mathbb{E}^x F\big((Y_t)_{t\in[0,T]}\big)m(dx) = \int\int_{-\infty}^{T} F\big((\eta_{T-t-s})_{t\in[0,T]}\big)ds\, Q(d\eta) \tag{4.45}$$

for all measurable functions $F\colon \mathrm{C}\big([0,T],I\big) \to [0,\infty)$ satisfying $F(\underline{0}) = 0$ and all $T \in [0,\infty)$.

Proof. Express the speed measure in terms of the excursion measure as in (4.38)

$$\begin{aligned}
\int_I \mathbb{E}^x F\big((Y_t)_{t\in[0,T]}\big)m(dx) &= \int_{-\infty}^{0}\int \mathbb{1}_{\eta_{-s}>0}\mathbb{E}^{\eta_{-s}} F\big((Y_t)_{t\in[0,T]}\big)Q(d\eta)\, ds\\
&= \int\int_{-\infty}^{\infty} \mathbb{1}_{\eta_{-s}>0} F\big((\eta_{t-s})_{t\in[0,T]}\big)ds\, Q(d\eta)\\
&= \int\int_{-\infty}^{T} F\big((\eta_{T-t-s})_{t\in[0,T]}\big)Q(d\eta)\, ds.
\end{aligned} \tag{4.46}$$

The last two steps are the Markov property (4.40) and Lemma 4.17, respectively. □

With Lemma 4.18 in hand, we now reverse time to prove a first version of the Poisson approximation (4.4).

Lemma 4.19. *Assume 4.1 and 2.2. Let $c, s \in [0,\infty)$ be real numbers. Let $(Y_{t,s}^{N,c})_{t\geq s}$ be the solution of (4.3) with $\zeta_N(\cdot) \equiv c$ and $Y_{s,s}^{N,c} = 0$ for every $N \in \mathbb{N}$ with $N \geq c/|I|$. Then*

$$\lim_{N\to\infty} N\mathbb{E}^0\big[f_N(Y_{t,s}^{N,c})\big] = c\int_s^t\int f_0(\chi_{t-r})Q(d\chi)\, dr \tag{4.47}$$

for all $t \in [s,\infty)$ and all functions $f_N\colon I \to \mathbb{R}$, $N \in \mathbb{N}_0$, with $\sup_{x\neq y\in I}\sup_{N\in\mathbb{N}} \frac{|f_N(x)-f_N(y)|}{|x-y|} < \infty$, with $\lim_{N\to\infty} N|f_N(0)| = 0$ and with $\lim_{N\to\infty} f_N(y) = f_0(y)$ for all $y \in I$.

Proof. If $c = 0$, then both sides of (4.47) are equal to zero. So for the rest of the proof we assume $c > 0$. Let $\varepsilon, \delta_0 \in (0, |I| \wedge \alpha)$ be such that $\sigma_N^2(y) \geq \varepsilon\sigma^2(y)$ for all $y \in [0, \delta_0]$ and such that $\int_0^{\delta_0} y/\sigma^2(y)\,dy < \infty$. Lemma 4.14 provides us with $N/m_N\big((0,\delta)\big) \to c$ as $N \to \infty$ for all $\delta \in (0, \delta_0)$. Thus

$$\begin{aligned}\lim_{N\to\infty} N\mathbb{E}^0\big[f_N(Y_{t,s}^{N,c})\big] &= \lim_{N\to\infty}\frac{N}{m_N\big((0,\delta)\big)}\lim_{N\to\infty}\int_0^\delta \mathbb{E}^0\big[f_N(Y_{t,s}^{N,c})\big]m_N(dx)\\ &=: c\lim_{N\to\infty}\int_0^\delta \mathbb{E}^x\big[f_N(Y_{t,s}^{N,c})\big]m_N(dx) + C(\delta)\end{aligned}\tag{4.48}$$

for all $\delta \in (0, \delta_0)$ and all $t \in [s, \infty)$ where $C\colon [0, \delta_0] \to \mathbb{R}$ is a suitable function. The term $C(\delta)$ converges to zero as $\delta \to 0$ according to Lemma 4.13.

The speed measure m_N is an invariant (non-probability) measure for $(Y_{t,s}^{N,c})_{t\geq 0}$, see e.g. (15.5.34) in [27]. Thus we may reverse time. As we let $N \to \infty$, we will exploit that $(Y_{t,s}^{N,c})_{t\geq 0}$ converges weakly to the solution process $(Y_t)_{t\geq s}$ of the SDE (1.5). In addition, $m_N(dy)$ converges vaguely to $m(dy)$ as $N \to \infty$ due to the dominated convergence theorem and Assumptions 4.1 and 2.2 as the densities converge. These observations imply that

$$\begin{aligned}&\lim_{N\to\infty}\int_0^\delta \mathbb{E}^x\big[f_N(Y_{t,s}^{N,c})\big]m_N(dx) = \lim_{N\to\infty}\int_I f_N(y)\mathbb{P}^y\big[Y_{t,s}^{N,c} \leq \delta\big]m_N(dy)\\ &= \int_I f_0(y)\mathbb{P}^y\big[Y_{t-s} \leq \delta\big]m(dy) = \int\int_{-\infty}^t f_0\big(\chi_{t-r}\big)\mathbb{1}_{\chi_{s-r}\leq\delta}dr\,Q(d\chi) \xrightarrow{\delta\to 0} \int\int_s^t f_0\big(\chi_{t-r}\big)dr\,Q(d\chi).\end{aligned}\tag{4.49}$$

The last but one step is Lemma 4.18 and the last step follows from the dominated convergence theorem together with Lemma 4.15. Putting (4.48) and (4.49) together completes the proof of Lemma 4.19. □

Next we use induction to generalize Lemma 4.19 to test functions which depend on finitely many time coordinates. For this let $\mathcal{E}_{s,T}$ be the following set of bounded functions on $\mathrm{C}\big([s,T], I\big)$ for $0 \leq s \leq T < \infty$ which depend on finitely many coordinates and which are globally Lipschitz continuous in every coordinate

$$\begin{aligned}\mathcal{E}_{s,T} := \Big\{&\mathrm{C}\big([s,T], I\big) \ni \eta \mapsto \prod_{i=1}^n f_i(\eta_{t_i}) \in \mathbb{R}\colon n \in \mathbb{N}, s \leq t_1 < \ldots < t_n \leq T,\\ &f_1, ..., f_n \in \mathrm{C}\big(I, \mathbb{R}\big) \text{ are bounded and globally Lipschitz continuous}\Big\}\end{aligned}\tag{4.50}$$

for every $T \in [s, \infty)$ and every $s \in [0, \infty)$. Due to the Lipschitz continuity and boundedness of $f_1, \ldots, f_n$, there exists a constant $L_F \in (0, \infty)$ such that

$$|F(\eta) - F(\bar\eta)| \leq L_F\sum_{j=1}^n |\eta_{t_j} - \bar\eta_{t_j}| \qquad \forall\, \eta, \bar\eta \in \mathrm{C}\big([s,T], I\big)\tag{4.51}$$

for all $F \in \mathcal{E}_{s,T}$ and all $0 \leq s \leq T < \infty$. Note that the set $\mathcal{E}_{s,T}$ is closed under multiplication and separates points in $\mathrm{C}([s,T], I)$ for all $0 \leq s \leq T < \infty$. Thus the linear span of $\mathcal{E}_{s,T}$ is an algebra which separates points in $\mathrm{C}([s,T], I)$ for all $0 \leq s \leq T < \infty$. According to Theorem 3.4.5 in [16] the linear span of $\mathcal{E}_{s,T}$ is distribution determining for measures on $\mathrm{C}([s,T], I)$ and so is $\mathcal{E}_{s,T}$ for all $0 \leq s \leq T < \infty$.

Lemma 4.20. *Assume 4.1 and 2.2. Let $0 \leq s \leq T < \infty$. Suppose that $\zeta\colon [s,\infty) \to [0,\infty)$ and that $\hat\zeta_N\colon [s,\infty) \to N{\cdot}I$ are square Lebesgue integrable and that $\int_s^T |\zeta(r) - \hat\zeta_N(r)|dr \to 0$ as $N \to \infty$. Let $(Y_{t,s}^{N,\hat\zeta})_{t\in[s,\infty)}$ satisfy* (4.3)*. Then*

$$\lim_{N\to\infty} N\mathbb{E}^0 F\big((Y_{t,s}^{N,\hat\zeta})_{t\in[s,T]}\big) = \int_s^\infty \zeta(r)\int F\big((\chi_{t-r})_{t\in[s,T]}\big)Q(d\chi)\,dr \in \mathbb{R}\tag{4.52}$$

for all functions $F \in \mathcal{E}_{s,T}$.

Proof. Let $L_F \in (0,\infty)$ be such that F satisfies (4.51) and let F be bounded by $C_F \in (0,\infty)$. Moreover, let $(Y_t)_{t\in[s,\infty)}$ be the pathwise unique stochastic process such that $(Y_t)_{t\in[s,\infty)}$ and $(Y_{t,s}^{N,\hat{\zeta}})_{t\in[s,\infty)}$ are solutions of (4.3) with respect to the same Brownian motion. Lemma 4.12 implies that

$$\begin{aligned}
&\Big|\lim_{N\to\infty} N\mathbb{E}^0 F\Big(\big(Y_{t,s}^{N,\hat{\zeta}}\big)_{t\in[s,T]}\Big) - \lim_{N\to\infty} N\mathbb{E}^0 F\Big(\big(Y_{t,s}^{N,\zeta}\big)_{t\in[s,T]}\Big)\Big| \\
&\leq \lim_{N\to\infty} N L_F \sum_{j=1}^{n} \mathbb{E}^0\Big|Y_{t_j,s}^{N,\hat{\zeta}} - Y_{t_j,s}^{N,\zeta}\Big| \\
&\leq L_T \cdot n e^{L_\mu T} \lim_{N\to\infty} \int_0^T \big|\hat{\zeta}_N(r) - \zeta(r)\big| dr = 0.
\end{aligned} \tag{4.53}$$

Therefore, it suffices to prove (4.52) with $\hat{\zeta}_N$ replaced by ζ. A similar argument shows that we may assume ζ to be bounded; otherwise replace ζ by $\min(\zeta, M)$ and let $M \to \infty$.

We begin with the case of ζ being a simple function. W.l.o.g. we consider $\zeta(\cdot) = \sum_{i=1}^n c_i \mathbb{1}_{[t_{i-1},t_i)}(\cdot)$ where $c_1, ..., c_n \geq 0$ and $t_0 = s$ as we may let F depend trivially on further time points. The proof of (4.52) is by induction on n. The case $n = 1$ has been settled in Lemma 4.19. For the induction step we split up the left-hand side of (4.52) into two terms according to whether the process at time t_1 is essentially zero or not. In order to formalize the notion "essentially zero", choose a function $\phi_\delta \in \mathrm{C}^2\big(I, [0,1]\big)$ such that $\phi_\delta(x) = 1$ for $x \geq 2\delta$ and $\phi_\delta(x) = 0$ for $x \leq \delta$ for every $\delta \in (0, |I|)$. Furthermore define $\bar{F}_2(\eta) := \prod_{i=2}^n f_i(\eta_{t_i})$ for all $\eta \in \mathrm{C}([t_1, T], I)$.

First we consider the case that the process is away from 0 at time t_1. The following equation (4.57) shows that we may discard immigration after time t_1. For this, note that the moment estimate of Lemma 4.12 implies that

$$\begin{aligned}
&\bigg|\Big|\mathbb{E}^y \bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big) - \mathbb{E}^y \bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big)\Big| - \Big|\mathbb{E}^z \bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big) - \mathbb{E}^z \bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big)\Big|\bigg| \\
&\leq \Big|\mathbb{E}^y \bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big) - \mathbb{E}^z \bar{F}_2\Big(\big(Y_{t-t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big)\Big| + \Big|\mathbb{E}^y \bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big) - \mathbb{E}^z \bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big)\Big| \\
&\leq 2L_F \sum_{j=2}^{n} e^{L_\mu(t_j - s)} |y - z|
\end{aligned} \tag{4.54}$$

for all $y, z \in I$ and all $N \in \mathbb{N}$. Consequently, the sequence of functions

$$I \ni y \mapsto \phi_\delta(y) f_1(y) \Big|\mathbb{E}^y \bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big) - \mathbb{E}^y \bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big)\Big| \in \mathbb{R}, \quad N \in \mathbb{N}, \tag{4.55}$$

is uniformly globally Lipschitz continuous and satisfies that

$$\lim_{N\to\infty} \phi_\delta(y) f_1(y) \Big|\mathbb{E}^y \bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big) - \mathbb{E}^y \bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big)\Big| = 0 \tag{4.56}$$

for all $y \in I$ and all $\delta \in (0, |I|)$. Lemma 4.19 thus implies that

$$\lim_{N\to\infty} N\mathbb{E}^0\bigg[\phi_\delta\big(Y_{t_1,s}^{N,c_1}\big) f_1\big(Y_{t_1,s}^{N,c_1}\big) \cdot \Big|\mathbb{E}^{Y_{t_1,s}^{N,c_1}} \bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big) - \mathbb{E}^{Y_{t_1,s}^{N,c_1}} \bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big)\Big|\bigg] = 0 \tag{4.57}$$

for all $\delta \in (0, |I|)$. If the process is essentially zero at time t_1 (that is $1 - \phi_\delta\big(Y_{t_1,s}^{N,\zeta}\big) \approx 1$), then we may restart the process at time t_1 in the state 0. The Lipschitz continuity of $f_1, \ldots, f_n$ together with the moment estimate of Lemma 4.12 provides us with

$$\begin{aligned}
&\limsup_{N\to\infty} N\mathbb{E}^0\bigg[\Big(1 - \phi_\delta\big(Y_{t_1,s}^{N,c_1}\big)\Big)\Big|f_1\big(Y_{t_1,s}^{N,c_1}\big)\mathbb{E}^{Y_{t_1,s}^{N,c_1}} \bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big) - f_1(0)\mathbb{E}^0 \bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big)\Big|\bigg] \\
&\qquad \leq \limsup_{N\to\infty} N\mathbb{E}^0\Big[\Big(1 - \phi_\delta\big(Y_{t_1,s}^{N,c_1}\big)\Big) n L_F e^{L_\mu t_n} Y_{t_1,s}^{N,c_1}\Big] \\
&\qquad = n L_F e^{L_\mu t_n} c_1 \int_s^{t_1} \int \Big[\Big(1 - \phi_\delta(\chi_{t_1 - r})\Big)\chi_{t_1 - r}\Big] Q(d\chi)\, dr
\end{aligned} \tag{4.58}$$

for all $\delta \in (0,|I|)$. The last step is Lemma 4.19. By the dominated convergence theorem together with Lemma 4.15, the right-hand side of (4.58) converges to zero as $\delta \to 0$. Therefore, using (4.57), (4.58), the Markov property, Lemma 4.19, $\lim_{N\to\infty} \mathbb{E}^0\big[\phi_\delta(Y_{t_1,s}^{N,c_1})\big] = 0$ for all $\delta > 0$ and applying the induction hypothesis

$$\begin{aligned}
&\lim_{N\to\infty} N\mathbb{E}^0\Big[F\Big(\big(Y_{t,s}^{N,\zeta}\big)_{t\in[s,T]}\Big)\Big] \\
&= \lim_{N\to\infty} N\mathbb{E}^0\Big[\Big(\phi_\delta\big(Y_{t_1,s}^{N,c_1}\big) + 1 - \phi_\delta\big(Y_{t_1,s}^{N,c_1}\big)\Big)F\Big(\big(Y_{t,s}^{N,\zeta}\big)_{t\in[s,T]}\Big)\Big] \\
&= \lim_{\delta\to 0}\lim_{N\to\infty} N\mathbb{E}^0\Big[\phi_\delta\big(Y_{t_1,s}^{N,c_1}\big)f_1\big(Y_{t_1,s}^{N,c_1}\big)\mathbb{E}^{Y_{t_1,s}^{N,c_1}}\bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big)\Big] \\
&\quad + \lim_{\delta\to 0}\lim_{N\to\infty} \mathbb{E}^0\Big[1 - \phi_\delta\big(Y_{t_1,s}^{N,c_1}\big)\Big]f_1(0)\lim_{N\to\infty} N\mathbb{E}^0\Big[\bar{F}_2\Big(\big(Y_{t,t_1}^{N,\zeta}\big)_{t\in[t_1,T]}\Big)\Big] \\
&= \lim_{\delta\to 0}\int_s^{t_1} c_1 \int \phi_\delta\big(\chi_{t_1-r}\big)f_1\big(\chi_{t_1-r}\big)\mathbb{E}^{\chi_{t_1-r}}\Big[\bar{F}_2\Big(\big(Y_{t-t_1}\big)_{t\in[t_1,T]}\Big)\Big]Q(d\chi)\,dr \\
&\quad + f_1(0)\int_{t_1}^\infty \zeta(r)\int \bar{F}_2\Big(\big(\chi_{t-r}\big)_{t\in[t_1,T]}\Big)Q(d\chi)\,dr \\
&= \int_s^\infty \zeta(r)\int F\Big(\big(\chi_{t-r}\big)_{t\in[s,T]}\Big)Q(d\chi)\,dr.
\end{aligned} \tag{4.59}$$

The last step follows from the pointwise convergence $\phi_\delta(x) \to 1$ as $\delta \to 0$ for every $x \in (0,|I|)$ together with the dominated convergence theorem and from the Markov property (4.40). Finally let ζ be integrable. Approximate ζ with simple functions $(\zeta_n)_{n\in\mathbb{N}}$. Applying Lemma 4.12, it is straight forward to show that equation (4.52) with ζ replaced by ζ_n converges to equation (4.52) as $n \to \infty$. ☐

Lemma 4.21. *Assume 4.1 and 2.2. Fix $s \in [0,\infty)$. Suppose that $\zeta\colon [s,\infty) \to [0,\infty)$ and $\hat{\zeta}_N\colon [s,\infty) \to N\cdot I$ are locally square integrable functions for $N \in \mathbb{N}$ and that $\hat{\zeta}_N \to \zeta$ as $N \to \infty$ in L^1_{loc}. Let $(Y_{t,s}^{N,\hat{\zeta}}(i))_{t\in[s,\infty)}$, $i \le N$, be independent solutions of* (4.3) *satisfying $Y_{s,s}^{N,\hat{\zeta}}(i) = 0$ for every $N \in \mathbb{N}$. Then we have that*

$$\Big(\sum_{i=1}^N \delta_{Y_{t,s}^{N,\hat{\zeta}}(i)}\Big)_{s\le t\le T} \xrightarrow{w} \Big(\int \delta_{\eta_{t-u}}\Pi(du,d\eta)\Big)_{s\le t\le T} \qquad \text{as } N\to\infty \tag{4.60}$$

where Π is a Poisson point process on $[s,\infty)\times U$ with intensity measure $\zeta(u)du \otimes Q(d\eta)$. Moreover let $\bar{F}$ be a continuous function from $\mathrm{C}\big([s,\infty),\mathbb{R}\big)$ to $\mathbb{R}$ satisfying the Lipschitz condition

$$\big|\bar{F}(\eta) - \bar{F}(\bar{\eta})\big| \le L_{\bar{F}}\sum_{j=1}^n \big|\eta_{t_j} - \bar{\eta}_{t_j}\big| \qquad \forall\, \eta,\bar{\eta} \in \mathrm{C}\big([s,\infty),\mathbb{R}\big) \tag{4.61}$$

for some $s \le t_1 \le \cdots \le t_n \le T$, some $n \in \mathbb{N}$ and some $L_{\bar{F}} \in (0,\infty)$. In addition let $\bar{f}\colon I \to \mathbb{R}$ be a continuous function satisfying $|\bar{f}(x)| \le L_{\bar{f}}x$ for all $x \in I$. Then we have that

$$\lim_{N\to\infty}\mathbb{E}\Big[\bar{F}\Big(\Big(\sum_{i=1}^N \bar{f}\big(Y_{t,s}^{N,\hat{\zeta}}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big] = \mathbb{E}\Big[\bar{F}\Big(\Big(\int \bar{f}\big(\eta_{t-u}\big)\Pi(du,d\eta)\Big)_{t\in[s,\infty)}\Big)\Big]. \tag{4.62}$$

Proof. Let $\mathcal{F} \subseteq \mathrm{C}_c^2\big(I\cap(0,\infty),\mathbb{R}\big)$ be a countable dense subset of $\mathrm{C}_c\big(I\cap(0,\infty),\mathbb{R}\big)$. Tightness of

$$\Big\{\Big(\sum_{i=1}^N \delta_{Y_{t,s}^{N,\hat{\zeta}}(i)}\Big)_{t\in[s,\infty)} : N \in \mathbb{N}\Big\} \tag{4.63}$$

follows from tightness of

$$\Big\{\Big(\sum_{i=1}^N f\big(Y_{t,s}^{N,\hat{\zeta}}(i)\big)\Big)_{t\in[s,\infty)} : N \in \mathbb{N}\Big\} \qquad \forall\, f \in \mathcal{F}. \tag{4.64}$$

This type of argument has been established in Theorem 2.1 of Roelly-Coppoletta (1986) for the weak topology and C_0. Following the proof hereof, one can show the analogous argument for the vague topology and C_c. Fix $f \in \mathcal{F}$ and define $S_{t,s}^N := \sum_{i=1}^N f\big(Y_{t,s}^{N,\hat{\zeta}}(i)\big)$ for all $t \in [s,\infty)$ and $N \in \mathbb{N}$. Note that f is globally Lipschitz continuous. For $K \in \mathbb{N}$ and fixed $t \in [s,\infty)$, global Lipschitz continuity of f implies that

$$\mathbb{P}\Big[\big|S_{t,s}^N\big| \geq K\Big] \leq \frac{L_f}{K}\mathbb{E}\Big[\sum_{i=1}^N Y_{t,s}^{N,\hat{\zeta}}(i)\Big] \leq \frac{L_f}{K}\sup_{M\in\mathbb{N}} M\mathbb{E}^0\Big[Y_{t,s}^{M,\hat{\zeta}}(1)\Big] \tag{4.65}$$

for some constant $L_f < \infty$ and for all $N \in \mathbb{N}$. The right-hand side is finite according to Lemma 4.12 and converges to zero as $K \to \infty$. This proves tightness of $S_{t,s}^N$, $N \in \mathbb{N}$, for every $t \in [s,\infty)$. For the second part of the Aldous criterion, fix $T \in [s,\infty)$ and let τ_N, $N \in \mathbb{N}$, be stopping times which values in $[s,T]$. In addition define

$$\mathcal{G}_t^N f(x) := \Big(\frac{\hat{\zeta}_N(t)}{N} - x + \tilde{\mu}_N(x)\Big)f^{'}(x) + \frac{1}{2}\sigma_N^2(x)f^{''}(x) \qquad \forall\, x \in I, t \in [s,\infty), N \in \mathbb{N}. \tag{4.66}$$

The functions $\tilde{\mu}_N$, σ_N and σ_N^2 are uniformly globally Lipschitz continuous on the support of f according to Assumption 4.1. Therefore there exists a constant $C_f \in [1,\infty)$ such that $|\mathcal{G}_t^N f(x)| \leq C_f\big(\frac{\hat{\zeta}_N(t)}{N} + x\big)$ and $\big(f^{'}\sigma_N\big)^2(x) \leq C_f^2 x$ for all $x \in I$, $t \in [s,\infty)$ and all $N \in \mathbb{N}$. For fixed $\eta > 0$ and $\bar{\delta} \in [0,1]$, we use Itô's formula to obtain that

$$\begin{aligned}
&\eta^2\mathbb{P}\Big[\big|S_{\tau_N+\delta,s}^N - S_{\tau_N,s}^N\big| > \eta\Big] \leq \mathbb{E}\Big[\Big(S_{\tau_N+\delta,s}^N - S_{\tau_N,s}^N\Big)^2\Big] \\
&= \mathbb{E}\Big[\Big(\sum_{i=1}^N \int_{\tau_N}^{\tau_N+\delta} \mathcal{G}_u^N f\big(Y_{u,s}^{N,\hat{\zeta}}(i)\big)\,du + \sum_{i=1}^N \int_{\tau_N}^{\tau_N+\delta} \big(f^{'}\cdot\sigma_N\big)\big(Y_{u,s}^{N,\hat{\zeta}}(i)\big)dB_u(i)\Big)^2\Big] \\
&\leq 3C_f^2\mathbb{E}\Big[\Big(\int_{\tau_N}^{\tau_N+\delta} \hat{\zeta}_N(u)du\Big)^2\Big] + 3C_f^2\mathbb{E}\Big[\Big(\sum_{i=1}^N \int_0^\delta Y_{\tau_N+u,s}^{N,\hat{\zeta}}(i)du\Big)^2\Big] \\
&\qquad + 3\sum_{i=1}^N \mathbb{E}\Big[\int_0^\delta \big(f^{'}\sigma_N\big)^2\big(Y_{\tau_N+u,s}^{N,\hat{\zeta}}(i)\big)du\Big] \\
&\leq 3C_f^2\delta\mathbb{E}\Big[\int_{\tau_N}^{\tau_N+\delta} \big(\hat{\zeta}_N(u)\big)^2 du\Big] + 3C_f^2\delta\mathbb{E}\Big[\int_0^\delta \Big(\sum_{i=1}^N Y_{\tau_N+u,s}^{N,\hat{\zeta}}(i)\Big)^2 du\Big] \\
&\qquad + 3C_f^2\mathbb{E}\Big[\int_0^\delta \sum_{i=1}^N Y_{\tau_N+u,s}^{N,\hat{\zeta}}(i)du\Big] \\
&\leq \bar{\delta}\cdot 3C_f^2 \sup_{M\in\mathbb{N}} \int_s^{T+1} \Big(\hat{\zeta}_M(u)\Big)^2 du + \bar{\delta}\cdot 6C_f^2 \sup_{M\in\mathbb{N}} \mathbb{E}^0\Big[\sup_{s\leq u\leq T+1}\Big(\sum_{i=1}^M Y_{u,s}^{M,\hat{\zeta}}(i)\Big)^2\Big] + \bar{\delta}3C_f^2
\end{aligned} \tag{4.67}$$

for all $\delta \leq \bar{\delta}$ and for all $N \in \mathbb{N}$. The right-hand side of (4.67) is finite by Lemma 4.11. Letting $\bar{\delta} \to 0$, the left-hand side of (4.67) converges to zero uniformly in $N \in \mathbb{N}$ and $\delta \leq \bar{\delta}$. This proves tightness of (4.63) according to the Aldous criterion, see Aldous (1978).

Next we prove convergence of finite-dimensional distributions. Let $n \in \mathbb{N}$, $f \in \mathcal{F}$ with $f \geq 0$, $s \leq t_1 \leq \cdots \leq t_n$ and $\lambda_1, \cdots, \lambda_n \in [0,\infty)$ be arbitrary. Using independence we obtain that

$$\begin{aligned}
\mathbb{E}\Big[e^{-\sum_{j=1}^n \lambda_j S_{t_j,s}^N}\Big] &= \mathbb{E}\Big[e^{-\sum_{i=1}^N\sum_{j=1}^n \lambda_j f\left(Y_{t_j,s}^{N,\hat{\zeta}}(i)\right)}\Big] = \prod_{i=1}^N \Big[\mathbb{E}e^{-\sum_{j=1}^n \lambda_j f\left(Y_{t_j,s}^{N,\hat{\zeta}}(i)\right)}\Big] \\
&= \Bigg(1 - \frac{N\mathbb{E}^0\Big[1 - e^{-\sum_{j=1}^n \lambda_j f\left(Y_{t_j,s}^{N,\hat{\zeta}}(1)\right)}\Big]}{N}\Bigg)^N
\end{aligned} \tag{4.68}$$

for all $N \in \mathbb{N}$. Note that $1 - e^{-\sum_{j=1}^n \lambda_j f(x_j)} = \sum_{j=1}^n \big(1 - e^{-\lambda_j f(x_j)}\big)e^{-\sum_{i=1}^{j-1}\lambda_i f(x_i)}$ for all $x_1, \ldots, x_n \in [0,\infty)$.

Applying Lemma 4.20 to each summand of this telescope sum we get that

$$\begin{aligned}\lim_{N\to\infty}\mathbb{E}\Big[e^{-\sum_{j=1}^n \lambda_j S^N_{t_j,s}}\Big] &= \exp\Big(-\lim_{N\to\infty} N\mathbb{E}^0\Big[1-e^{-\sum_{j=1}^n \lambda_j f\left(Y^{N,\hat{\zeta}}_{t_j,s}(1)\right)}\Big]\Big)\\ &= \exp\Big(-\int\int_s^\infty \zeta(r)\Big(1-e^{-\sum_{j=1}^n \lambda_j f\left(\eta_{t_j-r}\right)}\Big)drQ(d\eta)\Big)\\ &= \mathbb{E}\Big[\exp\Big(-\sum_{j=1}^n \lambda_j\int f\Big(\eta_{t_j-u}\Big)\Pi(du,d\eta)\Big)\Big].\end{aligned}\tag{4.69}$$

This proves convergence of finite-dimensional distributions.

It remains to prove (4.62), which includes non-bounded test functions. Let $\bar{F}$ and $\bar{f}$ as in (4.62). By the previous step and by the Skorokhod representation of weak convergence (e.g. Theorem II.86.1 in [39]) there exists a version of $\big\{\big(\sum_{i=1}^N \delta_{Y^{N,\hat{\zeta}}_{t,s}(i)}\big)_{t\in[s,\infty)}\colon N\in\mathbb{N}\big\}$ which converges almost surely as $N\to\infty$. For every $K\in\mathbb{N}$, let $h_K\colon[0,\infty)\to[0,1]$ be a continuous function which satisfies $h_K(x)=1$ for every $x\in[\frac{1}{K},K]$ and $h_K(x)=0$ for every $x\notin[\frac{1}{2K},2K]$. Then $\bar{f}\cdot h_K\in \mathrm{C}_c\big(I\cap(0,\infty),\mathbb{R}\big)$ for all $K\in\mathbb{N}$. Thus

$$\Big\{M\wedge\bar{F}\Big(\Big(\sum_{i=1}^N\big(\bar{f}\cdot h_K\big)\big(Y^{N,\hat{\zeta}}_{t,s}(i)\big)\Big)_{t\in[s,\infty)}\Big)\colon N\in\mathbb{N}\Big\}\tag{4.70}$$

converges almost surely as $N\to\infty$ for all $M,K\in\mathbb{N}$. Next we observe that

$$\begin{aligned}&\varlimsup_{N\to\infty}\Big|\mathbb{E}\Big[\bar{F}\Big(\Big(\sum_{i=1}^N\bar{f}\big(Y^{N,\hat{\zeta}}_{t,s}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big]-\mathbb{E}\Big[M\wedge\bar{F}\Big(\Big(\sum_{i=1}^N\bar{f}\big(Y^{N,\hat{\zeta}}_{t,s}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big]\Big|\\ &=\varlimsup_{N\to\infty}\Big|\mathbb{E}\Big[\bar{F}\Big(\Big(\sum_{i=1}^N\bar{f}\big(Y^{N,\hat{\zeta}}_{t,s}(i)\big)\Big)_{t\in[s,\infty)}\Big)\mathbb{1}_{\bar{F}\left((\sum_{i=1}^N\bar{f}(Y^{N,\hat{\zeta}}_{t,s}(i)))_{t\in[s,\infty)}\right)>M}\Big]\Big|\\ &\le\sup_{N\in\mathbb{N}}\frac{1}{M}\mathbb{E}\Big[\Big(\bar{F}\Big(\Big(\sum_{i=1}^N\bar{f}\big(Y^{N,\hat{\zeta}}_{t,s}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big)^2\Big]\\ &\le\frac{1}{M}2\bar{F}\big(\underline{0}\big)+\frac{1}{M}\sup_{N\in\mathbb{N}}2L_{\bar{F}}^2L_{\bar{f}}^2\mathbb{E}\Big[\Big(\sum_{j=1}^n\sum_{i=1}^N Y^{N,\hat{\zeta}}_{t_j,s}(i)\Big)^2\Big]\\ &\le\frac{1}{M}2\bar{F}\big(\underline{0}\big)+\frac{1}{M}2L_{\bar{F}}^2L_{\bar{f}}^2n^2\sup_{N\in\mathbb{N}}\mathbb{E}\Big[\sup_{t\in[s,t_n]}\Big(\sum_{i=1}^N Y^{N,\hat{\zeta}}_{t,s}(i)\Big)^2\Big]\end{aligned}\tag{4.71}$$

for all $M\in\mathbb{N}$. The right-hand side is finite due to Lemma 4.11 for every $M\in\mathbb{N}$ and converges to zero as $M\to\infty$. The Lipschitz condition (4.61) implies that

$$\begin{aligned}&\varlimsup_{N\to\infty}\Big|\mathbb{E}\Big[M\wedge\bar{F}\Big(\Big(\sum_{i=1}^N\bar{f}\big(Y^{N,\hat{\zeta}}_{t,s}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big]-\mathbb{E}\Big[M\wedge\bar{F}\Big(\Big(\sum_{i=1}^N\big(\bar{f}\cdot h_K\big)\big(Y^{N,\hat{\zeta}}_{t,s}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big]\Big|\\ &\le\varlimsup_{N\to\infty}\sum_{j=1}^n L_{\bar{F}}L_{\bar{f}}\mathbb{E}\Big[\sum_{i=1}^N Y^{N,\hat{\zeta}}_{t_j,s}(i)\Big(1-h_K\big(Y^{N,\hat{\zeta}}_{t_j,s}(i)\big)\Big)\Big]\\ &\le\varlimsup_{N\to\infty}\sum_{j=1}^n L_{\bar{F}}L_{\bar{f}}\mathbb{E}\Big[\sum_{i=1}^N Y^{N,\hat{\zeta}}_{t_j,s}(i)\Big(\mathbb{1}_{\sup_{t\in[s,t_n]}Y^{N,\hat{\zeta}}_{t,s}(i)\ge K}+\mathbb{1}_{Y^{N,\hat{\zeta}}_{t_j,s}(i)\le\frac{1}{K}}\Big)\Big]\\ &\le\frac{nL_{\bar{F}}L_{\bar{f}}}{K}\sup_{N\in\mathbb{N}}\mathbb{E}\Big[\sup_{t\in[s,t_n]}\Big(\sum_{i=1}^N Y^{N,\hat{\zeta}}_{t,s}(i)\Big)^2\Big]+L_{\bar{F}}L_{\bar{f}}\sum_{j=1}^n\varlimsup_{N\to\infty}N\mathbb{E}^0\Big[Y^{N,\hat{\zeta}}_{t_j,s}(1)\wedge\tfrac{1}{K}\Big]\\ &\le\frac{nL_{\bar{F}}L_{\bar{f}}}{K}\sup_{N\in\mathbb{N}}\mathbb{E}^0\Big[\sup_{t\in[s,t_n]}\Big(\sum_{i=1}^N Y^{N,\hat{\zeta}}_{t,s}(i)\Big)^2\Big]+L_{\bar{F}}L_{\bar{f}}\sum_{j=1}^n\int_s^{t_j}\zeta(r)\int\chi_{t_j-r}\wedge\tfrac{1}{K}\;Q(d\chi)\,dr\end{aligned}\tag{4.72}$$

for all $M,K\in\mathbb{N}$. For the last step, we applied Lemma 4.20. The right-hand side of (4.72) is finite according to Lemmas 4.11 and 4.15 and converges to zero as $K\to\infty$ for every $M\in\mathbb{N}$ according to the dominated convergence theorem. Thus

$$\begin{aligned}
&\lim_{N\to\infty}\mathbb{E}\Big[\bar F\Big(\Big(\sum_{i=1}^{N}\bar f\big(Y_{t,s}^{N,\hat\zeta}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big] = \lim_{M\to\infty}\lim_{N\to\infty}\mathbb{E}\Big[M\wedge\bar F\Big(\Big(\sum_{i=1}^{N}\bar f\big(Y_{t,s}^{N,\hat\zeta}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big]\\
&=\lim_{M\to\infty}\lim_{K\to\infty}\lim_{N\to\infty}\mathbb{E}\Big[M\wedge\bar F\Big(\Big(\sum_{i=1}^{N}\big(\bar f\cdot h_K\big)\big(Y_{t,s}^{N,\hat\zeta}(i)\big)\Big)_{t\in[s,\infty)}\Big)\Big]\\
&=\lim_{M\to\infty}\lim_{K\to\infty}\mathbb{E}\Big[M\wedge\bar F\Big(\Big(\int\big(\bar f\cdot h_K\big)\big(\eta_{t-u}\big)\Pi(du,d\eta)\Big)_{t\in[s,\infty)}\Big)\Big]\\
&=\mathbb{E}\Big[\bar F\Big(\Big(\int\bar f\big(\eta_{t-u}\big)\Pi(du,d\eta)\Big)_{t\in[s,\infty)}\Big)\Big].
\end{aligned}\tag{4.73}$$

The last equality follows from the dominated convergence theorem together with Assumption 2.2. This proves (4.62). □

Remark: Instead of referring to Theorem 2.1 of Roelly-Coppoletta (1986) one could alternatively apply Theorem 3.1 of Jakubowski (1986).

4.4 Convergence of the loop-free process

Recall the loop-free N-island process from (4.2). The following lemma shows that the loop-free N-island process converges to the virgin island model.

Lemma 4.22. *Assume 4.1, 2.2, 3.2 and 4.2. Then we have that*

$$\Big(\sum_{i=1}^{N}\sum_{k=0}^{\infty}\delta_{Z_t^{N,k}(i)}\Big)_{t\le T}\xrightarrow{w}\Big(\sum_{(s,\eta)\in\mathcal{V}}\delta_{\eta_{t-s}}\Big)_{t\le T}\qquad \text{as } N\to\infty \tag{4.74}$$

for every $T\in[0,\infty)$.

Proof. Fix $T\in[0,\infty)$. Let $\mathcal{F}\subseteq \mathrm{C}_c^2\big((0,\infty),\mathbb{R}\big)$ be a countable dense subset of $\mathrm{C}_c\big((0,\infty),\mathbb{R}\big)$. Tightness of the left-hand side of (4.74) in $N\in\mathbb{N}$ follows from tightness of

$$\Big\{\Big(\sum_{i=1}^{N}\sum_{k=0}^{\infty}f\big(Z_t^{N,k}(i)\big)\Big)_{t\le T}: N\in\mathbb{N}\Big\}\qquad\forall\, f\in\mathcal{F}. \tag{4.75}$$

This type of argument has been established in Theorem 2.1 of Roelly-Coppoletta (1986) for the weak topology and C_0. Following the proof hereof, one can show the analogous argument for the vague topology and C_c. Fix $f\in\mathcal{F}$ and define $S_t^N:=\sum_{i=1}^{N}\sum_{k=0}^{\infty}f\big(Z_t^{N,k}(i)\big)$ for all $t\le T$ and $N\in\mathbb{N}$. For fixed $t\in[0,\infty)$, global Lipschitz continuity of f implies that

$$\mathbb{P}\Big[\big|S_t^N\big|\ge K\Big]\le\frac{L_f}{K}\mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k=0}^{\infty}Z_t^{N,k}(i)\Big]\le\frac{L_f}{K}\sup_{M\in\mathbb{N}}\mathbb{E}\Big[\sum_{i=1}^{M}\sum_{k=0}^{\infty}Z_t^{M,k}(i)\Big] \tag{4.76}$$

for some constant $L_f<\infty$ and for all $K,N\in\mathbb{N}$. The right-hand side is finite according to Lemma 4.5. This proves tightness of (4.75) for every fixed time point. For the second part of the Aldous criterion, let τ_N, $N\in\mathbb{N}$, be stopping times which are uniformly bounded by T. In addition define $H(\underline{x}):=\sum_{i=1}^{N}\sum_{k=0}^{\infty}f(x_i^k)$ and

$$\mathcal{G}^N H(\underline{x}):=\sum_{i=1}^{N}\sum_{k=0}^{\infty}\Big(\frac{\mathbb{1}_{k>0}}{N}\sum_{j=1}^{N}x_j^{k-1}-x_i^k+\mathbb{1}_{k=0}\mu_N(0)+\tilde\mu_N(x_i^k)\Big)f'(x_i^k)+\sum_{i=1}^{N}\sum_{k=0}^{\infty}\frac{1}{2}\sigma_N^2(x_i^k)f''(x_i^k) \tag{4.77}$$

for all $\underline{x}=(x_i^k)_{i\le N,k\in\mathbb{N}_0}\in I^{N\times\mathbb{N}_0}$. Assumption 4.1 implies that the functions μ_N, σ_N and σ_N^2 are uniformly globally Lipschitz continuous on the support of f. Moreover $N\mu_N(0)$ is bounded by 2θ uniformly in $N\in\mathbb{N}$. Therefore there exists a constant $C_H\in[1,\infty)$ such that $|\mathcal{G}^N H(\underline{x})|\le C_H\big(2\theta+\sum_{i=1}^{N}\sum_{k=0}^{\infty}x_i^k\big)$ for all $\underline{x}=$

$(x_i^k)_{i\le N,k\in\mathbb{N}_0}$ and all $N\in\mathbb{N}$ and such that $\big((f^{'}\sigma_N)(y)\big)\le C_H y$ for all $y\in I$ and $N\in\mathbb{N}$. For fixed $\eta>0$ and $\bar\delta\in[0,1]$, we use Itô's formula to obtain that

$$\begin{aligned}
&\eta^2\mathbb{P}\Big[\big|S^N_{\tau_N+\delta}-S^N_{\tau_N}\big|>\eta\Big]\le\mathbb{E}\Big[\Big(S^N_{\tau_N+\delta}-S^N_{\tau_N}\Big)^2\Big]\\
&=\mathbb{E}\Big[\Big(\int_{\tau_N}^{\tau_N+\delta}\mathcal{G}^N H\big(Z_u^{N,\cdot}(\cdot)\big)\,du+\sum_{i=1}^N\sum_{k=0}^\infty\int_{\tau_N}^{\tau_N+\delta}\big(f^{'}\cdot\sigma_N\big)\big(Z_u^{N,k}(i)\big)dB_u^k(i)\Big)^2\Big]\\
&\le 2C_H^2\mathbb{E}\Big[\Big(\int_0^\delta 2\theta+\sum_{i=1}^N\sum_{k=0}^\infty Z^{N,k}_{\tau_N+u}(i)\,du\Big)^2\Big]+2\sum_{i=1}^N\sum_{k=0}^\infty\mathbb{E}\Big[\int_0^\delta\big(f^{'}\sigma_N\big)^2\big(Z^{N,k}_{\tau_N+u}(i)\big)du\Big]\\
&\le\bar\delta\cdot 4C_H^2\sup_{M\in\mathbb{N}}\mathbb{E}\Big[\sup_{t\le T+1}\Big(2\theta+\sum_{i=1}^M\sum_{k=0}^\infty Z_t^{M,k}(i)\Big)^2\Big]
\end{aligned}\tag{4.78}$$

for all $\delta\le\bar\delta$ and for all $N\in\mathbb{N}$. The second step follows from $(a+b)^2\le 2a^2+2b^2$ for $a,b\in\mathbb{R}$ and from Itô's isometry. In the last step we used that $\big(\int_0^\delta h(u)du\big)^2\le\delta\int_0^\delta(h(u))^2du$ for every integrable function h. The right-hand side of (4.78) is finite according to Lemma 4.7. Letting $\bar\delta\to 0$, the left-hand side of (4.78) converges to zero uniformly in $N\in\mathbb{N}$ and $\delta\le\bar\delta$. Tightness of (4.75) now follows from the Aldous criterion, see Aldous (1978).

It remains to identify the limit of the left-hand side of (4.74) by proving that

$$\lim_{N\to\infty}\mathbb{E}\exp\Big(-\sum_{j=1}^n\lambda_j\sum_{k=0}^m\sum_{i=1}^N f\big(Z^{N,k}_{t_j}(i)\big)\Big)=\mathbb{E}\exp\Big(-\sum_{j=1}^n\lambda_j\sum_{k=0}^m\sum_{(s,\eta)\in\mathcal{V}^{(k)}}f\big(\eta_{t_j-s}\big)\Big)\tag{4.79}$$

for all $\lambda_1,\dots,\lambda_n\ge 0$, for all $0\le t_1\le\cdots\le t_n$, for all $m,n\in\mathbb{N}_0$ and for all non-negative $f\in\mathrm{C}_c^2((0,\infty))$. Lemma 4.6 justifies to restrict the summation over k to finitely many summands. We prove (4.79) by induction on $m\in\mathbb{N}_0$ using the Poisson limit (4.60) for independent one-dimensional diffusions. Define $F(\eta):=\sum_{j=1}^n\lambda_j f(\eta_{t_j})$ for every $\eta\in\mathrm{C}([0,\infty),I)$. Note that F satisfies the Lipschitz condition (3.6) for some constant $L_F\in(0,\infty)$. Let $\big(Y_{t,0}^{N,\zeta}(i)\big)_{t\ge0}$ and $\big(\tilde Y_{t,0}^{N,\zeta}(i)\big)_{t\ge0}$ be solutions of (4.3) with respect to the Brownian motion $(B_t^0(i))_{t\ge0}$ such that $Y_{0,0}^{N,\zeta}(i)=X_0(i)$, $\tilde Y_{0,0}^{N,\zeta}(i)=0$ and $\zeta_N(\cdot)\equiv N\mu_N(0)$ for every $N\in\mathbb{N}$ and $i\in\mathbb{N}$. In addition let $\big(Y_t(i)\big)_{t\ge0}$, $i\in\mathbb{N}$, be independent solutions of (1.5) with $Y_0(\cdot)\equiv X_0(\cdot)$. Note that $Z^{N,0}_\cdot(i)$ is the solution of (4.3) with respect to the Brownian motion $(B_t^0(i))_{t\ge0}$ with $Z_0^{N,0}(i)=X_0^N(i)$ for every $i\le N$ and every $N\in\mathbb{N}$. Recall the random permutation π^N from Assumption 3.2. The first-moment estimate of Lemma 4.12 implies that

$$\begin{aligned}
&\Big|\mathbb{E}\Big[\exp\Big(-\sum_{i=1}^N F\Big(\big(Z_t^{N,0}(i)\big)_{t\ge0}\Big)\Big)\Big]-\mathbb{E}\Big[\exp\Big(-\sum_{i=1}^N F\Big(\big(Y_{t,0}^{N,\zeta}(i)\big)_{t\ge0}\Big)\Big)\Big]\Big|\\
&\le L_F\sum_{j=1}^n\sum_{i=1}^N\mathbb{E}\Big[\Big|Z^{N,\zeta}_{t_j,0}(\pi_i^N)-Y^{N,\zeta}_{t_j,0}(i)\Big|\Big]\le L_F e^{L_\mu t_n}\sum_{j=1}^n\mathbb{E}\Big[\sum_{i=1}^N\Big|X_0^N(\pi_i^N)-X_0(i)\Big|\Big]
\end{aligned}$$

for all $N\in\mathbb{N}$. of Lemma 4.12. Letting $N\to\infty$ the right-hand side converges to zero according to Assumption 3.2. The process $\big(Y^{N,\zeta}_{\cdot,0}(i)\big)_{i\le N}$ in turn is close to $\big(\tilde Y^{N,\zeta}_{\cdot,0}(i)\big)_{i\le N}$ except for islands with a significant amount of mass at time zero. Formalizing this we use Lemma 4.12 to obtain that

$$\begin{aligned}
&\lim_{K\to\infty}\lim_{N\to\infty}\Big|\mathbb{E}\Big[\exp\Big(-\sum_{i=K+1}^N F\Big(\big(Y_{t,0}^{N,\zeta}(i)\big)_{t\ge0}\Big)\Big)\Big]-\mathbb{E}\Big[\exp\Big(-\sum_{i=1}^N F\Big(\big(\tilde Y_{t,0}^{N,\zeta}(i)\big)_{t\ge0}\Big)\Big)\Big]\Big|\\
&\le\lim_{K\to\infty}L_F n e^{L_\mu t_n}\sum_{i=K+1}^\infty\mathbb{E}\big[X_0(i)\big]+\lim_{K\to\infty}\lim_{N\to\infty}\mathbb{E}\Big[1-\exp\Big(-\sum_{i=1}^K F\Big(\big(\tilde Y_{t,0}^{N,\zeta}(i)\big)_{t\ge0}\Big)\Big)\Big].
\end{aligned}$$

The first summand on the right-hand side is zero according to Assumption 3.2. Note that $\mu_N\to\mu$ and $\sigma_N\to\sigma$ as $N\to\infty$ by Assumption 4.1. Thus $\big(\tilde Y^{N,\zeta}_{t,0}(i)\big)_{t\ge0}$ converges in distribution to the zero function as

$N \to \infty$ for every fixed $i \in \mathbb{N}$. Consequently the second summand on the right-hand side is zero. Moreover $\big(Y_{t,0}^{N,\zeta}(i)\big)_{t\geq 0}$ converges in distribution to $\big(Y_t(i)\big)_{t\geq 0}$ as $N \to \infty$ for every fixed $i \in \mathbb{N}$. These observations imply that

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{N} F\Big(\big(Z_t^{N,0}(i)\big)_{t\geq 0}\Big)\Big)\Big] \\
&= \lim_{K\to\infty}\lim_{N\to\infty} \mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{K} F\Big(\big(Y_{t,0}^{N,\zeta}(i)\big)_{t\geq 0}\Big)\Big)\Big]\mathbb{E}\Big[\exp\Big(-\sum_{i=K+1}^{N} F\Big(\big(Y_{t,0}^{N,\zeta}(i)\big)_{t\geq 0}\Big)\Big)\Big] \\
&= \lim_{K\to\infty} \mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{K} F\Big(\big(Y_t(i)\big)_{t\geq 0}\Big)\Big)\Big]\lim_{N\to\infty}\mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{N} F\Big(\big(\tilde{Y}_{t,0}^{N,\zeta}(i)\big)_{t\geq 0}\Big)\Big)\Big] \\
&= \mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{\infty} F\Big(\big(Y_t(i)\big)_{t\geq 0}\Big)\Big)\Big]\mathbb{E}\Big[\exp\Big(-\sum_{(s,\eta)\in\Pi^\theta} F\big((\eta_{t-s})_{t\geq 0}\big)\Big)\Big] \\
&= \mathbb{E}\Big[\exp\Big(-\sum_{(s,\eta)\in\mathcal{V}^{(0)}} F\big((\eta_{t-s})_{t\geq 0}\big)\Big)\Big].
\end{aligned}$$

The last but one step follows from Lemma 4.21 with $\hat{\zeta}_N(\cdot) \equiv N\mu_N(0)$ and $\zeta(\cdot) \equiv \theta$. This proves (4.79) in the base case $m = 0$.

For the induction step $m \to m+1$, we observe that a version of $(Z_t^{N,m+1}(i))_{t\geq 0}$ conditioned on $\zeta_N(r) := \sum_{j=1}^{N} Z_r^{N,m}(j)$, $r \geq 0$, is given by the one-dimensional diffusion $(Y_{t,0}^{N,\zeta})_{t\geq 0}$ with vanishing immigration. Thus we may realize $(Z_t^{N,m+1}(i))_{t\geq 0}$ by choosing a suitable version of $(Z_t^{N,m}(j))_{t\geq 0}$, $j = 1,\ldots,N$, and by independently sampling a version of $(Y_{t,0}^{N,\zeta})_{t\geq 0}$ whose driving Brownian motion is independent of $\{(Z_t^{N,m}(j))_{t\geq 0} \colon j = 1,\ldots,N\}$. Tightness of $\big\{\big(\sum_{i=1}^{N}\sum_{k=0}^{\tilde{m}} \delta_{Z_t^{N,k}(i)}\big)_{t\leq T}\big\}$ together with the induction hypothesis implies that

$$\Big(\sum_{i=1}^{N}\sum_{k=0}^{\tilde{m}} \delta_{Z_t^{N,k}(i)}\Big)_{t\leq T} \xrightarrow{w} \Big(\sum_{k=0}^{\tilde{m}}\sum_{(s,\eta)\in\mathcal{V}^{(k)}} \delta_{\eta_{t-s}}\Big)_{t\leq T} \qquad \text{as } N \to \infty \tag{4.80}$$

for every $\tilde{m} \leq m$. Thanks to the Skorokhod representation of weak convergence (e.g. Theorem II.86.1 in [39]), we may assume that the convergence in (4.80) holds almost surely. As a consequence we obtain that

$$\Big(\sum_{i=1}^{N} \delta_{Z_t^{N,m}(i)}\Big)_{t\leq T} \longrightarrow \Big(\sum_{(s,\eta)\in\mathcal{V}^{(m)}} \delta_{\eta_{t-s}}\Big)_{t\leq T} \qquad \text{as } N \to \infty \tag{4.81}$$

holds almost surely. Using arguments from the proof of (4.62), one can deduce from this that

$$\Big(\sum_{i=1}^{N} Z_t^{N,m}(i)\Big)_{t\leq T} \xrightarrow{w} \Big(V_t^{(m)}\Big)_{t\leq T} \qquad \text{as } N \to \infty \tag{4.82}$$

holds almost surely where the total mass of the n-th generation of the virgin island model is defined as

$$V_t^{(n)} := \sum_{(s,\eta)\in\mathcal{V}^{(n)}} \eta_{t-s}, \quad t \geq 0, \tag{4.83}$$

for every $n \in \mathbb{N}_0$. Together with continuity of $\big(\eta_s\big)_{s\leq T} \mapsto \int_0^t |\eta_s - V_s^{(m)}|\,ds$, (4.82) implies that

$$\int_0^t \Big|\sum_{i=1}^{N} Z_s^{N,m}(i) - V_s^{(m)}\Big|\,ds \xrightarrow{N\to\infty} 0 \qquad \text{almost surely.} \tag{4.84}$$

Now the main step of the proof is Lemma 4.21 with $\hat{\zeta}_N(t) := \sum_{i=1}^{N} Z_t^{N,m}(i)$ and $\zeta(t) := V_t^{(m)}$ for all $t \geq 0$.

Lemma 4.21 implies that

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{N} F\Big(\big(Z_t^{N,m+1}\big)_{t\le T}\Big)\Big)\big|\big(Z_\cdot^{M,m}\big)_{M\in\mathbb{N}}\Big]\\
&= \mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{N} F\Big(\big(Y_{t,0}^{N,\zeta}\big)_{t\le T}\Big)\Big)\Big|\big(Z_\cdot^{M,m}\big)_{M\in\mathbb{N}}\Big]\\
&\longrightarrow \mathbb{E}\Big[\exp\Big(-\int F\big((\eta_{t-s})_{t\le T}\big)\Pi^{(m+1)}(ds,d\eta)\Big)\Big|\mathcal{V}^{(m)}\Big] \qquad \text{as } N\to\infty
\end{aligned} \tag{4.85}$$

almost surely. Here $\Pi^{(m+1)}$ conditioned on $\mathcal{V}^{(m)}$ is a Poisson point process on $[0,\infty)\times U$ with intensity measure

$$\begin{aligned}
&\mathbb{E}\Big[\Pi^{(m+1)}(ds,d\eta)|\mathcal{V}^{(m)}\Big] = V_s^{(m)}\,ds\otimes Q(d\eta)\\
&= \sum_{(r,\psi)\in\mathcal{V}^{(m)}} \psi_{s-r}ds\otimes Q(d\eta) = \sum_{(r,\psi)\in\mathcal{V}^{(m)}} \mathbb{E}\Pi^{(r,\psi)}(ds,d\eta).
\end{aligned} \tag{4.86}$$

Due to this decomposition of $\Pi^{(m+1)}$, we may realize $\Pi^{(m+1)}$ conditioned on $\mathcal{V}^{(m)}$ as the independent superposition of $\{\Pi^{(m,s,\psi)}\colon (s,\psi)\in\mathcal{V}^{(m)}\}$. In other words, $\Pi^{(m+1)}$ is equal in distribution to the $(m+1)$st-generation of the virgin island model. Therefore we get that

$$\begin{aligned}
&\lim_{N\to\infty}\mathbb{E}\Big[\exp\Big(-\sum_{k=0}^{m+1}\sum_{i=1}^{N} F\Big(\big(Z_t^{N,k}(i)\big)_{t\le T}\Big)\Big)\Big]\\
&= \mathbb{E}\Big[\lim_{N\to\infty}\exp\Big(-\sum_{k=0}^{m}\sum_{i=1}^{N} F\big(Z_\cdot^{N,k}(i)\big)\Big)\mathbb{E}\Big[\exp\Big(-\sum_{i=1}^{N} F\big(Z_\cdot^{N,m+1}(i)\big)\Big)\Big|\big(Z_\cdot^{M,m}\big)_{M\in\mathbb{N}}\Big]\Big]\\
&= \mathbb{E}\Big[\exp\Big(-\sum_{k=0}^{m}\sum_{(s,\eta)\in\mathcal{V}^{(k)}} F\big((\eta_{t-s})_{t\le T}\big)\Big)\mathbb{E}\Big[\exp\Big(-\sum_{(s,\eta)\in\mathcal{V}^{(m+1)}} F\big((\eta_{t-s})_{t\le T}\big)\Big)\Big|\mathcal{V}^{(m)}\Big]\Big]\\
&= \mathbb{E}\Big[\exp\Big(-\sum_{k=0}^{m+1}\sum_{(s,\eta)\in\mathcal{V}^{(k)}} F\big((\eta_{t-s})_{t\le T}\big)\Big)\Big]
\end{aligned}$$

which proves (4.79) and completes the proof of Lemma 4.22. □

4.5 Reducing to the loop-free (N,μ_N,σ_N^2)-process

Next we show that the (N,μ_N,σ_N^2)-process with migration levels and the loop-free (N,μ_N,σ_N^2)-process are identical in the limit $N\to\infty$. Our proof formalizes the following intuition. The individuals of a certain migration level are concentrated on essentially finitely many islands. That these finitely many islands are populated by migrants of a different migration level has a probability of order $\frac{1}{N}$. As a consequence, all individuals on one fixed island have the same migration level in the limit $N\to\infty$. This intuition is subject of Lemma 4.24.

First we show that a generation cannot be dispersed uniformly over all islands. To obtain this interpretation from the following lemma, assume $X_t^{N,k}(i)\approx\frac{1}{N}$ for all $i\le N$ and some time $t\ge 0$. Then the cutting operation in (4.87) has no effect for N large enough. However it is clear that the total mass of all individuals with migration level k does not tend to zero as $N\to\infty$. Thus, as a consequence of the following lemma, $X_t^{N,k}(i)\approx\frac{1}{N}$ cannot be true.

Lemma 4.23. *Assume 4.1, 3.2 and 4.2. Then any solution of* (4.1) *satisfies*

$$\sum_{k\ge 0}\sup_{N\in\mathbb{N}}\mathbb{E}\Big[\sum_{i=1}^{N}\Big(X_t^{N,k}(i)\wedge\delta\Big)\Big] \xrightarrow{\delta\to 0} 0 \tag{4.87}$$

for all $t\in[0,\infty)$*. The assertion is also true if* $X_t^{N,k}(i)$ *is replaced by* $Z_t^{N,k}(i)$*.*

Proof. Fix $T \in [0,\infty)$ and $t \in [0,T]$. According to Lemma 4.6, for each $\varepsilon > 0$, there exists a $k_0 \in \mathbb{N}$ such that

$$\sum_{k \geq k_0} \sup_{N \in \mathbb{N}} \sup_{s \leq T} \mathbb{E}\Big[\sum_{i=1}^{N} X_s^{N,k}(i)\Big] \leq \varepsilon. \tag{4.88}$$

Thus it suffices to prove convergence of every summand in (4.87). In addition, if we forget the migration levels in the (N, μ_N, σ_N^2)-process with migration levels, then we obtain the N-island model. More formally, Lemma 4.4 shows that

$$\bar{X}_t^N(i) := \sum_{m \geq 0} X_t^{N,m}(i) \qquad \forall\, i \leq N, t \geq 0 \tag{4.89}$$

defines an N-island model. Recall $\tilde{\tau}_K^N$ from (4.8) and fix $K \in \mathbb{N}$. According to Lemma 4.8 it suffices to prove (4.87) with expectation being restricted to the event $\{\tilde{\tau}_K^N > t\}$ for every $N \in \mathbb{N}$. Now Lemma 3.3 in [23] shows that, on the event $\{\tilde{\tau}_K^N > t\}$, $\bar{X}_t^N(i)$ is stochastically bounded above by $Y_{t,0}^{N,K+N\mu_N(0)}$. Hence we get for all $N, K \in \mathbb{N}$, $k \in \mathbb{N}_0$ and $\tilde{\delta} > 0$ that

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{i=1}^{N}\Big(X_t^{N,k}(i) \wedge \delta\Big)\mathbb{1}_{\tilde{\tau}_K^N > t}\Big] \\
&\leq \sum_{i=1}^{N} \mathbb{E}\left[\mathbb{1}_{\bar{X}_0^N(i) \leq \tilde{\delta}}\Big(X_t^{N,k}(i) \wedge \delta\Big)\mathbb{1}_{\tilde{\tau}_K^N > t}\right] + \delta \mathbb{E}\left[\sum_{i=1}^{N} \mathbb{1}_{\bar{X}_0^N(i) > \tilde{\delta}}\right] \\
&\leq \sum_{i=1}^{N} \mathbb{E}^{\bar{X}_0^N(i) \wedge \tilde{\delta}}\Big[Y_{t,0}^{N,K+N\mu_N(0)} \wedge \delta\Big] + \frac{\delta}{\tilde{\delta}}\mathbb{E}\Big[\sum_{i=1}^{N} \bar{X}_0^N(i)\Big] \\
&\leq N\mathbb{E}^0\Big[Y_{t,0}^{N,K+N\mu_N(0)} \wedge \delta\Big] + C_t \sum_{i=1}^{N} \mathbb{E}\Big[\Big(\bar{X}_0^N(i) \wedge \tilde{\delta}\Big)\Big] + \frac{\delta}{\tilde{\delta}} \sup_{M \in \mathbb{N}} \mathbb{E}\Big[\sum_{i=1}^{M} \bar{X}_0^M(i)\Big]
\end{aligned} \tag{4.90}$$

for some constant $C_t < \infty$. The last step follows from Lemma 4.12. Next we let $N \to \infty$ and $\delta \to 0$ in (4.90). Applying Lemma 4.21 with $\hat{\zeta}_N := K + N\mu_N(0)$ and using Assumption 3.2, we see that the limit of (4.90) as $N \to \infty$ and as $\delta \to 0$ is bounded above by

$$\lim_{\delta \to 0}(K+\theta)\int\int_0^t \big(\chi_{t-r} \wedge \delta\big)\, dr\, Q(d\chi) + C_t \mathbb{E}\Big[\sum_{i=1}^{\infty}\Big(\bar{X}_0(i) \wedge \tilde{\delta}\Big)\Big] \tag{4.91}$$

for every $K \in \mathbb{N}$ and $\tilde{\delta} > 0$. The first summand in (4.91) is zero by the dominated convergence theorem and Lemma 4.15. The last summand converges to zero as $\tilde{\delta} \to 0$ by the dominated convergence theorem and Assumption 3.2. This completes the proof of (4.87). The proof of (4.87) with $X_t^{N,k}(i)$ replaced by $Z_t^{N,k}(i)$ is similar □

Now we prove that all individuals on a fixed island have the same migration level in the limit $N \to \infty$. More precisely, we show that $X_t^{N,k}(i)$ and $\sum_{m \neq k} X_t^{N,m}(i)$ cannot be big at the same time for any $i \leq N$ or any $k \in \mathbb{N}_0$.

Lemma 4.24. *Assume 4.1, 3.2 and 4.2. Then any solution of* (4.1) *satisfies*

$$\lim_{N \to \infty} \sup_{s,t \leq T} \mathbb{E}\Big[\sum_{i=1}^{N} \sum_{k \geq 0} X_s^{N,k}(i) \sum_{m \neq k} X_t^{N,m}(i)\Big] = 0 \tag{4.92}$$

for every $T \in [0,\infty)$. The assertion is also true if $X_t^{N,k}(i)$ is replaced by $Z_t^{N,k}(i)$.

Proof. We begin with the case $s = t$. Fix $K \in \mathbb{N}$ and recall $\tilde{\tau}_K^N$ from (4.8). By the uniform local Lipschitz continuity of μ_N, there exists a finite constant L_K such that $\tilde{\mu}_N(x) \leq L_K x$ for all $x \leq K$ and all $N \in \mathbb{N}$.

Applying Itô's formula, we see that

$$\begin{aligned} d\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{t\wedge\tilde\tau^N_K}(i)\sum_{m\neq k} X^{N,m}_{t\wedge\tilde\tau^N_K}(i) &\leq 2\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{t\wedge\tilde\tau^N_K}(i)\mu_N(0)\,dt \\ &+\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{t\wedge\tilde\tau^N_K}(i)\sum_{m\neq k}\left(\frac{1}{N}\sum_{j=1}^{N} X^{N,m-1}_{t\wedge\tilde\tau^N_K}(j)\,dt + L_K X^{N,m}_{t\wedge\tilde\tau^N_K}(i)\,dt + d\,M^{N,m}_t(i)\right) \\ &+\sum_{i=1}^{N}\sum_{k\geq 0}\sum_{m\neq k} X^{N,m}_{t\wedge\tilde\tau^N_K}(i)\left(\frac{1}{N}\sum_{j=1}^{N} X^{N,k-1}_{t\wedge\tilde\tau^N_K}(j)\,dt + L_K X^{N,k}_{t\wedge\tilde\tau^N_K}(i)\,dt + d\,\tilde M^{N,k}_t(i)\right) \end{aligned} \tag{4.93}$$

where $(M^{N,m}_t(i))_{t\geq 0}$ and $(\tilde M^{N,k}_t(i))_{t\geq 0}$ are suitable martingales for each $i\leq N$ and $k,m\in\mathbb{N}_0$. Now take expectations to obtain that

$$\begin{aligned} \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{t\wedge\tilde\tau^N_K}(i)\sum_{m\neq k} X^{N,m}_{t\wedge\tilde\tau^N_K}(i)\Big] &\leq 2\mu_N(0)\int_0^t \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{s\wedge\tilde\tau^N_K}(i)\Big]\,ds \\ &+\int_0^t \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{s\wedge\tilde\tau^N_K}(i)\frac{1}{N}\sum_{j=1}^{N}\sum_{m\geq 0} X^{N,m-1}_{s\wedge\tilde\tau^N_K}(j)\Big]\,ds \\ &+\int_0^t \mathbb{E}\Big[\sum_{k\geq 0}\frac{1}{N}\sum_{j=1}^{N} X^{N,k-1}_{s\wedge\tilde\tau^N_K}(j)\sum_{i=1}^{N}\sum_{m\geq 0} X^{N,m}_{s\wedge\tilde\tau^N_K}(i)\Big]\,ds \\ &+2L_K\int_0^t \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{s\wedge\tilde\tau^N_K}(i)\sum_{m\neq k} X^{N,m}_{s\wedge\tilde\tau^N_K}(i)\Big]\,ds \end{aligned} \tag{4.94}$$

for every $N\in\mathbb{N}$ and $t\leq T$. Note that the right-hand side is finite. Using Gronwall's inequality, $\mu_N(0)\leq 2\theta/N$ and $\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{t\wedge\tilde\tau^N_K}(i)\leq K$, we conclude that

$$\sup_{t\leq T}\mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k\geq 0} X^{N,k}_{t\wedge\tilde\tau^N_K}(i)\sum_{m\neq k} X^{N,m}_{t\wedge\tilde\tau^N_K}(i)\Big]\leq\frac{1}{N}T(4\theta K+2K^2)\cdot e^{2L_K T} \tag{4.95}$$

for every $N\in\mathbb{N}$. Letting $N\to\infty$ proves (4.92) in the case $s=t$. For the case $s<t$, apply Itô's formula in

$$X^{N,k}_{s\wedge\tilde\tau^N_K}(i)\sum_{m\neq k}\int_{s\wedge\tilde\tau^N_K}^{t\wedge\tilde\tau^N_K} dX^{N,m}_r(i) \tag{4.96}$$

and use similar estimates as above. The case $s>t$ is analogous to the case $s<t$. □

Knowing that there is asymptotically at most one generation on every island, we are now in a position to prove that $\big(X^{N,k}_t(i)\big)_{t\geq 0}$ and $\big(Z^{N,k}_t(i)\big)_{t\geq 0}$ are close to each other.

Lemma 4.25. *Assume 3.1, 3.2 and 4.2. For each $N\in\mathbb{N}$, let*

$$\big\{\big(X^{N,k}_t(i),\bar B^k_t(i)\big)_{t\geq 0}\colon k\in\mathbb{N}_0, i\leq N\big\} \tag{4.97}$$

be a solution of (4.1) *and let*

$$\big\{(Z^{N,k}_t(i))_{t\geq 0}\colon k\in\mathbb{N}_0, i\leq N\big\} \tag{4.98}$$

be the unique solution of (4.2) *with* $\big\{(B^k_t(i))_{t\geq 0}\big\}$ *replaced by* $\big\{(\bar B^k_t(i))_{t\geq 0}\big\}$*. Then*

$$\mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k=0}^{\infty}\big|X^{N,k}_t(i)-Z^{N,k}_t(i)\big|\Big]\longrightarrow 0\qquad\text{as } N\to\infty \tag{4.99}$$

for every $t\in[0,\infty)$.

Proof. In the first step, assume that σ_N^2 and μ_N are uniformly globally Lipschitz continuous and bounded. The general case will later be handled by a stopping argument.

Define $x^+ := \max(x,0)$ and $x^- := (-x)^+$ for all $x \in \mathbb{R}$. We first prove (4.99) with $|x|$ replaced by x^+ and x^-, respectively, separately. As $\mathbb{R} \ni x \mapsto x^+$ is not differentiable, we will apply Itô's formula to ϕ_n being defined as follows. Let $1 = a_0 > a_1 > \cdots > a_n > \cdots > 0$ satisfy

$$\int_{a_1}^{1} \frac{1}{u}\,du = 1,\ \int_{a_2}^{a_1} \frac{1}{u}\,du = 2, ..., \int_{a_n}^{a_{n-1}} \frac{1}{u}\,du = n, ... \tag{4.100}$$

Note that $a_n \to 0$ as $n \to \infty$. For every $n \in \mathbb{N}$, there exists a continuous function $\psi_n \colon \mathbb{R} \to [0,\infty)$ with support in (a_n, a_{n-1}) such that

$$0 \le \psi_n(u) \le \frac{2}{nu} \text{ for all } u > 0 \text{ and } \int_{a_n}^{a_{n-1}} \psi_n(x)\,dx = 1. \tag{4.101}$$

Note that $\psi_n(u) \le \frac{2}{na_n}$ for all $u \ge 0$ and $n \in \mathbb{N}$. With this, we define the function

$$\phi_n(x) := \mathbb{1}_{x>0} \int_0^x dy \int_0^y \psi_n(u)\,du, \quad x \in \mathbb{R}, \tag{4.102}$$

for every $n \in \mathbb{N}$. These functions satisfy $\phi_n \in \mathrm{C}^2(\mathbb{R})$, $|\phi_n^{'}(x)| \le 1$, $\phi_n^{''}(x) = \mathbb{1}_{x>0}\psi_n(x)$, $\phi_n(x) \le x^+$ and $\phi_k(x) \to x^+$ as $k \to \infty$ for every $x \in \mathbb{R}$ and $n \in \mathbb{N}$.

Denote the difference process by $\Delta_t^k(i) := X_t^{N,k}(i) - Z_t^{N,k}(i)$ for all $i \in G$, $t \ge 0$, $k \in \mathbb{N}_0$ and $N \in \mathbb{N}$. The dependence on N is suppressed for the sake of a more compact notation. By definition of ϕ_n, $x^+ \le \phi_n(x) + a_{n-1} \wedge x^+$ for all $x \in \mathbb{R}$ and $n \in \mathbb{N}$. Thus

$$\begin{aligned} \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k=0}^{\infty}\big(\Delta_t^k(i)\big)^+\Big] &\le \sum_{k=K+1}^{\infty} \sup_{M\in\mathbb{N}} \sum_{i=1}^{M} \mathbb{E}\Big[X_t^{M,k}(i)\Big] \\ &\quad + \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k=0}^{K}\phi_{N^2}\big(\Delta_t^k(i)\big)\Big] + \sum_{k=0}^{\infty}\sup_{M\ge 1}\mathbb{E}\Big[\sum_{i=1}^{M}\Big(a_{N^2-1}\wedge X_t^{M,k}(i)\Big)\Big] \end{aligned} \tag{4.103}$$

for all $K \in \mathbb{N}$ and $N \in \mathbb{N}$. The first summand on the right-hand side converges to zero as $K \to \infty$ uniformly in $N \in \mathbb{N}$ according to Lemma 4.6. The last summand on the right-hand side converges to zero as $N \to \infty$ according to Lemma 4.23. Consequently

$$\limsup_{N\to\infty} \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k=0}^{\infty}\Big(\Delta_t^k(i)\Big)^+\Big] \le \limsup_{K\to\infty}\limsup_{N\to\infty}\mathbb{E}\Big[\sum_{i=1}^{N}\sum_{k=0}^{K}\phi_{N^2}\Big(\Delta_t^k(i)\Big)\Big]. \tag{4.104}$$

To estimate the right-hand side, we apply Itô's formula to obtain that

$$\begin{aligned} &d\sum_{i=1}^{N}\phi_{N^2}\big(\Delta_t^k(i)\big) = \\ &\quad \sum_{i=1}^{N}\phi_{N^2}^{'}\left(\sqrt{\frac{X_t^{N,k}(i)\cdot\sigma_N^2\Big(\sum_{m=0}^{\infty}X_t^{N,m}(i)\Big)}{\sum_{m=0}^{\infty}X_t^{N,m}(i)}} - \sqrt{\sigma_N^2\big(Z_t^{N,k}(i)\big)}\right)\,d\bar{B}_t^k(i) \\ &\quad + \sum_{i=1}^{N}\phi_{N^2}^{'}\big(\Delta_t^k(i)\big)\left(\frac{1}{N}\sum_{j=1}^{N}\Delta_t^{k-1}(j) - \Delta_t^k(i)\right)dt \\ &\quad + \sum_{i=1}^{N}\phi_{N^2}^{'}\big(\Delta_t^k(i)\big)\left(\frac{X_t^{N,k}(i)\cdot\tilde{\mu}_N\Big(\sum_{m=0}^{\infty}X_t^{N,m}(i)\Big)}{\sum_{m=0}^{\infty}X_t^{N,m}(i)} - \tilde{\mu}_N\Big(Z_t^{N,k}(i)\Big)\right)dt \\ &\quad + \sum_{i=1}^{N}\frac{\phi_{N^2}^{''}\big(\Delta_t^k(i)\big)}{2}\left(\sqrt{\frac{X_t^{N,k}(i)\cdot\sigma_N^2\Big(\sum_{m=0}^{\infty}X_t^{N,m}(i)\Big)}{\sum_{m=0}^{\infty}X_t^{N,m}(i)}} - \sqrt{\sigma_N^2\big(Z_t^{N,k}(i)\big)}\right)^2 dt \end{aligned} \tag{4.105}$$

for every $k \in \mathbb{N}_0$. Now we simplify the last summand on the right-hand side using the assumption of σ_N^2 being linear. This linearity implies that $x \cdot \sigma_N^2(y)/y = \sigma_N^2(x)$ for $x = X_t^{N,k}(i)$ and $y = \sum_{m=0}^{\infty} X_t^{N,m}(i)$. In addition take expectations, estimate $|\phi'_{N^2}(x)| \le 1$ and use $-\phi'_{N^2}(x) \cdot x \le 0$ for $x \in \mathbb{R}$. Thus we have that

$$\begin{aligned}
&\frac{d}{dt}\mathbb{E}\Big[\sum_{i=1}^{N} \phi_{N^2}\big(\Delta_t^k(i)\big)\Big] \\
&\le \mathbb{E}\Big[\sum_{j=1}^{N} \big|\Delta_t^{k-1}(j)\big|\Big] + \mathbb{E}\Big[\sum_{i=1}^{N} \big|\tilde{\mu}_N\big(X_t^{N,k}(i)\big) - \tilde{\mu}_N\big(Z_t^{N,k}(i)\big)\big|\Big] \\
&+ \mathbb{E}\Bigg[\sum_{i=1}^{N} \Big|\frac{X_t^{N,k}(i) \cdot \tilde{\mu}_N\Big(\sum_{m=0}^{\infty} X_t^{N,m}(i)\Big)}{\sum_{m=0}^{\infty} X_t^{N,m}(i)} - \tilde{\mu}_N\Big(X_t^{N,k}(i)\Big)\Big|\Bigg] \\
&+ \mathbb{E}\Bigg[\sum_{i=1}^{N} \frac{\phi''_{N^2}\big(\Delta_t^k(i)\big)}{2} \left(\sqrt{\sigma_N^2\big(X_t^{N,k}(i)\big)} - \sqrt{\sigma_N^2\big(Z_t^{N,k}(i)\big)}\right)^2\Bigg]
\end{aligned} \tag{4.106}$$

for every $t \ge 0$, $k \in \mathbb{N}_0$ and $N \in \mathbb{N}$. The last but one summand on the right-hand side is estimated as follows. Let $\delta > 0$. In case of $X_t^{N,k}(i) \ge \delta$, apply the inequality $|\frac{a}{a+b}\tilde{\mu}_N(a+b) - \tilde{\mu}_N(a)| \le 2L_\mu b$ with $a = X_t^{N,k}(i)$ and $b = \sum_{m \ne k}^{\infty} X_t^{N,m}(i)$. In case of $X_t^{N,k} < \delta$, use $|\tilde{\mu}_N(x)| \le L_\mu x$ for all $x \ge 0$. For the last summand on the right-hand side, use $(\sqrt{x} - \sqrt{y})^2 \le |x-y|$ and $|\sigma_N^2(x) - \sigma_N^2(y)| \le L_\sigma |x-y|$ for $x, y \in [0,\infty)$. Moreover recall $\phi''_{N^2}(x) \le \frac{2}{N^2 x^+}$ for all $x > 0$. Define $L := L_\mu \vee L_\sigma$. Hence, we get that

$$\begin{aligned}
&\frac{d}{dt}\mathbb{E}\Big[\sum_{i=1}^{N} \phi_{N^2}\big(\Delta_t^k(i)\big)\Big] \\
&\le \mathbb{E}\Big[\sum_{j=1}^{N} \big|\Delta_t^{k-1}(j)\big|\Big] + L\mathbb{E}\Big[\sum_{i=1}^{N} \big|X_t^{N,k}(i) - Z_t^{N,k}(i)\big|\Big] \\
&+ 2L\mathbb{E}\Bigg[\sum_{i=1}^{N} \mathbb{1}_{X_t^{N,k}(i) \ge \delta} \sum_{m \ne k} X_t^{N,m}(i)\Bigg] \\
&+ 2L\mathbb{E}\Bigg[\sum_{i=1}^{N} \Big(X_t^{N,k}(i) \wedge \delta\Big)\Bigg] + \mathbb{E}\Bigg[\sum_{i=1}^{N} \frac{L\big(\Delta_t^k(i)\big)^+}{N^2\big(\Delta_t^k(i)\big)^+}\Bigg]
\end{aligned} \tag{4.107}$$

for every $t \ge 0$, $k \in \mathbb{N}_0$ and $N \in \mathbb{N}$. Summing over $k \in \{0, \ldots, K\}$ leads to

$$\begin{aligned}
&\frac{d}{dt}\mathbb{E}\Bigg[\sum_{k=0}^{K}\sum_{i=1}^{N} \phi_{N^2}\big(\Delta_t^k(i)\big)\Bigg] \\
&\le (1+L)\mathbb{E}\Bigg[\sum_{k=0}^{\infty}\sum_{i=1}^{N} \Big|\Delta_t^k(i)\Big|\Bigg] + \frac{2L}{\delta}\mathbb{E}\Bigg[\sum_{k=0}^{\infty}\sum_{i=1}^{N} X_t^{N,k}(i) \sum_{m \ne k} X_t^{N,m}(i)\Bigg] \\
&+ 2L\mathbb{E}\Bigg[\sum_{k=0}^{\infty}\sum_{i=1}^{N} \Big(X_t^{N,k}(i) \wedge \delta\Big)\Bigg] + (K+1)\frac{L}{N}
\end{aligned} \tag{4.108}$$

for every $t \ge 0$ and $K, N \in \mathbb{N}$. The last three summands on the right-hand side of (4.108) converge to zero uniformly in $t \in [0,T]$ if we first let $N \to \infty$ and then $\delta \to 0$. This follows from Lemma 4.24 and Lemma 4.23. After inserting (4.108) into (4.104), we see that

$$\limsup_{N\to\infty} \mathbb{E}\Bigg[\sum_{i=1}^{N}\sum_{k=0}^{\infty} \Big(\Delta_t^k(i)\Big)^+\Bigg] \le \int_0^t (1+L) \limsup_{N\to\infty} \mathbb{E}\Bigg[\sum_{i=1}^{N}\sum_{k=0}^{\infty} \Big|\Delta_s^k(i)\Big|\Bigg]\, ds \tag{4.109}$$

for all $t \in [0,T]$. Note that the right-hand side is finite by Lemma 4.6. Similarly we obtain (4.109) with $(\cdot)^+$ replaced by $|\cdot|$. Finally apply Gronwall's inequality to arrive at (4.99).

In the second step, we consider functions σ_N^2 and μ_N which are not globally Lipschitz-continuous. For each $K > 0$ choose function $\sigma_{N,K}^2$ and $\mu_{N,K}$ which agree with σ_N and μ_N, respectively, on $[0,K]$ and which are uniformly globally Lipschitz continuous and uniformly bounded. Existence of such functions follows from the uniform local Lipschitz continuity of σ_N^2 and μ_N. Recall the stopping time $\tilde{\tau}_K^N$, $K,N \in \mathbb{N}$, from (4.8). The process $\{\big(Z_s^{N,m}(i)\big)_{s\le t}\colon i \le N, m \in \mathbb{N}_0\}$ agrees with the loop-free $(N,\sigma_{N,K}^2,\mu_{N,K})$-process on the event $\{\tilde{\tau}_K^N > t\}$. Furthermore the process $\{\big(X_s^{N,m}(i)\big)_{s\le t}\colon i \le N, m \in \mathbb{N}_0\}$ agrees with an $(N,\sigma_{N,K}^2,\mu_{N,K})$-process with migration levels on $\{\tilde{\tau}_K^N > t\}$. Therefore

$$\limsup_{N\to\infty} \mathbb{E}\Big[\sum_{i=1}^N \sum_{k=0}^\infty \Big|X_t^{N,k}(i) - Z_t^{N,k}(i)\Big| \mathbb{1}_{\tilde{\tau}_K^N > t}\Big] = 0 \tag{4.110}$$

for every $K > 0$ by the preceding step. Lemma 4.8 handles the event $\{\tilde{\tau}_K^N \le t\}$. This completes the proof of Lemma 4.25. □

4.6 Proof of Theorem 3.3

Proof of Theorem 3.3. First we prove Theorem 3.3 under the additional Assumption 4.2. This will be relaxed later.

We begin with convergence of finite-dimensional distributions. Recall $\mathcal{E}_{s,T}$ from (4.50). Let $F(\eta) = \prod_{j=1}^n f_j(\eta_{t_j}) \in \mathcal{E}_{s,T}$ satisfy the Lipschitz condition (4.51) with Lipschitz constant $L_F \in (0,\infty)$ and let F be bounded by $C_F < \infty$. Furthermore let the function $f\colon I \to \mathbb{R}$ have compact support in $(0,|I|)$, let f be bounded by $C_f < \infty$ and let f be globally Lipschitz continuous with Lipschitz constant $L_f \in (0,\infty)$. Recall the (N,μ_N,σ_N)-process with migration levels from (4.1). We will exploit below that all individuals on one island have the same migration level in order to show that

$$\lim_{N\to\infty} \mathbb{E}F\bigg(\Big(\sum_{i=1}^N f\big(X_t^N(i)\big)\Big)_{t\le T}\bigg) = \lim_{N\to\infty} \mathbb{E}F\bigg(\Big(\sum_{i=1}^N \sum_{k=0}^\infty f\big(X_t^{N,k}(i)\big)\Big)_{t\le T}\bigg). \tag{4.111}$$

Assuming (4.111) we now prove convergence of finite-dimensional distributions. We may replace the (N,μ_N,σ_N^2)-process with migration levels in (4.111) by the loop-free process because of Lemma 4.25 and the Lipschitz continuity of F and f. Hence (4.111) and Lemma 4.25 imply that

$$\begin{aligned}\lim_{N\to\infty} \mathbb{E}F\bigg(\Big(\sum_{i=1}^N f\big(X_t^N(i)\big)\Big)_{t\le T}\bigg) &= \lim_{N\to\infty} \mathbb{E}F\bigg(\Big(\sum_{i=1}^N \sum_{k=0}^\infty f\big(X_t^{N,k}(i)\big)\Big)_{t\le T}\bigg)\\ = \lim_{N\to\infty} \mathbb{E}F\bigg(\Big(\sum_{i=1}^N \sum_{k=0}^\infty f\big(Z_t^{N,k}(i)\big)\Big)_{t\le T}\bigg) &= \mathbb{E}F\bigg(\Big(\sum_{(s,\eta)\in\mathcal{V}} f\big(\eta_{t-s}\big)\Big)_{t\le T}\bigg).\end{aligned} \tag{4.112}$$

The last equality is the convergence of the loop-free (N,μ_N,σ_N^2)-process to the virgin island model and has been established in Lemma 4.22.

Next we prove (4.111). According to Lemma 4.4 if we ignore the migration levels in the (N,μ_N,σ_N^2)-process with migration levels, then we obtain a version of the (N,μ_N,σ_N^2)-process, that is,

$$\lim_{N\to\infty} \mathbb{E}F\bigg(\Big(\sum_{i=1}^N f\big(X_t^N(i)\big)\Big)_{t\le T}\bigg) = \lim_{N\to\infty} \mathbb{E}F\bigg(\Big(\sum_{i=1}^N f\Big(\sum_{m=0}^\infty X_t^{N,m}(i)\Big)\Big)_{t\le T}\bigg) \tag{4.113}$$

For proving (4.111) we observe that

$$\begin{aligned}\Big|1 - \sum_{k=0}^\infty \mathbb{1}_{x^k \ge \delta}\Big| &\le \mathbb{1}_{x^m \le \delta \ \forall\, m\in\mathbb{N}_0} + \mathbb{1}_{\exists m\ne l\colon x^m\ge\delta, x^l\ge\delta} \cdot \sum_{k=0}^\infty \mathbb{1}_{x^k\ge\delta}\\ &\le \mathbb{1}_{x^m \le \delta \ \forall\, m\in\mathbb{N}_0} + \frac{1}{\delta^2}\sum_{k=0}^\infty x^k \sum_{l\ne k} x^l\end{aligned} \tag{4.114}$$

for every sequence $(x^k)_{k\in\mathbb{N}_0}\subseteq[0,\infty)$ and every $\delta>0$. Thus we get that

$$\begin{aligned}&\mathbb{E}\Big|\sum_{i=1}^N f\Big(\sum_{m=0}^\infty X_t^{N,m}(i)\Big)\Big(1-\sum_{k=0}^\infty \mathbb{1}_{X_t^{N,k}(i)\geq\delta}\Big)\Big|\\&\leq L_f\sum_{i=1}^N\mathbb{E}\Big[\sum_{m=0}^\infty X_t^{N,m}(i)\wedge\delta\Big]+C_f\sum_{i=1}^N\frac{1}{\delta^2}\mathbb{E}\Big[\sum_{k=0}^\infty X_t^{N,k}(i)\sum_{l\neq k}X_t^{N,l}(i)\Big]\\&=:C(N,\delta,t)\end{aligned}\tag{4.115}$$

for all $N\in\mathbb{N}$, $t\geq 0$ and $\delta>0$. The second summand on the right-hand side converges to zero as $N\to\infty$ according to Lemma 4.24. The first summand on the right-hand side converges to zero as $\delta\to 0$ uniformly in $N\in\mathbb{N}$ according to Lemma 4.23. Using (4.115) we obtain that

$$\begin{aligned}&\frac{1}{L_F}\Big|\mathbb{E}F\Big(\Big(\sum_{i=1}^N f\Big(\sum_{m=0}^\infty X_t^{N,m}(i)\Big)\Big)_{t\leq T}\Big)-\mathbb{E}F\Big(\Big(\sum_{i=1}^N\sum_{k=0}^\infty f\big(X_t^{N,k}(i)\big)\Big)_{t\leq T}\Big)\Big|\\&\leq\sum_{j=1}^n\mathbb{E}\Big[\sum_{i=1}^N\sum_{k=0}^\infty\mathbb{1}_{X_{t_j}^{N,k}(i)\geq\delta}\Big|f\Big(\sum_{m=0}^\infty X_{t_j}^{N,m}(i)\Big)-f\Big(X_{t_j}^{N,k}(i)\Big)\Big|\Big]\\&\qquad+\sum_{j=1}^n C(N,\delta,t_j)+\sum_{j=1}^n\mathbb{E}\Big[\sum_{i=1}^N\sum_{k=0}^\infty\mathbb{1}_{X_{t_j}^{N,k}(i)<\delta}\Big|f\Big(X_{t_j}^{N,k}(i)\Big)\Big|\Big]\\&\leq\frac{L_f}{\delta}\sum_{j=1}^n\mathbb{E}\Big[\sum_{i=1}^N\sum_{k=0}^\infty X_{t_j}^{N,k}(i)\sum_{m\neq k}X_{t_j}^{N,m}(i)\Big]\\&\qquad+\sum_{j=1}^n C(N,\delta,t_j)+L_f\sum_{j=1}^n\mathbb{E}\Big[\sum_{i=1}^N\sum_{k=0}^\infty X_{t_j}^{N,k}(i)\wedge\delta\Big]\end{aligned}\tag{4.116}$$

for all $N\in\mathbb{N}$ and $\delta>0$. Letting first $N\to\infty$ and then $\delta\to 0$, the right-hand side converges to zero according to Lemmas 4.24 and 4.23 and according to the preceding step. Inserting this into (4.113) proves (4.111).

The next step is to prove tightness. This is analogous to the tightness proof in Lemma 4.21. Use Lemmas 4.8 and 4.16 instead of Lemma 4.11. So we omit this step.

It remains to prove Theorem 3.3 in the case when Assumption 4.2 fails to hold. Fix $T\in[0,\infty)$. Let $H\colon D_{\mathcal{M}_F(I)}([0,T])\to\mathbb{R}$ be a bounded continuous function on measure-valued càdlàg-paths. It follows from Assumption 3.2 that $\sum_{i=1}^N X_0^N(i)$ converges in L^1 and thus also in distribution to $\sum_{i\in G}X_0(i)$. By the Skorokhod representation theorem (e.g. Theorem II.86.1 in [39]), there exists a version of $\{X_0^N(\cdot)\colon N\in\mathbb{N}\}$ such that $\sum_{i=1}^N X_0^N(i)$ converges almost surely to $\sum_{i\in G}X_0(i)$ as $N\to\infty$. Now the previous step implies that

$$\mathbb{E}\Big[H\Big(\Big(\sum_{i=1}^N\delta_{X_t^N(i)}\Big)_{t\leq T}\Big)|X_0^N(\cdot)\Big]\xrightarrow{N\to\infty}\mathbb{E}\Big[H\Big(\Big(\sum_{(u,\eta)\in\mathcal{V}}\delta_{\eta_{t-u}}\Big)_{t\leq T}\Big)|X_0(\cdot)\Big]\tag{4.117}$$

almost surely. Taking expectations and applying the dominated convergence theorem results in

$$\mathbb{E}\Big[H\Big(\Big(\sum_{i=1}^N\delta_{X_t^N(i)}\Big)_{t\leq T}\Big)\Big]\xrightarrow{N\to\infty}\mathbb{E}\Big[H\Big(\Big(\sum_{(u,\eta)\in\mathcal{V}}\delta_{\eta_{t-u}}\Big)_{t\leq T}\Big)\Big]\tag{4.118}$$

almost surely. This finishes the proof of Theorem 3.3. □

4.7 McKean-Vlasov limit of the N-island model as $N \to \infty$

Proposition 4.29 below establishes convergence of the N-island model as $N \to \infty$ in the case of exchangeable initial configurations. First we prove estimate (4.122) which implies pathwise uniqueness and monotonicity of the solution. The following lemma is a special case of Lemma 4.12. Define $x^+ = \max(x, 0)$ for all $x \in \mathbb{R}$.

Lemma 4.26. *Let Assumption 2.1 be fulfilled, let $x, y \in I$, $s \in [0,\infty)$ and let $\zeta, \bar{\zeta} \colon [s,\infty) \to I$ be locally square Lebesgue integrable. Then there exists a unique strong solution $(Y_{t,s}^{\zeta,x}, \bar{Y}_{t,s}^{\bar{\zeta},y})_{t\in[s,\infty)}$ of*

$$dY_{t,s}^{\zeta,x} = \zeta(t)\,dt - Y_{t,s}^{\zeta,x}dt + \mu\big(Y_{t,s}^{\zeta,x}\big)\,dt + \sqrt{\sigma^2\big(Y_{t,s}^{\zeta,x}\big)}dB_t \tag{4.119}$$

$$d\bar{Y}_{t,s}^{\bar{\zeta},x} = \bar{\zeta}(t)\,dt - \bar{Y}_{t,s}^{\bar{\zeta},x}dt + \mu\big(\bar{Y}_{t,s}^{\bar{\zeta},x}\big)\,dt + \sqrt{\sigma^2\big(\bar{Y}_{t,s}^{\bar{\zeta},x}\big)}dB_t \tag{4.120}$$

having almost surely continuous paths and satisfying $Y_{s,s}^{\zeta,x} = x$, $\bar{Y}_{s,s}^{\bar{\zeta},y} = y$ for all $s \geq 0$. Moreover

$$\mathbb{E}\left[(Y_{t,s}^{\zeta,x} - \bar{Y}_{t,s}^{\bar{\zeta},y})^+\right] \leq e^{L_\mu(t-s)}\left(\int_s^t \big(\zeta(r) - \bar{\zeta}(r)\big)^+\,dr + \big(x-y\big)^+\right) \tag{4.121}$$

for all $t \in [s,\infty)$.

Lemma 4.27. *Let Assumption 2.1 be fulfilled. Moreover let $(M_t)_{t\geq 0}$ and $(\bar{M}_t)_{t\geq 0}$ be two solutions of the McKean-Vlasov equation* (1.4) *with respect to the same Brownian motion having almost surely continuous paths and satisfying $\mathbb{E}[|M_0| + |\bar{M}_0|] < \infty$. Then*

$$\sup_{t\in[0,T]} \mathbb{E}\left[(M_t - \bar{M}_t)^+\right] \leq e^{L_\mu T + Te^{L_\mu T}}\mathbb{E}\left[(M_0 - \bar{M}_0)^+\right] \tag{4.122}$$

for all $T \in [0,\infty)$.

Proof. Applying Lemma 4.26 and Jensen's inequality yields that

$$\begin{aligned}
\mathbb{E}\left[(M_t - \bar{M}_t)^+\right] &= \int\int \mathbb{E}\left[\left(Y_{t,0}^{E[M],x} - Y_{t,0}^{E[\bar{M}],y}\right)^+\right]\mathbb{P}(M_0 \in dx, \bar{M}_0 \in dy)\\
&\leq e^{L_\mu T}\left(\int_0^t \big(\mathbb{E}[M_r] - \mathbb{E}[\bar{M}_r]\big)^+\,dr + \mathbb{E}\big[\big(M_0 - \bar{M}_0\big)^+\big]\right)\\
&\leq e^{L_\mu T}\int_0^t \mathbb{E}\big[\big(M_r - \bar{M}_r\big)^+\big]\,dr + e^{L_\mu T}\mathbb{E}\big[\big(M_0 - \bar{M}_0\big)^+\big]
\end{aligned} \tag{4.123}$$

for all $t \in [0,T]$ and all $T \in [0,\infty)$. Therefore, Gronwall's inequality implies inequality (4.122). ☐

Lemma 4.28. *Let Assumption 2.1 be fulfilled and let M_0 be an I-valued random variable with $\mathbb{E}[|M_0|] < \infty$. Then the McKean-Vlasov equation* (1.4) *has a unique strong solution $(M_t)_{t\geq 0}$.*

Proof. Fix a standard Brownian motion $(B_t)_{t\geq 0}$ which is independent of M_0. Let the processes $(Z^{(k)})_{t\geq 0}$, $k \in \mathbb{N}_0$, be the unique strong solutions of

$$dZ_t^{(k)} = \mathbb{1}_{k\geq 1}\mathbb{E}\big[Z_t^{(k-1)}\big]\,dt - Z_t^{(k)}dt + \mu\big(Z_t^{(k)}\big)\,dt + \sqrt{\sigma^2\big(Z_t^{(k)}\big)}dB_t \tag{4.124}$$

and $Z_0^{(k)} = M_0$ for $k \in \mathbb{N}_0$. We show by induction on $k \in \mathbb{N}_0$ that $Z_t^{(k)} \leq Z_t^{(k+1)}$ for all $t \in [0,\infty)$ and $k \in \mathbb{N}_0$ almost surely. The base case $k = 0$ follows from $\mathbb{E}\big[Z_t^{(0)}\big] \geq 0$ for all $t \in [0,\infty)$ and from a time inhomogeneous version of the monotonicity result in Lemma 3.3 in [23]. For the induction step $k \to k+1$, note that the induction hypothesis implies that $\mathbb{E}\big[Z_t^{(k)}\big] \leq \mathbb{E}\big[Z_t^{(k+1)}\big]$ for all $t \in [0,\infty)$. Thus the induction step follows from a time inhomogeneous version of the monotonicity result in Lemma 3.3 in [23]. Consequently the process $(M_t)_{t\geq 0}$ defined through $M_t =\uparrow \lim_{k\to\infty} Z_t^{(k)}$ for $t \in [0,\infty)$ is a well-defined progressively measurable stochastic process with values in $I \cup \{\infty\}$. Note that Lemma 4.26 implies that

$$\mathbb{E}\big[Z_t^{(k)}\big] \leq e^{L_\mu t}\Big(\mathbb{E}\big[M_0\big] + \int_0^t \mathbb{1}_{k\geq 1}\mathbb{E}\big[Z_s^{(k-1)}\big]\,ds\Big) \leq e^{L_\mu t}\Big(\mathbb{E}\big[M_0\big] + \int_0^t \mathbb{E}\big[Z_s^{(k)}\big]\,ds\Big) \tag{4.125}$$

for all $k \in \mathbb{N}_0$ and all $t \in [0,\infty)$. By induction on $k \in \mathbb{N}_0$, we get that $\mathbb{E}[Z_t^{(k)}]$ is finite for all $k \in \mathbb{N}_0$ and all $t \in [0,\infty)$. Therefore the monotone convergence theorem and Gronwall's lemma result in

$$\mathbb{E}\big[M_t\big] = \mathbb{E}\big[\lim_{k\to\infty} Z_t^{(k)}\big] = \lim_{k\to\infty} \mathbb{E}\big[Z_t^{(k)}\big] \le e^{L_\mu T + Te^{L_\mu T}}\, \mathbb{E}\left[M_0\right] \tag{4.126}$$

for all $t \in [0,T]$ and all $T \in [0,\infty)$. Next we show that $(M_t)_{t\ge 0}$ solves the McKean-Vlasov equation (1.4). Define a stopping time $\tau_K := \inf(\{t \in [0,\infty)\colon M_t \ge K\} \cup \{\infty\}) \in [0,\infty]$ for every $K \in \mathbb{N}$. Doob's L^2 inequality (e.g. Theorem II.70.2 in [39]) implies that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{T\wedge\tau_K} \left|\mu(M_s) - \mu\big(Z_s^{(k)}\big)\right| ds\right] + \mathbb{E}\left[\sup_{t\in[0,T]} \left(\int_0^{t\wedge\tau_K} \sigma(M_s)dB_s - \int_0^{t\wedge\tau_K} \sigma\big(Z_s^{(k)}\big)dB_s\right)^2\right] \\
&\le \int_0^T \mathbb{E}\left[\left|\mu(M_{s\wedge\tau_K}) - \mu\big(Z_{s\wedge\tau_K}^{(k)}\big)\right| ds\right] + 4\int_0^T \mathbb{E}\left[\Big(\sigma\big(M_{s\wedge\tau_K}\big) - \sigma\big(Z_{s\wedge\tau_K}^{(k)}\big)\Big)^2\right] ds \\
&\le \left(\sup_{x\ne y\in[0,K]} \tfrac{|\mu(x)-\mu(y)|}{|x-y|} + 4\sup_{x\ne y\in[0,K]} \tfrac{|\sigma^2(x)-\sigma^2(y)|}{|x-y|}\right) \int_0^T \mathbb{E}\left[\left|M_{s\wedge\tau_K} - Z_{s\wedge\tau_K}^{(k)}\right|\right] ds
\end{aligned} \tag{4.127}$$

for all $T \in [0,\infty)$, $K \in \mathbb{N}$ and all $k \in \mathbb{N}_0$. The right-hand side converges to 0 as $k \to \infty$ for every $K \in \mathbb{N}$ and every $T \in [0,\infty)$ by the dominated convergence theorem. Thus, letting $k \to \infty$ in (4.124) and using (4.127), we conclude that $(M_t)_{t\ge0}$ solves the McKean-Vlasov equation (1.4) for all $t \in [0,\tau_K]$ almost surely for all $K \in \mathbb{N}$, that is, $(M_t)_{t\ge0}$ is a solution of the McKean-Vlasov equation (1.4). This proves existence of a solution.

Applying Lemma 4.27 twice implies that two solutions $(M_t)_{t\ge0}$ and $(\bar{M}_t)_{t\ge0}$ of (1.4) with respect to the same Brownian motion and with the same initial point have the same finite-dimensional distributions. Together with path continuity this yields pathwise uniqueness and – together with weak existence – that the SDE (1.4) is exact (Definition V.9.3 in [40]). Moreover Lemma 4.27 implies almost sure monotonicity and continuity of the solution in the initial point. More precisely if $M_0 \le \bar{M}_0$ almost surely then $M_t \le \bar{M}_t$ almost surely for all $t \in [0,\infty)$ and, due to path continuity, $M_t \le \bar{M}_t$ for all $t \in [0,\infty)$ almost surely. The proof of Theorem V.13.1 in [40] shows that if the solution of an exact SDE is almost surely continuous in the initial point, then there exists a unique strong solution. This finishes the proof of Lemma 4.28. □

Recall the (N,μ,σ^2)-island process $(X_t^N)_{t\ge0}$ from (1.2). Let $\mathcal{M}_1(I)$ be the set of probability measures on I equipped with the Prohorov metric, which induces weak convergence (see e.g. Theorem 3.3.1 in [16]). Moreover let $\mathcal{M}_1^2(I) = \{\nu \in \mathcal{M}_1(I)\colon \langle v, x^2\rangle = \int x^2\nu(dx) < \infty$ be the set of probability measures on I with finite second moments on which a sequence $(\nu_n)_{n\in\mathbb{N}}$ converges to $\nu \in \mathcal{M}_1^2(I)$ if and only if $\nu_n \to \nu$ as $n \to \infty$ in $\mathcal{M}_1(I)$ and $\sup_{n\in\mathbb{N}}\langle \nu_n, x^2\rangle < \infty$. The space $\mathrm{C}([0,\infty),\mathcal{M}_1(I))$ is endowed with the topology of uniform convergence and the subspace $\mathrm{C}([0,\infty),\mathcal{M}_1^2(I))$ is endowed with a topology such that $\nu_n \to \nu$ as $n\to\infty$ in $\mathrm{C}([0,\infty),\mathcal{M}_1^2(I))$ if and only if $\nu_n \to \nu$ as $n \to \infty$ in $\mathrm{C}([0,\infty),\mathcal{M}_1(I))$ and $\sup_{n\in\mathbb{N}}\sup_{t\in[0,T]}\langle \nu_n(t), x^2\rangle < \infty$ for every $T \in [0,\infty)$, for details see Appendix B in [18].

Proposition 4.29. *Let Assumption 2.1 be fulfilled and let M_0 be an I-valued random variable. Moreover let $X_0^N(j)$, $j \le N$, be exchangeable random variables with values in I for every $N \in \mathbb{N}$ such that $\sup_{N\in\mathbb{N}} \mathbb{E}[(X_0^N(1))^2] < \infty$, such that $X_0^N(1) \to M_0$ in distribution as $N \to \infty$ and such that $\frac{1}{N}\sum_{j=1}^N \delta_{X_0^N(j)} \to \mathbb{E}[\delta_{M_0}]$ in distribution in $\mathcal{M}_1^2(I)$ as $N \to \infty$. Let $(X_t^N)_{t\ge0}$ be the unique solution of* (1.2) *and let $(M_t)_{t\ge0}$ be the unique solution of* (1.4)*. Then $(X_t^N(i))_{t\ge0} \to (M_t)_{t\ge0}$ in distribution as $N \to \infty$ for every $i \in \mathbb{N}$ and $\big(\frac{1}{N}\sum_{i=1}^N \delta_{X_t^N(i)}\big)_{t\in[0,\infty)} \to \big(\mathbb{E}\big[\delta_{M_t}\big]\big)_{t\in[0,\infty)}$ in distribution in $\mathrm{C}([0,\infty),\mathcal{M}_1^2)$ as $N \to \infty$. Moreover we have that*

$$\sqrt{N}\mathbb{E}\Big[\big|X_t^N(1) - M_t\big|\Big] \le e^{(1+L_\mu)t}\left(\sqrt{N}\mathbb{E}\Big[\big|X_0^N(1) - M_0\big|\Big] + \int_0^t \big(\mathrm{Var}\left(M_s\right)\big)^{\frac{1}{2}}\, ds\right) \in [0,\infty) \tag{4.128}$$

for all $N \in \mathbb{N}$ and all $t \in [0,\infty)$.

Proof. Theorem 4.1 of Gärtner (1988) establishes an analogous assertion under general assumptions including the ellipticity assumption that (in our notation) $\sigma^2(x) > 0$ for all $x \in I$. This assumption is not satisfied in our situation. Nevertheless, parts of the proof of Theorem 4.1 of Gärtner (1988) carry over to our situation.

We only prove the assertion for $i = 1$, the general case being analogous. First we show that the distributions of the sequence $(\frac{1}{N}\sum_{k=1}^N \delta_{X_t^N(k)}, X_t^N(1))_{t\in[0,\infty)}$, $N \in \mathbb{N}$, are relatively sequentially weakly compact. For this, we introduce more notation. Let $\mathcal{A}(\nu)\colon \mathrm{C}^2(I) \to \mathrm{C}(I)$, $\nu \in \mathcal{M}_1^2(I)$, be operators defined through

$$\mathcal{A}(\nu)f(x) = f^{'}(x)\left(\int_I z\nu(dz) - x + \mu(x)\right) + \tfrac{1}{2}f^{''}(x)\sigma^2(x) \tag{4.129}$$

for all $x \in I$, $f \in \mathrm{C}^2(I)$ and all $\nu \in \mathcal{M}_1^2(I)$ and let the operators $\mathcal{G}^N\colon \mathrm{C}^2(I^N) \to \mathrm{C}(I^N)$, $N \in \mathbb{N}$, be defined through

$$\mathcal{G}^N f(x) = \sum_{k=1}^N \tfrac{\partial}{\partial x_k} f(x)\bigg(\tfrac{1}{N}\sum_{j=1}^N x_j - x_k + \mu(x_k)\bigg) + \tfrac{1}{2}\sum_{k=1}^N \tfrac{\partial^2}{\partial x_k^2} f(x)\sigma^2(x_k) = \sum_{k=1}^N \mathcal{A}_k\bigg(\tfrac{1}{N}\sum_{j=1}^N \delta_{x_j}\bigg)f(x)$$

for all $x \in I^N$, $f \in \mathrm{C}^2(I^N)$ and all $N \in \mathbb{N}$ where $\mathcal{A}_k$ is the operator $\mathcal{A}$ acting on the k-th variable. Note that the N-island process (1.2) solves the well-posed martingale problem for $\mathcal{G}^N$ for every $N \in \mathbb{N}$. The functions μ and σ are continuous so that Assumption (A) of [18] is satisfied except for $\sigma(0) = 0$ and $\sigma(I) = 0$ if $|I| < \infty$. Define functions $\lambda\colon [0,\infty) \ni x \mapsto 2(1 + L_\mu + 2\theta + L_\sigma)(1 + x) \in (0,\infty)$ and $\varphi\colon I \ni x \mapsto 1 + x^2 \in (0,\infty)$. Then $\int_1^\infty dx/\lambda(x) = \infty$ and $\langle \nu, \mathcal{A}(\nu)\varphi\rangle \le \lambda\big(\langle \nu,\varphi\rangle\big)$ for every $\nu \in \mathcal{M}_1^2(I)$ due to Assumption 2.1 so that Assumption (B) of [18] is satisfied. Lemma 4.9 and the inequality $a \le 1 + a^2$ for $a \in \mathbb{R}$ yield that

$$\begin{aligned}
&\lim_{r\to\infty}\sup_{N\in\mathbb{N}} \mathbb{P}\left[\sup_{t\in[0,T]} \tfrac{1}{N}\sum_{k=1}^N \big(X_t^N(k)\big)^2 \ge r\right] \le \lim_{r\to\infty}\frac{1}{r}\sup_{N\in\mathbb{N}} \mathbb{E}\left[\sup_{t\in[0,T]} \tfrac{1}{N}\sum_{k=1}^N \big(X_t^N(k)\big)^2\right]\\
&\le \lim_{r\to\infty}\frac{2}{r}\left(24T\theta^2 + 8L_\sigma T\Big(2\theta T + 1\Big)e^{L_\mu T} + (1 + 8L_\sigma T e^{L_\mu T})\sup_{N\in\mathbb{N}}\mathbb{E}\left[\big(X_0^N(1)\big)^2\right]\right)e^{40(1+T)(1+L_\mu+L_\sigma)^2 T}\\
&= 0
\end{aligned}$$

for every $T \in [0,\infty)$. This implies assumption (i) of Lemma 1.4 of [18]. The proof of assertion (ii) in the proof of Theorem 1.5 of [18] only requires that $\sup_{x\in I\cap[0,r]}(|\mu(x)| + |\sigma(x)|) < \infty$ for every $r \in [0,\infty)$ (which follows from Assumption 2.1) and carries over to our situation without further changes. This implies assumption (ii) of Lemma 1.4 of [18]. Lemma 1.4 of [18] thus yields that the distributions of the sequence $\big(\frac{1}{N}\sum_{k=1}^N \delta_{X_t^N(k)}\big)_{t\in[0,\infty)}$, $N \in \mathbb{N}$, are relatively sequentially weakly compact in $C([0,\infty),\mathcal{M}_1^2(I))$. In addition, the proof of Theorem 4.1 of [18] shows relative weak compactness of $(X_t^N(1))_{t\ge0}$, $N \in \mathbb{N}$, in $\mathrm{C}([0,\infty), I)$ (without using Assumptions (C) or (E) in [18]). This proves that the distributions of the sequence $\big(X_t^N(1), \frac{1}{N}\sum_{k=1}^N \delta_{X_t^N(k)}\big)_{t\in[0,\infty)}$, $N \in \mathbb{N}$, are relatively sequentially weakly compact in $C([0,\infty), I \times \mathcal{M}_1^2(I))$.

Next we identify the weak limit of the sequence $\big(X_t^N(1), \frac{1}{N}\sum_{k=1}^N \delta_{X_t^N(k)}\big)_{t\in[0,\infty)}$, $N \in \mathbb{N}$. Let $(M_t(j))_{t\ge0}$ be the unique strong solution of (1.4) with respect to the Brownian motion $(B_t(j))_{t\ge0}$ for every $j \in \mathbb{N}$ such that $M_0(j)$, $j \in \mathbb{N}$, are independent copies of M_0. As in Theorem 1 of Yamada and Watanabe (1971) [47], an approximation of the function $\mathbb{R} \ni x \to |x| \in [0,\infty)$ with C^2-functions (see also the proof of Lemma 4.25 for this approximation) results in

$$d\big|X_t^N(j) - M_t(j)\big| = \operatorname{sgn}\big(X_t^N(j) - M_t(j)\big)d\big(X_t^N(j) - M_t(j)\big) \tag{4.130}$$

for all $j \in \{1,\dots,N\}$ and all $N \in \mathbb{N}$. Taking expectations and using Assumption 2.1 yields that

$$\begin{aligned}
&\mathbb{E}\bigg[\frac{1}{N}\sum_{j=1}^N \big|X_t^N(j) - M_t(j)\big|\bigg] - \mathbb{E}\bigg[\frac{1}{N}\sum_{j=1}^N \big|X_0^N(j) - M_0(j)\big|\bigg]\\
&\le \frac{1}{N}\sum_{j=1}^N \int_0^t \mathbb{E}\bigg[\Big|\frac{1}{N}\sum_{k=1}^N X_s^N(k) - \mathbb{E}\big[M_s(j)\big]\Big| + \operatorname{sgn}\big(X_s^N(j) - M_s(j)\big)\Big(\mu\big(X_s^N(j)\big) - \mu\big(M_s(j)\big)\Big)\bigg]ds \qquad (4.131)\\
&\le \int_0^t \mathbb{E}\Big[\Big|\frac{1}{N}\sum_{k=1}^N M_s(k) - \mathbb{E}\big[M_s(1)\big]\Big|\Big]ds + (1 + L_\mu)\int_0^t \mathbb{E}\bigg[\frac{1}{N}\sum_{j=1}^N \Big|X_s^N(j) - M_s(j)\Big|\bigg]ds
\end{aligned}$$

for all $N \in \mathbb{N}$ and all $t \in [0,\infty)$. Therefore, Gronwall's inequality implies that

$$\begin{aligned}
\mathbb{E}\Big[\big|X_t^N(1) - M_t(1)\big|\Big] &= \mathbb{E}\Big[\frac{1}{N}\sum_{j=1}^N \big|X_t^N(j) - M_t(j)\big|\Big] \\
&\leq e^{(1+L_\mu)t}\bigg(\mathbb{E}\Big[\frac{1}{N}\sum_{j=1}^N \big|X_0^N(j) - M_0(j)\big|\Big] + \int_0^t \mathbb{E}\Big[\Big|\frac{1}{N}\sum_{k=1}^N M_s(k) - \mathbb{E}\big[M_s(1)\big]\Big|\Big]ds\bigg) \\
&\leq e^{(1+L_\mu)t}\bigg(\mathbb{E}\Big[\big|X_0^N(1) - M_0(1)\big|\Big] + \int_0^t \bigg(\mathrm{Var}\Big(\frac{1}{N}\sum_{k=1}^N M_s(k)\Big)\bigg)^{\frac{1}{2}} ds\bigg) \\
&= e^{(1+L_\mu)t}\bigg(\mathbb{E}\Big[\big|X_0^N(1) - M_0(1)\big|\Big] + \frac{1}{\sqrt{N}}\int_0^t \Big(\mathrm{Var}\big(M_s(1)\big)\Big)^{\frac{1}{2}} ds\bigg)
\end{aligned} \tag{4.132}$$

for all $N \in \mathbb{N}$ and all $t \in [0,\infty)$. This proves inequality (4.128). Moreover we infer that

$$\begin{aligned}
&\bigg|\mathbb{E}\Big[e^{-\sum_{j=1}^n \lambda_j X_{t_j}^N(1) - \sum_{j=1}^n \tilde{\lambda}_j \frac{1}{N}\sum_{k=1}^N f_j\left(X_{t_j}^N(k)\right)}\Big] - \mathbb{E}\Big[e^{-\sum_{j=1}^n \lambda_j M_{t_j}(1) - \sum_{j=1}^n \tilde{\lambda}_j \mathbb{E}\left[f_j\left(M_{t_j}(1)\right)\right]}\Big]\bigg| \\
&\leq \sum_{j=1}^n \lambda_j \mathbb{E}\Big[\Big|X_{t_j}^N(1) - M_{t_j}(1)\Big|\Big] + \sum_{j=1}^n \tilde{\lambda}_j \mathbb{E}\bigg[\Big|\frac{1}{N}\sum_{k=1}^N f_j\big(X_{t_j}^N(k)\big) - \mathbb{E}\Big[f_j\big(M_{t_j}(1)\big)\Big]\Big|\bigg] \\
&\leq \sum_{j=1}^n \Big(\lambda_j + \tilde{\lambda}_j \sup_{x\neq y\in I}\tfrac{|f_j(x)-f_j(y)|}{|x-y|}\Big)\mathbb{E}\Big[\Big|X_{t_j}^N(1) - M_{t_j}(1)\Big|\Big] + \sum_{j=1}^n \tilde{\lambda}_j \tfrac{1}{\sqrt{N}}\Big(\mathrm{Var}\Big(f_j\big(M_{t_j}(1)\big)\Big)\Big)^{\frac{1}{2}}
\end{aligned} \tag{4.133}$$

for all $N \in \mathbb{N}$, $0 \leq t_1 < t_2 < \ldots < t_n < \infty$, $\lambda_1, \ldots, \lambda_n, \tilde{\lambda}_1, \ldots \tilde{\lambda}_n \in [0,\infty)$, all globally Lipschitz continuous functions $f_1, \ldots, f_n \colon I \to [0,\infty)$ and all $n \in \mathbb{N}$. The right-hand side converges to 0 as $N \to \infty$ due to inequality (4.132). This identifies the limit and, together with tightness, implies that the sequence $\big(X_t^N(1), \frac{1}{N}\sum_{k=1}^N \delta_{X_t^N(k)}\big)_{t\in[0,\infty)}$, $N \in \mathbb{N}$, converges to $\big(M_t, \mathbb{E}\big[\delta_{M_t}\big]\big)_{t\in[0,\infty)}$ in distribution in $C([0,\infty), I \times \mathcal{M}_1^2(I))$. This finishes the proof of Proposition 4.29. ☐

5 Comparison with the virgin island model

As in Section 4, we define a loop-free process. Let $\big\{(Z_t^{(k)}(i))_{t\geq 0} \colon k \in \mathbb{N}_0, i \in G\big\}$ be the solution of

$$\begin{aligned}
dZ_t^{(k)}(i) =& \bigg(\sum_{j\in G} Z_t^{(k-1)}(j)m(j,i) - Z_t^{(k)}(i) + \mu\big(Z_t^{(k)}(i)\big)\bigg)\, dt \\
&+ \sqrt{\sigma^2\big(Z_t^{(k)}(i)\big)}\, dB_t^k(i), \quad Z_0^{(k)}(i) = \mathbb{1}_{k=0}X_0(i), \qquad i \in G, k \in \mathbb{N}_0,
\end{aligned} \tag{5.1}$$

where we agree on $Z_t^{(-1)}(i) := 0$ for $t \geq 0$ and $i \in G$. We will refer to this process as the loop-free (G, m, μ, σ^2)-process. The main two steps in our proof of Theorem 3.8 are as follows. Lemma 5.8 below shows that the total mass of the (G, m, μ, σ^2)-process is dominated by the total mass of the loop-free (G, m, μ, σ^2)-process. Lemma 5.10 then proves that the total mass of the loop-free (G, m, μ, σ^2)-process is dominated by the total mass of the virgin island model. Our proof of Lemma 5.10 exploits the hierarchical structure of the loop-free process. Note that conditioned on migration level $k-1$, the islands with migration level k are independent one-dimensional diffusions. We prepare this in Subsection 5.2 by studying the one-dimensional time-inhomogeneous diffusion

$$dY_{t,s}^{\zeta,x}(i) = \zeta(t)\, dt - Y_{t,s}^{\zeta,x}(i)dt + \mu\big(Y_{t,s}^{\zeta,x}(i)\big)\, dt + \sqrt{\sigma^2\big(Y_{t,s}^{\zeta,x}(i)\big)}dB_t(i) \tag{5.2}$$

where $Y_{s,s}^{\zeta,x} = x \in I$ and $s \geq 0$. The path $\zeta \in \mathrm{C}\big([0,\infty), I\big)$ will later represent the mass immigrating from lower migration levels.

The core of the comparison result is the following generator calculation which manifests the intuition that separating mass onto different islands increases the total mass. If $\zeta \equiv c$ is constant, then a formal generator of $(Y_{t,s}^{c,\cdot})_{t\geq s}$ is

$$\mathcal{G}^{(c)}f(x) := \big(c - x + \mu(x)\big)f^{'}(x) + \frac{1}{2}\sigma^2(x)f^{''}(x) \quad x \in I \tag{5.3}$$

where $f \in \mathrm{C}^2(I)$, see e.g. Section 5.3 in [16]. Recall $\mathcal{F}^{(1)}_{++}$ from (3.18).

Lemma 5.1. *Assume 2.1. Suppose that $c, c_1, c_2 \in I$ satisfy $c \le c_1 + c_2$. Assume μ to be subadditive, that is, $\mu(x+y) \le \mu(x) + \mu(y)$ for all $x, y \in I$ with $x + y \in I$. Let $x, y, x+y \in I$. If σ^2 is superadditive, then*

$$\mathcal{G}^{(c)} f(x+y) \le \big(\mathcal{G}^{(c_1)} f(\cdot + y)\big)(x) + \big(\mathcal{G}^{(c_2)} f(x + \cdot)\big)(y) \quad \forall\, f \in \mathcal{F}^{(1)}_{+-} \cap \mathrm{C}^2. \tag{5.4}$$

If σ^2 is subadditive, then

$$\mathcal{G}^{(c)} f(x+y) \le \big(\mathcal{G}^{(c_1)} f(\cdot + y)\big)(x) + \big(\mathcal{G}^{(c_2)} f(x + \cdot)\big)(y) \quad \forall\, f \in \mathcal{F}^{(1)}_{++} \cap \mathrm{C}^2. \tag{5.5}$$

If σ^2 is additive, then

$$\mathcal{G}^{(c)} f(x+y) \le \big(\mathcal{G}^{(c_1)} f(\cdot + y)\big)(x) + \big(\mathcal{G}^{(c_2)} f(x + \cdot)\big)(y) \quad \forall\, f \in \mathcal{F}^{(1)}_{+\pm} \cap \mathrm{C}^2. \tag{5.6}$$

Proof. For $f \in \mathcal{F}^{(1)}_{+-} \cap \mathrm{C}^2$, the first derivative is non-negative and the second derivative is nonpositive. Thus

$$\begin{aligned}
\mathcal{G}^{(c)} f(x+y) &= f^{'}(x+y)\Big(c - (x+y) + \mu(x+y)\Big) + \frac{1}{2} f^{''}(x+y)\sigma^2(x+y) \\
&\le f^{'}(x+y)\Big(c_1 - x + \mu(x)\Big) + \frac{1}{2} f^{''}(x+y)\sigma^2(x) \\
&\quad + f^{'}(x+y)\Big(c_2 - y + \mu(y)\Big) + \frac{1}{2} f^{''}(x+y)\sigma^2(y) \\
&= \big(\mathcal{G}^{(c_1)} f(\cdot + y)\big)(x) + \big(\mathcal{G}^{(c_2)} f(x + \cdot)\big)(y).
\end{aligned} \tag{5.7}$$

This is inequality (5.4). The proof of inequality (5.5) is analogous. If σ^2 is additive, then $\sigma^2(x+y) = \sigma^2(x) + \sigma^2(y)$ and no property of $f^{''}$ is needed in the above calculation. □

As a remark, we observe that the operator on the right-hand side of (5.4) is a formal generator of the superposition $\big(Y^{c_1,x}_{t,s} + \tilde{Y}^{c_2,y}_{t,s}\big)_{t \ge s}$ of two independent solutions of (5.2). This follows from Theorem 4.10.1 in [16].

We will lift inequality (5.4) between formal generators to an inequality between the associated semigroups. For this we use an integration by parts formula. For its formulation, let $\mathcal{G}_S$ and $\mathcal{G}_T$ be two generators associated with the semigroups $(S_t)_{t \ge 0}$ and $(T_t)_{t \ge 0}$, respectively. Then, for $t \in [0, \infty)$, we have that

$$S_t f - T_t f = \int_0^t T_{t-s}\big(\mathcal{G}_S - \mathcal{G}_T\big) S_s f \, ds \tag{5.8}$$

if $S_s f \in \mathcal{D}(\mathcal{G}_S) \cap \mathcal{D}(\mathcal{G}_T)$ for all $s \le t$, see p. 367 in Liggett (1985). The idea of using (5.8) for a comparison is borrowed from Cox et al. (1996). As the generator inequality (5.4) holds for functions in $\mathcal{F}^{(1)}_{+-}$, we need to show that the semigroup of $(Y^{x,\cdot}_{t,s})_{t \ge s}$ preserves $\mathcal{F}^{(1)}_{+-}$. This is subject of the following subsection.

5.1 Preservation of convexity

We write $\underline{x}_n := (x_1, \ldots, x_n)$ for $x_1, \ldots, x_n \in \mathbb{R}$ and $n \in \mathbb{N}$. The i-th unit row vector is denoted as e_i for every $i \in \mathbb{N}$. Recall $\mathcal{F}^{(n)}_{++}$ from (3.18).

Lemma 5.2. *For every $n \in \mathbb{N}$ and $f \in \mathcal{F}^{(n)}_{++}\big([0,\infty)\big)$, we have that*

$$f\Big(\underline{x}_n + \sum_{j=1}^k h_j e_{i_j}\Big) - f\Big(\underline{x}_n + \sum_{j=2}^k h_j e_{i_j}\Big) \ge f\Big(\underline{x}_n + h_1 e_{i_1}\Big) - f\big(\underline{x}_n\big) \tag{5.9}$$

for all $\underline{x}_n \in [0,\infty)^n$, all $h_1, \ldots, h_k \ge 0$, all $(i_1, \ldots, i_k) \in \{1, \ldots, n\}^k$ and all $k \in \mathbb{N}$. The reverse inequality holds if $\mathcal{F}^{(n)}_{++}$ is replaced by $\mathcal{F}^{(n)}_{+-}$.

Proof. The proof is by induction on $k \in \mathbb{N}$. The base case $k=1$ is trivial. Now assume that (5.9) holds for some $k \geq 1$. Fix $\underline{x}_n \in [0,\infty)^n$, $h_1,\ldots,h_k \geq 0$ and $(i_1,\ldots,i_k) \in \{1,\ldots,n\}^k$. Applying the induction hypothesis at location $\underline{x}_n + \sum_{j=2}^k h_j e_{i_j}$ to the index tuple (i_1, i_{k+1}), we obtain that

$$\begin{aligned} & f\Big(\underline{x}_n + \sum_{j=2}^k h_j e_{i_j} + h_1 e_{i_1} + h_{k+1} e_{i_{k+1}}\Big) - f\Big(\underline{x}_n + \sum_{j=2}^k h_j e_{i_j} + h_{k+1} e_{i_{k+1}}\Big) \\ & \geq f\Big(\underline{x}_n + \sum_{j=2}^k h_j e_{i_j} + h_1 e_{i_1}\Big) - f\Big(\underline{x}_n + \sum_{j=2}^k h_j e_{i_j}\Big) \geq f\Big(\underline{x}_n + h_1 e_{i_1}\Big) - f\big(\underline{x}_n\big). \end{aligned} \tag{5.10}$$

The last step is again the induction hypothesis. □

Lemma 5.3. *Let $n \in \mathbb{N}$, $c \in [0,\infty)$ and $f \in \mathcal{F}_{++}^{(n+1)}\big([0,\infty)\big)$. Then the two functions*

$$\tilde{f}\colon [0,\infty)^n \to \mathbb{R}, \underline{x}_n \to f(x_1,\ldots,x_n,x_n) \tag{5.11}$$

and

$$\bar{f}\colon [0,\infty)^n \to \mathbb{R}, \underline{x}_n \to f(x_1,\ldots,x_n,c) \tag{5.12}$$

are elements of $\mathcal{F}_{++}^{(n)}\big([0,\infty)\big)$. This is also true if $\mathcal{F}_{++}$ is replaced by $\mathcal{F}_{+-}$ and $\mathcal{F}_{+\pm}$, respectively.

Proof. The functions $\tilde{f}$ and $\bar{f}$ are non-decreasing and either bounded or non-negative. It is clear that $\bar{f}$ is again (i,j)-convex for $1 \leq i,j \leq n$ and that $\tilde{f}$ is (i,j)-convex for $1 \leq i,j \leq n-1$. It remains to prove (i,n)-convexity of $\tilde{f}$ for $1 \leq i \leq n$. Applying Lemma 5.2 at location $\underline{x}_{n+1}$ to the index tuple $(i,n,n+1)$, we obtain for all $h_1,h_2 \geq 0$ that

$$\begin{aligned} & f\big(\underline{x}_n + h_1 e_i + h_2 e_n, x_n + h_2\big) = f\big(\underline{x}_{n+1} + h_1 e_i + h_2 e_n + h_2 e_{n+1}\big)|_{x_{n+1}=x_n} \\ & \geq \Big(f\big(\underline{x}_{n+1} + h_2 e_n + h_2 e_{n+1}\big) + f\big(\underline{x}_{n+1} + h_1 e_i\big) - f\big(\underline{x}_{n+1}\big)\Big)|_{x_{n+1}=x_n} \\ & = f\big(\underline{x}_n + h_2 e_n, x_n + h_2\big) + f\big(\underline{x}_n + h_1 e_i, x_n\big) - f\big(\underline{x}_n, x_n\big), \end{aligned} \tag{5.13}$$

that is, $\tilde{f}$ is (i,n)-convex. □

Lemma 5.4. *Assume 2.1. Let $n \in \mathbb{N}$ and $f \in \mathcal{F}_{+\pm}^{(n+1)}\big([0,\infty)\big)$. Then the function*

$$(x_1,\ldots,x_n) \mapsto \mathbb{E} f\big(x_1,\ldots,x_n,Y_{t,s}^{\zeta,x_n}\big) \tag{5.14}$$

is an element of $\mathcal{F}_{+\pm}^{(n)}\big([0,\infty)\big)$ for every $0 \leq s \leq t$. If μ is concave, then this property still holds if $\mathcal{F}_{+\pm}^{(n+1)}$ is replaced by $\mathcal{F}_{++}^{(n+1)}$ and if $\mathcal{F}_{+\pm}^{(n)}$ is replaced by $\mathcal{F}_{++}^{(n)}$, respectively.

Proof. Fix $0 \leq s \leq t$ and $n \in \mathbb{N}$. We only prove the case of μ being concave and $f \in \mathcal{F}_{++}^{(n+1)}$ as the remaining cases are similar. According to Lemma 5.3, it suffices to prove that

$$\tilde{f}\colon [0,\infty)^{n+1} \to \mathbb{R}, (x_1,\ldots,x_{n+1}) \mapsto \mathbb{E} f\big(x_1,\ldots,x_n,Y_{t,s}^{\zeta,x_{n+1}}\big) \tag{5.15}$$

is an element of $\mathcal{F}_{++}^{(n+1)}\big([0,\infty)\big)$. Let $(Y_{t,s}^{\zeta,x})_{t\geq s}$, $x \in I$, be solutions of (5.2) with respect to the same Brownian motion. It is known that $Y_{t,s}^{\zeta,x} \leq Y_{t,s}^{\zeta,x+h}$ holds almost surely for all $x \leq x+h \in I$, see e.g. Theorem V.43.1 in [40] for the time-homogeneous case. Thus the function $\tilde{f}$ is again non-decreasing. Moreover $\tilde{f}$ inherits (i,j)-convexity from f for every $1 \leq i,j \leq n$. It remains to show that $\tilde{f}$ is $(i,n+1)$-convex for $1 \leq i \leq n+1$. If $i \leq n$, then $(i,n+1)$-convexity of f at the point $\big(x_1,\ldots,x_n,Y_{t,s}^{\zeta,x_{n+1}}\big)$ implies that

$$\begin{aligned} & \tilde{f}\big(\underline{x}_{n+1} + h_1 e_i + h_2 e_{n+1}\big) = \mathbb{E} f\big(\underline{x}_n + h_1 e_i, Y_{t,s}^{\zeta,x_{n+1}} + Y_{t,s}^{\zeta,x_{n+1}+h_2} - Y_{t,s}^{\zeta,x_{n+1}}\big) \\ & \geq \mathbb{E} f\big(\underline{x}_n + h_1 e_i, Y_{t,s}^{\zeta,x_{n+1}}\big) + \mathbb{E} f\big(\underline{x}_n, Y_{t,s}^{\zeta,x_{n+1}+h_2}\big) - \mathbb{E} f\big(\underline{x}_n, Y_{t,s}^{\zeta,x_{n+1}}\big) \\ & = \tilde{f}\big(\underline{x}_{n+1} + h_1 e_i\big) + \tilde{f}\big(\underline{x}_{n+1} + h_2 e_{n+1}\big) - \tilde{f}\big(\underline{x}_{n+1}\big) \end{aligned} \tag{5.16}$$

for every $h_1, h_2 \geq 0$ and $(x_1, \ldots, x_n) \in [0,\infty)^n$, that is, $(i, n+1)$-convexity of $\tilde{f}$ in the case $i \leq n$. One can establish convexity of

$$y \mapsto \tilde{f}(x_1, \ldots, x_n, y) \tag{5.17}$$

as in Lemma 6.1 of [23] (this Lemma 6.1 shows concavity if f is $(n+1, n+1)$-concave and smooth; for the general case, approximate f, μ and σ with smooth functions and exploit that convexity is preserved under pointwise limits). This step uses concavity of μ. Consequently, $\tilde{f}$ is $(n+1, n+1)$-convex. This completes the proof of $\tilde{f} \in \mathcal{F}_{++}^{(n+1)}\big([0,\infty)\big)$. □

Lemma 5.4 extends Proposition 16 of Cox et al. (1996) [11]. This Proposition 16 is used in [11] to establish a comparison result between diffusions with different diffusion functions, see Theorem 1 in [11]. Using the above Lemma 5.4, this comparison result can be extended to more general test functions.

5.2 Decomposition of a one-dimensional diffusion with immigration into subfamilies

Feller's branching diffusion with immigration can be decomposed into independent families which originate either from an individual at time zero or from an immigrant, see e.g. Theorem 1.3 in Li and Shiga (1995). A diffusion does in general not agree with its family decomposition if individuals interact with each other, e.g. if the branching rate depends on the population size. If the drift function is subadditive and if the branching function is superadditive, however, then we get at least a comparison result. In that situation, the diffusion is dominated by its family decomposition. More precisely, the total mass increases in the order $\leq_{\mathbb{F}_{+-}}$ if we let all subfamilies evolve independently, see Lemma 5.7 below. The following lemma is a first step in this direction.

Lemma 5.5. *Assume 2.1. Let $x, y, x+y \in I$ and let $\zeta, \tilde{\zeta}\colon [0,\infty) \to [0,\infty)$ be locally Lebesgue integrable. If μ is concave and σ^2 is superadditive, then*

$$\left(Y_{t,s}^{\zeta+\tilde{\zeta},x+y}\right)_{t\geq s} \leq_{\mathbb{F}_{+-}([0,\infty))} \left(Y_{t,s}^{\zeta,x} + \tilde{Y}_{t,s}^{\tilde{\zeta},y}\right)_{t\geq s}, \qquad \forall\, s \geq 0, \tag{5.18}$$

where $\left(Y_{t,s}^{\zeta,x}\right)_{t\geq s}$ and $\left(\tilde{Y}_{t,s}^{\tilde{\zeta},x}\right)_{t\geq s}$ are independent processes. If μ is concave and σ^2 is subadditive, then inequality (5.18) *holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{++}$. If μ is subadditive and σ^2 is additive, then inequality* (5.18) *holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{+\pm}$.*

Proof. Let $F_n(\eta) = f_n(\eta_{t_1}, \ldots, \eta_{t_n}) \in \mathbb{F}_{+-}([0,\infty))$ where $f_n \in \mathcal{F}_{+-}^{(n)}([0,\infty))$. We begin with the case of $\zeta, \tilde{\zeta}$ being simple functions. W.l.o.g. we consider $\zeta(t) = \sum_{i=1}^n c_i \mathbb{1}_{[t_{i-1},t_i)}(t)$ and $\tilde{\zeta}(t) = \sum_{i=1}^n \tilde{c}_i \mathbb{1}_{[t_{i-1},t_i)}(t)$ where $c_1, ..., c_n, \tilde{c}_1, \ldots, \tilde{c}_n \geq 0$, $t_0 = s$ and $t_{n+1} = \infty$ as we may let F_n depend trivially on further time points. We will prove by induction on $n \in \mathbb{N}$ that

$$\mathbb{E}F_n\Big(\big(Y_{t,s}^{\zeta+\tilde{\zeta},x+y}\big)_{t\geq s}\Big) \leq \mathbb{E}F_n\Big(\big(Y_{t,s}^{\zeta,x} + \tilde{Y}_{t,s}^{\tilde{\zeta},y}\big)_{t\geq s}\Big). \tag{5.19}$$

For the base case $n = 1$ additionally assume $f_1 \in \mathbf{C}^2$. Approximate σ and μ with functions $\sigma_l, \mu_l \in \mathbf{C}^\infty(\mathbb{R})$ having the following properties. All derivatives $\sigma_l^{(k)}, \mu_l^{(k)}$, $k \in \mathbb{N}_0$, are bounded, μ_l is concave and σ_l^2 is superadditive. Both functions vanish at zero. If $|I| < \infty$, then $\mu_l(|I|) \leq 0 = \sigma_l^2(|I|)$. Moreover $\mu_l(x) \to \mu(x)$ and $\sigma_l^2(x) \to \sigma^2(x)$ as $l \to \infty$ for all $x \in I$. Let $(Y_{t,s}^{\zeta,x,l})_{t\geq s}$, $x \in I$, be solutions of (5.2) with σ^2 and μ replaced by σ_l^2 and μ_l, respectively, and let $(\tilde{Y}_{t,s}^{\tilde{\zeta},y,l})_{t\geq s}$, $y \in I$, be an independent version hereof. Then $x \mapsto S_t f_1(x) := \mathbb{E}f_1\big(Y_{t,s}^{c_1,x,l}\big)$ is twice continuously differentiable for every $t \geq s$, see Theorem 8.4.3 in Gikhman and Skorokhod (1969). In addition, Lemma 5.4 proves $S_t f_1 \in \mathcal{F}_{+-}^{(1)}$ for all $t \in [s, t_1]$. Consequently, we may apply Lemma 5.1 to $S_t f_1 \in \mathcal{F}_{+-}^{(1)} \cap \mathrm{C}^2$ for every $t \in [s, t_1]$ and the integration by parts formula (5.8) yields that

$$\mathbb{E}\Big[f_1\big(Y_{t_1}^{c_1+\tilde{c}_1,x+y,l}\big)\Big] \leq \mathbb{E}\Big[f_1\big(Y_{t_1}^{c_1,x,l} + \tilde{Y}_{t_1}^{\tilde{c}_1,y,l}\big)\Big]. \tag{5.20}$$

Now as $l \to \infty$, $(Y_t^{c_1,x,l})_{t\geq 0}$ converges weakly to $(Y_t^{c_1,x})_{t\geq 0}$ for every $x \in I$, see Lemma 19 in Cox et al. (1996) for a sketch of the proof. Therefore letting $l \to \infty$ in (5.20) proves (5.19) for $n = 1$ if $f_1 \in \mathrm{C}^2$. The case of general $f_1 \in \mathcal{F}_{+-}^{(1)}$ follows by approximating f_1 with smooth functions in $\mathcal{F}_{+-}^{(1)}$. For the induction step $n \to n+1$, define

$$\tilde{f}_n(y_1, \ldots, y_n) := \mathbb{E} f_{n+1}\Big(y_1, \ldots, y_n, Y_{t_{n+1}, t_n}^{\zeta+\tilde{\zeta}, y_n}\Big) \qquad \forall\ (y_1, \ldots, y_n) \in I^n. \tag{5.21}$$

Note that the induction hypothesis implies that

$$\tilde{f}_n(x_1 + y_1, \ldots, x_n + y_n) \le \mathbb{E} f_{n+1}\Big(x_1 + y_1, \ldots, x_n + y_n, Y_{t_{n+1}, t_n}^{\zeta, x_n} + \tilde{Y}_{t_{n+1}, t_n}^{\tilde{\zeta}, y_n}\Big) \tag{5.22}$$

and that Lemma 5.4 implies that $\tilde{f}_n \in \mathcal{F}_{+-}^{(n)}$. Therefore, using the Markov property and the induction hypothesis, we get that

$$\begin{aligned}
&\mathbb{E} F_{n+1}\Big(Y_{\cdot, s}^{\zeta+\tilde{\zeta}, x+y}\Big) = \mathbb{E} \tilde{f}_n\Big(Y_{t_1, s}^{\zeta+\tilde{\zeta}, x+y}, \ldots, Y_{t_n, s}^{\zeta+\tilde{\zeta}, x+y}\Big) \\
&\le \mathbb{E} \tilde{f}_n\Big(Y_{t_1, s}^{\zeta, x} + \tilde{Y}_{t_1, s}^{\tilde{\zeta}, y}, \ldots, Y_{t_n, s}^{\zeta, x} + \tilde{Y}_{t_n, s}^{\tilde{\zeta}, y}\Big) \\
&\le \mathbb{E}\bigg[\mathbb{E}\bigg[f_{n+1}\Big(x_1 + y_1, \ldots, x_n + y_n, Y_{t_{n+1}, t_n}^{\zeta, x_n} + \tilde{Y}_{t_{n+1}, t_n}^{\tilde{\zeta}, y_n}\Big)\Big|_{x_i = Y_{t_i, s}^{\zeta, x},\ y_i = \tilde{Y}_{t_i, s}^{\tilde{\zeta}, y}}\bigg]\bigg] \\
&= \mathbb{E} F_{n+1}\bigg(\Big(Y_{t, s}^{\zeta, x} + \tilde{Y}_{t, s}^{\tilde{\zeta}, y}\Big)_{t \ge s}\bigg)
\end{aligned} \tag{5.23}$$

for all $x, y \in I$ satisfying $x + y \in I$. The last step follows from the Markov property and from independence of the two processes. This proves (5.19).

In case of general functions ζ and $\tilde{\zeta}$, approximate ζ and $\tilde{\zeta}$ with simple functions ζ_l and $\tilde{\zeta}_l$, $l \in \mathbb{N}$, respectively. The process $Y_{\cdot, s}^{\zeta_l, x}$ converges in the sense of finite-dimensional distributions in L^1, see Lemma 4.10, and due to tightness also weakly to the process $Y_{\cdot, s}^{\zeta, x}$. This completes the proof as the remaining cases are analogous. ☐

Lemma 5.6. *Assume 2.1 and 2.2. Then we have that*

$$\bigg(\sum_{i=1}^{N} \delta_{Y_{t, s}^{0, \frac{x}{N}}(i)}\bigg)_{s \le t \le T} \xrightarrow{w} \bigg(\int_0^x \int \delta_{\eta_{t-s}} \Pi(dy, d\eta)\bigg)_{s \le t \le T} \qquad \text{as } N \to \infty \tag{5.24}$$

for all $x \ge 0$ and all $s \le T$ where Π is a Poisson point process on $[0, \infty) \times U$ with intensity measure $\mathrm{Leb} \otimes Q$.

Proof. The proof is analogous to the proof of Lemma 4.21. For convergence of finite-dimensional distributions use the convergence (2.4) instead of Lemma 4.21. Tightness follows from an estimate as in (4.67) together with boundedness (see Lemma 9.9 in [21]) of second moments. ☐

Finally we prove the main result of this subsection. The following lemma shows that the total mass increases if we let all subfamilies evolve independently. In the special case of μ and σ^2 being linear, inequality (5.25) is actually an equality according to the classical family decomposition of Feller's branching diffusion with immigration.

Lemma 5.7. *Assume 2.1. Let $\zeta\colon [0, \infty) \to [0, \infty)$ be locally Lebesgue integrable and let $x \in I$. If the drift function μ is concave and the diffusion function σ^2 is superadditive, then*

$$\Big(Y_{t, s}^{\zeta, x}\Big)_{t \ge s} \le_{\mathbb{F}_{+-}([0, \infty))} \bigg(\int_0^x \int \eta_{t-s} \Pi(dy, d\eta) + \int_s^\infty \int \eta_{t-u} \tilde{\Pi}(du, d\eta)\bigg)_{t \ge s} \tag{5.25}$$

for every $s \ge 0$ where Π is a Poisson point process on $[0, \infty) \times U$ with intensity measure $\mathrm{Leb} \otimes Q$ and where $\tilde{\Pi}$ is an independent Poisson point process on $[0, \infty) \times U$ with intensity measure $\zeta(s)\, ds \otimes Q$. If μ is concave and σ^2 is subadditive, then (5.25) holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{++}$. If μ is subadditive and σ^2 is additive, then (5.25) holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{+\pm}$.

Proof. The idea is to split the initial mass and the immigrating mass into smaller and smaller pieces. Fix $s \geq 0$. Let μ be concave and let σ^2 be superadditive. According to Lemma 5.5

$$\left(Y_{t,s}^{\zeta,x}\right)_{t\geq s} \leq_{\mathbb{F}_{+-}([0,\infty))} \left(\sum_{i=1}^{N} Y_{t,s}^{0,\frac{x}{N}}(i) + \sum_{i=1}^{N} \tilde{Y}_{t,s}^{\frac{\zeta}{N},0}(i)\right)_{t\geq s} \tag{5.26}$$

for every $N \in \mathbb{N}$ where all processes are independent of each other. Letting $N \to \infty$ in (5.26), the right-hand side of (5.26) converges to the right-hand side of (5.25), see Lemma 4.21 and Lemma 5.6. The remaining cases are analogous. □

5.3 The (G, m, μ, σ^2)-process is dominated by the loop-free (G, m, μ, σ^2)-process

Lemma 5.8. *Assume 2.1. If μ is concave and σ^2 is superadditive, then*

$$\left(X_t\right)_{t\geq 0} \leq_{\mathbb{F}_{+-}(G,[0,\infty))} \left(\sum_{k=0}^{\infty} Z_t^{(k)}\right)_{t\geq 0}. \tag{5.27}$$

If μ is concave and σ^2 is subadditive, then inequality (5.27) *holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{++}$. If μ is subadditive and σ^2 is additive, then inequality* (5.27) *holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{+\pm}$.*

Proof. Assume that μ is subadditive and that σ^2 is superadditive. We follow the proof of Lemma 5.5 and begin with a generator calculation similar to Lemma 5.1. Let $\mathcal{G}^X$ and $\mathcal{G}^Z$ denote the formal generators of (G, m, μ, σ^2)-process and of the loop-free (G, m, μ, σ^2)-process, respectively. Assume that $f \in \mathrm{C}^2\big([0,\infty)^G\big) \cap \mathcal{F}_{+-}^{(1)}\big([0,\infty)^G\big)$ depends only on finitely many coordinates. Associated with this test function is

$$\tilde{f}\Big((x_i^{(k)})_{i\in G,k\in\mathbb{N}_0}\Big) := f\left(\sum_{k=0}^{\infty} x_\cdot^{(k)}\right) \tag{5.28}$$

where $(x_i^{(k)})_{i\in G,k\in\mathbb{N}_0} \in I^{G\times\mathbb{N}_0}$. Note that $\tilde{f} \in \mathfrak{D}(\mathcal{G}^Z) = \mathrm{C}^2\big(I^{G\times\mathbb{N}_0}\big)$. The first partial derivatives $f_i := \big(\frac{\partial}{\partial x_i}\big) f\big(\sum_{k=0}^{\infty} x_\cdot^{(k)}\big)$, $i \in G$, are non-negative and the second partial derivatives

$$f_{ii} := \Big(\tfrac{\partial^2}{\partial x_i^2}\Big) f\left(\sum_{k=0}^{\infty} x_\cdot^{(k)}\right), \qquad i \in G, \tag{5.29}$$

are nonpositive. Thus we see that

$$\begin{aligned}
&\Big(\mathcal{G}^X f\Big)\Big(\sum_{k=0}^{\infty} x_\cdot^{(k)}\Big)\\
&= \sum_{i\in G} f_i \Big[\sum_{j\in G}\Big(\sum_{k=0}^{\infty} x_j^{(k)} m(j,i) - \sum_{k=0}^{\infty} x_i^{(k)} + \mu\Big(\sum_{k=0}^{\infty} x_i^{(k)}\Big)\Big)\Big] + \frac{1}{2}\sum_{i\in G} f_{ii}\cdot\sigma^2\Big(\sum_{k=0}^{\infty} x_i^{(k)}\Big)\\
&\leq \sum_{k=0}^{\infty}\Big[\sum_{i\in G} f_i\Big(\sum_{j\in G} x_j^{(k)} m(j,i) - x_i^{(k)} + \mu\big(x_i^{(k)}\big)\Big) + \frac{1}{2}\sum_{i\in G} f_{ii}\cdot\sigma^2\big(x_i^{(k)}\big)\Big]\\
&= \sum_{k=0}^{\infty}\Big[\sum_{i\in G} f_i\Big(\sum_{j\in G} x_j^{(k-1)} m(j,i) - x_i^{(k)} + \mu\big(x_i^{(k)}\big)\Big) + \frac{1}{2}\sum_{i\in G} f_{ii}\cdot\sigma^2\big(x_i^{(k)}\big)\Big]\\
&= \mathcal{G}^Z \tilde{f}\Big(\big(x_i^{(k)}\big)_{i\in G,k\in\mathbb{N}_0}\Big)
\end{aligned} \tag{5.30}$$

for every $(x_i^{(k)})_{i\in G,k\in\mathbb{N}_0} \in I^{G\times\mathbb{N}_0}$. Now we wish to apply the integration by parts formula (5.8). In order to guarantee $x \mapsto \mathbb{E}^x f(X_s) \in \mathfrak{D}(\mathcal{G}^X)$, approximate f with smooth functions in $\mathcal{F}_{+-}^{(1)}$, approximate μ and σ as in the proof of Lemma 5.5 and approximate G with finite sets. Moreover in order to exploit the generator inequality (5.30) in the integration by parts formula (5.8), we note that $x \mapsto \mathbb{E}^x f(X_s) \in \mathcal{F}_{+-}^{(1)}$ (see Lemma 6.1 in [23]). Therefore the integration by parts formula (5.8) together with inequality (5.30) implies that

$$\mathbb{E}^{\sum_{k=0}^{\infty} x_\cdot^{(k)}}\Big(f\big(X_t\big)\Big) \leq \mathbb{E}^{x_\cdot^{(\cdot)}}\left(f\Big(\sum_{k=0}^{\infty} Z_t^{(k)}\Big)\right) \qquad \forall\, f \in \mathcal{F}_{+-}^{(1)}\big([0,\infty)^G\big) \tag{5.31}$$

for all $(x_i^{(k)})_{i\in G,k\in\mathbb{N}_0} \in I^{G\times\mathbb{N}_0}$. In addition note that $(X_t)_{t\geq 0}$ is stochastically non-decreasing in its initial configuration, see Lemma 3.3 in [23]. As stochastic monotonicity is the only input into the proof of Lemma 5.4, the assertion of Lemma 5.4 holds also for $(X_t)_{t\geq 0}$. Thus, for all $f \in \mathcal{F}_{+-}^{(n+1)}\big([0,\infty)^G\big)$, we have that

$$\big(I^G\big)^n \ni (x_1,\ldots,x_n) \mapsto \mathbb{E}^{x_n} f\big(x_1,\ldots,x_n,X_t\big) \in \mathcal{F}_{+-}^{(n)}\big(I^G\big). \tag{5.32}$$

Using this, the assertion follows as in the proof of Lemma 5.5 by induction on the number of arguments of $F \in \mathbb{F}_{+-}\big([0,\infty)^G\big)$. □

5.4 The loop-free process is dominated by the virgin island model

We show in Lemma 5.10 below that the total mass of the loop-free process is dominated by the total mass of the virgin island model. In the proof of this lemma, we use that the Poisson point processes appearing in the definition of the virgin island model preserve convexity in a suitable way. This is subject of the following lemma.

Lemma 5.9. *For every vector $\underline{z} = (z_1,\ldots,z_m) \in [0,\infty)^m$, $m \in \mathbb{N}$, let $\Pi^{(\underline{z})}$ be a Poisson point process on a Polish space S with intensity measure $\sum_{i=1}^m z_i\mu_i$ where $\mu_1,\ldots,\mu_m$ are fixed measures on S. If $f \in \mathcal{F}_{+-}^{(n+1)}\big([0,\infty)\big)$, $n \in \mathbb{N}$, then the function*

$$\tilde{f}\colon [0,\infty)^{n+m} \to \mathbb{R},\ (\underline{x},\underline{z}) \mapsto \mathbb{E} f\big(\underline{x},\langle g,\Pi^{(\underline{z})}\rangle\big) \tag{5.33}$$

is an element of $\mathcal{F}_{+-}^{(n+m)}\big([0,\infty)\big)$ for every measurable test function $g\colon S \to [0,\infty)$ satisfying $\langle g,\mu_i\rangle = \int g\,d\mu_i < \infty$ for $i = 1,\ldots,m$. Analogous results hold if $\mathcal{F}_{+-}$ is replaced by $\mathcal{F}_{++}$ and $\mathcal{F}_{+\pm}$, respectively.

Proof. The function $\tilde{f}$ is non-decreasing in the first n variables and (i,j)-concave for all $1 \leq i,j \leq n$. Furthermore $\tilde{f}$ is non-decreasing in the last m variables as $\Pi^{(\underline{z})}$ is stochastically non-decreasing in $\underline{z}$. Fix $1 \leq i,j \leq m$ and $h_1,h_2 \geq 0$. Let the Poisson point processes $\Pi^{(\underline{z})}$, $\Pi^{(h_1e_i)}$ and $\Pi^{(h_2e_j)}$ be independent. Fix $\underline{z} \in [0,\infty)^m$, $\underline{x} \in [0,\infty)^n$ and a measurable test function $g\colon S \to [0,\infty)$. Note that $\Pi^{(\underline{z})} + \Pi^{(h_1e_i)} + \Pi^{(h_2e_j)}$ has the same distribution as $\Pi^{(\underline{z}+h_1e_i+h_2e_j)}$. Therefore, using $(n+1,n+1)$-concavity of f in the point $(\underline{x},\langle g,\Pi^{(\underline{z})}\rangle)$, we obtain that

$$\begin{aligned}
&\mathbb{E}f\Big(\underline{x},\big\langle g,\Pi^{(\underline{z}+h_1e_i+h_2e_j)}\big\rangle\Big) + \mathbb{E}f\Big(\underline{x},\big\langle g,\Pi^{(\underline{z})}\big\rangle\Big)\\
&= \mathbb{E}f\Big(\underline{x},\big\langle g,\Pi^{(\underline{z})}\big\rangle + \big\langle g,\Pi^{(h_1e_i)}\big\rangle + \big\langle g,\Pi^{(h_2e_j)}\big\rangle\Big) + \mathbb{E}f\Big(\underline{x},\big\langle g,\Pi^{(\underline{z})}\big\rangle\Big)\\
&\leq \mathbb{E}f\Big(\underline{x},\big\langle g,\Pi^{(\underline{z})}\big\rangle + \big\langle g,\Pi^{(h_1e_i)}\big\rangle\Big) + \mathbb{E}f\Big(\underline{x},\big\langle g,\Pi^{(\underline{z})}\big\rangle + \big\langle g,\Pi^{(h_2e_j)}\big\rangle\Big)\\
&= \tilde{f}(\underline{x},\underline{z}+h_1e_i) + \tilde{f}(\underline{x},\underline{z}+h_2e_j).
\end{aligned} \tag{5.34}$$

For the last step, note that $\Pi^{(\underline{z})} + \Pi^{(h_1e_i)}$ has the same distribution as $\Pi^{(\underline{z}+h_1e_i)}$ and that $\Pi^{(\underline{z})} + \Pi^{(h_2e_j)}$ has the same distribution as $\Pi^{(\underline{z}+h_2e_j)}$. This proves $(n+i,n+j)$-concavity of $\tilde{f}$ for all $1 \leq i,j \leq m$. Similar arguments prove $(i,n+j)$-concavity for $1 \leq i \leq n$ and $1 \leq j \leq m$. □

Recall the total mass process $(V_t^{(n)})_{t\geq 0}$ of the n-th generation of the virgin island model from (4.83) for every $n \in \mathbb{N}_0$. Define $|Z_t^{(k)}| := \sum_{i\in G} Z_t^{(k)}$ for $t \in [0,\infty)$ and $k \in \mathbb{N}_0$.

Lemma 5.10. *Assume 2.1 and 3.7. If μ is concave and σ^2 is superadditive, then*

$$\Big(\big(|Z_t^{(k)}|\big)_{k=0,\ldots,k_0}\Big)_{t\geq 0} \leq_{\mathbb{F}_{+-}\left(\{0,\ldots,k_0\},[0,\infty)\right)} \Big(\big(V_t^{(k)}\big)_{k=0,\ldots,k_0}\Big)_{t\geq 0} \tag{5.35}$$

for every $k_0 \in \mathbb{N}_0$. If μ is concave and σ^2 is subadditive, then inequality (5.35) *holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{++}$. If μ is subadditive and σ^2 is additive, then inequality* (5.35) *holds with $\mathbb{F}_{+-}$ replaced by $\mathbb{F}_{+\pm}$.*

Proof. We prove (5.35) by induction on $k_0 \in \mathbb{N}_0$. The base case $k_0 = 0$ follows from $|Z_\cdot^{(0)}| \stackrel{d}{=} V_\cdot^{(0)}$. We apply Lemma 5.7 for the induction step $k_0 \to k_0+1$. Fix

$$F_{n+1}(\chi) = f_{n+1}(\chi_{s_1}^{(k_1)},\ldots,\chi_{s_{n+1}}^{(k_{n+1})}) \in \mathbb{F}_{+-}\big(\{0,\ldots,k_0+1\},[0,\infty)\big) \tag{5.36}$$

where $k_1,\dots,k_{n+1}\in\{0,\dots,k_0+1\}$ and $0\le s_1\le\dots\le s_{n+1}$. Let $\Pi^{i,\zeta}$, $i\in G,\zeta\in \mathrm{C}\big([0,\infty),[0,\infty)\big)$, be independent Poisson point processes on $[0,\infty)\times U$, independent of $\{Z^{(0)},\dots,Z^{(k_0)}\}$ and with $\Pi^{i,\zeta}\stackrel{d}{=}\Pi^{\zeta}$ where Π^{ζ} has intensity measure

$$\mathbb{E}\Pi^{\zeta}(du,d\eta)=\zeta(u)du\otimes Q(d\eta). \tag{5.37}$$

Note that conditioned on $\zeta_i^{(k_0)}(\cdot):=\sum_{j\in G}Z_\cdot^{(k_0)}(j)m(j,i)$, the law of $Z_\cdot^{(k_0+1)}(i)$ is equal to the law of $Y_{\cdot,0}^{\zeta_i^{(k_0)},0}(i)$ (defined in (5.2)) where $Y_{\cdot,0}^{\zeta_i^{(k_0)},0}(i)$, $i\in G$, are independent of each other. Thus conditioning on $\{Z^{(0)},\dots,Z^{(k_0)}\}$ and applying Lemma 5.7 with $x=0$ and $s=0$, we obtain that

$$\begin{aligned}
&\mathbb{E}\Big[F_{n+1}\Big(\big(|Z^{(k)}|\big)_{k=0,\dots,k_0+1}\Big)\Big]\\
&=\mathbb{E}\Big\{\mathbb{E}\Big[F_{n+1}\Big(\big(|Z^{(k)}|\big)_{k=0,\dots,k_0},\sum_{i\in G}Y_{\cdot,0}^{\zeta_i^{(k_0)},0}(i)\Big)\;\Big|\;\big(Z^{(k)}\big)_{k=0\dots k_0}\Big]\Big\}\\
&\le\mathbb{E}\Big[F_{n+1}\Big(\big(|Z^{(k)}|\big)_{k=0,\dots,k_0},\sum_{i\in G}\int_0^\infty\int\eta_{\cdot-u}\Pi^{i,\zeta_i^{(k_0)}}(du,d\eta)\Big)\Big]\\
&=\mathbb{E}\Big[F_{n+1}\Big(\big(|Z^{(k)}|\big)_{k=0,\dots,k_0},\int_0^\infty\int\eta_{\cdot-u}\Pi^{\sum_{i\in G}\zeta_i^{(k_0)}}(du,d\eta)\Big)\Big]\\
&\le\mathbb{E}\Big[F_{n+1}\Big(\big(|Z^{(k)}|\big)_{k=0,\dots,k_0},\int_0^\infty\int\eta_{\cdot-u}\Pi^{|Z^{(k_0)}|}(du,d\eta)\Big)\Big].
\end{aligned}\tag{5.38}$$

The last inequality follows from $\sum_{i\in G}\zeta_i^{(k_0)}(\cdot)=\sum_{j\in G}Z_\cdot^{(k_0)}\sum_{i\in G}m(j,i)\le|Z_\cdot^{(k_0)}|$, where we used the inequality $\sum_{i\in G}m(j,i)\le1$ from Assumption 3.7, and where we used that F_{n+1} is non-decreasing. Next we would like to apply the induction hypothesis. However the right-hand side of (5.38) depends on $(|Z_t^{(k_0)}|)_{t\ge0}$ through a continuum of time points and not only through finitely many time points. To remedy this, we approximate the Poisson point process on the right-hand side of (5.38) by approximating $(|Z_t^{(k_0)}|)_{t\ge0}$ with simple functions. For each $\mathbb{N}\ni m\ge\max(s_n,2n)$, choose a discretization $\{t_0,\dots,t_{m^2}\}$ of $[0,m]$ of maximal width $\frac{2}{m}$ such that $t_0=0$, $t_{m^2}=m$ and $\{s_1,\dots,s_n\}\subset\{t_0,\dots,t_{m^2}\}$. Define $l_i:=k_i$ if $t_i\in\{s_1,\dots,s_n\}$ and $l_i:=0$ otherwise. For a path $(\chi_t)_{t\ge0}\in\mathrm{C}\big([0,\infty),[0,\infty)\big)$, define

$$\big(D_m\chi\big)(t):=\sum_{i=1}^{m^2}\chi_{t_{i-1}}\mathbb{1}_{[t_{i-1},t_i)}(t)\quad\text{for } t\ge0. \tag{5.39}$$

Note that $(D_m\chi)(t)\to\chi_t$ for every $t\ge0$ as $m\to\infty$. Thus the intensity measure $\mathbb{E}\Pi^{D_m\chi^{(k_0)}}(du,d\eta)$ converges to $\mathbb{E}\Pi^{\chi^{(k_0)}}(du,d\eta)$ as $m\to\infty$. This convergence of the intensity measures implies weak convergence of the Poisson point process $\Pi^{D_m\chi^{(k_0)}}$ to the Poisson point process $\Pi^{\chi^{(k_0)}}$. Due to Lemma 5.9, the function $\bar f\colon[0,\infty)^{m^2+1}\to\mathbb{R}$ defined through

$$\bar f\Big(\chi_{t_0}^{(l_0)},\dots,\chi_{t_{m^2}}^{(l_{m^2})}\Big)=\mathbb{E}f_{n+1}\Big(\chi_{s_1}^{(k_1)},\dots,\chi_{s_n}^{(k_n)},\int_0^\infty\int\eta_{\cdot-u}\Pi^{D_m\chi^{(k_0)}}(du,d\eta)\Big) \tag{5.40}$$

is an element of $\mathcal{F}_{+-}^{m^2+1}\big([0,\infty)\big)$. Now we apply the induction hypothesis and obtain that

$$\begin{aligned}
&\mathbb{E}\Big[F\Big(\big(|Z^{(k)}|\big)_{k=0,\dots,k_0},\int_0^\infty\int\eta_{\cdot-u}\Pi^{|Z^{(k_0)}|}(du,d\eta)\Big)\Big]\\
&=\lim_{m\to\infty}\mathbb{E}\Big[F\Big(\big(|Z^{(k)}|\big)_{k=0,\dots,k_0},\int_0^\infty\int\eta_{\cdot-u}\Pi^{D_m|Z^{(k_0)}|}(du,d\eta)\Big)\Big]\\
&\le\lim_{m\to\infty}\mathbb{E}\Big[F\Big(\big(V^{(k)}\big)_{k=0,\dots,k_0},\int_0^\infty\int\eta_{\cdot-u}\Pi^{D_mV^{(k_0)}}(du,d\eta)\Big)\Big]\\
&=\mathbb{E}\Big[F\Big(\big(V^{(k)}\big)_{k=0,\dots,k_0},\int_0^\infty\int\eta_{\cdot-u}\Pi^{V^{(k_0)}}(du,d\eta)\Big)\Big]\\
&=\mathbb{E}\Big[F\Big(\big(V^{(k)}\big)_{k=0,\dots,k_0},V^{(k_0+1)}\Big)\Big].
\end{aligned}\tag{5.41}$$

Putting (5.38) and (5.41) together completes the induction step. □

5.5 Proof of the comparison result of Theorem 3.8

Proof of Theorem 3.8. We prove the case of μ being concave and σ^2 being superadditive. The remaining two cases are analogous. According to Lemma 5.8 we have that

$$\Big(\sum_{i\in\Lambda} X_t(i)\Big)_{t\geq 0} \leq_{\mathbb{F}_{+-}([0,\infty))} \Big(\sum_{i\in\Lambda}\sum_{k=0}^{\infty} Z_t^{(k)}(i)\Big)_{t\geq 0} \tag{5.42}$$

for every finite subset $\Lambda \subseteq G$. Letting $\Lambda \nearrow G$, we see that the total mass of the (G, m, μ, σ^2)-process is dominated by the total mass of the loop-free (G, m, μ, σ^2)-process. Now we get from Lemma 5.10 that

$$\Big(\sum_{k=0}^{k_0} \big|Z_t^{(k)}\big|\Big)_{t\geq 0} \leq_{\mathbb{F}_{+-}([0,\infty))} \Big(\sum_{k=0}^{k_0} V_t^{(k)}\Big)_{t\geq 0} \tag{5.43}$$

for every $k_0 \in \mathbb{N}$. Letting $k_0 \to \infty$, we obtain that the total mass of the loop-free (G, m, μ, σ^2)-process is dominated by the total mass of the virgin island model. Therefore, the total mass of the (G, m, μ, σ^2)-process is dominated by the total mass of the virgin island model. □

References

[1] Romain Abraham and Jean-François Delmas. Changing the branching mechanism of a continuous state branching process using immigration. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(1):226–238, 2009. MR-2500236

[2] David Aldous. Stopping times and tightness. *Ann. Probability*, 6(2):335–340, 1978. MR-0474446

[3] Roland Alkemper and Martin Hutzenthaler. Graphical representation of some duality relations in stochastic population models. *Electron. Comm. Probab.*, 12:206–220, 2007. MR-2320823

[4] Jan Bergenthum and Ludger Rüschendorf. Comparison of semimartingales and Lévy processes. *Ann. Probab.*, 35(1):228–254, 2007. MR-2303949

[5] Jean Bertoin. The structure of the allelic partition of the total population for Galton-Watson processes with neutral mutations. *Ann. Probab.*, 37(4):1502–1523, 2009. MR-2546753

[6] Jean Bertoin. Asymptotic regimes for the partition into colonies of a branching process with emigration. *Ann. Appl. Probab.*, 20(6):1967–1988, 2010. MR-2759725

[7] Jean Bertoin. A limit theorem for trees of alleles in branching processes with rare neutral mutations. *Stochastic Process. Appl.*, 120(5):678–697, 2010. MR-2603059

[8] Matthias Birkner and Andrej Depperschmidt. Survival and complete convergence for a spatial branching system with local regulation. *Ann. Appl. Probab.*, 17(5-6):1777–1807, 2007. MR-2358641

[9] Jochen Blath, Alison Etheridge, and Mark Meredith. Coexistence in locally regulated competing populations and survival of branching annihilating random walk. *Ann. Appl. Probab.*, 17(5-6):1474–1507, 2007. MR-2358631

[10] Benjamin M. Bolker and Stephen W. Pacala. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theor. Popul. Biol.*, 52(3):179–197, 1997.

[11] J. Theodore Cox, Klaus Fleischmann, and Andreas Greven. Comparison of interacting diffusions and an application to their ergodic theory. *Probab. Theory Related Fields*, 105(4):513–528, 1996. MR-1402655

[12] Donald Dawson and Andreas Greven. Multiscale analysis: Fisher-Wright diffusions with rare mutations and selection, logistic branching system. `http://arxiv.org/abs/1007.5462`, 2010.

[13] Donald Dawson and Andreas Greven. Invasion by rare mutants in a spatial two-type Fisher-Wright system with selection. `http://arxiv.org/abs/1104.0253`, 2011.

[14] Donald A. Dawson and Zenghu Li. Construction of immigration superprocesses with dependent spatial motion from one-dimensional excursions. *Probab. Theory Related Fields*, 127(1):37–61, 2003. MR-2006230

[15] Alison M. Etheridge. Survival and extinction in a locally regulated population. *Ann. Appl. Probab.*, 14(1):188–214, 2004. MR-2023020

[16] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes: Characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. MR-0838085

[17] Nicolas Fournier and Sylvie Méléard. A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.*, 14(4):1880–1919, 2004. MR-2099656

[18] Jürgen Gärtner. On the McKean-Vlasov limit for interacting diffusions. *Math. Nachr.*, 137:197–248, 1988. MR-0968996

[19] Iosif I. Gikhman and Anatoli V. Skorokhod. *Introduction to the theory of random processes*. Translated from the Russian by Scripta Technica, Inc. W. B. Saunders Co., Philadelphia, Pa., 1969. MR-0247660

[20] R. C. Griffiths and Anthony G. Pakes. An infinite-alleles version of the simple branching process. *Adv. in Appl. Probab.*, 20(3):489–524, 1988. MR-0955502

[21] Martin Hutzenthaler. The virgin island model. *Electron. J. Probab.*, 14:no. 39, 1117–1161 (electronic), 2009. MR-2511279

[22] Martin Hutzenthaler and Jesse E. Taylor. Time reversal of some stationary jump diffusion processes from population genetics. *Advances in Applied Probability*, 42(4):1147–1171, 2010. MR-2796680

[23] Martin Hutzenthaler and Anton Wakolbinger. Ergodic behavior of locally regulated branching populations. *Ann. Appl. Probab.*, 17(2):474–501, 2007. MR-2308333

[24] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1981. MR-0637061

[25] Adam Jakubowski. On the Skorokhod topology. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(3):263–285, 1986. MR-0871083

[26] Mark Kac. *Probability and related topics in physical sciences*, volume 1957 of *With special lectures by G. E. Uhlenbeck, A. R. Hibbs, and B. van der Pol. Lectures in Applied Mathematics. Proceedings of the Summer Seminar, Boulder, Colo.* Interscience Publishers, London-New York, 1959. MR-0102849

[27] Samuel Karlin and Howard M. Taylor. *A second course in stochastic processes*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1981. MR-0611513

[28] Achim Klenke. *Probability theory*. Universitext. Springer-Verlag London Ltd., London, 2008. A comprehensive course, Translated from the 2006 German original. MR-2372119

[29] Amaury Lambert. The allelic partition for coalescent point processes. *Markov Process. Related Fields*, 15(3):359–386, 2009. MR-2554367

[30] Christian Léonard. Une loi des grands nombres pour des systèmes de diffusions avec interaction et à coefficients non bornés. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(2):237–262, 1986. MR-0850759

[31] Zenghu Li and Tokuzo Shiga. Measure-valued branching diffusions: immigrations, excursions and limit theorems. *J. Math. Kyoto Univ.*, 35(2):233–274, 1995. MR-1346228

[32] Thomas M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985. MR-0776231

[33] Henry P. McKean, Jr. 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)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967. MR-0233437

[34] Carl Mueller and Roger Tribe. A phase transition for a stochastic PDE related to the contact process. *Probab. Theory Related Fields*, 100(2):131–156, 1994. MR-1296425

[35] Karl Oelschläger. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.*, 12(2):458–479, 1984. MR-0735849

[36] Etienne Pardoux and Anton Wakolbinger. From exploration paths to mass excursions – variations on a theme of Ray and Knight. In *Surveys in Stochastic Processes, Proceedings of the 33rd SPA Conference in Berlin, 2009, J. Blath, P. Imkeller, S. Roelly (eds.), EMS 2010* (to appear). Available via *http://www.cmi.univ-mrs.fr/~pardoux/survey070410.pdf*, 2010.

[37] Jim Pitman and Marc Yor. A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete*, 59(4):425–457, 1982. MR-0656509

[38] Sylvie Roelly-Coppoletta. A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics*, 17(1-2):43–65, 1986. MR-0878553

[39] L. C. G. Rogers and David Williams. *Diffusions, Markov processes and martingales. Vol. 1*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Foundations, Reprint of the second (1994) edition. MR-1796539

[40] L. C. G. Rogers and David Williams. *Diffusions, Markov processes and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[41] L. Rüschendorf. Solution of a statistical optimization problem by rearrangement methods. *Metrika*, 30(1):55–61, 1983. MR-0701979

[42] Rinaldo B. Schinazi and Jason Schweinsberg. Spatial and non-spatial stochastic models for immune response. *Markov Process. Related Fields*, 14(2):255–276, 2008. MR-2437531

[43] Moshe Shaked and J. George Shanthikumar. Parametric stochastic convexity and concavity of stochastic processes. *Ann. Inst. Statist. Math.*, 42(3):509–531, 1990. MR-1093966

[44] Moshe Shaked and J. George Shanthikumar. *Stochastic orders and their applications*. Probability and Mathematical Statistics. Academic Press Inc., Boston, MA, 1994. MR-1278322

[45] Jan M. Swart. Duals and thinnings of some relatives of the contact process. http://arxiv.org/abs/math.PR/0604335, 2006.

[46] Ziad Taïb. *Branching processes and neutral evolution*, volume 93 of *Lecture Notes in Biomathematics*. Springer-Verlag, Berlin, 1992. MR-1176317

[47] Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971. MR-0278420

Acknowledgments. I thank Anton Wakolbinger and Don Dawson for inspiring discussions and valuable remarks. Moreover I am indebted to two anonymous referees for very helpful comments and suggestions which led to a considerable improvement of the presentation and of the proofs.