Invasion and adaptive evolution for individual-based spatially structured populations

Nicolas Champagnat[1], Sylvie Méléard[2,3]

December 6, 2006

Abstract

The interplay between space and evolution is an important issue in population dynamics, that is particularly crucial in the emergence of polymorphism and spatial patterns. Recently, biological studies suggest that invasion and evolution are closely related. Here we model the interplay between space and evolution starting with an individual-based approach and show the important role of parameter scalings on clustering and invasion. We consider a stochastic discrete model with birth, death, competition, mutation and spatial diffusion, where all the parameters may depend both on the position and on the phenotypic trait of individuals. The spatial motion is driven by a reflected diffusion in a bounded domain. The interaction is modelled as a trait competition between individuals within a given spatial interaction range. First, we give an algorithmic construction of the process. Next, we obtain large population approximations, as weak solutions of nonlinear reaction-diffusion equations. As the spatial interaction range is fixed, the nonlinearity is nonlocal. Then, we make the interaction range decrease to zero and prove the convergence to spatially localized nonlinear reaction-diffusion equations. Finally, a discussion of three concrete examples is proposed, based on simulations of the microscopic individual-based model are given. These examples illustrate the strong effects of the spatial interaction range on the emergence of spatial and phenotypic diversity (clustering and polymorphism) and on the interplay between invasion and evolution. The simulations focus on the qualitative differences between local and nonlocal interactions.

MSC 2000 subject classifications: primary 60J85, 60K35, 92D15; secondary 92D25, 35K60.

Key words and phrases. Spatially structured population, adaptive evolution, stochastic individual-based model, birth-and-death point process, reflected diffusion, mutation and selection, nonlinear reaction-diffusion equation, nonlocal interaction and local interaction, clustering, polymorphism, invasion and evolution.

[1]Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, Berlin, champagn@wias-berlin.de

[2]Université Paris 10, MODAL'X, 200 av. de la République, 92000 Nanterre, sylvie.meleard@u-paris10.fr

[3]Fonctionnement et évolution des systèmes écologiques, UMR 7625, ENS, 46 rue d'Ulm, 75005 Paris

1 Introduction

The spatial aspect of population dynamics is a fundamental ecological issue that has been extensively studied (Murray [33], Durrett and Levin [14], Tilman and Kareiva [46], McGlade [28], Dieckmann et al. [11]). In particular, the spatial organization of populations is a crucial question, for example concerning spatial invasions and epidemics (Mollison [32], Murray [33], Rand et al. [38], Tilman and Kareiva [46], Lewis and Pacala [27]), or clustering and agglomeration of a population, i.e. its organization as isolated patches (Hassel and May [21], Hassel and Pacala [22], Niwa [34], Flierl et al. [16], Young et al. [47]). The combination of spatial motion and mutation-selection processes is also known for a long time to have important effects on population dynamics (Mayr [29], Endler [15]). In particular, it can strongly favor the diversity and the emergence and stability of polymorphism in a population (Durrett and Levin [13], Dieckmann and Doebeli [9]). More recently, several biological studies (Thomas et al [45], Phillips et al. [35]) observed that classical models could underestimate the invasion speed and suggested that evolution and invasion are closely related. Namely, the evolution of morphology can have strong impact on the expansion of invading species, such as insect species ([45]) or cane toads ([35]). In this context, the study of space-related traits, such as dispersal speed (Prévost [36], Desvillettes et al. [8]), or sensibility to heterogeneously distributed resources (Bolker and Pacala [3], Grant and Grant [20]), is fundamental.

In this paper, we propose and construct stochastic and deterministic population models describing the interplay between evolution and spatial structure. We show how helpful these models can reveal to understand and predict several specific behaviors related to clustering and invasion.

We study the dynamics of a spatially structured asexual population with adaptive evolution, in which individuals can move, reproduce with possible phenotypic mutation, or die of natural death or because of the competition between individuals. The individuals are characterized both by their position and by one or several phenotypical adaptive traits (such as body size, rate of food intake, age at maturity or dispersal speed). The interaction is modelled as a trait competition between individuals in some spatial range. Our approach is based on a stochastic microscopic description of the individual mechanisms, involving both space and traits. This approach has already been developed in simpler ecological contexts. For populations undergoing dispersal, Bolker and Pacala [2, 3] and Dieckmann and Law [10], offered the first microscopic heuristics and simulations. Their individual-based model has been rigourously constructed in Fournier and Méléard [17]. In a mathematical point of view, dispersion in the physical space can be interpreted as mutation in a trait space. Such evolutionary models with general mutation and competition are studied in Champagnat, Ferrière and Méléard [7, 6]. Different large population deterministic or stochastic approximations have been obtained, depending on several scalings on the ecological parameters.

Here we are concerned by the combination of spatial motion and evolutionary dynamics. In contrast with the references above, we model spatial motion as a diffusion. This assumption is relevant, when the time scale of the spatial motion is much smaller than the ecological time scale. Then we are led to define a birth-and-death diffusion process, in which the diffusion, birth, mutation and death parameters of each individual depend on its position and trait. An offspring, appearing at the same position as its progenitor, usually inherits the trait value of the latter, except when a mutation causes the offspring to take an instantaneous mutation step at birth to new trait values. As soon as it is alive, an individual moves in the spatial domain according to a diffusion process reflected on the boundary of the domain. Moreover, each individual dies because of natural death or is eliminated because of the competition (selecting the fittest traits) between individuals living in a given spatial range $\delta > 0$.

We are interested in studying the interplay between different scalings of the parameters: population size, mutation amplitude, spatial range and diffusion coefficient. Firstly we consider a large population assumption and derive a macroscopic deterministic approximation where the spatial interaction is nonlocal with range δ. Secondly we are interested in local spatial interaction, ($\delta = 0$). Then we study in details three concrete examples. Two of them are theoretical and the third one is more empirical. Each case is numerically studied, in order to focus on the interplay between the different parameters. A particular emphasis is put on the range of spatial interaction. The results of the simulations show how rich and intricate are these space and trait dynamics.

The model is described in Section 2 and an algorithmic construction is given in Section 3. The dynamics is mathematically described by a stochastic Markov point process whose generator captures the individual migration and ecological mechanisms in the population. The existence of this measure-valued stochastic process and its martingale properties are obtained.

Next (Section 4), we study approximations of this model based on large-population limits. We consider a large number N of individuals at initial time and assume that a fixed amount of available resources has to be partitioned between individuals. When N tends to infinity, the conveniently normalized point process converges to a deterministic finite measure, solution of a nonlinear nonlocal integro-differential equation, with Neumann's boundary conditions, parameterized by the spatial range. The proof is based on the martingale properties of the process and on limit theorems for measure-valued jump processes. Moreover, we prove that, for sufficiently smooth and non-degenerate diffusion coefficients and assuming that the initial condition has a density, the limiting measure has at each time a density with respect to the Lebesgue measure. This means that no singularity can appear in finite time in the distribution of the population. This is due to the regularizing effect of the reflected diffusion process. The proof is based on the mild formulation of the limiting nonlinear equation.

In Section 5, we study the behavior of this density function as the interaction range tends to 0. We show its convergence to the solution of a spatially local nonlinear integro-differential equation with Neumann's boundary conditions. This equation has been introduced and studied in Prévost [36] from an analytic viewpoint (see also Desvillettes et al. [8]). In their case, numerical simulations by finite element methods are given and show the influence of diffusion and mutation parameters on the invasion of the domain by the population.

In Section 6, we give simulations of the microscopic process illustrating the time-dependent interplay between space and adaptation. We address the effect of the population size, and the crucial role of the interaction range with respect to spatial organization (clustering) and polymorphism. We focus on the qualitative differences between nonlocal and local interactions. In a first example, it is assumed that an different optimal trait value corresponds to each spatial position, as a result of non-homogeneously distributed resources. The simulations show a local adaptation to the optimal trait value. Moreover, when migrations and mutations are not too strong, a large interaction range induces a spatial organization of the population as a finite set of isolated clusters, as assumed in classical metapopulation models ([13]). Such a spatial organization is related to the ecological notion of "niches" (different types of individuals settle different regions of space, Roughgarden [41]). Conversely, for sufficiently small interaction range, the clustering phenomenon is no more observed. Our second example also illustrates spatial clustering, this time depending on the balance between the spatial interaction range and the size of the space region with high resources. In this case, simulations suggest that a quite sharp phase transition occurs at some critical value of the spatial interaction range. When it is bigger, there is only one cluster, whereas when it is smaller, two (or more) clusters coexist. In our last example, we investigate a model describing the invasion of a species with evolving dispersal speed (as in [8]). The diffusion coefficient and the trait are assumed to be proportional. One specific feature of our simulations is the triangular invasion pattern which indicates that the invasion front is composed of faster individuals. Moreover, mutation produces faster and faster individuals at the invasion front. This implies a progressive acceleration of the invasion, in accordance with the observations for cane toads in Australia ([35]).

Notation

The individuals live in the closure $\bar{\mathcal{X}}$ of a bounded open domain $\mathcal{X}$ of $\mathbb{R}^d$ of class C^3 and their trait values belong to a compact set $\mathcal{U}$ of $\mathbb{R}^q$.

- For $x \in \partial\mathcal{X}$, we denote by $n(x)$ the outward normal to the boundary $\partial\mathcal{X}$ at point x.
- For a sufficiently smooth function f and $(x,u) \in \partial\mathcal{X} \times \mathcal{U}$, we denote by $\partial_n f(x,u)$ the scalar product $\nabla_x f(x,u) \cdot n(x)$.
- We denote by $C_0^{2,b}$ the space of measurable functions $f(x,u)$ of class C^2 in x and bounded in u satisfying $\partial_n f(x,u) = 0$ for all $(x,u) \in \partial\mathcal{X} \times \mathcal{U}$ and by $C_0^{2,0}$ the subspace of functions $f(x,u)$ which are moreover continuous in u.
- For each $p \geq 1$, the L^p-norm on $\bar{\mathcal{X}} \times \mathcal{U}$ is denoted by $\| \cdot \|_p$.
- We denote by $M_F(\bar{\mathcal{X}} \times \mathcal{U})$ the set of finite measures on $\bar{\mathcal{X}} \times \mathcal{U}$, endowed by the weak topology, and by $\mathcal{M}$ the subset of $M_F(\bar{\mathcal{X}} \times \mathcal{U})$ composed of all finite point measures, that is
$$\mathcal{M} = \left\{ \sum_{i=1}^{n} \delta_{(x^i,u^i)},\ n \in \mathbb{N},\ x^1, \ldots, x^n \in \bar{\mathcal{X}}\ ,\ u^1, ..., u^n \in \mathcal{U} \right\}$$
where $\delta_{(x,u)}$ denotes the Dirac measure at (x,u). For any $\nu \in M_F(\bar{\mathcal{X}} \times \mathcal{U})$ and for any measurable function f on $\bar{\mathcal{X}} \times \mathcal{U}$, we write indifferently $\langle \nu, f \rangle$ or $\int_{\bar{\mathcal{X}} \times \mathcal{U}} f d\nu$. If $\nu = \sum_{i=1}^n \delta_{(x^i,u^i)}$, then $\langle \nu, f \rangle = \sum_{i=1}^n f(x^i, u^i)$ and the total mass $\langle \nu, 1 \rangle$ of ν is equal to the number n of Dirac masses composing ν.
- We denote by $\mathbb{D}([0,\infty), M_F(\bar{\mathcal{X}} \times \mathcal{U}))$ the Skorohod space of left limited and right continuous functions from $\mathbb{R}_+$ to $M_F(\bar{\mathcal{X}} \times \mathcal{U})$, endowed with the Skorohod topology.
- The constant C will be a constant which can change from line to line.

Remark 1.1 *Let us remark that the space of $C^2(\bar{\mathcal{X}})$-functions with a vanishing normal derivative is dense, for the uniform norm, in $C(\bar{\mathcal{X}})$. Indeed, let us consider the Cauchy problem for the parabolic differential equation $\frac{\partial u}{\partial t}(t,x) = \Delta u(t,x)\ ;\ t > 0\ ;\ x \in \mathcal{X}$ with the boundary condition $\frac{\partial u}{\partial n}(t,x) = 0\ ;\ t > 0\ ;\ x \in \partial\mathcal{X}$. Since $\mathcal{X}$ is of class C^3, we may apply Sato-Ueno [42] Theorem 2.1. There exists a smooth fundamental solution $q(t,x,y)$ to this* *problem and each $f \in C(\bar{\mathcal{X}})$ is the uniform limit of the sequence $\int_{\bar{\mathcal{X}}} q(t,x,y)f(y)dy$ of $C^2(\bar{\mathcal{X}})$-functions with vanishing normal derivative, as t tends to 0.*

We easily extend this result and show that the space $C_0^{2,0}$ is dense in the space of continuous functions on $\bar{\mathcal{X}} \times \mathcal{U}$.

2 The model

2.1 Ecological Parameters

Let us now describe the evolutionary process we are interested in. The population will be described at any time by the finite point measure $\nu \in \mathcal{M}$ given by the sum of the Dirac masses at the pair composed of the position and the trait values of each individual in the population. Each individual, characterized by its position and trait (x,u), may move, give birth or die, as described below. The reader can see three concrete examples of parameters in Section 6.

1. The **migration** is described as a diffusion process normally reflected at the boundary of the domain $\mathcal{X}$. Biologists usually assume that the random behavior is isotropic, so the diffusion matrix is chosen with the form $m(x,u)\mathrm{Id}$ (Id is the identity matrix on $\mathbb{R}^d$) and the nonnegative coefficient $m(x,u)$ (depending on the position x and the trait value u), is the diffusion coefficient. We moreover model the environment heterogeneity (resources, topography, external effects,...) by a drift term driven by a $\mathbb{R}^d$-vector $b(x,u)$. We assume that the coefficients $m(x,u)$ and $b(x,u)$ depend Lipschitz continuously on the position and measurably on the trait, and there exist constants $m^\star > 0$ and $b^* > 0$ such that for all $(x,u) \in \bar{\mathcal{X}} \times \mathcal{U}$
$$\begin{aligned} 0 \le m(x,u) \le m^\star \\ |b(x,u)| \le b^*. \end{aligned} \tag{2.1}$$

2. **Births and mutations**. We consider a population with asexual reproduction. An individual with position x and trait u can give birth either to a clonal child at rate $\lambda(x,u)$, or to a mutant with trait v according to the transition measure $M(x,u,v)dv$, both at position x. That means that a mutation occurs at rate $\int M(x,u,v)dv$, and the mutant trait has law $\frac{M(x,u,v)}{\int M(x,u,v)dv}dv$. A typical example is given by $\lambda(x,u) = \beta(x,u)(1-p(x,u))$ and $\int M(x,u,v)dv = \beta(x,u)p(x,u)$, where β is the individual birth rate and p is the mutation probability.

   It is natural from a biological point of view to assume that all birth rates are bounded:

there exists λ^* such that

$$0 \leq \lambda(x,u) \leq \lambda^*, \quad \forall (x,u) \in \bar{\mathcal{X}} \times \mathcal{U}. \tag{2.2}$$

The kernel M is assumed to be nonnegative and symmetric in (u,v) for each $x \in \bar{\mathcal{X}}$ and

$$\sup_{x \in \bar{\mathcal{X}}, u \in \mathcal{U}} M(x,u,v) = M^*(v) \in L^1(\mathcal{U}). \tag{2.3}$$

3. The **death** rate μ of an individual depends on its position x and trait u and on the spatial and phenotypic interaction with the individuals located around x. We assume that the effects of spatial distance and trait are multiplicative. The strength of the trait competition is governed by a nonnegative and bounded kernel W. We also introduce a nonnegative and bounded spatial kernel I^δ such for each $x \in \bar{\mathcal{X}}$,

$$\int_{\mathcal{X}} I^\delta(x-y)dy = 1. \tag{2.4}$$

Below, the parameter $\delta > 0$ will be interpreted as the **range** of the spatial interaction. For a population $\nu = \sum_{i=1}^n \delta_{(x^i,u^i)} \in \mathcal{M}$, the death rate is given by

$$\begin{aligned}\mu(x,u,I^\delta W \star \nu(x,u)) &= \mu\Big(x,u,\int_{\mathcal{X}\times\mathcal{U}} I^\delta(x-y)W(u-v)\nu(dy,dv)\Big)\\ &= \mu\Big(x,u,\sum_{i=1}^n I^\delta(x-x^i)W(u-u^i)\Big).\end{aligned}$$

The decoupling between spatial and phenotypic interactions is realistic in many biological situations. This particular choice of interaction modelling will allow us to scale physical space independently of the phenotype space.

We assume that the function $\mu(x,u,r)$ is measurable on $\mathcal{X} \times \mathcal{U} \times \mathbb{R}$ and that there exists a positive constant μ^* such that

$$\forall (x,u,r) \in \mathcal{X} \times \mathcal{U} \times \mathbb{R}, \quad 0 \leq \mu(x,u,r) \leq \mu^*(1+|r|) \tag{2.5}$$

A typical example of competition is the so-called logistic competition case, where $\mu(x,u,r) = \mu_0(x,u) + \mu_1(x,u)r$.

A typical example for the spatial interaction is to choose I^δ proportional to $1_{\{|x|\leq\delta\}}$. In this case, the interpretation of δ as the spatial interaction range is clear, and (2.4) means that the interaction is proportional to the area in $\mathcal{X}$ around x. This is a natural biological assumption, especially if x lies on the boundary of $\mathcal{X}$. We will later assume that the measure $I^\delta(y)dy$ weakly converges to the Dirac measure at 0 as δ tends to 0. This limit corresponds to the limit of local spatial interaction.

Hypotheses (2.1), (2.2), (2.3), (2.4), (2.5) are referred as **Hypotheses (H)** and will be assumed in all the sequel.

They imply in particular that for each $\nu \in M_F(\bar{\mathcal{X}} \times \mathcal{U})$ and each $(x,u) \in \bar{\mathcal{X}} \times \mathcal{U}$,

$$\mu(x,u,I^\delta W \star \nu(x,u)) \leq \mu^*(1 + \|I^\delta W\|_\infty \langle \nu, 1\rangle) \tag{2.6}$$

which yields

$$\begin{aligned}\mu(x,u,I^\delta W \star \nu(x,u)) + \lambda(x,u) + \int_{\mathcal{U}} M(x,u,v)dv \\ \leq \mu^*(1 + \|I^\delta W\|_\infty \langle \nu, 1\rangle) + \lambda^* + \|M^*\|_1 \leq C_\delta(\langle \nu, 1\rangle + 1)\end{aligned} \tag{2.7}$$

and the total jump rate for a population ν is bounded by

$$C_\delta \langle \nu, 1\rangle(\langle \nu, 1\rangle + 1). \tag{2.8}$$

Several explicit expressiosn for the parameters λ, μ, I^δ, W, b and m, motivated by various ecological questions, are given in Section 6.

2.2 The Dynamics

We are interested in the evolution of the stochastic point process (ν_t), taking its values in $\mathcal{M}$ and describing the evolution of the population at each time t. We define

$$\nu_t = \sum_{i=1}^{N_t} \delta_{(X_t^i, U_t^i)} \ ,$$

$N_t \in \mathbb{N}$ standing for the number of living individuals at time t, $X_t^1, ..., X_t^{N_t}$ describing their position (in $\bar{\mathcal{X}}$) and $U_t^1, ..., U_t^{N_t}$ their trait values (in $\mathcal{U}$).

The dynamics of the population can be roughly summarized as follows. The initial population is characterized by a measure $\nu_0 \in \mathcal{M}$ at time $t = 0$, and any individual located at $x \in \bar{\mathcal{X}}$ with trait u at time t has three independent exponential clocks: a "clonal reproduction" clock with parameter $\lambda(x,u)$, a "mutant reproduction" clock with parameter $M(x,u,v)$, and a "mortality" clock with parameter $\mu(x,u,\sum_{j=1}^{N_t} I^\delta(x - X_t^j)W(u - U_t^j))$. If the "mortality" clock of an individual rings, then this individual disappears; if the "clonal reproduction" clock of an individual rings, then it produces at the same location an individual with the same trait as itself; if the "mutant reproduction" clock of an individual rings, then it produces at the same location an individual with characteristics (x,v).

All the living individuals move in the domain $\bar{\mathcal{X}}$ according to independent diffusion processes with diffusion coefficient $m(x,u)$ and drift $b(x,u)$, normally reflected at the boundary of $\mathcal{X}$.

The measure-valued process $(\nu_t)_{t\geq 0}$ is a Markov process whose law is characterized by its infinitesimal generator L which captures the dynamics described above. This generator is the sum of a jump part L_1 corresponding to the phenotypic evolution and of a diffusion part L_2. The generator L_1 is defined for bounded and measurable functions ϕ from $\mathcal{M}$ into $\mathbb{R}$ and for $\nu = \sum_{i=1}^{n} \delta_{(x^i,u^i)}$ by

$$\begin{aligned} L_1\phi(\nu) &= \sum_{i=1}^{\langle \nu,1\rangle} \lambda(x^i,u^i)(\phi(\nu+\delta_{(x^i,u^i)})-\phi(\nu)) \\ &+ \int_{\mathcal{U}} \sum_{i=1}^{\langle \nu,1\rangle} (\phi(\nu+\delta_{(x^i,v)})-\phi(\nu))M(x^i,u^i,v)dv \\ &+ \sum_{i=1}^{\langle \nu,1\rangle} (\phi(\nu-\delta_{(x^i,u^i)})-\phi(\nu))\mu(x^i,u^i,I^\delta W\star\nu(x^i,u^i)). \end{aligned} \tag{2.9}$$

This part of the generator captures all the potential changes of the values of ϕ due to each possible transition from state ν, weighted by the corresponding transition rate.

A standard class of cylindrical functions generating the set of bounded and measurable functions from $M_F(\bar{\mathcal{X}}\times\mathcal{U})$ into $\mathbb{R}$ is the class of functions

$$F_f(\nu) = F(\langle \nu, f\rangle), \tag{2.10}$$

for bounded and measurable functions F and f.

For such functions F_f, with $F\in C_b^2(\mathbb{R})$ and $f\in C_0^{2,0}$, the diffusive part L_2 of the generator can easily be deduced from Itô's formula. Its form is similar to the one obtained in the whole space for branching diffusing processes (cf. Roelly-Rouault [40]) and is given by

$$L_2F_f(\nu) = \langle \nu, m\Delta_x f + b.\nabla_x f\rangle F'(\langle\nu,f\rangle) + \langle \nu, m|\nabla_x f|^2\rangle F''(\langle \nu, f\rangle). \tag{2.11}$$

Hence,

$$\begin{aligned} LF_f(\nu) &= L_1F_f(\nu) + L_2F_f(\nu) \\ &= \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{ \lambda(x,u)\big(F(\langle\nu,f\rangle + f(x,u)) - F(\langle\nu,f\rangle)\big) \\ &\qquad + \int_U \big(F(\langle\nu,f\rangle + f(x,v)) - F(\langle \nu,f\rangle)\big)M(x,u,v)dv \\ &\qquad + \mu(x,u,I^\delta W\star\nu(x,u))\big(F(\langle\nu,f\rangle - f(x,u)) - F(\langle\nu,f\rangle)\big) \\ &\qquad + \big(m(x,u)\Delta_x f(x,u) + b(x,u).\nabla_x f(x,u)\big)F'(\langle\nu,f\rangle)) \\ &\qquad + m(x,u)|\nabla_x f(x,u)|^2 F''(\langle\nu,f\rangle)\Big\}\ \nu(dx,du). \end{aligned} \tag{2.12}$$

3 Construction of the particle system and martingale properties

In this section, an algorithmic construction of the Markov process with infinitesimal generator L on the path space $\mathbb{D}([0,\infty), M_F(\bar{\mathcal{X}} \times \mathcal{U}))$ is given. This iterative construction, which gives an effective simulation algorithm of the population process, if combined with a diffusion simulation step such as an Euler scheme for reflected diffusions (see Lépingle [26], Gobet [19], Bossy, Gobet, Talay [4] and Section 6). Then, some martingale properties satisfied by this process are proved. They are the key points to obtain the large population approximations of the next section.

3.1 The algorithm

Since the individual jump rates depend on the trait and position, we can not use directly a standard birth and death construction. However, thanks to the upper-boundedness of the total population jump rate, our dynamics is stochastically dominated by a Poissonian one. This observation suggests to use an acceptance-rejection scheme in order to simulate the birth and death part of our dynamics. We will construct a sequence of potential birth or death times. Between two of these jump times, the population size and each individual's trait values are fixed. Their positions evolve according to independent diffusion with reflection on the boundary of the space domain. At one potential jump time, we randomly choose an individual in the population, say with position x and trait u. A simulated uniform random variable will help to decide what happens to this individual: its death, a clonal birth, a birth with birth or nothing. If death occurs, the individual is killed and the population sizes decreases by one; if clonal birth occurs, a clone is added at (x,u) and the population size increases by one; if a mutation occurs, the mutant trait is also computed by the acceptance-rejection procedure, a new individual with this mutant trait is added at spatial position x and the population size increases by one. It is proved below that this procedure allows us to obtain, after a finite number of steps, the state of the population at any fixed time.

More formally, we will inductively construct the sequence $(T_k)_k$ of potential birth and death times, the numbers $(N_k)_k$ of individuals in the population at these times, and the sequence of vectors of individual position and traits $(X_{T_k})_k$ and $(U_{T_k})_k$ at these times ($(X_{T_k}, U_{T_k}) = (X^i_{T_k}, U^i_{T_k})_{1\leq i\leq N_k} \in (\bar{\mathcal{X}} \times \mathcal{U})^{N_k}$ for any $k \geq 0$). The initial number of individuals is equal to some fixed integer $N_0 = N \in \mathbb{N}^*$ and the vector of random variables $(X_0, U_0) \in (\bar{\mathcal{X}} \times \mathcal{U})^N$ denotes the initial positions and traits of these individuals. For each pair of consecutive jump times T_k and T_{k+1}, we denote by $X_t \in \bar{\mathcal{X}}^{N_k}$ the vector of positions of the individuals in the population at time $t \in [T_k, T_{k+1})$.

Let us introduce the following sequences of independent random variables, independent of (X_0, U_0).

- $(B^{j,k})_{j,k\in\mathbb{N}^*}$ are d-dimensional Brownian motions
- $(\theta_k)_k$ are uniform random variables on $[0,1]$
- $(V_k)_k$ take values in $\mathcal{U}$ with law $\frac{M^*(v)}{\|M^*\|_1}dv$
- $(\tau_k)_k$ are exponential random variables with law $C_\delta e^{-C_\delta t}1_{t\geq 0}$. (The constant C_δ is defined in (2.7)).

The system is obtained inductively for $k \geq 1$ as described below. We set $T_0 = 0$ and $N_0 = N$. Assume that $(T_{k-1}, N_{k-1}, X_{T_{k-1}}, U_{T_{k-1}})$ are given. If $N_{k-1} = 0$, then we put $\nu_t = 0$ for all $t \geq T_{k-1}$. If not, let

- $T_k = T_{k-1} + \frac{\tau_k}{N_{k-1}(N_{k-1}+1)}$. The form $\frac{\tau_k}{N_{k-1}(N_{k-1}+1)}$ for the time between two possible jumps comes from the upper-bound $C_\delta N_{k-1}(N_{k-1}+1)$ obtained in (2.8) for the total jump rate in a population composed of N_{k-1} individuals.
- On the time-interval $[T_{k-1}, T_k)$, the number of particles remains equal to N_{k-1} and their trait values to $U^j_{T_{k-1}}, 1 \leq j \leq N_{k-1}$. Their positions $(X^j_t,\ 1 \leq j \leq N_{k-1})$ evolve according to the stochastic differential equations with normal reflection: $\forall t \in [T_{k-1}, T_k]$,

$$\begin{aligned}
&X^j_t \in \bar{\mathcal{X}},\\
&X^j_t = X^j_{T_{k-1}} + \int_{T_{k-1}}^t \sqrt{2m(X^j_s, U^j_{T_{k-1}})}dB^{j,k}_s + \int_{T_{k-1}}^t b(X^j_s, U^j_{T_{k-1}})ds - k^j_t\ ,\\
&|k^j|_t = \int_{T_{k-1}}^t \mathbf{1}_{\{X^j_s\in\partial\mathcal{X}\}}d|k^j|_s\ ;\ k^j_t = \int_{T_{k-1}}^t n(X^j_s)d|k^j|_s.
\end{aligned} \tag{3.1}$$

- At time T_k, an individual $I_k = i$ is chosen uniformly at random among the N_{k-1} individuals living during the time-interval $[T_{k-1}, T_k)$. Its position and trait are $(X^i_{T_k}, U^i_{T_{k-1}})$.
  - If $0 \leq \theta_k \leq \frac{\mu(X^i_{T_k}, U^i_{T_{k-1}}, \sum_{j=1}^{N_{k-1}} I^\delta(X^i_{T_k} - X^j_{T_k})W(U^i_{T_{k-1}} - U^j_{T_{k-1}}))}{C_\delta(N_{k-1}+1)} =: \theta^i_1(X_{T_k}, U_{T_{k-1}})$, then the individual i dies (its characteristics are removed from the vector $(X_{T_k}, U_{T_{k-1}})$) and $N_k = N_{k-1} - 1$.

- If $\theta_1^i(X_{T_k}, U_{T_{k-1}}) < \theta_k \leq \theta_1^i(X_{T_k}, U_{T_{k-1}}) + \frac{\lambda(X^i_{T_k}, U^i_{T_{k-1}})}{C_\delta(N_{k-1}+1)} =: \theta_2^i(X_{T_k}, U_{T_{k-1}})$, then the individual i gives birth to an offspring with characteristics $(X^i_{T_k}, U^i_{T_{k-1}})$ (these coordinates are added to the vector $(X_{T_k}, U_{T_{k-1}})$) and $N_k = N_{k-1} + 1$.
- If $\theta_2^i(X_{T_k}, U_{T_{k-1}}) < \theta_k \leq \theta_2^i(X_{T_k}, U_{T_{k-1}}) + \frac{M(X^i_{T_k}, U^i_{T_{k-1}}, V_k)\|M^*\|_1}{M^*(V_k) C_\delta(N_{k-1}+1)} =: \theta_3^i(X_{T_k}, U_{T_{k-1}}, V_k)$, then the individual i gives birth to a mutant offspring with trait V_k at the position $X^i_{T_k}$ ($(X^i_{T_k}, V_k)$ is added to the vector $(X_{T_k}, U_{T_{k-1}})$), and $N_k = N_{k-1} + 1$.
- If $\theta_k > \theta_3^i(X_{T_k}, U_{T_{k-1}}, V_k)$, nothing happens and $N_k = N_{k-1}$.

Following this procedure, the state of the population at time $t \leq \sup_k T_k$ can be recovered as follows: the total number N_t of individuals at time t is equal to $N_t = \sum_{k\geq 0} 1_{\{T_k \leq t < T_{k+1}\}} N_k$, and

$$\nu_t = \sum_{k\geq 0} 1_{\{T_k \leq t < T_{k+1}\}} \sum_{i=1}^{N_k} \delta_{(X^i_t, U^i_{T_k})} = \sum_{i=1}^{N_t} \delta_{(X^i_t, U^i_t)}.$$

3.2 Mathematical formulation and Martingale Properties

In this more technical subsection, we give the mathematical justification of the algorithm described above, and the basic moment and martingale properties of the process, which are the key points of the proofs of the macroscopic approximations of Section 4.

The stochastic individual-based process ν can be rigorously expressed as solution of a stochastic differential equation driven by d-dimensional Brownian motions $(B^i)_{i\in\mathbb{N}^*}$ and the $\mathbb{R}_+ \times \mathbb{N} \times [0,1] \times \mathcal{U}$-valued multivariate point process

$$Q(dt, di, d\theta, dv) = \sum_{k\geq 1} \delta_{(T_k, I_k, \theta_k, V_k)}(dt, di, d\theta, dv)$$

associated with the birth, mutation and death of individuals. We will deduce from moment properties its existence for any time t in $\mathbb{R}^+$.

Let $\nu_0 \in \mathcal{M}$. For each $C_0^{2,0}$-function f, we define the process $\langle \nu_t, f\rangle$ as a solution of the stochastic differential equation, which reproduces formally the algorithm of the previous subsection,

$$\langle \nu_t, f\rangle = \langle \nu_0, f\rangle + \int_0^t \langle \nu_r, m(x,u)\Delta_x f + b(x,u).\nabla_x f\rangle dr$$
$$+ \int_0^t \sum_{i=1}^{\langle \nu_{r-},1\rangle} \sqrt{2m(X_r^i, U_r^i)} \nabla_x f(X_r^i, U_r^i) dB_r^i$$
$$+ \int_{[0,t]\times\mathbb{N}\times[0,1]\times U^2} \Big\{ - f(X_r^i, U_r^i)\mathbf{1}_{\{\theta\leq\theta_1^i(X_r,U_r)\}} + f(X_r^i, U_r^i)\mathbf{1}_{\{\theta_1^i(X_r,U_r)<\theta\leq\theta_2^i(X_r,U_r)\}}$$
$$+ f(X_r^i, v)\mathbf{1}_{\{\theta_2^i(X_r,U_r)<\theta\leq\theta_3^i(X_r,U_r,v)\}} \Big\} Q(dr, di, d\theta, dv). \qquad (3.2)$$

The first two integrals in (3.2) corresponds to the diffusion of each particle and the last integral to the birth, mutation and death jumps in the process. The existence of a solution can be proved inductively following the algorithmic construction described in Subsection 3.1.

By Remark 1.1, the knowledge of $\langle \nu_t, f\rangle$ for $f \in C_0^{2,0}$ is enough to characterized the finite measure-valued process ν.

We introduce the canonical filtration

$$\mathcal{F}_t = \sigma\{\nu_0;\ B_r^j, j\in\mathbb{N}^*;\ Q([0,r]\times A),\ A\in\mathcal{P}(\mathbb{N})\otimes\mathcal{B}([0,1]\times\mathcal{U}),\ r\leq t\},$$

where $\mathcal{B}([0,1]\times\mathcal{U})$ is the Borel σ-field on $[0,1]\times\mathcal{U}$.

Lemma 3.1 *The measure*

$$q(dt,di,d\theta,dv) = C_\delta \sum_{k\geq 0} 1_{\{T_k<t\leq T_{k+1}\}}(N_k+1)\sum_{j=1}^{N_k}\delta_j(di)dtd\theta\frac{M^*(v)}{\|M^*\|_1}dv$$
$$= C_\delta(N_t+1)\sum_{j=1}^{N_t}\delta_j(di)dtd\theta\frac{M^*(v)}{\|M^*\|_1}dv$$

is the (predictable) compensator of the multivariate point process Q.

Proof. For $k\geq 0$, a regular version of the conditional law of $(T_{k+1}, I_{k+1}, \theta_{k+1}, V_{k+1})$ with respect to $\sigma\{\nu_0,\ (B^j_.),\ j\in\mathbb{N}^*,\ (T_p,I_p,\theta_p,V_p),\ 1\leq p\leq k\}$ is given by the measure

$$C_\delta(N_k+1)1_{\{T_k<t\}}e^{-C_\delta N_k(N_k+1)(t-T_k)}\sum_{j=1}^{N_k}\delta_j(di)dtd\theta\frac{M^*(v)}{\|M^*\|_1}dv.$$

The conclusion is thus a consequence of [23] Theorem 1.33 p.136. □

Using Lemma 3.1 and Itô's formula, it can immediately be shown that any solution ν of (3.2), such that $E(\sup_{t\leq T}\langle \nu_t, 1\rangle^2) < +\infty$, is a Markov process with infinitesimal generator L defined by (2.12). Moreover, we also deduce the following existence, moment and martingale properties.

Proposition 3.2 *1) Assume (H) and that $E(\langle \nu_0, 1\rangle) < +\infty$.*

Then $E(\sup_{t\leq T}\langle \nu_t, 1\rangle) < +\infty$ for each $T > 0$ and the process ν defined by (3.2) is well defined on $\mathbb{R}^+$.

2) If furthermore for some $p \geq 1$, $E(\langle \nu_0, 1\rangle^p) < +\infty$, then for each $T > 0$

$$E(\sup_{t\leq T}\langle \nu_t, 1\rangle^p) < +\infty.$$

Proof. We firstly prove 2). Fix $T \geq 0$. For each integer k, define $S_k = \inf\{t \geq 0,\, \langle \nu_t, 1\rangle \geq k\}$. A simple computation using (3.2), and dropping the non-positive death term, gives, for any $t \leq T$,

$$\begin{aligned} E(\sup_{s\in[0,t\wedge S_k]}\langle \nu_s, 1\rangle^p) &\leq CE\left(\langle \nu_0, 1\rangle^p + \int_0^{t\wedge S_k}(1+\langle \nu_s, 1\rangle^p)ds\right) \\ &\leq C\left(1 + E\left(\int_0^t \langle \nu_{s\wedge S_k}, 1\rangle^p ds\right)\right), \end{aligned}$$

where the constants C depend on T. Gronwall's lemma implies the existance of a constant C independent of k, such that $E(\sup_{t\in[0,T\wedge S_k]}\langle \nu_t, 1\rangle^p) \leq C$. Therefore, S_k tends a.s. to infinity when k tends to infinity and Fatou's lemma yields $E(\sup_{t\in[0,T]}\langle \nu_t, 1\rangle^p) < +\infty$.

Point 1) is a consequence of point 2). Indeed, the solution $(\nu_t)_{t\geq 0}$ can be built step by step. To conclude, it is only necessary to check that the sequence of jump times $(T_k)_k$ goes to infinity a.s. as k tends to infinity. But this follows from $E(\sup_{t\leq T}\langle \nu_t, 1\rangle) < +\infty$.
□

The following martingale properties are the key point to study large population approximations (Section 4).

Theorem 3.3 *Assume (H) and that for some $p \geq 2$, $E(\langle \nu_0, 1\rangle^p) < +\infty$.*

1) Then, for F and $f \in C_0^{2,0}$ such that for all $\nu \in \mathcal{M}$, $|F_f(\nu)| + |LF_f(\nu)| \leq C(1 + \langle \nu, 1\rangle^p)$, the process

$$F_f(\nu_t) - F_f(\nu_0) - \int_0^t LF_f(\nu_s)ds$$

is a left limited and right continuous martingale starting from 0. It is in particular true for $F(y) = y^{p-1}$.

2) The process Z^f defined for $f \in C_0^{2,0}$ by

$$Z_t^f = \langle \nu_t, f\rangle - \langle \nu_0, f\rangle - \int_0^t \int_{\bar{\mathcal{X}}\times U} \Big\{ m(x,u)\Delta_x f(x,u) + b(x,u).\nabla_x f(x,u)$$
$$+ \big(\lambda(x,u) - \mu(x,u,I^\delta W \star \nu_s(x,u))\big) f(x,u) + \int_U f(x,v)M(x,u,v)dv \Big\} \nu_s(dx,du)ds \quad (3.3)$$

is a left limite and right continuous L^2-martingale starting from 0 with predictable quadratic variation

$$\langle Z^f\rangle_t = \int_0^t \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{ 2m(x,u)|\nabla_x f|^2 + \big(\lambda(x,u) + \mu(x,u,I^\delta W \star \nu_s(x,u))\big) f^2(x,u)$$
$$+ \int_{\mathcal{U}} f^2(x,v)M(x,u,v)dv \Big\} \nu_s(dx,du)ds \quad (3.4)$$

Proof. Point 1) immediately follows from Itô's formula and Proposition 3.2. For point 2), we first assume that $E(\langle \nu_0, 1\rangle^3) < +\infty$. Applying point 1) with $F(y) = y$ (or (3.2) and Lemma 3.1) leads to Z^f. Applying point 1) again with $F(y) = y^2$ yields that the process

$$\langle \nu_t, f\rangle^2 - \langle \nu_0, f\rangle^2 - \int_0^t \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{ 2(m(x,u)\Delta_x f(x,u) + b(x,u).\nabla_x f(x,u))\langle \nu_s, f\rangle$$
$$+ 2m(x,u)|\nabla_x f|^2 + \lambda(x,u)(2\langle \nu_s, f\rangle f(x,u) + f^2(x,u))$$
$$+ \int_{\mathcal{U}} (2f(x,v)\langle \nu_s, f\rangle + f^2(x,v))M(x,u,v)dv$$
$$+ \mu(x,u,I^\delta W \star \nu_s(x,u))(-2\langle \nu_s, f\rangle f(x,u) + f^2(x,u)) \Big\} \nu_s(dx,du)ds \quad (3.5)$$

is a left limited and right continuous martingale. On another hand, $\langle \nu_t, f\rangle^2$ can be computed from (3.3): by Itô's formula, the process

$$\langle \nu_t, f\rangle^2 - \langle \nu_0, f\rangle^2 - \int_0^t \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{ 2(m(x,u)\Delta_x f(x,u) + b(x,u).\nabla_x f(x,u))\langle \nu_s, f\rangle$$
$$+ 2(\lambda(x,u) - \mu(x,u,I^\delta W \star \nu_s(x,u)))\langle \nu_s, f\rangle f(x,u)$$
$$+ \int_{\mathcal{U}} 2f(x,v)\langle \nu_s, f\rangle M(x,u,v)dv \Big\} \nu_s(dx,du)ds - \langle Z^f\rangle_t \quad (3.6)$$

is a left limited and right continuous martingale. Comparing (3.5) and (3.6) leads to (3.4). The extension to the case where $E(\langle \nu_0, 1\rangle^2) < +\infty$ is straightforward, noticing that $E(\langle Z^f\rangle_t) < +\infty$ $\forall t \geq 0$ for such an initial condition (Proposition 3.2).. □

4 Large population approximation for a fixed interaction range

We are now interested in deterministic approximations of the population point process when the size of the population increases. We assume in this section that the interaction range $\delta > 0$ is fixed.

4.1 Convergence result

We introduce a parameter N in order to scale the population size. This parameter has the following biological interpretation. In a case of fixed amount of available global resources, a large system composed of N individuals may only exist if the biomass of each interacting individual scales as $\frac{1}{N}$, which implies that the interaction effect between two individuals scales as $\frac{1}{N}$ as well. In the case of competition for resources, the parameter N can also be interpreted as scaling the amount of available resources, so that an increase of N implies a decrease of the strength of competition between two individuals for resources.

With this interpretation in mind, we introduce the following scaling of the parameters. For each $N \in \mathbb{N}^*$, we keep all parameters (m, b, λ, M) unchanged, except the competition kernel: we assume that for each N,

$$\mu_N(x, u, r) = \mu(x, u, \frac{r}{N}). \tag{4.1}$$

We consider a sequence of initial measures $(\nu_0^N)_{N \in \mathbb{N}^*}$ belonging to $\mathcal{M}$, and we assume that the sequence $\frac{\nu_0^N}{N}$ converges, as N tends to infinity. The size $\langle \nu_0^N, 1 \rangle$ of the population is then of order N and will stay at this order (or at a smaller order) during finite time-intervals, since birth rates are bounded. Hence, our aim is to study the asymptotic behavior, as N tends to infinity, of the left limited and right continuous process

$$\Lambda_t^N = \frac{1}{N} \sum_{i=1}^{N_t} \delta_{(X_t^i, U_t^i)} = \frac{1}{N} \nu_t^N, \tag{4.2}$$

taking values in $\mathcal{M}^N = \{\frac{1}{N}\nu, \nu \in \mathcal{M}\}$.

The process $(\Lambda_t^N)_{t \geq 0}$ is a Markov process whose generator L_N, as in (2.12), can be decomposed into the sum of a diffusion part $L_{N,1}$ and a jump part $L_{N,2}$. An easy computation gives that, for $F \in C^2(\mathbb{R})$ and $f \in C_0^{2,0}$,

$$L_{N,2} F_f(\nu) = \langle \nu, m(.)\Delta_x f + b(.).\nabla_x f \rangle F'(\langle \nu, f \rangle) + \langle \nu, \frac{m(.)}{N} |\nabla_x f|^2 \rangle F''(\langle \nu, f \rangle) \tag{4.3}$$

and (using (4.1))

$$L_{N,1}F_f(\nu) = N\int_{\bar{\mathcal{X}}\times\mathcal{U}}\Big\{\lambda(x,u)\Big[F\Big(\langle\nu,f\rangle+\frac{1}{N}f(x,u)\Big)-F(\langle\nu,f\rangle)\Big]$$

$$+\mu(x,u,I^\delta W\star\nu(x,u))\Big[F\Big(\langle\nu,f\rangle-\frac{1}{N}f(x,u)\Big)-F(\langle\nu,f\rangle)\Big]$$

$$+\int_{\mathcal{U}}\Big[F\Big(\langle\nu,f\rangle+\frac{1}{N}f(x,v)\Big)-F(\langle\nu,f\rangle)\Big]M(x,u,v)dv\Big\}\nu(dx,du) \qquad (4.4)$$

Similarly, we deduce from Theorem 3.3 the following martingale properties, which will be crucial in the proof of the convergence of Λ^N (see Theorem 4.2 below).

Lemma 4.1 *Let $N\geq 1$ be fixed and assume that for some $p\geq 2$, $E\left(\langle\Lambda_0^N,1\rangle^p\right)<\infty$. For all $C_0^{2,0}$-function f, the process*

$$Z_t^{N,f} = \langle\Lambda_t^N,f\rangle-\langle\Lambda_0^N,f\rangle-\int_0^t\int_{\bar{\mathcal{X}}\times\mathcal{U}}\Big\{m(x,u)\Delta_x f(x,u)+b(x,u).\nabla_x f(x,u)$$

$$+\big(\lambda(x,u)-\mu(x,u,I^\delta W\star\Lambda_s^N(x,u))\big)f(x,u)+\int_{\mathcal{U}}f(x,v)M(x,u,v)dv\Big\}\Lambda_s^N(dx,du)ds \qquad (4.5)$$

is a left limited and right continuous L^2 martingale starting from 0 with predictable quadratic variation

$$\langle Z^{N,f}\rangle_t = \frac{1}{N}\int_0^t\int_{\bar{\mathcal{X}}\times\mathcal{U}}\Big\{2m(x,u)|\nabla_x f|^2+\big(\lambda(x,u)+\mu(x,u,I^\delta W\star\Lambda_s^N(x,u))\big)f^2(x,u)$$

$$+\int_{\mathcal{U}}f^2(x,v)M(x,u,v)dv\Big\}\Lambda_s^N(dx,du)ds \qquad (4.6)$$

We assume

Assumption (H1):

1) The initial measures Λ_0^N converge in law and for the weak topology on $M_F(\bar{\mathcal{X}}\times\mathcal{U})$ to some deterministic finite measure $\xi_0\in M_F(\bar{\mathcal{X}}\times\mathcal{U})$, and $\sup_N E(\langle\Lambda_0^N,1\rangle^3)<+\infty$.

2) All the parameters of the model are assumed to be continuous, either on $\bar{\mathcal{X}}\times\mathcal{U}$, or on $\bar{\mathcal{X}}\times\mathcal{U}\times\mathbb{R}$.

3) There exists a constant k_μ such that

$$\forall x\in\mathcal{X},\ u\in\mathcal{U},\ r_1,r_2\in\mathbb{R}_+,\quad |\mu(x,u,r_1)-\mu(x,u,r_2)|\leq k_\mu|r_1-r_2|. \qquad (4.7)$$

By the law of large numbers, Assumption (H1-1) is for example satisfied for $\Lambda_0^N=\frac{1}{N}\sum_{i=1}^N\delta_{(X_0^i,U_0^i)}$, when $(X_0^i,U_0^i)_{\{1\leq i\leq N\}}$ are independent random variables with law ξ_0

with finite 3rd-order moment.

Let us recall that the parameters of diffusion (m, b), birth (λ) and mutation (M) associated with Λ^N are independent of N, whereas the death and interaction parameter μ_N is defined by (4.1). The following result shows that, when the population size increases, the renormalized population process Λ^N converges to a macroscopic deterministic limit.

Theorem 4.2 *Assume (H) and (H1), and consider the sequence of processes Λ^N defined by (4.2). Then for all $T>0$, the sequence (Λ^N) converges in law, in $\mathbb{D}([0,T],M_F(\bar{\mathcal{X}}\times\mathcal{U}))$, to a deterministic continuous function ξ^δ belonging to $C([0,T],M_F(\bar{\mathcal{X}}\times\mathcal{U}))$.*

This measure-valued function ξ^δ is the unique weak solution satisfying $\sup_{t\in[0,T]}\langle\xi^\delta_t,1\rangle<+\infty$ of the following (weak form of a) nonlinear integro-differential equation. For all function $f\in C^{2,b}_0$,

$$\langle\xi^\delta_t,f\rangle=\langle\xi_0,f\rangle+\int_0^t\int_{\bar{\mathcal{X}}\times\mathcal{U}}\Big\{m(x,u)\Delta_xf(x,u)+b(x,u).\nabla_xf(x,u)$$
$$+\big(\lambda(x,u)-\mu(x,u,I^\delta W\star\xi^\delta_s(x,u))\big)f(x,u)+\int_{\mathcal{U}}f(x,v)M(x,u,v)dv\Big\}\ \xi^\delta_s(dx,du)ds \tag{4.8}$$

Remark 4.3 *In the case where $\xi^\delta_t(dx,du)$ has a density $g_t(x,u)$ with repsect to the Lebesgue measure $dxdu$ on $\bar{\mathcal{X}}\times\mathcal{U}$ for any $t\geq 0$, then $g_t(x,u)$ is weak solution to the nonlocal and nonlinear partial integro-differential equation*

$$\begin{aligned}
&\partial_t g^\delta_t=\Delta_x(m(x,u)g^\delta_t(x,u))-\nabla_x(b(x,u)g^\delta_t(x,u))\\
&\qquad+\big(\lambda(x,u)-\mu(x,u,I^\delta W\star g^\delta_t(x,u))\big)g^\delta_t(x,u)+\int_{\mathcal{U}}g^\delta_t(x,v)M(x,u,v)dv\ ;\\
&g^\delta_0(x,u)=g_0(x,u)\ ;\\
&\nabla_x g^\delta(t,x,u).n(x)=0\quad\forall(t,x,u)\in\mathbb{R}_+\times\partial\mathcal{X}\times\mathcal{U}.
\end{aligned}\tag{4.9}$$

This equation has Neumann's boundary condition because of the underlying spatial motion is a reflected diffusion. Its weak form is given by (4.8). See Subsection 4.3 for conditions ensuring the existence of a density for ξ^δ_t.

Remark 4.4 *Applying (4.8) to the constant function equal to 1, the positivity of μ and Hypotheses (H) gives $\langle\xi^\delta_t,1\rangle\ \leq\ \langle\xi_0,1\rangle+C\int_0^t\langle\xi^\delta_s,1\rangle ds$. We conclude by Gronwall's lemma that any solution ξ^δ of (4.8) is bounded on every finite time interval $[0,T]$:*

$$\sup_{t\in[0,T]}\langle\xi^\delta_t,1\rangle\ \leq\ \langle\xi_0,1\rangle e^{CT}.$$

4.2 Proof of Theorem 4.2

As a first step in the proof of Theorem 4.2, we now give a mild formulation for solutions of (4.8). To this aim, and for each fixed trait $u \in \mathcal{U}$, we denote by P^u the semigroup of the diffusion process normally reflected at the boundary of $\mathcal{X}$, with diffusion matrix $m(\cdot, u)\mathrm{Id}$ and drift coefficient $b(\cdot, u)$.

Lemma 4.5 *Let us consider a solution ξ^δ of (4.8). Then, for each continuous and bounded function φ defined on $\bar{\mathcal{X}} \times \mathcal{U}$,*

$$\langle \xi^\delta_t, \varphi\rangle = \langle \xi_0, P^u_t \varphi\rangle + \int_0^t \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{ \big(\lambda(x,u) - \mu(x,u, I^\delta W \star \xi^\delta_s(x,u))\big) P^u_{t-s}\varphi(x,u) + \int_{\mathcal{U}} P^v_{t-s}\varphi(x,v) M(x,u,v)dv \Big\}\, \xi^\delta_s(dx,du)ds. \quad (4.10)$$

Proof. We may classically derive from (4.8) a space-time weak equation for measurable functions $\psi_s(x,u) = \psi(s,x,u)$ which are of class $C^{1,2}$ on $[0,t]\times\bar{\mathcal{X}}$, measurable and bounded in u and such that $\partial_n \psi = 0$ on $[0,t] \times \partial\mathcal{X} \times \mathcal{U}$:

$$\langle \xi^\delta_t, \psi_t\rangle = \langle \xi_0, \psi_0\rangle + \int_0^t \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{ \partial_s\psi_s(x,u) + m(x,u)\Delta_x\psi_s(x,u) + b(x,u).\nabla_x\psi_s(x,u) + \big(\lambda(x,u) - \mu(x,u, I^\delta W \star \xi^\delta_s(x,u))\big)\psi_s(x,u) + \int_{\mathcal{U}} \psi_s(x,v)M(x,u,v)dv \Big\}\, \xi^\delta_s(dx,du)ds \quad (4.11)$$

Let us now consider a continuous function φ on $\bar{\mathcal{X}} \times \mathcal{U}$ and fix a time $t \in [0,T]$. Let us define for $(s,x,u) \in [0,t] \times \bar{\mathcal{X}} \times \mathcal{U}$,

$$\psi_s(x,u) = P^u_{t-s}\varphi(x,u).$$

Then ψ is solution of the boundary value problem

$$\begin{aligned} &\partial_s\psi_s(x,u) + m(x,u)\Delta_x\psi_s(x,u) + b(x,u).\nabla_x\psi_s(x,u) = 0 \quad \text{on } [0,T]\times\mathcal{X}\times\mathcal{U} \\ &\partial_n\psi_s(x,u) = 0 \quad \text{on } [0,T]\times\partial\mathcal{X}\times\mathcal{U} \\ &\psi_t(x,u) = \varphi(x,u) \quad \text{on } \bar{\mathcal{X}}\times\mathcal{U}. \end{aligned}$$

Equation (4.11) applied to this function ψ yields (4.10). □

Proof of Theorem 4.2. Fix $T > 0$.

We firstly prove the uniqueness of solutions ξ of (4.8). By Remark 4.4 and Lemma 4.5, it is equivalent to prove the uniqueness of solutions of (4.10). Let us consider two such solutions $(\xi_t)_{t\geq 0}$ and $(\bar{\xi}_t)_{t\geq 0}$ and compute the quantity $|\langle \xi_t - \bar{\xi}_t, \varphi\rangle|$, for each continuous function φ such that $\|\varphi\|_\infty \leq 1$.

Using (4.10), we obtain for $t \leq T$

$$
\begin{aligned}
|\langle \xi_t - \bar{\xi}_t, \varphi \rangle| \leq & \int_0^t \Big| \int_{\bar{\mathcal{X}} \times \mathcal{U}} \Big\{ \big(\lambda(x,u) - \mu(x,u,I^\delta W \star \xi_s(x,u))\big) P^u_{t-s}\varphi(x,u) \\
& \qquad + \int_{\mathcal{U}} P^v_{t-s}\varphi(x,v) M(x,u,v) dv \Big\} \big(\xi_s(dx,du) - \bar{\xi}_s(dx,du)\big) \Big| ds \\
+ \int_0^t \int_{\bar{\mathcal{X}} \times \mathcal{U}} & \Big| \big(\mu(x,u,I^\delta W \star \bar{\xi}_s(x,u)) - \mu(x,u,I^\delta W \star \xi_s(x,u))\big) P^u_{t-s}\varphi(x,u) \Big| \bar{\xi}_s(dx,du) ds
\end{aligned}
$$

Now, using Assumption (H), Remark 4.4 and the fact that $\|\varphi\|_\infty \leq 1$, there exists a positive constant C_1 such that for all $(x,u) \in \bar{\mathcal{X}} \times \mathcal{U}$ and all $0 < s \leq t \leq T$,

$$
\begin{aligned}
|\lambda(x,u) P^u_{t-s}\varphi(x,u) + \int_{\mathcal{U}} P^v_{t-s}\varphi(x,v) M(x,u,v) dv| & \leq C_1, \\
|\mu(x,u,I^\delta W \star \bar{\xi}_s(x,u)) P^u_{t-s}\varphi(x,u)| \leq \mu_0(1 + \|I^\delta W\|_\infty \langle \bar{\xi}_s, 1 \rangle) & \leq C_1
\end{aligned}
$$

while thanks to (H1-2),

$$
|\mu(x,u,I^\delta W \star \bar{\xi}_s(x,u)) - \mu(x,u,I^\delta W \star \xi_s(x,u))| \leq k_\mu \|I^\delta W\|_\infty \sup_{\|\varphi\|_\infty \leq 1} |\langle \xi_s - \bar{\xi}_s, \varphi \rangle|.
$$

Then

$$
\begin{aligned}
\Big| \int_{\bar{\mathcal{X}} \times \mathcal{U}} \Big(\mu(x,u,I^\delta W \star \bar{\xi}_s(x,u)) - \mu(x,u,I^\delta W \star \xi_s(x,u)) \Big) P^u_{t-s}\varphi(x,u) \bar{\xi}_s(dx,du) \Big| \\
\leq C_2 \sup_{\|\varphi\|_\infty \leq 1} |\langle \xi_s - \bar{\xi}_s, \varphi \rangle|
\end{aligned}
$$

where C_2 is a positive constant. Therefore, there exists $C > 0$ such that

$$
|\langle \xi_t - \bar{\xi}_t, \varphi \rangle| \leq C \int_0^t \sup_{\|\varphi\|_\infty \leq 1} |\langle \xi_s - \bar{\xi}_s, \varphi \rangle| ds
$$

and Gronwall's lemma yields $\sup_{\|\varphi\|_\infty \leq 1} |\langle \xi_t - \bar{\xi}_t, \varphi \rangle| = 0$ for all $t \leq T$, where the supremum is taken over all continuous functions. Since $\xi_s - \bar{\xi}_s$ is a measure on $\mathbb{R}^{d+q}$, we can apply Lusin's theorem (see [5]) to deduce that the same equality holds when the supremum is taken over all $L^1(\xi_s + \bar{\xi}_s) \cap L^\infty$ functions. Therefore, the total variation norm of $\xi_t - \bar{\xi}_t$ is zero for all $t \leq T$ and uniqueness holds.

Let us next prove that for all $T > 0$,

$$
\sup_{N \in \mathbb{N}^*} E\left(\sup_{[0,T]} \langle \Lambda^N_t, 1 \rangle^3 \right) < +\infty \tag{4.12}
$$

Define $S^N_k = \inf\{t \geq 0, \langle \Lambda^N_t, 1 \rangle \geq k\}$ for $k \in \mathbb{N}^*$. A simple computation using the specific form of $L_{N,1}F_f$ and $L_{N,2}F_f$ with $f = 1$ and $F(y) = y^3$, dropping the negative death term, yields

$$E\left(\sup_{s\le t\wedge S_k^N}\left\langle \Lambda_s^N,1\right\rangle^3\right)\le E(\left\langle \Lambda_0^N,1\right\rangle^3)+CE\left(\int_0^{t\wedge S_k^N}(\langle \Lambda_s^N,1\rangle+\left\langle \Lambda_s^N,1\right\rangle^3)ds\right)$$

where C is a positive constant independent of k and N. Then Assumption (H1-1) and Gronwall's lemma imply that there exists a constant C_T independent of k and N such that $E\left(\sup_{s\le T\wedge S_k^N}\left\langle \Lambda_s^N,1\right\rangle^3\right)\le C_T$. We deduce that the sequence $(S_k^N)_k$ tends a.s. to infinity and finally obtain (4.12) by Fatou's lemma.

Now, we will follow a classical method to prove our convergence result: we will first prove the uniform tightness of the sequence of the laws Q^N of the processes Λ^N (i.e. the compactness of this sequence), and next, we will prove that any accumulation point of this sequence has the required limit law. Following Roelly [39] and using Remark 1.1, in order to prove the tightness of the sequence $(Q^N)_N$ in $\mathcal{P}(\mathbb{D}([0,T],M_F(\bar{\mathcal{X}}\times\mathcal{U})))$, where M_F is endowed with the *vague* topology, it is sufficient to prove that the laws of the processes $\left\langle \Lambda^N,f\right\rangle$ form a tight sequence in $\mathcal{P}(\mathbb{D}([0,T],\mathbb{R}))$ for any function $f\in C_0^{2,0}$. By Aldous' [1] and Rebolledo's [24] criteria, this tightness follows from

$$\sup_{N\in\mathbb{N}^*}E(\sup_{[0,T]}|\left\langle \Lambda_s^N,f\right\rangle|)<\infty, \tag{4.13}$$

and from the tightness of the laws of $\langle Z^{N,f}\rangle$ and of the drift part of the semimartingales $\left\langle \Lambda^N,f\right\rangle$.

Clearly, since f is bounded, (4.13) is a consequence of (4.12). Let us now consider stopping times (S,S') satisfying a.s. $0\le S\le S'\le S+\delta\le T$. Thanks to Doob's inequality, Lemma 4.1, and (4.12), we get

$$E\left(\langle Z^{N,f}\rangle_{S'}-\langle Z^{N,f}\rangle_S\right)\le E\left(C\int_S^{S+\delta}\left(\langle \Lambda_s^N,1\rangle+\left\langle \Lambda_s^N,1\right\rangle^2\right)ds\right)\le C\delta.$$

Similar arguments prove that the expectation of the finite variation part of $\left\langle \Lambda_{S'}^N,f\right\rangle-\left\langle \Lambda_S^N,f\right\rangle$ is bounded by $C\delta$. Therefore, by Aldous' criterion [1], the sequence $(Q^N)_N$ is uniformly tight for the vague topology.

Let us now denote by Q the limiting law in $\mathcal{P}(\mathbb{D}([0,T],M_F(\bar{\mathcal{X}}\times\mathcal{U})))$ of a subsequence of Q^N, still denoted by Q^N for simplicity. By construction, almost surely,

$$\sup_{t\in[0,T]}\sup_{||f||_\infty\le 1}|\langle \Lambda_s^N,f\rangle-\langle \Lambda_{s-}^N,f\rangle|\le 1/N.$$

We deduce immediately that each process Λ with law Q is a.s. strongly continuous. Let us finally prove that it is the unique solution of (4.8).

For $t \leq T$, $f \in C_0^{2,0}$ and $\nu \in \mathbb{D}([0,T], M_F(\bar{\mathcal{X}} \times \mathcal{U}))$, we define

$$\Psi_{t,f}(\nu) = \langle \nu_t, f\rangle - \langle \nu_0, f\rangle - \int_0^t \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{ m(x,u)\Delta_x f(x,u) + b(x,u).\nabla_x f(x,u)$$
$$+ \big(\lambda(x,u) - \mu(x,u,I^\delta W \star \nu_s(x,u))\big) f(x,u) + \int_{\mathcal{U}} f(x,v)M(x,u,v)dv \Big\} \nu_s(dx)ds.$$

We want todeduce that, for any $t \leq T$,

$$E\left(|\Psi_{t,f}(\Lambda)|\right) = 0, \tag{4.14}$$

from the fact that (see Lemma 4.1)

$$Z_t^{N,f} = \Psi_{t,f}(\Lambda^N). \tag{4.15}$$

A fair computation using Lemma 4.1, Assumptions (H) and (H1), and (4.12) shows that

$$E\left(|Z_t^{N,f}|^2\right) = E\left(\langle Z^{N,f}\rangle_t\right) \leq \frac{C_f}{N} E\left(\int_0^t \left\{1 + \langle \Lambda_s^N, 1\rangle^2\right\} ds\right) \leq \frac{C_{f,t}}{N} \tag{4.16}$$

which goes to 0 as N tends to infinity. On another hand, since Λ is a.s. strongly continuous and $f \in C_0^{2,0}$, the function $\Psi_{t,f}$ is a.s. continuous at Λ. Furthermore, for any $\nu \in \mathbb{D}([0,T], M_F(\bar{\mathcal{X}} \times \mathcal{U}))$,

$$|\Psi_{t,f}(\nu)| \leq C_{t,f} \sup_{[0,t]} \left(1 + \langle \nu_s, 1\rangle^2\right),$$

and (4.12) implies that the sequence $(\Psi_{t,f}(\Lambda^N))_N$ is uniformly integrable. Thus

$$\lim_N E\left(|\Psi_{t,f}(\Lambda^N)|\right) = E\left(|\Psi_{t,f}(\Lambda)|\right) \tag{4.17}$$

Combining (H1-1), (4.15), (4.16) and (4.17), we conclude that (4.14) holds and that (4.8) is satisfied for any $f \in C_0^{2,0}$.

Then Λ is a.s. equal to the deterministic process ξ^δ, and the sequence (Λ^N) converges to ξ^δ in $\mathbb{D}([0,T], M_F(\bar{\mathcal{X}} \times \mathcal{U}))$, where $M_F(\bar{\mathcal{X}} \times \mathcal{U})$ is endowed with the vague topology. To extend this result to the weak topology, we use a criterion proved in [30]. Since the limiting process is continuous, it suffices to prove that the sequence $(\langle \Lambda^N, 1\rangle)_N$ converges in law to $\langle \xi^\delta, 1\rangle$ in $\mathbb{D}([0,T], \mathbb{R})$. We may apply what has been done above with $f \equiv 1$. Theorem 4.2 is proved. □

4.3 Existence of a density for the limit

In the next section, we will be interested in the limit of small spatial interaction range δ. A preliminary result, also interesting by itself, consists in proving the existence of a density for the measure ξ_t^δ, $t \geq 0$. We make the additional

Assumption (H2):

1) The diffusion coefficient $m(x,u)$ is of class C^2 in $x = (x_1, \ldots, x_d)$ and the second derivatives of m with respect to $x_1, \ldots, x_d$ are α-Hölderian, uniformly in u, with $\alpha > 0$. Moreover, m is assumed to be positive. Hence, since $\bar{\mathcal{X}} \times \mathcal{U}$ is a compact set, there exists $m_* > 0$ such that for all $(x,u) \in \bar{\mathcal{X}} \times U$,

$$m(x,u) \geq m_* > 0.$$

2) The drift coefficient $b(x,u)$ is of class C^1 in x and the derivatives of b with respect to $x_1, \ldots, x_d$ are α-Hölderian, uniformly in u, with $\alpha > 0$.

Let us prove that, if $\xi_0^\delta(dx,du)$ has a density with respect to the Lebesgue measure $dxdu$ (i.e. the measure ξ_0^δ has no singularity, such as Dirac masses), then this property propagates in time. In other words, on bounded time intervals, modes can appear, but no singularity.

Theorem 4.6 *Assume (H), (H1) and (H2) and that $\xi_0(dx,du) = g_0(x,u)dxdu$. Then for each time t, the measure $\xi_t^\delta(dx,du)$ has a density $g_t^\delta \in L^\infty([0,T], L^1)$ with respect to the Lebesgue measure on $\bar{\mathcal{X}} \times \mathcal{U}$. Moreover, g^δ is weak solution to the nonlocal nonlinear partial integro-differential equation (4.9) and, for each t and u, the function $g_t^\delta(.,u)$ is continuous on $\bar{\mathcal{X}}$.*

The end of this section is devoted to the proof of this result.

Proof. As a preliminary result, assumptions (H) and (H2) and the smoothness of $\partial\mathcal{X}$ allow us to adapt Sato-Ueno [42] (Theorem 2.1 and Appendix) to obtain the following lemma.

Lemma 4.7 *There exists a unique function $p_t(x,u,y)$ defined on $\mathbb{R}_+ \times \bar{\mathcal{X}} \times \mathcal{U} \times \bar{\mathcal{X}}$, continuous in (t,x,y) and which is a density function in $y \in \bar{\mathcal{X}}$ such that, for each continuous function φ defined on $\bar{\mathcal{X}} \times U$ and each $(x,u) \in \bar{\mathcal{X}} \times \mathcal{U}$,*

$$P_t^u \varphi(x,u) = \int_{\bar{\mathcal{X}}} p_t(x,u,y)\varphi(y,u)dy. \tag{4.18}$$

Let us come back to the equation (4.10) satisfied by ξ^δ. Using basic results on linear parabolic equations, we construct by induction a sequence of functions $(g_n)_n$ satisfying in a weak sense the following semi-implicit scheme

$$
\begin{aligned}
&\partial_t g_t^{n+1}(x,u) = \Delta_x(m(x,u)g_t^{n+1}(x,u)) - \nabla_x(b(x,u)g_t^{n+1}(x,u)) \\
&\qquad + \lambda(x,u)g_t^n(x,u) + \int_{\mathcal{U}} g_t^n(x,v)M(x,u,v)dv - \mu(x,u,I^\delta W \star g_t^n(x,u))g_t^{n+1}(x,u) \\
&g_0^{n+1}(x,u) = g_0(x,u) \\
&\nabla_x g^{n+1}(t,x,u).n(x) = 0 \quad \forall (t,x,u) \in \mathbb{R}_+ \times \partial\mathcal{X} \times \mathcal{U}.
\end{aligned}
\tag{4.19}
$$

Thanks to the nonnegativity of g_0, μ, λ and M, and applying the maximum principle, we can show that the functions g_n are nonnegative (see [8]). By symmetry of M, Equation (4.19) (understood in the weak sense) means that for all $C_0^{2,b}$-function φ from $\bar{\mathcal{X}} \times \mathcal{U}$ into $\mathbb{R}$,

$$
\begin{aligned}
\langle g_t^{n+1}, \varphi\rangle = \langle g_0, \varphi\rangle + \int_0^t \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{ &\Big(m(x,u)\Delta_x\varphi(x,u) + b(x,u).\nabla_x\varphi(x,u)\Big) g_s^{n+1}(x,u) \\
&+ \Big(\lambda(x,u)\varphi(x,u) + \int_{\mathcal{U}} \varphi(x,v)M(x,u,v)dv\Big) g_s^n(x,u) \\
&- \mu(x,u,I^\delta W \star g_s^n(x,u))\varphi(x,u)g_s^{n+1}(x,u)\Big\} dx\, du\, ds.
\end{aligned}
\tag{4.20}
$$

Similarly as in in Lemma 4.5, the corresponding mild equation writes as foolows. For each continuous function φ,

$$
\begin{aligned}
\langle g_t^{n+1}, \varphi\rangle = &\int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big(\int_{\bar{\mathcal{X}}} p_t(x,u,y)\varphi(y,u)dy\Big) g_0(x,u)dxdu \\
&+ \int_0^t \int_{\bar{\mathcal{X}}\times\mathcal{U}} \Big\{\Big[\lambda(x,u)\Big(\int_{\bar{\mathcal{X}}} p_{t-s}(x,u,y)\varphi(y,u)dy\Big) \\
&+ \int_{\mathcal{U}} \Big(\int_{\bar{\mathcal{X}}} p_{t-s}(x,v,y)\varphi(y,v)dy\Big) M(x,u,v)dv\Big] g_s^n(x,u) \\
&- \mu(x,u,I^\delta W \star g_s^n(x,u))\Big(\int_{\bar{\mathcal{X}}} p_{t-s}(x,u,y)\varphi(y,u)dy\Big) g_s^{n+1}(x,u)\Big\} dxduds.
\end{aligned}
\tag{4.21}
$$

the assumptions on the coefficients allow us to apply Fubini's theorem. This yields, for each $(y,u) \in \bar{\mathcal{X}} \times \mathcal{U}$,

$$
\begin{aligned}
g_t^{n+1}(y,u) = &\int_{\bar{\mathcal{X}}} p_t(x,u,y)g_0(x,u)dx \\
&+ \int_0^t \int_{\bar{\mathcal{X}}} \Big\{(\lambda(x,u)p_{t-s}(x,u,y)g_s^n(x,u) + \int_{\mathcal{U}} p_{t-s}(x,u,y)g_s^n(x,v)M(x,u,v)dv \\
&- \mu(x,u,I^\delta W \star g_s^n(x,u))p_{t-s}(x,u,y)g_s^{n+1}(x,u)\Big\} dxds.
\end{aligned}
\tag{4.22}
$$

Thanks to the nonnegativity of g^{n+1}, we get

$$0 \leq g_t^{n+1}(y,u) \leq \int_{\bar{\mathcal{X}}} p_t(x,u,y) g_0(x,u) dx + \int_0^t \int_{\bar{\mathcal{X}}} \Big\{ (\lambda(x,u) p_{t-s}(x,u,y) g_s^n(x,u) \\ + \int_{\mathcal{U}} p_{t-s}(x,u,y) g_s^n(x,v) M(x,u,v) dv \Big\} dx ds. \quad (4.23)$$

Integrating over $y \in \bar{\mathcal{X}}$, using Fubini's Theorem, the symmetry of M and Gronwall's Lemma, we easily deduce that there exists a constant C independent of δ such that

$$\sup_{n\in\mathbb{N}} \sup_{t\leq T} \|g_t^n\|_1 \leq \|g_0\|_1 e^{CT}. \quad (4.24)$$

Let us now show the convergence of the sequence g^n in $L^\infty([0,T],L^1)$ to a function g^δ. A straightforward computation using (4.22), Hypotheses (H), (H1) and (H2), and similar arguments as above yields

$$\sup_{s\leq t} \|g_s^{n+1} - g_s^n\|_1 \leq C \int_0^t \left(\sup_{u\leq s} \|g_u^{n+1} - g_u^n\|_1 + \sup_{u\leq s} \|g_u^n - g_u^{n-1}\|_1 \right) ds$$

where C is a positive constant. It follows from Gronwall's Lemma that for each $0 \leq t \leq T$ and each n, $\sup_{s\leq t} \|g_s^{n+1} - g_s^n\|_1 \leq C \int_0^t \sup_{u\leq s} \|g_u^n - g_u^{n-1}\|_1 ds$.

Picard's Lemma then yields $\sum_{n\in\mathbb{N}} \sup_{t\in[0,T]} \|g_t^{n+1} - g_t^n\|_1 < +\infty$ for any $T > 0$. Therefore, the sequence $(g^n)_n$ converges in $L^\infty([0,T],L^1)$ to a function g^δ which satisfies, by (4.24),

$$\sup_{\delta>0} \sup_{t\leq T} \|g_t^\delta\|_1 \leq \|g_0\|_1 e^{CT}. \quad (4.25)$$

Moreover, the function g^δ is weak solution of (4.9). Hence, by the uniqueness result proved in Theorem 4.2, $\xi^\delta(dx,du) = g^\delta(x,u) dx du$. Moreover, Lemma 4.5 implies that g^δ is also solution of the mild equation

$$g_t^\delta(y,u) = \int_{\mathcal{X}} p_t(x,u,y) g_0(x,u) dx \\ + \int_0^t \int_{\mathcal{X}} \Big\{ \big(\lambda(x,u) - \mu(x,u,I^\delta W \star g_s^\delta(x,u))\big) p_{t-s}(x,u,y) g_s^\delta(x,u) \\ + \int_{\mathcal{U}} p_{t-s}(x,u,y) g_s^\delta(x,v) M(x,u,v) dv \Big\} dx ds. \quad (4.26)$$

Using (4.26), the continuity of $y \mapsto g_t^\delta(y,u)$ follows immediately from the continuity of $(t,x,y) \mapsto p_t(x,u,y)$ and the boundedness of g^δ, λ, μ and M. $\square$

5 Convergence of the number density when the interaction range decreases

In this section, our aim is to scale space in order to observe what happens when the spatial interaction range is very local ($\delta \to 0$). We obtain the convergence of the solution g^δ of (4.9) to a nonlinear partial intergo-differential equation with localized spatial interaction. In this equation, the space and trait convolution with respect to the kernel $I^\delta W$ in the nonlinear term is replaced by a convolution in the trait space only with respect to W. Our motivation is to recover from an individual-based model a class a models studied by Prévost [36] and Desvillettes et al. [8].

In order to control the nonlinear term in the equation (4.9) satisfied by g^δ, we need L^∞-estimates, independent on δ, on g^δ. To this end, we make the following additional assumption.

(H3) *The initial density $g_0(x,u)$ is independent of δ and bounded on $\mathcal{X} \times \mathcal{U}$.*

Proposition 5.1 *Assume (H), (H1), (H2), (H3). Then there exists a positive constant C_T, such that*

$$\sup_{\delta>0} \sup_{t\in[0,T]} \|g_t^\delta\|_\infty \leq C_T \|g_0\|_\infty. \tag{5.1}$$

Proof. Again, we consider the sequence $(g^n)_n$ approximating g^δ introduced in the proof of Theorem 4.6. The maximum principle implies that $\sup_n \sup_{t\leq T} \|g_t^n\|_\infty \leq C\|g_0\|_\infty$, where $C > 0$ is a constant only depending on T, λ^* and M^* (and independent of δ). This property is preserved when $n \to +\infty$, and (5.1) is proved. (For details on the maximum principle, see [8]) $\square$

Let us now prove the convergence of g^δ, as the interaction range δ goes to 0, to a locally nonlinear partial integro-differential equation with Neumann's boundary condition (again, because of the reflected diffusion governing the motion of individuals). The convergence of the interaction range to 0 translates in this result into the assumption that $I^\delta(y)dy$ converges weakly to the Dirac measure δ_0 at 0. One can think for example to the case where $I^\delta(x) = C_\delta \mathbf{1}_{\{|x|\leq\delta\}}$.

Theorem 5.2 *Assume (H), (H1), (H2), (H3) and that the measure $I^\delta(y)dy$ weakly converges to δ_0 as δ tends to 0. Then the sequence $(g^\delta)_{\delta>0}$ converges in $L^\infty([0,T], L^1)$ as δ tends to 0, to the unique function $g \in L^\infty([0,T], L^1 \cap L^\infty(\bar{\mathcal{X}} \times \mathcal{U}))$ with initial condition*

g_0 *satisfying (in a weak sense)*

$$
\begin{aligned}
\partial_t g_t(x,u) &= \Delta_x(m(x,u)g_t(x,u)) - \nabla_x(b(x,u)g_t(x,u)) \\
&\qquad + \big(\lambda(x,u) - \mu(x,u,\rho_g(t,x,u))\big)g_t(x,u) + \int_{\mathcal{U}} g_t(x,v)M(x,u,v)dv \ ; \\
\nabla_x g(t,x,u).n(x) &= 0 \quad \forall (t,x,u) \in \mathbb{R}_+ \times \partial\mathcal{X} \times \mathcal{U}
\end{aligned}
\tag{5.2}
$$

where ρ_g describes the (local) interaction in x. It is defined for $(x,u) \in \bar{\mathcal{X}} \times \mathcal{U}$ by

$$
\rho_g(t,x,u) = \int_{\mathcal{U}} W(u-v)g_t(x,v)dv.
$$

Moreover, for each t and u, the function $g_t(.,u)$ is continuous on $\mathcal{X}$.

Proof. As before, we will use the mild formulation of (5.2): a solution g of (5.2) in $L^\infty([0,T], L^1 \cap L^\infty(\bar{\mathcal{X}} \times \mathcal{U}))$ satisfies, for each $y,u \in \bar{\mathcal{X}} \times \mathcal{U}$,

$$
\begin{aligned}
g_t(y,u) &= \int_{\mathcal{X}} p_t(x,u,y)g_0(x,u)dx \\
&\quad + \int_0^t \int_{\mathcal{X}} \Big\{ \big(\lambda(x,u) - \mu(x,u,\rho_g(s,x,u))\big)p_{t-s}(x,u,y)g_s(x,u) \\
&\qquad\qquad + \int_{\mathcal{U}} p_{t-s}(x,u,y)g_s(x,v)M(x,u,v)dv \Big\} dxds.
\end{aligned}
\tag{5.3}
$$

One can easily prove the existence and uniqueness of the integrable and bounded function g solution of (5.3) by adapting the proofs of Theorem 4.6 and Proposition 5.1. The continuity of $y \to g_t(y,u)$ is obtained as in the proof of Theorem 4.6, and we can show as in the proof of Proposition 5.1 that

$$
\sup_{t\in[0,T]} \|g_t\|_\infty \le C_T \|g_0\|_\infty.
\tag{5.4}
$$

Let us write

$$
\begin{aligned}
g_t^\delta(y,u) - g_t(y,u) &= \int_0^t \int_{\mathcal{X}} \Big\{ (\lambda(x,u)p_{t-s}(x,u,y)\big(g_s^\delta(x,u) - g_s(x,u)\big) \\
&\quad + \int_{\mathcal{U}} p_{t-s}(x,u,y)\big(g_s^\delta(x,v) - g_s(x,v)\big)M(x,u,v)dv \\
&\quad - \Big[\mu(x,u,I^\delta W \star g_s^\delta(x,u))g_s^\delta(x,u) - \mu(x,u,\rho_g(s,x,u))g_s(x,u)\Big]p_{t-s}(x,u,y)\Big\} dxds
\end{aligned}
\tag{5.5}
$$

We want to compute an upper bound for the norm of $g_t^\delta(y,u) - g_t(y,u)$ that allows us to apply Gronwall's lemma. Because of (5.1) and (5.4), the unique term which deserves attention is the term $\mu(x,u,\rho_g(x,u)) - \mu(x,u,I^\delta W \star g^\delta(x,u))$. By (4.7), we have

$$\begin{aligned}
&\int_{\mathcal{X}} |\mu(x,u,\rho_g(t,x,u)) - \mu(x,u,I^\delta W \star g_t^\delta(x,u))| dx \\
&\le k_\mu \int_{\mathcal{X}} \left| \int_{\mathcal{U}} W(u-v) g_t(x,v) dv - \int_{\mathcal{X}\times\mathcal{U}} I^\delta(x-z) W(u-v) g_t^\delta(z,v) dz dv \right| dx \\
&\le k_\mu \int_{\mathcal{X}} \Bigg(\left| \int_{\mathcal{U}} W(u-v) g_t(x,v) dv - \int_{\mathcal{X}\times\mathcal{U}} I^\delta(x-z) W(u-v) g_t(z,v) dz dv \right| \\
&\qquad + \int_{\mathcal{X}\times\mathcal{U}} I^\delta(x-z) W(u-v) \left| g_t(z,v) - g_t^\delta(z,v) \right| dz dv \Bigg) dx
\end{aligned}$$

Let us fix our attention on the first term in the right-hand side of this inequality, that we will call $A_\delta(t,u)$. Since $I^\delta(y)dy$ weakly converges to δ_0, $\int_{\mathcal{X}} I^\delta(x-z) g_t(z,v) dz$ converges to $g_t(x,v)$ as δ goes to 0. Because of (5.4), this convergence holds in a bounded pointwise sense with respect to $t \le T$, $x \in \mathcal{X}$ and $v \in \mathcal{U}$. Then Lebesgue's theorem implies that $A_{\delta,T} := \int_{\mathcal{U}} \int_0^T A_\delta(t,u) dt du$ tends to 0 as δ tends to 0.

Therefore, integrating (5.5) with respect to $dy\, du$, we get

$$\sup_{s\le t} \|g_s^\delta - g_s\|_1 \le C_T A_{\delta,T} + C'_T \int_0^t \sup_{u \le s} \|g_u^\delta - g_u\|_1 ds.$$

We conclude using Gronwall's lemma. □

The zero interaction range equation (5.2) has been numerically studied in Prévost [36]. A lot of simulations based on finite element schemes are given, studying the simultaneous effects of diffusion, mutation and selection parameters on the invasion of the space and trait domain by the population. The simulations show that the coefficient which seems to affect the most the invasion aptitude is the mutation size coefficient. However, their model is restricted to local interactions.

In the next section, we wish to illustrate, by simulations of the stochastic discrete model, the effect of the spatial interaction range on the interplay between invasion and evolution, and on the emergence of spatial and phenotypic diversity (clustering and polymorphism). Our simulations focus on the qualitative differences between local and nonlocal interactions.

6 Simulations

This section presents simulations based on the algorithm of Section 3 of several examples motivated by theoretical and empirical biological issues. The Euler scheme used to simulate reflected diffusions is detailed in Section 6.1, as well as some numerically very efficient simplifications of the algorithm described in Section 3 in the specific case of linear death rates (logistic competition). Next, we give simulations of three biologically relevant examples. In the first one (Section 6.2), the parameters inducee an advantage for an optimal trait value that depends linearly on the spatial position. The simulations show a local adaptation at this optimal trait value, but not necessarily uniformly over space. When migrations and mutations are not too strong, a large spatial interaction range induces a spatial organization of the population as a finite set of isolated clusters. In Section 6.3, we propose a second example, where a similar phase transition occurs between clustering and non-clustering. In this example, clustering happens when the width of the spatial region of high growth rate is bigger than the spatial interaction range, and the critical interaction range can be identified quite precisely. In our last example (Section 6.4), we investigate a model describing the invasion of a species with evolving dispersal speed (the trait is proportional to the migration speed). The model is similar to the one numerically studied in [8], but we adapt it to the setting of [35]. The behavior of our model agrees with the qualitative observations of this article: the invasion speed increases progressively, due to evolution, and the invasion front is composed of faster individuals.

6.1 Algorithm for logistic interaction and Euler scheme

There is a particular interaction form for which the complexity of the algorithm can be considerably reduced. In Section 3, one needs to compute $I^\delta W * \nu(x,u)$ at some point $(x,u) \in \mathcal{X} \times \mathcal{U}$ at each time step, which involves a sum over all individuals in the population. In the case of *logistic competition* (linar death rate), where

$$\mu(x,u,r) = \mu_0(x,u) + \mu_1(x,u)r, \tag{6.1}$$

one can use the following algorithm, that avoids to compute such a sum at each time step.

Fix a constant C_δ in a similar way as in (2.7) such that

$$\mu_0(x,u) + \lambda^* + \|M^*\|_1 \leq C_\delta \text{ and } \mu_1(x,u)\|I^\delta W\|_\infty \leq C_\delta.$$

Take the Brownian motions $(B^{j,k})_{j,k\in\mathbb{N}}$ and the random variables $(\theta_k)_{k\in\mathbb{N}}$, $(V_k)_{k\in\mathbb{N}}$ and $(\tau_k)_{k\in\mathbb{N}}$ as in Section 3. Set $T_0 = 0$ and $N_0 = N$ (the initial number of individuals). Assume that $(T_{k-1}, N_{k-1}, X_{T_{k-1}}, U_{T_{k-1}})$ are given, where N_{k-1} is the number of individuals at time T_{k-1}. At this time, their positions and traits are the coordinates of the vectors $X_{T_{k-1}} = (X^i_{T_{k-1}})_{1\leq i\leq N_{k-1}}$ and $U_{T_{k-1}} = (U^i_{T_{k-1}})_{1\leq i\leq N_{k-1}}$.

The two first steps of the algorithm are the same: the next (possible) jump time is given by $T_k = T_{k-1} + \tau_k/N_{k-1}(N_{k-1}+1)$ and the motion of each particle is governed by the SDE with normal reflection (3.1).

The third step deals with the different events that may happen at time T_k. We start with the same procedure. Choose one individual $I_k = i$ uniformly at random among the N_{k-1} individuals living during the time interval $[T_{k-1}, T_k)$. Its position and trait are $(X^i_{T_k}, U^i_{T_{k-1}})$. The event occurring at time T_k is decided by comparing θ_k with constants related to the rate of each kind of event. The only difference with the algorithm of Section 3 is in the first sub-step, that has to be divided in two steps as follows:

- If $0 \leq \theta_k < \frac{N_{k-1}}{N_{k-1}+1} =: \theta^i_0(X_{T_k}, U_{T_{k-1}})$, then let $j \in \{1, \ldots, N_{k-1}\}$ be such that $\frac{j-1}{N_{k-1}+1} \leq \theta_k < \frac{j}{N_{k-1}+1}$. If $\theta_k - \frac{j-1}{N_{k-1}+1} < \frac{\mu_1(X^i_{T_k}, U^i_{T_{k-1}}) I^\delta(X^i_{T_k} - X^j_{T_k}) W(U^i_{T_{k-1}} - U^j_{T_{k-1}})}{C_\delta}$, then the individual i dies from competition with individual j and $N_k = N_{k-1} - 1$. Otherwise, nothing happens and $N_k = N_{k-1}$. This step amounts to choose j at random between 1 and N_{k-1} and next to do the acceptance-rejection procedure to determine if the competition exerted by the individual j "kills" i.

- If $\theta^i_0(X_{T_k}, U_{T_{k-1}}) \leq \theta_k \leq \theta^i_0(X_{T_k}, U_{T_{k-1}}) + \frac{\mu_0(X^i_{T_k}, U^i_{T_{k-1}})}{C_\delta(N_{k-1}+1)} =: \theta^i_1(X_{T_k}, U_{T_{k-1}})$, then the individual i dies by natural death and $N_k = N_{k-1} - 1$. This step is similar to the birth step without mutation.

The three other sub-steps are the same.

The main difference with the algorithm of Section 3 is that we no longer have to compute $\sum_{j=1}^{N_{k-1}} I^\delta(X^i_{T_k} - X^j_{T_k}) W(U^i_{T_{k-1}} - U^j_{T_{k-1}})$ in the first sub-step, but it suffices to compute $I^\delta(X^i_{T_k} - X^j_{T_k}) W(U^i_{T_{k-1}} - U^j_{T_{k-1}})$ for chosen i and j. Moreover, we do not need to compute the position of each individual in the population at each time step. The algorithm described above only requires to compute the position of at most two particles at time T_k (the particles numbered i, and possible j).

Next, as mentionned in Section 3, the reflected diffusion which drives the spatial motion of the particles can be simulated with an Euler scheme. We will assume in this subsection and in the following examples that $\mathcal{X} = (\alpha, \beta) \subset \mathbb{R}$ and we will use the scheme of Lépingle [26] (see also [19]). Fix $x \in [\alpha, \beta]$ and $u \in \mathcal{U}$. On any time interval where its trait is constant, an indivual at (x, u) moves according to the reflected diffusion

$$
\begin{aligned}
X_t &= x + \int_0^t \sqrt{2m(X_s, u)} dB_s + \int_0^t b(X_s, u) ds - k_t \ , \\
|k|_t &= \int_0^t \mathbf{1}_{\{X_s \in \{\alpha, \beta\}\}} d|k|_s \ ; \ k_t = \int_0^t \left(\mathbf{1}_{\{X_s = \beta\}} - \mathbf{1}_{\{X_s = \alpha\}} \right) d|k|_s,
\end{aligned}
\tag{6.2}
$$

where B is a one-dimensional Brownian motion.

If m and b are Lipschitz with respect to the first variable, this diffusion can be simulated on $[0,T]$ as follows. Fix $h>0$, $\bar\alpha$ and $\bar\beta$ such that $\alpha<\bar\alpha<\bar\beta<\beta$, and let n be the first integer greater than T/h. For $\rho\in\{0,1,\dots,n-1\}$ and $\rho h<t\le(\rho+1)h$, let

$$\begin{aligned}
\tilde X_0 &= x,\\
\tilde X_t &= \max\Big\{\alpha,\min\Big[\beta,\tilde X_{\rho h}+b(\tilde X_{\rho h},u)(t-\rho h)+\sqrt{2m(\tilde X_{\rho h},u)}(B_t-B_{\rho h})\\
&\qquad\qquad\qquad +\mathbf{1}_{\{\tilde X_{\rho h}<\bar\alpha\}}\max(0,A^\rho_t-(\tilde X_{\rho h}-\alpha))\\
&\qquad\qquad\qquad -\mathbf{1}_{\{\tilde X_{\rho h}>\bar\beta\}}\max(0,B^\rho_t+(\tilde X_{\rho h}-\beta))\Big]\Big\},
\end{aligned}$$

where

$$\begin{aligned}
A^\rho_t &= \sup_{\rho h\le s\le t}\Big\{-b(\tilde X_{\rho h},u)(s-\rho h)-\sqrt{2m(\tilde X_{\rho h},u)}(B_s-B_{\rho h})\Big\},\\
B^\rho_t &= \sup_{\rho h\le s\le t}\Big\{b(\tilde X_{\rho h},u)(s-\rho h)+\sqrt{2m(\tilde X_{\rho h},u)}(B_s-B_{\rho h})\Big\}.
\end{aligned}$$

Then, there exists a constant C independent of h such that for any function f on $[\alpha,\beta]$ with finite variation bounded by 1, $\sup_{0\le t\le T}|E(f(X_t)-f(\tilde X_t))|\le C\sqrt{h}$.

In each step of this scheme, one has to simulate simultaneously B_t and $S_t:=\sup_{s\le t}(aB_s+bs)$ for fixed constants a,b. This can be done as follows (Shepp [43]). Let U be a Gaussian centered random variable with variance t, and let V be an exponential random variable with parameter $1/(2t)$ independent of U. Put

$$Y=\frac{1}{2}\left(aU+bt+\sqrt{a^2V+(aU+bt)^2}\right).$$

Then the vectors (B_t,S_t) and (U,Y) have the same distribution.

Note that this scheme can be easily generalized to state spaces of the form $(\alpha_1,\beta_1)\times\ldots\times(\alpha_d,\beta_d)\subset\mathbb{R}^d$, as explained in [26].

We assume in the following examples a logistic competition of the form (6.1) and a physical space of the form $\mathcal{X}=(\alpha,\beta)$. Our simulations are realized with the previous algorithm.

6.2 Example 1. Spatial clustering

We consider here a set of parameters similar to the one of [12] and [36], in which, for each spatial position x, the growth rate is maximal for the trait value $u=x$. This can represent the effect of a gradual spatial distribution of different resources, involving a gradual distribution of traits. For example, for some bird species, a linearly spatially varying seed size determines a linear variation of the beak sizes (Grant and Grant [20]).

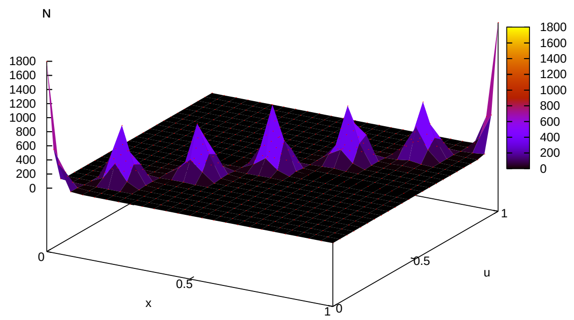
$$\mathcal{X} = (0,1), \quad \mathcal{U} = [0,1], \quad m(x,u) \equiv m, \quad b(x,u) \equiv 0,$$
$$\lambda(x,u) = 2 - 20(x-u)^2 \text{ if } |x-u| \leq 1/\sqrt{10};\ 0 \text{ otherwise},$$
$$\mu_N(x,u,r) = 1 + \frac{r}{N}, \quad I^\delta(y) = C_\delta \mathbf{1}_{\{|y|\leq\delta\}}, \quad W(v) \equiv 1.$$

Moreover, $M(x,u,v) = 0.1 \times k_s(u,v)$ where 0.1 is the mutation rate and $k_s(u,v)$ is the probability density of a Gaussian random variable with mean u and standard deviation s conditioned on staying in $\mathcal{U} = [0,1]$.

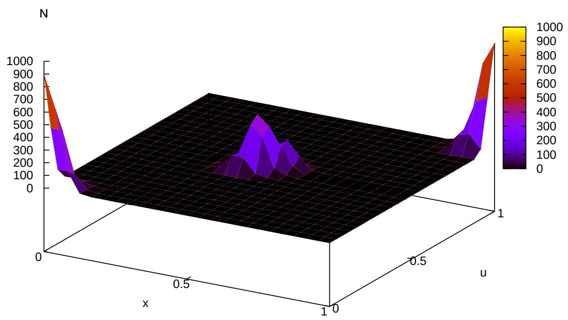
We assume in this model that there is no spatial heterogeneity ($b = 0$ and m constant) and that competition occurs independently of the trait ($W = 1$) between individuals distant of less than δ from each other. The birth rate λ is chosen such that it is maximal for $u = x$, and the death rate of an individual is assumed independent of its position and traits, with logistic form. There are four free parameters in this model, m, δ, s and the population size N. The initial population in our simulations is composed of N individuals at $(0.5, 0.5)$.

The simulations of this model show, as in [36], that the invasion of space occurs along the diagonal $x = u$, and, as in [12], that polymorphism (stable coexistence of several sub-populations with different typical traits) may occur in this model, accompanied with a spatial specialization. Several different population clusters may coexist at different position, with trait values located around the corresponding optimal traits. We have investigated in our simulations the effect of the different parameters on the clustering and polymorphism of the population. We give pictures of the seemingly stable state of the population (Fig. 1).

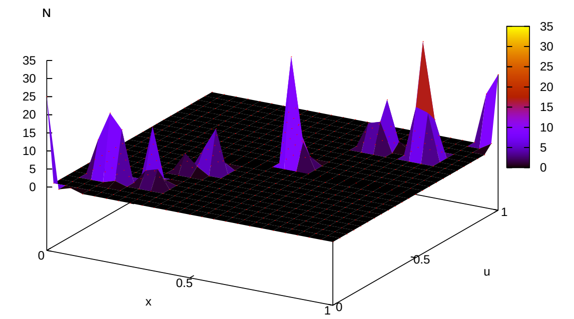
On the one hand, we have investigated the effect of the interaction range δ. The main result is that the existence, the number of clusters and the distance between clusters strongly depend on the interaction range δ. As shown in Fig. 1(a,b,d), the number of clusters decreases with δ and the distance between two population peaks is roughly 1.5δ, which is exactly the width of the interaction interval. This observation can be explained as follows. The emergence of population clusters is mainly a consequence of local births ([47]). Indeed, since the progeny of an individual is close to its original location, each individual's progeny can create a colony with stable position on short timescales. Once an individual is at a distance greater than δ from the main part of the population, it experiences very little competition and it can create a new colony. When several colonies appear, they organize in a way to minimize the competition between them and to maximize the growth rate. Moreover, if two population peaks are distant of more than 2δ, then there is a zone between them with in which an individual alone feels no competition at all. Hence, this



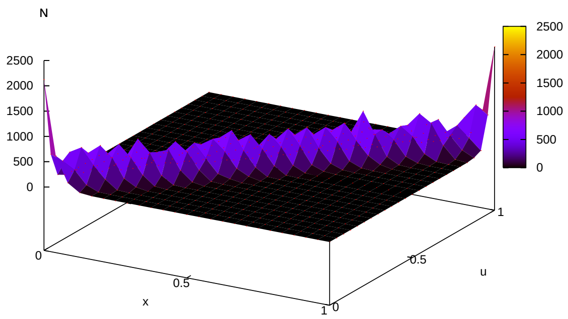
(a) $N = 3000, s = 0.01, m = 0.01, \delta = 0.1.$



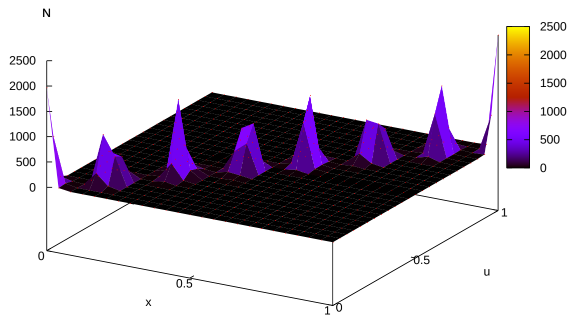
(b) $N = 3000, s = 0.01, m = 0.01, \delta = 0.3.$



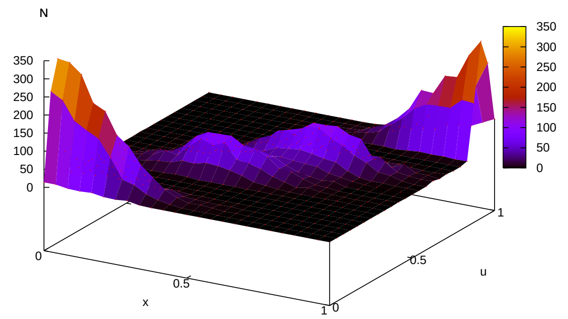
(c) $N = 50, s = 0.01, m = 0.01, \delta = 0.1.$



(d) $N = 3000, s = 0.01, m = 0.01, \delta = 0.03.$



(e) $N = 3000, s = 0.01, m = 0.003, \delta = 0.1.$



(f) $N = 3000, s = 0.01, m = 0.1, \delta = 0.1.$

Figure 1: Simulations of Example 1 for various parameters. All of them are taken at time 4000.

individual can create a new colony without any competitive pressure. The exact distance between two cluster is therefore between δ and 2δ. It also depends on the "width" of the clusters: if they are wide, then the distance between the center of each cluster has to bigger so that no competition is felt by a individual of one cluster from the individuals of the others. It is hard to have a clear idea of the proper definition of this notion of width of clusters. However it seems crucial to understand the phenomenon of clustering. Indeed, the width of a cluster should increase when the motion parameters m (diffusion strength) and s (range of mutations) increase. As observed in Fig. 1(d), if δ is sufficiently small with respect to this width, we observe a flat distribution of the population, and thus a qualitative difference with respect to cases (a,b). As an explanation, decreasing δ increases the number of clusters, and the width of a cluster stays constant when m and s are fixed. As δ decreases, the clusters eventually overlap and no distinct colony can be observed.

We also investigated the effect of the population size N. It appears that this parameter has very little qualitative effect on the clustering of the population, except for small N (Fig. 1(c)), where the width of each clusters is reduced, and we observe much more fluctuations in the population distribution. However, we still can observe a similar pattern of population clusters than in Fig. 1(a).

The impact of the diffusion parameter m and the mutation size s on the width of the clusters is further illustrated in Fig. 1(e) and (f), where all the parameters are the same as in Fig. 1(a), except the diffusion parameter. In Fig. (e), the diffusion parameter is smaller and two main consequences can be observed. First, clusters are narrower, as suggested by the previous heuristics; second, the number of coexisting clusters increases of 1 with respect to Fig. (a), illustrating the impact of the clusters'width on the distance between clusters. When a cluster is narrower, the other clusters next to it can be closer, because the individuals in these nearby clusters have to be at a distance δ from the boundary of the central cluster in order to feel no competitive pressure from it.

In Fig. (f), the diffusion parameter is bigger and four main consequences can be observed. First, the width of the clusters is bigger. Second, only four clusters can be distinguished for a similar reason as above. Third, the cluster has no longer a circular form because only m has been increased. The clusters are wider is the spatial direction than in the trait direction. Finally, the balance between the spatial interaction range δ and the motion parameters m and s is nearly critical for clustering, in the sense that we observe a population distribution which is intermediate between Fig. 1(a), where the clusters are clearly separated, and Fig 1(d), where the population distribution is flat and no cluster can be observed. In Figure (f), m and s are big enough with respect to δ so that quick movements mix the population. The two central clusters are overlapping and can hardly be distinguished, and no cluster at all can be observed if one looks at the spatial distribution only. Therefore, this case rather corresponds to a phenotypic clustering than to a spatial clustering. When the diffusion parameter m is kept constant and the mutation size s varies, qualitatively similar results can be observed in simulations (except that the roles of the physical space and the trait space has to be exchanged in our observations).

We conclude the study of this example with the general observation that the clusters are more concentrated at the boundary of the domain. This is a consequence of the fact that the reflected diffusion governing the motion of individuals is not isotropic close to this boundary, so that the population density is bigger.

6.3 Example 2. The role of spatial competition for clustering

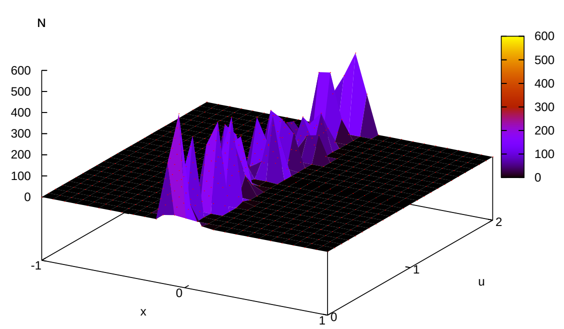
As we have seen above, the balance between the spatial competition range δ and the motion parameters s and m has a crucial effect on clustering in the population. In this section, we want to address the balance between the range of competition and the size of the region of space with high birth rate. For this purpose, we consider the following model, inspired by the adaptive dynamics model of [9]:

$$\mathcal{X} = (-1, 1), \quad \mathcal{U} = [0, 2], \quad m(x, u) \equiv m, \quad b(x, u) \equiv 0,$$
$$\lambda(x, u) = \exp(-x^2/2\rho^2), \quad \mu_N(x, u, r) = 1 + \frac{r}{N},$$
$$I^\delta(y) = C_\delta \exp(-y^2/2\delta^2), \quad W(v) = \exp(-v^2/0.02).$$

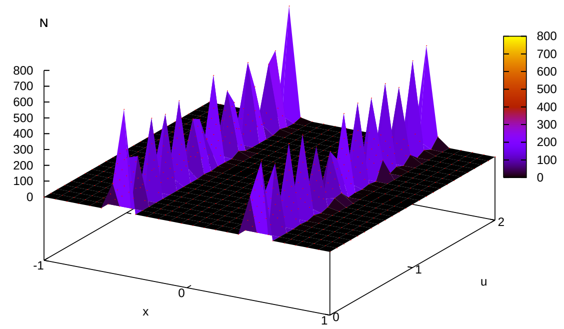
and the same mutation kernel as above. Here again, space is homogeneous, except for the birth rate, which is maximal at $x = 0$. The width of the space region with significant birth rate is ρ (which can also be interpreted as the width of the space region with high concentration of resources). The spatial competition kernel is of Gaussian shape, with standard deviation δ. The trait competition is also of Gaussian shape, with fixed standard deviation of 0.1. Therefore, trait competition is local. Observe that, in this example, the trait has no effect on the growth rate, so that the trait structure is neutral (all individual's parameters are equal, independent of the trait, in absence of interaction).

This example has five free parameters m, δ, s, N, and ρ. The initial population in our simulations is composed of N individuals at $(0, 1)$.

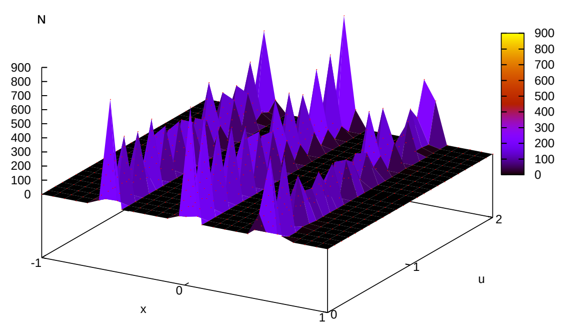
If we consider the space $\mathcal{X}$ as a trait space, this model is similar to the one of [9]. In particular, the biological theory of adaptive dynamics [31, 18] suggests that evolutionary branching, i.e. the split of the population into two sub-populations with different traits stably coexisting (translating in our model into spatial clustering) occurs when the range of



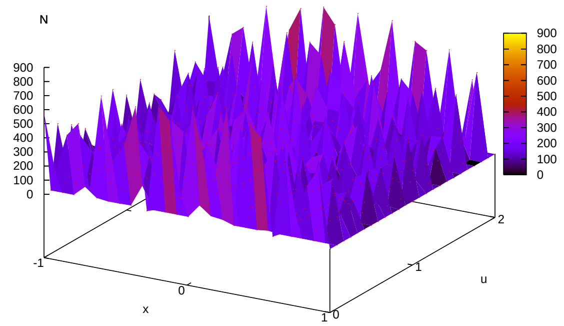
(a) $\delta = 1.1$, $\rho = 1$.



(b) $\delta = 0.9$, $\rho = 1$.



(c) $\delta = 0.5$, $\rho = 1$.



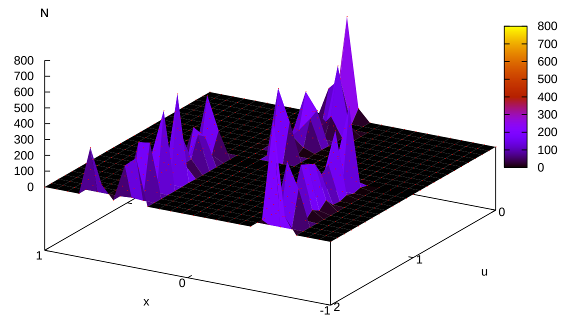
(d) $\delta = 0.1$, $\rho = 1$.

Figure 2: Simulations of Example 2. Neutral case. All the figures are taken at time 5000, except the last one, taken at time 10000 (more time is needed to fill the whole space). In all the simulations, $N = 1000$, $s = 0.003$ and $m = 0.003$.

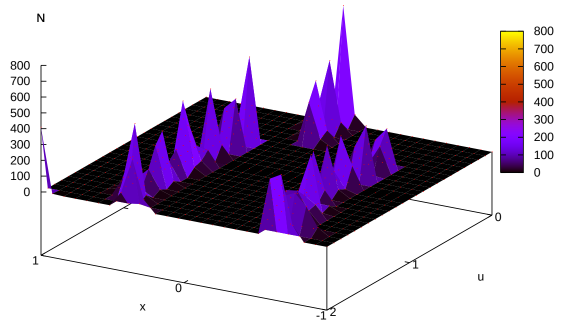
interaction δ is smaller than the width of the growth rate ρ. This is illustrated by Fig. 2(a) and (b), where, in (a), $\delta < \rho$ and the population stabilizes around two distinct positions (branching occurs) and in (b), $\delta > \rho$ and the population stabilizes around position 0 (there is no branching). This branching criterion reveals quite sharp in simulations.

Figures (c) and (d) prove that other phase transitions occur for smaller δ, leading to the coexistence of three clusters or more. As in Example 1, we notice in Figure (d) that a very small δ leads to a distribution without distinct clusters. This can be explained by similar arguments.

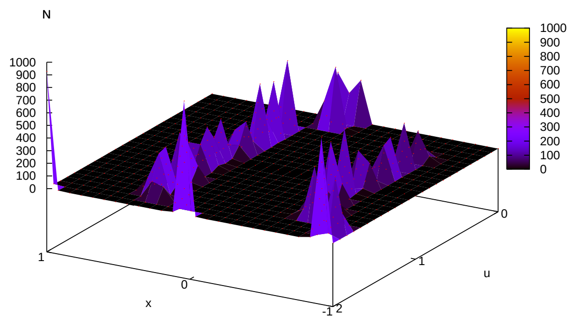
It is possible to add some phenotypic structure to this example by assuming that the growth rate depends on the trait u, in a way such that spatial branching occurs for some



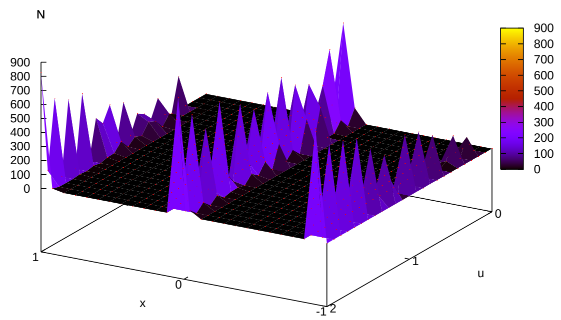
(a) $t = 10000$.



(b) $t = 20000$.



(c) $t = 45000$.



(d) $t = 80000$.

Figure 3: Simulations of Example 2. Trait-dependent case. In this simulation, $N = 1000, s = 0.003, m = 0.003$ and $\delta = 1$. These simulations are seen from a backward point of view for a better visualisation of the branches.

traits but not for others, according to the above branching criterion. We take the same parameters, except for the birth rate, which has the following form.

$$\lambda(x, u) = \exp(-x^2/2(u + 0.1)).$$

The parameter ρ is then replaced by $\sqrt{u + 0.1}$, so that branching occurs if $\sqrt{u + 0.1} > \delta$. The constant 0.1 has been added in order to avoid a null standard deviation for the Gaussian shape of the birth rate.

Actually, this is what happens for small times (Fig. 3(a)), but after a longer time (Fig. 3(b,c)), the two clusters spread over the trait space because of the mutation phenomenon. Eventually, we actually observe the appearance and the spread of three spatial clusters, colonizing all the trait space (Fig. 3(d)).

6.4 Example 3. Invasion and evolution of migration speed

Here, we investigate a model describing the invasion of a species with evolving dispersal speed (as in [8]). This can model phenomena such as the invasion of Australia by cane toads, for which an adaptation to high invasion speeds has been recently detected (Phillips et al. [35]). In particular, it has been observed that the size of the legs of the toads, which is thought to be related to their speed, is larger at the front of the invasion, and that the speed of the invasion front increases progressively over years, because of the adaptation of toads to higher speed.

The parameters are as follows.

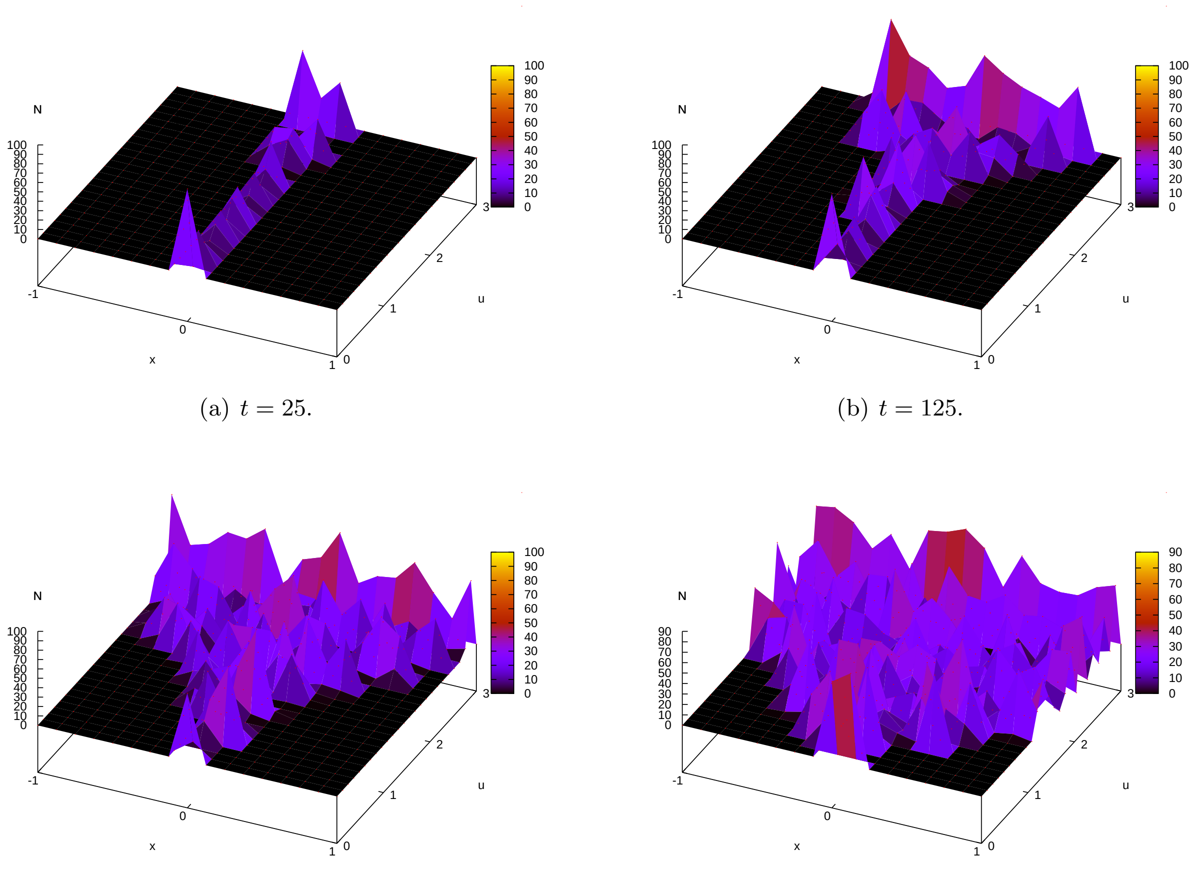
$$\mathcal{X} = (-1,1), \quad \mathcal{U} = [0,3], \quad m(x,u) \equiv m(u+0.1), \quad b(x,u) \equiv 0,$$
$$\lambda(x,u) = 1, \quad \mu_N(x,u,r) = 1 + \frac{r}{N},$$
$$I^\delta(y) = C_\delta \mathbf{1}_{\{|y|\leq\delta\}}, \quad W(v) = \exp(-10v^2).$$

and the same mutation kernel as above. Here we study invasion into an homogeneous space and trait domain (λ is constant, $b = 0$ and μ is independent of x and u). The trait competition kernel W is of Gaussian form, with standard deviation $1/2\sqrt{5} \approx 0.22$. Individuals interact only if they are distant of less than δ from each other. The diffusion rate m is proportional (up to a constant, in order to avoid a null diffusion constant when $u = 0$) to the trait u. Thus, individuals with large u move fast. The trait u can be a morphological trait responsible for the speed of dispersal (e.g. the length of legs for toads, [35]). This example has four free parameters, the diffusion coefficient m, the interaction range δ, the standard deviation of mutations s and the population size N.

In Fig. 4 and 5, we present two extreme cases with respect to the initial trait distribution, but with identical parameters. In the first one, all individuals are at (physical) position 0, and with traits regularly distributed in $\mathcal{U} = [0,3]$. In the second one, all individuals are initially located at a single point $(0,0)$.

In both figures, we observe a triangular invasion pattern indicating that the invasion front is composed of faster individuals. In Fig. 5, we also observe a simultaneous invasion in physical and trait spaces, and a slower spread of the population. This can be explained by the fact that the population, initially composed of slow individuals, has first to colonize the trait space before invading the physical space. Because of the progressive appearance of larger traits, the invasion speed increases over time (compare the different time values in Fig. 5). This agrees with the observations of [35].

When the parameters vary, the simulations of this microscopic model can show very different patterns of colonization. As an illustration of the richness of this model, we give



(a) $t = 25$. (b) $t = 125$.

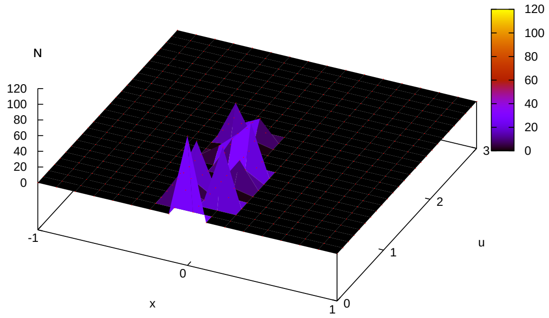
(c) $t = 250$. (d) $t = 500$.

Figure 4: Simulations of Example 3. The parameters are $N = 100, s = 0.03, m = 0.003$ and $\delta = 0.1$. The initial condition is composed of N individuals located at 0 and with trait values $3i/N$ for $1 \leq i \leq N$.

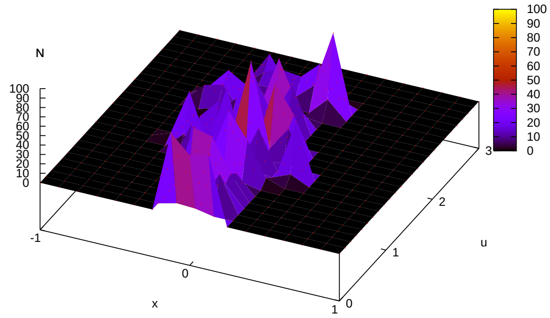
an example (Fig. 6) where the interaction range δ is bigger. The parameters N and m are chosen such that two clusters appear for large traits and spread over the trait space in a short time. The initial condition is the same as in Fig. 4. We can observe two branches linking the initial cluster with the two extreme valued clusters (Fig. 6(c,d)).

7 Conclusion

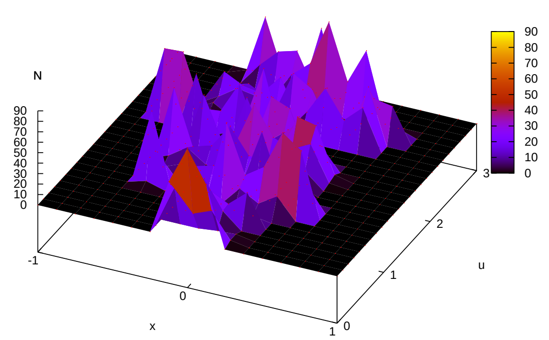
We have proposed different levels of models which describe the interplay between evolution and space. We started from a microscopic detailled stochastic model, in which each individual's birth and death are described, and which include ecological interactions between individuals. We have also given an efficient algorithm allowing us to construct and sim-



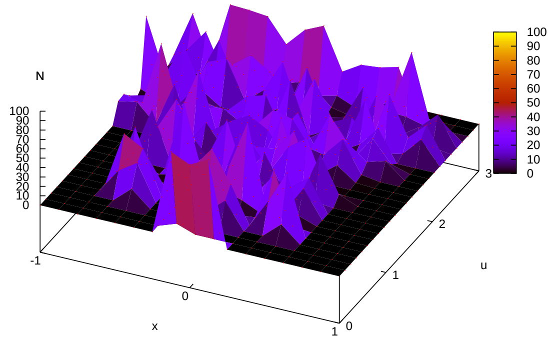
(a) $t = 500$.



(b) $t = 750$.

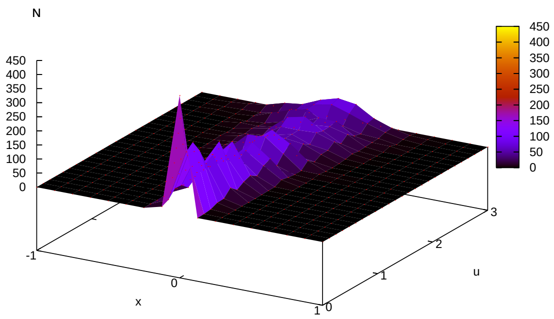


(c) $t = 850$.

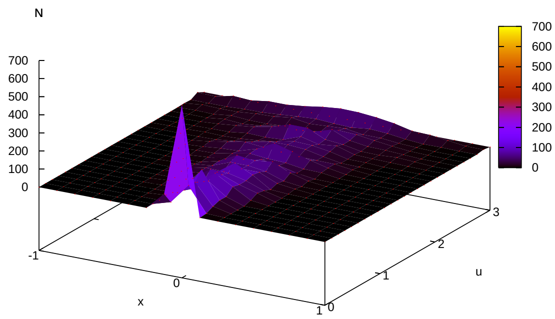


(d) $t = 1000$.

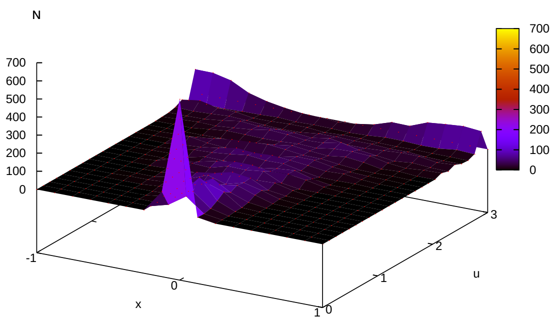
Figure 5: Simulations of Example 3. The parameters are $N = 100, s = 0.03, m = 0.003$ and $\delta = 0.1$. The initial condition is composed of N individuals located at $(0, 0)$.



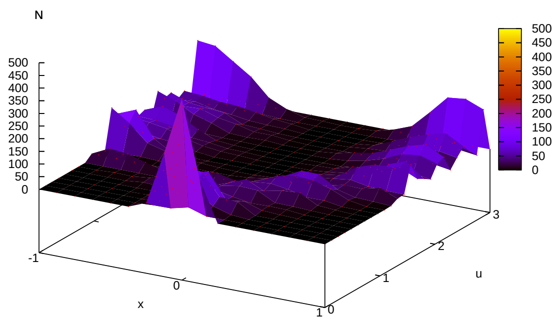
(a) $t = 5$.



(b) $t = 10$.



(c) $t = 20$.



(d) $t = 50$.

Figure 6: Simulations of Example 3. The parameters are $N = 1000, s = 0.03, m = 0.03$ and $\delta = 1$. The initial condition is composed of N individuals located at 0 and with trait values $3i/N$ for $1 \leq i \leq N$.

ulate such a population process. We then obtained a simpler, macroscopic deterministic model in the form of a nonlocal and nonlinear partial integro-differential equation. In order to recover deterministic models already studied in [36] where the spatial interaction is local, we then studied the limit of small spatial interaction range, and obtained a spatially localized nonlinear partial intergro-differential equation. The local and nonlocal models show very different behavior.

We then made the numerical study of several biological examples, in order to illustrate the role of the spatial interaction range. More precisely, our first two examples illustrated the local adaptation to the maximal resources exploitation (Example 1) and the interplay between the various parameters in terms of spatial clustering (Examples 1 and 2). Specifically, we have shown that the number of clusters decreases when the spatial interaction range increases and that the width of clusters increases when the speed of spatial motion and the range of mutation increase. When these two parameters are too big, then, no cluster is observed, and the population distribution is flat. The impact of the population size on clustering seems quite weak. We also shown that the width of the space region with high growth rate has to be sufficiently large in order to obtain clustering. In our second example, we were able to identify quite precisely the critical range of interaction for which clustering appears. We also studied another class of models (Example 3) describing the interplay between invasion and evolution. We were in particular able to recover in simulations the experimental observations on the invasion of Australia by toads obtained in [35]. Our numerical study of this model shows a wide variety of patterns of invasion, depending on the parameters.

Further works can be developed in many directions. From a mathematical point of view, several other macroscopic limits of the stochastic individual process can be studied, in a similar way as in [6]. This would allow to obtain macroscopic models where the mutation term has a reaction-diffusion form, and to include is these models a demographic stochasticity. Such stochastic partial differential equations are biologically very interesting, but their mathematical study and simulation are quite difficult. Many other questions can be asked about the clustering phenomenon, and its evolutionary counterpart, called "evolutionary branching". One could study the stationary distribution of the individual process, and look at its form. Under suitable assumptions, for example when motion is slow, the stationary distribution could converge to a sum of Dirac masses (clusters). A challenging problem would then be to determine the distance between each clusters, and their size. Another interesting question concerns the size of the clusters, and the way they depend on the motion parameters. We have seen in our first example that this notion is important to determine the distance between cluster. The dynamical study of branching and clustering (i.e. the transition from one cluster to two or more clusters) is also an important problem. The adaptive dynamics theory provides a first progress in this direction [31, 18]. From a biological point of view, the clustering phenomenon is related to the notion of ecological niches, and it is the basis of the metapopulation models, which are extensively studied. It is therefore important to have criterions (for example on the motion parameters and the range of competition, or the width of the space region with high birth rate) allowing to validate such models. Invasion phenomena are also extensively studied. With our example, we wanted to point out the fact that evolution has also a crucial role in these phenomena. Our model is a first attempt to describe the interplay between evolution and invasion from a microscopic point af view. The richness of the dynamics we obtain shows that the study of this problem is difficult. A first step could be to try to list all type of behavior that can be expected in our model and to identify the parameters that are responsible for each particular type of invasion.

Acknowledgments. The authors would like to thank Laurent Desvillettes who pointed out to our attention the interest of combining space and traits, and the article [35]. They also thank the other participants of the ACI "Structured Populations" and more specifically Régis Ferrière for fruitful discussions.

References

[1] Aldous, D. (1978). Stopping Times and Tightness. Ann. Prob. 6, 335–340.

[2] Bolker B., Pacala, S.W. (1997). Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. Theor. Popul. Biol. 52, 179–197.

[3] Bolker, B.M., Pacala, S.W. (1999). Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. Am. Nat. 153, 575–602.

[4] Bossy, M., Gobet, E., Talay, D. (2004). A symmetrized Euler scheme for an afficient approximation of reflected diffusions. J. Appl. Probab. 41, 877–889.

[5] Rudin, W. (1987). Real and complex analysis. Third edition, McGraw-Hill Book Co., New York.

[6] Champagnat, N., Ferrière, R., Méléard, S. (2006). Unifying evolutionary dynamics: from individual stochastic processes to macroscopic models. To appear in Theor. Popul. Biol.

[7] Champagnat, N., Ferrière, R., Méléard, S. (2006). Individual-based probabilistic models and various time scaling approximations in adaptive evolution. To appear in the "Progress in Probability" Series, Birkhäuser.

[8] Desvillettes, L., Ferrière, R., Prevost, C. (2004). Infinite dimensional Reaction-Diffusion for Population Dynamics. Preprint CMLA, ENS Cachan.

[9] Dieckmann, U., Doebeli, M. (1999). On the origin of species by sympatric speciation. Nature 400, 354–357.

[10] Dieckmann, U., Law, R. (2000). Relaxation projections and the method of moments. In: Dieckmann, U., Law, R., Metz. J.A.J. (Eds.), The Geometry of Ecological Interactions: Symplifying Spatial Complexity, 412–455. Cambridge University Press, Cambridge.

[11] Dieckmann, U., Law, R., Metz, J.A.J. (2000). The Geometry of Ecological Interactions: Symplifying Spatial Complexity. Cambridge University Press, Cambridge.

[12] Doebeli, M., Dieckmann, U. (2003). Speciation along environmental gradients. Nature 421, 259–263.

[13] Durrett, R., Levin, S. (1994). The importance of being discrete (and spatial). Theor. Pop. Biol. 46, 363–394.

[14] Durrett, R., Levin, S. (1994). Stochastic spatial models: a user's guide to ecological applications. Phil. Trans. Roy. Soc. London, 343, 329–350.

[15] Endler, J.A. (1977). Geographic variation, speciation, and clines. Princeton university Press.

[16] Flierl, G., Grünbaum, D., Levin, S., Olson, D. (1999). From individuals to aggregations: the interplay between behaviour and physics. J. Theor. Biol. 196, 397–454.

[17] Fournier, N., Méléard, S. (2004). A microscopic probabilistic description of a locally regulated population and macroscopic approximations. Ann. Appl. Probab. 14, 1880–1919.

[18] Geritz, S.A.H., Metz, J.A.J., Kisdi, E., Meszena, G. (1997). The dynamics of adaptation and evolutionary branching. Phys. Rev. Lett. 78, 2024–2027.

[19] Gobet, E. (2001). Euler schemes and half-space approximations for the simulation of diffusion in a domain. ESAIM PS 5, 261–293.

[20] Grant, P.R., Grant, B.R. (2002). Unpredictable evolution in a 30-year study of Darwin's finches. Science 296, 707–711.

[21] Hassel, M.P., May, R.M. (1974). Aggregation in predators and insect parasites and its effects on stability. J. Animal Ecol. 43, 567–594.

[22] Hassel, M.P., Pacala, S.W. (1990). Heterogeneity and the dynamics of host parasitoid interactions. Phil. Trans. Roy. Soc. London B 330, 203–220.

[23] Jacod, J., Shiryaev, A.N. (1987). Limit Theorems for Stochastic Processes. Springer-Verlag.

[24] Joffe, A., Métivier, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. Adv. Appl. Prob. 18, 20–65.

[25] Law, R., Murrell, D.J., Dieckmann, U. (2003). Population growth in space and time: spatial logistic equations. Ecology 84 (1), 252–262.

[26] Lépingle, D. (1995). Euler scheme for reflected stochastic differential equations. Math. Comp. Simul. 38, 119–126.

[27] Lewis, M.A., Pacala, S. (2000). Modeling and analysis of stochastic invasion processes. J. Math. Biol. 41, 387–429.

[28] McGlade, J. (1999). Advanced Ecological Theory: Principles and Applications. Blackwell Science, Oxford.

[29] Mayr, E. (1963). Animal Species and Evolution. Harvard University Press, Cambridge.

[30] Méléard, S., Roelly, S. (1993). Sur les convergences étroite ou vague de processus à valeurs mesures. C. R. Acad. Sci. Paris Sr. I Math. 317, 785–788.

[31] Metz, J.A.J., Geritz, S.A.H., Meszéna, G., Jacobs, F.A.J., van Heerwaarden, J.S. (1996). Adaptive Dynamics, a geometrical study of the consequences of nearly faithful reproduction. In: van Strien, S.J., Verduyn Lunel, S.M. (Eds.), Stochastic and Spatial Structures of Dynamical Systems. North Holland, Amsterdam, pp. 183–231.

[32] Mollison, D. (1977). Spatisl contact models for ecological and epidemic spread. J. Roy. Statist. Soc. B 39, 283–326.

[33] Murray, J.D. (1989). Mathematical Biology. Biomathematics texts 19, Springer-Verlag, Berlin.

[34] Niwa, H.S. (1994). Self-organizing dynamic-model of fish schooling. J. Theor. Biol. 171, 123–136.

[35] Phillips, B.L., Brown, G.P., Webb, J.K., Shine, R. (2006). Invasion and the evolution of speed in toads. Nature, 439, 803.

[36] Prévost, C. (2004). Applications des équations aux dérivées partielles aux problèmes de dynamique des populations et traitement numérique. PhD thesis, Université d'Orléans (France).

[37] Rainey, P.B., Travisano, M. (1998). Adaptive radiation in heteregeneous environment. Nature 394, 69–72.

[38] Rand, D.A., Keeling, M.J., Wilson, H.B. (1995). Invasion, stability and evolution to criticality in spatially extended, artificial host-pathogen ecologies. Proc. Roy. Soc. London B 259, 55–63.

[39] Roelly-Coppoletta, S. (1986). A criterion of convergence of measure-valued processes: application to measure-valued branching processes. Stochastics 17, 43–65.

[40] Roelly, S., Rouault, A. (1990). Construction et propriétés de martingales des branchements spatiaux interactifs. International Statistical Review 58 (2), 173–189.

[41] Roughgarden, J. (1972). Evolution of niche width. Am. Nat. 106, 683–718.

[42] Sato, K., Ueno, T. (1965). Multi-dimensional diffusion and the Markov process on the boundary. J. Math. Kyoto Univ. 4(3) 529–605.

[43] Shepp, L.A. (1979). The joint density of the maximum and its location for a Wiener process with drift. J. Appl. Probab. 16, 423–427.

[44] Slatkin, M. (1978). Spatial patterns in the distribution of phylogenetic characters. J. Theor. Biol. 197, 149–162.

[45] Thomas, C.D., Bodsworth, E.J., Wilson, R.J., Simmons, A.D., Davies, Z.G., Musche, M., Conradt, L. (2001). Ecological and evolutionary processes at expanding range margins. Nature 411, 577–581.

[46] Tilman, D., Kareiva, P. (1996). Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions. Princeton University Press, Princeton, NJ.

[47] Young, W.R., Roberts, A.J., Stuhne, G. (2001). Reproductive pair correlations and the clustering of organisms. Nature 412, 328–331.