# Rare mutations in the spatial Lambda-Fleming-Viot model in a fluctuating environment and SuperBrownian Motion

Jonathan Chetwynd-Diggle *
Department of Mathematics
Oxford University
Radcliffe Observatory Quarter
Oxford OX2 6GG
UK
chetwynd-diggle@maths.ox.ac.uk

Aleksander Klimek †‡
Max Planck Institute
for Mathematics in the Sciences
Inselstrasse 22
04103 Leipzig
DE
klimek@mis.mpg.de

January 14, 2019

**Abstract**

We investigate the behaviour of an establishing mutation which is subject to rapidly fluctuating selection under the Lambda-Fleming-Viot model and show that under a suitable scaling it converges to the Feller diffusion in a random environment. We then extend to a population that is distributed across a spatial continuum. In this setting the scaling limit is the SuperBrownian motion in a random environment. The scaling results for the behaviour of the rare allele are achieved via particle representations which belong to the family of 'lookdown constructions'. This generalises the results obtained for the neutral version of the model by Chetwynd-Diggle and Etheridge (2018), which was proved using a duality argument. To our knowledge this is the first instance of the application of the lookdown approach in which other techniques seem unavailable.

**Key words:** Spatial Lambda Fleming-Viot model, Fluctuating selection, SuperBrownian motion, lookdown construction, scaling limits

**MSC 2010 Subject Classification:** Primary: 60F05, 60G57, 60J25, 60J68, 60J80, Secondary: 60G51, 60G55, 60J75, 92D10, 92D15

# Contents

**1 Introduction** **2**
1.1 Statement of main results . . . 4
1.2 Structure of the paper . . . 7

**2 Scaling limits of the LFV - dynamics of the rare type** **8**
2.1 Scaling . . . 11
2.2 Main result of this section . . . 11
2.2.1 Projected model . . . 13
2.3 Convergence of the non-spatial model . . . 14

*Supported by EPSRC grant number EP/L015811/1
†Supported by EPSRC grant number EP/L015811/1
‡Funding for this work has been provided by the Alexander von Humboldt Foundation in the framework of the Sofja Kovalevskaja Award endowed by the German Federal Ministry of Education and Research

2.3.1 Intensity estimate . . . 17
2.3.2 Neutral model - proof of Theorem 2.5.1 . . . 19
2.3.3 Selective model - proof of Theorem 2.5.2 . . . 29
2.3.4 Model with selection in fluctuating environment - proof of Theorem 2.4.1 . . 34

**3 SuperBrownian motion in a random environment** **34**
3.1 Definitions . . . 34
3.2 Lookdown representation for BBMRE and SBMRE . . . 37

**4 Scaling limits of the SLFV - dynamics of the rare type** **40**
4.1 Scaling and statement of main results . . . 42
4.2 Intensity estimate . . . 43
4.3 Sketch of the proof of Theorem 1.3.1 . . . 45

**A Poisson random measures** **48**

**B Markov Mapping Theorem** **49**

**C Kurtz-Rodrigues' Martingale Lemma** **50**

**D Proofs of Theorem 3.7.1 and Theorem 3.7.2** **50**

**E Lookdown construction of the Spatial Lambda-Fleming-Viot model** **55**

**References** **58**

# 1 Introduction

We address a question of some biological interest: how the frequency of a rare mutation evolves in a spatially distributed population if the direction of selection on that mutation fluctuates in time? This type of question is particularly relevant in the context of the 'Court Jester hypothesis' (see Barnosky (2001), Benton (2009)), which states that long term improvements in fitness may not occur, since populations must constantly evolve to keep pace with changes in the environment.

The simplest mathematical framework in which this question can be addressed is given by observation of a single genetic locus. For simplicity, we consider a population with two genetic types, the 'common type', $\kappa_c$, (often referred to as a wild type in biological literature) and the 'rare' type, $\kappa_r$. We assume that the rare type forms only a small fraction of the total population. This simple setting may be interpreted as the model for a new mutation before it establishes itself within the population.

Consider for a moment a simple model without selection (for example, a Moran model or a Wright-Fisher model). In the absence of spatial structure, the absolute number of rare individuals will evolve approximately according to a branching process, which, under appropriate scaling, converges to the Feller diffusion. We may ask whether a similar phenomenon occurs in the presence of selection, especially when the direction of selection fluctuates rapidly in time. One expects that in the latter case the evolution approximately follows a branching process in a random environment, and, under suitable scaling, converges to the Feller diffusion in a random environment.

It seems natural to try to establish an analogous result for a spatially distributed population. It is well known that there are serious difficulties when trying to construct models which incorporate genetic drift in higher dimensional spatial continua, see Barton et al. (2013) for a review. A framework which allows us to overcome these difficulties has been found in the spatial Lambda-Fleming-Viot process, introduced in Etheridge (2008) and Barton et al. (2010). Diffusion approximations of this model lead to a limit of the Fisher-KPP type, (see e.g. Etheridge et al. (2018), Forien and Penington (2017)) which is consistent with the behaviour of its non-spatial counterpart.

Therefore, the spatial Lambda-Fleming-Viot model provides a reasonable framework to study the behaviour of the establishment of a mutation. Recent work by Chetwynd-Diggle and Etheridge (2018) for the model without selective advantage or disadvantage for the rare mutation shows convergence. The limiting object is a superBrownian motion, a measure-valued process introduced independently by Watanabe (1968) and Dawson (1975), which is the spatial counterpart of the Feller diffusion. There is some evidence that superBrownian motion (sometimes referred to as a Dawson-Watanabe superprocess) is a universal scaling limit of critical interacting particle systems, see e.g. Cox et al. (2000), Bramson et al. (2001) and van der Hofstad et al. (2017) and references therein.

Recently Biswas et al. (2018) studied a diffusion approximation of the spatial Lambda-Fleming-Viot with selection in a fluctuating environment. In contrast to their work we are interested in the behaviour of an establishing mutation within a large population rather than two established populations of comparable size. In an analogy to the non-spatial case, we show that the limiting process is the superBrownian motion in a random environment, introduced and studied in Mytnik (1996). The work of Nakashima (2015) shows that superBrownian motion in a random environment is the scaling limit of a model of branching random walks on a lattice in random environment, introduced by Birkner et al. (2005). We conjecture that superBrownian motion in a random environment is a universal scaling limit for the critical interacting particle systems in random environments.

The proof of the scaling result in Chetwynd-Diggle and Etheridge (2018) is based on a duality method. However, as discussed in detail in Section 5 of Biswas et al. (2018), a useful dual process seems to not be available in our setting. The techniques of Biswas et al. (2018) are also not available to us as we are considering a rare mutation. Therefore, we use a different approach, based on a particle representation which belongs to the family of lookdown constructions. We build on the work of Kurtz and Rodrigues (2011) and provide a new lookdown construction of superBrownian motion in a random environment. We then use ideas from Etheridge and Kurtz (2018) and a slightly modified version of their construction of the spatial Lambda-Fleming-Viot lookdown which can incorporate a random environment.

For other techniques discussed above many difficulties arise with the introduction of spatial continua. However, the vast majority of our work in the non-spatial result transfers to the spatial result without difficulty, see Remark 1.4 for an explanation. For this reason, the majority of this paper is devoted to the rigorous derivation of the non-spatial result. Our proof technique is based on four main ingredients: lookdown representation, an averaging trick due to Kurtz (1973), a perturbation result due to Kurtz (1992) (which we recall as Theorem 2.7.1) and the Markov Mapping Theorem of Kurtz (1998) (which we recall in Appendix B). Previous results using the lookdown approach have shown either the existence of processes under weak conditions, or convergence of processes which had previously been shown to converge through other means. This is, to our knowledge, the first proof of convergence using the lookdown approach in which other techniques are not available.

Lookdown constructions were introduced in Donnelly and Kurtz (1996, 1999). This approach has proved to be particularly fruitful in applications to population models. In this setting, each individual in the population is assigned a 'level' (taking values in either the integers, as in the original paper of Donnelly and Kurtz (1996), or the reals, as introduced in Kurtz (2000)). Levels typically carry information about genealogical relations between individuals. The name 'lookdown' is used as individuals usually determine their parents by 'looking down' at the sub-population with levels lower than their own.

From a practical perspective, one of the most useful properties of lookdown constructions is that when passing from individual based models to their high density limits, or continuous approximation, the genealogies are preserved. The importance of this can be seen in the examples of systems of individual based models approximated by the same diffusion processes with very different genealogies obtained in Taylor (2009). For further examples in the context of Lambda-Fleming-Viot models we refer the reader to Miller (2015). For an excellent explanation of the general principle of lookdown constructions we refer to Kurtz and Rodrigues (2011), particularly their death process example in Section 2.1.

In the context of the Spatial Lambda-Fleming-Viot we would like to point to two different approaches. The first one is a the construction developed in Etheridge and Kurtz (2018), Section 4.1.3. This construction forms the basis for our construction of the SLFV with selection in a fluctuating environment. We recall a special case of this construction in Appenidix E. The second construction, which was the first lookdown construction for the SLFV, was presented in Véber and Wakolbinger (2015), and was developed using a different approach in Etheridge and Kurtz (2018), Section 4.1.1. The later construction is much closer in the flavour to the original ideas of Donnelly and Kurtz (1996).

## 1.1 Statement of main results

We suppose that the population, which is distributed across $\mathbb{R}^d$, is subdivided into two genetic types. We will denote the space of types by $\mathcal{K} = \{\kappa_c, \kappa_r\}$. Formally, the state of the population at time $t$ is described by a measure $M_t \in \mathcal{M}$. Where $\mathcal{M}$ is the space of measures whose first marginal is Lebesgue measure on $\mathbb{R}^d \times \mathcal{K}$. We note $\mathcal{M}$ is compact when equipped with the topology of weak convergence.

At any fixed time there is a density $w(t,\cdot) : \mathbb{R}^d \to [0,1]$ such that

$$M_t(dx, d\kappa) = (w(t,x)\delta_{\kappa_r}(d\kappa) + (1-w(t,x))\delta_{\kappa_c}(d\kappa))\,\mathrm{d}x.$$

We interpret $w(t,x)$ as the proportion of population of type $\kappa_r$ at location $x$ at time $t$. It is defined only up to Lebesgue null set. For what follows, it is convenient to fix a representative of $M_0$ and update it according to a procedure described below. We consider two types of events - neutral and selective events. Selective events are influenced by the state of the environment.

We begin with a description of the environment, which is used for all models in this section. Our environment is modelled through a simple random field.

**Definition 1.1.** *Let $\Pi^{env}$ be a Poisson process with intensity $E$, dictating the times of the changes in the environment. Let $q(x,y)$ be a covariance function which belongs to $C_0\left(\mathbb{R}^d \times \mathbb{R}^d\right)$ (continuous functions vanishing at infinity) and let $\{\xi^{(m)}(\cdot)\}_{m\geq 0}$ be a family of identically distributed random fields on $\mathbb{R}^d$ such that*

$$\begin{aligned}\mathbb{P}\left[\xi^{(m)}(x) = -1\right] =& \frac{1}{2} = \mathbb{P}\left[\xi^{(m)}(x) = +1\right],\\ \mathbb{E}\left[\xi^{(m)}(x)\xi^{(m)}(y)\right] =& q(x,y).\end{aligned}$$

*Set $\tau_0 = 0$ and write $\{\tau_m\}_{m\geq 1}$ for the points in $\Pi^{env}$ and define*

$$\zeta(t,\cdot) := \sum_{m=0}^{\infty} \xi^{(m)}(\cdot)\mathbf{1}_{[\tau_m,\tau_{m+1})}(t).$$

For the construction of the random field in Definition 3.1 we refer to Ma (2009), especially Example 1. We observe that the generator of the process describing the evolution of the environment, $A^{env}$, is given by

$$A^{env} f(\zeta) = \mathbb{E}_\pi[f(\zeta)] - f(\zeta), \tag{1.1}$$

where $\pi$ is the stationary distribution of the random field $\zeta$.

The following version of the Spatial Lambda-Fleming-Viot is a slight modification of the process discussed in Biswas et al. (2018).

**Definition 1.2** (Spatial Lambda-Fleming-Viot process with fluctuating selection (SLFVFS))**.** *Let $\mu$ be a $\sigma$-finite measure on $(0,\infty)$ and for each $r \in (0,\infty)$, let $\nu_r$ be a probability measure on $[0,1]$, such that the mapping $r \to \nu_r$ is measurable and*

$$\int_{(0,\infty)} r^d \int_{[0,1]} u \,\nu_r(\mathrm{d}u)\mu(\mathrm{d}r) < \infty.$$

*Further, fix $\mathbf{s} \in [0,1]$ and let $\Pi^{neu}$, $\Pi^{fsel}$, be independent Poisson point processes on $\mathbb{R}_+ \times \mathbb{R}^d \times (0,\infty) \times [0,1]$ with intensity measures $(1-\mathbf{s})\mathrm{d}t \otimes \mathrm{d}x \otimes \mu(\mathrm{d}r)\nu_r(\mathrm{d}u)$ and $\mathbf{s}\mathrm{d}t \otimes \mathrm{d}x \otimes \mu(\mathrm{d}r)\nu_r(\mathrm{d}u)$ respectively. Let $\Pi^{env}$ be a Poisson process of Definition 1.1, evolving independently of $\Pi^{neu}, \Pi^{sel}$. Let $\sigma(\kappa,\xi) : \big(\mathcal{K} \times \{-1,1\}\big) \to \mathbb{R}$ be a function which satisfies the symmetry condition*

$$\mathbb{E}_\pi \left[ \frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)} - 1 \right] = 0, \tag{1.2}$$

*The* spatial Lambda-Fleming-Viot process with fluctuating selection (SLFVFS) *with driving noises $\Pi^{neu}$, $\Pi^{fsel}$, $\Pi^{env}$, is the $\mathcal{M}_\lambda$-valued process $M_t$ with dynamics described as follows. Let $w(t_-,\cdot)$ be a representative of the density of $M_{t-}$ immediately before an event $(t,x,r,u)$ from $\Pi^{neu}$ or $\Pi^{fsel}$. Then the measure $M_t$ immediately after the event has density $w(t,\cdot)$ determined by:*

1. *If $(t,x,r,u) \in \Pi^{neu}$, a neutral event occurs at time $t$ within the closed ball $B(x,r)$. Then*
   - *(a) Choose a parental location $l$ according to the uniform distribution on $B(x,r)$.*
   - *(b) Choose the parental type $\kappa \in \{\kappa_r, \kappa_c\}$ according to distribution*

     $$\mathbb{P}[\kappa = \kappa_r] = w(t_-,l), \quad \mathbb{P}[\kappa = \kappa_c] = 1 - w(t_-,l).$$

   - *(c) A proportion $u$ of the population within $B(x,r)$ dies and is replaced by offspring with type $\kappa$. Therefore, for each point $y \in B(x,r)$,*

     $$w(t,y) = w(t_-,y)(1-u) + u\mathbf{1}_{\{\kappa=\kappa_r\}}.$$

2. *If $(t,x,r,u) \in \Pi^{fsel}$, a selective event occurs at time $t$ within the closed ball $B(x,r)$. Then*
   - *(a) Choose a parental location $l$ according to the uniform distribution on $B(x,r)$.*
   - *(b) Choose the parental type $\kappa \in \{\kappa_r, \kappa_c\}$ according to*

     $$\mathbb{P}[\kappa = \kappa_r] = \frac{\sigma(\kappa_r,\zeta)w(t_-,l_i)}{\sigma(\kappa_c,\zeta)w(t_-,l_i) + \sigma(\kappa_r,\zeta)(1-w(t_-,l_i)},$$
     $$\mathbb{P}[\kappa = \kappa_c] = \frac{\sigma(\kappa_r,\zeta)(1-w(t_-,l_i))}{\sigma(\kappa_c,\zeta)w(t_-,l_i) + \sigma(\kappa_r,\zeta)(1-w(t_-,l_i)}.$$

*(c) A proportion $u$ of the population within $B(x, r)$ dies and is replaced by offspring with type $\kappa$. Therefore, for each point $y \in B(x, r)$,*

$$w(t, y) = w(t_-, y)(1 - u) + u\mathbf{1}_{\{\kappa=\kappa_r\}}.$$

The existence of the process follows from the methods of Etheridge et al. (2018) or results of Etheridge and Kurtz (2018), Section 4.1.2. We note the symmetry condition is not required for existence of the model. This condition can be relaxed if $\sigma$ depends on $N$, however for simplicity, we assume fixed $\sigma$ and so require condition (1.2) to be satisfied. Furthermore, we will fix both the impact and radius of events in our models. We assume, with a slight abuse of notation that $\mu = \delta_r$ and $\nu_r = \delta_u$.

We discuss the lookdown representation of the non-spatial version of the model in Section 2, while Section 2.2.1 explains how the lookdown relates to the underlying process. The spatial version of the lookdown representation for this process is described in Section 4.

We shall define the limiting process, which is a variant of superBrownian motion in a random environment with a drift, in terms of the generator. Let $\mathcal{M}_F(\mathbb{R}^d)$ we denote the space of finite measures on $\mathbb{R}^d$, again equipped with a topology of weak convergence.

**Definition 1.3** (SuperBrownian motion in a random environment)**.** *Let $q(x, y) \in C_0\left(\mathbb{R}^d \times \mathbb{R}^d\right)$ be a covariance function. The superBrownian motion in a random environment with a diffusion parameter $m$, a growth parameter $b$ and a quadratic variation parameter $(a, c)$ is the (unique) process $\mu$, taking values in $\mathcal{M}_f(\mathbb{R}^d)$, characterised by the generator (specified for $f \in \bar{C}^2(R_+), \phi \in \mathcal{D}(\Delta)$)*

$$\begin{aligned}\mathcal{L}f(\langle\phi, X_t\rangle) = f'(\mu(\phi))\big[\frac{m}{2}\langle\Delta\phi, X_t\rangle + b\langle\phi, X_t\rangle\big] \\ + \frac{1}{2}f''(\langle\phi, X_s\rangle)\left(a\langle\phi^2, X_s\rangle + c\int_{\mathbb{R}^d\times\mathbb{R}^d} q(x, y)\phi(x)\phi(y)X_s(\mathrm{d}x)X_s(\mathrm{d}y)\mathrm{d}s\right).\end{aligned} \tag{1.3}$$

This model is discussed in more detail in Section 3. In particular, we provide a new lookdown construction of the model, which we derive from the lookdown construction of branching Brownian motion in a random environment.

Our results describe the scaling limit of a sequence of processes. At the $N$th stage of our scaling, the local population density will be $K = K(N)$. We shall denote the representative of the density of the scaled SLFVSRE by $w^N$ and the population of rare individuals by $X^N = Kw^N$, which is defined Lebesgue almost everywhere. We shall think of $X^N$ as a measure-valued process and abuse notation by writing, for any Borel measurable $\phi$,

$$\left\langle X_t^N, \phi\right\rangle = K\int_{\mathbb{R}^d} \phi(x)w_t^N(x)\mathrm{d}x = \int_{\mathbb{R}^d} \phi(x)X_t^N(x)\mathrm{d}x.$$

The scaling for the neutral part of the model is nearly the same as in Chetwynd-Diggle and Etheridge (2018). The modifications required by the presence of selection and fluctuations in the environment are inspired by Biswas et al. (2018). For the scaled process, time is sped up by a factor $N$, space is shrunk by $M(N)$, and the impact of each event is reduced by a factor $J(N)$. The rate of environmental changes is multiplied by $\widehat{S}(N)^2$ and the proportion of selective events is multiplied by $\widehat{S}(N)/S(N)$. Scaling of the selective events is motivated by our inclination to model short burst of strong selection. In the model weak selection limit the rate of selective events is scaled by $1/S(N)$. The additional portion of selective events prevents the action of selection to average out in the diffusive limit. The Central Limit Theorem suggests the relation between the rate of additional selective events and rate of changes of the environment. See also Biswas et al. (2018), Section 3.2 and Section 4.2. Let $C(d) := \int_{\mathcal{B}_1(0)} x^2\mathrm{d}x$. We are now in position to state the main result.

**Theorem 1.3.1.** *Suppose that $X_0^N$ is absolutely continuous with respect to Lebesgue measure with support* $\mathrm{supp}(X_0^N) \subseteq D$, *where $D$ is a compact subset of $\mathbb{R}^d$ independent of $N$, and $X_0^N$ converges weakly to $X_0$. Moreover, suppose that, as $N$ tends to infinity,*

$$\frac{C_d u r^{d+2} N}{JM^2} \to C_1;\ J, K, M, S, \widehat{S} \to \infty;\ \frac{K}{JM^d} \to 0;\ \frac{u^2 V_R N K}{J^2 M^d} \to a;$$
$$\frac{N^2}{KJ^2M^d} \to 0;\ \mathbb{E}_\pi \left[ \left\{ \frac{suNV_R}{SJ} \left( \frac{\sigma(\kappa_r, \zeta)}{\sigma(\kappa_c, \zeta)} - 1 \right) \right\}^2 \right] \to b^2;\ \frac{\widehat{S}}{S} \to 0.$$

*If there exists an $n$ such that, as $N$ tends to infinity, $N(K/J)^n \to 0$ then the sequence $X_N(t)$ converges weakly to superBrownian motion in a random environment with initial condition $X_0$, diffusion parameter $C_1$, growth parameter $b^2$, quadratic variation parameter $(a, b^2)$.*

**Remark 1.4.** *We note that the neutral case of our model is analysed in Chetwynd-Diggle and Etheridge (2018). Their scaling requires*

$$\frac{N}{JM^2} \to C_1;\ J, K, M \to \infty;\ \frac{NK}{J^2M^d} \to C_2,$$

*along with a 'sparsity' condition which we discuss below. Their paper discusses the heuristic reason for their scaling which also gives a solid justification for our choice of scaling.*

*We note that our 'sparsity' condition is the requirement $K/JM^d \to 0$, which is stronger than the one present in Chetwynd-Diggle and Etheridge (2018). Our set of conditions implies $M^2/J \to 0$ in comparison to theirs*

$$\frac{M}{J} \to 0 \ \text{if}\ d = 1;\ \frac{\log M}{J} \to 0 \ \text{if}\ d = 2;\ \frac{1}{J} \to 0 \ \text{if}\ d \geq 3.$$

*It is due to the fact that the lookdown construction 'sees' the Hausdorff dimension of the support of superBrownian motion in a way that previous work did not, since our set of test functions does not smooth out the support of the process, in sharp contrast to their approach. We require the intensity of levels within the ball of radius $r/M$ to tend to infinity in order to allow the limit to have Hausdorff dimension two. This observation explains why the proof of our results in the spatial case does not differ significantly from the proof in the non-spatial case.*

## 1.2 Structure of the paper

The rest of the paper is structured as follows. In Section 2 we discuss the Lambda-Fleming-Viot model, studying its scaling limits and discussing the lookdown representation. In particular, Section 2.3 contains the bulk of our proof. In Section 3, we discuss the lookdown construction for a version of Branching Brownian motion in a random environment and the lookdown representation of the superBrownian motion in a random environment. In Section 4, we discuss how to extend the result of Section 2 to the spatial setup. Appendix A contains some information of Poisson random measures, which are used extensively throughout the paper. Appendix C briefly recalls a Lemma A.13 from Kurtz and Rodrigues (2011) which ensures our projected process is a solution to the correct martingale problem. Appendix B discusses the Markov Mapping Theorem. Appendix D contains some of the proofs of Theorems of Section 3. In Appendix E, we recall the original construction of Etheridge and Kurtz (2018).

## 2 Scaling limits of the LFV - dynamics of the rare type

In this section we are interested in describing the evolution of a subpopulation with a rare mutation within a population which evolves according to the Λ-Fleming-Viot model with selection in a fluctuating environment (LFVSFE). We will again consider a type space with two types, rare and common. Which of those two types has higher fitness changes with the environment. We show that the evolution of the rare subpopulation follows a Feller diffusion in a random environment.

We provide a new description of the LFVSFE in terms of a lookdown construction. This construction is inspired by the lookdown construction of the neutral model in Section 4.1.3 of Etheridge and Kurtz (2018).

Let us begin with a description of the model which gives some insight into its construction before defining the model precisely in Definition 2.1. Each of the individuals in the population is assigned a *genetic type* $\kappa$ from the set $\mathcal{K}$ and a *level* $l \in \mathbb{R}^+ \cup \{0\}$. We restrict our attention to $\mathcal{K} = \{\kappa_c, \kappa_r\}$, which we refer to as the 'common' and 'rare' type, respectively. Since in this section we consider a model without spatial structure, the state of the population can be represented as a collection of points, $\eta = \{(l, \kappa)\}$, or as a measure which can be written, with a slight abuse of notation, as

$$\eta = \sum_{(l,\kappa)\in\eta} \delta_{(l,\kappa)}, \quad \text{where } (l,\kappa) \in \left(\mathbb{R}^+ \cup \{0\}, \mathcal{K}\right).$$

We assume that the process with levels $\eta$ is always a conditionally Poisson system with Cox measure $m_{leb} \times \Xi$, where $m_{leb}$ is Lebesgue measure on $\mathbb{R}^+$. This is because, in our lookdown model, when our initial condition has this form then our process will have this form for all subsequent times. This is a key feature in how the lookdown construction relates to its underlying model.

For lookdown representations we consider test functions of the form

$$f(\eta) = \prod_{l\in\eta} g(l),$$

with the additional requirement that there exists a $\lambda_g > 0$ such that $g(l) = 1$ for all $l > \lambda_g$. This set of test functions will be used throughout this paper for computations involving lookdown representations.

We are interested in the situation when the type which is selectively advantageous depends on the environment. We write $\zeta$ for the random process which models the state of the environment. Therefore the full state of the model at time $t$ is given by a pair $(\eta_t, \zeta_t)$.

We now proceed to carefully state our non-spatial model, the scaling and the non-spatial results. The following statements are almost identical to those found in Section 1.1 but we include them for completeness. This further demonstrates the similarity between the spatial and non-spatial methods when using the lookdown construction and motivates why our spatial proof mainly lies within Proposition 4.6 and Proposition 4.7.

The evolution of the population is determined by reproduction events of two types - neutral and selective. We assume that a proportion, $s$, of events are selective and favour one of the two types. Events are driven by independent Poisson processes $\Pi^{neu}$ and $\Pi^{sel}$. For simplicity we assume that the impact of the events is fixed and equal to $u$.

Both neutral and selective events are composed of two elements - discrete births and thinning. The birth phase of the events differs between neutral and selective events, while the thinning phase is the same.

Whenever $t \in \Pi^{neu}$, a birth event produces offspring, with levels distributed according to an independent Poisson point process with intensity $u$. Let $v^*$ be the smallest of the new levels. Let $(l^*_{neu}, \kappa^*)$ denote the element of $\eta$ with the smallest level greater than $v^*$, that is

$$l^*_{neu} = \min\{l : (l, \kappa) \in \eta, l > v^*\}.$$

The individual $(l^*_{neu}, \kappa^*)$ is chosen as the parent of the event and removed from the population. All new individuals are assigned type $\kappa^*$, the type of the parent. The levels of all old individuals in the population are changed. If the level of the individual was smaller than $v^*$ it remains unaffected by the birth. If the level of the individual was larger than $v^*$, it is moved to $l - l^*_{neu} + v^*$.

If the event is selective, the situation is more complicated. We introduce an additional function $\sigma(\kappa, \zeta)$, which influences the likelihood of an individual of type $\kappa$ being parent, given the state of the environment, $\zeta$. When the environment is in state $\zeta$, the higher the value of $\sigma(\kappa, \zeta)$, the more likely an individual of type $\kappa$ is to be selected as a parent during selective events in environment $\zeta$.

Whenever $t \in \Pi^{sel}$, a birth event produces offspring, with levels distributed according to an independent Poisson process with intensity $u$. As before, let $v^*$ be the smallest of the new levels. Let $(l^*_{sel}, \kappa^*)$ denote the element of $\eta$ which obtains the minimum

$$\min\left\{\frac{l_i - v^*}{\sigma(\kappa_i, \zeta)} : (l_i, \kappa_i) \in \eta, l_i > v^*\right\}. \tag{2.1}$$

The individual $(l^*_{sel}, \kappa^*)$ is chosen as the parent of the event and removed from the population. All new individuals are assigned the parent's type as in neutral events. The levels of all old individuals in the population are then changed. If the level of the individual was smaller than $v^*$ it remains unaffected by the birth. If the level of the individual was larger than $v^*$, it is moved to $\sigma(\kappa, \zeta)(l - l^*_{neu} + v^*)/\sigma(\kappa^*, \zeta)$.

For both neutral and selective events once the parent has been selected, and the old levels moved to their new locations, thinning takes place. Thinning does not affect the new individuals. The thinning takes the level of each individual which is neither a child or the parent of the event present within the population and multiplies it by $1/(1-u)$. The combined effect of this movement is defined explicitly through the function $\mathcal{J}_{neu}$, defined in (2.3), for neutral events and $\mathcal{J}_{sel}$, defined in (2.5), for selective events. We note that instead of removing the parent from the population you can consider the parent to be the lowest offspring instead. The movement of the levels is chosen in this way to maintain levels with a conditionally Poisson system after any event.

We now define the main process of interest.

**Definition 2.1** (Lookdown representation of LFVSFE)**.** *Fix $s \in (0,1)$. Let $\Pi^{neu}, \Pi^{sel}$ be a pair of independent Poisson point processes with intensity measures $(1-s)\mathrm{d}t \otimes \nu(\mathrm{d}u)$ and $s\mathrm{d}t \otimes \nu(\mathrm{d}u)$ respectively on $\mathbb{R}^+ \times (0,1)$. Moreover, let $\Pi^{env}$ be a Poisson process with rate $E$, independent of $\Pi^{neu}, \Pi^{sel}$. Let $\sigma : \mathcal{K} \times \{-1, 1\} \to \mathbb{R}$ be a function.*

*The lookdown representation of LFVSRE is the process taking values in purely atomic measures on $\mathbb{R} \times \mathcal{K} \times \{-1, 1\}$ with dynamics described as follows.*

1. *If $(t, u) \in \Pi^{neu}$*

   *(a) a group of new individuals with levels $(v_1, v_2, \dots)$ is added to the population. Their levels are distributed according to a Poisson process with intensity $u$.*

   *(b) Let $v^* = \min\{v_1, v_2, \dots\}$. The type of the new individuals is chosen to be the same as the type of the individual $(\kappa^*, l^*)$ whose level is the lowest above $v^*$, that is*

$$l^*_{neu} = \min\{l : (l, \kappa) \in \eta, l > v^*\}. \tag{2.2}$$

(c) *As a result of an event the levels with position $l$ before an event will have new position given by*

$$\mathcal{J}_{neu}(l, l^*_{neu}, v^*) = \begin{cases} \frac{1}{1-u}(l - (l^*_{neu} - v^*)) & \text{if } l > l^*_{neu}, \\ \frac{1}{1-u} l & \text{if } l < l^*_{neu}, \\ v^* & \text{if } l = l^*_{neu}. \end{cases} \tag{2.3}$$

2. *If* $(t, u) \in \Pi^{sel}$

(a) *a group of new individuals with levels* $(v_1, v_2, \dots)$ *is added to the population. Their levels are distributed according to a Poisson process with intensity* $u$.

(b) *Let* $v^* = \min\{v_1, v_2, \dots\}$. *The type of the new individuals is chosen to be the same as the type of the individual* $(l^*, \kappa^*)$ *whose level minimizes*

$$\left\{ \frac{l_i - v^*}{\sigma(\kappa_i, \zeta)} : (l_i, \kappa_i) \in \eta, l_i > v^* \right\}. \tag{2.4}$$

(c) *As a result of an event the levels with position $l$ and type $\kappa$ before an event will maintain their type but will have new position given by*

$$\mathcal{J}_{sel}((l, \kappa), (l^*_{sel}, \kappa^*), \zeta, v^*) = \begin{cases} v^* & \text{if } l = l^*_{sel}, \\ \frac{1}{1-u}\left(l - (l^* - v^*)\frac{\sigma(\kappa,\zeta)}{\sigma(\kappa^*,\zeta)}\right) & \text{if } l \neq l^*_{sel}, l > v^*, \\ \frac{1}{1-u} l & \text{if } l < v^*. \end{cases} \tag{2.5}$$

3. *If* $t \in \Pi^{env}$, *the environmental variable* $\zeta_t$ *is resampled uniformly from* $\{-1, 1\}$.

As in the spatial case our results require a symmetry condition on the values of $\sigma(\kappa, \zeta)$. To be more precise, we require that

$$\mathbb{E}_\pi \left[ \frac{\sigma(\kappa_r, \zeta)}{\sigma(\kappa_c, \zeta)} - 1 \right] = 0, \tag{2.6}$$

where $\pi$ is the stationary distribution of $\zeta$.

We note that the generator of the process from Definition 2.1 is given by

$$Af(\eta, \zeta) = A_{neu} f(\eta, \zeta) + A_{sel} f(\eta, \zeta) + A_{env} f(\eta, \zeta), \tag{2.7}$$

where

$$\begin{aligned} &A_{neu} f(\eta, \zeta) \\ &\quad = (1-s) \int_0^\infty \left[ u e^{-uv^*} g(\kappa^*, v^*) e^{-u\int_{v^*}^\infty (1-g(\kappa^*, v))\mathrm{d}v} \prod_{(\kappa,l)\in\eta, l\neq l^*} g\left(\kappa, \mathcal{J}_{neu}(l, l^*_{neu}, v^*)\right) \right] \mathrm{d}v^* - f(\eta), \\ &A_{sel} f(\eta, \zeta) = s \int_0^\infty \Bigg[ u e^{-uv^*} g(\kappa^*, v^*) e^{-u\int_{v^*}^\infty (1-g(\kappa^*, v))\mathrm{d}v} \\ &\qquad\qquad \times \prod_{(\kappa,l)\in\eta, l\neq l^*_{sel}} g\left(\kappa, \mathcal{J}_{sel}((l,\kappa), (l^*_{sel}, \kappa^*), \zeta, v^*)\right) \Bigg] \mathrm{d}v^* - f(\eta), \\ &A_{env} f(\eta, \zeta) = \mathbb{E}_\pi[f(\eta, \zeta)] - f(\eta, \zeta). \end{aligned}$$

Observe that the only differences between neutral and selective events are the choice of the parent $l^*$ and the movement of levels.

## 2.1 Scaling

We are interested in the evolution of the subpopulation of a rare type within the population evolving according to LFVSFE. To quantify the rarity, we consider a population with total density $K$ (which is equal to one for the usual Lambda-Fleming-Viot model), and will let $K$ tend to infinity. We wish the rare type to make up about $\mathcal{O}(1/K)$ of the population at each level of the scaling. This is represented in the look-down process by the intensity of levels. The levels of individuals are initially Poisson distributed with intensity $K$ with individuals given the rare type with probability $\mathcal{O}(1/K)$ and given the common type otherwise.

In order to recover the correct scaling limit we need to readjust our parameters. Let $w^N$ denote the proportion of individuals of the rare type at the $N^{th}$ stage of scaling. Let $X^N = Kw^N$ denote the size of the population of the rare type. Our scaling limit describes the behaviour of $X^N$. The right scaling of the other parameters is suggested by Chetwynd-Diggle and Etheridge (2018) and Biswas et al. (2018) as discussed in Remark 1.4. We speed up the reproduction rate by $N$ and increase the total population size to $K(N)$, but scale down both the impact and the selection coefficient. The impact of an event at the $N^{th}$ stage of the approximation will be given by $u/J(N)$. The selection coefficient in the presence of fluctuations will be $s\widehat{S}(N)/S(N)$ and $s/S(N)$ in the absence of fluctuations. In order to simplify the notation we drop the explicit dependence of scaling parameters on $N$ in what follows. For our results to hold, certain relations between the parameters need to be satisfied. In Theorem 2.4.1, Theorem 2.5.1 and Theorem 2.5.2 we specify scaling limits for the model with fluctuating selection, the model where the direction of selection does not change and the neutral model respectively. In particular, our results require a specific relation between the rate of the changes in the environment and the rate of selection. We therefore assume that the rate of the environmental events is $\widehat{S}^2$.

The generator of the scaled process can then be written as

$$A^N f(\eta,\zeta) = A^N_{neu} f(\eta,\zeta) + \widehat{S} A^N_{sel} f(\eta,\zeta) + \widehat{S}^2 A^N_{env} f(\eta,\zeta), \tag{2.8}$$

where

$$\begin{aligned}
A^N_{neu} f(\eta,\zeta) &= N\Bigg( \int_0^\infty \Bigg[ \frac{uK}{J} e^{-\frac{uK}{J} v^*} g(\kappa^*, v^*) e^{-\frac{uK}{J}\int_{v^*}^\infty (1-g(\kappa^*,v))\mathrm{d}v} \\
&\qquad \times \prod_{(\kappa,l)\in\eta, l\neq l^*} g\left(\kappa, \mathcal{J}_{neu}(l,l^*,v^*)\right) \Bigg] \mathrm{d}v^* - f(\eta) \Bigg), \\
A^N_{sel} f(\eta,\zeta) &= \frac{sN}{S}\Bigg( \int_0^\infty \Bigg[ \frac{uK}{J} e^{-\frac{uK}{J} v^*} g(\kappa^*, v^*) e^{-\frac{uK}{J}\int_{v^*}^\infty (1-g(\kappa^*,v))\mathrm{d}v} \\
&\qquad \times \prod_{(\kappa,l)\in\eta, l\neq l^*_{sel}} g\left(\kappa, \mathcal{J}_{sel}(l,l^*_{sel},v^*)\right) \Bigg] \mathrm{d}v^* - f(\eta) \Bigg), \\
A^N_{env} f(\eta,\zeta) &= \mathbb{E}_\pi[f(\eta,\zeta)] - f(\eta,\zeta).
\end{aligned}$$

## 2.2 Main result of this section

We now state the main results of this section. We recall the definitions of some of the classical models in terms of their generators, in order to state the results formally.

**Definition 2.2** (Feller diffusion)**.** *Let $a, b > 0$. The Feller diffusion is the process taking values in*

$\mathbb{R}$ *with generator* $C_{fd}$*, defined for every* $f \in C^\infty(\mathbb{R})$*, given by*

$$C_{fd}f(y) = ayf''(y) + byf'(y).$$

*Its lookdown representation is characterised by the process with generator* $A_{fd}$ *given by*

$$A_{fd}f(\eta) = f(\eta)\sum_i 2a\int_{l_i}^\infty \big(g(v)-1\big)\mathrm{d}v + f(\eta)\sum_i \big(al_i^2 - bl_i\big)\frac{g'(l_i)}{g(l_i)}. \tag{2.9}$$

**Definition 2.3** (Feller diffusion in random environment)**.** *Let* $a, b > 0$*. The Feller diffusion in a random environment is the process taking values in* $\mathbb{R}$ *with generator* $C_{fdr}$*, defined for every* $f \in C^\infty(\mathbb{R})$*, of the form*

$$C_{fdr}f(y) = \big(ay + b^2y^2\big)f''(y) + b^2yf'(y).$$

*Its lookdown representation is characterised by the process with generator* $A$ *given by*

$$\begin{aligned}Af(\eta) = f(\eta)\sum_i 2a\int_{l_i}^\infty \big(g(v)-1\big)\mathrm{d}v + f(\eta)\sum_i \big(al_i^2 - b^2l_i\big)\frac{g'(l_i)}{g(l_i)} \\ + b^2f(\eta)\sum_j\left(\sum_{i\neq j} l_jl_i\frac{g'(l_i)g'(l_j)}{g(l_i)g(l_j)} + l_j^2\frac{g''(l_i)}{g(l_i)}\right).\end{aligned} \tag{2.10}$$

For a detailed discussion of the lookdown constructions for the Feller diffusion and the Feller diffusion in a random environment we refer to Kurtz and Rodrigues (2011), Section 2.

**Remark 2.4.** *We observe that our choice of test functions guarantees that since*

$$\sum_i\left(g'(l_i)\prod_{j\neq i}g(l_j)\right) = f(l)\sum_i\left(\frac{g'(l_i)}{g(l_i)}\right),$$

*all terms appearing in* (2.9) *and* (2.10) *are well-defined, even if* $g(l_i) = 0$*.*

**Theorem 2.4.1.** *Let* $X^N(t)$ *denote the total intensity of individuals of rare type at time* $t$*. Suppose that* $X_0^N$ *converges to* $X_0$*. Moreover, suppose that, as* $N$ *tends to infinity,*

$$\begin{aligned}J, K, S, \widehat{S} \to \infty; \quad \frac{K}{J} \to 0; \quad \frac{u^2NK}{J^2} \to 2a; \quad \frac{N^2}{KJ^2} \to 0; \\ \mathbb{E}_\pi\left[\left\{\frac{suN}{SJ}\left(\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)} - 1\right)\right\}^2\right] \to b^2; \quad \frac{\widehat{S}}{K} \to 0; \quad \frac{\widehat{S}}{S} \to 0.\end{aligned} \tag{2.11}$$

*In addition, assume that there exists an* $m$ *such that, as* $N \to \infty$*,* $NK^m/J^m \to 0$*. Then the sequence* $X_N(t)$ *converges weakly to the Feller diffusion in a random environment with initial condition* $X_0$ *and parameters* $2a$*,* $b^2$*.*

**Example 2.5.** *Fix* $\epsilon \in (0, 1/4)$*,* $\beta \in (0, 1/4 - \epsilon)$ *and* $\gamma \in (0, \beta)$*. The conditions of Theorem 2.4.1 are satisfied if*

$$J = N^{\frac{3}{4}+\epsilon} \quad S = N^\beta \quad K = N^{\frac{1}{2}+2\epsilon} \quad \widehat{S} = N^\gamma.$$

As a by-product of our technique, we prove an analogous result for the neutral model and the model with selection.

**Theorem 2.5.1.** *Let $X^N(t)$ denote the total intensity of individuals of rare type at time $t$. Suppose that $X_0^N$ converges to $X_0$ and the intensity of selective events is zero. Moreover, suppose that, as $N$ tends to infinity,*

$$J, K \to \infty; \quad \frac{K}{J} \to 0; \quad \frac{N^2}{KJ^2} \to 0; \quad \frac{u^2 NK}{J^2} \to 2a. \tag{2.12}$$

*In addition, assume that there exists an $m$ such that, as $N \to \infty$, $NK^m/J^m \to 0$. Then the sequence $X_N(t)$ converges weakly to the critical Feller diffusion ($b = 0$) with initial condition $X_0$ and variance parameter $2a$.*

**Theorem 2.5.2.** *Let $X^N(t)$ denote the total intensity of individuals of rare type at time $t$. Suppose that $X_0^N$ converges to $X_0$, $\sigma$ does not depend on the environment and $\hat{S} = 1$. Moreover, suppose that, as $N$ tends to infinity,*

$$J, K, S \to \infty; \quad \frac{K}{J} \to 0; \quad \frac{N^2}{KJ^2} \to 0; \quad \frac{u^2 NK}{J^2} \to 2a; \quad \frac{suN}{SJ}\left(\frac{\sigma(\kappa_r)}{\sigma(\kappa_c)} - 1\right) \to b \tag{2.13}$$

*In addition, assume that there exists an $m$ such that, as $N \to \infty$, $NK^m/J^m \to 0$. Then the sequence $X_N(t)$ converges weakly to the Feller diffusion with initial condition $X_0$ and parameters $2a$ and $b$.*

**Remark 2.6.** *We stress that our proofs do not guarantee convergence of the lookdown representations, but only convergence of the projected models.*

### 2.2.1 Projected model

We follow the work in Etheridge and Kurtz (2018) to show the connection between our lookdown construction and the standard LFV.

The neutral generator is simply the non-spatial counterpart of the lookdown construction from Etheridge and Kurtz (2018) and so we do not treat it here. However, in order to clarify how our form of selection acts from the perspective of the underlying process we investigate further. This section follows Etheridge and Kurtz (2018) and all the techniques in this section are taken from there, but formulated in our notation. The part of the generator of our process which describes selective events is of the form

$$A_{sel}f(\eta) = s\Bigg(\int_0^\infty \Bigg[ue^{-uv^*} g(v^*, \kappa^*) e^{-u\int_{v^*}^\infty (1-g(v,\kappa^*))\mathrm{d}v} \times \prod_{(l,\kappa)\in\eta, l\neq l^*_{sel}} g\left(\mathcal{J}_{sel}(l, l^*_{sel}, v^*), \kappa\right)\Bigg]\mathrm{d}v^* - f(\eta)\Bigg).$$

We define $h(\kappa) = \int_0^\infty (1 - g(l, \kappa))\mathrm{d}l$ and note that integration by parts gives

$$\int_0^\infty ue^{-uv^*} g(v^*, \kappa^*) e^{-\frac{uK}{J}\int_{v^*}^\infty (1-g(v,\kappa^*))\mathrm{d}v}\mathrm{d}v^* = e^{-uh(\kappa^*)}.$$

We also note that if $\eta$ is formed from a Poisson point process with intensity measure $m_{leb} \otimes \Xi(\mathrm{d}\kappa)$ then

$$\{\mathcal{J}_{sel}(l, l^*_{sel}, v^*) : l \neq l^*_{sel}\},$$

is a Poisson point process with intensity measure $m_{leb} \otimes (1-u)\Xi(\mathrm{d}\kappa)$. Consider the distribution of $\kappa^*$ conditioned on $\Xi(\mathrm{d}\kappa)$. We note that

$$\mathbb{P}[\kappa^* = \kappa_r] = \frac{\sigma(\kappa_r)\Xi(\{\kappa_c\})}{\sigma(\kappa_r)\Xi(\{\kappa_c\}) + \sigma(\kappa_c)\Xi(\{\kappa_r\})},$$

and so when we average our selective generator we get

$$\begin{aligned}\alpha A^N_{sel} f(\Xi) = se^{-\int_{\mathcal{K}} h(\kappa)\Xi(\mathrm{d}\kappa)} \\ \times \Bigg(\Bigg[&\frac{\sigma(\kappa_r)\Xi(\{\kappa_c\})}{\sigma(\kappa_r)\Xi(\{\kappa_c\}) + \sigma(\kappa_c)\Xi(\{\kappa_r\})} e^{-uh(\kappa_r)} \\ &+ \frac{\sigma(\kappa_c)\Xi(\{\kappa_r\})}{\sigma(\kappa_r)\Xi(\{\kappa_c\}) + \sigma(\kappa_c)\Xi(\{\kappa_r\})} e^{-uh(\kappa_c)}\Bigg] e^{u\int_{\mathcal{K}} h(\kappa)\Xi(\mathrm{d}\kappa)} - 1\Bigg).\end{aligned}$$

We finally note that our selection can now be seen as a simple weighted choice of parent in the (non-spatial) Lambda Fleming-Viot process. The same calculation can be done for the spatial case.

**Remark 2.7.** *We notice that our way of modelling selection differs from the approach of Etheridge et al. (2018) and Biswas et al. (2018). However, a quick generator calculation shows that this type of selection leads to the same diffusion approximation as in Etheridge et al. (2018) and Biswas et al. (2018).*

## 2.3 Convergence of the non-spatial model

We are interested in convergence of the model with selection in a fluctuating environment. Recall that we study the behaviour of an establishing mutation under this model. We take that into account by considering test functions which are unaffected by the individuals of the common type, $\kappa_c$.

We recall the generator of the rescaled process takes the form

$$A^N f(\eta,\zeta) = A^N_{neu} f(\eta,\zeta) + \widehat{S} A^N_{sel} f(\eta,\zeta) + \widehat{S}^2 A^N_{env} f(\eta,\zeta), \tag{2.14}$$

where $A^N_{neu}$ is the part of the generator which describes neutral events, $A^N_{sel}$ is the part of the generator which describes selective events, and $A^N_{env}$ describes the evolution of the environment.

In Sections 2.3.2 and 2.3.3, we show that the terms $A^N_{neu}$ and $A^N_{sel}$ converge to well-defined limits. However, since $\widehat{S} \to \infty$ as $N$ tends to infinity, it may seem that as $N$ tends to infinity, (2.14) will not converge to a non-trivial limit. However, the naive limiting procedure does not take into account the changes in the direction of selection. In order to identify the correct limit, we apply a 'separation of timescales' trick due to Kurtz (1973).

By the calculations in the proof of Theorem 2.5.2 the operator $A^N_{sel}$ can be written as

$$A^N_{sel} f(\eta) = f(\eta)\frac{suN}{JS}\left(\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)} - 1\right)\sum_j l_j \frac{g'(\kappa_j, l_j, \xi)}{g(\kappa_j, l_j, \xi)} + \mathcal{O}\left(\frac{1}{S} + \frac{1}{K}\right) \tag{2.15}$$

We consider a test function $\widetilde{f}$ of the form

$$\widetilde{f}(\eta,\zeta) = f(\eta) + \frac{1}{\hat{S}} f(\eta) \frac{suN}{JS}\left(\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)} - 1\right) \sum_j \frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)} l_j =: f(\eta) + \frac{1}{\hat{S}} f_1(\eta,\zeta).$$

Observe that since $\hat{S} \to \infty$, we can prove the test function $\widetilde{f}$ will tend to $f(\eta)$ as $N \to \infty$. We apply the generator (2.14) to $\widetilde{f}$. This leads to

$$\begin{aligned} A^N \widetilde{f}(\eta,\zeta) &= A^N_{neu} f(\eta) + \widehat{S} A^N_{sel} f(\eta) + \frac{1}{\hat{S}} A^N_{neu} f_1(\eta,\zeta) + A^N_{sel}\left(f_N(\eta,\xi)\right) - \widehat{S} f_N(\eta,\xi) \\ &= A^N_{neu} f(\eta) + A^N_{sel}\left(f_N(\eta,\xi)\right) + \mathcal{O}\left(\frac{1}{\widehat{S}} + \frac{\widehat{S}}{S} + \frac{\widehat{S}}{K}\right) \end{aligned}$$

where we have used (2.15), that $f$ does not depend on $\xi$ (to see $A_{env} f(\eta) = \mathbb{E}_\pi[f(\eta)] - f(\eta) = 0$) and (2.6), (to see that, as $\mathbb{E}_\pi[f_N(\eta,\zeta)] = 0$, $A_{env} f_N(\eta,\zeta) = -f_N(\eta,\zeta)$). Therefore, at least heuristically, the identification of $A^N_{sel}\left(f_N(\eta,\xi)\right)$ should lead to the correct limit.

To make this argument rigorous, we shall use a theorem due to Kurtz (1992) which we recall here. Let us introduce some notation. For a metric space $E$, let $l_m(E)$ be the space of measures on $[0,\infty) \times E$ such that $\mu \in l_m(E)$ if and only if $\mu([0,t) \times E) = t$.

**Theorem 2.7.1** (Kurtz (1992), Theorem 2.1)**.** *Let $E_1$, $E_2$ be complete separable metric spaces, and set $E = E_1 \times E_2$. For each $n$, let $\{(X_n, Y_n)\}$ be a stochastic process with sample paths in $D_E([0,\infty))$ adapted to a filtration $\{\mathcal{F}^n_t\}$. Assume that $\{X_n\}$ satisfies the compact containment condition, that is, for each $\epsilon > 0$ and $T > 0$, there exists a compact $K \subset E$ such that*

$$\inf_n \mathbb{P}[X_n(t) \in K, t \le T] \ge 1 - \epsilon, \tag{2.16}$$

*and assume that $\{Y_n(t) : t \ge 0, n = 1, 2, \ldots\}$ is relatively compact (as a collection of $E_2$-valued random variables). Suppose that there is an operator $A : \mathcal{D}(A) \subset \overline{C}(E_1) \to C(E_1 \times E_2)$ such that for $f \in \mathcal{D}(A)$ there is a process $\epsilon^f_n$ for which*

$$f(X_n(t)) - \int_0^t Af(X_n(s), Y_n(s))\mathrm{d}s + \epsilon^f_n(t) \tag{2.17}$$

*is an $\{\mathcal{F}^n_t\}$-martingale. Let $\mathcal{D}(A)$ be dense in $\overline{C}(E_1)$ in the topology of uniform convergence on compact sets. Suppose that for each $f \in \mathcal{D}(A)$ and each $T > 0$, there exists $p > 1$ such that*

$$\sup_n \mathbb{E}\left[\int_0^T |Af(X_n(t), Y_n(t)|^p \mathrm{d}t\right] < \infty \tag{2.18}$$

*and*

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\le T} |\epsilon^f_n(t)|\right] = 0. \tag{2.19}$$

*Let $\Gamma_n$ be the $l_m(E_2)$-valued random variable given by*

$$\Gamma_n\left([0,t] \times B\right) = \int_0^t \mathbf{1}_B(Y_n(s))\mathrm{d}s.$$

*Then $\{(X_n, \Gamma_n)\}$ is relatively compact in $D_{E_1}[0,\infty) \times l_m(E_2)$, and for any limit point $(X, \Gamma)$ there exists a filration $\{\mathcal{G}_t\}$ such that*

$$f(X(t)) - \int_0^t \int_{E_2} Af(X(s), y)\Gamma(\mathrm{d}s \times \mathrm{d}y) \tag{2.20}$$

*is a $\{\mathcal{G}_t\}$-martingale for each $f \in \mathcal{D}(A)$.*

We observe that

$$f(\eta_t)-\int_0^t Af(\eta_t,\zeta_t)\mathrm{d}s+(\widehat{f}(\eta_t,\zeta_t)-f(\eta_t))+\int_0^t Af(\eta_s,\zeta_s)-A^N\widehat{f}(\eta_s,\zeta_s)\mathrm{d}s$$
$$=\widetilde{f}(\eta_t,\zeta_t)-\int_0^t A^N\widetilde{f}(\eta_s,\zeta_s)\mathrm{d}s,$$

where $A$ is given by (2.10). Since $\widetilde{f}-\int_0^t A^N\widetilde{f}(s)\mathrm{d}s$ is a martingale, we have written our problem in the form (2.17) with

$$\epsilon_N^f(t)=(\widetilde{f}(\eta_t,\zeta_t)-f(\eta_t))+\int_0^t Af(\eta_s,\zeta_s)-A^N\widetilde{f}(\eta_s,\zeta_s)\mathrm{d}s \tag{2.21}$$

To check that the assumptions of Theorem 2.7.1 are satisfied, we work with both the lookdown representation and the projected model. The projected model allows us to check the compact containment condition (2.16) and prove the $L^p$ estimate (2.18). Both are achieved via an intensity estimate given by the following Lemma.

**Lemma 2.8.** *Let $X^N=Kw^N$ denote the total intensity of individuals of the rare type. Assume that $\mathbb{E}[X^N(0)]<\infty$. Then for any $T>0$*

$$\sup_{t\le T}\sup_N \mathbb{E}[X^N(t)]<\infty, \tag{2.22}$$

$$\lim_{H\to\infty}\sup_N \mathbb{P}\left[\sup_{t\le T}X^N(t)>H\right]=0. \tag{2.23}$$

We discuss the proof in Section 2.3.1.

The part of the argument which allows us to identify the correct limit and shows that condition (2.19) is satisfied, that is

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{t\le T}|\epsilon_N^f(t)|\right]=0, \tag{2.24}$$

is more involved, and requires the use of the lookdown representation. We will first look at the behaviour of

$$\int_0^t A^Nf(\eta_s,\zeta_s)-Af(\eta_s,\zeta_s)\mathrm{d}s.$$

The terms involving $A^N_{neu}$ and $A^N_{sel}$ are tackled separately.

Let $\eta_t^r$ be the process obtained from $\eta_t$ by only considering the individuals of the rare type, $\eta_t^r:=\{l:(l,\kappa_r)\in\eta_t\}$.

**Proposition 2.9.** *Under the conditions of Theorem 2.5.1,*

$$\mathbb{E}\left[\sup_{t\le T}\left|\int_0^t A^N_{neu}f(\eta_s)\right.\right.$$
$$\left.\left.-\left(f(\eta_s^r)\sum_{l_i(t)\in\eta_t^N}al_i^2\frac{g'(l_i(t))}{g(l_i(t))}+2af(\eta_s^r)\sum_{l_i(t)\in\eta_s^r}\int_{l_i(t)}^\infty\big(1-g(\kappa_i,v)\big)\mathrm{d}v\right)\mathrm{d}s\right|\right]\to 0.$$

This proposition will be proved via Taylor's formula through Proposition 2.15 and Proposition 2.16 in Section 2.3.2. An analogous proposition applies to the terms involving $A^N_{sel}$.

**Proposition 2.10.** *Under the conditions of Theorem 2.5.2 for any $T \in \mathbb{R}$*

$$\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t A^N_{sel}f(\eta^N_s) - f(\eta^N_s)\left(\sum_{l_i(t)\in\eta^N_t} -bl_i\frac{g'(l_i)}{g(l_i)}\right)\mathrm{d}s\right|\right] \to 0.$$

We discuss the proof of this proposition, along with calculations which allow us to fully justify the separation of timescales procedure in Section 2.3.3. These three propositions will allow us to conclude Theorem 2.5.1 and Theorem 2.5.2. We then proceed to use these results to prove Theorem 2.4.1 in Section 2.3.4. We notice that if the random perturbation $Y$ appearing in the statement of Theorem 2.7.1 is trivial, the statement itself reduces to the usual condition for relative compactness of the sequence of stochastic processes, see, for example, Ethier and Kurtz (1986), Theorem 3.9.1 and Theorem 3.9.4. Therefore as a by-product of our construction we give a proof of Theorem 2.5.1 and Theorem 2.5.2. The proof of Theorem 2.5.1 is the most technical of this section.

**Remark 2.11.** *Our proof guarantees the relative compactness of the sequences of scaled lookdown representations and that limit points must satisfy a martingale problem. However, we do not have a proof of uniqueness of the martingale problem characterizing the limiting equation. We will use the Markov Mapping Theorem to deduce the relative compactness of the sequence of projected models and a martingale problem characterising limit points. Lemma A.13 from Kurtz and Rodrigues (2011) guarantees that every projection given by the Markov Map solves the projected martingale problem. Since this projected martingale problem has unique solutions, the relative compactness is enough to guarantee convergence of the sequence of projected models.*

**Remark 2.12.** *We are interested in the behaviour of a rare subpopulation. As we have discussed earlier, we would like it to form $\mathcal{O}(1/K)$ of the population. For technical reasons, instead of the process $X^N(t)$, it is sometimes convenient to consider a stopped process $X^N(t\wedge\tau^N)$, where*

$$\tau^N := \inf\{t > 0 : X^N_t > Z^N\}.$$

*We require the sequence of real numbers $Z^N$ to be finite for each $N$ and to tend to infinity as $N$ tends to infinity. This requirement coupled with Lemma 2.8 guarantees that the convergence of the stopped processes translates directly into convergence of the unstopped processes. We shall see that technical assumptions will require that $Z^N \to \infty$ sufficiently slowly. However, these assumptions will not change our proof.*

### 2.3.1 Intensity estimate

This subsection is devoted to the proof of our intensity estimate.

*Proof of Lemma 2.8.* The generator of the projected process can be written as

$$\begin{aligned}\mathcal{L}f(w,\zeta) &= \{wf((1-u)w+u,\zeta) + (1-p)f((1-u)w,\zeta) - f(w,\zeta)\}\\ &+ s\left[\frac{\sigma(\kappa_r,\zeta)w}{\sigma(\kappa_r,\zeta)w+\sigma(\kappa_s,\zeta)(1-w)}pf((1-u)w+u,\zeta)\right.\\ &+\left.\frac{\sigma(\kappa_c,\zeta)(1-w)}{\sigma(\kappa_r,\zeta)w+\sigma(\kappa_s,\zeta)(1-w)}wf((1-u)w+u,\zeta) - f(w,\zeta)\right]\end{aligned}$$

$$+ \mathcal{L}^{env} f(p,\zeta).$$

Recall that we are interested in an estimate for the intensity of the process describing the evolution of the rare individuals. We therefore substitute $X = Kw$. Under the assumptions of Theorem 2.4.1 this leads to the generator

$$\begin{aligned}
\mathcal{L}f(X,\zeta) = N &\left\{ \frac{X}{K} f\left( \left(1-\frac{u}{J}\right) X + K\frac{u}{J}, \zeta \right) + \left(1-\frac{X}{K}\right) f\left( \left(1-\frac{u}{J}\right) X, \zeta \right) - f(X,\zeta) \right\} \\
&+ N\widehat{S}\frac{s}{S}\left[ \frac{\sigma(\kappa_r,\zeta)\frac{X}{K}}{\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_s,\zeta)\left(1-\frac{X}{K}\right)} f\left( \left(1-\frac{u}{J}\right) X + K\frac{u}{J}, \zeta \right) \right. \\
&+ \left. \frac{\sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)}{\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_s,\zeta)\left(1-\frac{X}{K}\right)} f\left( \left(1-\frac{u}{J}\right) X, \zeta \right) - f(X,\zeta) \right] \\
&+ \mathcal{L}^{env} f(X,\zeta).
\end{aligned}$$

This means that for any $f \in C^\infty(\mathbb{R})$,

$$f(X^N(T),\zeta_T) = f(X_0,\zeta_0) + \int_0^T \mathcal{L}(X^N(s),\zeta_s)\mathrm{d}s + M(T),$$

where $M$ is a martingale. Substituting $f(x,\zeta) = x$ leads to

$$\begin{aligned}
X(T) =& X(0) + N\widehat{S}\frac{s}{S}\frac{u}{J}\int_0^T \left[ \frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)} \frac{X}{K}(K-X) \right. \\
&+ \left. \frac{-\sigma(\kappa_c,\zeta)}{\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)} \left(1-\frac{X}{K}\right) X \right] \mathrm{d}s + M(T) \\
=& X(0) + N\widehat{S}\frac{s}{S}\frac{u}{J}\int_0^T \left[ \frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)} X \right. \\
&- \left. \left( \frac{\sigma(\kappa_r,\zeta)\frac{X}{K}}{\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)} + \frac{\sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)}{\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)} \right) X \right] \mathrm{d}s + M(T).
\end{aligned}$$

Since

$$\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)} = \frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)} + \frac{X}{K}\frac{\sigma(\kappa_c,\zeta)\sigma(\kappa_r,\zeta) - \sigma^2(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)\left(\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)\right)}, \tag{2.25}$$

this expression can be written as

$$\begin{aligned}
X(T) = X(0) + N\widehat{S}\frac{s}{S}\frac{u}{J} &\left\{ \int_0^T \left[ \left( \frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)} - 1 \right) X \right] \mathrm{d}s \right. \\
&+ \left. \int_0^T \left[ \frac{X^2}{K} \frac{\sigma(\kappa_c,\zeta)\sigma(\kappa_r,\zeta) - \sigma^2(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)\left(\sigma(\kappa_r,\zeta)\frac{X}{K} + \sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)\right)} \right] \mathrm{d}s \right\} + M(T).
\end{aligned} \tag{2.26}$$

Consider the stopping time $\tau = \inf\{t \geq 0 : X^N > H\}$ and the stopped process $\widehat{X}^N(t) = X^N(t \wedge \tau)$. Since (2.26) holds at a bounded stopping time, taking the expectation and using the symmetry condition (2.6) leads to

$$\mathbb{E}[\widehat{X}^N(T)] = \mathbb{E}[\widehat{X}(0)] + N\widehat{S}\frac{s}{S}\frac{u}{J}\mathbb{E}\left[\int_0^T \left\{\left(\frac{\sigma(\kappa_c,\zeta)}{\sigma(\kappa_r,\zeta)} - 1\right)\widehat{X}^N(s)\right\}\mathrm{d}s\right.$$
$$\left. + \int_0^T \left\{\frac{(\widehat{X}^N(s))^2}{K}\frac{\sigma^2(\kappa_r,\zeta) + \sigma(\kappa_c,\zeta)\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)\left(\sigma(\kappa_r,\zeta)\frac{\widehat{X}^N(s)}{K} + \sigma(\kappa_c,\zeta)\left(1 - \frac{\widehat{X}^N(s)}{K}\right)\right)}\right\}\mathrm{d}s\right]$$
$$\leq \mathbb{E}[\widehat{X}(0)] + N\frac{H}{K}\frac{s\widehat{S}}{S}\frac{u}{J}C_\sigma\int_0^T \mathbb{E}[\widehat{X}^N(s)]\mathrm{d}s,$$

where $C_\sigma$ is a constant depending on $\sigma$. By Grönwall's inequality

$$\mathbb{E}[\widehat{X}^N(T)] \leq \mathbb{E}[\widehat{X}^N(0)]\exp\left(N\frac{H}{K}\frac{s\widehat{S}}{S}\frac{u}{J}C_\sigma T\right).$$

If one takes $H = Z^N$ this, combined with the conditions from Theorem 2.4.1, concludes the proof of (2.22). To see that (2.23) holds, it is enough to observe that by Markov's inequality

$$\mathbb{P}\left[\sup_{0\leq t\leq T} X_t \geq H\right] = \mathbb{P}[\widehat{X}_T \geq H] \leq \frac{\mathbb{E}[\widehat{X}_T]}{H}, \tag{2.27}$$

and for any $T$, $\varepsilon > 0$ we can choose $H_1$, $N_1$ such that for $H \geq H_1$, $N \geq N_1$ and $t \leq T$ the right hand side of (2.27) is less than $\varepsilon$. □

We also take note of a simple corollary which we will use later in our main proof.

**Corollary 2.13.** *The lookdown process satisfies the compact containment condition,* (2.16) *in Theorem 2.7.1.*

*Proof.* We note that the characterisation of convergence given in Theorem A.2.1 ensures that a sequence of lookdown processes satisfies the compact containment condition if and only if the projected processes also satisfy compact containment condition. □

**Remark 2.14.** *We do not give details but a simpler calculation also proves Lemma 2.8 under the assumptions of Theorem 2.5.1 and Theorem 2.5.2.*

### 2.3.2 Neutral model - proof of Theorem 2.5.1

Even though this subsection contains the proof of the result for the least complicated model, the proof itself is the most involved one. The relative simplicity of the proof for the more complicated model demonstrates the power of this lookdown method. Proofs of Theorem 2.5.2 and Theorem 2.4.1 heavily rely on technical observations from this subsection. We recall that the generator of the neutral part of the process takes the form

$$A^N_{neu}f(\eta) = N\left(\int_0^\infty \left[\frac{uK}{J}e^{-\frac{uK}{J}v^*}g(\kappa^*,v^*)e^{-\frac{uK}{J}\int_{v^*}^\infty(1-g(\kappa^*,v))\mathrm{d}v}\right.\right.$$
$$\left.\left.\times \prod_{(\kappa,l)\in\eta, l\neq l^*} g\left(\kappa, \mathcal{J}_{neu}(l,l^*,v^*)\right)\right]\mathrm{d}v^* - f(\eta)\right), \tag{2.28}$$

where $(\kappa^*, v^*)$ denotes the parent and $\mathcal{J}_{neu}$ is defined as in (2.3), that is

$$\mathcal{J}_{neu}(l, l^*_{neu}, v^*) = \begin{cases} \frac{1}{1-u}(l - (l^*_{neu} - v^*)) & \text{if } l > l^*_{neu}, \\ \frac{1}{1-u} l & \text{if } l < l^*_{neu}, \\ v^* & \text{if } l = l^*_{neu}. \end{cases}$$

Before stating the propositions and lemmas which prove Theorem 2.5.1, let us rewrite (2.28) in a more convenient form.

We start by observing that a Taylor expansion of $g$ gives

$$\begin{aligned} g(\kappa^*, v^*) &\prod_{(\kappa,l)\in\eta, l\neq l^*} g\left(\kappa, \mathcal{J}(l, l^*, v^*)\right) - f(\eta) \\ &= f(\eta) \sum_l \frac{g'(\kappa,l)}{g(\kappa,l)} (\mathcal{J}(l,l^*,v^*) - l) + \sum_{l,\hat{l}} \frac{g'(\hat{\kappa},\hat{l})}{g(\hat{\kappa},\hat{l})} \frac{g'(\kappa,l)}{g(\kappa,l)} \mathcal{O}\big((\mathcal{J}(l,l^*,v^*) - l)(\mathcal{J}(\hat{l},l^*,v^*) - \hat{l})\big) \\ &\qquad + \sum_l \frac{g''(\kappa,l)}{g^2(\kappa,l)} \mathcal{O}\big((\mathcal{J}^2(l,l^*,v^*) - l)\big), \end{aligned} \tag{2.29}$$

and a Taylor expansion of the exponential function about 0 leads to

$$e^{-\frac{uK}{J}\int_{v^*}^{\infty}(1-g(\kappa^*,v))\mathrm{d}v} = 1 - \frac{uK}{J}\int_{v^*}^{\infty}(1 - g(\kappa^*,v))\mathrm{d}v + \mathcal{O}\left(\frac{uK}{J}\int_{v^*}^{\infty}(1 - g(\kappa^*,v))\mathrm{d}v\right)^2. \tag{2.30}$$

Applying (2.30) we may rewrite (2.28) as

$$A^N_{neu} = A^N_{neu,1} + A^N_{neu,2} + \mathcal{O}\left(\frac{K}{J} A^N_{neu,2}\right),$$

where

$$A^N_{neu,1} = N\left(\int_0^{\infty} \frac{uK}{J} e^{-\frac{uK}{J}v^*} \left[g(\kappa^*, v^*) \prod_{(\kappa,l)\in\eta, l\neq l^*} g\left(\kappa, \mathcal{J}(l,l^*,v^*)\right) - f(\eta)\right] \mathrm{d}v^*\right), \tag{2.31}$$

$$\begin{aligned} A^N_{neu,2} = N\Bigg( &- \int_0^{\infty} \frac{u^2K^2}{J^2} e^{-\frac{uK}{J}v^*} \int_{v^*}^{\infty}(1 - g(\kappa^*,v))\mathrm{d}v \; g(\kappa^*,v^*) \\ &\times \prod_{(\kappa,l)\in\eta, l\neq l^*} g\left(\kappa, \mathcal{J}(l,l^*,v^*)\right) \mathrm{d}v^* \Bigg). \end{aligned} \tag{2.32}$$

We begin with statements of the propositions that identify limits of (2.31) and (2.32) separately. The proofs appear later in this section.

**Proposition 2.15.** *Under the conditions of Theorem 2.5.1,*

$$\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t A^N_{neu,1} f(\eta_s) - f(\eta^r_s)\left(\sum_{l_i(t)\in\eta^r_t} a l_i^2 \frac{g'(l_i(t))}{g(l_i(t))}\right)\mathrm{d}s\right|\right] \to 0.$$

**Proposition 2.16.** *Under the conditions of Theorem 2.5.1,*

$$\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t \left\{A^N_{neu,2} f(\eta_s) - 2af(\eta^r_s)\sum_{l_i(t)\in\eta^r_s}\int_{l_i(t)}^{\infty}\big(1 - g(\kappa_i,v)\big)\mathrm{d}v\right\}\mathrm{d}s\right|\right] \to 0.$$

The proof of Proposition 2.15 will use the following elementary lemma:

**Lemma 2.17.** *Let $\Gamma$ be a stochastic process. Assume that the second moment of $\Gamma$ is bounded uniformly for all times up to time $T$ by $\epsilon$, that is $\mathbb{E}[(\Gamma(s))^2] \leq \epsilon$, for $0 \leq s \leq T$. Then*

$$\mathbb{E}\left[\sup_{t<T}\left|\int_0^t \Gamma(s)\mathrm{d}s\right|\right] < \sqrt{T\epsilon}.$$

*Proof.* By Jensen's inequality

$$\mathbb{E}\left[\sup_{t\leq T}\left(\int_0^t \Gamma(s)\mathrm{d}s\right)^2\right] \leq \mathbb{E}\left[\sup_{t\leq T}\, t\int_0^t \Gamma^2(s)\mathrm{d}s\right] = T\mathbb{E}\left[\int_0^T \Gamma^2(s)\mathrm{d}s\right].$$

The inequality follows by assumption and a final application of Jensen's inequality. □

*Proof of Proposition 2.15.* By Lemma 2.17, it suffices to prove that, conditioned on the postion of rare levels,

$$\mathbb{E}\Bigg[\Bigg\{N\Bigg(\int_0^\infty \frac{uK}{J}e^{-\frac{uK}{J}v^*}\Big\{g(\kappa^*,v^*) \times \prod_{(\kappa(t),l(t))\in\eta, l(t)\neq l^*(t)} g\left(\kappa,\mathcal{J}(l(t),l^*(t),v^*)\right) - f(\eta_s)\Big\}\mathrm{d}v^*\Bigg) - f(\eta_s^r)\left(\sum_i a l_i^2 \frac{g'(l_i(t))}{g(l_i(t))}\right)\Bigg\}^2\Bigg] \to 0,$$

uniformly for $0 \leq s \leq T$.

For ease of notation we let $l_0 = 0$. We will use the ordering $l_i$ and (2.29) to observe that $A^N_{neu,1}f(\eta)$ can be approximated by

$$\begin{aligned}
&Nf(\eta)\Bigg(\sum_i \int_{l_{i-1}}^{l_i}\frac{uK}{J}e^{-\frac{uK}{J}v^*}\left[\frac{g'(\kappa_i,l_i)}{g(\kappa_i,l_i)}(v^*-l_i) + \sum_{l\neq l_i}\frac{g'(\kappa,l)}{g(\kappa,l)}\left(\frac{l\frac{u}{J}}{1-\frac{u}{J}} - \mathbf{1}_{l>l_i}\frac{l_i-v^*}{1-\frac{u}{J}}\right)\right]\mathrm{d}v^*\Bigg)\\
&= Nf(\eta)\sum_i\Bigg[\frac{g'(\kappa_i,l_i)}{g(\kappa_i,l_i)}\left(-(l_i-l_{i-1})e^{-\frac{uK}{J}l_{i-1}} + \frac{J}{uK}\left(e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i}\right)\right)\\
&\qquad + \sum_{j\neq i}\left(\frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)}\frac{l_j\frac{u}{J}}{1-\frac{u}{J}}\left(e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i}\right)\right)\\
&\quad + \sum_{j\neq i}\sum_{j<i}\frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)}\frac{1}{1-\frac{u}{J}}\left(-(l_i-l_{i-1})e^{-\frac{uK}{J}l_{i-1}} + \frac{J}{uK}\left(e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i}\right)\right)\Bigg],
\end{aligned} \tag{2.33}$$

where the second line follows from integration. At this stage we note that we can drop the factor $1-u/J$ at the cost of an error of order $NK/J^3$, which tends to zero as $N$ tends to infinity. We also swap the order of summation to rewrite (2.33) as

$$Nf(\eta)\sum_j \frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)}\Bigg[\sum_{i\leq j}\left(-(l_i-l_{i-1})e^{-\frac{uK}{J}l_{i-1}} + \frac{J}{uK}\left(e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i}\right)\right)$$

$$+ l_j \frac{u}{J} \sum_{i \neq j} \left( e^{-\frac{uK}{J} l_{i-1}} - e^{-\frac{uK}{J} l_i} \right) \Bigg].$$

We observe that since the second and the third terms in the inner sum are telescoping sums and $l_0 = 0$ we simplify our expression to

$$f(\eta) \sum_j \frac{g'(\kappa_j, l_j)}{g(\kappa_j, l_j)} \big( P_1 + P_2 + P_3 \big) :=$$

$$N f(\eta) \sum_j \frac{g'(\kappa_j, l_j)}{g(\kappa_j, l_j)} \Bigg[ \frac{J}{uK} \left( 1 - e^{-\frac{uK}{J} l_j} \right) + l_j \frac{u}{J} \left( 1 - e^{-\frac{uK}{J} l_{j-1}} + e^{-\frac{uK}{J} l_j} \right)$$
$$+ \sum_{i \leq j} \left( -(l_i - l_{i-1}) e^{-\frac{uK}{J} l_{i-1}} \right) \Bigg]. \quad (2.34)$$

Now we treat each term in the new sum separately. Using the Taylor expansion of the exponential function about 0 we observe that

$$P_1 = \frac{NJ}{uK} \left( 1 - e^{-\frac{uK}{J} l_j} \right) = N \sum_{1 \leq k \leq n+1} \frac{(-1)^{k-1}}{k!} l_j^k \left( \frac{uK}{J} \right)^{k-1} + \mathcal{O} \left( \frac{N (uK)^{n+1}}{J^{n+1}} \right),$$
$$P_2 = l_j \frac{uN}{J} \left( 1 - e^{-\frac{uK}{J} l_{j-1}} + e^{-\frac{uK}{J} l_j} \right) = \frac{uN l_j}{J} + \mathcal{O} \left( \frac{NK}{J^2} (l_j - l_{j-1}) \right),$$
$$P_3 = N \sum_{i \leq j} \left( -(l_i - l_{i-1}) e^{-\frac{uK}{J} l_{i-1}} \right) = N \sum_{1 \leq k \leq n+1} \frac{(-1)^k}{(k-1)!} \left( \frac{uK}{J} \right)^{k-1} \sum_{i \leq j} (l_i - l_{i-1}) l_{i-1}^{k-1}$$
$$+ \mathcal{O} \left( \frac{N (uK)^{n+1}}{J^{n+1}} \right).$$

We focus our attention on $P_3$. We investigate the terms corresponding to different values of $k$ separately. We observe that the first three terms involve

$$\sum_{i \leq j} (l_i - l_{i-1}) = l_j, \quad (2.35)$$

$$\sum_{i \leq j} (l_i - l_{i-1}) l_{i-1} = \frac{1}{2} \left( \sum_{i \leq j} (l_i^2 - l_{i-1}^2) - \sum_{i \leq j} (l_i - l_{i-1})^2 \right)$$
$$= \frac{1}{2} \left( l_j^2 - \sum_{i \leq j} (l_i - l_{i-1})^2 \right) \quad (2.36)$$

$$\sum_{i \leq j} \left( (l_i - l_{i-1}) l_{i-1}^2 \right) = \frac{1}{3} \sum_{i \leq j} \left( l_i^3 - l_{i-1}^3 \right) - \frac{1}{3} \sum_{i \leq j} (l_i - l_{i-1})^3 - \sum_{i \leq j} (l_i - l_{i-1})^2 l_{i-1}$$
$$= \frac{1}{3} l_j^3 - \frac{1}{3} \sum_{i \leq j} (l_i - l_{i-1})^3 - \sum_{i \leq j} (l_i - l_{i-1})^2 l_{i-1}, \quad (2.37)$$

respectively. We are therefore interested in $\sum_i (l_i - l_{i-1})^2$ conditioned on the locations of the rare levels, that is, conditioned on $\eta^r$.

Recall that conditioned on the locations of the levels of the rare type, the levels of the common type will be Poisson distributed with intensity $K - Z^N$. Conditioned on the number of levels of the rare type between 0 and $l_j$ these levels will be independent uniformly distributed on $[0,l_j]$. We denote $n$ independent uniformly distributed random variables on $[0,1]$ by $u_1,\dots,u_n$ and define the order statistics by letting $u_{(i)}=u_k$ if and only if $\#\{\hat{k}:u_{\hat{k}}\le u_k\}=i$ for $i\in\{1,\dots n\}$. We note that $n$ points uniformly distributed on $[0,1]$ can be identified with $n+1$ points uniformly distributed on the unit circle with one of these points chosen at random to be a reference point corresponding to both 0 and 1. This then leads us to see that $u_{(i)}-u_{(i-1)}$ is equal in distribution to $u_{(1)}$ for $i\in\{1,\dots,n+1\}$, where, by convention, $u_{(n+1)}:=1$ and $u_{(0)}:=0$. From this one can see that

$$\mathbb{E}\left[\sum_{i=1}^{j}(l_i-l_{i-1})^2|l_0=0,j=n+1,l_j\right]=l_j^2\mathbb{E}\left[\sum_{i=1}^{n+1}(u_{(i)}-u_{(i-1)})^2\right]=l_j^2(n+1)\frac{2}{(n+1)(n+2)}.$$

We then use that the number of levels of the rare type within $[0,l_j]$ will be Poisson distributed to see

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{j}(l_i-l_{i-1})^2|l_0=0,l_j\right]=&\sum_{i=0}^{\infty}\frac{2l_j^2}{n+2}\frac{\left(l_j(K-Z^N)\right)^n\exp(-\left(l_j(K-Z^N)\right))}{n!}\\=&\frac{2l_j}{(K-Z^N)}+\mathcal{O}\left(\exp\left(-l_j(K-Z^N)\right)\right).\end{aligned}$$

Identical calculations show

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{i=1}^{j}(l_i-l_{i-1})^2\right)^2|l_0=0,l_j\right]=&\frac{4l_j}{(K-Z^N)^3}\left(3+x(K-Z^N)\right)\\&+\mathcal{O}\left((K-Z^N)\exp\left(-l_j(K-Z^N)\right)\right),\qquad(2.38)\\\mathrm{Var}\left(\sum_{i=1}^{j}(l_i-l_{i-1})^2|l_0=0,l_j\right)=&\mathcal{O}\left(\frac{l_j}{(K-Z^N)^3}\right).\end{aligned}$$

We note that we are considering $j$ to be a random variable throughout this corresponding to the level of a given rare individual.

From this calculation since we multiply (2.36) by $NK/J$ in we see that we require $\frac{N^2}{KJ^2}\to 0$. Therefore, we may approximate $\left(\frac{uK}{J}\right)\sum_{i\le j}(l_i-l_{i-1})^2$ by $\frac{ul_j}{J}$. For $k=3$ we again condition on the number of levels of the rare type beneath $l_j$ and see that

$$\mathbb{E}\left[\sum_{i=1}^{j}(l_i-l_{i-1})^2l_{i-1}|l_0=0,j=n+1,l_j\right]=(n+1)\mathbb{E}\left[(l_I-l_{I-1})^2l_{I-1}|l_0=0,j=n+1,l_j\right],$$

where $I$ is chosen uniformly at random from $(1,\cdots,n+1)$. We then again consider the levels as $n+1$ points chosen uniformly at random from a circle to see that this will be $\frac{l_j^3}{n+2}$. This then gives us

$$\mathbb{E}\left[\sum_{i\le j}\left(l_i-l_{i-1}\right)^2l_{i-1}|l_0=0,l_j\right]=\frac{l_j^2}{K-Z^N}+\mathcal{O}\left(\exp(-l_j(K-Z^N))\right).\qquad(2.39)$$

We see that for $k\ge 3$

$$\mathbb{E}\left[\sum_{i\le j}\left((l_i-l_{i-1})l_{i-1}^k\right)-\frac{1}{k+1}l_j^{k+1}\right]=\mathcal{O}\left(\frac{1}{K}\right),\qquad(2.40)$$

$$\mathrm{Var}\left(\sum_{i\leq j}\left((l_i-l_{i-1})l_{i-1}^k\right)-\frac{1}{k+1}l_j^{k+1}\right)=\mathcal{O}\left(\frac{1}{K^3}\right), \tag{2.41}$$

which will suffice as each of these terms will be multiplied by $\frac{NK^{k-1}}{J^{k-1}}$ in (2.34). We then note that $\frac{NK}{J^2}$ is bounded and $\frac{K}{J}\to 0$. We combine (2.35), (2.36), (2.37), (2.40) to see that, conditioned on $l_j$,

$$\begin{aligned}\mathbb{E}&\left[\sum_{i\leq j}\left(-(l_i-l_{i-1})e^{-\frac{uK}{J}l_{i-1}}\right)\right]\\&=\sum_{0\leq k\leq n}\frac{(-1)^{k+1}}{(k+1)!}\left(\frac{uK}{J}\right)^k l_j^{k+1}-\frac{u}{J}l_j+\frac{1}{2}\frac{u^2K}{J^2}l_j^2+\mathcal{O}\left(\frac{K^2}{J^3}\right)+\mathcal{O}\left(\frac{K}{J}\right)^{n+1}.\end{aligned}$$

Finally, we observe that all these approximations and cancellations allow us to approximate (2.34) by

$$f(\eta)\sum_j\frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)}\left(\frac{1}{2}\frac{u^2NK}{J^2}l_j^2+\mathcal{O}\left(\frac{NK^2}{J^3}\right)+\mathcal{O}\left(\frac{NK^{n+1}}{J^{n+1}}\right)+\mathcal{O}\left(\frac{NK}{J^2}(l_j-l_{j-1})\right)\right),$$

plus a random, mean zero, correction term with variance $\mathcal{O}(N^2K^2/(J^2(K-Z)^3))$. □

The proof of Proposition 2.16 is more involved than that of Proposition 2.15. Once again we begin with an elementary lemma.

**Lemma 2.18.** *Let $\Gamma(s)$ be a stochastic process. For any partition $0=t_0<t_1<\cdots<t_m=T$,*

$$\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t\Gamma(s)\mathrm{d}s\right|\right]\leq\sum_{j=1}^m\sqrt{\mathbb{E}\left[\left(\int_{t_{j-1}}^{t_j}\Gamma(s)\mathrm{d}s\right)^2\right]}+\mathbb{E}\left[\sup_j\int_{t_{j-1}}^{t_j}|\Gamma(s)|\mathrm{d}s\right].$$

*Proof.* We observe that

$$\begin{aligned}\mathbb{E}\left[\sup_{t\leq T}\left|\int_0^t\Gamma(s)\mathrm{d}s\right|\right]&\leq\mathbb{E}\left[\sum_{j=1}^m\left|\int_{t_{j-1}}^{t_j}\Gamma(s)\mathrm{d}s\right|+\sup_j\int_{t_{j-1}}^{t_j}|\Gamma(s)|\mathrm{d}s\right]\\&\leq\sum_{j=1}^m\mathbb{E}\left[\left|\int_{t_{j-1}}^{t_j}\Gamma(s)\mathrm{d}s\right|\right]+\mathbb{E}\left[\sup_j\int_{t_{j-1}}^{t_j}|\Gamma(s)|\mathrm{d}s\right]\\&\leq\sum_{j=1}^m\sqrt{\mathbb{E}\left[\left(\int_{t_{j-1}}^{t_j}\Gamma(s)\mathrm{d}s\right)^2\right]}+\mathbb{E}\left[\sup_j\int_{t_{j-1}}^{t_j}|\Gamma(s)|\mathrm{d}s\right].\end{aligned}$$

which concludes the proof. □

We require a few more computations to transform $A^N_{neu,2}$ into a more convenient form. First of all, we observe that as $(1-g)$ is bounded, the following approximation is valid:

$$\int_{l_{i-1}}^{l_i}\int_{v^*}^{\infty}\big(1-g(\kappa_i,v)\big)\mathrm{d}v\,\mathrm{d}v^* = \int_{l_{i-1}}^{\infty}\int_{l+i-1}^{l_i\wedge v}\big(1-g(\kappa_i,v)\big)\mathrm{d}v^*\,\mathrm{d}v$$
$$= \int_{l_i}^{\infty}\int_{l+i-1}^{l_i}\big(1-g(\kappa_i,v)\big)\mathrm{d}v^*\,\mathrm{d}v + \int_{l_{i-1}}^{l_i}\int_{l+i-1}^{v}\big(1-g(\kappa_i,v)\big)\mathrm{d}v^*\,\mathrm{d}v$$
$$= (l_i-l_{i-1})\int_{l_i}^{\infty}\big(1-g(\kappa_i,v)\big)\mathrm{d}v + \mathcal{O}\left((l_i-l_{i-1})^2\right). \quad (2.42)$$

For convenience, we order the individuals present in the system according to their level, that is we consider $\eta = \{(\kappa_i, l_i)\}_{i\geq 1}$ where $l_i < l_{i+1}$. We observe that the ordering leads to the following simplification:

$$A^N_{neu,2}f(\eta) = -\frac{u^2NK}{J}\sum_i\left(\int_{l_{i-1}}^{l_i}\frac{uK}{J}e^{-\frac{uK}{J}v^*}f(\eta)\int_{v^*}^{\infty}\big(1-g(\kappa_i,v)\big)\mathrm{d}v\,\mathrm{d}v^*\right).$$

Since $\frac{K}{J}\to 0$, we may use (2.42) to further simplify $A^N_{neu,2}$ to

$$A^N_{neu,2}f(\eta) = -\frac{u^2NK^2}{J^2}f(\eta)\sum_i(l_i-l_{i-1})\left(\int_{l_i}^{\infty}\big(1-g(\kappa_i,v)\big)\mathrm{d}v\,\mathrm{d}v^*\right)\left(1+\mathcal{O}\left(\frac{K}{J}+\frac{1}{K^2}\right)\right).$$

By this calculation and Lemma 2.18, it is clear that in order to prove Proposition 2.16, it is enough to show that for any partition $0 = t_0 < t_1 < \cdots < t_m = T$,

$$\sum_{j=1}^{m}\sqrt{\mathbb{E}\left[\left(\int_{t_{j-1}}^{t_j}\Gamma^N(s)\mathrm{d}s\right)^2\right]} + \mathbb{E}\left[\sup_j\int_{t_j-1}^{t_j}|\Gamma^N(s)|\mathrm{d}s\right]\to 0, \quad (2.43)$$

where

$$\Gamma^N(s) = \sum_i\left(\frac{u^2NK^2}{J^2}(l_i(t)-l_{i-1}(t))-2a\right)f(\eta_t)\int_{l_i(t)}^{\infty}\big(1-g(\kappa_i,v)\big)\mathrm{d}v\,\mathrm{d}v^*\mathrm{d}t. \quad (2.44)$$

We first turn our attention to the parts without a supremum which we calculate directly.

**Lemma 2.19.** *Conditioned on the locations of the rare levels,*

$$\mathbb{E}\left[\left(\int_0^T\sum_i\left(\frac{u^2NK^2}{J^2}(l_i(t)-l_{i-1}(t))-2a\right)f(\eta_t)\int_{l_i(t)}^{\infty}\big(1-g(\kappa_i,v)\big)\mathrm{d}v\,\mathrm{d}v^*\mathrm{d}t\right)^2\right]$$
$$= \mathcal{O}\left(\frac{LT}{N}\right)+o(T^2), \quad (2.45)$$

*for any $L(N)$ such that $\frac{L}{J}\to\infty$.*

*Proof.* We recall that there exists $\lambda_g$ such that for all $v\geq\lambda_g$, $g(v)=1$. Therefore

$$f(\eta_t)\int_{l_i(t)}^{\infty}\big(1-g(\kappa_i,v)\big)\mathrm{d}v$$

is bounded for each test function. As $\eta_t$ is constant between events we condition on exactly $n$ reproductive events occurring within $[0,T]$. We denote the time of events by $t_i$ for $i\in\{1,\ldots,n\}$.

Observe that $t_i$ are independent random variables uniformly distributed on $[0,T]$. For ease of notation we let $t_0 = 0$ and $t_{n+1} = T$. We will use $\hat{l}_k$ to denote a rare type individual's level and $\tilde{l}_k$ to be the highest level across the population (that is, from individuals of both rare and common type) below $\hat{l}_k$ and consider

$$\mathbb{E}\left[\sum_{i=0}^{n}\sum_{j=0}^{n}(t_{i+1}-t_i)(t_{j+1}-t_j)\right.$$
$$\left.\left(\frac{u^2NK^2}{J^2}(\hat{l}_k(t_i)-\tilde{l}_k(t_i))-2a\right)\left(\frac{uNK^2}{J^2}(\hat{l}_k(t_j)-\tilde{l}_k(t_j))-2a\right)\right]. \quad (2.46)$$

We note that the times of events are independent of the level process and so we use

$$\mathbb{E}[(t_{i+1}-t_i)(t_{j+1}-t_j)] = \begin{cases} \frac{T^2}{(n+1)(n+2)} & \text{if } i \neq j, \\ \frac{2T^2}{(n+1)(n+2)} & \text{if } i = j. \end{cases}$$

We also use the Cauchy–Schwarz inequality to see

$$\mathbb{E}[(\hat{l}_k(t_i)-\tilde{l}_k(t_i))(\hat{l}_k(t_j)-\tilde{l}_k(t_j))] \leq \frac{2}{(K-Z^N)^2},$$

where we note that we are still looking at the stopped process.

We denote the intensity, after $n$ events, of the levels of all individuals which were offspring in one of those $n$ events by $I_n$. It is then easy to see $I_{n+1} = I_n(1-\frac{u}{J})+\frac{uK}{J}$ and $I_0 = 0$. This recurrence equation has a solution given by

$$I_n = K\left(1-\left(1-\frac{u}{J}\right)^n\right).$$

We now introduce some notation in order to simplify the calculations presented in this section. For $i < j$ we define $\tilde{l}_k^{new}(t_j)$ as the highest level below $\hat{l}_k$ of an individual (of either rare or common type) which has been born since $t_i$. Then we can see that

$$(\hat{l}_k(t_j)-\tilde{l}_k(t_j)) \leq (\hat{l}_k(t_j)-\tilde{l}_k^{new}(t_j)),$$

and so

$$\mathbb{E}[(\hat{l}_k(t_i)-\tilde{l}_k(t_i))(\hat{l}_k(t_j)-\tilde{l}_k(t_j))] \leq \frac{1}{K-Z^N}\frac{1}{I_{j-i}}.$$

We then also see that

$$\begin{aligned}\mathbb{E}[(\hat{l}_k(t_i)-\tilde{l}_k(t_i))(\hat{l}_k(t_j)-\tilde{l}_k(t_j))] \geq& \mathbb{E}[(\hat{l}_k(t_i)-\tilde{l}_k(t_i))(\hat{l}_k(t_j)-\tilde{l}_k(t_j))\mathbf{1}_{\tilde{l}_k(t_j)=\tilde{l}_k^{new}(t_j)}] \\ =& \mathbb{E}[(\hat{l}_k(t_i)-\tilde{l}_k(t_i))(\hat{l}_k(t_j)-\tilde{l}_k^{new}(t_j))] \\ &- \mathbb{E}[(\hat{l}_k(t_i)-\tilde{l}_k(t_i))(\hat{l}_k(t_j)-\tilde{l}_k^{new}(t_j))\mathbf{1}_{\tilde{l}_k(t_j)\neq\tilde{l}_k^{new}(t_j)}] \\ \geq& \frac{1}{I_{j-i}}\frac{1}{K-Z^N} - \sqrt{\frac{2}{(K-Z^N)^2}\frac{2}{I_{j-i}^2}\mathbb{P}[\tilde{l}_k(t_j)\neq\tilde{l}_k^{new}(t_j)]},\end{aligned}$$

where we have used the Cauchy-Schwartz inequality for the final line.

We note that

$$\mathbb{P}[\tilde{l}_k(t_j) \neq \tilde{l}_k^{new}(t_j)] \leq \frac{K\left(1-\frac{u}{J}\right)^{j-i}}{K\left(1-\frac{u}{J}\right)^{j-i}+I_{j-i}}.$$

We now consider a function $L(N)$. By splitting (2.46) into parts with $|i-j| \leq L$ and $|i-j| > L$, we bound it above by

$$\begin{aligned}
&\frac{4(n+1)LT^2}{(n+1)(n+2)}\left(\frac{N^2K^4}{J^4}\frac{2}{(K-Z^N)^2}+4a^2\right)\\
&\quad+\frac{C(n+1)(n+1-L)T^2}{(n+1)(n+2)}\Bigg\{\frac{N^2K^4}{J^4}\frac{1}{K-Z^N}\frac{1}{I_L}-\left(\frac{NK^2}{J^2}\frac{1}{K-Z^N}+\frac{NK^2}{J^2}\frac{1}{I_L}\right)2a+4a^2\\
&\qquad+\mathcal{O}\left(\frac{N^2K^4}{J^4}\frac{1}{K-Z^N}\frac{1}{I_L}\sqrt{\frac{K\left(1-\frac{u}{J}\right)^L}{K\left(1-\frac{u}{J}\right)^L+I_L}}\right)\Bigg\}.
\end{aligned}$$

We observe that

$$\begin{aligned}
\sum_{n=1}^{\infty}\frac{n+1}{(n+1)(n+2)}\frac{(NT)^ne^{-NT}}{n!} &= \frac{1}{NT}-\frac{1}{(NT)^2}+\frac{e^{-NT}}{(NT)^2}-\frac{e^{-NT}}{2},\\
\sum_{n=1}^{\infty}\frac{(n+1)^2}{(n+1)(n+2)}\frac{(NT)^ne^{-NT}}{n!} &\leq 1.
\end{aligned}$$

We may combine the two results above to see

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^T\sum_i\left(\frac{u^2NK^2}{J^2}(l_i(t)-l_{i-1}(t))-2a\right)f(\eta_t)\int_{l_i(t)}^{\infty}\big(1-g(\kappa_i,v)\big)\mathrm{d}v\ \mathrm{d}v^*\mathrm{d}t\right)^2\right]\\
&\quad=\mathcal{O}\Bigg(\frac{LT}{N}\left(\frac{N^2K^4}{J^4}\frac{2}{(K-Z)^2}+4a^2\right)\\
&\qquad+T^2\Bigg\{\frac{N^2K^4}{J^4}\frac{1}{K-Z}\frac{1}{I_L}-\left(\frac{NK^2}{J^2}\frac{1}{K-Z}+\frac{NK^2}{J^2}\frac{1}{I_L}\right)2a\\
&\qquad+4a^2+\mathcal{O}\left(\frac{N^2K^4}{J^4}\frac{1}{K-Z}\frac{1}{I_L}\sqrt{\frac{K\left(1-\frac{u}{J}\right)^L}{K\left(1-\frac{u}{J}\right)^L+I_L}}\right)\Bigg\}\Bigg).
\end{aligned}$$

Therefore, in order to conclude that (2.45) converges to 0, it is enough to find $L$ which satisfies

$$\frac{L}{N}\to 0;\quad \frac{NK^2}{J^2}\frac{1}{I_L}\to 2a;\quad \frac{K\left(1-\frac{u}{J}\right)^L}{K\left(1-\frac{u}{J}\right)^L+I_L}\to 0.$$

Since $\frac{N}{J}\to\infty$, this is achieved by any $L(N)$ such that $\frac{L}{J}\to\infty$ and $\frac{L}{N}\to 0$. For example, we may choose $L(N)=\sqrt{NJ}$.

□

We now turn our attention back to (2.43). We consider $t_j = j\delta(N)$ and see that to ensure the first part of (2.43) converges to zero, we need choices of $L$ and $\delta$ such that

$$\frac{L}{N} \to 0; \quad \frac{L}{J} \to \infty; \quad \frac{L}{N\delta} \to 0. \tag{2.47}$$

To show that the second term of (2.43) converges to zero, it is enough to show boundedness of $\mathbb{E}[\sup_{t\leq T} \Gamma^N(t)]$, provided that there exists a $\delta$ satisfying (2.47) such that $\delta \to 0$. The latter is satisfied by taking $\delta = \sqrt{L/N}$.

To show $\mathbb{E}[\sup_{t\leq T} \Gamma^N(t)]$ is bounded we consider an auxiliary process $\beta^N$, defined by

$$\beta_t^N = \frac{NK^2}{J^2} \sum_{l(t)\in\eta_t^r} \left(l(t) - \hat{l}(t)\right) h_\lambda(l(t)), \tag{2.48}$$

where $h$ is a decreasing, positive cut-off function (that is, we assume that there exists a $\lambda$ such that $h(v) = 0$ for $v > \lambda$) and $\hat{l}(t) \in \eta_t$ is the first level below $l(t)$. To show the required bound on the expectation of the supremum of the integral of $X$, (and therefore conclude the proof), we show that $\beta^N$ is dominated by a bounded submartingale. We note that under the conditions of Theorem 2.4.1 $\frac{NK^2}{J^2} = \mathcal{O}(K)$.

**Lemma 2.20.** *Define $\beta^N$ as in (2.48). Then $\beta^N$ is dominated by a bounded submartingale.*

*Proof.* We recall that the offspring in an event can be ordered $(v_1, v_2, \dots)$ with $v_1 = v^*$. Therefore, as the parent is thought of as moving to $v^*$ we will refer to $(v_i)_{i\geq 2}$ as the 'children' after an event. Recall that within the period of time $[0, t]$ with probability $\mathcal{O}(t^2)$ we will see two or more events. Therefore

$$\begin{aligned}
&\frac{\mathbb{E}\left[\sum_{l(t)\in\eta_t^r}\left(l(t)-\hat{l}(t)\right)h(l(t)) - \sum_{l(0)\in\eta_0^r}\left(l(0)-\hat{l}(0)\right)h(l(0))\right]}{t} \\
&\quad = N\int_0^\infty \frac{uK}{J} e^{-\frac{uK}{J}v^*} \Bigg\{ \sum_{l(0)\in\eta_t^r} \Bigg\{ \left[\left(\mathcal{J}(l(0)) - \mathcal{J}(\hat{l}(0))\right) - \left(l(0) - \hat{l}(0)\right)\right] h(l(0)) \\
&\qquad\qquad + \left(\mathcal{J}(\hat{l}(0)) - \hat{l}(t)\right) h(l(0)) \\
&\quad + \left(l(t) - \hat{l}(0)\right)\left(\mathcal{J}(l(0)) - l(0)\right)\mathcal{O}(||h'||_\infty)\Bigg\} + \sum_{i=2}^\infty (v_i - \hat{v}_i)\, h(v_i)\mathbf{1}_{\{v_i \text{ is rare}\}}\Bigg\} \mathrm{d}v^* + \mathcal{O}(t),
\end{aligned}$$

where we are using $\hat{v}_i$ to denote the highest level below the $i$th 'child', $v_i$. We now require that $h(l) = 0$ for all $l \geq \lambda_h$. We note that in the proof of Proposition 2.15 we have shown that the parts involving $\mathcal{J}(l(0)) - l(0)$ converge in $L^2$ to $al(0)^2$ and so we can see that this is approximated by

$$\begin{aligned}
&\sum_{l(0)\in\eta_t^r} \left\{ \left(al^2(0) - a\hat{l}^2(0)\right) h(l(0)) + \left(l(0) - \hat{l}(0)\right) al^2(0)\mathcal{O}(||h'||_\infty) \right\} \\
&\quad + N\int_0^\infty \frac{uK}{J} e^{-\frac{uK}{J}v^*} \Bigg\{ \sum_{l(0)\in\eta_t^r} \left(\mathcal{J}(\hat{l}(0)) - \hat{l}(t)\right) h(l(0)) + \sum_{i=2}^\infty (v_i - \hat{v}_i)\, h(v_i)\mathbf{1}_{\{v_i \text{ is rare}\}} \Bigg\} \mathrm{d}v^*.
\end{aligned}$$

We note that since the ordering of the $l(t) \in \eta_t$ which are not the parent is unaffected by an event, the only possibilities that will force $\mathcal{J}(\hat{l}(0)) < \hat{l}(t)$ is when a 'child' has been born between the new locations of levels or when $\hat{l}(0)$ is chosen as the parent. However we use the last sum appearing in the previous equation and the fact that $h$ is decreasing to see that the generator applied to $\beta_t^N$, for large enough $N$, is bounded below by

$$-C_h\beta_t^N + NS\frac{NK^2}{J^2}\sum_{l_i\in\eta_t^r}\int_{l_{i-2}}^{\frac{l_{i-2}}{1-\frac{u}{J}}}\frac{uK}{J}e^{-\frac{uK}{J}v^*}\left(v^*-\frac{l_{i-2}}{1-\frac{u}{J}}\right)\mathrm{d}v^*$$
$$\geq -C_h\beta_t^N - \sum_{l_i\in\eta_t^r}\mathcal{O}\left(\frac{NK^2}{J^3}l_{i-2}^2\right),$$

where $C_h = \lambda_h||h'||_\infty$. We therefore see that

$$\beta_t^N + \int_0^t C_h\beta_s^N + C\frac{NK^2Z}{J^3}\mathrm{d}s,$$

is a sub-martingale. Combined with the conditions of Theorem 2.4.1 this concludes the proof of the lemma. □

To complete the proof of Proposition 2.16 we see that, by using Lemma 2.19,

$$\mathbb{E}\left[\left(\beta_T^N + \int_0^T C_h\beta_s^N + C\frac{NK^2Z}{J^3}\mathrm{d}s\right)^2\right],$$

is bounded. Therefore, by Jensen's inequality,

$$\mathbb{E}[\sup_{t\leq T}\beta_t^N] \leq \sqrt{\mathbb{E}\left[\left(\sup_{t\leq T}\beta_t^N\right)^2\right]}$$

which is also bounded by Doob's martingale inequality. Proposition 2.16 therefore follows by an application of Lemma 2.18.

### 2.3.3 Selective model - proof of Theorem 2.5.2

The generator of the model presented in Theorem 2.5.2 is of the form

$$A^N f(\eta) = A_{neu}^N f(\eta) + A_{sel}^N f(\eta).$$

The analysis for the neutral part of the generator, $A_{neu}^N$, is the same as in Section 2.3.2, with the exception of a slight modification of Lemma 2.20, which we discuss in Lemma 2.23. We therefore turn our attention to $A_{sel}^N$, which is the part of the generator describing the selection. It can be written as

$$A_{sel}^N f(\eta) = \frac{sN}{S}\Bigg(\int_0^\infty\Bigg[\frac{uK}{J}e^{-\frac{uK}{J}v^*}g(\kappa^*,v^*)e^{-\frac{uK}{J}\int_{v^*}^\infty(1-g(\kappa^*,v))\mathrm{d}v}$$
$$\times\prod_{(\kappa,l)\in\eta,l\neq l_{sel}^*}g\left(\kappa,\mathcal{J}_{sel}(l,l_{sel}^*,v^*)\right)\Bigg]\mathrm{d}v^* - f(\eta)\Bigg), \quad (2.49)$$

where we recall that the movement of the levels $\mathcal{J}_{sel}$ is specified by

$$\mathcal{J}_{sel}((l,\kappa),(l^*_{sel},\kappa^*),v^*) = \begin{cases} v^* & \text{if } l = l^*_{sel}, \\ \frac{1}{1-\frac{u}{J}}\left(l - (l^*_{sel} - v^*)\frac{\sigma(\kappa)}{\sigma(\kappa^*)}\right) & \text{if } l \neq l^*_{sel}, l > v^*, \\ \frac{1}{1-\frac{u}{J}} l & \text{if } l < v^*. \end{cases}$$

As in Section 2.3.2, we transform $A^N_{sel}$ into a more convenient form. We use the Taylor approximation (2.30) to approximate the generator by

$$A^N_{sel} = A^N_{sel,1} + A^N_{sel,2} + \mathcal{O}\left(\frac{K}{J} A^N_{sel,2}\right),$$

where

$$\begin{aligned} &A^N_{sel,1} f(\eta) \\ &= \frac{sN}{S}\left(\int_0^\infty \frac{uK}{J} e^{-\frac{uK}{J}v^*}\left[g(\kappa^*,v^*)\prod_{(\kappa,l)\in\eta, l\neq l^*_{sel}} g\left(\kappa, \mathcal{J}_{sel}(l,l^*_{sel},v^*)\right) - f(\eta)\right]\mathrm{d}v^*\right), \qquad (2.50) \\ &A^N_{sel,2} f(\eta) = \frac{N}{S}\left(-\int_0^\infty \frac{u^2K^2}{J^2} e^{-\frac{uK}{J}v^*}\int_{v^*}^\infty (1-g(\kappa^*,v))\mathrm{d}v\; g(\kappa^*,v^*)\right. \\ &\qquad\qquad \left.\times \prod_{(\kappa,l)\in\eta, l\neq l^*_{sel}} g\left(\kappa, \mathcal{J}_{sel}(l,l^*_{sel},v^*)\right)\mathrm{d}v^*\right). \qquad (2.51) \end{aligned}$$

Once again, we treat $A^N_{sel,1}$ and $A^N_{sel,2}$ separately. We order the levels $\eta = \{(\kappa_i, l_i)\}_{i\geq 1}$ present in the system by requiring that for each $i$, $l_i < l_{i+1}$. Arguing in the same way as for the neutral case, we may conclude that the term involving $A^N_{sel,2}$ tends to 0 as $N \to \infty$, as the requirements $NK/J^2 \to C$ combined with $1/S \to 0$ imply that $NK/J^2S \to 0$. We turn our attention to $A^N_{sel,1}$. Recalling the difference between $l^*_{neu}$ and $l^*_{sel}$, we note that for large $N$ we will very rarely see $l^*_{sel} \neq l^*_{neu}$ and so we consider

$$\begin{aligned} &A^N_{sel,3} f(\eta) \\ &\quad = \frac{sN}{S}\left(\int_0^\infty \frac{uK}{J} e^{-\frac{uK}{J}v^*}\left[g(\kappa^*,v^*)\prod_{(\kappa,l)\in\eta, l\neq l^*_{neu}} g\left(\kappa, \mathcal{J}_{sel}(l,l^*_{neu},v^*)\right) - f(\eta)\right]\mathrm{d}v^*\right). \qquad (2.52) \end{aligned}$$

We will prove Proposition 2.10 by showing that $A^N_{sel,3}$ satisfies a suitable estimate and then showing that $A^N_{sel,1} - A^N_{sel,3}$ converges to 0 by virtue of Lemma 2.22.

**Lemma 2.21.** *Under the conditions of Theorem 2.5.2,*

$$\begin{aligned} \mathbb{E}\Bigg[\Bigg| &\frac{sN}{S}\left(\int_0^\infty \frac{uK}{J} e^{-\frac{uK}{J}v^*}\left[g(\kappa^*,v^*)\prod_{(\kappa,l)\in\eta_t, l\neq l^*_{neu}} g\left(\kappa, \mathcal{J}_{sel}(l,l^*_{neu},v^*)\right) - f(\eta)\right]\mathrm{d}v^*\right) \\ &\qquad - b\left(f(\eta^r_t)\sum_j \frac{g'(\kappa_i,l_i)}{g(\kappa_i,l_i)}\frac{suN}{JS} l_i\right)\Bigg|\Bigg] \to 0, \end{aligned}$$

*Proof.* Just as in Proposition 2.15, we approximate (2.52) using (2.29) by

$$
\begin{aligned}
&\frac{sN}{S} f(\eta)\Bigg( \sum_i \int_{l_{i-1}}^{l_i} \frac{uK}{J} e^{-\frac{uK}{J}v^*} \Bigg[ \frac{g'(\kappa_i,l_i)}{g(\kappa_i,l_i)}(v^*-l_i) \\
&\qquad + \sum_{l\neq l_i} \frac{g'(\kappa,l)}{g(\kappa,l)} \left( \frac{l\frac{u}{J}}{1-\frac{u}{J}} - \mathbf{1}_{l>l_i} \frac{\sigma(\kappa)}{\sigma(\kappa^*)} \frac{l_i - v^*}{1-\frac{u}{J}} \right) \Bigg] \mathrm{d}v^* \Bigg) \\
&= N f(\eta) \sum_i \Bigg[ \frac{g'(\kappa_i,l_i)}{g(\kappa_i,l_i)} \left( -(l_i - l_{i-1}) e^{-\frac{uK}{J}l_{i-1}} + \frac{J}{uK}\left( e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i} \right) \right) \\
&\qquad + \sum_{j\neq i} \left( \frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)} \frac{l_j\frac{u}{J}}{1-\frac{u}{J}} \left( e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i} \right) \right) \\
&\quad + \sum_{j\neq i}\sum_{j<i} \frac{\sigma(\kappa_j)}{\sigma(\kappa_i)} \frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)} \frac{1}{1-\frac{u}{J}} \left( -(l_i - l_{i-1}) e^{-\frac{uK}{J}l_{i-1}} + \frac{J}{uK}\left( e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i} \right) \right) \Bigg] \\
&=: \frac{\mathcal{S}_1}{S} + \frac{sN}{S} f(\eta) \sum_i \sum_{j<i} \left( \frac{\sigma(\kappa_j)}{\sigma(\kappa_i)} - 1 \right) \frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)} \frac{1}{1-\frac{u}{J}} \Big( -(l_i - l_{i-1}) e^{-\frac{uK}{J}l_{i-1}} \\
&\qquad\qquad + \frac{J}{uK}\left( e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i} \right) \Big),
\end{aligned}
$$

where we have used integration by parts and $\mathcal{S}_1$ is equal to the terms appearing in (2.33). Now we notice that $\mathcal{S}_1$ is multiplied by $1/S$ and since $S \to \infty$ it can be neglected. Consider then the second term. We can neglect the factor $1 - u/J$ at the cost of an error of order $NK/J^3$, which tends to zero. We change the order of summation to approximate it by

$$
\frac{sN}{S} f(\eta) \left( \frac{\sigma(\kappa_r)}{\sigma(\kappa_c)} - 1 \right) \sum_j \frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)} \sum_{i<j} \Bigg( -(l_i - l_{i-1}) e^{-\frac{uK}{J}l_{i-1}} + \frac{J}{uK}\left( e^{-\frac{uK}{J}l_{i-1}} - e^{-\frac{uK}{J}l_i} \right) \Bigg).
$$

Now we proceed precisely as in the proof of Proposition 2.15. The final approximation of $A^N_{sel,3}$ then takes the form

$$
f(\eta) \left( \frac{\sigma(\kappa_r)}{\sigma(\kappa_c)} - 1 \right) \sum_j \frac{g'(\kappa_j,l_j)}{g(\kappa_j,l_j)} \left( \frac{suN}{JS} l_j + \mathcal{O}\left(\frac{1}{S}\right) \right).
$$

□

We now turn our attention to $A^N_{sel,1} - A^N_{sel,3}$. We must treat the cases when the rare type is favoured or unfavoured separately. However, we show in both cases that $\mathbb{E}[\left|A^N_{sel,1} - A^N_{sel,3}\right|] \to 0$ and conclude with Jensen's inequality.

**Lemma 2.22.** *Let $l^*_{neu}$ be the lowest level above $v^*$. Let $l^*_{sel}$ be the level above $v^*$ which minimises $\frac{l-v^*}{\sigma}$. Then*

$$
\mathbb{E}\left[ \left| \frac{sN}{S} f(\eta) \int_0^\infty \frac{uK}{J} e^{-\frac{uK}{J}v^*} g(v^*,\kappa^*) \sum_l \frac{g'(l)}{g(l)} \left( \mathcal{J}_{sel}(l, l^*_{sel}, v^*) - \mathcal{J}_{sel}(l, l^*_{neu}, v^*) \right) \mathrm{d}v^* \right| \right] \to 0.
$$

*Proof.* We note that for $l < v^*$, $\mathcal{J}_{sel}(l, l^*_{sel}, v^*) = \mathcal{J}_{sel}(l, l^*_{neu}, v^*)$ as

$$\mathcal{J}_{sel}(l, l^*, v^*) = \begin{cases} \frac{1}{1-\frac{u}{J}}\left(l - (l^* - v^*)\frac{\sigma(\kappa)}{\sigma(\kappa^*)}\right) & l > v^*,\ l \neq l^*, \\ v^* & l = l^*, \\ \frac{1}{1-\frac{u}{J}} l & l < v^*, \end{cases}$$

and so we need only consider

$$\frac{sN}{S} f(\eta) \int_0^\infty \frac{uK}{J} e^{-\frac{uK}{J}v^*} g(v^*, \kappa^*) \sum_{l > v^*} \frac{g'(l)}{g(l)} \left(l^*_{neu} - l^*_{sel}\right) \frac{\sigma(\kappa)}{\sigma(\kappa^*)} \mathrm{d}v^*. \tag{2.53}$$

We proceed as in Section 2.3.2 and consider $v^*$ in the interval $[l_{i-1}, l_i]$. We note that $l^*_{neu} = l_i$ and if $l_i$ is of the favoured type then $l^*_{sel} = l_i$ also. If $l_i$ is unfavoured we denote the lowest favoured type above $l_i$ by $l^s_i$ then $l^*_{sel} \neq l_i$ if and only if

$$l_{i-1} \le v^* \le \left(l_i - \frac{\sigma_w}{\sigma_s - \sigma_w}(l^s_i - l_i)\right) \wedge l_{i-1},$$

where $\sigma_w = \sigma(\kappa_w)$, $\sigma_s = \sigma(\kappa_s)$, $\kappa_w$ is the unfavoured type and $\kappa_s$ is the favoured type. Therefore we see that (2.53) can be written as

$$\frac{sN}{S} f(\eta) \sum_{(l_i,\kappa_i)\in\eta, \kappa_i = \kappa_w} \int_{l_{i-1}}^{(l_i - \frac{\sigma_w}{\sigma_s - \sigma_w}(l^s_i - l_i)) \wedge l_{i-1}} \frac{uK}{J} e^{-\frac{uK}{J}v^*} g(v^*, \kappa^*) \sum_{l > v^*} \frac{g'(l)}{g(l)} \left(l^*_n - l^*_s\right) \frac{\sigma(\kappa)}{\sigma(\kappa^*)} \mathrm{d}v^*.$$

We can then integrate and use a Taylor expansion to see that we need only show

$$\mathbb{E}\left[\frac{sN}{S} f(\eta) \sum_j \frac{g'(l_j)}{g(l_j)} \sum_{(l_i,\kappa_i)\in\eta, l_i < l_j, \kappa_i = \kappa_w} \frac{uK}{J}(l^s_i - l_i)\left(l_i - l_{i-1} - \frac{\sigma_w}{\sigma_s - \sigma_w}(l^s_i - l_i)\right) \wedge 0\right] \to 0.$$

If the rare type is unfavoured we can then conclude by noting that $\mathbb{E}\left[\frac{suNK}{SJ}(l^s_i - l_i)(l_i - l_{i-1})\right] = \mathcal{O}\left(\frac{NK}{SJ}\frac{1}{(K-Z)^2}\right)$.

If the rare type is favoured we may bound it by

$$\sum_{\hat{l}_{k-1} < l_i < \hat{l}_k} \frac{suNK}{JS}(l_i - l_{i-1})^2 \mathbf{1}_{\{l_i - l_{i-1} - \frac{\sigma_w}{\sigma_s - \sigma_w}(\hat{l}_k - l_i) > 0\}}, \tag{2.54}$$

where we recall that $\hat{l}_k$ is the $k$th highest rare level. As before, we condition on there being $n$ individuals of the common type with levels between $\hat{l}_k$ and $\hat{l}_{k-1}$. We can then consider the $n$ levels as independent uniformly distributed rather than the ordered levels $l_i$. Suppose that $z, z_1, \ldots, z_{n-1}$ are independent uniformly distributed random variables on $[0,1]$, $Y = 1 - z$ and $X = \min\left(\{z\}, \{(z - z_i)\}_{z_i < z}\right)$. We then see that (2.54) is bounded by

$$n\frac{suKN}{JS}\mathbb{E}\left[X^2 \mathbf{1}_{X > cY}\right].$$

We now look at

$$\mathbb{E}[X^2 \mathbf{1}_{X > cY}] = \int_0^1 \mathbb{E}[X^2 \mathbf{1}_{X > c(1-u)} | z = u] \mathrm{d}u \tag{2.55}$$

$$
\begin{aligned}
&= \int_{\frac{c}{1+c}}^{1} \mathbb{E}[X^2 \mathbf{1}_{X>c(1-u)} | z = u] \mathrm{d}u \\
&= \int_{\frac{c}{1+c}}^{1} \int_{0}^{u} x^2 \mathbf{1}_{x>c(1-u)} \mathrm{d}\mathbb{P}[X = x | z = u] \mathrm{d}u \\
&= \int_{\frac{c}{1+c}}^{1} \int_{c(1-u)}^{u} x^2 \mathrm{d}\mathbb{P}[X = x | z = u] \mathrm{d}u.
\end{aligned}
$$

Since

$$
\begin{aligned}
\mathbb{P}[X > x | z = u] =& \mathbb{P}\big[[u - x, u] \cap \{z_i\} = \emptyset\big] = (1 - x)^{n-1}, \\
\mathrm{d}\mathbb{P}[X = x | z = u] =& (n - 1)(1 - x)^{n-2},
\end{aligned}
$$

we may bound (2.55) by

$$
\begin{aligned}
\mathbb{E}[X^2 \mathbf{1}_{X>cY}] \leq& \int_{\frac{c}{1+c}}^{1} \int_{c(1-u)}^{1} x^2 (n - 1)(1 - x)^{n-2} \mathrm{d}x \, \mathrm{d}u \\
=& \int_{\frac{c}{1+c}}^{1} \int_{0}^{1-c+cu} (n - 1)(1 - v)^2 v^{n-2} \mathrm{d}v \, \mathrm{d}u \\
\leq& \frac{n-1}{c} \left[ \frac{v^n}{n(n-1)} - 2\frac{v^{n+1}}{n(n+1)} + \frac{v^{n+2}}{(n+1)(n+2)} \right]_0^1 \\
=& \frac{6}{cn(n+1)(n+2)}.
\end{aligned}
$$

We can then conclude by noting that

$$
\sum_{n=1}^{\infty} \frac{1}{(n+1)(n+2)} \frac{(K - Z^N)^n e^{-(K-Z^N)}}{n!} = \mathcal{O}\left( \frac{1}{(K - Z^N)^2} + \exp(-(K - Z^N)) \right),
$$

which again leads to an error of order $\mathcal{O}\left(\frac{NK}{SJ}\frac{1}{(K-Z)^2}\right)$, completing the proof of the Lemma. □

*Proof of Proposition 2.10.* By an application of Lemma 2.17, Proposition 2.10 follows from a combination of Lemma 2.21 and Lemma 2.22. □

**Lemma 2.23.** *Define $\beta^N$ as in (2.48). Then $\beta^N$ is a dominated by a bounded submartingale.*

**Remark 2.24.** *Although the statements of Lemma 2.20 and Lemma 2.23 are the same, the behaviour of the process $\beta^N$ is subtly different, as in Lemma 2.23 the movement of the levels is affected by selective events. This difference does not effect the proof however we include them as seperate lemmas for completeness.*

*Proof.* As in the proof of Lemma 2.20 (and using the notation introduced there)

$$
\begin{aligned}
&\frac{\mathbb{E}\left[\sum_{l(t)\in\eta_t^r} \left(l(t) - \hat{l}(t)\right) h(l(t)) - \sum_{l(0)\in\eta_0^r} \left(l(0) - \hat{l}(0)\right) h(l(0))\right]}{t} \\
&\quad = N \int_0^\infty \left(1 + \frac{s}{S}\right) \frac{uK}{J} e^{-\frac{uK}{J} v^*} \Bigg\{ \sum_{l(0)\in\eta_t^r} \Bigg\{ \left[\left(\mathcal{J}(l(0)) - \mathcal{J}(\hat{l}(0))\right) - \left(l(0) - \hat{l}(0)\right)\right] h(l(0))
\end{aligned}
$$

$$+\left(\mathcal{J}(\hat{l}(0)) - \hat{l}(t)\right) h(l(0))$$

$$+\left(l(t) - \hat{l}(0)\right) \left(\mathcal{J}(l(0)) - l(0)\right) \mathcal{O}(||h'||_\infty)\Bigg\} + \sum_{i=2}^{\infty} (v_i - \hat{v}_i)\, h(v_i) \mathbf{1}_{\{v_i \text{ is rare}\}}\Bigg\} \mathrm{d}v^* + \mathcal{O}(t).$$

Considering combination of Proposition 2.15 and Proposition 2.10 we know that the parts involving $\mathcal{J}(l(0)) - l(0)$ converge to $al(0) - bl^2(0)$. Since $\frac{s}{S} \to 0$ and observing Lemma 2.22, we may proceed as in the proof of Lemma 2.20 to conclude. □

### 2.3.4 Model with selection in fluctuating environment - proof of Theorem 2.4.1

*Proof.* We begin by identifying the limit. Recall that the rescaled generator takes the form

$$A^N f(\eta, \zeta) = A^N_{neu} f(\eta, \zeta) + \hat{S} A^N_{sel} f(\eta, \zeta) + \hat{S}^2 A_{env} f(\eta, \zeta), \tag{2.56}$$

where $A^N_{neu}$ is defined as in (2.28), $A^N_{sel}$ is defined as in (2.49), and

$$A_{env} f(\eta, \xi) = \mathbb{E}_\pi[f(\eta, \xi)] - f(\eta, \xi).$$

We also use (2.15) and that $A_{neu} f_1$ is of order 1 by previous calculations. To identify the correct limit it is therefore enough to evaluate $A^N_{sel}\left(f_N(\eta, \xi)\right)$, which can be approximated as

$$A^N_{sel}\left(f_N(\eta, \xi)\right) = A^N_{sel}\left(f(\eta) \frac{suN}{JS}\left(\frac{\sigma(\kappa_r, \zeta)}{\sigma(\kappa_c, \zeta)} - 1\right) \sum_j l_j \frac{g'(\kappa_j, l_j)}{g(\kappa_j, l_j)}\right)$$

$$= \left[\frac{suN}{JS}\left(\frac{\sigma(\kappa_r, \zeta)}{\sigma(\kappa_c, \zeta)} - 1\right)\right]^2 f(\eta)$$

$$\times \left\{\left(\sum_{i \neq j} l_i l_j \frac{g'(\kappa_i, l_i,) g'(\kappa_j, l_j)}{g(\kappa_i, l_i) g(\kappa_j, l_j)}\right) + \left(\sum_j l_j \frac{g'(\kappa_j, l_j) + l_j g''(\kappa_j, l_j)}{g(\kappa_j, l_j)}\right)\right\} + \mathcal{O}\left(\frac{1}{S} + \frac{1}{K^{1/2}}\right)$$

The bound (2.19) follows directly from proofs of Propositions 2.10, 2.15, 2.16, with Lemma 2.20 modified in a way analogous to Lemma 2.23. □

# 3 SuperBrownian motion in a random environment

In this section we present a precise definition of superBrownian motion in a random environment. We begin by defining Branching Brownian Motion in a random environment and recalling the original definition of the corresponding superprocess from Mytnik (1996). Then we describe a lookdown construction for both of these models based on the ideas in Kurtz and Rodrigues (2011).

## 3.1 Definitions

Branching Brownian Motion in a random environment (BBMRE) can be described as follows. Imagine a collection of particles on $\mathbb{R}^d$. Each particle moves according to independent standard Brownian motions. Each particle, if alive, gives birth to one new particle at a time, at rate $a$. The initial offspring location is the same as that of the parent. After the birth, the offspring moves and reproduces independently of all other particles. The particles die at instantaneous rate $a - \zeta_t(x)b$, where $\zeta_t(x)$, taking values in $\{-1, 1\}$ as before, models the random environment and $x$ is the current location of the particle. We assume that $a > b > 0$. If $\zeta_t(x)$ is positive, the particle is less likely to die, if it is negative, it is more likely to do so. The evolution of the environment and particles are independent. We give a more formal definition below.

We begin by recalling the description of our environment, which is used for all models in this section. Our environment is modelled through a simple random field.

**Definition 3.1.** *Let $\Pi^{env}$ be a Poisson process with intensity $E$, dictating the times of the changes in the environment. Let $q(x,y)$ be an element of $C_0\left(\mathbb{R}^d \times \mathbb{R}^d\right)$ (continuous functions vanishing at infinity) and let $\{\xi^{(m)}(\cdot)\}_{m\geq 0}$ be a family of identically distributed random fields on $\mathbb{R}^d$ such that*

$$\mathbb{P}\left[\xi^{(m)}(x) = -1\right] = \frac{1}{2} = \mathbb{P}\left[\xi^{(m)}(x) = +1\right],$$
$$\mathbb{E}\left[\xi^{(m)}(x)\xi^{(m)}(y)\right] = q(x,y).$$

*Set $\tau_0 = 0$ and write $\{\tau_m\}_{m\geq 1}$ for the points in $\Pi^{env}$ and define*

$$\zeta(t,\cdot) := \sum_{m=0}^{\infty} \xi^{(m)}(\cdot)\mathbf{1}_{[\tau_m,\tau_{m+1})}(t).$$

Since the exact labelling of our particles is not important, we identify the particle component of the process with a counting measure, that is for a vector $\overline{x} = (x_1, \dots x_n)$

$$\mu_{\overline{x}} = \sum_i \delta_{x_i} \quad x_i \in \mathbb{R}^d.$$

We are now ready to state the definition of Branching Brownian motion in a random environment.

**Definition 3.2** (Branching Brownian motion in a random environment (BBMRE))**.** *Branching Brownian motion in the random environment $\zeta$ is the stochastic process taking values in purely atomic measures on $\{-1,1\} \times \mathbb{R}^d$ whose evolution consists of four ingredients.*

1. ***Spatial motion*** *The location of each particle, $x_i$, evolves according to a standard Brownian motion, independently of all other particles.*

2. ***Birth events*** *At exponential rate $a$ (independent for each particle), a particle gives birth to a new particle (a new particle is added to the system). The location of the offspring is the same as the location of their parent. The behaviour of the new particle after the birth is independent of all other particles.*

3. ***Death events*** *Each particle dies (is removed from the system) at instantaneous rate $a - \zeta_t(x_i)b$, where $x_i$ is its location.*

4. ***Environment changes*** *The environment evolves as described in Definition 3.1.*

Alternatively, we may define the BBMRE by the means of the generator. For a counting measure $\mu = \sum_i \delta_{x_i}$, define

$$f(\mu) = \sum_i f(x_i).$$

Let $f(\zeta,\mu) = f_0(\zeta)f_1(\mu) = f_0(\zeta)\pi_i h(x_i)$ be a function such that $h \in C_c(\mathbb{R}^d)$, that is $h$ is a continuous function with a compact support. We define the generator of the BBMRE as

$$\begin{aligned}\mathcal{L}f(\zeta,\mu) \\ &= f_1(\mu)A^{env}f_0(\zeta) + f_0(\zeta)\Bigg(\sum_{i=1}^{n} Bf_1(\mu) + \sum_i a\big(f_1(\mu_{b(x|x_i)}) - f_1(\mu)\big) \\ &\qquad + \sum_i \big(a - \zeta(x_i)b(f_1(\mu_{d(x|x_i)}) - f_1(\mu)\big)\Bigg),\end{aligned}$$

where $\mu_{b(x|x_i)}$ denotes the addition of a particle at location $x_i$, and $\mu_{d(x|x_i)}$ denotes a removal of the particle at location $x_i$.

It is well known that the high density limit for the Branching Brownian Motion gives rise to a SuperBrownian motion (see e.g. Etheridge (2000)). An analogous result (under certain scaling of the environment) was established by Mytnik (1996) for branching random walk in a random environment. The limiting object is SuperBrownian motion in a random environment (SBMRE).

**Definition 3.3** (SuperBrownian motion in a random environment (SBMRE))**.** *Let $q(x,y)$ be a covariance function which belongs to $C_0\left(\mathbb{R}^d \times \mathbb{R}^d\right)$. The superBrownian motion in a random environment is the (unique) process for which, for all $\phi \in \mathcal{D}(\Delta)$,*

$$X_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t \frac{1}{2} X_s(\Delta\phi)\mathrm{d}s \tag{3.1}$$

*is a square-integrable martingale with quadratic variation given by*

$$\langle X(\phi)\rangle_t = \int_0^t X_s(\phi^2)\mathrm{d}s + \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} q(x,y)\phi(x)\phi(y)X_s(\mathrm{d}x)X_s(\mathrm{d}y)\mathrm{d}s.$$

An equivalent characterisation of the SBMRE can be given in terms of the generator, see Theorem 4.8 in Mytnik (1996). Namely, for $f \in \bar{C}^2(R_+), \phi \in \mathcal{D}(\Delta)$ and $q(x,y)$ as in Definition 3.3, the generator is given by

$$\mathcal{L}f(\mu(\phi)) = f'(\mu(\phi))\mu(\Delta\phi) + \frac{1}{2}f''(\mu(\phi))\left(\mu(\phi^2) + \int_{\mathbb{R}^d\times\mathbb{R}^d} q(x,y)\phi(x)\phi(y)\mu(\mathrm{d}x)\mu(\mathrm{d}y)\mathrm{d}s\right).$$

We would like to point out that the process of Definition 1.3 which is obtained as a limiting behaviour of the scaled SLFV in a random environment differs from the one from Definition 3.3 by a presence of a drift term.

**Remark 3.4.** *Uniqueness of solutions to the martingale problem of Definition 3.3 is not immediately clear. It was established by Mytnik (1996) using a novel approximate duality technique, and later re-proved by the means of the log-Laplace transform by Crisan (2004). A uniqueness result for the process of Definition 1.3 is a simple consequence of Dawson's Girsanov Theorem, see Dawson (1978) and Etheridge (2000), Chapter 7.*

**Remark 3.5.** *A model similar to that in Mytnik (1996) was studied in Sturm (2003). The main difference between the two is in the behaviour of the environment. For the limiting model in Sturm (2003),* (3.1) *is again a martingale, but with quadratic variation of the form*

$$\langle X(\phi)\rangle_t = \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} q(x,y)\phi(x)\phi(y)X_s(\mathrm{d}x)X_s(\mathrm{d}y)\mathrm{d}s.$$

*In this case, the density of the process can be described as a solution to an SPDE in all dimensions, whereas the SBMRE has a density only in dimension one and the analogous SPDE has no solution in dimensions $d \geq 2$.*

**Remark 3.6.** *An alternative construction of the SBMRE has been suggested by Nakashima (2015). The construction in this paper is based on the model introduced by Birkner et al. (2005).*

## 3.2 Lookdown representation for BBMRE and SBMRE

In this section we describe a new construction of the SBMRE, inspired by the constructions of Kurtz and Rodrigues (2011). As a by-product of this construction, we provide a lookdown construction for the BBMRE. A precise statement of the results is given in Theorem 3.7.1 and Theorem 3.7.2. We will use the lookdown representation of SBMRE in Section 4 to describe the behaviour of the 'rare' type (by which we mean a new mutation establishing in the population) in the Spatial Lambda-Fleming-Viot model with fluctuating selection.

In order to motivate what follows, let us informally describe the construction of SBMRE from Mytnik (1996). The SBMRE is constructed via a series of approximations. Consider a sequence of mean 0 random fields $\zeta_k$, taking values in $\{-1,1\}$, with correlation as in Definition 3.1 (Mytnik (1996) considered a more general class of random fields, but this one is sufficient for our purposes). At stage $n$ of the approximation, we start with a population with size of order $n$. Over each time interval $(k/n,(k+1)/n), k \in \mathbb{N}$, each individual, if alive, moves independently of all the others, according to a standard Brownian motion. At times $k/n$, each individual either splits into two with probability $1/2+\zeta_k(x_i)/\sqrt{n}$, or dies with probability $1/2-\zeta_k(x_i)/\sqrt{n}$. The state of the environment is resampled with each reproduction event. If for each set B we define

$$X_t^n(B) = \frac{\text{number of particles in } B \text{ alive at time } t}{n},$$

and assume that $X_0^n$ converges to some $X_0$, then passing to the limit as $n$ tends to infinity the process $X^n$ converges to a SBMRE with initial condition $X_0$. Intuitively, this procedure corresponds to increasing the rate of the branching events in BBMRE by $n$, while scaling down the impact of the environment by $\sqrt{n}$. Observe that Mytnik (1996) considers a model with non-overlapping generations. This is for purely technical reasons.

We now move to a description of a lookdown construction for BBMRE (with overlapping generations). Fix $\lambda \in \mathbb{R}$. Consider a system of particles in the geographical space $\mathbb{R}^d$. Each particle moves independently, according to a standard Brownian motion. Each particle $i$ is assigned a level, $l_i$. The level takes values in $[0,\lambda]$.

In order to take environmental fluctuations into account we use the process $\zeta_t(x) \in \{-1,1\}$, introduced in Definition 3.1, to model the environment. The branching rate of the particles depends on both their level and the state of the environment. We assume that every particle gives birth at instantaneous rate $2a(\lambda - l_i(t))$, with the location of offspring being the same as the location of the parent. The initial level of the offspring is distributed uniformly on the interval $[l_i,\lambda]$. The levels of the particles evolve according to an ODE with a random coefficient

$$\frac{\mathrm{d}l_i}{\mathrm{d}t} = al_i^2 - \zeta(x_i)\sqrt{\lambda}bl_i.$$

The particle dies when its level reaches $\lambda$. We are now ready to define the process of interest.

**Definition 3.7** (Lookdown representation of Branching Brownian motion in a random environment)**.** *The lookdown representation of BBMRE is the process taking values in* $(\mathbb{R}^d\times\mathbb{R})^\infty\times\{-1,1\}$ *with dynamics specified by four components.*

1. ***Spatial motion*** *The spatial location of each particle,* $x_i \in \mathbb{R}^d$*, evolves according to a standard Brownian motion with generator* $B=\frac{1}{2}\Delta$*.*

2. ***Birth events*** *Each particle gives birth at instantaneous rate $2a(\lambda - l_i(t))$, where $l_i$ is the level of the particle. The spatial location of the offspring is the same as the location of their parent, $x_i$. The level of the offspring is chosen uniformly at random from $[l_i, \lambda]$.*

3. ***Level movement*** *The level of each of the particles evolves according to the equation*

$$\frac{\mathrm{d}l_i}{\mathrm{d}t} = al_i^2 - \zeta(x_i)\sqrt{\lambda}bl_i.$$

4. ***Environment changes*** *The environment evolves as described in Definition 3.1.*

A formal definition, via the generator of the process, is given in (3.4). It is convenient to identify the process with the counting measure

$$\mu_{\overline{x},\overline{l}} = \sum_i \delta_{x_i,l_i}.$$

We now state the first result of this section.

**Theorem 3.7.1** (Lookdown construction for Branching Brownian Motion in random environment)**.** *Let $\mu_0$ be the measure associated with the initial state of the population. Let $\gamma : \mathcal{N}\big(\mathbb{R}^d \times [0,\lambda)\big) \to \mathcal{M}(\mathbb{R}^d)$ be given by*

$$\gamma\left(\sum_i \delta_{x_i,l_i}\right) = \frac{1}{\lambda}\sum_{l_i} \delta_{x_i}.$$

*The lookdown process of Definition 3.7 corresponds to BBMRE, in the sense that if $\eta(\overline{x},\overline{l})$ is a solution to the martingale problem for the process defined in Definition 3.7, then $\gamma\left(\eta(\overline{x},\overline{l})\right)$ is the solution to the martingale problem for the process of Definition 3.2.*

Our next aim is to write down the generator of the process obtained by passing to the limit $\lambda \to \infty$, which would correspond to passing to the limit with $n$ passing to infinity in the sequence of approximations described at the beginning of this section. In order to formulate the result for the limiting process, we need to consider a special class of test functions of the form

$$f(\zeta, x, l) = f_0(\zeta)f_1(x,l) = f_0(\zeta)\prod_i g(x_i, l_i), \tag{3.2}$$

with the additional requirement

$$g(x_i, l_i) = 1 \text{ for } l_i > \lambda_g,$$

which ensures that we ignore all individuals with levels above $\lambda_g$. We are now in a position to state the main result of this section.

**Theorem 3.7.2** (Lookdown construction for superBrownian motion in random environment)**.** *Consider the process with generator given by*

$$A_\infty(f_1) = f_1(x,l)\sum_i \frac{Bg(x_i,l_i)}{g(x_i,l_i)} + f_1(x,l)\sum_i 2a\int_{l_i}^{\lambda}(g(x_i,v)-1)\mathrm{d}v + f_1(x,l)\sum_i (al_i^2 - b^2 l_i)\frac{\partial_l g(x_i,l_i)}{g(x_i,l_i)}$$

$$- f_1(x,l)\sum_i b^2 l_i \left( \sum_{j\neq i} q(x_i,x_j) l_j \frac{\partial_l g(x_i,l_i)\partial_l g(x_j,l_j)}{g(x_i,l_j)g(x_i,l_i)} + \frac{l_i \partial_l^2 g(x_i,l_i)}{g(x_i,l_i)} \right). \quad (3.3)$$

*Let $\mu_0$ be the measure associated with the initial state of the population in the limit. Define the map $\gamma : \mathcal{N}\big(\mathbb{R}^d \times [0,\infty)\big) \to \mathcal{M}(\mathbb{R}^d)$ by*

$$\gamma\left(\sum_i \delta_{x_i,l_i}\right) = \begin{cases} \lim_{\lambda\to\infty} \frac{1}{\lambda}\sum_{l_i\leq\lambda}\delta_{x_i} & \text{if the limit exists,} \\ \mu_0 & \text{otherwise .} \end{cases}$$

*The process described by the limiting generator (3.3) corresponds to the SBMRE of Definition 1.3, in the sense that if $\eta(\overline{x},\overline{l})$ is a solution to the martingale problem for the process described by the limiting generator (3.3), then $\gamma(\eta(\overline{x},\overline{l}))$ is a solution to the martingale problem for the process in Definition 1.3.*

We now wish to specify the generator for the process in Definition 3.7. In general, we consider the set of test functions of the form (3.2). To make the functional setup more precise, we need to borrow the following condition from Kurtz and Rodrigues (2011), which guarantees that the assumptions of the Markov Mapping Theorem are satisfied.

**Condition 3.8** (Based on Kurtz and Rodrigues (2011), Condition 3.1)**.** *We assume that the following conditions on the operator $B$, the coefficients $a, b$ and the test functions $g$ are satisfied.*

1. *The operator $B$ is defined on a subset of the space of bounded continuous functions, and its domain is closed under multiplication and separating.*

2. *The test functions are of the form (3.2), where*

$$g(x,l) = \prod_{j=1}^{m} (1 - g_1^j(x) g_2^j(l))$$

   *and $g_1^j \in \mathcal{D}(B)$ and $g_2^j$ are twice differentiable with support in $[0,\lambda]$. Moreover,*

$$0 \leq g_1^j g_2^j < \rho_g < 1.$$

3. *There exists a continuous, non-negative function $\psi_B$ such that for every test function $g$, and for every $x \in \mathbb{R}^d$*

$$\sup_l |Bg(x,l)| \leq c_g \psi_B(x)$$

   *for some constant $c_g$ which depends only on the function $g$.*

4. *The following bound holds for the test functions:*

$$\int_0^\infty |g(x,l) - 1| \mathrm{d}l + \sup_l \{l + l^2\} \partial_l g(x,l) \leq c_g \psi_B(x)$$

   *for some constant $c'_g$ which depends only on the function $g$.*

5. *$a > 0$, $\lambda a - \sqrt{\lambda} b > 0$.*

**Remark 3.9.** *We notice that the conditions for the generator of the motion $B$ are satisfied by the Laplacian, the generator of the Brownian motion. This is the only generator that we consider. We provide the general construction for the sake of completeness and to highlight potential extensions of this work.*

The generator $A_\lambda$ of the process of Definition 3.7 can be written as

$$\begin{aligned} A_\lambda f(\zeta, x, l) = f_1(x, l) A_\lambda^{env} f_0(\zeta) + f(\zeta, x, l) \sum_i \frac{Bg(x_i, l_i)}{g(x_i, l_i)} \\ + f(\zeta, x, l) \sum_i 2a \int_{l_i}^{\lambda} (g(x_i, v) - 1) \mathrm{d}v \\ + f(\zeta, x, l) \sum_i (a l_i^2 - \sqrt{\lambda} \zeta(x_i) b l_i) \frac{\partial_{l_i} g(x_i, l_i)}{g(x_i, l_i)}. \end{aligned} \tag{3.4}$$

A naive limiting procedure does not lead to a well defined object, since the form of the generator does not take into account the cancellations coming from fluctuations in the environment. In order to identify the correct limit, we once again use the separation of timescales trick. Proofs of Theorem 3.7.1 and Theorem 3.7.2 are presented in Appendix D.

**Remark 3.10.** *Our computations do not lead to any surprising results - we show that the SBMRE can be obtained as a scaling limit of BBMRE. Our arguments, combined with tightness of the sequence of projected processes, guarantees convergence of the sequence of projected models. We therefore provide a new construction for the SBMRE. However, we still refer to results described in Remark 3.4 to guarantee the uniqueness of solutions to the projected martingale problem. The question of uniqueness of solutions to the martingale problem for the limiting process with levels (which, by the Markov Mapping Theorem would guarantee uniqueness of solutions to the projected model) will be pursued elsewhere.*

## 4 Scaling limits of the SLFV - dynamics of the rare type

In this section we are interested in a spatial analogue of the results from Section 2. We show that under a certain scaling, the dynamics of the subpopulation with a rare mutation, which is a part of a population evolving according to a version of the Spatial-Lambda-Fleming-Viot model with selection in a fluctuating environment, is given by a superBrownian motion in a random environment. Since under our scaling the proof of the result does not differ significantly from the one discussed in Section 2, our discussion will be rather brief and focus on highlighting the differences and required modifications.

We begin with a description of the model. As before, we consider a population with two genetic types, rare and common, which we denote by $\kappa_r$ and $\kappa_c$, respectively. We also consider the random field $\zeta$, specified by Definition 3.1. Since in our model we consider a population with a countable number of individuals, the state of the population can be represented as

$$\eta = \sum_{(x,\kappa,l)} \delta_{x,\kappa,l}.$$

We assume that $\eta$ is a conditionally Poisson system with Cox measure $\Xi(\mathrm{d}x, \mathrm{d}\kappa) \times m_{leb}(\mathrm{d}l)$.

The evolution of the population is determined by reproduction events of two types - neutral and selective driven by independent Poisson point processes $\Pi^{neu}$ and $\Pi^{sel}$, which specify the time, location, impact and radius of the events. They are analogous to those in Section 2, but now we assume that an event has a location and radius, only individuals within the ball of given radius centred at the location of the event are affected. A rigorous definition of the model follows.

**Definition 4.1** (Lookdown representation of SLFVSRE)**.** *Let $\mu$ be a measure on $(0,\infty)$ and for each $r \in (0,\infty)$, let $\nu_r$ be a probability measure on $[0,1]$, such that the mapping $r \mapsto \nu_r$ is measurable and*

$$\int_{(0,\infty)} r^d \int_{[0,1]} u \, \nu_r(\mathrm{d}u)\mu(\mathrm{d}r) < \infty. \tag{4.1}$$

*Fix $s \in [0,1]$. Let $\Pi^{neu}, \Pi^{sel}$ be a pair of independent Poisson processes with intensity measures $(1-s)\mathrm{d}t \otimes \mathrm{d}y \otimes \mu(\mathrm{d}r)\nu_r(\mathrm{d}u)$ and $s\mathrm{d}t \otimes \mathrm{d}y \otimes \mu(\mathrm{d}r)\nu_r(\mathrm{d}u)$ respectively. Let $\Pi^{env}$ be a Poisson process independent of $\Pi^{neu}, \Pi^{sel}$.*

*The lookdown representation of SLFVSRE is a process taking values in purely atomic measures on $\mathbb{R}^d \times \mathbb{R} \times \{\kappa_r, \kappa_c\} \times \{-1, 1\}$ with dynamics described as follows.*

1. *If $(t, y, r, u) \in \Pi^{neu}$*
   (a) *a group of new individuals with levels $(v_1, v_2, \dots,)$ is added to the population within the ball $B_r(y)$. Their levels are distributed according to a Poisson process with intensity $u$.*
   (b) *Let $v^* = \min\{v_1, v_2, \dots,\}$. The type of the new individuals is chosen to be the same as the type of the individual with the lowest level above $v^*$ within the ball $B_r(y)$.*
   (c) *As a result of an event the individual originally with level, $l$, and position, $x$, has a new level given by*

$$\mathcal{J}_{neu}(l, l^*, v^*, x, (y,r)) = \begin{cases} l & \text{if } x \notin B_r(y), \\ \frac{1}{1-\frac{u}{J}}(l - (l^* - v^*)) & \text{if } l > l^*, x \in B_r(y), \\ \frac{1}{1-\frac{u}{J}} l & \text{if } l < l^*, x \in B_r(x), \\ v^* & \text{if } l = l^*, x \in B_r(y). \end{cases}$$

2. *If $(t, y, r, u) \in \Pi^{sel}$*
   (a) *a group of new individuals with levels $(v_1, v_2, \dots,)$ is added to the population. Their levels distributed according to a Poisson process with intensity $u$.*
   (b) *Let $v^* = \min\{v_1, v_2, \dots,\}$. The type of the new individuals is chosen to be the same as the type of the individual with level above $v*$ minimizing $(l_i - v^*)/\sigma(\kappa_i, \zeta)$ within the ball $B(x,r)$.*
   (c) *As a result of an event the individual originally with level, $l$, and position, $x$, has a new level given by*

$$\mathcal{J}_{sel}(l, l^*, v^*, x, (y,r)) = \begin{cases} l & \text{if } x \notin B_r(y), \\ \frac{1}{1-\frac{u}{J}}(l - (l^* - v^*)\frac{\sigma(\kappa,\zeta)}{\sigma(\kappa^*,\zeta)}) & \text{if } l > l^*, x \in B_{r/M}(y), \\ \frac{1}{1-\frac{u}{J}} l & \text{if } l < l^*, x \in B_r(y), \\ v^* & \text{if } l = l^*, x \in B_r(y). \end{cases}$$

3. *The dynamics of $\Pi^{env}$ are specified by Definition 3.1.*

Definition 4.1 is more general than we require, however, we include it to underline the possibility of extending our results. However, in the interest of keeping our notation as simple as possible, from now on we shall specialise to fix the radius and impact of reproduction events.

**Assumption 4.2.** *From now on, fix $R \in (0, \infty)$ and $\bar{u} \in (0,1)$ and take*

$$\mu(dr) = \delta_R(\mathrm{d}r), \qquad \nu_r(\mathrm{d}u) = \delta_{\bar{u}}(\mathrm{d}u).$$

The integrability condition (4.1) is trivially satisfied for our model with fixed radius and impact.

### 4.1 Scaling and statement of main results

As in Section 2 we record two theorems which are a by-product of our technique.

**Theorem 4.2.1.** *Suppose that $X_0^N$ is absolutely continuous with respect to Lebesgue measure, that the support* $\mathrm{supp}(X_0^N) \subseteq D$*, where $D$ is a compact subset of $\mathbb{R}^d$ (independent of $N$), and that $X_0^N$ converges weakly to $X_0$. Furthermore suppose that the intensity of selective events is* 0 *and as $N$ tends to infinity,*

$$\frac{C_d u r^{d+2} N}{JM^2} \to C_1; \quad J, K, M \to \infty; \quad \frac{K}{JM^d} \to 0; \quad \frac{N^2}{M^d K J^2} \to 0; \quad \frac{u^2 V_R N K}{J^2 M^d} \to a,$$

*In addition, assume that there exists an $n$ such that, as $N$ tends to infinity, $N(K/J)^n \to 0$. Then the sequence $X_N(t)$ converges weakly to superBrownian motion without drift, initial condition $X_0$, diffusion parameter $C_1$ and quadratic variation parameter $2a$.*

**Theorem 4.2.2.** *Suppose that $X_0^N$ is absolutely continuous with respect to Lebesgue measure, that the support* $\mathrm{supp}(X_0^N) \subseteq D$*, where $D$ is a compact subset of $\mathbb{R}^d$ (independent of $N$), and that $X_0^N$ converges weakly to $X_0$. Furthermore, suppose that $\widehat{S} = 1$, $\sigma(\kappa, \zeta) = \sigma(\kappa)$ and as $N$ tends to infinity,*

$$\frac{N}{JM^2} \to C_1 \quad J, K, M \to \infty; \quad \frac{K}{JM^d} \to 0; \quad \frac{N^2}{M^d K J^2} \to 0; \quad \frac{u^2 V_R N K}{J^2 M^d} \to a;$$
$$\frac{suNV_R}{JS}\left(\frac{\sigma(\kappa_r)}{\sigma(\kappa_c)} - 1\right) \to b,$$

*In addition, assume that there exists an $n$ such that, as $N$ tends to infinity, $N(K/J)^n \to 0$. Then the sequence $X_N(t)$ converges weakly to a critical superBrownian motion with initial condition $X_0$, diffusion parameter $C_1$, growth rate $b$ and quadratic variation parameter $2a$.*

The strategy of the proof is analogous to that laid out in Section 2.3. We therefore focus on describing the differences. In contrast to the situation described in the Remark 2.11 we do not claim to show any results on the limits of sequences of lookdown representations. We focus on the result for the projected version of the model. We divide the generator of the lookdown representation into two separate parts - one part describes the spatial movement of the particles, the other describes the evolution of the levels. We show that the selective events (and therefore the fluctuations in the direction of selection) do not affect the movement of the particles in the limit. For the convergence of the neutral model we refer to results of Chetwynd-Diggle and Etheridge (2018). We then use the lookdown representation to deduce the right form of the limiting generator and to justify the separation of timescales trick for the selective part of the generator. The part of the proof which deals with the evolution of the levels is analogous to that of Section 2, and we do not repeat it here. For the readers convenience we include intensity estimate and discuss the splitting of the generator.

Once again we consider a stopped process, see Remark 2.12. The domain of the generator is specified by (E.2), (E.3).

## 4.2 Intensity estimate

As in Section 2.3.1 we observe that the generator of the projected process is given by

$$
\begin{aligned}
\mathcal{L}^N f(\langle\phi, X_0^N\rangle) = NM^d\Bigg[&\int_{\mathbb{R}^d}\int_{B_r^M(x)} \frac{1}{|B_r^M|} w_0^N(z) f\Bigg(K\frac{u}{J}\int_{B_r^M(x)}\phi(y)\mathrm{d}y \\
&+ K\left(1-\frac{u}{J}\right)\int_{B_r^M(x)}\phi(y)w_0^N(y)\mathrm{d}y\Bigg) - f\Bigg(K\int_{B_r^M(x)}\phi(y)w_0^N(y)\mathrm{d}y\Bigg)\mathrm{d}z\mathrm{d}x \\
&+\int_{\mathbb{R}^d}\int_{B_r^M(x)}\left(1-w_0^N(z)\right)f\Bigg(K\left(1-\frac{u}{J}\right)\int_{B_r^M(x)}\phi(y)w_0^N(y)\mathrm{d}y\Bigg) \\
&- f\Bigg(K\int_{\mathbb{R}^d}\phi(y)w_0^N(y)\mathrm{d}y\Bigg)\mathrm{d}z\ \mathrm{d}x\Bigg] \\
+ N\frac{s\widehat{S}}{S}M^d\Bigg[&\int_{\mathbb{R}^d}\int_{B_r^M(x)}\frac{1}{|B_r^M|}\frac{\sigma(\kappa_r,\zeta)w_0^N(z)}{\sigma(\kappa_r,\zeta)w_0^N(z)+\sigma(\kappa_s,\zeta)\left(1-w_0^N(z)\right)} f\Bigg(K\frac{u}{J}\int_{B_r^M(x)}\phi(y)\mathrm{d}y \\
&+ K\left(1-\frac{u}{J}\right)\int_{B_r^M(x)}\phi(y)w_0^N(y)\mathrm{d}y\Bigg) - f\Bigg(K\int_{B_r^M(x)}\phi(y)w_0^N(y)\mathrm{d}y\Bigg)\mathrm{d}z\mathrm{d}x \\
&+\int_{\mathbb{R}^d}\int_{B_r^M(x)}\frac{\sigma(\kappa_c,\zeta)(1-w_0^N(z))}{\sigma(\kappa_r,\zeta)w_0^N(z)+\sigma(\kappa_s,\zeta)\left(1-w_0^N(z)\right)} f\Bigg(K\left(1-\frac{u}{J}\right)\int_{B_r^M(x)}\phi(y)w_0^N(y)\mathrm{d}y\Bigg) \\
&- f\Bigg(K\int_{\mathbb{R}^d}\phi(y)w_0^N(y)\mathrm{d}y\Bigg)\mathrm{d}z\ \mathrm{d}x\Bigg]. \qquad (4.2)
\end{aligned}
$$

The first two terms represent the part of the generator describing the effect of the neutral events and is the same as in Chetwynd-Diggle and Etheridge (2018). The last two terms represent the effects of selective events. Substituting $Kw_0^N = X_0^N$ this becomes

$$
\begin{aligned}
\mathcal{L}^N f(\langle\phi, X_0^N\rangle) &= \mathcal{L}^N_{neu} f(\langle\phi, X_0^N\rangle) + \mathcal{L}^N_{sel} f(\langle\phi, X_0^N\rangle) \\
NM^d\Bigg[&\int_{\mathbb{R}^d}\int_{B_r^M(x)}\frac{1}{|B_r^M|}\frac{X_0^N(z)}{K} f\Bigg(K\frac{u}{J}\int_{B_r^M(x)}\phi(y)\mathrm{d}y \\
&+\left(1-\frac{u}{J}\right)\int_{B_r^M(x)}\phi(y)X_0^N(y)\mathrm{d}y\Bigg) - f\Bigg(\int_{B_r^M(x)}\phi(y)X_0^N(y)\mathrm{d}y\Bigg)\mathrm{d}z\mathrm{d}x \\
&+\int_{\mathbb{R}^d}\int_{B_r^M(x)}\left(1-\frac{X_0^N(z)}{K}\right) f\Bigg(\left(1-\frac{u}{J}\right)\int_{B_r^M(x)}\phi(y)X_0^N(y)\mathrm{d}y\Bigg) \\
&- f\Bigg(\int_{\mathbb{R}^d}\phi(y)X_0^N(y)\mathrm{d}y\Bigg)\mathrm{d}z\ \mathrm{d}x\Bigg] \\
+ N\frac{s\widehat{S}}{S}M^d\Bigg[&\int_{\mathbb{R}^d}\int_{B_r^M(x)}\frac{1}{|B_r^M|}\frac{\sigma(\kappa_r,\zeta)\frac{X_0^N(z)}{K}}{\sigma(\kappa_r,\zeta)\frac{X_0^N(z)}{K}+\sigma(\kappa_s,\zeta)\left(1-\frac{X_0^N(z)}{K}\right)} f\Bigg(K\frac{u}{J}\int_{B_r^M(x)}\phi(y)\mathrm{d}y \\
&+\left(1-\frac{u}{J}\right)\int_{B_r^M(x)}\phi(y)X_0^N(y)\mathrm{d}y\Bigg) - f\Bigg(\int_{B_r^M(x)}\phi(y)X_0^N(y)\mathrm{d}y\Bigg)\mathrm{d}z\mathrm{d}x
\end{aligned}
$$

$$+ \int_{\mathbb{R}^d}\int_{B_r^M(x)} \frac{\sigma(\kappa_c,\zeta)\left(1-\frac{X_0^N(z)}{K}\right)}{\sigma(\kappa_r,\zeta)\frac{X_0^N(z)}{K}+\sigma(\kappa_s,\zeta)\left(1-\frac{X_0^N(z)}{K}\right)} f\Bigg(\left(1-\frac{u}{J}\right)\int_{B_r^M(x)}\phi(y)X_0^N(y)\mathrm{d}y\Bigg)$$
$$- f\Bigg(\int_{\mathbb{R}^d}\phi(y)X_0^N(y)\mathrm{d}y\Bigg)\mathrm{d}z\,\mathrm{d}x\Bigg]. \quad (4.3)$$

We state useful lemmas from Chetwynd-Diggle and Etheridge (2018), which describe the form and estimates for the neutral part of the generator.

**Definition 4.3.** *We denote by $\mathcal{A}^N$ the operator*

$$\mathcal{A}^N(\phi) := \frac{C(d)Nur^{d+2}}{JM^2}\Delta\phi,$$

*with $C(d) := \int_{\mathcal{B}_1(0)} x^2\mathrm{d}x$.*

**Lemma 4.4** (Chetwynd-Diggle and Etheridge (2018), Lemma 4.2)**.** *For $f(x,\zeta) = x$ and $\phi_s(x) : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R} \in C_0^{2,3}$,*

$$\int_0^t \mathcal{L}_{neu}^N(\langle\phi, X^N(s)\rangle)\mathrm{d}s = \int_0^t \left\langle X_s^N, \dot{\phi}_s\right\rangle + \left\langle X_s^N, \mathcal{A}^N(\phi_s)\right\rangle\mathrm{d}s + N_t^N(\phi), \quad (4.4)$$

*where*

$$|N_t^N(\phi)| \le \mathcal{O}\left(\frac{N\sup_{0\le s\le t}\|\phi_s\|_{C^3}}{JM^3}V_R\right)\int_0^t \left\langle X_s^N, \mathbf{1}\right\rangle\mathrm{d}s. \quad (4.5)$$

We proceed to the proof of Lemma 4.5. We notice that the proof is a simple combination of our proof of Lemma 2.8 and the proof of Chetwynd-Diggle and Etheridge (2018), Lemma 5.6.

**Lemma 4.5.** *Let $X = Kw$ denote the total intensity of individuals of the rare type. Assume that $\mathbb{E}[X^N(0)] < \infty$. Then for any $T > 0$*

$$\sup_{t\le T}\sup_N \mathbb{E}[\langle X^N, 1\rangle] <\infty, \quad (4.6)$$

$$\lim_{H\to\infty}\sup_N \mathbb{P}\left[\sup_{t\le T}\langle X^N, 1\rangle \ge H\right] =0. \quad (4.7)$$

*Proof.* We observe that with the test functions chosen as in Lemma 4.4, the part of the generator describing the change in the population resulting from a selective event can be written as

$$\mathcal{L}_{sel}^N(\langle\phi, X^N(s)\rangle) =$$
$$N\frac{us\widehat{S}}{SJ}M^d\Bigg[\int_{\mathbb{R}^d}\frac{1}{|B_r^M|}\int_{B_r^M(x)}\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_r,\zeta)\frac{X}{K}+\sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)}\frac{X(z)}{K}\int_{B_r^M(x)}(K-X(y))\,\phi(y)\mathrm{d}y\mathrm{d}z\mathrm{d}x$$
$$-\int_{\mathbb{R}^d}\frac{1}{|B_r^M|}\int_{B_r^M(x)}\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_r,\zeta)\frac{X}{K}+\sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)}\left(1-\frac{X(z)}{K}\right)\int_{B_r^M(x)}X(y)\phi(y)\mathrm{d}y\mathrm{d}z\mathrm{d}x\Bigg]$$
$$= N\frac{us\widehat{S}}{SJ}M^d\Bigg[\int_{\mathbb{R}^d}\frac{1}{|B_r^M|}\int_{B_r^M(x)}\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_r,\zeta)\frac{X}{K}+\sigma(\kappa_c,\zeta)\left(1-\frac{X}{K}\right)}X(z)\int_{B_r^M(x)}\phi(y)\mathrm{d}y\mathrm{d}z\mathrm{d}x$$

$$- \int_{\mathbb{R}^d} \frac{1}{|B_r^M|} \int_{B_r^M(x)} \int_{B_r^M(x)} X(y)\phi(y)\mathrm{d}y\mathrm{d}z\mathrm{d}x \Bigg]. \quad (4.8)$$

We observe that by Taylor's Theorem $\phi(y)$ can be locally approximated by

$$\phi(y) = \phi(z) + \nabla\phi(z)(y-z) + \|\phi\|_{C^2(\mathbb{R}^d)}\mathcal{O}\left(|y-z|^2\right).$$

Therefore by using a calculation analogous to (2.25) and the fact that $|y-z| < R/M$ within a ball $B_r(x)$ we may approximate (4.8) by

$$\mathcal{L}_{sel}^N(\langle\phi, X^N(s)\rangle) = \\ N\frac{us\widehat{S}}{SJM^{2d}}\left(\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)} - 1\right)\langle X,\phi\rangle + \frac{1}{K}\langle X,\phi\rangle^2\mathcal{O}\left(C_\sigma N\frac{us\widehat{S}}{SJM^{2d}}\right) + \mathcal{O}\left(\frac{N\widehat{S}}{SJM}\right), \quad (4.9)$$

where $C_\sigma$ is a constant depending on $\sigma$.

Let $\chi^N(t) = \mathbb{E}[\langle X_t^N, 1\rangle]$. Let $h_R$ denote a sequence of smooth functions such that $h_R$ is supported on the ball of radius $2R$ centred at zero and is equal to 1 on the ball of radius $R$ centred at zero.

Assume in addition that the sequence $h_R$ satisfies

$$\Delta h_R \leq \epsilon; \quad h_R \leq h_{R+1}.$$

We combine (4.4) and (4.9) and take expectation in (4.3) to obtain

$$\begin{aligned}\mathbb{E}\left[\langle X_t^N,\phi\rangle\right] &= \mathbb{E}\left[\langle X_0^N,\phi\rangle\right] + \mathbb{E}\left[\int_0^t \left\langle X_s^N, \dot{\phi}_s\right\rangle + \left\langle X_s^N, \mathcal{A}^N(\phi_s)\right\rangle \mathrm{d}s\right] \\ &\quad + N\frac{us\widehat{S}}{SJM^{2d}}\mathbb{E}\left[\left(\frac{\sigma(\kappa_r,\zeta)}{\sigma(\kappa_c,\zeta)} - 1\right)\int_0^t \langle X,\phi\rangle + \frac{C_\sigma}{K}\langle X,\phi\rangle^2 \mathrm{d}s\right] + \mathbb{E}\left[N_t^N(\phi)\right] \\ &\leq \chi^N(0) + C\|\Delta h_R\| + \mathcal{O}\left(\frac{N\sup_{0\leq s\leq t}\|\phi_s\|_{C^3}}{JM^3}V_R + N\frac{us\widehat{S}C_\sigma}{SJM^{2d}}\right)\chi^N(0), \quad (4.10)\end{aligned}$$

where we have used the properties of $h_R$. Letting $R$ tend to infinity and using the Monotone Convergence Theorem, we arrive at

$$\chi^N(t) \leq \chi^N(0) + C\|\Delta h_R\| + \mathcal{O}\left(\frac{N\sup_{0\leq s\leq t}\|\phi_s\|_{C^3}}{JM^3}V_R + N\frac{us\widehat{S}C_\sigma}{SJM^{2d}}\right)\chi^N(0).$$

We now apply Grönwall's inequality to conclude. The second part of the statement follows exactly as the second part of the proof of Lemma 2.8. □

### 4.3 Sketch of the proof of Theorem 1.3.1

The generator of the SLFVRE process is given by

$$A^N f(\eta) = NM^d \int_{\mathbb{R}^d}\left(\int_0^\infty \left[\frac{uK|B_r|}{JM^d} e^{-\frac{uK|B_r|}{JM^d}v^*} \hat{g}_{y,r/M}(\kappa^*, v^*) e^{-\frac{uK|B_r|}{JM^d}\int_{v^*}^\infty (1-\hat{g}_{y,r/M}(\kappa^*,v))\mathrm{d}v} \right.\right. \\ \left.\left. \prod_{(x,\kappa,l)\in\eta, l\neq l^*} g\Big(x,\kappa,\mathcal{J}_{neu}\big(l,l^*,v^*,x,(y,r/M)\big)\Big)\right]\mathrm{d}v^* - f(\eta)\right)\mathrm{d}y$$

$$+ N\frac{s\widehat{S}}{S}M^d \int_{\mathbb{R}^d} \left( \int_0^\infty \left[ \frac{uK|B_r|}{JM^d} e^{-\frac{uK|B_r|}{JM^d}v^*} \hat{g}_{y,r/M}(\kappa^*, v^*) e^{-\frac{uK|B_r|}{JM^d}\int_{v^*}^\infty (1-\hat{g}_{y,r/M}(\kappa^*,v))\mathrm{d}v} \right.\right.$$
$$\left.\left. \prod_{(x,\kappa,l)\in\eta, l\neq l^*} g\Big(x,\kappa,\mathcal{J}_{sel}\big(l,l^*,v^*,x,(y,r/M)\big)\Big)\right]\mathrm{d}v^* - f(\eta)\right)\mathrm{d}y + A^{env}, \quad (4.11)$$

where $A^{env}$ is specified as in (1.1) and $\hat{g}_{y,r/M}(\kappa,l) := \frac{1}{|B_{r/M}|}\int_{B_{r/M}(y)} g(z,\kappa,l)\mathrm{d}z$.

We split the generator into three parts, by adding and subtracting $g(y,\kappa^*,v)$ inside the integral. The first two parts describes the movement of the levels.

$$A^N_{full,neu}f(\eta)$$
$$= NM^d \int_{\mathbb{R}^d} \left( \int_0^\infty \left[ \frac{uK|B_r|}{JM^d} e^{-\frac{uK|B_r|}{JM^d}v^*} g(x^*,\kappa^*,v^*) e^{-\frac{uK|B_r|}{JM^d}\int_{v^*}^\infty (1-\hat{g}_{y,r/M}(\kappa^*,v))\mathrm{d}v} \right.\right.$$
$$\left. \prod_{(x,\kappa,l)\in\eta_{B_{r/M}(y)}, l\neq l^*} g\Big(x^*,\kappa,\mathcal{J}_{neu}\big(l,l^*,v^*,x,(y,r/M)\big)\Big) \prod_{(x,\kappa,l)\notin\eta_{B_{r/M}(y)}, l\neq l^*} g\Big(x,\kappa,l\Big)\right]\mathrm{d}v^*$$
$$\left. - f(\eta)\right)\mathrm{d}y, \quad (4.12)$$

where $\eta_{B_{r/M}(y)} := \{(x,\kappa,l)\in\eta : x\in B_{r/M}(y)\}$ and

$$A^N_{full,sel}f(\eta)$$
$$= N\frac{s\widehat{S}}{S}M^d \int_{\mathbb{R}^d} \left( \int_0^\infty \left[ \frac{uK|B_r|}{JM^d} e^{-\frac{uK|B_r|}{JM^d}v^*} g(x^*,\kappa^*,v^*) e^{-\frac{uK|B_r|}{JM^d}\int_{v^*}^\infty (1-\hat{g}_{y,r/M}(\kappa^*,v))\mathrm{d}v} \right.\right.$$
$$\left. \prod_{(x,\kappa,l)\in\eta_{B_{r/M}(y)}, l\neq l^*} g\Big(x^*,\kappa,\mathcal{J}_{sel}\big(l,l^*,v^*,x,(y,r/M)\big)\Big) \prod_{(x,\kappa,l)\notin\eta_{B_{r/M}(y)}, l\neq l^*} g\Big(x,\kappa,l\Big)\right]\mathrm{d}v^*$$
$$\left. - f(\eta)\right)\mathrm{d}y, \quad (4.13)$$

Observe that $\widehat{g})_{y,R/M}$ is symmetrical with respect to $y$. Therefore (4.12), (4.13) can be treated using arguments which are analogous to those which we applied to $A_{neu,1}, A_{neu,2}, A_{sel,1}$ and $A_{sel,2}$ in Section 2. The second two parts are given by

$$A^N_{full,neu,2}f(\eta) = NM^d \int_{\mathbb{R}^d} \left( \int_0^\infty \left[ \frac{uK|B_r|}{JM^d} e^{-\frac{uK|B_r|}{JM^d}v^*} \Big(\hat{g}_{y,r/M}(\kappa^*,v^*) - g(y,\kappa^*,v^*)\Big) \right.\right.$$
$$\times\, e^{-\frac{uK|B_r|}{JM^d}\int_{v^*}^\infty (1-\hat{g}_{y,r/M}(\kappa^*,v))\mathrm{d}v}$$
$$\left.\left. \times \prod_{(x,\kappa,l)\in\eta, l\neq l^*} g\Big(x,\kappa,\mathcal{J}_{neu}\big(l,l^*,v^*,x,(y,r/M)\big)\Big)\right]\mathrm{d}v^*\right)\mathrm{d}y \quad (4.14)$$

and

$$A^N_{full,sel,2}f(\eta) = N\frac{s\widehat{S}}{S}M^d \int_{\mathbb{R}^d} \left( \int_0^\infty \left[ \frac{uK|B_r|}{JM^d} e^{-\frac{uK|B_r|}{JM^d}v^*} \Big(\hat{g}_{y,r/M}(\kappa^*,v^*) - g(y,\kappa^*,v^*)\Big) \right.\right.$$

$$\times\, e^{-\frac{uK|B_r|}{JM^d}\int_{v^*}^{\infty}(1-\hat{g}_{y,r/M}(\kappa^*,v))\mathrm{d}v}$$
$$\times \prod_{(x,\kappa,l)\in\eta,l\neq l^*} g\Big(x,\kappa,\mathcal{J}_{sel}\big(l,l^*,v^*,x,(y,r/M)\big)\Big)\Bigg]\mathrm{d}v^*\Bigg)\mathrm{d}y. \quad (4.15)$$

We observe that our scaling implies that if the contribution from (4.14) is non-negligible, the contribution from (4.15) vanishes in the limit. For that reason the results of Chetwynd-Diggle and Etheridge (2018) are sufficient to deduce the behaviour of the spatial movement of the generator of all cases of interest. From

$$A^N_{full,sel,2}f(\eta) = NM^d\int_{\mathbb{R}^d}\int_0^\infty \frac{uK|B_r|}{JM^d}e^{-\frac{uK|B_r|}{JM^d}v^*}\left(\hat{g}_{y,r/M}(\kappa^*,v^*)-g(x^*,\kappa^*,v^*)\right)$$
$$\times\, e^{-\frac{uK|B_r|}{JM^d}\int_{v^*}^{\infty}(1-\hat{g}_{y,r/m}(\kappa^*,v))\mathrm{d}v}\prod_{(x,\kappa,l)\in\eta,l\neq l^*}g(x,\kappa,\mathcal{J})\mathrm{d}v^*\mathrm{d}y,$$

we can see as in the calculations in Section 2 that this leads to

$$\alpha A^N_{full,sel,2}h(\Xi)$$
$$= NM^d\int_{\mathbb{R}^d\times\mathcal{K}}\frac{M^d}{|B_r|}\int_{B_{r/M}(x^*)} e^{-\int_{\mathbb{R}/B_{r/M}(y)}h(z,\kappa)\Xi(\mathrm{d}z,\mathrm{d}\kappa)}e^{-\left(1-\frac{u}{J}\right)\int_{B_{r/M}(y)}h(z,\kappa)\Xi(\mathrm{d}z,\mathrm{d}\kappa)}$$
$$\times\left[e^{-\frac{uK|B_r|}{JM^d}\hat{h}_{y,r/M}(\kappa^*)}-e^{-\frac{uK|B_r|}{JM^d}h(x^*,\kappa^*)}\right]\mathrm{d}y\;\Xi(\mathrm{d}x^*,\mathrm{d}\kappa^*).$$

Performing a simple Taylor expansion and noting this is identical to calculations appearing in Chetwynd-Diggle and Etheridge (2018), Section 4.3 we see that this will give us

$$\alpha A^N_{full,sel,2}f(\Xi) = \exp(-\langle h,\Xi\rangle)\left\langle\frac{C(d)ur^dN}{JM^2}\Delta h+\mathcal{O}(1/M),\Xi\right\rangle,$$

where we recall that $C(d) := \int_{B_1}x^2\mathrm{d}x$ and $\langle h,\Xi\rangle := \int_{\mathbb{R}^d\times\mathcal{K}}h(x,\kappa)\Xi(\mathrm{d}x,\mathrm{d}\kappa)$.

For the sake of completeness we state the propositions describing the evolution of the levels. We observe that both (4.12) and (4.13) are integrals over compact sets (recall that $g(x,l)$ is equal to 1 outside of a compact set) of their non-spatial counterparts studied in Section 2.3. The task of analysing this generator is then a simple expansion of arguments in that section. In particular, the sparsity condition $K/M^d\to\infty$ leads to following analogous of Proposition 2.9 and Proposition 2.10.

**Proposition 4.6.** *Under the conditions of Theorem 4.2.1,*

$$\mathbb{E}\left[\sup_{t\le T}\left|\int_0^t A^N_{full,neu}f(\eta_s)\right.\right.$$
$$\left.\left.-\left(f(\eta^r_s)\sum_{l_i(s)\in\eta^N_s}al_i^2\frac{\partial_l g(x_i,l_i(s))}{g(x_i,l_i(s))}+2af(\eta^r_s)\sum_{l_i(s)\in\eta^r_s}\int_{l_i(s)}^\infty\big(1-g(x_i,v)\big)\mathrm{d}v\right)\mathrm{d}s\right|\right]\to 0.$$

**Proposition 4.7.** *Under the conditions of Theorem 4.2.2 for any $T\in\mathbb{R}$*

$$\mathbb{E}\left[\sup_{t\le T}\left|\int_0^t A^N_{full,sel}f(\eta^N_s)-f(\eta^N_s)\left(\sum_{l_i(t)\in\eta^N_t}-bl_i\frac{g'(x_i,l_i)}{g(x_i,l_i)}\right)\mathrm{d}s\right|\right]\to 0.$$

The proofs of Theorem 1.3.1, Theorem 4.2.1 and Theorem 4.2.2 now follow by applying Theorem 2.7.1 to the process characterised by the generator of the projected version of $A^N_{full,neu}$, $A^N_{full,sel}$ and $A_{env}$.

# A Poisson random measures

In this section we present some facts about Poisson random measures. Most of the facts presented in this section have been stated in the papers on lookdown constructions, see, for example Kurtz and Rodrigues (2011), Etheridge and Kurtz (2018). We present them again as they are useful for many calculations involving lookdown constructions.

**Lemma A.1.** *Let $\xi$ be a Poisson random measure with mean measure $\nu$. Let $f \in L^1(\mathbb{R}^d, \nu)$. Then*

$$\mathbb{E}\left[\exp\left(\int f(z)\xi(\mathrm{d}z)\right)\right] = \exp\left(\int (e^{f(x)} - 1)\nu(\mathrm{d}x)\right)$$

*Similarly, the expected value and variance of the integral with respect to a Poisson random measure is given by*

$$\mathbb{E}\left[\int f(z)\xi(\mathrm{d}z)\right] = \int f(x)\nu(\mathrm{d}x) \quad Var\left[\int f(z)\xi(\mathrm{d}z)\right] = \int f^2(x)\nu(\mathrm{d}x)$$

**Definition A.2** (Conditionally Poisson system)**.** *Consider a counting measure $\xi$ on $\mathbb{R}^d$. Let $\Xi$ be a locally finite random measure on $\mathbb{R}^d$. We say that $\xi$ is conditionally Poisson with Cox measure $\Xi$ if, conditioned on $\Xi$, $\xi$ is a Poisson random measure with mean measure $\Xi$.*

Conditionally Poisson systems are sometimes referred to as Cox processes. We notice that to check that a Poisson random measure $\xi$ is actually a Cox process, it is enough to check that

$$\mathbb{E}\left[\exp\left(-\int_{\mathbb{R}^d} f \mathrm{d}\xi\right)\right] = \mathbb{E}\left[\exp\left(-\int_{\mathbb{R}^d} (1 - e^f)\mathrm{d}\Xi\right)\right]$$

for all positive Borel-measurable functions $f$. Our main application of the presented theory is to show convergence of the particles systems to their high intensity limits. We need some more definitions, as the convergence of sequences of conditionally Poisson systems requires a rather exotic topology.

Consider a family of continuous functions $h_k : \mathbb{R}^d \to [0, 1]$ such that

$$\bigcup_k S_{h_k} = \mathbb{R}^d,$$

where $S_f$ denotes the support of $f$. Let $\mathcal{M}_{h_k}(\mathbb{R}^d)$ be the collection of Borel measures on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} h_k f \mathrm{d}\nu < \infty.$$

Let $\mathrm{d}\nu^k = h_k \mathrm{d}\nu$. The space $\mathcal{M}_{h_k}(\mathbb{R}^d)$ endowed with the topology of weak convergence of $\mathrm{d}\nu^k$ is metrizable. We observe that checking convergence in $\mathcal{M}_{h_k}(\mathbb{R}^d)$ is equivalent to checking convergence of $\int_{\mathbb{R}^d} f \mathrm{d}\nu_n$ for all bounded and continuous functions which satisfy

$$\int_{\mathbb{R}^d} f \mathrm{d}\nu^k < \infty \text{ for } f \leq ch_k$$

for some constant $c > 0$. The space $\mathcal{M}_{h_k}(\mathbb{R}^d \times [0, \infty))$ can be defined in a similar way.

**Theorem A.2.1** (Kurtz and Rodrigues (2011), Theorem A.9)**.** *Let $\xi_n$ be a sequence of conditionally Poisson random measures on $\mathbb{R}^d \times [0, \infty)$ with Cox measures $\{\Xi_n \times \Lambda\}$. Then $\xi_n \Rightarrow \xi$ in $\mathcal{M}_{h_k}(\mathbb{R}^d \times [0, \infty))$ if and only if $\Xi_n \Rightarrow \Xi$ in $\mathcal{M}_{h_k}(\mathbb{R}^d)$ . If the limit exists, $\xi$ is a conditionally Poisson random measure with Cox measure $\Xi \times \Lambda$.*

# B Markov Mapping Theorem

We recall some basic definitions and introduce the necessary notation. For a detailed account of this introductory material we refer to Ethier and Kurtz (1986), Chapter 1, and Lunardi (2012).

Let $(E, d)$, $(E_0, d_0)$ be a pair of complete, separable metric spaces (with metrics $d$ and $d_0$, respectively). Let $B(E)$ be the space of bounded measurable functions on $E$. We notice that equipped with the usual supremum norm $\| \cdot \|_\infty$ $B(E)$ forms a Banach space. Let $C(E) \subset B(E)$ denote the subspace of continuous functions on $E$. A subspace $A$ of $B(E) \times B(E)$ is a multivalued linear operator. It domain is given by $\mathcal{D} = \{f : (f, g) \in A\}$ and its range by $\mathcal{R} = \{g : (f, g) \in A\}$.

**Definition B.1** (Dissipative operator)**.** *We say that the operator $A$ is dissipative if for each $(f, g) \in A$ and $\lambda > 0$*

$$\|\lambda f - g\| \geq \lambda \|f\|.$$

**Definition B.2** (Graph separable pre-generator)**.** *We say that an operator $A \subset B(E) \times B(E)$ is a pre-generator if it is dissipative and there exists a sequence of functions $\mu_n$ mapping $E$ to the set of probability measures over $\mathcal{P}(E)$, and a sequence of $\lambda_n \in E$, such that for each $(f, g) \in A$*

$$g(x) = \lim_{n \to \infty} \lambda_n \int_E (f(y) - f(x))\, \mu_n(x, \mathrm{d}y).$$

*If in addition there exists a countable subset $\{f_k\} \subset \mathcal{D}(A) \bigcup C(E)$ such that the graph of $A$ is contained in the closure of the linear span of $(f_n, Af_n)$, we say that it is graph-separable.*

We notice that the generators of Markov process are graph-separable pre-generators. Let $D_E[0, \infty)$ denote the space if càdlàg functions and $M_E[0, \infty)$ denote the space of Borel measurable functions from $[0, \infty)$ taking values in $E$.

**Theorem B.2.1** (Kurtz and Rodrigues (2011), Theorem A.15)**.** *Let $A \subset \overline{C}(E) \times C(E)$ and let $\psi$ be a continuous function taking values in $\mathbb{R}$ such that $\psi \geq 1$. Suppose that for each $f \in \mathcal{D}(A)$ there exists a constant $c_f > 0$ such that*

$$|Af(x)| \leq c_f \psi(x).$$

*Let $A_0$ be defined as*

$$A_0 f(x) = \frac{Af(x)}{\psi(x)}.$$

*Suppose that $A_0$ is graph-separable pre-generator and suppose that $\mathcal{D}(A) = \mathcal{D}(A_0)$ is closed under multiplication and separating. Let $\gamma : E \to E_0$ be Borel measurable, and let $\alpha$ be a transition function from $E_0$ into $E$ satisfying $\alpha(y, \gamma^{-1}(y)) = 1$. Assume that for each $y \in E_0$*

$$\widetilde{\psi} = \int_E \psi(y, z) \alpha(y, \mathrm{d}z) < \infty.$$

*and define*

$$C = \left\{ \left( \int_E f(z) \alpha(\cdot, \mathrm{d}z), \int_E Bf(z) \alpha(\cdot, \mathrm{d}z) \right) : f \in \mathcal{D}(B) \right\}.$$

*Let $\mu_0 \in \mathcal{P}(E_0)$, and define $\nu_0(y) = \int_{\mathbb{R}} \alpha(y, \cdot) \mu_0(\mathrm{d}y)$.*

1. *Let $\tilde{Y}$ be a solution of a martingale problem for $(C, \mu_0)$. Assume that it satisfies the moment condition*

$$\int_0^t \mathbb{E}\left[\widetilde{\psi}(\tilde{Y}(s))\right] \mathrm{d}s < \infty \quad \forall t \geq 0. \tag{B.1}$$

*Then there exists a solution $X$ of the martingale problem for $(A, \nu_0)$ such that $\tilde{Y}$ has the same distribution on $\mathcal{M}_{E_0}[0, \infty)$ as $\widetilde{Y} = \gamma \circ Y$.*

2. *If, in addition, uniqueness holds for the martingale problem for $(A, \nu_0)$, then uniqueness holds for the martingale problem for $(C, \mu_0)$.*

The original proof of the theorem was inspired by the proofs of generalisations of Burke's Output Theorem appearing in Kliemann et al. (1990) and the proof of equivalence of martingale problems for the Moran model and its lookdown representation in Donnelly and Kurtz (1996).

Since then the Markov Mapping Theorem has been a useful tool in mathematical population genetics (Etheridge and Kurtz (2018), Kurtz and Rodrigues (2011)), mathematical biology (Gupta (2012)), mathematical finance (Stockbridge (2002)) and analysis of infinite dimensional stochastic differential equations (Kurtz (2010)).

The main power of the Markov Mapping Theorem comes in simplifications of proofs of equivalence of seemingly different martingale problems. The main source of the power is in exploitation of properties of exchangeable process and conditionally Poisson systems.

# C Kurtz-Rodrigues' Martingale Lemma

The following Lemma plays an important role in our applications of the Markov Mapping Theorem. Intuitively, it clarifies why the averaged process is a solution to a projected martingale problem.

**Lemma C.1** (Kurtz and Rodrigues (2011), Lemma A.13)**.** *Let $\{\mathcal{F}_t\}$ and $\{\mathcal{G}_t\}$ be filtrations with $\mathcal{G}_t \subset \mathcal{F}_t$. Suppose that for each $t \geq 0$*

$$\mathbb{E}\big[|X_t| + \int_0^t |Y_s| \mathrm{d}s\big] < \infty.$$

*and that*

$$M_t = X_t - \int_0^t Y_s \mathrm{d}s$$

*is an $\mathcal{F}_t$-martingale. Then*

$$\widehat{M_t} = \mathbb{E}\big[X_t|\mathcal{G}_t\big] - \int_0^t \mathbb{E}\big[Y_s|\mathcal{G}_s\big] \mathrm{d}s$$

*is a $\{\mathcal{G}_t\}$martingale.*

# D Proofs of Theorem 3.7.1 and Theorem 3.7.2

*Proof of Theorem 3.7.1.* Before we can proceed, additional objects need to be defined. Let $\alpha_\lambda(n, l)$ be the joint distribution of $n$ i.i.d. uniformly distributed random variables on $[0, \lambda]$. Recall that

$\mu$ denotes a point measure representing positions of individuals. For a test function $f_1(x_i, l_i)$, we define the projection onto type space $\hat{f}$ as

$$\hat{f}(\mu) = \prod_i \hat{g}(x_i) = e^{-\sum_i \mathcal{I}(g(x_i))},$$

where the average for a single level is defined as

$$e^{-\mathcal{I}(g(x_i))} = \hat{g}(x_i) = \frac{1}{\lambda}\int_0^\lambda g(x_i, z)\mathrm{d}z.$$

To calculate the generator of the projected model (the generator averaged over the distribution of the levels), we need to evaluate

$$\int A_\lambda f(\zeta, x, l)\alpha_\lambda(\mathrm{d}l).$$

Let us integrate the four terms appearing in $A_\lambda$ separately. We begin with the two terms which are least involved - the movement of particles and the environment. Since both of those terms do not depend on the levels, integrals with respect to them do not alter our projections, namely

$$\int f(\zeta, x, l)\sum_i \frac{Bg(x_i, l_i)}{g(x_i, l_i)}\alpha_\lambda(\mathrm{d}l) = \sum_{i=1}^n B\hat{f}(\zeta, \mu) = nB\hat{f}(\zeta, \mu). \tag{D.1}$$

Analogously,

$$\int \lambda f_1(x, l)A_\lambda^{env} f_0(\zeta, x)\alpha_\lambda(n, \mathrm{d}l) = \lambda\hat{f}_1(\mu)A_\lambda^{env} f_0(\zeta, \mu). \tag{D.2}$$

In order to evaluate terms describing births and movement of the levels, which both do depend on the exact value of the level, it is convenient to note that (here we follow the calculation on p. 492 in Kurtz and Rodrigues (2011))

$$\lambda^{-1}2a\int_0^\lambda g(x, z)\int_z^\lambda (g(x, v) - 1)\mathrm{d}v\mathrm{d}z = a\lambda e^{-\mathcal{I}_g} - 2a\lambda^{-1}\int_0^\lambda g(x, z)(\lambda - z)\mathrm{d}z, \tag{D.3}$$

where we have used Fubini's Theorem, and

$$\begin{aligned}\lambda^{-1}\int_0^\lambda (az^2 - \zeta bz)g'(x, z)\mathrm{d}z &= -\lambda^{-1}\int_0^\lambda (2az - \zeta b)(g(z) - 1)\mathrm{d}z \\ &= \lambda^{-1}2a\int_0^\lambda zg(x, z)\mathrm{d}z + a\lambda + b(e^{-\mathcal{I}_g} - 1),\end{aligned} \tag{D.4}$$

where we have integrated by parts. It will also be useful to describe the changes in our system due to births and deaths. Whenever a birth event occurs, the new individual is located at the same place as the parent. If a death occurs, the individual is just removed from the system. Therefore, if we denote the new collection of particles after a birth at location $y$ by $(b(\overline{x}|y))$ and the new collection of particles after a death at location $x_j$ by $d(\overline{x}|x_j)$, we see that

$$\mu_{b(\overline{x}|y)} = \delta_y + \sum_{i=1}^n \delta_{x_i}, \quad \mu_{d(\overline{x}|x_j)} = -\delta_{x_j} + \sum_{i=1}^n \delta_{x_i}.$$

Armed with these observations and identities (D.3), (D.4) we proceed to evaluate the remaining terms. A simple calculation shows that

$$\begin{aligned}
&\int f(\zeta,x,l)\left\{\sum_i 2a\int_{l_i}^{\lambda}(g(x_i,v)-1)\mathrm{d}v+\sum_i(al_i^2-\sqrt{\lambda}\zeta(x_i)bl_i)\frac{\partial_{l_i}g(x_i,l_i)}{g(x_i,l_i)}\right\}\alpha_\lambda(\mathrm{d}l)\\
&\qquad=\sum_j e^{\sum_{i\neq j}\mathcal{I}_g(x_i)}\left\{2a\lambda e^{-\mathcal{I}_g(x_j)}+2a\lambda^{-1}\int_0^\lambda g(x_j,z)\mathrm{d}z\right.\\
&\qquad\left.-2a\lambda^{-1}\int_0^\lambda g(x_j,z)\mathrm{d}z+a\lambda+\sqrt{\lambda}b\zeta(x_j)(e^{\mathcal{I}_g(x_j)}-1)\right\}\\
&=a\lambda e^{\sum_i\mathcal{I}_g(x_i)}\sum_j(e^{-\mathcal{I}_g(x_j)}-1)+\sum_j(\lambda a-\sqrt{\lambda}b\zeta(x_j))e^{\sum_{i\neq j}\mathcal{I}_g(x_i)}(1-e^{\mathcal{I}_g(x_j)})\\
&\qquad=\sum_i\lambda a(\hat{f}(\mu_{b(x|x_i)})-\hat{f}(\mu))+\sum_i(\lambda a-\sqrt{\lambda}\zeta(x_i)b)(\hat{f}(\mu_{d(x|x_i)})-\hat{f}(\mu)).
\end{aligned}\tag{D.5}$$

Combining (D.1), (D.2) and (D.5) we have established that the projected generator can be written as

$$\begin{aligned}
\mathcal{L}_\lambda\hat{f}(\zeta,\mu)=\lambda\hat{f}_1(\mu)A_\lambda^{env}f_0(\xi)+\sum_i B_{x_i}\hat{f}(\mu)+\sum_i\lambda a\big(\hat{f}(\mu_{b(x|x_i)})-\hat{f}(\mu)\big)\\
+\sum_i\big(\lambda a-\zeta(x_i)\sqrt{\lambda}b\big)\big(\hat{f}(\mu_{d(x|x_i)})-\hat{f}(\mu)\big),
\end{aligned}$$

which is the generator of the BBMRE with birth rate $\lambda a$ and death rate $(\lambda a-\sqrt{\lambda}\zeta b)$, as claimed.

Now we only need to check that all assumptions of the Markov Mapping Theorem are satisfied. Fortunately our Condition 3.8 has been imposed to guarantee just that. The map $\gamma:\mathcal{N}(\mathbb{R}^d\times[0,\lambda))\to\mathcal{M}(\mathbb{R}^d)$ (mapping counting measures to measures on $\mathbb{R}^d$) is given by

$$\gamma\left(\sum_i\delta_{x_i,l_i}\right)=\frac{1}{\lambda}\sum_{l_i}\delta_{x_i}$$

The moment condition (B.1) is satisfied if we consider $\psi$ of the form

$$\psi(x,l)=1+\sum_i\psi_B(x_i)\,(1+a+b)\,e^{-l_i},\tag{D.6}$$

so that the averaged $\tilde{\psi}$ is of the form

$$\tilde{\psi}(x)=1+\sum_i\psi_B(x_i)\,(1+a+b)\,(1-e^{-\lambda}).\tag{D.7}$$

We note that the 1 appearing in the definitions of $\psi$ and $\tilde{\psi}$ has been added only to ensure that both of these functions are greater than or equal to 1. □

*Proof of Theorem 3.7.2.* We define a test function of the form

$$h_1(\zeta,x,l)=-f_1(x,l)b\sum_i\zeta(x_i)l_i\frac{\partial_l g(x_i,l_i)}{g(x_i,l_i)}$$

and apply the generator (3.4) to a test function of the form $G = f_1 + \frac{1}{\sqrt{\lambda}}h_1$. This leads to

$$
\begin{aligned}
A_\lambda\left(f_1(x,l) + \frac{1}{\sqrt{\lambda}}h_1\right) &= f_1(x,l)\sum_i \frac{Bg(x_i,l_i)}{g(x_i,l_i)} \\
&+ f_1(x,l)\sum_i 2a\int_{l_i}^{\lambda}(g(x_i,v)-1)\mathrm{d}v + f(x,l)\sum_i (al_i^2 - \sqrt{\lambda}\zeta(x_i)bl_i)\frac{\partial_l g(x_i,l_i)}{g(x_i,l_i)} \\
&+ \sqrt{\lambda}f_1(x,l)\sum_i \zeta(x_i)l_i\frac{\partial_l g(x_i,l_i)}{g(x_i,l_i)} \\
&- \frac{1}{\sqrt{\lambda}}\left\{f_1(x,l)b\zeta(x_i)\sum_i\left[\frac{Bg(x_i,l_i)}{g(x_i,l_i)} + \frac{B\partial_l g(x_i,l_i)}{g(x_i,l_i)}\right]\right. \\
&+ \left[\sum_i \zeta(x_i)l_i\frac{\partial_l g(x_i,l_i)}{g(x_i,l_i)}\right]f_1(x,l)\sum_i 2a\int_{l_i}^{\lambda}(g(x_i,v)-1) \\
&+ f_1(x,l)\sum_i[al_i^2 - \sqrt{\lambda}b\zeta(x_i)l_i]\left(\sum_{j\neq i} l_j\zeta(x_j)\frac{\partial_l g(x_i,l_i)\partial_l g(x_j,l_j)}{g(x_i,l_j)g(x_i,l_i)}\right. \\
&\left.\left. + \frac{\zeta(x_i)\partial_l g(x_i,l_i) + \zeta(x_i)l_i\partial_{l^2}^2 g(x_i,l_i)}{g(x_i,l_i)}\right)\right\}, \qquad \text{(D.8)}
\end{aligned}
$$

where we have used the fact that $f_1$ does not depend on the environment and that $\mathbb{E}_\pi[h_1] = 0$, (where $\mathbb{E}_\pi$ is the expected value over the stationary distribution for the environment) since $\mathbb{E}_\pi[\zeta] = 0$. Passing to the limit in (D.8) as $\lambda$ tends to infinity we obtain

$$
\begin{aligned}
A_\infty(f_1) &= f_1(x,l)\sum_i \frac{Bg(x_i,l_i)}{g(x_i,l_i)} \\
&+ f_1(x,l)\sum_i 2a\int_{l_i}^{\infty}(g(x_i,v)-1)\mathrm{d}v + f_1(x,l)\sum_i(al_i^2 - b^2 l_i)\frac{\partial_{l_i} g(x_i,l_i)}{g(x_i,l_i)} \\
&- f_1(x,l,n)\sum_i b^2\zeta(x_i)l_i\left(\sum_{j\neq i}\zeta(x_j)l_j\frac{\partial_{l_i} g(x_i,l_i)\partial_{l_j} g(x_j,l_j)}{g(x_i,l_j)g(x_i,l_i)} + \frac{\zeta(x_i)l_i\partial_{l_i^2}^2 g(x_i,l_i)}{g(x_i,l_i)}\right).
\end{aligned}
$$

Conditions of Theorem 2.7.1 are satisfied if we consider $A = A_\infty(f_1)$ which would lead too $\varepsilon_n^f = \mathcal{O}\left(\frac{1}{\lambda}\right)$. We now show that if we average the levels of the limiting generator, we obtain the generator of the SBMRE of Definition 3.3. The general principle is the same as for the proof of Theorem 3.7.1 - we average out the levels and refer to the Markov Mapping Theorem to show that the distribution of the projected process is the distribution of the SBMRE.

We consider a Poisson random measure with distribution $\alpha(\mu, \mathrm{d}x \times \mathrm{d}l)$ on $\mathbb{R}^d \times \mathbb{R}_+$ with mean measure $\mu \times m_{leb}$, where $m_{leb}$ is Lebesgue measure. Just as in the Branching Brownian motion case, we consider a special set of test functions of the form

$$h(x) = \int_0^\infty (1 - g(x,l))\mathrm{d}l.$$

In this setup, for a test function $f$, the projected (averaged) test function, $\hat{f}$, takes the form

$$\hat{f}(\mu) = \alpha f(\mu) = \int f(x,v)\alpha(\mu, \mathrm{d}x\times\mathrm{d}v) = e^{\int_{\mathbb{R}^d}\int_0^\infty (1-g(x,v))\mathrm{d}v\mu(\mathrm{d}x)} = e^{-\langle h,\mu\rangle},$$

which is a simple consequence of properties of Poisson random measures. Once again, we integrate the groups of terms that behave similarly separately. Also, to make the calculations easier to read, we write the averaging 'level by level' - performing the computation for a single level wherever possible.

Since the part of the generator which describes the movement of the particles does not depend on the value of the level $l$ the averaging is simply

$$\alpha\left(f_1(x,l)\sum_i \frac{Bg(x_i,l_i)}{g(x_i,l_i)}\right) = \int_{\mathbb{R}^d} -Bh(y)\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle}, \tag{D.9}$$

We now turn our attention to the terms which behave in a very similar fashion to those in (D.5). The computation is analogous.

$$\begin{aligned}
&\int_{\mathbb{R}^d}\int_0^\infty avg(y,z)\int_v^\infty (1-g(y,z))\mathrm{d}z\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} \\
&\quad + \int_{\mathbb{R}^d}\int_0^\infty (av^2-b^2v)\partial_v g(y,v)\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} \\
&= \int_{\mathbb{R}^d}\int_0^\infty avg(y,z)\int_v^\infty (1-g(y,z))\mathrm{d}z\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} \\
&\quad - \int_{\mathbb{R}^d}\int_0^\infty (2av-b^2)g(y,v)\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} \\
&= \int_{\mathbb{R}^d}\int_0^\infty avg(y,z)\int_v^\infty (1-g(y,z))\mathrm{d}z\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} \\
&\quad - \int_{\mathbb{R}^d}\int_0^\infty 2a\int_v^\infty g(z,v)\mathrm{d}z\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} + \int_{\mathbb{R}^d}\int_0^\infty b^2 g(y,v)\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} \\
&= \int_{\mathbb{R}^d} a\left(\int_0^\infty [g(y,v)-1]\mathrm{d}v\right)^2\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} + \int_{\mathbb{R}^d}\int_0^\infty b^2 g(y,v)\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} \\
&= \int_{\mathbb{R}^d}\{ah^2(y)+b^2h(y)\}\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle},
\end{aligned} \tag{D.10}$$

where we have integrated by parts and used the analogues of identities (D.3), (D.4).

Finally, the projections of the terms which are a direct consequence of separation of timescales lead to

$$\begin{aligned}
&\int_{\mathbb{R}^d\times\mathbb{R}} b^2\int_0^\infty\int_0^\infty \zeta(y_1)\zeta(y_2)v\partial_v g(y,v)z\partial_z g(y,z)\mathrm{d}v\mathrm{d}z\mu(\mathrm{d}y_1)\mu(\mathrm{d}y_2)e^{-\langle h,\mu\rangle} \\
&\quad = -\int_{\mathbb{R}^d\times\mathbb{R}^d} b^2\zeta(y_1)\zeta(y_2)\int_0^\infty (g(y,v)-1)\mathrm{d}v\int_0^\infty (g(y,z)-1)\mathrm{d}v\mathrm{d}z\mu(\mathrm{d}y_1)\mu(\mathrm{d}y_2)e^{-\langle h,\mu\rangle} \\
&\quad = \int_{\mathbb{R}^d\times\mathbb{R}^d} b^2\zeta(y_1)\zeta(y_2)h(y_1)h(y_2)\mu(\mathrm{d}y_1)\mu(\mathrm{d}y_2)e^{-\langle h,\mu\rangle},
\end{aligned} \tag{D.11}$$

where we have integrated by parts, and

$$\begin{aligned}
\int_{\mathbb{R}^d} b^2\int_0^\infty v^2\partial_v^2 g(y,v)\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} &= -\int_{\mathbb{R}^d} 2b^2\int_0^\infty v\partial_v g(y,v)\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} \\
&= -\int_{\mathbb{R}^d} 2b^2\int_0^\infty (1-g(y,v))\mathrm{d}v\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle} = -\int_{\mathbb{R}^d} 2b^2h(y)\mu(\mathrm{d}y)e^{-\langle h,\mu\rangle},
\end{aligned} \tag{D.12}$$

where we have integrated by parts twice.

Combining the calculations (D.9), (D.10), (D.11), (D.12) and appealing to Theorem 2.7.1 to average over the environment we arrive at

$$\mathcal{L}\hat{f}(\mu) = \mathcal{L}e^{\langle f,\mu\rangle} \int_{\mathbb{R}^d} \left\{-Bh(y) + ah^2(y) - b^2h(y)\right\} \mu(\mathrm{d}y)e^{-\langle h,\mu\rangle}$$
$$+ \int_{\mathbb{R}^d\times\mathbb{R}^d} b^2 q(y_1,y_2)h(y_1)h(y_2)\mu(\mathrm{d}y_1)\mu(\mathrm{d}y_2)e^{-\langle h,\mu\rangle},$$

which is the generator of the SBMRE.

As before, to ensure that the solution of the martingale problem for the lookdown process gives us information about the solution of the martingale problem for the projected process, we need to specify the Markov map $\gamma$ and check that the conditions of the Markov Mapping Theorem are satisfied. Once again we appeal to Condition 3.8.

The Markov map $\gamma$ is given by

$$\gamma\left(\sum_i \delta_{x_i,l_i}\right) = \begin{cases} \lim_{\lambda\to\infty} \frac{1}{\lambda}\sum_{l_i\le\lambda} \delta_{x_i} & \text{if measures converge,} \\ \mu_0 & \text{otherwise .}\end{cases}$$

Our class of test functions is separating over the counting measures, and closed under multiplication. The moment condition B.1 is satisfied if we consider $\psi$ of the form (D.6) and the averaged $\tilde{\psi}$ of the form (D.7). □

# E Lookdown construction of the Spatial Lambda-Fleming-Viot model

In this section we describe a construction of SLFV model, which is a special case of the construction developed in Etheridge and Kurtz (2018), Section 4.1.3. This construction forms the basis for the construction of the SLFV with selection in a fluctuating environment, whose scaling limits we investigate in Section 4. We restrict our attention to the neutral model and do not intend to present any proofs or details.

Let us recall the key elements of the SLFV process. We consider a population living in a geographical space, which, for simplicity, we choose to be $\mathbb{R}^d$. Each individual is assigned a type from a typespace $\mathcal{K}$. Let $\mu = m_{leb}\times\nu^1(w,\mathrm{d}u)\times\nu^2(\mathrm{d}w)$ be a measure on $\mathbb{R}^d\times[0,1]\times[0,\infty)$, where $m_{leb}$ is $d$-dimensional Lebesgue measure, $\nu^1$ is a measure which determines *impacts* of the events and $\nu^2$ is a $\sigma$-finite measure of event radii which satisfy conditions which are specified in (E.1). Evolution of the population is driven by a Poisson point process $\Pi$ on $[0,\infty)\times\mathbb{R}^d\times[0,1]\times[0,\infty)$ with mean measure $m_{leb}\times\mu$. Whenever $(t,x,u,r)\in\Pi$, a reproduction event occurs at time $t$ in the closed ball $B_r(x)$ (a ball of radius $r$ centred at $x$) with *impact* $u$. The impact of the event determines the proportion of the individuals within the ball $B_r(x)$ that are replaced during the event by the offspring of a parent chosen from the ball $B_r(x)$ just before the event. The locations of new individuals are distributed uniformly over $B_r(x)$. For the construction to be valid, we assume that

$$\int_{[0,1]\times(0,\infty)} uw^d\nu^1(w,\mathrm{d}u)\nu^2(\mathrm{d}w) < \infty. \tag{E.1}$$

For simplicity, we only describe the construction for a fixed impact $u$ and assume that the radius of the reproduction events is always fixed and equal to $r$.

In the spirit of lookdown constructions, in addition to a location in geographical space and a type in the typespace, $\mathcal{K}$, each individual is equipped with a level $l \in \mathbb{R}_+ \cup \{0\}$. The value of the level impacts the choice of the parent during a reproduction event.

As was the case for the models of Section 3, it is convenient to consider our model as a counting measure on $\mathbb{R}^d \times \mathcal{K} \times [0, \infty)$, where the first component encodes the geographical space, the second encodes the type of the individual and the third encodes the level of the individual. The state of the population is given by

$$\eta = \sum_i \delta_{x_i, \kappa_i, l_i},$$

and the single individual $i$ is described by a triple $(x_i, \kappa_i, l_i)$, where $x_i$ is the location of the individual, $\kappa_i$ is their type and $l_i$ is their level. In our particular case the levels will always be a conditionally Poisson system with Cox measure $\Xi \times m_{leb}$, where $m_{leb}$ is the Lebesgue measure on $\mathbb{R}$. The measure $\Xi$ is then nothing else but the distribution of locations and types of individuals.

We specify the model in terms of generator. Let us describe the domains on which our generator are defined, which turns out to be useful not only for formalizing the constructions in this section but also will serve as a functional setup for the considerations in Section 4. Define

$$\mathcal{D}_\lambda = \Bigg\{ f(\eta) = \prod_{x,\kappa,l \in \eta} g(x, \kappa, l) : \\ 0 \leq g(x, \kappa, l) \leq 1, g(\cdot, \kappa, l) \in C^2(\mathbb{R}^d), \|\partial_l g(x, \kappa, l)\| < \infty \\ \exists \text{ compact } K_g \in \mathbb{R}^d, 0 < l_g \leq \lambda \\ g(x, \kappa, l) = 1 \text{ for } (x, l) \notin K_g \times [0, l_g] \Bigg\}, \tag{E.2}$$

and

$$\mathcal{D}_\infty = \bigcup_\lambda \mathcal{D}_\lambda. \tag{E.3}$$

Our test functions are specified by (E.3).

**Remark E.1.** *Notice that our restrictions on domains of the generators are very similar to those in Condition 3.8. This is due to the fact that once again we will apply the Markov Mapping Theorem.*

The evolution of the population is based on events which are composed of two elements - discrete births and so-called thinning of the population. Whenever $(t, x) \in \Pi$, if the number of individuals within $B_r(x)$ is greater than zero, the birth event produces offspring, with levels distributed on $[0, \infty)$ according to independent Poisson point processes with intensity $\alpha_z = uV_r$ (recall that $V_r$ denotes the volume of ball of radius $r$). The levels of new particles are denote by $(v_1, v_2, \dots)$. Their locations are distributed uniformly over $B_r(x)$. Let $v^*$ be the minimum of $(v_1, v_2, \dots)$.

Let $(x^*, \kappa^*, l^*)$ denote the element in $\eta$ such that $x^* \in B_r(y)$ with the smallest level greater than $v^*$. The individual $(x^*, \kappa^*, l^*)$ is chosen as the parent of the event and removed from the population. All new individuals are assigned a type which is same as the type of the parent. The levels of old individuals in the population are changed. If the level of the individual was smaller than $v^*$, it remains unaffected by the birth part of the event. If the level of the individual was larger than $v^*$, it is moved to $l - l^* + v^*$. The thinning occurs after the movement of the levels due to birth event has been accounted for. The new individuals are not affected by thinning. The thinning takes the new level of each individual present within the ball $B_r(x)$ just before the event (apart from the parent), and multiplies it by $1/(1-u)$.

We note that instead of removing the parent from the population we can identify the parent with the lowest offspring (with level $v^*$). The choice between those two options is a matter of convenience and does not affect the model. In our considerations in Section 2 and Section 4 we find it more convenient to identify the parent with the lowest offspring.

Let $v_{y,r}$ denote the density of the uniform distribution on $B_r(y)$. Let $\mathcal{J}_{EK}$ denote the expected value of the test function evaluated immediately after an event centred at $y$. It is given by

$$\begin{aligned}\mathcal{J}_{EK}(g,\eta) = \prod_{(x,\kappa,l)\in\eta, x\notin B_r(y)} & g(x,\kappa,l) \\ & \times \int_0^\infty \Big[\alpha_z e^{-\alpha_z v^*} \int g(x',\kappa^*,v^*) v_{y,r}(\mathrm{d}x') \\ & \times \exp\left(-\alpha_z \int_{v^*}^\infty \left(1 - \int g(x',\kappa^*,v^*) v_{y,r}(\mathrm{d}x')\right)\mathrm{d}v^*\right) \\ & \times \prod_{(x,\kappa,l)\in\eta, x\in B_r(y), l>l^*} g(x,\kappa,\frac{1}{1-u}(l-l^*+v^*)) \\ & \times \prod_{(x,\kappa,l)\in\eta, x\in B_r(y), l<l^*} g\left(x,\kappa,\frac{1}{1-u}l\right)\Big]\,\mathrm{d}v^*.\end{aligned}$$

The generator of the lookdown representation of the SLFV can be now written as

$$A_{EK}f(\eta) = \int_{\mathbb{R}^d} \mathbf{1}_{\eta(B_r(y)\times[0,\infty))>0}\left\{\mathcal{J}_{EK}(g,\eta) - f(\eta)\right\}\mathrm{d}y. \tag{E.4}$$

Recall that $\eta$ is a conditionally Poisson process with Cox measure $(\Xi(s)\times m_{leb})$. To average the generator over the distribution of the levels, we define

$$h(x,\kappa) = \int_0^\infty (1 - g(x,\kappa,l))\mathrm{d}l \tag{E.5}$$

and

$$h^*_{y,r}(\kappa) = \int\left(1 - \int g(x',\kappa,l) v_{y,r}(\mathrm{d}x')\right)\mathrm{d}l. \tag{E.6}$$

Observe that, by integration by parts,

$$\begin{aligned}\int_0^\infty \Big\{\alpha_z e^{-\alpha_z v^*} & \int g(x',\kappa^*,v^*) v_{y,r}(\mathrm{d}x') \\ & \times \exp\left(-\alpha_z \int_{v^*}^\infty \left(1 - \int g(x',\kappa^*,v^*) v_{y,r}(\mathrm{d}x')\right)\mathrm{d}v\right)\Big\}\,\mathrm{d}v^* \\ & = e^{-\alpha_z h^*_{y,r}(\kappa^*)}\end{aligned}$$

Therefore if we average out the levels in generator (E.4) we obtain

$$\alpha A_{EK} f(\Xi) = \exp\left(\int_{\mathbb{R}^d} h(x,\kappa)\Xi(\mathrm{d}x,\mathrm{d}\kappa)\right)$$

$$\times \int_{\mathrm{R}^d} \left\{ \mathbb{H}_2(h^*_{y,r}, \Xi) \exp\left( u \int_{B_r(x)\times\mathcal{K}} h(x,\kappa)\Xi(\mathrm{d}x, \mathrm{d}\kappa) \right) - 1 \right\},$$

where

$$\mathbb{H}_2(h^*_{y,r}, \Xi) = \frac{1}{\Xi(B_r(x)\times\mathcal{K})} \int_{B_r(x)\times\mathcal{K}} \exp\left(-uV_r h^*_{y,r}(\kappa)\right) \Xi(\mathrm{d}x, \mathrm{d}\kappa).$$

The averaged model is simply the usual Spatial Lambda-Fleming-Viot model, see Etheridge and Kurtz (2018). We also observe that if $\Xi(0, \mathrm{d}x\times\mathcal{K})$ is Lebesgue measure then $\Xi(t, \mathrm{d}x\times\mathcal{K})$ is Lebesgue measure, for arbitrary $t$. If this is the case, $\mathbb{H}_2$ can be written as

$$\mathbb{H}_2(h^*_{y,r}) = \frac{1}{\Xi(B_r(x)\times\mathcal{K})} \int_{B_r(x)\times\mathcal{K}} \exp\left(-uh^*_{y,r}(\kappa)\Xi(B_r(x)\times\mathcal{K})\right)\right) \Xi(\mathrm{d}x, \mathrm{d}\kappa).$$

# References

A. Barnosky. Distinguishing the effects of the red queen and court jester on miocene mammal evolution in the northern rocky mountains. *Journal of Vertebrate Paleontology*, 21(1):172–185, 2001.

N. H. Barton, A. M. Etheridge, and A. Véber. A new model for evolution in a spatial continuum. *Electron. J. Probab.*, 15:162–216, 2010.

N. H. Barton, A. M. Etheridge, and A. Véber. Modelling evolution in a spatial continuum. *J. Stat. Mech.*, page PO1002, 2013.

J. Benton. The red queen and the court jester: Species diversity and the role of biotic and abiotic factors through time. *Science*, 323(5915):728–732, 2009.

M. Birkner, J. Geiger, and G. Kersting. Branching processes in random environment - a view on critical and subcritical cases. In *Interacting stochastic systems*, pages 269–291. Springer, 2005.

N. Biswas, A. Etheridge, and A. Klimek. The spatial Lambda-Fleming-Viot process with fluctuating selection. *ArXiv e-prints*, February 2018.

M. Bramson, J. Cox, and J. Le Gall. Super-Brownian limits of voter model clusters. *Ann. Probab.*, 29(3):1001–1032, 2001.

J. Chetwynd-Diggle and A. Etheridge. Superbrownian motion and the spatial lambda-fleming-viot process. *Electron. J. Probab.*, 23:36 pp., 2018.

J. T. Cox, R. Durrett, and E. A. Perkins. Rescaled voter models converge to super-Brownian motion. *Ann. Probab.*, 28:185–234, 2000.

D. Crisan. Superprocesses in a Brownian environment. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 460(2041):243–270, 2004. ISSN 1364-5021.

D. A. Dawson. Stochastic evolution equations and related measure processes. *J. Multivar. Anal.*, 5:1–52, 1975.

D. A. Dawson. Geostochastic calculus. *Can. J. Statistics*, 6:143–168, 1978.

P. Donnelly and T. G. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.*, 24:698–742, 1996.

P. Donnelly and T. G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.*, 27:166–205, 1999.

A. M. Etheridge. *An introduction to superprocesses*, volume 20 of *University lecture notes*. Amer. Math. Soc., 2000.

A. M. Etheridge. Drift, draft and structure: some mathematical models of evolution. *Banach Center Publ.*, 80:121–144, 2008.

A. M. Etheridge and T. G. Kurtz. Genealogical constructions of population models. *Annals of Probability*, 2018.

A. M. Etheridge, A. Véber, and F. Yu. Rescaling limits of the spatial Lambda-Fleming-Viot process with selection. *arXiv:1406.5884*, 2018.

S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence.* Wiley, 1986.

R. Forien and S. Penington. A central limit theorem for the spatial lambda -fleming-viot process with selection. *Electronic Journal of Probability*, 22, 2017.

A. Gupta. Stochastic model for cell polarity. *The Annals of Applied Probability*, 22(2):827–859, 2012.

R. van der Hofstad, M. Holmes, and E. A. Perkins. A criterion for convergence to super-Brownian motion on path space. *Ann. Probab.*, 45:278–376, 2017.

W. H. Kliemann, G. Koch, and F. Marchetti. On the unnormalized solution of the filtering problem with counting process observations. *IEEE Transactions on Information Theory*, 36(6):1415–1425, 1990.

T. G. Kurtz. A limit theorem for perturbed operator semigroups with applications to random evolutions. *Journal of Functional Analysis*, 12(1):55–67, 1973.

T. G. Kurtz. Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis*, volume 177. Springer, Berlin, 1992.

T. G. Kurtz. Martingale problems for conditional distributions of Markov processes. *Elect. J. Prob.*, 3, paper 9, 1998.

T. G. Kurtz. Particle representations for measure-valued population processes with spatially varying birth rates. In *Stochastic models (Ottawa, ON, 1998)*, volume 26 of *CMS Conf. Proc.*, pages 299–317. Amer. Math. Soc., Providence, RI, 2000.

T. G. Kurtz. Equivalence of stochastic equations and martingale problems. In *Stochastic analysis 2010*, pages 113–130. Springer, 2010.

T. G. Kurtz and E. Rodrigues. Poisson representations of branching markov and measure-valued branching processes. *Ann. Probab.*, 39(3):939–984, 05 2011.

A. Lunardi. *Analytic semigroups and optimal regularity in parabolic problems.* Springer Science & Business Media, 2012.

C. Ma. Construction of non-gaussian random fields with any given correlation structure. *Journal of Statistical Planning and Inference*, 139(3):780–787, 2009.

L. R. Miller. *Evolution of highly fecund organisms.* PhD thesis, University of Oxford, 2015.

L. Mytnik. Superprocesses in random environments. *Ann. Probab.*, 24:1953–1978, 1996.

M. Nakashima. Branching random walks in random environment and super-brownian motion in random environment. *Ann. Inst. H. Poincar Probab. Statist.*, 51(4):1251–1289, 11 2015.

R. Stockbridge. Portfolio optimization in markets having stochastic rates. In Bozenna Pasik-Duncan, editor, *Stochastic Theory and Control*, pages 447–458, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.

A. Sturm. On convergence of population processes in random environments to the stochastic heat equation with colored noise. *Electron. J. Probab.*, 8:39 p., 2003.

J. E. Taylor. The genealogical consequences of fecundity variance polymorphism. *Genetics*, 182(3): 813–837, 2009.

A. Véber and A. Wakolbinger. The spatial Lambda-Fleming-Viot process: an event based construction and a lookdown representation. *Ann. Inst. H. Poincaré*, 51:570–598, 2015.

S. Watanabe. A limit theorem of branching processes and continuous state branching processes. *J. Math. Kyoto Univ.*, 8:141–167, 1968.

**Acknowledgements** The authors would like to thank Alison Etheridge and Tom Kurtz for useful discussions and insights.